

# Analysis and Forecast of the Capesize Bulk Carriers Shipping Market using Artificial Neural Networks

By

Athanasios V. Voudris

B.S. in Naval Architecture and Marine Engineering (2005)

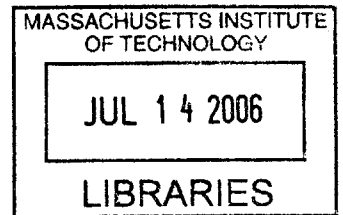
National Technical University of Athens

Submitted to the Department of Mechanical Engineering  
in Partial Fulfillment of the Requirement for the Degree of  
**Master of Science in Ocean Systems Management**

At the

**Massachusetts Institute of Technology**

June 2006



© 2006 Athanasios V. Voudris  
All rights reserved

**BARKER**

The author hereby grants to MIT permission to reproduce and to distribute publicly  
paper and electronic copies of this thesis document in whole or in part.

Signature of Author.....

A handwritten signature in black ink, appearing to be "Athanasios V. Voudris".

.....  
Department of Mechanical Engineering  
May 12, 2006

Certified by.....

A handwritten signature in black ink, appearing to be "Henry S. Marcus".

.....  
Henry S. Marcus  
Professor of Marine Systems  
Thesis Supervisor

Accepted by.....

.....  
Lallit Anand  
Professor of Mechanical Engineering  
Chairman, Department Committee on Graduate Students

# **Analysis and Forecast of the Capesize Bulk carriers Shipping Market using Artificial Neural Networks**

By

Athanasios V. Voudris

Submitted to the Department of Mechanical Engineering  
in Partial Fulfillment of the Requirement for the Degree of  
Master of Science in Ocean Systems Management

## **ABSTRACT**

Investing in the bulk carrier market constitutes a rather risky investment due to the volatility of the bulk carrier freight rates. In this study it is attempted to uncover the benefits of using Artificial Neural Networks (ANNs) in forecasting the Capesize Ore Voyage Rates from Tubarao to Rotterdam with a 145,000 dwt Bulk carrier. Initially, market analysis allows the assessment of the relation of some parameters of the dry bulk market with the evolution of freight rates. Subsequently, ANNs with an appropriate architecture are constructed and sufficient data, in terms of quantity and quality, are collected and organized so as to establish both the training and the testing data sets. The use of ANNs along with genetic algorithms allows the prediction of bulk freight rates with considerable accuracy for as long as eighteen months ahead and this is quantified by calculating the relative and absolute errors. It is concluded that ANNs offer a promising approach to forecasting the bulk market when coupled with efficient market modeling.

Thesis Supervisor: Henry S. Marcus

Title: Professor of Marine Systems

## **Acknowledgments**

I would like to express my thanks and appreciation to Prof. Henry S. Marcus for his valuable advice and guidance during my academic career at MIT. His encouragement always provided me with a new perspective in the field of Ocean Systems Management. The expertise that he shared with me will remain a tremendous source of professional growth to me.

I am also thankful to all of my professors at MIT, who with their expertise, professionalism, interest and assistance played an important role in the progress of my studies.

Special thanks to my friends Nikolas Veniamis, Aristides Miroyannis and Nicholas Petrakakos for making this year in Cambridge exciting and so unique.

Foremost, I would like to thank my family, Vasilis, Yvonne and Kostis, who have always encouraged me to follow my dreams and provided me with invaluable guidance and inspiration.

Finally, I would like to dedicate this work to my Grandmother and Georgina.

# Contents

<b>1 Introduction.....</b>	<b>13</b>
1.1 Aims and Objectives.....	14
1.2 Perspective.....	15
1.3 Organization of the study.....	16
<b>2 Economics of the Shipping market - Dry Bulk market.....</b>	<b>18</b>
2.1 Shipping cycles and shipping risk.....	18
2.1.1 The stages in the shipping market cycle.....	20
2.1.2 The frequency of shipping cycles.....	23
2.2 The four shipping markets.....	25
2.3 The freight rate mechanism.....	28
2.3.1 The supply and demand functions.....	29
2.3.2 Equilibrium and the importance of time.....	32
2.3.3 Momentary equilibrium.....	33
2.3.4 The short run equilibrium.....	35
2.3.5 The long run.....	36
2.4 The five major dry bulks.....	40
2.4.1 The seaborne iron ore trade.....	41
2.4.2 The transport system for iron ore.....	46
2.5 Dry Bulk Trade Outlook April 2006.....	48
2.6 Iron Ore – Australia and Brazil the dominant players in 2006.....	54
2.6.1 China Driving Demand.....	54
2.6.2 Concentrating the Supply Side.....	55
2.6.3 Trading Partners.....	56
2.6.4 Ore Shipping Demand.....	56
<b>3 Artificial Neural Networks.....</b>	<b>57</b>
3.1 Introduction.....	57

3.2	Brief History of Artificial Neural Networks.....	58
3.3	Biological Neural Networks .....	60
3.4	Definition .....	63
3.5	Memorization and Generalization.....	64
3.6	Acquiring Knowledge by Learning .....	65
3.7	An Artificial Neuron.....	68
3.8	A Single Layer Network .....	70
3.9	Multilayer Neural Networks .....	70
3.10	The Concept of Modularity.....	76
3.11	Modular Artificial Neural Networks.....	77
3.11.1	Model Complexity Reduction.....	79
3.11.2	Robustness .....	79
3.11.3	Scalability .....	80
3.11.4	Learning .....	80
3.11.5	Computational Efficiency .....	81
3.11.6	Learning Capacity.....	81
3.11.7	Economy of Learning .....	82
3.11.8	Knowledge Integration.....	82
3.12	Artificial Neural Networks as tools of forecasting .....	82
<b>4</b>	<b>Construction of Database – Data Analysis .....</b>	<b>84</b>
4.1	Introduction.....	84
4.2	“Clarksons” .....	85
4.3	“Lloyd's Shipping Economist” .....	86
4.4	Tables with data from “Clarksons” and the “Lloyd's Shipping Economist” .....	88
4.5	Available and exploitable Time series.....	97
4.6	The selection of independent time series .....	98
4.6.1	Theoretical approach.....	99
4.6.2	Autocorrelation and cross-correlation with time lag .....	99
4.7	Pearson Correlation - Definition.....	106
4.7.1	Pearson correlations for the current time series.....	107

4.7.2	Conclusions.....	112
4.7.3	Rejection of strong correlated independent variables.....	112
4.8	Evaluation of the database that will be used in the Neural Networks .....	112
4.9	Import of data in the program NeuroSolutions 4.0 .....	113
<b>5</b>	<b>Selection of the Artificial Neural Network .....</b>	<b>117</b>
5.1	Introduction.....	117
5.2	Layers - Number of Neurons – Transfer Function.....	117
5.3	Weights Adjustment - Error Function - Training Algorithms .....	120
5.4	Stop Training .....	122
5.5	Construction of the Neural Network.....	123
5.5.1	Selection of the type of the Neural Network .....	124
5.5.2	Import of Training Data.....	125
5.5.3	Import of desired response.....	129
5.5.4	Import of cross validation and test data .....	129
5.5.5	Selection of the number of hidden layers, their parameters and the network topology.. .....	131
5.5.6	Termination of the training.....	135
5.6	Analytical explanation of the neural network topology.....	137
5.6.1	Layer No 1 .....	138
5.6.2	Hidden layer No 1 .....	138
5.6.3	Hidden layer No 2.....	139
5.6.4	Layer No 2 .....	140
5.7	Performance Measuring.....	141
<b>6</b>	<b>Neural Networks Results.....</b>	<b>146</b>
6.1	Introduction.....	146
6.2	Results Presentation .....	146
6.2.1	Training Data .....	147
6.2.2	Cross-Validation Data.....	148
6.2.3	Testing Unknown Data .....	148

6.3	Performance of the four Neural Networks.....	149
6.4	Evaluation of the four Neural Networks.....	152
6.4.1	Training Data Evaluation.....	153
6.4.2	Cross-Validation Data Evaluation.....	159
6.4.3	Testing Unknown Data Evaluation.....	165
<b>7</b>	<b>Conclusions.....</b>	<b>182</b>
7.1	Introduction.....	182
7.2	Conclusions.....	182
7.3	Additional conclusions.....	185
7.4	Future Work.....	186
7.4.1	Input Data.....	186
7.4.2	Learning Algorithms.....	186
7.4.3	Use of more sophisticated genetic algorithms.....	187
7.4.4	Alternative methods of stopping training.....	187
	<b>References.....</b>	<b>188</b>
	<b>Appendix.....</b>	<b>190</b>

## List of Figures

Figure 2.1: Stages in a dry market cargo cycle.....	21
Figure 2.2: Twelve dry freight cycles 1869-1995.....	24
Figure 2.3: Dry cargo time charter rates .....	25
Figure 2.4: The four markets which control shipping.....	28
Figure 2.5: Shipping supply and demand functions.....	30
Figure 2.6: Momentary equilibrium in the VLCC market.....	34
Figure 2.7: Short run equilibrium. ....	35
Figure 2.8: Long-term adjustments of supply and demand 1980-92.....	38
Figure 2.9: Iron Ore Imports 1962-95.....	44
Figure 2.10: Major Iron Ore exporters and ports 2005.....	45
Figure 2.11a: OECD Industrial Prod. & Leading Indicator.....	48
Figure 2.11b: OECD Growth/Dry Trade Growth.....	48
Figure 2.12: Bulker Spot Earnings (\$/Day) .....	51
Figure 2.13: Total Iron Ore Imports from 1984-2006 .....	53
Figure 2.14: Steel Production and Capesize Freight Rates from 1993-2006.....	54
Figure 2.15: China's Iron Ore Imports from 2000-2006 .....	55
Figure 3.1: Simplified Biological Neurons.....	62
Figure 3.2: Learning and the Problem of Overfitting. ....	66
Figure 3.4: Examples of Multilayer Neural Network Architectures.....	71
Figure 3.5: The Backpropagation Network. ....	72
Figure 3.6: The influence of the Learning Rate on the Weight Changes. ....	75
Figure 4.1: Demand, Supply and Surplus of bulk Carriers from 1985 until February 2006.....	90
Figure 4.2: 1 Year Timecharter rates for 127,500 DWT and 120K DWT Capesize Newbuilding prices.....	92
Figure 4.3: Capesize Scrap Value and Bulk Carrier Orderbook.....	94

Figure 4.4: Capesize Ore Voyage Rates from Tubarao to Rotterdam for a 145,000 DWT Bulk Carrier. ....	96
Figure 4.5: Autocorrelation of time series that will be forecasted.....	105
Figure 4.6: Partial autocorrelation of time series that will be forecasted. ....	105
Figure 4.7a: Cross-correlation of Capesize ore rates with Total Supply. ....	109
Figure 4.7b: Cross-correlation of Capesize ore rates with Total Surplus. ....	109
Figure 4.7c: Cross-correlation of Capesize ore rates with 1 year Timecharter. ....	110
Figure 4.7d: Cross-correlation of Capesize ore rates with Scrap Value. ....	110
Figure 4.7e: Cross-correlation of Capesize ore rates with Total Order book. ....	111
Figure 4.7f: Cross-correlation of Capesize ore rates with New Building prices. ....	111
Figure 5.1: An analytical description of a neuron.....	119
Figure 5.2: The training procedure of a Neural Network. ....	121
Figure 5.4: A conventional Genetic Algorithm. ....	127
Figure 5.5: Example of how neurons and hidden layers are arranged in two modules. ...	136
Figure 5.6: The mapping for the PE of every Axon. ....	138
Figure 6.1: ANN Training Data with 3 months Delay.....	155
Figure 6.2: ANN Training Data with 6 months Delay.....	156
Figure 6.3: ANN Training Data with 12 months Delay.....	157
Figure 6.4: ANN Training Data with 18 months Delay.....	158
Figure 6.5: Distribution of the ANN Training data with 3 months of delay. ....	159
Figure 6.6: Distribution of the ANN Training data with 6 months of delay. ....	159
Figure 6.7: Distribution of the ANN Training data with 12 months of delay. ....	159
Figure 6.8: Distribution of the ANN Training data with 18 months of delay. ....	159
Figure 6.9: ANN Cross-Validation Data with 3 months Delay. ....	161
Figure 6.10: ANN Cross-Validation Data with 6 months Delay. ....	162
Figure 6.11: ANN Cross-Validation Data with 12 months Delay. ....	163
Figure 6.12: ANN Cross-Validation Data with 18 months Delay. ....	164
Figure 6.13a: Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 3 months Delay (in Lines).....	168

Figure 6.13b: Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 3 months Delay (in Bars).....	169
Figure 6.14a: Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 6 months Delay (in Lines).....	170
Figure 6.14b: Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 6 months Delay (in Bars).....	171
Figure 6.15a: Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 6 months Delay (in Lines).....	172
Figure 6.15b: Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 6 months Delay (in Bars).....	173
Figure 6.16a: Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 6 months Delay (in Lines).....	174
Figure 6.16b: Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 6 months Delay (in Bars).....	175
Figure 6.17: Forecast of March 2006 Capesize Ore Voyage Rate from all four ANN. ...	176
Figure 6.18: Forecast of April 2006 Capesize Ore Voyage Rate from all four ANN. ....	177
Figure 6.19: May 2006 Capesize Ore Voyage Rates Forecast from ANN with 3 months Delay.....	178
Figure 6.20: August 2006 Capesize Ore Voyage Rates Forecast from ANN with 6 months Delay.....	179
Figure 6.21: February 2007 Capesize Ore Voyage Rates Forecast from ANN with 12 months Delay.....	180
Figure 6.22: August 2007 Capesize Ore Voyage Rates Forecast from ANN with 18 months Delay.....	181

## List of Tables

Table 2.1: The five “major” bulk commodities shipped by sea (mt).....	40
Table 2.2: Seaborne Trade Data April 2006 .....	49
Table 2.3: No. of Vessels.....	51
Table 2.4: Trade growth in Million mt and % change yoy.....	51
Table 2.5: Fleet in Million DWT and % change yoy.....	51
Table 2.6: Iron Ore Imports to Europe from 1998-2006 .....	51
Table 2.7: Iron Ore Imports to Asia from 1998-2006.....	52
Table 2.8: Iron Ore Imports to Other Countries from 1998-2006 .....	52
Table 2.9: Total Iron Ore Imports from 1998-2006.....	52
Table 2.10: Iron Ore Exports from 2001-2006 .....	53
Table 4.1: Data of demand, supply and surplus of bulk carriers from 1985 until February 2006 as they have been collected from the Lloyd's Shipping Economist. (All data are shown in the appendix) .....	89
Table 4.2: Data for one year bulk carrier Timecharter rates of 127,500 DWT and newbuilding prices of a 120K DWT bulk carrier from 1985 until February 2006 as they have been collected from Clarksons. (All data are shown in the appendix).....	91
Table 4.3: Data for the capesize scrap value of a bulk carrier and the total order book of bulk carriers from 1985 until February 2006 as they have been collected from Clarksons and Lloyd's Shipping Economist. (All data are shown in the appendix). .....	93
Table 4.4: Data for the capesize ore voyage rates from Tubarao to Rotterdam for a 145,000 DWT bulk carrier 1985 until February 2006 as they have been collected from Clarksons. (All data are shown in the appendix).....	95
Table 4.5a: Time series of capesize ore voyage rates with number (1).....	97
Table 4.5b: Time series of supply, demand and surplus with numbers (2-4).....	97
Table 4.5c: Time series for Timecharter and New buildings with number (5-6).....	97
Table 4.5d: Time series for Scrap value and order book with numbers (7-8).....	98

Table 4.6: Autocorrelation for the variable “Capesize Ore Voyage Rates Tubarao/Rotterdam. ....	103
Table 4.7: Partial autocorrelations of the capesize ore voyage rates. ....	104
Table 4.8: Pearson correlations for the 7 timeseries. ....	108
Table 4.9: Excel columns with eight time series. The eighth column represents the desirable forecasting time series with 3, 6, 12 and 18 months delay. ....	114
Table 4.10: Database with the three different types of data. ....	115
Table 6.1a: Training Data Performance for all four ANNs. ....	150
Table 6.1b: Cross-Validation Data Performance for all four ANNs. ....	150
Table 6.1c: Unknown Data Performance for all four ANNs. ....	151
Table 6.1d: March 2006 Performance for all four ANNs. ....	152
Table 6.1e: March 2006 Performance for all four ANNs. ....	152
Table 6.2: Forecast of March and April 2006 from the four ANNs. ....	165
Table 6.3: Capesize Ore Voyage Rates Forecasts from all four ANNs. ....	166

# Chapter 1

## Introduction

Shipping is undoubtedly one of the most fascinating industries in the world, and a number of arguments can be put forward to support this statement. To begin with, shipping was the first truly global industry long before globalization became a buzzword. The fact that ships and shipping companies operate in a very international environment with owners, charterers, assets, crew and financial institutions sometimes thousands of miles away, gives someone the opportunity to get acquainted with other cultures through extensive communication and traveling as well as establish contacts all over the world.

In addition, the complexity of the industry and its dependence on world economic conditions require a wealth of knowledge and skills in order to cope with day-to-day operations and events that keep routine away. This complexity and skills requirement make ship-owners some of the most respected entrepreneurs in the world that can flourish in almost anything they do besides shipping. Names such as Onassis, Pao and McKinsey-Maersk are testimony to this statement, while the fact that Easyjet, the airline that revolutionized the industry through budget air travel belongs to a ship-owner, Hadjioannou, speaks for itself.

Finally, another major point is the industry's highly volatile market. Shipping is full of stories about fortunes built and/or lost overnight. At the heart of this volatility is

the bulk shipping market. By being one of the best examples of perfect competition thereby distinguishing it from liner shipping with its oligopolistic structure and high barriers to entry, bulk shipping provides to nearly anyone, poor or wealthy, well-educated or not, from different backgrounds to strike it rich or even super rich. All this person needs is one or more of the following important characteristics:

- Ambition.
- Determination
- Foresight
- Patience
- Good public relations skills
- A lot of luck, as in any efficient market full of inefficiencies and,
- Enough cash reserves especially during bad times.

## **1.1 Aims and Objectives**

*The aim of this study is to attempt to forecast the Capesize Ore Voyage Rates from Tubarao to Rotterdam for a 145,000 DWT Bulk carrier and especially the spot rate in \$ per Ton for this specific route. The study will measure and highlight the reactions of Artificial Neural networks in various modifications of their structure like the use of genetic algorithms and advanced learning algorithms.*

The objective of this study is to extract valuable conclusions about the use of Artificial Neural Networks as an econometrical tool of forecasting shipping figures. In addition, it will present an extensive outlook of the Bulk carrier market and it will provide data and information of the Iron Ore market.

## 1.2 Perspective

Artificial Neural Networks is a cognitive subject which is still in a developing stage. The statistical forecast is a small application of ANN's capabilities, which have driven the interest of economists and researchers. "*London Business School*" and the "*University College London*" have created a research center, the "*Neuro-Forecasting Center*", with a sole objective to develop decision-making models for economic purposes, based on Artificial Neural Networks.

In addition, Artificial Neural Networks are constantly enriched with new innovative and progressive ideas. This can be confirmed by searching the internet or recent publications, where a lot of new ideas and proposals from researchers all over the world appear regularly. These people have constructed a network faster than any one existed, more capable to generalize, more reliable and generally speaking with the all qualifications that can econometrics bring forth. It is therefore an area open to experiments, trials and efforts for specialized or generalized applications aiming at the accomplishment of a procedure more rapidly or more effectively. For these reasons, the current study is not only a model to forecast shipping figures, but also an attempt to further foster research and experiment on a science which appears prospective and leads the future of knowing what might appear tomorrow.

### 1.3 Organization of the study

The complementation of the aims and objectives of this study as described earlier takes place through seven chapters. Here we present a brief outline of the content of each chapter:

**Chapter 1** starts with a brief review of the shipping industry and evaluates the characteristics of a ship-owner. The chapter concludes by defining the aim and the objectives of the thesis and its structure.

**Chapter 2** refers to the economics of the shipping market and gives further attention to the dry bulk market. As a theoretical part of the study, there is an extensive analysis of the shipping market cycles, the four shipping markets, the freight rate mechanism and the five major bulk commodities. The chapter wraps up with an overview of the current situation in the Dry Bulk market and the Iron Ore industry.

**Chapter 3** contains information about Artificial Neural Networks generally, starting with a brief historical analysis. We present the resemblances of ANNs with the Biological Neural Networks and we give various definitions. There is also an extensive report of all the required mathematical equations, and the chapter ends with a presentation of the neural network that will be used in the current study.

**Chapter 4** refers to the construction of the database and its analysis. It is constituted by two basic parts. In the first one, the process of data collection and data evaluation with the process of cross-correlation of the time series is described. In the second part, we provide the all the details of the construction of the database that will be used.

In **Chapter 5** the construction of the Artificial Neural Network and the selection of the parameters, like genetic algorithms, that will be used are described. Our objective is to create a methodology that could be followed in every forecast we will perform. At the end of this chapter there is an analytical explanation of the elements of the neural network that are used and the six values that are used to measure the performance of the network.

In **Chapter 6** the results from the four neural networks are presented. All the forecasts are presented in figures and tables and we evaluate any misleading values. Our forecasts demonstrate the Capesize Ore Voyage Rates from Tubarao to Rotterdam even for a period of 18 months from now.

Finally **Chapter 7** summarizes the finding of this study as well as the conclusions we have drawn. Finally it presents some of our suggestions for future work on the field of forecasting shipping figures using Artificial Neural Networks.

## Chapter 2

### Economics of the Shipping market - Dry Bulk market

#### 2.1 Shipping cycles and shipping risk

The market cycle pervades the shipping industry. As one ship-owner put it "*When I wake up in the morning and freight rates are high I feel good. When they are low I feel bad.*" Just as the weather dominates the lives of seafarers, so the waves of the shipping cycle ripple through the financial lives of ship-owners.

Considering the sums of money involved, it is hardly surprising that the shipping cycle is so prominent. If we take the transport of grain from the US Gulf to Rotterdam as an example, a Panamax bulk carrier of 65,000 dwt trading on the spot market could have earned, after operating expenses, about \$1 million in 1986, \$3.5 million in 1989, \$1.5 million in 1992 and \$2.5 million in 1995. The ship itself, a five-year-old Panamax, would have cost \$6 million in 1986 and \$22 million in 1989. Yet in 1994 it was still worth \$22 million. In such a volatile environment the timing of decisions about buying, selling and chartering ships are crucial.

This is not just a problem for ship-owners. Businesses with cargo to transport face the same risks. The cost of transporting about half a mt of grain from the US Gulf to Japan increased from \$5.2 million in 1986 to \$12.7 million in 1989, a very substantial increase.

An important first step in understanding the shipping cycle is to recognize that it is there for a purpose. Cycles play a central part in the economics of the shipping industry by managing the risk of shipping investment in a business where there is great uncertainty about the future.

The whole process starts from the question: "*Who takes the shipping risk?*" A merchant ship is a large and expensive item of capital equipment. In a world where the volume of trade is constantly changing, someone has to decide when to order new ships and when to scrap old ones. If ships are not built but trade grows, eventually business will grind to a halt. Oil companies could not ship their oil, steel mills run out of iron ore and manufactured exports would pile up in the factories and ports. The lucky owners of the few available ships would auction them to the highest bidder and make their fortunes. However, if ships are built and trade does not grow, it is a very different story. With no cargo, the expensive ships sit idle while the unfortunate investors watch their investment rust away.

This, in essence, is "*shipping risk*" and it is what the shipping cycle is all about. When the risk is taken by the cargo owner this leads to an "*industrial shipping*" business in which ship-owners are subcontractors and cost minimizers. When the "*shipping risk*" is left to the ship-owner, the business becomes highly speculative. It is the world's biggest poker game, in which the ships are the chips. The analogy with poker is in some ways very appropriate. However, winning at the shipping game, like poker, also depends on probability, strategy, psychology and luck.

### 2.1.1 The stages in the shipping market cycle

Hampton (1991) goes on to argue that market sentiment plays an important part in determining the structure of cycles and that this can help to explain why the market repeatedly over-reacts to the prices signals.

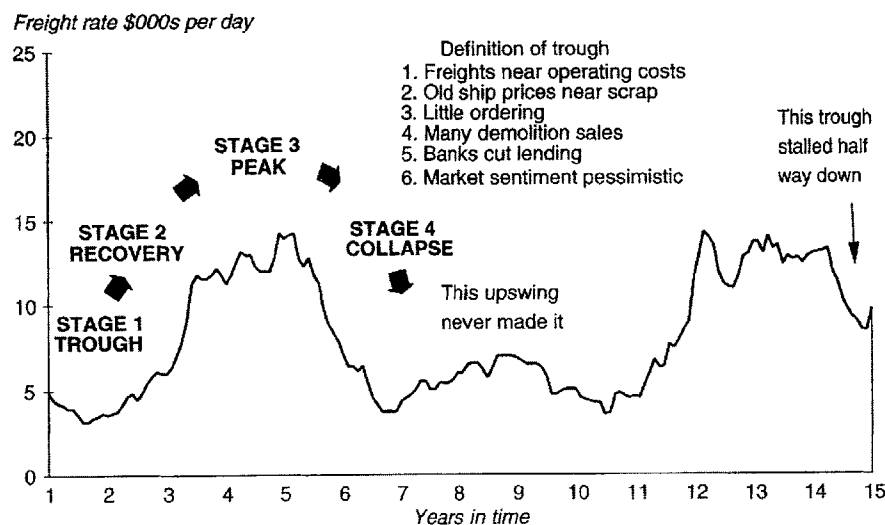
*“In any market including the shipping market, the participants are caught up in a struggle between fear and greed. Because we are human beings, influenced to varying degrees by those around us, the psychology of the crowd feeds upon itself until it reaches an extreme that cannot be sustained. Once the extreme has been reached, too many decisions have been made out of emotions and a blind comfort which comes from the following the crowd rather than objective fact.”*

A shipping cycle is a mechanism devoted to remove imbalances in the supply and demand for ships. If there is too little supply, the market rewards investors with high freight rates until more ships are ordered. When there are too many ships, it squeezes the cashflow until owners give up the struggle and ships are scrapped. Looked at in this way the length of the cycles is incidental. They last as long as is necessary to do the job. It is possible to classify them by length, but this is not very helpful as a forecasting aid. If investors decide that an upturn is due and decide not to scrap their ships, the cycle just lasts longer. Since ship-owners are constantly trying to second-guess the cycle, crowd psychology gives each cycle a distinctive character. The four stages in the shipping market cycle are:

**Stage 1: Trough.** We can identify three characteristics of a trough stage. *First*, there will be evidence of surplus shipping capacity. Ships queue up at loading points and vessels at

sea slow steam to save fuel and delay arrival. *Secondly* freight rates fall to the operating cost of the least efficient ships in the fleet which move into lay-up. *Thirdly*, sustained low freight rates and tight credit create a negative net cashflow which becomes progressively greater. Shipping companies short of cash are forced to sell ships at distress prices, since there are few buyers. The price of old ships falls to the scrap price, leading to active demolition market.

**Stage 2: Recovery.** As supply and demand move towards balance, the first positive sign of a recovery is positive increase in freight rates above operating costs, followed by a fall in laid up tonnage. Market sentiment remains uncertain and unpredictable. Spells of optimism alternate with profound doubts about whether a recovery is really happening (sometimes the pessimists are right, as shown by the false recovery in periods 7 to 9 in *Figure 2.1*). As liquidity improves, second-hand prices rise and sentiment firms.



*Figure 2.1: Stages in a dry market cargo cycle.*

Source: Martin Stopford, 1997

**Stage 3: Peak/Plateau.** When all the surplus has been absorbed, the market enters a phase where supply and demand are in tight balance. Freight rates are high, often two or three times operating costs. The peak may last a few weeks or several years, depending on the balance of supply/demand pressures. Only untradeable ships are laid up; the fleet operates at full speed; owners become very liquid; banks are keen to lend; the press reports the prosperous shipping business; there are public flotations of shipping companies. Secondhand prices move above “book value” and prompt modern ships may sell for more than the new building price. The shipbuilding order book expands, slowly at first, then more rapidly.

**Stage 4: Collapse.** When supply overtakes demand the market moves into the collapse phase. Although the downturn is generally caused by fundamental factors such as the business cycle, the clearing of port congestion and the delivery of vessels ordered at the top of the market, all of which take time, sentiment can accelerate the collapse into a few weeks. Spot ships build up in key ports. Freight rates fall, ships reduce operating speed and the least attractive vessels have to wait for cargo. Liquidity remains high. Sentiment is confused, changing with each rally in rates.

Therefore, the shipping cycle is a mechanism which co-ordinates supply and demand in the shipping market. It is a financial switchbox which regulates investment. As shown in *Figure 2.1* a complete cycle has four stages. A market trough (*stage 1*) is followed by a recovery (*stage 2*), leading to a market peak (*stage 3*), followed by a collapse (*stage 4*). The cycles are “*episodic*”, with no firm rules about the timing of each

stage. Regularity is not a necessary part of the process.

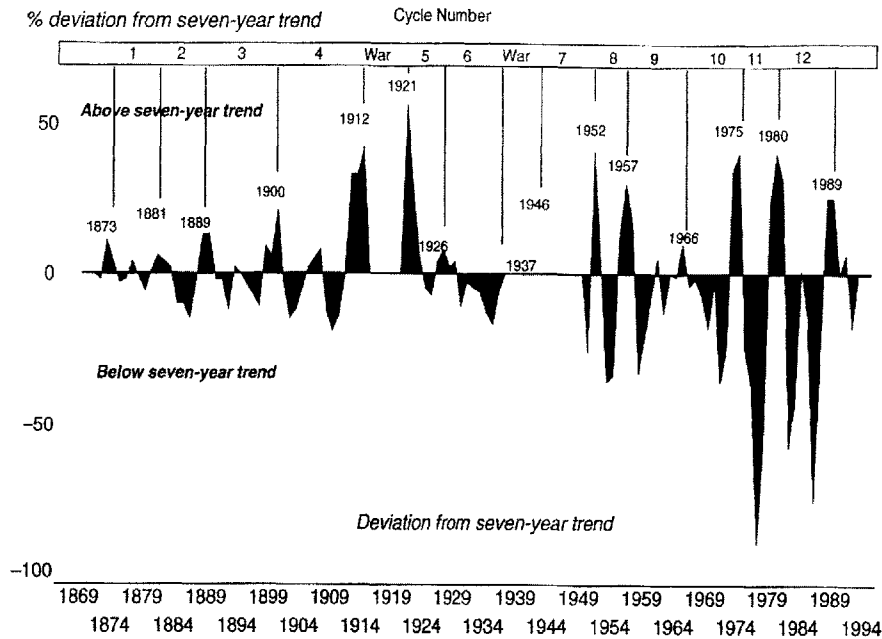
Furthermore, there is no simple formula for predicting the “*shape*” of the next cycle. Recoveries can stall half-way and slump back into recession. Market collapses may be reversed before they reach the trough. Troughs may last six months or six years. Peaks may last a month or a year. Sometimes the market gets stuck in the middle ground between trough and recession.

In short, shipping cycles, like ship-owners, are unique. In each “*cycle*” supply lurches after demand like a drunk walking a line that he cannot see very clearly, governed by the market fundamentals.

### **2.1.2 The frequency of shipping cycles**

Between 1869, the beginning of modern shipping, and 1994 there were twelve dry cargo freight cycles. If we represent each cycle as the deviation from a seven year moving average in *Figure 2.2* any thought of regularity is immediately dispelled. We are clearly dealing with a sequence of random fluctuations whose only common features are a beginning, middle and an end.

However a word of caution is needed. Because cycles are so irregular, identifying them is a matter of judgment. Some cycles are clearly defined, but others leave room for doubt. *Does a minor improvement in a single year, as happened in 1877, 1894 or 1986, count as a cycle?* There are five of these “*mini-cycles*” where the freight rates moved slightly above trend. The definition of the twelve cycles was arrived at by checking the timing of statistical peaks and troughs shown in the freight rate statistics against comments in contemporary market reports.

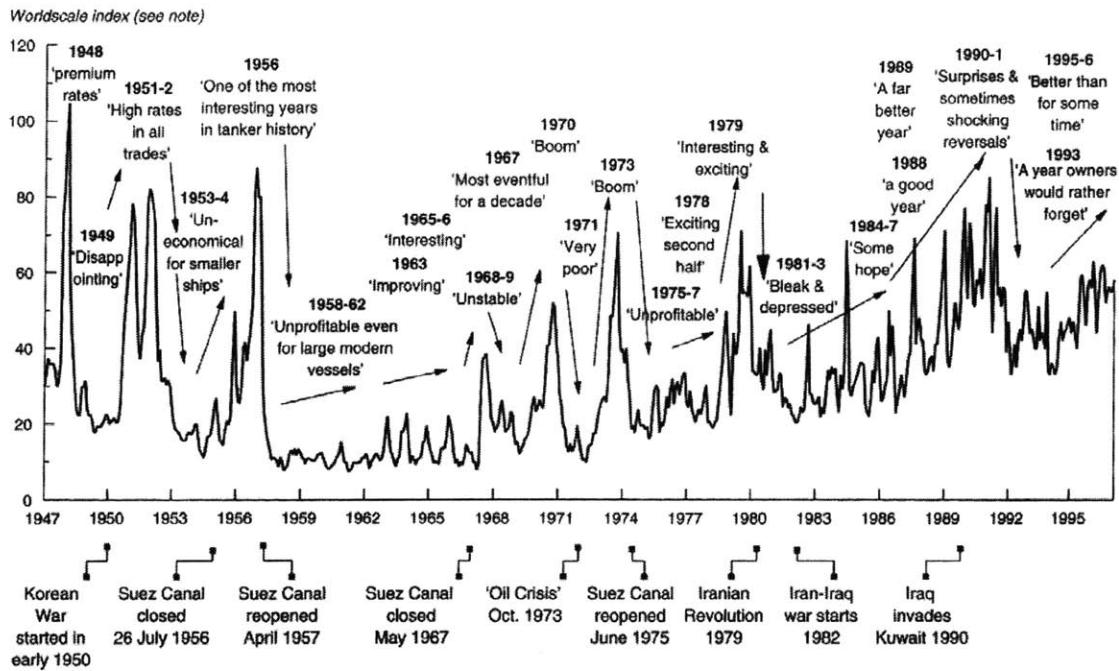


**Figure 2.2:** Twelve dry freight cycles 1869-1995.

Source: Martin Stopford, 1997

The 116-year period is divided by two world wars, the first from 1914-18 and the second from 1940-45. In both cases government intervention in the shipping industry thoroughly disrupted the market mechanism, so the period during and immediately after each war is excluded from the analysis. The wars provide a convenient breaking point for subdividing the analysis of freight rates into three time periods, 1869-1914, 1920-38, and 1945-89. Each period has its own character which pervades the cyclical process.

In the last fifty-year period, following the Second World War, there were five freight market cycles of about 7.4 years each. For the dry cargo vessels, the cycles are clearly defined, in *Figure 2.3*. Since freight rates do not tell the whole story, the graph is annotated to show the terms in which shipbrokers were describing the market at each point.



*Figure 2.3: Dry cargo time charter rates*

*Source: Martin Stopford, 1997*

## 2.2 The four shipping markets

In shipping there are four shipping markets trading in different commodities. *The freight market trades sea transport, the sale and purchase market trades second-hand ships, the new building market trades new ships and the demolition market deals in scrap ships.* Beyond this there is no formal structure. While this analysis provides guidance on how the markets operate, we are not dealing with immutable laws. The fact that market traders have behaved in a particular way in the past is no guarantee that they will do so in future. Because markets consist of people going about their business, the best commercial opportunities often arise when the market behaves inconsistently. For example, ordering ships at the top of the market cycle is usually bad business, but if for some reason few ships are ordered, the rule will not apply. Commercial judgments must be based on an

understanding of market dynamics, not economic principles taken out of context.

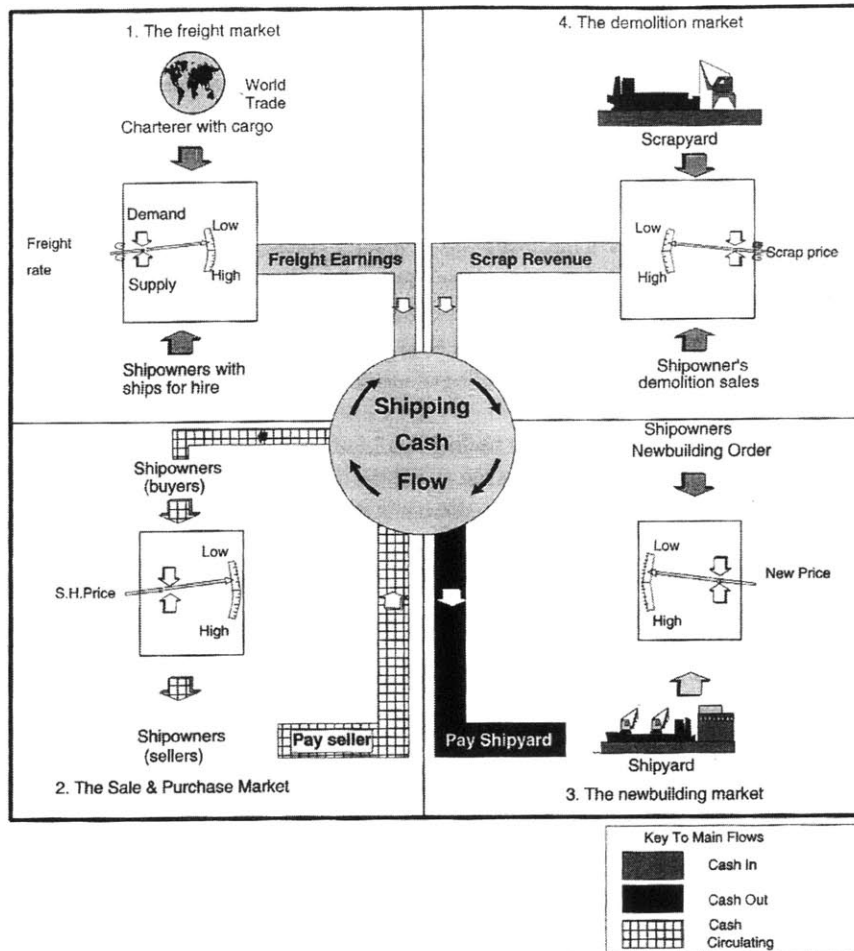
Because the same ship-owners are trading in all four shipping markets their activities are closely correlated. When freight rates rise or fall the changing sentiment ripples through into the sale and purchase market and from there into the new building market. The markets are also linked by cash. The relationship is shown graphically in *Figure 2.4*, cashflows back and forth between the industry's bank account (represented by the circle) and the four shipping markets (represented by the squares). The cashflow into the shipping companies' bank account is shown by the light shaded bars, while the black bars show outflows. The hatched bars indicate cash which changes hands from one ship-owner to another, but does not change the cash balance of the industry as a whole.

The main cash inflow is freight revenue. This goes up and down with freight rates and is the primary mechanism driving the activities of shipping investors. The other cash inflow comes from the demolition market. Old or obsolete vessels sold to scrap dealers provide a useful source of cash, especially during recessions. The sale and purchase (S&P) market has a more subtle role. Investing in a second-hand ship involves a transaction between a ship-owner and an investor. Because the investor is usually another ship-owner, money changes hands, but the transaction does not affect the amount of cash held by the industry. The sale of a tanker for \$20 million only transfers \$20 million cash from one shipping bank account to another, leaving the aggregate cash balance unchanged. In this sense the sale and purchase market is a zero sum game. For every winner there is a loser. The only real source of wealth is trading cargo in the freight market. In the case of the new building market the cashflow (shown in black) is in the opposite direction. Cash spent on new ships flows out of the shipping industry because

the shipyard uses it to pay for materials, labor and profit.

Waves of cashflowing between the four markets drive the shipping market cycle. At the beginning of the cycle freight rates rise and cash starts to pour in, allowing ship-owners to pay higher prices for second-hand ships. As prices are bid up, investors turn to the new building market which now looks better value. With the confidence created by bulging wallets they order many new ships. A couple of years later the ships arrive on the market and the whole process goes into reverse. Falling freight rates squeeze the cash inflow just as investors start paying for their new buildings. Financially weak owners who cannot meet their day-to-day obligations are forced to sell ships on the second-hand market. This is the point at which the assets play market start for those shipowners with strong balance sheets. In extreme circumstances like 1932 or 1986 modern ships change hands at bargaining prices. For older ships, there will be no offers from trading buyers, so hard pressed owners are obliged to sell for demolition. As more ships are scrapped, the supply falls, freight rates are bid up and the whole process starts again.

The whole commercial process is controlled and coordinated by cashflows between markets. Cash is the "*stick and carrot*" which the market uses to drive activity in the required direction. Whether they like it or not, shipowners are part of a process which controls the price of the ships they trade and the revenues they earn.



**Figure 2.4:** *The four markets which control shipping.*

*Source: Martin Stopford, 1997*

### 2.3 The freight rate mechanism

The *freight market* is the adjustment mechanism linking supply and demand. The way it operates is simple enough. Shipowners and shippers negotiate to establish a freight rate which reflects the balance of ships and cargoes available in the market. If there are too many ships, the freight rate is low while if there are too few ships, it will be high. Once this freight rate is established, shippers and shipowners adjust to it and eventually this brings supply and demand into balance.

### 2.3.1 The supply and demand functions

The supply function for an individual ship, shown in *Figure 2.5a* is a hockey stick shaped curve describing the amount of transport the owner provides at each level of freight rates. The ship in this example is a *280,000 dwt VLCC*. When the freight rate falls below \$155 per mtm the owner puts it into lay-up, offering no transport. As freight rates rise past \$155 per mtm he breaks lay-up but, to save fuel, steams at the lowest viable speed of 11 knots per hour. If he trades loaded with cargo at this speed for 137 days per annum, he will supply 10.1 btm of transport in a year (i.e.  $11 \times 24 \times 137 \times 280,000$ ). At higher freight rates he speeds up until at about \$220 per mtm the ship is at full speed of 15 knots and supplying 13.8 btm of sea transport per year (a lot of transport for just one ship!). Thus by increasing freight rates the market has obtained an extra 38 per cent supply.

Economic theory can help to define the shape of the supply curve. Provided the market is perfectly competitive, the shipowner maximizes his profit by operating his ship at the speed at which marginal cost (i.e. the cost of providing an additional ton-mile of transport) equals the freight rate. The relationship between speed and freight rates can be defined as follows:

$$s = \sqrt{\frac{R}{3pkd}}$$

Where  $s$  = optimum speed in miles per day

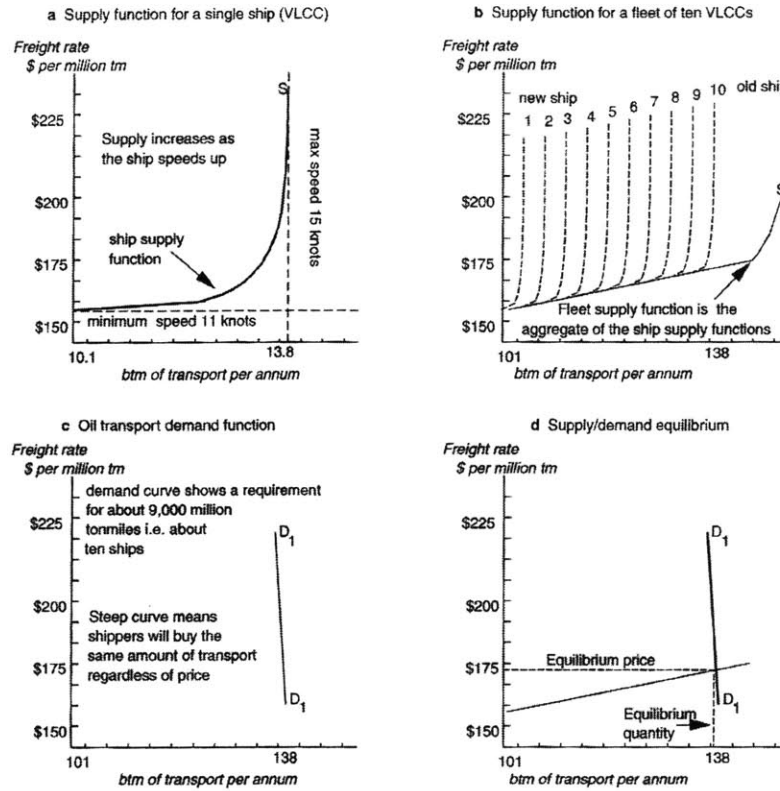
$R$  = voyage freight rate

$p$  = price of fuel

$k$  = the ship's fuel constant

$d$  = distance

This equation defines the shape of the supply curve. In addition to freight rates the optimum speed depends on the price of fuel, the efficiency of the ship and the length of the voyage.



**Figure 2.5:** Shipping supply and demand functions.

Source: Martin Stopford, 1997

In reality the supply function is more complex than the simple speed-freight rates relationship described in the previous paragraphs. Speed is not the only way supply responds to freight rates. The owner may take advantage of a spell of low freight rates to put his ship into dry dock, or fix a short-term storage contract. At higher rates he may decide to ballast back to the Arabian Gulf through the shorter Suez Canal route rather than taking the longer “free passage” round the Cape. All of these decisions affect supply. Similarly freight rates are not the only way the market adjusts shipowners’

revenue. During periods of surplus ships have to wait for cargo or accept small cargo parcels. This reduces the operating revenue in just the same way as a fall in freight rates, a factor often forgotten by owners and bankers doing cashflow forecasts on old ships. They may predict freight rates correctly but end up with an embarrassing cash deficit due to waiting time and part cargoes.

The next step is to show how the market adjusts the supply provided by a *fleet of ships*. To illustrate this process the supply function for a fleet of 10 VLCCs is shown in *Figure 2.5b*. The *fleet supply curve (S)* is built up from the supply curves of individual ships of varying age and efficiency. In this example the age distribution of the fleet ranges from two years old to twenty years old in intervals of two years. Ship 1 (newest ship) has low daily operating costs and its lay-up point is \$155 per mtm. Ship 10 (oldest ship) has high operating costs and its lay-up point is \$165 per mtm.

The fleet supply function works by moving ships in and out of service in response to freight rates. If freight rates fall below the operating costs of ship 10, it goes into lay-up and supply is reduced by one ship. Ship 9 breaks even and the other eight ships make a margin over their fixed expenses, depending on how efficient they are. If shippers only need five ships they can drop their offer to \$160 per mtm, the lay-up point of ship 5. In this way supply responds to movements in freight rates. Over a longer period the supply can be increased by building new more efficient ships and reduced by scrapping old ones.

The slope of the short term supply curve depends on three factors which determine the lay-up cost of the marginal ship. First, old ships generally have higher operating costs so the lay-up point will occur at a higher freight rate. Second, bigger ships have lower transport costs per ton of cargo than small ships, so if big and small ships are

competing for the same cargo, the bigger ship will have a lower lay-up point and will generally drive the smaller ships into lay-up during recessions. If the size of ships has been increasing over time, as has happened for most of the last century, the size and age will be correlated and there will be quite a steep slope to the supply curve which becomes very apparent during recessions. Third, the relationship between speed and freight rates.

The demand function shows how charterers adjust to changes in price. The *demand curve (D)* in *Figure 2.5c* is almost vertical. This is mainly supposition, but there are several reasons why this shape is likely for most bulk commodities. The most convincing is the lack of any competing transport mode. Shippers need the cargo and, until they have time to make alternative arrangements, must ship it regardless of cost. Conversely cheap rates will not tempt shippers to take an extra ship. The fact that freight generally accounts for only a small proportion of material costs, reinforces this argument.

### **2.3.2 Equilibrium and the importance of time**

The supply and demand curves intersect at the equilibrium price. At this point buyers and sellers have found a mutually acceptable price. In *Figure 2.5d* the equilibrium price is \$170 per mtm. At this price buyers are willing to hire ten ships and owners are prepared to make ten ships available. The equation balances.

However that is not the end of the story. If our aim is to understand why freight rates behave the way they do, it is just the beginning. We must be precise about time frame. In the real world the price at which buyers and sellers are prepared to trade depends on how much time they have to adjust their positions. There are three time periods to consider; *the momentary equilibrium* when the deal must be done immediately;

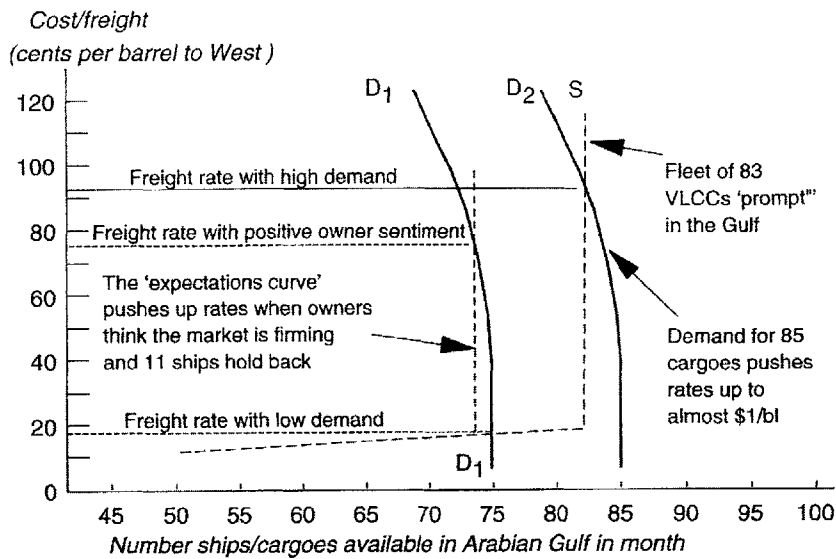
*the short run*, when there is time to adjust supply by short-term measures such as lay-up, reactivation, combined carriers switching markets or operating ships at a faster speed; and there is *the long run*, when shipowners have time to take delivery of new ships and shippers have time to rearrange their supply sources.

### 2.3.3 Momentary equilibrium

*Momentary equilibrium* describes the freight rate negotiated for “prompt” ships and cargoes. The ships are ready to load; the cargoes are awaiting transport and a deal must be done. The shipowner is in the same position as the farmer when he arrives at market with his pig. Within this time frame the shipping market is highly fragmented, falling into the regions so familiar in brokers’ reports - the Arabian Gulf, the Caribbean, USAC (United States Atlantic Coast), etc. Local shortages and surpluses build up, creating temporary peaks and troughs which show up as spikes on the freight chart. This is the market owners are constantly trying to anticipate when selecting their next cargo, or deciding whether to risk a ballast voyage to a better loading point.

Once these decisions are taken and the ship is in position, the options are very limited. The owner can “fix” at the rate on offer, or sit and lose money. Charterers with cargoes face the same choice. The two parties negotiate to find a price at which supply equals demand. *Figure 2.6* illustrates how this works out in practice.

Suppose there are about 75 cargoes on offer during the month. The *demand curve*, marked  $D_i$ , intercepts the horizontal axis at 75 cargoes, but as freight rises it curves to the left because at very high freight rates some cargoes will be withdrawn.



**Figure 2.6:** Momentary equilibrium in the VLCC market.

Source: Martin Stopford, 1997

The supply curve S shows there are 83 ships available to load, so there are more ships than cargoes. Since the alternative to fixing is earning nothing, rates fall to operating costs which equates to 18 cents a barrel, shown by the intersection of S and D<sub>1</sub>. If the number of cargoes increases to 85 (D<sub>2</sub>), there are more cargoes than ships. Charterers bid desperately to find a ship and freight rate shoots up to almost \$1 per barrel. A swing of 10 cargoes is quite common, but the effect on rates is dramatic.

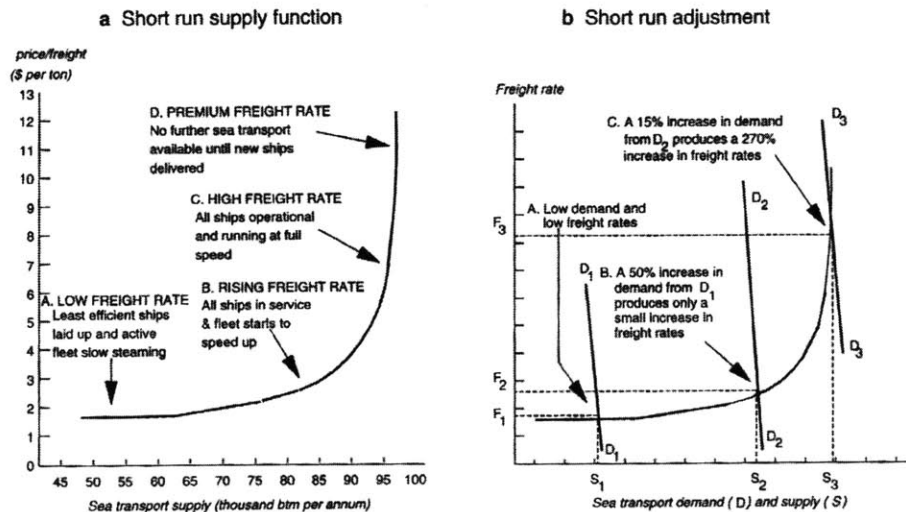
In this very short-term situation market sentiment can make rates very volatile. If there are more ships than cargoes, but owners believe that rates are rising, they may decide to wait (sometimes owners attempt to hide their ships from charterers by reporting the presence of only one ship in their fleet, or waiting outside the loading area). Suddenly there are more cargoes than ships and rates rise, at which point the reticent owners enter the market and fix at “last done”. This is shown by the “expectation curve” in Figure 2.6. If the surplus of ships persists, the ships which hold back may be unable to fix at all

and rates quickly collapse.

### 2.3.4 The short run equilibrium

In the “short run” there is more time for owners and charterers to respond to price changes by moving ships in and out of lay-up, so the analysis is a little different.

The short run supply curve shown in *Figure 2.7a*, plots, for a given size of fleet, the ton miles of transport available at each level of freight rates. The transport supply is measured in thousand btm per annum and the freight rate in dollars per thousand ton miles of cargo transported.



Source: Martin Stopford, 1997

**Figure 2.7:** Short run equilibrium.

At *point A*, the supply offered is only 5,000 btm per annum because the least efficient ships are laid up; at *point B*, all ships are back in operation and the supply has risen to about 8,500 btm per annum; at *point C*, the fleet is at maximum speed and the whole fleet is at sea; finally, at *point D*, no further supply is obtained by increasing freight rates and the supply curve becomes almost vertical. Very high freight rates may

tempt out a few remaining unutilized ships. For example, during the 1956 boom “A number of vessels half a century old and barely seaworthy obtained freights of up to five times the rate obtained a year earlier.” (Source: M. Hampton “Long & short shipping Cycles: The Rhythms and Psychology of Shipping Markets”. 1991)

If we now bring the short run demand curve into the picture, we can explain how freight rates are determined. The market settles at the freight rate at which supply equals demand. Consider the three different equilibrium points marked *A*, *B*, and *C* in *Figure 2.7b*. At point *A* demand is low and the freight rate settles at *point F<sub>1</sub>*. A major increase in demand to point *B* only pushes the freight rate up slightly because ships immediately come out of lay up to meet increasing demand. However a small increase in demand to point *C* is sufficient to treble the level of freight rates because the market rate is now set by the oldest and least efficient ships which need very high freight rates to tempt them into service. Finally, with no more ships available charterers bid against each other for the available capacity. Depending on how badly they need transport, rates can go to any level. However this is an unstable situation. Shippers look for cheaper supply sources and the high freight rates almost always trigger frenzied investment activity by owners and shippers.

### **2.3.5 The long run**

Finally, we must consider *the long run* during which the size of the fleet can be adjusted by ordering new ships and scrapping old ones. The longer-term adjustment mechanism balances supply and demand through the three other markets, the sale and purchase market, the new building market and the demolition market. As freight rates fall

during a recession, the profitability of ships and, consequently, their second-hand value also falls. Eventually the price of the least efficient ships falls to the scrap price. Ships are scrapped, removing them permanently from the market and reducing the surplus. Falling second-hand prices also make new uses of the surplus tonnage financially viable; the use of supertankers for oil storage or bulk carriers as trans-shipment facilities is examples. In these ways the price mechanism gradually reduces the supply of ships to the market. Conversely, when a shortage of ships pushes up freight rates this works through to the sale and purchase market. Shipowners are keen to add to their fleets and, because there is a shortage of ships, shippers may decide to expand their own shipping operations. With more buyers than sellers, second-hand prices rise until used ships become more expensive than new buildings. Frustrated shipowners turn to the new building market and the order book expands rapidly. Two or three years later the fleet starts to grow.

To illustrate this process we can take the example of the adjustment of the tanker market over the period 1980 and 1991. *Figure 2.8* shows four charts, illustrating the position of the supply demand chart in a different year, 1980 (chart a), 1985 (chart b), 1991 (chart c) and 1992 (chart d). The freight rate is shown on the vertical axis measured in \$000s per day and as an indicator of transport supply the tanker fleet is shown on the horizontal axis, measured in m.dwt. Neither of these units of measurement is strictly correct but they are adequate for an illustration. In the centre of *Figure 2.8* is a freight chart (chart e) which shows the level of freight rates in each of the four years. Our aim is to explain how the supply and demand curves moved between the 4 years. In 1980 (chart a) freight rates were moderately high at \$15,000 per day, with the demand curve intersecting the 'kink' of the supply curve. By 1985 (chart b) the supply curve has moved

to the left as heavy scrapping reduced the tanker fleet from 320 m.dwt to 251 m.dwt, but demand had fallen even more to below 150 m.dwt due to the collapse in the crude oil trade after the oil price rises in 1979. This left 60 m.dwt of tankers laid up, extensive slow steaming, and the demand curve intersecting the supply curve way down its span at  $D_{85}$ . Freight rates averaged about \$7,000 per day, close to operating costs.

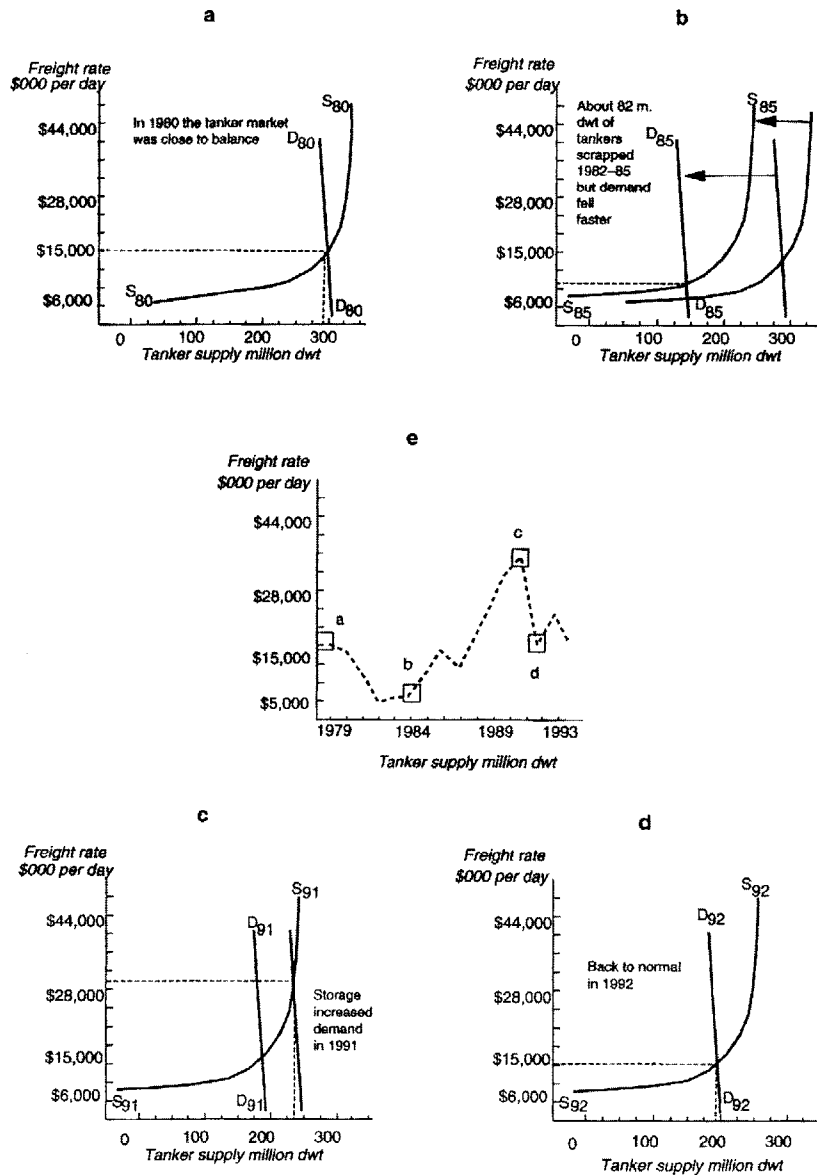


Figure 2.8: Long-term adjustments of supply and demand 1980-92.

Source: Martin Stopford, 1997

Between 1985 and 1991 (chart c), despite heavy scrapping, the tanker fleet fell by only 7 m.dwt, due to increased new building in the late 1980s. As a result the supply curve moved very slightly to the left to  $S_{91}$ , but a growing oil trade increased demand by 30 per cent to  $D_{91}$ , suggesting an equilibrium freight rate of about \$15,000 per day. However, in 1991 another factor intervened. After the invasion of Kuwait in August 1990 oil traders used tankers as temporary storage, moving the demand curve temporarily to the right, shown by the dotted line in *Figure 2.8c*. Freight rates increased to \$29,000 per day. Then in 1992 supply increased due to heavy deliveries and the demand curve moved back to its “normal” position as the temporary storage marker disappeared. This was enough to drive freight rates down to \$15,000 per day (chart d).

It is the combination of volatile demand and a significant time-lag before supply adjusts to demand which creates the framework for shipping market cycles. Shipowners tend to base investment on the current state of the market - they order more ships when freight rates are high and fewer when freight rates are low. The delay in delivering these ships means, however, that demand may have changed by the time the ships are delivered so any cyclical tendency is amplified. It takes two or three years for new orders to be delivered, two or three years for scrapping to catch up, and two or three years for the market to build up a head of steam for the next round of ordering. Of course it never happens exactly this way, but it makes a neat generalization.

## 2.4 The five major dry bulks

If oil is the energy of modern industrial society, the *five major bulks* are the building blocks from which it is constructed, Iron ore and coal are the raw materials of steelmaking and steel is the principal material used in the construction of industrial and domestic buildings, motor cars, merchant ships, machinery and the great majority of industry products. The staple foods of the modern industrial society are bread and meat, both of which require large quantities of grain - for baking and as the raw material of modern factory farming for the production of meat. Bauxite and alumina are the raw materials of aluminium making, the second most important structural metal in modern industrial society, while phosphate rock is the principal bulk fertilizer used in crop production. It follows that in discussing the five major bulk trades we are concerned with the whole material development of the world economy that uses these materials.

Because of their volume, the five major bulk trades are the driving force behind the dry bulk carrier market. In 2005 the trade totaled 1.7 bt, accounting for more than one-quarter of total seaborne cargo, and in terms of tonnage about half the crude oil trade. The tonnage of cargo in each commodity and its historic growth are shown in *Table 2.1*.

<i>Commodity</i>	<i>1965</i>	<i>1975</i>	<i>1985</i>	<i>1995</i>	<i>2005</i>
Iron Ore	152	292	321	402	651
Coal	59	127	272	403	682
Grain	70	137	181	216	271
Bauxite and Alumina	21	41	40	52	68
Phosphate rock	26	38	43	30	30
Total	328	635	857	1,103	1,702

*Table 2.1: The five "major" bulk commodities shipped by sea (mt).*

*Source: Fearnleys World Bulk Trade*

It is immediately apparent from the historical trade statistics that each of these five trades followed its own distinctive growth pattern during the four decades 1965-

2005. Iron ore grew rapidly until the mid-1970s, after which the trade has grown slowly. Coal followed a similar pattern until 1975, but continued to grow. It more than doubled between 1975 and 1985 and tripled in volume between 1985 and 2005. Grain grew fairly steadily until 1985, and then stagnated, though with substantial year-to-year fluctuations. Bauxite grew rapidly during the 1960s, but stagnated during the 1970s, then revived, increasing by 25 per cent between 1985 and 1995. Phosphate rock was unique in showing almost no growth over the 30 years. It grew, and then declined again.

Such major changes in the pattern of growth of the main dry bulk commodity trades are clearly of great importance to the shipping industry. There is no simple pattern; each commodity has its own distinctive industrial characteristics, growth trends and impact upon the dry bulk shipping industry.

#### **2.4.1 The seaborne iron ore trade**

Iron ore is the largest of the five major bulk commodity trades and the principal raw material of the steel industry. Like crude oil, the iron ore trade is determined by the location of the processing plant in relation to raw material supplies. During the Industrial Revolution, steel plants were located on sites close to major sources of raw materials, notably iron ore, coal and limestone. Access to these materials was a major concern in the economics of the steel industry. However, as transport technology developed it became clear that the distance over which the materials were shipped was less important than the freight-rate structure, the transport service and the quality of the raw materials.

By the early twentieth century, developments in bulk shipping technology meant that steel plants located near to raw material supplies no longer had a cost advantage,

particularly when land transport was required. For example, in the United Kingdom, Northamptonshire ores were trebled in cost by transport to Middleborough, whereas the cost of Lorraine ores in the Ruhr was hardly more than doubled for a much longer journey owing to the availability of water transport. As the demand for steel expanded in the twentieth century, the industry started to gravitate towards coastal steel plants, which could import raw materials at minimum cost by using a carefully planned integrated bulk shipping operation. This had the advantage that, with the resources of the world accessible by sea, it was possible to find higher-quality raw materials than were available locally, particularly in the traditional steelmaking areas of Western Europe where the better-quality ores were already depleted.

The prototype for the modern integrated dry bulk transport operation was probably the steel plant built by Bethlehem Steel at Sparrow's Point, Baltimore, in the early 1920s. This plant was designed specifically to import iron ore by sea from Cruz Grande in Chile, taking advantage of the newly-opened Panama Canal. To service the trade, a contract was placed with the Brostrom group, which ordered two ore carriers of 22,000 dwt. At the time these were two of the world's largest ocean-going cargo ships. Details of the shipping operation are recorded as follows (*Source: Dunn 1973*):

The contract, signed in 1922, called for two ships to carry ore from Chile through the Panama Canal to Bethlehem Steel Company's plant at Sparrow's Point, Baltimore. The ships had no conventional cargo handling gear, and hinged corrugated steel hatch covers. These were the full width of the holds, weighed 8 tons apiece and were clamped down to thick rubber gaskets. The *Sveiand* was delivered on 9th April 1925 and *Americaland* on 29th June and they promptly entered their designed service between

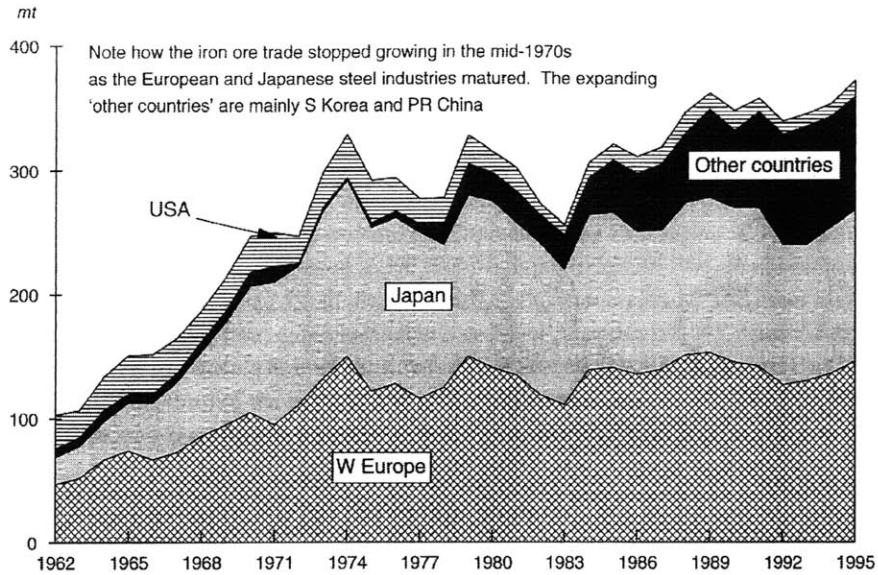
Cruz Grande and Baltimore. It was an exacting schedule and the average time spent at sea each year was 320-330 days. At Cruz Grande the 22,000 tons cargo was normally loaded in two hours, though the record was 48 minutes. Discharging at the other end required about 24 hours. Routine engine maintenance was carried out at sea, one of the two engines being shut down for eight hours per trip. Painting was also carried out while underway.

This pattern of using large, specially designed ships on a shuttle service between the mine and the steel plant became standard practice for coastal steel plants. Although the size of ship used increased during the next 50 years, reaching 120,000 dwt in the 1960s and 300,000 dwt in the 1980s, the basic operational principles remain the same.

The East Coast development of the US steel industry proved something of a false start, and the major portion of US steelmaking continued to be concentrated around the Great Lakes, using locally produced ores supplemented by imports from Canada via the St Lawrence Seaway when the Labrador iron ore fields were developed. As a result, the United States did not figure prominently in the post-war overseas iron ore trade, and the principal importers of iron ore were Western Europe and Japan, as can be seen in *Figure 2.9*.

During the post-war period of industrial expansion, steel demand grew rapidly. In Europe and Japan this growth was met by building modern integrated coastal steel plants using imported raw materials. In Japan there was little choice since there were no domestic reserves of iron ore, but even in Europe where extensive iron ore reserves are available these were of lower quality than the imported variety. For new developments, the shorter land transit leg offered little cost advantage over seaborne transport using

large bulk carriers.



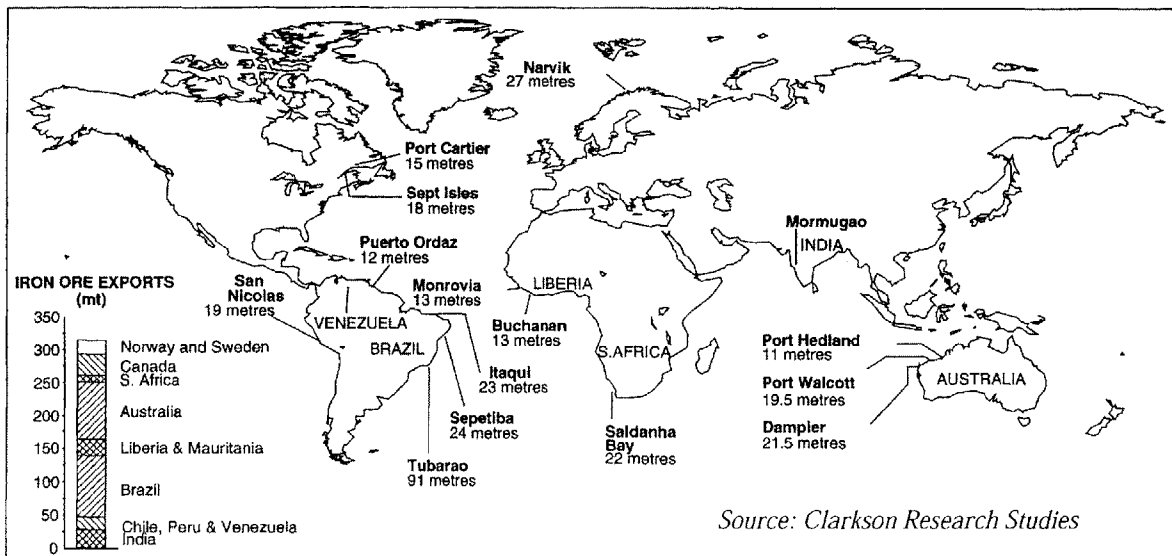
*Figure 2.9: Iron Ore Imports 1962-95.*

*Source: Clarkson Research Studies*

It was the rapid expansion of iron ore imports by the steel industry that underpinned the bulk carrier boom of the 1960s. The Japanese and European steel companies were prepared to offer long time charters to meet the regular raw material requirements of the new coastal steel plants. These charters provided many growing bulk shipping companies with the stable foundation on which to base their fleet development strategy. In the early 1970s, however, the period of growth came to an end. After a decade of steadily expanding ship demand, the steel companies found themselves facing excess capacity and 1974 proved to be a turning point for iron ore imports, as can be seen in *Figure 2.9*. The explanation of this change is that in these two areas steel demand had reached saturation point in the early 1970s: between 1975 and 1995, West European steel fell from 170 mt to 161 mt; during the same period Japanese production was static at 101 mt.

There are many reasons for this radical change of trend, but the most important was that the industries that are intensive users of steel (principally construction, vehicles and shipbuilding) had all reached a plateau in their output. As a result, the growth had been removed from the largest iron ore importers. In their place the newly industrializing countries such as South Korea started to make an impact in the 1990s, a trend that is visible in *Figure 2.9*.

Although we have concentrated on the demand for seaborne imports of iron ore, the trade also depends crucially upon the development of a global network of iron ore supplies, and *Figure 2.10* shows the pattern that developed. Generally at the initiative of the steel companies, iron ore resources were identified across the globe and the necessary capital rose to develop the mines and install the requisite transport infrastructure.



*Figure 2.10: Major Iron Ore exporters and ports 2005.*

*Note: Notes against each port indicate the maximum draught in Ports of the World.*

By far the largest iron ore exporters are Brazil and Australia. The first Brazilian iron ore reserves to be developed were located in the famous Iron Quadrangle of Minas Guerais and exported through the ports of Sepetiba and Tubarao. In 1986, the first

cargoes were exported from Carajas, a major iron ore development in the Para region of Northern Brazil with port facilities at Itaqui geared to 300,000 dwt bulk carriers. In 2005, Brazil exported 225 mt of iron ore, accounting for almost one-third of the iron ore trade. The other major iron exporter is Australia, from mines located in Northwest Australia. In 2005, Australia exported 241 mt of iron ore, mainly through the three ports of Port Hedland, Dampier and Port Walcott. The remaining third of the iron ore trade is supplied from a variety of smaller exporters, of whom the most important are Sweden, South Africa, Liberia, India, Peru and Mauritania.

#### **2.4.2 The transport system for iron ore**

Iron ore transport by sea is one of the great successes of industrial bulk shipping. The iron ore generally has a *stowage factor of 0.3 cubic meters per ton* and is almost always transported in bulk and in full shiploads. Over the past decade there has been great competition between suppliers in the Atlantic and Pacific, leading to increasing distance between source and markets and the employment of the largest ships possible. Generally iron ore ports serve as transfer terminals linking two modes of transport, with storage areas at the port providing a surge capability between the more or less continuous overland movement and the intermittent ocean shipment.

At the mine large earth-moving equipment removes the ore from open pits and transfers it to special trains or trucks that transfer it to port, where large cranes or automatic dumping mechanisms place it in storage areas from which it is transferred by means of gravity or by cranes to the ship. The ship then steams to a port or coastal steel mill where the process is reversed. The entire system is geared to anticipate mill needs

with a continuous flow of the ore from mine to mill.

Although the economies of scale which can be achieved through the use of large bulk vessels were well known in the 1950s, the transition from small vessels to the larger sizes was not immediate. In 1965 80 per cent of all iron ore was carried in vessels below 40,000 dwt, and over half the trade is still carried in ships less than 80,000 dwt. The process of introducing large ships was thus a gradual one, with the size of vessels built for use in the iron ore trade increasing steadily from around 30,000 dwt in the early 1960s to 60,000 dwt in 1965; 100,000 dwt in 1969; and 150,000 dwt plus in the early 1970s and 300,000 dwt in the 1990s. For example the Bergeland delivered in 1991 was a 300,000 dwt vessel designed exclusively for the carriage of iron ore. In fact the size of ship has grown with the volume of trade and the improvements in port facilities, though many small vessels built in previous periods continue to be used.

## 2.5 Dry Bulk Trade Outlook April 2006

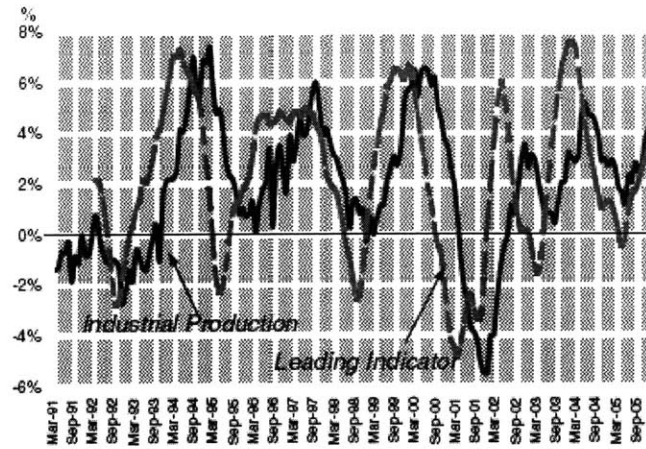


Figure 2.11a: OECD Industrial Prod. & Leading Indicator

Source: Clarkson Research Studies

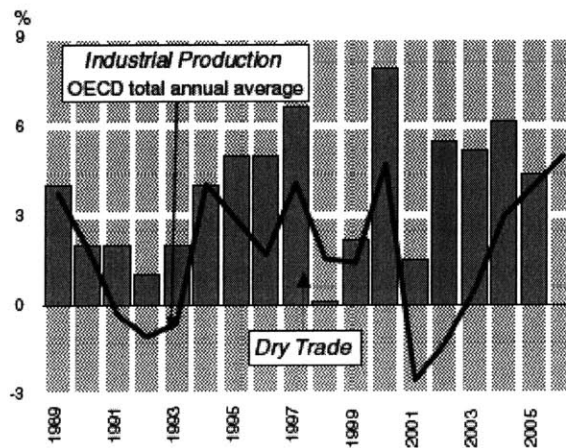


Figure 2.11b: OECD Growth/Dry Trade Growth

Source: Clarkson Research Studies

**Seaborne Trade at a Glance**  
Million Tones

Dry Bulks	Million Tones									Imported trend	
	1998	1999	2000	2001	2002	2003	2004	Est. 2005	F'cast 2006*	Next Year vs. This Year	
Iron Ore	427	402	448	451	481	519	589	651	699	<b>STRONG!</b>	7%
Total Coal	458	466	524	556	579	632	654	682	708	<b>FIRM....</b>	4%
Steam Coal	289	304	350	387	407	453	474	498	512	<b>STEADY</b>	3%
Coking Coal	169	162	174	169	173	179	180	184	196	<b>STRONG!</b>	7%
Grains	226	247	264	260	271	264	274	271	278	<b>STEADY</b>	3%
Wheat/Coarse Grain	187	205	214	207	214	205	214	205	212	<b>FIRM....</b>	3%
Soyabean	40	42	50	54	57	58	60	65	66	<b>STEADY</b>	1%
Bauxite/Alumina	55	54	54	52	55	63	66	68	69	<b>STEADY</b>	1%
Phosphate Rock	31	30	28	29	30	29	30	30	30	<b>STEADY</b>	0%
<b>5 MAJOR BULKS</b>	<b>1,197</b>	<b>1,199</b>	<b>1,318</b>	<b>1,347</b>	<b>1,416</b>	<b>1,507</b>	<b>1,613</b>	<b>1,702</b>	<b>1,784</b>	<b>FIRM...</b>	<b>5%</b>
% change	0%	0%	10%	2%	5%	6%	7%	6%			
Sugar	38	40	37	41	44	45	46	46	48	<b>FIRM....</b>	3%
Agribulks	92	93	90	92	94	96	90	93	95	<b>STEADY</b>	3%
Fertilizer	66	69	70	72	74	75	76	78	79	<b>STEADY</b>	2%
Scrap	50	54	62	68	70	80	93	94	94	<b>STEADY</b>	1%
Cement	45	45	46	46	45	47	60	60	65	<b>STRONG!</b>	8%
Coke	20	18	24	23	23	24	24	25	25	<b>STEADY</b>	1%
Pig Iron	14	13	13	13	12	9	17	17	17	<b>STEADY</b>	0%
Forest Product	156	158	161	164	160	163	166	170	174	<b>STEADY</b>	2%
Steel Product	189	174	184	193	198	205	211	217	218	<b>STEADY</b>	1%
Others	33	34	36	37	37	39	40	42	44	<b>FIRM...</b>	4%
<b>MINOR BULKS</b>	<b>703</b>	<b>697</b>	<b>723</b>	<b>748</b>	<b>756</b>	<b>784</b>	<b>824</b>	<b>841</b>	<b>859</b>	<b>STEADY</b>	<b>2%</b>
% change	-2%	-1%	4%	3%	1%	4%	5%	2%	-2%	<b>FIRM....</b>	<b>3%</b>
<b>TOTAL BULK</b>	<b>1,900</b>	<b>1,896</b>	<b>2,042</b>	<b>2,095</b>	<b>2,172</b>	<b>2,291</b>	<b>2,437</b>	<b>2,543</b>	<b>2,642</b>	<b>FIRM....</b>	<b>4%</b>
% change	-1%	0%	8%	3%	4%	5%	6%	4%			
<b>OTHER BULK</b>	<b>1,362</b>	<b>1,438</b>	<b>1,557</b>	<b>1,557</b>	<b>1,679</b>	<b>1,761</b>	<b>1,866</b>	<b>1,949</b>	<b>2,051</b>	<b>STRONG!</b>	<b>5%</b>
% change	2%	6%	8%	0%	8%	5%	6%	4%			
<b>Total Dry Trade</b>	<b>Million Tones</b>									<b>Imported trend</b>	
	1998	1999	2000	2001	2002	2003	2004	Est. 2005	F'cast 2006*	Next Year vs. This Year	
<b>TOTAL DRY</b>	<b>3,262</b>	<b>3,334</b>	<b>3,598</b>	<b>3,652</b>	<b>3,851</b>	<b>4,052</b>	<b>4,303</b>	<b>4,492</b>	<b>4,694</b>	<b>FIRM...</b>	<b>4%</b>
% change	0%	2%	8%	1%	5%	5%	6%	4%			

Table 2.2: Seaborne Trade Data April 2006

Source: Clarkson Research Studies

## *The Market Outlook*

### *The Freight Indices March – April '06*

**Capesize:** During the past month, the BCI showed a weakening trend. It was 3,442 on April 12, compared to 3,867 on March 13. The average spot earnings of a modern Capesize vessel was \$36,212/day in the first week of April compared to \$40,498/day a month before.

**Panamax:** The BPI also weakened from the second half of March. On March 14, it was reported as 2,543. By April 12 it had slipped down to 2,257. Average spot earnings for a modern Panamax stood at \$14,437/day in the first week of April, down from \$18,002/day a month before.

**Handymax:** In contrast, the BSI firmed during the past month. On April 12 it was at 1,828, up from 1,705 on March 13. The one-year timecharter rate for a 52,000 dwt vessel moved in the same way and averaged at \$16,750/day in the week ending April 7.

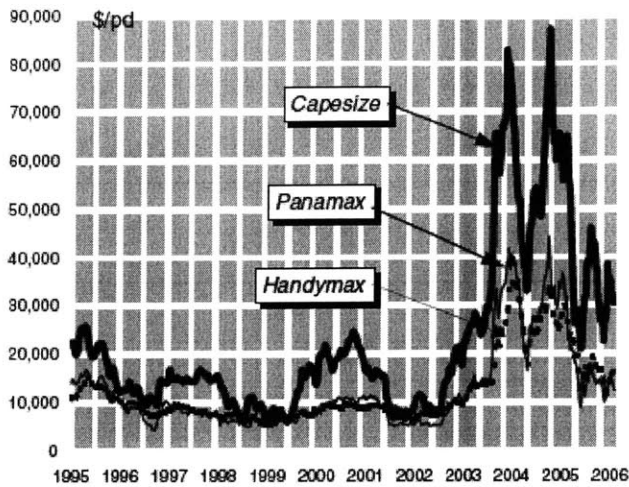


Figure 2.12: Bulker Spot Earnings (\$/Day)

	2005	2006
Demand Growth	321	301
Supply Growth	336	321
Supply/Demand Balance	15	20
<b>FREIGHT COSTS</b>	<b>SOFT</b>	<b>SOFT</b>

Table 2.3: No. of Vessels.

End Year	2005		2006	
Iron Ore	63	11%	48	7%
Coking Coal	3	2%	12	7%
Steam Coal	25	5%	13	3%
Grain"	-3	-1%	8	3%
Minor Bulk^	19	2%	18	2%
<b>Bulk Demand Growth</b>	<b>106</b>	<b>4%</b>	<b>99</b>	<b>4%</b>

Note: "incl. soyabean, ^including alumina/bauxite/phosphate

Table 2.4: Trade growth in Million mt and % change yoy.

End Year	2005		2006	
Handysize	74.1	2%	74.2	0%
Orderbook % Fleet	6%			
Handymax	66.4	9%	70.9	7%
Orderbook % Fleet	18%			
Panamax inc. Combo	94.6	8%	102.4	8%
Orderbook % Fleet	23%			
Capesize inc. Combo	114.3	7%	122.6	7%
Orderbook % Fleet	24%			
<b>Total Availability</b>	<b>349.3</b>	<b>7%</b>	<b>370.2</b>	<b>6%</b>
Orderbook % Fleet	19%			

Table 2.5: Fleet in Million DWT and % change yoy.

Source: Clarkson Research Studies

### Seaborne Iron Ore Trade

Imports to Europe	Million Tones										Imported trend	
	1998	1999	2000	2001	2002	2003	2004	Est. 2005	F'cast 2006*	Next Year vs. This Year		
Belg-Lux	12.1	10.9	10.6	10.1	11.0	10.4	11.0	9.9	10.4	<i>FIRM</i>	5%	
France	20.7	20.3	19.7	16.7	19.0	19.0	19.4	18.7	19.7	<i>FIRM</i>	6%	
Germany	53.6	38.9	47.6	40.1	44.3	38.9	39.7	38.1	40.0	<i>FIRM</i>	5%	
Italy	16.4	16.3	17.5	15.8	14.9	15.2	15.9	16.1	16.1	<i>STEADY</i>	0%	
Netherlands	8.4	7.6	7.0	7.4	7.0	7.8	7.5	7.5	7.5	<i>STEADY</i>	0%	
Spain	6.9	6.3	6.3	6.6	6.2	5.8	6.4	6.5	6.5	<i>STEADY</i>	0%	
UK	20.7	17.0	16.7	15.4	13.3	16.1	16.3	16.3	17.1	<i>FIRM</i>	5%	
Finland	2.1	2.0	2.0	2.0	1.9	1.7	1.7	1.7	1.8	<i>STRONG!!</i>	7%	
<b>EU15</b>	<b>143.7</b>	<b>121.6</b>	<b>130.1</b>	<b>116.5</b>	<b>120.1</b>	<b>117.5</b>	<b>120.3</b>	<b>117.6</b>	<b>122.0</b>	<i>FIRM</i>	4%	
Turkey	3.9	3.0	4.1	3.6	5.8	5.5	5.6	5.7	5.8	<i>STEADY</i>	2%	
<b>TOTAL W. EUROPE</b>	<b>147.6</b>	<b>125.7</b>	<b>134.5</b>	<b>120.0</b>	<b>126.0</b>	<b>124.6</b>	<b>128.5</b>	<b>126.4</b>	<b>131.0</b>	<i>FIRM</i>	4%	
<b>% change</b>	<b>11%</b>	<b>-15%</b>	<b>7%</b>	<b>-11%</b>	<b>5%</b>	<b>-1%</b>	<b>3%</b>	<b>-2%</b>				

Table 2.6: Iron Ore Imports to Europe from 1998-2006

Source: Clarkson Research Studies

Imports to Asia	Million Tones									Imported trend	
	1998	1999	2000	2001	2002	2003	2004	Est. 2005	F'cast 2006*	Next Year vs. This Year	
Japan	120.8	120.1	131.7	126.3	129.1	132.1	134.9	132.3	133.0	STEADY	1%
% change	-5%	-1%	10%	-4%	2%	2%	2%	-2%			
P.R.China	51.8	55.3	70.0	92.4	111.5	148.2	208.1	275.0	315.0	STRONG!!	15%
R.o.Korea	33.6	35.5	38.9	45.9	43.3	41.3	41.6	41.2	41.2	STEADY	0%
Taiwan	14.2	13.3	14.9	15.6	15.2	15.6	15.7	14.6	15.0	STEADY	3%
India	0.6	0.6	0.5	0.3	0.3	0.4	0.4	0.5	0.4	WEAKER	-12%
Pakistan	1.3	1.5	1.4	1.4	1.4	0.9	0.9	0.9	0.9	FIRM	5%
Malaysia	1.1	1.4	1.7	1.6	1.3	1.6	1.7	1.7	1.7	STEADY	2%
Indonesia	1.2	1.9	2	1.2	1.2	1.5	1.6	1.6	1.6	STEADY	-2%
Philippines	4.8	4	4.3	4.3	4.7	4.3	4.3	4.3	4.3	STEADY	0%
TOTAL excl. Japan	108.5	113.5	133.7	162.5	178.8	213.8	274.3	339.8	380.1	STRONG!!	12%
% change	-9%	5%	18%	22%	10%	20%	28%	24%			

Table 2.7: Iron Ore Imports to Asia from 1998-2006

Source: Clarkson Research Studies

Imports to Others	Million Tones									Imported trend	
	1998	1999	2000	2001	2002	2003	2004	Est. 2005	F'cast 2006*	Next Year vs. This Year	
Egypt	2.2	2.2	3.4	3.3	3.5	4.4	4.7	4.7	5.2	STRONG!!	11%
Libya	1.6	1.5	1.9	1.4	1.5	1.7	2.1	2.2	2.2	STEADY	0%
TOTAL AFRICA	5.0	3.7	5.5	4.9	5.2	6.2	6.7	6.8	7.3	STRONG!!	7%
Bahrain	2.0	1.8	3.3	1.8	2.9	3.1	4.0	4.3	4.4	STEADY	2%
Iran	1.7	1.8	2.4	3.0	3.0	1.5	1.7	3.0	2.4	WEAKER	-19%
Saudi	3.1	3.2	4.2	4.1	4.7	4.7	4.9	5.5	5.7	FIRM	4%
Qatar	1.0	1.1	1.1	1.2	1.2	1.2	1.2	1.2	1.2	STEADY	0%
TOTAL M. EAST	7.8	7.9	11.0	10.1	11.8	10.4	11.8	14.0	13.7	STEADY	-2%
USA (excl. Canada)	8.5	7.4	7.7	6.2	6.9	5.6	5.9	5.4	5.9	STRONG!!	9%
TOTAL N. AMERICA	12.3	11.3	11.2	9.3	10.5	9.1	9.8	8.8	9.6	STRONG!!	9%
TOTAL S. AMERICA	6.7	6.1	5.7	4.7	5.4	7.2	7.6	7.5	8.0	FIRM	6%

Table 2.8: Iron Ore Imports to Other Countries from 1998-2006

Total Imports	Million Tones									Imported trend	
	1998	1999	2000	2001	2002	2003	2004	Est. 2005	F'cast 2006*	Next Year vs. This Year	
TOTAL	427	402	448	451	481	519	589	651	699	STRONG!!	7%
% change	0%	-6%	11%	1%	7%	8%	13%	11%			

Table 2.9: Total Iron Ore Imports from 1998-2006

Source: Clarkson Research Studies

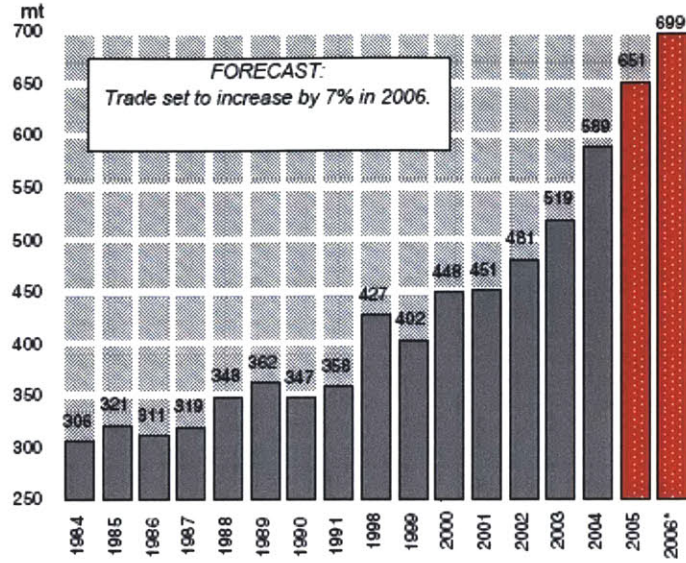


Figure 2.13: Total Iron Ore Imports from 1984-2006

Source: Clarkson Research Studies

Exports	Million Tones					
	2001	2002	2003	2004	2005	2006
<b>PACIFIC...</b>						
<i>Australia</i>	157.1	165.8	187.6	211.0	241.0	270.0
<i>India</i>	37.3	46.6	55.0	68.1	82.9	78.0
<i>Peru</i>	4.4	4.6	5.2	5.4	4.5	5.0
<b>ATLANTIC...</b>						
<i>Brazil</i>	155.7	170.0	184.4	204.7	225.0	249.0
<i>Canada~</i>	17.5	20.3	20.2	20.7	21.0	22.0
<i>Sweden</i>	13.7	14.5	16.1	17.4	17.8	18.0
<i>South</i>	23.5	25.1	24.1	26.5	28.9	29.0
<i>Mauritania</i>	10.1	10.5	9.6	10.6	11.0	12.0
<b>TOTAL</b>	<b>451</b>	<b>481</b>	<b>519</b>	<b>589</b>	<b>651</b>	<b>699</b>
<b>% change</b>	<b>1%</b>	<b>7%</b>	<b>8%</b>	<b>13%</b>	<b>11%</b>	<b>7%</b>

Table 2.10: Iron Ore Exports from 2001-2006

Source: Clarkson Research Studies

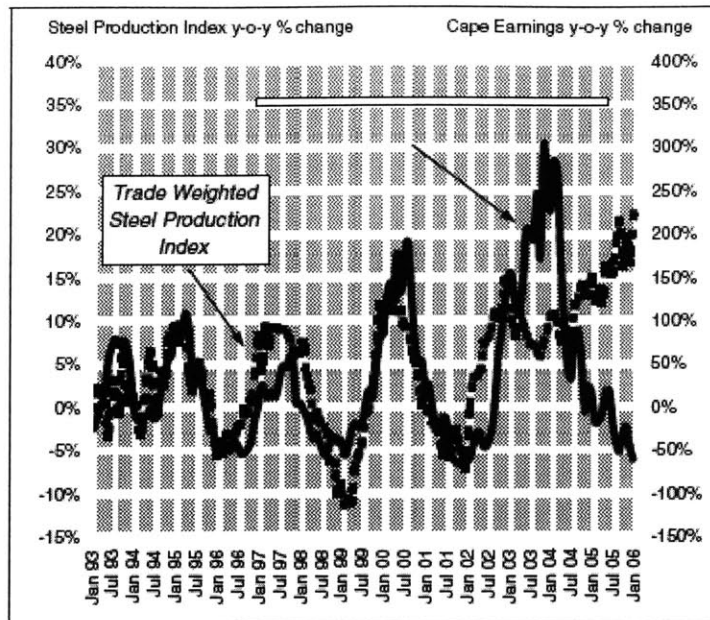


Figure 2.14: Steel Production and Capesize Freight Rates from 1993-2006

Source: Clarkson Research Studies

## 2.6 Iron Ore – Australia and Brazil the dominant players in 2006

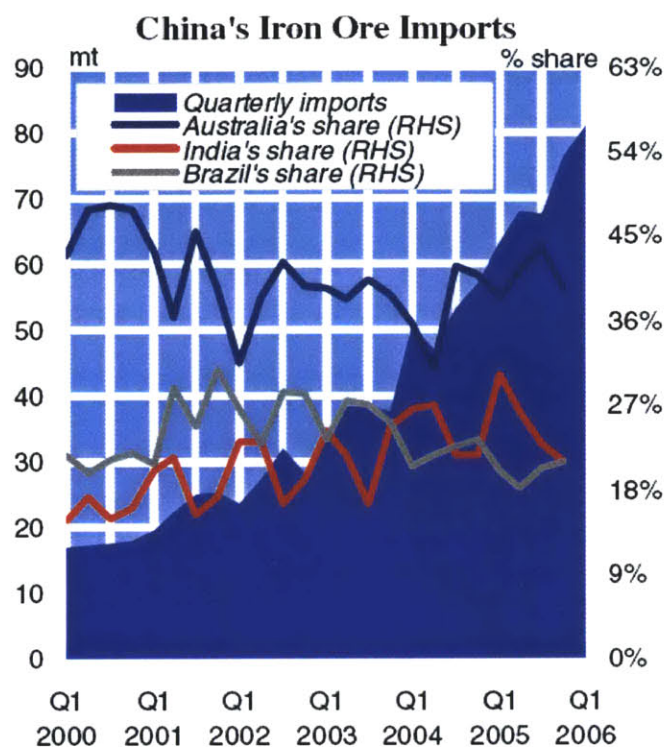
In the past few years, the iron ore market has been tight with rising ore prices a regular feature. Several brokers from annual iron ore contract talks suggest yet another increase this year. This would imply that the suppliers are still optimistic about demand growth, at least in the short run.

### 2.6.1 China Driving Demand

Most of the recent iron ore trade growth has come from China. Rapid steel production growth there has led to a surge in iron ore imports, as can be seen in Figure 2.15, which shows China's quarterly imports and major exporters' shares in this market. Australia, India and Brazil are currently the top three ore suppliers to China.

As more new capacity comes on stream and the old remains reluctant to be removed, Chinese ore demand will continue to grow in the short run. China might also

indirectly impact the pattern of the global ore trade. Its declining steel imports and increasing exports could eat into other countries' steel exports and lead to a reduction of ore demand outside China.



**Figure 2.15:** China's Iron Ore Imports from 2000-2006  
Source: Clarkson Research Studies

### 2.6.2 Concentrating the Supply Side

Australia, Brazil and India account for 85% of world iron ore exports. The top two, in which the world's top three miners BHP, Rio Tinto and CVRD are based, account for over 70%. Their expansion plans are very important to the world iron ore trade prospects. The high level of mining investment in Australia and Brazil means that expansions of production there are likely to match expected world demand growth. Consequently, there would be less room for India to expand its exports.

Moreover, the growth in domestic steel demand is likely to absorb the major part of whatever expansion to mine capacity is taking place in India. As a result it is not expected that Indian ore exports will grow, indeed they may even contract in the years to come. So it is possible that Australia and Brazil will become more dominant in world ore markets in the near future.

### **2.6.3 Trading Partners**

In the future, Australia and Brazil seem to be the most likely sources for meeting rising demand. For China, Australian ore will have the advantage of shorter shipping distances over Brazilian ore. As India and China work toward long term ore trade supply contracts to gain a more stable trade pattern, India and Brazil could still fight for second place in China's market. Over the past couple of years India has gained market share on Brazil, but as mentioned this may not last.

### **2.6.4 Ore Shipping Demand**

The outlook for seaborne iron ore demand is positive in the short run, but longer term the health of the steel market will determine whether it grows or slows. The vulnerability of the shipping markets to such shifts will be mitigated by the greater concentration of iron ore demand in China. Its long haul suppliers and the increase in tone-miles will help underpin the demand for tonnage.

# Chapter 3

## Artificial Neural Networks

### 3.1 Introduction

The simulation of human intelligence using machines was and is always a challenge to ingenuity. In the middle of this century a research discipline calling itself “*Artificial Intelligence*” (AI) emerged. The definition for the term AI is very indistinct; this has its major reason in the fact that there is no commonly accepted definition for “*intelligence*”. The most comprehensive definition for AI includes all research aimed to simulate intelligent behavior.

Despite the availability of massive computational power the results of current research are far from the aims proposed in the enthusiasm of the 1950s and 60s. Nevertheless systems built to simulate intelligence in a limited domain, such as expert systems in medicine or forecasting applications are already successfully used.

Artificial Neural Networks (ANNs) are of major research interest at present, involving researchers of many different disciplines. Subjects contributing to this research include biology, computing, electronics, mathematics, medicine, physics, economics and psychology. The approaches to this topic are very diverse, as are the aims. The basic idea is to use the knowledge of the nervous system and the human brain to design intelligent artificial systems.

On one side biologists and psychologists are trying to model and understand the brain and parts of the nervous system and searching for explanations for human behavior and reasons for the limitations of the brain.

On the other hand, computer scientists and computer engineers are searching for efficient ways to solve problems for which conventional computers are currently used. Biological and psychological models and ideas are often the resource of inspiration for these scientists.

### **3.2 Brief History of Artificial Neural Networks**

The origins of the concept of artificial neural networks can be traced back more than a century as a consequence of man's desire for understanding the brain and emulating its behavior. A century old observation by William James that the *amount of activity at any given point in the brain cortex is the sum of the tendencies of all other points to discharge into it*, has been reflected subsequently in the work of many researchers exploring the field of artificial neural networks. About half a century ago the advances in neurobiology allowed researchers to build mathematical models of the neurons in order to simulate the working model of the brain. This idea enabled scientists to formulate one of the first abstract models of a biological neuron, reported in 1943. The formal foundation for the field of artificial neural networks was laid down by McCulloch and Pitts. In their paper the authors proposed a computational model based on a simple neuron-like logical element. Later, a learning rule for incrementally adapting the connection strength of these artificial neurons was proposed by Donald Hebb. The Hebb rule became the basis for many artificial neural network research models.

The main factor responsible for the subsequent decline of the field of artificial neural networks was the exaggeration of the claims made about the capabilities of early models of artificial neural networks which cast doubt on the validity of the entire field. This factor is frequently attributed to a well known book "*The Perceptrons*" by Minsky and Papert MIT Press, 1968 who reported that perceptrons were not capable of learning moderately difficult problems like the XOR problem. However, theoretical analyses of that book were not alone responsible for the decline of research in the field of artificial neural networks. This fact, amplified by the frustration of researchers, when their high expectations regarding artificial neural network based artificial intelligent systems were not met, contributed to the decline and discrediting of this new research field. Also, at the same time logic based intelligent systems which were mostly based on symbolic logic reached a high degree of sophistication. Logic based systems were able to capture certain features of human reasoning, like default logic, temporal logic and reasoning about beliefs. All these factors lead to the decline of research in the field of artificial neural network to a point where it was almost abandoned.

Since the early 1980s there has been a renewed interest in the field of artificial neural networks that can be attributed to several factors. The shortcomings of the early simpler neuron network models were overcome by the introduction of more sophisticated artificial neural network models along with new training techniques. Availability of high speed desktop digital computers made the simulation of complex artificial neural network models more convenient. During the same time frame significant research efforts of several scientists helped in restoring the lost confidence in this field of research. Hopfield with his excellent research efforts is responsible for revitalization of ANN field. The term

connectionist was made popular by Feldman and Ballard. Confidence in the artificial neural network field was greatly enhanced by the research efforts of Rumelhart, Hinton and Williams who developed a generalization of Widrow's delta rule which would make multi-layered artificial neural network learning possible. This was followed by demonstrations of artificial neural network learning of difficult tasks in the fields of speech recognition, language development, control systems and pattern recognition, and others.

Since the mid 1980s research in the area of artificial neural networks has experienced an extremely rapid growth for different reasons which are evident by its interdisciplinary nature. These disciplines range from physics and engineering to physiology and psychology. In recent years there has been a lot of progress in the development of new learning algorithms, network structures and VLSI implementations for artificial neural networks.

### **3.3 Biological Neural Networks**

Artificial Neural Networks draw much of their inspiration from the biological nervous system. It is therefore very useful to have some knowledge of the way this system is organized.

Most living creatures, which have the ability to adapt to a changing environment, need a controlling unit which is able to learn. Higher developed animals and humans use very complex networks of highly specialized neurons to perform this task.

The control unit -or brain- can be divided in different anatomic and functional sub-units, each having certain tasks like vision, hearing, motor and sensor control. The brain is connected by nerves to the sensors and actors in the rest of the body.

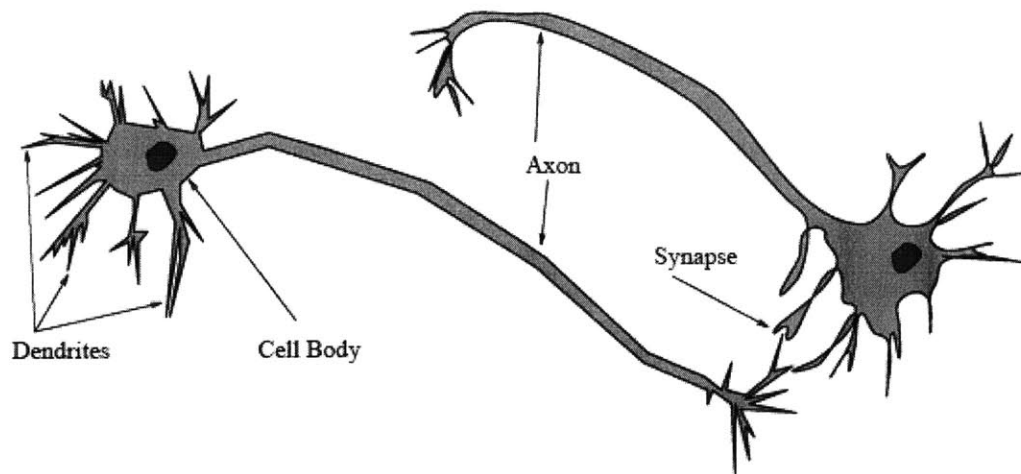
The brain consists of a very large number of neurons, about  $10^{11}$  in average. These can be seen as the basic building bricks for the central nervous system (CNS). The neurons are interconnected at points called synapses. The complexity of the brain is due to the massive number of highly interconnected simple units working in parallel, with an individual neuron receiving input from up to 10,000 others.

The neuron contains all structures of an animal cell. The complexity of the structure and of the processes in a simple cell is enormous. Even the most sophisticated neuron models in artificial neural networks seem comparatively toy-like.

Structurally the neuron can be divided in three major parts: the cell body (soma), the dendrites, and the axon, see *Figure 3.1* for an illustration. The cell body contains the organelles of the neuron and also the “dendrites” are originating there. These are thin and widely branching fibers, reaching out in different directions to make connections to a larger number of cells within the cluster. Input connections are made from the axons of other cells to the dendrites or directly to the body of the cell. These are known as axon-dendritic and axon-somatic synapses.

There is only one axon per neuron. It is a single and long fiber, which transports the output signal of the cell as electrical impulses (action potential) along its length.

The end of the axon may divide in many branches, which are then connected to other cells. The branches have the function to fan out the signal to many other inputs.



**Figure 3.1:** *Simplified Biological Neurons*

*Source: Murray L. Barr & John A. Kiernan the Human Nervous System. 1988*

There are many different types of neuron cells found in the nervous system. The differences are due to their location and function.

The neurons perform basically the following function: all the inputs to the cell, which may vary by the strength of the connection or the frequency of the incoming signal, are summed up. The input sum is processed by a threshold function and produces an output signal. The processing time of about *1ms per cycle* and transmission speed of the neurons of about *0.6 to 120 m/s* are extremely slow to a modern computer.

The brain works in both a parallel and serial way. The parallel and serial nature of the brain is readily apparent from the physical anatomy of the nervous system. That there is serial and parallel processing involved can be easily seen from the time needed to perform tasks. For example a human can recognize the picture of another person in about *100 ms*. Given the processing time of *1 ms* for an individual neuron this implies that a certain number of neurons, but less than 100, are involved in serial, whereas the complexity of the task is evidence for a parallel processing, because a difficult

recognition task can not be performed by such a small number of neurons. This phenomenon is known as the *100-step-rule*.

Biological neural systems usually have a very high fault tolerance. Experiments with people with brain injuries have shown that damage of neurons up to a certain level does not necessarily influence the performance of the system, though tasks such as writing or speaking may have to be learned again. This can be regarded as re-training the network.

### **3.4 Definition**

In the literature a wide variety of definitions and explanations for the term “*Artificial Neural Network*” can be found. The following definition by Laurene Fausett is considered the most accurate and descriptive one:

*“An artificial neural network is an information-processing system that has certain performance characteristics in common with biological neural networks. Artificial neural networks have been developed as generalizations of mathematical models of human cognition or neural biology, based on the assumption that (Laurence Fausett 1994):*

- 1. Information processing occurs at many simple elements called neurons.*
- 2. Signals are passed between neurons over connection links.*
- 3. Each connection link has an associated weight, which, in a typical neural net, multiplies the signal transmitted.*
- 4. Each neuron applies an activation function (usually nonlinear) to its net input (sum of weighted input signals) to determine its output signal.”*

### 3.5 Memorization and Generalization

To simulate intelligent behavior the abilities of memorization and generalization are essential. These are basic properties of artificial neural networks. The following definitions are according to the Collins English Dictionary:

*to memorize*: to commit to memory; learn so as to remember.

*to generalize*: to form general principles or conclusions from detailed facts, experience, etc.

*Memorizing*, given facts, is an obvious task in learning. This can be done by storing the input samples explicitly, or by identifying the concept behind the input data, and memorizing their general rules.

The ability to identify the rules, *to generalize*, allows the system to make predictions on unknown data. Despite the strictly logical invalidity of this approach, the process of reasoning from specific samples to the general case can be observed in human learning. Generalization also removes the need to store a large number of input samples. Features common to the whole class need not to be repeated for each sample; instead the system needs only to remember which features are parts of the sample. This can dramatically reduce the amount of memory needed, and produce a very efficient method of memorization.

### 3.6 Acquiring Knowledge by Learning

A vital property of neural networks is that they can learn the desired response from a set of examples in the domain. This contrasts with most other approaches in computing where an algorithm or rules are used to store the knowledge.

The advantage of learning from examples is that there is no need to explicitly form a rule system for the task. To extract rules from the knowledge in the domain implies that there is some expert interpretation. This process is often difficult, especially if the experts have different opinions on the problem. From abstract point of view training a neural network can be seen as an automatic process of extracting rules from a data set.

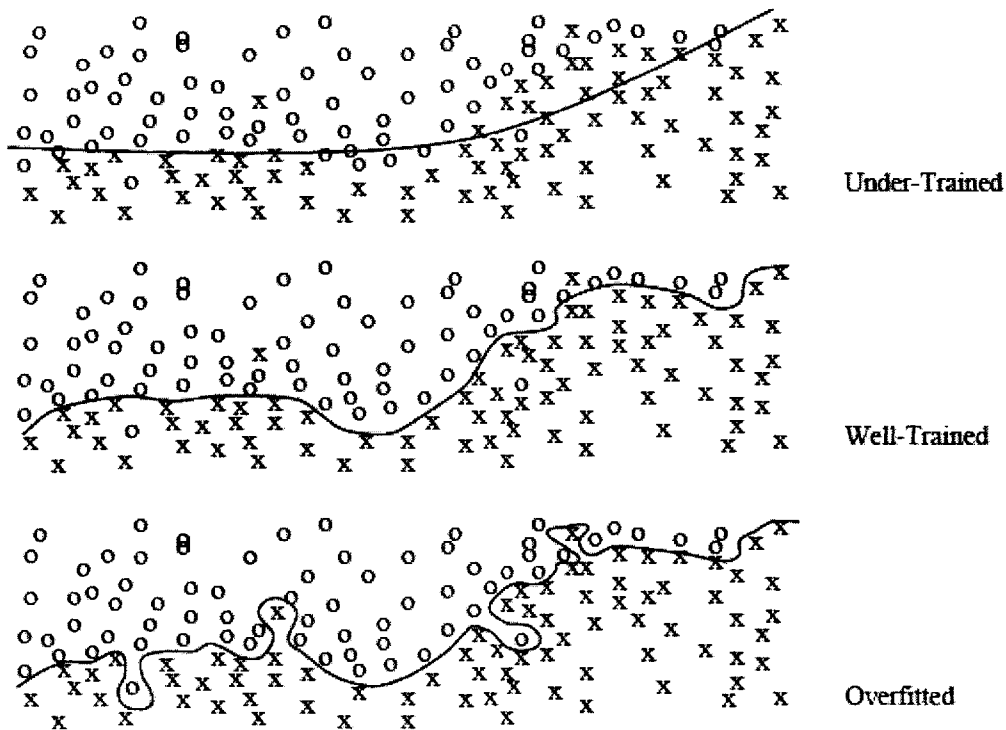
There are two basic paradigms of learning, *supervised and unsupervised*, both of which have their models in biology.

*Supervised learning* at its most general is a process where both information about the environment (e.g. the sensory stimuli) and the desired reaction of the system (e.g. the motor response) is given. It is analogous to human learning with a teacher who knows all the answers.

In an ANN context supervised learning is a process of memorizing vector pairs. The input vector together with the desired output vector is known. This method is often referred to as memorizing a mapping from the input space to the output space. A special case is auto-associative mapping where the input pattern and the output pattern are equal (often written as a single vector). Auto-associative memories are often used to retrieve the original pattern for a distorted input.

Many algorithms for supervised learning work on a comparison between the calculated response of the network during training and the target values. There are also learning techniques where the input-output pair is directly used to calculate the weights in the network (e.g. the bidirectional associative memory).

A variant of supervised learning is called *reinforcement learning*. In this method the required output is not provided; the response by the teacher is only whether the calculated result is “right” or “wrong”.



*Figure 3.2: Learning and the Problem of Overfitting.*

Source: Dan W. Patterson. *Artificial Neural Networks 1996*

In supervised learning it is often difficult to determine when the learning process should be terminated. A network with a small error (the overall difference between the calculated and desired output) does not necessarily show a good performance on new data from the same domain. The problem is called *overfitting*. If the training process

went on too long the network is biased towards the training set and the generalization ability of the network is decreased. If the process is stopped too early the decision is very rough. In *Figure 3.2* this is illustrated for the separation of two sets.

**Unsupervised learning** works only on the input vectors, and the desired output is not specified. This learning method can be compared to the process of categorization, discovering regularities, or adaptation to specific features.

The following sets explain the basic concept;  $S$  is the original set, the task is to split the set into two groups (here  $S_1$  and  $S_2$ ):

$$S = \{\sqcap, \vee, \sqcup, \cap, \wedge, \sqcap, \cup\}$$

$$S_1 = \{\sqcap, \wedge, \cap\}$$

$$S_2 = \{\vee, \sqcup, \cup\}$$

In this case the solution found, is a categorization according to the opening of the symbol. There are certainly other possible groupings, but the chosen one is very obvious. Considering another set, difficulties with the unsupervised learning appear. If there are different ways to split the set, the process is not straight forward anymore. Assuming  $S = \{a, A, b, B\}$ , the following categories make equal sense:  $S_1 = \{a, b\}$  and  $S_2 = \{A, B\}$  or  $S_1 = \{a, A\}$  and  $S_2 = \{b, B\}$ .

In many unsupervised models the categorization occurs according to the distance between the input vectors. An example for a measure used is the Hamming distance for binary vectors. Generalization based on this approach groups input vectors in a way to minimize the distance between the members of a category for all categories.

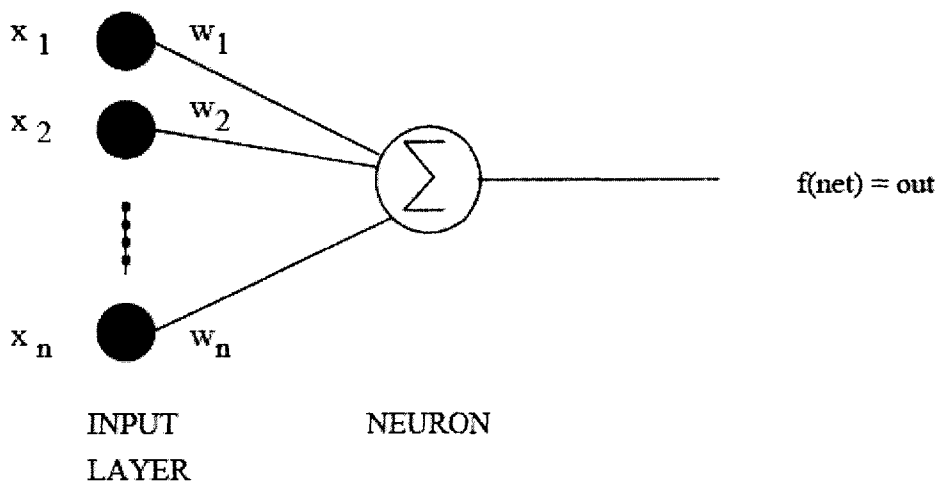
If using an unsupervised model it is very important to analyze whether the clustering done by the network is the right way of grouping the data for the given problem.

### 3.7 An Artificial Neuron

The artificial neuron shown in *Figure 3.3* is a very simple processing unit. The neuron has a fixed number of *inputs*  $n$ , each input is connected to the neuron by a *weighted link*  $w_i$ . The neuron sums up the *net* input according to the equation:  $net = \sum_{i=1}^n x_i \cdot w_i$  or expressed as vectors  $net = \vec{x}^T \cdot \vec{w}$ . To calculate the output an *activation function*  $f$  is applied to the net input of the neuron. This function is either a simple threshold function or a continuous non linear function. Two often used activation functions are:

$$f_c(net) = \frac{1}{1 - e^{-net}}$$

$$f_T(net) = \begin{cases} \text{if } a > \theta \text{ then } 1 \\ \text{else } 0 \end{cases}$$



*Figure 3.3: An Artificial Neuron.*

Source: Dan W. Patterson. *Artificial Neural Networks* 1996

The artificial neuron is an abstract model of the biological neuron. The strength of a connection is coded in the weight. The intensity of the input signal is modeled by using a real number instead of a temporal summation of spikes. The artificial neuron works in discrete time steps; the inputs are read and processed at one moment in time.

There are many different learning methods possible for a single neuron. Most of the supervised methods are based on the idea of changing the weight in a direction that the difference between the calculated output and the desired output is decreased. Examples of such rules are the *Perceptron Learning Rule*, the *Widrow-Hoff Learning Rule*, and the *Gradient descent Learning Rule*.

The *Gradient descent Learning Rule* operates on a differentiable activation function. The weight updates are a function of the input vector  $\bar{x}$ , the calculated output  $f(net)$ , the derivative of the calculated output  $f'(net)$ , the desired output  $d$ , and the learning constant  $n$ .

$$net = \bar{x}^T \cdot \bar{w}$$
$$\Delta w = n \cdot f'(net) \cdot (d - f(net)) \cdot \bar{x}$$

The delta rule changes the weights to minimize the error. The error is defined by the difference between the calculated output and the desired output. The weights are adjusted for one pattern in one learning step. This process is repeated with the aim to find a weight vector that minimizes the error for the entire training set.

A set of weights can only be found if the training set is linearly separable. This limitation is independent of the learning algorithm used; it can be simply derived from the structure of the single neuron.

To illustrate this consider an artificial neuron with two inputs and a threshold activation function  $f_T$ ; this neuron is intended to learn the XOR-problem (see table). It

can easily be shown that there are no real numbers  $w_1$  and  $w_2$  to solve the equations, and hence the neuron can not learn this problem.

Input Vector	Desired Output	Weight Equation
0 0	1	$0 \cdot w_1 + 0 \cdot w_2 > \theta \Rightarrow 0 > \theta$
1 0	0	$1 \cdot w_1 + 0 \cdot w_2 < \theta \Rightarrow w_1 < \theta$
0 1	0	$0 \cdot w_1 + 1 \cdot w_2 < \theta \Rightarrow w_2 < \theta$
1 1	1	$1 \cdot w_1 + 1 \cdot w_2 > \theta \Rightarrow w_1 + w_2 > \theta$

*Source: M. Minsky and S. Papert. Perceptrons. MIT Press 1969*

### 3.8 A Single Layer Network

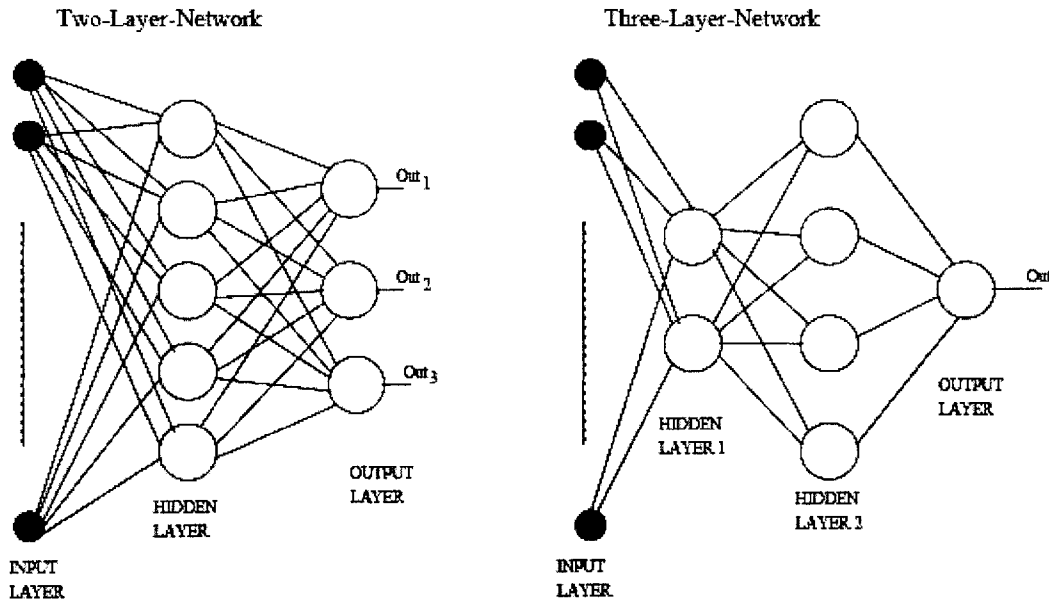
A single layer network is a simple structure consisting of  $m$  neurons each having  $n$  inputs. The system performs a mapping from the  $n$ -dimensional input space to the  $m$ -dimensional output space. To train the network the same learning algorithms as for a single neuron can be used.

This type of network is widely used for linear separable problems, but like a neuron, single layer network are not capable of classifying non linear separable data sets. One way to tackle this problem is to use multilayer network architecture.

### 3.9 Multilayer Neural Networks

*Multilayer networks* solve the classification problem for non linear sets by employing *hidden layers*, whose neurons are not directly connected to the output. The additional hidden layers can be interpreted geometrically as additional hyper-planes,

which enhance the separation capacity of the network. *Figure 3.4* shows typical multilayer network architectures.



**Figure 3.4:** Examples of Multilayer Neural Network Architectures.

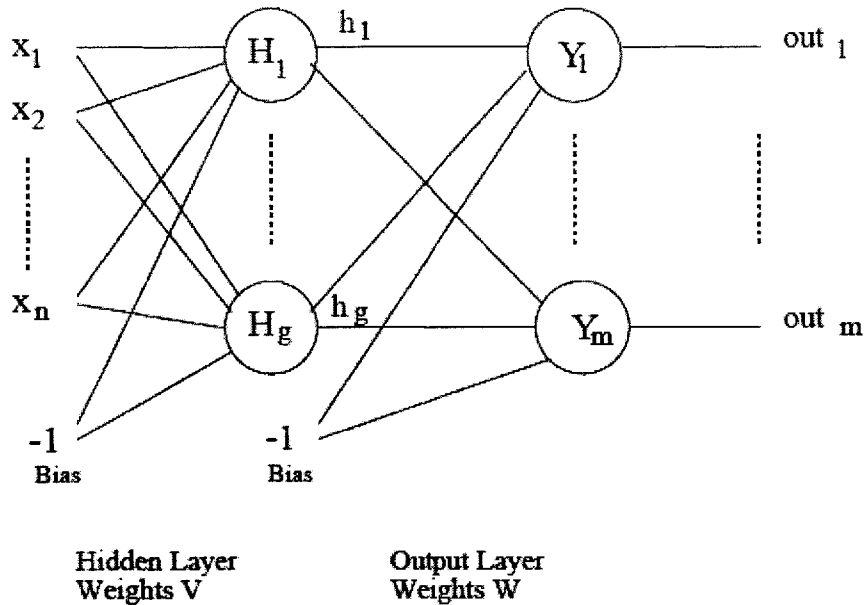
Source: Dan W. Patterson. *Artificial Neural Networks* 1996

This new architecture introduces a new question: “*how to train the hidden units for which the desired output is not known*” The ***Backpropagation algorithm*** offers a solution to this problem.

The training occurs in a supervised style. The basic idea is to present the input vector to the network; calculate in the forward direction the output of each layer and the final output of the network. For the output layer the desired values are known and therefore the weights can be adjusted as for a single layer network; in the case of the BP algorithm according to the gradient decent rule.

To calculate the weight changes in the hidden layer the error in the output layer is backpropagated to these layers according to the connecting weights. This process is repeated for each sample in the training set. One cycle through the training set is called

an *epoch*. The number of epochs needed to train the network depends on various parameters, especially on the error calculated in the output layer.



*Figure 3.5: The Backpropagation Network.*

*Source: D.E. Rumelhart, G.E. Hinton and R.J. Williams MIT Press 1986*

The assumed architecture is depicted in *Figure 3.5*. The input vector has  $n$  dimensions; the output vector has  $m$  dimensions, the bias (the used constant input) is  $-1$ , there is one hidden layer with  $g$  neurons. The matrix  $V$  holds the weights of the neurons in the hidden layer. The matrix  $W$  defines the weights of the neurons in the output layer. The learning parameter is  $n$  and the momentum is  $a$ .

The used unipolar activation function and its derivative are given by:

$$f(net) = \frac{1}{1 - e^{-\lambda \cdot net}}$$

$$f'(net) = \frac{e^{-\lambda \cdot net}}{(1 + e^{-\lambda \cdot net})^2}$$

The training set consists of pairs where  $\bar{x}^p$  is the input vector and  $\bar{t}^p$  is the desired output vector.

$$T = \{(\bar{x}, \bar{t}), \dots, (\bar{x}^p, \bar{t}^p)\}$$

1. Initialize the weights  $V$  and  $W$  randomly with numbers from a suitable interval (e.g. -1 to 1). Select the parameters  $n$  and  $a$ .
2. Randomly take one unmarked pair  $(\bar{x}, \bar{t})$  of the training set for the further steps and mark it as used.
3. Do the forward calculation. The notation  $\bar{x}'$  is the input vector  $\bar{x}$  enlarged by the bias,  $\bar{h}'$  is the enlarged hidden layer output vector.

$$\begin{aligned} net_{Hi} &= V^T \cdot \bar{x}' \\ h_i &= f(net_{Hi}) \\ net_{yi} &= W^T \cdot \bar{h}' \\ out_i &= f(net_{yi}) \end{aligned}$$

4. Do the backward calculation.

$$\begin{aligned} \delta_{out_i} &= f'(net_{yi}) \cdot (t_i - out_i) \\ \delta_{Hi} &= f'(net_{Hi}) \cdot \sum_{j=1}^m w_{ij} \cdot \delta_{out_j} \end{aligned}$$

5. Calculate the weight changes and update the weights.

$$\begin{aligned} \Delta W^T(t) &= n \cdot \delta_{out} \cdot h'^T \\ \Delta V^T(t) &= n \cdot \delta_{Hi} \cdot x'^T \\ W(t+1) &= W(t) + \Delta W(t) + a \cdot \Delta W(t-1) \\ V(t+1) &= V(t) + \Delta V(t) + a \cdot \Delta V(t-1) \end{aligned}$$

6. REPEAT from step 2 WHILE there are unused pairs in the training.
7. Set all pairs in the training set to unused.
8. REPEAT from step 2 UNTIL the stop condition is TRUE.

Training continues until the overall error in one training cycle is sufficiently small; this stop condition is given by:

$$E_{max} > E$$

This acceptable error  $E_{max}$  has to be selected very carefully, if  $E_{max}$  is too large the network is under-trained and lacks in performance, if  $E_{max}$  is selected too small the network will be biased towards the training set (it will be over-fitted). One measure for the  $E$  is the root mean square error calculated by:

$$E = \frac{1}{P \cdot m} \cdot \sqrt{\sum_{p=1}^P \sum_{i=1}^m (t_i^p - out_i^p)^2}$$

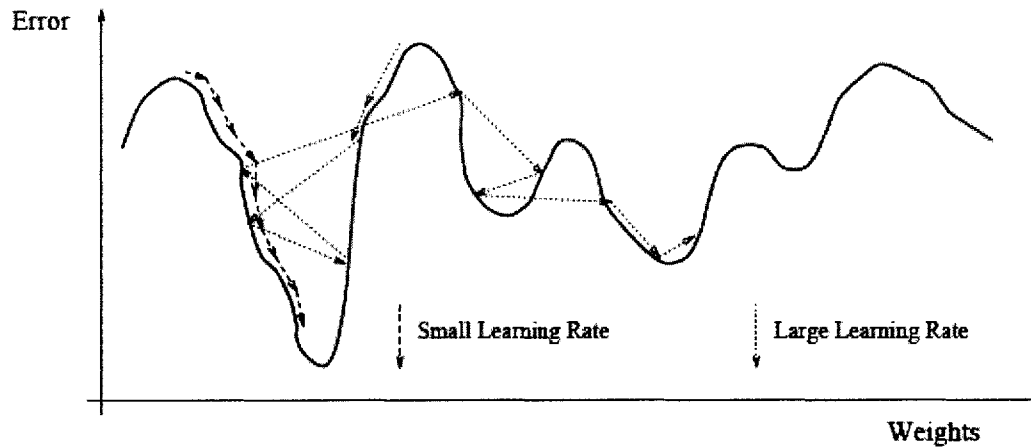
The selection of the parameters for the Backpropagation algorithm and the initial settings of the weight influence the learning speed as well as the convergence of the algorithm.

The initial **Weights** chosen determine the starting point in the error landscape, which controls whether the learning process will end up in a local minimum or the global minimum. The easiest method is to select the weights randomly from a suitable range, such as between (-0.1,0.1) or (-2,2).

If the weight values are too large, the net value will be large as well; this causes the derivative of the activation function to work in the saturation region and the weight changes to be near zero. For small initial weights the changes will also be very small, which causes the learning process to be very slow and might even prevent convergence.

More sophisticated approaches to select the weights, such as the *Nguyen-Widrow Initialization* which calculates the interval from which the weights are taken in accordance with the number of input neurons and the number of hidden neurons, can improve the learning process.

The *Learning Coefficient*  $n$  determines the size of the weight changes. A small value for  $n$  will result in a very slow learning process. If the learning coefficient is too large the large weight changes may cause the desired minimum to be missed.



**Figure 3.6:** *The influence of the Learning Rate on the Weight Changes.*

*Source: Laurence Fausett Fundamentals of Neural Networks 1994*

A useful range is between  $0.05$  and  $2$  dependent on the problem. The influence of the  $n$  on the weight changes is shown in *Figure 3.6*.

An improved technique is to use an adaptive learning rate. A large initial learning coefficient should help to escape from local minima, while reducing  $n$  later should prevent the learning process from overshooting the reached minimum.

The *Momentum*  $a$  causes the weight changes to be dependent on more than one input pattern. The change is a linear combination of the current gradient and the previous gradient. The useful range for this parameter is between  $0$  and  $1$ . For some data sets the momentum makes the training faster, while for others there may be no improvement. The momentum usually makes it less likely that the training process will get stuck in a local minimum.

In recent years an enormous number of publications on refinements and improvements of the Backpropagation algorithms have been published. However most of the suggested improvements are only useful if the problem meets certain conditions.

Nevertheless the multilayer feedforward networks trained with the Backpropagation method are probably the most practically used networks for real world applications.

### **3.10 The Concept of Modularity**

In general, a computational system can be considered to have a modular architecture if it can be split into two or more subsystems in which each individual subsystem evaluates either distinct inputs or the same inputs without communicating with other subsystems. The overall output of the modular system depends on an integration unit which accepts outputs of the individual subsystems as its inputs and combines them in a predefined fashion to produce the overall output of the system. In a broader sense modularity implies that there is a considerable and visible functional or structural division among the different modules of a computational system. The modular system design approach has some obvious advantages, like simplicity and economy of design, computational efficiency, fault tolerance and better extendibility.

The concept of modularity is an extension of the principle of *divide and conquers*. This principle has no formal definition but is an intuitive way by which a complex computational task can be subdivided into simpler subtasks. The simpler subtasks are then accomplished by a number of the specialized local computational systems or models. Each local computational model performs an explicit, interpretable and relevant

function according to the mechanics of the problem involved. The solution to the overall complex task is achieved by combining the individual results of specialized local computational systems in some task dependent optimal fashion. The overall task decomposition into simpler subtasks can be either a *soft-subdivision* or *hard-subdivision*.

In general, the modules exhibit the following characteristics:

- 1. The modules are domain specific and have specialized computational architectures to recognize and respond to certain subsets of the overall task.*
- 2. Each module is typically independent of other modules in its functioning and does not influence or become influenced by other modules.*
- 3. The modules generally have a simpler architecture as compared to the system as a whole. Thus a module can respond to given input faster than a complex monolithic system.*
- 4. The responses of the individual modules are simple and have to be combined by some integrating mechanism in order to generate the complex overall system response.*

### **3.11 Modular Artificial Neural Networks**

The obvious advantages of modularity in learning systems, particularly as seen in the existence of the functional and architectural modularity in the brain, has made it a main stream theme in cognitive neuroscience research areas. Specifically, in the field of artificial neural network research, which derives its inspiration from the functioning and structure of the brain, modular design techniques are gaining popularity. The use of modular neural networks for the purpose of regression and classification can be

considered as a competitor to conventional monolithic artificial neural networks, but with more advantages. Two of the most important advantages are a close neurobiological basis and greater flexibility in design and implementation. Another motivation for modular neural networks is to extend and exploit the capabilities and basic architectures of the more commonly used artificial neural networks that are inherently modular in nature.

Monolithic artificial neural networks exhibit a special sort of modularity and can be considered as hierarchically organized systems in which synapses interconnecting the neurons can be considered to be the fundamental level. This level is followed by neurons which subsequently form the layers of neurons of a multi layered neural network. The next natural step to extend the existing level of hierarchical organization of an artificial neural network is to construct an ensemble of neural networks arranged in some modular fashion in which an artificial neural network comprised of multiple layers is considered as a fundamental component. This rationale along with the advances in neurobiological sciences has provided researchers a justification to explore the paradigm of modularity in design and training of neural network architectures.

The modular neural networks are comprised of modules which can be categorized on the basis of both distinct structure and functionality which are integrated together via an integrating unit. With functional categorization, each module is a neural network which carries out a distinct identifiable subtask. Also, using this approach different types of learning algorithms can be combined in a seamless fashion. These algorithms can be neural network related, or otherwise. This leads to an improvement in artificial neural network learning because of the integration of the best suited learning algorithms for a given task (when different algorithms are available). On the other hand, structural

modularization can be viewed as an approach that deviates from the conventional thinking about neural networks as non-parametric models, learning from a given data set.

In structural modularization *a priori* knowledge about a task can be introduced into the structure of a neural network which gives it a meaningful structural representation. Generally, the functional and structural modularization approaches are used in conjunction with each other in order to achieve an optimal combination of modular network structure and learning algorithm.

The following subsections highlight some of the important motivations which make the modular neural network design approach more attractive than a conventional monolithic global neural network design approach.

### **3.11.1 Model Complexity Reduction**

The model complexity of global monolithic neural networks drastically increases with an increase in the task size or difficulty. The rise in the number of weights is quadratic with respect to the increase in neural network models size. Modular neural networks on the other hand, can circumvent the complexity issue, as the specialized modules have to learn only simpler and smaller tasks in spite of the fact that the overall task is complex and difficult.

### **3.11.2 Robustness**

The homogeneous connectivity in monolithic neural networks may result in a lack of stability of representation and is susceptible to interference. Modular design of neural

network adds additional robustness and fault tolerance capabilities to the neural network model. This is evident from the design of the visual cortex system which is highly modular in design and is comprised of communicating functionally independent modules. Damage to a part of visual cortex system can result in a loss of some of the abilities of the visual system, but, as a whole the system can still function partially.

### **3.11.3 Scalability**

Scalability is one of the most important characteristics of modular neural networks which sets them apart from the conventional monolithic neural networks. In global or unitary neural networks there is no provision for incremental learning, i.e., if any additional incremental information is to be stored in a neural network, it has to be retrained using the data for which it was trained initially along with the new data set to be learned. On the other hand, modular neural networks present an architecture which is suitable for incremental addition of modules that can store any incremental addition to the already existing learned knowledge of the modular neural network structure without having to retrain all of the modules.

### **3.11.4 Learning**

Modular neural networks present a framework of integration capable of both supervised and unsupervised learning paradigms. Modules can be pre-trained individually for specific subtasks and then integrated via an integration unit or can be trained along with an integrating unit. In the latter situation, there is no indication in the training data as

to which module should perform which subtask and during training individual modules compete, or cooperate to accomplish the desired overall task. This learning scheme is a combined function of both supervised as well as unsupervised learning paradigms.

### **3.11.5 Computational Efficiency**

If the processing can be divided into separate, smaller and possibly parallel subtasks, then the computational effort will in general be greatly reduced [43]. A modular neural network can learn a set of functional mappings faster than a corresponding global monolithic neural network because each individual module in a modular neural network has to learn a, possibly simpler, part of the overall functional mapping. Also, modular networks have an inherent capability of decomposing the decomposable tasks into a set of simpler tasks, thus enhancing the learn-ability and reducing the learning time.

### **3.11.6 Learning Capacity**

Embedding modularity into neural network structures leads to many advantages compared to a single global neural network. For example, introduction of integrated local computational models of neural networks increases the learning capacity of a modular neural network model, and thus permits their use for more complex large-scale problems which ordinarily cannot be handled by global neural network models. Also, a complex behavior may require different types of knowledge and processing techniques to be integrated together which is not possible without any structural or functional modularity.

### **3.11.7 Economy of Learning**

To enable continued survival of biological systems, new functionalities are integrated into already existing systems along with continued learning and adaptation to changing environments. Using the same analogy, modularity enables learning economy in a way that if the operating conditions change, then only those parts of the modular neural network need to be modified that do not conform to the new environment, rather than the entire system. In addition, it is also possible to reuse some of the existing specialist modules for different tasks of the same nature instead of learning again the parts common to the two tasks.

### **3.11.8 Knowledge Integration**

Modularity is a way of embedding *a priori* knowledge in a neural network architecture that is important to improve the neural network learning. The motivation for integration of *a priori* knowledge about the task at hand is that it might be the optimal way to design an appropriate neural network system for the available training data. This may include the possibility to hybridize the neural network architecture. In modular neural network architecture it is possible to use and integrate different neural functions, different neural structures or different kind of learning algorithms, depending on the task at hand.

## **3.12 Artificial Neural Networks as tools of forecasting**

The use of Artificial Neural Networks as forecasting tools, present advantages but also disadvantages in comparison to other methods. The network is trained on historical

data using time series as inputs and providing as an output the desirable forecasting time series. In simple words, it tries to locate the importance of certain historical rates in order to approach a future rate, with the support of certain not linear functions.

Two basic disadvantages that Artificial Neural Networks have when they are used to forecast time series are the following:

- The required database for a well-trained network is big enough in comparison to other forecasting methods.
- It is extremely difficult to analyze the reasons for mislaid forecasts.

However, two significant advantages that characterize Artificial Neural Networks are:

- The construction of a forecasting Artificial Neural Network is an automatic procedure. The network is supplied with an appropriate database and it is trained on its own using a personal computer. The user determines only the essential parameters of the network.
- The Artificial Neural Networks are capable of not being influenced by elements of the database that would lead other methods to mislaid forecasts. For example, an extraordinary behavior that one of the time series presented in the past, will slightly affect the network's training, due to the not linear processing of the data.

Therefore, the use of *Artificial Neural Networks*, and especially a topology of a *Modular Multi Layer Perceptron* is expected to give fine results. The use of genetic algorithms will further foster the training of the network, giving even improved results.

## Chapter 4

### Construction of Database – Data Analysis

#### 4.1 Introduction

Based on what we have mentioned in the previous chapters, the appropriate form of data needed to forecast the *Capesize ore voyage rates* with the use of neural networks is as much as possible bigger numerical time series of various figures that will constitute the independent input variables and of course a time series that will constitute the forecasted output variable.

It is generally difficult to find data for the bulk carrier market that are not modified or interrupted in a time interval bigger than 10 years, due to changes that demand presents during this period. Therefore, the data were collected from two sources. The first source is the “*Clarksons Research Studies*” (*CRS*) which provides a statistical and research service to Clarkson brokers, their clients and the shipping world in general. The other source is the monthly magazine *Lloyd's Shipping Economist* that is published from 1979.

The data from *Clarksons* were collected online through the *Shipping Intelligent Network* database. On the other hand, the *Lloyd's Shipping Economist* magazines exist only in the Library of the Massachusetts Institute of Technology.

## 4.2 “Clarksons”

*Clarksons*, the world's leading shipping services provider, is a dynamic organization at the forefront of change within the industry, an industry estimated to be worth in excess of \$100 billion p.a. Through its global network of offices, it is able to provide its clients with unique access to ship-broking services across the full range of dry bulk, tanker and sale and purchase sectors. The group's dry cargo business, which is based in London, is the largest group of dry cargo chartering brokers in the world, whilst the tanker business is involved in the chartering of ships carrying crude oil and other petroleum products with these vessels ranging from 20,000 tons up to 500,000 tons. The latter operations operate globally from three main centers of activity, London, Houston and Singapore. The sale and purchase activities are involved in newly built ships, second hand vessels and demolition brokering.

More recently, the group has expanded its range of services so that it is now a fully integrated shipping services provider covering activities such as research and consultancy, shipping publications, shipping derivatives, shipping finance advice, shipping logistics and ship valuations. The logistics business comprises Channel Freight Ferries, which provides a daily service between Southampton and Radicatel and the more recently purchased MV Pacific Dhow, which will transport jet fuel to Kong Kong International Airport. Potentially of more importance to the group though is the futures/derivatives business - given the volatility of freight markets, industry participants have been slow to hedge their physical positions and so there is considerable scope for expansion in this market.

The *Shipping Intelligent Network* is an online database offering data for all kinds of vessels in daily, weekly, monthly and yearly base. From the above database the following time-series were extracted in monthly frequency:

- *120K DWT capesize Bulkcarrier Newbuilding Prices in \$ million*
- *Capesize Scrap value in \$ million*
- *1 year Bulkcarrier Timecharter Rates 127,500 dwt in \$/day*
- *Capesize ore Voyage Rates Tubarao/Rotterdam 145,000 dwt in \$/Ton*

### **4.3 “Lloyd's Shipping Economist”**

This magazine was published for the first time in February 1979 and it is refers to those who want to be aware of the developments in the shipping space. From the very first issue, it has a column for certain categories of vessels - one of them is the bulk carriers - that has a title “*STATISTICS*” and contains tables with data of the shipping market for the supply, demand, exports, imports etc.

From January 1981, the column was enriched with data for the prices (spot and charter market), the prices of new buildings, the prices of second-hand vessels and the scrap prices. The data for the supply and demand are presented in a centralized table called “*SUPPLY AND DEMAND DATA*”. The data referring to bulk-carriers were collected from those tables, and used for the current analysis of the bulk-carrier shipping market. It is important that the all data are collected from the same source so that they have a consistency.

The time series that were extracted from the magazine are the following:

- *Total supply of the Bulk carrier fleet in million DWT*

- *Total demand of the Bulk carrier fleet in million DWT*
- *Total Size of the Bulk carrier fleet that is in slow steaming situation, Surplus in million DWT*
- *Total number of Bulk carrier's order book in million DWT*

For all these time-series we have made the following assumptions:

- The data were not collected solely from *Clarksons* since the *Shipping Intelligent Network* does not provide any data for the supply, demand and surplus of the bulk carriers. The supply, demand and surplus were crucial data for the current analysis and could not be replaced by any other figures. Therefore, they were extracted from the *Lloyd's Shipping Economist* which has the same validity and uses the same assumptions with *Clarksons* to collect all these information. The same thing occurs to the total number of Bulk carrier's order book. *Clarksons* was providing numbers for this time series that were starting from 1995 and for the benefit of the analysis, past rates were required.
- The total number of Bulk carrier's order book is referred to periodical prices when the new buildings are ordered and recorded and not when the ships are delivered. This element should be taken into serious consideration, since there is an interval of about two years from the order to the delivery. Also the data for the order book includes new orders that took place with extremely good contract terms.

- The data for supply and surplus of the fleet in situation of slow steaming that is either laid-up or inactive are measured in tons of dead weight (dwt ton). The size of dead weight characterizes every commercial vessel. However it should not be confused with its pay load capacity. The dead weight is composed of: the pay load, the weight of fuels, the weight of supplies, the weight of lubricants, the weight of potable water and the crew weight. The ratio of pay load to dead weight depends a lot on the service range of every vessel. We must also point out that there is no relation between the dead weight and the weight of hull that is sold in the demolition facilities.

#### **4.4 Tables with data from “Clarksons” and the “Lloyd’s Shipping Economist”**

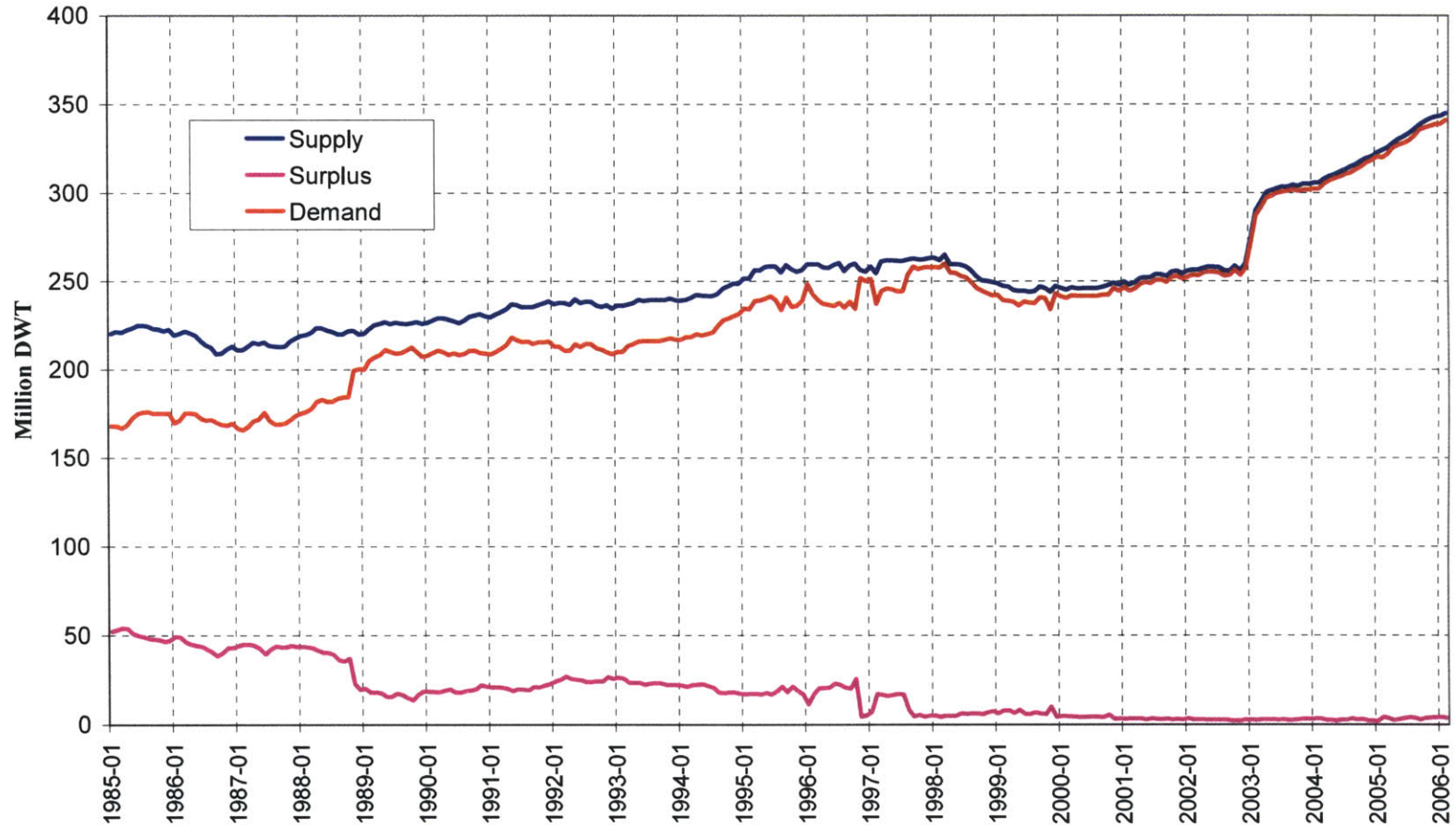
At this point examples with the tables are showed and are collected by Clarksons and Lloyds Shipping Economist. They can be categorized in three different groups.

## SUPPLY AND DEMAND DATA ON DRY BULK CARRIERS

	Date	Total Supply (Mill.DWT)	Total Surplus (Mill.DWT)	Total Demand (Mill.DWT)
1985	1985-01	220.2	52.2	168.0
	1985-02	221.3	53.2	168.1
	1985-03	221.0	54.2	166.8
	1985-04	222.4	53.8	168.6
	1985-05	223.3	50.9	172.4
	1985-06	224.8	49.9	174.9
	1985-07	225.0	49.3	175.7
	1985-08	224.4	48.2	176.2
	1985-09	223.1	48	175.1
	1985-10	222.8	47.5	175.3
	1985-11	221.7	46.6	175.1
	1985-12	222.6	47.3	175.3
2005	2005-01	322.90	2.1	320.8
	2005-02	324.42	4.3	320.1
	2005-03	325.81	3.7	322.1
	2005-04	328.19	2.5	325.7
	2005-05	330.31	3.1	327.2
	2005-06	331.92	3.7	328.3
	2005-07	333.58	4.2	329.4
	2005-08	335.93	3.8	332.1
	2005-09	338.39	2.8	335.6
	2005-10	340.38	3.7	336.7
	2005-11	341.84	4.1	337.8
	2005-12	342.97	4.2	338.8
2006	2006-01	343.35	4.3	339.0
	2006-02	344.87	3.8	341.1

*Table 4.1: Data of demand, supply and surplus of bulk carriers from 1985 until February 2006 as they have been collected from the Lloyd's Shipping Economist. (All data are shown in the appendix).*

**Figure 4.1:** Demand, Supply and Surplus of bulk Carriers from 1985 until February 2006.



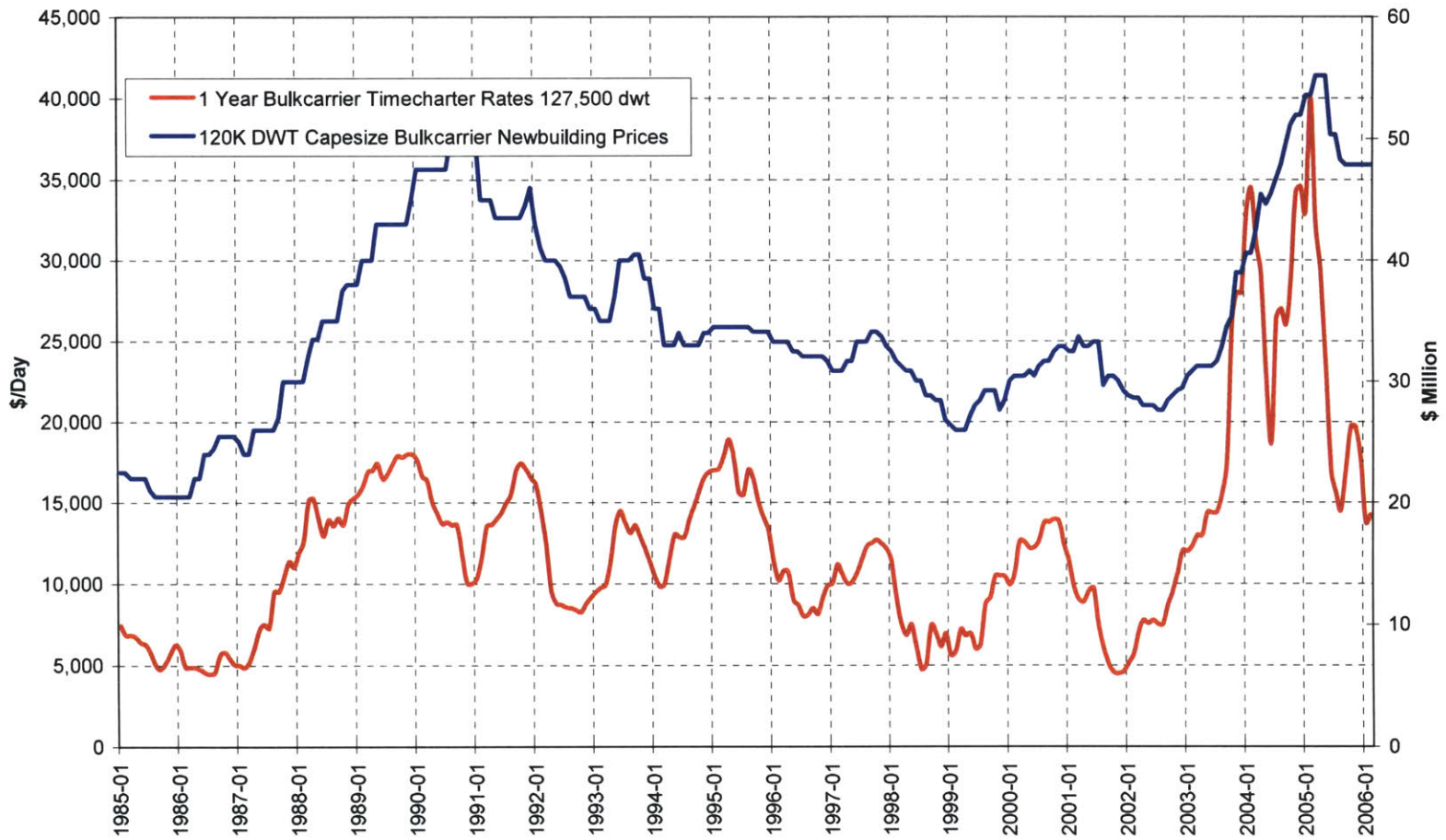
Source: Lloyd's Shipping Economist 1985-2006

The following tables shows the Timecharter prices for one year for a capesize bulk carrier of 127,500 DWT and the new building prices for a capesize bulk carrier of 120,000 DWT.

		RATES	
	Date	1 Year Bulkcarrier Timecharter Rates 127,500 dwt (\$/day)	120K DWT Capesize Bulk carrier Newbuilding prices (\$ Million)
1985	1985-01	7,438	22.5
	1985-02	6,850	22.5
	1985-03	6,850	22.0
	1985-04	6,725	22.0
	1985-05	6,375	22.0
	1985-06	6,250	22.0
	1985-07	5,750	21.0
	1985-08	5,000	20.5
	1985-09	4,760	20.5
	1985-10	5,125	20.5
	1985-11	5,725	20.5
	1985-12	6,250	20.5
2005	2005-01	33,000	53.58
	2005-02	40,000	53.58
	2005-03	32,500	55.2
	2005-04	29,400	55.2
	2005-05	23,500	55.2
	2005-06	17,000	50.33
	2005-07	15,600	50.33
	2005-08	14,500	48.30
	2005-09	17,100	47.89
	2005-10	19,750	47.89
	2005-11	19,625	47.89
	2005-12	17,700	47.89
2006	2006-01	13,750	47.89
	2006-02	14,250	47.89

**Table 4.2:** Data for one year bulk carrier Timecharter rates of 127,500 DWT and newbuilding prices of a 120K DWT bulk carrier from 1985 until February 2006 as they have been collected from Clarksons. (All data are shown in the appendix).

Figure 4.2: 1 Year Timecharter rates for 127,500 DWT and 120K DWT Capesize Newbuilding prices.



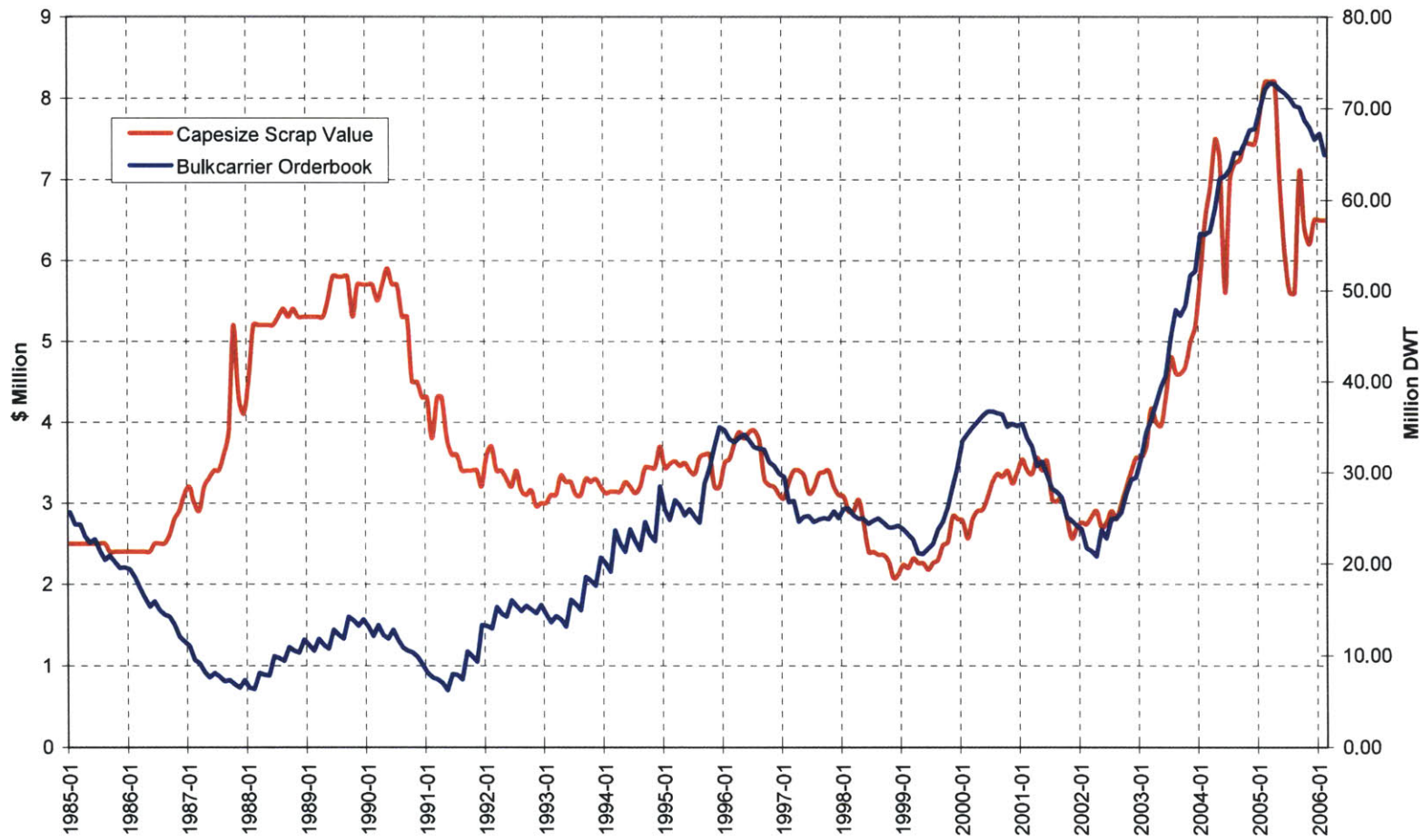
Source: Clarkson Research Studies

Moreover, the following table contains information about the capesize scrap value and the order book of new bulk carriers.

<b>PRICES</b>			
	<b>Date</b>	<b>Capesize Scrap Value (\$ Million)</b>	<b>Total Bulk carrier Order book (Mill. DWT)</b>
1985	1985-01	2.5	25.60
	1985-02	2.5	24.30
	1985-03	2.5	24.30
	1985-04	2.5	23.00
	1985-05	2.5	22.30
	1985-06	2.5	22.70
	1985-07	2.5	21.40
	1985-08	2.5	20.40
	1985-09	2.4	20.90
	1985-10	2.4	20.20
	1985-11	2.4	19.50
	1985-12	2.4	19.60
2005	2005-01	7.8	69.84
	2005-02	8.2	72.08
	2005-03	8.2	72.78
	2005-04	8.2	72.66
	2005-05	7	72.09
	2005-06	6.1	71.67
	2005-07	5.6	71.08
	2005-08	5.6	70.24
	2005-09	7.1	70.08
	2005-10	6.4	68.69
	2005-11	6.2	67.87
	2005-12	6.5	66.57
2006	2006-01	6.5	67.26
	2006-02	6.5	64.90

**Table 4.3:** Data for the capesize scrap value of a bulk carrier and the total order book of bulk carriers from 1985 until February 2006 as they have been collected from Clarksons and Lloyd's Shipping Economist. (All data are shown in the appendix).

Figure 4.3: Capesize Scrap Value and Bulk Carrier Orderbook.



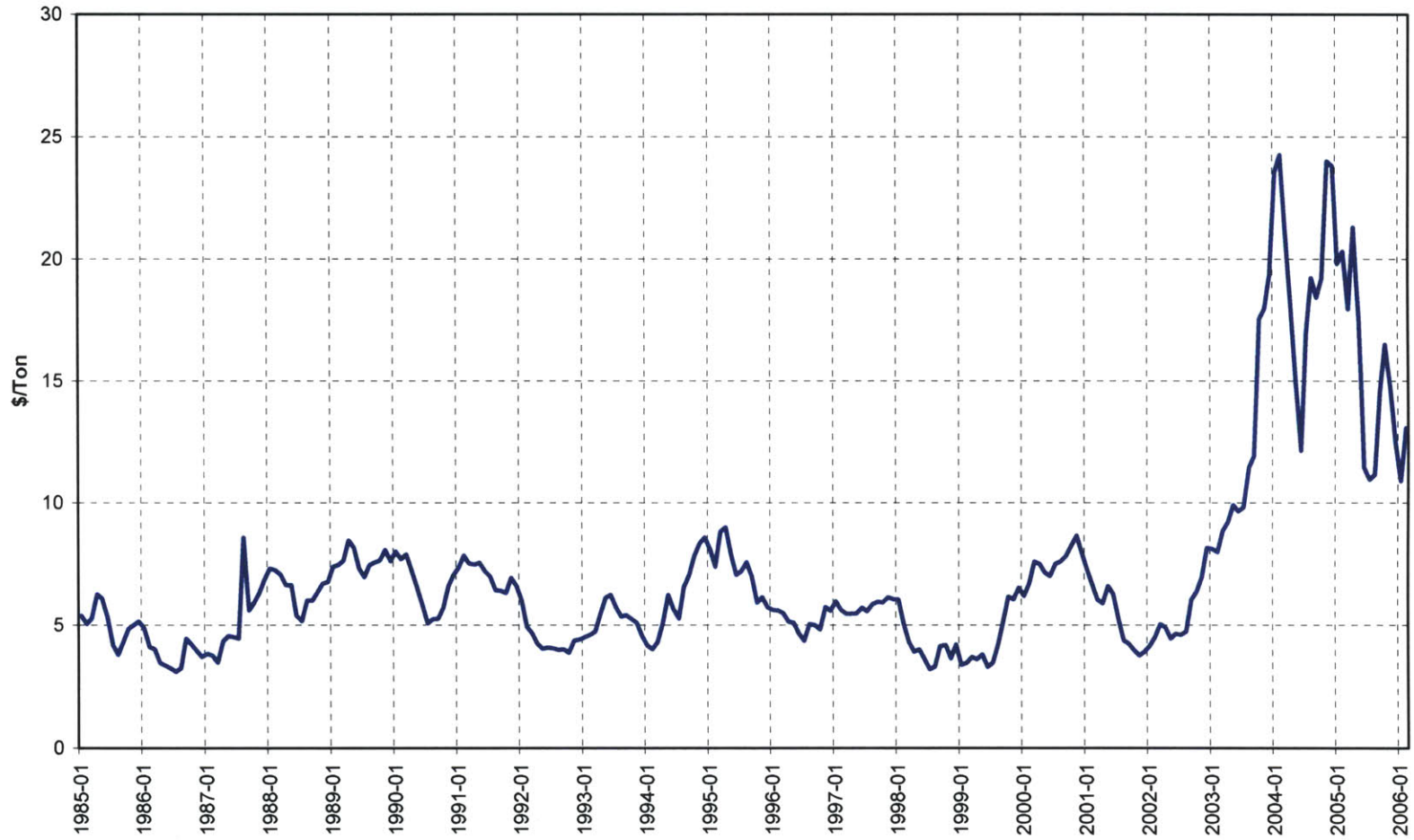
Source: Clarkson Research Studies and Lloyd's Shipping Economist

Finally, the last table shows the capesize ore voyage Rates from Tubarao to Rotterdam for a capesize bulk carrier of 145,000 DWT.

		<b>Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t (\$/Ton)</b>
	<b>Date</b>	
1985	1985-01	5.38
	1985-02	5.03
	1985-03	5.26
	1985-04	6.25
	1985-05	6.09
	1985-06	5.31
	1985-07	4.18
	1985-08	3.78
	1985-09	4.30
	1985-10	4.83
	1985-11	4.98
	1985-12	5.14
2005	2005-01	19.81
	2005-02	20.31
	2005-03	17.94
	2005-04	21.3
	2005-05	17.46
	2005-06	11.44
	2005-07	10.94
	2005-08	11.14
	2005-09	14.54
	2005-10	16.50
	2005-11	14.73
	2005-12	12.43
2006	2006-01	10.88
	2006-02	13.06

*Table 4.4: Data for the capesize ore voyage rates from Tubarao to Rotterdam for a 145,000 DWT bulk carrier 1985 until February 2006 as they have been collected from Clarksons. (All data are shown in the appendix).*

**Figure 4.4:** Capesize Ore Voyage Rates from Tubarao to Rotterdam for a 145,000 DWT Bulk Carrier.



Source: Clarkson Research Studies

#### 4.5 Available and exploitable Time series

From all the available data from the two sources, not all of them are usable. Having as a criterion the theoretical background of how the bulk carrier shipping market operates and the integrity of the past data, we have concluded to the following time series that we think can make the analysis and forecast of the capesize ore voyage rates more accurate. All the time series are numbered so that their use in the softwares *SPSS*, *EXCEL* and *NeuroSolutions* is more practical.

<b>Time series of Capesize ore rates</b>		
<b>No</b>	<b>Time series</b>	<b>Units</b>
1	<i>Capesize ore Voyage Rates</i> from Tubarao to Rotterdam 145,000t	Million Dollars per ton (\$/ Ton)

*Table 4.5a: Time series of capesize ore voyage rates with number (1).*

<b>Time series of Supply, Demand and Surplus</b>		
<b>No</b>	<b>Time series</b>	<b>Units</b>
2	Total <i>supply</i> of the Bulk carrier fleet.	Millions tones of dead weight (dwt)
3	Total Size of the Bulk carrier fleet that is in slow streaming situation, <i>Surplus</i>	Millions tones of dead weight (dwt)
4	Total <i>demand</i> of the Bulk carrier fleet.	Millions tones of dead weight (dwt)

*Table 4.5b: Time series of supply, demand and surplus with numbers (2-4).*

<b>Time series for Timecharter and New buildings</b>		
<b>No</b>	<b>Time series</b>	<b>Units</b>
5	1 Year Bulk carrier <i>Timecharter</i> Rates for 127,500 dwt	Thousand Dollars per day (\$/day)
6	120K DWT Capesize Bulk carrier <i>New building</i> Prices	Million Dollars (\$)

*Table 4.5c: Time series for Timecharter and New buildings with number (5-6).*

Time series for Scrap Value and Order book		
No	Time series	Units
7	Capesize <i>Scrap</i> Value	Million Dollars (\$)
8	Total Bulk carrier <i>Order book</i>	Millions tones of dead weight (dwt)

*Table 4.5d: Time series for Scrap value and order book with numbers (7-8).*

This is the total number of time series that will be analyzed and used for the forecast of the capesize voyage ore rates. **I will attempt to design and train an artificial neural network that will forecast the price of the first (1) variable, using data of entry, historical prices of the first time series along with the other seven time series.**

#### 4.6 The selection of independent time series

The *time series "1"* shows the capesize ore voyage rates from Tubarao to Rotterdam. For the final selection of the independent time series, a specific process is followed. This process is standard and can be implemented to any additional forecast. This process is completed in three stages:

- **Theoretical approach.** The forecast model is initially studied from a theoretical standpoint. The aim of this approach is to distinguish all the variables that can give information for forecasting *variable 1* through historical data.
- **Autocorrelation and cross-correlation with time lag.** From the variables that were selected from the theoretical approach, only the time series that finally offer information for the future price of the required time series are selected. This takes

place with the process of cross-correlation between dependent and independent variables along with the method of autocorrelation.

- **Reject strong correlated independent variables.** From the time series that were selected from the 2nd stage, they are rejected those who present strong cross-correlation with each other.

Afterwards, it is possible to build a table of entry data ( $n \times k$ ) for every forecast model, for every time period we wish to make forecasts.

#### **4.6.1 Theoretical approach**

From the 80 independent variables of various time series that Clarksons and Lloyd's Shipping Economist provide us for the bulk carrier's market, the 7 independent variables that were described in the previous tables are selected. For these variables, it is believed that they hide information for the future price of the capesize ore voyage rates (the one which we will attempt to forecast).

#### **4.6.2 Autocorrelation and cross-correlation with time lag**

In this stage we take advantage of how important "*time*" is. In other words, we have time series and not independent correlated data. In order for the future forecast to be valid, it is essential to cross-correlate the time series we will attempt to forecast with historical independent time series. At the end of this stage, we will be able to decide which of the seven time series present a strong cross-correlation with the actual rate of

the time series 1. For this reason we will use the method of *autocorrelation and cross-correlation* with time lags.

The diagrams that give us the autocorrelation function and the confidence limit result from “*SPSS 12.0 for Windows*”.

The confidence limit depends on the range of data that we have in our disposal for every correlation. Therefore, the control of historical rates of eighteen months has a bigger confidence limit, due to the existence of many extreme rates that are not utilizable for the process.

Autocorrelation is used for two major purposes: *first of all to track all those data that are correlated with each other and secondly to track the suitable data time series that will be used for an example model presupposed that the data are correlated with each other*. If the measures  $Y_1, Y_2, \dots, Y_N$  are known in the certain moments  $X_1, X_2, \dots, X_N$ , autocorrelation *ACF* is defined as:

$$ACF_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y}) \cdot (Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}$$

Even though the *variety of time X* isn't shown in the definition of autocorrelation, this variety is referred to equal periods of time, and to be more specific, in the model we study, it refers to a period of 18 months.

*Autocorrelation* is a correlation coefficient. However, instead of correlating two different varieties, we have a correlation between two different prices of the same variety  $Y_i$  during the periods of time  $X_i$  and  $X_{i+k}$ . If autocorrelation refers to a specific element of

the data, we are only interested in the first component of the series, which means that  $\kappa = 1$ , while if it refers to a time series, we are interested and for the other components of the  $r_k$  ( $\kappa > 1$ ).

Subsequently, we present the autocorrelations for the Capesize Ore voyage Rates from Tubarao to Rotterdam for a 145,00dwt vessel for the period of time between January 1985 and February 2006.

Given a time series, which is the white noise, the autocorrelation ACF has already being defined and the properties of the certain function (ACF) can be studied from the analysis of the specific model.

A way of analyzing this problem is to study the values of the ACF in a time series and to develop an error type, which will examine when a value of the ACF is significantly different to 0.

Theoretically all autocorrelation coefficients ACF of random number time series must be zero. But since samples are finite, autocorrelation coefficients are not zero. It has been proven that the autocorrelation coefficients ACF of a white noise model, are approximately normally distributed (mean is 0 and steady error equal to  $\frac{1}{\sqrt{n}}$  , where n is the number of data entries of the time series. That's why when we draw the graphic of an autocorrelation function (ACF) we also draw the line  $\pm 1.96/\sqrt{n}$ , while the boundaries are known as critical values.

The white noise model is described by the following equation.

$$Y_t = c + e_t$$

where  $Y_t$  is the variety, which is equal to a sum of a *constant*  $c$  with a *random error*  $e_t$ , which differs from time period to time period.

We notice, that by using the autocorrelation function (ACF), all the coefficient values are bigger than the critical value (in absolute terms)  $1.96/\sqrt{n}$ , which means that by *95% our data aren't white noise*.

*Partial Autocorrelation* is the autocorrelation between the  $Y_t$  and  $Y_{t-k}$  values, by not taking into account the other values between the above mentioned (values) - (1,2,3,...,k-1).

Partial Autocorrelation  $\Phi_{kk}$  is defined as:

$$\Phi_{kk} = \frac{r_k - \sum_{j=1}^{k-1} \Phi_{k-1,j} r_{k-j}}{1 - \sum_{j=1}^{k-1} \Phi_{k-1,j} r_j}$$

where  $r_k$  is the autocorrelation function. The range of the partial autocorrelation values fluctuates between the -1 and +1, with values *spotted* near the boundaries (-1 and +1) having a much more powerful correlation.

By using for one more time the critical value, which is derived by the  $1.96/\sqrt{n}$  function, we notice the following: it is confirmed that the variables aren't white noise; furthermore they tend to be correlated for at least one step of the method. Finally from the statistical procedure we can't say safely that we have seasonality of the varieties.

# ACF

Autocorrelations: caprates Capesize ore Voyage Rates Tubarao/Rotterd

Lag	Auto- Corr.	Stand. Err.	-1	-.75	-.5	-.25	0	.25	.5	.75	1	Box-Ljung	Prob.
1	.958	.062					.	*****				235.699	.000
2	.906	.062					.	*****				447.690	.000
3	.862	.062					.	*****				640.149	.000
4	.821	.062					.	*****				815.378	.000
5	.784	.062					.	*****				975.722	.000
6	.746	.062					.	*****				1121.584	.000
7	.717	.062					.	*****				1256.825	.000
8	.696	.062					.	*****				1384.926	.000
9	.687	.061					.	*****				1510.151	.000
10	.660	.061					.	*****				1626.397	.000
11	.619	.061					.	*****				1728.916	.000
12	.581	.061					.	*****				1819.470	.000
13	.535	.061					.	*****				1896.628	.000
14	.483	.061					.	*****				1959.871	.000
15	.419	.061					.	*****				2007.661	.000
16	.351	.060					.	*****				2041.265	.000
17	.298	.060					.	****				2065.629	.000
18	.265	.060					.	***				2085.042	.000

Plot Symbols: Autocorrelations \* Two Standard Error Limits .

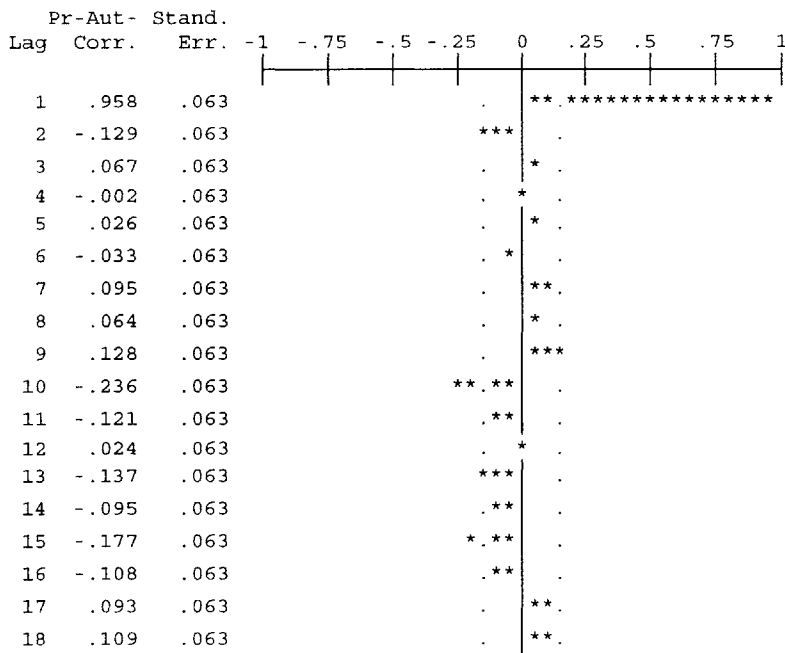
Total cases: 254 Computable first lags: 253

*Table 4.6: Autocorrelation for the variable "Capesize Ore Voyage Rates Tubarao/Rotterdam."*

The partial autocorrelations of the Capesize ore voyage rates follow.

## Partial Autocorrelations:

caprates Capesize ore Voyage Rates Tubarao/Rotter

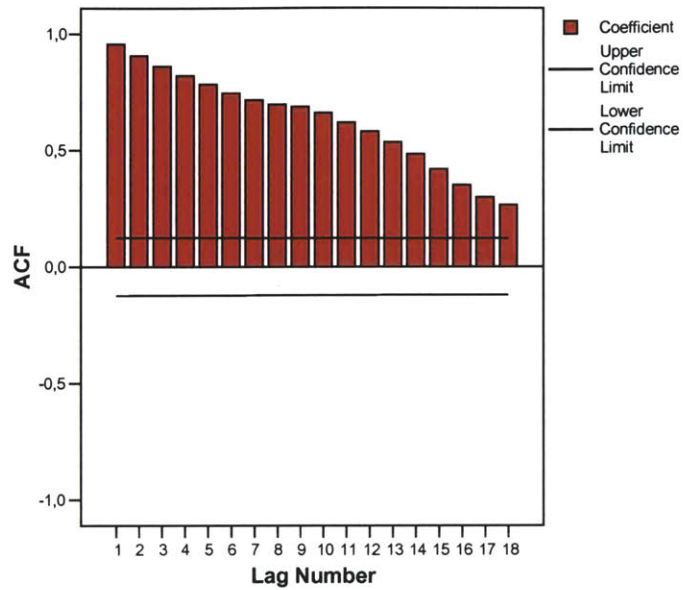


Plot Symbols: Autocorrelations \* Two Standard Error Limits .

Total cases: 254 Computable first lags: 253

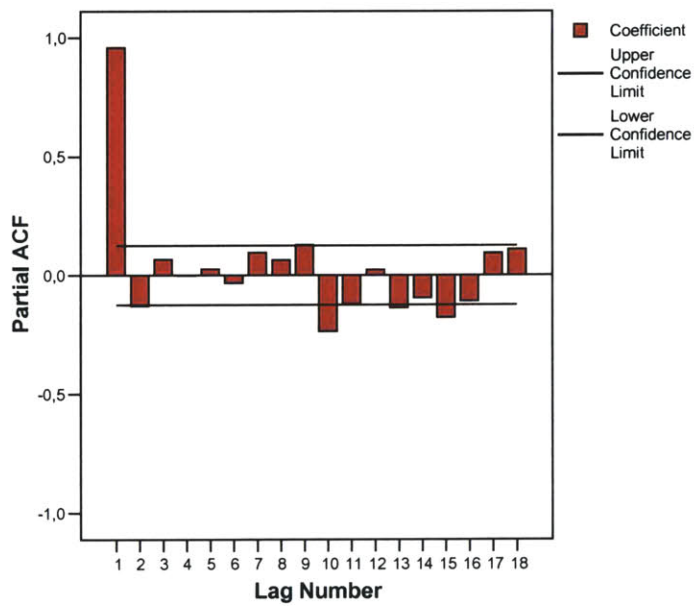
*Table 4.7: Partial autocorrelations of the capesize ore voyage rates.*

**Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t**



*Figure 4.5: Autocorrelation of time series that will be forecasted.*

**Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t**



*Figure 4.6: Partial autocorrelation of time series that will be forecasted.*

## 4.7 Pearson Correlation - Definition

The correlation between two variables portrays the degree that the variables are related with each other. The most usual meter of correlation is the Pearson Product Moment Correlation, which is generally called Pearson correlation. When the Pearson correlation is used in a “population” it is symbolized with the Greek letter  $\rho$ . On the contrary, when it is used in a “sample”, as the model we examine, it is symbolized with the Latin letter  $r$ .

### **Definition of Pearson correlation.**

For two given time series:  $p^x_t, p^y_t, t = 1, \dots, N$ , the Pearson correlation is defined as:

$$r_t = \frac{(N - 1)^2 \sum_i (p_i^x - \overline{p^x})(p_i^y - \overline{p^y})}{N^2 \sigma_x^n \sigma_y^n}$$

where  $\overline{p^x}$  is the mean of the time series and  $\sigma_x$  is the standard deviation. The same things also apply to the time series  $y$ .

The Pearson correlation depicts the degree of linear relation between two variables. It oscillates from -1 to +1. A correlation +1 means that there is a perfect positive linear relation between the variables. A relation is characterized as positive, when high prices of variable  $X$ , are accompanied by high prices of the other variable  $Y$ . A correlation -1 means that there is a perfect negative linear relation between the variables. A relation is characterized as negative, when high prices of variable  $X$ , are connected with low prices of the other variable  $Y$ . In addition, a 0 correlation means that there is no linear relation between the two variables, a non-zero price for the  $r$  or  $\rho$

*(except for the +1 and - 1) does not mean that the variables are dependent with each other. Everyday examples are likely to show important statistical prices for the  $r$  or  $\rho$ , but they are substantially insignificant.*

#### **4.7.1 Pearson correlations for the current time series**

The cross-correlations for all the time series that will be used are presented in the following table. The table contains the following time series:

- *Capesize ore Voyage Rates Tubarao/Rotterdam 145,000 dwt in \$/Ton*
- *Total supply of the Bulk carrier fleet in million DWT*
- *Total Size of the Bulk carrier fleet that is in slow steaming situation, Surplus in million DWT*
- *1 year Bulkcarrier Timecharter Rates 127,500 dwt in \$/day*
- *120K DWT capesize Bulkcarrier Newbuilding Prices in \$ million*
- *Capesize Scrap value in \$ million*
- *Total number of Bulk carrier's order book in million DWT*

*from January 1985 until February 2006.*

*To calculate the Pearson correlations we have used "SPSS 12.0 for Windows".*

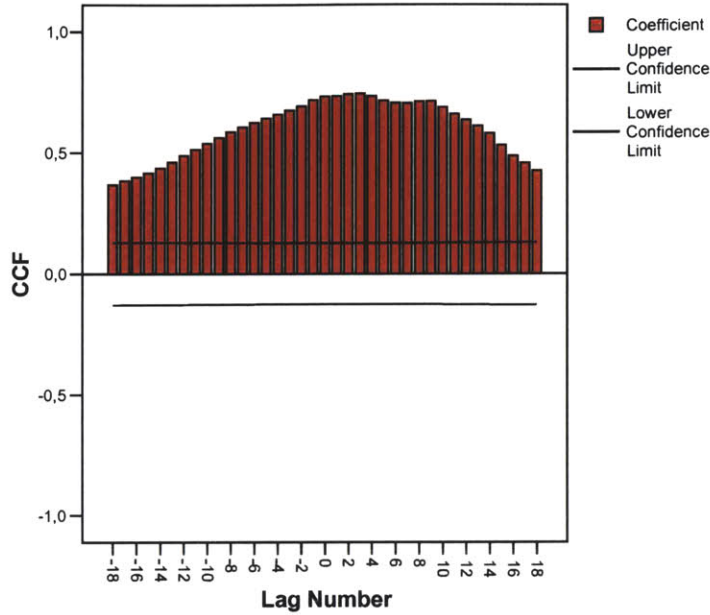
### Correlations

		120K DWT Capesize Bulkcarrier Newbuilding Prices	Capesize Scrap Value	Dry Bulkcarrier Total Supply	Dry Bulkcarrier Total Surplus	Bulkcarrier orderbook	1 Year Bulkcarr ier Timecha rter Rates 127,500 dwt	Capesize ore Voyage Rates Tubarao/Ro tterdam 145,000t
120K DWT Capesize Bulkcarrier Newbuilding Prices	Pearson Correlation	1	.793(**)	.462(**)	-.377(**)	.353(**)	.728(**)	.597(**)
	Sig. (2-tailed)	.	.000	.000	.000	.000	.000	.000
	N	254	254	254	254	254	254	254
Capesize Scrap Value	Pearson Correlation	.793(**)	1	.532(**)	-.236(**)	.508(**)	.822(**)	.792(**)
	Sig. (2-tailed)	.000	.	.000	.000	.000	.000	.000
	N	254	254	254	254	254	254	254
Dry Bulkcarrier Total Supply	Pearson Correlation	.462(**)	.532(**)	1	-.678(**)	.927(**)	.581(**)	.733(**)
	Sig. (2-tailed)	.000	.000	.	.000	.000	.000	.000
	N	254	254	254	254	254	254	254
Dry Bulkcarrier Total Surplus	Pearson Correlation	-.377(**)	-.236(**)	-.678(**)	1	-.586(**)	.369(**)	-.401(**)
	Sig. (2-tailed)	.000	.000	.000	.	.000	.000	.000
	N	254	254	254	254	254	254	254
Bulkcarrier orderbook	Pearson Correlation	.353(**)	.508(**)	.927(**)	-.586(**)	1	.575(**)	.761(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.	.000	.000
	N	254	254	254	254	254	254	254
1 Year Bulkcarrier Timecharter Rates 127,500dwt	Pearson Correlation	.728(**)	.822(**)	.581(**)	-.369(**)	.575(**)	1	.910(**)
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.	.000
	N	254	254	254	254	254	254	254
Capesize ore Voyage Rates Tubarao/Rott terdam 145,000t	Pearson Correlation	.597(**)	.792(**)	.733(**)	-.401(**)	.761(**)	.910(**)	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	.000	.
	N	254	254	254	254	254	254	254

\*\* Correlation is significant at the 0.01 level (2-tailed).

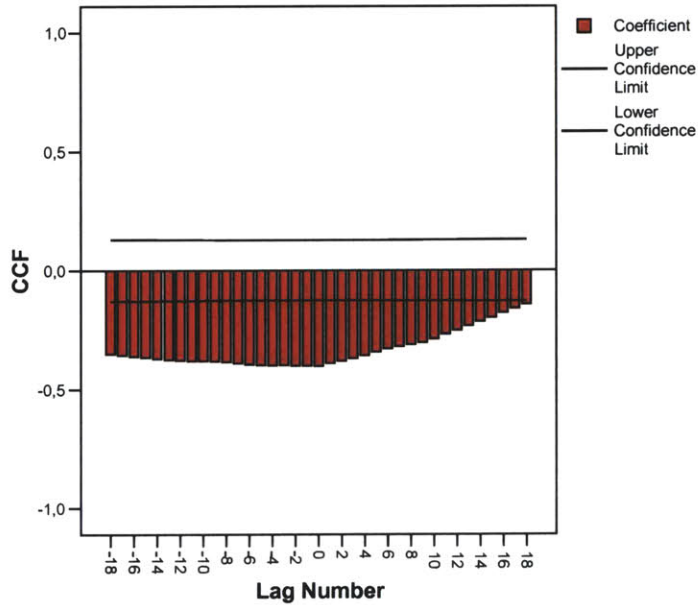
*Table 4.8: Pearson correlations for the 7 timeseries.*

**Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t with Dry Bulkcarrier Total Supply**



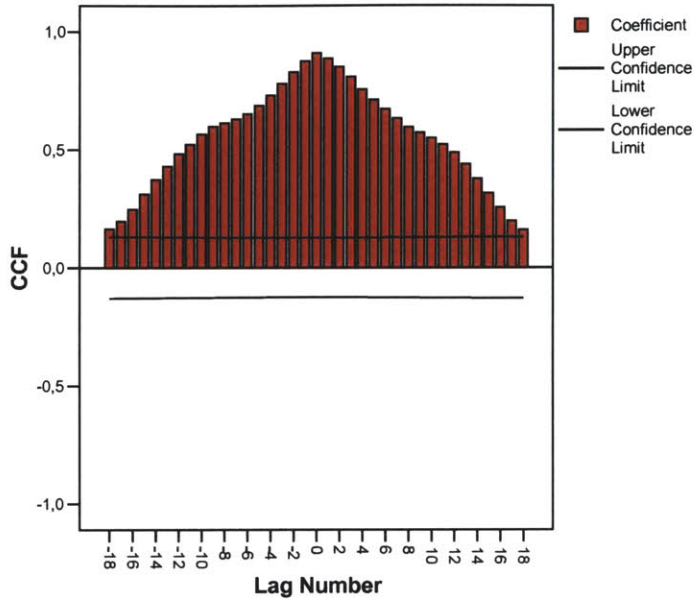
*Figure 4.7a: Cross-correlation of Capesize ore rates with Total Supply.*

**Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t with Dry Bulkcarrier Total Surplus**



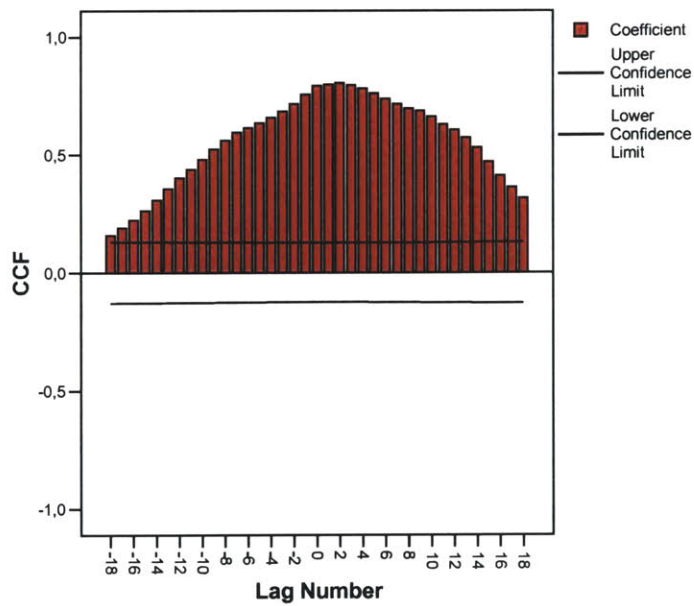
*Figure 4.7b: Cross-correlation of Capesize ore rates with Total Surplus.*

**Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t with  
1 Year Bulkcarrier Timecharter Rates 127,500dwt**



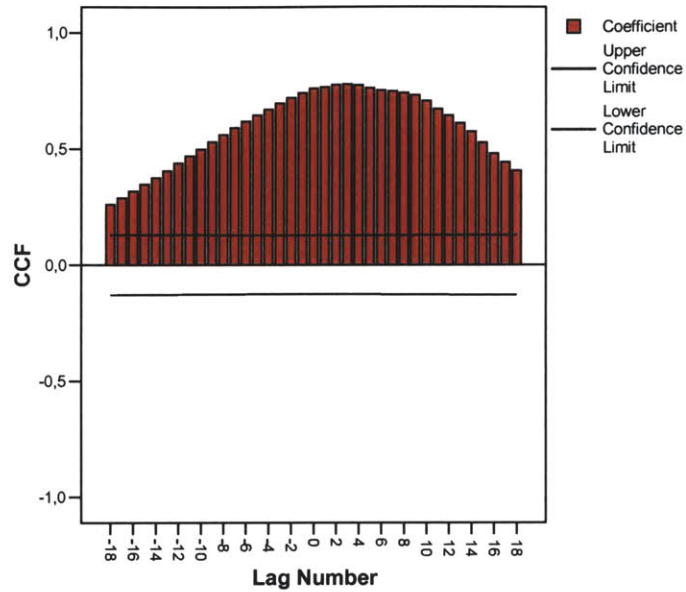
*Figure 4.7c: Cross-correlation of Capesize ore rates with 1 year Timecharter.*

**Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t with  
Capesize Scrap Value**



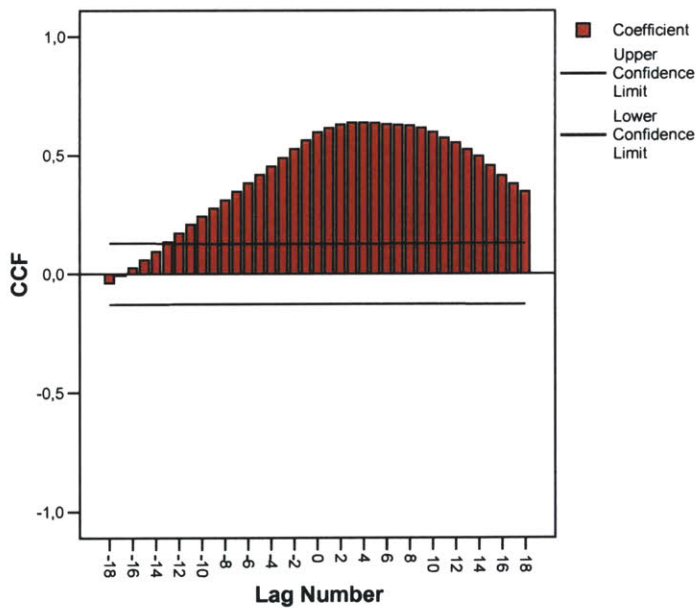
*Figure 4.7d: Cross-correlation of Capesize ore rates with Scrap Value.*

**Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t with Bulkcarrier orderbook**



*Figure 4.7e: Cross-correlation of Capesize ore rates with Total Order book.*

**Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t with 120K DWT Capesize Bulkcarrier Newbuilding Prices**



*Figure 4.7f: Cross-correlation of Capesize ore rates with New Building prices.*

#### **4.7.2 Conclusions**

*As we have expected there is a positive cross-correlation between the Capesize ore rates and the Total Supply, the 1-year Timecharter, the Scrap Value, the Order book and the New Building prices. On the contrary, there is a negative cross-correlation with the total Surplus.*

#### **4.7.3 Rejection of strong correlated independent variables**

From all the time series that were analyzed, only a few of the prices of some of the variables present a strong correlation with the capesize ore voyage rates time series. This confirms that our selection was accurate, while the *Clarksons' database* along with the *Lloyd's Shipping Economist's data* contained utilizable information. In addition, *neural networks* and mainly the *genetic algorithms (GA)* that will be used for the training of the forecasting model *have the ability to undermine and exclude data of entry that have strong correlation with each other and orient the model to undesirable outcomes. Therefore, any further attempt to manually reject any data is considered unnecessary, since it could lead to a loss of useful variables.*

#### **4.8 Evaluation of the database that will be used in the Neural Networks**

It is reminded that from the eight time series that we have initially selected to evaluate, we will only use the seven ones. The seven variables were correlated with the time series we will forecast, *the capesize ore voyage rates*, with a time lag of up to 18 months. The time series of *Total Demand* was rejected, since its data and information were included in the time series of *Total Supply and Total Surplus*.

The use of *genetic algorithms* optimizes the forecasting process, through the selection of optimum variables that will give the best results.

Finally, all the time series that will be used to forecast the *Capesize Ore Voyage Rates from Tubarao to Rotterdam for 145,000 DWT vessels* are the following:

#### **Forecasting of Capesize Ore Voyage Rates with 3, 6, 12 and 18 months delay:**

- *Capesize ore Voyage Rates Tubarao/Rotterdam 145,000 dwt in \$/Ton*
- *Total supply of the Bulk carrier fleet in million DWT*
- *Total Size of the Bulk carrier fleet that is in slow streaming situation, Surplus in million DWT*
- *1 year Bulkcarrier Timecharter Rates 127,500 dwt in \$/day*
- *120K DWT capesize Bulkcarrier Newbuilding Prices in \$ million*
- *Capesize Scrap value in \$ million*
- *Total number of Bulk carrier's order book in million DWT*

#### **4.9 Import of data in the program NeuroSolutions 4.0**

In order for the data to be inserted to the Neural Network, they should be first formed accordingly in Excel. We create nine different columns as they appear in the following table:

Date	Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t	120K DWT Capesize Bulk carrier New building Prices	Capesize Scrap Value	Dry Bulk carrier Total Supply	Dry Bulk carrier Total Surplus	Bulk carrier Order book	1 Year Bulk carrier Timecharter Rates 127,500 dwt	Capesize ore Rates 145,000t+3 or +6 or +12 or +18 months
1985-01	5.38	22.5	2.5	220.2	52.2	25.60	7,438	6.25
2006-02	13.06	47.89	6.5	344.87	3.8	64.90	14,250	-

*Table 4.9: Excel columns with eight time series. The eighth column represents the desirable forecasting time series with 3, 6, 12 and 18 months delay.*

Then, a new *active worksheet* will be created that will contain all the above variables in **Randomized Rows**. *This step is crucial for the best training of the networks. In other words, the neural networks will not be confused and untuned by the sequence of the data, but it will rather try to optimize the solution and minimize the forecasting error using random data.* The next step is to suitably tag all the time series so that the networks can read them correctly. The columns that will be used as entry data (*Columns as Inputs*) are the following:

- *Capesize ore Voyage Rates Tubarao/Rotterdam 145,000 dwt in \$/Ton*
- *120K DWT capesize Bulkcarrier Newbuilding Prices in \$ million*
- *Capesize Scrap value in \$ million*
- *Total supply of the Bulk carrier fleet in million DWT*
- *Total Size of the Bulk carrier fleet that is in slow steaming situation, Surplus in million DWT*
- *Total number of Bulk carrier's order book in million DWT*
- *1 year Bulkcarrier Timecharter Rates 127,500 dwt in \$/day*

The column that will be used as desirable data (Columns as Desired) is:

- *Capesize ore Voyage Rates Tubarao/Rotterdam 145,000 dwt in \$/Ton with 3, 6, 12 and 18 months delay.*

*Finally, we have 250 lines of data, of which 230 are selected for Training, 20 of them are selected for Cross Validation and the remaining for testing the network in unknown rates.*

Hence, the database format appears in the following table:

	Capesize Rates	Capesize Newbuilding Prices	Capesize Scrap	Bulkcarrier Supply	Bulkcarrier Surplus	Bulkcarrier Orderbook	Timecharter Rates	Capesize Rates+threemonth
2003-04	9.21	31.25	4	300.45	3.07	37.59033	13000	<b>9.81</b>
1994-02	4	36	3.14	239.5	21.2	19.1	9825	<b>6.23</b>
1997-06	5.71	33.28	3.12	261.4	17.2	25.23632	10713	<b>5.94</b>
1987-06	4.5	26	3.4	215.4	39.6	8.1	7500	<b>5.59</b>

*Rows as Training*

1990-02	7.68	47.5	5.7	227.8	18.4	12.1	16600	<b>6.5</b>
1993-06	6.23	40	3.26	238.8	22.5	16.1	14500	<b>5.4</b>
2002-05	4.44	28.01	2.7	258.1	2.9	23.71248	7550	<b>4.74</b>
1994-07	5.26	33	3.12	241.7	20.8	22.5	12890	<b>7.83</b>
1990-11	6.57	50	4.5	231.5	22.2	9.8	10050	<b>7.83</b>
1998-05	4	30.85	2.78	259.7	4.9	24.95804	7500	<b>3.31</b>

*Rows as Cross Validation*

2005-12	12.43	47.89	6.5	342.97	4.15	66.56706	17700	
2006-01	10.88	47.89	6.5	343.35	4.33	67.25721	13750	
2006-02	13.06	47.89	6.5	344.87	3.82	64.90428	14250	

*Rows as Production*

**Table 4.10:** Database with the three different types of data.

We export the following files that will be used in the construction of the four neural networks.

- *Training Input*
- *Training Desired*
- *Cross Validation Input*
- *Cross Validation Desired*
- *Testing Input*
- *Testing Desired*
- *Production Input*

# Chapter 5

## Selection of the Artificial Neural Network

### 5.1 Introduction

The selection of the artificial neural network that will be used is not defined by certain rules. In the bibliography there are only indications and conclusions of past trials.

In this particular project, we aim, for the first time, to use **Modular Neural Networks**, which constitute from parallel Multilayer Perception models and use a combination of **genetic algorithms** to optimize the input data. The topology of every network is selected from four different choices so that the operation of each sub-module is further strengthened.

In the beginning, the number of hidden layers will be decided, then the number of neurons that every layer will have and finally we will select the transfer function and the training algorithms. However, the process is not so fixed, since all the parameters are weighted by the use of genetic algorithms.

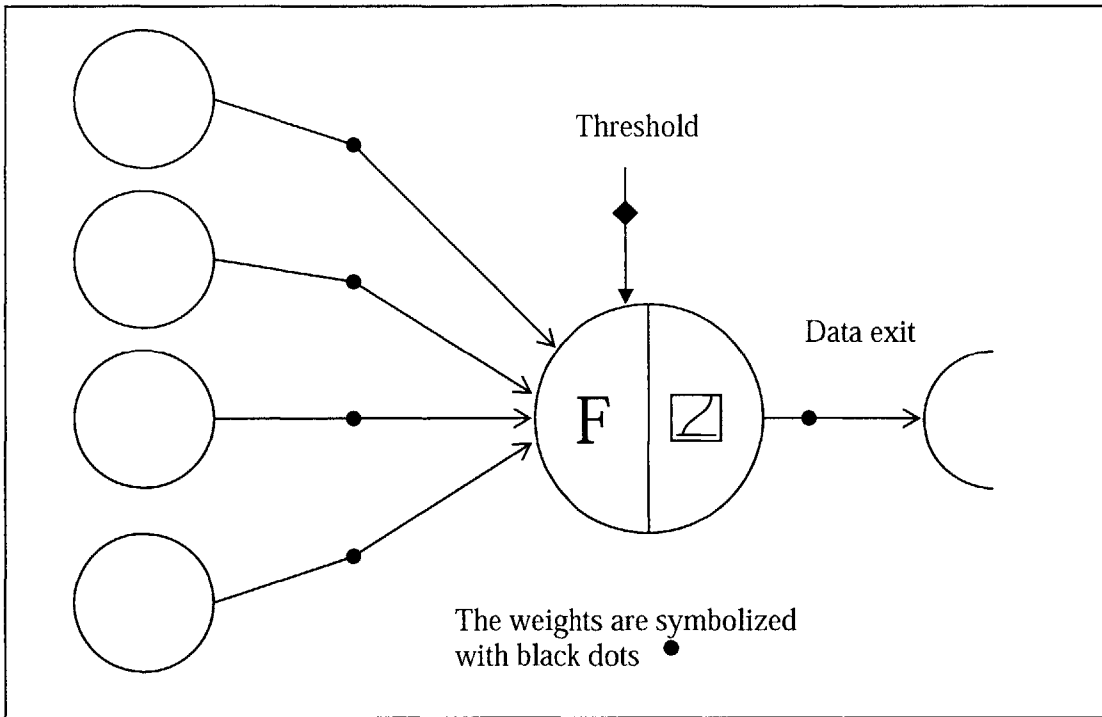
### 5.2 Layers - Number of Neurons – Transfer Function

When the neural networks are used for the forecasting of financial figures, three and very rarely four layers are chosen. The neurons of a network are distributed across layers.

There are three types of layers the *input layer*, the *hidden layers*, and the *output layer*. Each network has exactly one input and one output layer. The number of hidden layers can vary from 0 to any number. The input layer is the only layer that does not contain transfer functions. Also, all of them apart from the input layer consist of neurons. The number of inputs to the neural network equals to the dimension of our input samples, while the number of the outputs we want from the neural network defines the number of neurons in the output layer. In our case the output layer will have exactly one neuron since the only output we want from the network is the forecasting of tomorrow's capesize ore voyage rates. The mass of hidden layers as well as the mass of neurons in each hidden layer is proportional to the ability of the network to approximate more complicated functions. Of course this does not imply by any means that networks with complicated structures will always perform better. The reason for this is that the more complicated a network is the more sensitive it becomes to noise or else, it is easier to learn apart from the underlying function the noise that exists in the input data. Therefore it is clear that there is a trade off between the representational power of a network and the noise it will incorporate.

Lippmann claims that a neural network with only one hidden layer gives fairly good results. As far as the number of neurons is concerned, he proposes that it must be at least one time more than the twofold of the input data. In other words, if we have  $N$  input data, the neurons of every hidden layer must be  $(2N + 1)$ . However, various observations have shown that many neural networks present better results when they are structured like a "bottle neck". To make things more comprehensible, the second hidden layer of the network will have less neurons than first one, while the network will progressively end up

with only one neuron in the final layer. Other researchers propose that the size of the sample must be dependable of the number of weights, something which can only be applied if the database is very big and does not have any restrictions.



**Figure 5.1:** An analytical description of a neuron.

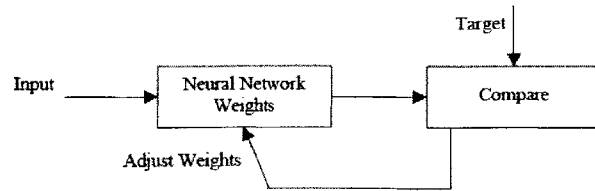
Each *neuron* apart from the number of its inputs is characterized by the function  $f$  known as *transfer function*. The most commonly used transfer functions are: the *hard limit*, the *pure linear*, the *sigmoid* and the *tan sigmoid* function. The preference on these functions derives from their characteristics. *Hard limit* maps any value that belongs to  $(-\infty, +\infty)$  into two distinct values  $\{0, 1\}$ , thus it is preferred for networks that perform classification tasks. *Sigmoid* and *tan sigmoid*, known as squashing functions, map any value from  $(-\infty, +\infty)$  to the intervals  $[0, 1]$  and  $[-1, 1]$  respectively. Lastly *pure linear* is used due to its ability to return any real value and is mostly used at the neurons that are

related with the output of the network. Researchers have come to the conclusion that the use of linear transfer functions drastically decreases the ability of the network to be efficiently trained, which mostly applies to the shipping market which does not behave in a linear way. The transfer functions with which the network presents the best results were the sigmoid and tan sigmoid, and in this case we will use the tan sigmoid.

### 5.3 Weights Adjustment - Error Function - Training Algorithms

The power of neural networks models lies in the way that their *weights* (inter unit-connection strengths) is adjusted. The procedure of adjusting the weights of a neural networks based on a specific dataset is referred to as the training of the network on that set (*training set*). The basic idea behind training is that the network will be adjusted in a way that will be able to learn the patterns that lie in the *training set*. Using the adjusted network in future situations (unseen data) it will be able based on the patterns that learnt to generalize giving us the ability to make inferences. In our case we will train neural networks models on a part of our time series (*training set*) and we will measure their ability to generalize on the remaining part (*test set*). The size of the test set is usually selected to be 10% of the available samples.

The way that a network is trained is depicted by the following figure. Each sample consists of two parts the input and the target part (*supervised learning*). Initially the weights of the network are assigned random values (usually within $[-1,1]$ ). Then the input part of the first sample is presented to the network. The network computes an output based on: the values of its weights, the number of its layers and the type and mass of neurons per layer.



*Figure 5.2: The training procedure of a Neural Network.*

This output is compared with the target value of the sample and the weights of the network are adjusted in a way that a metric that describes the distance between outputs and targets is minimized.

There are two major categories of network training the *incremental* and the *batch* training. During the *incremental* training the weights of the network are adjusted each time that each one of the input samples are presented to the network, while in *batch* mode training the weights are adjusted only when all the training samples have been presented to the network. The number of times that the training set will be fed to the network is called *number of epochs*.

Issues that arise and are related to the training of a network are: what exactly is the mechanism by which weights are updated, when does this iterative procedure cease, which metric is to be used to calculate the distance between targets and outputs?

The *error function* or the *cost function* is used to measure the distance between the targets and the outputs of the network. The weights of the network are updated in the direction that makes the error function minimum. The most common error functions are the *MSE* and the *MAE*. In our case study the networks will be trained and tested using the *MSE* function.

The mechanism of weights update is known as *training algorithm*. There are several training algorithms proposed in the literature. We will give a brief description of

those that are related with the purposes of our study. The algorithms described here are related to *feed-forward* networks. A neural network is characterized as *feed-forward* network “if it is possible to attach successive numbers to the inputs and to all of the hidden and output units such that each unit only receives connections from inputs or units having a smaller number”. All these algorithms use the gradient of the *cost function* to determine how to adjust the weights to minimize the *cost function*. The gradient is determined using a technique called backpropagation, which involves performing computations backwards through the network. Then the weights are adjusted in the direction of the negative gradient.

## 5.4 Stop Training

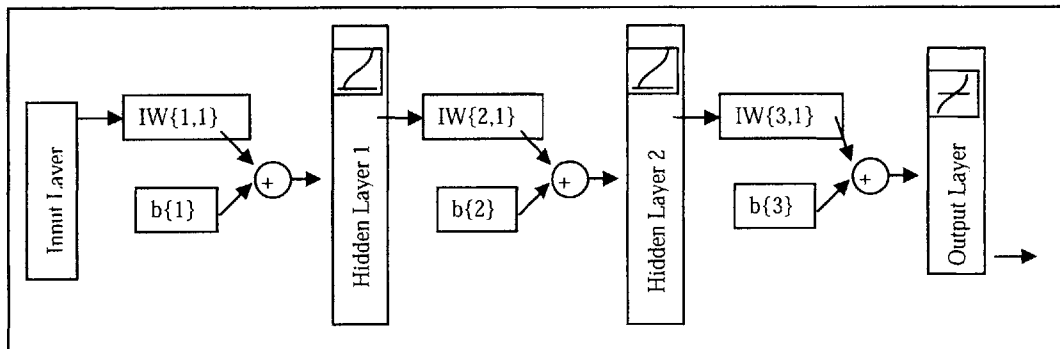
A significant decision related with the training of a neural network is the time on which its weight adjustment will be ceased. Over-trained networks become over-fitted to the training set and they are useless in generalizing and inferring from unseen data. While under-trained networks do not manage to learn all the patterns in the underlying data and due to this reason under perform on unseen data. Therefore there is a tradeoff between over-training and under-training our networks.

The methodology that is used to overcome this problem is called *validation* of the trained network. Apart from the *training set* a second set, the *validation set*, which contains the same number of samples, is used. The weights of the network are adjusted using the samples in the *training set* only. Each time that the weights of the network are adjusted its performance (in terms of error function) is measured on the *validation set*. During the initial period of training both the errors on *training* and *validation* sets are

decreased. This is due to the fact that the network starts to learn the patterns that exist in the data. From a number of iterations of the training algorithm and beyond the network will start to over-fit to the training set. If this is the case, the error in the validation set will start to rise. In the case that this divergence continues for a number of iterations the training is ceased. The output of this procedure would be a not over-fitted network.

After describing the way that a neural network works and the parameters that are related to its performance we select these parameters in a way that will allow us to achieve optimum performance in the task we are aiming to accomplish.

Therefore, the basic structure of our neural network will be as the one appeared in the following table.



*Figure 5.3: Basic structure of the neural network that will be used.*

## 5.5 Construction of the Neural Network

Below, we present analytically the construction of the network and every of its parameters that will be used.

### 5.5.1 Selection of the type of the Neural Network

The process begins with selecting the type of the neural network we will use. Based on previous research and references in the bibliography, we will attempt for the first time, to use a **Modular Multilayer Perceptron** model which is a sub-category of a **Multilayer Perceptron** model.

*Multilayer perceptrons (MLPs)* are layered feed-forward networks typically trained with static *backpropagation*. These networks have found their way into countless applications requiring static pattern classification. Their main advantage is that they are easy to use, and that they can approximate any input/output map. The key disadvantages are that they train slowly, and require lots of training data (typically three times more training samples than network weights).

On the contrary, *Modular feed-forward networks* are a special class of MLP. These networks process their input using several parallel MLPs, and then recombine the results. This tends to create some structure within the topology, which will foster specialization of function in each sub-module. In contrast to the MLP, modular networks do not have full interconnectivity between their layers. Therefore, a smaller number of weights are required for the same size network (i.e. the same number of processing elements, neurons). This tends to speed up training times and reduce the number of required training exemplars.

### 5.5.2 Import of Training Data

In this stage the training data are imported in the network. The data were analyzed and extracted from Excel as column-formatted ASCII files. The training data are:

- *Capesize Rates*
- *Capesize New building Prices*
- *Capesize Scrap*
- *Bulk carrier Supply*
- *Bulk carrier Surplus*
- *Bulk carrier Order book*
- *Timecharter Rates*

with *3, 6, 12 or 18 months delay* depending on the forecasting period that our network is built for.

At this point, we also select the use of *genetic algorithms (GA)*, which means that an input might be skipped or not, based on the performance of multiple training runs. The combination of inputs that produces the lowest error across these training runs will be used for the final model. The *genetic algorithm* that is being used is a conventional genetic algorithm which is inspired by the mechanism of natural selection where stronger individuals are likely the winners in a competing environment.

A *conventional genetic algorithm* has three major components. The first component is related with the creation of an *initial population* of  $m$  randomly selected individuals. The *initial population* shapes the first *generation*. The second component inputs  $m$  individuals and gives as output an evaluation for each of them based on an objective function known as *fitness function*. This evaluation describes how close to our demands each one of these  $m$  individuals is. Finally the third component is responsible for the

formulation of the next generation. A new generation is formed based on the fittest individuals of the previous one. This procedure of evaluation of generation N and production of generation N+1 (based on N) is iterated until a performance criterion is met. The creation of offspring based on the fittest individuals of the previous generation is known as *breeding*. The *breeding* procedure includes three basic genetic operations: *reproduction*, *crossover* and *mutation*.

*Reproduction* selects probabilistically one of the fittest individuals of generation N and passes it to generation N+1 without applying any changes to it. On the other hand, *crossover* selects probabilistically two of fittest individuals of generation N; then in a random way chooses a number of their characteristics and exchanges them in a way that the chosen characteristics of the first individual would be obtained by the second and vice versa. Following this procedure creates two new offspring that both belong to the new generation. Finally the *mutation* selects probabilistically one of the fittest individuals and changes a number of its characteristics in a random way. The offspring that comes out of this transformation is passed to the next generation. The way that a conventional GA works by combining the three components described above is depicted in the following flowchart.

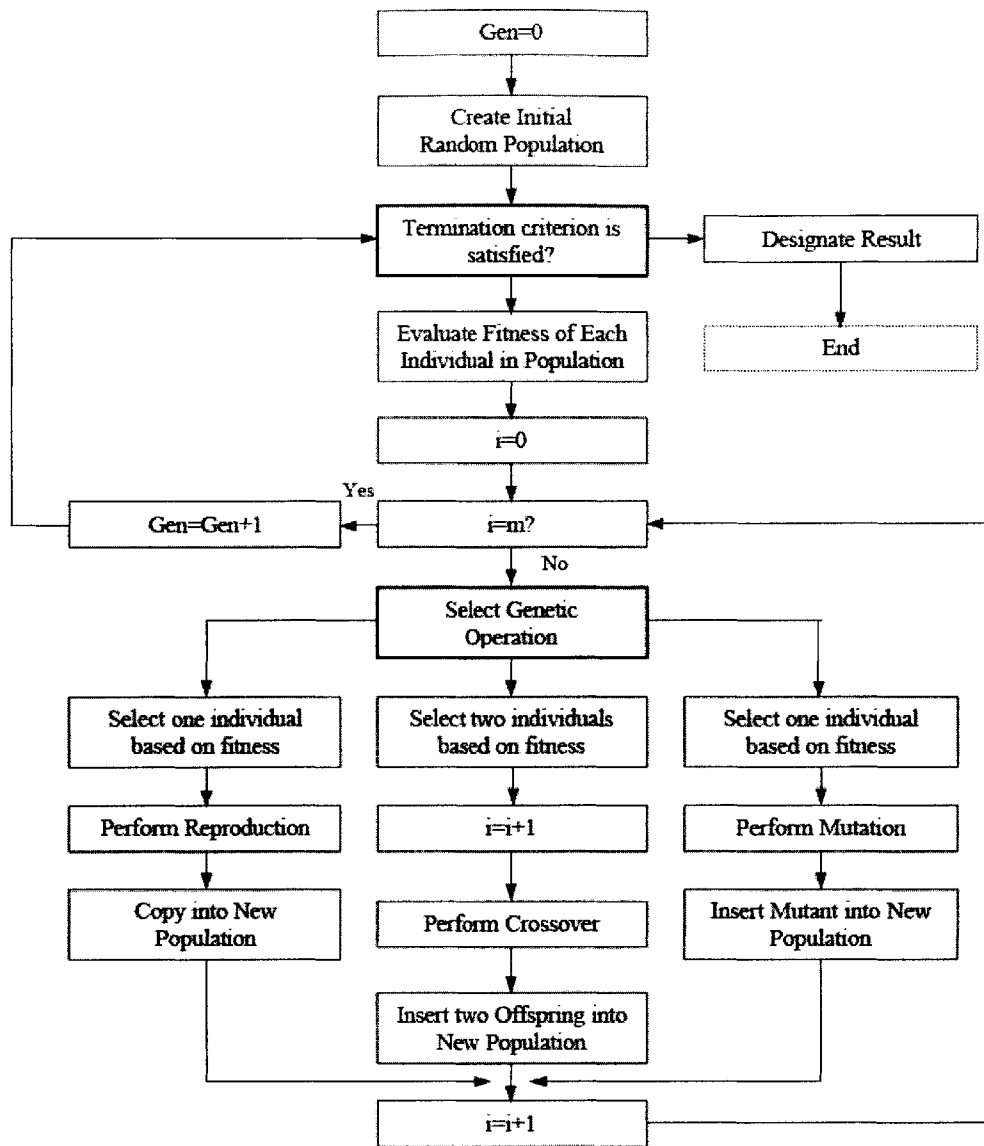


Figure 5.4: A conventional Genetic Algorithm.

As it has been stated each one of the individuals has a certain number of characteristics. For these characteristics the term *genes* is used. Furthermore according to the biological paradigm the set of all *genes* of an individual form its *chromosome*. Thus each individual is fully depicted by its *chromosome* and each generation can be fully described by a set of  $m$  *chromosomes*.

It is clear from the flowchart of the GA that each member of a new generation comes either from a *reproduction*, *crossover* or *mutation* operation. The operation that will be applied each time is selected based upon a probabilistic schema. Each one of the three operations is related with a probability  $P_{reproduction}$ ,  $P_{crossover}$ , and  $P_{mutation}$  in a way that

$$P_{reproduction} + P_{crossover} + P_{mutation} = 1$$

Therefore the number of offspring that come from reproduction, crossover or mutation is proportional to  $P_{reproduction}$ ,  $P_{crossover}$ , and  $P_{mutation}$  respectively.

Relative to the way that the selection of an individual (or two in the case of crossover) according to its fitness is done, again the selection is based on a probabilistic method. The selection is implemented by a scheme known in literature as *roulette wheel*. In GAs the higher the fitness value of an individual the better the individual.

Based upon this fact a *roulette wheel* is created by the following steps:

- Place all population members in a specific sequence
- Sum the fitness off all population members  $F_{sum}$ .
- Generate a random number ( $r$ ) between 0 and  $F_{sum}$ .
- Return the first population member whose fitness value added to the fitness of the preceding population members, is greater than or equal to ( $r$ ).

In case we want to select two individuals (crossover) we create the *roulette wheel* twice, the first time using all fitness values and the second time using all apart from the fitness value that corresponds to the chromosome selected from the firstly created *roulette wheel*. This guarantees us that we will not crossover between the same

chromosomes, which would mean that the crossover operation would be equivalent to a reproduction operation twice on the same chromosome.

The termination criterion is the number of evolution cycles (generations), which in this project is defined as 100 generations.

### **5.5.3 Import of desired response**

In this stage we provide the network with the desired response, since the network requires supervised training. Once again, the desired response was analyzed and extracted from Excel as column-formatted ASCII files. The desired response for our network is:

- *Capesize Rates with 3, 6, 12 or 18 months delay*

The desired columns cannot be tagged for genetic optimization. In every supervised learning process, the desired output values are known for each input pattern. At each instant of time, when an input pattern is applied to an artificial neural network, the parameters of the neural network are adapted according to the difference between the desired value and the neural network output.

### **5.5.4 Import of cross validation and test data**

*Cross validation* is a highly recommended method for stopping network training. The cross validation set is used to determine the level of generalization produced by the training set. Cross validation is executed in concurrence with the training of the network. Every so often, the network weights are frozen, the cross validation data is fed through

the network, and the results are reported. The stop criteria of the controller can be based on the error of the cross validation set instead of the training set to insure this generalization. The *testing set* is used to test the performance of the network. Once the network is trained the weights are then frozen, the testing set is fed into the network and the network output is compared with the desired output.

Once again, the cross validation and test data were analyzed and extracted from Excel as column-formatted ASCII files. For the cross validation data we have two separate files, one for the input data and one for the desired response. The cross validation files for input and desired response for our network are:

- *Capesize Rates*
- *Capesize New building Prices*
- *Capesize Scrap*
- *Bulk carrier Supply*
- *Bulk carrier Surplus*
- *Bulk carrier Order book*
- *Timecharter Rates*
  
- *Capesize Rates for 3, 6, 12 or 18 months delay (desired response)*

with *3, 6, 12 or 18 months delay* depending on the forecasting period that our network is built for.

At this point, we will not provide any test data to the network, but this will happen when the network is trained and the weights have taken their final rates. Then we will test and evaluate the performance of the network to the unknown rates of *March and April 2006*.

### 5.5.5 Selection of the number of hidden layers, their parameters and the network topology

*Multilayer perceptrons* are an extension of Rosenblatt's perceptron, a device that was invented in the '50s for optical character recognition. The perceptron only had an input and an output layer (each with multiple processing elements). It was shown that the perceptron would only solve pattern recognition problems where the classes could be separated by hyper planes (an extension of a plane for more than two dimensions). A lot of problems in practice do not fit this description. Multilayer perceptrons (MLPs) extend the perceptron with hidden layers, i.e., layers of processing elements that are not connected to the external world.

There are two important characteristics of the multilayer perceptron. First, its *processing elements* (PEs) are nonlinear. The nonlinearity function must be smooth (the logistic function and the hyperbolic tangent are the most widely utilized). Second, *they are massively (fully) interconnected* such that any element of a given layer feeds all the elements of the next layer.

The perceptron and the multilayer perceptron are trained with error correction learning, which means that the desired response for the system must be known. This is normally the case with pattern recognition. Error correction learning works in the following way: From the system response at PE  $i$  at iteration  $n, \dots$ , and the desired response for a given input pattern an instantaneous error is defined by:

$$e_i(n) = d_i(n) - y_i(n)$$

Using *the theory of gradient descent learning*, each weight in the network can be adapted by correcting the present value of the weight with a term that is proportional to the present input at the weight and the present error at the weight:

$$w_{ij}(n+1) = w_{ij}(n) + n \cdot \delta_i(n) \cdot x_j(n)$$

The local error can be directly computed from  $e_i(n)$  at the output PE or can be computed as a weighted sum of errors at the internal PEs. The constant  $h$  is called the step size. This procedure is called the *backpropagation algorithm*.

Backpropagation computes the sensitivity of the output with respect to each weight in the network, and modifies each weight by a value that is proportional to the sensitivity. The beauty of the procedure is that it can be implemented with local information and is efficient because it requires just a few multiplications per weight. However, since it is a gradient descent procedure and only uses the local information, it can get caught in a local minimum. Moreover, the procedure is a little noisy since we are using a poor estimate of the gradient, so the convergence can be slow.

*Momentum learning* is an improvement to the straight gradient descent in the sense that a memory term (the past increment to the weight) is utilized to speed up and stabilize convergence. In momentum learning the equation to update the weights becomes:

$$w_{ij}(n+1) = w_{ij}(n) + n \cdot \delta_i(n) \cdot x_j(n) + a \cdot (w_{ij}(n) - w_{ij}(n-1))$$

where  $a$  is the momentum. Normally  $a$  should be set between 0.1 and 0.9.

The training can be implemented in two ways: Either we present a pattern and update the weights (on-line learning), or we present all the patterns in the input file (an

epoch), store the weight update for each pattern, and then update the weights with the average weight update (batch learning). They are equivalent theoretically, but the former sometimes has advantages in tough problems (ones with many similar input-output pairs).

To start backpropagation, an initial value needs to be loaded for each weight (normally a small random value), and proceed until some stopping criteria is met. The three most common criteria are:

1. *The number of iterations*
2. *The mean square error of the training set*
3. *The mean squared error of the cross validation set.*

**Cross validation** is the most powerful of the three since it stops the training at the point of optimal generalization (i.e., the error of the cross validation set is minimized). To implement cross validation, one must put aside a small part of the training data (10% is recommended) and use it to determine how well the trained network is learning. When the performance starts to degrade in the cross validation set, the training should be stopped.

In this project, we will use *two hidden layers*, while the topology of our network will be a **modular multi-layer Perceptron model**. **Modular multi-layer Perceptron networks** are a special class of MLPs. These networks process their input using several parallel MLPs, and then recombine the results. This tends to create some structure within the topology, which will foster specialization of function in each sub-module.

Since modular networks *do not have full interconnectivity* between their layers, a smaller number of weights are required for the same size network (i.e. the same number

of processing elements). This tends to speed up training times and reduce the number of required training exemplars.

Neural networks are vector based for efficiency. This implies that each layer contains a vector of processing elements and that the *parameters* selected apply to the entire vector. The parameters are dependent on the neural model, but all require a nonlinearity function to specify the behavior of the processing elements.

Every layer has an associated *learning rule and learning parameters*. Learning from the data is the essence of neuron-computing. Every processing element that has an adaptive parameter must change it according to some pre-specified procedure. Momentum is by far the most common form of learning. Here it is sufficient to say that the weights are changed based on their previous value and a correction term. The learning rule is the means by which the correction term is specified. Once the momentum rule is selected, we specify how much correction should be applied to the weights, referred to as the learning rate. If the learning rate is too small, then learning takes a long time. On the other hand, if it is set too high, then the adaptation diverges and the weights are unusable.

Therefore, we have set *seven processing elements*, since we have seven input processing elements, *momentum as a learning rule* and a *tan sigmoid transfer function* with which the network presents the best results.

In order to make sure that the network will try to find the optimum setting for the above parameters, we will also use *genetic algorithms*. Genetic algorithms will be applied in every parameter, so that optimization is achieved. On the other hand, this form of optimization requires that the network is trained multiple times in order to find the

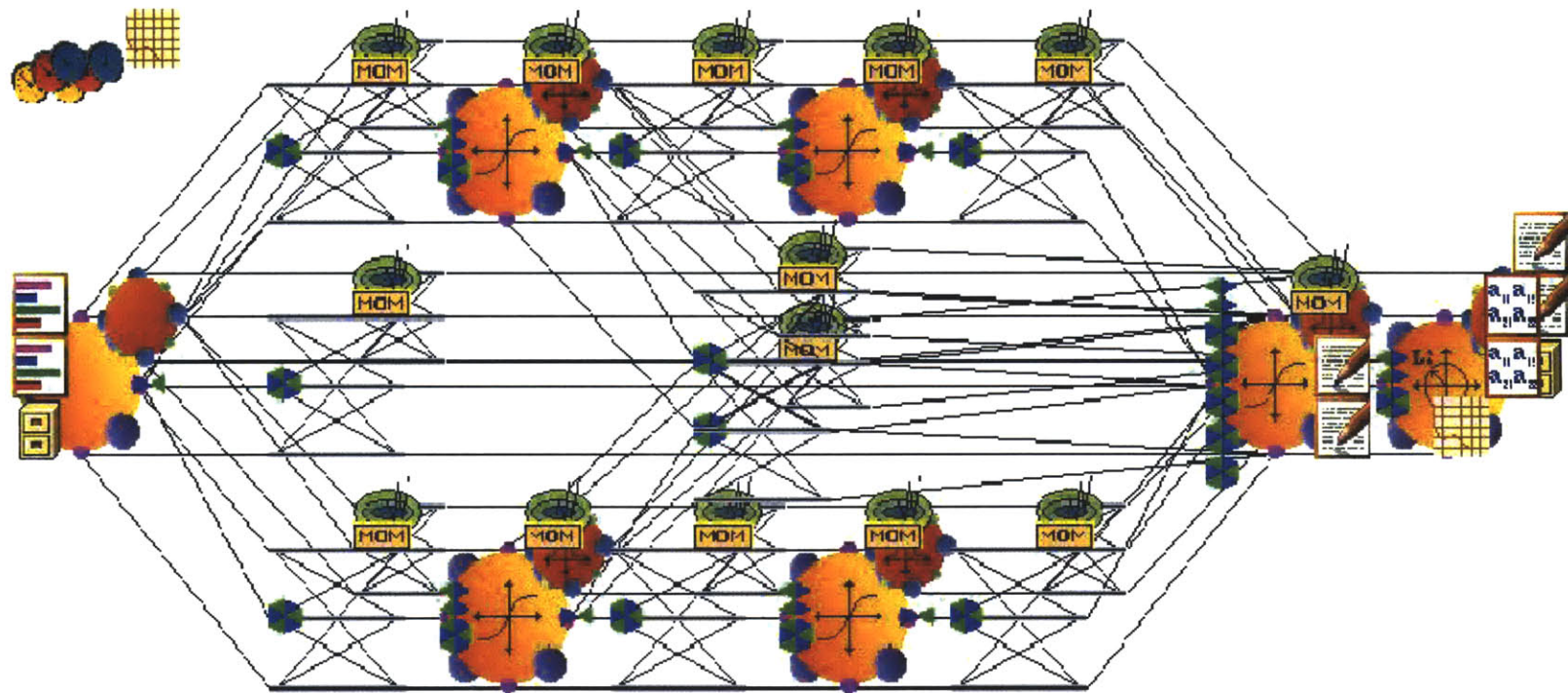
settings that produce the lowest error. The combination of parameter settings that produces the lowest error across these training runs will be used for the final model.

Finally, the *Output Layer* panel is the same as the previous panel except that the number of PEs is fixed to the number of output, which in our case is only one. The default Step Size is one magnitude smaller (set to 0.1) than the previous layer. This is because the error attenuates as it is backpropagated through the network. Since the error is largest towards the output of the network, the output layer requires a smaller step size than that of the hidden layer in order to balance the weight updates.

#### **5.5.6 Termination of the training**

In the final stage, we set the *Maximum Epochs*, in other words the number of training iterations. We set this number to 1,000 Maximum Epochs, even though the network may learn the problem (i.e., have a small error) in fewer iterations than this.

We use the cross validation set to calculate the *MSE* and decide when to terminate the training process. As we have mentioned earlier, this tends to be a good indicator of the level of generalization that the network has achieved. At the beginning the MSE is constantly falling, but when the MSE starts again to increase the training process is terminated. This is an indication that the network has begun to over-train. Overtraining is when the network simply memorizes the training set and is unable to generalize the problem. When the network is being trained, we can then test it to a series of unknown data and evaluate it accordingly. The network presents the following figure:



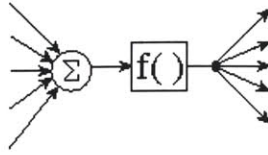
*Figure 5.5: Example of how neurons and hidden layers are arranged in two modules.*

## 5.6 Analytical explanation of the neural network topology

Artificial neural networks are constructed by interconnecting processing elements (PEs) which mimic the biological nerve cell, or neuron. Our network divides the functionality of a neuron into two disjoint operations: *a nonlinear instantaneous map*, which mimics the neuron's threshold characteristics; and *a linear map applied across an arbitrary discrete time delay*, which mimics the neuron's synaptic interconnections. The *Axon* implements common variations on the nonlinear instantaneous maps employed by neural models. *Each axon represents a layer, or a vector of PEs*. All axons are equipped with a summing junction at their input and a splitting node at their output. This allows multiple components to feed an axon, which then processes their accumulated activity. For generality, an axon's map may actually be either linear or nonlinear. However, components in the Axon family typically apply a nonlinear instantaneous map, as given by:

$$y_i(t) = f(x_i(t), w_i)$$

where  $y_i(t)$  is the axon's output,  $x_i(t)$  is an accumulation of input activity from other components,  $w_i$  is an internal weight or coefficient and  $f: R^n \rightarrow R^n$  represents an arbitrary functional map. We call  $f: R^n \rightarrow R^n$  the activation function. All Axons accumulate input from, and provide output to, an arbitrary number of activation components. In other words, each axon has a summing junction at its input and a splitting node at its output. This functionality is illustrated by the following diagram:



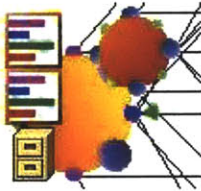
**Figure 5.6:** The mapping for the PE of every Axon.

Each of the axons and the hidden layers in the previous figure constitute a vital piece for the neural network topology. Starting from the left side of the network we will describe it in a more quantitative and coherent way.

### 5.6.1 Layer No 1

This *Axon (Layer No 1)* simply performs an identity map between its input and output activity. The activation function of this axon is

$$f(x_i, w_i) = x_i$$



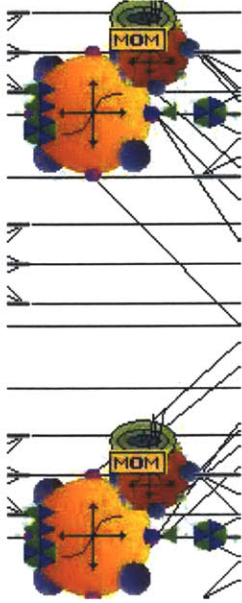
In the beginning, this axon has seven processing elements (neurons), since there are seven input vectors. Consequently, it will also have seven output vectors. With the use of *genetic algorithms*, the input and output vectors have been modified after the termination of the training process, to optimize the performance of the network.

### 5.6.2 Hidden layer No 1

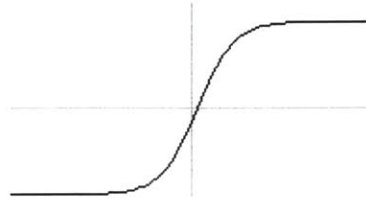
The *TanhAxon (hidden layer No 1)* applies a tanh function to each neuron in the layer. This will squash the range of each neuron in the layer to between -1 and 1. Such nonlinear elements provide a network with the ability to make soft decisions. The activation function of this axon is

$$f(x_i, w_i) = \tanh[x_{i,i}^{lin}]$$

where  $[x_{i,i}^{lin}] = \beta \cdot x_i$  is the scaled and offset activity inherited from the LinearAxon. The LinearAxon implements a linear axon with

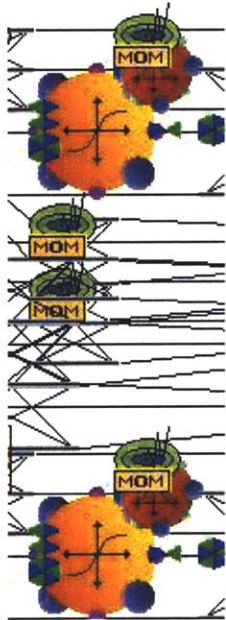


slope and offset control. The slope is controlled by an additional parameter  $b$ , which is not adaptive. The transfer function is shown in the following figure:



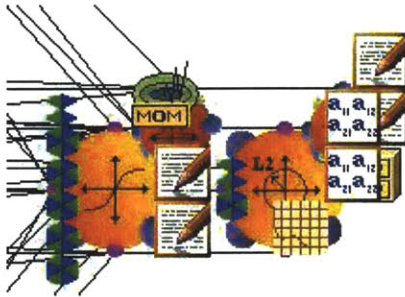
Once again this axon initially contains seven processing elements (neurons), since there are seven input vectors coming from the first layer. But with the use of genetic algorithms, the input and output vectors are modified after the termination of the training process so that network optimization is being achieved.

### 5.6.3 Hidden layer No 2



Respectively, the second hidden layer is a *TanhAxon* with the same activation function and works in the exact same way that the hidden layer No 1 works. As we have already mentioned, modular networks do not have full interconnectivity between their layers. Therefore, a smaller number of weights are required for the same size network (i.e. the same number of processing elements, neurons). This tends to speed up training times and reduce the number of required training exemplars.

#### 5.6.4 Layer No 2



The last *TanhAxon* (*layer No 2*) contains only one **neuron**, one input and one output processing element, which is the desired output of the capesize ore voyage rates. At this point of the network's topology, the network is being supervised and the performance of the network compared to the desired response is measured.

Supervised learning requires a metric, a measure of how the network is doing. **Error Criteria** monitor the output of a network, compare it with some desired response and report any error to the appropriate learning procedure. In gradient descent learning, the metric is determined by calculating the sensitivity that a cost function has with respect to the network's output. This **cost function  $J$** , is normally positive, but should decay towards zero as the network approaches the desired response. The literature has presented several cost functions, but the quadratic cost function is by far the most widely applied. The cost function is defined as:

$$J(t) = \frac{1}{2} \sum (d_i(t) - y_i(t))^2$$

where  $d(t)$  and  $y(t)$  are the desired response and network's output, respectively. The error function is defined as:

$$\varepsilon_i(t) \equiv \frac{\partial J(t)}{\partial y_i(t)}$$

Each Error Criteria component accepts its desired response through the Desired Signal access point, and reports the total cost between weight updates to the Average

Cost access point. Error Criteria components are responsible for determining an error that is used by the backpropagation plane to calculate the gradient information.

The neural network implements supervised learning procedures using component planes. Each component used for implementing the activation plane has a single dual component that is used to implement the backpropagation plane. Components of the backpropagation plane are responsible for computing the weight gradients and backpropagating the sensitivities. Error Criteria components are responsible for determining the error used for the backpropagation.

## 5.7 Performance Measuring

During the network's training, the Error Criterion component provides six values that are used to measure the performance of the network to a particular data set.

The six different values are:

- **Mean Squared Error "MSE"**

The mean squared error is simply two times the average cost. The formula for the mean squared error is:

$$MSE = \frac{\sum_{j=0}^P \sum_{i=0}^N (d_{ij} - y_{ij})^2}{N \cdot P}$$

where  $P = \text{number of output processing elements}$

$N = \text{number of exemplars in the data set}$

$y_{ij}$  = network output for exemplar  $i$  at processing element  $j$

$d_{ij}$  = desired output for exemplar  $i$  at processing element  $j$

- **Nominalized Mean Squared Error “NMSE”**

To Nominalized mean squared error is defined by the following formula:

$$NMSE = \frac{P \cdot N \cdot MSE}{\sum_{j=0}^P \frac{N \cdot \sum_{i=0}^N d_{ij}^2 - \left( \sum_{i=0}^N d_{ij} \right)^2}{N}}$$

where  $P$  = number of output processing elements

$N$  = number of exemplars in the data set

$MSE$  = mean squared error

$d_{ij}$  = desired output for exemplar  $i$  at processing element  $j$

- **Correlation Coefficient “r”**

The size of the mean square error (MSE) can be used to determine how well the network output fits the desired output, but it doesn't necessarily reflect whether the two sets of data move in the same direction. For instance, by simply scaling the network output, we can change the MSE without changing the directionality of the data. The **correlation coefficient  $r$**  solves this problem. By definition, the correlation coefficient between a *network output  $x$*  and a *desired output  $d$*  is:

$$r = \frac{\sum_i (x_i - \bar{x}) \cdot (d_i - \bar{d})}{\sqrt{\frac{\sum_i (d_i - \bar{d})^2}{N}} \cdot \sqrt{\frac{\sum_i (x_i - \bar{x})^2}{N}}}$$

where  $N$  = number of exemplars in the data set

$x_i = \text{network output}$

$d_i = \text{desired output for exemplar } i$

The correlation coefficient is confined to the range [-1, 1]. When  $r = 1$  there is a perfect positive linear correlation between  $x$  and  $d$ , they co-vary, which means that they vary by the same amount. When  $r = -1$ , there is a perfectly linear negative correlation between  $x$  and  $d$ , they vary in opposite ways (when  $x$  increases,  $d$  decreases by the same amount). When  $r = 0$  there is no correlation between  $x$  and  $d$ , i.e. the variables are called uncorrelated. Intermediate values describe partial correlations. For example a correlation coefficient of 0.88 means that the fit of the model to the data is reasonably good.

- **Percentage Error “% Error”**

The percent error is defined by the following formula:

$$\%Error = \frac{100}{N \cdot P} \cdot \sum_{j=0}^P \sum_{i=0}^N \frac{|dy_{ij} - dd_{ij}|}{dd_{ij}}$$

where  $P = \text{number of output processing elements}$

$N = \text{number of exemplars in the data set}$

$dy_{ij} = \text{denormalized network output for exemplar } i \text{ at processing element } j$

$dd_{ij} = \text{denormalized desired output for exemplar } i \text{ at processing element } j$

On the other hand, it is important to mention that this value can easily be misleading. For example, say that our output data is in the range of 0 to 100. For one

exemplar our desired output is 0.1 and our actual output is 0.2. Even though the two values are quite close, the percent error for this exemplar is 100.

- **Akaike's Information Criterion "AIC"**

Akaike's information criterion (AIC) is used to measure the tradeoff between training performance and network size. The goal is to minimize this term to produce a network with the best generalization:

$$AIC(k) = N \cdot \ln(MSE) + 2k$$

where  $k = \text{number of network weights}$

$N = \text{number of exemplars in the training set}$

$MSE = \text{mean squared error}$

- **Minimum description length "MDL"**

Rissanen's minimum description length (MDL) criterion is similar to the AIC in that it tries to combine the model's error with the number of degrees of freedom to determine the level of generalization. The goal is to minimize this term:

$$MDL(k) = N \cdot \ln(MSE) + 0.5 \cdot k \cdot \ln(N)$$

where  $k = \text{number of network weights}$

*N = number of exemplars in the training set*

*MSE = mean squared error*

We have presented the topology and the parameters of the artificial neural network model we will use to forecast the capesize ore voyage rates. The network is now ready to forecast time series using, from them beginning to the end of the process, an automated artificial model, which can be adapted to any problem.

# Chapter 6

## Neural Networks Results

### 6.1 Introduction

After systematic research and further analysis of the topology of the network, we will present the results obtained from the four neural networks. My objective is to present the forecasting of the *Capesize ore Voyage Rates from Tubarao to Rotterdam for 145,000 dwt vessel with 3, 6, 12 and 18 months delay* in figures and diagrams and evaluate their results accordingly.

### 6.2 Results Presentation

The results will be divided in three main parts:

- *Training Data*
- *Cross-Validation Data*
- *Testing Unknown Data*

### 6.2.1 Training Data

Training is the process by which the free parameters of the network (i.e. the weights) get optimal values. With supervised learning, the network is able to learn from the input and the error (the difference between the output and the desired response). The ingredients for supervised learning are therefore the input, the desired response, the definition of error, and a learning law. Error is typically defined through a cost function. Good network performance should result in a small value for the cost. A learning law is a systematic way of changing the weights such that the cost is minimized. In supervised learning the most popular learning law is backpropagation.

The network is trained in an attempt to find the optimal point on the performance surface, as defined by the cost definition. Backpropagation changes each weight of the network based on its localized portion of the input signal and its localized portion of the error. The change has to be proportional (a scaled version) of the product of these two quantities. When this algorithm is used for weight change, the state of the system is doing gradient descent; moving in the direction opposite to the largest local slope on the performance surface. In other words, the weights are being updated in the direction of down.

In addition, genetic algorithms are used to skip inputs that cross-correlate with each other or deteriorate the performance of the network.

We present the results from the training process of the network in *Figures 6.1-6.4*. The training inputs vary in total from **216 to 231 inputs** in all four neural networks. This represents a percentage of *85-90% of the total data available*, which are very satisfactory

for the training process. Last but not least, all the training inputs are organized in a random arrangement, so that the network and its weights do not converge in one area of inputs, but consider the problem as random events.

### 6.2.2 Cross-Validation Data

Cross validation computes the error in a test set at the same time that the network is being trained with the training set. It is known that the MSE will keep decreasing in the training set, but may start to increase in the test set. This happens when the network starts “*memorizing*” the training patterns.

A practical way to find a point of better generalization is to set aside a small percentage (*around 10-15%*) of the training set and use it for cross validation. One should monitor the error in the training set and the validation set. When the error in the validation set increases, the training should be stopped because the point of best generalization has been reached. Cross validation is one of the most powerful methods to stop the training.

Using that percentage of *10-15% of the training set*, we have selected **20 training inputs as cross-validation data**. We present the results from the cross-validation process of the network in the *Figures 6.5-6.12*.

Interview

### 6.2.3 Testing Unknown Data

Once we have trained the network and we have determined that the network adequately models our data, we want to test it on data that the network was not trained

with. In doing so, we will only have input data but no desired data introduced to the model. The desired data will be used afterwards, to evaluate the outputs.

For this purpose, we will use the *Capesize ore voyage rates for March and April 2006*. In *Figures 6.13-6.16*, the forecasts of March and April 2006 are presented as extracted from the four neural networks. The *Figures 6.17-6.18* also compare the results of the four neural networks in forecasting March and April 2006.

In addition, every neural network will provide its forecasts for the upcoming **3, 6, 12 and 18 months**. In *Figures 6.19-6.22*, for each of these rates the forecasting was based on data taken with 3, 6, 12 and 18 months delay. This is the reason why every neural network presents a different rate for one specific month. These figures can be evaluated in the future and may provide a picture of how the market might look like up to 18 months from now.

### **6.3 Performance of the four Neural Networks**

The *error function* or the *cost function* is used to measure the distance between the targets and the outputs of the network. The weights of the network are updated in the direction that makes the error function minimum. As we have mentioned, we will use six values to measure the performance of the network.

The six values of performance are presented for the training data, the cross-validation data and the unknown data. These values correspond to all four neural networks individually.

To begin with the training data, *Table 6.1a* illustrates all the error values for all four networks.

<b>Training Data Performance</b>				
	<i>Delay</i>			
	<b>3 Months</b>	<b>6 Months</b>	<b>12 Months</b>	<b>18 Months</b>
<b>MSE</b>	0.01572	0.015791	0.016145	0.024614
<b>NMSE</b>	0.100912	0.098624	0.141214	0.237683
<b>r</b>	0.967009	0.956789	0.906529	0.892958
<b>% Error</b>	18.46332	18.330399	19.855789	20.41568
<b>AIC</b>	438.943573	635.033752	780.161621	481.911041
<b>MDL</b>	307.886658	454.7677	480.538452	342.317841

*Table 6.1a: Training Data Performance for all four ANNs.*

The cross-validation data have their performance, as presented in *Table 6.1b*. Occasionally the process was stopped and the weights were tested so that the actual outputs fit with the desired outputs.

<b>Cross-Validation Data Performance</b>				
	<i>Delay</i>			
	<b>3 Months</b>	<b>6 Months</b>	<b>12 Months</b>	<b>18 Months</b>
<b>MSE</b>	0.006582	0.011637	0.013078	0.00811
<b>NMSE</b>	0.042251	0.072681	0.060771	0.078312
<b>R</b>	0.978704	0.969293	0.969973	0.961898
<b>% Error</b>	12.381732	15.401151	11.897898	13.774767
<b>AIC</b>	421.531097	534.929077	1987.26416	1231.706177
<b>MDL</b>	290.474182	378.263306	1466.551392	898.289368

*Table 6.1b: Cross-Validation Data Performance for all four ANNs.*

Finally, the performance for the unknown data is shown in *Table 6.1c*. At this phase, we have not provided the network with the desired rates of *March and April 2006*, but we tried to forecast these rates using historical data. Therefore, we do not have any values for the *MSE*, *NMSE* and *r*. The calculations of their functions require a desired response.

Unknown Data Performance				
	<i>Delay</i>			
	3 Months	6 Months	12 Months	18 Months
MSE	-	-	-	-
NMSE	-	-	-	-
<b>r</b>	-	-	-	-
<b>% Error</b>	7.512208	32.350433	31.172344	38.467579
AIC	507.645386	614.141907	1563.134277	1301.195679
MDL	129.014282	269.656403	1077.558472	932.799133

*Table 6.1c: Unknown Data Performance for all four ANNs.*

The *MSE*, which is two times the average cost function and determines how well the network output fits the desired response, should decay towards zero as the network approaches the desired response. For all the ANNs the MSE is below 0.025 and is increasing as we increase the delay of every network.

The *correlation coefficient* “*r*” is confined for all networks to the range [-1, 1]. There is a perfect positive linear correlation between the network output and the desired response. In other words, they co-vary and the fit of every model to the data is reasonably good.

The *percentage error* “% Error” is very low and does not exceed 21% for the training data, 16% for the cross-validation data, and 39% for the unknown data. If for example the desired output was 0.1 and the actual output was 0.2, even though the two values are quite close, the percent error for this example is 100%.

Last but not least, the *AIC* and *MDL* functions which try to measure the tradeoff between training performance and network size are kept in reasonable values.

Finally, for *March and April 2006*, we have the following values, as presented in *Table 6.1d* and *Table 6.1e*.

March 2006 Performance				
	<i>Delay</i>			
	3 Months	6 Months	12 Months	18 Months
<b>MSE</b>	0.113786	0.045789	0.027416	0.225569
<b>NMSE</b>	0.091754	0.055219	0.049234	0.473386

*Table 6.1d: March 2006 Performance for all four ANNs.*

April 2006 Performance				
	<i>Delay</i>			
	3 Months	6 Months	12 Months	18 Months
<b>MSE</b>	0.013452	0.094631	0.134902	0.294182
<b>NMSE</b>	0.029136	0.235748	0.378291	0.584715

*Table 6.1e: March 2006 Performance for all four ANNs.*

The *MSE* and *NMSE* values are very low and for all networks, they do not exceed 0.3 and 0.6 respectively. This can be easily explained in the figures presented later in this chapter, by the extremely fine results that the networks gave as outputs compared to the actual rates of the Tubarao/Rotterdam Capesize Ore route.

## 6.4 Evaluation of the four Neural Networks

The *Figures 6.1-6.22* present all the data extracted from the four neural networks. Their evaluation will give significant and imperative information about the use of artificial neural networks in the shipping market. They compose a complete picture of the Capesize Bulk Carrier market and they specialize in the route Tubarao-Rotterdam.

The most substantial and constructive conclusions derive from the unsuccessful forecasts of the networks. My objective is to clarify the reasons for these unsuccessful forecasts and provide suggestions for the improvement of the current methodology.

An unsuccessful forecast may happen due to various factors. One basic factor is an unexpected and extraordinary turn of the market. The network has been trained with data

that occurred in an interval of twenty-one years. If for example, there was a unique and unusual reaction in the market during these twenty-one years that would cause another reaction; the network might give completely different results. In addition, events that occurred only once or twice in the market, can also hardly be forecasted. But all these events are a challenge for neural networks. The current thesis is trying to cover as much of these unexpected conditions and give as many accurate results and represent the capsize market in a realistic way.

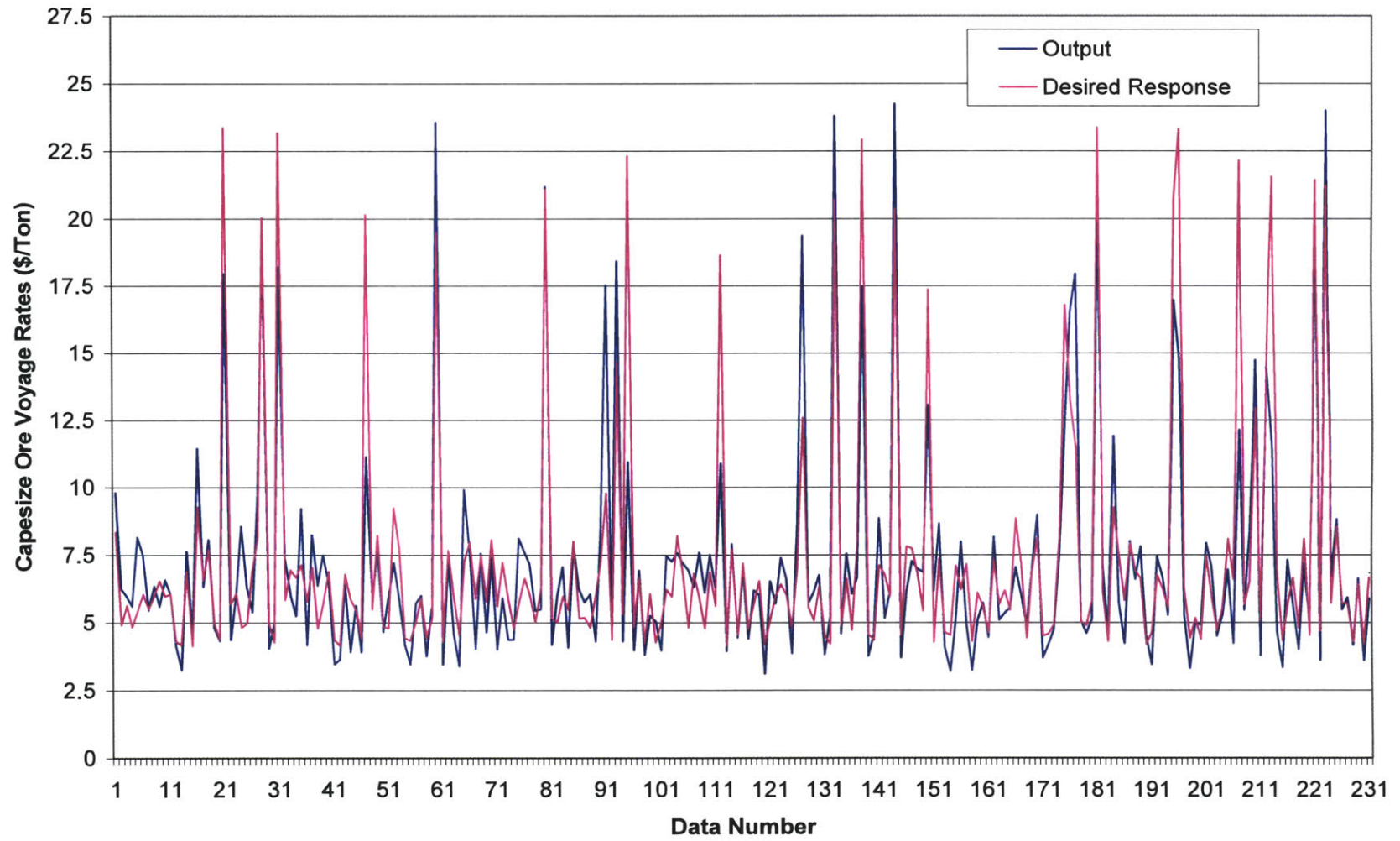
#### **6.4.1 Training Data Evaluation**

The *Figures 6.1-6.4* show that the four artificial neural networks have been trained remarkably well. The forecasted rates (outputs) follow the stream of the real prices (desired responses); while the extremums of the two curves almost coincide. *Genetic algorithms* have considerably contributed, with the exceptional and careful selection of the input vectors. The genetic operations of reproduction, crossover and mutation have created an offspring of inputs that do not correlate with each other and have selected the parameters that apply to every vector of processing elements.

In addition *momentum* which was used as a learning rule and *tan sigmoid transfer function*, have trained the weights and locked their values so that the MSE is minimized. The *two modulars* have created a structure within the topology that fosters specialization of function in each sub-module. A smaller number of weights are required since the two modulars do not have full interconnectivity. This tends to speed up training times and reduce the number of required training exemplars.

Finally, *the training was properly stopped*, and the network was not over-trained or under-trained. If the network was under-trained, we would not have achieved the extreme rates that the capesize ore voyage rates have taken, and the output would not have followed the stream of the desired response. On the contrary, if the network was over-trained, the output would have far more extreme rates than the desired response has.

*Figure 6.1: ANN Training Data with 3 months Delay.*



*Figure 6.2: ANN Training Data with 6 months Delay.*

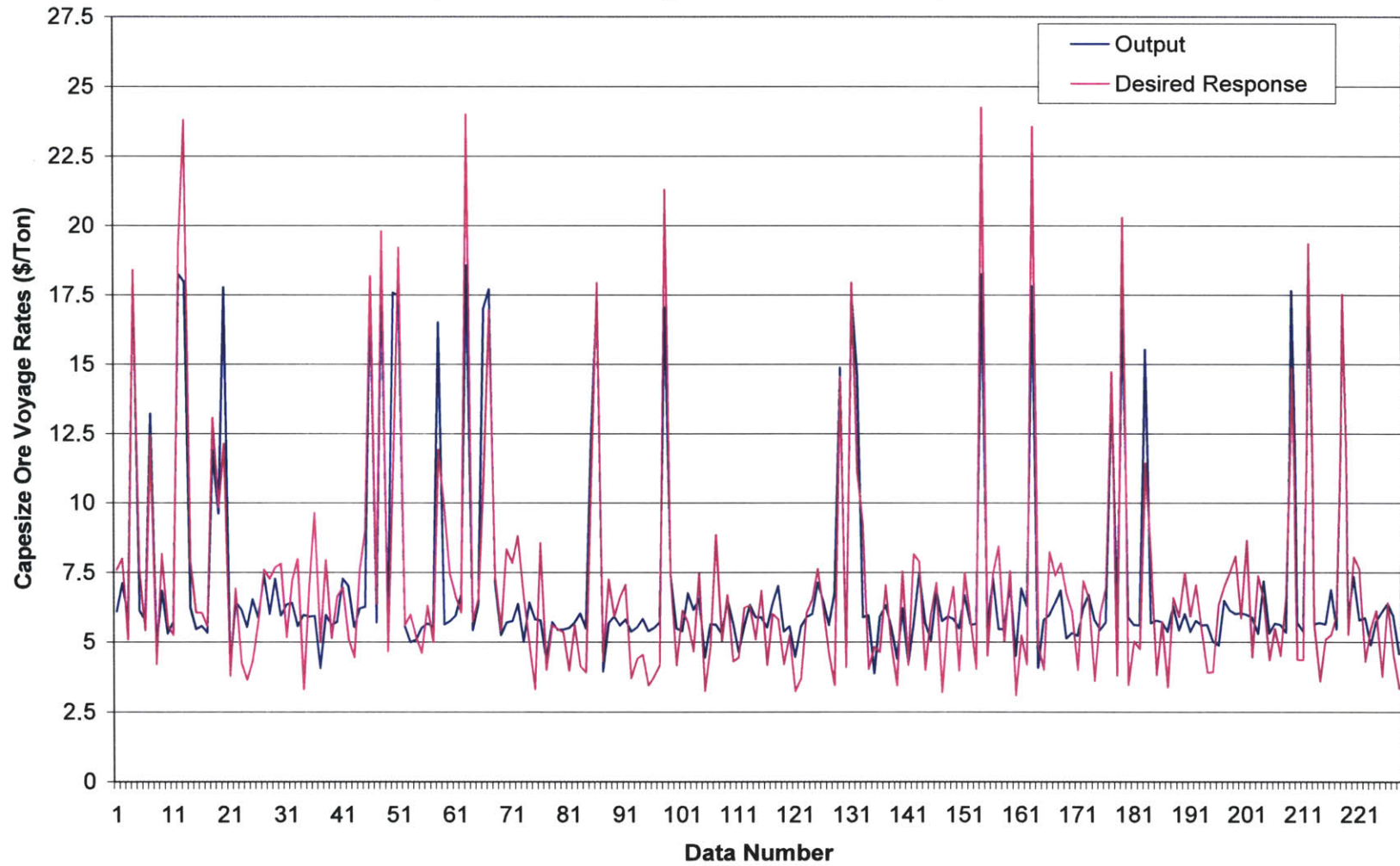
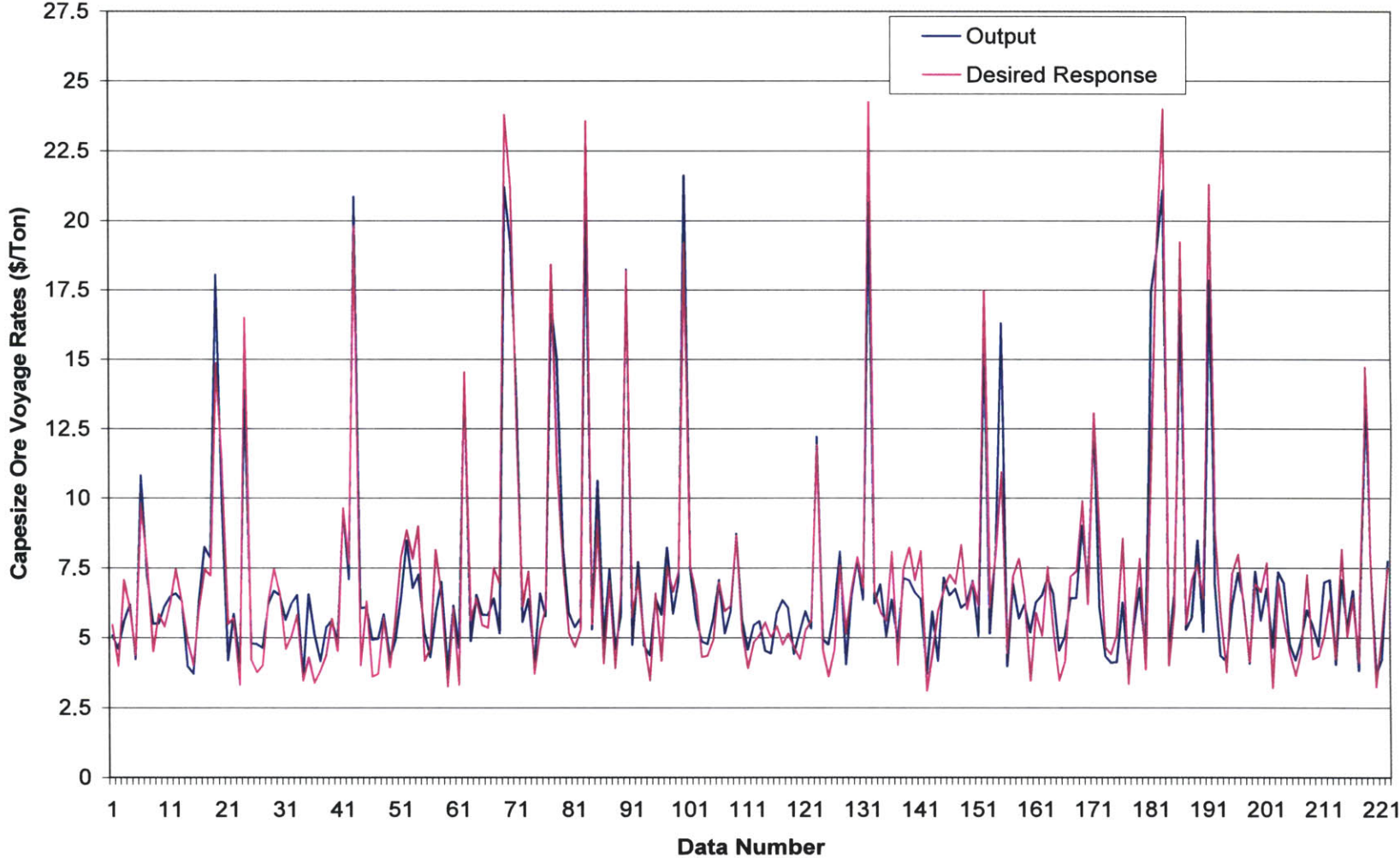
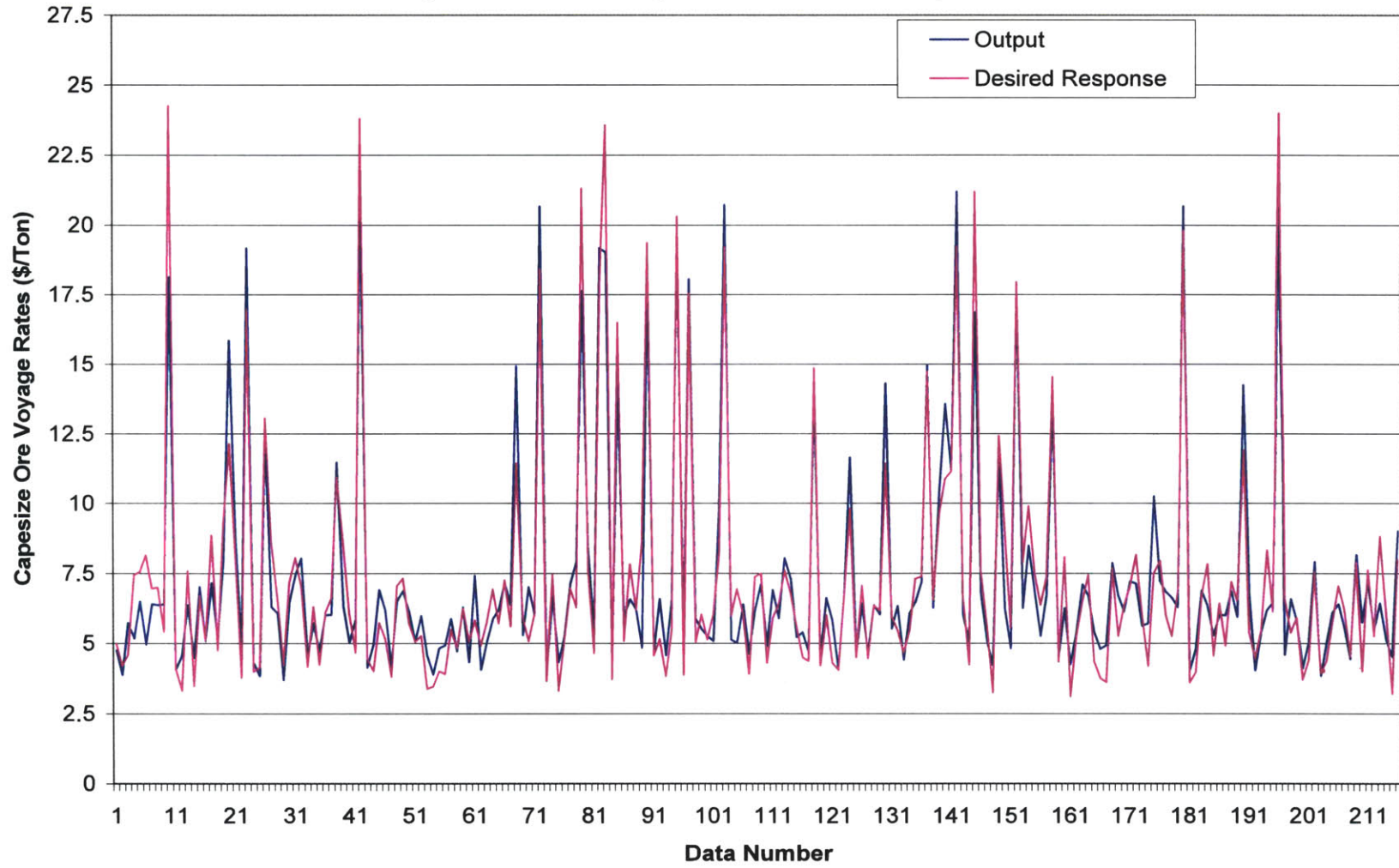


Figure 6.3: ANN Training Data with 12 months Delay.

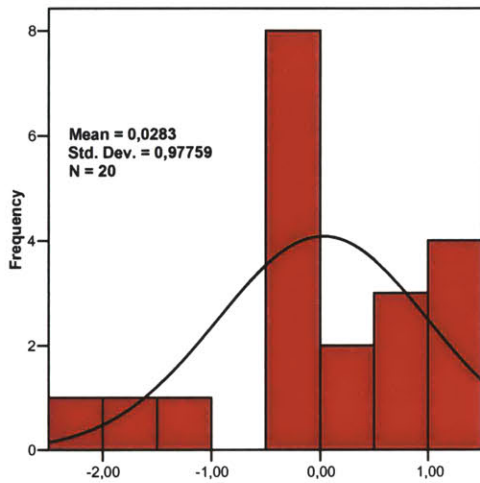


*Figure 6.4: ANN Training Data with 18 months Delay.*

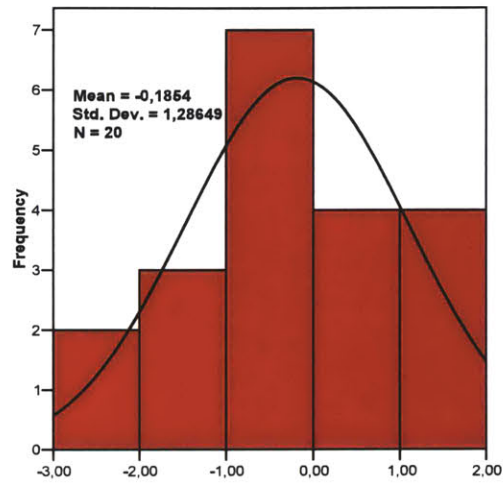


## 6.4.2 Cross-Validation Data Evaluation

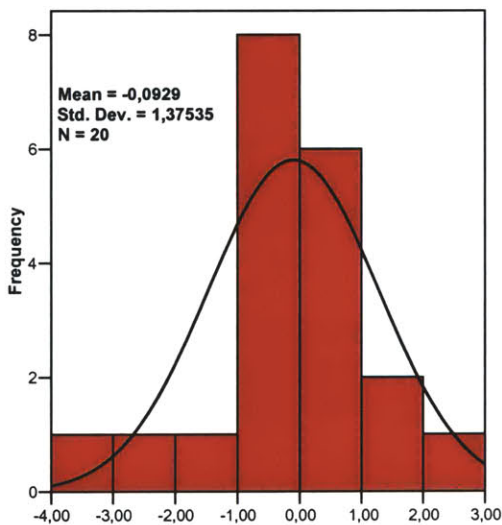
The following *Figures 6.5-6.8* show the normal distributions of the difference between the desired rates and the output rates for all four neural networks.



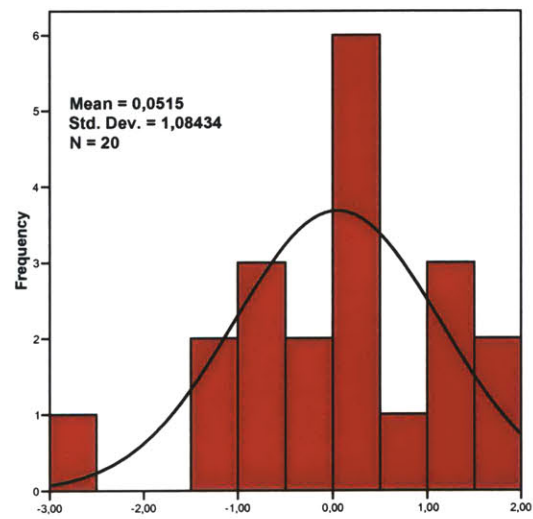
*Figure 6.5: Distribution of the ANN Training data with 3 months of delay.*



*Figure 6.6: Distribution of the ANN Training data with 6 months of delay.*



*Figure 6.7: Distribution of the ANN Training data with 12 months of delay.*



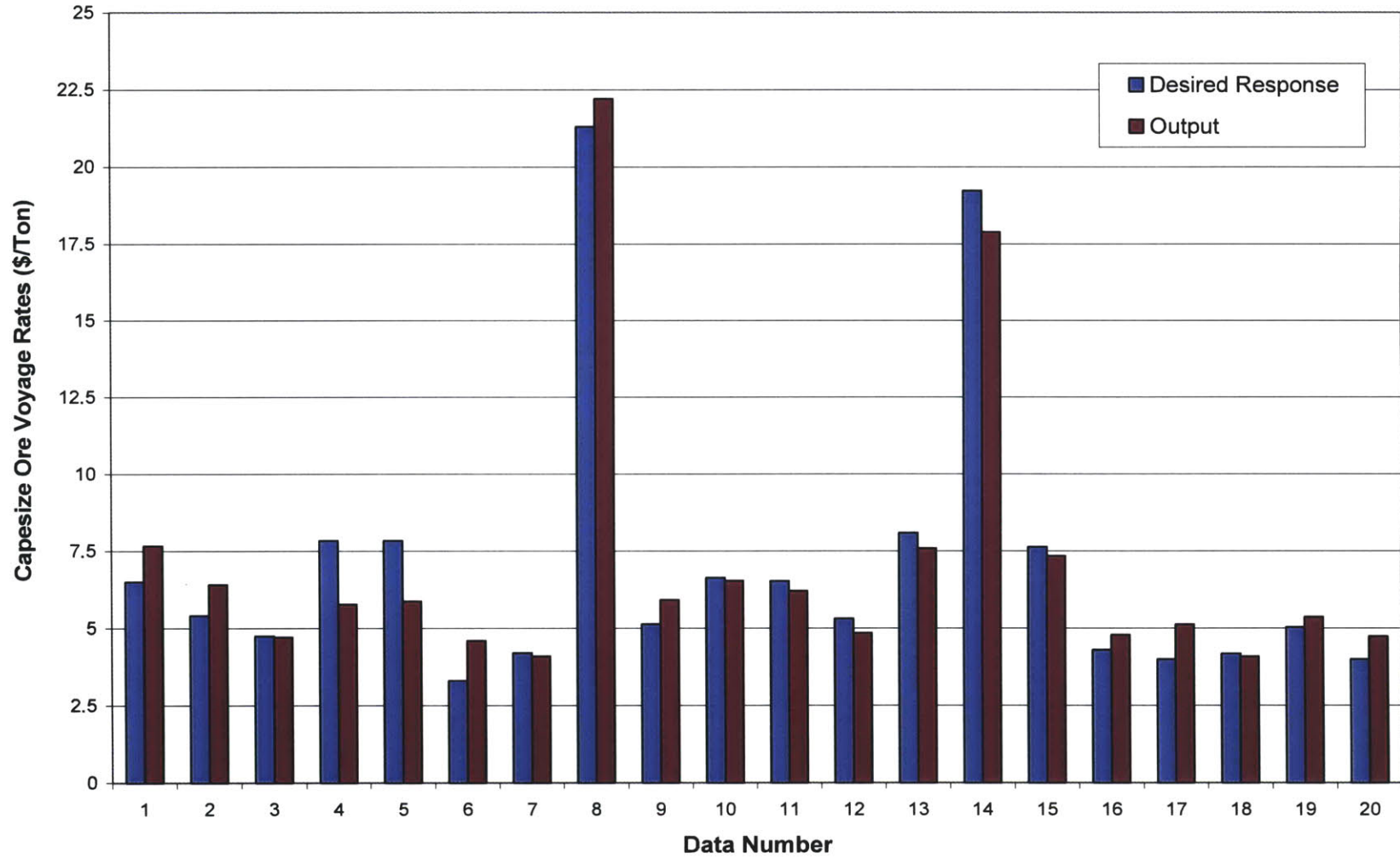
*Figure 6.8: Distribution of the ANN Training data with 18 months of delay.*

The mean of every distribution is *less than*  $|0.2|$  and similarly their standard deviations does not exceed  $1.4 \text{ \$/Ton}$ . It can be clearly seen that  $68.26\%$  (*from*  $-1\sigma$  to  $+1\sigma$ ) of the data are around the *mean number* of every distribution.

In addition, *Figures 6.9-6.12* show the 20 training inputs that have been used as cross-validation data. The blue bar depicts the desired response and the red one the outputs of the neural networks. None of the outputs illustrates extreme divergences, confirming that the network was not over-trained or under-trained.

*Cross validation* is a highly recommended method for stopping network training. The cross validation set is used to determine the level of generalization produced by the training set. It is executed in concurrence with the training of the network. Every so often, the network weights are frozen, the cross validation data is fed through the network, and the results are reported. The stop criteria of the controller can be based on the error of the cross validation set instead of the training set to insure this generalization. When the performance starts to degrade in the cross validation set, the training should be stopped. Hence, as the *Figures 6.5-6.12* indicate, the cross-validation data were first sufficient to test the networks, and secondly able to stop the training avoiding over-training and under-training.

*Figure 6.9: ANN Cross-Validation Data with 3 months Delay.*



*Figure 6.10: ANN Cross-Validation Data with 6 months Delay.*

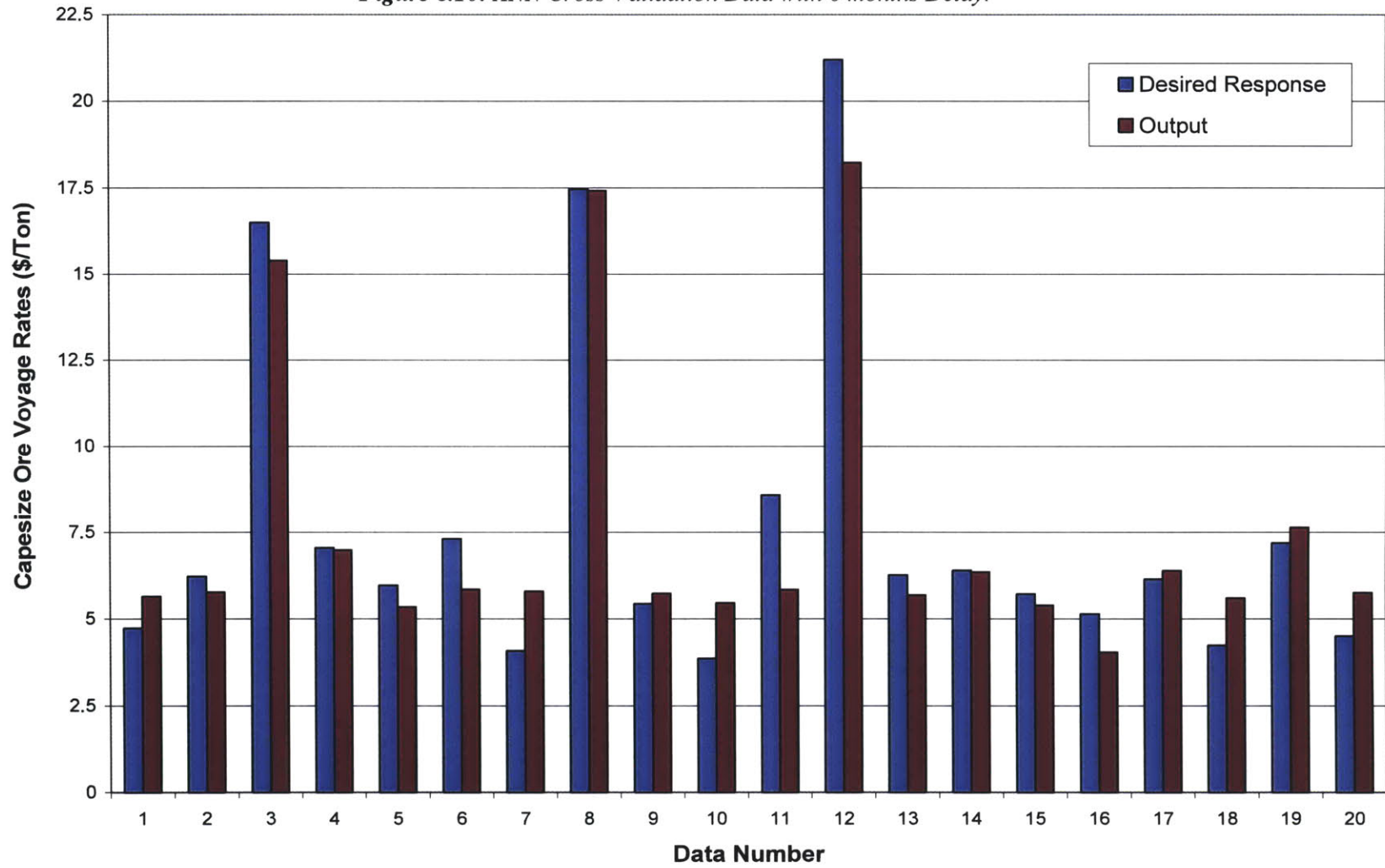
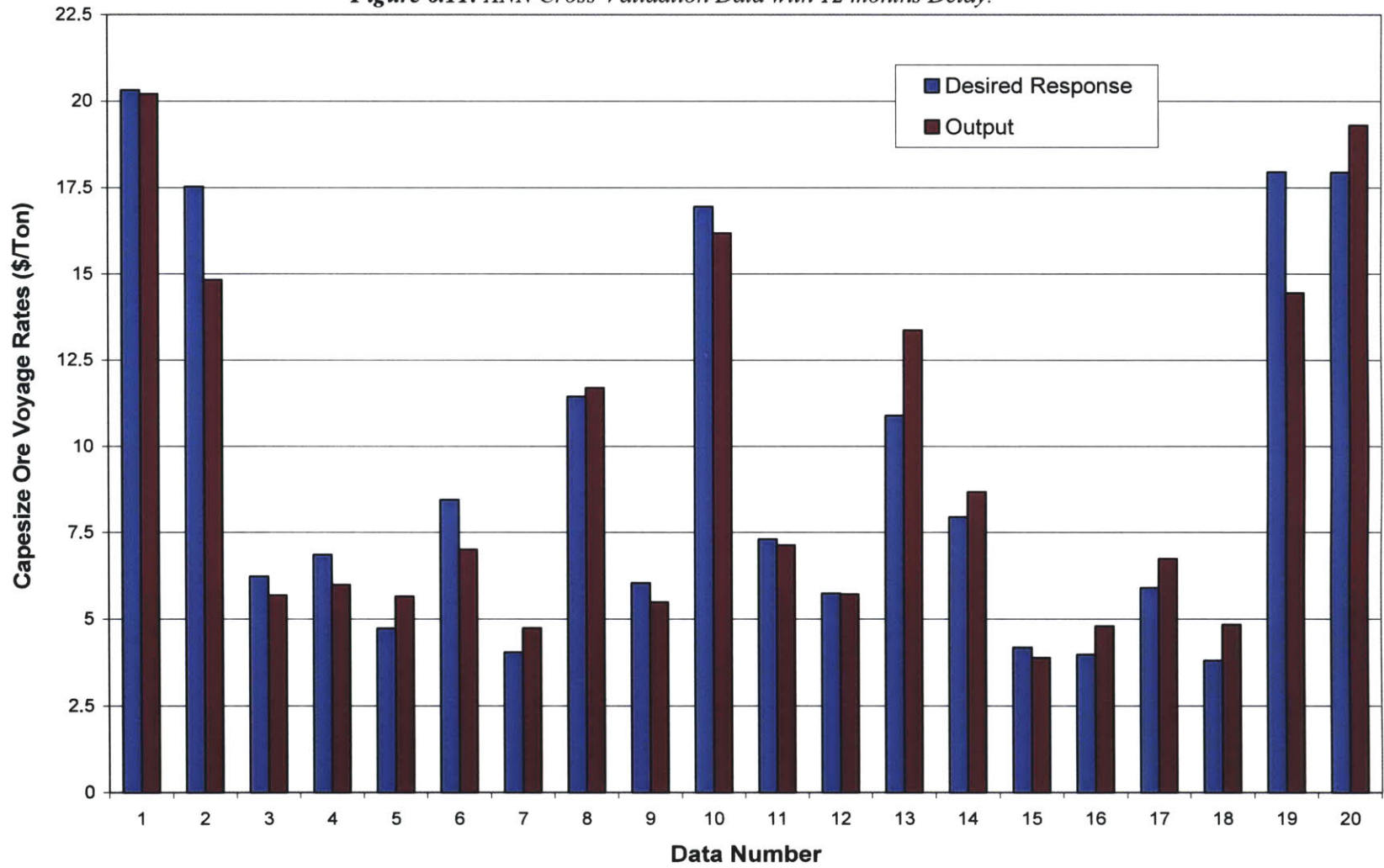
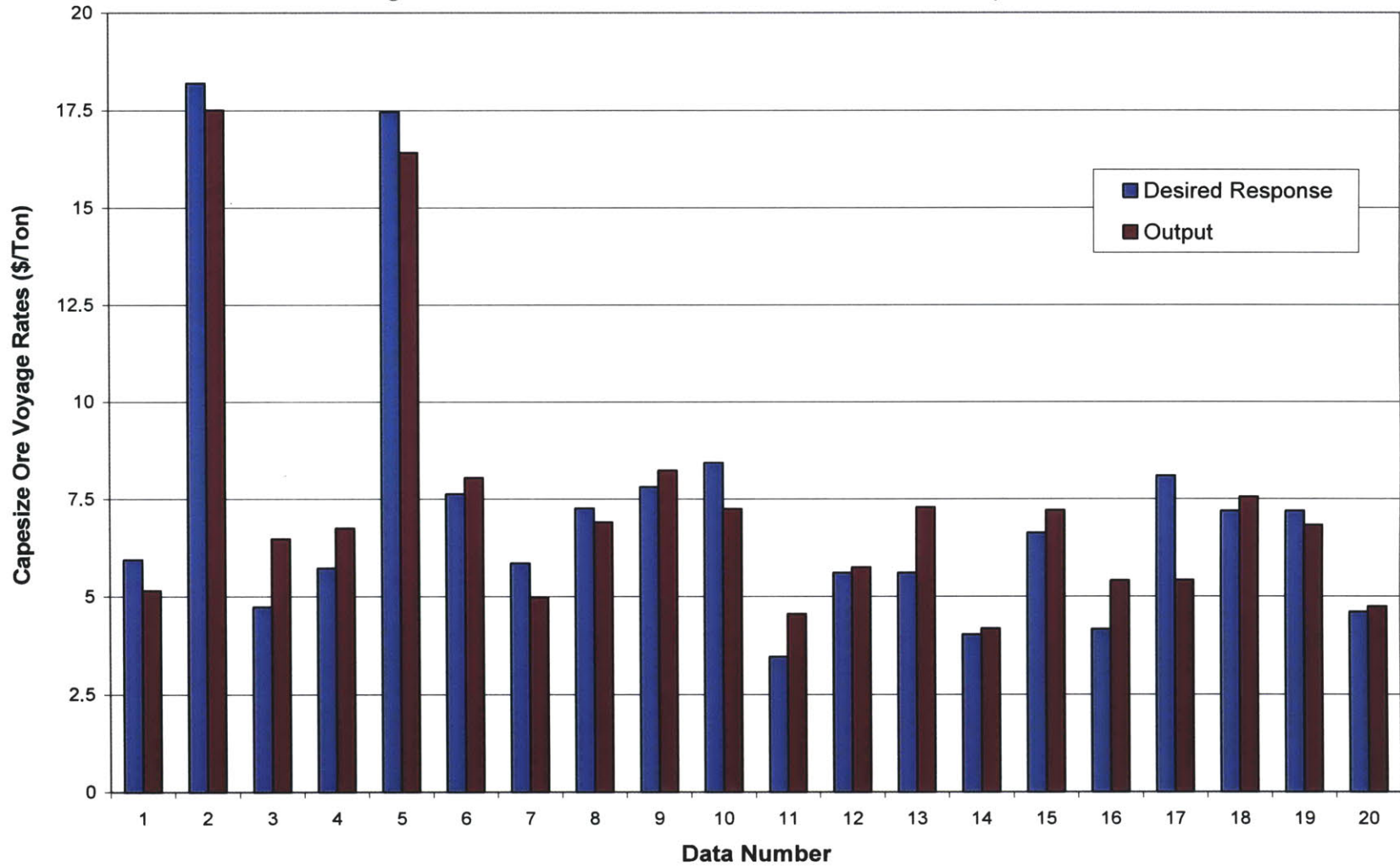


Figure 6.11: ANN Cross-Validation Data with 12 months Delay.



*Figure 6.12: ANN Cross-Validation Data with 18 months Delay.*



### 6.4.3 Testing Unknown Data Evaluation

The testing set is used to test the performance of the network. Once the network is trained the weights are then frozen, the testing set is fed into the network and the network output is compared with the desired output. For this purpose, we will use the Capesize ore voyage rates for *March and April 2006*.

In *Figure 6.13-6.18*, we present the forecasts of March and April 2006 from all four networks. For an easier evaluation, the following table presents the rates for *March and April 2006* collectively.

March and April 2006 Capesize Ore Voyage Rates (\$/Ton)					
Date	Desired Response	Delay			
		3 Months	6 Months	12 Months	18 Months
<i>2006-03</i>	<i>13.07</i>	15.078	12.074	12.190	11.038
<i>2006-04</i>	<i>12.11</i>	12.631	10.719	10.501	10.038

*Table 6.2: Forecast of March and April 2006 from the four ANNs.*

The forecasted rates (outputs) follow the stream of the real prices (desired responses) for a period of one year; while the extremums of the two curves almost agree. The forecasted prices are lower than the real prices for the 6, 12 and 18 months delay ANNs. On the contrary, they are higher for the ANN with 3 months delay. It is important to mention that all four networks are four independent models, which have different weights and different training and cross-validation data. In any case, the forecasted rates do not exceed *2.1 \$/Ton* for March 2006 and *2.2\$/Ton* for April 2006.

In addition, every neural network will provide its forecasts for the upcoming 3, 6, 12 and 18 months. In *Figures 6.19-6.22*, for each of these rates the forecasts were based

on data taken with 3, 6, 12 and 18 months delay. The following table presents the rates until February 2007.

<b>Capesize Ore Voyage Rates (\$/Ton) until August 2007</b>					
<i>Date</i>	<b>Desired Response</b>	<i>Delay</i>			
		<b>3 Months</b>	<b>6 Months</b>	<b>12 Months</b>	<b>18 Months</b>
<i>2006-03</i>	<i>13.07</i>	15.078	12.074	12.190	11.038
<i>2006-04</i>	<i>12.11</i>	12.631	10.719	10.501	10.038
<i>2006-05</i>	-	10.311	10.030	8.940	9.805
<i>2006-06</i>	-	-	9.661	5.879	8.966
<i>2006-07</i>	-	-	9.607	4.683	7.088
<i>2006-08</i>	-	-	9.415	4.557	5.757
<i>2006-09</i>	-	-	-	6.840	5.848
<i>2006-10</i>	-	-	-	7.308	6.581
<i>2006-11</i>	-	-	-	6.616	4.982
<i>2006-12</i>	-	-	-	5.498	4.869
<i>2007-01</i>	-	-	-	4.624	4.713
<i>2007-02</i>	-	-	-	4.491	4.138
<i>2007-03</i>	-	-	-	-	9.099
<i>2007-04</i>	-	-	-	-	7.609
<i>2007-05</i>	-	-	-	-	8.509
<i>2007-06</i>	-	-	-	-	9.207
<i>2007-07</i>	-	-	-	-	7.145
<i>2007-08</i>	-	-	-	-	9.538

*Table 6.3: Capesize Ore Voyage Rates Forecasts from all four ANNs.*

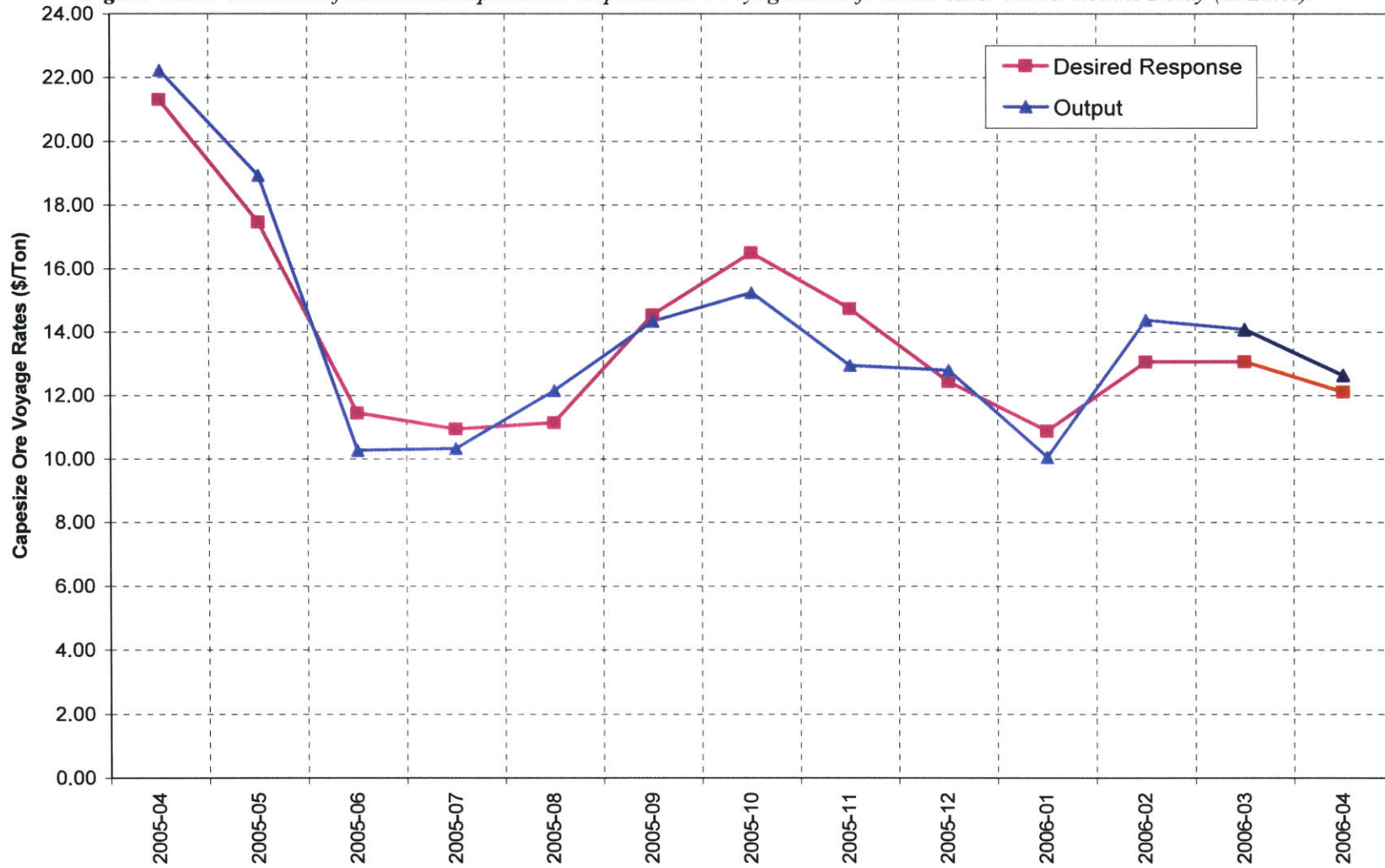
Every neural network presents a different rate for one specific month, since every forecast was based on data collected 3, 6, 12 and 18 months ago. Our forecasts indicate that the rates will continue to fall for the next 6 months.

One explanation for this reaction is that new capacity comes on stream and the old one remains reluctant to be removed. Furthermore, the Brazilian exports are forecasted to present an increase from 225.0 Million mt in 2005 to 249.0 Million mt in 2006; an increase of 10.6%, while the imports in Europe and especially in the Netherlands are forecasted to remain the same.

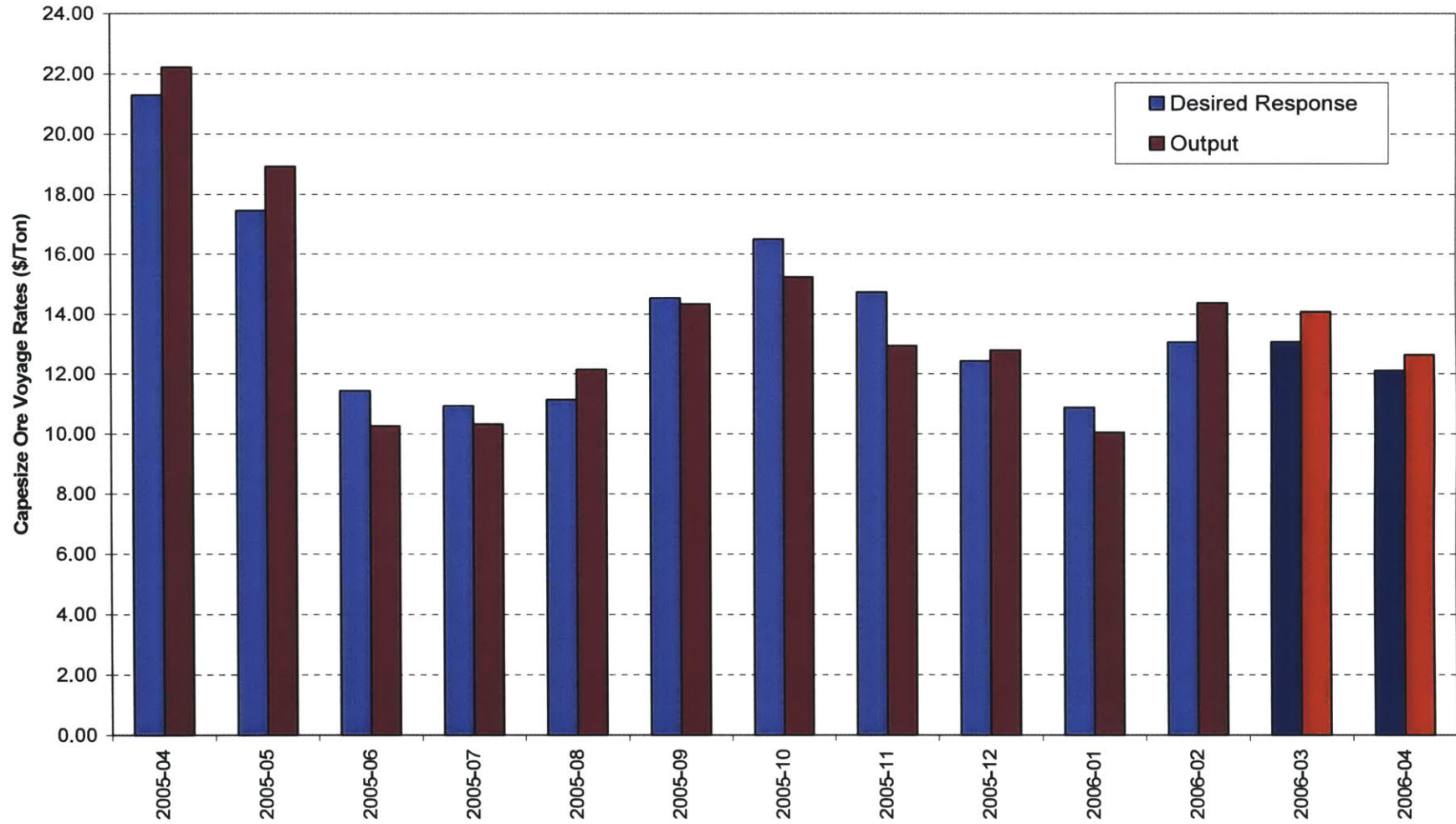
According to the market forecasts, only China will continue to drive global iron ore demand with an import total of over 350 Million mt in 2007, compared to 275 Million mt in 2005.

Therefore, it is our firm belief that the Capesize Ore Voyage rates from Tubarao to Rotterdam will continue to drop.

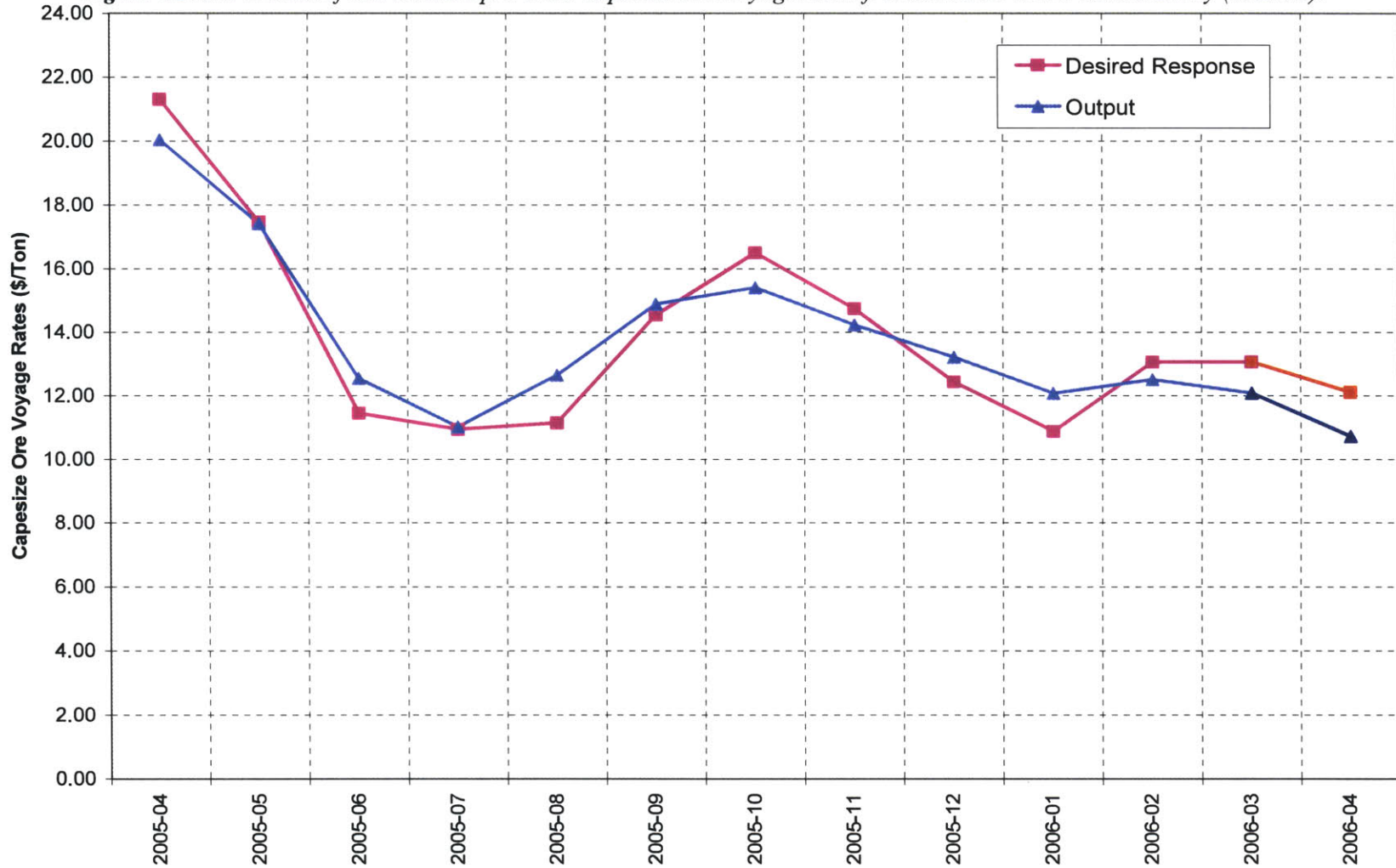
**Figure 6.13a:** Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 3 months Delay (in Lines).



*Figure 6.13b: Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 3 months Delay (in Bars).*



**Figure 6.14a:** Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 6 months Delay (in Lines).



**Figure 6.14b:** Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 6 months Delay (in Bars).

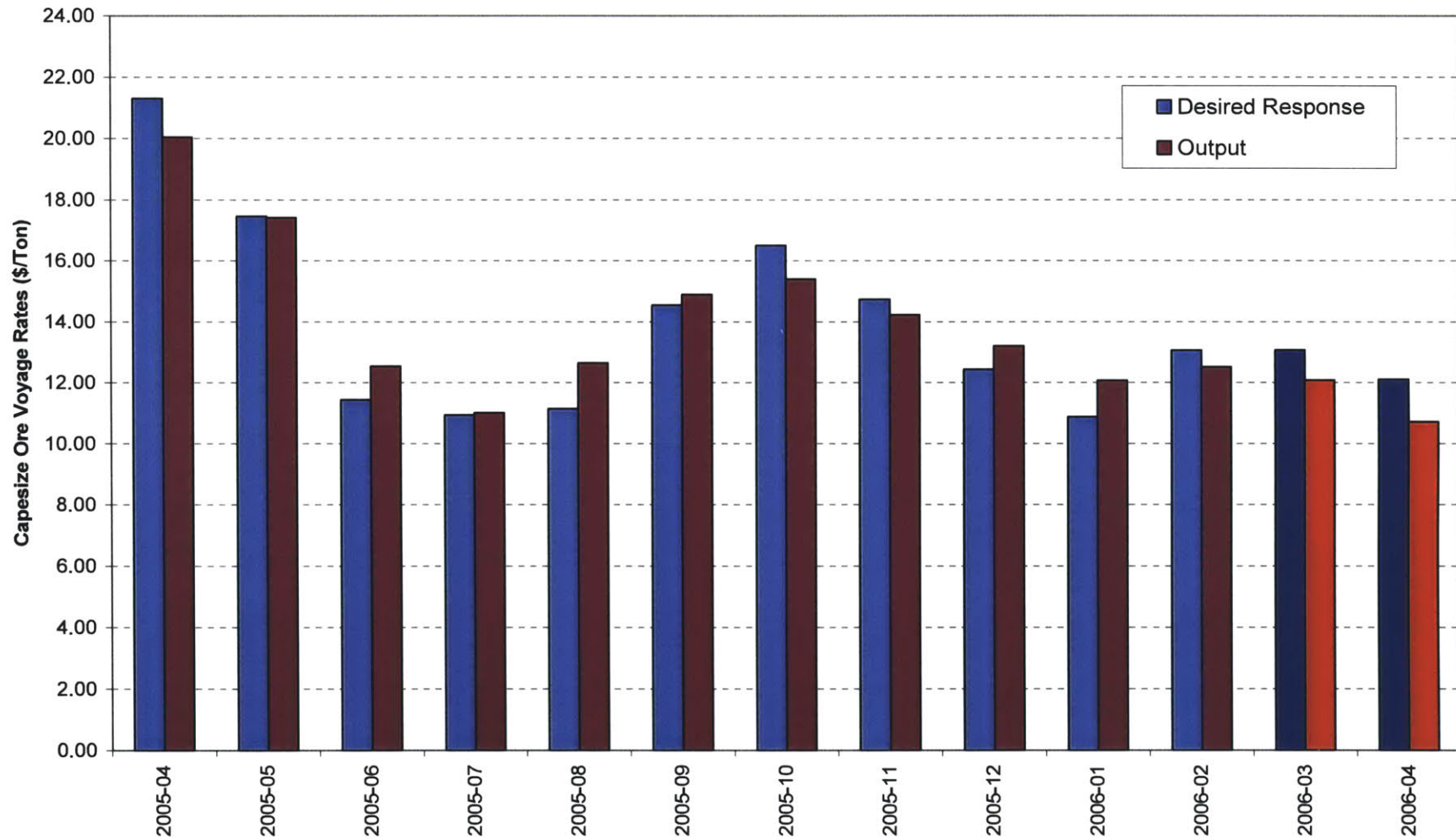
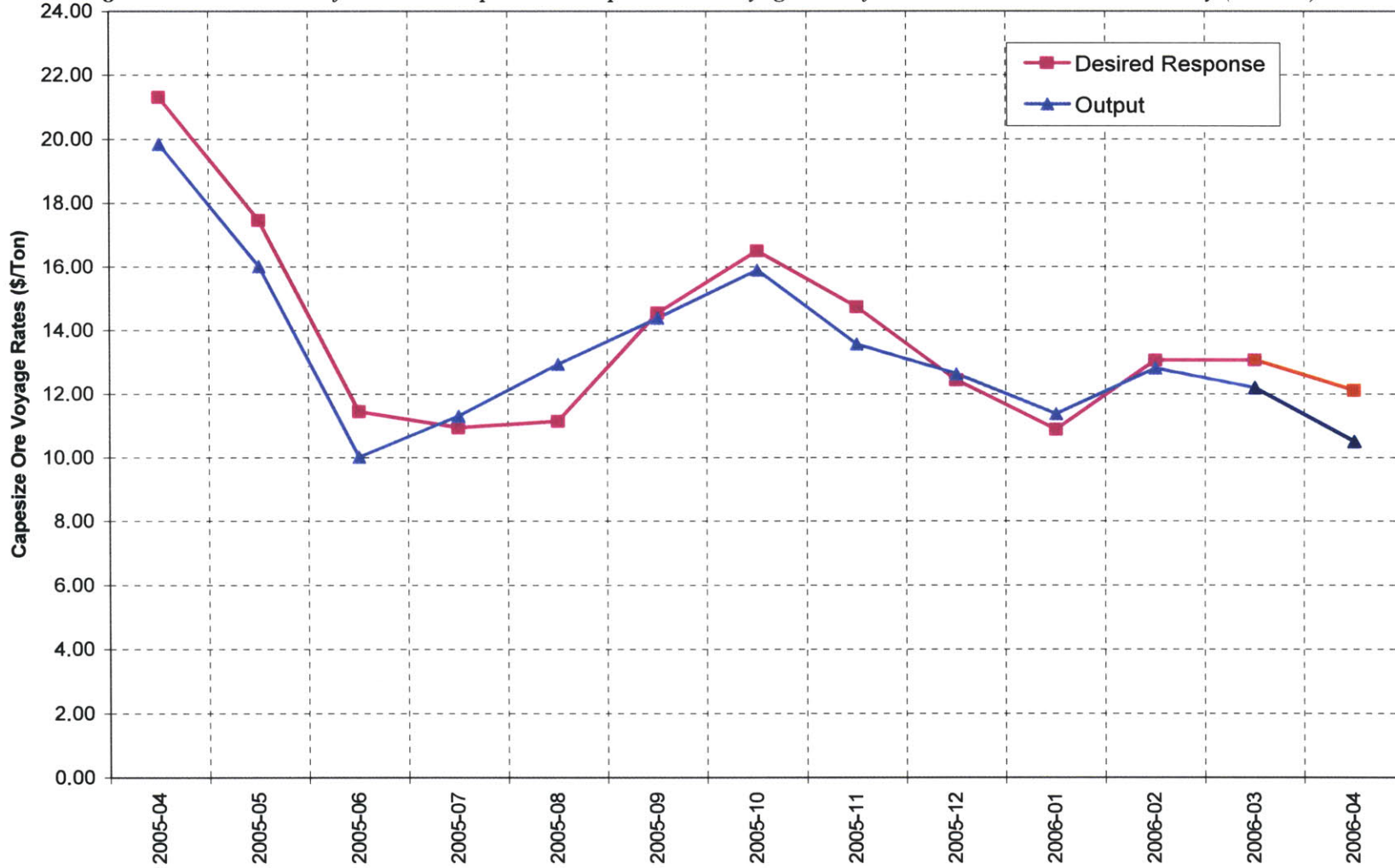
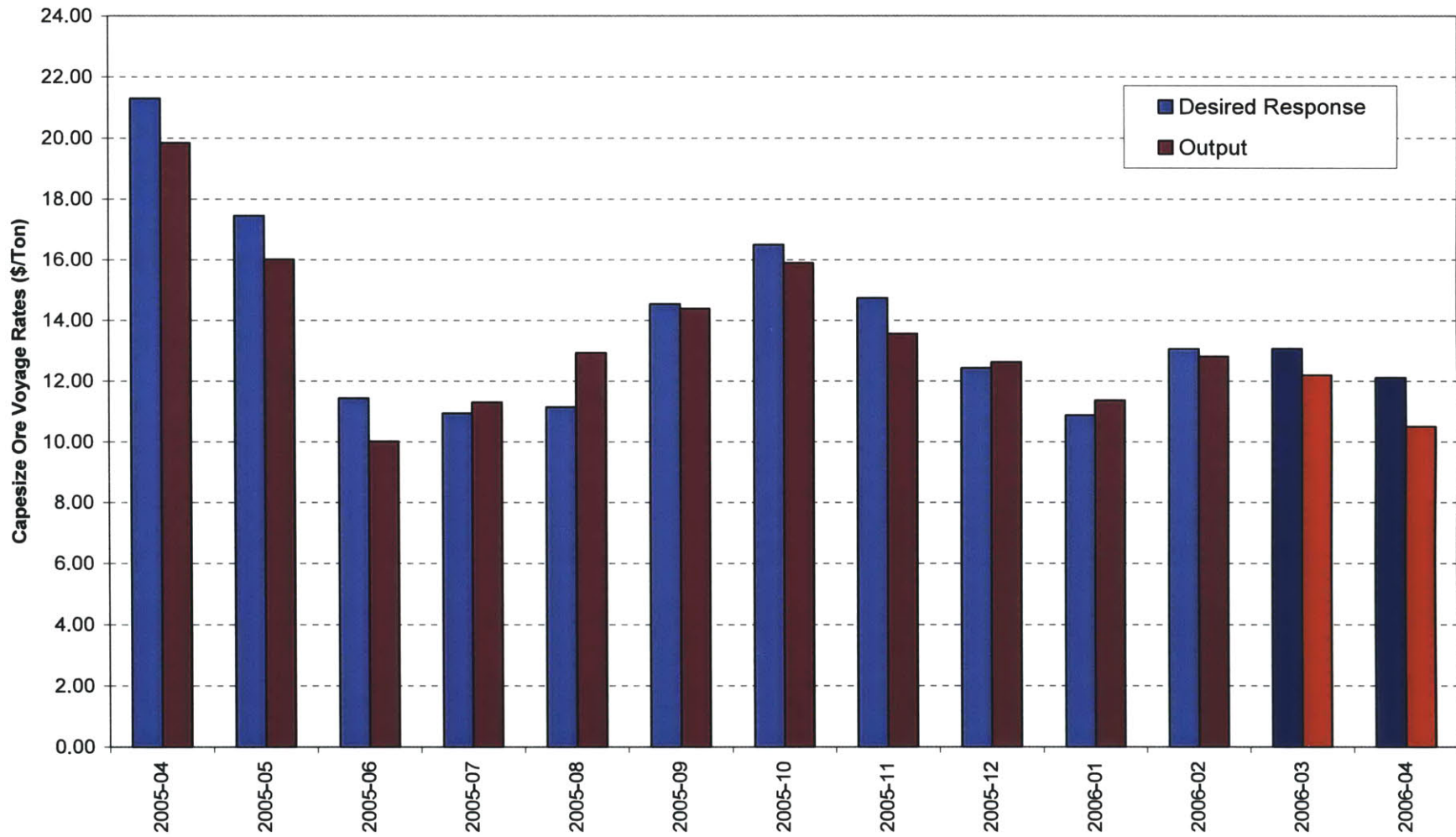


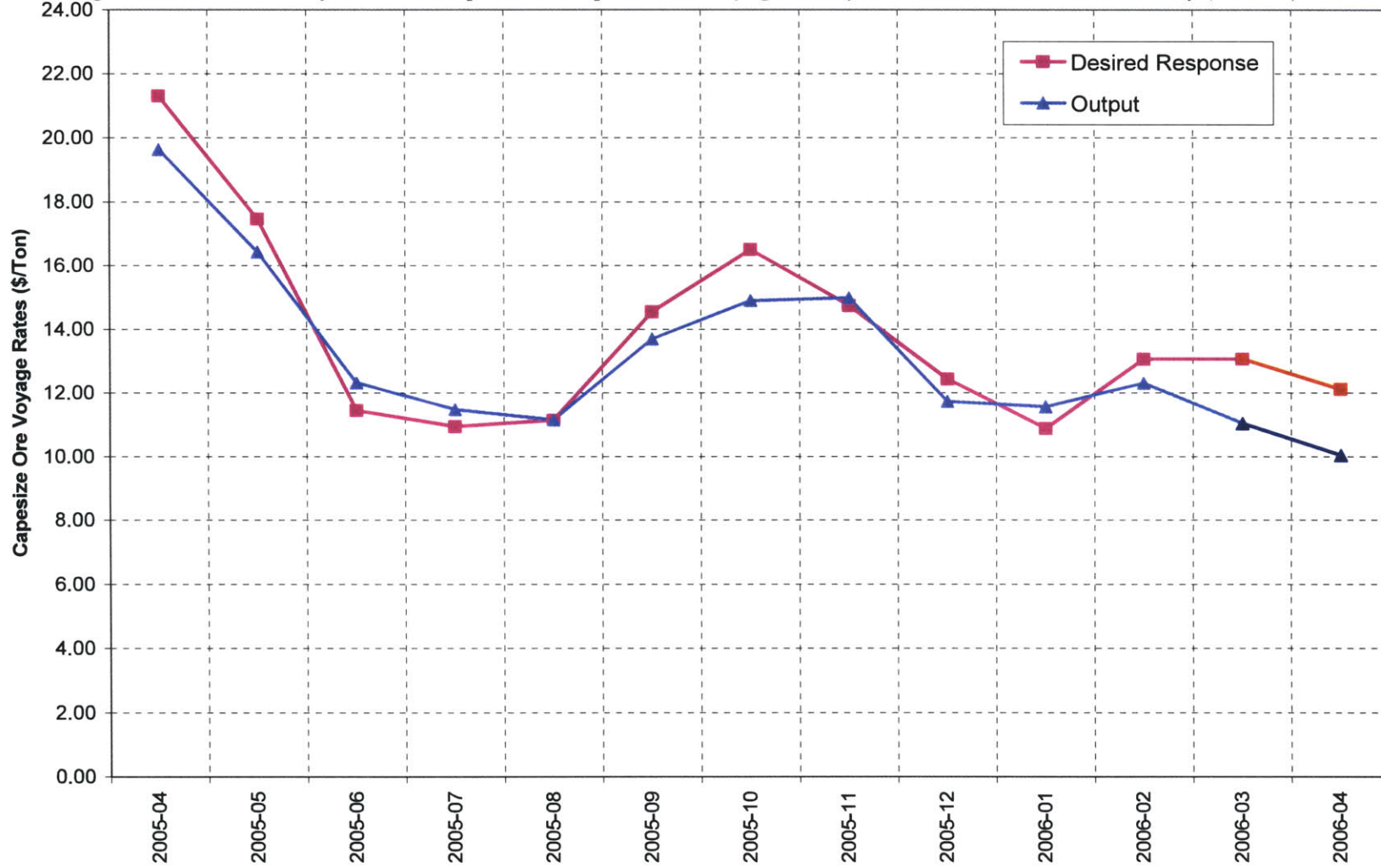
Figure 6.15a: Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 6 months Delay (in Lines).



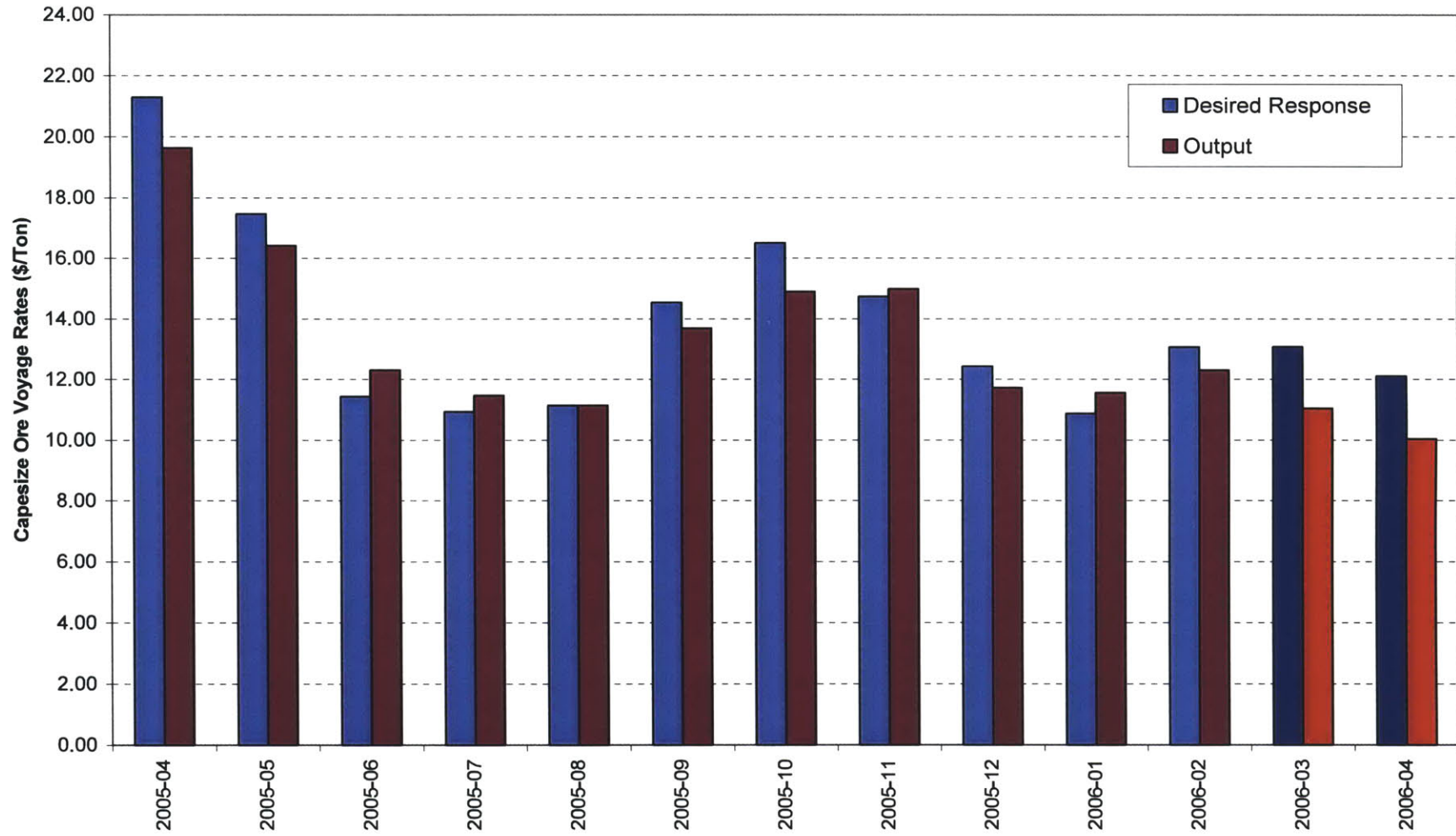
**Figure 6.15b:** Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 6 months Delay (in Bars).



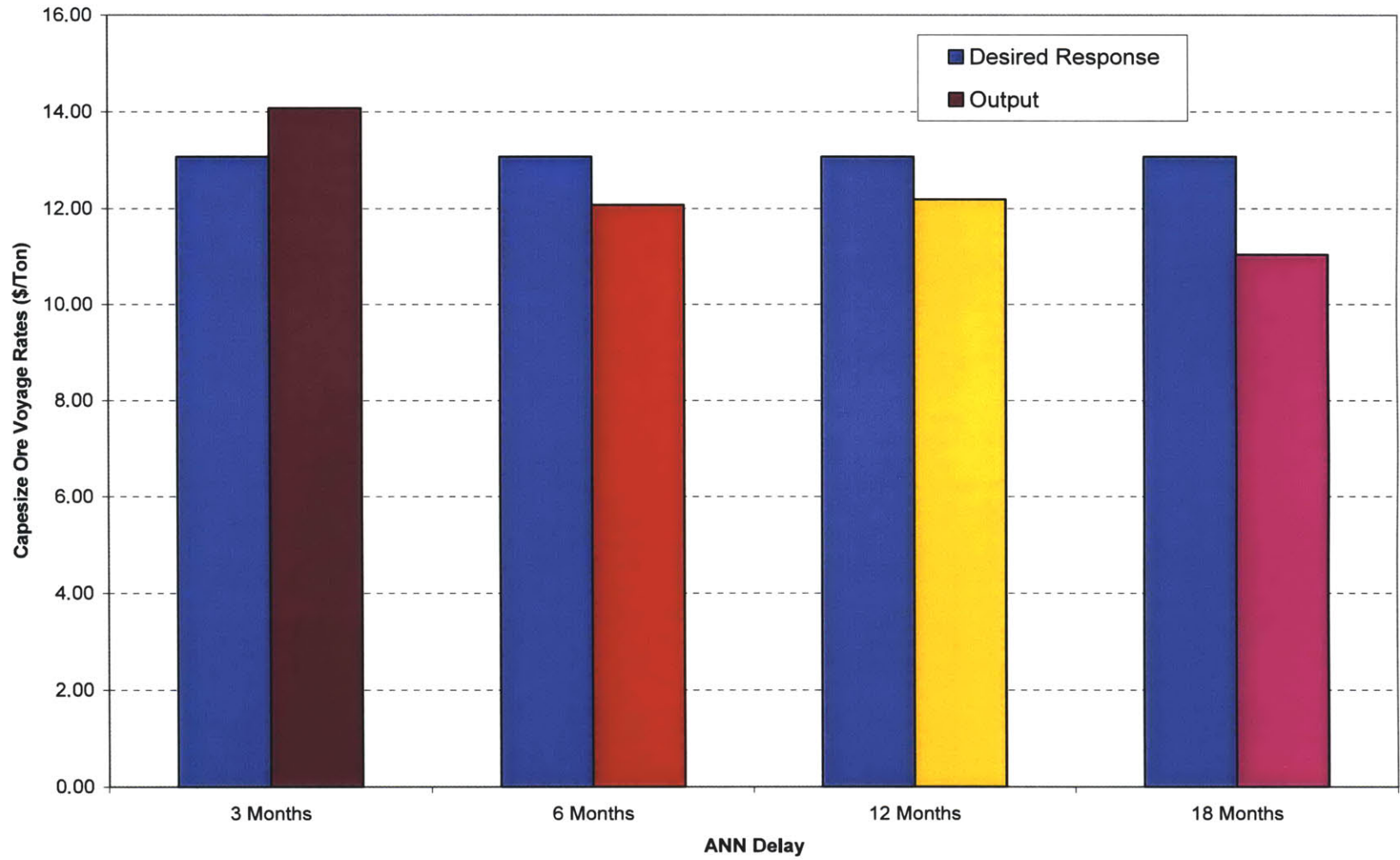
**Figure 6.16a:** Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 6 months Delay (in Lines).



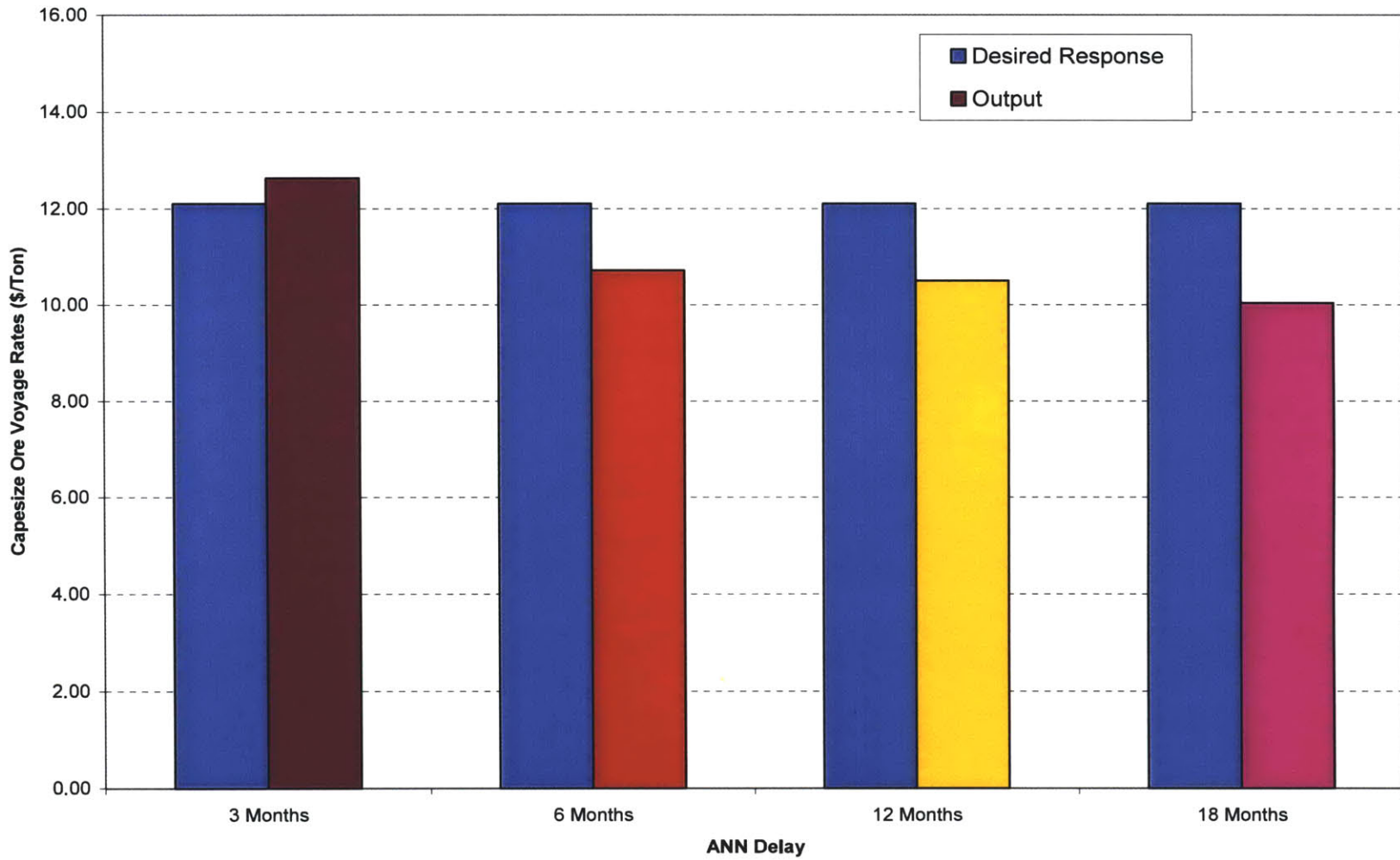
**Figure 6.16b:** Forecast of March and April 2006 Capesize Ore Voyage Rates from the ANN with 6 months Delay (in Bars).



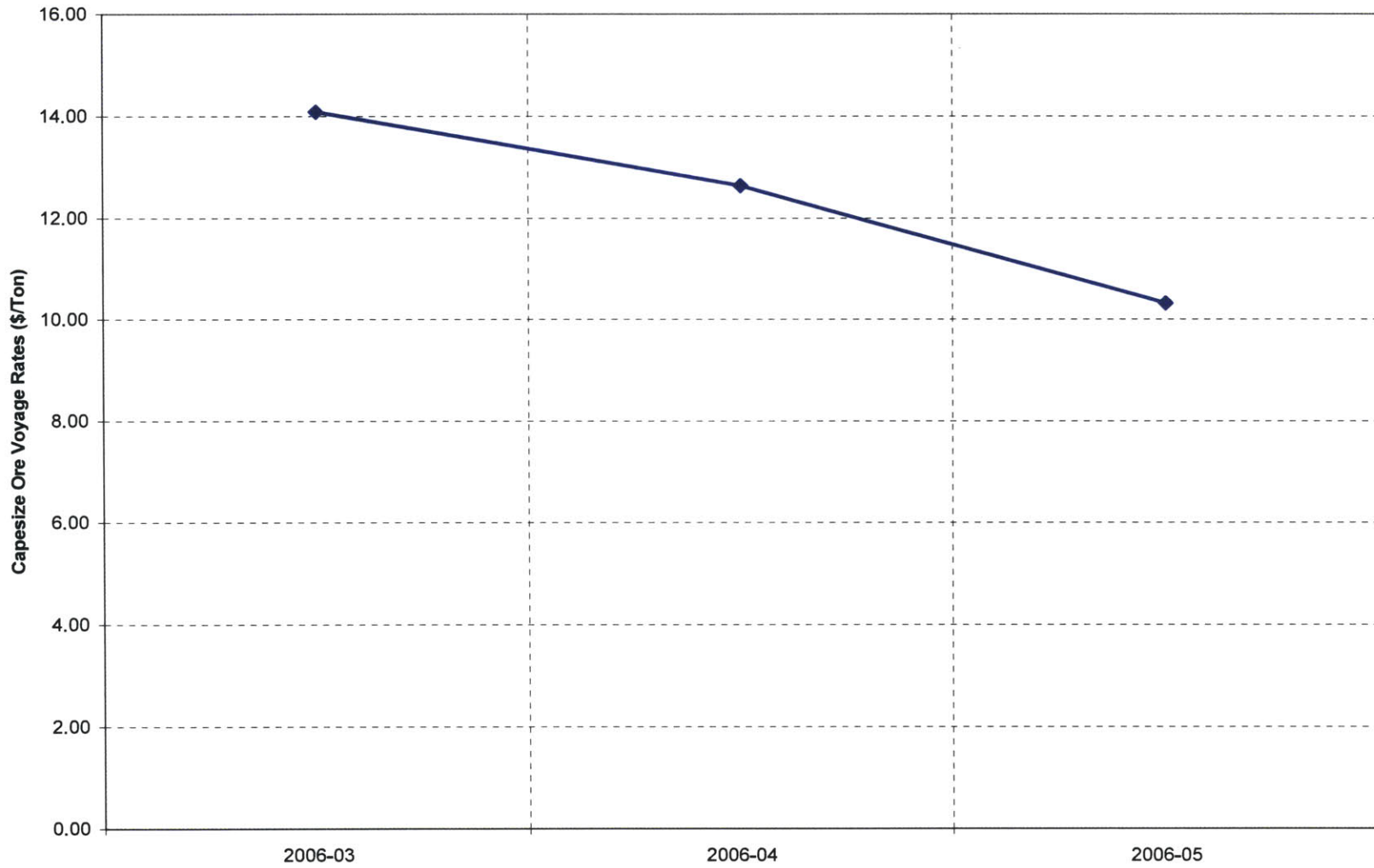
*Figure 6.17: Forecast of March 2006 Capesize Ore Voyage Rate from all four ANN.*



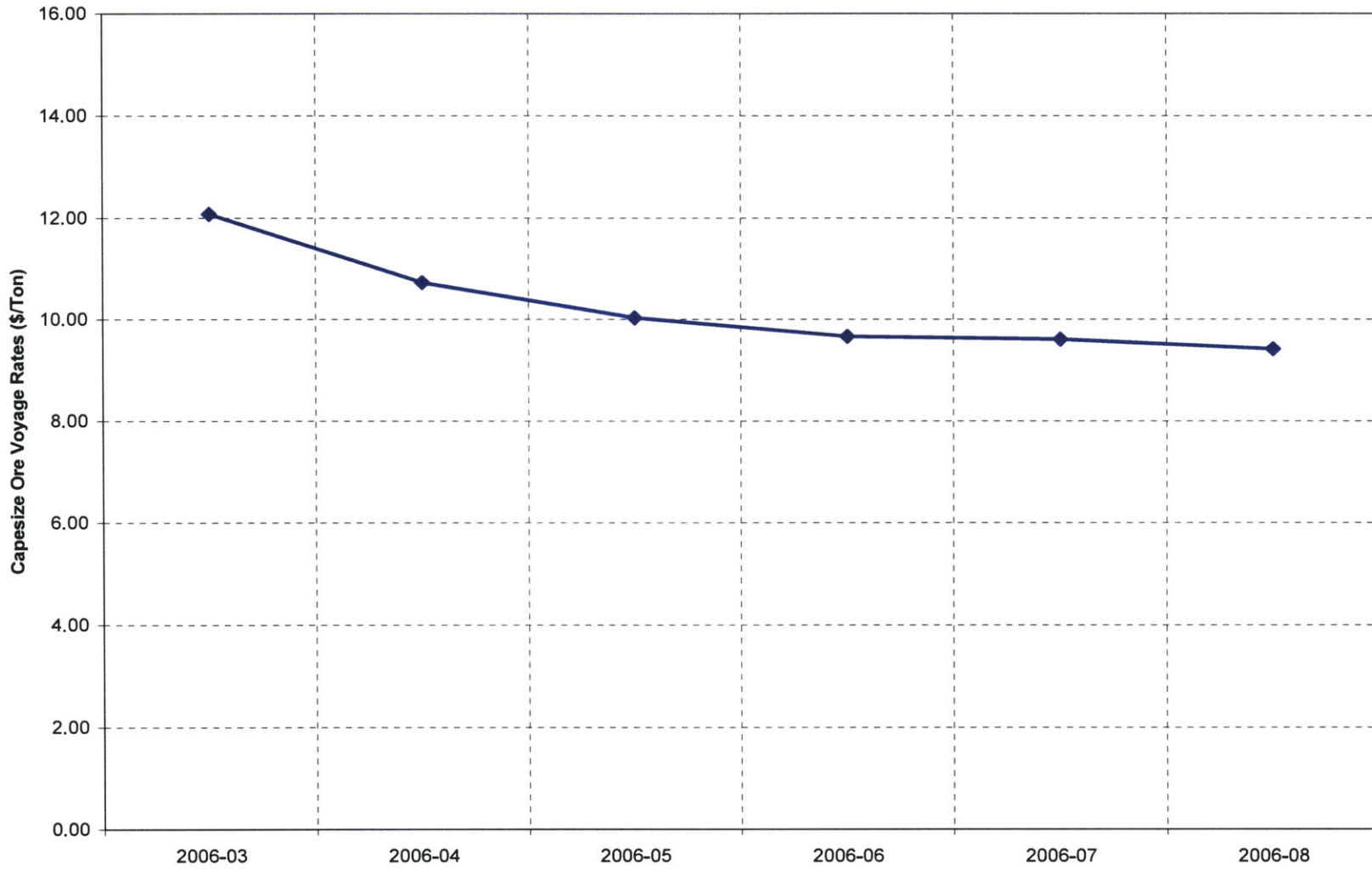
*Figure 6.18: Forecast of April 2006 Capesize Ore Voyage Rate from all four ANN.*



*Figure 6.19: May 2006 Capesize Ore Voyage Rates Forecast from ANN with 3 months Delay.*



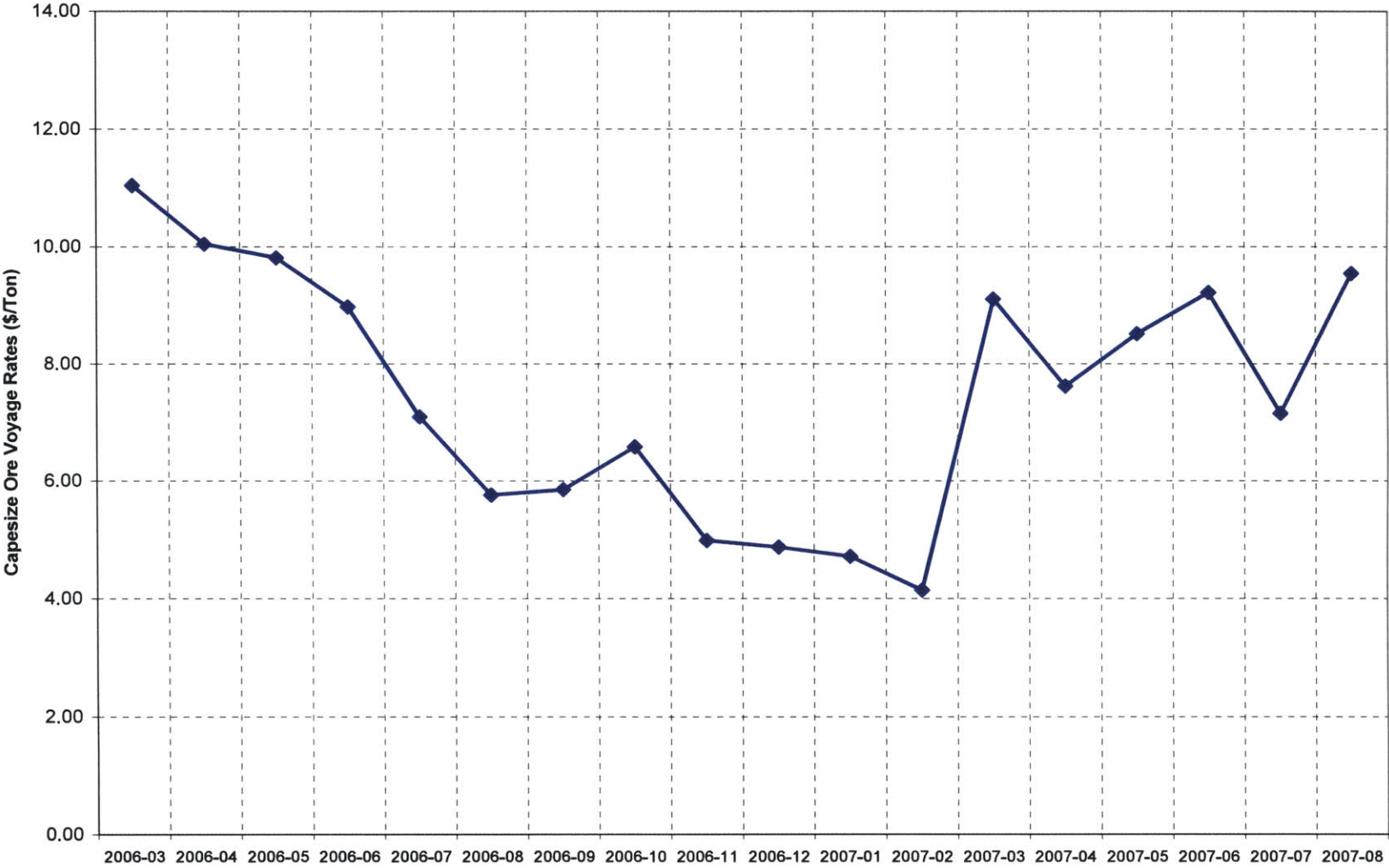
*Figure 6.20: August 2006 Capesize Ore Voyage Rates Forecast from ANN with 6 months Delay.*



*Figure 6.21: February 2007 Capesize Ore Voyage Rates Forecast from ANN with 12 months Delay.*



Figure 6.22: August 2007 Capesize Ore Voyage Rates Forecast from ANN with 18 months Delay.



# Chapter 7

## Conclusions

### 7.1 Introduction

The forecasting of financial figures always constitutes a challenge for researchers worldwide. The artificial neural networks provide this option since they are considered as the most reliable and effective tools of forecasting financial figures. Past and recent scientific papers showed that the use of Multi Layers Perceptron (MLPs) Networks is offering satisfactory results when they are used to forecast time series. In the current study, three important innovations were introduced, in an attempt to evolve MLP networks and develop new biologically modular artificial neural network architectures that accurately represent the shipping market.

### 7.2 Conclusions

To begin with, the forecast was accomplished by the use of a *Modular MLP network*, which is a special class of an MLP. This type of network processes its input using several parallel MLPs, and then recombines the results. This tends to create some

structure within the topology, which will foster specialization of function in each sub-module. In addition, modular networks do not have full interconnectivity between their layers. Therefore, a smaller number of weights are required for the same size network (i.e. the same number of processing elements, neurons). This tends to speed up training times and reduce the number of required training exemplars; an element that was extremely important since every time series had a population of 254 values.

However, a Modular MLP network requires several experiments so that the network's topology and the processing elements can provide accurate and consistent forecasts. This major disadvantage is resolved with the use of *genetic algorithms (GA)*. A genetic algorithm is used to optimize one or more parameters within the neural network. The most common parameters to optimize are the input columns, the number of hidden PEs, number of memory taps, and the learning rates. Therefore, a genetic algorithm will try various permutations of includes and skips among inputs, based on the performance of multiple training runs. The combination of inputs that produces the lowest error across these training runs will be used for the final model. The genetic algorithm that is being used in the current model is a conventional genetic algorithm which is inspired by the mechanism of natural selection where stronger individuals are likely the winners in a competing environment. Each one of the individuals has a certain number of characteristics. For these characteristics the term *genes* is used. According to the biological paradigm the set off all *genes* of an individual form its *chromosome*. A genetic algorithm creates an initial population (a collection of chromosomes) and then evaluates this population by training the neural network for each chromosome. It then evolves the population through multiple generations using the three basic genetic

operations of reproduction, crossover and mutation, in the search for the best network parameters.

Finally, we have attempted to improve the use of the *Backpropagation algorithm* with implementing the *Momentum Learning algorithm* that offers simplicity and efficiency to the training of the network. Backpropagation computes the sensitivity of the output with respect to each weight in the network, and modifies each weight by a value that is proportional to the sensitivity. The perfection of the procedure is that it can be implemented with local information and is efficient because it requires just a few multiplications per weight. However, since it is a gradient descent procedure and only uses the local information, it can get caught in a local minimum. The procedure is a little noisy since we are using a poor estimate of the gradient, so the convergence can be slow. Momentum learning is the improvement to the straight gradient descent in the sense that a memory term (the past increment to the weight) is utilized to speed up and stabilize convergence. In addition, it provides the gradient descent with some inertia, so that it tends to move along a direction that is the average estimate for down. The amount of inertia (i.e., how much of the past to average over) is dictated by the momentum parameter. The higher the momentum, the more it smoothes the gradient estimate and the less effect a single change in the gradient has on the weight change. In the end, the major benefit is the added ability to break out of local minima that the backpropagation might otherwise get caught in.

### 7.3 Additional conclusions

- Artificial Neural Networks, as forecasting models of time series, exhibit remarkable behavior. In addition, they present many advantages as well as a few disadvantages which have to do with the complexity of non linear forecasting models.
- The use of the above methodology is a revolutionary technique that finds direct application in econometrical systems such as the bulk carrier charter and spot market.
- The qualitative and quantitative enrichment of the database is expected to constitute a decisive factor for more precise forecasts. Qualitative enrichment means more time series, while quantitative means bigger in size and time of available values.
- The modeling of the forecasting process with the simultaneously use of genetic algorithms in Artificially Neural Networks is feasible. Moreover, it will give the researchers the possibility to export reliable forecasts in a quick rate.
- Finally, it will constitute a radical tool for decision-making for financial institutions as well as a strategic tool that may give answers to ship-owners.

## 7.4 Future Work

The performance of the proposed Artificial Neural Networks highlights the powerful potential of these approaches in the fields of forecasting shipping figures. However, there are still some aspects of these modes that can be improved. Some of these are outlined in the following as suggested future research work.

### 7.4.1 Input Data

In our study we made the assumption that all the information we need in order to forecast the Capesize Ore voyage rates from Tubarao to Rotterdam is included in the selected time series. *But is this assumption valid or there are other forms of input data that can offer extra information to our models?* Although we have gathered all those data we suspect that can influence the market, we could test the reaction of the network to different variations of these specific time series, like the subtraction of one value from another value or the variation of one time series from another one. In this way, we could observe the reaction of the network to small or big changes in one time series.

### 7.4.2 Learning Algorithms

An interesting and challenging idea would be the use of *different learning algorithms* for the different specialist modules. One example would be in the case where a major portion of a task can be learned offline and due to some change in the operating environments, a continuous adaptation of the weights of a module is required to compensate for the change in the operating environment. Then, the offline training of the

specialist modules can be carried out using efficient offline learning algorithms like Levenberg-Marquardt, whereas for the module for which a continuous adaptation is required, an online algorithm like standard backpropagation can be used.

#### **7.4.3 Use of more sophisticated genetic algorithms**

The use of more *advanced and sophisticated genetic algorithms* will offer to the neural networks better characteristics of training and will improve considerably the results of the forecast.

#### **7.4.4 Alternative methods of stopping training**

An important issue that would improve the precision of the forecasts is the development of an *alternative methodology for stopping the training*. More accurate results would be extracted and less time would be necessary for the accomplishment of the forecasting process.

## References

1. Azoff E. M. "*Neural network time series forecasting of financial markets.*" Chichester: John Wiley and Sons 1994.
2. Mitchell M.T. "*Machine learning.*" New York: The McGraw-Hill Companies 1997.
3. Gurney K. "*An introduction to neural networks.*" London: UCL Press 1997.
4. Bishop M. C. "*Neural networks for pattern recognition.*" New York: Oxford University Press 1996.
5. Goldberg E. D. "*Genetic algorithm in search, optimization, and machine learning.*" New York: Addison-Wesley 1989.
6. Mitchell M. "*An introduction to genetic algorithms.*" Cambridge Massachusetts: MIT Press 1997.
7. Koza R. J. "*Genetic programming, on the programming of computers by means of natural selection.*" Cambridge Massachusetts: MIT Press 1992.
8. Man F. K., Tang S. K. & Kwong S. "*Genetic algorithms: Concepts and designs.*" Heidelberg: Springer-Verlang 1999.
9. Igor Aleksander and Hellen Morton. "*An Introduction to Neural Computing.*" Second Edition. Chapman & Hall 1995.
10. Murray L. Barr & John A. Kiernan. "*The Human Nervous System. An Anatomical Viewpoint.*" Fifth Edition. Harper International 1988.

11. Laurene Fausett. *"Fundamentals of Neural Networks, Architectures, Algorithms, and Applications."* Prentice-Hall, New Jersey 1994.
12. Mohamad H. Hassoun. *"Fundamentals of Artificial Neural Networks."* MIT Press 1995.
13. Robert L. Harvey. *"Neural Network Principles."* Prentice Hall. New Jersey 1994.
14. Dan W. Patterson. *"Artificial Neural Networks. Theory and Applications."* Prentice Hall. Singapore 1996.
15. D.E. Rumelhart, G.E. Hinton, and R.J. Williams. *"Learning internal representations by error propagation."* In: D.E. Rumelhart, and J.L.McClelland (Eds.), *"Parallel distributed processing."* Volume I: Foundations. Cambridge, MA: MIT Press 1986.
16. Patrick K. Simpson. *"Artificial Neural Systems. Foundations, Paradigms, Applications, and Implementations."* Pergamon Press 1990.
17. Andreas Zell. *"Simulation Neuronaler Netze."* Addison-Wesley 1994.
18. N. K. Perugini and W. E Engeler. *"Neural network learning time: Effects of network and training set size."* Proceedings of the International Joint conference on neural networks, 2:395-401, 1989.
19. Martin Stopford *"Maritime Economics."* Second Edition, Routledge London 1997.
20. *"Lloyd's Shipping Economist Magazines."* London 1985-2006.
21. Clarkson Research Services Limited Ltd. *"Shipping Intelligence Network 2005."* (<http://www.clarksons.net/markets/>) London England 2006.
22. User's Manual *"NeuroSolutions"*, NeuroDimension, Inc. Gainesville, 2005.

Date	120K DWT Capesize Bulkcarrier Newbuilding Prices \$ Million	Capesize Scrap Value \$ Million	Dry Bulkcarrier Total Supply Million DWT	Dry Bulkcarrier Total Surplus Million DWT	Dry Bulkcarrier Total Demand	Bulkcarrier Orderbook Million DWT	1 Year Bulkcarrier Timecharter Rates 127,500 dwt \$/Day	Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t \$/Ton
1985-01	22.5	2.5	220.2	52.2	168.0	25.60	7,438	5.38
1985-02	22.5	2.5	221.3	53.2	168.1	24.30	6,850	5.03
1985-03	22	2.5	221	54.2	166.8	24.30	6,850	5.26
1985-04	22	2.5	222.4	53.8	168.6	23.00	6,725	6.25
1985-05	22	2.5	223.3	50.9	172.4	22.30	6,375	6.09
1985-06	22	2.5	224.8	49.9	174.9	22.70	6,250	5.31
1985-07	21	2.5	225	49.3	175.7	21.40	5,750	4.18
1985-08	20.5	2.5	224.4	48.2	176.2	20.40	5,000	3.78
1985-09	20.5	2.4	223.1	48	175.1	20.90	4,760	4.3
1985-10	20.5	2.4	222.8	47.5	175.3	20.20	5,125	4.83
1985-11	20.5	2.4	221.7	46.6	175.1	19.50	5,725	4.98
1985-12	20.5	2.4	222.6	47.3	175.3	19.60	6,250	5.14
1986-01	20.5	2.4	219.3	49.5	169.8	19.40	5,938	4.88
1986-02	20.5	2.4	220.3	49.1	171.2	18.50	4,900	4.1
1986-03	20.5	2.4	221.6	46.1	175.5	17.40	4,900	4
1986-04	22	2.4	220.5	45.1	175.4	16.30	4,863	3.45
1986-05	22	2.4	219.2	44.3	174.9	15.30	4,750	3.34
1986-06	24	2.5	216.1	43.8	172.3	15.90	4,550	3.24
1986-07	24	2.5	213.7	42.4	171.3	14.90	4,475	3.1
1986-08	24.5	2.5	212.4	40.8	171.6	14.40	4,560	3.24
1986-09	25.5	2.6	208.7	38.6	170.1	14.20	5,600	4.44
1986-10	25.5	2.8	209.1	40.3	168.8	13.30	5,810	4.2
1986-11	25.5	2.9	211.6	43.1	168.5	12.00	5,400	3.93
1986-12	25.5	3.1	212.9	43.3	169.6	11.50	5,033	3.7
1987-01	25	3.2	210.9	44.3	166.6	11.00	5,000	3.82
1987-02	24	3	211	45.1	165.9	9.50	4,850	3.76
1987-03	24	2.9	212.9	45.1	167.8	9.10	5,225	3.46
1987-04	26	3.2	215.2	44.3	170.9	8.20	6,133	4.33
1987-05	26	3.3	214.5	42.7	171.8	7.60	7,180	4.55
1987-06	26	3.4	215.4	39.6	175.8	8.10	7,500	4.5
1987-07	26	3.4	213.3	42.1	171.2	7.70	7,280	4.45
1987-08	26	3.6	213.1	43.9	169.2	7.20	9,500	8.56

Date	120K DWT Capesize Bulkcarrier Newbuilding Prices \$ Million	Capesize Scrap Value \$ Million	Dry Bulkcarrier Total Supply Million DWT	Dry Bulkcarrier Total Surplus Million DWT	Dry Bulkcarrier Total Demand	Bulkcarrier Orderbook Million DWT	1 Year Bulkcarrier Timecharter Rates 127,500 dwt \$/Day	Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t \$/Ton
1987-09	27	3.9	212.7	43.5	169.2	7.30	9,500	5.59
1987-10	30	5.2	213.1	43.5	169.6	6.90	10,400	5.9
1987-11	30	4.3	215.9	44.3	171.6	6.50	11,375	6.31
1987-12	30	4.1	217.6	43.6	174.0	7.30	11,000	6.85
1988-01	30	4.5	219	43.8	175.2	6.50	11,875	7.31
1988-02	30	5.2	219.6	43.5	176.1	6.30	12,625	7.25
1988-03	32	5.2	220.9	43	177.9	8.10	15,125	7.06
1988-04	33.5	5.2	223.5	41.7	181.8	7.90	15,250	6.63
1988-05	33.5	5.2	223.4	40.5	182.9	7.80	14,000	6.63
1988-06	35	5.2	222.1	40.3	181.8	9.90	12,938	5.38
1988-07	35	5.3	221.3	39.4	181.9	9.70	13,950	5.15
1988-08	35	5.4	219.9	36.2	183.7	9.40	13,563	6
1988-09	35	5.3	220	35.7	184.3	10.90	14,050	5.99
1988-10	37.5	5.4	221.7	37.1	184.6	10.50	13,625	6.34
1988-11	38	5.3	222.1	22.8	199.3	10.30	14,875	6.69
1988-12	38	5.3	219.9	19.6	200.3	11.70	15,250	6.75
1989-01	38	5.3	220.3	20.3	200.0	11.10	15,500	7.38
1989-02	40	5.3	223	18.1	204.9	10.50	16,000	7.44
1989-03	40	5.3	225.2	18.2	207.0	11.80	16,900	7.62
1989-04	40	5.3	225.9	17.8	208.1	11.10	17,000	8.44
1989-05	43	5.5	226.8	15.7	211.1	10.70	17,438	8.16
1989-06	43	5.8	225.8	15.7	210.1	12.80	16,470	7.3
1989-07	43	5.8	226.6	17.6	209.0	12.20	16,750	6.95
1989-08	43	5.8	226	16.7	209.3	11.80	17,375	7.44
1989-09	43	5.8	225.8	15	210.8	14.20	17,900	7.55
1989-10	43	5.3	226.3	13.7	212.6	13.80	17,813	7.63
1989-11	43	5.7	226.8	16.9	209.9	13.20	18,000	8.06
1989-12	45	5.7	225.9	18.7	207.2	13.90	18,000	7.6
1990-01	47.5	5.7	226.4	18.6	207.8	13.10	17,625	8
1990-02	47.5	5.7	227.8	18.4	209.4	12.10	16,600	7.68
1990-03	47.5	5.5	229	18.3	210.7	13.30	16,380	7.88

Date	120K DWT Capesize Bulkcarrier Newbuilding Prices \$ Million	Capesize Scrap Value \$ Million	Dry Bulkcarrier Total Supply Million DWT	Dry Bulkcarrier Total Surplus Million DWT	Dry Bulkcarrier Total Demand	Bulkcarrier Orderbook Million DWT	1 Year Bulkcarrier Timecharter Rates 127,500 dwt \$/Day	Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t \$/Ton
1990-04	47.5	5.7	229	19.2	209.8	12.20	15,000	7.19
1990-05	47.5	5.9	228.2	20	208.2	11.80	14,313	6.5
1990-06	47.5	5.7	227.4	18.2	209.2	12.80	13,700	5.81
1990-07	47.5	5.7	226.3	18	208.3	11.70	13,813	5.06
1990-08	50	5.3	227.8	19	208.8	10.80	13,600	5.24
1990-09	50	5.3	230	19.3	210.7	10.50	13,625	5.25
1990-10	50	4.5	230.7	20	210.7	10.30	11,875	5.69
1990-11	50	4.5	231.5	22.2	209.3	9.80	10,050	6.57
1990-12	50	4.3	230.2	21.3	208.9	8.90	10,003	7.05
1991-01	50	4.3	229.6	21	208.6	8.10	10,350	7.34
1991-02	45	3.8	231.2	21.1	210.1	7.60	11,625	7.83
1991-03	45	4.3	232.6	20.7	211.9	7.40	13,550	7.5
1991-04	45	4.3	234.2	20.3	213.9	7.00	13,638	7.46
1991-05	43.5	3.8	237.1	18.9	218.2	6.20	13,960	7.54
1991-06	43.5	3.6	236.5	20	216.5	8.00	14,375	7.19
1991-07	43.5	3.6	235.2	19.7	215.5	7.90	15,000	6.98
1991-08	43.5	3.4	235.4	19.4	216.0	7.40	15,550	6.42
1991-09	43.5	3.4	235.6	21.2	214.4	10.40	17,025	6.4
1991-10	43.5	3.4	236.5	20.9	215.6	9.90	17,438	6.3
1991-11	44.5	3.4	237.4	22	215.4	9.30	17,060	6.92
1991-12	46	3.2	238.7	22.7	216.0	13.30	16,500	6.59
1992-01	43	3.6	237.1	24.2	212.9	13.20	16,110	6
1992-02	41	3.7	237.9	25.1	212.8	12.90	14,375	4.91
1992-03	40	3.4	237.7	27.2	210.5	15.30	12,413	4.66
1992-04	40	3.4	236.6	25.8	210.8	14.50	9,550	4.23
1992-05	40	3.3	239.8	25.3	214.5	14.20	8,780	4.03
1992-06	39.5	3.2	237.7	25	212.7	16.00	8,700	4.08
1992-07	38.5	3.4	238.6	24	214.6	15.40	8,540	4.04
1992-08	37	3.16	238.5	24.1	214.4	14.80	8,500	3.98
1992-09	37	3.1	236.4	24.4	212.0	15.40	8,375	4
1992-10	37	3.16	235.4	24.2	211.2	15.00	8,240	3.86

Date	120K DWT Capesize Bulkcarrier Newbuilding Prices \$ Million	Capesize Scrap Value \$ Million	Dry Bulkcarrier Total Supply Million DWT	Dry Bulkcarrier Total Surplus Million DWT	Dry Bulkcarrier Total Demand	Bulkcarrier Orderbook Million DWT	1 Year Bulkcarrier Timecharter Rates 127,500 dwt \$/Day	Capesize ore Voyage Rates Tubarao/Rotte rdam 145,000t \$/Ton
1992-11	37	2.96	236.5	26.9	209.6	14.60	8,750	4.36
1992-12	36	3	234.5	25.8	208.7	15.50	9,125	4.4
1993-01	36	3	236.4	26.4	210.0	14.40	9,500	4.5
1993-02	35	3.1	236.1	26.1	210.0	13.60	9,763	4.58
1993-03	35	3.1	236.9	23.6	213.3	14.30	9,925	4.73
1993-04	35	3.34	237.7	23.4	214.3	13.90	11,300	5.44
1993-05	37	3.26	239.5	23.6	215.9	13.10	13,563	6.1
1993-06	40	3.26	238.8	22.5	216.3	16.10	14,500	6.23
1993-07	40	3.1	239.4	23.2	216.2	15.60	13,800	5.72
1993-08	40	3.1	239.5	23.4	216.1	14.90	13,125	5.34
1993-09	40.5	3.3	239.6	23.3	216.3	18.60	13,613	5.4
1993-10	40.5	3.26	239.4	22.4	217.0	18.20	12,950	5.24
1993-11	38.5	3.3	240.3	22.5	217.8	17.60	12,188	5.09
1993-12	38.5	3.2	239.1	22.2	216.9	20.70	11,320	4.55
1994-01	36	3.12	239	22.1	216.9	20.10	10,450	4.16
1994-02	36	3.14	239.5	21.2	218.3	19.10	9,825	4
1994-03	33	3.14	240.7	22.2	218.5	23.70	10,000	4.3
1994-04	33	3.14	242.4	22.3	220.1	22.20	11,670	5.07
1994-05	33	3.26	241.9	22.7	219.2	21.30	13,038	6.23
1994-06	34	3.2	241.8	21.8	220.0	23.80	12,875	5.68
1994-07	33	3.12	241.7	20.8	220.9	22.50	12,890	5.26
1994-08	33	3.2	242.8	18.1	224.7	21.50	13,975	6.53
1994-09	33	3.44	245.8	17.8	228.0	24.60	14,780	7.06
1994-10	33	3.44	247.2	18.3	228.9	23.10	15,700	7.83
1994-11	34	3.44	248.6	18.3	230.3	22.50	16,550	8.33
1994-12	34	3.7	248.8	17.4	231.4	28.50	16,890	8.58
1995-01	34.5	3.44	251.7	17.1	234.6	25.90	17,013	8.08
1995-02	34.5	3.48	251.4	17.3	234.1	24.80	17,088	7.38
1995-03	34.5	3.52	256.2	17.3	238.9	27.00	17,870	8.81
1995-04	34.5	3.46	256	16.9	239.1	26.40	18,925	8.99
1995-05	34.5	3.5	258.1	17.9	240.2	25.30	17,875	7.89
1995-06	34.5	3.4	258.4	16.9	241.5	26.00	15,600	7.04

Date	120K DWT Capesize Bulkcarrier Newbuilding Prices \$ Million	Capesize Scrap Value \$ Million	Dry Bulkcarrier Total Supply Million DWT	Dry Bulkcarrier Total Surplus Million DWT	Dry Bulkcarrier Total Demand	Bulkcarrier Orderbook Million DWT	1 Year Bulkcarrier Timecharter Rates 127,500 dwt \$/Day	Capesize ore Voyage Rates Tubarao/Rotte rdam 145,000t \$/Ton
1995-07	34.5	3.36	258.2	18.8	239.4	25.10	15,500	7.2
1995-08	34.5	3.56	254.9	21.3	233.6	24.50	17,063	7.56
1995-09	34.1	3.6	259.1	18.3	240.8	28.80	16,450	6.97
1995-10	34.1	3.6	256.8	21.3	235.5	30.60	15,000	5.9
1995-11	34.1	3.2	255.2	19	236.2	32.90	14,125	6.13
1995-12	34.1	3.2	256.2	17	239.2	35.04	13,300	5.72
1996-01	33.28	3.5	259.7	11.4	248.3	34.69	11,438	5.61
1996-02	33.28	3.54	259.6	17.1	242.5	33.70	10,200	5.59
1996-03	33.28	3.74	259.6	20.5	239.1	33.40	10,790	5.48
1996-04	33.28	3.88	257.7	20.5	237.2	33.82	10,750	5.13
1996-05	32.47	3.8	257.4	20.8	236.6	34.18	9,000	5.08
1996-06	32.47	3.88	259.1	23.2	235.9	33.57	8,688	4.64
1996-07	32.06	3.9	260.3	22.4	237.9	32.80	8,013	4.35
1996-08	32.06	3.78	255.7	20.7	235.0	32.62	8,070	5.03
1996-09	32.06	3.3	258.8	20.3	238.5	32.54	8,500	4.99
1996-10	32.06	3.22	260.1	25.8	234.3	31.17	8,125	4.81
1996-11	32.06	3.2	256.4	4.6	251.8	30.74	9,120	5.74
1996-12	31.66	3.1	255.3	5.1	250.2	29.88	9,913	5.58
1997-01	30.85	3.06	258.3	7.1	251.2	29.52	10,040	5.97
1997-02	30.85	3.26	254.42	17.2	237.2	26.73	11,188	5.61
1997-03	30.85	3.4	261.2	16.7	244.5	26.90	10,575	5.45
1997-04	31.66	3.4	261.8	16	245.8	24.58	10,000	5.45
1997-05	31.66	3.34	261.9	16.5	245.4	25.16	10,100	5.46
1997-06	33.28	3.12	261.4	17.2	244.2	25.24	10,713	5.71
1997-07	33.28	3.18	261.5	17	244.5	24.56	11,500	5.55
1997-08	33.28	3.36	262.1	8.2	253.9	24.84	12,310	5.85
1997-09	34.09	3.38	262.8	4.5	258.3	25.03	12,500	5.94
1997-10	34.09	3.4	262.3	5.5	256.8	24.89	12,700	5.92
1997-11	33.69	3.2	262.3	4.6	257.7	25.75	12,425	6.13
1997-12	32.88	3.1	262.9	4.9	258.0	25.01	12,138	6.05
1998-01	32.47	3.08	263.3	5.3	258.0	26.01	11,350	6.04
1998-02	31.66	2.9	261.9	4.3	257.6	26.06	9,063	5.05

Date	120K DWT Capesize Bulkcarrier Newbuilding Prices \$ Million	Capesize Scrap Value \$ Million	Dry Bulkcarrier Total Supply Million DWT	Dry Bulkcarrier Total Surplus Million DWT	Dry Bulkcarrier Total Demand	Bulkcarrier Orderbook Million DWT	1 Year Bulkcarrier Timecharter Rates 127,500 dwt \$/Day	Capesize ore Voyage Rates Tubarao/Rotterdam 145,000t \$/Ton
1998-03	31.25	2.9	265	5	260.0	25.31	7,438	4.3
1998-04	30.85	3.04	259.6	4.9	254.7	24.93	6,813	3.91
1998-05	30.85	2.78	259.7	4.9	254.8	24.96	7,500	4
1998-06	30.04	2.4	259.3	6.3	253.0	24.37	6,063	3.61
1998-07	30.04	2.4	258.2	6	252.2	24.73	4,750	3.2
1998-08	28.82	2.36	256	6.4	249.6	24.96	5,063	3.31
1998-09	28.82	2.36	252.6	6.4	246.2	24.51	7,438	4.13
1998-10	28.41	2.28	250.4	5.8	244.6	23.98	6,900	4.18
1998-11	28.41	2.08	250.1	6.9	243.2	24.00	6,125	3.64
1998-12	26.79	2.12	249.3	7.4	241.9	24.18	6,950	4.2
1999-01	26.38	2.24	248.8	6.4	242.4	23.81	5,625	3.38
1999-02	25.98	2.2	247.2	8	239.2	23.22	5,875	3.45
1999-03	25.98	2.32	246.8	8	238.8	22.64	7,188	3.7
1999-04	25.98	2.26	244.9	6.5	238.4	21.19	6,800	3.6
1999-05	27.19	2.26	244.4	8.3	236.1	21.09	6,938	3.8
1999-06	28.01	2.18	244.6	6	238.6	21.65	6,000	3.3
1999-07	28.41	2.26	244	6.1	237.9	22.17	6,250	3.46
1999-08	29.22	2.3	244.2	6.7	237.5	23.75	8,688	4.19
1999-09	29.22	2.48	247	6.2	240.8	24.60	9,188	5.15
1999-10	29.22	2.52	246.2	5.8	240.4	26.20	10,500	6.15
1999-11	27.6	2.84	244	10	234.0	28.50	10,500	6.05
1999-12	28.41	2.8	247.3	4.3	243.0	30.58	10,450	6.52
2000-01	30.04	2.78	246.1	4.8	241.3	33.50	9,938	6.18
2000-02	30.44	2.56	245.1	4.7	240.4	34.28	10,688	6.69
2000-03	30.44	2.8	246.4	5	241.9	34.95	12,650	7.58
2000-04	30.44	2.9	245.9	4.3	241.6	35.57	12,563	7.49
2000-05	30.85	2.92	245.8	4.3	241.5	36.25	12,188	7.14
2000-06	30.44	3.06	246	4	241.6	36.69	12,250	6.99
2000-07	31.25	3.26	245.9	4.5	241.4	36.71	12,688	7.5
2000-08	31.66	3.36	246.1	4	241.8	36.54	13,813	7.59
2000-09	31.66	3.32	246.7	4.4	242.3	36.42	13,800	7.81
2000-10	32.47	3.4	247.7	6	242.2	35.06	14,000	8.23

Date	120K DWT Capesize Bulkcarrier Newbuilding Prices \$ Million	Capesize Scrap Value \$ Million	Dry Bulkcarrier Total Supply Million DWT	Dry Bulkcarrier Total Surplus Million DWT	Dry Bulkcarrier Total Demand	Bulkcarrier Orderbook Million DWT	1 Year Bulkcarrier Timecharter Rates 127,500 dwt \$/Day	Capesize ore Voyage Rates Tubarao/Rotte rdam 145,000t \$/Ton
2000-11	32.88	3.24	249.1	3	246.0	35.40	13,813	8.65
2000-12	32.88	3.38	247.9	3.4	244.5	35.12	12,400	7.94
2001-01	32.47	3.54	249.7	3	246.5	35.39	11,375	7.26
2001-02	32.47	3.4	248	4	244.4	33.81	9,813	6.64
2001-03	33.69	3.36	249.2	3.5	245.7	32.86	9,020	6.03
2001-04	32.88	3.56	251.6	3	248.2	30.77	8,900	5.88
2001-05	32.88	3.4	252.1	3	249.4	31.19	9,563	6.59
2001-06	33.28	3.52	252.1	3.4	248.7	29.76	9,750	6.27
2001-07	33.28	3.04	254	3	250.8	28.23	7,313	5.26
2001-08	29.63	3.02	253.8	3	250.8	27.86	6,000	4.37
2001-09	30.44	3.06	252.8	3.3	249.5	27.30	5,063	4.23
2001-10	30.44	2.8	255.3	2.9	252.4	25.09	4,563	3.97
2001-11	30.04	2.56	255.8	3.1	252.7	24.75	4,500	3.76
2001-12	29.22	2.7	254.3	2.9	251.4	24.15	4,625	3.91
2002-01	28.82	2.76	255.8	3.4	252.4	23.78	5,125	4.16
2002-02	28.61	2.74	256.5	2.9	253.6	21.73	5,688	4.51
2002-03	28.61	2.84	256.2	3.0	253.2	21.38	7,000	5.04
2002-04	28.01	2.9	257.4	2.7	254.7	20.79	7,750	4.9
2002-05	28.01	2.7	258.1	2.9	255.2	23.71	7,550	4.44
2002-06	28.01	2.76	258	2.9	255.1	22.77	7,750	4.65
2002-07	27.6	2.9	257.7	2.9	254.8	24.89	7,500	4.6
2002-08	27.6	2.8	255.9	2.9	253.0	24.95	7,500	4.74
2002-09	28.41	3	255.8	2.4	253.4	25.61	8,750	6.04
2002-10	28.82	3.2	258.7	2.2	256.5	27.71	9,563	6.36
2002-11	29.22	3.4	255.9	2.4	253.5	29.27	10,700	6.93
2002-12	29.43	3.56	260	2.7	257.3	29.44	12,063	8.15
2003-01	30.44	3.56	275	2.8	272.2	31.66	12,000	8.1
2003-02	30.85	3.7	290	2.5	287.5	34.46	12,375	7.98
2003-03	31.25	4.16	295	3.0	292.0	35.71	13,000	8.85
2003-04	31.25	4	300.45	3.1	297.4	37.59	13,000	9.21
2003-05	31.25	3.96	301.45	3.1	298.4	39.40	14,400	9.9
2003-06	31.25	4.3	302.66	2.7	300.0	40.64	14,375	9.64

Date	120K DWT Capesize Bulkcarrier Newbuilding Prices \$ Million	Capesize Scrap Value \$ Million	Dry Bulkcarrier Total Supply Million DWT	Dry Bulkcarrier Total Surplus Million DWT	Dry Bulkcarrier Total Demand	Bulkcarrier Orderbook Million DWT	1 Year Bulkcarrier Timecharter Rates 127,500 dwt \$/Day	Capesize ore Voyage Rates Tubarao/Rotte rdam 145,000t \$/Ton
2003-07	31.66	4.8	303.55	3.0	300.5	44.74	14,438	9.81
2003-08	32.88	4.6	303.5	2.5	301.1	47.87	15,580	11.44
2003-09	34.5	4.6	304.59	2.6	302.0	47.27	17,625	11.91
2003-10	35.31	4.7	304.04	3.1	301.0	48.34	25,800	17.53
2003-11	38.96	5	305.17	3.3	301.8	51.69	28,000	17.95
2003-12	38.96	5.2	305.1	3.2	301.9	52.15	28,000	19.36
2004-01	40.59	5.8	305.8	3.3	302.5	56.30	33,000	23.57
2004-02	40.59	6.5	305.96	3.6	302.4	56.18	34,500	24.25
2004-03	42.62	6.9	308.07	2.7	305.4	56.53	31,000	21.19
2004-04	45.46	7.5	309.65	2.6	307.0	59.03	29,000	18.19
2004-05	44.65	7.24	310.64	2.2	308.5	62.28	22,250	14.86
2004-06	45.46	5.6	312.13	2.9	309.3	62.61	18,750	12.13
2004-07	46.68	7	313.53	2.6	310.9	63.37	26,400	16.96
2004-08	47.89	7.2	315.03	3.6	311.5	65.14	27,000	19.23
2004-09	49.52	7.24	316.29	2.8	313.5	65.11	26,000	18.41
2004-10	51.14	7.44	318.27	3.1	315.2	66.16	28,600	19.2
2004-11	51.95	7.44	319.82	2.4	317.4	67.60	34,250	24
2004-12	51.95	7.44	320.61	2.0	318.6	67.76	34,600	23.8
2005-01	53.58	7.8	322.90	2.1	320.8	69.84	33,000	19.81
2005-02	53.58	8.2	324.42	4.3	320.1	72.08	40,000	20.31
2005-03	55.2	8.2	325.81	3.7	322.1	72.78	32,500	17.94
2005-04	55.2	8.2	328.19	2.5	325.7	72.66	29,400	21.3
2005-05	55.2	7	330.31	3.1	327.2	72.09	23,500	17.46
2005-06	50.33	6.1	331.92	3.7	328.3	71.67	17,000	11.44
2005-07	50.33	5.6	333.58	4.2	329.4	71.08	15,600	10.94
2005-08	48.3	5.6	335.93	3.8	332.1	70.24	14,500	11.14
2005-09	47.89	7.1	338.39	2.8	335.6	70.08	17,100	14.54
2005-10	47.89	6.4	340.38	3.7	336.7	68.69	19,750	16.5
2005-11	47.89	6.2	341.84	4.1	337.8	67.87	19,625	14.73
2005-12	47.89	6.5	342.97	4.2	338.8	66.57	17,700	12.43
2006-01	47.89	6.5	343.35	4.3	339.0	67.26	13,750	10.88
2006-02	47.89	6.5	344.87	3.8	341.1	64.90	14,250	13.06

2484-89