FORECASTING OF THE SHIP DEMOLITION MARKET USING ARTIFICIAL NEURAL NETWORKS

FARSHID KHALILI

A THESIS SUBMITTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
AT

NEWCASTLE UNIVERSITY

SCHOOL OF MARINE SCIENCE AND TECHNOLOGY

JUNE 2008

For my wife Marjan

ABSTRACT

Each section of the shipping market including the Newbuilding, Freight, Second-hand and Demolition markets has its own unique structure and individual internal parameters. Internal parameters can influence one or more parameters in their own and other markets. This makes the shipping markets, and each of their sections, a complex environment. Additionally, some external elements, such as inflation, political issues and economic policies, will affect certain outcomes. In such an environment, the main problem for creation of a "market model" is to recognise the most effective and influential input parameters on a set of desired outputs whilst considering the time-dependant nature of the data.

In this study, the traditional multivariate analysis methods have been implemented to try and create the best model of the Demolition market and use the created model to forecast the market. However, the accuracy of the model is poor. Then a new approach, based on the Artificial Neural Networks (ANN) methodology, has been implemented to model the market and consequently forecast the market.

Both static and dynamic ANNs were implemented, trained and tested for various internal and external inputs and the desired outputs of the Demolition market to find out the best combination of various elements. Performance of the network, in terms of Mean Square Error (MSE) and correlation coefficient (r), has been measured and compared for every individual structure and consequently the best functional relationship has been identified. In addition, the sensitivity of different parameters has been identified and the effectiveness of the input parameters demonstrated.

The results of the studies indicate that it is feasible to implement a suitable Neural Network architecture to map the inputs and outputs accurately and establish a usable "Ship Demolition Model". The model produced good results and can explain the complex structure of the Demolition market and identify and validate the main inputs which can alter market trends. The performance of the model has also been measured for forecasting three months ahead of the market and it shows a reasonable accuracy.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

The author is grateful for the advice, help and support from the Department of Marine Science and Technology at the University of Newcastle upon Tyne.

The author is particularly indebted to:

Prof. G. J. Bruce	for supervising this research, his assistance, encouragement and support. His knowledge of the shipping industry has been a great asset in this research.
Prof. E. Mesbahi	for supervising this research, his help and guidance throughout this research.
Prof. R. Birmingham	for his support, guidance and involvement with many aspects of this research.
Prof. A. Incecik	for his invaluable support during times of personal crises.
My office	for always being there for me and offering friendship and support during my research.

Finally, I would like to thank my wife and my family for all their love, patients and support during the years of research.

CHAPTER 1

INTRODUCTION

1.1 INTRODUCTION

The maritime market, like any other market, is difficult to predict and the number of variables and unclear relationship between different variables in the market makes it more complicated to be as simple as a predictable rule, or defined as an equation together with some inputs and outputs.

In general, the demolition market, as a section of the maritime market, deals with the ships which are obsolete. Their useful life is over and they can not be employed for other purposes e.g. conversion or donation. They have to be scrapped in a shipbreaking yard, usually located in Asia. The condition of these shipbreaking yards is rather harsh for the people who are working there and it is also unpleasant for the environment. Different hazardous substances are released during the process of scrapping and are devastating to the environment. They damage not only local environment, but also the whole wider ecosystem. This is a serious threat worldwide.

Currently, most of the scrapping processes are done manually with no particular attention to serious pollutants and no concern about working health and safety. Environmental issues are likely to lead toward having a more efficient scrapping industry. Reforms of the scrapping technology will be necessary to prevent more pollution and minimise the damage, but they also make the industry more restricted and controlled. Therefore, a reliable plan is needed to take into account of all aspects of the demolition including environment, safety, location of the scrapyards and time management for scrapping activities.

Reforms require investment and a proper action plan can minimise capital investment and other expenditure. Any information on the future scrapping market could lead to a more accurate and realistic plan. As an example, if the number of ships which might be sent to a particular scrapyard could be identified, the volume of hazardous waste could be calculated for a period of time in the future. It would then be possible to build a proper storage tank to use as a temporary storage space for those hazardous wastes.

Reform of the scrapping industry will affect the whole maritime market. The demolition market is going to be more structured and consequently the maritime market will be changed as a result.

This research has aimed to forecast the future of the demolition market which is a section of the maritime market. The main focus of interest of this research is to predict the monthly ship scrapping rate and also the price of ship scrapping for up to three months ahead.

In this respect, various parameters of the maritime market, which may influence the two above variables, are taken into account to find out a suitable model for the market.

Firstly, the conventional multivariate statistical methods of prediction are implemented to investigate the market and produce a suitable model for prediction. Secondly, the Artificial Neural Networks (ANN), as a powerful computer tool, are used to learn the fundamentals and relations of the various elements of the market. In this way, different network architectures with various specifications are trained, tested and compared to find out the best ANN model for each case. Subsequently, the performance of the obtained ANN model for future predictions are measured and compared with previous multivariate methods. Furthermore, sensitivity analyses are also carried out to identify the most influential parameters to each ANN model produced.

The first part of this chapter is dedicated to address some of the fundamental issues in the shipping industry and also to illustrate the position of ship demolition in the maritime market. This is followed by identification of the main characteristics of ship scrapping and explanation of the difficulties and its environmental problems in the second part. It shows the importance of having a proper plan for ship demolition as well. The third part is provided to explain the overall abilities and flexibility of the ANNs to tackle such problems and find out how it is suitable to be used, as a tool, to predict the future of the demolition market.

1.2 BACKGROUND

The shipping industry is influenced by many factors. Some of these factors are internal and dependent on physical and technical criteria. Others are external which can influence the shipping industry but they are not inside the industry, e.g. economic policies, inflation and political issues. Collating all the factors into a model to predict the future is complicated and difficult because some of the factors like political issues are unpredictable.

There are also a few time lag parameters which are notably considerable in the shipping industry: For a ship owner, the time lag from ordering a new ship to its delivery is varied and depends on the ship type; for a shipbuilder, the time lag from ordering a required amount of the steel to the time that it is ready and delivered. In both situations a realistic estimation of the time lags can determine the success or failure of the business plan.

The shipping market has been divided into four individual markets (Stopford, 2003):

- Newbuilding market
- Freight market
- Sale and purchase market
- Demolition market

Changing a parameter in each of these four markets can not only affect the other factors in the same market, but also can influence one or more variables in the other markets. The newbuiding market deals with the orderbooks and new ships, and is highly dependant on steel production as the basic building material of every ship. Almost 90% of a ship's weight is steel. Therefore, newbuilding prices can be varied as a result of the steel price variation.

The current fleet and the number of ships recruited can affect the freight rate itself and the freight rate market: if there are a limited number of ships to transfer the commodities in a particular route, freight rates go up due to lack of cargo space, and vice versa.

The challenge in the sales and purchase market for the owners is to distinguish the best time to sell the ship, considering the second-hand prices, freight rate and the national and international regulations and surveys. Similarly, the challenge for buyers is to find out the right time to buy a new or a second-hand ship, taking into account of the time lag, ships' age and the freight rate.

The demolition market deals with the scrapped steel which is convertible for new steel production. Therefore, the demolition market can drive the rate of steel production and, as a result, the steel price. For example, during a recession if the freight rates are low and the operating cost of the ship is high, due to the bad economic situation, it is possible that a large number of the owners prefer to send their ships to the breaking yards to earn money instead of losing it gradually. This event can increase the production of the scrap steel dramatically. Technical and economical obsolescence are other reasons to send a ship to a breaking yard, because an old ship with old technology can not compete with the new high technology ship in terms of operating costs, cargo space, cargo handling and the time spent in ports to load and discharge cargo. Therefore, a technology leap in cargo handling or any other related matter could be a reason to have more ships in the breaking yards.

1.3 PURPOSE

There are two separate reasons for carrying out this study which are both financially important to all the investors. They are also important for the future environmental reformation of the demolition industry and scrapyards.

1.3.1 IN THE MARKET

There has been a lift in the underlying long-term growth in the world economy, up from 3.3 percent per year on average in the 1990s to 3.9 percent so far in this decade (Platou, 2006). World economy growth can be a reason for the growth of the seaborne trade and larger maritime market as a result. Also the distribution of this growth is in favour of the seaborne trades because of the presence and the location of countries like China. A significant part of the shipbuilding costs is related to the steel price which has been booming in the last two years because of the higher demand of the steel. The price index for heavy steel plates peaked in May 2005 at $667 per ton and ended the year 15 percent below at $575/mt (Platou, 2006).

Compared to the other shipping markets, the ship demolition market is a less glamorous but essential part of the business (Stopford, 2003). In 2004, a massive scrapping rate had been anticipated for 2005, because of the International Maritime Organisation (IMO) regulation for the cut-off date of Category 1 tankers in April 2005. It did not happen and only 6.4 million dead weight (dwt) were sold for demolition, compared with 11.8 million dwt in 2004. On the tanker side zero VLCCs, 3 Suezmaxes and 18 Aframaxes were scrapped compared to 4 VLCCs, 9 Suezmaxes and 27 Aframaxes in 2004. This was because of the strong tanker market, resulting with many vessels being converted to Category 2 (Platou, 2006).

The shipowners come to the demolition market to offer a ship which they cannot sell as second-hand for continued trading. The customers, in this market, are shipbreakers who buy and then transfer the ship to a scrap yard. It is clear that both the shipowner and the shipbreaker are looking for the most profitable deal and, considering the high price of the ships, a slight change of the contract could be a big change for the financial success or failure. They must have extensive knowledge not only about the demolition market, but also about the other shipping markets, to find out the most profitable moment for buying or selling the ship.

This needs a comprehensive study in all aspects of the business including the different parameters of all the markets, then combining all the variables to find out the results and making the final decision. Furthermore, there are some external factors that should be considered as well. These parameters are generally unpredictable and almost impossible to forecast.

In this way, time is also a very important variable. Mostly, there is not much time to think because the deal will be cancelled. Therefore, it is vital to carry out a decision in time. All the people who are engaged in this process must decide as soon as they can, and the person with the least decision time will win the deal.

In such a competitive environment the winners are people who have more accurate and quicker tools to calculate all the possibilities in a short time. Furthermore, if they could have an acceptable estimation for the future of the market, i.e. the price or the rate of scrapping, they can arrange a more profitable plan for their company. A shipowner can lay up his/her ship for a few months if he/she knows that there will be a better price in the coming months, or vice versa. A shipbreaker may be able to wait a couple of months to get a cheaper deal which can be key to survival in the business.

1.3.2 ENVIRONMENTAL PURPOSE

A ship has a huge structure and its decommissioning generally takes place in different stages on the shallow water, shore and the beach. Scrap steel represents the largest recyclable fraction from the vessel and is commonly classed as ferrous scrap. There are non-ferrous materials as well, e.g. the accommodation which it could be possible to sell or to be reused.

During the shipbreaking process various chemicals and hazardous wastes are generated e.g. poly chlorinated biphenyls (PCBs), tributyl tins (TBTs), mercury and asbestos. All the toxic substances released into the environment cause severe damage and harm, and pollute the landfill.

Figure 1-1: A shipbreaking yard in Bangladesh; this picture shows the manual process of ship scrapping and its harsh working condition. It carries out in a highly polluted and dangerous site. (October 2005 ©RUBEN DAO/GP/FIDH)

PCBs are a mixture of individual chemicals which are either oily liquids or solids that are colourless to pale yellow. Some PCBs can exist as a vapour in air. Health effects that have been associated with exposure to PCBs include acne-like skin conditions in adults and neurobehavioral and immunological changes in children. PCBs are known to cause cancer in animals (European Commission, 2003).

TBT is one of the most poisonous substances being released into the aquatic environment today and has been used in most of the world's marine paints to keep barnacles, seaweed and other fouling organisms from clinging to ships. It is also used in wood treatment and preservation, water and refrigeration systems. The following damages associated with the use of TBTs (European Commission, 2003):

- The reduction in shellfish stocks on a widespread geographic basis around the world
- The documented discovery of imposex[1] in as many as 150 species of marine snails, with the exact number of organisms affected unknown
- Shell deformity effects and larval mortality in aquatic organisms

[1] A pseudo-hermaphroditic condition in female gastropods (snails) caused by TBTs.

- Corresponding financial losses suffered by the aquaculture industry and costs imposed on the harbour authorities

Mercury is a naturally occurring heavy metal. At ambient temperature and pressure, mercury is a silvery-white liquid that readily vaporises and may stay in the atmosphere for up to a year. Mercury is highly toxic, especially when metabolised into methyl mercury. It may be fatal if inhaled and harmful if absorbed through the skin. Around 80% of the inhaled mercury vapour is absorbed in the blood through the lungs. It may cause harmful effects to the nervous, digestive, respiratory, immune systems and to the kidneys, besides causing lung damage (Agency for Toxic Substances and Disease Registry, 1999).

Asbestos fibres can enter the air or water from the breakdown of manufactured asbestos products. Asbestos fibres do not evaporate into air or dissolve in water. Small diameter fibres and particles may remain suspended in the air for a long time and be carried long distances by wind or water before settling down. Exposure to asbestos usually occurs by breathing contaminated air in a shipbreaking yard. Asbestos exposure can cause serious lung problems and cancer (Agency for Toxic Substances and Disease Registry, 2001).

The work conditions in a shipbreaking yard, with such a harsh environment, are commonly below the standard because there are numerous safety issues which are very difficult to observe and the workers put their life at risk to work in a shipbreaking yard. Therefore, ships that are ready for scrapping are often sent to demolition under conditions that would not be accepted in developed countries with respect to environmental and health and safety conditions of the work. They are sent to Asian countries like Pakistan, India, Bangladesh or China to dismantle. But the environmental problems are not local or national problems and hazardous wastes can move from site to site and move around the globe to contaminate the land, air and water. If a ship has been transferred to the other part of the world to scrap at its end of life, it is just like to moving the release point of those hazardous substances but they still exist.

Producing steel from the scrap is a sustainable process in terms of energy efficiency and recycling, but it is a serious threat for the environment and future generations. It is not rational or possible to stop scrapping activities but something has to be done to protect the environment and minimise the damaging effect of scrapping on the environment.

There are various regulations, legislations, national conventions and international conventions to define a framework to dispose these materials in a proper manner and also to restrict the dangerous and unnecessary scrapping activities. Because of the importance of these considerations, there are a growing number of legislations which are getting tougher and creating more restrictions for the scrapyards.

Demolition technology has to improve and adapt to the present and future circumstances. Manual scrapping process, as an inefficient and dangerous procedure, should convert to a set of mechanised and automated demolition procedures. In this respect, significant reforms have to be carried out to control all aspects of the demolition process without destruction of the demolition market and the other maritime markets.

A clear idea about the market and the monthly amount of the scrapping or any other linked information can be a guiding star throughout this development. It is obvious that any technical changes or technological reforms require short-term and long-term investments. But how should investment be done? How is it possible to distinguish the proper time for investment? Where is the best place to use the capital investment? In the light of prediction, the answers to such questions are more accurate and precise i.e. the prediction of the scrapping rate makes it possible to justify the investment.

To tackle the environmental problems of scrapping, it is important to have a clear and realistic plan for the future of the demolition with respect to the quantity and location of the demolitions. Supply-demand combination should be studied carefully, considering the demolition market, scrap prices and the ship's age. It is vital to look at the present and the future capacity of the shipbreaking yards, and find out if there is not a proper balance between the capacity of the shipbreaking yard and the demand for the demolition. Both sides of the balance cause serious problems for the environment. Lack of scrapyards can cause various problems due to storage of the old ships and the excess of them is hard to control and facilitate. More data and information about the various parameters, which are affecting the demolition market and the balance between supply and demand in the market, means a better plan for scrapping and less environmental problems as a result.

1.4 Artificial Neural Networks (ANN)

ANN is a technique which can be used as a powerful computational tool in a diversity of applications including: function approximation, pattern recognition, classification and time series forecasting. The literature on market forecasting using ANNs falls into two groups. The first group reports that ANNs were not useful and were unable to improve the explanation of the residual variance as compared to traditional linear models (Faber and Sidorowich, 1988). The second group reports that ANNs provided excellent techniques for enhanced forecasting. In recent years, many studies have come to a conclusion that the relationship between the financial and economic variables in a market is nonlinear, and that ANNs can be accurately used to model problems involving nonlinearities (Abhyankar et al., 1997). Nevertheless, not many successes were reported in full detail, probably due to commercial reasons or competitive advantage factors. This study will examine the demolition market, using ANN methodology, to investigate the market and obtain an appropriate model. Consequently, the model is evaluated and the performance of this technique compared against the conventional multivariate statistical methods.

Compared with the traditional statistical prediction methods Artificial Neural Networks can be applied to time series modelling without assuming a priori function forms of models. ANNs, with hidden units, are universal approximators which mean that they are capable of learning an arbitrarily accurate approximation to any unknown function, provided that they increase in complexity proportional to the size of the training data (Hu and Hwang, 2003).

In the real world, many time series are generated by nonlinear mechanisms and therefore in many applications the choice of a nonlinear model may be necessary to achieve an acceptable performance. Forecasting of the various parameters in the demolition market is an example of a nonlinear problem which is challenging due to the large number of the influential parameters, high noise, non-linearity and non-stationary environment. Hence, the choice of the model has crucial importance[1] and practical applications have shown that nonlinear models can offer a better prediction performance than their linear counterparts (Mandic and Chambers, 2001). They also reveal rich dynamical behaviour, such as limit cycles, bifurcations[2] and fixed points (points that are mapped to themselves by the functions) that cannot be captured by linear models (Gershenfeld and Weigend, 1993). The model-free nature of the ANN besides the essentially nonlinear structure of neural networks is particularly useful for capturing the complex underlying relationship of the demolition market. ANNs are versatile methods for forecasting applications in that not only can they find nonlinear structures in a problem, they can also model linear processes.

The choice of which neural networks to employ to represent a nonlinear physical process depends on the dynamics and complexity of the network that is best for representing the problems in such a market. The demolition market is being influenced by a large number of parameters which are generally unpredictable or very difficult to forecast. ANN can be a convenient computational tool to model such a market and find out the relation between the various inputs and outputs but it is important to choose a correct structure for the neural networks, based on the nature of the data and the time series. Each ANN has various variables which are sensitive and need to be adjusted carefully before and during the learning process. Changing these variables can change the general structure of the neural network and therefore different outputs will be produced. The variable selection is a critical factor in maximising a neural network's forecasting performance. A well trained ANN is also able to identify dependant and independent variables of the model and compare sensitivity of the inputs for producing favourable output variables.

---

[1] System identification, for instance, consists of choice of the model, model parameter estimation and model validation.

[2] Bifurcation occurs when a small smooth change made to the parameter values of a system causes a change in its long-term dynamical behaviour.

1.5 OVERVIEW OF THE RESEARCH

The demolition market, like other maritime markets, has its own internal factors and parameters, e.g. labour costs or scrap price, which directly affect the structure of the market. In addition, there are external factors that affect the market e.g. the political issues, inflation or the governmental decisions. The interaction between these internal and external elements is not the whole system which can define the demolition market. There are also some additional complex extra factors in the other maritime markets, including newbuildings, second-hand and freight rate market, which can affect the demolition market. For example, the freight rates or second-hand prices can change the scrapping price or the number of newly built tankers could have an effect on the demolition market.

The IMO regulations and legislations are the other issues that make the maritime market more complicated. These rules cause more restrictions in all sections of the market, especially the demolition market, because all scrapyards will face more pressure, in terms of the environment and safety issues, in the near future. These kinds of rules are getting tougher as the environmental issues are getting more serious. Therefore, the interaction between the different complex factors of the maritime markets and the external parameters make complex connections which are impossible to model. However it is important to identify the relation between the various variables to find out the end result. ANNs are a new approach to this sort of complex problem and their use can be a convenient method to tackle these difficulties. ANNs can learn the structure of the markets based on the past data. Since they have learnt the fundamentals, they are capable of prediction. It is important to use suitable neural networks, with a proper architecture, for every specific reason i.e. it is not possible to use one ANN, with a particular specification, for any purpose.

The numbers of 33 different time series are used for the investigations and analyses in this research. They are monthly data from January 1995 to the end of 2004. This means there are 120 patterns of data available for each analytical method. Inputs and outputs in each stage are varied and depend on the individual prediction purposes. The favourable outputs are monthly scrapped tonnage and scrap price in two locations, Far-East and Subcontinent.

Conventional multivariate statistical methods are employed to investigate the above data and produce a suitable model for prediction. Then the Artificial Neural Networks are used to produce the ANN model for the same data. Both static and dynamic ANNs are implemented, trained and their performances tested for various inputs and the desired outputs. The best designed architecture in each stage is identified and the performance of the model for forecasting the three months ahead of the market is measured in terms of Mean Square Error and correlation coefficient. Sensitivity analysis is also carried out for each model to identify the most influential variables. Finally the obtained model for both multivariate and ANN methods are compared.

CHAPTER 2

SHIP DISPOSAL

2.1 Introduction

In this chapter, the process of the ship scrapping is explained and its various methods are introduced. Scrap steel, as the main reason for the ship scrapping, has the highest percentage amongst the other elements in a scrapped ship. There are also some unwanted materials due to the scrapping activities. For example, there are various poisonous substances which will be released into the environment due to the scrapping process and are highly dangerous for both the environment and humans, including the people who work for the scrapyards. In this chapter, these substances and their risks are identified and their environmental issues are explained.

The growth of the world economy is related to the seaborne trade and the maritime market. The maritime market has different sections, including the demolition market. They are highly intercorrelated and changes in one of them can be a reason for changes in the others. Fundamental issues of these markets are illustrated separately and the possibilities of affecting each other are studied in this chapter.

There are some international regulations which can affect the demolition process as well as the demolition market. For example, these sorts of regulations can force ships to go for scrapping in a certain time. Therefore, it is important for both the ship owners and the scrapyards to take these regulations into account. In this chapter some of the organisations, which can affect the scrapping activities and related regulations are explained.

2.2 Growth of the World Fleet

One of the main business areas for the maritime industry is to provide the transportation services. In the relation between seaborne trade and global economic growth, maritime transportation has always been the dominant support of global trade (Figure 2-1).

Different commodities i.e. clothes or toys and bulk shipments like coal, iron ore, crude oil or grain are transferred all around the world using different types of ships.

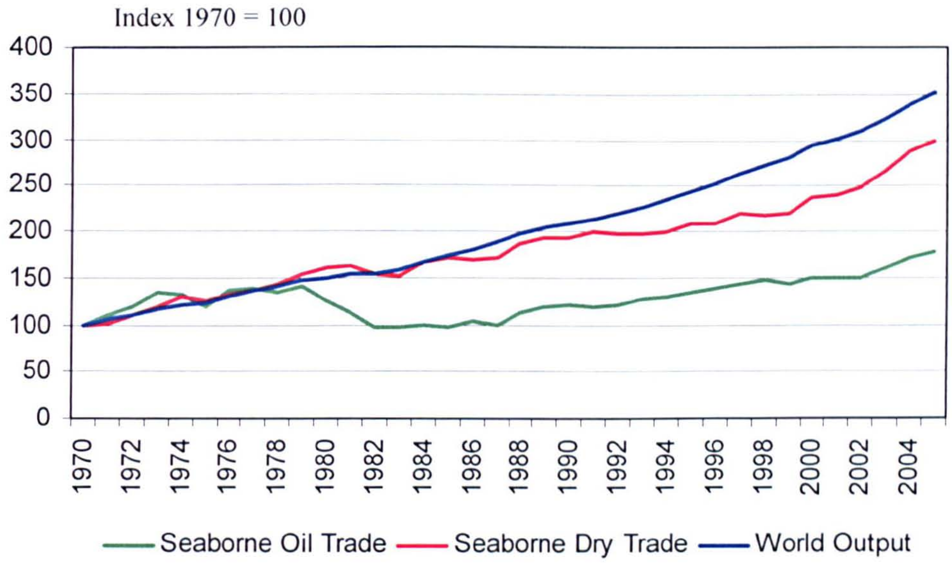


Figure 2-1: World seaborne trade and economic grouth (Platou, 2006).

Global seaborne trade has dramatically changed over the last few decades. According to the UN review of maritime transport (UNCTAD, 2005), 2,504 million tonnes of different goods were loaded for international trade in 1970, but this amount was raised to 5,983 million tonnes, which is more than twice as many as in 1970, in 2000 (Table 2-1). Over the three decades the growth of the seaborne trade averaged 3.1% annually but the growth continued faster over the last few years and 6,758 million tonnes traded in 2004.

Year	Oil		Dry Cargo	Total all goods
	Crude	Products		
1970	1,109	232	1,162	2,504
1980	1,527	344	1,833	3,704
1990	1,287	468	2,253	4,008
2000	1,665	498	3,821	5,983
2001	1,678	499	3,844	6,020
2002	1,637	509	3,981	6,127
2003	1,690	533	4,257	6,480
2004	1,770	546	4,442	6,758

Table 2-1: World Seaborne trade (goods loaded) by types of cargo (million Tonnes) (UNCTAD 2005).

The distribution of the world economy is also in favour of increasing the seaborne trades because of the developing economies of East and South-East Asian, countries like China (Table 2-2). In 2004, 27,635 billion tonne-miles various commodities e.g. grain, coal, iron oar and oil have been traded all around the world which shows growth of 1,791 billion of tonne-miles compared with 2003 and a significant growth more than 10 billion compared with 1999.

Year	Oil		Iron ore	Coal	Grain	Main Dry bulks	Other Dry Cargoes	World total
	Crude	Products						
1970	5,597	890	1,093	481	475	2,049	2,118	10,654
1975	8,882	845	1,471	621	734	2,826	2,810	15,363
1980	8,385	1,020	1,613	952	1,087	3,652	3,720	16,777
1985	4,007	1,150	1,675	1,479	1,004	4,480	3,428	13,065
1990	6,261	1,560	1,978	1,849	1,073	5,259	4,041	17,121
2000	8,180	2,085	2,545	2,509	1,244	6,638	6,790	23,693
2001	8,074	2,105	2,575	2,552	1,322	6,782	6,930	23,891
2002	7,848	2,050	2,731	2,549	1,241	6,789	7,395	24,172
2003	8,390	2,190	3,025	2,810	1,273	7,454	7,810	25,844
2004	8,910	2,325	3,415	2,965	1,325	8,065	8,335	27,635

Table 2-2: World seaborne trade in billion of tonne-miles (Fearnleys Review 2004).

In line with the growth of the seaborne trade, the number and the size of the world fleet has grown to provide the appropriate support for all types of cargo transportation. At some points new ship types were needed to fulfil the requirements of having a suitable transportation pattern. For example, new ship types, like LNG-tankers or Containerships, were introduced in the 1960s and 1970s and changed the previous transportation pattern. Since then, the importance of these kinds of ships has increased gradually and they have technically changed to meet the new criteria. Later in 1990s double-hulled oil tankers were introduced and with regard to the environmental problems of the single-hulled tankers and recent IMO regulations, they are now gaining in importance. In contrast, other ship types may have gradually been seen to decline their importance in the new transport pattern.

One of the most outstanding features of the world merchant fleet during the last 30 years, in which particularly there has been a rapid escalation of ship size, is the bulk sector of the fleet. In the tanker market there was a steady increase in the average size of tankers until the early 1980s when the size structure stabilised. In the bulk carriers there was a similar upward movement in ship size, but the pattern was more evenly spread between the different ship size groups with the fleet of Handy, Panamax and large bulk carriers over 80,000 dwt all expanding. Larger and more efficient ships have progressively pushed their way into the market and depressed rates for smaller sizes. At the same time investment for specialisation, as in the case of car carriers, and chemical tankers played an important part in the development of the world fleet (Stopford, 2003b).

The volume of the world fleet in 1995 was 661.5 million dwt. The share of bulk carriers and tankers were 229.9 million dwt and 270.9 million dwt respectively which is 75.7% of the world fleet. Ten years later, in 2005, the volume of the world fleet has been increased to 855 million dwt but the bulk carriers and tankers shares together is 73.1% of the world fleet, which slightly decreased by 2.6%. The volume of the chemical carriers doubled during this period (Table 2-3).

Year	Tankers	Chemical carriers	Bulk carriers	Combined carriers	Others	Total
1995	270.9	0	229.9	25.9	134.8	661.5
1996	261.0	9.5	241.3	20.7	140.9	673.4
1997	265.1	10.0	250.0	17.3	149.1	691.5
1998	268.5	11.0	260.7	16.9	155.3	712.4
1999	273.2	11.9	260.4	16.1	160.9	722.6
2000	276.0	13.5	264.8	15.2	166.7	736.2
2001	281.3	15.0	274.0	14.6	169.3	754.3
2002	274.9	15.0	287.4	13.8	174.7	765.9
2003	278.8	15.4	295.0	12.6	181.2	783.0
2004	287.9	17.3	303.3	12.2	189.6	810.3
2005	304.1	18.0	320.7	11.7	200.5	855.0

Table 2-3: World Fleet Development (million dwt) (Platou, 2006).

In the period 2001-2005, the average growth rate for the total dwt tonnage supply was 3.3 percent; this represents in absolute terms an increase of 109.3 million dwt of the world fleet (ISL, 2006).

In 2005, the world supply tonnage dramatically increased by 5.7%, which is more than the above average, and shows a quick rise in the world supply of tonnage in the last year (Figure 2-2).

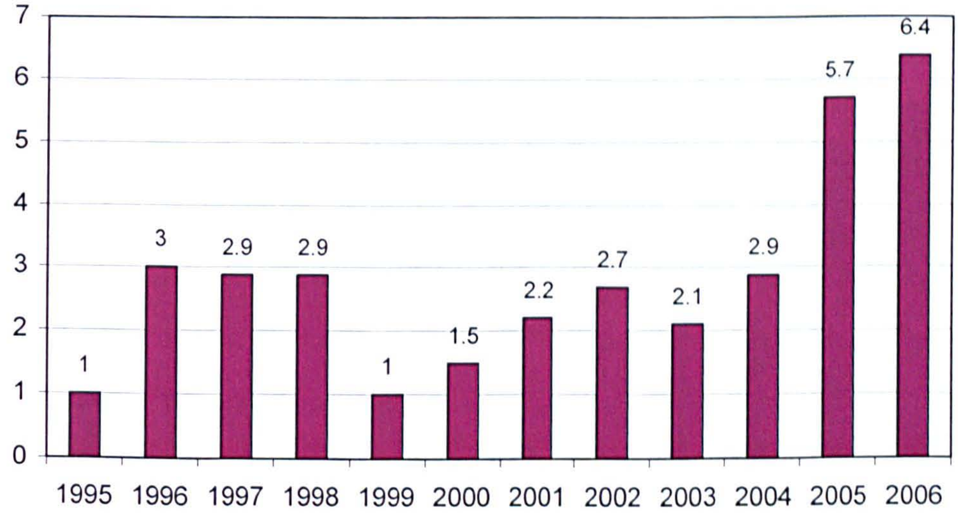


Figure 2-2: Annual tonnage changes in the world merchant fleet (dwt % change) (ISL, 2004&2006).

This is just because of the growing economies all over the world in recent years (Table 2-4). The minimum rate of growth of 1% occured in 1995 and 1999.

	2004	2005	2006
USA	4.2	3.6	3.4
Japan	2.3	2.5	2.3
EU	2.3	1.6	2.1
C. and Europe	6.5	4.3	4.6
Russia	7.2	5.5	5.3
Africa	4.5	5.9	5.5
China	10.1	9.8	8.7
India	6.9	7.6	7.1
Other Asia	6.0	5.7	5.7
M. East	5.5	5.4	5.0
World	5.1	4.5	4.3

Table 2-4: : Percentage change of economic growth in real GDP from previous year (Platou, 2006).

One of the main indicators of operational productivity of the world fleet, tonne-miles per deadweight tonne, is shown in (Figure 2-3). According to the UNCTAD calculations thousands of tonne-miles performed per deadweight tonne, after a significant increase in 2003, increased from 30.2 in 2003 to 3.8 in 2004. This increase resulted from the increased carriage distance of seaborne trade.

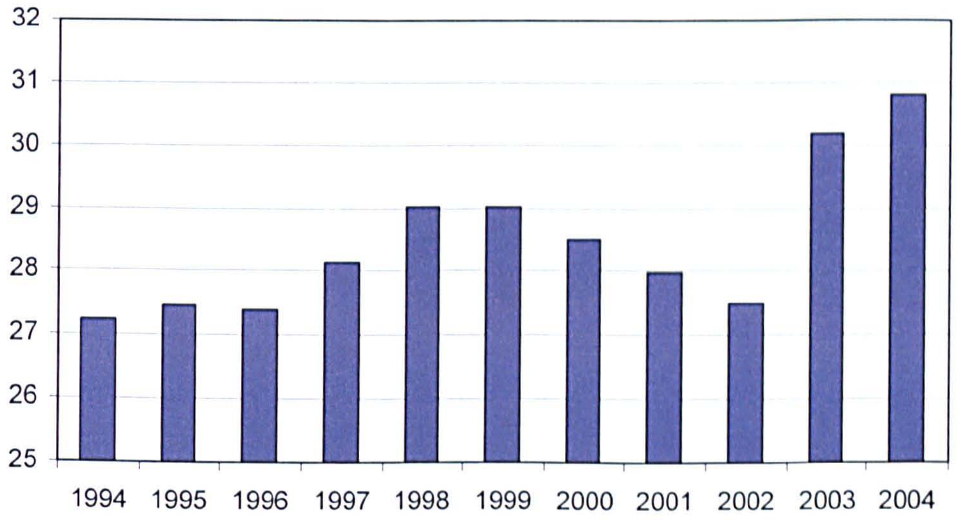


Figure 2-3: Thousand tonne-miles performed per dwt tonne of total world fleet (UNCTAD, 2005).

Tonnes of cargo carried per deadweight tonne is the other indicator of the world fleet productivity. As shown in Table 2-5, after an increase in 2003 to 7.6 it has slightly decreased in 2004 to 7.5. This represent the faster fleet expansion relative to the cargo carried.

Year	World fleet Million dwt	Total cargo Million tonnes	Total tonne-miles performed (Billion tonne-miles)	Tonnes carried per dwt
1990	658.4	4,008.0	17,121	6.1
1995	734.9	4,651.0	20,188	6.3
2000	808.4	5,871.0	23,016	7.3
2003	857.0	6,479.5	25,844	7.6
2004	895.8	6,758.3	27,635	7.5

Table 2-5: Cargo carried per deadweight tonne of the total fleet (UNCTAD, 2005).

Each ship has a particular life cycle, it varies between 20 to 50 years (usually 25 years), and it depends on several factors e.g. the ship type. At the end of its life cycle, the ship should go out of the world fleet and it is usually decided to sell to scrap yards. Occasionally, it can be donated to a non-profit organisation for use as an historical memorial or museum. Therefore, the total number of ships in the world fleet in each year is a reaction between adding the newbuildings and reducing the old ships. As indicated in Figure 2-4, in 2003 fleet additions, in terms of tonnage, exceeded demolitions by approx 21.8 million dwt and this amount has been increased in the next years. During 2005, additions to the merchant fleet reached a volume of 70.1 million dwt, with the number of 1,627 merchant vessels, compared with 62.4 million dwt in 2004, with 1.341 merchant vessels.

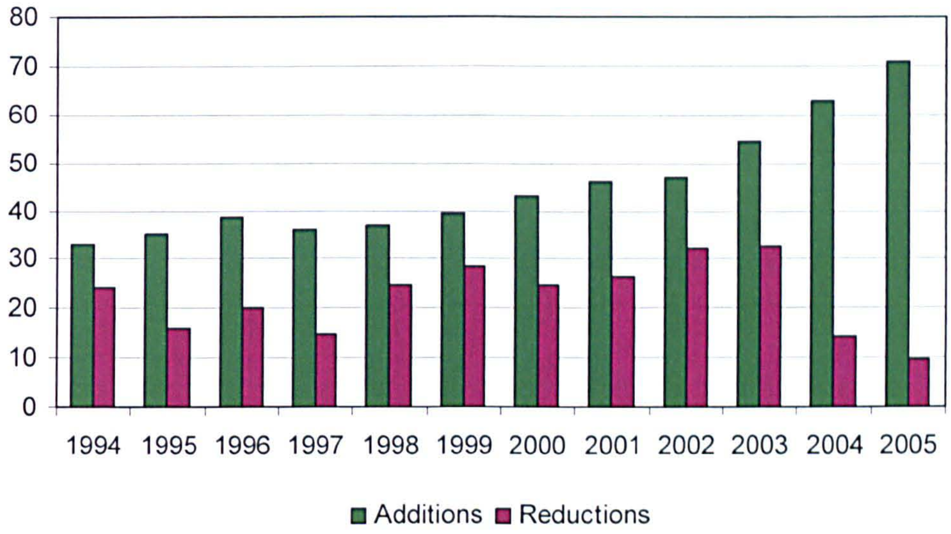


Figure 2-4: : World tonnage addtions and reductions (mill dwt) (ISL, 2006).

At the beginning of 2006, the tanker fleet for ships of 300 gt and over comprised 10,401 tankers totalling 387.7 million dwt (Figure 2-5). This represents, in terms of dwt, 5.2 percent increase over 2005. Also, 7,635 tankers equal to 73.4 per cent of all tankers representing in terms of dwt 15.3 percent of the total world tanker tonnage were attributable to size classes below 40,000 dwt. The crude oil tanker fleet consisted of 471 VLCC tankers (200,000–320,000 dwt) and 10 ULCC tankers (320,000 and above) of which 4 units were single-hull tankers. In the period 2002-2006 the average size of the tanker fleet increased by 3.3 million dwt. Crude oil tankers showed the highest raise amongst the other ship types from 63.4 to 72.3 million dwt i.e., 8.9 million dwt increase.

Ship type	2002		2006		Average growth rate '02-06'		Average size (million dwt)	
	Number	million dwt	Number	million dwt	Number	million dwt	2002	2006
Crude oil	3672	232.9	3556	257.1	-0.8	2.5	63.4	72.3
Product	2302	47.6	2467	58.8	1.7	5.4	20.7	23.9
Oil/Chemical	1337	22.7	1840	37.6	8.3	13.4	17.0	20.4
Chemical	1291	8.5	1354	9.9	1.2	4.0	6.6	7.3
Liquid gas	1114	19.0	1184	24.2	1.5	6.3	17.1	20.5
Total	9716	330.7	10401	387.7	1.7	4.1	34.0	37.3

Figure 2-5: World tanker fleet by type as of January 1st, 2002 and 2006 (ISL, 2006a).

Subsequently, the average ship size of the world merchant fleet is increasing. Within the last five years period the average size increased from 19,970 to 22,970 dwt. At the beginning of 2006, about 28,700 ships equal to 75.8 per cent (2002: 74.1 per cent) of all merchant ships (300 gt and over) belonged to the size segment up to 19,999 dwt. The majority of ships in these size classes are general cargo ships. Moreover, 9,944 of all merchant ships and 48.5 per cent of the total deadweight tonnage aggregated to size classes between 20,000-99,999 dwt (ISL, 2006).

2.3 Ship Recycling

When a ship becomes obsolete for the market it serves or non-compliant for other reasons it reaches the end of its useful life. There are few alternatives at the end of the ship's life:

- Lay up
- Conversion
- Donation
- Sold for Scrapping

Essentially, lay up only postpones the scrapping issue. Donation and also Conversion to other uses presents only a very limited number of alternatives; such as storage facilities, breakwaters or tourist attractions. Hence, demolition at a scrap yard is the most profitable and common option. In principle, the process of ship scrapping consists of a sequential chain of operations undertaken at different locations at a scrap yard (Andersen, 2001):

- Offshore. Prior to beaching tanks are discharged and valuables (uncontaminated oil product and saleables such as electronic equipment) are removed.
- Inter-tidal zone. The vessel is beached under its own power and demolition is initiated (in a certain sequence).
- The beach. Further cutting into manageable sizes, extraction of components and sorting for transport to respective receivers are carried out.
- Shore. Supply of second-hand equipment and components to market and remanufacturing/recycling into new products/components.

From the statistical point of view, the total tonnage removed had a peak in 2001 with 31.7 million dwt. Consequently, first there was a decrease to 25.9 million dwt until 2003 and then it almost halved each year for the next two years (Table 2-6).

It is remarkable that the tankers removal represented 18.9 million dwt in 2003 but the bulk carrier deletion represented only 3.5 million dwt in this year. This shows a significant difference between tankers and bulk carriers (or any other ship types) removal. The same pattern of the elimination repeated in 2005 with 5.1 and 1.0 million dwt removals for tankers and bulk carriers respectively.

Start	Tankers	Bulk Carriers	Combined Carriers	Others	Total
1994	12,4	4,6	3,3	1,1	21,4
1995	10,9	2,6	1,7	0,5	15,7
1996	6,7	8,5	1,9	0,7	17,9
1997	3,6	7,9	2,3	2,5	16,4
1998	7,0	11,8	1,3	3,0	23,1
1999	16,3	9,1	0,9	3,9	30,3
2000	13.9	4,4	0,6	3,1	22,1
2001	19,5	7,2	0,8	4,0	31,7
2002	18,9	6,0	1,2	3,9	30,4
2003	18,9	3,5	0,7	2,8	25,9
2004	10,2	0,8	0,5	1,0	12,7
2005	5,1	1,2	-	1,0	7,6

Table 2-6: Sold for scrapping, lost and other removals (million dwt) (Platou, 2006).

At the end of sailing life, ships are sold for their valuable recyclable steel. The ship type is important in determining the price offered by the ship breaker. Large ships with easily accessible surfaces, such as tankers are easier to cut in pieces and are therefore more valuable and profitable. Generally, Tankers have large flat steel panels which are easy to cut, so the steel utilisation percentage in a tanker is higher than the other ship types. Steel scrap obtained from shipbreaking process has comparatively high quality, especially from tankers because of its large flat panels. All of these factors make tankers more profitable for scrapping. The scrap steel provides most of the value of the ship. The percentage of the steel varies and it depends on ship type and size but there has been an estimate of roughly 74.4% for a standard tanker with 120,000 dwt and 63.15% dwt for a standard bulk carrier with 52,000 dwt (Table 2-7).

Element	Standard Tanker	Standard Bulk Carrier
Steel	74.4	63.15
Copper	0.01	0.04
Zinc	0.03	0.04
Special Bronze	0.03	0.04
Machinery	14	19
Electrical/Electronic Equipment	2.5	5
Joinery - Related products	5	6
Minerals	0.5	2.5
Plastics	0.5	1.2
Liquids	2	1
Chemicals and gases	0.03	0.03
Other miscellaneous	1	2
Total	100	100

Table 2-7: Percentage of various elements for a standard ship (DNV, 2001).

In 1980s there was the worst recession in maritime history (Stopford, 2003b). Hence, the scrapping rate peaked in 1985 with 43.4 million dwt (Figure 2-6). On the supply side, 61.8 million dwt were delivered in 1975 and shipyards were at the peak of the 1970s fleet replacement cycle in last few years. Actually, it was the last phase of 1970s scrapping with an orderbook of 168.4 million dwt at the beginning of the 2004 (Platou, 2004).

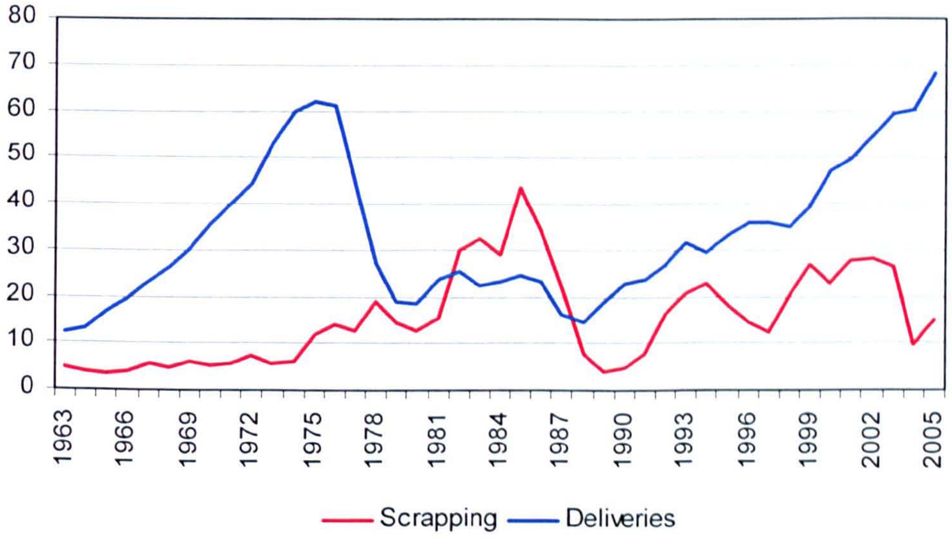


Figure 2-6: Deliveries and Scrapping rates (million dwt) (Clarkson research, 2005).

According to the age profile of the world fleet on January the first 2006, 39.2 million dwt of tankers and 90.5 million dwt of bulk carriers were built in 1985 or earlier (Table 2-8). This means they are already almost 20 years old and potentially ready for scrapping (it will depend on trade and need to have double hull). Moreover, there are 38.0 and 45.2 million dwt of tankers and bulk carriers respectively which were built between 1986 and 1990 and will be ready for scrapping in the near future.

Tankers	-85	86-90	91-95	96-00	01-2005	Total
10-69,999	23,3	8,9	8,7	13,5	24,4	78,9
70-119,999	6,6	10,6	11,5	15,4	33,5	77,5
120-199,999	5,0	4,6	11,5	12,3	18,1	51,5
200,000+	4,2	14,0	36,8	34,2	49,0	138,2
Total	39,2	38,0	68,5	75,4	117,2	346,1
Bulk Carriers	-85	86-90	91-95	96-00	01-2005	Total
10-59,999	53,4	19,2	13,7	25,3	31,2	142,9
60-79,999	20,1	8,2	11,7	20,8	24,8	85,6
80,000+	17,0	17,7	19,3	27,8	31,6	113,4
Total	90,5	45,2	44,8	74,0	78,6	341,9

Table 2-8: Tankers and Bulk Carriers age profile in January the first 2006 (million dwt) (Platou, 2006).

2.4 ENVIRONMENTAL ASPECTS

Shipbreaking is one of the roughest and labour intensive forms of work. Most of the ship scrapping industry uses manual labour to break ships. Although it is possible to increase profitability by using mechanised shipbreaking methods, it requires especial investment which is not easy to manage.

There are a few shipbreaking methods, for example dry dock and afloat (or beaching). A combination of methods can be used in some cases as well. The dry dock has been designed primarily for ship construction and repair and it provides more flexibility and better containment of debris. However, it is an expensive process and needs more capital assets. In this method, workers immediately begin to remove large sections or modules of the ship, transferring them to other project areas for environmental abatement, separation and cutting. The Afloat or beaching method has lower facility cost, but presents the greatest challenge in containing debris and controlling ship stability. With the ship in the water, workers begin by moving through doors and hatches to extract interior parts and strip out compartments. Then they cut and remove the ship's structure above the waterline. As the work progresses the ship gets lighter and it is gradually pulled onto a beach, or earth ramp, for final dismantling of the bottom hull (Association of scientist and engineer, 2000).

In environmental terms the hierarchy of demolition waste is (DNV, 2001):

- ✓ Re-use
- ✓ Recycle
- ✓ Disposal

First, reusable items should be extracted including pumps, motors, engines, repair parts, electronic items, cables and any other equipment. Then it is the time to scrap the residual material. Scrap steel is the most important in recyclables. Steel production from scrap is a sustainable process in that it achieves a far better environmental performance in light of energy efficiency and the preservation of non-renewable resources in comparison with the alternative ore-based production. The energy balance between the two approaches may differ by up to 70% (DNV, 1999). Disposals[1] including asbestos, batteries, plastics, radiation sources, lead and minerals can cause a threat to human health and the environment.

Many of the vessels currently designated for scrapping were built in the 1950s, 1960s, and 1970s using dangerous materials in their construction. Many of these materials are currently classified as hazardous, e.g. asbestos, PCBs, lead, chromates, mercury, and cadmium (Table 2-9). In addition to the above mentioned environmental hazards due to the scrapping activities, there are numerous safety and security issues for the

[1] Disposal means that goods which may once have had a residual value when they reach the end of their working life such as motor vehicles.

people who work in a scrap yard. These workers are threatening their health and lives by working in such a harsh condition (Figure 2-7).

Figure 2-7: Shipbreaking harsh working condition in Chittagong, Bangladesh 2000 (Picture: Edward Burtynsky).

As explained in chapter one, health effects that have been associated with exposure to PCBs include acne-like skin conditions in adults and neurobehavioral and immunological changes in children. Also, TBT, which is one of the most poisonous substances, has been used in most of the world's marine paints to keep fouling organisms from clinging to ships. It is also used in wood treatment and preservation, water and refrigeration systems. TBT paint is due to be phased out of ships by 2008.

Hazardous or Harmful Factors in Ship Scrapping
Asbestos
Polychlorinated Biphenyls (PCBs)
Lead
Chromates
Mercury
Fumes of welding & cuffing
Radiation
Noise
Vibration
Air pollution
Low-level radium sources
Organic liquids (Benzene etc.)
Battery, Compressed gas cylinders, fire fighting liquids, etc.
Chemical materials
Work using plasma and gas torches
Explosive(s)
Work using cranes and lifting equipment
Saws, Grinders and Abrasive cutting wheels
Accident factors: falling, upsetting, electric shock, etc.

Table 2-9: Identifiable hazards associated with ship-breaking and existing ILO standards (Bailey, 2000).

The forecast of annual production for some of the waste materials due to the scrapping over the period of 2001 – 2015 in OECD countries has shown in Table 2-10.

Waste Stream	OECD Europe	Geographical Europe
Steel	860,000	1,480,000
Copper	115	197
Zinc	345	591
Special Bronze	345	591
Machinery	161,000	275,800
Electrical/Electronic Equipment	28,750	49,250
Joinery – related products	57,500	98,500
Minerals	5,750	9,850
Plastics	5,750	9,850
Liquids	23,000	39,400
Chemicals and gases	345	591
Other miscellaneous	11,500	19,700
Total	1,154,400	1,984,320

Table 2-10: The annual production of wastes due to the scrapping over the period of 2001 – 2015 in OECD countries (tonnes) (DNV, 2001).

In the 1970s shipbreaking activities were concentrated in Europe. They were performed at docks and were highly mechanised industrial operation. But the costs of upholding environmental, health and safety standards gradually increased and the scrapping industry moved to poorer Asian states which had few health and safety standards (Greenpeace, 2004). First it had been shifted to regions such as Taiwan and South Korea, but then moved on to new areas within the same region where labour costs traditionally had been even lower (DNV, 1999).

From a technical point of view, the choice of the location for establishment of scrapping sites is based upon some prerequisites and it may summarised as follows (Andersen, 2001):

- A long uniform inter-tidal zone with sufficient tidal difference (allowing vessels of a range of sizes to be dry-beached);
- Minimum exposure (coastal protection) and stable weather conditions;
- Availability of low-cost labour;
- A certain level of infrastructure for disposals.

In 1992 and 1993, half of all ocean going ships were being scrapped in China (Drewry Shipping Consultant, 1998) but a few years later their share dramatically decreased to only 1 per cent in 1997 (Table 2-11) and they were nearly eliminated from the market.

Scrap location	Unit	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003*	Un-known	Total
Bangladesh	mdwt	3.1	3.9	4.6	3.2	5.8	7.2	4.2	9.5	8.7	4.1	0.0	54.2
	Number	25	31	61	63	66	65	61	123	69	39	0	603
India	mdwt	6.5	6.1	8.8	7.7	10.0	10.6	8.1	8.1	11.1	7.6	0.1	84.7
	Number	107	148	262	293	360	340	274	298	326	229	1	2,638
Pakistan	mdwt	3.7	3.1	2.0	0.9	3.4	4.3	1.2	3.7	1.7	1.0	0.0	24.9
	Number	19	20	16	14	40	34	16	26	13	14	1	213
China	mdwt	2.8	0.9	0.3	0.1	2.1	5.4	5.7	5.7	5.9	8.2	0.1	37.1
	Number	34	19	13	6	48	72	77	76	90	79	9	523
Vietnam	mdwt	0.3	0.4	0.2	0.5	0.4	0.3	0.1	0.0	0.0	0.0	0.0	2.3
	Number	3	2	6	3	5	5	4	1	1	0	0	30
Other Asia	mdwt	0.0	0.3	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.6
	Number	0	2	1	1	1	4	0	1	1	1	0	12
EU	mdwt	0.0	0.0	0.0	0.0	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.3
	Number	2	1	1	3	7	3	3	2	4	1	0	27
Turkey	mdwt	0.0	0.1	0.2	0.2	0.3	0.6	0.2	0.3	0.3	0.1	0.0	2.3
	Number	2	5	10	12	15	18	14	16	21	12	0	125
North America	mdwt	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.0	0.0	0.2
	Number	0	0	0	0	0	0	0	1	4	1	0	6
South America	mdwt	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
	Number	1	1	0	0	2	6	8	1	0	1	0	9
Mexico	mdwt	0.0	0.0	0.0	0.0	0.1	0.2	0.0	0.0	0.0	0.0	0.0	0.3
	Number	0	1	0	2	6	8	1	0	1	0	0	19
Other	mdwt	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	Number	0	0	0	0	2	0	0	0	0	1	0	3
Unknown	mdwt	0.8	0.1	1.0	1.1	0.6	1.5	1.7	0.2	0.6	0.7	4.0	12.3
	Number	14	8	22	32	23	41	34	14	24	15	140	367
Total	mdwt	18.2	15.5	17.6	15.3	23.9	30.8	21.8	28.2	28.5	21.9	4.2	226.0
	Number	213	240	401	447	599	600	492	565	556	394	151	4,658

Table 2-11: Ship demolition by location, 1994-2003* (Jan-Sep 03) (European Commission, 2004).

On the contrary, India had a growth in scrapping rate during this period and 7.7 million dwt were being scrapped in 1997 which was more than 50% of the whole world fleet. According to the statistics for the year 2001, India breaks 42% of the vessels that are dismantled every year, Bangladesh 7%, Pakistan 6%, china 4% and the rest of the world 41% (UNEP, 2001). EU, North America and South America did not scrap ships in this year, a trend that continued in subsequent years.

There have been considerable variations market shares of the major shipbreaking nations over the years. Nowadays two thirds or more of the old ships are dismantled on the Indian subcontinent, with Bangladesh currently holding the largest share of the market. The figure below shows that Bangladesh accounts for the largest share in 2006, while only 5 years back in time India was the world's largest shipbreaking nation (Figure 2-8).

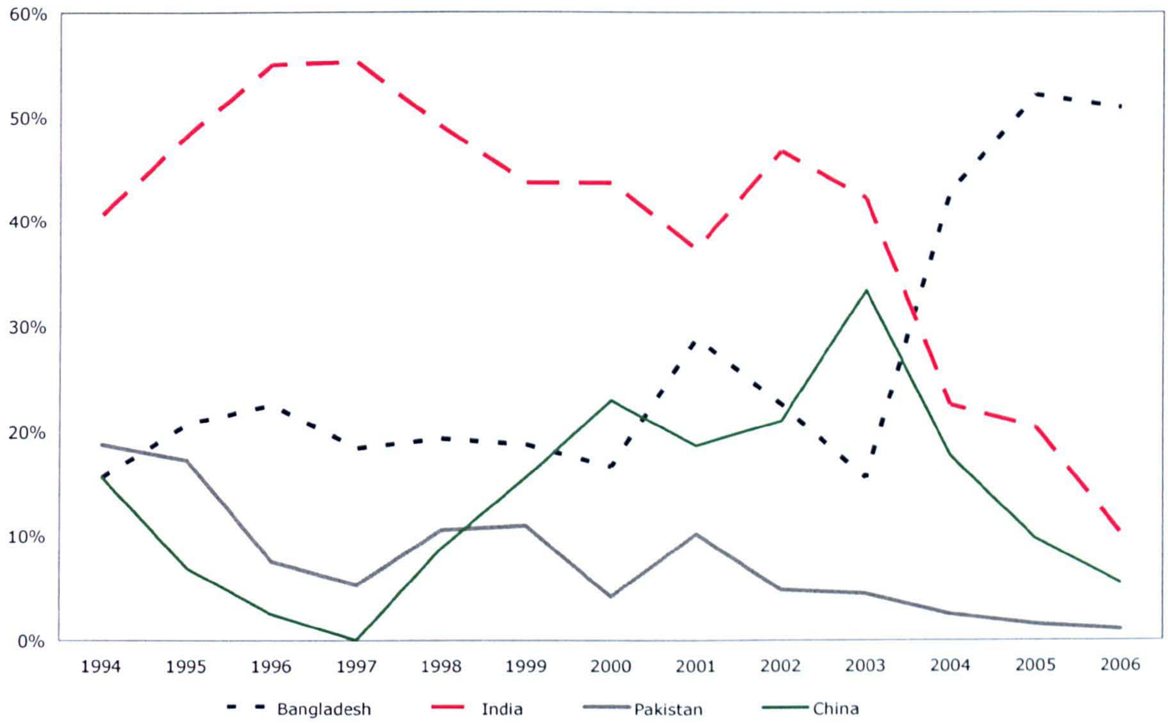


Figure 2-8: Market share of main ship-breaking nations, 1994-2006 (Commission of the European Communities, 2006)

Generally, bulk carriers and tankers have bigger share of the scrapping among the other ship types. For example, according to DNV studies, during the period of 1992 - 2000, on average the number of 363 vessels have been scrapped each year, averaging 19,570 in terms of dead weight. The Tankers' share was 50% and the bulk carriers' share was 31% (Table 2-12). It represents a high percentage of the scrapping market about 81% of the total sum of demolition of vessels. Therefore, bulk carriers and tankers are dominant to the scrapping industry.

Year	Number/ dwt/Age	Tankers	Bulk Carriers	Combos	Gas vessels	Other dry	All Vessels
1992	Number	94	67	11	4	64	240
	dwt	10,22	3,913	1,296	0,011	0,775	16,215
	Age	23,8	23,6	20,8	26,8	24,7	23,9
1993	Number	110	50	15	10	129	314
	dwt	10,685	2,557	2,27	0,111	1,398	17,021
	Age	23,1	24,2	21,9	24,9	29,4	25,9
1994	Number	87	70	18	7	112	294
	dwt	12,558	4,351	2,421	0,018	1,234	20,582
	Age	22,6	24	21,9	26,3	26,5	24,5
1995	Number	93	33	9	1	91	227
	dwt	10,794	2,093	1,229	0,002	1,195	15,313
	Age	25,2	25,2	22,4	30	27,2	25,9
1996	Number	72	128	15	5	168	388
	dwt	6,829	7,297	1,904	0,021	1,967	18,018
	Age	25,3	25	23,1	27,9	27,2	26
1997	Number	40	161	6	6	187	400
	dwt	3,611	7,707	0,746	0,075	2,596	14,735
	Age	28,3	25,5	23,6	28,4	26,5	26,3
1998	Number	52	236	10	6	191	495
	dwt	7,547	11,666	1,416	0,028	3,181	23,838
	Age	25	25	22.8	27.5	25.5	25.2
1999	Number	113	194	9	6	226	548
	dwt	17,114	9,385	1,130	0,019	3,185	30,833
	Age	24.9	24.9	24.3	31.4	25.2	25.1
2000*	Number	55	29	4	1	45	134
	dwt	7,234	1,353	393	18	641	9,639
	Age	26.1	27.1	25	31.7	25.7	26.2
Average 92-99	Number	83	117	12	6	146	363
	dwt	9,920	6,120	1,550	40	1,940	19,570
	Age	24.4	24.8	22.4	26.9	26.3	25.3

Table 2-12: Vessels (> 10,000 dwt) sold for scrapping 1992 -2000* (Jan-Mar), (000 dwt) (DNV, 2001).

2.5 INTERNATIONAL REGULATIONS

In addition to the national statutes and regulations, which are applicable for the ship scrapping industry in each particular country, there are also a number of international agencies which monitor the different aspects of the demolition process and address the various topics of the ship scrapping activities, including:

- IMO
- ILO
- United Nation Commission on Human Rights
- United Nation Environment Programme – The Basel Convention
- Commission of European community

IMO is responsible for coordinating issues associated with ship recycling. It is responsible for monitoring issues arising during ship design, building and operation which might impact on recycling including preparations for recycling onboard. ILO is responsible for establishing standards of operation in shore-based industries involved in ship recycling, concentrating on considering the application of its already existing standards and recommendations to ship recycling and developing guidance for the ship recycling industry and also, to take the lead on working conditions in and around vessels once they have been beached for the scrapping (Andersen, 2001). A central goal of The Basel Convention, under the administration of UNEP is Environmentally Sound Management (ESM), is to protect human health and the environment by minimising hazardous waste production whenever possible. ESM means addressing the issue through an "integrated life-cycle approach" which involves strong controls from the initial generation of a hazardous waste to its storage, transport, treatment, reuse, recycling, recovery and final disposal (Basel, 2004). The Basel Convention has recently been redefined to cover ships being sold "intended to be disposed of" from rich nations to developing nations. The broad effect is to make shipowners in countries that adhere to the convention liable for removing or dealing with toxic materials, or those that pose a risk to the environment, from the ship before it is sold.

The International Convention for the Prevention of Marine Pollution (MARPOL 73/78) is a legislation to control the maritime environment and prevent from any kind of pollution e.g. pollution by oil, sewage, chemicals, harmful substances, noxious liquid substances and air pollution. MARPOL was initially adopted in 1973 and it has completed in 1978. IMO held a Conference on Tanker Safety and Pollution Prevention in February 1978. This conference adopted measures affecting tanker design and operation which were incorporated into the 1973 International Convention for the Prevention of Pollution from Ships. MARPOL 73/78 finally entered into force on October the second 1983 (IMO, 2004). Since then, according to the industrial progresses and also the environmental conditions, The Marine Environment Protection Committee (MEPC) of the International Maritime Organization has adopted additional amendments to the MARPOL 73/78, to keep it efficient, and also up to date.

One of the amendments which have influenced scrapping industry was the 1992 amendment. According to this amendment and its regulation 13G tankers that are 25 years old and which were not constructed according to the requirements of the 1978 Protocol to MARPOL 73/78 have to be fitted with double sides and double bottoms (IMO, 2004). In addition, it revised requirements in the 2002 amendment and brings in a new global timetable for accelerating the phase-out of the single-hull oil tankers. The timetable will see most single-hull oil tankers eliminated by 2015 or earlier. Double-hull tankers offer greater protection of the environment from pollution in certain types of accident. All new oil tankers built since 1996 are required to have double hulls (IMO, 2004). Although the new phase-out timetable sets 2015 as the principal cut-off date for all single-hull tankers, the flag state administration may allow for some newer single hull ships registered in its country that conform to certain technical specifications to continue trading until the 25th anniversary of their delivery. The revised regulation identifies three categories of tankers, as follows (IMO, 2004a):

- "Category 1 oil tanker" means oil tankers of 20,000 tonnes deadweight and above carrying crude oil, fuel oil, heavy diesel oil or lubricating oil as cargo, and of 30,000 tonnes deadweight and above carrying other oils, which do not comply with the requirements for protectively located segregated ballast tanks (commonly known as Pre-MARPOL tankers).

- "Category 2 oil tanker" means oil tankers of 20,000 tonnes deadweight and above carrying crude oil, fuel oil, heavy diesel oil or lubricating oil as cargo, and of 30,000 tonnes deadweight and above carrying other oils, which do comply with the protectively located segregated ballast tank requirements (MARPOL tankers), while
- "Category 3 oil tanker" means an oil tanker of 5,000 tonnes deadweight and above but less than the tonnage specified for Category 1 and 2 tankers.

The oil tanker phase out original timetable by IMO category including their hull type and the year of phase out has been shown in Table 2-13

Phase out year	CAT1			CAT2			CAT3			Total
	DB/DS	SS	CAT1 total	DB/DS	SS	CAT2 total	DB/DS	SS	CAT2 total	
2003	0.6	1.5	2.0	0.0	0.0	0.0	0.1	1.0	1.1	3.1
2004	0.3	4.3	4.6	0.0	0.0	0.0	0.2	0.6	0.8	5.4
2005	0.8	8.4	9.2	0.0	0.0	0.0	0.1	0.4	0.5	9.7
2006	2.3	9.8	12.1	0.0	0.0	0.0	0.2	0.4	0.5	12.6
2007	1.8	5.3	7.1	0.0	0.0	0.0	0.5	0.5	1.0	8.2
2008	0.0	0.0	0.0	1.3	3.9	5.2	0.2	0.3	0.5	5.7
2009	0.0	0.0	0.0	1.2	3.2	4.4	0.3	0.2	0.5	4.9
2010	0.0	0.0	0.0	0.7	2.4	3.1	0.3	0.3	0.6	3.7
2011	0.0	0.0	0.0	1.3	2.2	3.5	0.4	0.4	0.7	4.3
2012	0.0	0.0	0.0	1.6	4.4	5.9	0.2	0.2	0.4	6.3
2013	0.0	0.0	0.0	1.5	3.1	4.6	0.1	0.2	0.3	4.8
2014	0.0	0.0	0.0	1.8	4.1	5.9	0.1	0.2	0.3	6.2
2015	0.0	0.0	0.0	2.8	45.2	48.1	0.2	1.1	1.2	49.3
2016	0.0	0.0	0.0	1.6	0.0	1.6	0.2	0.0	0.2	1.7
2017	0.0	0.0	0.0	1.8	0.0	1.8	0.0	0.0	0.0	1.9
2018	0.0	0.0	0.0	1.4	0.0	1.4	0.1	0.0	0.1	1.4
2019	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.1
2020	0.0	0.0	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.2
2021	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Total	5.7	29.3	35.0	17.2	68.6	85.7	3.2	5.6	8.8	129.5

Table 2-13: Oil tanker phase out by IMO category, (million dwt) (European Commission, 2004).

As it has been noted earlier, tankers and bulkers are dominant to the scrapping industry. Hence, the phase out of the tankers can change the standards and criteria for both the scrapping industry and the demolition market. Regulation (EC) No 1726/2003 of the European Parliament and of the Council of European Union of 22 July 2003 amending Regulation (EC) No 417/2002 on the accelerated phasing-in of double-hull or equivalent design requirements of the MARPOL 73/78 Convention to single-hull oil tankers. This regulation aims to reduce the risks of accidental oil pollution in European waters by establishing a scheme for accelerating the phasing in of double-hull or equivalent design requirements for single-hull oil tankers (European Union, 2003). In the second paragraph of this amendment it is written that:

> "No oil tanker shall be allowed to operate under the flag of a Member State, nor shall any oil tanker, irrespective of its flag, be allowed to enter into ports or offshore terminals under the jurisdiction of a Member State after the anniversary of the date of delivery of the ship in the year specified hereafter, unless such tanker is a double hull oil tanker:
> (a) For category 1 oil tankers:
> — 2003 for ships delivered in 1980 or earlier,
> — 2004 for ships delivered in 1981,
> — 2005 for ships delivered in 1982 or later;
> (b) For category 2 and 3 oil tankers:
> — 2003 for ships delivered in 1975 or earlier,
> — 2004 for ships delivered in 1976,
> — 2005 for ships delivered in 1977,
> — 2006 for ships delivered in 1978 and 1979,
> — 2007 for ships delivered in 1980 and 1981,
> — 2008 for ships delivered in 1982,
> — 2009 for ships delivered in 1983,
> — 2010 for ships delivered in 1984 or later;"

Comparisons of phase out schemes for single hull oil tankers are represented in Figure 2-9. As it appears in this figure, according to the EC 1726/2003 amendment there will be a massive scrap in 2010 with 66.7 million dwt. The next peak for the scrapping, according to the IMO regulations, will be in 2015 with 49.3 million dwt.

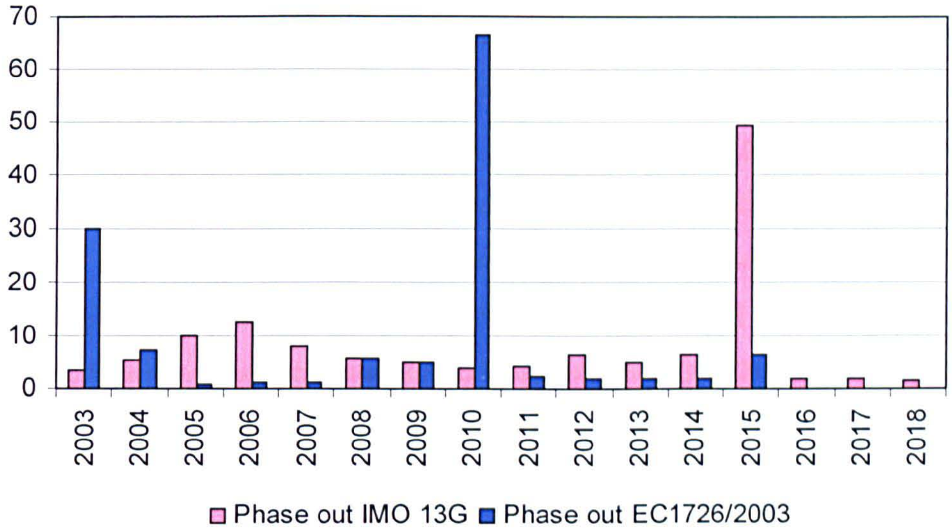


Figure 2-9: Comparison of phase out schemes for single hull oil tankers. (million dwt) (European Commission, 2004).

IMO were compiling some guidelines for ship recycling to define procedures for new ships and also existing ships related to ship recycling. According to these guidelines, shipowners should prepare the Green Passport for existing ships which is a document facilitating the application of these guidelines providing information with regard to materials known to be potentially hazardous utilised in the construction of the ship, its equipment and systems (IMO, 2004).

2.6 Demolition Market

Ship demolition provides a large amount of recyclable materials. Some 95% of an average merchant ship will be re-used, from the steel to the non-ferrous metals and pipework of the ship which will be re-used. The scrap price of ships is volatile and depends upon the demand for steel from this source. A ship is sold on its lightweight displacement (ldt) on a price per tonne basis. Lightweight is a measure of the weight of the ship when it does not contain oil, water, fuel, cargo, crew and so on. Scrap steel price is heavily influenced by the demand for constructional steel from the building industry. A surge in building, such as has been apparent for the past several years in China has pushed the price of ship scrap up to very high levels (BIMCO, 2004).

As mentioned in section 2-1, all the sections of shipping markets are highly interconnected and changing a parameter in each of these four markets not only affect other factors in the same market, but also influence one or more variables in the other markets. This makes the whole shipping market more complex. The demolition market deals with ships which are generally obsolete, their useful life is over and they can not be employed for the other purposes. In this market, the price is affected by the availability of ships for demolition which itself is governed by the freight market conditions. If freights are high, few ships will be available for scrap, and prices will be at their highest point. If there are many vessels being offered at a time of poor freight rates, then the scrap prices will also be low (BIMCO, 2004). It is all a matter of supply and demand, as the price of ships for demolition decreases shipbreakers will demand more vessels to scrap and vice versa.

The interaction of the supply and demand curves determines the market price and quantity of scrapped vessels (Figure 2-10). It means that the equilibrium or balance in a competitive market is the point where the supply and demand curves cross. This is the price (p*) and the quantity (q*) of vessels that will be sold for demolition (European Commission, 2004). A change of price causes a move along the demand curve. The same goes for the supply side.

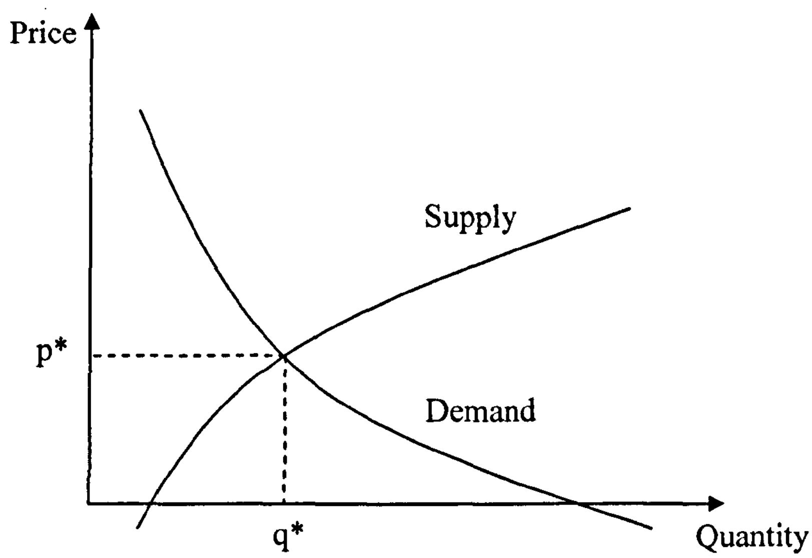


Figure 2-10: Supply and Demand for ship demolition. p is the price and q is the quantity of vessels that will be sold for demolition (European Commission, 2004).

With the strong tanker and dry bulk market in 2003, the numbers of vessels sold for demolition were significantly down from the massive 2002 level. At total of 25.9 million dwt were sold for scrapping compared with 30.4 million dwt in 2002 and 31.7 million dwt in 2001 (Platou., 2006). In this year, the strong freight rate and also increased steel prices pushed scrap prices up to levels never seen before and this consequently continued to 2004 (Figure 2-11). When the market peaked at about \$460 per ldt in February 2005, Bangladeshi buyers were by far the most active buyers, and with the exception of a few sales, paid the highest prices for standard tonnage. After a short-lived cartel formed by Bangladeshi buyers in June, a new cartel was formed in September, successfully forcing prices down 10 percent from about \$400/ldt to \$360/ldt. The cartel showed signs of weakness towards the end of the year, with some breakers purchasing outside the cartel (Platou, 2006).

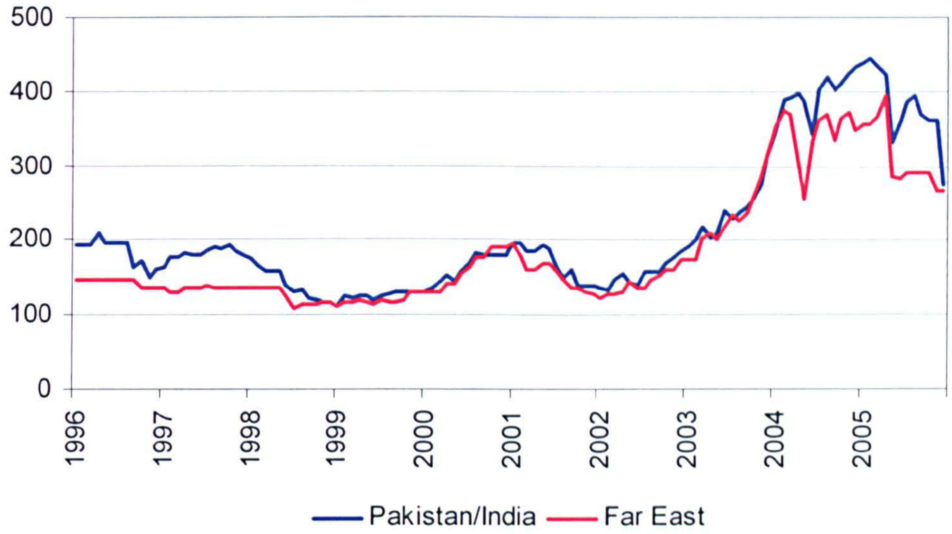


Figure 2-11: Demolition Prices (ldt/$) (Platou, 2006).

Ship demolition prices have been influenced by some internal and external factors. Internal factors are the variables and parameters inside the demolition market itself and the external factors are the parameters in the other markets which affecting the relative parameters in the demolition market, e.g. Freight rates, steel price and operating costs of ships. In the following sections some of the drivers and the influential parameters of the demolition market are discussed.

2.6.1 FREIGHT RATES

In the past few years, the tanker fleet increased much more strongly than in many years, with deliveries of 28 million dwt, while scrapping and other removals amounted to no more than 5 million dwt in year 2005. The fleet growth was as high as 7 percent as an annual average, resulting in a drop in the utilisation rate from 91.5 percent in 2004 to 88.5 percent in 2005 (Platou, 2006). Consequently, freight rate in recent years showed a significant growth comparing to the past years. The year 2003 was the fourth freight market spike of tankers, as the dominant feature of the scrapping market, subsequent to 1991, 1997 and 2000 (Figure 2-12). Freight rate of VLCCs achieved an average of $50,000 per day in 2003, lower than $53,000 in 2000. These two years represented the best years in the tanker market since 1973. Despite the high annual average in 2003, there were large fluctuations over the year with $20,000 per day in July and August and between $80,000 and $90,000 per day in November and December. In 2004 the average rate was $87,000, which was the best year since the 1970s, and this reached to its highest peak with $181,000 per day in November. Since then, freight rates started to decrease but actually 2005 was also a profitable year for tankers owner and recorded as the second best year since the 1970s.

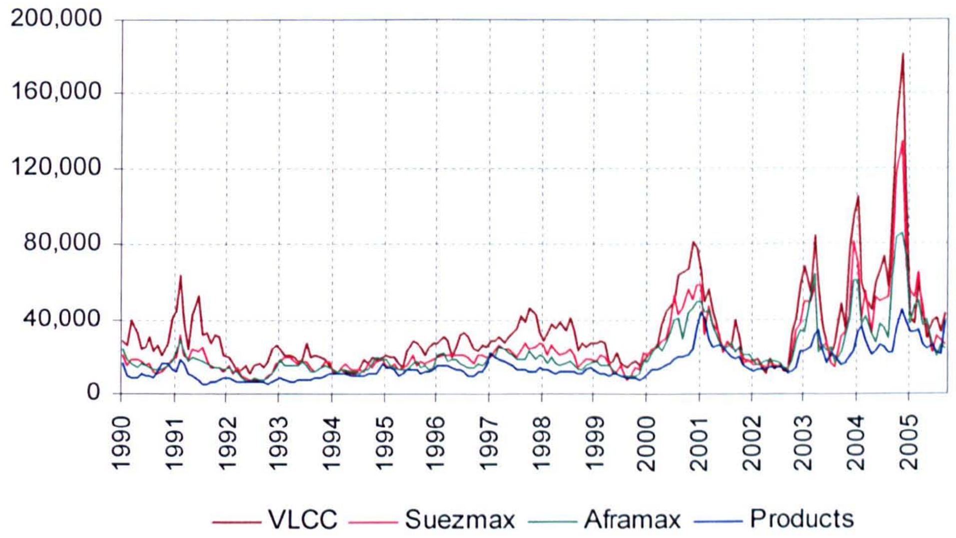


Figure 2-12: Tankers Freight rates (1990- Sept. 2005) ($ per day earning) (Clarkson Research, 2005).

Like tankers, dry bulk market has been through a series of cycles over the last years (Figure 2-13). In the spike of 2003, the freight rate of Capesizes fluctuated between $18,000 and $80,000, averaging $35,600 per day in 2003, up from $2,800 in 2002. Also the average for Panamax rate was $20,300 per day, compared with $8,000 in the previous year (Platou, 2004). As can be seen in the below figure, average freight rates for the Capesize in 2005, with $47,200 per day, dropped significantly from the all time high levels of $62,000 per day in 2004. This trend also happened for the Panamax and the Handymax in the same period. In 2006 the freight rate of Capesizes marginally decreased to $43,400 per day but in 2007 it sharply rose to $99,700 per day.

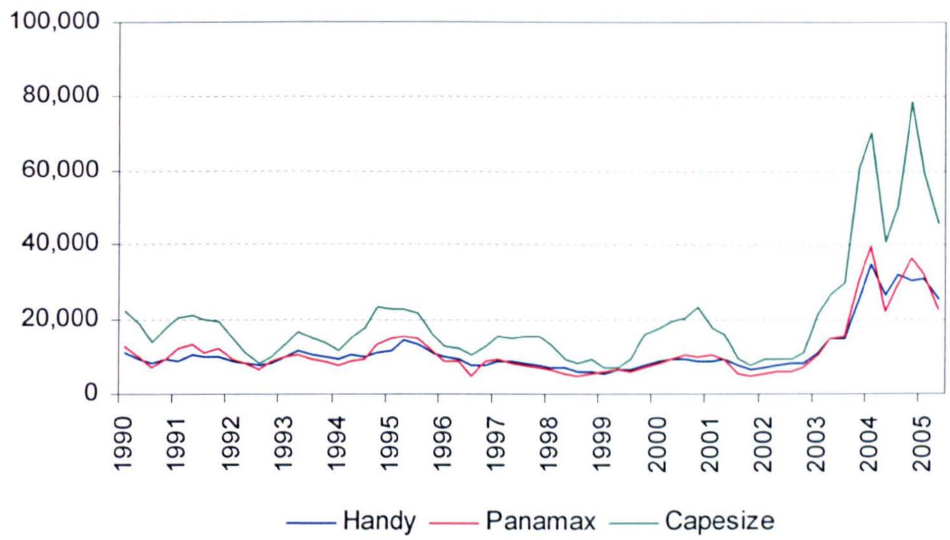


Figure 2-13: Bulk carriesr Freight rates (1990- Sept. 2005) ($ per day earning) (Clarkson Research, 2005).

With the strong tanker and dry bulk market in 2003, the numbers of vessels sold for demolition were significantly down from the massive 2001 level. A total of 25.9 million dwt were sold for scrapping compared with 30.4 million dwt in 2002 and 31.7 million dwt in 2001 (Platou, 2004).

2.6.2 Steel Prices

A significant part of the cost for a shipbuilder is closely linked to the steel market which has been booming in the past two years following the global economic upturn which created higher demand for steel (Platou, 2006). Steel scrap obtained from the shipbreaking process has comparatively high quality, especially from tankers because of its large flat panels. The scrap steel provides most of the value of the ship. The crude steel industry and the ship demolition industry have direct interaction because the scrap steel obtained from shipbreaking process has comparatively high quality. Moreover, as mentioned earlier, steel production from scrap is a sustainable process in that it achieves a far better environmental performance and the preservation of non-renewable resources in comparison with the alternative ore-based production (DNV, 1999).

Since 2003, world demand for steel has been intense resulting in record growth in steel output and strong increases in steel prices (Figure 2-14). China has been the main driver for the upturn in steel demand, because its strong economic growth has been very steel intensive (Platou, 2006). In October 2004, the steel price reached to its highest point with $591 per tonne and after a few same months started to decrease since then.

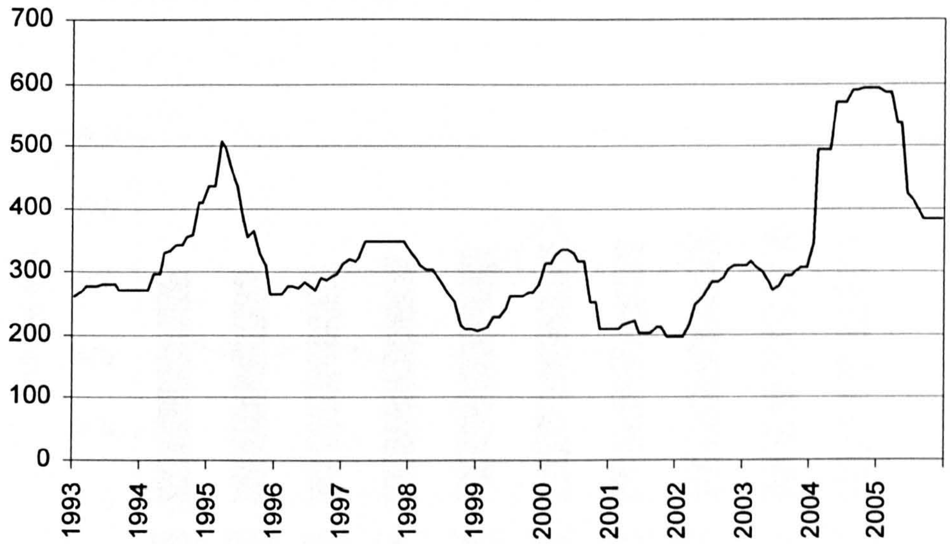


Figure 2-14: Steel Price ($/tonne) (Platou, 2006).

2.6.3 NEWBUILDINGS

The constant increase over the years in the size and the number of the world fleet has led to a general increase in the supply of ships to the ship scrapping industry. The trend in volumes of ship scrapping has followed the increasing trend in the size of the fleet. However, it is also evident that there have been large variations over the years. These variations are determined by the developments in the key drivers of supply and demand (European Commission, 2004).

The capacity of the ship scrap yards and the size of the world fleet have been shown in Figure 2-15. Since 1994, the size of the world fleet has always increased year by year but the capacity of the scrap yards had a peak in 1999 and then decreased to 2000. This also remained constant between 2001 and 2002. In January 2003 the fleet size was 776 million dwt and the capacities of the scarp yards were 22 million dwt which is 2.8% of the fleet size. The capacity and also the location of the shipbreaking yards are parameters which require careful thought. Lack of scrapyards can cause ship owners to lay up their obsolete ships and subsequently cause environmental problems. An excess of scrapyards would cause a crisis because they would not have enough supply to carry on their business, especially if the scrapyard is highly mechanised and costs a lot of money to be established.

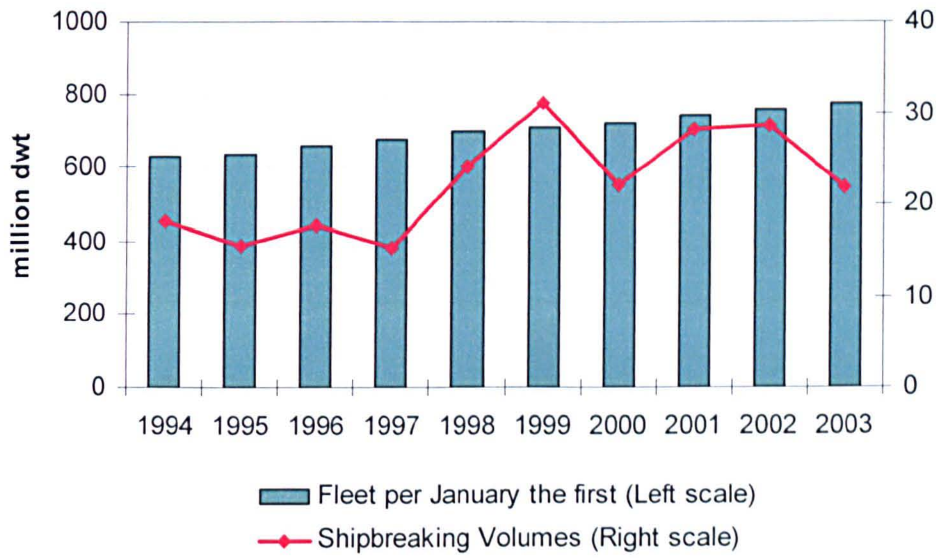


Figure 2-15: Ship breaking volumes and the size of the world fleet (European Commission, 2004).

The obsolete ships of the world fleet are sent to the scrapyards. Obsolescence not only deals with the aged ships, but also a technological leap can be a cause for a ship to become obsolete in the market it serves. For example new containers with a very large capacity compared with the old ones, or a new technology for faster cargo handling could be a reason to have more obsolete ships in the market.

The prices of the new building ships are other parameters which may affect the demolition prices indirectly. If the newbuilding prices are too high, the ship owners may postpone the sale of their ships and this would affect the second hand market and subsequently the demolition market.

2.6.4 SHIPS' AGE PROFILE

The obsolete vessels are the suppliers of the scrapping industry. Obsolete ships are two types: technologically obsolete and age obsolete. The age profile of the world fleet has an important role for the demolition market because it represents the number of the aged vessels and consequently an overview of the future scrapping industry. The age structure of the world fleet by ship types is represented in Figure 2-16. As it appears in this figure, 13% of tankers were built in 1985 or before, 30% between 1986 and 1995 and 56% between 1996 and 2005. So, 13% of the tankers have already 20 years or more age and have potential to be sent for the demolition.

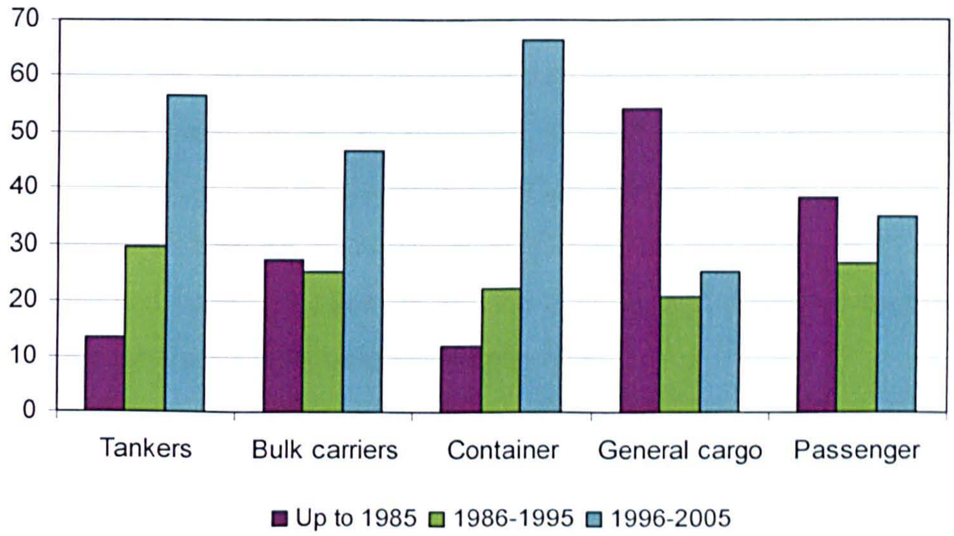


Figure 2-16: Age structure of the world fleet by ship types (dwt percent share) (ISL, 2006).

As of January the 1st, 2006 the following age profiles for major ship types can be highlighted (ISL 2006):

- ✓ 7,037 general cargo ships representing 42.5 per cent of all general cargo ships were older than 25 years (built before 1981).
- ✓ 4,017 oil tankers (including products tankers) equal to a share of 51 per cent of the total number of oil tankers were older than 15 years.
- ✓ 28.7 per cent of all container ships were built during the last five years.

2.6.5 Other Factors

There are also a lot of additional parameters which are capable of making changes to the demolition market structure, e.g. oil production, oil prices, second-hand ship prices and operating costs of a ship. The impact of these factors could be either direct or indirect. Variables like crude oil production and their prices can influence the tankers' freight rates and changing the freight rate market and subsequently the demolition prices. This is an indirect influence. Moreover, the price of the oil can affect the bunker price, as the main element for the operating cost of a ship, and change the demolition prices. As an example, if the operating cost of a ship is relatively high and the freight rates are not able to cover that, the owner may stop using the ship and then sell the ship to a shipbreaker.

Second-hand prices are the other variables which may change the scrapping prices. In a strong scrapping market and poor freight market, high scrap prices compared with the second hand prices could be a reason for selling ships to shipbreakers. Subsequently, the high number of ships which are ready for scrap could bring the scrap prices down after a while.

Political issues are the other influential factors in the ship demolition market. These issues are normally unpredictable and may cause rapid changes to the prices. For example, a sudden war can easily change the market and cause a rapid decrease or increase in steel or oil prices and afterwards all the maritime market structure, including the demolition market, will be changed. Inflation, recession and the economic troughs are the other concerns which can affect the maritime market as well.

2.7 CONCLUSION

The ship demolition process is highly dangerous and causes severe contamination to the environment, including the land, air and water (See section 2-3). It is also threatening to the health and lives of the people who work in this industry (See section 2-4). Therefore, international agencies and organisations are trying to write more regulations and legislations to monitor the process of ship scrapping and restrict the industry to protect the environment and people (See section 2-5). In this way, more accurate information about the future of scrapping, including the volume of scrapped materials and substances and location of the scrapping activities helps them to have a realistic plan. Moreover, these regulations cause some difficulties for the scrapyards and the ship owners which can change the foundation of the industry. Therefore, the scrapping industry and the international organisations should move forward carefully with a reasonable plan.

The scrapping industry and demolition market are highly related to each other and as explained in this chapter (See section 2-2), the demolition market plays an important role for the maritime market and consequently for the world economy. So, changing the scrapping process can change the demolition market and world economy as a result. This shows the importance of having a proper plan to improve the scrapping industry and demolition market and it explains that the more accurate the information is, the more realistic the plan will be.

CHAPTER 3

STATISTICAL APPROACHES

3.1 INTRODUCTION

The fundamentals of time series analysis in a statistical point of view are explained in this chapter. Both the linear and non-linear regression modelling methods are illustrated and their equations and parameters explained. Afterwards, univariate and, more complicated, multivariate forecasting methods are explained. Multivariate analysis deals with issues related to the observations of many variables on units of a selected random sample. The different methods of multivariate analysis are introduced and demonstrated separately. These multivariate statistical analysis methods are concerned with analysing and understanding data in high dimensions and can be suitable to model the demolition market.

3.2 TIME SERIES ANALYSIS

A time series is a collection of observations made sequentially in time which is viewed as a relation of an underlying random (stochastic[1]) process. In a time series, observations are variables and time intervals between the observations are constants. In Mathematics, a "variable" often represents an unknown quantity which can be changed and, in contrast, a "constant" is known and remains steady during the analysis. Variation in a time series can be decomposed into components or "signals" (like trend, seasonal variation or cycle changes) and remaining irregular fluctuations called "noise". Different combinations of these components and noise create various patterns for time series. Behaviour of a component or combination of components is often easy to study but the problem is the remaining random fluctuation which is generally unpredictable.

In general, a trend can be referring to a long term movement in the mean level of a time series. This movement may be upward or downward. For example, the oil production of the non-OPEC countries (countries which are not members of OPEC -

[1] A stochastic process can be described as a statistical phenomenon that evolves in time according to probabilistic laws.

the Organization of the Petroleum Exporting Countries- e.g. US, UK, Russia, Mexico, China and Canada) between February 1995 and December 2004 (Figure 3-1) showed an upward trend. The black line represents the linear trend of the time series.

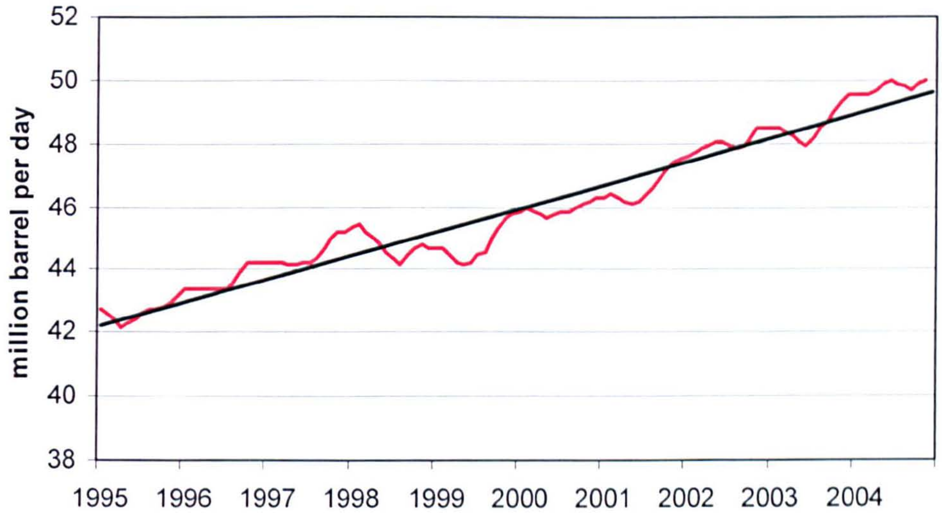


Figure 3-1: Oil Production of the non-OPEC countries shows an upward trend (Platou R.S. 2006).

Some time series, such as retail sales figures or temperature readings, demonstrate variations which have any regular fluctuations with a period of less than one year. These variations, called "seasonality", are easy to study and can be measured explicitly. Also, if the period of these kinds of variations is annual they can be referred to as "seasonal effects". Apart from seasonality and seasonal effects some time series may exhibit other kinds of variations at a fixed period of time. These cyclical components describe any regular fluctuations which vary in a recognisable cycle. Furthermore, it is possible for these cycles to occur at different levels in a time series. In addition to the above mentioned regular mechanism in a time series, there are some irregular components which can affect a series as well. Some of these irregularities can be studied and recognised. For example, a special situation, such as inflation for a financial time series, may affect a time series over a short period and change its regular pattern. The general level will shift because of special circumstances over a certain period of time.

Since all the regular components have been removed from a time series, a set of data including all the residuals will remain. In fact, the main concern in every time series is to deal with the irregular fluctuations or residuals. Various methods and techniques are represented to analyse time series to see if these irregularities may be explained in terms of probability methods. Regression analysis is used to model relationships between variables and determine the magnitude of those relationships. These models can be used to make predictions.

3.3 Regression Analysis

Regression analysis is a statistical tool for the investigation or modelling of relationships between variables. It is important to determine the causal effect of one variable on another e.g. the effect of increased demand on freight rate or changes in second-hand prices on the scrapping rate in the shipping market.

To explore such issues, the data should assemble the underlying variables of interest and regression employed to estimate the quantitative effect of the causal variables upon the variable that they influence. Typically, the "statistical significance" of the estimated relationships should assess, that is the degree of confidence that the real relationship is close to the estimated relationship.

Regression analysis with a single explanatory variable is termed “simple regression”. In reality, any effort to quantify the effects of a single explanatory variable upon the other variables without careful attention to the other factors could create serious statistical difficulties. At the beginning of any regression study, there are several hypotheses about the relationship between the variables. To investigate these hypotheses, the information can be plotted for all of the individuals in the sample using a two-dimensional diagram, conventionally termed a “scatter diagram”. Figure 3-2 is a sample of scatter diagram and shows the relationship between two variables, scrap prices in Far-East and Subcontinent, by plotting one against the other. Each point in the diagram represents an individual in the sample. The value of scrap price in Far-East for each observation is plotted on the y-axis against the value for scrap price in Subcontinent on the x-axis.

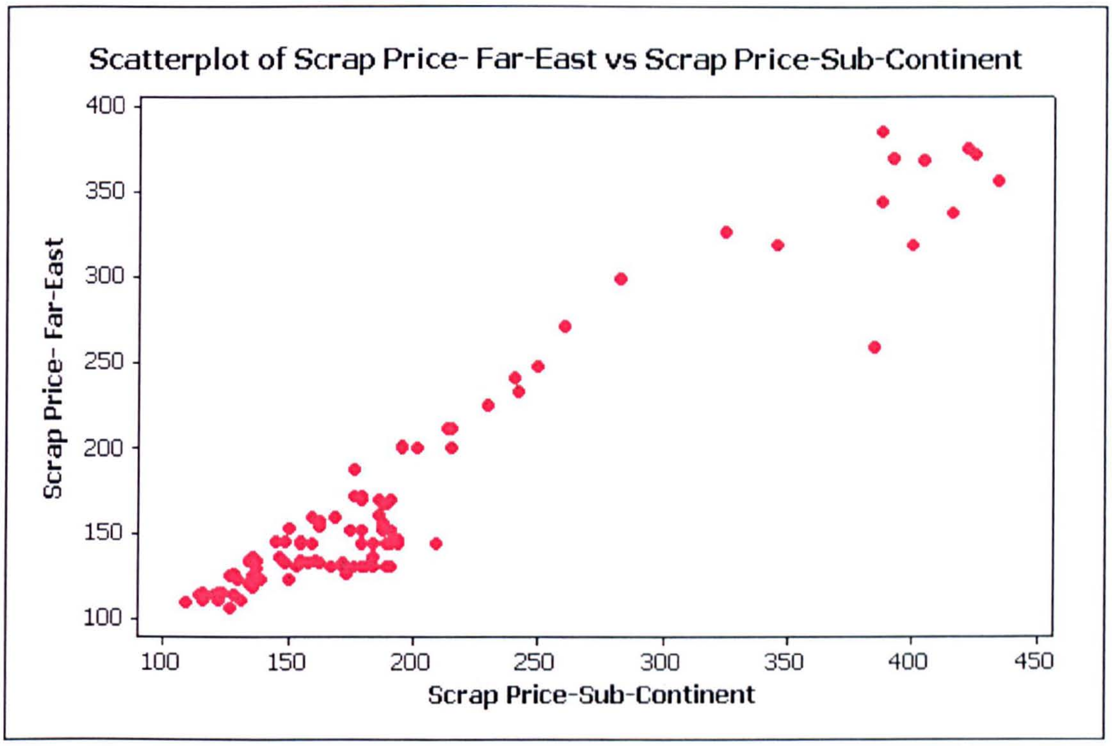


Figure 3-2: Scatter diagram for the regression between ship scrap prices in two different locations

According to this diagram, there might be correlation between the two variables, resulting in the clustering of data points along a line, so it is possible to determine that there is a relationship between the two. However, this is not necessarily true because it is also possible that both could be related to some third variable that explains their behaviour.

"Multiple regression" or "multivariate regression" is a technique that allows additional factors to enter the analysis separately, so that the effect of each can be estimated. Furthermore, with other similar techniques can quantify the impact of various simultaneous influences upon a single dependent variable[1]. The bias due to omitted variables with simple and multiple regressions is often essential even when researcher is only interested in the effects of one of the independent variables.

Statistical models can be used to describe populations of interest. The models are defined in terms of parameters and they provide relations between variables of interest. A very simple model for variables y is:

---

[1] When there is more than one single dependent variable, then it is called "multivariate multiple regression".

$$y = \mu + \varepsilon \qquad (3\text{-}1)$$

Where μ can be regarded as population mean or average and ε is random error. Generally, y will never be equal to μ; this means that the probability is zero that a sample will ever arise in which y is exactly equal to μ.

Once a sample of data is collected the mean μ of the population, which is typically denoted by $\hat{\mu}$, can be estimated. The estimation of μ would be different in each different sample. However, using results from statistical inference, how much the estimator of μ will vary from sample to sample can be quantified.

The sample mean of a random sample of n observations $y_1, y_2, \ldots, y_n$ is given by the ordinary arithmetic average:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad (3\text{-}2)$$

The variance of the population σ^2 is defined as the average squared deviation from the mean and is thus an indication of the extent to which the values of y are spread or scattered.

$$\sigma^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n} \qquad (3\text{-}3)$$

Similarly, the variance of the error term ε is a measure of spread of the distribution. It means if σ^2 is equal to zero, then there will be no variability and the samples will be exactly the same.

The sample variance is defined as:

$$s^2 = \frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n} \qquad (3\text{-}4)$$

The sample variance s^2 is generally never equal to the population variance σ^2 (the probability of such an occurrence is zero) but it is an unbiased estimator for σ^2; that is $E(s^2) = \sigma^2$ (Rencher, 2002). The notation $E(s^2)$ indicates the mean of all possible sample variances. The square root of either the population variance or sample variance is called the "standard deviation".

When the model in Equation 3-1 has been specified, it concerns the probability distribution of the random error ε. Often in practice it is assumed that ε has a normal distribution when the variable y is continuous. Although this normality assumption is often approximately valid, many times it is not. The data should always be examined (usually graphically) to verify that the distribution of the error (and hence y) is approximately normal.

3.4 Linear Regression Model

Many statistical applications deal with a kind of modelling: how a variable y, called a response or dependent variable, depends on another variable x which is called the independent or predictor variable (also called the regressor variable). The simplest way to model a relation between two variables is via a linear function:

$$y = \beta_0 + \beta_1 x \qquad \textbf{(3-5)}$$

Where β_0 and β_1 are the y-intercept and the slope of the line (rate of change in y for a unit change in x) respectively, so by tuning both of them, y can be found as a linear function.

The problem with this model is that it is completely deterministic. For the data collected on any two variables from an experiment, even if there is a linear relationship between the variables, the data points will not fall exactly on a line. Thus, a probabilistic model is needed to account for the variability of points about the line. This can be achieved by adding a random error into the above linear relationship (Equation 3-6).

There are many reasons for the vast popularity of regression models in statistical practice. One of the leading reasons is that regression models allow the relation of variables together in a mathematical form which can provide insight into the relationships between variables of interest. Related to this reason, regression models allow the determination statistically whether a response variable is related to one or more other explanatory variables.

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \qquad \textbf{(3-6)}$$
$$\textit{for: } i=1, 2, \ldots, n.$$

The index i represent the observation number. The error ε_i is a random variable with zero mean and finite variance σ^2 and x is the single predictor variable. It is assumed that the random errors are independent of each other and that they all have the same variance. The model given in Equation 3-6 is called the "simple linear regression" model. The ε_i is unobservable quantity introduced to the model to account for the failure of the observed values to fall upon a single straight line. The x_i and y_i are observed and these data are used to obtain estimates of the unknown parameters β_0 and β_1.

In the simple linear regression equation, the first statistical problem, when analysing the data, is to estimate the parameters β_0 and β_1 referred to as a "linear adjustment". The estimation problem of these parameters can be motivated by the simple model in the Equation 3-1 with:

$$y_i = \mu + \varepsilon_i \qquad \textbf{(3-7)}$$
$$\textit{for: } i=1, 2, \ldots, n.$$

One criterion for estimating μ from the data is to determine the value of μ that minimises the sum of squares:

$$SS = \sum_{i=1}^{n} (y_i - \mu)^2 \qquad \textbf{(3-8)}$$

The value of μ that minimises this sum of squares is $\hat{\mu} = \bar{y}$, the sample mean which is typically used to estimate the mean. In the simple linear regression framework, the same criterion is also used, that is finding the values of β_0 and β_1 that minimise the sum of squares:

$$SS = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_i))^2 \qquad \textbf{(3-9)}$$

This method is known as "least squares", since finding the estimates of the parameters that make the sum of squares take the least possible value.
Most practical applications of regression analysis utilises models with more than one predictor variable. Probabilistic models that include more than one predictor variable are called "multiple regression" model. General form of the multiple regression model is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_k x_k + \varepsilon \qquad \textbf{(3-10)}$$

Where, $x_1, x_2, \ldots, x_k$ are predictor variables, and $\beta_1, \beta_2, \ldots, \beta_k$ are parameters. Multiple linear regression attempts to model the relationship between two or more explanatory variables and a response variable by fitting a linear equation to observed data. Every value of the independent variable x is associated with a value of the dependent variable y.

In order to perform statistical inference using a multiple regression model, several assumptions are made. It is required that the n observations are independent and that the variability of the error ε is constant for all values of the regressors. In addition, as mentioned before, many of the statistical tests require that the error distribution is normal. If the error distribution deviates somewhat from normality, the inference procedures will remain approximately valid. The slope parameters β_k are estimated using least squares, the same as in simple linear regression, by determining the values of the β_k s that minimise the error sum of squares:

$$ESS = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik}))^2 \qquad \textbf{(3-11)}$$

A special case of the multivariate regression model useful for situations, where the relation between a response y and a predictor x appears nonlinear, is a "polynomial regression" model. It attempts to model a situation where a response y is related to a single regressor variable x but the relationship is nonlinear:

$$y = f(x) + \varepsilon \qquad \textbf{(3-12)}$$

In practice, for some nonlinear functions the functional relationship between x and y

is unknown. However, if the function f is "well behaved" then $f(x)$ can be approximated fairly well by a polynomial due to Taylor's theorem[1].

Therefore, the following polynomial model often works well in practice:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + ... + \beta_k x^k + \varepsilon \qquad \textbf{(3-13)}$$

Note that this is just a special case of a multiple regression model with regressors $x, x^2, ..., x^k$. This model is still called a linear model because it is linear in the parameters $\beta_0, \beta_1, ..., \beta_k$ even though the relationship between y and x may be nonlinear.

3.5 NONLINEAR REGRESSION MODEL

The basic idea of nonlinear regression is the same as that of linear regression which is to relate a response y to a vector of predictor variables x . Nonlinear regression is characterised by the fact that the prediction equation depends nonlinearly on one or more unknown parameters. Whereas linear regression is often used for building a purely empirical model, nonlinear regression usually arises when there are physical reasons for believing that the relationship between the response and the predictors follows a particular functional form (Smyth, 2002). The general form of a nonlinear regression model is:

$$y = f(x; \theta_1, ..., \theta_k) + \varepsilon \qquad \textbf{(3-14)}$$

[1] The precise statement of Taylor's theorem is: if $n \geq 0$ is an integer and f is a function which is n times continuously differentiable on the closed interval *[a,x]* and *n+1* times differentiable on the open interval *(a,x)*, then:

$$f(x) = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f^{(2)}(a)}{2!}(x-a)^2 + ... + \frac{f^{(n)}(a)}{n!}(x-a)^n + R$$

where the response variable y is related to the regressor variable x, via an unknown nonlinear function f, which depends on unknown parameters $\theta_1,...,\theta_k$. The unknown parameter vector θ in the nonlinear regression model is estimated from the data by minimising a suitable goodness-of-fit expression with respect to θ. The most popular criterion is the sum of squared residuals:

$$RSS = \sum_{i=1}^{n}(y_i - f(x_i;\theta_1,...,\theta_k))^2 \qquad \textbf{(3-15)}$$

The estimation based on this criterion is known as "nonlinear least squares. The definition of nonlinearity relates to the unknown parameters and not to the relationship between the covariates and the response. For example, based on Equation 3-13, the quadratic regression model:

$$y = \beta_0 + \beta_1 x + \beta_2 x^2 + \varepsilon \qquad \textbf{(3-16)}$$

is considered to be linear rather than nonlinear because the regression function is linear in the parameters β_i and the model can be estimated by using classical linear regression methods (Ratkowsky, 1983).

One of the most common nonlinear models is the "exponential decay" or "exponential growth model":

$$f(x,\theta) = \theta_1 \exp(-\theta_2 x) \qquad \textbf{(3-17)}$$

This model can be characterised by the fact that the function f satisfies the first-order differential equation, leading to higher-order exponential function models, of the form:

$$f(x,\theta) = \theta_1 + \sum_{j=1}^{k}\theta_{2j}\exp(-\theta_{2j+1}x) \qquad \textbf{(3-18)}$$

Where, k is the order of the differential equation (Smyth, 2002).
Another common form of model is the rational function (the ratio of two polynomials):

$$f(x,\theta) = \frac{\sum_{j=1}^{k}\theta_j x^{j-1}}{1+\sum_{j=1}^{m}\theta_{k+j}x^j} \qquad \textbf{(3-19)}$$

Rational functions are very flexible in form and can be used to approximate a wide variety of functional shapes (Kecman, 2001).

In many applications the systematic part of the response is known to be monotonic increasing in x, where x might represent time. Nonlinear regression models with this property are called "growth models". The simplest growth model is the exponential growth model (Equation 3-17). A more generally useful growth curve is the logistic curve:

$$f(x,\theta)=\frac{\theta_1}{1+\theta_2\exp(-\theta_3 x)} \qquad \textbf{(3-20)}$$

Where θ_1, θ_2 and θ_3 are the parameters. Despite these interpretations, it can often be difficult in practice to isolate the interpretations of individual parameters in a nonlinear regression model because of high correlations between the parameter estimators. After obtaining data $(x_1,y_1),\ldots,(x_n,y_n)$ from any regression model setup, the parameters $\theta_1,\ldots,\theta_k$ can be estimated. Least square is used most frequently for parameter estimation.

Correlation is a unitless measure of the amount of linear relationship between two variables. It is computed as the covariance between the two variables divided by the square root of the product of their variances. It varies from -1 to +1. Positive correlation indicates a positive link between the two variables, i.e. when one of them increases, the other has a tendency to increase too. The closer to +1 represents the stronger positive link between the two variables. Negative correlation indicates a negative link between the two variables, i.e. when one of them increases, the other has a tendency to decrease. Similarly, the closer to -1 represents the stronger negative link between them.

3.6 MULTIVARIATE ANALYSIS

There are two basic types of time series forecasting: univariate and multivariate. Univariate models, like Box-Jenkins, contain only one variable in their equation.

Box-Jenkins is a complicated process of fitting data to appropriate model parameters. Multivariate models are univariate models expanded to discover factors that affect the behaviour of the data (Van Eyden, 1996). As the name suggests, these models contain more than one variable in their equations. In fact, multivariate model has been frequently compared with artificial neural networks (Rencher, 2002).

Multivariate analysis deals with issues related to the observations of many, usually correlated, variables on units of a selected random sample. The observations are gathered as vectors, for each selected unit corresponds to a vector of observed variables (Bilodeau and Brenner, 1999). In multivariate analysis, the variables can be examined simultaneously in order to access the key features of the process that produced them. This approach allows firstly exploring the joint performance of the variables and secondly determining the effect of each variable in the presence of the others. Multivariate analysis provides either "descriptive" or "inferential" procedures which mean that the data can search for patterns or test hypotheses about patterns of a priori interest (Rencher, 2002).

In general, multivariate statistical analysis is concerned with analysing and understanding data in high dimensions. For a given a data set $\{x_i\}_{i=1}^{n}$ of n observations of a variable vector X in $\Re^p$, each observation x_i have p dimensions:

$$x_i = (x_{i1}, x_{i2}, \ldots, x_{ip})$$

This is an observed value of a variable vector $\mathbf{X} \in \Re^p$. Therefore, X is composed of p random variables:

$$X = (X_1, X_2, \ldots X_p)$$

where X_j , for $j = 1,2,\ldots,p$, is a one-dimensional random variable.

Covariance is a measure of dependency between random variables. Given two random variables X and Y the theoretical covariance is defined by:

$$\sigma_{XY} = Cov(X,Y) \qquad \textbf{(3-21)}$$

If X and Y are independent of each other, the covariance $Cov(X,Y)$ is necessarily equal to zero (the converse is not true). The covariance of X with itself is the variance:

$$\sigma_{XX} = Var(X) = Cov(X,X) \tag{3-22}$$

As it is mentioned above, assume that the random variable X is p-dimensional multivariate in the form of:

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix}$$

then the theoretical covariance among all the elements are put into another matrix form, i.e., the covariance matrix:

$$\Sigma = \begin{pmatrix} \sigma_{X_1X_1} & \cdots & \sigma_{X_1X_p} \\ \vdots & \ddots & \vdots \\ \sigma_{X_pX_1} & \cdots & \sigma_{X_pX_p} \end{pmatrix} \tag{3-23}$$

Empirical versions of these quantities are:

$$s_{XY} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \tag{3-24}$$

$$s_{XX} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{3-25}$$

For small n ($n \le 20$) the factor $1/n$ in both above equations should replaced by $1/(n-1)$ in order to correct for a small bias (Hardle and Simar, 2003). For a p-dimensional random variable the covariance matrix is:

$$S = \begin{pmatrix} s_{X_1X_1} & \cdots & s_{X_1X_p} \\ \vdots & \ddots & \vdots \\ s_{X_pX_1} & \cdots & s_{X_pX_p} \end{pmatrix} \tag{3-26}$$

Eigen values and Eigen vectors play an important role in multivariate techniques.

For every square matrix **A**, i.e. $(p \times p)$, a scalar λ and a non-zero vector γ can be found such that (Rencher, 2002):

$$\mathbf{A}\gamma = \lambda\gamma \qquad \textbf{(3-27)}$$

λ is called an Eigen value of **A** and γ is an Eigen vector of **A** corresponding to λ. To find λ and γ, Equation 3-27 can be written as the form of:

$$(\mathbf{A} - \lambda\mathbf{I})\gamma = 0 \qquad \textbf{(3-28)}$$

where **I** is the identity matrix. It can be proven that an Eigen value λ is a root of the *p-th* order polynomial, which means that:

$$|\mathbf{A} - \lambda\mathbf{I}_p| = 0 \qquad \textbf{(3-29)}$$

Therefore, there are up to *p* Eigen values $\lambda_1, \lambda_2, ..., \lambda_p$ of vector **A**. For each Eigen value λ_j, there exists a corresponding eigenvector *j* given by Equation 3-27. After finding $\lambda_1, \lambda_2, ..., \lambda_p$, the accompanying eigenvectors $\gamma_1, \gamma_2, ..., \gamma_p$ can be found subsequently.

The eigenvectors are positioned along the directions of greatest data variance. They are found from the covariance matrix of the input dataset. An Eigen value $\lambda_p, i = 1, ..., P$ is associated with each eigenvector. Every input data vector is then represented by a linear combination of eigenvectors (Mandic and Chambers, 2001). Eigen values measure how much variance of the data set the eigenvectors account for. The larger the Eigen values, the better the eigenvectors represent the data set. The Principal Components (PCs) are the vectors that minimise the Mean Square Error (MSE) between the actual points in the data set and the points described by a smaller number of components (Kasabov, 1998).

There are sets of methods dedicated to statistical multivariate analysis. These methods use information in the correlation structure amongst response variables (*Y*-variables) which often increases the power of the statistical analysis to detect treatment difference as compared to the corresponding univariate methods. The correlation *Y*-variables can be utilised to potentially increase the power of detecting differences above univariate statistical procedures. Each response variable *Y* adds another dimension to the analysis problem. Many of these techniques were developed recently because of the existence of modern computers and their high computational capabilities. Abdi (2003) classified these methods according to the number of data sets to analyse: one data set and two (or more) data sets. With two data sets, two cases have been considered: in the first case, one set of data plays the role of predictors, or independent, variables (*X*-variables) and the second set of data corresponds to measurements or dependent variables (*Y*-variables). In the second case, the different sets of data correspond to different sets of *Y*-variables.

Methods like Principal Component Analysis (PCA), Correspondence Analysis (CA), Multiple Correspondence Analysis (MCA) and Multidimensional Scaling (MDS) can be classified as having one data set. If there are two or more data sets techniques like Multiple Linear Regression analysis (MLR), Partial Least Square (PLS), Partial Least Square Regression (PLSR), Principal Component Regression (PCR), Multivariate Analysis of Variance (MANOVA), Discriminant Analysis (DA) and Reduced Rank Regression (RRR) are suitable for the mentioned first case. In addition, Canonical Correlation Analysis (CC), Multiple Factor Analysis (MFA), Multiple Correspondence Analysis (MCA) techniques are appropriate for the mentioned second case. A few of these techniques are introduced in the following sections.

3.6.1 Multiple Linear Regression (MLR)

Multiple Linear Regression (MLR) is a multivariate analysis method which relates the variations in a response variable (Y-variable) to the variations of several predictors (X-variables), with explanatory or predictive purposes. In this method, which is based on ordinary least squares regression, several *Y*-variables are measured corresponding to each set of *X*-variables. Each of $y_1, y_2, \cdots, y_p$ is to be predicted by all of $x_1, x_2, \cdots, x_q$. The *n* observed values of the vector of *Y*-variables can be listed as rows in the following matrix (Rencher 2002):

$$\mathbf{Y} = \begin{pmatrix} y_{11} & y_{12} & \cdots & y_{1p} \\ y_{21} & y_{22} & \cdots & y_{2p} \\ \vdots & \vdots & & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{np} \end{pmatrix} = \begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{pmatrix}$$

Each row of **Y** contains the values of the p dependent variables measured on a subject. Each column of **Y** consists of the n observations on one of the p variables. The n values of $x_1, x_2, \cdots, x_q$ can be placed in a matrix that turns out to be the same as the **X** matrix in the multiple regression formulation:

$$\mathbf{X} = \begin{pmatrix} 1 & x_{11} & \cdots & x_{1q} \\ 1 & x_{21} & \cdots & x_{2q} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nq} \end{pmatrix}$$

Since each of the p Y-variables will depend on the X-variables in its own way, each column of **Y** will need different βs. Thus there is a column of βs for each column of **Y**, and these columns form a matrix $\mathbf{B} = (\beta_1, \beta_2, \cdots, \beta_p)$. Therefore, the multivariate model is:

$$\mathbf{Y} = \mathbf{XB} + \boldsymbol{\varepsilon} \qquad \textbf{(3-30)}$$

$\boldsymbol{\varepsilon}$ is a noise term for the model which has the same dimensions as **Y**. The model coefficient **B** called the least squares estimator and can be calculated by the equation (Rencher 2002):

$$\mathbf{B} = (\mathbf{X'X})^{-1}\mathbf{X'Y} \qquad \textbf{(3-31)}$$

An assumption for the MLR method is that the X-variables are linearly independent. When the X-variables carry common information, problems can arise due to linear relationship between variables called Collinearity. Two variables are collinear if the value of one variable can be computed from the other, using a linear relation. Three or more variables are collinear if one of them can be expressed as a linear function of the others. The MLR operation involves a matrix inversion, which leads to collinearity problems if the variables are not linearly independent. This is the reason why the predictors are called independent variables in this method. The ability to vary independently of each other is a crucial requirement to variables used as predictors with this method. MLR also requires more samples than predictors or the matrix cannot be inverted.

In MLR, all the X-variables are supposed to participate in the model independently of each other. Their co-variations are not taken into account, so X-variance is not meaningful in MLR. Thus the only relevant measure of how well the model performs is provided by the Y-variances. Residual is a measure of the variation that is not taken into account by the model and the residual variance of a variable is the mean square of its residuals for all model components. The ratio between explained variance and residual variance called the F-ratio. It shows how large the effect of the predictor is, as compared with random noise. The F-ratio associated to every tested effect is computed as the ratio of mean squared of the model to mean squared of error. These ratios, which compare structured variance to residual variance, have a statistical distribution which is used for significance testing. In other words the higher the ratio, the more important the effect.

3.6.2 PRINCIPAL COMPONENT ANALYSIS (PCA)

In MLR, when the X-variables carry common information, problems can arise due to exact or approximate collinearity. Variables which are not collinear are called "linearly independent". Collinearity, i.e. very strong correlation, is the major cause of trouble for MLR models, whereas projection methods like PCA, PCR and PLS handle collinearity well (Abdi, 2003).

The main objective of PCA is to reduce the dimension of the observations or dimensionality (Everitt & Dunn, 2001; Maddala, 1997). In fact, the goal of PCA is to decompose a data table with correlated measurements into a new set of uncorrelated (i.e., orthogonal) variables. These variables are called, depending upon the context, principal components, factors, eigenvectors, singular vectors, or loadings. Each unit is also assigned a set of scores which correspond to its projection on the components (Abdi, 2003).

Principal component analysis is the amount of variance in a variable that is shared by all the variables in the analysis. The analysis summarises and groups nearly all of the original information into a smaller number of factors that can then be used for estimation purposes. This involves a mathematical procedure that transforms a number of correlated variables into a smaller number of uncorrelated variables called "principal components". The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. The components are chosen based upon their Eigen values i.e. components with Eigen values greater than the average were chosen (Rabunal and Dorado, 2006).

In matrix algebra, the transpose of a matrix A, denoted by $\mathbf{A}^T$, is obtained from A by interchanging rows and columns. Also, a simple function of a $p \times p$ matrix A is the trace, denoted by "tr A" and defined as the sum of the diagonal elements of A; that is:

$$tr\mathbf{A} = \sum_{i=1}^{p} a_{ii} \tag{3-32}$$

The trace is a scalar.

A matrix $\mathbf{H} \in \Re_p^p$ is said to be orthogonal if the columns (or rows) of H form an orthonormal basis of $\Re_n$, i.e. $\mathbf{H}^T\mathbf{H} = \mathbf{I} = \mathbf{H}\mathbf{H}^T$. The group of orthogonal matrices in $\Re_p^p$ will be denoted by:

$$\mathbf{O}_p = \{\mathbf{H} \in \Re_p^p : \mathbf{H}\mathbf{H}^T = \mathbf{I}\}$$

According to the above definitions, the total variance of variable x is defined as (Bilodeau and Brenner, 1999):

$$E\,|\,x-\mu\,|^2 = \sum_{i=1}^{p} \operatorname{var} x_i = \sum_{i=1}^{p} \sigma_{ii} = tr\,\Sigma \tag{3-33}$$

If $\Sigma \geq 0$ can be written as[1]:

$$\Sigma = \mathbf{HDH}^T \tag{3-34}$$

where $\mathbf{H} = (\mathbf{h}_1, \ldots, \mathbf{h}_p) \in \mathbf{O}_p$ and $\mathbf{D} = diag(\lambda_1, \ldots, \lambda_p)$ and $\lambda_1 \geq \ldots \geq \lambda_p$ are the ordered Eigen values of Σ. Since the "*var x*" is interested, it can be assumed that $\mu = 0$, so:

---

[1] In the general case where Σ is not diagonal, there exists $\mathbf{H} \in \mathbf{O}_p$ such that: $\Sigma = \mathbf{H}diag(d_1\mathbf{I}_{p1}, \ldots, d_k\mathbf{I}_{pk})\mathbf{H}^T = \mathbf{HDH}^T$

$$y = \mathbf{H}^T x = \begin{pmatrix} h^T{}_1 x \\ \vdots \\ h^T{}_p x \end{pmatrix} \qquad \textbf{(3-35)}$$

$\operatorname{var} y = \mathbf{D}$, then

$$\sum_{i=1}^{p} \operatorname{var} y_i = \sum_{i=1}^{p} \lambda_i = tr\,\Sigma \qquad \textbf{(3-36)}$$

So x and y have the same "total variance". Moreover, the variables y_i s are uncorrelated.

$$\operatorname{cov}(h_i^T x, h^T{}_j x) = h_i^T h_j = \lambda_j h_i^T h_j = \lambda_j \delta_{ij} \qquad \textbf{(3-37)}$$

The variables $y_i = h_i^T x$, $i = 1, \ldots, p$, are, by definition, the principal components of x. When the ratio $\sum_{i=1}^{k} \lambda_i \Big/ tr\,\Sigma$ is close to 1, then $(y_1, \ldots, y_k)^T$ can effectively replace x without losing much in terms of "total variance."

In PCA, the first mode finds a straight line approximation to a given dataset, which accounts for the maximum amount of variance in the data. In fact, it finds a line which passes through the "middle" of the data cluster (Hsieh, 2004).

Non Linear PCA, denoted as NLPCA, was introduced by Kramer in 1991(Kramer, 1991). He proposed a model where the straight line is replaced by a continuous open curve for approximating the data. The fundamental difference between NLPCA and PCA is that PCA only allows a linear mapping between variable x and the principal component u while NLPCA allows a nonlinear mapping.

Principal Component Regression (PCR) is a two-step procedure which first decomposes the X-matrix by PCA, then fits a MLR model, using the PCs instead of the original X-variables as predictors.

In both PCA and PCR weighting of the X-variables has an important role to the accuracy of the final model. In fact, the weighting determines the effect of each variable to the model and it can change the influence of variables on the model. To have equal chances for all X-variables to influence the model the considered initial weighting for every individual variable can be: 1/ Variable Standard Deviation (Hsieh, 2004). This gives equal influence to all the X-variables at the beginning of the modelling. Since the modelling has been completed it is necessary to ensure the obtained model is accurate. Validation is important to check whether the model will make a good fit on future data not used in the original computations. A good model should generally describe data similar to that available when building the model. Cross Validation (CV) is a method to verify the accuracy of the model and simulate test set validation. In this method a few samples are left out from the calibration data set and the model is calibrated on the remaining data points. Then the values for the left-out samples are predicted and the prediction residuals are computed. There are a few validating methods including: full CV, random CV and Systematic CV. The difference between these methods is the way that they choose the validation sets (Hsieh, 2004). For example, Random Cross Validation method picks the samples at random. It is may be time consuming but it can assesses the stability of the PCR results.

3.6.3 PARTIAL LEAST SQUARES (PLS)

Partial Least Squares (PLS) regression (World *et al.*, 1984) is an extension of the mentioned Multiple Linear Regression model (MLR).Therefore, the main purpose of PLS regression is to build a linear model as the Equation 3-30.

PLS is a method for relating the variations in one or several response variables (Y-variables) to the variations of several predictors (X-variables), with explanatory or predictive purposes. PLS is a bilinear modelling method where information in the original X-data is projected onto a small number of underlying variables, called PLS components or latent variables or scores (Hsieh, 2004).

PLS, like PCR, reduces the dimension of the observations or dimensionality but the reduction and regression are performed simultaneously, i.e. PLS outputs the matrix of least square estimator **B** (Equation 3-31) as well as the loadings and weight and latent component matrices.

PLS can be seen as methods to construct a matrix of latent components **T** as a linear transformation of **X** (Boulesteix and Strimmer, 2006):

$$\mathbf{T} = \mathbf{XW} \qquad (3\text{-}38)$$

where **W** is a $p \times c$ matrix of weights. The latent variables are then used for prediction in place of the original variables. The Y-data are actively used in estimating the latent variables to ensure that the first components are those that are most relevant for predicting the Y-variables. Interpretation of the relationship between X-data and Y-data is then simplified as this relationship in concentrated on the smallest possible number of components. PLS models both the X- and Y-matrices simultaneously to find the latent variables in X that will best predict the latent variables in Y. These PLS components are similar to Principal Components.

Loading weights are specific to PLS (they have no equivalent in previous methods) and express how the information in each X-variable relates to the variation in Y. To summarising the variation in the X- or Y-space, two different sets of components can be considered in PLS method:

- *t*-scores are the new coordinates of the data points in the X-space, computed in such a way that they capture the part of the structure in X which is most predictive for Y.
- *u*-scores summarise the part of the structure in Y which is explained by X along a given model component.

In fact, the relationship between t- and u-scores is a summary of the relationship between X and Y along a specific model component. This method performs particularly well when the various X-variables express common information, i.e. when there is a large amount of correlation, or even collinearity.

There are two versions of the PLS algorithms:

- PLS1: deals with only one response variable at a time;
- PLS2: handles several responses simultaneously.

PCR and PLS regression differ in the methods used in extracting factor scores. PCR produces the weight matrix reflecting the covariance structure between the predictor variables, while PLS regression produces the weight matrix reflecting the covariance structure between the predictor and response variables.

3.6.4 Multivariate Analysis of Variance (MANOVA)

The analysis of variance is a classical method to assess the significance of effects by decomposition of a response's variance into explained parts, related to variations in the predictors, and a residual part which summarises the experimental error. It can be divided into two groups: Univariate Analysis of Variance called ANOVA and Multivariate Analysis of Variance (MANOVA).

In ANOVA, it is assumed that the average values of the response variable y are induced by one simple factor. Suppose that this factor takes on p values and that for each factor level, there are $m = n/p$ observations. The sample is of the form given in Table 3-1, where all of the observations are independent (Hardle and Simar, 2003).

Sample Elements	Factor Levels l				
1	y_{11}	$\cdots$	y_{1l}	$\cdots$	y_{1p}
2	$\vdots$		$\vdots$		$\vdots$
$\vdots$	$\vdots$		$\vdots$		$\vdots$
k	y_{k1}	$\cdots$	y_{kl}	$\cdots$	y_{kp}
$\vdots$	$\vdots$		$\vdots$		$\vdots$
$m = {}^{n}\!/_{p}$	y_{m1}	$\cdots$	y_{ml}	$\cdots$	y_{mp}

Table 3-1: Observation structure of a simple ANOVA (Hardle and Simar, 2003).

The objective of a simple ANOVA is to analyse the observation structure: $y_{kl} = \mu_l + \varepsilon_{kl}$ for $k = 1,\dots,m$ and $l = 1,\dots,p$. Each factor has a mean value μ_l. Each observation y_{kl} is assumed to be a sum of the corresponding factor mean value μ_l and a zero mean random error ε_{kl}. The linear regression model falls into this scheme with $m = 1, p = n$ (Equation 3-7) and $\mu_i = \alpha + \beta x_i$, where x_i is the *i-th* level value (Hardle and Simar, 2003).

In MANOVA, the X-variables have the same structure as in a standard ANOVA and are used to predict a set of Y-variables. MANOVA computes a series of ordered orthogonal linear combinations of the Y-variables with the constraint that the first factor generates the largest **F** if used in an ANOVA. The sampling distribution of this **F** is adjusted to take into account its construction (Abdi, 2003). There are three basic variations of MANOVA (Wendorf, 2004):

- *Hotelling's T*: This is the MANOVA analogue of the two group *T*-test situation; in other words, one dichotomous independent variable and multiple dependent variables.
- *One-Way MANOVA*: This is the MANOVA analogue of the one-way *F* situation; in other words, one multi-level nominal independent variable and multiple dependent variables.

- *Factorial MANOVA*: This is the MANOVA analogue of the factorial ANOVA design; in other words, multiple nominal independent variables, and multiple dependent variables.

All the above methods have one feature in common: they form linear combinations of the Y-variables which best discriminate among the groups in the particular experimental design. In other words, MANOVA is a test of the significance of group differences in some m-dimensional space where each dimension is defined by linear combinations of the original set of dependent variables (Wendorf, 2004).

In the multivariate analysis, it is assumed that k independent random samples of size n are obtained from p-variate normal populations with equal covariance matrices, as in the following layout (Table 3-2) for one-way multivariate analysis of variance (Rencher, 2002).

	Sample 1 from $N_p(\mu_1, \Sigma)$	Sample 2 from $N_p(\mu_2, \Sigma)$	...	Sample k from $N_p(\mu_k, \Sigma)$
	y_{11}	y_{21}	...	y_{k1}
	y_{12}	y_{22}	...	y_{k2}
	$\vdots$	$\vdots$		$\vdots$
	y_{1n}	y_{2n}	...	y_{kn}
Total	$y_{1.}$	$y_{2.}$	...	$y_{k.}$
Mean	$\bar{y}_{1.}$	$\bar{y}_{2.}$	...	$\bar{y}_{k.}$

Table 3-2: the observation structure of a one-way MANOVA (Rencher, 2002) .

Totals and means are defined as follows:

- Total of the *i-th* sample: $y_{i.} = \sum_{j=1}^{n} y_{ij.}$

- Overall total: $y_{..} = \sum_{i=1}^{k}\sum_{j=1}^{n} y_{ij.}$

- Mean of the *i-th* sample: $\bar{y}_i = {y_{i.}}/{n.}$

- Overall mean: $y_{..} = {y_{..}}/{kn.}$

Also, the model for each observation vector is (Rencher, 2002):

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij} = \mu_i + \varepsilon_{ij} \qquad \text{(3-39)}$$
$$\text{for: } i = 1,\ldots,k \text{ and } j = 1,\ldots,n$$

The above equation in terms of the *p* variables in y_{ij} becomes:

$$\begin{pmatrix} y_{ij1} \\ y_{ij2} \\ \vdots \\ y_{ijp} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix} + \begin{pmatrix} \alpha_{i1} \\ \alpha_{i2} \\ \vdots \\ \alpha_{ip} \end{pmatrix} + \begin{pmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \vdots \\ \varepsilon_{ijp} \end{pmatrix} = \begin{pmatrix} \mu_{i1} \\ \mu_{i2} \\ \vdots \\ \mu_{ip} \end{pmatrix} + \begin{pmatrix} \varepsilon_{ij1} \\ \varepsilon_{ij2} \\ \vdots \\ \varepsilon_{ijp} \end{pmatrix} \qquad \text{(3-40)}$$

So that the model for the *r-th* variable ($r = 1,2,\ldots,p$) in each vector y_{ij} is:

$$y_{ijr} = \mu_r + \alpha_{ir} + \varepsilon_{ijr} = \mu_{ir} + \varepsilon_{ijr} \qquad \text{(3-41)}$$

3.6.5 Canonical Correlation Analysis (CCA)

Canonical Correlation Analysis (CCA) technique was originally developed by Hotelling (1935). The aim of CCA is to identify and quantify the relations between a *p*-dimensional random X-variable and a *q*-dimensional random Y-variable. In fact, this is a technique that can be used to study the relationship between two sets of variables, each of which might contain several variables. Its purpose is to summarise or explain the relationship between two sets of variables by finding a small number of linear combinations from each set of variables that have the highest correlation possible between the sets. The associations between two sets of variables may be identified and quantified by canonical correlation analysis. Canonical correlation combines the Y-variables to find pairs of new variables (called canonical variables, or CV, one for each data table) which have the highest correlation. However, the CVs, even when highly correlated, do not necessarily explain a large portion of the variance of the original tables. This make the interpretation of the CV sometimes difficult but CC is nonetheless an important theoretical tool because most multivariate techniques can be interpreted as a specific case of canonical correlation (Abdi, 2003).

CCA finds a pair of linear transformations such that the correlation coefficient between extracted features is maximised (Akaho, 2001) (Figure 3-3). For given two random variables $X \in \Re^q$ and $Y \in \Re^p$, the idea is to find an index describing a link between X and Y i.e. measure the overall correlation between X and Y.

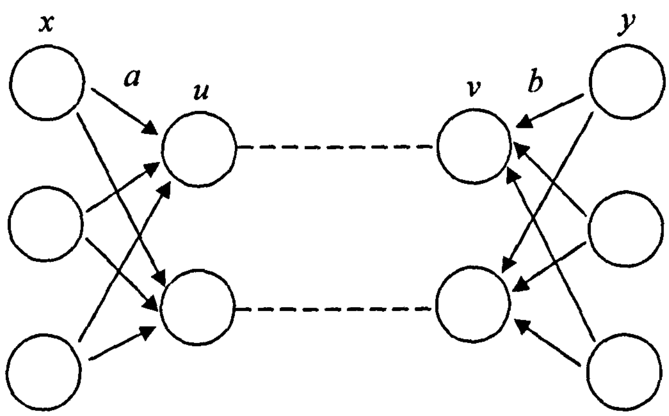


Figure 3-3: Canonical Correlation Analysis (Akaho, 2001)

CCA is based on linear indices, i.e. linear combinations $a^T X$ and $b^T Y$ of the random variables. It searches for vectors a and b such that the relation of the two indices $a^T X$ and $b^T Y$ is quantified in some interpretable way. Then the aim is to look for the "most interesting" projections a and b in the sense that they maximise the below correlation between the two indices (Hardle and Simar, 2003):

$$\rho(a,b) = \rho a^T X b^T Y \qquad \textbf{(3-42)}$$

To study the correlation $\rho(a,b)$ between the two projections in more details, suppose that:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \approx \left(\begin{pmatrix} u \\ v \end{pmatrix}, \begin{pmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{pmatrix} \right)$$

where the sub-matrices of this covariance structure are given by

$$\mathrm{var}(X) = \Sigma_{XX}(q \times q) \tag{3-43}$$

$$\mathrm{var}(Y) = \Sigma_{YY}(p \times p) \tag{3-44}$$

$$\mathrm{cov}(X,Y) = E(X-u)(Y-v)^T = \Sigma_{XY} = \Sigma_{YX}^T(q \times p) \tag{3-45}$$

Considering general correlation between variables X and Y

$$\rho_{XY} = \frac{\mathrm{cov}(X,Y)}{\sqrt{\mathrm{var}(X)\,\mathrm{var}(Y)}} \tag{3-46}$$

and the facts that

$$\mathrm{cov}(AX,BY) = A\,\mathrm{cov}(X,Y)B^T \tag{3-47}$$

and

$$\rho(a,b) = \frac{a^T \Sigma_{XY} b}{(a^T \Sigma_{XX} a)^{1/2}(b^T \Sigma_{YY} b)^{1/2}}, \tag{3-48}$$

$\rho(ca,b) = \rho(a,b)$ for any $c \in \Re^+$. Given the invariance of scale, rescale projections a and b can be rescaled and thus it is possible to equally solve (Hardle and Simar, 2003)

$$\max_{a,b} = a^T \Sigma_{XY} b \tag{3-49}$$

under the constraints:

$$a^T \Sigma_{XX} a = 1$$

$$b^T \Sigma_{YY} b = 1.$$

For this problem, define:

$$\kappa = \Sigma_{XX}^{-1/2}\Sigma_{XY}\Sigma_{YY}^{-1/2} \tag{3-50}$$

The matrix κ may be decomposed as[1]:

$$\kappa = \Gamma\Lambda\Delta^T \tag{3-51}$$

with:

$$\Gamma = (\gamma_1,\ldots,\gamma_k)$$
$$\Delta = (\delta_1,\ldots,\delta_k)$$
$$\Lambda = diag(\lambda_1^{1/2},\ldots,\lambda_k^{1/2})$$

where $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$ are the non-zero Eigen values of $N_1 = \kappa\kappa^T$ and $N_2 = \kappa\kappa^T$ and γ_i and δ_j are the standardised eigenvectors of N_1 and N_2 respectively.
Define now for $i = 1,\ldots,k$ the vectors:

$$a_i = \Sigma_{XX}^{-1/2}\gamma_i \tag{3-52}$$
$$b_i = \Sigma_{YY}^{-1/2}\delta_i \tag{3-53}$$

which are called the "canonical correlation vectors". Using these canonical correlation vectors to define the canonical correlation variables:

$$\eta_i = a_i^T X \tag{3-54}$$
$$\varphi_i = b_i^T Y \tag{3-55}$$

The quantities $\rho_i = \lambda_i^{1/2}$ for $i = 1,\ldots,k$ are called the "canonical correlation coefficients". For simplicity, if assume that the average of x and y are 0, and the

---

[1] Each matrix $A(n \times p)$ with rank r can be decomposed as $A = \Gamma\Lambda\Delta^T$ where $\Gamma(n \times r)$ and $\Delta(p \times r)$. Both Γ and Δ are column orthonormal i.e. $\Gamma^T\Gamma = \Delta^T\Delta = I_r$.

dimensionality of feature is 1, then by the transformations $u = \{a, x\}$ and $v = \{b, y\}$ it is favourable to find the transformation *a, b* that maximises (Akaho, 2001)

$$\rho = \frac{E \mid uv \mid}{\sqrt{\text{var}[u]\,\text{var}[v]}} \qquad \textbf{(3-56)}$$

where $\{a, x\}$ represents the inner product. The maximum ρ is the maximum canonical correlation.

3.7 CONCLUSION

Multivariate analysis methods have been frequently compared with the ANN methods (Rencher, 2002 and Abdi, 2003). Therefore, several multivariate analysis methods have been introduced and their fundamentals have been explained in this chapter.

It has been shown that statistical techniques can model the relation between dependant and independent variables. Each of them can perform better and more accurately in a particular situation. There is a problem that there is considerable difficulty in selecting the appropriate technique. For example, as explained in section 3-6-1, it is considered that MLR performs better when there is not collinearity between variables but, as explained in section 3-6-2, projection methods like PCR and PLS are able to handle collinearity. The correct analysis method should be chosen for each particular situation as the problem of technique selection become significant when the number of variables is large and the nature of some variables is not clear.

Later in chapter 5, these methods will be used to model the demolition market.

CHAPTER 4

ANN APPROACHES

4.1 INTRODUCTION

In this chapter, the fundamentals of Biological and Artificial Neural Networks (ANNs) are explained and comparisons are made. This is followed by a historic review of the development of ANNs and an explanation of basic features and parameters. These are explained in detail to clarify the concepts and structures. In addition, after demonstrating various network topologies and learning methods, major applications of ANNs are introduced and different models and architectures are demonstrated.

In the last part of this chapter, some case studies are reviewed to show how ANNs are used in various fields of interest. Furthermore, the disadvantages of using ANNs are explained and solutions to problems discussed.

4.2 NEURAL NETWORKS CONCEPT

Neural Networks are a network of neurons which are interconnected and operate together for computation purposes. The neuron, as the fundamental component of the network, is a processing element and the combination of all the neurons in a network function together to produce the favourable output. Neurons have the ability to process individually and each one is connected to a large number of others, hence the whole information processing in a network is carried out on a parallel basis. This gives a relatively high progression speed compared with serial systems. Also, malfunctions of a few neurons would not harm the global procedure. In general, neural networks are able to learn and then solve the problems if they can be trained correctly.

4.2.1 BIOLOGICAL NEURAL NETWORKS

The brain, as an information processing system, is a biological neural network which consists of almost 10 billion processing elements. In fact, nerve cells are the processing element for a brain, called "neurons". Each neuron is connected to the others through synapses. The number of those depends on the local neuroanatomy but it is roughly around 10,000 synapses for each neuron. Non-nervous cells, such as muscles, require signals from the nervous system in order to operate. Therefore, synapses also connect these cells to neurons. This makes an enormous parallel information system. The brain can learn from the experiences and solve problems that it has never encountered before (Madan et al. 2003).

Figure 4-1 represents the major components of a typical neuron in the central nervous system. Four basic components are known by their biological names: dendrites, soma, axon and synapses. Dendrites are hair-like extensions of the soma which act like input channels. These input channels receive their inputs through the synapses of other neurons. The soma then processes these incoming signals over time. The soma then turns that processed value into an output which is sent out to other neurons.

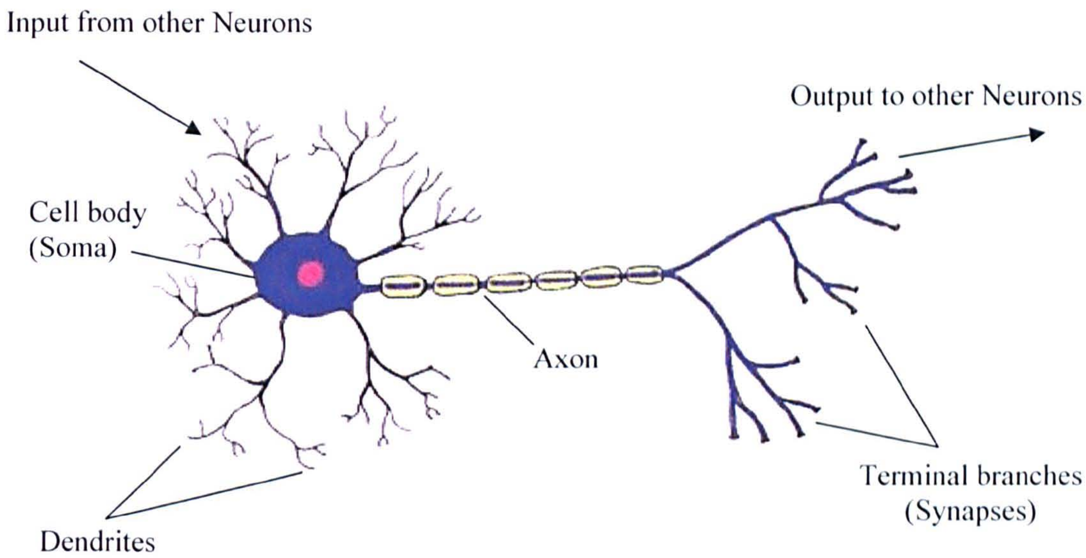


Figure 4-1: The major structures of a typical neuron (Madan et al. 2003)

4.2.2 ARTIFICIAL NEURAL NETWORKS

The artificial neurons are elementary information processing units for the ANN. It may also be called as a unit, node or a Processing Element (PE). The structure is almost the same as a biological neuron which means a neuron receives a number of inputs and then sums them to create the desirable output. Compared with the biological neuron, inputs represent dendrites and outputs represent synapses. Typically the sum of each neuron is weighted, and the sum is passed through a non-linear function known as an activation or transfer function. An artificial neuron with a threshold activation function is shown in Figure 4-2.

The connections (synapses) w_i transfer the signals u_i into the neuron. Where w_i can be interpreted as a weight representing the "importance" of that specific input x_i. Inside the neuron the sum of the weighted inputs $w_i x_i$ is calculated. Given that the sum x is greater than an externally applied threshold θ, the neuron emits an output y. y which is either continuous or binary valued, depending on the activation function. In most cases one chooses an activation function that limits the range of the neuron's output to the interval *[0, 1]* or *[−1, 1]*.

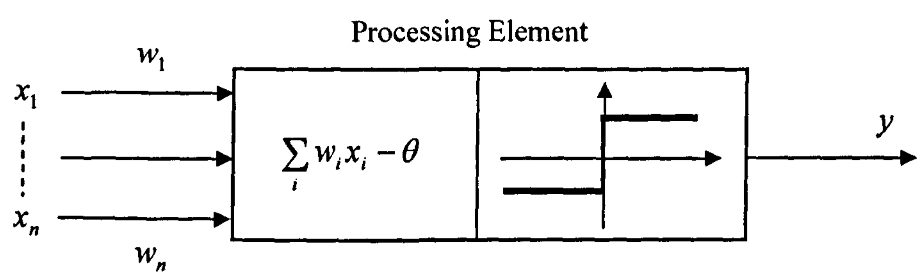


Figure 4-2: The artificial neuron with a threshold function. (Nygren, 2004)

In mathematical terms the following equations gives a dense description of the neuron (Haykin, 1994):

$$u = \sum_{i=1}^{n} w_i x_i - \theta \qquad \textbf{(4-1)}$$

and

$$y = f(u)$$

where x is the network input and $f(\cdot)$ the activation function. There are several common activation functions like the sigmoid, linear and hyperbolic functions which are discussed later in this chapter. Individual PEs are linked together in different ways to create different ANN architectures for various applications.

4.3 FUNDAMENTALS OF ARTIFICIAL NEURAL NETWORKS

4.3.1 HISTORIC REVIEW

ANNs were gradually created over time and previous research was continuously built on by scientists. Various types of networks were created during this process and different methods of operation analysed.

The first step in this direction came in 1943 when neurophysiologist Warren McCulloch and mathematician Walter Pitts, wrote a paper on how neurons might work. In fact, they modeled a simple neural network with electrical circuits. They studied the potential of the interconnection of a model of a neuron and proposed a computational model based on a simple neuron-like element (McCulloch and Pitts, 1943). Reinforcing this concept of neurons and how they work was a book written in 1949 by Donald Hebb, "The Organization of Behavior: A Neuro-psychological Theory". It pointed out that neural pathways are strengthened each time that they are used. He devised a learning rule for adapting the connections within artificial neurons (Hebb, 1949). In 1958, Rosenblatt created the name "perceptron". Based upon the perceptron, he developed the theory of statistical separability (Rosenblatt, 1958). In 1959, Bernard Widrow and Marcian Hoff of Stanford developed models they called ADALINE (ADAptive LINear Elements) and MADALINE (Multiple ADAptive LINear Elements). MADALINE was the first neural network to be applied to a real-world problem. It was an adaptive switching circuit which eliminated echoes on phone lines. The next major progress was the new formulation of learning rules by Widrow and Hoff in their ADALINE model (Widrow and Hoff 1960). In 1969, Minsky and Papert provided a rigorous analysis of the perceptron (Minsky and Papert, 1969).

Grossberg proposed several new architectures of nonlinear dynamical systems and introduced adaptive resonance theory (ART) (Grossberg, 1974) which is a real-time ANN that performs supervised and unsupervised learning of categories, pattern classification and prediction. He continued his previous work in 1976 based on biological and psychological evidence.

In 1971 Werbos developed a backpropagation learning algorithm which he published in his doctoral thesis (Werbos, 1974). Rumelhart et al. rediscovered this technique in 1986 (Rumelhart et al. 1986).

The idea of an associative memory network introduced by Kohonen in 1977 (Kohonen, 1977). He used faces to illustrate the potential of a linear autoassociative network (where the input and output patterns are identical) to act as a parallel distributed memory for images. Later in 1982 he also introduced Self-Organised Maps (SOM) network (Kohonen, 1982), which is a type of unsupervised learning, for pattern recognition (Burr 1993). The first application area of the SOM was speech recognition, or perhaps more accurately, speech-to-text transformation (Kohonen et al., 1984).

In the early 1980s, researchers showed renewed interest in neural networks. Recent work includes Boltzmann machines, Hopfield nets, competitive learning models, multilayer networks, and adaptive resonance theory models (Burr 1993).
By 1985 the American Institute of Physics began what has become an annual meeting - Neural Networks for Computing. The historical notes are presented in Figure 4-3.

ANN progression over time

1943 – McCulloch and Pits:
A Logical Calculus of the Ideas Immanent in Nervous Activity

1949 – Hebb:
The Organization of Behavior: A Neuro-psychological Theory.

1958 - Rosenblatt:
The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain.

1960 - Widrow, Hoff:
Adaptive Switching Circuits.

1969 - Minsky, Papert:
Perceptrons.

1974 - Werbos:
Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Science.

1977 - Kohonen:
Associative Memory - A System Theoretic Approach.

1982 - Hopfield:
Neural Networks and Physical Systems with Emergent Collective Computational Abilities.

1984 - Hinton, Sejnowski, Ackley:
Boltzmann Machines: Constraint Satisfaction Networks that Learn.

1986 - Rumelhart, McClelland:
Parallel Distributed Processing.

Figure 4-3: The Artificial Neural Networks historical notes

4.3.2 ANN CONSTRUCTION

Artificial Neural Networks exploits an analogy to the human brain. The idea behind ANN was to transfer the idea of parallel distributed processing, as found in the brain, to the computer in order to take advantage of the processing features of the brain (Magnani and Nersessian, 2002).

As mentioned in the previous section, the brain consists of large numbers of neurons connected to each other by synapses. The output from the neuron is a function of its inputs from many other neurons, which are "weighted" at the receiving synapses. This output is a nonlinear function of its input and the strength of the connection in the synapses can be modified by activity; in other words, the brain (the collection of neurons) learns (changes its synaptic weights) from experience. This is the behaviour which an ANN attempts to model algorithmically. The assumption that learning occurs in the brain when modifications are made to the effective coupling between one cell and another at a synaptic junction is simulated mathematically in artificial systems through positive or negative reinforcement of connections (Bailer-Jones, 2001). This forms the basis of the analogy exploited in artificial neural networks. A schematic pattern of the ANN architecture is shown in Figure 4-4 .

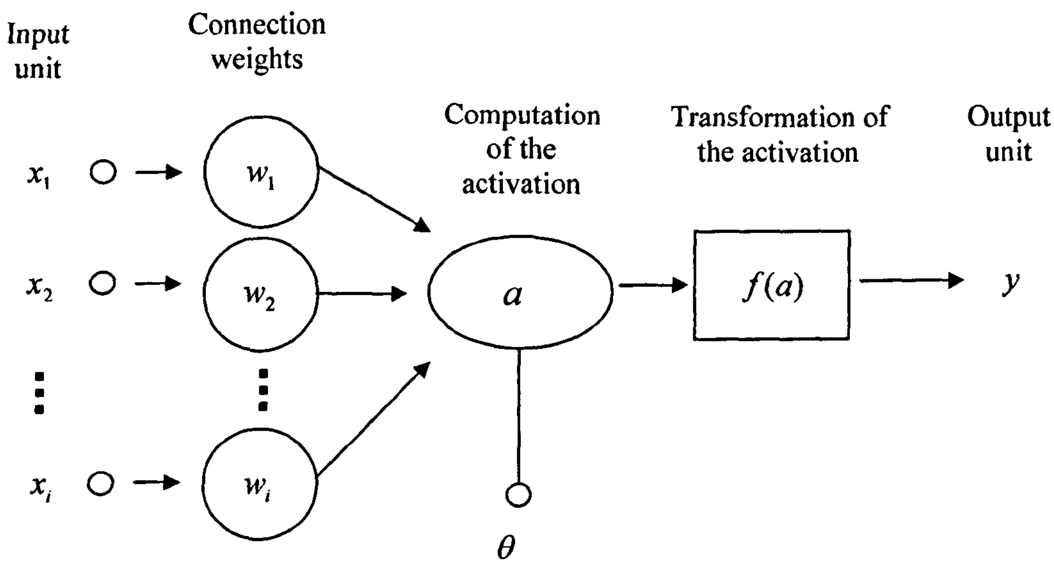


Figure 4-4: Schematic ANN Architecture (Haykin, 1994).

Neural network units receive weighted inputs from the original data or from an adjoining unit. Each unit integrates incoming information, usually by computing the weighted sum of all inputs to determine the level of activation. Formally, if each input is denoted x_i, and each weight w_i then the activation is equal to:

$$a = \sum_{i=1}^{n} x_i w_i + \theta \qquad \textbf{(4-2)}$$

Where n is the dimension of the input space and θ is the bias. The difference between the network expected value and the true value of the output being estimated is called the bias.

The response of the unit is then determined by an activation function *f(a)*. This transformation involves two steps: First, the activation of the neuron is computed as the weighted sum of inputs. Second, this activation is transformed into a response by using a transfer function. Hence the output from each unit is based on the weighted sum of all inputs, and is ultimately defined by an activation function. Any function whose domain is within the real numbers can be used as a transfer function (Abdi, 1999), which can be a linear, Gaussian, hyperbolic or sigmoid functions (also called the logistic function).

When ANNs are used for data analysis, it is important to distinguish between ANN models and ANN algorithms. Many ANN models are similar or identical to popular statistical techniques i.e. generalised linear models or polynomial regression, especially where the emphasis is on prediction of complicated matter rather than on explanation. ANNs can be trained more efficiently by standard numerical optimisation algorithms such as those used for nonlinear regression. Various models can be displayed as network diagrams (Figure 4-5), which illustrates ANN and statistical terminology for a simple linear regression model. Neurons are represented by circles and boxes, while the connections between neurons are shown as arrows.

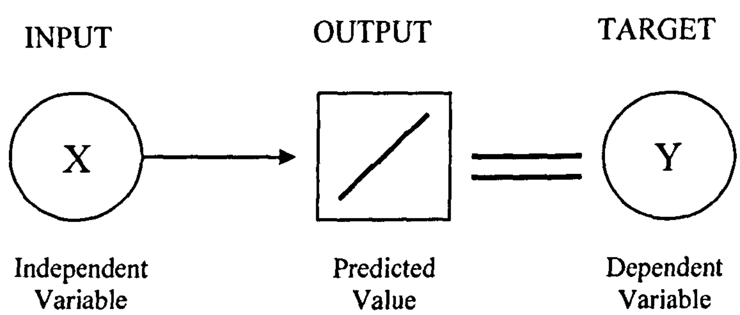


Figure 4-5: Simple Linear Regression (Warren and Cary, 1994).

Circles represent observed variables, with the names shown inside the circle and boxes represent values computed as a function of one or more arguments. The symbol inside the box indicates the type of function e.g. linear. Most boxes also have a corresponding parameter called a bias. Arrows indicate that the source of the arrow is an argument of the function computed at the destination of the arrow. Furthermore, each arrow usually has a corresponding weight or parameter to be estimated. Two long parallel lines indicate that the values at each end are to be fitted by least squares, maximum likelihood or some other estimation criterion.

The linear combination of inputs, called the net input, is computed by a perceptron and then a possibly nonlinear activation function is applied to the net input to produce the output. An activation function maps any real input into a bounded range, usually 0 to 1 or -1 to 1. Some common activation functions are:

- Linear or identity: act(x) = x
- Hyperbolic tangent: act(x) = $\tanh(x)$ = $\dfrac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$
- Logistic (Sigmoid): act(x) = $(1+e^{-x})^{-1}$ = $(\tanh(x/2)+1)/2$
- Threshold: act(x) = 0 if x<0 , 1 otherwise
- Gaussian: act(x) = $e^{-x^2/2}$

The Hyperbolic tangent function (tanh) (Equation 4-3) will give an output in the range [-1, 1].

$$\tanh(\beta x)=\frac{e^{\beta x}-e^{-\beta x}}{e^{\beta x}+e^{-\beta x}} \qquad \textbf{(4-3)}$$

where $\beta \in \Re$. For $\beta = 1$ the behaviour of the hyperbolic tangent function is shown in Figure 4-6. It compresses the combinations of the inputs within the interval [-1, 1], rather than [0, 1] in the sigmoid function.

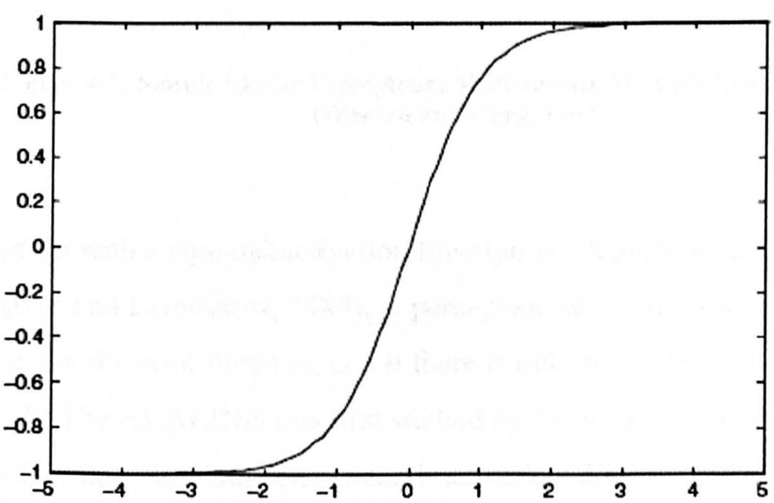


Figure 4-6: Hyperbolic tangent function

A perceptron can have one or more outputs. Each output has a separate bias and set of weights. Usually the same activation function is used for each output, although it is possible to use different activation functions.

Perceptrons are most often trained by least squares i.e. by attempting to minimise $\sum\sum r_j^{\,2}$, where r_j is the error or residual and as noted before it is equal to a predicted value (output value) p_j, subtracted from dependant variable (training values) y_j. A perceptron with a linear activation function is thus a linear regression model possibly multiple or multivariate, as shown in Figure 4-7 (Myers, 1986).

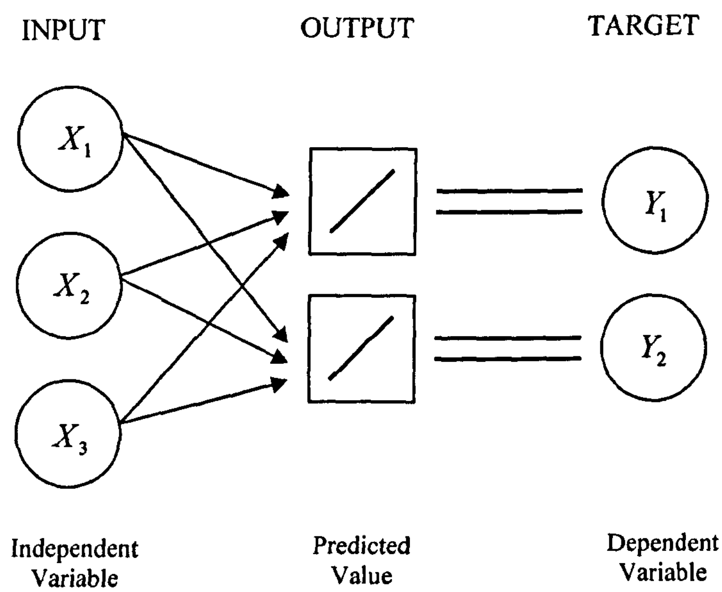


Figure 4-7: Simple Linear Perceptron; Multivariate Multiple Linear Regression (Warren and Cary, 1994).

A perceptron with a sigmoid activation function is a logistic regression model, Figure 4-8 (Hosmer and Lemeshow, 1989). A perceptron with a threshold activation function is a linear discriminant function, and if there is only one output, it is also called an ADALINE. The ADALINE was first studied by Widrow and Hoff in the 1960s and is the basic building block for many neural networks (Widrow and Hoff, 1988).

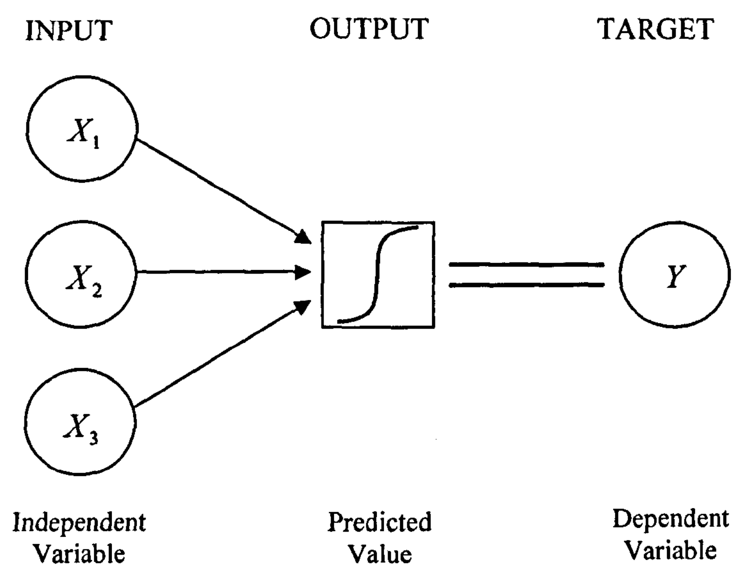


Figure 4-8: Simple Nonlinear Perceptron; Logistic Regression (Warren and Cary, 1994).

Instead of a threshold activation function, it is often more useful to use a multiple logistic function to estimate the conditional probabilities of each class (McLachlan, 1992). The activation function in a perceptron is analogous to the inverse of the link function in a generalised linear model (GLM) (McCullagh and Nelder, 1989). Polynomial regression can be represented by a diagram of the form shown in Figure 4-8, in which the arrows from the inputs to the polynomial terms would usually be given a constant weight of 1. In NN terminology, this is a type of functional link network (Pao, 1989). In general, functional links can be transformations of any type, that do not require extra parameters, and the activation function for the output is the identity, so the model is linear in the parameters.

A functional link network introduces an extra hidden layer of neurons, but there is still only one layer of weights to be estimated. If the model includes estimated weights between the inputs and the hidden layer, and the hidden layer uses nonlinear activation functions such as the logistic function, the model becomes nonlinear, i.e. nonlinear in the parameters. The resulting model is called a Multilayer Perceptron or MLP. The Multilayer Perceptron was first introduced by Minskey and Papert in 1969. A multilayer perceptron network has three distinctive characteristics (Haykin, 1994).

First, the network consists of a set of source nodes that constitute the input layer, one or more layers of hidden neurons, and the output layer. Second, the model of each neuron in the MLP includes a differentiable nonlinearity at the output end. A commonly used form of nonlinearity that satisfies this requirement is a sigmoidal nonlinearity defined by the logistic (sigmoid) function. Third, the network exhibits a high degree of connectivity determined by the weights of the network. A change in the connectivity of the network requires a change in the population of network weights. Perceptrons are arranged in layers with no connections inside a layer. Each layer is fully connected to its preceding and following layers. The first layer is the input layer which just presents the input feature vector (does not perform processing); the last layer is the output layer and its outputs are the output of the network. Other layers are hidden layers. The input layer is usually not counted when the number of layers of a network is specified. Thus, a two-layer perceptron has one hidden and an output layer. A two-layer MLP for simple nonlinear regression is shown in Figure 4-9.

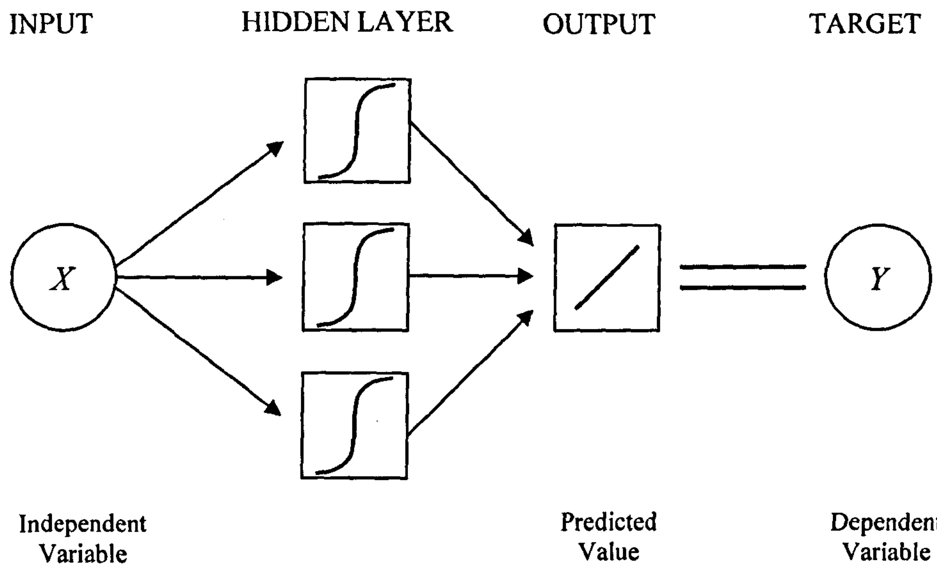


Figure 4-9: Multilayer Perceptron; Simple Nonlinear Regression (Warren and Cary, 1994).

The MLP can also have multiple inputs and outputs, as shown in Figure 4-10.
MLPs are general-purpose, flexible, nonlinear models that, given enough hidden neurons and enough data, can approximate virtually any function to any desired degree of accuracy. MLPs can be used when there is little knowledge about the form of the relationship between the independent and dependent variables (White, 1992).

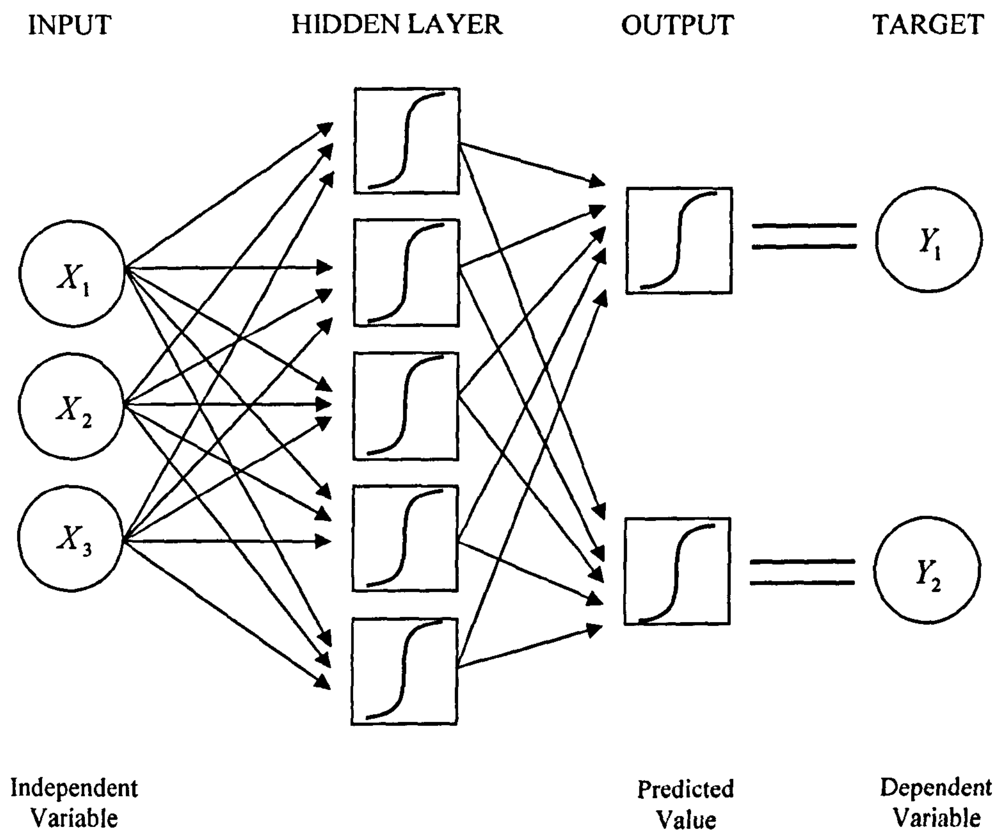


Figure 4-10: Multilayer Perceptron; Multivariate Multiple Nonlinear Regression (Warren and Cary, 1994).

4.4 Network Topologies

The arrangement of neural processing elements (PEs) and their interconnections can have a profound impact on the processing capabilities of the Artificial Neural Networks. All neural networks have some set of PEs that receives inputs from the outside, which was referred to as the input unit before, and many of them also have one or more layers of hidden PEs that receive inputs only from other processing units. A layer of processing elements receives a vector of data or the outputs of a previous layer and processes them in parallel. The set of PEs that represents the final result of the neural network computation is designated as the output units. There are some major connection topologies that define how data flows between the input, hidden, and output processing elements.

4.4.1 FEED-FORWARD NETWORKS

A network is "feed-forward" if all of the hidden and output neurons receive inputs from the preceding layer only. Actually, feed-forward is a definition of connection topology and data flow. The input is presented to the input layer and it is propagated forwards through the network. Output never forms a part of its own input (Figure 4-11).

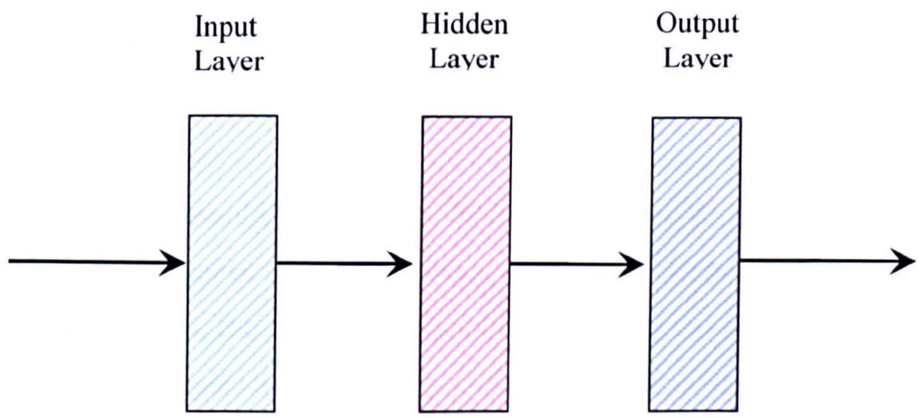


Figure 4-11: Data flow in a Feed-forward Artificial Neural Networks.

Each PE combines all of the input signals coming into the unit along with a threshold value. This total input signal is then passed through an activation function to determine the actual output of the processing element, which in turn becomes the input to another layer of units in a multilayer network. Therefore, feed-forward networks have one-way connections from input to output layers. The most typical activation function used in neural networks is the sigmoid function (Figure 4-12). This function converts an input value to an output ranging from 0 to 1.

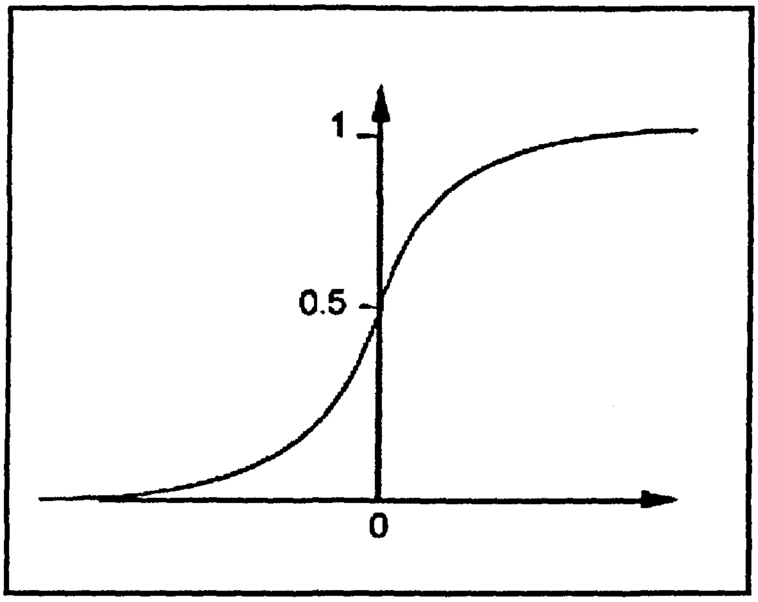


Figure 4-12: The sigmoid activation function which converts an input to an output ranging between 0 and 1.

The effect of the threshold weights is to shift the curve right or left, thereby making the output value higher or lower, depending on the sign of the threshold weight. As shown in Figure 4-11, the data flows from the input layer through zero, one, or more succeeding hidden layers and then to the output layer. In most networks, the elements from one layer are fully connected to the elements in the next layer. However, this is not a requirement of feed-forward neural networks. In some cases, especially when the neural network connections and weights are constructed from a particular rule, there could be less connection weights than in a fully connected network. There are also techniques for pruning unnecessary weights from a neural network after it is trained. Feed-forward networks are commonly used for prediction, pattern recognition, and nonlinear function fitting.

The mathematical representation of the feed-forward network with the *tanh* activation function is given by the following system (McNelis, 2005):

$$n_{k,t} = w_{k,0} + \sum_i w_{k,i} x_{i,t} \qquad \text{(4-4)}$$

$$N_{k,t} = T(n_{k,t}) = \frac{e^{n_{k,t}} - e^{-n_{k,t}}}{e^{n_{k,t}} + e^{-n_{k,t}}} \quad (4\text{-}5)$$

$$y_t = \gamma_0 + \sum_k \gamma_k N_{k,t} \quad (4\text{-}6)$$

where $T(n_{k,t})$ is the mentioned *tanh* activation function for the input neuron $n_{k,t}$.

The mean of the hyperbolic tangent function is centred around zero. Therefore, in order to perform efficient prediction, the range of the input data, their mean and variance, should match the range of the chosen activation function. There are several operations that could be performed on the input data, such as normalisation, to achieve this.

4.4.2 Partial/Fully Recurrent Networks

In contrast, there is another type of network called "feed-back" networks. Feed-back networks can have signals travelling in both directions by introducing loops in the network. Feed-back networks are very powerful and can be extremely complicated. They are dynamic which means their state is changing continuously until they reach an equilibrium point. They remain at the equilibrium point until the input changes and a new equilibrium needs to be found. Feed-back architectures are also referred to as interactive or recurrent (Bigus, 1996).

Recurrent networks are used in situations when current information exists to give to the network, but the sequence of inputs is important, and the neural networks need to store a record of the prior inputs and factor them in with the current data to produce an answer. In recurrent networks, information about past inputs is fed back into and mixed with the inputs through recurrent or feed-back connections for hidden or output units. In this way, the neural network contains a memory of the past inputs via the activations (Figure 4-13).

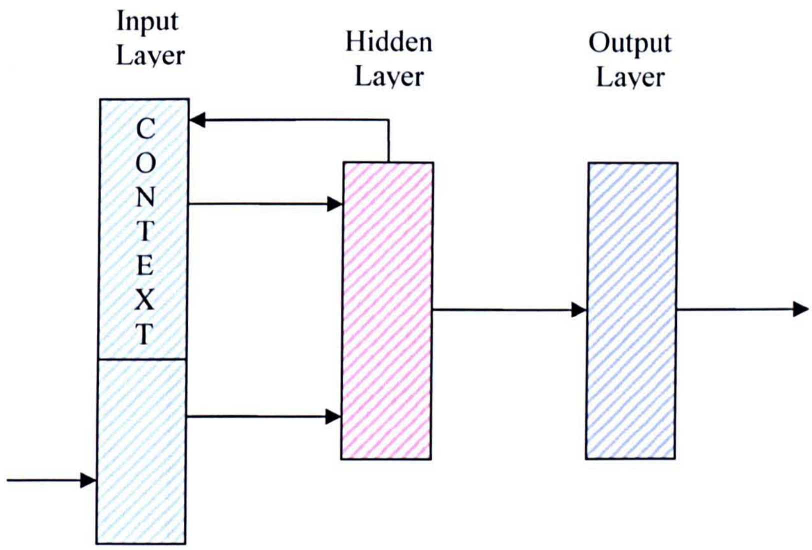


Figure 4-13: Data flow in a Partial recurrent neural networks.

A network can be limited or fully recurrent; two major architectures for limited recurrent networks are widely used. Elman suggested a recurrent network which allowed feed-back from the hidden units to a set of additional inputs called "context units". The connections are mainly feed-forward but also include a set of carefully chosen feed-back connections that let the network remember cues from the recent past. The input layer is divided into two parts: the true input units and the context units that hold a copy of the activations of the hidden units from the previous time step. The network is able to recognise sequences and also to produce short continuations of known sequences (Elman, 1990).

Earlier, Jordan described a network with feed-back from the output units back to a set of context units. This form of recurrence is a compromise between the simplicity of a feed-forward network and the complexity of a fully recurrent neural network because it still allows the popular back propagation training algorithm (described in the following sections) to be used (Jordan, 1986).

Fully recurrent networks provide two-way connections between all processors in the neural networks. A subset of the units is designated as the input processors, and they are assigned or clamped to the specified input values. The data then flows to all adjacent connected units and circulates back and forth until the activation of the units stabilises. Figure 4-14 shows the input units feeding into the hidden units and the output units. The activations of the hidden and output units then are recomputed until the neural networks stabilise. At this point, the output values can be read from the output layer of processing units. Fully recurrent networks are used primarily for optimisation problems and as associative memories (Bigus, 1996).

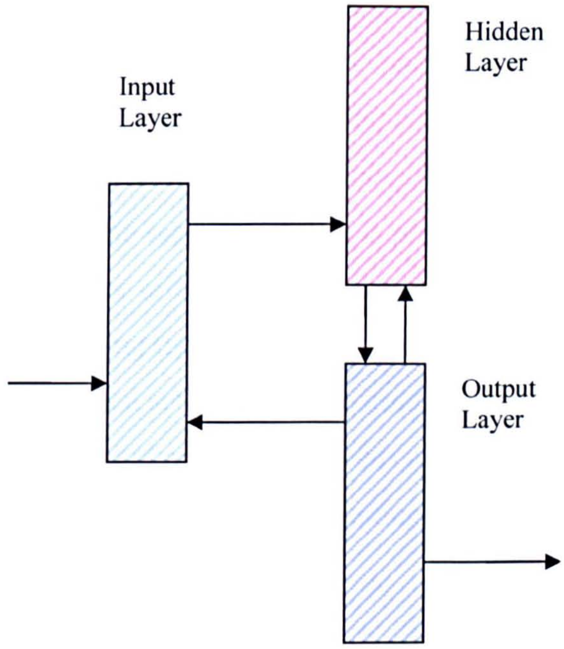


Figure 4-14: Data flow in a fully recurrent neural networks.

4.5 Network Learning

When a network has been structured for a particular application, it is ready to be trained. To start this process the initial weights are chosen randomly. Then, the training, or learning, begins. There are three approaches to training:

- Supervised
- Unsupervised
- Reinforcement

Supervised training involves a mechanism of providing the network with the desired output either by manually grading the performance of the network or by providing the desired outputs with the inputs (Figure 4-15). In supervised training, both the inputs and the outputs are provided. The network then processes the inputs and compares its resulting outputs against the desired outputs. Errors are then propagated back through the system, causing the system to adjust the weights which control the network. This process occurs over and over as the weights are continually adjusted (Demuth and Beale, 2004). Training of a network proceeds by making weight and bias changes based on an entire set (batch) of input vectors. Incremental training changes the weights and biases of a network as needed after presentation of each individual input vector. The set of data which enables the training is called the "training set". During the training of a network the same set of data is processed many times as the connection weights are ever refined. To monitor the network to determine if the system is simply memorising its data in some non-significant way, supervised training needs to hold back a set of data to be used to test the system after it has undergone its training. Neural networks have been trained to perform complex functions in various fields of application including pattern recognition and identification, classification, etc.

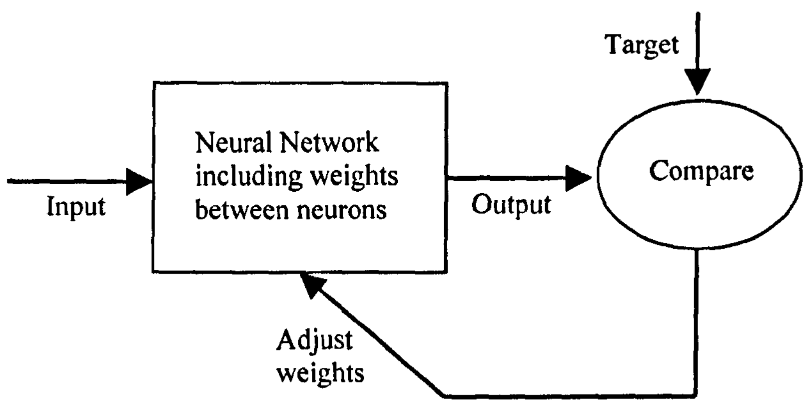


Figure 4-15: Supervised training of a Neural Network (Demuth and Beale, 2004).

In supervised learning, the learning rule is provided with the training set of proper network behaviour:

$$\{p_1,t_1\},\{p_2,t_2\},\ldots,\{p_q,t_q\}$$

Where p is an input to the network, and t is the corresponding correct output (target). As the inputs are applied to the network, the network outputs are compared to the targets. The learning rule is then used to adjust the weights and biases of the network

in order to move the network outputs closer to the targets. The perceptron learning rule falls in this supervised learning category (Demuth and Beale, 2004).
Supervised learning assumes the availability of a labelled set of training data made up of N input-output examples:

$$T = \{x_i, d_i\}_{i=1}^{N} \qquad \textbf{(4-7)}$$

Where x_i is the input vector of the ith example, d_i desired (target) response of ith example, assumed to be scalar for convenience of presentation and N is the sample size.
Given the training sample T, the requirement is to compute the free parameters of the neural networks so that the actual output y_i of the neural network due to x_i is close enough to d_i for all i in a statistical sense. For example, we may use the Mean Square Error (MSE) (Haykin, 1999):

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(d_i - y_i)^2 \qquad \textbf{(4-8)}$$

as the index of performance to be minimised.

Unsupervised training is where the network has to make sense of the inputs without outside help, a learning process in which changes in weights of a network and biases are not due to the intervention of any external teacher. Commonly changes are a function of the current network input vectors, output vectors, and previous weights and biases. Actually, unsupervised training is used to perform some initial characterisation on inputs. In unsupervised training, the network is provided with inputs but not with desired outputs, and the weights and biases are modified in response to network inputs only. There are no target outputs available. Most of these algorithms perform clustering operations. They categorise the input patterns into a finite number of classes. The system itself must then decide what features it will use to group the input data. This is often referred to as "self-organisation" or "adaption". A famous neural network model based on unsupervised training is the Kohonen network, named after its inventor, Teuvo Kohonen from University of Helsinki (first presented in 1982). The Kohonen neural network contains only an input and output layer of neurons (Figure 4-16).

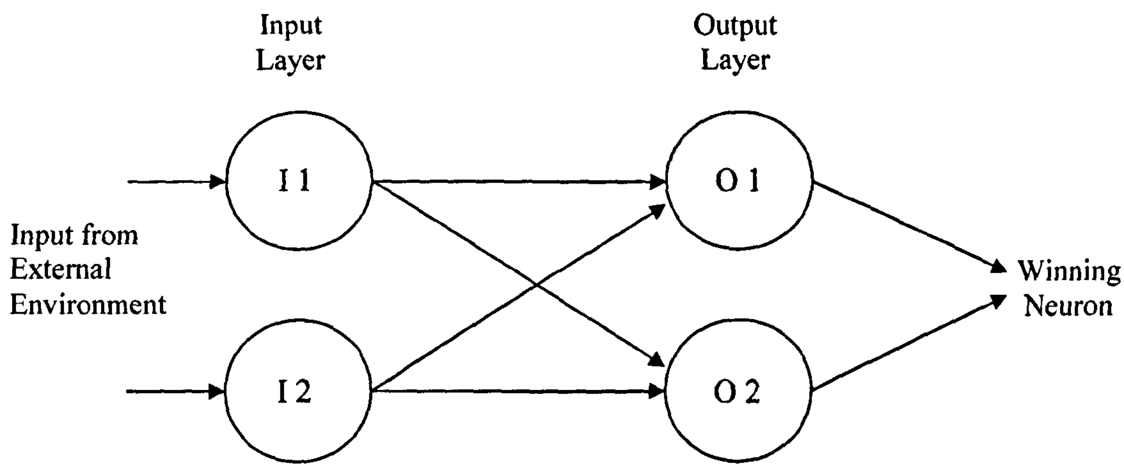


Figure 4-16: The Kohonen neural network (Warren and Cary, 1994).

There is no hidden layer in the network. Output from the network does not consist of the output of several neurons. When a pattern is presented to a Kohonen network one of the output neurons is selected as a "winner". This "winning neuron" is the output from the network (Heaton, 2003). Unlike supervised networks, the inputs and outputs are not presented at the same time to the network. The Kohonen network relies on a type of learning called "competitive learning", where neurons compete for the privilege of learning, and the correct output is not known.

Another type of unsupervised learning is found in the "cognitron", introduced by Fukushima as early as 1975. This network, with primary applications in pattern recognition, was improved at a later stage to incorporate scale, rotation, and translation invariance resulting in the neocognitron (Fukushima, 1988).

The third Neural Network training approach is called reinforcement learning, or sometimes called reward-penalty learning, which is a combination of the above two learning approaches. It is based on presenting input vector x to a neural network and looking at the output vector calculated by the network. If it is considered "good", then a "reward" is given to the network in the sense that the existing connection weights are increased; otherwise the network is "punished", the connection weights, being considered as "not appropriately set", decrease (Kasabov, 1998). Actually, the reinforcement learning describes a form of "semi-supervised" learning where the network is not provided with an explicit form of error at each time step but rather receives only generalised reinforcement which yields little immediate indication of how any neuron should change its behaviour (Arbib, 2003).

4.5.1 Learning Rate

All training methods discussed above result in an adjustment of the weights of the connections between units, according to some modification rules. Virtually all learning rules for models of this type can be considered as a variant of the Hebbian learning rule suggested by Hebb (Hebb, 1949). The basic idea is that if two units, *j* and *k*, are active simultaneously, their interconnection must be strengthened. If *j* receives input from *k*, the simplest version of Hebbian learning prescribes modification of the weight W_{jk} by the synaptic weight change Δw_{jk} (Kros and Van Der Smagt, 1996). The synaptic weight change Δw_{jk} is then calculated as a function of the product of the two activation values y_j and y_k of the neurons *j* and *k*:

$$\Delta w_{jk} = \gamma y_j y_k \qquad (4\text{-}9)$$

Where γ is a positive constant of proportionality representing the learning rate. Learning rate is a training parameter that controls the size of weight and bias changes during the ANN training process and adaptive learning rate is a learning rate that can be adjusted according to an algorithm during the training to minimise training time (Rabuñal and Dorado, 2006). The general process of learning in neural networks is described by a characteristic called "convergence". The network reacts better and better to the same training example *x*, the more it is introduced to it through training, eventually ending up with the desired output *y*. The learning rate of the network is adjusted during training process and can improve the convergence of the network and increase the convergence speed.

In the first step, the training coefficient is selected as a large number, so the resulting error values are large. However, the error will be decreased as the training progresses, due to the decrease in the learning rate (Zilouchian and Jamshidi, 2001). Also, while low values of the learning parameters avoid oscillations, they may unnecessarily prolong the convergence process (McNelis, 2005).

After the network has stopped learning the training examples, the synaptic weights do not change any more, that is, $\Delta w_{jk} = 0$ for every connection *(j, k)* in the network when training examples from the training set are further presented (Kasabov, 1998).

Another common rule uses not the actual activation of unit *k* but the difference between the actual and desired activation for adjusting the weights (Kros and Van Der Smagt, 1996):

$$\Delta w_{jk} = \gamma y_j (d_k - y_k) \qquad \textbf{(4-10)}$$

In which d_k is the desired activation provided by a teacher. This is often called the Widrow-Hoff rule or the delta rule.

A network can stop learning for two reasons (Kasabov, 1998):

1. The network has learned the training examples, and
2. The network has become saturated.

If the network can learn the training examples it would be a well trained network and can be used to generate favourable outputs. In contrast, saturation is a serious condition of a network which must be prevented. According to the Grasberg's saturation theorem, large input signals saturate a neuron when it is sensitive to small input signals, but if it is sensitive to large input signals, small signals are ignored as noise. Small random values introduced as noise during a learning process tend to increase the robustness of the performance of the neural network. In this case the Hebbian learning law will take the form of:

$$\Delta w_{jk} = \gamma y_j y_k + \varepsilon \qquad \textbf{(4-11)}$$

where ε is a noise signal. When noise is presented during learning, the neural network reaches a convergence when weights change within the magnitude of the noise (stochastic equilibrium):

$$\Delta w_{jk} \leq \varepsilon$$

The state of convergence may also be reached in a so-called oscillatory mode, that is, the synaptic weights oscillate between two or more states.

4.5.2 Sensitivity Analysis

Sensitivity Analysis is a technique for deriving how and how much the solution to a given problem depends on data (Castillo et al., 2006). In neural networks, sensitivity analysis is a tool to explore the embedded knowledge of an ANN model and evaluate the effectiveness of the network learning. The concept of the sensitivity analysis comes from a calculus called the chain rule. The chain rule explains how to compute the partial derivative of a variable with respect to another when a functional form links the two. If $y = f(x)$ and the goal is to compute $\partial y/\partial x$, the sensitivity of y with respect to x, as long as f is differentiable, is (Principe *et al.* 2000):

$$\frac{\partial y}{\partial x} = \frac{\partial y}{\partial f}\frac{\partial f}{\partial x} \qquad \textbf{(4-12)}$$

Equation 4-12 computes how much a change in x is reflected in y i.e., how sensitive y is to change in x. The above sensitivity for a single neuron of an ANN, with multiple weights connected to its input, is illustrated in Figure 4-17.

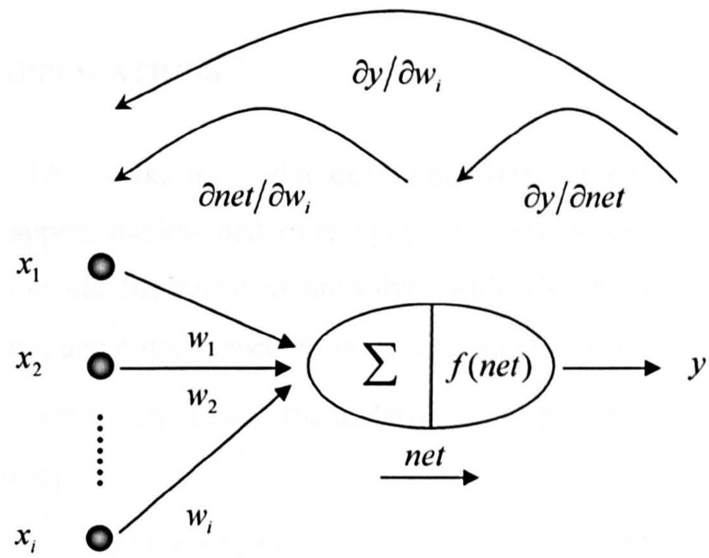


Figure 4-17: Illustration of the sensitivity computation through a non-linear neuron (or PE) of a neural networks (Principe *et al.* 2000).

In order to do the sensitivity analysis for a neural network, after training the network and fixed the weights w_i, each input vector should change around its mean value. The change in the input is normally done by adding a random value of a known variance to each sample and computing the output. The sensitivity for input k is expressed as (Principe *et al.* 2000):

$$S_k = \frac{\sum_{p=1}^{P}\sum_{i=1}^{o}(y_{ip} - \bar{y}_{ip})^2}{\sigma_k^{\ 2}} \quad \textbf{(4-13)}$$

where $\bar{y}_{ip}$ is the ith output obtained with the fixed weight for the pth pattern, o is the number of network outputs, P is the number of patterns and $\sigma_k^{\ 2}$ is the variance of the input changes. This is a common way to measure how much a change in a given input affects the output across the training data set.

Inputs with large sensitivities have more influence in the mapping of the network and the inputs with small sensitivities have less (or no) importance for the model and can be discarded.

4.6 ANN APPLICATIONS

Artificial Neural Networks are used to deal with different kinds of problems such as classification, approximation, pattern recognition, signal processing, prediction, feature extraction, etc. Most of them are solved with ANN by learning of the mapping between the input and output space for given data sets $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$, where (x_i, y_i) is input-output pair. The underlying mapping can be written as (Jankowski, 1999):

$$f(x_i) = y_i + \varepsilon \quad \text{for } i=1,...,n \qquad (4\text{-}14)$$

where ε is a zero mean white noise with variance σ_{ns}^2.

Neural network problems can generally be categorised as one of four types:

- *Classification*;

  procedure in which individual items are placed into groups based on quantitative information on one or more characteristics inherent in the items, including: Target Recognition, Pattern Recognition, Character Recognition.
- *Function Approximation*;

  is using to select a function that closely approximates a target function in a specific way, including: Process Modelling, Process Control, Data Modelling, Machine Diagnostics.
- *Time Series Prediction*;

  including: Dynamic Modelling System, Financial Forecasting.
- *Data Mining*;

  involves sorting through large amounts of data and picking out relevant information including: Clustering, Data visualisation, Data extraction.

4.6.1 CLASSIFICATION

Classification problems are those where the goal is to label each input pattern as belonging to a certain class. In statistics and machine learning, classification is a type of statistical algorithm, which takes a feature representation of an object or concept and maps it to a classification label. Typically, a classification algorithm computes a posterior probability: the probability of a class label, given that the feature input was observed (Fausett, 1994). Traditional statistical classification methods usually try to find a clear cut boundary to divide the pattern space into some classification regions based on some pre-defined criterion, such as maximising deviation between groups divided by deviation within groups in the linear discriminant analysis (Lachenbruch, 1975). However, it is impossible to provide information of degree of uncertainty for a particular example for this method since the error rate estimate is a statistical result of the entire sample set (Wang and Archer, 1991). In fact, classification systems are systems that automatically identify objects based on their measured properties. For classification problems, the input attributes are mapped to the desired classification categories. The training of the neural network amounts to setting up the correct set of discriminate functions to correctly classify the inputs. With this viewpoint, artificial neural networks also can be a classification system. The existing neural networks that can be served as classifiers are grouped into several categories or their variations e.g. Adaptive Resonance Theory (ART), Radial Basis Functions (RBF) and Probabilistic Neural Networks (PNN).

A simple example of a classification problem is one where the goal of the neural networks is to label each person as male or female (the two classes) based on their height and weight. The input into the neural network would be the height and weight measurements and the desired output would be their gender. For classification of a large number of objects the unsupervised strategy seems to be more efficient than supervised one (Kohonen, 1997). Unsupervised neural networks are trained by letting the network continually adjust itself to new inputs. They find relationships within data as it is presented and can automatically define classification schemes. There is a type of unsupervised neural networks employed for classification problems called Competitive layer recognition. Competitive layers recognise similar groups of input vectors. By using these groups, the network automatically sorts the inputs into categories. It means, after training, the final weights of the "winning neuron" were proportional to the probabilities of the corresponding classes to which the unknown object should belong (Kohonen, 1997).

4.6.2 FUNCTION APPROXIMATION

Function approximation is a general class of problem solving where a function is created that approximates an unknown function. Function approximation methods fall into two broad categories: global and local. Global approximations can be made with many different function representations, e.g. polynomials, rational approximation, and MLP (Farmer and Sidorowich, 1988). To approximate a function *f*, a model must be able to represent its many possible variations. The dependence on representation can be reduced using local approximation where the domain of *f* is broken into local neighbourhoods and a separate model is used for each neighbourhood (Farmer and Sidorowich, 1988).

Artificial neural networks can be used as a function approximation system which tries to produce the desired output for each training input. The task performed by a trained network to respond to inputs with an approximation of a desired function. The ANN then creates a map through input data sets to desired data. This map can be a function between data groups. In the test phase, this map produces function approximation using adjusted weight coefficients and transfer functions (Quing *et. al.*, 1997).
As with other transfer functions the sigmoid function provides linear, near-linear, and non-linear approximations for a given set of inputs (Berry and Linoff, 1997). In the field of supervised learning, the most popular form of the feed-forward neural networks, the multi-layer perceptrons (MLP) have been proven to approximate smooth functions very well (Barron, Yang, and Yu, 1994). Many application problems use the MLPs as a model for identifying and controlling nonlinear complex dynamic systems (Fausett, 1994).

MLPs are suitable for high-dimensional function approximation. MLP network can be used for a function approximation problem in which the inputs to the network are equivalent to the predictor variables in the regression model and the output of the network is equivalent to the predicted value. For a given problem, there is a cost function ε_T (Haykin, 1994), which is similar to the error sum of squares:

$$\varepsilon_T = \sum_{i=1}^{n} (y_i - (\beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik}))^2 \qquad \textbf{(4-15)}$$

for the regression model, as the measure of training set learning performance. The objective of the learning process is to adjust the weights of the neural networks to minimise the least squares output error cost function ε_T.

A popular training algorithm known as the back-propagation algorithm is generally used to adjust the network weights until a stop criterion is reached (Hush and Horne, 1993) e.g. a reasonable error rate. Back-propagation can train multilayer feed-forward neural networks with differentiable transfer functions to perform function approximation. The term back-propagation refers to the process by which derivatives of network error, with respect to network weights and biases, can be computed. This process can be used with a number of different optimisation strategies which will be explained in section 4-7-1.

4.6.3 Time Series Prediction

A time series is a sequence of vectors, $x(t)$, for $t = 0,1,\ldots$, where t represents elapsed time. Theoretically, x may be a value which varies continuously with t. In practice, for any given physical system, x will be sampled to give a series of discrete data points, equally spaced in time. The rate at which samples are taken dictates the maximum resolution of the model; however, it is not always the case that the model with the highest resolution has the best predictive power. Hence, those superior results may be obtained by employing only every *n*th point in the series. Statistical methods and neural networks are commonly used for time series prediction. Artificial neural networks are reliable for modelling non-linear and dynamic signals (Guido, 1994). An ANN attempts to capture the dynamics of the system which underlies the data series by training to take as input a representation of the current state of the system and to output a prediction of the state of the system at some point in the future. An ANN has concentrated on forecasting future developments of the time series from values of x up to the current time. Formally this can be stated as: find a function f to obtain an estimate of x at time $t + d$, from the N time steps back from time t, so that (Frank, Davey and Hunt, 1997):

$$x(t+d) = f(x(t), x(t-1), \ldots, x(t-N+1)) \qquad \textbf{(4-16)}$$

For if d is equal to 1, f will forecast the next value of x.
The standard neural networks method of performing time series prediction is to induce the function f using any feed-forward function approximating neural network architecture. Examples are standard MLP, RBF architecture or a cascade correlation mode (Gershenfeld and Weigend, 1993), using a set of inputs and a single output as the target value of the network. This method is often called the sliding window technique as the input slides over the full training set. Figure 4-18 gives the basic architecture.

This technique can be seen as an extension of auto-regressive time series modelling, in which the function f is assumed to be a linear combination of a fixed number of previous series values. Such a restriction does not apply with the non-linear neural network approach as such networks are general function approximators (Dorffner, 1996).

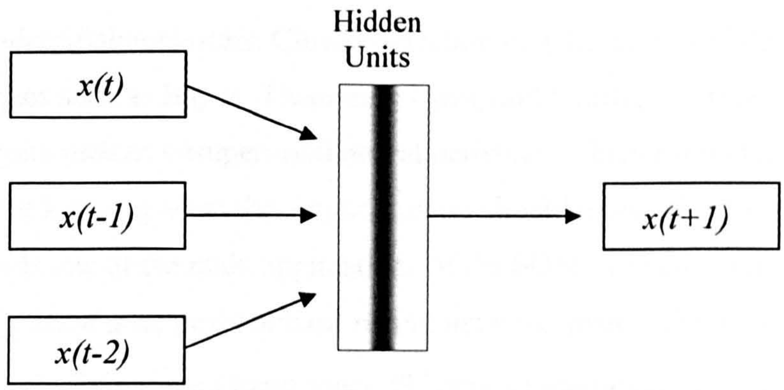


Figure 4-18: The standard method of performing time series prediction (Frank, Davey and Hunt, 1997).

Another type of ANN-based approach to time series forecasting, is to use recurrent networks. Elman in 1990 showed how an otherwise feed-forward network with a recurrent context layer which took a copy of the network's hidden layer at time *t-1* and re-applied it in addition to the input vector at time *t* was able to learn temporal dependencies. Also extra context layers could be added during training in order to allow a recurrent network to be trained on several examples of a time series from the same source system (Swingler, 1994).

A recurrent network with one input unit representing the value of the time series at time *t*, one output unit representing the value of the time series at time *t+1* and a recurrent layer to store and re-apply the state of the hidden layer from time *t* can forecast one step ahead along a time series. By taking the network output and feeding it back in as input, this method can be extended to multiple steps forward.

4.6.4 Data Mining

Data mining is the process of sifting through and analysing rich sets of domain specific data and then extracting the information and knowledge in the form of new relationships, patterns or clusters for decision-making purposes (Watterson, 1995). Data mining is a form of knowledge discovery essential for solving problems in a specific domain (Dankel and Gonzalez, 1993). Many data mining applications make use of clustering, i.e. the process of finding similarities in the data and then grouping similar data into identifiable clusters. Cluster detection may be employed through statistical techniques such as Bayes' Theorem[1] (Lwin, and Maritz, 1989) or non-statistical techniques such as unsupervised neural networks, which form clusters on the data set without knowing what the output clusters should model (Kohonen, 1990). Clustering of data is one of the main applications of the SOM (Vesanto and Alhoniemi, 2000). SOM is an unsupervised neural network model which provides a mapping from a high-dimensional input space $\Re^n$ onto a two-dimensional map space (Figure 4-19). This capability allows an intuitive analysis and exploration of unknown data spaces, with its applications ranging from the analysis of data (Kohonen, 1995).

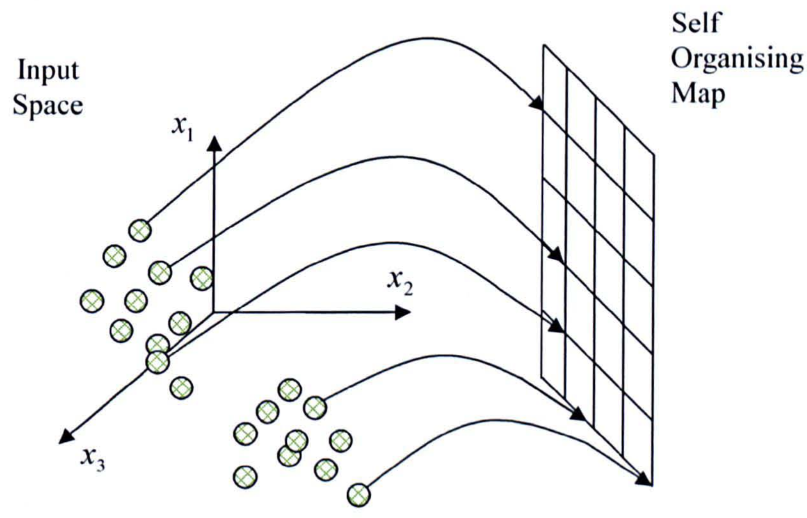


Figure 4-19: Input space and Self Organizing Map (SOM).

The SOM consist of two layers of neurons. The input layer consists of *N* neurons corresponding to the real-valued input vector of dimension *N*. These units are connected to a second layer of neurons *U*. By means of lateral connections, the neurons in *U* form a lattice structure (output neurons arranged in rows and columns) of dimensionality *M*. Typically *M* is much smaller than *N*. When an input data vector is presented to the network, it responds with the location of a unit in *U*, which

[1] Bayes' theorem (also known as Bayes' rule or Bayes' law) is a result in probability theory, which relates the conditional and marginal probability distributions of random variables. In some interpretations of probability Bayes' theorem tells how to update or revise beliefs in light of new evidence.

corresponds most closely to the presented input. This is called the best-matching neuron. As a result of the learning process, i.e. the presentation of all input vectors and the adaptation of the weight vectors, the SOM generates a mapping from the input space $\Re^n$ onto the lattice U with the property, that the topological relationships in input space are preserved in U as well as possible (Ritter and Schulten, 1986; Kohonen, 1989). Similar input data should correspond to best-matches in U that are close together on the lattice. The structure of the trained SOM is then visualised by means of the U Matrix for the purpose of clustering.

4.7 ANN Models

Since the 1940s different network types have been initiated and introduced by scientists and consequently, various network architectures were created over time. In 1990 Kohonen proposed a classification for different neural networks (Kohonen, 1990). According to this classification, network architectures can be categorised to three main types:

- Feed-forward networks,
- Recurrent networks (feed-back networks),
- Self-organising networks.

Later, in 1994, Haykin divided the neural networks into four classes (Haykin, 1994):

- Single-layer feed-forward networks,
- Multilayer feed-forward networks,
- Recurrent networks, and
- Lattice structures.

A lattice network is a feed-forward network, which has output neurons arranged in rows and columns.

Nowadays, the varieties of the networks are dramatically increased compared with the last decade. So, various classifications of the neural networks may be found in different literatures. Without considering the classification of the networks, different models like ART, RBF, PNN and Recurrent Back-propagation Neural Networks are the examples of the recently developed neural networks. In the following sections some of related neural networks are introduced and examined.

4.7.1 BACK-PROPAGATION NETWORKS

A back-propagation network uses a feed-forward topology, supervised learning and the back-propagation learning algorithm which is a general purpose learning algorithm. In 1986, Rumelhart *et. al.* popularised the back-propagation algorithm which they called the generalised delta rule. This algorithm is a learning algorithm applied on multilayer feed-forward networks consisting of non-linear units. The learning strategy in this algorithm is to modify the weights of connections between units towards decreasing the total sum of prediction errors.

Theoretically, the back-propagation algorithm performs gradient descent on the total error only if the weights are adjusted after the full set of learning patterns has been presented; more often than not the learning rule is applied to each pattern separately. The training of an MLP is usually accomplished by using a back-propagation algorithm that involves two phases (Werbos 1974; Rumelhart *et. al.* 1986):

- *Forward Phase.* During this phase the free parameters of the network are fixed, and the input signal is propagated through the network layer by layer. The forward phase finishes with the computation of an error signal e_k in the Equation 4-10 as:

$$e_k = d_k - y_k \tag{4-17}$$

  where d_k is the desired response and y_k is the actual output produced by the network in response to the input x_i.

- *Backward Phase.* During this second phase, the error signal e_k is propagated through the network in the backward direction, hence the name of the algorithm. It is during this phase that adjustments are applied to the free parameters of the network so as to minimise the error e_k in a statistical sense.

The data set is split into a test and a training set. The training set is used to train the network or to get the optimum weights and later is tested using the test set. Most ANN studies in the literature use convenient ratio of splitting for training and testing samples such as 70%:30%, 80%:20%, or 90%:10%. It is important to note that in data splitting, the issue is not about what proportion of data should be allocated in each sample but it is about sufficient data points in each sample to ensure adequate learning, validation, and testing. Granger (1993) suggests that for non-linear modelling at least 20% of the data should be held back for an out-of-sample evaluation. Hoptroff (1993) recommends that at least 10 data points should be in the test sample while Ashley (2003) suggests that a much larger out-of-sample size is necessary in order to achieve statistically significant improvement for forecasting problems.

Training starts with random initial values as weights and biases. The network is activated using activation function and the output is obtained. The output is compared with the desired values. The difference between the actual and desired output values is measured and the initial weights are changed by back-propagating the error. The process is continued until the difference reaches the global minimal. Gradient descent can be used to adjust the gain of the node and increases its learning speed and subsequently the overall learning speed of the network (Kruschke and Movellan 1991).

There are a few important issues which must be identified for training with the back-propagation algorithm. For example, it is important to determine the best time to stop training the network and select the size of individual hidden layers of the MLP. The answers to these important issues can be found through the use of a statistical technique known as cross-validation, which proceeds as follows (Haykin 1999):

- The set of training examples is split into two parts:
- The estimation subset is used for training of the model

- The validation subset is used for evaluating the model performance
- The network is finally tuned by using the entire set of training examples and then tested on test data not seen before.

To find the weight matrices of a single neuron MLP (Figure 4-20), with two inputs x_1 and x_2, the output y is to be compared with the desired target value d and their difference, the error e (Equation 4-17), will be computed.

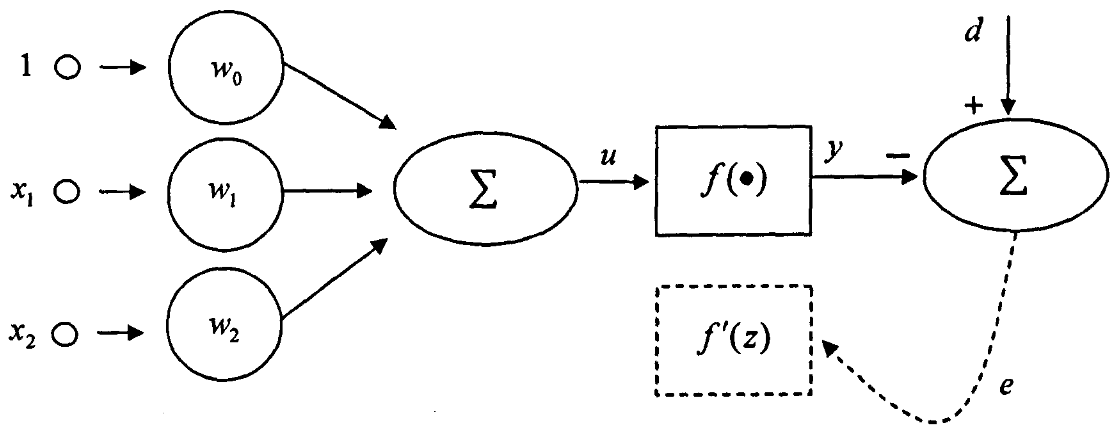


Figure 4-20: Back-propagation training in a single neuron MLP.

There are two inputs $[x_1, x_2]$ with corresponding weights w_1 and w_2. The input labelled with a constant 1 represents the bias term θ shown in Figure 4-4. Here, the bias term is labelled w_0. The net function is computed as (Hu and Hwang, 2003):

$$u = \sum_{i=0}^{2} w_i x_i = Wx \qquad \textbf{(4-18)}$$

where $x_0 = 1$, $W = [w_0 \quad w_1 \quad w_2]$ is the weight matrix, and $x = [1 \quad x_1 \quad x_2]'$ is the input vector.

Given a set of training samples $\{(x_{(k)}, d_{(k)}); 1 \le k \le K\}$, the error back-propagation training begins by feeding all K inputs through the MLP network and computing the corresponding output $\{y_{(k)}; 1 \le k \le K\}$.

Here, an initial guess for the weight matrix W is used. Then a sum of squares (Hu and Hwang, 2003):

$$E = \sum_{k=1}^{K}[e_k]^2 = \sum_{k=1}^{K}[d_k - y_k]^2 = \sum_{k=1}^{K}[d_k - f(Wx_k)]^2 \qquad \textbf{(4-19)}$$

The objective is to adjust the weight matrix W to minimise the error E. This leads to a non-linear least square optimisation problem. There are numerous non-linear optimisation algorithms available to solve this problem. Basically, these algorithms adopt a similar iterative formulation:

$$w(t) = w(t-1) + \Delta w(t) \qquad \textbf{(4-20)}$$

where $\Delta w(t)$ is the correction made to the current weights $w(t)$. Different algorithms, e.g. conjugate-gradient method, Newton's method or steepest descent gradient method, differ in the form of $\Delta w(t)$.

Despite the practical success, the error back-propagation learning algorithm has training problems and suffers from slow convergence (Specht, 1991). Moreover, it is known for the possibility of getting stuck in a local minimum on the error surface. There are a number of techniques designed to overcome this problem. The most useful technique is to use momentum. Momentum is a technique often used to make it less likely for a back-propagation ANN to get caught in a shallow minimum. In fact, one effective solution for speeding up the process of back-propagation toward convergence is to add a momentum term to the above process. Mathematically, this happens by inserting a momentum term α to the weights update Equation 4-20 w(t) to obtain the new form: (Medsker and Jain, 2001):

$$\Delta w'(t) = \alpha \Delta w(t-1) + (1-\alpha)\Delta w(t) \qquad \textbf{(4-21)}$$

The momentum term keeps the movement over the weight error space for some time, even if the network has fallen into a local minimum.

4.7.2 Radial Basis Function networks

Radial Basis Function (RBF) networks are feed-forward networks trained using a supervised learning algorithm. They are typically configured with a single hidden layer of neurons (whereas MLPs can have any number of hidden layers) whose activation function is selected from a class of functions called "basis functions" (Figure 4-21).

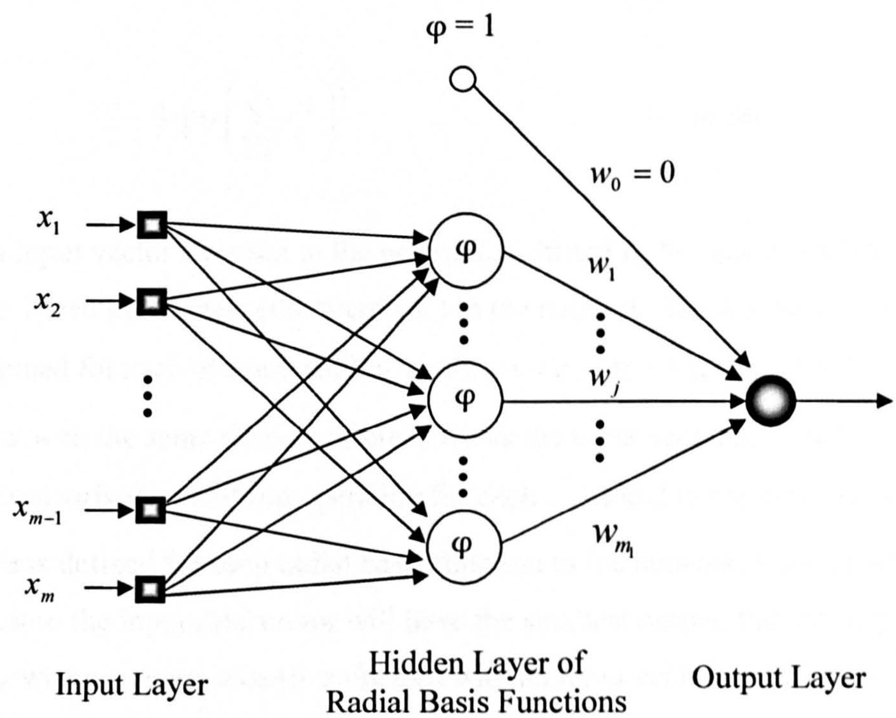


Figure 4-21: Radial-basis function network.

The first layer of this network consists of *m* inputs. They are fully connected to the neurons in the second layer. A hidden node has a radial basis function as an activation function. The RBF network generally consists of two weight layers, the hidden layer and the output layer and they can be described by the following equation (Hu and Hwang, 2003):

$$y = w_0 + \sum_{i=1}^{m_h} w_i f(\|x - c_i\|) \qquad \textbf{(4-22)}$$

where *f* is the radial basis functions, w_i is the output layer weight, $w(\cdot)$ is the output offset, *x* is the input to the network, c_i is the centre associated with the basis functions, m_h is the number of the basis functions in the network, and $\|\cdot\|$ denotes the Euclidean norm ($\sqrt{x.x}$). Given the vector

$$x = [x_1, x_2, ..., x_n]' \qquad \textbf{(4-23)}$$

on $\Re^n$, the Euclidean norm on this space measures the size of the vector in a general sense and is defined as:

$$\|x\| = \left(\sum_{i=1}^{n} x_i^{\,2}\right)^{\frac{1}{2}} \qquad \textbf{(4-24)}$$

Each input vector x passed to the network is shifted in $\Re^n$ space, according to some stored parameters (the "centres") in the network. The Euclidean norm is computed for each of these shifted vectors $x - c_j$ for $j = 1,...,n_h$. Each c_j is a vector with the same number of elements as the input vector x. Note that there is one comparison or shifting operation for each c_j stored in the network, and one centre is defined for each radial basis function in the network. Centres which are closest to the input data vector will have the smallest output, the limiting case being where a centre exactly coincides with an input vector. In this case, the Euclidean distance is zero (Hu and Hwang, 2003).

One of the famous radial basis functions is the Gaussian function:

$$f(x) = e^{\frac{-(x-c)^2}{r^2}} \qquad \textbf{(4-25)}$$

where $c \in \Re$ is the centre of the basis function which has radius r. The Gaussian radial basis function with centre $c = 0$ and radius $r = 1$ is shown in Figure 4-22.

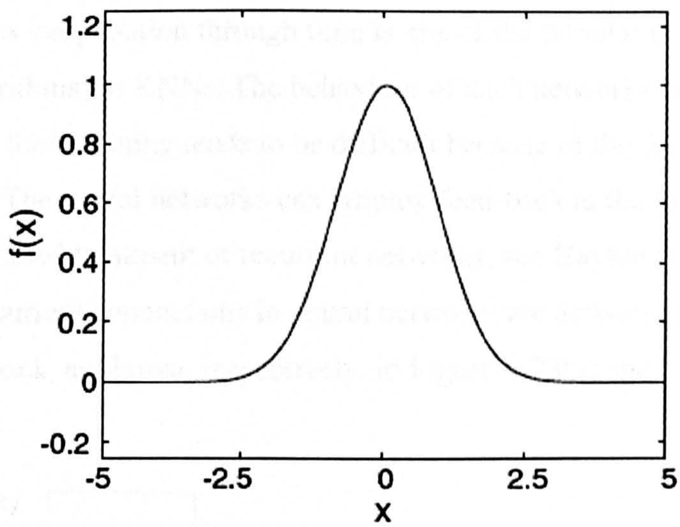
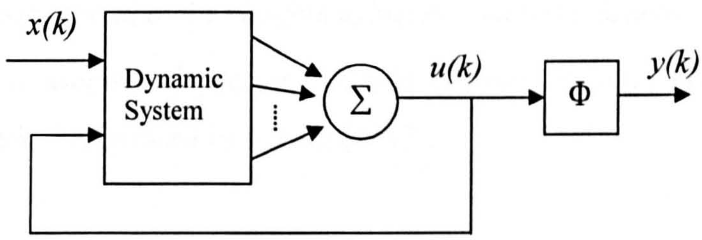


Figure 4-22: The Gaussian radial basis function with center c = 0 and radius r = 1.

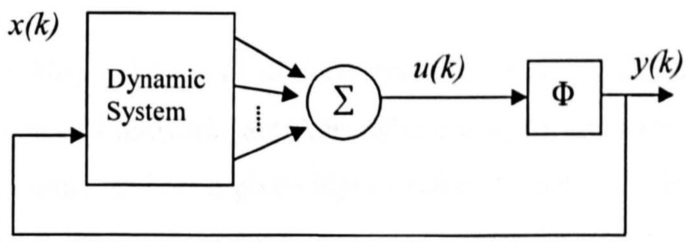
4.7.3 Recurrent Back-propagation Networks

As mentioned in section 4.4, a neural network is feed-forward if all of the hidden and output neurons receive inputs from the preceding layer only. The input is presented to the input layer and it is propagated forwards through the network. Output never forms a part of its own input. A recurrent network has at least one feed-back loop, i.e. cyclic connection, which means that at least one of its neurons feed its signal back to the inputs of all the other neurons.
Recurrent back-propagation is a back-propagation network with feed-back or the recurrent connections. Recurrent Neural Networks (RNN) have feedback connections from neurons in one layer to neurons in a previous layer. Different modifications of such networks have been developed and explored over time.

A typical recurrent network has concepts bound to the neurons whose output values feed-back as inputs to the network. So the next state of a network depends not only on the connection weights and the currently presented input signals but also on the previous states of the network. The network leaves a trace of its behaviour, i.e. the network keeps a memory of its previous states. They address the temporal relationship of inputs by maintaining internal states that have memory. Back-propagation through time is one of the popular learning and training algorithms for RNNs. The behaviour of such networks may be extremely complex and their training tends to be difficult because of the feedback connections. The neural networks can employ feed-back at the local or global level. For detailed treatment of recurrent networks, see Haykin (1999). Two ways to include recurrent connections in neural networks are activation feed-back and output feed-back, as shown, respectively, in Figure 4-23(a) and Figure 4-23(b).



(a)



(b)

Figure 4-23: feed-back schemes; (a) Activation feed-back (b) Output feed-back (Medsker and Jain, 2001).

The output of a neuron shown in Figure 4-23 (a) can be expressed as (Mandic and Chambers, 2001):

$$u(k) = \sum_{i=0}^{M} w_{x,i}(k)x(k-i) + \sum_{j=1}^{N} w_{u,j}(k)u(k-j) \qquad \textbf{(4-26)}$$

and:

$$y(k) = \Phi(u(k))$$

where $w_{x,i}$ and $w_{u,j}$ correspond to the weights associated with x and u, respectively. The transfer function of a neuron shown in Figure 4-23 (b) can be expressed as (Mandic and Chambers, 2001):

$$u(k) = \sum_{i=0}^{M} w_{x,i}(k)x(k-i) + \sum_{j=1}^{N} w_{y,j}(k)y(k-j) \qquad \textbf{(4-27)}$$

and:

$$y(k) = \Phi(u(k))$$

where $w_{y,j}$ correspond to the weights associated with the delayed outputs. A comprehensive account of types of synapses and short-term memories in dynamical neural networks is provided by Mozer (1993).

4.8 NN ARCHITECTURES

Artificial Neural Networks can be divided into two classes: static and dynamic networks. Static neural networks are those whose outputs are linear, or non-linear, functions of its inputs, and for a given input vector, the network always generate the same output vector. In static networks, a given set of input variables at time t are used to forecast a target output variable at time t. These networks are suitable to identify input-output relationships and also categorise the input-output maps to find out the most accurate one. In contrast, dynamic neural networks are capable of implementing memories which give them the possibility of retaining information to be used later. The dynamic neural networks can generate diverse output vectors in response to the same input vector. The response may also depend on the actual state of the existing memories, by their inherent characteristic of memorising past information, for long or short-term periods. Moreover, from the perspective of the network connection patterns, static networks have no loops but in dynamic networks loops occur because of the feedback connections in the networks architecture.

The Multilayer Perceptron (MLP) networks (Haykin, 1999) are widespread examples of static neural networks architecture. They carry no memory but their universal approximation capability has been proved (Hornik et al, 1989). The choice of which neural networks to employ to represent a nonlinear physical process depends on the dynamics and complexity of the network that is best for representing the problems. The Multilayer Perceptron (Figure 4-24) and RBF based neural networks (see section 4-6-2) are the two most commonly used types of static feed-forward networks. A fundamental difference between the two is the way in which hidden units combine values at their inputs. Feed-forward networks might not be powerful enough to capture the dynamic of the underlying non-linear dynamical system but the non-linear static map generated by the static network can adequately represent the system's dynamical behaviour in the ranges of interest for a particular application (Mandic and Chambers, 2001). However, wherever possible the material has been presented starting from feed-forward networks and building up to the dynamic ANN.

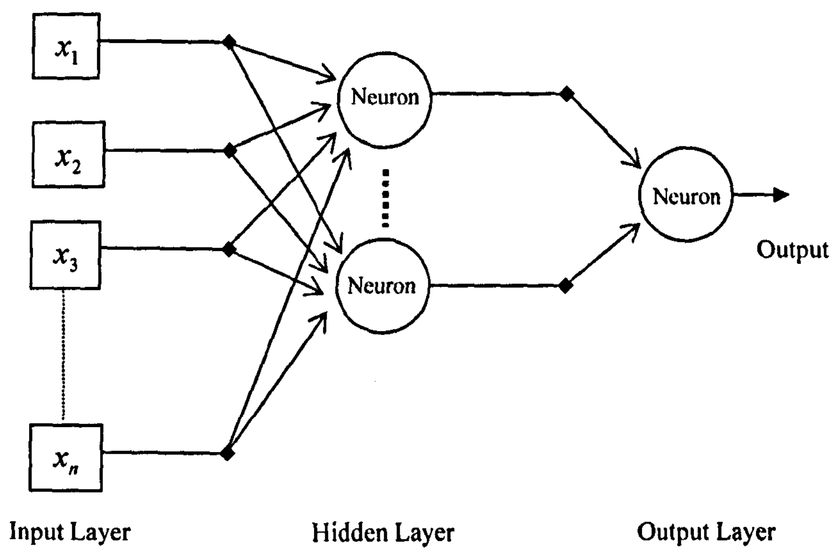


Figure 4-24: A static multilayer feed-forward neural networks.

The traditional multilayer perceptron neural network can only learn a static mapping of input variables to a forecast value. This means that MLP alone cannot model the temporal or dynamic aspects of the system where the time is not a constant. For example, if data in the sample are not homogeneous or the underlying data generating process in a time series changes over time, then a larger sample may damage the performance of the static neural networks as well as other traditional methods. But the output of a dynamic neural unit depends on the present inputs as well as the past neural states and lets the information be temporally memorised by the networks (Kung and Hwang, 1998).

Dynamic neural network architectures were categorised by Tsoi and Back in 1997. Earlier, Werbos introduced the back-propagation through time approach, approximating the time evolution of a dynamic neural network as a sequence of static networks using gradient methods (Werbos, 1990).

There are also other similar architectures able to process dynamic data, usually by inducing dynamics in existing static models with recurrent connections among neurons (global memory) or implementing dynamics in neurons. The latter types of architectures vary with regard to the place of the dynamics in the weights, in the activation function, or both (Lawrence, 1995; Tsoi, 1994).

The provision of feedback, with delay, introduces memory to the network and so is appropriate for prediction. The feedback within dynamic neural networks can be achieved in a global manner (Figure 4-25). The global feedback is produced by the connection of the network output to the network input. Inter-neuron connections can also exist in the hidden layers.

The Time Lagged Recurrent Network (TLRN) closely resembles a MLP extended with memory structures. In fact, TLRNs have been shown to be appropriate models for processing time-varying information (Principe *et al.* 2000). Moreover, a sequential framework can be added to the MLP feed-forward networks by adding short-term memory in the form of a delay line. A section of the time series (called the time window) of the form $[x(k), x(k-1),..., x(k-p)]$, and $[y(k), y(k-1),..., y(k-p)]$ are used as input for the feed-forward network. The delay line is in order of p, and the desired output is $y(k+1)$. This network can be referred to as the Time Delay Neural Network (TDNN). Furthermore, the final choice of structure for each case depends upon the dynamics of the signal, learning algorithm and ultimately the prediction performance. There is no rule as to the best structure to use for a particular problem (Personnaz and Dreyfus 1998).

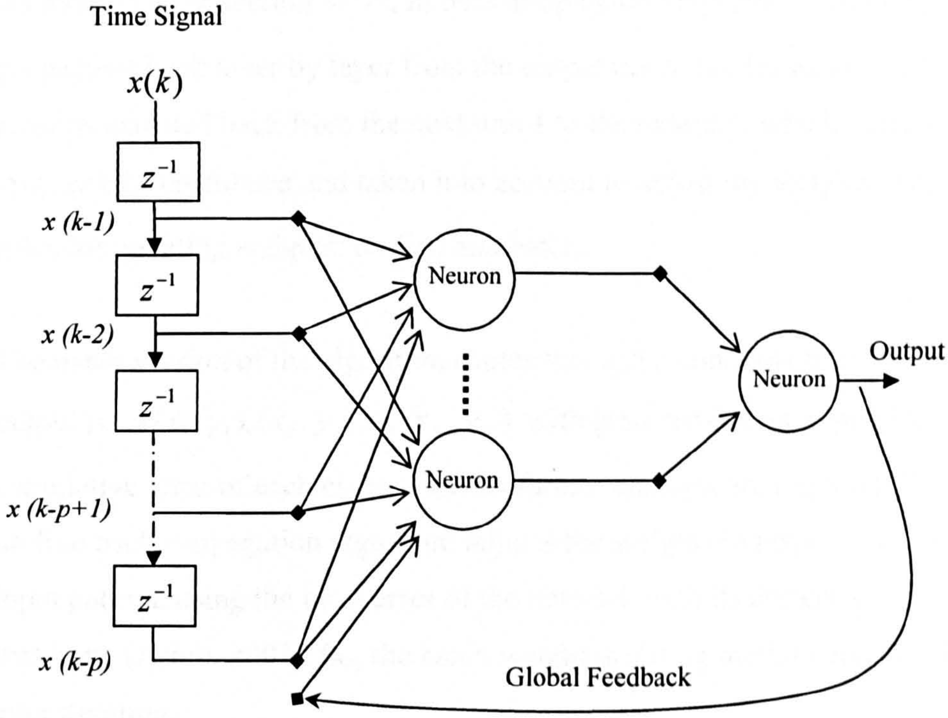


Figure 4-25: The Structure of a dynamic neural networks with global feedback (Mandic and Chambers, 2001).

The correlation coefficient (r) is a method to measure the performance of the above neural networks. By definition, the correlation coefficient between a network output y and a desired output d is:

$$r = \frac{\dfrac{\sum_i (y_i - \bar{y})(d_i - \bar{d})}{N}}{\sqrt{\dfrac{\sum_i (d_i - \bar{d})^2}{N}}\sqrt{\dfrac{\sum_i (y_i - \bar{y})^2}{N}}} \tag{4-28}$$

The correlation coefficient is confined to the range [-1, 1]. When r =1 there is a perfect positive correlation between x and d and r = -1 represents a perfectly negative correlation between x and d. Also, when r = 0 there is no correlation between x and d.

4.8.1 Weight updating

As mentioned in section 4-7-1, in back-propagation algorithm, the error signal ε_i propagates back layer by layer from the output units. So, for example, unit *i* receives error propagated back from the next unit *k* to the extent to which *i* affects *k*. Then Δw_{ki} will be calculated and taken into account to adjust the weights. There are two rules for updating weights: on-line and batch.

The batch version of the algorithm cycles through a complete training set of input-output pairs $(x_1, y_1), (x_2, y_2), ..., (x_N, y_N)$ with gradient decent applied to the cumulative error of each cycle, until no further changes are required. On the contrary, on-line back-propagation algorithm adjusts the weights in response to each single input pattern, using the local error of the network with its current weight settings for that input (Arbib, 2003). So, the batch weight updating method updates the weights after iteration.

4.9 CASE STUDIES

It is important to point out that building successful ANN is a combination of art and science; software alone is not sufficient to solve all problems in the process (Zhang, 2004). There are no precise rules to find the best architecture of an ANN. It is critical to understand the key issues surrounding the model-building process. There are different parameters and connections (how the neurons in each layer are connected to neurons in adjacent layers) which need to be identified in each particular case. Parameters like the number of hidden layers, the number of neurons (or processing elements) in each layer, the activation functions and the learning rate of the networks have to be identified carefully. For both ANN models, static and recurrent, the existence of a vast amount of design possibilities allows experimentation but also sets out the doubt about which might be the best of combinations among design and training parameters. Unfortunately, there is not a mathematical basis that might back the selection of a specific architecture, and only few works (Lapedes & Farber, 1987; Cybenko, 1989) have shown lower and upper limits for PE number at some models and with restricted types of problems (Rabuñal and Dorado, 2006).

The major decisions a neural networks forecaster must make include data preparation, input variable selection, choice of network type and architecture, transfer function, and training algorithm, as well as model validation, evaluation and selection. Some of these can be solved during the model building process while others must be considered before actual modelling starts (Zhang, 2004). Moreover, the types of connections also play an important role for the performance of the networks. Because of the problems such as over-fitting or over-training of the neural networks (see section 4-9) the strategies to train and consequently test the network need to be chosen carefully.

A few examples of using artificial neural networks in different areas are presented in Table 4-1. The first four rows are networks which are implemented to predict different stock markets. As it appears in the table there is not a unique architecture for the purpose of the forecasting. There are some common characteristics but different types of activations, various numbers of layers, various numbers of neurons in each layer and some other factors that make different prediction abilities. The challenge is to find out the special type of ANN and its specifications for each particular purpose. For example, the specifications of the networks for the two studies of the Swedish stock market, which used different delay strategy, are different from the Tokyo or Australia stock market networks. There are some common characteristics between an ANN to predict the stock market and the river system in Spain or the number of the lynx in Canada (Table 4-1). When a network is properly configured, it offers outstanding abilities that can be used for prediction. But ANNs have numerous parameters and architectures, and no confirmed theory has been established for specifying the optimal network architecture. This can be observed by a wide variety of research methods reported in the literature.

Application	Time Series	Training Algorithm	NN Architecture	Inputs	Outputs	Success	Software
Stock Market Prediction - Swedish Stock Exchange (Nygren, K.2004)	Daily Data; January 1993 – June 2000	Back-propagation 30 Days Windowing	ECNN (Error Correction Neural Network)	*Internal:* Highest Price Lowest Price Closing Price Volume *External:* Intrest Rate DowJonesInd.Gold Price (Total Ext.10)	Swedish Stock Index (SXGE)	56.8%	SENN (Simiulation Environmentfor Neural Networks)
Prediction of Two Major Stock, Erricsson B and Volvo B - Swedish Stock Exchange (Nygren, K.2004)	Weekly Data; January 1993 - December 1999	Back-propagation 12 Weeks Windowing	ECNN (Error Correction Neural Network)	*Internal:* Highest Price Lowest Price Closing Price Volume *External:* Intrest Rate DowJonesInd.Swedish Ind. (Total Ext.6)	Erricsson B and Volvo B Indices	Good	SENN (Simiulation Environmentfor Neural Networks)
Prediction of the Tokyo Stock Exchange Prices (Kulkarni, 1998)	Daily Data; Over 33 Months	Back-propagation (Moving the Target Learning)	Recurrent - MNN (Modular Neural Network; 4 BP Modules)	Turnover, Interest Rate, Foreign Exchange rate, New York DowJones Average (Total 7)	Tokyo Stock Exchange Prices Index (TOPIX)	93.8%	N/A
Security Selection in the Australian Stock market (Vanstone et al., 2004)	1994-2003, (22,944 Row) 8 Year Training and 2 Years Test	Back-propagation (Logistical Sigmoid Function)	Feed-forward (Various Number of Neurons Tested Finally used 7)	P/E Ratio, Payout Ratio, Total Current Assets, Earnings per Share, Year End Share Price (Total 14)	Securities (determine which securities are likely to have the best chances)	Very Good	N/A
Modelling for the simulation of Managing the upper Tagus River system in Spain. (Ocho-Rivera et al., 2003)	Monthly Data; For 53 years	Back-propagation (Bipolar Sigmoid Function)	MLP Feed-forward (With an added Multivariate Random component) (3 Layer, 6 Hidden Nodes; [14-6-7])	Stream Flow for 7 Different River Station (Total 14)	Generation of Discharge Data Series	Very Good	N/A
Ecological research - Lynx Population in North-West Canada. (Katijani et al., 2005)	Annual Data; 1820 - 1934	Back-propagation	Feed-forward [1-2-2]	Numbers of Lynx Trapped	Magnitude of the Lynx Population	Very Good	SPSS; Neural Connection

Table 4-1: Neural Networks application in some related studies.

4.10 ANN DISADVANTAGES

Besides all the positive aspects of the ANNs there are also some disadvantages.

In general, neural networks are data-driven techniques. Therefore, data preparation is a critical step in building a successful neural networks model. Without an adequate and representative data set, it is impossible to develop a useful, predictive ANN model (Zhang, 2004). Thus, the reliability of ANN models depends to a large extent on the quality of data.

The functioning of ANNs is also hardly transparent. This limitation is due to the fact that they function as a black box. The knowledge acquired during network training is encoded within the internal structure of the ANNs. Currently, there is no explicit technique to extract the relationship between input and output variables, causing difficulty in interpreting results from the networks (Zhang, 2004). It is suggested that a careful modelling decision for training and testing should have been made based on a particular problem.

Another common problem which is associated with the neural networks in general is the over-fitting problem. ANN training should stop at an appropriate point to avoid over-training. Over-training of the networks causes the over-fitting of the network and consequently lack of accuracy of the networks. Figure 4-26 explains this issue and, as it appears in the plot, the best point to stop training is the minimum of the test error curve. There are also some other reasons for the over-fitting problem e.g. noise in the data, insufficient data or too many inputs.

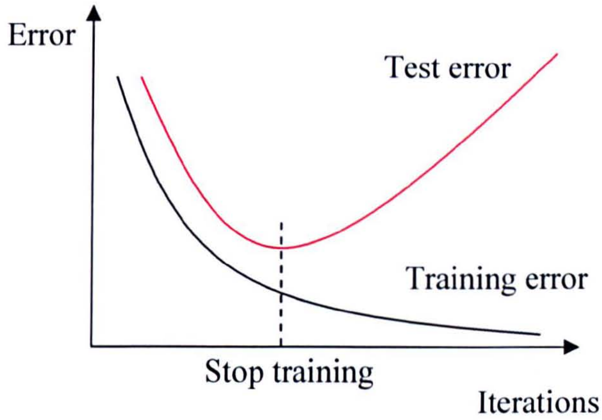


Figure 4-26: Over-training problem.

The performance of ANNs is usually not very stable since the training process may depend on the choice of a random start. The degree of success may fluctuate from one training pass to another. This needs a careful training strategy and subsequently a reliable procedure to handle the produced data and the obtained information of each network.

One of the disadvantages of back-propagation algorithm, which has been used for this study, is that the training phase is very time consuming. This issue is getting worse when the number of neurons in different layers increases and sometimes it needs few hours to train the networks depending on the abilities of the computer. As it is explained before in this chapter, a solution for speeding up the process of back-propagation toward convergence is to add a momentum term to the above process. The extensions of the back-propagation method also allow different learning rates for different parameters. However, efficient as back-propagation may be, it still suffers from the trap of local rather than global minima, or saddle point convergence. Moreover, while low values of the learning parameters avoid oscillations, they may unnecessarily prolong the convergence process (McNelis, 2005). Therefore, one of the problems with back-propagation is that both convergence speed toward the solution and the possibility of reaching a solution depend on the choice of the learning rate γ and the proportionality constant of momentum α .

4.11 Conclusion

The different ANN models and architectures have been demonstrated and their variables introduced and explained in detail in this chapter. It has been shown that ANNs are able to capture the dynamics of the data (see section 4-6-3 and 4-8) and identify the most effective inputs on a set of desired outputs (see section 4-5-2). These criteria make them a proper tool to analyse complex environments, like the shipping market, and establish a proper model to map the inputs and the outputs accurately (see section 4-6-2). In parallel, considering the fast growth of computers, in terms of speed and the power of calculations, ANNs can now perform more accurate and faster. Therefore, using ANN methodology to model and consequently forecast markets is a growing topic and the shipping market is not an exception. Implementing the ANN for a specific market is a high maintenance process and needs intensive study. The structure and architecture of ANNs vary for every network (see section 4-9). There are different factors and parameters which need to be identified and evaluated accurately to achieve realistic results as a lack of accuracy can cause convergence to irrational results.

CHAPTER 5

STATISTICAL MODELLING OF THE DEMOLITION MARKET

5.1 INTRODUCTION

In this chapter, different statistical analytical methods, which are generally multivariate analysis methods, are using to model the demolition market. The scrap price and the scrapped tonnage, as the main pillars of the demolition market, are investigated and consequently predicted in each method. The MLR, PCR and PLS modelling methods are used to model these variables using The Unscrambler® 9.1 and SPSS® 13.0 software. Firstly, the mathematical structure of the demolition market is investigated and secondly, the future of the market is predicted using the above statistical methods. It means these three statistical approaches have been used to forecast the future of the demolition market for a given period of time. The focus of the study is on the tanker market as one of the main drivers of the demolition market. As mentioned before, tankers have big flat panels which are generally easy to access and subsequently cut into small pieces and these criteria make them favourable for scrap yards. Hence, the objective is to predict: first monthly scrapped tonnage and second monthly demolition prices (in two different locations i.e. Subcontinent and Far-East scrapyards) based on the tanker market. In fact, the prediction of these two parameters plays an important role having a realistic plan for the future of the ship scrapping.

In addition to the internal market inputs there are also some external factors which have influence over the demolition market including importantly steel price and oil production. To obtain the most accurate models for each method both the internal and external factors are taken into account for the modelling.

5.2 DATA STRUCTURE

The data, which have been used for this study, are monthly data from January 1995 till December 2004. This means each time series has 120 variables. A total number of 33 inputs/outputs have been considered for this study which was split into two different groups of internal and external factors. Internal factors are variables which are applicable to the tanker market including tanker new building prices, second-hand tanker market values, tanker freight rates and tanker fleet supply and demand. External factors are variables which may influence the tanker market but they are not in the tanker market itself including oil production, steel production, steel price and exchange rates. A list of all variables can be found in Table 5-1.

There are also important variables which can directly influence the market. For example, the old tanker fleet (about 25 years old or more) are likely to be scrapped in near future and, as demolition is labour intensive, wages also play an important role to determine scrap prices. In general, the price which has been offered for scrapping a ship will be estimated as a function of (Buxton, 1991):

- value of realisable materials (tonnage outturn of various material categories),
- delivery cost to scrapyard and
- cost of demolition.

Hence, labour cost can determine the cost of demolition. In addition to the above, the second-hand values for the older ships would also be beneficial to the model. In this research, however, none of the above variables were available on a monthly scale to be entered into the ANN.

All timeseries are based on monthly data, so the prediction of the final model will be based on monthly forecasting as well. The aim of the research, as explained before, is to predict three steps ahead of the market which means 3 months ahead. It was possible to use data in, for example, the average of quarterly period to have a smoother graph but the number of patterns for the neural networks would decrease by three quarters and therefore there will not be enough patterns for accurate modelling.

In order to use this model for the period longer than 3 months, mean 6 or 9 months, it is possible to use the produced model for 3 months and put them in the model for a second time. However, it is not recommended as the error will increase for the secondary and any subsequent modelling.

No.	External Variables	Unit
1	Crude Steel Production- USA	million tonnes
2	Crude Steel Production- EU	million tonnes
3	Crude Steel Production- China	million tonnes
4	Crude Steel Production- Japan	million tonnes
5	Crude Steel Production- S. Korea	million tonnes
6	Average Steel Price	USD per tonne
7	Scrap Price- Subcontinent	USD per ldt
8	Scrap Price- Far-East	USD per ldt
9	Oil Production- OPEC	million Barrel per day
10	Oil Production- non-OPEC	million Barrel per day
11	Oil World Trade	million Barrel per day
12	Exchange Rate- Euro/USD	Index 1995/1=100
13	Exchange Rate- Won/USD	Index 1995/1=100
14	Exchange Rate- YEN/$	Index 1995/1=100
15	Bunkers- 380 CST in Rotterdam	USD per million tonnes
No.	**Internal Variables**	**Unit**
1	MR Product Tankers' Building Price	million USD
2	Aframax Tanker DH Building Price	million USD
3	Suezmax Tanker DH Building Price	million USD
4	VLCC DH Building Price	million USD
5	Crude Carrier 105000 dwt- Freight Rate Single Voyage	kUSD per day
6	Crude Carrier 150000 dwt- Freight Rate Single Voyage	kUSD per day
7	Crude Carrier 300000 dwt- Freight Rate Single Voyage	kUSD per day
8	Clean Carrier 70/85000 dwt- Freight Rate Single Voyage	kUSD per day
9	MR Product Tanker DS/DH- 5Years Market Value	million USD
10	Aframax Tanker DS/DH- 5Years Market Value	million USD
11	Suezmax Tanker SH/DH- 5Years Market Value	million USD
12	VLCC SH/DH- 5Years Market Value	million USD
13	Clean Carrier 40/45000 dwt DB/DH- 10 Years SH Value	million USD
14	Tanker Fleet 10000+ dwt Supply	million dwt
15	Tanker Fleet 10000+ dwt Demand	million dwt
16	Tanker Fleet 10000+ dwt Utilisation Rate	percent
17	Tanker Order Book in Percent of Existing Fleet	percent
18	Tankers Sold for Scrapping	million dwt

Table 5-1: List of all the variables used in the following studies

5.3 STATISTICAL ANALYSIS OF THE DATA

In this section, multivariate statistical methods, as explained in Chapter 3, have been used to investigate the time series. The main objective, in each method, is to find out the structure of the inputs and favourable output (or outputs) and consequently model the demolition market.

To distinguish the character of all the variables in this study, standard deviation and mean for every individual time series is shown in (Figure 5-1).This plot displays the average value and the standard deviation of all the variables together. The vertical bar is the average value for the time series, and the standard deviation is shown as an error bar around the average. The standard deviation is a measure of the spread of the variable around that average. Comparing the standard deviations confirms that they vary a lot from one variable to another, so it is recommended to pre-process the data and standardise the variables before the examination in some of the multivariate analyses e.g. PCA, PCR or PLS.

The Box-plots which correspond to all the variables are also shown in Figure 5-2. This plot contains one Box-plot for each variable, either over the whole sample set, or for different subgroups. It shows the minimum, the 25% percentile[1] (lower quartile), the median, the 75% percentile (upper quartile) and the maximum. This plot is a good summary of the distributions of all variables. It demonstrates the total range of variation of each variable, and act as a check whether all variables are within the expected range. The distance between minimum and maximum shows the spread of the variable.

[1] The X% percentile of an observed distribution is the variable value that splits the observations into X% lower values, and 100-X% higher values.

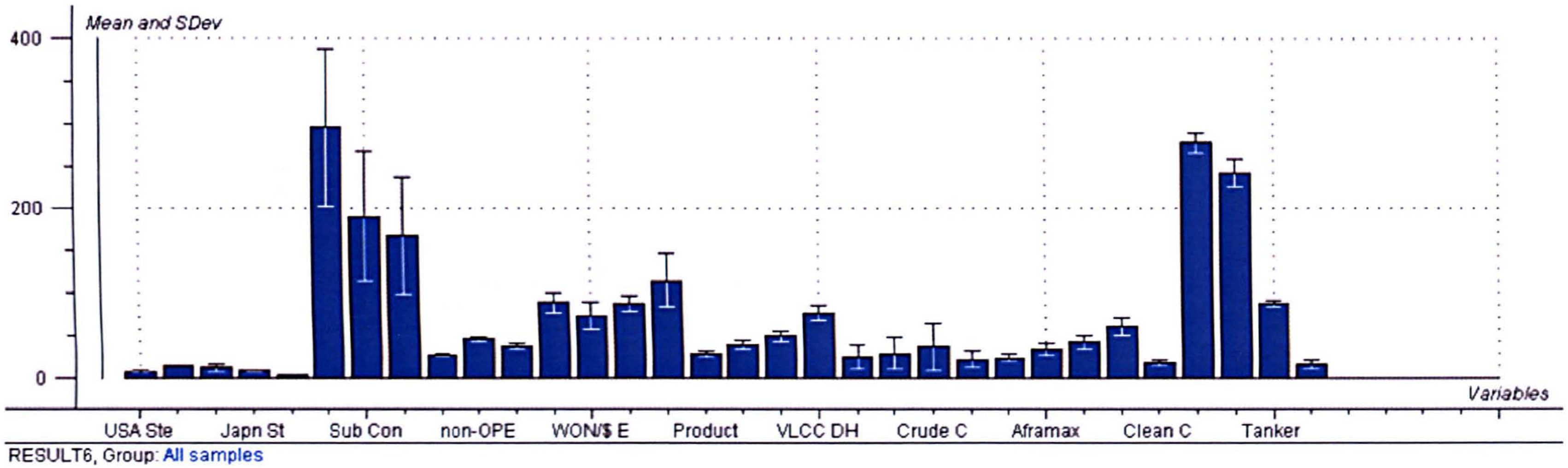


Figure 5-1: Standard deviation and Mean for every individual variable. Standard deviation varies a lot from one variable to another.

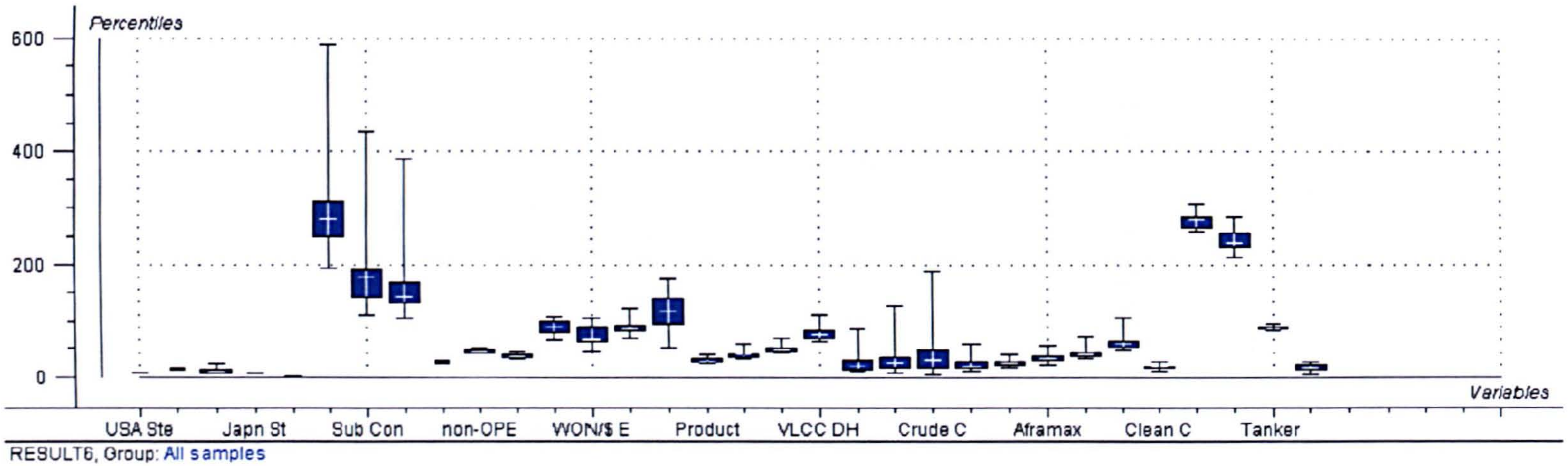


Figure 5-2: The Box-plot for every individual variable shows the minimum, the 25% percentile (lower quartile), the median, the 75% percentile (upper quartile) and the maximum.

5.3.1 MLR

The fundamental of the MLR method is explained in Chapter 3. The aim of this section is to model the monthly scrapped ships and scrap prices for two different locations, the Subcontinent and Far-East, based on the MLR modelling method. The first part of this section is dedicated to the scrap tonnage prediction and the second part to the scrapping prices prediction.

5.3.1.1 Scrap Tonnage Prediction Using MLR

There are 32 *X*-variables for this part of the study and the only *Y*-variable is the scrapped tanker tonnage (Figure 5-3). The overall correlation for all the raw variables of the study is shown in Table 5-2. As it appears, the correlation between the scrapped tanker tonnage, as the only *Y*-variable, and the rest of the variables, as *X*-variables, shows a small correlation. The highest correlation in the table (which is a negative correlation) belongs to the OPEC oil production with -0.6. Also, Won/USD and Yen/USD exchange rates have the highest positive correlation with 0.4. As mentioned in section 3-5-1, generally, the MLR method can be used when the *X*-variables are not correlated or they have only a small correlation. Hence, MLR seems to be a proper method to model the scrapped tanker tonnage.

X-variables

$x_1(t)$: USA Steel Production
$x_2(t)$: EU Steel Production
$x_3(t)$: China Steel Production
$x_4(t)$: Japan Steel Production
$x_5(t)$: South Korea Steel Production
$x_6(t)$: Steel Price
$x_7(t)$: Subcontinent Scrap Price
$x_8(t)$: Far-East Scrap Price
$x_9(t)$: OPEC Oil Production
$x_{10}(t)$: Non-OPEC Oil Production
$x_{11}(t)$: Oil World Trade
$x_{12}(t)$: EUR/$ Ex. Rate
$x_{13}(t)$: WON/$ Ex. Rate
$x_{14}(t)$: YEN/$ Ex. Rate
$x_{15}(t)$: Bunkers Price
$x_{16}(t)$: Product Tankers Building Price
$x_{17}(t)$: Aframax DH Building Price
$x_{18}(t)$: Suezmax DH Building Price
$x_{19}(t)$: VLCC DH Building Price
$x_{20}(t)$: Crude Carrier 105000dwt FRSingle Voyage
$x_{21}(t)$: Crude Carrier 150000dwt FRSingle Voyage
$x_{22}(t)$: Crude Carrier 300000dwt FRSingle Voyage
$x_{23}(t)$: Clean Carrier 70/85000dwt FRSingle Voyage
$x_{24}(t)$: Product Tankers DS/DH 5Years Market Value
$x_{25}(t)$: Aframax DS/DH 5Years Market Value
$x_{26}(t)$: Suezmax SH/DH 5Years Market Value
$x_{27}(t)$: VLCC SH/DH 5Years Market Value
$x_{28}(t)$: Clean Carrier- 40/45000dwt DB/DH 10 Years
$x_{29}(t)$: Tanker Fleet- 10000 DWT+ Supply
$x_{30}(t)$: Tanker Fleet- 10000 DWT+ Demand
$x_{31}(t)$: Tanker Fleet- 10000 DWT+ Util. Rate
$x_{32}(t)$: Tanker Order Book in Percent of Existing Fleet

Y-variable

$y(t)$: Scrapped Tankers

Figure 5-3: All the X and Y-variables of the model for the scrapping tonnage

A model based on MLR has been made (Figure 5-4) and subsequently the full cross validation method has been chosen to validate the accuracy of the model. In addition, as explained in section 3-6-2, all the weights are set to 1/Stdv for each *X*-variable. The prediction of the mentioned model is shown in Figure 5-5.

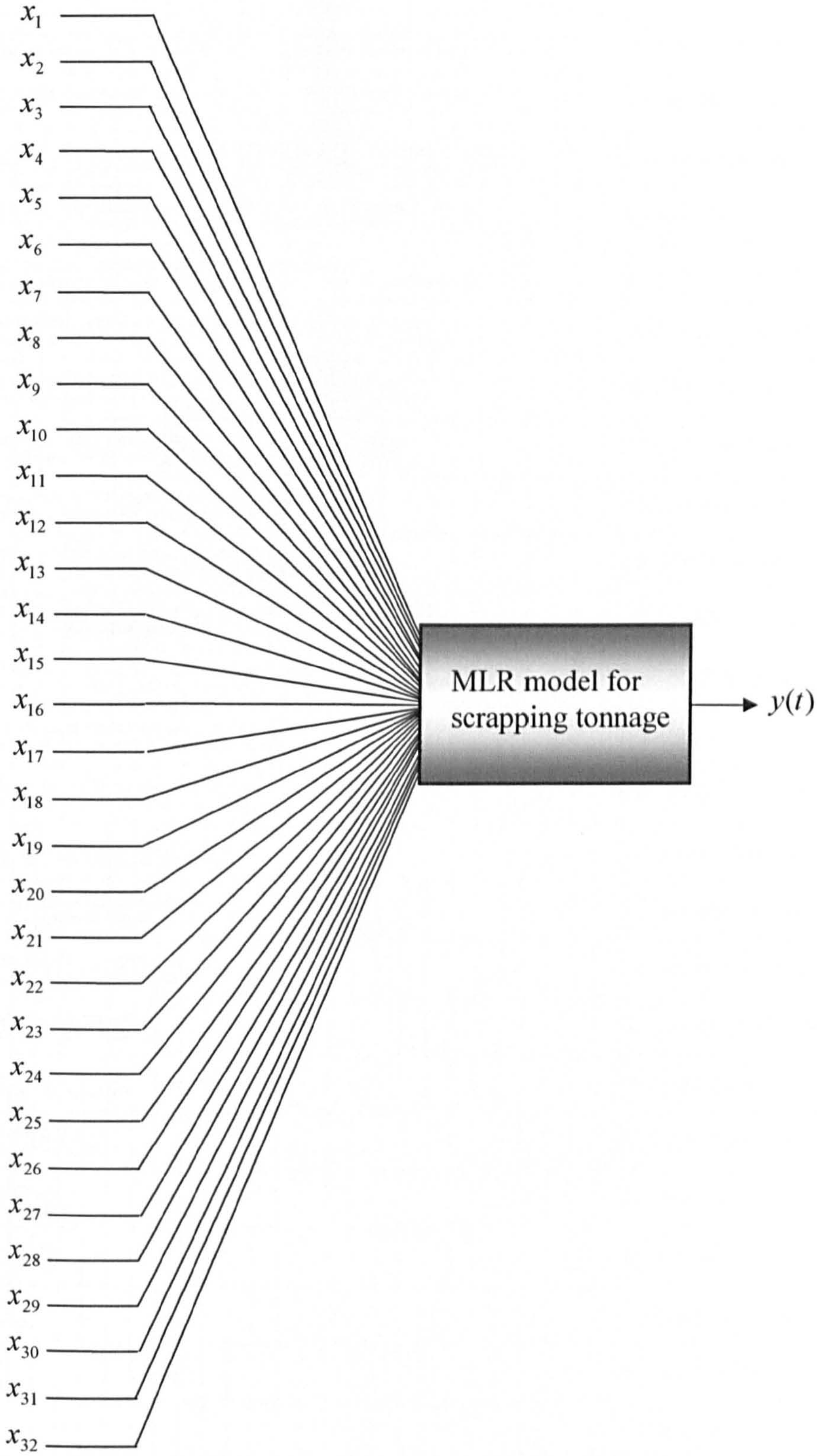


Figure 5-4: the MLR model for the monthly scrapped tonnage

	1- USA Steel Prod.	2- EU Steel Prod.	3- ChinaSteel Prod.	4- Japn Steel Prod.	5- S.Korea Steel Prod.	6- Steel Price	7- Sub Continent Scrap Price	8- FarEast Scrap Price	9- OPEC Oil Prod.	10- non-OPEC Oil Prod.	11- Oil World Trade	12- EUR/$ Ex. Rate	13- WON/$ Ex. Rate	14- YEN/$ Ex. Rate	15- Bunkers Price	16- Product Tankers Building Price	17- Aframax DH Building Price	18- Suezmax DH Building Price	19- VLCC DH Building Price	20- Crude Carrier 105000dwt FRSingle V.	21- Crude Carrier 150000dwt FRSingle V.	22- Crude Carrier 300000dwt FRSingle V.	23- Clean Carrier 70/85000dwt FRSingle	24- Product Tankers DS/DH 5Years Market	25- Aframax DS/DH 5Years Market Value	26- Suezmax SH/DH 5Years Market Value	27- VLCC SH/DH 5Years Market Value	28- Clean Carrier- 40/45000dwt DB/DH 10 Y	29- Tanker Fleet- 10000 DWT+ Supply	30- Tanker Fleet- 10000 DWT+ Demand	31- Tanker Fleet- 10000 DWT+ Util. Rate	32- Tanker Order Book in Percent of Existing	33- Tankers Sold for Scrapping
1	1.0																																
2	0.5	1.0																															
3	0.1	0.6	1.0																														
4	0.0	0.6	0.7	1.0																													
5	0.0	0.6	0.8	0.7	1.0																												
6	0.4	0.6	0.6	0.6	0.4	1.0																											
7	0.1	0.5	0.8	0.7	0.5	0.8	1.0																										
8	0.1	0.6	0.8	0.7	0.6	0.8	1.0	1.0																									
9	0.5	0.6	0.5	0.3	0.4	0.5	0.5	0.5	1.0																								
10	0.1	0.6	0.9	0.7	0.9	0.5	0.6	0.7	0.5	1.0																							
11	0.1	0.7	0.9	0.6	0.9	0.5	0.6	0.8	0.7	0.9	1.0																						
12	0.1	0.2	0.0	0.0	0.4	0.4	0.4	0.3	0.1	0.3	0.2	1.0																					
13	0.1	0.6	0.4	0.1	0.7	0.0	0.0	0.1	0.5	0.6	0.7	0.6	1.0																				
14	0.1	0.1	0.1	0.2	0.4	0.2	0.3	0.2	0.2	0.3	0.3	0.5	0.7	1.0																			
15	0.1	0.4	0.7	0.7	0.7	0.5	0.6	0.6	0.1	0.7	0.7	0.2	0.2	0.2	1.0																		
16	0.2	0.1	0.1	0.3	0.2	0.7	0.7	0.5	0.1	0.1	0.1	0.7	0.6	0.5	0.1	1.0																	
17	0.2	0.3	0.4	0.5	0.1	0.8	0.9	0.7	0.4	0.2	0.3	0.5	0.4	0.4	0.3	0.9	1.0																
18	0.2	0.3	0.4	0.5	0.1	0.8	0.8	0.7	0.3	0.2	0.3	0.5	0.4	0.4	0.3	0.9	1.0	1.0															
19	0.3	0.1	0.1	0.3	0.2	0.7	0.7	0.5	0.2	0.1	0.0	0.6	0.5	0.5	0.0	1.0	0.9	0.9	1.0														
20	0.2	0.5	0.6	0.6	0.5	0.5	0.6	0.7	0.7	0.6	0.7	0.1	0.2	0.0	0.5	0.3	0.5	0.5	0.3	1.0													
21	0.3	0.6	0.7	0.6	0.6	0.6	0.7	0.7	0.7	0.6	0.7	0.0	0.3	0.0	0.5	0.3	0.6	0.5	0.3	0.9	1.0												
22	0.3	0.6	0.7	0.6	0.5	0.6	0.7	0.7	0.8	0.6	0.7	0.0	0.3	0.0	0.4	0.3	0.6	0.5	0.4	0.9	0.9	1.0											
23	0.1	0.4	0.5	0.6	0.4	0.3	0.5	0.6	0.5	0.5	0.5	0.1	0.2	0.0	0.4	0.2	0.5	0.5	0.3	0.8	0.7	0.7	1.0										
24	0.2	0.4	0.6	0.7	0.4	0.8	0.9	0.8	0.5	0.4	0.5	0.4	0.1	0.2	0.4	0.8	0.9	0.9	0.8	0.7	0.7	0.7	0.6	1.0									
25	0.2	0.6	0.6	0.7	0.5	0.7	0.8	0.8	0.7	0.6	0.6	0.1	0.1	0.0	0.4	0.6	0.8	0.8	0.6	0.7	0.8	0.7	0.7	0.9	1.0								
26	0.2	0.6	0.8	0.6	0.6	0.7	0.8	0.8	0.7	0.7	0.8	0.0	0.3	0.0	0.5	0.4	0.7	0.7	0.5	0.8	0.8	0.8	0.7	0.9	0.9	1.0							
27	0.3	0.6	0.7	0.6	0.5	0.8	0.8	0.8	0.7	0.6	0.7	0.1	0.2	0.0	0.5	0.5	0.8	0.8	0.6	0.8	0.8	0.8	0.7	0.9	0.9	1.0	1.0						
28	0.3	0.3	0.3	0.5	0.1	0.7	0.7	0.6	0.4	0.1	0.2	0.4	0.4	0.3	0.2	0.9	0.9	0.9	0.9	0.5	0.5	0.6	0.5	0.9	0.8	0.7	0.8	1.0					
29	0.0	0.6	0.9	0.6	0.8	0.5	0.6	0.7	0.6	0.9	1.0	0.3	0.6	0.2	0.7	0.1	0.2	0.2	0.1	0.6	0.7	0.6	0.5	0.4	0.6	0.8	0.7	0.1	1.0				
30	0.2	0.7	0.8	0.7	0.8	0.6	0.7	0.8	0.8	0.8	0.9	0.2	0.5	0.2	0.6	0.1	0.4	0.4	0.1	0.8	0.8	0.8	0.7	0.7	0.8	0.9	0.8	0.4	0.9	1.0			
31	0.5	0.6	0.5	0.7	0.5	0.5	0.6	0.6	0.8	0.5	0.6	0.0	0.3	0.1	0.3	0.3	0.5	0.5	0.3	0.8	0.8	0.8	0.7	0.7	0.7	0.8	0.7	0.6	0.5	0.8	1.0		
32	0.1	0.7	0.9	0.6	0.9	0.4	0.6	0.7	0.5	0.9	0.9	0.3	0.7	0.3	0.6	0.1	0.2	0.2	0.1	0.6	0.6	0.6	0.5	0.4	0.6	0.7	0.6	0.1	0.9	0.8	0.5	1.0	
33	0.3	0.3	0.2	0.1	0.3	0.2	0.2	0.2	0.6	0.2	0.4	0.1	0.4	0.4	0.0	0.0	0.1	0.1	0.0	0.3	0.4	0.5	0.3	0.3	0.3	0.4	0.4	0.2	0.2	0.4	0.5	0.3	1.0

Table 5-2: Correlation between all variables. Negative numbers are in blue cells.

Figure 5-5: The MLR model prediction of the monthly scrapped tanker

The model and its Analysis Of Variance (ANOVA) are shown in Table 5-3 . This table includes Sum of Squares (SS), number of Degrees of Freedom (DF), Mean Square (MS=SS/DF), F-value, p-value, B-coefficient and Standard Error of the b-coefficients (STDerr).

	SS	DF	MS	F-ratio	p-value	B-coefficient	STDerr
Summary							
Model	91.145	32	2.848	5.111	0		
Error	48.482	87	0.557				
Adjusted Total	139.627	119	1.173				
Variable							
Intercept	0.681	1	0.681	1.222	0.2719	29.357	26.552
USA Steel Prod.	5.08E-02	1	5.08E-02	9.12E-02	0.7634	-0.218	0.723
EU Steel Prod.	1.129	1	1.129	2.026	0.1582	-0.659	0.463
ChinaSteel Prod.	1.389	1	1.389	2.492	0.1181	0.464	0.294
Japn Steel Prod.	6.18E-02	1	6.18E-02	0.111	0.74	0.245	0.735
S.Korea Steel Pr	0.394	1	0.394	0.707	0.4026	1.382	1.644
Steel Price	0.802	1	0.802	1.44	0.2334	5.18E-03	4.31E-03
Sub Continent Sc	7.29E-02	1	7.29E-02	0.131	0.7184	2.92E-03	8.06E-03
FarEast Scrap Pr	0.373	1	0.373	0.669	0.4155	5.41E-03	6.62E-03
OPEC Oil Prod.	1.24E-02	1	1.24E-02	2.23E-02	0.8818	5.29E-02	0.355
non-OPEC Oil Pro	1.954	1	1.954	3.507	0.0645	-0.705	0.376
Oil World Trade	5.371	1	5.371	9.638	0.0026	-1.055	0.34
EUR/$ Ex. Rate	9.93E-02	1	9.93E-02	0.178	0.674	-1.57E-02	3.71E-02
WON/$ Ex. Rate	5.17E-02	1	5.17E-02	9.28E-02	0.7614	-8.45E-03	2.77E-02
YEN/$ Ex. Rate	2.519	1	2.519	4.521	0.0363	5.78E-02	2.72E-02
Bunkers Price	1.67E-02	1	1.67E-02	3.00E-02	0.8628	-1.30E-03	7.50E-03
Product Tankers	1.351	1	1.351	2.424	0.1231	0.427	0.274
Aframax DH Build	1.274	1	1.274	2.285	0.1342	-0.524	0.347
Suezmax DH Build	0.282	1	0.282	0.507	0.4785	-0.187	0.263
VLCC DH Building	0.257	1	0.257	0.461	0.4987	0.122	0.179
Crude Carrier 10	0.182	1	0.182	0.326	0.5694	1.40E-02	2.46E-02
Crude Carrier 15	0.489	1	0.489	0.878	0.3512	1.67E-02	1.78E-02
Crude Carrier 30	2.769	1	2.769	4.968	0.0284	-2.26E-02	1.01E-02
Clean Carrier 70	0.334	1	0.334	0.6	0.4408	1.48E-02	1.91E-02
Product Tankers	1.78	1	1.78	3.194	0.0774	-0.307	0.172
Aframax DS/DH 5Y	1.632	1	1.632	2.929	0.0905	0.214	0.125
Suezmax SH/DH 5Y	1.158	1	1.158	2.079	0.153	-0.174	0.121
VLCC SH/DH 5Year	0.33	1	0.33	0.592	0.4436	7.05E-02	9.16E-02
Clean Carrier- 4	3.37E-02	1	3.37E-02	6.06E-02	0.8062	6.18E-02	0.251
Fleet- Supply	1.558	1	1.558	2.796	0.0981	0.105	6.28E-02
Fleet-Demand	2.817	1	2.817	5.055	0.0271	9.26E-02	4.12E-02
Fleet- Util. Rate	0.151	1	0.151	0.271	0.6037	-9.51E-02	0.183
Tanker Order Book	1.563	1	1.563	2.805	0.0976	0.157	9.37E-02

Table 5-3: the Analysis of variance for the scrapped tonnage MLR model

The b-coefficients are the values of the regression coefficients are displayed for each variable of the model and each regression coefficient is estimated with a certain precision, measured as a standard error (STDerr in the table).

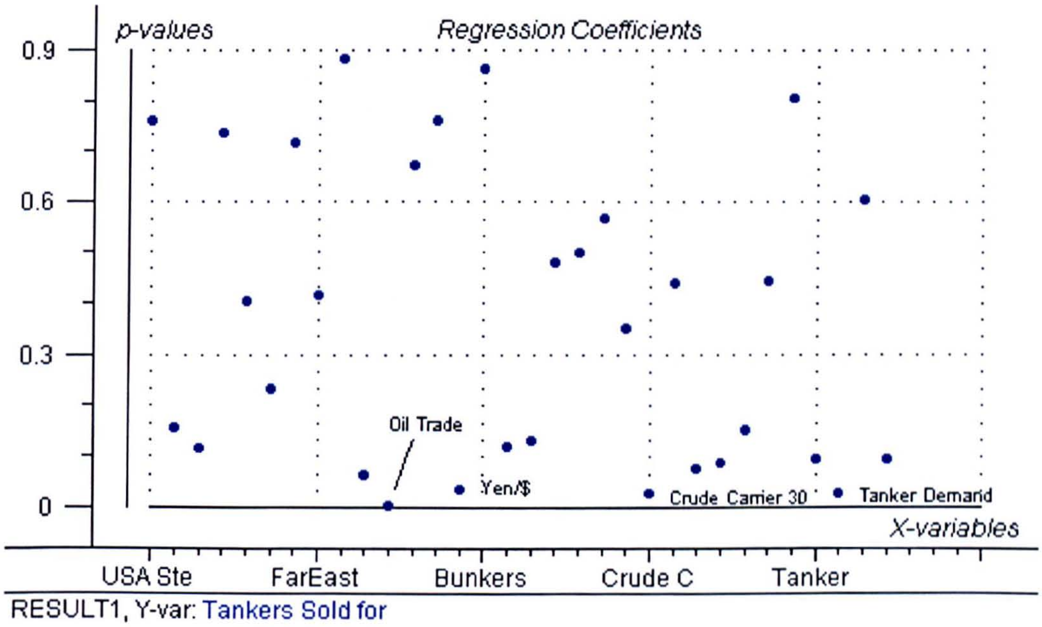


Figure 5-6: The p-values correspond to each *X*-variable.

As mentioned in section 3-5-1, the only relevant measure of how well the model performs is provided by the *Y*-variances. Residual *Y*-variance for the present model is shown in Figure 5-7 as well.

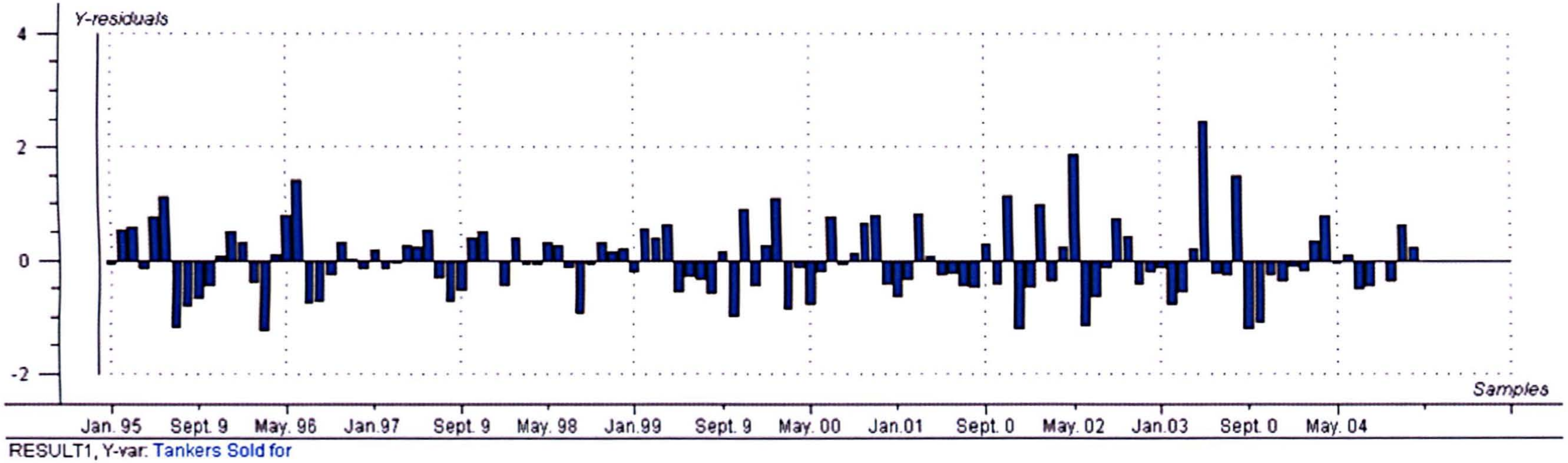


Figure 5-7: The residuals versus sample number for selected variables for the MLR model, used to detect outliers and lack of model fit.

5.3.1.2 Scrap Price Prediction Using MLR

For this part of the study, similarly, a model is implemented to forecast the scrap prices for both Subcontinent and Far-East scrapyards. There are two *Y*-variables compared with the one *Y*-variable for the previous model (Figure 5-8).

X-variables

$x_1(t)$: USA Steel Production

$x_2(t)$: EU Steel Production

$x_3(t)$: China Steel Production

$x_4(t)$: Japan Steel Production

$x_5(t)$: South Korea Steel Production

$x_6(t)$: Steel Price

$x_7(t)$: OPEC Oil Production

$x_8(t)$: Non-OPEC Oil Production

$x_9(t)$: Oil World Trade

$x_{10}(t)$: EUR/$ Ex. Rate

$x_{11}(t)$: WON/$ Ex. Rate

$x_{12}(t)$: YEN/$ Ex. Rate

$x_{13}(t)$: Bunkers Price

$x_{14}(t)$: Product Tankers Building Price

$x_{15}(t)$: Aframax DH Building Price

$x_{16}(t)$: Suezmax DH Building Price

$x_{17}(t)$: VLCC DH Building Price

$x_{18}(t)$: Crude Carrier 105000dwt FRSingle Voyage

$x_{19}(t)$: Crude Carrier 150000dwt FRSingle Voyage

$x_{20}(t)$: Crude Carrier 300000dwt FRSingle Voyage

$x_{21}(t)$: Clean Carrier 70/85000dwt FRSingle Voyage

$x_{22}(t)$: Product Tankers DS/DH 5Years Market Value

$x_{23}(t)$: Aframax DS/DH 5Years Market Value

$x_{24}(t)$: Suezmax SH/DH 5Years Market Value

$x_{25}(t)$: VLCC SH/DH 5Years Market Value

$x_{26}(t)$: Clean Carrier- 40/45000dwt DB/DH 10 Years

$x_{27}(t)$: Tanker Fleet- 10000 DWT+ Supply

$x_{28}(t)$: Tanker Fleet- 10000 DWT+ Demand

$x_{29}(t)$: Tanker Fleet- 10000 DWT+ Util. Rate

$x_{30}(t)$: Tanker Order Book in Percent of Existing Fleet

$x_{31}(t)$: Scrapped Tankers

Y-variable

$y_1(t)$: Subcontinent Scrap Price

$y_2(t)$: Far-East Scrap Price

Figure 5-8: All the X and Y-variables of the model for the scrapping price

The rest of the parameters remain the same i.e. weights and cross validation method. The overall model connections is represented in Figure 5-9. The prediction curves of this model are shown in Figure 5-10 and Figure 5-11 beside the actual measured samples for both regions.

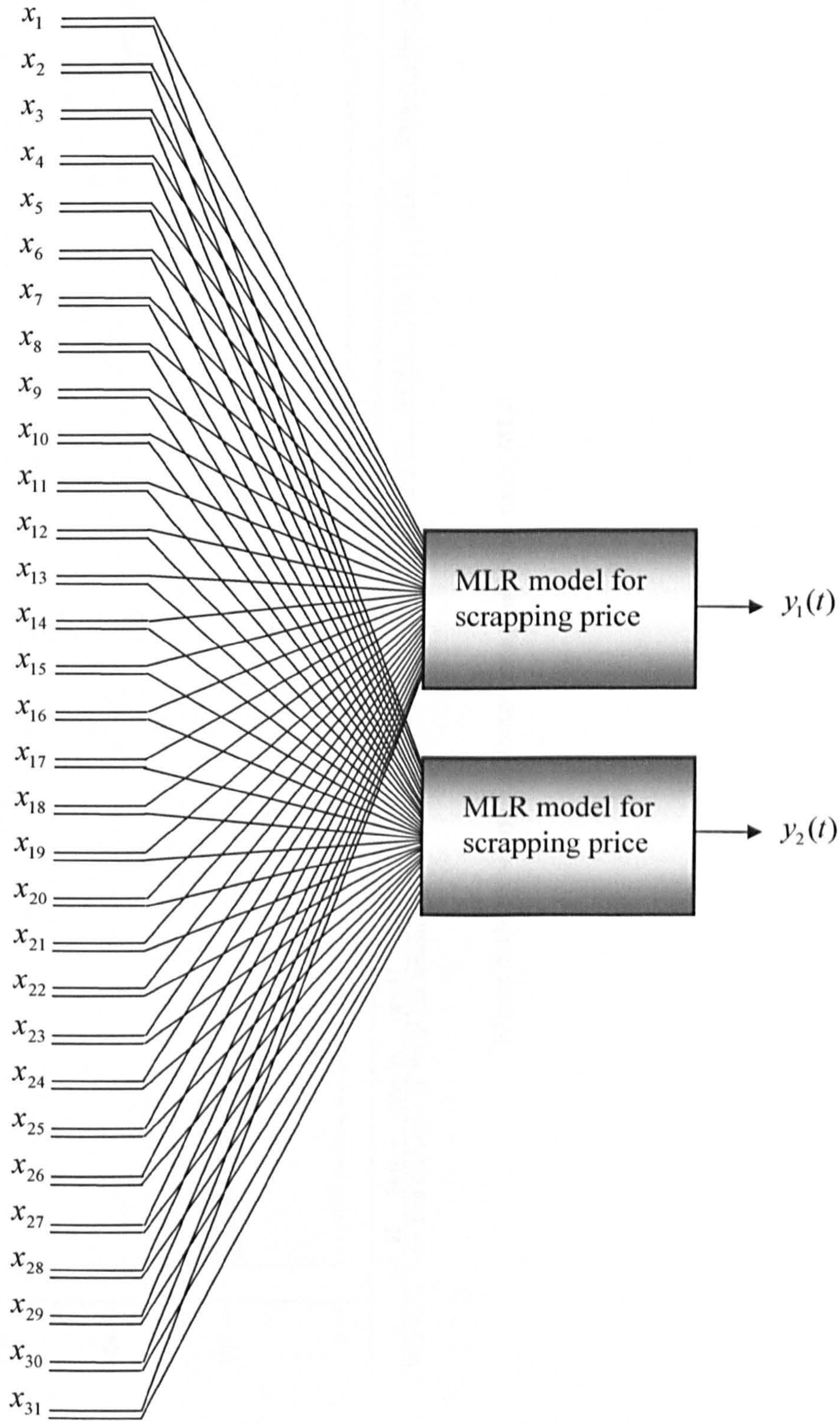


Figure 5-9: the MLR model structure for the monthly scrap prices

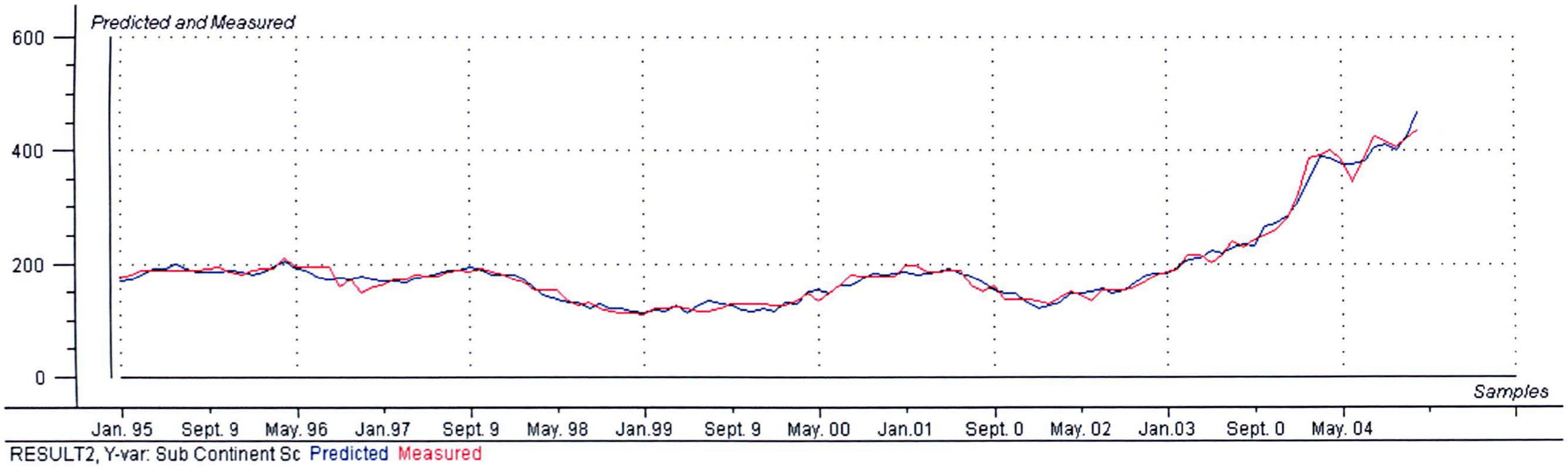


Figure 5-10: Subcontinent scrap price prediction using MLR

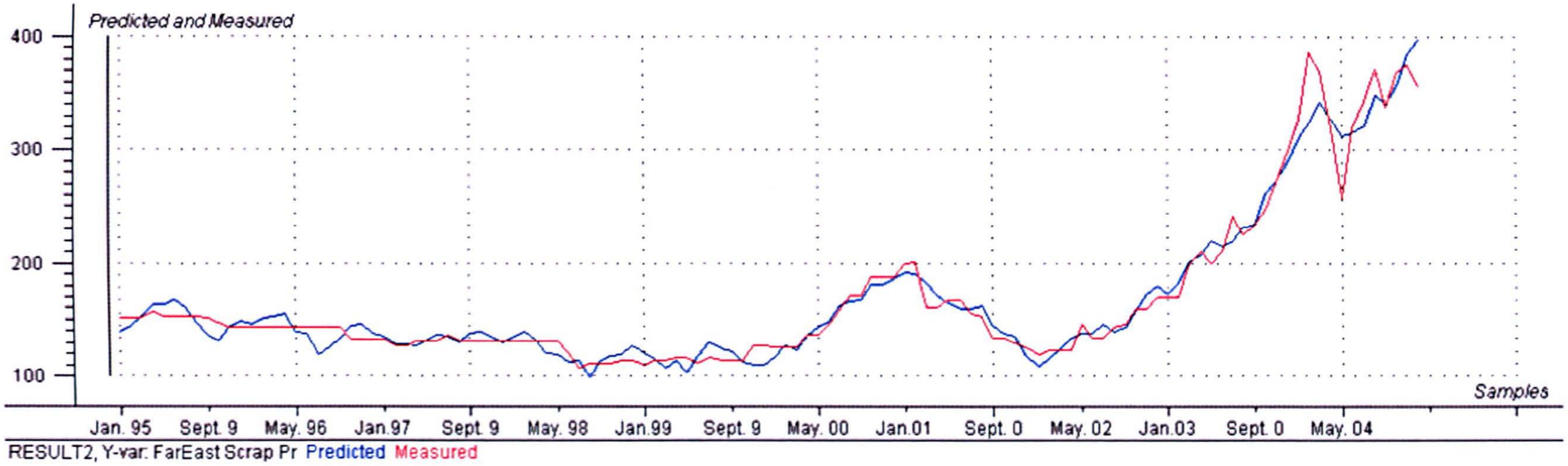


Figure 5-11: Far-East scrap prices prediction using MLR

Consequently, the analysis of variance of this model is shown in Table 5-4 and Table 5-5.

	SS	DF	MS	F-ratio	p-value	B-coefficient	STDerr
Summary							
Model	6.74E+05	31	2.17E+04	144.586	0		
Error	1.32E+04	88	150.348				
Adjusted Total	6.87E+05	119	5.77E+03				
Variable							
Intercept	528.002	1	528.002	3.512	0.0642	805.308	429.727
USA Steel Prod.	90.228	1	90.228	0.6	0.4406	-9.175	11.844
EU Steel Prod.	9.488	1	9.488	6.31E-02	0.8022	1.929	7.678
ChinaSteel Prod.	1.15E+03	1	1.15E+03	7.677	0.0068	12.96	4.678
Japn Steel Prod.	642.842	1	642.842	4.276	0.0416	-24.302	11.753
S.Korea Steel Pr	178.901	1	178.901	1.19	0.2783	-29.29	26.851
Steel Price	1.87E+03	1	1.87E+03	12.435	0.0007	0.229	6.49E-02
OPEC Oil Prod.	351.186	1	351.186	2.336	0.13	8.788	5.75
non-OPEC Oil Pro	306.025	1	306.025	2.035	0.1572	-8.883	6.227
Oil World Trade	2.661	1	2.661	1.77E-02	0.8945	-0.757	5.693
EUR/$ Ex. Rate	5.693	1	5.693	3.79E-02	0.8462	-0.119	0.61
WON/$ Ex. Rate	0.161	1	0.161	1.07E-03	0.974	-1.49E-02	0.456
YEN/$ Ex. Rate	3.714	1	3.714	2.47E-02	0.8755	7.16E-02	0.456
Bunkers Price	456.362	1	456.362	3.035	0.085	0.208	0.119
Product Tankers	287.953	1	287.953	1.915	0.1699	-6.241	4.51
Aframax DH Build	1.31E+03	1	1.31E+03	8.68	0.0041	16.106	5.467
Suezmax DH Build	41.003	1	41.003	0.273	0.6028	-2.183	4.18
VLCC DH Building	227.384	1	227.384	1.512	0.2221	-3.595	2.923
Crude Carrier 10	30.223	1	30.223	0.201	0.655	0.18	0.401
Crude Carrier 15	13.862	1	13.862	9.22E-02	0.7621	8.87E-02	0.292
Crude Carrier 30	461.65	1	461.65	3.071	0.0832	-0.293	0.167
Clean Carrier 70	60.231	1	60.231	0.401	0.5284	-0.197	0.312
Product Tankers	960.659	1	960.659	6.39	0.0133	6.963	2.755
Aframax DS/DH 5Y	848.8	1	848.8	5.646	0.0197	4.794	2.018
Suezmax SH/DH 5Y	679.216	1	679.216	4.518	0.0363	-4.137	1.947
VLCC SH/DH 5Year	467.058	1	467.058	3.107	0.0815	-2.606	1.479
Clean Carrier- 4	57.805	1	57.805	0.384	0.5368	2.523	4.069
Tanker Fleet- 10	678.889	1	678.889	4.515	0.0364	-2.161	1.017
Tanker Fleet- 10	1.44E+03	1	1.44E+03	9.589	0.0026	2.045	0.66
Tanker Fleet- 10	192.493	1	192.493	1.28	0.2609	-3.37	2.978
Tanker Order Boo	4.222	1	4.222	2.81E-02	0.8673	0.262	1.563
Tankers Sold for	171.505	1	171.505	1.141	0.2884	1.862	1.743

Table 5-4: The analysis of variance for the Subcontinent MLR model

	SS	DF	MS	F-ratio	p-value	B-coefficient	STDerr
Summary							
Model	5.40E+05	31	1.74E+04	78.535	0		
Error	1.95E+04	88	221.848				
Adjusted Total	5.60E+05	119	4.70E+03				
Variable							
Intercept	80.853	1	80.853	0.364	0.5476	315.132	522.001
USA Steel Prod.	43.94	1	43.94	0.198	0.6574	-6.403	14.387
EU Steel Prod.	78.745	1	78.745	0.355	0.5529	5.556	9.326
China Steel Prod.	238.456	1	238.456	1.075	0.3027	5.891	5.682
Japan Steel Prod.	793.096	1	793.096	3.575	0.0619	-26.993	14.276
S.Korea Steel Prod.	329.84	1	329.84	1.487	0.226	-39.77	32.616
Steel Price	2.962	1	2.962	1.34E-02	0.9083	9.11E-03	7.89E-02
OPEC Oil Prod.	305.726	1	305.726	1.378	0.2436	8.2	6.985
non-OPEC Oil Pro	57.542	1	57.542	0.259	0.6118	-3.852	7.564
Oil World Trade	798.331	1	798.331	3.599	0.0611	13.118	6.915
EUR/$ Ex. Rate	2.054	1	2.054	9.26E-03	0.9236	-7.13E-02	0.741
WON/$ Ex. Rate	0.856	1	0.856	3.86E-03	0.9506	3.44E-02	0.554
YEN/$ Ex. Rate	137.011	1	137.011	0.618	0.4341	0.435	0.554
Bunkers Price	1.35E+03	1	1.35E+03	6.075	0.0156	0.357	0.145
Product Tankers	34.603	1	34.603	0.156	0.6938	-2.163	5.478
Aframax DH Build	161.975	1	161.975	0.73	0.3952	5.674	6.641
Suezmax DH Build	656.388	1	656.388	2.959	0.0889	8.734	5.077
VLCC DH Building	284.263	1	284.263	1.281	0.2607	-4.02	3.551
Crude Carrier 10	249.307	1	249.307	1.124	0.292	0.517	0.488
Crude Carrier 15	47.973	1	47.973	0.216	0.6431	-0.165	0.355
Crude Carrier 30	3.787	1	3.787	1.71E-02	0.8964	-2.65E-02	0.203
Clean Carrier 70	97.262	1	97.262	0.438	0.5096	0.251	0.379
Product Tankers	1.35E+03	1	1.35E+03	6.097	0.0155	8.263	3.346
Aframax DS/DH 5Y	151.377	1	151.377	0.682	0.411	2.025	2.451
Suezmax SH/DH 5Y	1.07E+03	1	1.07E+03	4.843	0.0304	-5.204	2.365
VLCC SH/DH 5Year	38.578	1	38.578	0.174	0.6777	-0.749	1.796
Clean Carrier- 4	146.08	1	146.08	0.658	0.4193	-4.011	4.942
Tanker Fleet- 10	952.127	1	952.127	4.292	0.0412	-2.559	1.235
Tanker Fleet- 10	555.467	1	555.467	2.504	0.1172	1.269	0.802
Tanker Fleet- 10	236.539	1	236.539	1.066	0.3046	-3.735	3.618
Tanker Order Boo	17.167	1	17.167	7.74E-02	0.7815	0.528	1.898
Tankers Sold for	375.291	1	375.291	1.692	0.1968	2.754	2.117

Table 5-5: The analysis of variance for the Far-East MLR model

Residual Y-variance for the present model for Subcontinent and Far-East scrap prices are shown in Figure 5-12 and Figure 5-13 respectively.

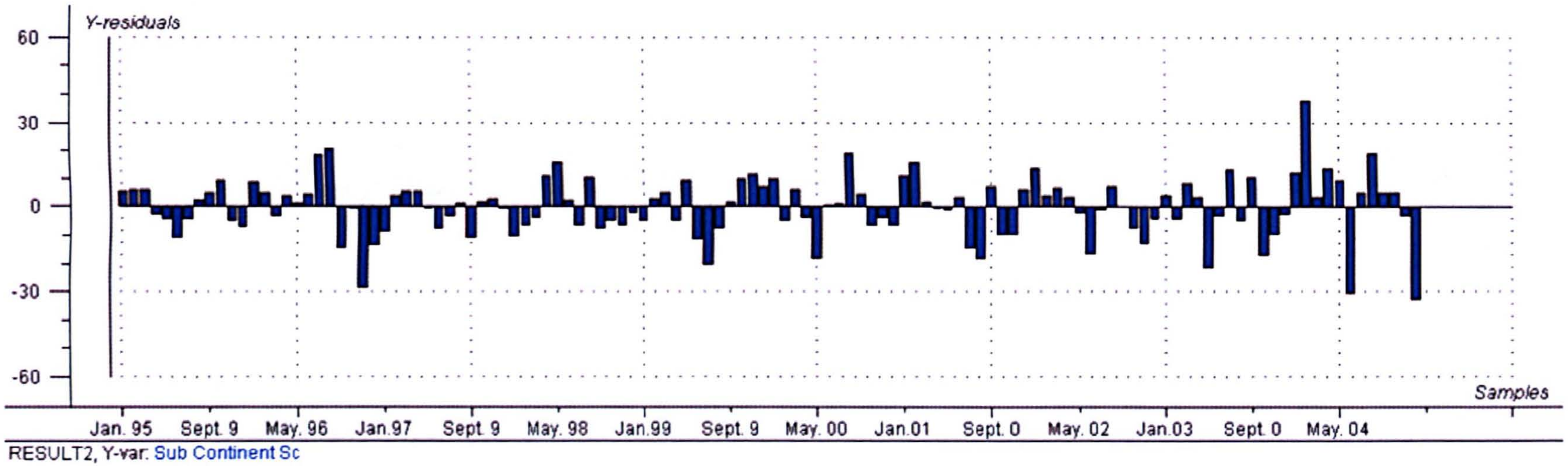


Figure 5-12: Subcontinent residual for MLR model

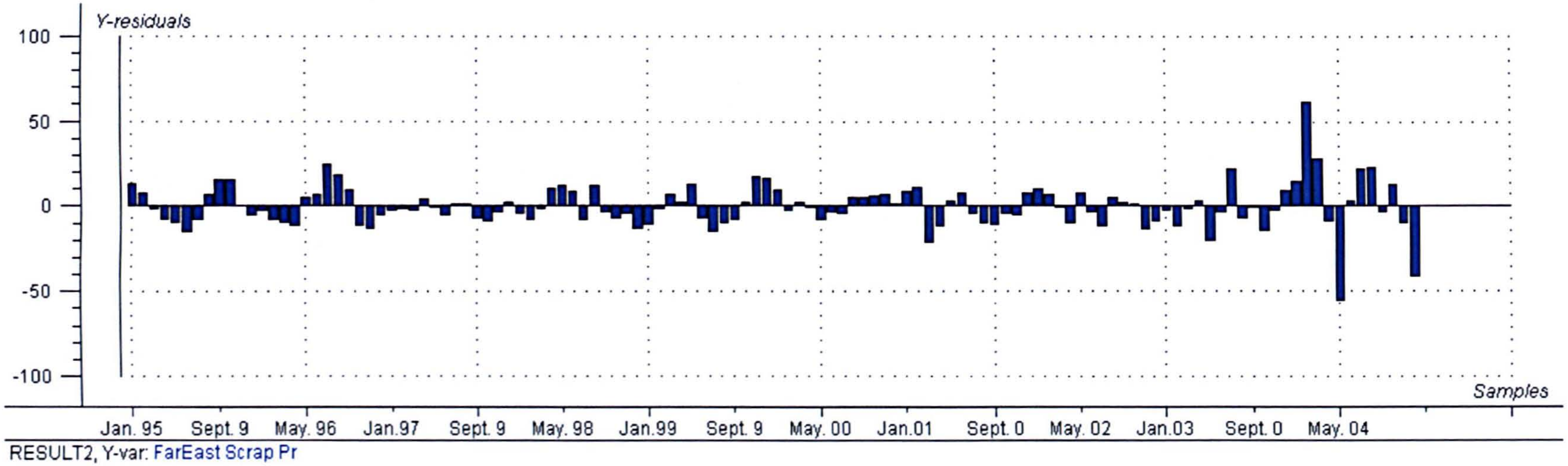


Figure 5-13: Far-East residual for MLR model

5.3.1.3 MLR MODELLING RESULTS

5.3.1.3.1 SCRAPED TONNAGE

For the MLR modelling which is carried out for the monthly scrapped tonnage, the first part of the analysis of variance table (Table 5-3) is a summary of the significance of the global model. The p-value for the global model is 0.00 which is far smaller than 0.05. This means that the model explains more of the variations of the response variable than could be expected from random phenomena. In other words, the model is significant at the 5% level. The second part of this table deals with each *X*-variable of the model separately. If the p-value for an individual *X*-value is smaller than 0.05, it means that the corresponding source of variation explains more of the variations of the response variable than could be expected from random phenomena or the effect is significant at the 5% level. The p-values are also shown in Figure 5-6 and as it appears, oil world trade, Yen/USD exchange rate, crude carrier 300k dwt freight rate (single voyage) and tanker fleet demand have the lowest p-values amongst the other *X*-variables for the model. The F-ratios can confirm the previous p-value analysis results. The F-ratio for the world oil trade shows the maximum value with 9.6 and then 5.0, 4.9 and 4.5 for the Yen/USD exchange rate, crude carrier 300k dwt freight rate (single voyage) and tanker fleet demand respectively.

In general, all the B-coefficients, in the table, have small values. The *X*-values with higher regression coefficients have more influence on the obtained model, so this model is a function of mainly:

- South Korean steel production
- Oil world trade
- Non-OPEC oil production
- EU steel production

The highest positive regression belongs to South Korean steel production with 1.382 which represents a small regression between this input and the final MLR model. The highest negative regression belongs to oil world trade with 1.055. Based on this model, there is no significant difference between inputs and it is not possible to reduce the dimension of the input space in order to analyse the market.

The F-ratio represents a small value for the global MLR model. Based on this model, there is no significant difference between all the influences of the inputs. It means that we cannot highlight the most influential inputs to the model.

Residual *Y*-variance for this model is shown in Figure 5-7 which is the variance of the *Y*-residuals and expresses how much variation remains in the observed response if the modelled part has been take out. It is an overall measure of the misfit. In overview, since September 2001 the variation of the residuals has increased and May 2003, May 2002 and August 2003 represent the highest residuals respectively. There is a possibility that there are some special events which happened in those months to influence the market but in general model shows the maximum misfit in these particular dates.

5.3.1.3.2 SCRAP PRICES

For the MLR modelling which is carried out for the monthly scrap prices, the first part of analysis of variance table (Table 5-4) shows a summary of the significance of the global model. The p-value for the global model is 0.00 which is far smaller than 0.05. This means that the model explains more of the variations of the response variable than could be expected from random phenomena. In other words, the model is significant at the 5% level. The second part of the table deals with each *X*-variable. Similarly, if the p-value for an individual *X*-value is smaller than 0.05, it means that the correspondent source of variation explains more of the variations of the response variable than could be expected from random phenomena or the effect is significant at the 5% level.

The p-values for the Subcontinent X-variables represent small p-values for China and Japan Steel Production, Steel price, Aframax Double Hull building price, Suezmax Single/Double Hull 5 years old market value, Tanker Fleet supply and demand.

The X-values with higher regression coefficients have more influence on this model, so this model is a function of mainly:

- South Korean steel production
- Japan steel production
- Aframax double hull building price
- China steel production
- USA steel production

South Korea and Japan steel production have the highest and second highest negative regressions respectively. China steel production has the second highest positive regression. It means that the increment of steel production in China will increases the scrap prices in Subcontinent scrap yards but the increment of steel production in either South Korea or Japan decreases the prices significantly. The Aframax double hull building price is also represents a high influence to the final MLR model as its B-coefficient is calculated 16.106.

Similarly, based on Table 5-5 for the Far-East X-variables, Bunkers Price, Product tankers, Suezmax Single/Double Hull 5 years old market value and tankers supply have small p-values. This model is a function of mainly:

- South Korean steel production
- Japan steel production
- Oil world trade
- Suezmax double hull building price

Similar to the previous model, the highest negative regressions belong to South Korea with 39.77 and the second highest is Japan steel production with 26.99. It means that the increment of production in each of these countries will decrease the scrap prices sharply. The highest positive regression belongs to the oil world trade with 13.11

which is far smaller than the previous two. China steel production has a positive value of 5.89 to this model.

The F-ratio of the global model for both modelling is relatively high with 144.6 and 78.5 respectively. As mentioned earlier, the F-ratio compares structured variance to residual variance and has a statistical distribution which is used for significance testing. Hence, the first model for Subcontinent is more effective than the second one for Far-East. Furthermore, these models are much more effective compared with the monthly scrapped tonnage model in the last section.

Based on these models for both locations, Korean steel production plays the most important role to determine the prices. It also shows that higher steel production in South Korea decreases the scrap prices but an increase in steel production in China increases the scrap prices. These negative and positive effects on scrap price due to the increment of the steel production in different counties are not rational.

Based on the residual *Y*-variance of the MLR model for both Subcontinent and Far-East scrap prices, which are shown in Figure 5-12 and Figure 5-13 respectively, the highest variation for both prices happened in February 2004 which represents the most misfit of the model for this month. There is also a relatively high misfit in December and June 2004 for the Subcontinent prices and in May and December for the Far-East prices. There is a possibility that there are some special events which happened in those months to influence the market but in general model shows the maximum misfit in these particular dates.

5.3.2 PCA AND PCR

The fundamentals of the PCA and PCR methods are explained in Chapter 3. In this section different models, based on these two methods, are built and their performances are tested. The first part of this section is the modelling of the monthly scrapped tonnage and the second part is the modelling of the monthly scrap prices for Subcontinent and Far-East.

5.3.2.1 SCRAP TONNAGE PREDICTION USING PCR

Similar to the previous study of the MLR method, there are 32 X-variables and the only Y-variable is the scrapped tanker tonnage (Figure 5-3). The connections of the variables with the PCR model are illustrated in Figure 5-14.

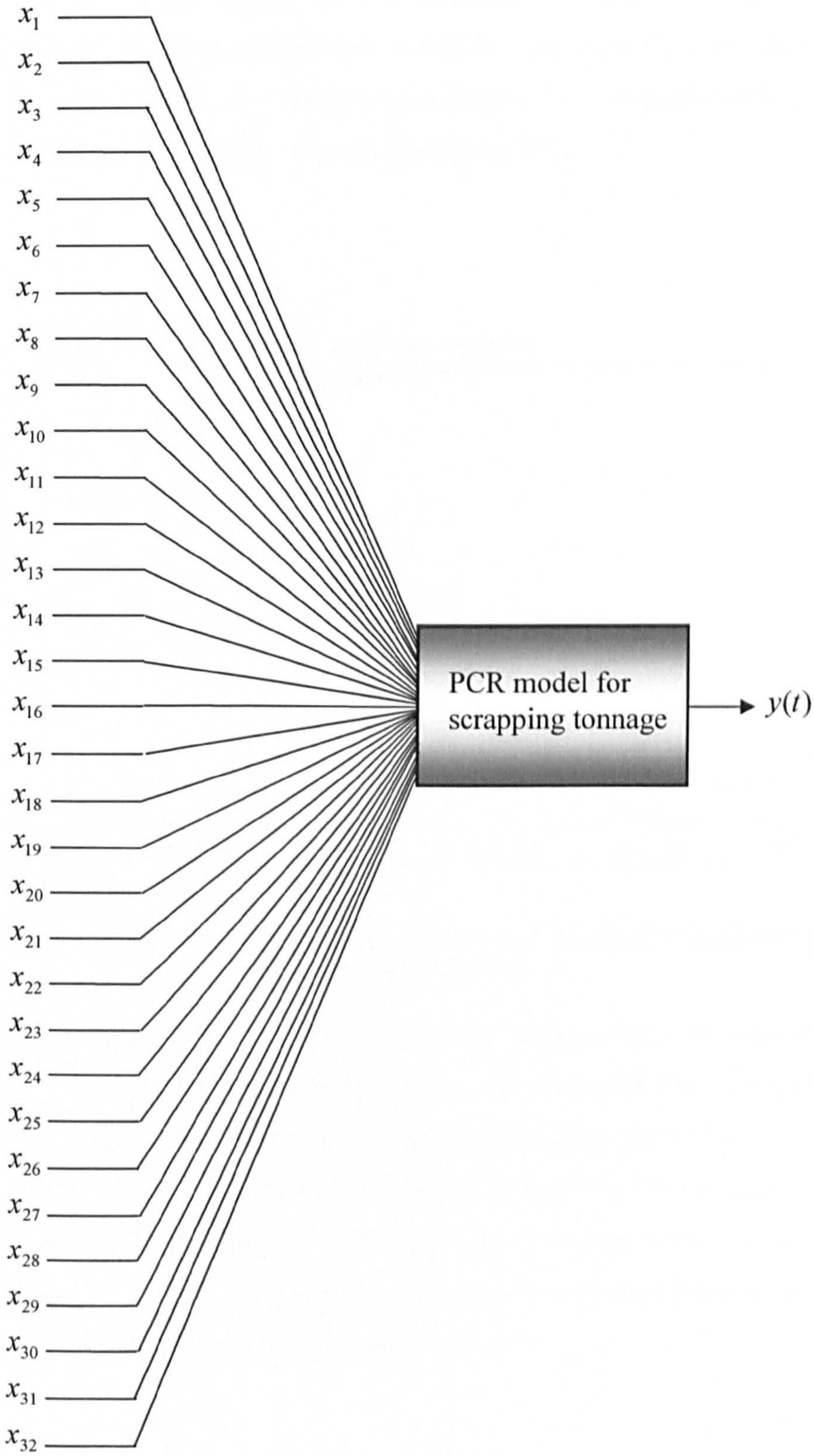


Figure 5-14: The PCR model for the monthly scrapped tonnage

The first step of this modelling is to find out the number of the processing components for the model. For this reason various numbers of PCs are considered to identify the most efficient combination. The variance curve (Figure 5-15) explains how much information is being modelled for various numbers of PCs.

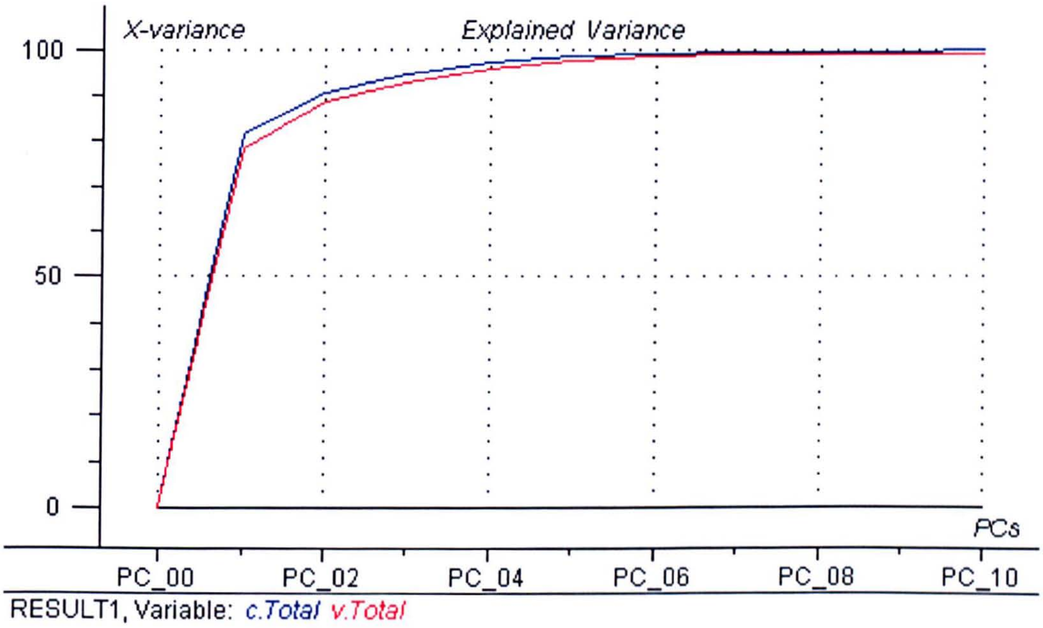


Figure 5-15: Variance curve for different numbers of PCs of the PCR model for the monthly scrapped tonnage

A model with one PC is able to explain 81.5% of the information in the data table and the model with four PCs explain about 96% of the information in the data table. The models including five PCs or more show saturation i.e. increasing 4 PCs to 5 PCs over-fits the model. Thus, a model including 4 PCs seems to be the most accurate choice and consequently its prediction will be the most accurate prediction amongst the others. Figure 5-16 shows the predicted tonnage of the scrapped tankers versus the actual measurements based on a PCR model with 4 PCs.

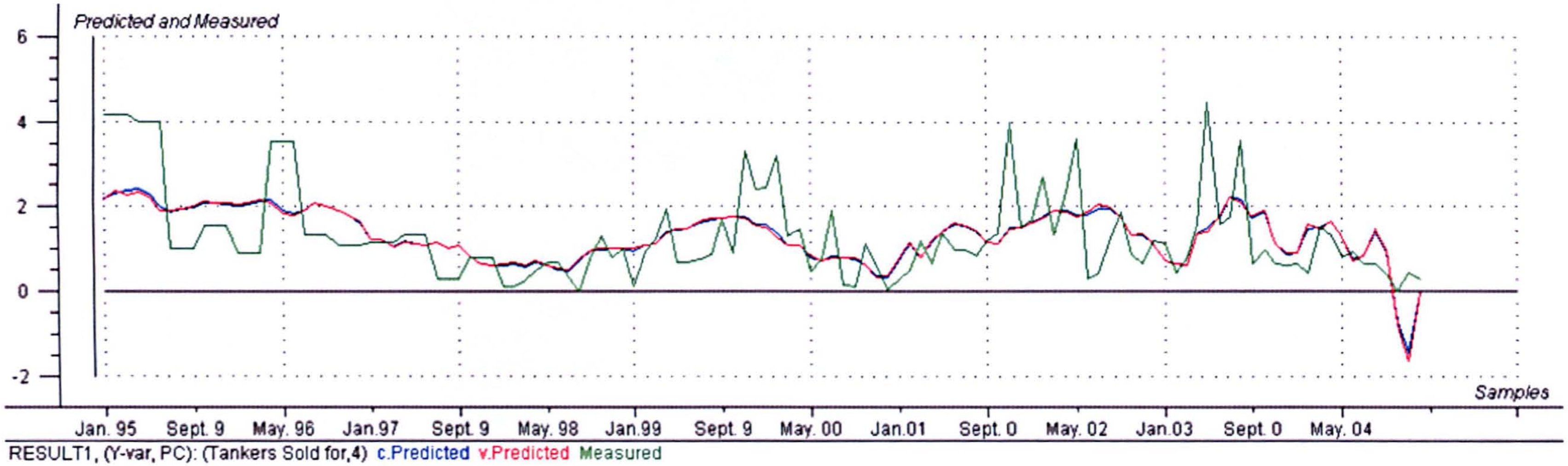


Figure 5-16: Monthly prediction of the scrapped tanker versus actual measurements based on PCR model with four Processing Components (PCs)

As explained in section 3-6-2, random Cross Validation (CV) method is used to simulate test set validation for this model. It is may be time consuming but it can assesses the stability of the PCR results. Validation curve is shown in red in Figure 5-16.

The average error of the model for different stages is shown in (Figure 5-17). This is a plot of the average prediction error, for either the calibration or the validation samples. In this plot the Root Mean Square Error of Calibration (RMSEC) and Root Mean Square Error of Prediction (RMSEP) plotted against the number of components used in the model. The average modelling error for a model with 4 PCs is 0.881 and the average prediction error that can expect for future prediction is 0.916.

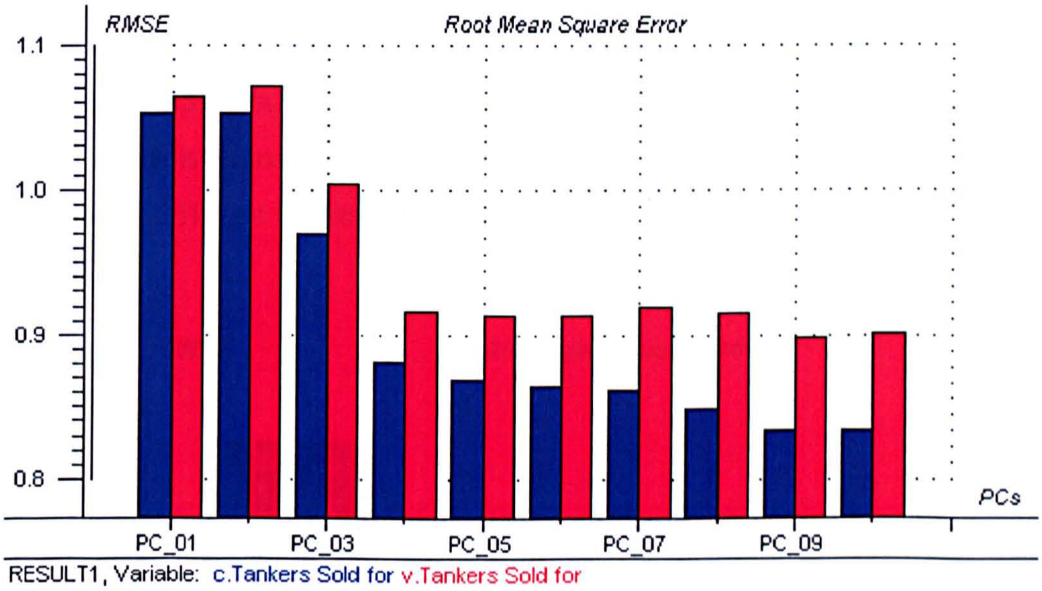


Figure 5-17: The RMSE for prediction with the model with different number of PCs.

The regression analysis can explain which *X*-variables are most important to predict the *Y*-variable for a model. The scrapped tanker tonnage is the only *Y*-variable of this model. The regression coefficients of all the X-variables for the above model are represented in Figure 5-18.

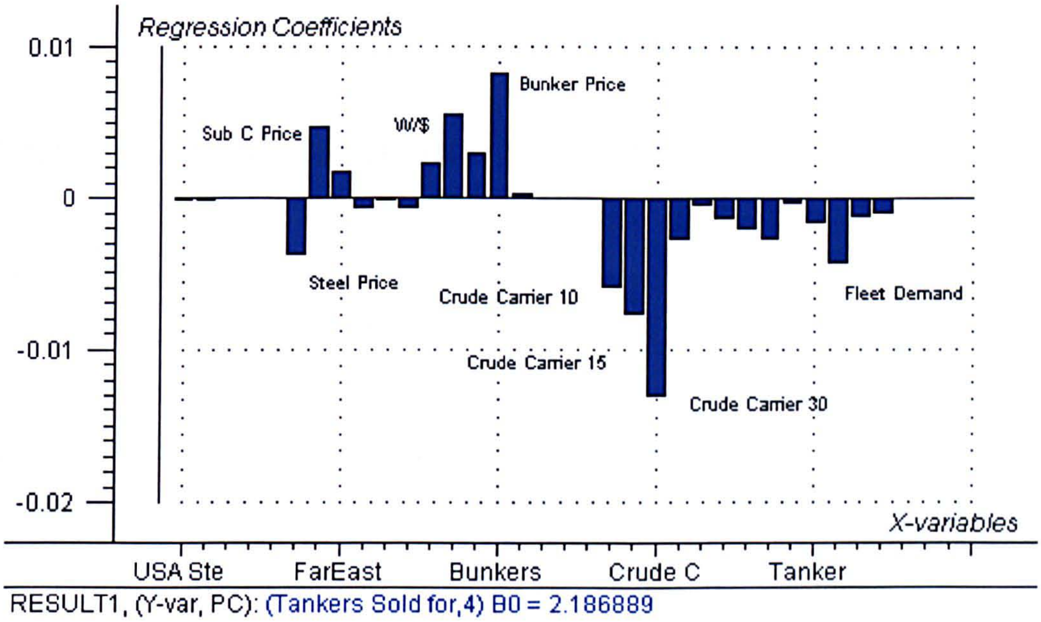


Figure 5-18: Regression Coefficient analysis of the scrapped tanker tonnage.

To increase the accuracy of the model using recent obtained results for regression coefficients, the model can be recalculated with only the *X*-variables which have the highest coefficients.

The new RMSE results (Figure 5-19) illustrate that the average modelling error is 0.894 which is almost the same for the recalculated model. It means that there is no significant different between the two models.

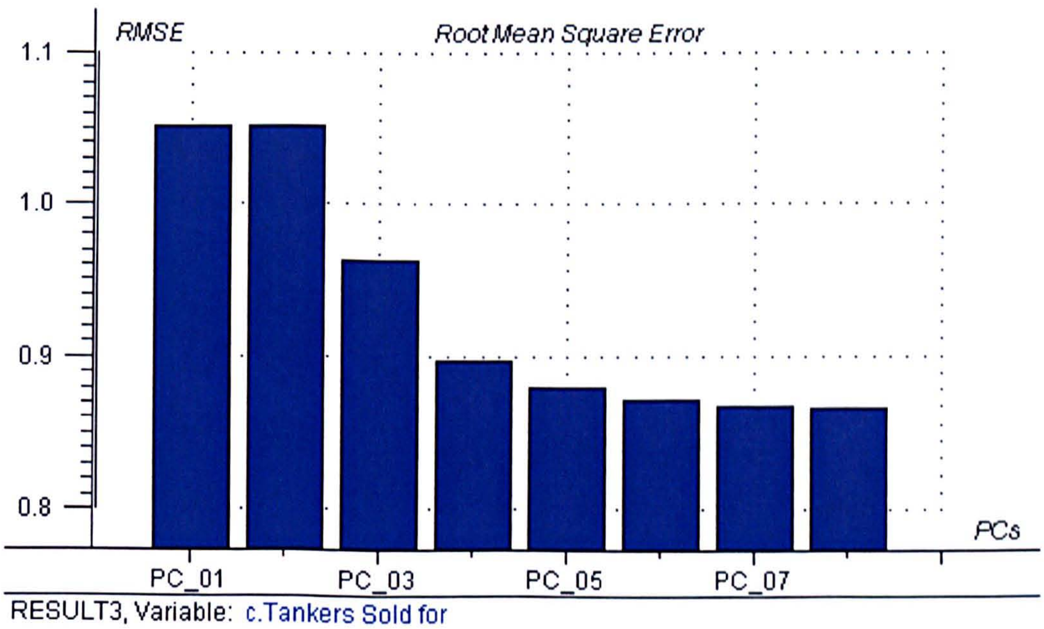


Figure 5-19: RMSE results for recalculated model

5.3.2.2 SCRAP PRICE PREDICTION USING PCR

To model the scrap prices, there are 31 *X*-variables and two *Y*-variables, Far-East and Subcontinent scrap prices, as illustrated in Figure 5-8. All the X and Y-variables are monthly data so they are functions of time. The fundamental of this modelling is similar to the previous method but this time PCR is used to model the prices. The *X* and *Y*-variables and their connections to the PCR model are represented in Figure 5-20.

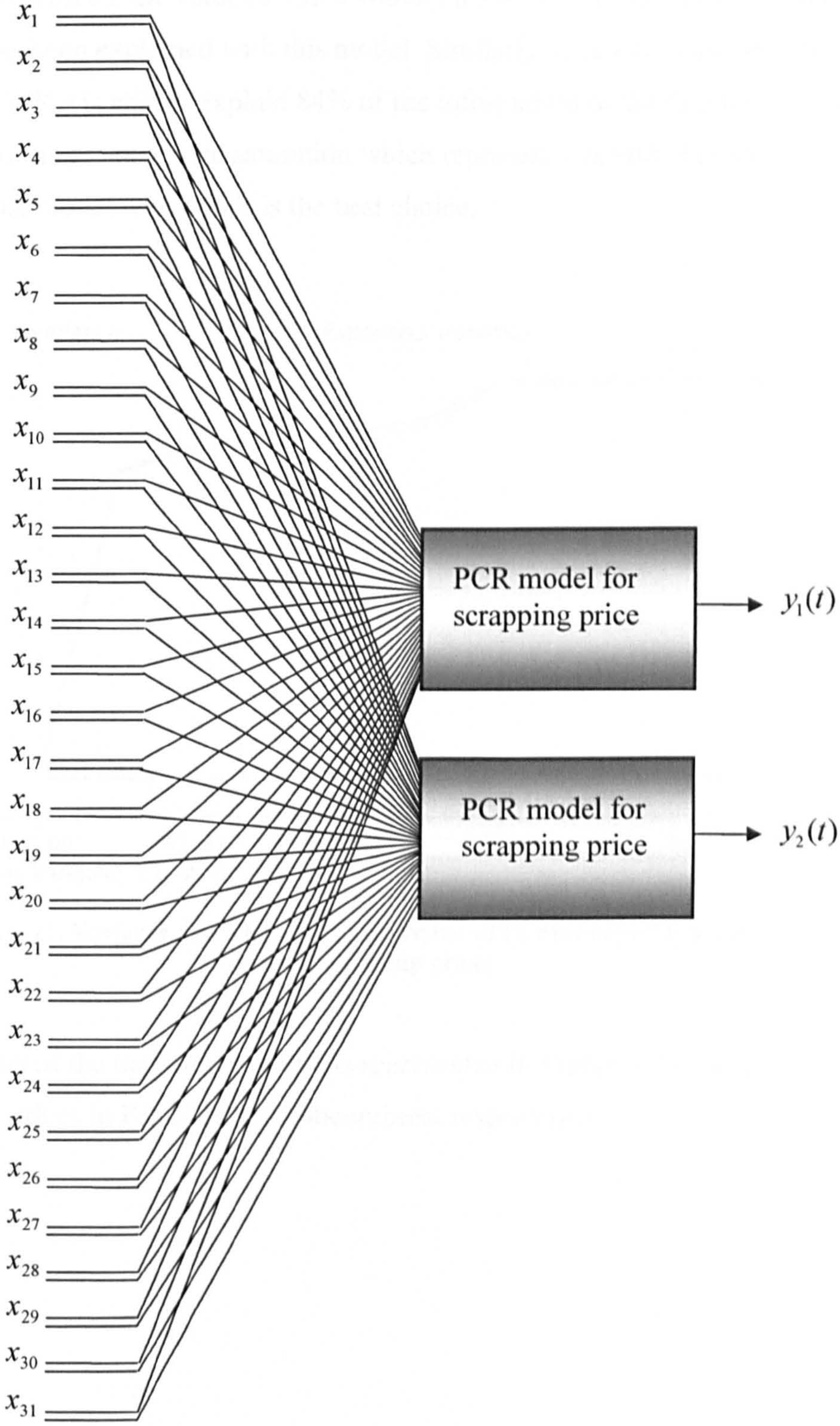


Figure 5-20: The PCR model structure for the monthly scrap prices in Subcontinent and Far-East

The first step is to distinguish the best combination of the PCs for this model. The variance curve is shown in Figure 5-21 which explains how much information is being modelled for various numbers of PCs. As it appears in this figure, the model with 6 PCs represents the value of 93.38 which means, 93% of the information in the data table has been explained with this model. Similarly, it can be found that the model with 5 PCs is able to explain 84% of the information in the data table. Models with 7 PCs or more are shown saturation which represent over-fitted models. Therefore, the model with 6 PCs is the best choice.

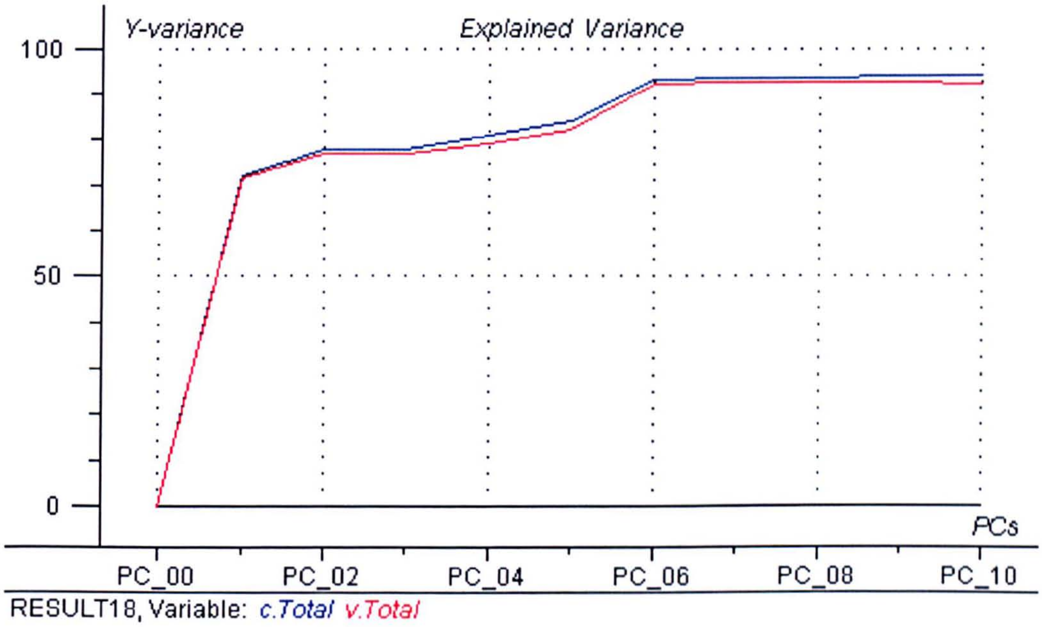


Figure 5-21: Variance curve for different number of PCs for the PCR model of the monthly scrap prices

The prediction of the mentioned model is represented in Figure 5-22 and Figure 5-23 for the scrap prices in Far-East and Subcontinent respectively.

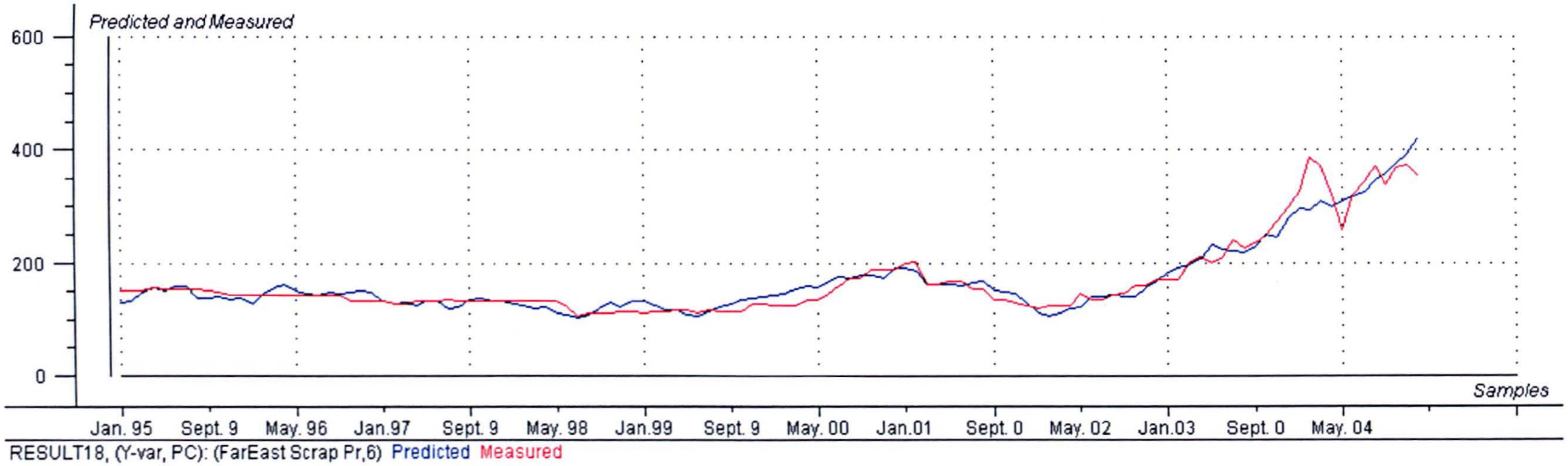


Figure 5-22: Far-East scrap price predictions based on the PCR model

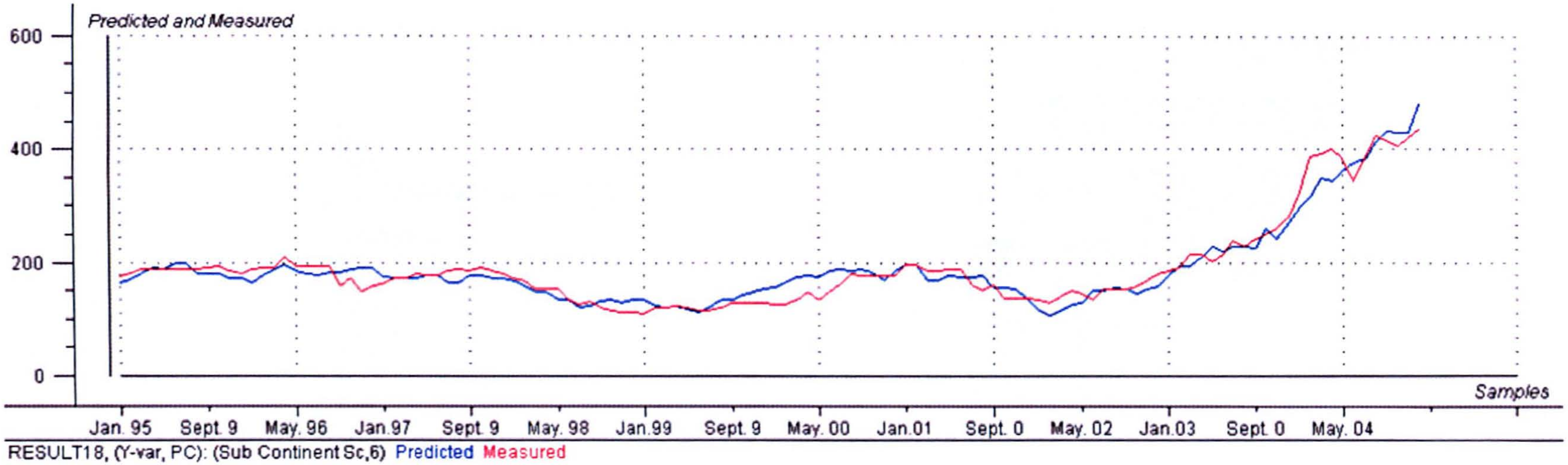


Figure 5-23: Subcontinent scrap price predictions based on the PCR model

To identify the average prediction error for both Y-variables in the above model, RMSE results are presented in Figure 5-24.

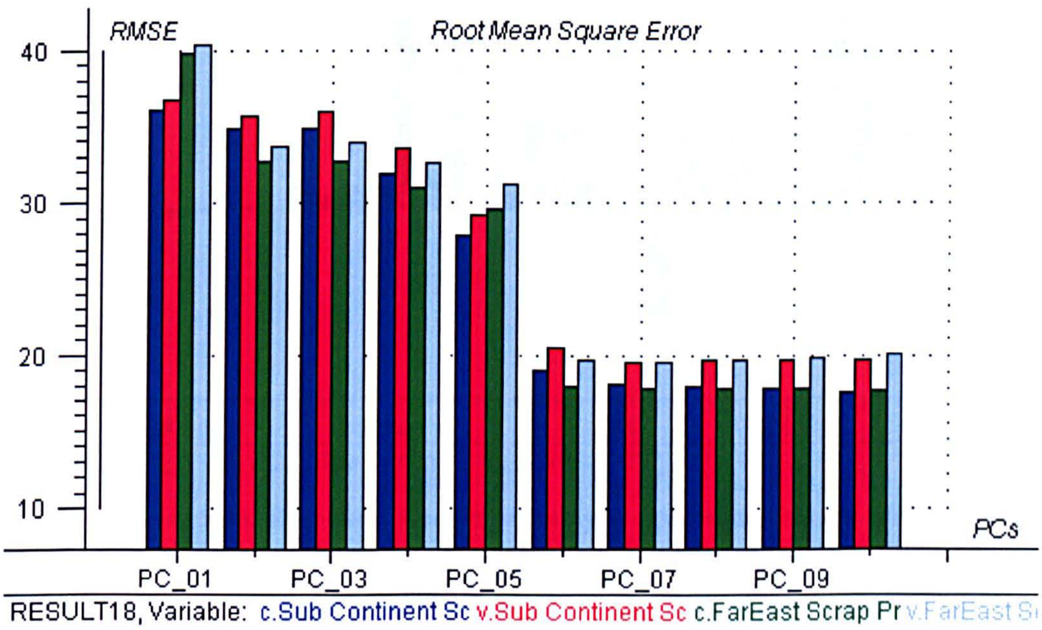


Figure 5-24: RMSE for the prediction of the PCR model with 6 PCs.

Regression analyses for both models are represented in Figure 5-25 and Figure 5-26.

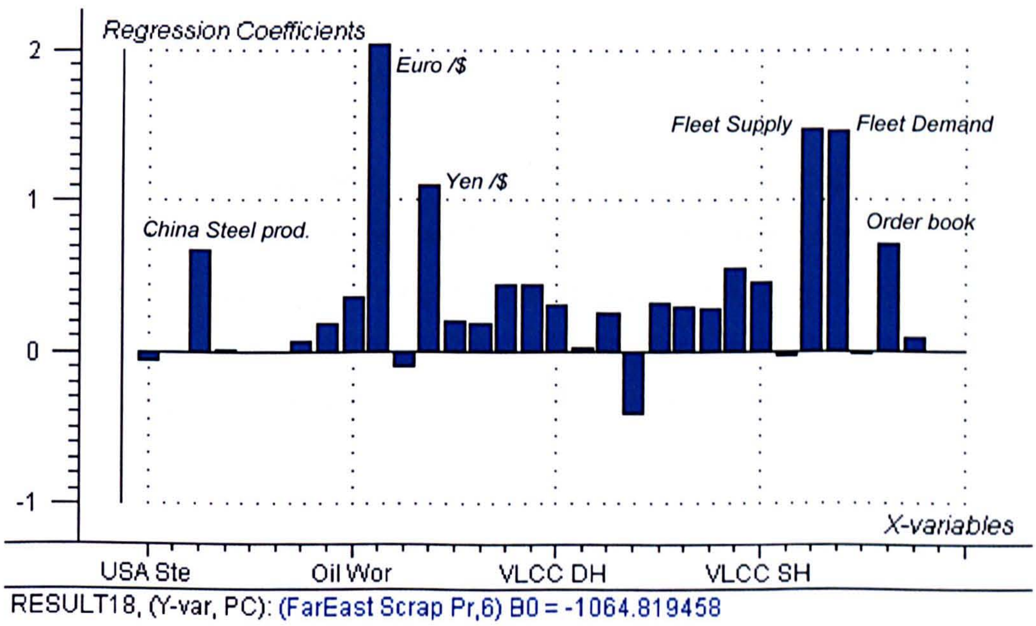


Figure 5-25: Regression Coefficient analysis of the scrap price in Far-East

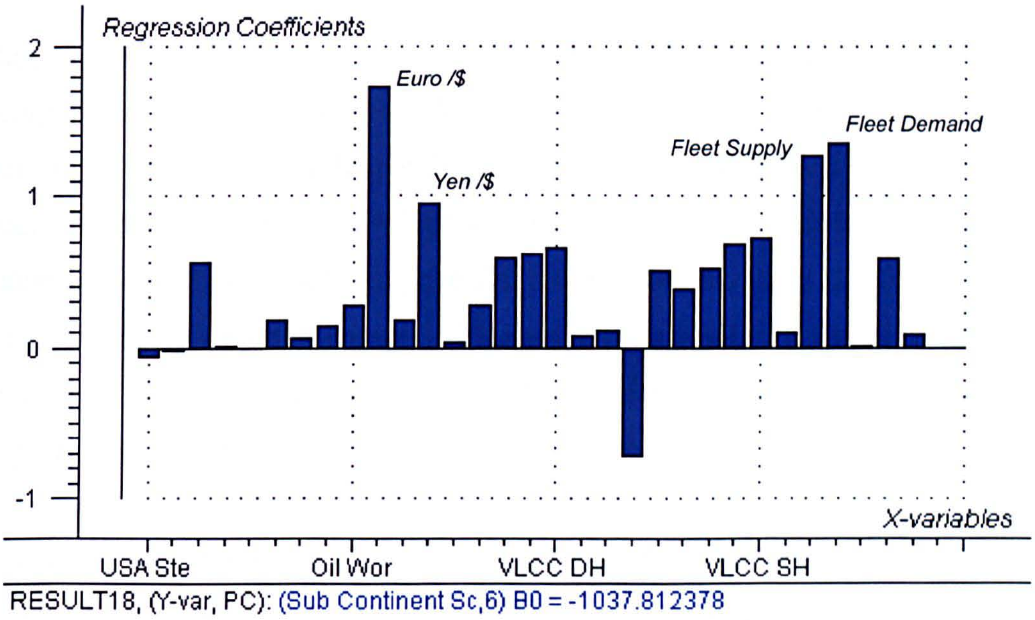


Figure 5-26: Regression Coefficient analysis of the scrap price in Subcontinent

5.3.2.3 PCR Modelling Results

5.3.2.3.1 Scrap Tonnage

For the PCR modelling, which is carried out for the monthly scrap tonnage, Figure 5-16 represents that the overall model is not able to analyse the variation of the data. Regression analysis of the obtained model illustrates which variables are most important to predict the *Y*-variable (Figure 5-18). The *X*-values with higher regression coefficients have more influence on the obtained PCR model, so this model is a function of mainly:

- Steel price
- Subcontinent scrap price
- Crude carrier 300k dwt freight rates
- Won/USD exchange rate

- Bunker price

Regression coefficients is also summarise the modelled relationships between scrapped tanker tonnage and each *X*-variable of this model. All the above variables have relatively higher coefficient than the others. Freight rates, in general, have the negative influences to the model. Crude carrier 300k dwt freight rate has the most value amongst the others. Steel price also shows a negative regression but its value is less than the freight rates. Bunker price has the most positive influence to the model and Won/USD exchange rate and Subcontinent scrap price have the second and third highest positive values.

The average modelling error (Figure 5-17) for the PCR model with 4 PCs is 0.881 and the average prediction error that can expect for future prediction is 0.916. It means error has increased for the prediction stage.

5.3.2.3.2 Scrap Prices

For the PCR modelling of the monthly scrap prices, the regression analysis is carried out separately for each location. Regression analysis of the model for the Far-East location (Figure 5-25) shows that the model is a function of mainly:

- Euro/USD exchange rate
- Yen/USD exchange rate
- Tanker fleet supply
- Tankers fleet demand
- China steel production
- Tankers order book

Similarly, for the Subcontinent's model (Figure 5-26) the model is a function of mainly:

- Euro/USD exchange rate

- Yen/USD exchange rate
- Tankers fleet supply
- Tankers fleet demand

Both the above PCR models represent that Euro/USD and Yen/USD exchange rates have the most positive influence to the model. It means that a change in these exchange rates can significantly change the scrap prices in both locations. Moreover, tankers fleet supply and demand are the next influential variables to the model. These results do not reflect the importance of some internal or external variables and do not include the variables which are the fundamentals of the demolition market. Hence, this model is not reliable to analyse the structure of the demolition market.

The average error of the model, for both the prediction and the modelling stages, is represented in Figure 5-24. For Subcontinent scrap prices the average modelling error is 19.04 and the average prediction error that can expect for future prediction is 20.53. For Far-East prices the average modelling error is 18.00 and the average prediction error is 19.72.

5.3.3 PLS

As explained in Chapter 3, there are two types of PLS modelling. In the following sections the PLS1 modelling is used to model the scrapped tanker tonnage and the PLS2 modelling is employed to model the scrap prices in Subcontinent and Far-East.

5.3.3.1 Scrap Tonnage Prediction Using PLS1

The *X* and *Y*-variables remain the same as previous modelling (Figure 5-3). It means that there are 32 *X*-variables and one *Y*-variable for this model. The structure of the model is same but the PLS1 method is used to model the data (Figure 5-27).

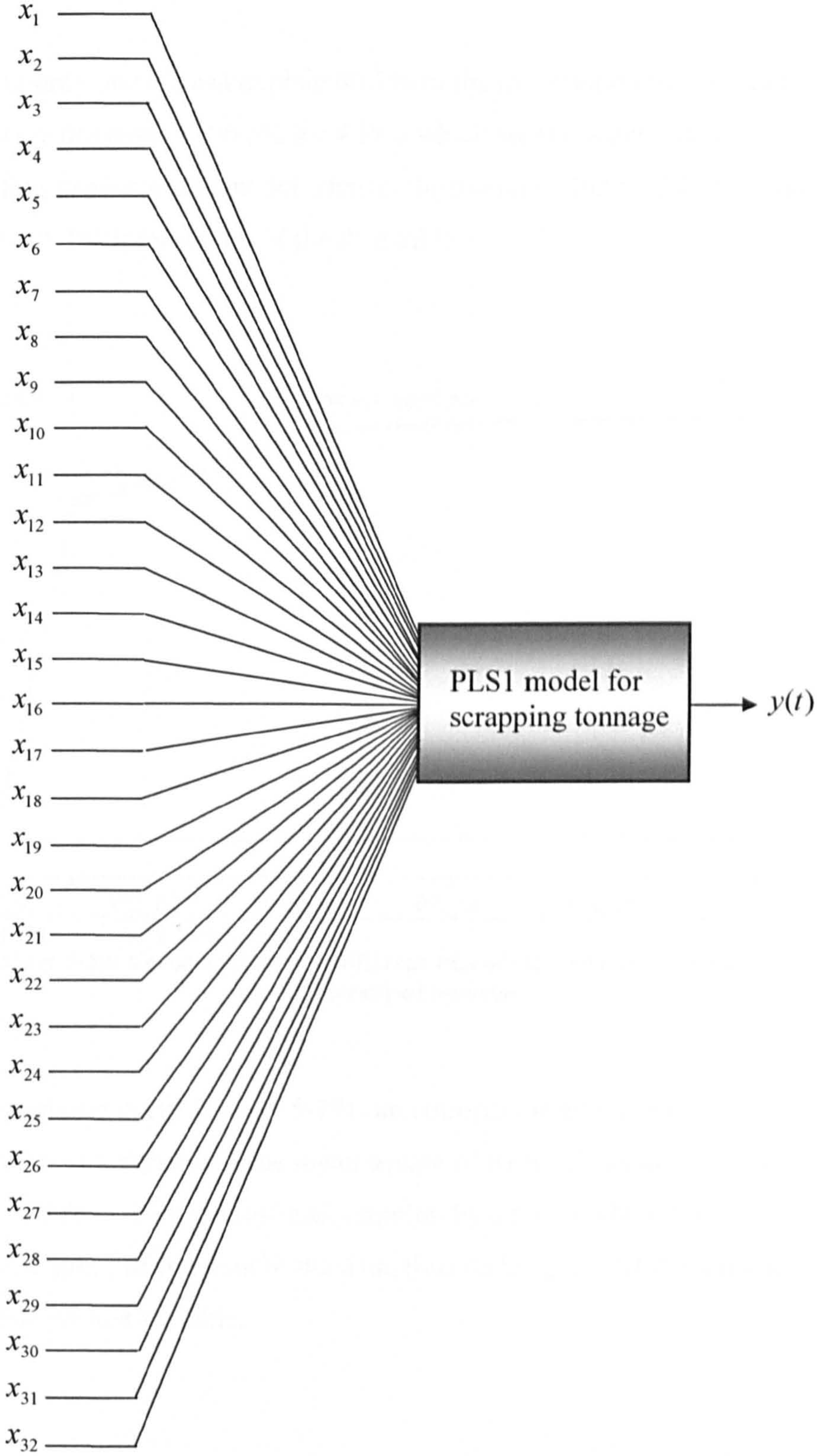


Figure 5-27: ThePLS1 model structure for the monthly scrapped tonnage

The first step of this modelling is to find out the best combination of PCs for the model. For this reason different number of PCs has been studied and the variance curve is prepared to distinguish the best model (Figure 5-28).

A model with only one PC can explain 80.7% of the information in the data table. This percentage increases to 96.5% for 4 PCs which shows better data explanation, but it is leading to over fit the model. Hence, the model including 3 PCs is chosen which covers 91.1% information of the data table.

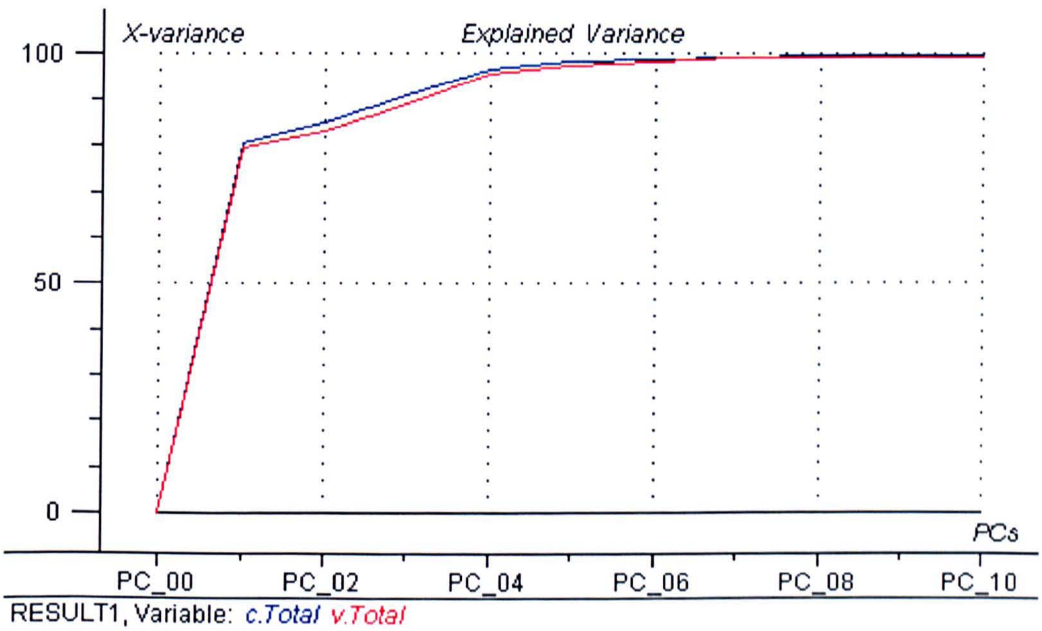


Figure 5-28: Variance curve for different PCs of the PLS1 model for the monthly scrapped tonnage

The residual variance curve (Figure 5-29) can confirm the above decision. The residual variance of a variable is the mean square of its residuals for all model components. It differs from the residual variation by a factor which takes into account the remaining degrees of freedom in the data, thus making it a valid expression of the modelling error for that variable.

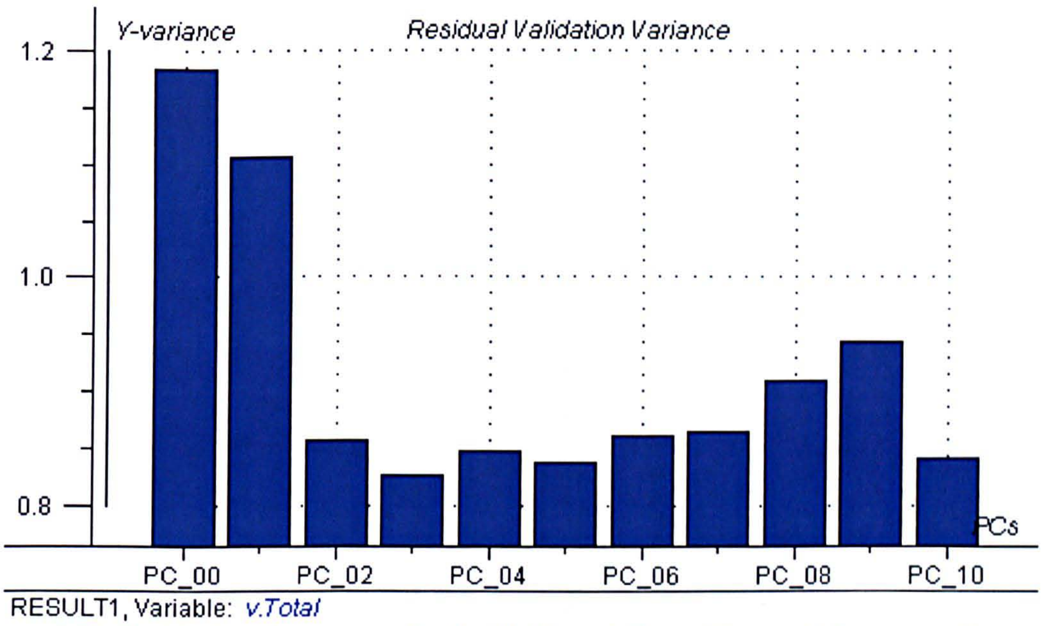


Figure 5-29: Residual variance curve for the PLS1 modelling of the monthly scrapped tonnage

The corresponding residual value for the model including 3 PCs is 0.83 which shows a smaller value compared with the model including 4 PCs with 0.85. Moreover, 3 PCs show the minimum value of the residual variance amongst all the other models.

The prediction of this model for scrapped tankers is shown in Figure 5-30 besides the actual measurements.

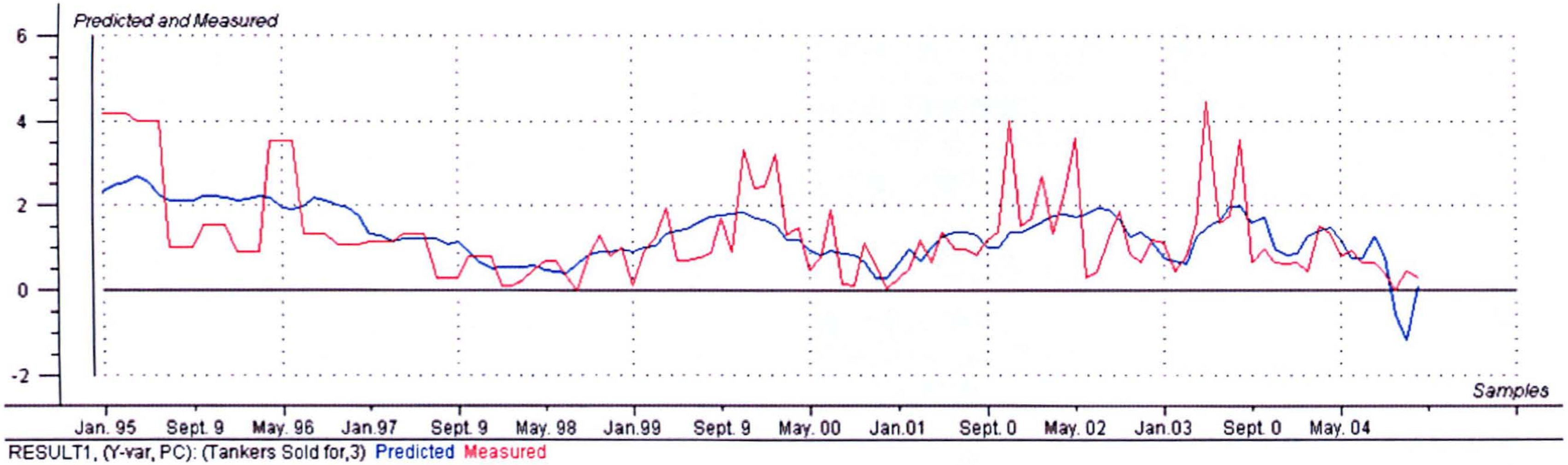


Figure 5-30: Monthly prediction of the scrapped tanker versus actual measurements based on PLS1 modelling.

Root Mean Square Error is used to discover the error of the above PLS model for the monthly scrapped tonnage. The average error of the model for different stages is shown in (Figure 5-31).

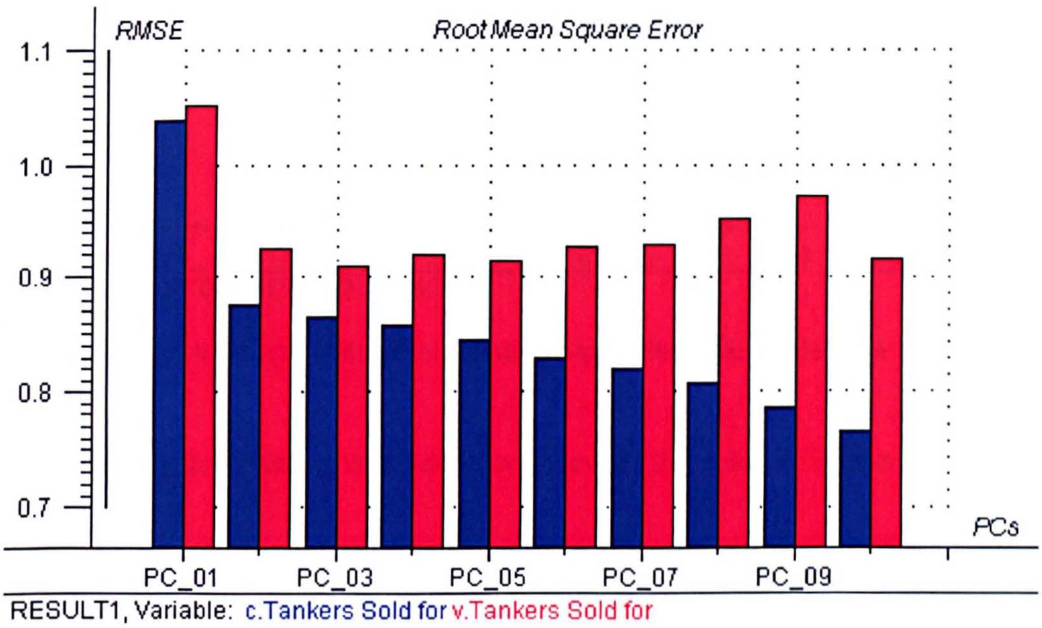


Figure 5-31: Root Mean Square Error of the PSL1model for the scrapped tanker tonnage

As mentioned earlier in Chapter 3, u-scores are the scores found by PLS in the *Y*-matrix and t-scores are the scores found by PLS (and also in PCA and PCR) in the *X*-matrix. Also, as explained in section 3-6-3, the relationship between t- and u-scores is a summary of the relationship between *X* and *Y* along a specific model component. Figure 5-32 shows u- and t-scores relationship for the 3 PCs model. The samples should lie as close to each other as possible along a straight line through most of the samples. Samples that stick out from this line are possible outliers. To interpret this plot a regression line is also drawn between the data points using the least squares algorithm. The slope of the line is 0.005 and the offset (or intercept) is 8.9e-8. The value of the correlation for the samples in this model is 0.15 which is quite small.

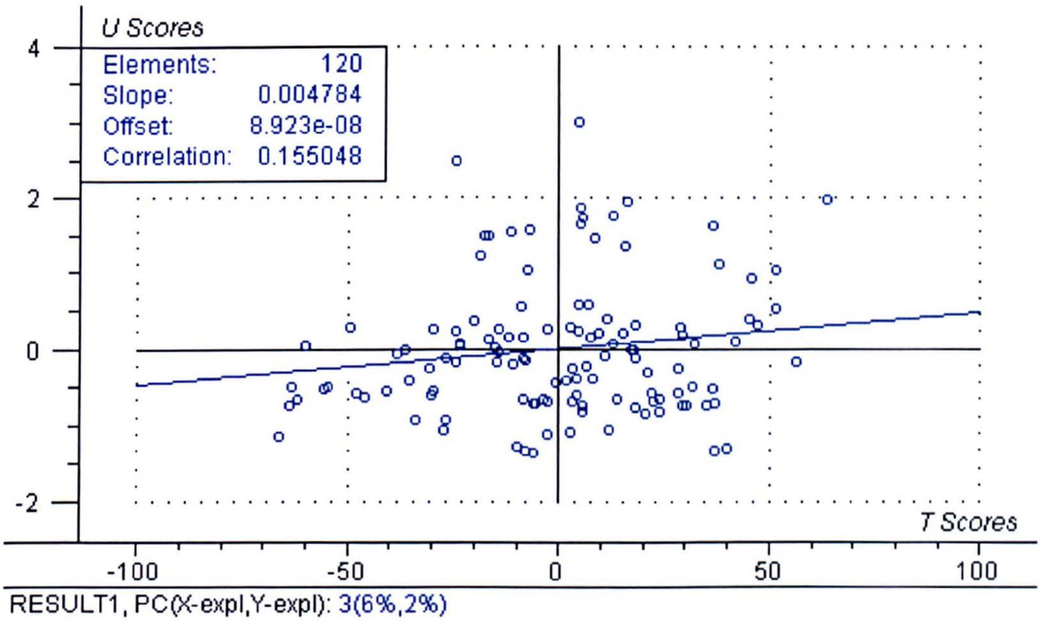


Figure 5-32: u- and t-scores relationship for the PLS1 model with 3 PCs

The Y-residual curve can give a better overview of the possible outliers (Figure 5-33).

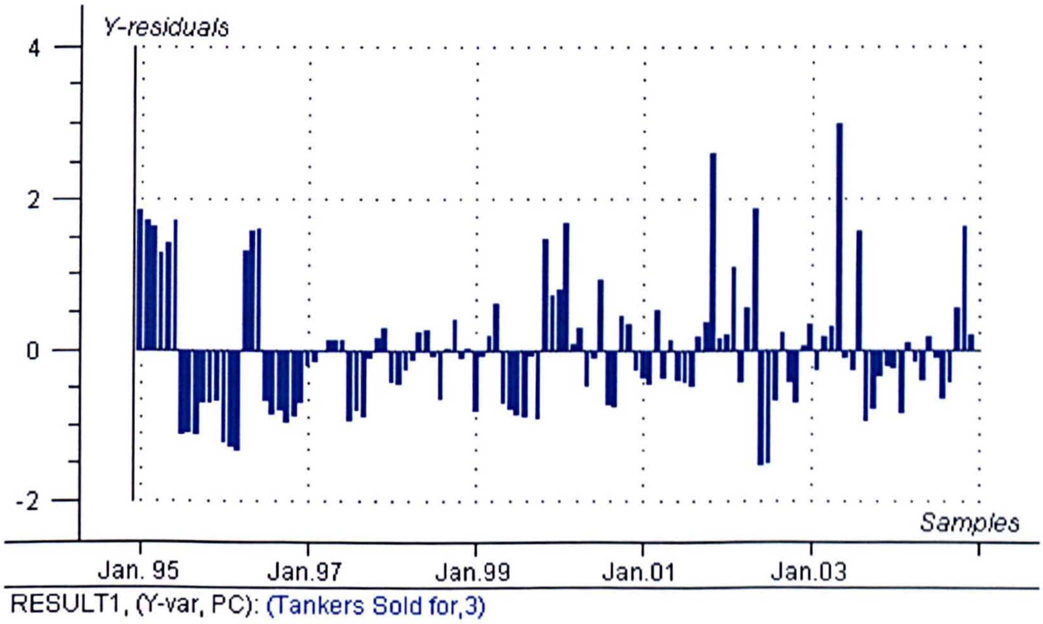


Figure 5-33: Y-residuals of the PLS1 model for all the samples

To indicate the sensitivity of different X-variables for this particular model, regression coefficients are shown in Figure 5-34.

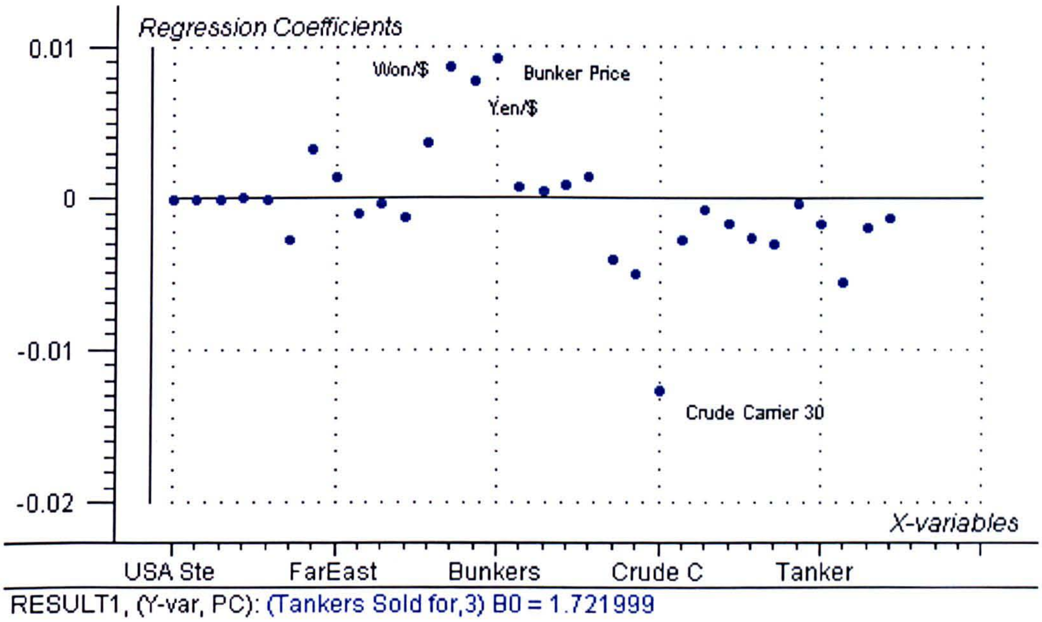


Figure 5-34: Regression coefficient analysis of the scrapped tanker model

5.3.3.2 Scrap Price Prediction Using PLS2

PLS2 handles several responses simultaneously so in this modelling a model is built for two *Y*-variables, i.e. Far-East and Subcontinent scrap prices, simultaneously (Figure 5-8). The structure of this modelling is illustrated in Figure 5-35. To find out the most effective model for the sake of this study, different combination of principal components are implemented and subsequently the variance of each combination is measured (Figure 5-36).

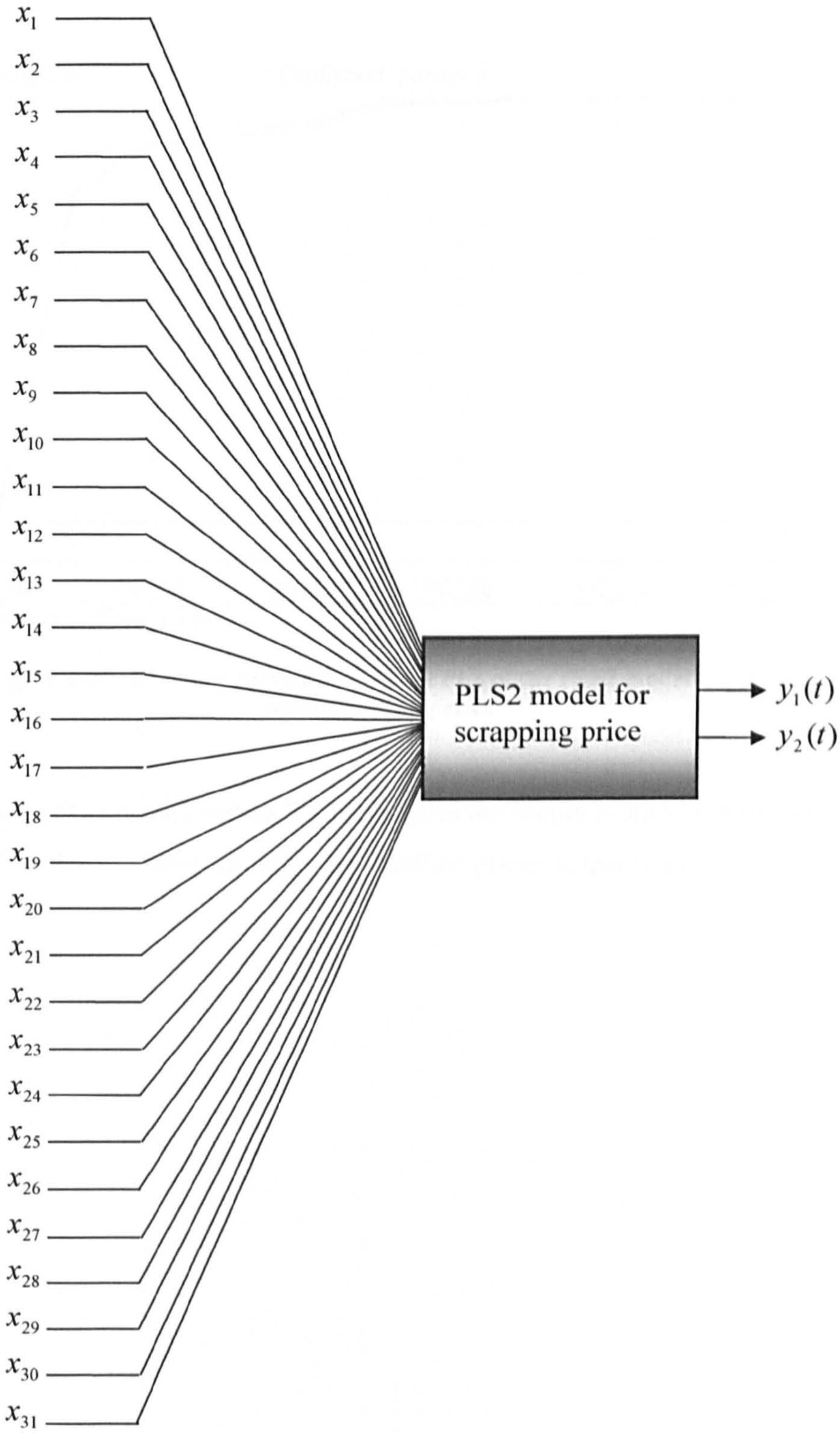


Figure 5-35: The PLS2 model structure for the monthly scrap prices

As represented in the variance curve, a model includes 4 PCs shows the best sample variance with the value of 93.38. It means that this model is able to explain 93.4% of the information of the data table. The models with 5 PCs, or more, clearly show the over-fit for the correspondent model.

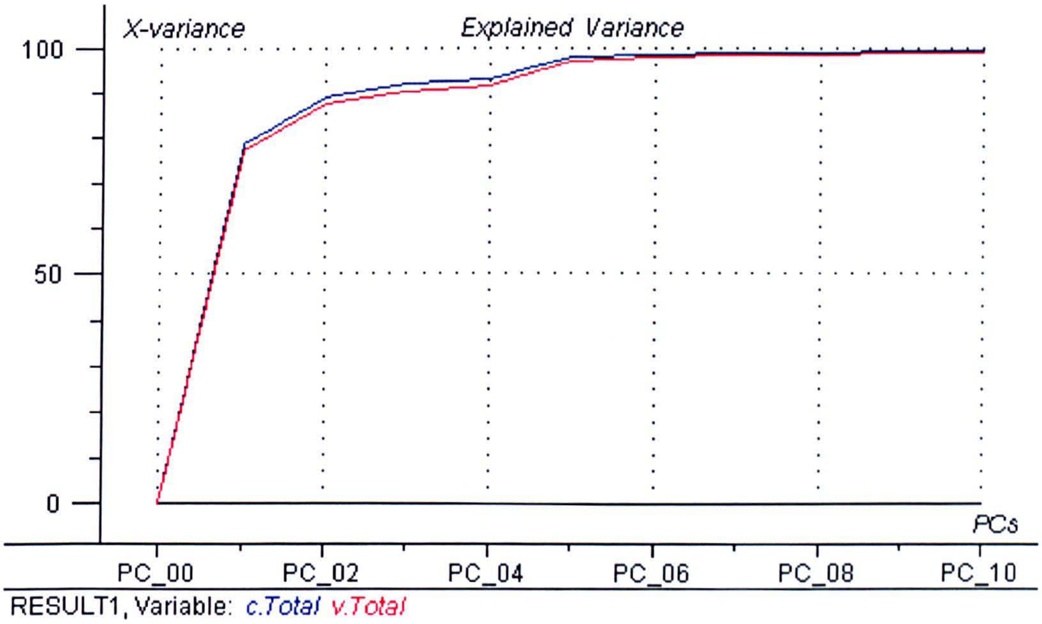


Figure 5-36: Variance curve for different PCs of the PLS2 model for the monthly scrap prices

The prediction of the model versus the actual measurements is shown in Figure 5-37 and Figure 5-38 for the Subcontinent and Far-East prices respectively.

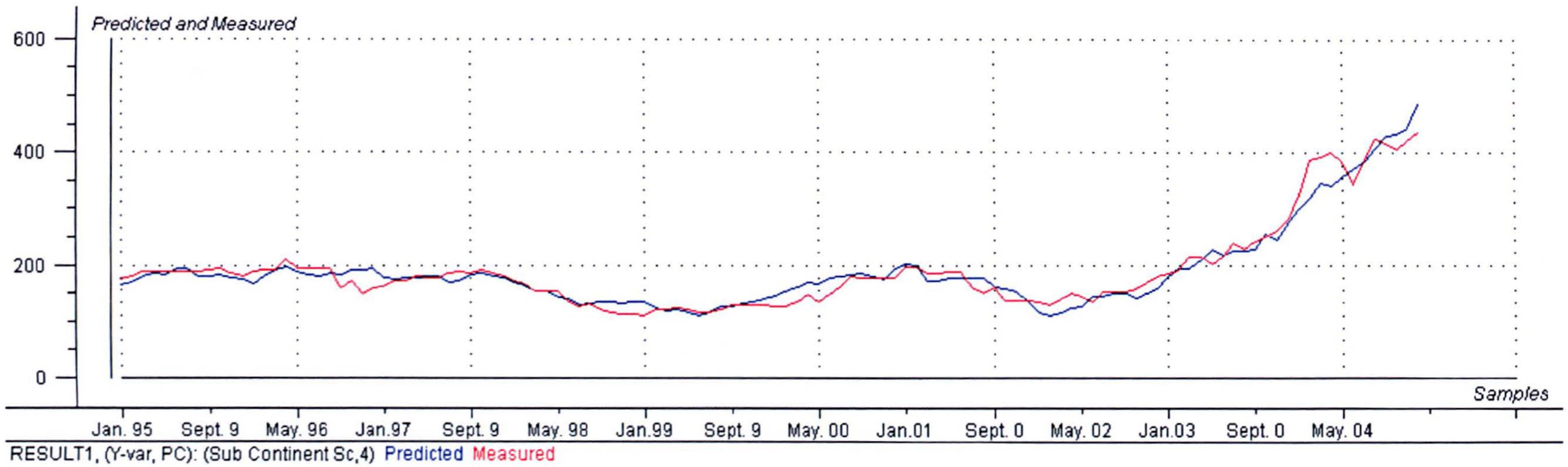


Figure 5-37: PLS2 model predictions for the Subcontinent scrap prices

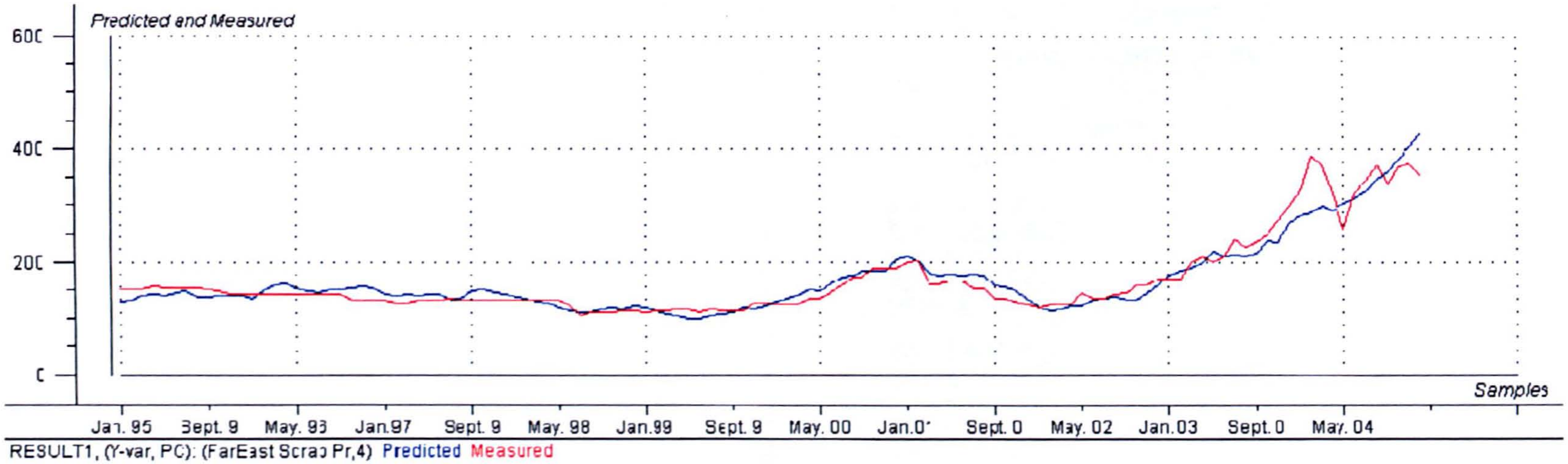


Figure 5-38: PLS2 model predictions for the Far-East scrap prices

RMSE values of the above PLS2 model are represented in Figure 5-39. It includes the average prediction error for the model for either the calibration or the validation samples.

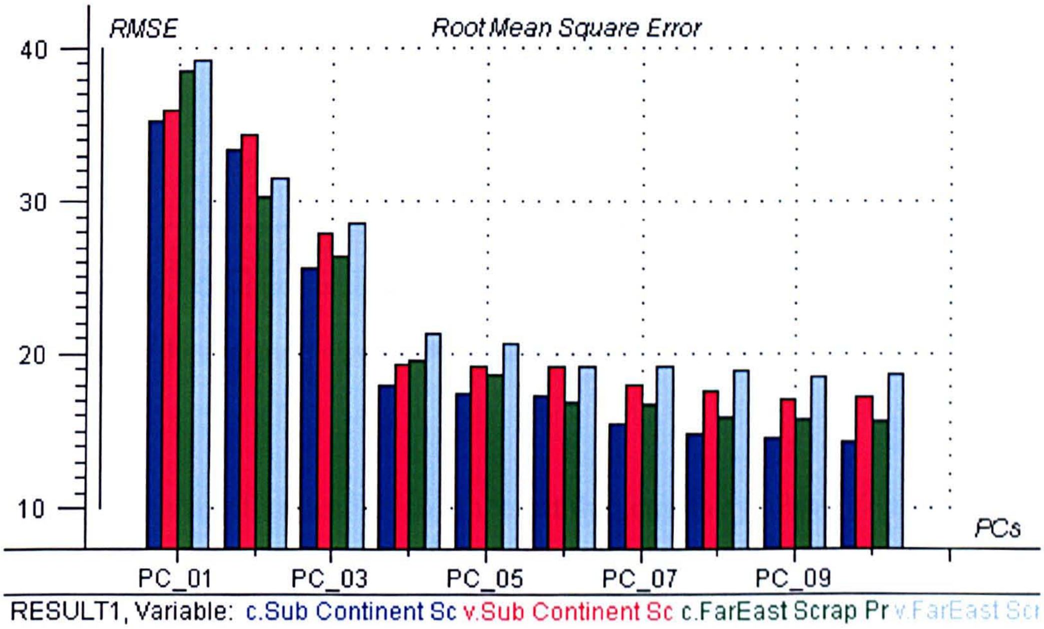


Figure 5-39: Root Mean Square Error for the prediction of the model

The u-scores and t-scores relationship for the above 4 PCs model, based on PLS2 method, is represented in Figure 5-40.

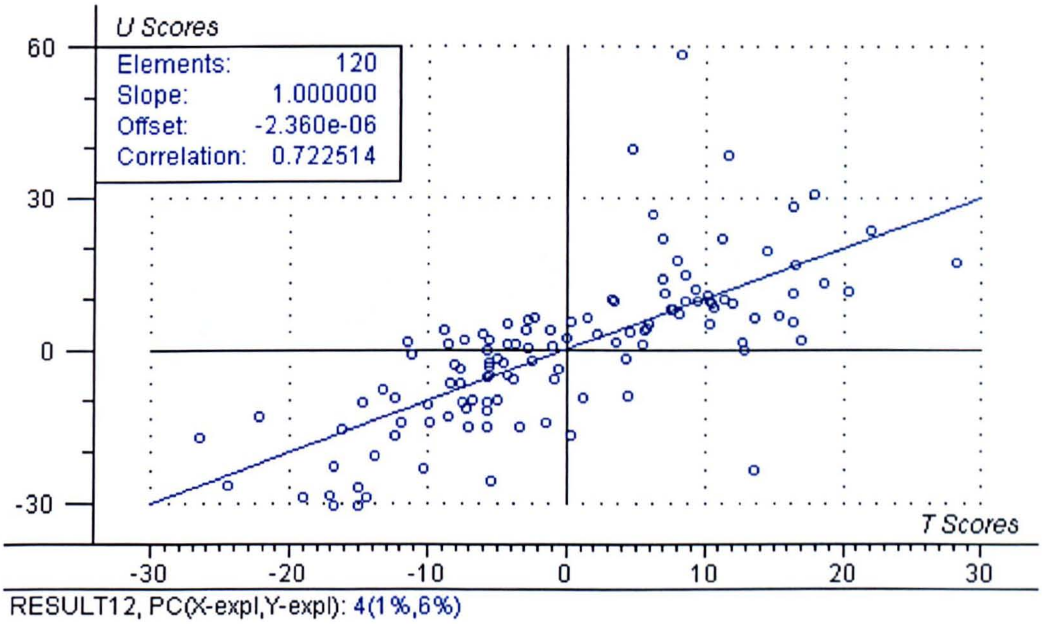


Figure 5-40: u- and t-scores relationship for the model based on PLS2 with 4 PCs

The regression coefficients, which represent the sensitivity of the *X*-variables, are shown in Figure 5-41 and Figure 5-42 for the two *Y*-variables.

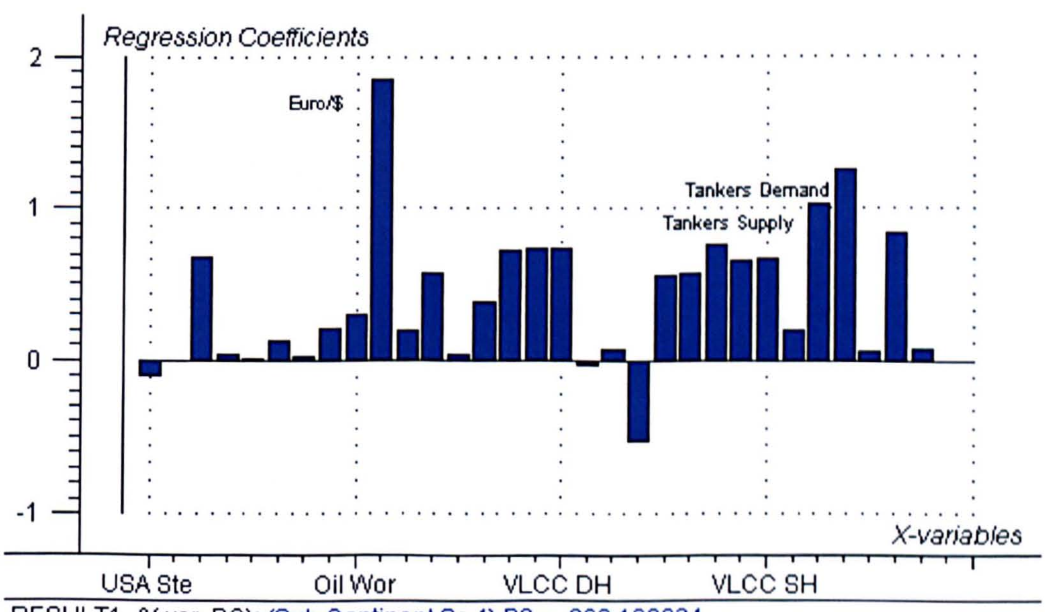


Figure 5-41: Regression Coefficient of the Subcontinent PLS2 model

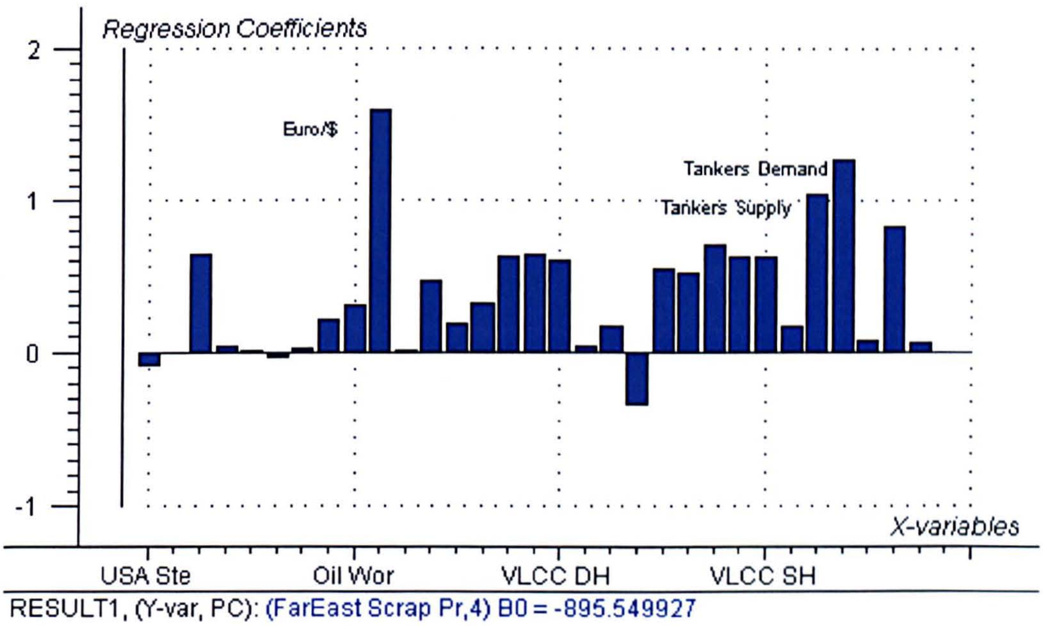


Figure 5-42: Regression Coefficient of the Far-East PLS2 model

The *Y*-residuals plot for the above model is shown in Figure 5-16.

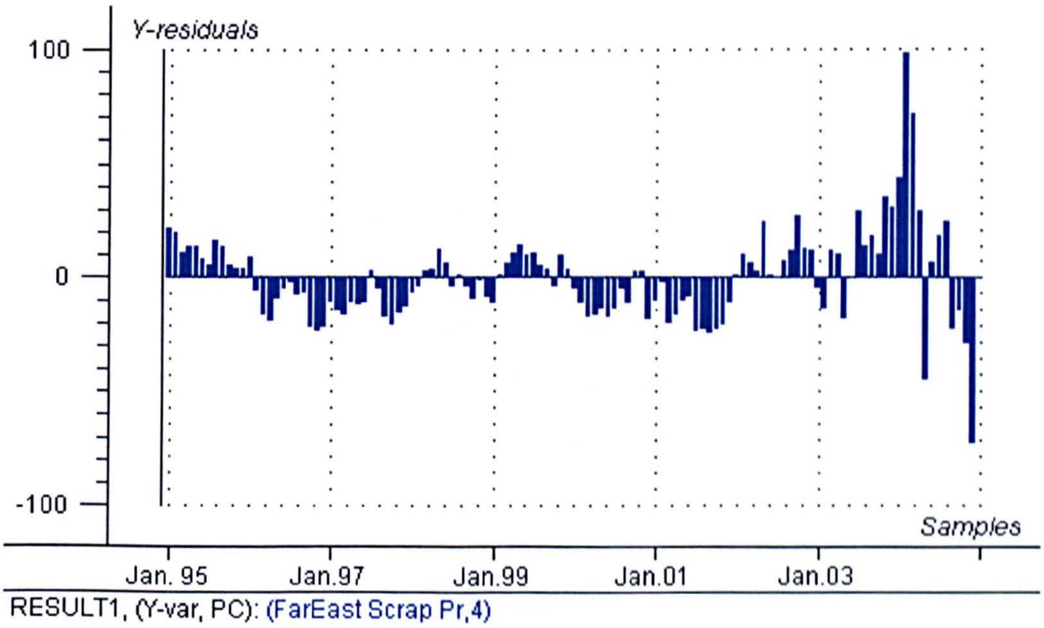


Figure 5-43: Residuals of the PLS2 model

5.3.3.3 PLS Modelling Results

5.3.3.3.1 Scrap Tonnage

For the PLS1 modelling of the monthly scrapped tonnage, as it appears in Figure 5-30, the obtained model is not able to analyse the variation in the data accurately. According to Figure 5-29, the average modelling error is 0.86 and the average prediction error that can expect for future prediction is 0.91. Regression analysis of the model (Figure 5-34) shows that below variables have the highest sensitivities for the obtained model:

- Crude carrier 300k dwt freight rate
- Bunker price
- Won/USD exchange rate
- Yen/USD exchange rate

Crude carrier 300k dwt freight rate have the highest regression value amongst the others. It has a negative regression but the others have positive regressions. Bunker price has the maximum value of positive regression to the model. The influence of the steel production and steel price is negligible (near zero) in the obtained model. Hence, the relation between the scrap steel (obtained from the demolished tankers) and steel production does not take into account in this model which is not sensible.

Figure 5-32 represents that the modelling is not effective as the correlation is 0.155. There are also a few outliers for the model. Figure 5-33 represents the *Y*-residuals of the model and confirms the existence of the outliers e.g. May 2003 or November 2001.

5.3.3.3.2 SCRAP PRICES

For the PLS2 modelling for the scrap prices, Figure 5-39 illustrates that the average modelling error is 17.89 and the average prediction error that can expect for future prediction is 19.30 for the Subcontinent scrap prices. Similarly, for Far-East scrap prices the average modelling error is 19.60 and the average prediction error is 21.26. Hence, there are no significant differences in prediction between the two variables in the model. In addition, the error in prediction mode is higher than the modelling stage for both variables.

Figure 5-40 illustrates t and u scores relationships. The slope of the regression line and the intercept is measured by is 1.0 and -2.36e-06 respectively. The value of the correlation for all the samples is 0.72 which represents more accurate modelling compared with the previous PLS1 model (for scrap tonnage). According to this plot the samples in February, March, April and December 2004 are outliers of the model. Regression analysis of the model for both locations, Figure 5-41 and Figure 5-42, show that the below variables have the most values amongst the other X-variables:

- Euro/USD exchange rate
- Tankers fleet demand
- Tankers fleet supply
- Tankers order book

Based on the obtained model, Euro/USD exchange rate has the most influence to the scrap prices. Tankers fleet demand and supply have the second and third highest regressions. This model does not take into account of the other fundamental variables to the demolition market.

The Y-residuals plot of this model (Figure 5-43) explains that the misfit rate of the PLS2 model is getting bigger in recent data i.e. 2004. This is just because of the higher variation of the data in this period compared with the previous years.

5.4 DISCUSSION AND COMPARISON BETWEEN THE THREE METHODS

Three different modelling methods (MLR, PCR, PLS) were used to investigate the demolition market and their prediction abilities for two different purposes, monthly scrapped tonnage and scrap prices, are tested. In this section the three discrete methods, which discussed independently earlier, will be compared.

To investigate accuracy of the modelling for the examined methods, Root Mean Square Error (RMSE) of the modelling and standard error of the prediction for each individual model has been calculated. Standard error of prediction is computed as the standard deviation of the residuals. The RMSE values show the average uncertainty that can be expected when predicting *Y*-values for new samples, expressed in the same units as the *Y*-variable. The results of future predictions can then be presented as: Predicted values ± 2 * RMSE. This measure is valid provided that the new samples are similar to the ones used for calibration, otherwise, the prediction error might be much higher.

The performances of the three modelling methods for the monthly scrapped tonnage are represented in Table 5-6. The modelling error for MLR is lower than the others with 0.83 (million dwt) which shows more accurate modelling than PCR and PLS. Standard error for MLR is also shown a smaller value than the others which confirms the RMSE results. The PCR method is slightly more accurate than the PLS method in this case.

Modelling Method	RMSE	Stde	Key Variables
MLR	0.83	0.94	South Korean steel production Oil world trade Non-OPEC oil production EU steel production
PCR	0.90	0.97	Steel price Subcontinent scrap price Crude carrier 300k dwt freight rate Won/USD exchange rate Bunker price
PLS	0.92	0.98	Crude carrier 300k dw freight rate Bunker price Won/USD exchange rate Yen/USD exchange rate

Table 5-6: Comparison of the three modelling methods for the monthly scrapped tonnage

Table 5-6 is also represents the key variables in each of the modelling methods. As the MLR model is performed more accurate than the others, this suggests that its key variables are more reliable. As explained in section 5-3-3-3-1, key variables in the PLS model do not include some fundamental variables of the demolition market e.g. steel production or steel price. Hence, the error of the PLS model is greater than the others.

Similarly, the performances of the three modelling methods for the monthly scrap prices are represented in Table 5-7. For the scrap prices in Subcontinent RMSE comparison shows more accuracy for the PLS method with 19.30. For the Far-East prices the PCR method shows more accuracy with 19.72. Therefore, the PLS model is slightly performs better than the PCR in Subcontinent prices but PCR model performs better for the Far-East prices.

	Modelling Method	RMSE	Stde	Key Variables
Subcontinent				
	MLR	20.50	28.79	S Korean steel Production Japan steel production Aframax DH building price China steel production USA steel production
	PCR	20.53	29.59	Euro/USD exchange rate Yen/USD exchange rate Tankers fleet supply Tankers fleet demand
	PLS	19.30	25.26	Euro/USD exchange rate Tankers fleet demand Tankers fleet supply Tankers order book
Far-East				
	MLR	22.75	26.75	South Korean steel production Japan steel production Oil world trade Suezmax DH building price
	PCR	19.72	24.51	Euro/USD exchange rate Yen/USD exchange rate Tanker fleet supply Tankers fleet demand China steel production Tankers order book
	PLS	21.26	25.01	Euro/USD exchange rate Tankers fleet demand Tankers fleet supply Tankers order book

Table 5-7: Comparison of the three modelling methods for the scrapping prices in Subcontinent and Far-East

As represented in Table 5-7, the most influential variables to the model for both PCR and PLS methods, for both locations, are Euro/USD exchange rate. This can be true but the decision making of the model is based on a variable which, in reality, can not influence the demolition market that much. Exchange rates can influence the prices but not as the main influential variable to the model. The combination of the key variables for the MLR method is better than the other two but the modelling error is greater this time.

CHAPTER 6

ARTIFICIAL NEURAL NETWORKS MODELLING OF THE DEMOLITION MARKET

6.1 INTRODUCTION

In the previous chapter the demolition market was analysed and investigated using statistical methodology. Different statistical multivariate methods were implemented, tested and subsequently compared with each other to evaluate each method.

The aim of this chapter is to re-evaluate the modelling and the issues identified in the previous chapter, concerning the demolition market, using Artificial Neural Networks methodology. Tonnage of the monthly scrapped ships and their prices for two locations are favourable for this study. To model each of these parameters, various architectures are designed and their specifications identified. Subsequently, ANNs are trained and their performances tested to find out the best neural networks layouts to model the market. Both static and dynamic networks are created for each study, using NeuroSolutions 5.0 software, in order to analyse the model inputs and build the most accurate model for the demolition market. Then the prediction abilities of each ANN are measured by forecasting three months ahead of the market and compared with each other. The main objectives of the above procedures are summarised below:

- Model the monthly scrapped tonnage using ANNs
- Model the monthly scrap prices for Subcontinent and Far-East scrap yards
- Identify the most influential inputs to the above models
- Use the obtained models to forecast the market

6.2 DEVELOPMENT OF THE ANN MODELS FOR SCRAPPED TONNAGE

As mentioned in the last chapter, the data, which are used for this study, are monthly data from January 1995 till December 2004 (120 patterns). Two different types of inputs are considered for this study, Internal and External. Internal inputs are variables which are applicable to the tanker market. External inputs are variables which may influence the tanker market but they are not in the tanker market itself like oil production, steel production, steel price and exchange rates. All the inputs are listed below:

$x_1(t)$: USA Steel Production
$x_2(t)$: EU Steel Production
$x_3(t)$: China Steel Production
$x_4(t)$: Japan Steel Production
$x_5(t)$: South Korea Steel Production
$x_6(t)$: Steel Price
$x_7(t)$: Subcontinent Scrap Price
$x_8(t)$: Far-East Scrap Price
$x_9(t)$: OPEC Oil Production
$x_{10}(t)$: Non-OPEC Oil Production
$x_{11}(t)$: Oil World Trade
$x_{12}(t)$: EUR/\$ Ex. Rate
$x_{13}(t)$: WON/\$ Ex. Rate
$x_{14}(t)$: YEN/\$ Ex. Rate
$x_{15}(t)$: Bunkers Price
$x_{16}(t)$: Product Tankers Building Price
$x_{17}(t)$: Aframax DH Building Price
$x_{18}(t)$: Suezmax DH Building Price
$x_{19}(t)$: VLCC DH Building Price
$x_{20}(t)$: Crude Carrier 105000dwt FRSingle Voyage
$x_{21}(t)$: Crude Carrier 150000dwt FRSingle Voyage
$x_{22}(t)$: Crude Carrier 300000dwt FRSingle Voyage
$x_{23}(t)$: Clean Carrier 70/85000dwt FRSingle Voyage
$x_{24}(t)$: Product Tankers DS/DH 5Years Market Value
$x_{25}(t)$: Aframax DS/DH 5Years Market Value
$x_{26}(t)$: Suezmax SH/DH 5Years Market Value
$x_{27}(t)$: VLCC SH/DH 5Years Market Value
$x_{28}(t)$: Clean Carrier- 40/45000dwt DB/DH 10 Years
$x_{29}(t)$: Tanker Fleet- 10000 DWT+ Supply
$x_{30}(t)$: Tanker Fleet- 10000 DWT+ Demand
$x_{31}(t)$: Tanker Fleet- 10000 DWT+ Util. Rate
$x_{32}(t)$: Tanker Order Book in Percent of Existing Fleet

The structure of the inputs, output and patterns are represented in Table 6-1 and Table 6-2. To increase the accuracy of the result, all the time series are gathered from various resources in R.S. Platou Economic Research and they are sometimes obtained by digitising in graphical form.

To find out the most accurate ANN model in the following section, the number of neurons and consequently the best point to stop the training of the network, i.e. the iteration number, are identified. Then the learning rates of the hidden and the output layers are identified and finally the momentum is added to the neural network to obtain the most accurate network. The schematic diagram (Figure 6-1) represents the above stages.

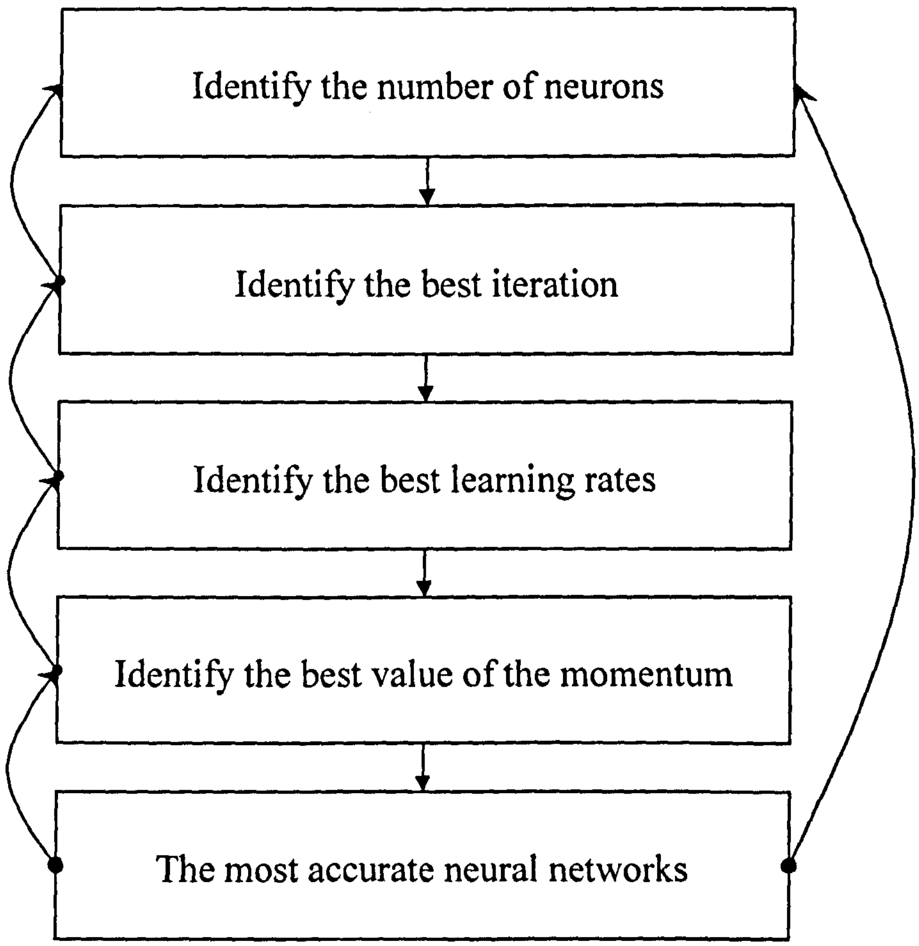


Figure 6-1: The schematic diagram of the different stages of the modelling

Pattern Number	Patterns	Input (1) USA Steel Production	Input (2) EU Steel Production	Input (3) China Steel Production	Input (4) Japan Steel Production	Input (5) South Korea Steel Production	Input (6) Steel Price	Input (7) Subcontinent Scrap Price	Input (8) Far-East Scrap Price	Input (9) OPEC Oil Prod.	Input (10) Non-OPEC Oil Prod.	Input (11) Oil World Trade	Input (12) EUR/$ Ex. Rate	Input (13) WON/$ Ex. Rate	Input (14) YEN/$ Ex. Rate	Input (15) Bunkers Price	Input (16) Product Tankers Building Price	Input (17) Aframax DH Building Price
01	January 1995	7.73	13.99	7.36	8.56	3.04	258.70	175.00	151.56	24.78	42.72	31.74	100.00	100.00	100.00	104.33	32.22	41.94
02	February 1995	7.73	14.17	7.36	8.56	3.04	258.70	179.69	151.56	24.78	42.61	31.85	102.46	100.00	101.85	106.30	32.78	42.22
03	March 1995	7.91	14.26	7.64	8.65	3.04	258.70	187.50	151.56	24.89	42.39	31.96	107.38	101.85	108.62	105.32	33.06	42.50
04	April 1995	7.91	14.26	7.64	8.56	3.13	269.57	187.50	156.25	24.89	42.17	32.07	107.08	103.38	120.00	107.28	33.33	42.78
05	May 1995	7.91	14.45	7.64	8.65	3.13	269.57	187.50	153.13	25.00	42.28	32.50	105.23	104.00	117.23	94.49	33.33	42.78
06	June 1995	7.91	14.63	7.64	8.65	3.13	295.65	187.50	153.13	25.00	42.39	32.83	107.38	104.00	118.77	82.68	33.33	42.78
07	July 1995	7.82	14.45	7.73	8.65	3.13	295.65	187.50	153.13	25.11	42.61	33.37	108.00	104.00	114.46	85.63	33.33	42.78
08	August 1995	7.82	14.36	7.73	8.65	3.13	284.78	187.50	153.13	25.22	42.72	33.15	103.08	102.15	104.92	87.60	33.06	42.78
09	September 1995	7.82	14.26	7.73	8.56	3.13	280.43	190.63	151.56	25.22	42.72	32.93	104.62	103.08	99.38	85.63	32.78	42.78
10	October 1995	7.82	14.26	7.73	8.47	3.13	269.57	193.75	146.88	25.22	42.83	32.61	104.92	103.38	99.38	87.60	32.78	42.78
11	November 1995	7.82	13.90	7.73	8.37	3.13	265.22	184.38	143.75	25.43	42.93	33.04	102.77	102.46	97.85	104.33	32.50	42.50
12	December 1995	7.82	13.44	7.82	8.37	3.13	254.35	179.69	143.75	25.65	43.15	33.48	103.08	102.15	96.92	101.38	32.22	42.22
13	January 1996	7.91	13.34	7.91	8.28	3.22	254.35	189.06	143.75	25.98	43.37	34.02	99.38	100.31	93.54	97.44	32.22	41.94
14	February 1996	7.91	13.34	7.91	8.10	3.22	250.00	190.63	143.75	25.87	43.37	34.02	101.23	100.62	96.00	107.28	32.22	41.67
⋮	⋮																	
115	July 2004	8.10	15.92	20.89	9.29	4.05	569.57	387.50	343.75	28.80	49.89	44.67	97.23	67.38	88.92	161.42	34.17	50.00
116	August 2004	8.37	15.74	21.53	9.39	3.96	569.57	425.00	371.88	29.13	49.78	45.00	97.23	68.31	90.77	173.23	34.44	51.67
117	September 2004	8.37	15.74	21.90	9.39	3.96	589.13	415.63	337.50	29.57	49.67	45.11	98.77	68.31	89.85	143.70	34.72	52.78
118	October 2004	8.47	15.74	22.64	9.48	3.96	589.13	404.69	368.75	29.67	49.89	45.33	102.15	70.15	93.54	154.53	36.39	55.00
119	November 2004	8.47	15.74	23.37	9.48	4.05	589.13	421.88	375.00	29.89	50.00	45.43	107.38	75.38	96.92	163.39	37.50	56.94
120	December 2004	8.47	15.64	24.20	9.48	4.05	589.13	434.38	356.25	30.00	50.22	45.54	108.92	76.00	96.92	174.21	39.44	59.72
	Maximum	8.83	15.92	24.20	9.48	4.05	589.13	434.38	385.94	30.00	50.22	45.54	108.92	104.00	120.00	174.21	39.44	59.72
	Minimum	6.99	13.25	7.36	7.36	3.04	193.48	109.38	106.25	24.57	42.17	31.74	68.00	47.08	69.54	52.17	25.00	33.33

Table 6-1: The scrapped tonnage model inputs (1 to 17)

Pattern Number	Patterns	Input (18)	Input (19)	Input (20)	Input (21)	Input (22)	Input (23)	Input (24)	Input (25)	Input (26)	Input (27)	Input (28)	Input (29)	Input ((30)	Input (31)	Input (32)	Output(1)
		Suezmax DH Building Price	VLCC DH Building Price	Crude Carrier 105000dwt FRSingle V.	Crude Carrier 150000dwt FRSingle V.	Crude Carrier 300000dwt FRSingle V.	Clean Carrier 70/85000dwt FRSingle	Product Tankers DS/DH 5Years Market	Aframax DS/DH 5Years Market Value	Suezmax SH/DH 5Years Market Value	VLCC SH/DH 5Years Market Value	Clean Carrier- 40/45000dwt DB/DH 10 Y	Tanker Fleet- 10000 DWT+ Supply	Tanker Fleet- 10000 DWT+ Demand	Tanker Fleet- 10000 DWT+ Util. Rate	Tanker Order Book in Percent of Existing	Tankers Sold for Scrapping
01	January 1995	53.06	86.11	14.81	14.81	11.11	17.83	21.89	29.81	33.21	52.22	16.76	260.87	218.48	83.91	9.2	4.18
02	February 1995	53.33	86.11	12.96	12.96	8.02	16.59	22.26	29.81	33.96	51.85	17.31	259.78	216.30	83.48	9.1	4.18
03	March 1995	53.61	86.11	12.35	13.58	9.88	16.10	22.64	29.81	34.34	51.85	17.85	258.70	215.22	83.04	9.0	4.18
04	April 1995	53.89	86.11	14.20	12.35	8.02	13.87	23.02	29.81	34.34	51.85	18.18	258.70	213.04	82.61	8.8	3.98
05	May 1995	53.89	86.11	11.73	12.96	6.79	13.37	23.77	29.81	34.34	51.85	18.73	258.70	215.22	83.91	8.6	3.98
06	June 1995	53.89	86.11	12.96	13.58	13.58	17.34	24.15	29.81	34.34	51.85	19.05	258.70	218.48	85.00	8.4	3.98
07	July 1995	53.61	86.11	14.20	19.75	19.14	19.32	24.15	29.81	34.72	52.22	19.05	258.70	220.65	85.87	8.1	1.02
08	August 1995	53.33	86.67	14.20	14.81	18.52	18.82	24.15	30.19	35.47	53.33	19.05	258.70	220.65	85.87	7.9	1.02
09	September 1995	53.06	86.94	14.20	14.81	14.81	18.33	24.15	30.57	36.23	53.70	19.05	258.70	220.65	85.87	7.6	1.02
10	October 1995	52.78	87.50	14.20	14.81	11.11	20.06	24.15	30.57	36.23	53.70	19.05	258.70	220.65	86.09	7.2	1.55
11	November 1995	52.50	86.67	13.58	16.05	16.67	21.80	24.15	30.57	36.23	53.70	19.05	258.70	221.74	86.52	7.0	1.55
12	December 1995	52.50	85.83	14.81	16.05	16.67	22.04	24.15	30.57	36.23	53.70	19.05	258.70	222.83	86.96	6.8	1.55
13	January 1996	52.22	85.00	17.90	17.28	16.05	21.05	24.15	30.94	37.36	54.81	19.05	258.70	223.91	87.39	6.5	0.89
14	February 1996	51.94	84.72	17.28	17.28	19.75	20.06	24.15	30.94	38.49	55.93	19.05	259.78	230.43	86.96	6.3	0.89
⋮	⋮																
115	July 2004	61.11	93.33	32.72	52.47	79.01	30.71	32.45	49.81	60.00	87.41	23.31	301.09	271.74	90.43	26.8	0.67
116	August 2004	63.89	96.94	32.10	45.68	62.35	26.75	33.58	51.32	62.26	90.74	24.07	302.17	275.00	91.30	26.9	0.65
117	September 2004	66.39	99.72	32.72	46.30	63.58	26.75	34.72	53.21	66.79	93.33	25.71	303.26	278.26	92.17	27.0	0.40
118	October 2004	68.06	103.06	80.25	107.41	136.42	40.37	36.60	54.72	71.32	101.48	26.91	304.35	280.43	93.04	27.1	0.10
119	November 2004	68.89	106.11	85.19	125.93	189.51	56.97	39.62	56.98	72.45	105.19	28.11	305.43	282.61	93.70	26.9	0.44
120	December 2004	69.72	109.44	68.52	77.78	125.31	53.00	39.62	56.98	72.45	105.19	28.11	306.52	284.78	94.35	26.5	0.29
	Maximum	69.72	109.44	85.19	125.93	189.51	60.43	39.62	56.98	72.45	105.19	28.11	306.52	284.78	94.35	27.14	4.46
	Minimum	41.94	63.89	9.88	9.26	6.79	10.15	17.36	23.02	33.21	48.89	12.07	258.70	213.04	82.61	6.02	0.10

Table 6-2: The scrapped tonnage model inputs (18 to 32) and the output

6.2.1 APPLICATION OF STATIC ANNS FOR MODELLING OF THE MONTHLY SCRAPPED TONNAGE

A feed-forward MLP network with error back-propagation learning algorithm and the batch weight updating method has been implemented to model the data.

As explained in section 4-7-1, regarding the data splitting, it is important to note that the accuracy is not about what proportion of data should be allocated in each sample. But it is about sufficient data points in each sample to ensure adequate learning, validation, and testing. For this study as Granger (1993) suggests 20% of the data is used for testing the network. Therefore, the data is randomised in first place and then 70% of the patterns (84 months observation) in each time series is used to train the neural networks.10% of the patterns (12 months observation) are used to validate the networks. After the training and validation, the rest of the data, i.e. 20% of the patterns (24 months observation), is used to test the performance of each particular neural networks.

This network includes one hidden layer with the hyperbolic tangent activation function (Equation 4-3), which will give an output in the range [-1, 1]. As explained in section 4-3-2, to achieve to an accurate prediction, normalisation has been performed as the pre-processing of the data. Figure 6-2 illustrates the general structure of the ANN and the system of inputs to the ANN and the corresponding output.

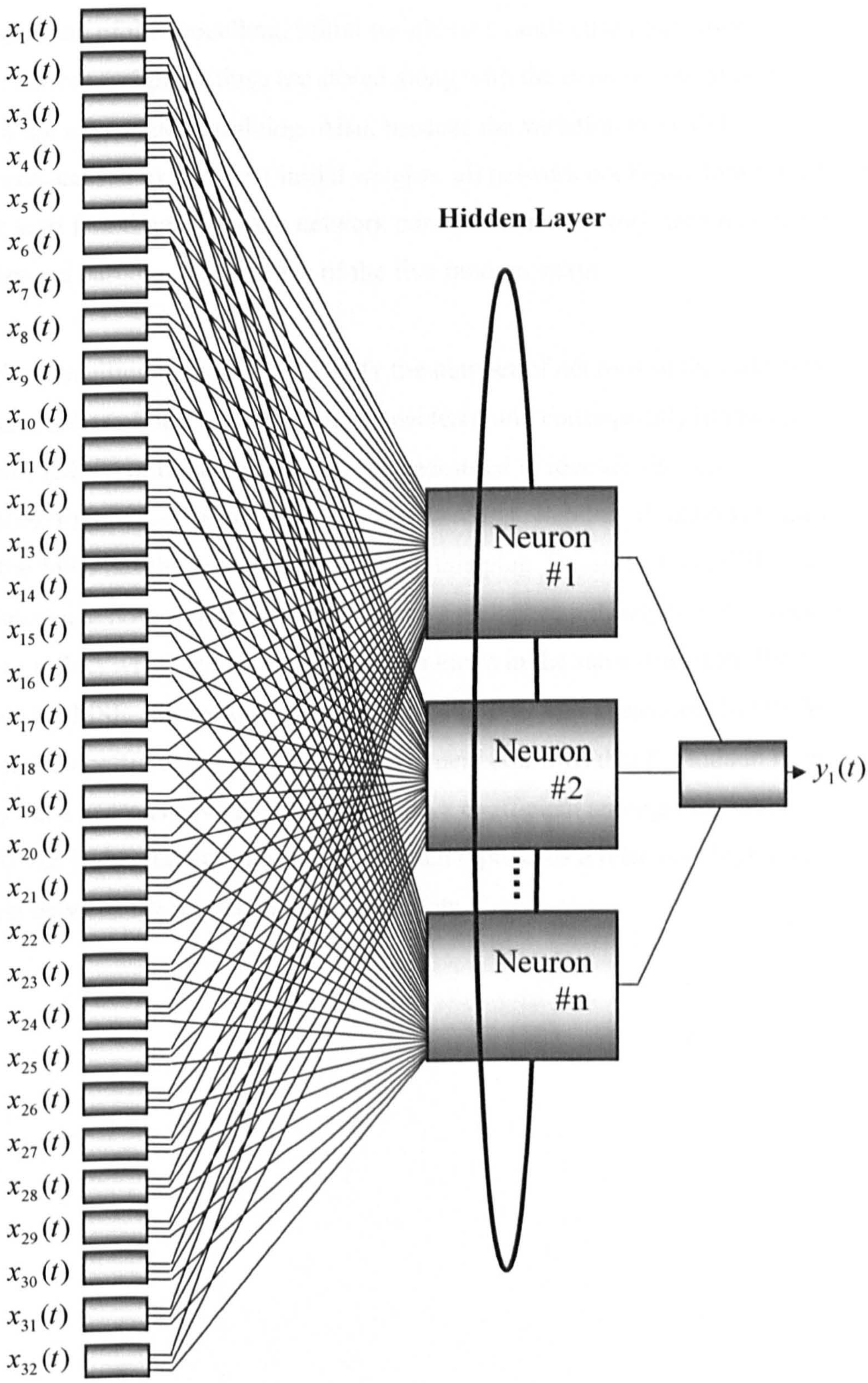


Figure 6-2: The static ANN model for the monthly scrapped tonnage

At the beginning of the modelling, initial weights are randomised but, after each stage, the current weight settings are stored along with the components to be re-loaded for the next stage of training. Also, because the variation in model performance, caused by different initial weights, all network configurations have been run for at least five times and each network parameter and network architecture has been optimised based on the average of the five random starts.

To start the modelling process and identify the number of neurons in the hidden layer, a different number of neurons have been considered and consequently the mean square error (MSE) of the network has been measured to identify the best combination. Figure 6-3 suggests that the most accurate number of neurons is three, because the MSE has the least value (0.161) in this plot. The size of the MSE can be used to determine how well the network output fits the desired output, but it does not necessarily reflect whether the two sets of data move in the same direction. Hence, In addition to the MSE, the correlation coefficient has also been measured. In this figure, the correlation coefficient of the similar experiment confirms that the hidden layer including three neurons have the highest positive coefficient amongst the others. The value of 0.928 is the maximum of the plot which represents a relatively high positive correlation between the model and the actual data.

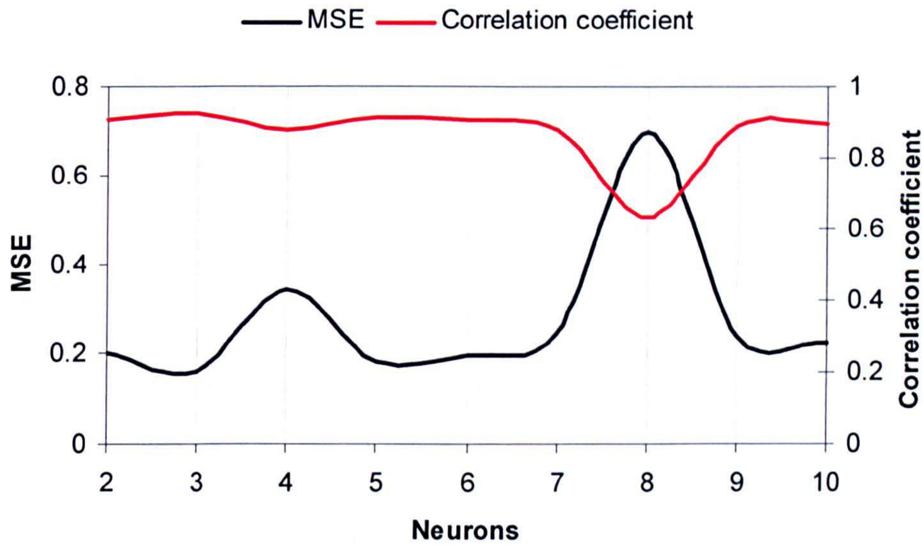


Figure 6-3: Impact of the neurons on the static ANN model for scrapped tonnage

To find out the best point in time to stop the training of the above network various iterations have been tried and their performances are tested. The training error measures how well the ANN models the data and the testing error is a measure of how well the model follows the common pattern. Figure 6-4 illustrates both training and testing error of the above model. In the first section of the plot the testing error begins to learn the pattern in the time series, indicated by the decreasing error. At the point which the testing error is minimal the training of the ANN is in its optimal level. Beyond this minimum in the next section, the training error continues to decrease but the testing error increases.

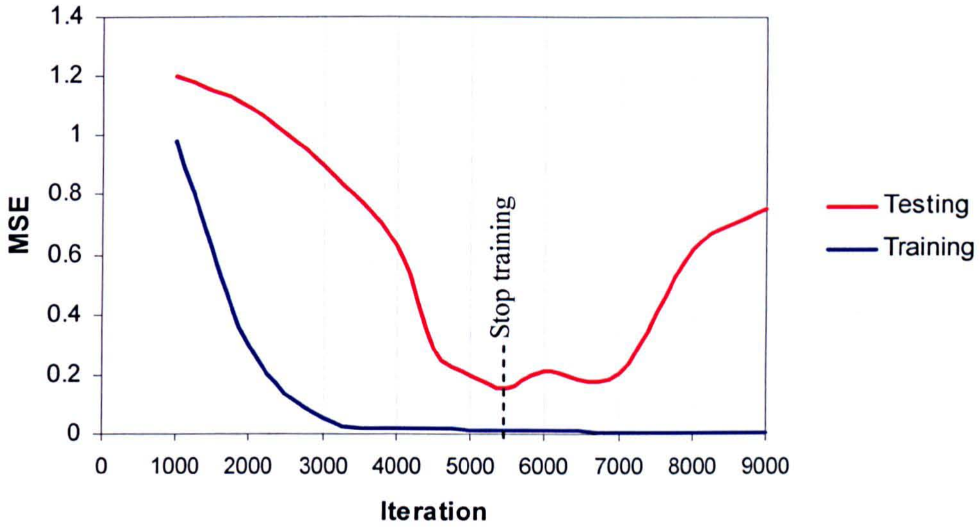


Figure 6-4: Mean Square Error of the training and testing stages of the model for different iterations

As it appears in this figure, there is a global minimum for the error of the model at the iteration 5500 which shows the most adequate training. There is also a local minimum at the iteration 7000.

So far, the best number of neurons and iterations of the network has been identified. At this point, the different learning rates γ are examined to find out the best values for each layers of the network. Various learning rates are considered for both hidden and output layers and subsequently the error and the correlation coefficient are measured. Figure 6-5-a represents the measured learning rates for the hidden layer and similarly Figure 6-5-b plots various learning rates for the output layer. The first plot suggests that the most accurate model with the lowest error and also the highest positive correlation coefficient takes place at the value of 0.02 for the output layer. Subsequently, the second plot suggests that the most accurate model, corresponding to the previous 0.02 value of the output layer, happens at 0.2 for the hidden layer. Therefore, the final decision about the learning rates in both mentioned layer would be 0.02 and 0.2 respectively.

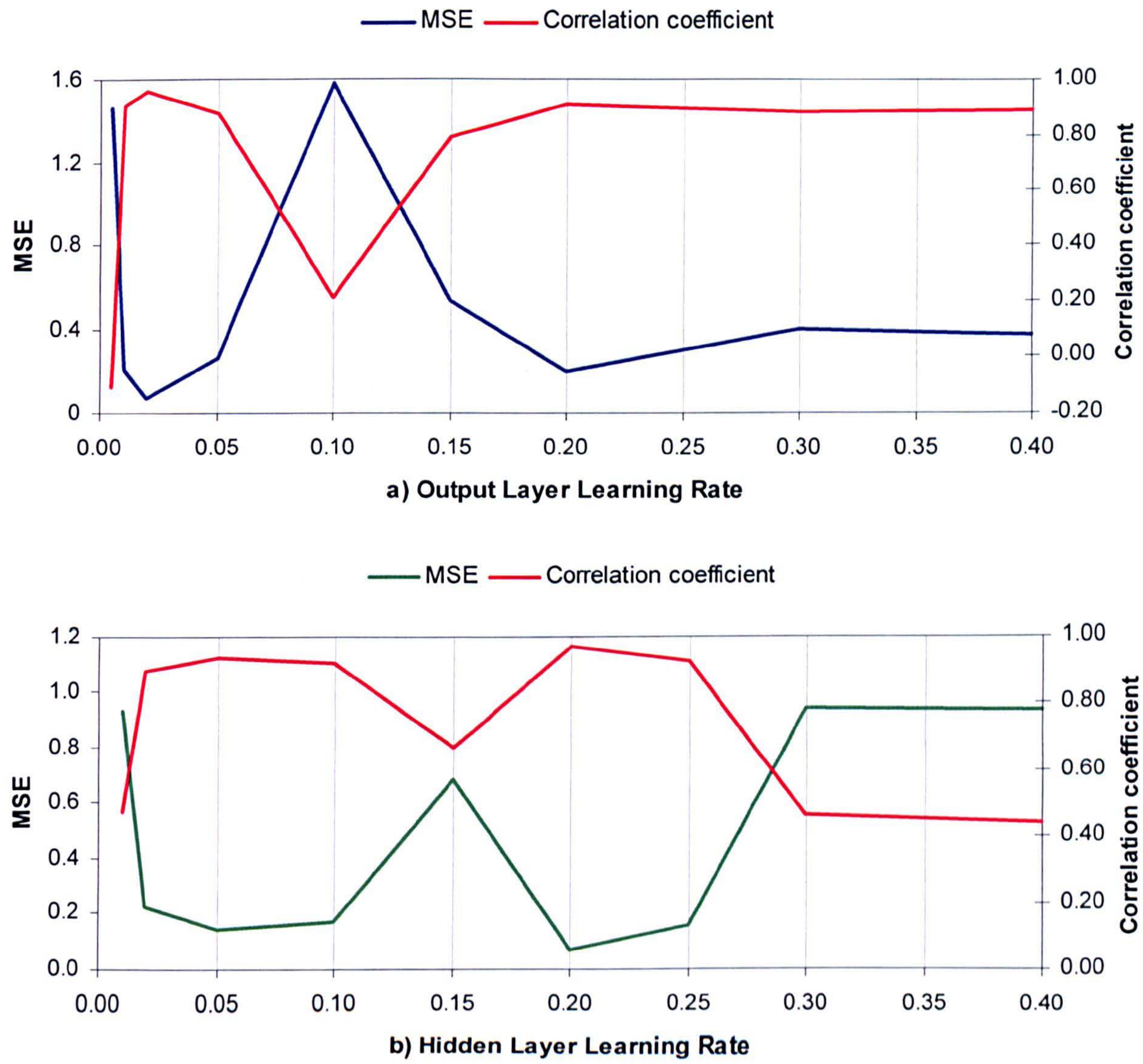


Figure 6-5: a) Impact of different learning rates in the output layer on MSE and correlation coefficient for the static ANN for scrapped tonnage
b) Impact of different learning rates in the hidden layer on MSE and correlation coefficient for the static ANN for the scrapped tonnage

The last part of the modelling identifies the momentum α of the network. Different momentums are tried and then the error of the modelling is measured to see the behaviour of changing momentum on the ANN (Figure 6-6). In addition, the correlation coefficient is calculated in each stage to clarify the corresponding MSE results. The minimum error of 0.069 occurs when $\alpha = 0.7$. The error value is also quiet near to the global minimum when $\alpha = 0.9$, with 0.084. The corresponding correlation coefficient values are 0.963 and 0.962 respectively which are shown a relatively high positive correlation between the model and the actual numbers.

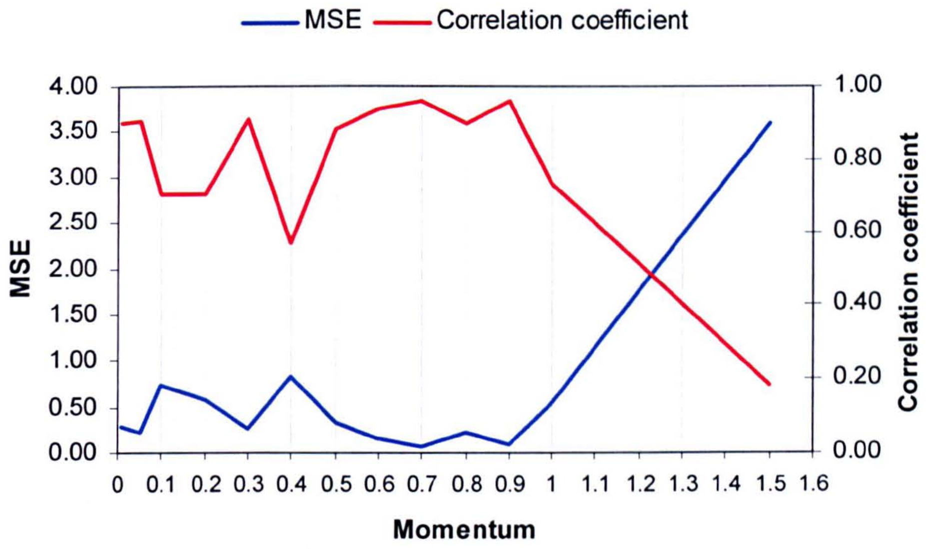


Figure 6-6: Impact of adding different momentums to the static ANN model for monthly scrapped tonnage

According to the above studies for various static ANN model parameters, the best ANN model and the training is the one which includes:

- one hidden layer
- three neurons (or PEs)
- the hyperbolic tangent activation function.
- the learning rate of 0.2 for the hidden layer $(\gamma = 0.2)$
- the learning rate of 0.02 for the output layer $(\gamma = 0.02)$
- the momentum of 0.7 for the networks $(\alpha = 0.7)$

6.2.1.1 SENSITIVITY ANALYSIS OF THE STATIC ANN MODEL FOR THE MONTHLY SCRAPPED TONNAGE

As the architecture of the static ANN for the monthly scrapped tonnage is identified and the training of the neural networks is completed (see section 6-2-1), it is possible to determine the sensitivity of the out put with respect to each individual input. Sensitivity analysis can determine the effect that each of the network inputs is having on the network output. The sensitivity analysis, which is employed for this model, has been explained in section 4-5-2. Figure 6-7 represents the sensitivity analysis of the static ANN model for the monthly scrapped tonnage.

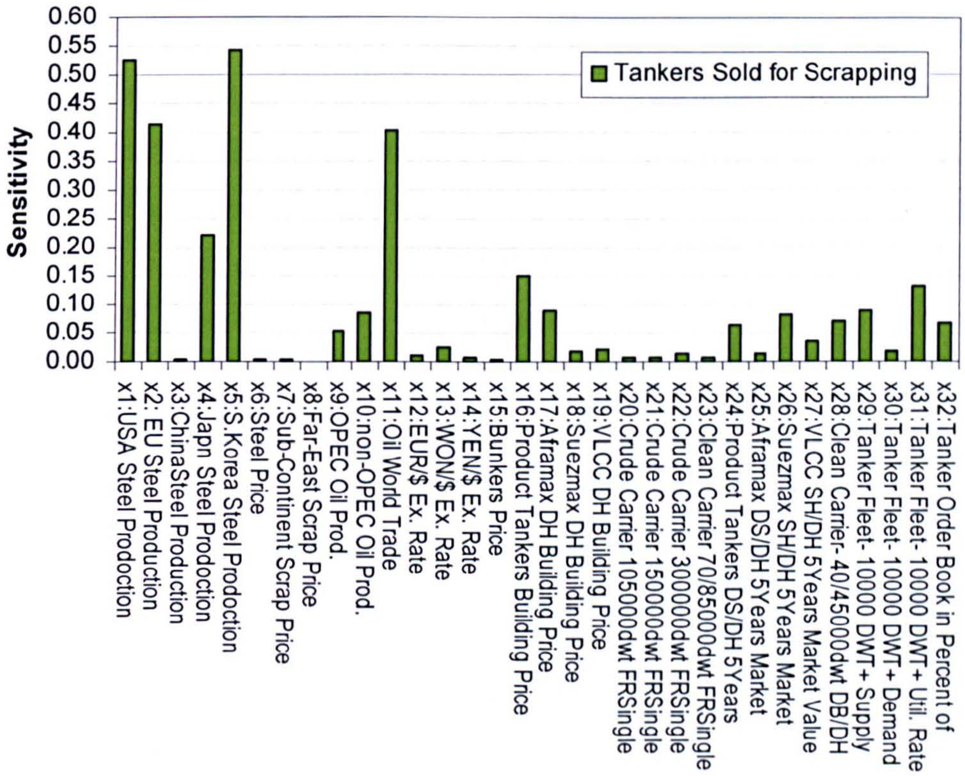


Figure 6-7: Sensitivity about the mean for all the inputs to the static ANN model for the monthly scrapped tonnage

Figure 6-7 explains that steel production, in general, has the most influence to the model. South-Korean steel production has the highest value (0.542), and USA's steel production (0.525) is the second most valuable. EU's and Japan's have the third and forth highest values respectively (0.413 and 0.220) but China's steel production does not have a significant influence on the model (0.004). There are a few possibilities to justify these results. Firstly, it is possible that the public intuition about the influence of the Chinese steel production is incorrect as the model shows its influence is quite low. Secondly, the low influence of the Chinese steel production is due to the fact that the steel demand inside the country is quiet high because of the booming economy of China i.e. consumers use most of the steel production inside China and the rate of the exportation is not enough to affect the demolition market. On the contrary, USA affects the market because they send their steel production to the international market.

Alternatively, it is possible that the necessary data is not accessible for the model to carry on the accurate sensitivity analysis.

Oil world trade has also shown a relatively high influence on the model. Compared with the above steel production sensitivities oil production sensitivities are small but non-OPEC oil production shows a slightly more influence. Product Tanker building prices and the tanker fleet utilisation rate have shown a small influence on the model. On the contrary, freight rates, bunker price (as the operating cost of the ship), second-hand prices and exchange rates have very small impacts on the model. There are a few reasons for the low influence of the above variables, like freight rates, to the model. Firstly, it is possible that their influence is offset by influence of another variable to the model or, secondly, there is not enough data to carry out an accurate sensitivity analysis.

Based on the above static ANN studies for the monthly scrapped tonnage, the most sensitive parameters that drive the output for this model, as Figure 6-7 explains, are:

1. x_5 : South Korea steel production
2. x_1 : USA steel production
3. x_2 : EU steel production
4. x_{11} : Oil world trade
5. x_4 : Japan steel production
6. x_{16} : Product tankers building price
7. x_{31} : Tanker fleet utilisation rate

There are also few parameters with negligible inputs to the model including:

1. x_3 : China steel production
2. x_6 : Steel price
3. x_7 : Subcontinent scrap price
4. x_8 : Far-East scrap price
5. x_{15} : Bunker price (represents the operating cost for this study)

The obtained static ANN model is based on the sensitive inputs and the inputs with the low number of sensitivities are not taken into account to produce the desire output.

6.2.2 Development of the Dynamic ANN model for the scrapped tonnage Prediction

For this part of the study, a dynamic ANN is implemented to forecast the monthly demolition tonnage. The architecture of a dynamic network is different from a static network (see section 4-7). For example, setting the parameters such as learning rate, activations or weight updates might result in a significant difference compared to the static ANN training. Figure 6-8 illustrates the system of inputs to the dynamic ANN and the corresponding output.

A feed-forward MLP neural network with an error back- propagation learning algorithm has been created. Then a sequential framework has been added to the network by adding a short-term memory mechanism in the form of a delay line. A section of the time series of the form $[x(t), x(t-1), ..., x(t-p)]$, and $[y(t-1), ..., y(t-p)]$ are used as inputs for the network. The delay line in order of p, and the desired output is $y(t)$. An integer value is substituted for p and several ANNs have been trained. The best performing ANN is noted and then the process is repeated with a different integer value for p. When sufficient p values have been investigated the MSE values are compared and the minimum is chosen. Consequently, forecast errors are measured in order to judge how good that model is in terms of its prediction abilities.

Similar to the previous study, as illustrated in Figure 6-1, the number of neurons and iterations of the neural networks are identified. Then, learning rates and momentum are measured.

Unlike the static mode, data is split into two groups, one of training and the other validating data. As explained in the previous section and as Granger (1993) suggests 80% of the patterns (94 observations) in each time series is used to train the neural networks and 20% (23 observations) to validate the networks. The three last months, in each time series, are not taken into account for the training purposes because their data is needed to check the prediction performance of the final ANN. All neural networks which are considered in this part of the study are included one hidden layer with the hyperbolic tangent activation function (*tanh*), Equation 4-3, in their hidden layers.

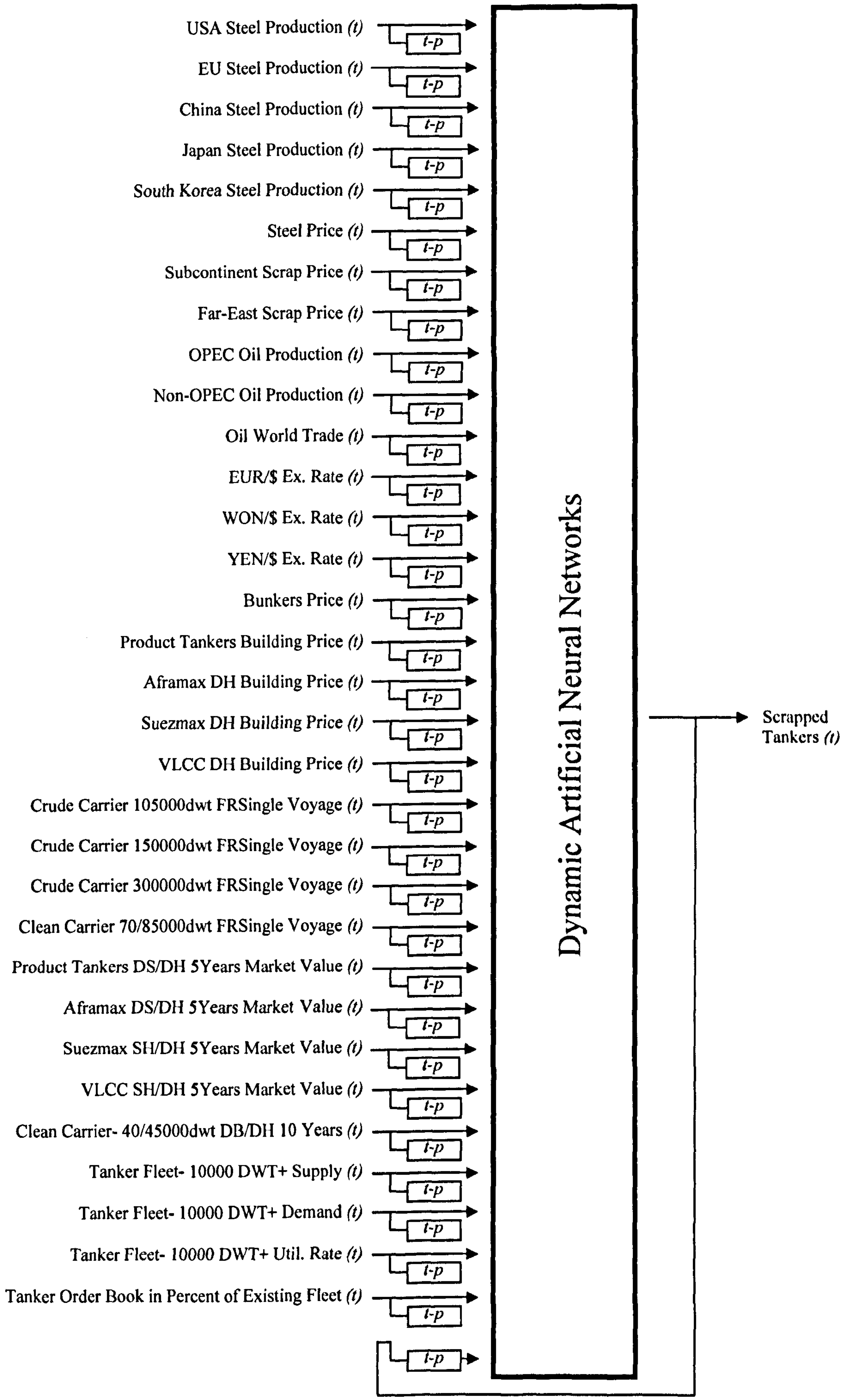


Figure 6-8: The system of inputs to the dynamic ANN and the corresponding output

In order to start the modelling and thus investigate the performance of the different dynamic ANN configurations, various numbers of neurons and iterations have been changed simultaneously. Then the mean square error of the testing stage for each ANN has been measured to find out the most accurate combination.

Figure 6-9 represents the error surface of the above analysis. The minimum measured MSE of the error surface, which is shown in Figure 6-9, is 0.167 and it occurs at iteration 16000 and at a number of neurons of 9.

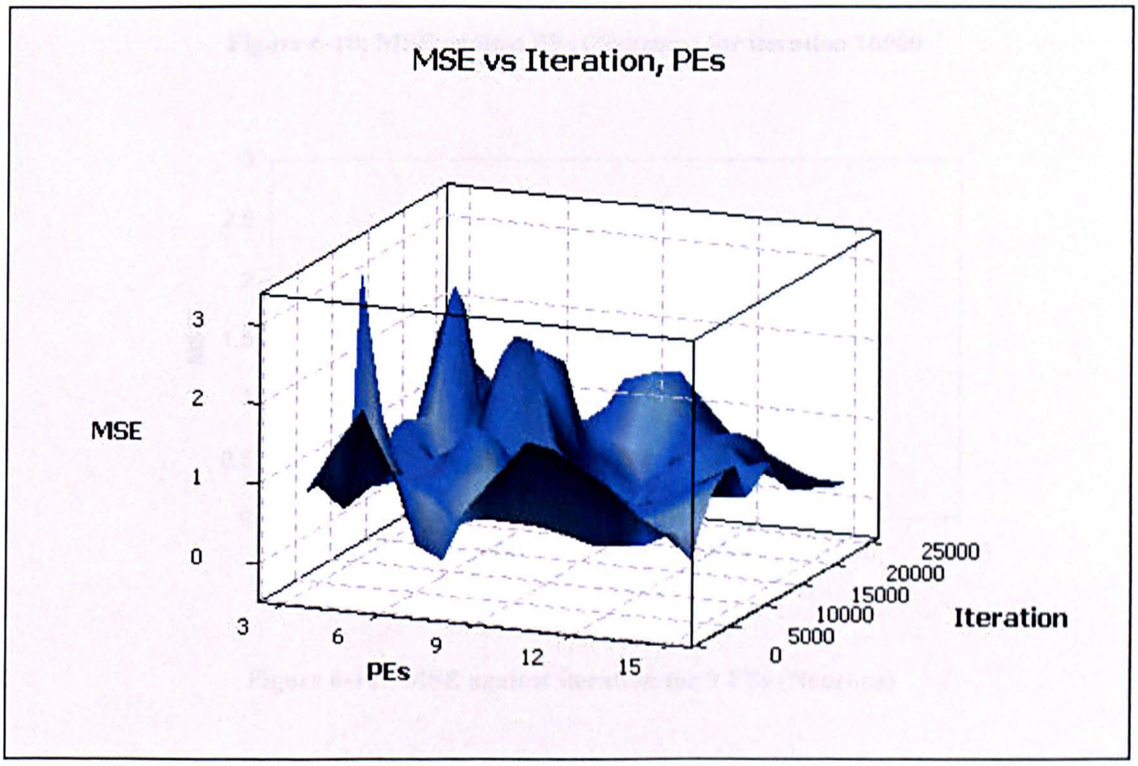


Figure 6-9: The error surface of the designed dynamic ANN for the monthly scrapped tonnage with one month delay (*p*=1).

Two cross sections of the above figure are represented in Figure 6-10 and Figure 6-11 to clarify these results.

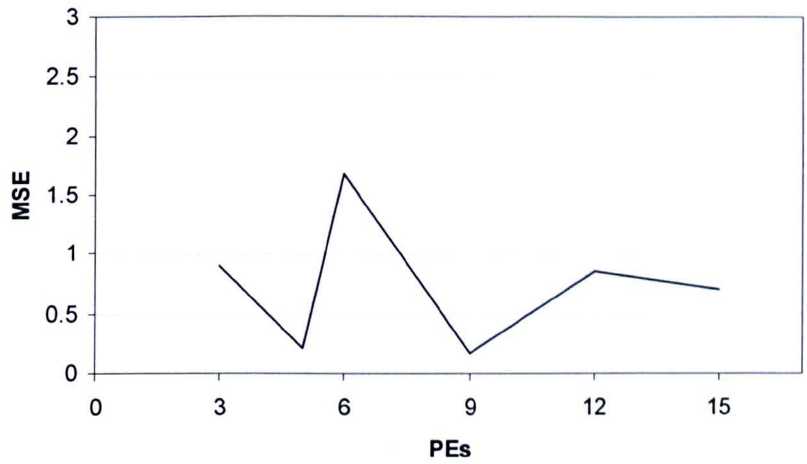


Figure 6-10: MSE against PEs (Neurons) for iteration 16000

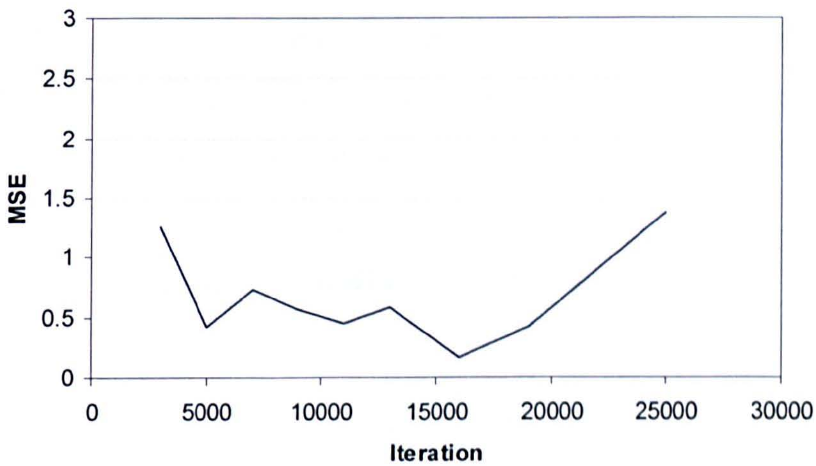


Figure 6-11: MSE against iteration for 9 PEs (Neurons)

It is important to note that the mean square error of the training stage is quite satisfactory. The average MSE of the training stage for the above dynamic ANN is measured 0.00141. Also, the standard deviation of this training error is 0.000100. Figure 6-12 is shown the mean square error of the network in training stage for different iterations.

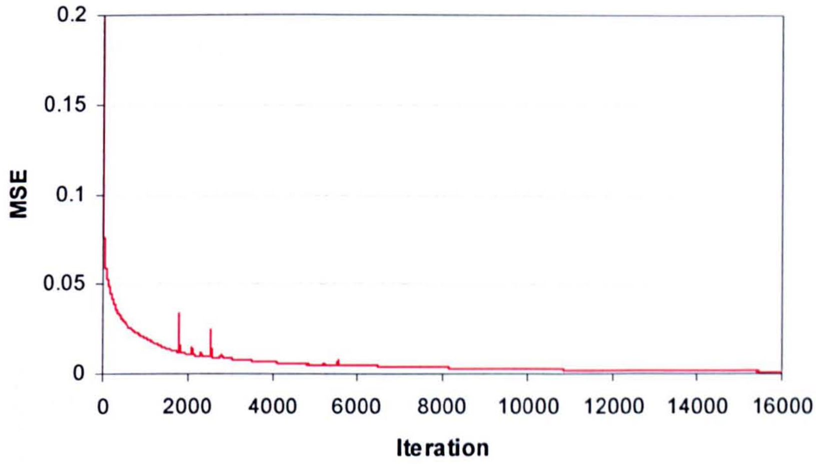


Figure 6-12: Training error of the chosen dynamic ANN for the monthly scrapped tonnage in its best run

The above result can also be confirmed with the correlation coefficient values. Figure 6-13 shows the correlation surface of the same experiment. As it appears, the maximum of 0.95 occurs at iteration 16000 and at a number of neurons of 9.

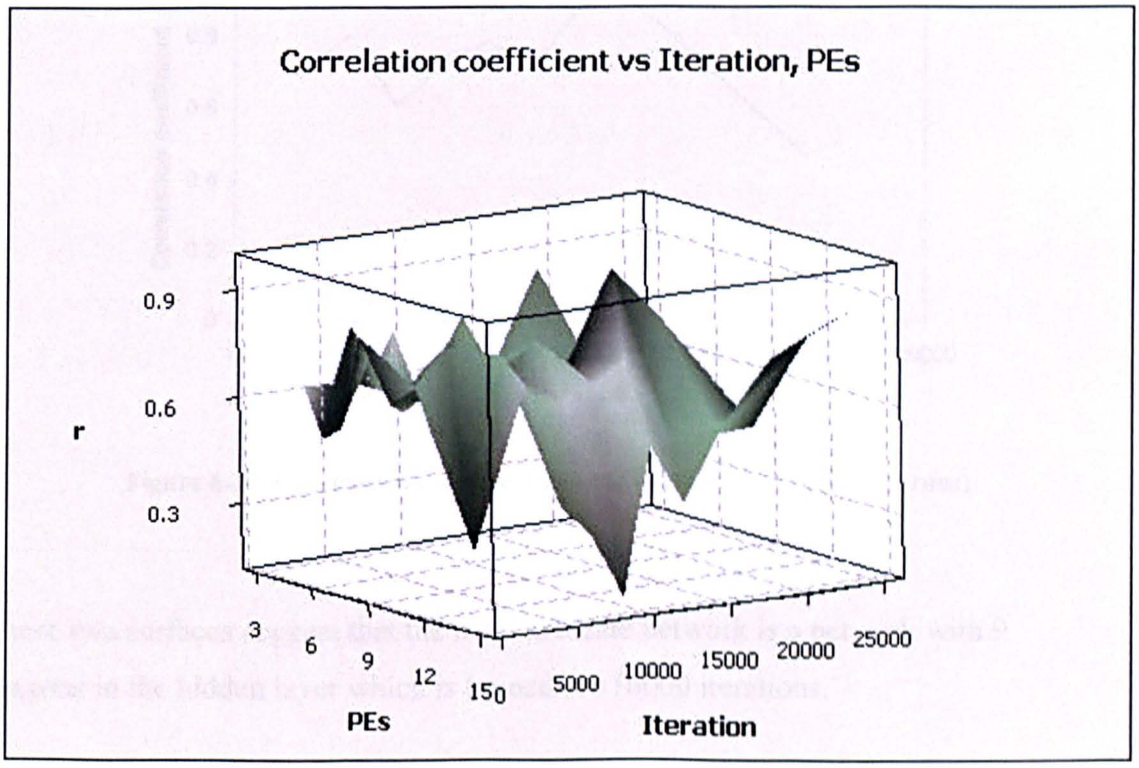


Figure 6-13: The correlation surface of the designed dynamic ANN for the monthly scrapped tonnage with one month delay (p=1)

Two cross sections of the above figure are represented in Figure 6-14 and Figure 6-15 to clarify these results.

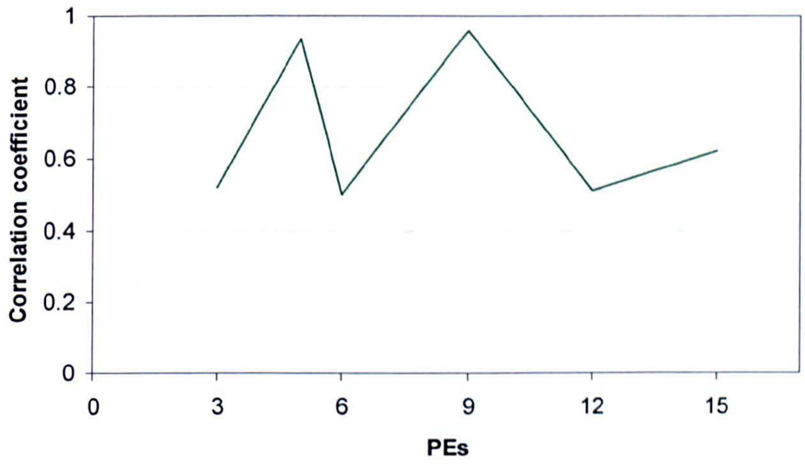


Figure 6-14: Correlation coefficient against PEs (Neurons) for iteration 16000

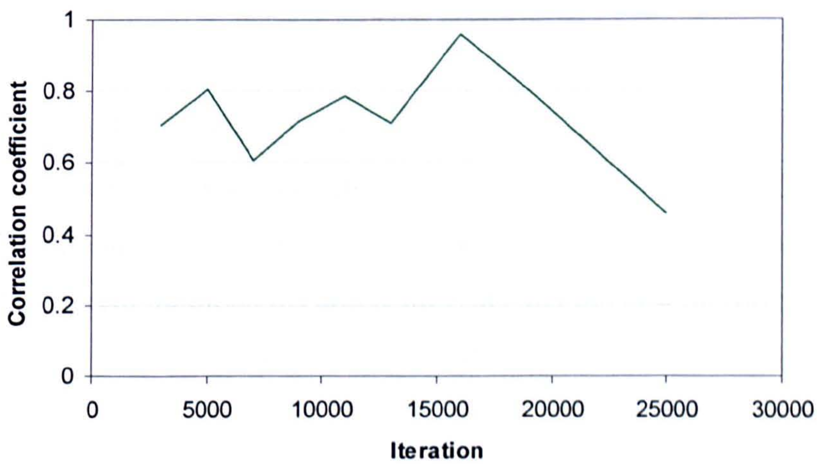


Figure 6-15: Correlation coefficient against iteration for 9 PEs (Neurons)

These two surfaces suggest that the most accurate network is a network with 9 neurons in the hidden layer which is trained for 16000 iterations.

To identify the best learning rate γ for the network training, the mean square error is measured for different rates. The criterion to choose the best learning rate for this network is to measure MSE of the testing stage. Figure 6-16 represents the MSE (and also the corresponding correlation coefficient) for various learning rates. It is shown that the minimum MSE occurs at $\gamma = 0.1$ with the value of 0.167. Simultaneously, correlation coefficient has the maximum of its value with 0.957 in the same point.

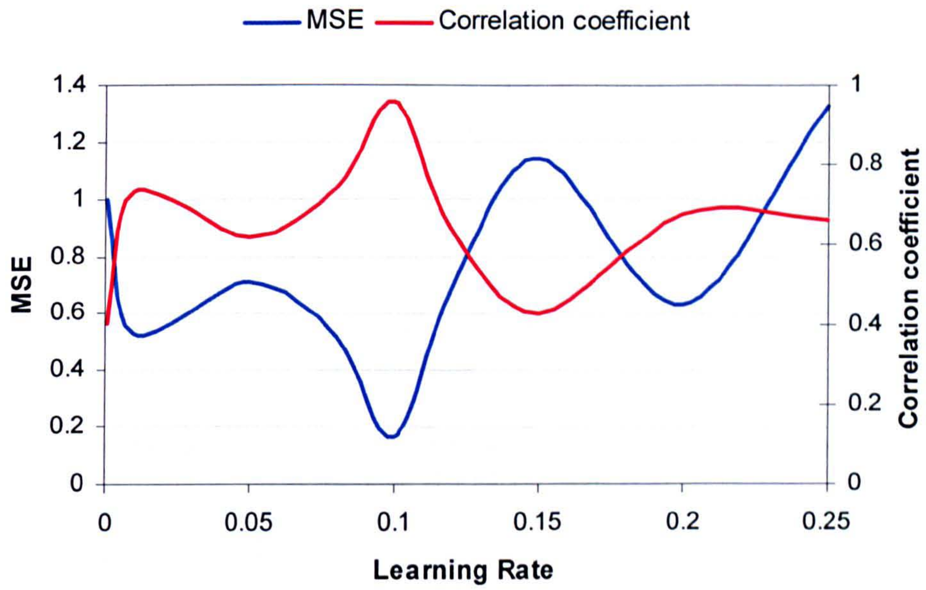


Figure 6-16: Impact of the different learning rate on the dynamic ANN for the monthly scrapped tonnage with one month delay

To investigate the effect of adding momentum α to the neural networks, different momentums added and subsequently the testing errors are calculated to find out the best performing dynamic ANN model. Figure 6-17 represents the performance of the network with different momentums and it suggests that the best performance happens at momentum 0.01. The value of the error is 0.080 and the corresponding correlation coefficient is 0.971 which explains a high positive correlation for the model.

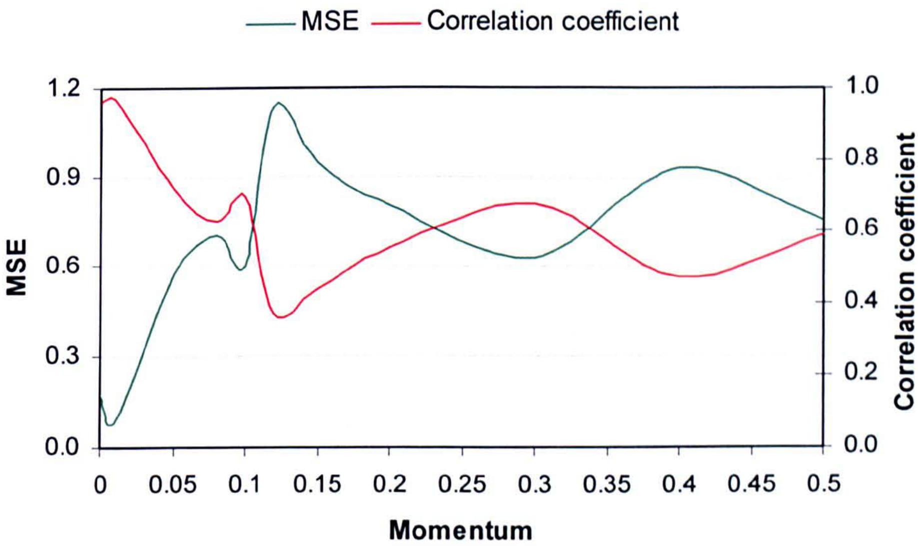


Figure 6-17: Impact of the different momentums on the dynamic ANN for the monthly scrapped tonnage with one month delay (t-1)

So far, the parameters of the dynamic ANN model with one month time delay (p=1) are specified.

- one hidden layer
- nine neurons (or PEs)
- 16000 iterations for the network
- the hyperbolic tangent activation function
- the learning rate of 0.1 for both layers ($\gamma = 0.1$)
- the momentum of 0.01 for the networks ($\alpha = 0.01$)

The next step is to expand the time delay (p= 2, 3 …) and investigate the possibility of reaching more accuracy (Figure 6-18). Expanding the delay window can increases the experience and the power of the decision making of the network because it has access to earlier data as well as data from one month previous. This may give more accuracy to the networks. According to the system of inputs to the above dynamic ANN and the corresponding output (Figure 6-8), which is explained earlier in this section, changing the value of the delay p will change the structure of the neural network. A bigger value of p means more accessibility to the past data.

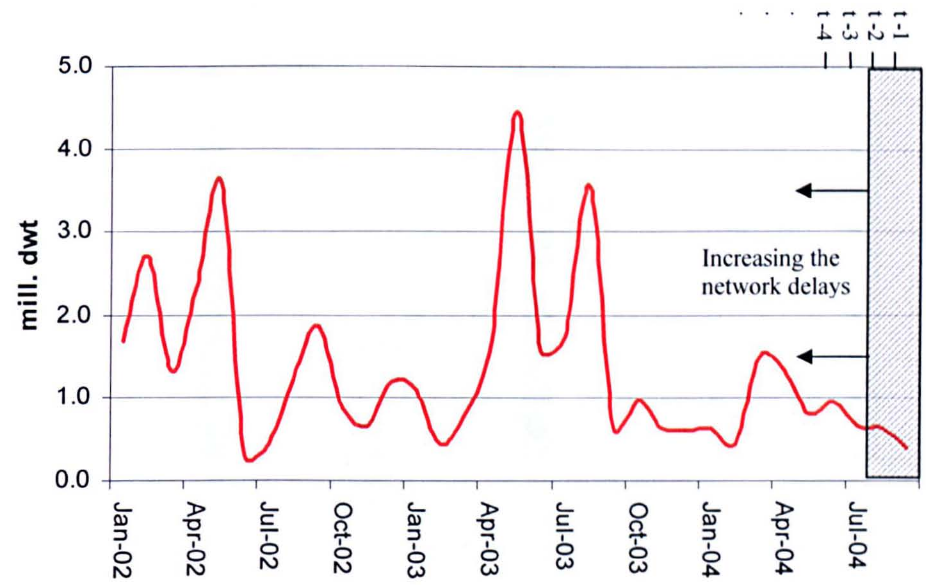


Figure 6-18: The time delay window for the ship demolition tonnage time series. The size of the window can be varied by changing the structure of the neural networks.

Based on the previous analysis and the dynamic ANN models, new models are implemented to investigate the impact of increasing the delay windows to two, three, four and five months, i.e p=2, p=3, p=4 and p=5. The number of iterations and neurons are changed simultaneously to obtain the error surface of each neural network. The MSE results of these analyses, for the best ANN chosen for each delay, are represented in Figure 6-19. This figure indicates that ANN's with 2 and 4 months delays have more accuracy than ANNs with 1, 3 or 5 months delay but as there is no correlation between the networks in each time delay this is just a random behaviour of the model.

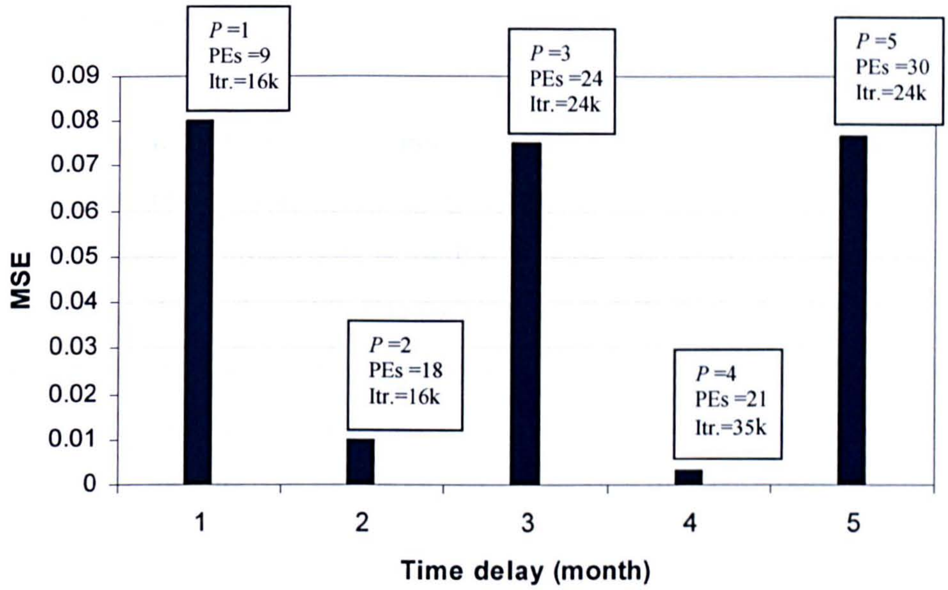


Figure 6-19: MSE against time delay of the best obtained dynamic ANN model for the monthly scrapped tonnage. Box annotations indicate the specification of the ANN for each time delay.

Table 6-3 is a comparison between the best performing dynamic ANNs in each particular time delay. It is shown that the minimum mean square error happens for the neural networks with four months time delay (t-4).

Delay window	PEs	Iteration	MSE
One month (t-1)	9	16,000	0.080
Two months (t-2)	18	16,000	0.010
Three months (t-3)	24	24,000	0.075
Four months (t-4)	21	35,000	0.003
Five months (t-5)	30	24,000	0.077

Table 6-3: The comparison between different window sizes for training and development of a Dynamic ANN model.

6.2.2.1 MONTHLY SCRAPPED TONNAGE PREDICTION USING DYNAMIC ANN MODEL

The best performing dynamic ANN model for the monthly scrapped tonnage is identified in previous section. In this section, the performance of the above neural networks is analysed. Unseen data is used to evaluate and verify the predictability of this ANN model. As explained before (see section 6-2-2), in the beginning of the modelling process the last three months of data is pulled out of the training and testing stages. Therefore, these data are available to put in the dynamic ANN model and check the performance of that model by comparing the prediction with the real values.

Figure 6-20 represents the actual measurements versus prediction of the dynamic ANN model for monthly scrapped tonnage in October, November and December.

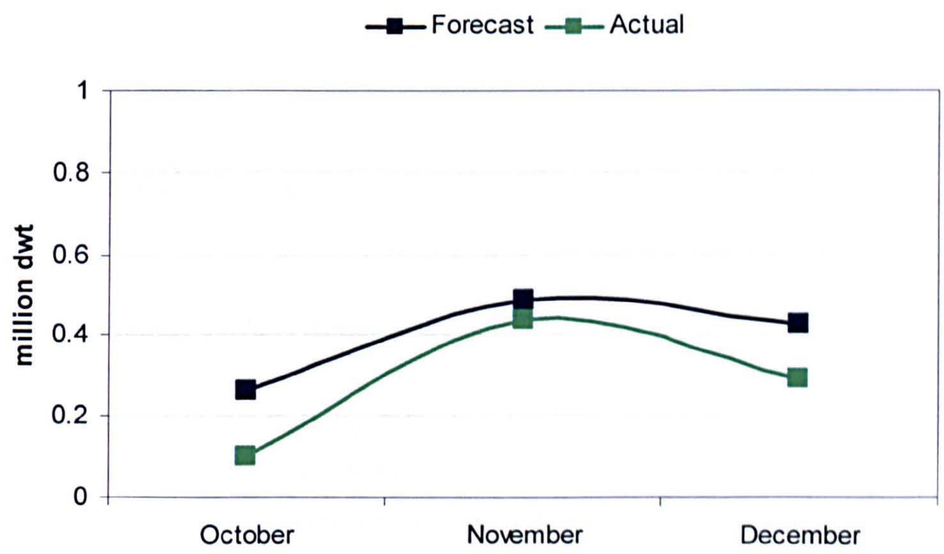


Figure 6-20: Forecast vs. Actual scrapped tonnage for three months period.

Calculated RMSE for the above prediction is 0.097 which represents the certainty of the model for the future prediction. The correlation coefficient between predicted and actual value for the above ANN models is 0.961, which shows a relatively high value of correlation for the model prediction. Therefore, the ANN approach is effective.

6.3 DEVELOPMENT OF THE ANN MODELS FOR SCRAP PRICES

For this part of the study, the structure of the data and the time series are similar to the previous study but the inputs and outputs have changed. The scrapped tanker tonnage was the only output in the previous model but there are two outputs now: the scrap prices in two different locations, Subcontinent and Far-East. All the inputs are listed below:

$x_1(t)$: USA Steel Production
$x_2(t)$: EU Steel Production
$x_3(t)$: China Steel Production
$x_4(t)$: Japan Steel Production
$x_5(t)$: South Korea Steel Production
$x_6(t)$: Steel Price
$x_7(t)$: OPEC Oil Production
$x_8(t)$: Non-OPEC Oil Production
$x_9(t)$: Oil World Trade
$x_{10}(t)$: EUR/$ Ex. Rate
$x_{11}(t)$: WON/$ Ex. Rate
$x_{12}(t)$: YEN/$ Ex. Rate
$x_{13}(t)$: Bunkers Price
$x_{14}(t)$: Product Tankers Building Price
$x_{15}(t)$: Aframax DH Building Price
$x_{16}(t)$: Suezmax DH Building Price
$x_{17}(t)$: VLCC DH Building Price
$x_{18}(t)$: Crude Carrier 105000dwt FRSingle Voyage
$x_{19}(t)$: Crude Carrier 150000dwt FRSingle Voyage
$x_{20}(t)$: Crude Carrier 300000dwt FRSingle Voyage
$x_{21}(t)$: Clean Carrier 70/85000dwt FRSingle Voyage
$x_{22}(t)$: Product Tankers DS/DH 5Years Market Value
$x_{23}(t)$: Aframax DS/DH 5Years Market Value
$x_{24}(t)$: Suezmax SH/DH 5Years Market Value
$x_{25}(t)$: VLCC SH/DH 5Years Market Value
$x_{26}(t)$: Clean Carrier- 40/45000dwt DB/DH 10 Years
$x_{27}(t)$: Tanker Fleet- 10000 DWT+ Supply
$x_{28}(t)$: Tanker Fleet- 10000 DWT+ Demand

$x_{29}(t)$: Tanker Fleet- 10000 DWT+ Util. Rate
$x_{30}(t)$: Tanker Order Book in Percent of Existing Fleet
$x_{31}(t)$: Scrapped Tankers

The structure of the inputs, outputs and patterns are represented in Table 6-4 and Table 6-5.

Similarly, to find out the best performing ANN models and training procedure in each section, the number of neurons and the best number of iterations are identified in the first place. Then learning rates of layers are identified and finally the momentum is added to obtain the most accurate network. Figure 6-1 illustrates the above stages.

Pattern Number	Patterns	Input (1) USA Steel Production	Input (2) EU Steel Production	Input (3) China Steel Production	Input (4) Japan Steel Production	Input (5) South Korea Steel Production	Input (6) Steel Price	Input (7) OPEC Oil Prod.	Input (8) Non-OPEC Oil Prod.	Input (9) Oil World Trade	Input (10) EUR/$ Ex. Rate	Input (11) WON/$ Ex. Rate	Input (12) YEN/$ Ex. Rate	Input (13) Bunkers Price	Input (14) Product Tankers Building Price	Input (15) Aframax DH Building Price	Input (16) Suezmax DH Building Price	Input (17) VLCC DH Building Price
01	January 1995	7.73	13.99	7.36	8.56	3.04	258.70	24.78	42.72	31.74	100.00	100.00	100.00	104.33	32.22	41.94	53.06	86.11
02	February 1995	7.73	14.17	7.36	8.56	3.04	258.70	24.78	42.61	31.85	102.46	100.00	101.85	106.30	32.78	42.22	53.33	86.11
03	March 1995	7.91	14.26	7.64	8.65	3.04	258.70	24.89	42.39	31.96	107.38	101.85	108.62	105.32	33.06	42.50	53.61	86.11
04	April 1995	7.91	14.26	7.64	8.56	3.13	269.57	24.89	42.17	32.07	107.08	103.38	120.00	107.28	33.33	42.78	53.89	86.11
05	May 1995	7.91	14.45	7.64	8.65	3.13	269.57	25.00	42.28	32.50	105.23	104.00	117.23	94.49	33.33	42.78	53.89	86.11
06	June 1995	7.91	14.63	7.64	8.65	3.13	295.65	25.00	42.39	32.83	107.38	104.00	118.77	82.68	33.33	42.78	53.89	86.11
07	July 1995	7.82	14.45	7.73	8.65	3.13	295.65	25.11	42.61	33.37	108.00	104.00	114.46	85.63	33.33	42.78	53.61	86.11
08	August 1995	7.82	14.36	7.73	8.65	3.13	284.78	25.22	42.72	33.15	103.08	102.15	104.92	87.60	33.06	42.78	53.33	86.67
09	September 1995	7.82	14.26	7.73	8.56	3.13	280.43	25.22	42.72	32.93	104.62	103.08	99.38	85.63	32.78	42.78	53.06	86.94
10	October 1995	7.82	14.26	7.73	8.47	3.13	269.57	25.22	42.83	32.61	104.92	103.38	99.38	87.60	32.78	42.78	52.78	87.50
11	November 1995	7.82	13.90	7.73	8.37	3.13	265.22	25.43	42.93	33.04	102.77	102.46	97.85	104.33	32.50	42.50	52.50	86.67
12	December 1995	7.82	13.44	7.82	8.37	3.13	254.35	25.65	43.15	33.48	103.08	102.15	96.92	101.38	32.22	42.22	52.50	85.83
13	January 1996	7.91	13.34	7.91	8.28	3.22	254.35	25.98	43.37	34.02	99.38	100.31	93.54	97.44	32.22	41.94	52.22	85.00
14	February 1996	7.91	13.34	7.91	8.10	3.22	250.00	25.87	43.37	34.02	101.23	100.62	96.00	107.28	32.22	41.67	51.94	84.72
⋮	⋮																	
115	July 2004	8.10	15.92	20.89	9.29	4.05	569.57	28.80	49.89	44.67	97.23	67.38	88.92	161.42	34.17	50.00	61.11	93.33
116	August 2004	8.37	15.74	21.53	9.39	3.96	569.57	29.13	49.78	45.00	97.23	68.31	90.77	173.23	34.44	51.67	63.89	96.94
117	September 2004	8.37	15.74	21.90	9.39	3.96	589.13	29.57	49.67	45.11	98.77	68.31	89.85	143.70	34.72	52.78	66.39	99.72
118	October 2004	8.47	15.74	22.64	9.48	3.96	589.13	29.67	49.89	45.33	102.15	70.15	93.54	154.53	36.39	55.00	68.06	103.06
119	November 2004	8.47	15.74	23.37	9.48	4.05	589.13	29.89	50.00	45.43	107.38	75.38	96.92	163.39	37.50	56.94	68.89	106.11
120	December 2004	8.47	15.64	24.20	9.48	4.05	589.13	30.00	50.22	45.54	108.92	76.00	96.92	174.21	39.44	59.72	69.72	109.44
	Maximum	8.83	15.92	24.20	9.48	4.05	589.13	30.00	50.22	45.54	108.92	104.00	120.00	174.21	39.44	59.72	69.72	109.44
	Minimum	6.99	13.25	7.36	7.36	3.04	193.48	24.57	42.17	31.74	68.00	47.08	69.54	52.17	25.00	33.33	41.94	63.89

Table 6-4: The scrap price model inputs (1 to 17)

Pattern Number	Patterns	Input (18)	Input (19)	Input (20)	Input (21)	Input (22)	Input (23)	Input (24)	Input (25)	Input (26)	Input (27)	Input (28)	Input (29)	Input ((30)	Input (31)	Output (1)	Output (2)
		Crude Carrier 105000dwt FRSingle V.	Crude Carrier 150000dwt FRSingle V.	Crude Carrier 300000dwt FRSingle V.	Clean Carrier 70/85000dwt FRSingle	Product Tankers DS/DH 5Years Market	Aframax DS/DH 5Years Market Value	Suezmax SH/DH 5Years Market Value	VLCC SH/DH 5Years Market Value	Clean Carrier- 40/45000dwt DB/DH 10 Y	Tanker Fleet- 10000 DWT+ Supply	Tanker Fleet- 10000 DWT+ Demand	Tanker Fleet- 10000 DWT+ Util. Rate	Tanker Order Book in Percent of Existing	Tankers Sold for Scrapping	Subcontinent Scrap Price	Far-East Scrap Price
01	January 1995	14.81	14.81	11.11	17.83	21.89	29.81	33.21	52.22	16.76	260.87	218.48	83.91	9.2	4.18	175.00	151.56
02	February 1995	12.96	12.96	8.02	16.59	22.26	29.81	33.96	51.85	17.31	259.78	216.30	83.48	9.1	4.18	179.69	151.56
03	March 1995	12.35	13.58	9.88	16.10	22.64	29.81	34.34	51.85	17.85	258.70	215.22	83.04	9.0	4.18	187.50	151.56
04	April 1995	14.20	12.35	8.02	13.87	23.02	29.81	34.34	51.85	18.18	258.70	213.04	82.61	8.8	3.98	187.50	156.25
05	May 1995	11.73	12.96	6.79	13.37	23.77	29.81	34.34	51.85	18.73	258.70	215.22	83.91	8.6	3.98	187.50	153.13
06	June 1995	12.96	13.58	13.58	17.34	24.15	29.81	34.34	51.85	19.05	258.70	218.48	85.00	8.4	3.98	187.50	153.13
07	July 1995	14.20	19.75	19.14	19.32	24.15	29.81	34.72	52.22	19.05	258.70	220.65	85.87	8.1	1.02	187.50	153.13
08	August 1995	14.20	14.81	18.52	18.82	24.15	30.19	35.47	53.33	19.05	258.70	220.65	85.87	7.9	1.02	187.50	153.13
09	September 1995	14.20	14.81	14.81	18.33	24.15	30.57	36.23	53.70	19.05	258.70	220.65	85.87	7.6	1.02	190.63	151.56
10	October 1995	14.20	14.81	11.11	20.06	24.15	30.57	36.23	53.70	19.05	258.70	220.65	86.09	7.2	1.55	193.75	146.88
11	November 1995	13.58	16.05	16.67	21.80	24.15	30.57	36.23	53.70	19.05	258.70	221.74	86.52	7.0	1.55	184.38	143.75
12	December 1995	14.81	16.05	16.67	22.04	24.15	30.57	36.23	53.70	19.05	258.70	222.83	86.96	6.8	1.55	179.69	143.75
13	January 1996	17.90	17.28	16.05	21.05	24.15	30.94	37.36	54.81	19.05	258.70	223.91	87.39	6.5	0.89	189.06	143.75
14	February 1996	17.28	17.28	19.75	20.06	24.15	30.94	38.49	55.93	19.05	259.78	230.43	86.96	6.3	0.89	190.63	143.75
⋮	⋮																
115	July 2004	32.72	52.47	79.01	30.71	32.45	49.81	60.00	87.41	23.31	301.09	271.74	90.43	26.8	0.67	387.50	343.75
116	August 2004	32.10	45.68	62.35	26.75	33.58	51.32	62.26	90.74	24.07	302.17	275.00	91.30	26.9	0.65	425.00	371.88
117	September 2004	32.72	46.30	63.58	26.75	34.72	53.21	66.79	93.33	25.71	303.26	278.26	92.17	27.0	0.40	415.63	337.50
118	October 2004	80.25	107.41	136.42	40.37	36.60	54.72	71.32	101.48	26.91	304.35	280.43	93.04	27.1	0.01	404.69	368.75
119	November 2004	85.19	125.93	189.51	56.97	39.62	56.98	72.45	105.19	28.11	305.43	282.61	93.70	26.9	0.44	421.88	375.00
120	December 2004	68.52	77.78	125.31	53.00	39.62	56.98	72.45	105.19	28.11	306.52	284.78	94.35	26.5	0.29	434.38	356.25
	Maximum	85.19	125.93	189.51	60.43	39.62	56.98	72.45	105.19	28.11	306.52	284.78	94.35	27.14	4.46	434.38	385.94
	Minimum	9.88	9.26	6.79	10.15	17.36	23.02	33.21	48.89	12.07	258.70	213.04	82.61	6.02	0.01	109.38	106.25

Table 6-5: The scrap price model inputs (18 to 31) and the two outputs

6.3.1 APPLICATION OF STATIC ANNS FOR MODELLING OF THE MONTHLY SCRAP PRICES

To implement the first part of this study, a feed-forward MLP network with error back- propagation learning algorithm has been developed to model the monthly scrap prices for both Subcontinent and Far-East scrapyards. Therefore, there are 31 inputs to this ANN model and 2 outputs. Figure 6-21 illustrates the system of inputs to the above ANN and the corresponding outputs.

After initial randomisation of the patterns, with the same argument as section 6-2-1, 20% of the patterns (24 months observation) in each time series are used to test the performance of each particular neural network and 70% of the patterns (84 months observation) in each time series are used to train the ANN. The rest of the patterns 10% (12 months observation) are used to validate the networks. This network includes one hidden layer. The activation function is used for this study is hyperbolic tangent function *(tanh)*, Equation 4-3, which will give an output in the range [-1, 1]. For this study, normalisation has been performed as the pre-processing of the data.

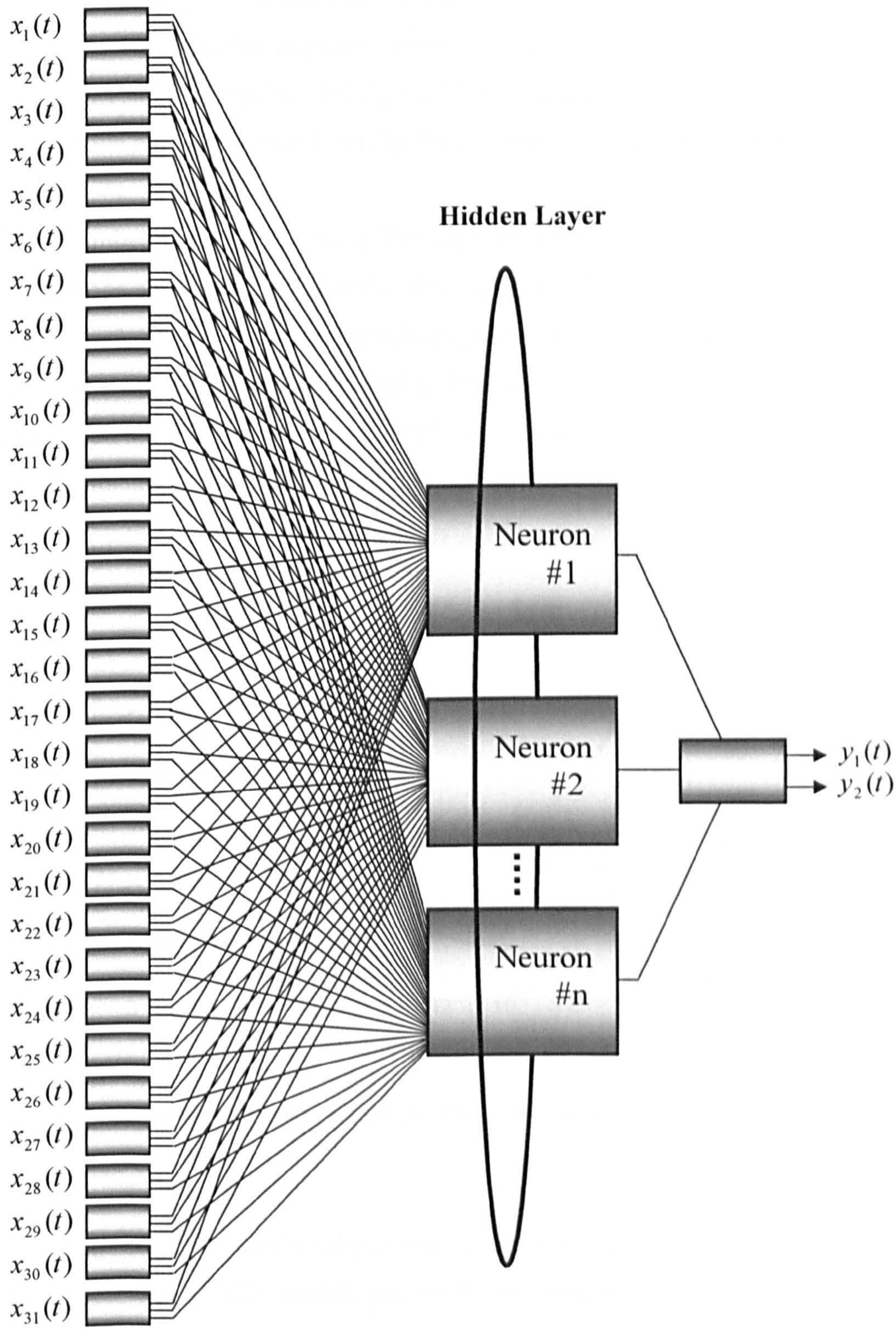


Figure 6-21: The static ANN model for the monthly scrap prices

Similarly, at the beginning of the modelling initial weights are randomised but, after each stage, the current weight settings are stored along with the components to be re-loaded for the next stage of training. All the ANN architectures and their configurations have been optimised based on the average of the five random starts.

To decide the number of neurons in the hidden layer of the static ANN, various numbers of neurons have been considered and subsequently the mean square error of each neural network has been measured to identify the most accurate model. According to Figure 6-22, the minimum testing MSE of 32.067 happens in the static ANN model which has 5 neurons in its hidden layer. The value of the corresponding correlation coefficient for this network is 0.997.

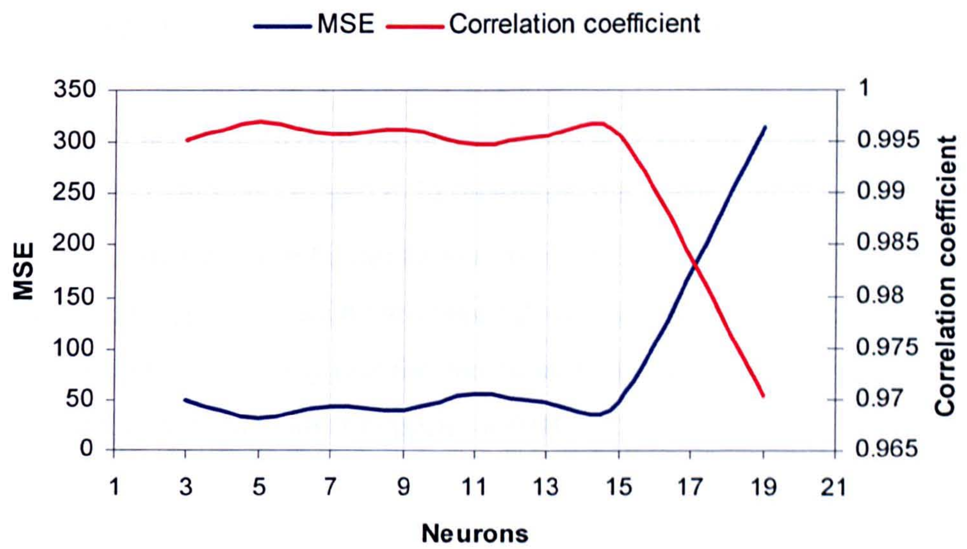


Figure 6-22: Impact of the neurons on the static ANN model for scrap prices

The next step is to find out the best point in time to stop the training of the ANN. To ensure the best performing ANN model, various iterations have been tried and their performances tested. The error of the model and its corresponding correlation coefficient for different iterations is presented in Figure 6-23. As this plot suggests, the overall positive correlation coefficient of the model is relatively high. The maximum value of this coefficient belongs to iteration 9000. Simultaneously, the minimum MSE has happened in this point. The MSE of the training stage is measured as 0.000662 and the standard deviation is measured as 0.00030 for the above static ANN.

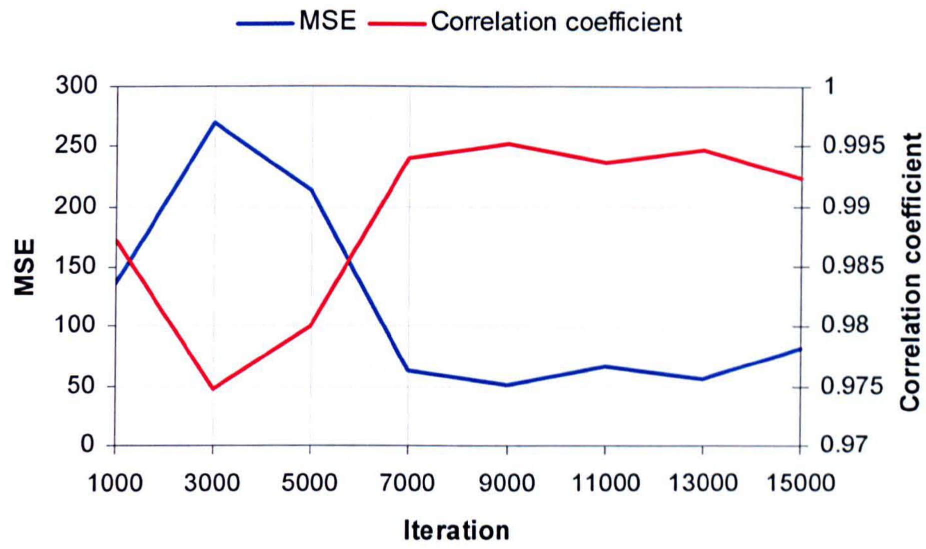
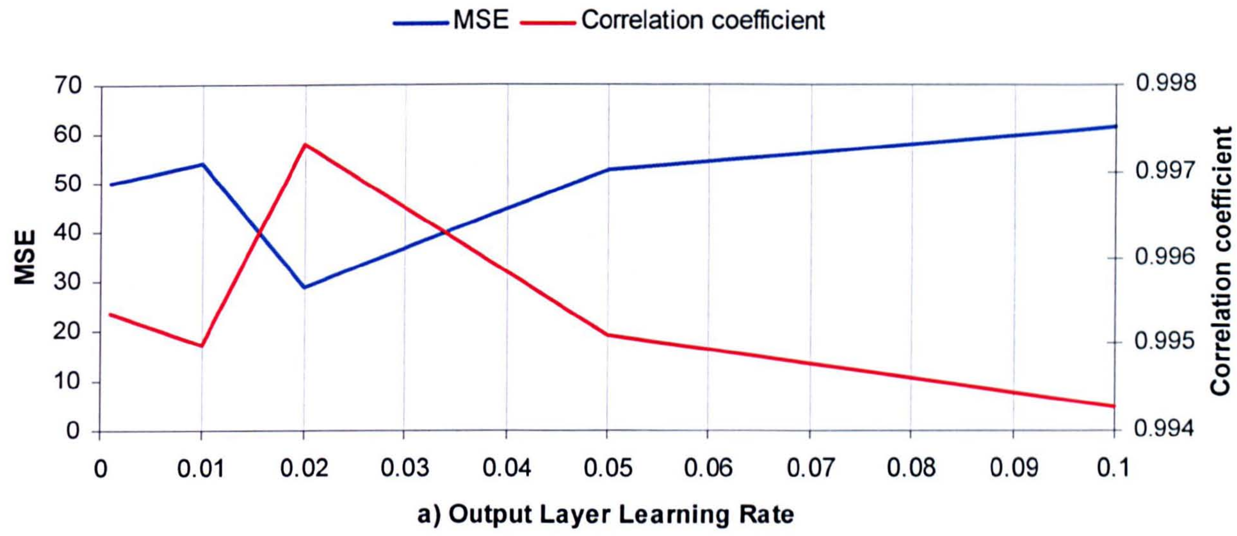


Figure 6-23: Impact of the different iterations to the static ANN model for monthly scrap prices

So far, the most accurate ANN is the one with 5 neurons in the hidden layer which is trained for 9000 iterations. Figure 6-24 represents the impact of the learning rate γ variations for both hidden and output layer. MSE and correlation coefficient are measured simultaneously in each particular rate for both layers. These plots suggest that the most accurate learning rate for the output layer occurs at 0.02 and the best learning rate for the hidden layer happens at 0.01.



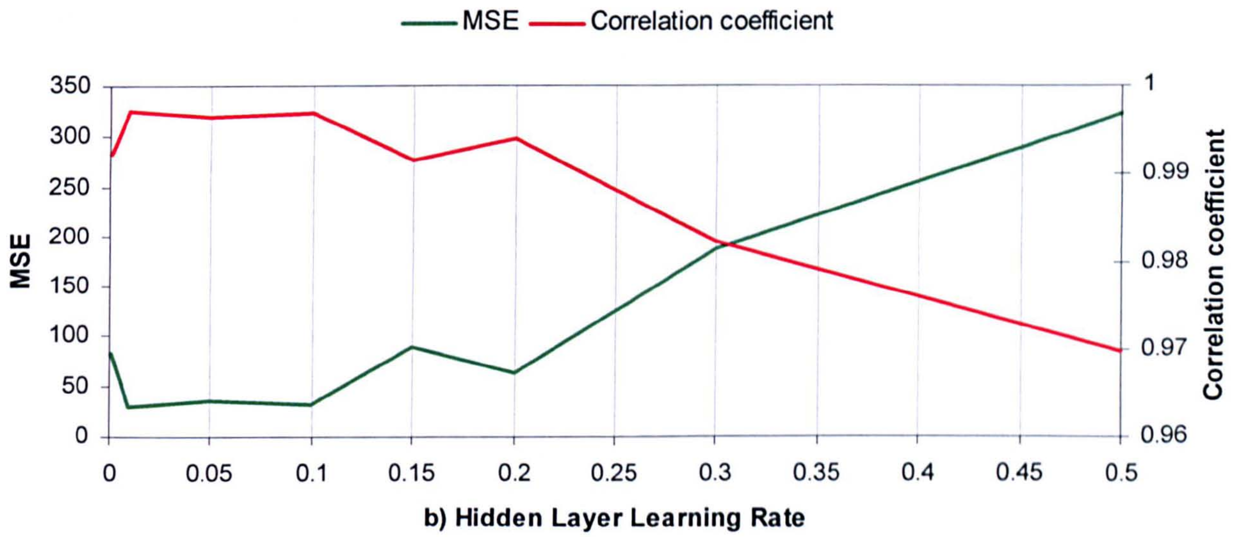


Figure 6-24: a) Impact of different learning rates in the output layer on MSE and correlation coefficient for the static ANN model for scrap prices
b) Impact of different learning rates in the hidden layer on MSE and correlation coefficient for the static ANN model for scrap prices

The last variable of the above ANN, which needs to be adjusted correctly, is the momentum α . For this reason, various momentums are considered, for the above static ANN model, and their performances tested. Figure 6-25 explains the variation of the mean square errors and the correlation coefficients for different momentum values. As the plot suggests the best momentum for the networks occur at 0.7.

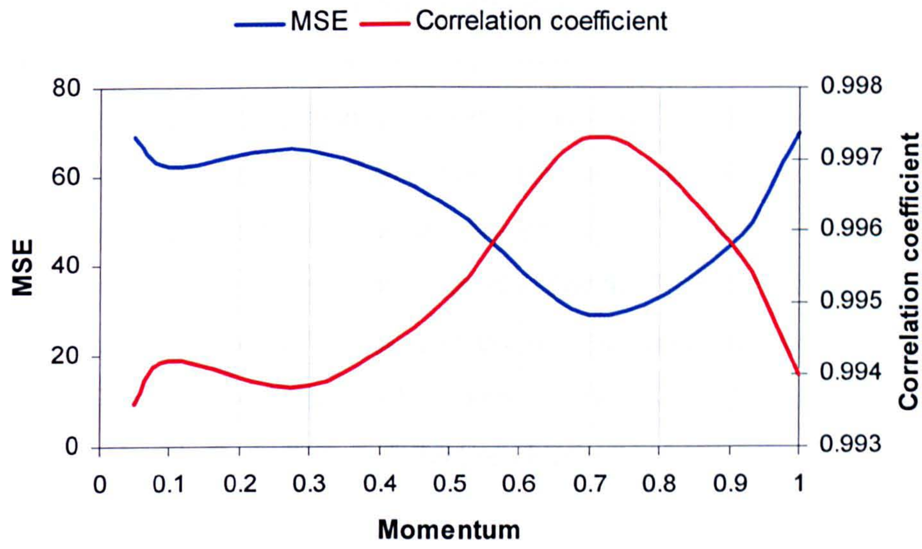


Figure 6-25: Impact of adding different momentums to the static ANN model for monthly scrap prices

According to the above studies for various static ANN model parameters, the best ANN model and the training is the one which includes:

- one hidden layer
- five neurons (or PEs)
- the hyperbolic tangent activation function.
- the learning rate of 0.01 for the hidden layer $(\gamma = 0.01)$
- the learning rate of 0.02 for the output layer $(\gamma = 0.02)$
- the momentum of 0.7 for the networks $(\alpha = 0.7)$

6.3.1.1 Sensitivity Analysis of the static ANN model for the monthly scrap prices

As the architecture of the static ANN for the monthly scrap prices is identified and the training of the neural networks are completed (see section 6-3-1), it is possible to determine the sensitivity of the outputs with respect to each individual input. The sensitivity analysis (see section 4-5-2), which is employed for this model, is based on Equation 4-13.

Figure 6-26 represents the sensitivity analysis of the static ANN model for the monthly scrap prices. It explains that steel production has the most influence on the model. For the Subcontinent prices, South-Korean steel production has the highest value, and Japan's steel production is the second most valuable. EU's and China's have the third and forth highest values respectively but USA's steel production does not have a significant influence on the model. For the Far-East prices, the pattern is almost the same but the EU's steel production is the second most valuable and Japan's steel production is the third. In both case, the South-Korea's steel production is more than double that of Japan's steel production.

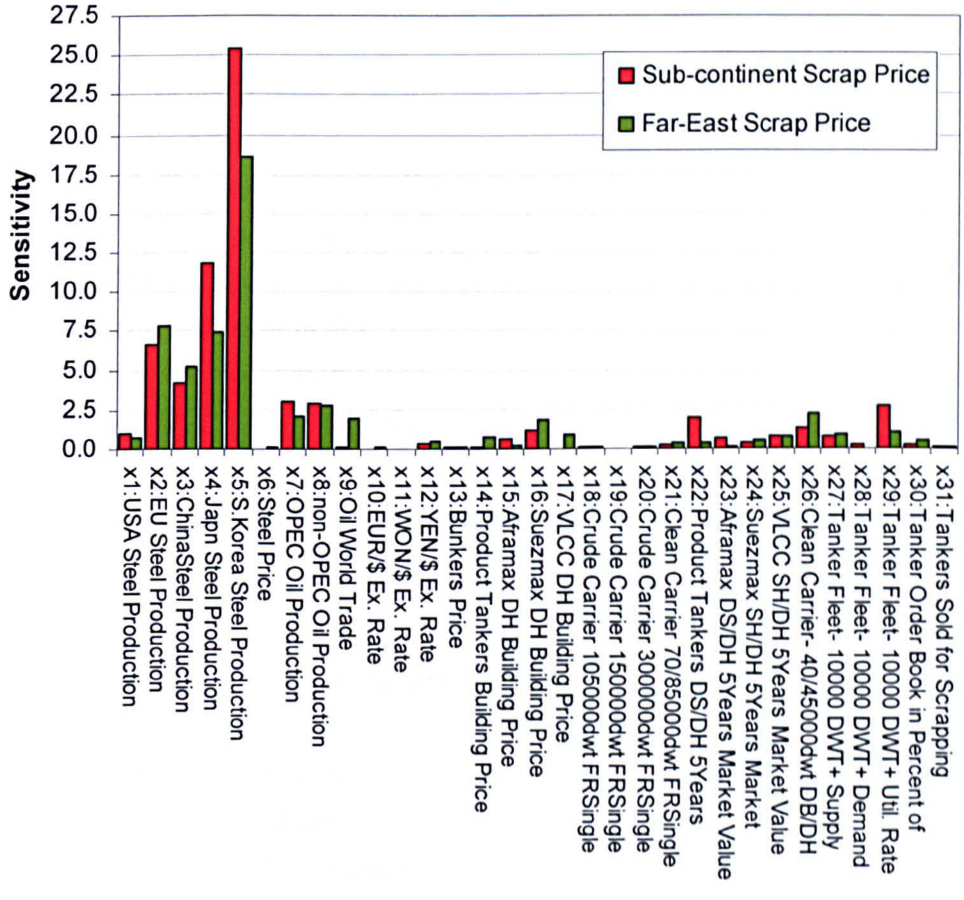


Figure 6-26: Sensitivity about the mean for all the inputs to the static ANN model for the monthly scrap prices

OPEC and non-OPEC oil production have also shown an influence on the model. The influence of the OPEC is slightly higher than the non-OPEC oil production for the Subcontinent prices but it is lower for the Far-East prices. Product Tanker building prices and the tanker fleet utilisation rate have shown an influence on the model for Subcontinent. Suezmax (double hull) Tanker building prices and Clean Carrier (40/45000 dwt) 10 years Second-hand values have shown an influence on the model for Far-East. On the contrary, tankers sold for scrapping, freight rates, bunker price (as the operating cost of the ship) and exchange rates have very small impacts on the model. There are a few reasons for the low influence of the above variables, like freight rates or steel price, to the model. Firstly, it is possible that their influence is offset by influence of another variable to the model or, secondly, there is not enough data to carry out an accurate sensitivity analysis.

Similarly, according to the studies which are carried out for the static ANN model for the monthly scrap prices, the most sensitive parameters of the model for the Subcontinent prices are (Figure 6-26):

1. x_5 : South Korea steel production
2. x_4 : Japan steel production
3. x_2 : EU steel production
4. x_3 : China steel production
5. x_7 : OPEC oil production
6. x_8 : Non-OPEC oil production
7. x_{29} : Tanker fleet utilisation rate
8. x_{22} : Product tankers 5 years market value

and for the Far-East prices are:

1. x_5 : South Korea steel production
2. x_2 : EU steel production
3. x_4 : Japan steel production
4. x_3 : China steel production
5. x_8 : Non-OPEC oil production
6. x_9 : Oil world trade
7. x_{16} : Suezmax Double Hull building price
8. x_{26} : Clean carriers 10 years market value

Negligible inputs to the above models are:

1. x_6 : Steel price

2. x_{18} to x_{21} : Freight rates
3. x_{10} to x_{12} : Exchange rates (a small sensitivity for Yen/USD)
4. x_{13} : Bunker price
5. x_{31} : Tankers sold for scrapping
6. x_{28} : Tankers demand

6.3.2 DEVELOPMENT OF THE DYNAMIC ANN MODEL FOR THE SCRAP PRICE PREDICTION

A dynamic ANN is implemented to forecast the monthly scrap prices for both Subcontinent and Far-East scrapyards. The overall structure of the dynamic neural networks for this study is illustrated in Figure 6-27. It also explains the system of inputs to the network and the corresponding outputs.

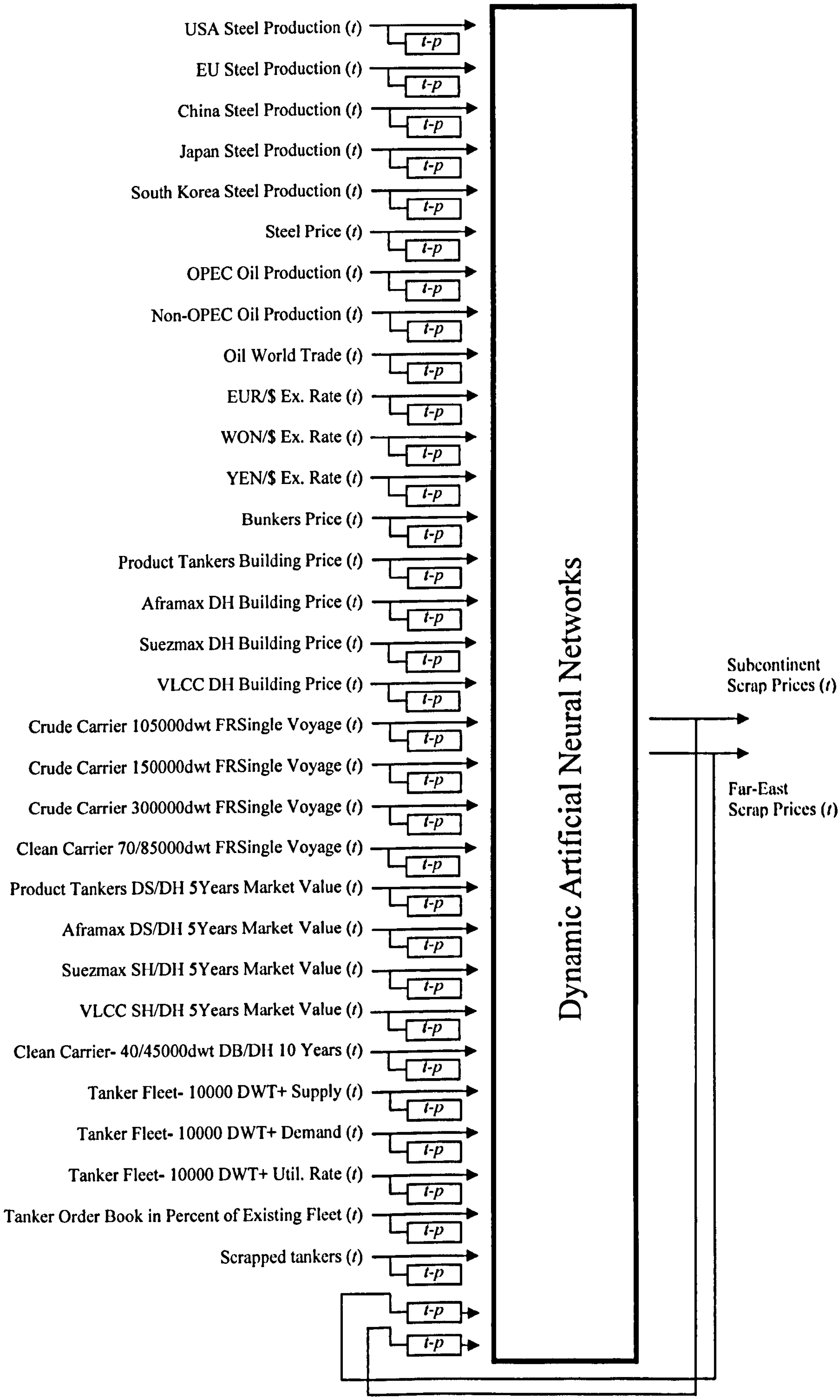


Figure 6-27: The system of inputs to the dynamic ANN and the corresponding output

To develop this dynamic ANN model, a feed-forward MLP network with error back-propagation learning algorithm has been created. Then a sequential framework has been added to that by adding short-term memory in the form of a delay line, as before. A section of the time series of the form $[x(t), x(t-1),..., x(t-p)]$, and $[y(t-1),..., y(t-p)]$ are used as input for the network. The delay line is order of p, and the desired outputs are $y(t)$ for each output. Similar to the previous modelling, an integer value is substituted for p and several ANNs have been trained. The best performing ANN is noted and then the process is repeated with a different integer value for p. When sufficient p values have been investigated the MSE are compared and the minimum is chosen. Consequently, forecast errors are measured in order to judge how good that models are in terms of their prediction abilities.

To start the modelling, the data has been split into two groups: training and validation. In the previous dynamic ANN study, section 6-2-2, 80% of the data (94 months) in each time series had been chosen to train the neural networks and 20% (23 months) for validation. But for this dynamic ANN, the above combination of the categories could not produce an accurate and well performing model because of the special condition in the last part of the output time series (Figure 6-28). The training set is extended to cover a part of the significant changes of the last part of the output time series (point T in the figure). Using an 85% training (100 patterns) and 15% validation (17 months) combination puts the sharp increase in the training section and the ANN can learn the process of such an increase. This can also provide a visible indication of the quality of training because if this was poor, the sharp increase would not materialise.

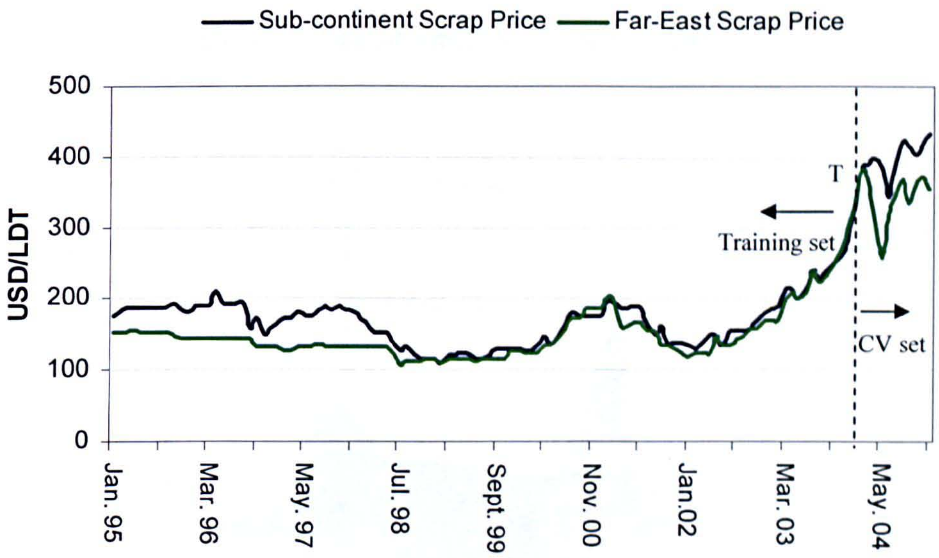


Figure 6-28: The actual outputs (scrap price in Subcontinent and Far-East) time series for the dynamic ANN investigations. They are split into two sets of Training and Cross Validation.

Three last months, in each time series, are not taken into account for the training purposes because their data is needed to check the prediction performance of the final neural networks.

All the dynamic ANNs which are considered for this modelling are included one hidden layer with the hyperbolic tangent (*tanh*) activation function, Equation 4-3, in their hidden layers. The procedure of this modelling is illustrated in Figure 6-1.

To investigate the performance of the different dynamic ANN configurations, various numbers of neurons and iterations, for a network with one month delay (t-1), have been changed simultaneously. Then the mean square error of the testing stage for each ANN has been measured to find out the most accurate combination. Figure 6-29 and Figure 6-30 represent the mean square error surface of the above measurements.

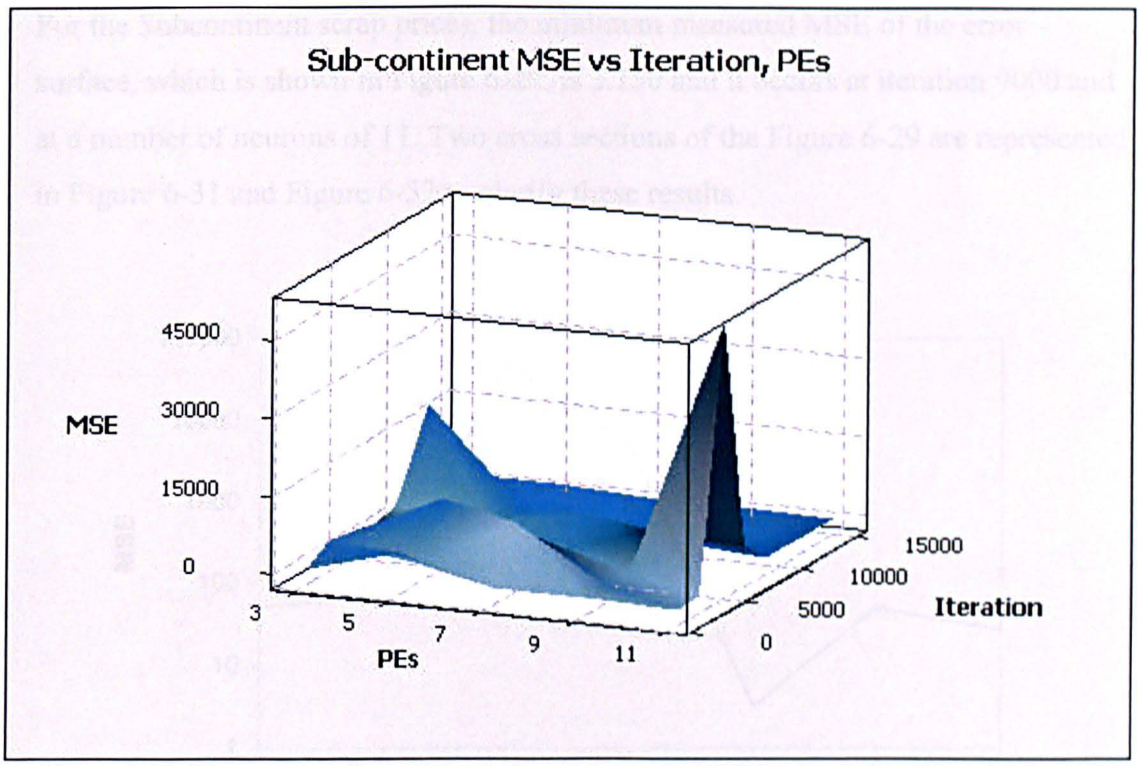


Figure 6-29: The error surface of the designed dynamic ANN for the Subcontinent price with one month delay (*p*=1)

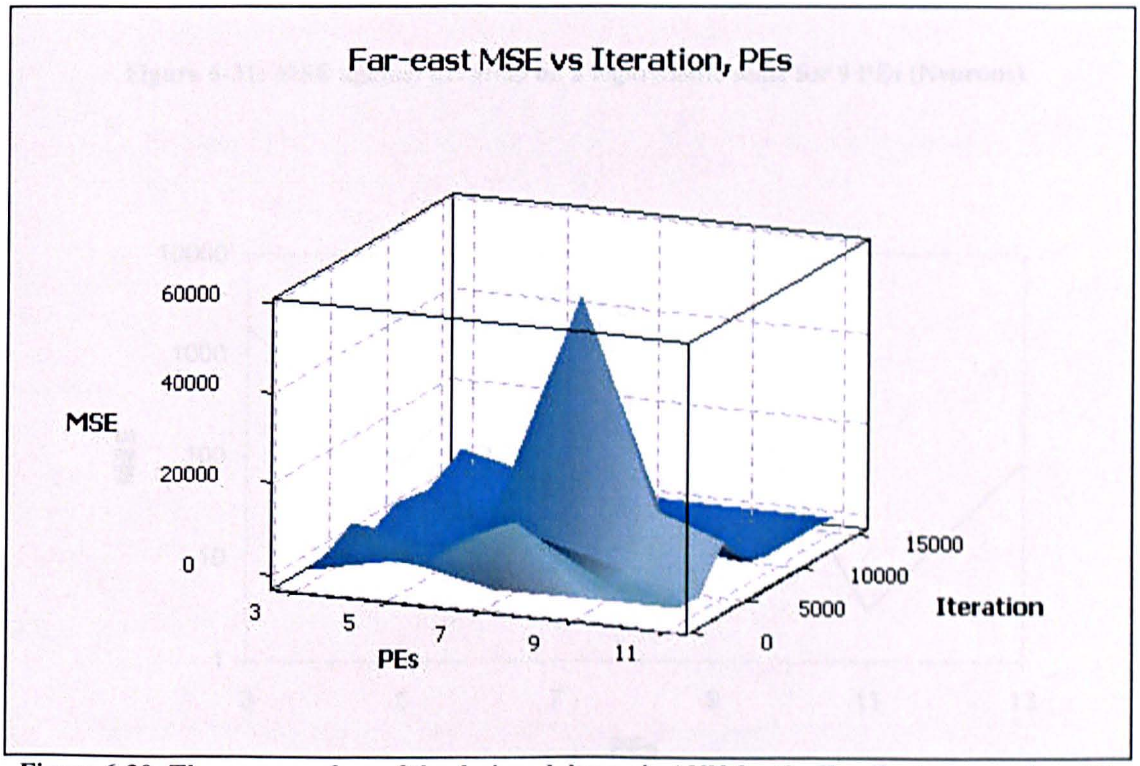


Figure 6-30: The error surface of the designed dynamic ANN for the Far-East scrap price with one month delay (*p*=1)

For the Subcontinent scrap prices, the minimum measured MSE of the error surface, which is shown in Figure 6-29, is 3.150 and it occurs at iteration 9000 and at a number of neurons of 11. Two cross sections of the Figure 6-29 are represented in Figure 6-31 and Figure 6-32 to clarify these results.

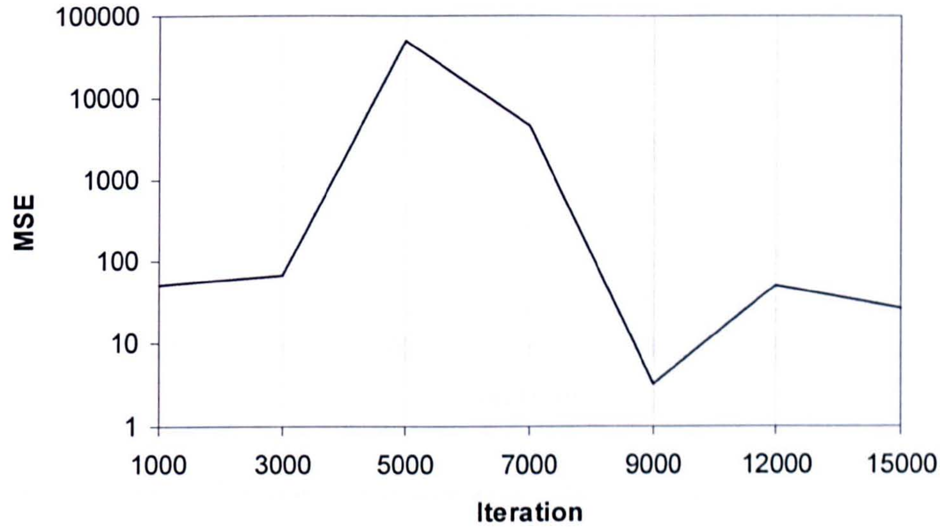


Figure 6-31: MSE against iteration on a logarithmic scale for 9 PEs (Neurons)

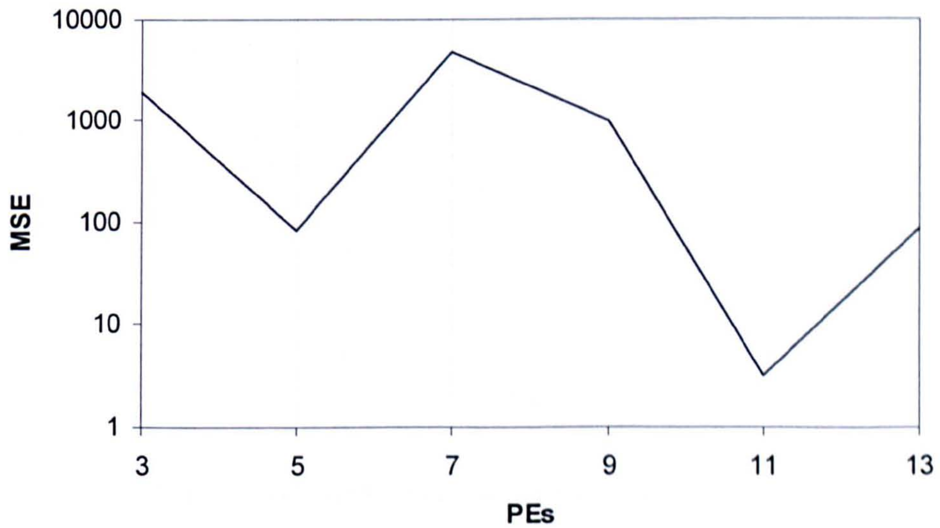


Figure 6-32: MSE against PEs (Neurons) on a logarithmic scale for iteration 9000

For the Far-East scrap prices, the minimum measured MSE of the error surface, which is shown in Figure 6-30, is 3.519 and it occurs at iteration 12000 and at a number of neurons of 5. Similarly, two cross sections of the Figure 6-30 are represented in Figure 5-33 and Figure 5-34 to clarify these results.

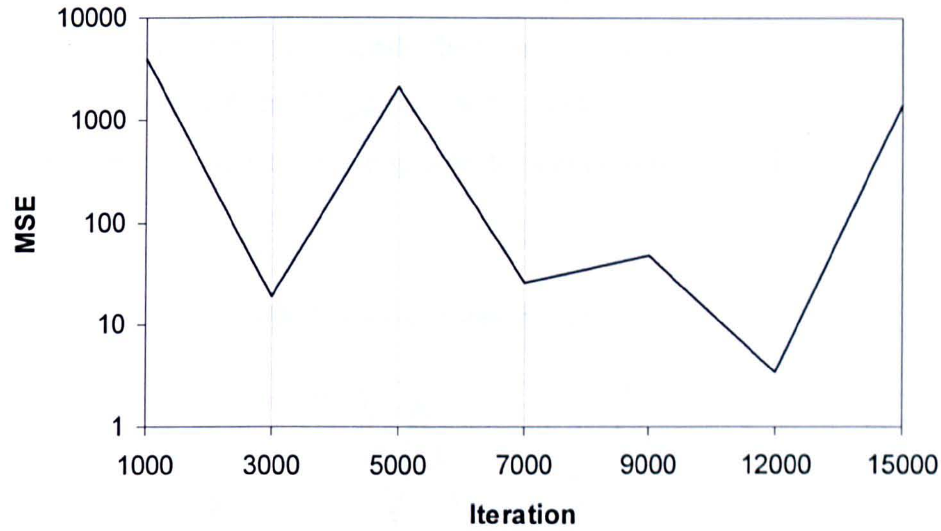


Figure 6-33: MSE against iteration on a logarithmic scale for 5 PEs (Neurons)

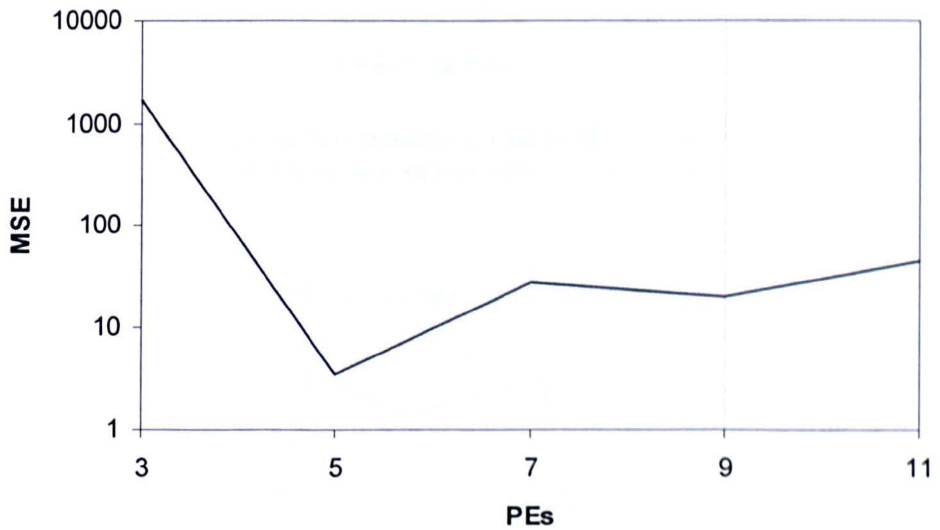


Figure 6-34: MSE against PEs (Neurons) on a logarithmic scale for iteration 12000

The average MSE of the training stage for the above ANNs are measured 0.030 and 0.0007 respectively. Also, the standard deviations of these training errors are 0.052

and 0.0002. These values suggest that the second ANN has lower training error than the first one.

To identify the best learning rate γ for the training of the above dynamic ANN, different rates are set and consequently the mean square errors of the testing stages are measured. Figure 6-35 and Figure 6-36 are shown that the minimum error occur at $\gamma = 0.1$ which means the best learning rate, for both outputs, is 0.1.

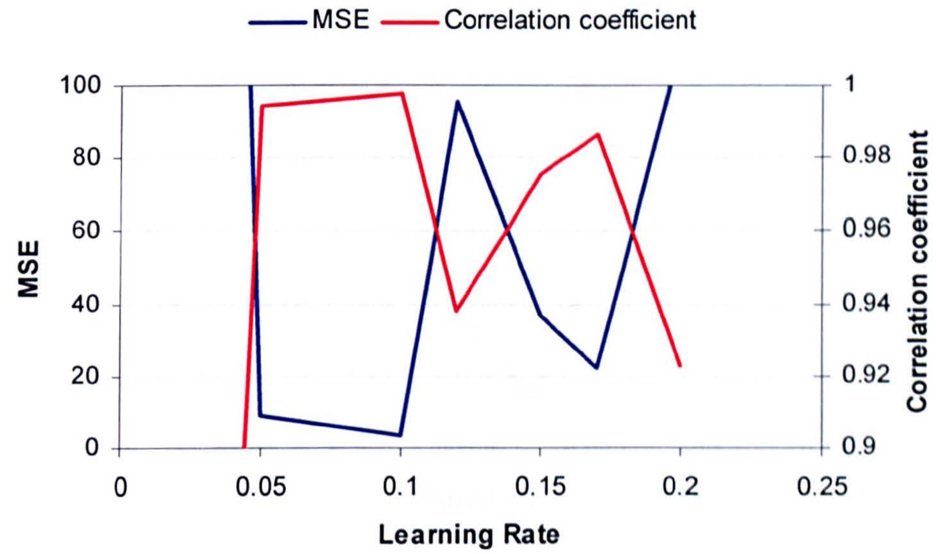


Figure 6-35: Impact of the different learning rate on the dynamic ANN with one month delay for the Subcontinent scrap prices

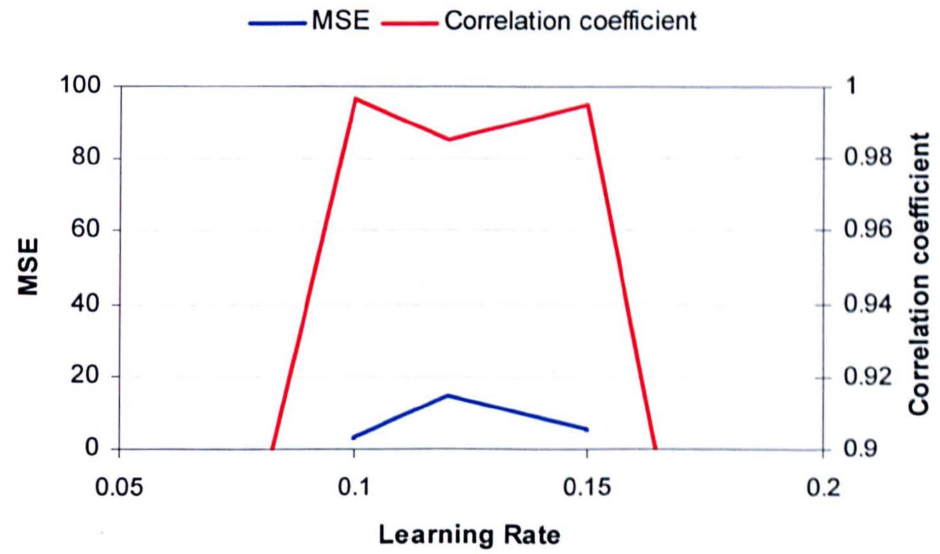


Figure 6-36: Impact of the different learning rate on the dynamic ANN with one month delay for the Far-East scrap prices

To investigate the effect of adding momentum to the above neural networks, different momentums are added and subsequently the mean square error of the trainings calculated to find out the best performing dynamic ANN. Figure 6-37 and Figure 6-38 represent the behaviour of the networks with different momentums and they suggest that the best networks happens at momentum zero for both Subcontinent and Far-East.

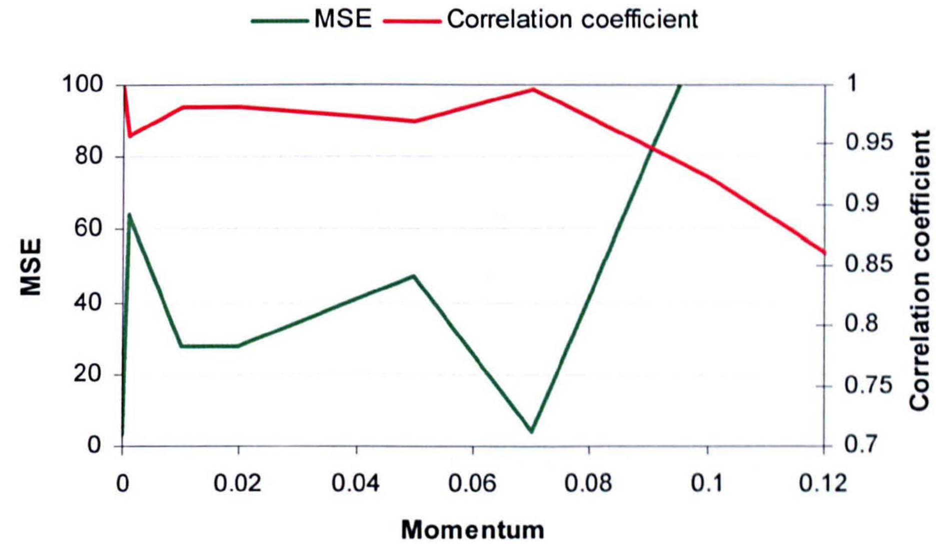


Figure 6-37: Various Momentums of the dynamic ANN model with one month delay for the Subcontinent scrap prices

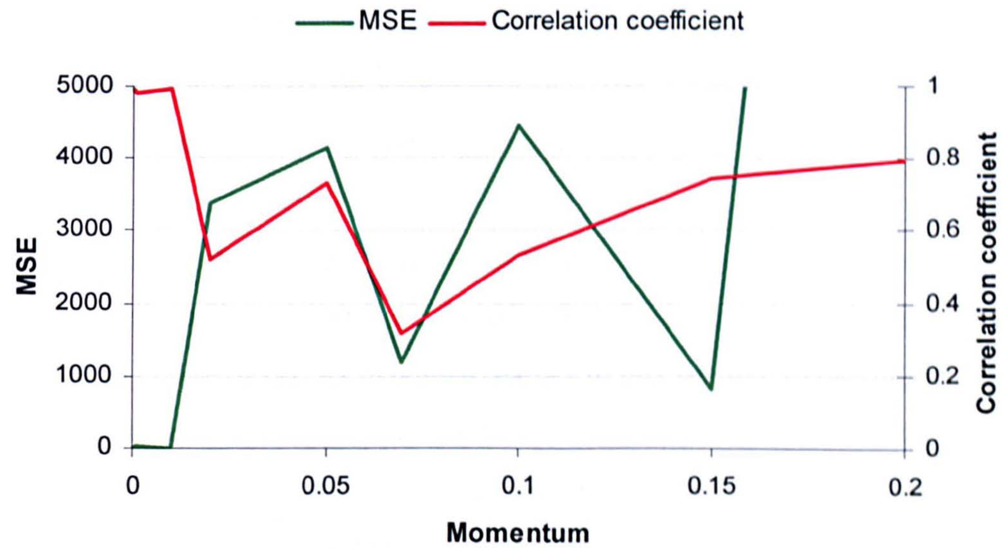


Figure 6-38: Various Momentums of the dynamic ANN model with one month delay for the Far-East scrap prices

The values of the mean square errors are 3.150 and 3.519 and the corresponding correlation coefficients are 0.997 and 0.996 respectively which explain a high positive correlation for these models.

So far, the parameters of the dynamic ANN model and training specifications with one month time delay (p=1) for each location are specified. The ANN model and training specifications for the Subcontinent prices are:

- one hidden layer
- 11 neurons (or PEs)
- 9000 iterations for the network
- the hyperbolic tangent activation function
- the learning rate of 0.1 for both layers $(\gamma = 0.1)$
- the momentum of zero for the networks $(\alpha = 0)$

and the ANN model and training specifications for the Far-East scrap prices are:

- one hidden layer
- 5 neurons (or PEs)
- 12000 iterations for the network
- the hyperbolic tangent activation function
- the learning rate of 0.1 for both layers $(\gamma = 0.1)$
- the momentum of zero for the networks $(\alpha = 0)$

In the next section, the impacts of different time delays for each of the above ANN models are analysed to understand if there is a possibility to achieve better performance (Figure 6-39). As explained earlier in this section, expanding the delay window can increases the experience and the power of the decision making of the network because it has access to earlier data as well as data from one month previous. Therefore, according to the system of inputs to the above dynamic ANN and the corresponding outputs (Figure 6-27) changing the value of the delay p will change the structure of the neural network. A bigger value of p means more accessibility to the past data.

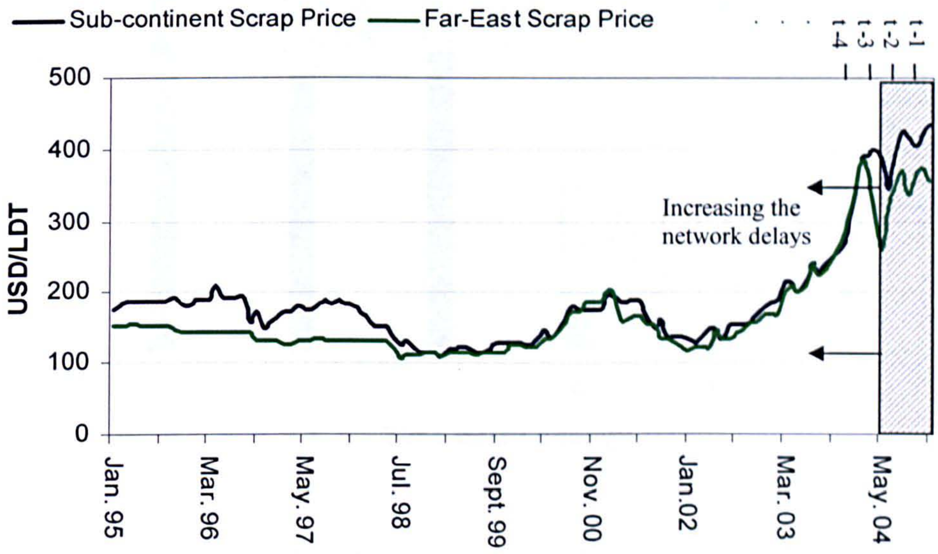


Figure 6-39: The time delay window for the dynamic ANN models of the scrap prices. The size of the delay window can be varied by changing the structure of the neural networks.

Based on the previous analysis, new dynamic ANN models are implemented to analyse the impact of 2, 3, 4, 5 and 6 months delay (p=2, p=3, p=4, p=5 and p=6). The number of iterations and neurons are changed simultaneously to plot the mean square error surface of each dynamic ANN for both locations. The MSE results of these analyses, for the best ANN chosen for each delay, are represented in Figure 6-40 and Figure 6-41 for the Subcontinent and Far-East prices respectively. These figures indicate that ANN with 5 months delays has more accuracy than the others in both locations. There is no correlation between the most accurate ANN in each time delay and other ANNs as their specifications, e.g. the number of PEs or iterations, are different.

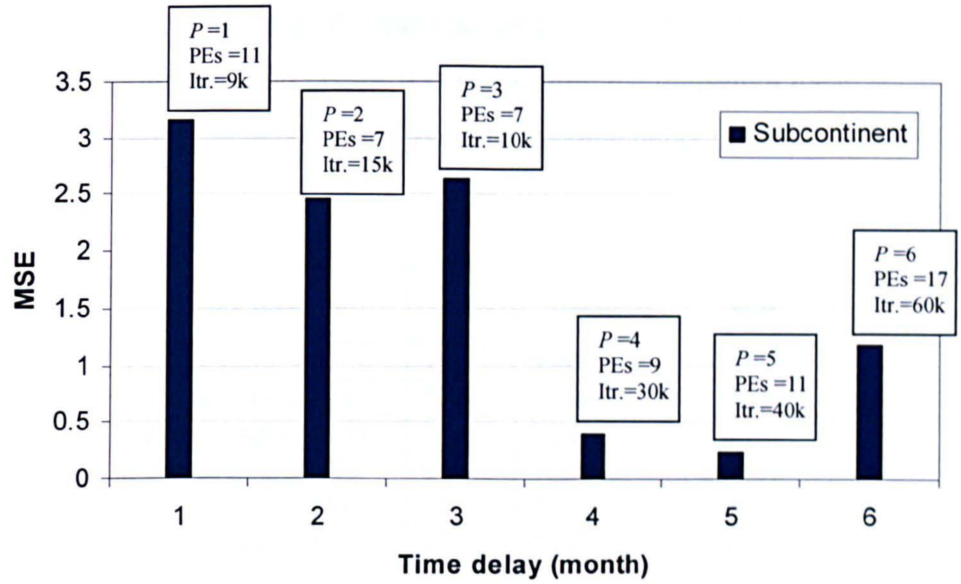


Figure 6-40: MSE against time delay of the best obtained dynamic ANN model for the monthly scrap prices in the Subcontinent scrapyards. Box annotations indicate the specification of the ANN for each time delay.

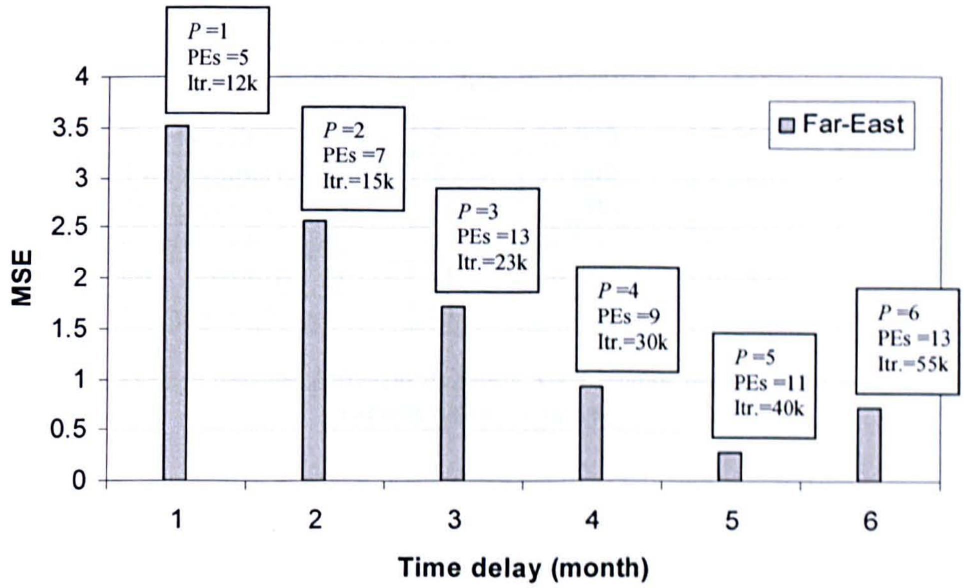


Figure 6-41: MSE against time delay of the best obtained dynamic ANN model for the monthly scrap prices in the Far-East scrapyards. Box annotations indicate the specification of the ANN for each time delay.

Table 6-6 illustrates the specifications of the best neural network for each time delay. It is also a comparison between the most accurate ANNs in each particular time delay model for Subcontinent scrap prices. It is shown that the minimum mean square error happens for the neural networks with five months time delay i.e. p=5.

Delay window	PEs	Iteration	MSE
One month (t -1)	11	9,000	3.1501
Two months (t - 2)	7	15,000	2.4548
Three months (t - 3)	7	10,000	2.6359
Four months (t - 4)	9	30,000	0.3927
Five months (t - 5)	11	40,000	0.2422
Six months (t - 6)	17	60,000	1.16788

Table 6-6: Comparison between different ANN models for the Subcontinent prices with various delay windows

Similarly, for the Far-East scrap prices, Table 6-7 shows that the minimum mean square error happens for the neural networks with five months time delay i.e. p=5.

Delay window	PEs	Iteration	MSE
One month (t -1)	5	12,000	3.5193
Two months (t - 2)	7	15,000	2.5681
Three months (t - 3)	13	23,000	1.7238
Four months (t - 4)	9	30,000	0.9345
Five months (t - 5)	11	40,000	0.2844
Six months (t - 6)	13	55,000	0.7323

Table 6-7: Comparison between different dynamic ANN models for the Far-East prices with various delay windows

6.3.2.1 MONTHLY SCRAP PRICES PREDICTION USING DYNAMIC ANN MODEL

The best performing dynamic ANN model for the monthly scrap prices (in Subcontinent and Far-East scrapyards) is identified in previous section. In this section, the performance of the above neural networks is analysed. Similar to the section 6-2-2-1, for the scrapped tonnage modelling, unseen data is used to evaluate and verify the predictability of this ANN model.

As explained before (see section 6-3-2), in the beginning of the modelling process the last three months of data is pulled out of the training and testing stages. Therefore, these data are available to put in the obtained dynamic ANN model and check the performance of the model by comparing the prediction with the real values.

Figure 6-42 and Figure 6-43 represent the actual measurements versus prediction of the dynamic ANN model of scrap prices for both locations, Subcontinent and Far-East respectively.

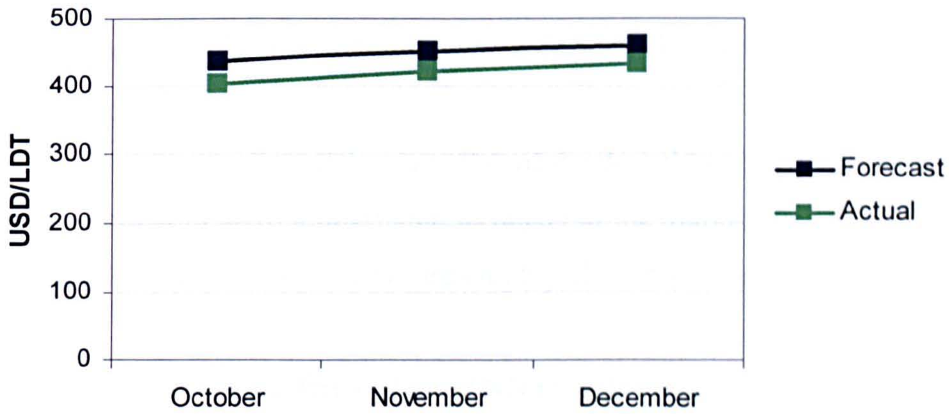


Figure 6-42: The Subcontinent Forecast vs. Actual scrap price for three months period

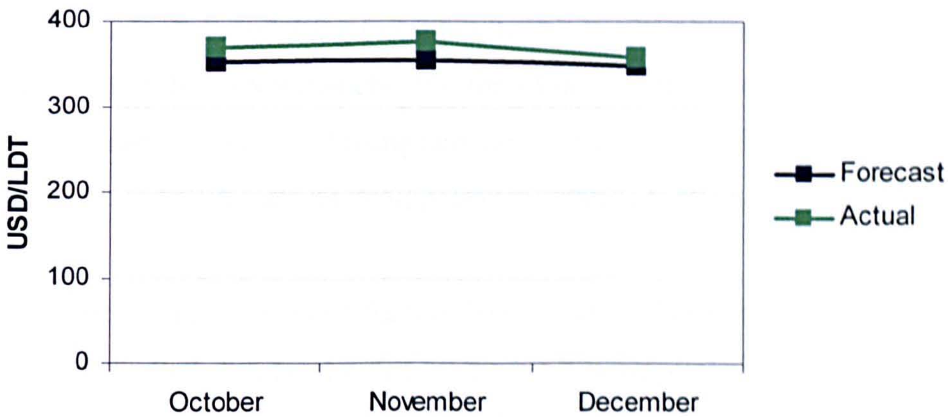


Figure 6-43: The Far-East Forecast vs. Actual scrap price for three months period

Similarly, for the obtained dynamic ANN model for the scrap prices in Subcontinent and Far-East, calculated RMSEs are 15.57 and 15.61 respectively. These numbers show that the above model is slightly more accurate when it predicts the Far-East prices. To have a better understanding of how accurate the model is, for both locations, percentage prediction error is calculated.

The correlation coefficients between predicted and actual values for both models are 0.996 and 0.992 respectively, which show relatively high values of correlation for the model predictions.

6.4 CONCLUSION

Two different Artificial Neural Networks (ANN) approaches have been used, in this chapter, to:

1- Identify the most influential parameters of the demolition market,

2- Build a model to predict three months ahead of the market.

The static and dynamic ANN methods are employed to meet these criteria.

To meet the first goal, two different static ANN models were built for the monthly scrapped tonnage and scrap price and the order of the influential parameters for each model identified separately (sections 6-2-1-1 and 6-3-1-1). At the beginning of the study, all the possible variables (inputs) which might influence the demolition market were considered for the modelling (33 inputs). The results obtained identify the most influential inputs to the ANN models, in order of importance, for each particular model. These orders are deduced using sensitivity analyses. Each input is assigned a number (or factor) which reflects its importance relative to the other inputs.

As explained before (see section 2-5), the demolition market is a section of the maritime market which can be affected by the other parameters in other markets sections e.g. newbuildings, second-hand and freight rate market. In such a complex environment, the considerable reduction in the number of inputs to the ANN models causes a consequent reduction of the free parameters which can affect the demolition market. Based on the above static ANN investigations, the dimension of the input space for the monthly scrapped tonnage reduces to 7 inputs. Similarly, for the monthly scrap prices in the Subcontinent and Far-East, the dimensions of the input space reduce to 8 and 9 respectively.

It is also shown that if a dynamic ANN is trained properly for a specific purpose in the demolition market, it is able to predict accurately even in situations with complex data. As explained in section 6-2-2, the most accurate ANN model for prediction of the monthly scrapped volume is as follows:

- one hidden layer
- 9 neurons (or PEs)
- 16000 iterations for the network
- the hyperbolic tangent activation function
- the learning rate of 0.1 for both layers $(\gamma = 0.1)$
- the momentum of 0.01 for the networks $(\alpha = 0.01)$
- 4 months time delay $(p = 4)$

Also, the specification of the most accurate dynamic ANN for price prediction in Subcontinent scrapyards is as follows:

- one hidden layer
- 11 neurons (or PEs)
- 9000 iterations for the network
- the hyperbolic tangent activation function
- the learning rate of 0.1 for both layers $(\gamma = 0.1)$
- the momentum of zero for the networks $(\alpha = 0)$
- 5 months time delay $(p = 5)$

For the Far-East scrapyards the structure of the most accurate ANN is:

- one hidden layer
- 5 neurons (or PEs)

- 12000 iterations for the network
- the hyperbolic tangent activation function
- the learning rate of 0.1 for both layers ($\gamma = 0.1$)
- the momentum of zero for the networks ($\alpha = 0$)
- 5 months time delay ($p = 5$)

The prediction of the monthly scrapped tonnage is more accurate than the previous multivariate predictions. Compared wit the results obtained from the statistical multivariate methods in Chapter 5, for the monthly scrapped tonnage, (Table 5-6), the ANN model shows smaller value of RMSE which represents more accuracy for the prediction, Table 6-8.

Modelling Method	RMSE	Key Variables
MLR	0.83	South Korean steel production Oil world trade Non-OPEC oil production EU steel production
PCR	0.90	Steel price Subcontinent scrap price Crude carrier 300k dwt freight rate Won/USD exchange rate Bunker price
PLS	0.92	Crude carrier 300k dw freight rate Bunker price Won/USD exchange rate Yen/USD exchange rate
ANN	0.12	South Korea steel production USA steel production EU steel production Oil world trade Japan steel production Product tankers building price Tanker fleet utilisation rate

Table 6-8: Performance comparison of the different modelling methods for the monthly scrapped tonnage

The predictions of the ANN dynamic models for the scrap prices in both locations are satisfactory. The prediction of the ANN for the monthly scrap price in Far-East scrapyards is slightly more accurate than the model for the Subcontinents scrap yards in terms of the error and the correlation coefficient. Similarly, compared with the obtained results for the monthly scrap prices using multivariate analysis in Chapter 5, (Table 5-7), the ANN models show smaller error for the prediction in terms of RMSE, Table 6-9.

	Modelling Method	RMSE	Key Variables
Subcontinent			
	MLR	20.50	South Korea steel Production Japan steel production Aframax DH building price China steel production USA steel production
	PCR	20.53	Euro/USD exchange rate Yen/USD exchange rate Tankers fleet supply Tankers fleet demand
	PLS	19.30	Euro/USD exchange rate Tankers fleet demand Tankers fleet supply Tankers order book
	ANN	15.75	South Korea steel production Japan steel production EU steel production China steel production OPEC oil production Non-OPEC oil production Tanker fleet utilisation rate Product tankers 5 years market value
Far-East			
	MLR	22.75	South Korea steel production Japan steel production Oil world trade Suezmax DH building price
	PCR	19.72	Euro/USD exchange rate Yen/USD exchange rate Tanker fleet supply Tankers fleet demand China steel production Tankers order book
	PLS	21.26	Euro/USD exchange rate Tankers fleet demand Tankers fleet supply Tankers order book
	ANN	15.61	South Korea steel production Japan steel production EU steel production China steel production OPEC oil production Non-OPEC oil production Tanker fleet utilisation rate Product tankers 5 years market value

Table 6-9: Performance comparison of the different modelling methods for the monthly scrap prices

CHAPTER 7

CONCLUSION AND RECOMMENDATIONS FOR FUTURE WORKS

7.1 CONCLUSION

Forecasting the demolition market is essential for investors and environmental policy makers. Investors and financial decision makers want to be able to predict the market to aid financial planning. It is also important to predict the demolition market so that environmental policy makers can employ adaptive management to changes in the market and act upon environmental concerns and thereby minimise the environmental impacts of ship demolition. However, the identification of the main inputs to the demolition market that can alter market trends is complicated because of the range of variables, which have an influence on the demolition market. Also the complex situation of the whole shipping market including the Newbuilding, Freight, Second-hand and Demolition markets adds to difficulties in predicting the market. In addition, external elements such as inflation, political issues and economic policies can also affect the market. This complex situation can be a reason for the uncertainty of the conventional forecasting analysis methods to model the Demolition market as they can not take into account of all the variables.

This research is expected to be beneficial for investors in ship demolition market. An investor needs to make a decision about buying or selling a ship in a short time and the ANN model, as shown in this research, can help him/her to have a realistic plan about the future which can be key to survival in the business. The produced ANN model is also beneficial for financial decision makers who are focusing on the shipping markets (specially the demolition market) and its related variables to minimise the capital investments. It is also beneficial for the environmental policy makers like people who are writing new legislations and regulations surrounding the environmental aspects of shipping market, specially the demolition market. They can have a realistic prospect of the demolition market using the ANN model. They can create an ANN based on their needs and the time lag they are looking to the market e.g. six months or annual.

This research aimed to model the demolition market and forecast the short-term trend of prices and the volume of the scrapped ships for scrapyards in the Far-East and Subcontinent. The main objectives of this thesis are:

- Implement and validate the conventional multivariate analysis method to produce a model for the demolition market and forecasting the demolition market using the obtained models.
- Implement various ANN techniques to model the demolition market and consequently forecast the demolition market using the obtained models.
- Measure the performance of the obtained model for both the conventional and ANN modelling methods in order to identify the most accurate model.
- Analyse the structure of the demolition market using the best obtained model and identify the inputs that can alter the market trends.
- Forecast three steps ahead of the market using obtained ANN models for both volume and prices of the scrapped ships for scrapyards in the Far-East and Subcontinent separately.

All the above objectives have been met within the chapter 1 to 6 of this thesis and the following main conclusions are acquired:

- the accuracy of the obtained ANN models for both the volume and price are greater than the conventional multivariate statistical forecasting methods but ANNs can only accurately predict the demolition market if they train correctly and also given the proper data upon which to train.
- The large number of internal and external variables (inputs) to the demolition market is a reason for the uncertainty of the conventional statistical methods. They are not able to perform accurately and model the market accurately in comparison to the new ANN modelling approach.
- When compared to multivariate analysis methods, ANNs are a better tool for the following main reasons:

  - When using multivariate methods, the governing regression assumptions must be true and, because of the large amount of variables in the demolition market, it is difficult to have a correct assumption which covers all the variables and conditions. The linearity assumption itself may not hold in many cases. ANNs can model both linear and non-linear systems and does not need any initial assumptions.

- When using multivariate analysis methods, the researcher must have a deep understanding of statistics to ensure only the necessary independent variables are used but it is not possible to distinguish the dependant and independent variables in such a complex market. There are also other considerations which can affect the multivariate model and some of them are not easy to discover especially when the number of variables is large.
- For the models studied in this research, ANNs are significantly more accurate than multivariate analysis methods because they can capture the dynamics of the demolition market, which is non-linear in reality, trough time and identify the input-output maps in order to find the main inputs to the market.

- Considering the complex situation of the shipping market, implementing ANNs are high maintenance and needs an in depth understanding of various network parameters. Small changes of a parameter within a network can significantly change the model and consequently the results.

There are also some conclusions regarding the design of the static and dynamic ANNs for the demolition market. The specifications of different ANNs in each stage have been identified in detail throughout the 6th chapter but it is important to note that:

- Normalisation of the data, as pre-processing, is necessary to equalise the effects of various input data on the model and consequently achieve the adequate model.
- To identify the input-output maps, hyperbolic tangent (tanh), which gives an output in the range between [-1, +1], is performing better than sigmoid activation which gives an output in the range between [0, +1]. With the

sigmoid activation function ANN do not perform well when it is working close to the lower limit, zero, and it seems that the network is down for half of the training period.

- There is no need to increase the complexity of the networks in both static and dynamic modes by adding hidden layers. ANNs with one hidden layers can analyse the data in the market.
- In static ANNs, different learning rates for different layers increases the accuracy of the networks but it is not the case for the dynamic ANNs and equal learning rates can be considered for all the layers.
- In dynamic ANNs the amount of the added momentum to the networks is very low (near zero). ANNs with greater momentums can not converge properly and therefore the network will not be stable and the level of the error rises significantly.
- The batch learning method performs more accurately than the on-line method.
- The amount of the data in each data set is important when dividing the data into separate train, cross validation and test sets. It depends on the data variations for example; there should be some data in the train set to represent peaks and troughs to let ANNs learn the different situations on the market.
- A dynamic ANN should have access the past data to capture the dynamics of the market but there is a limitation in using past data. Having more patterns than needed inhibits the decision making process of the network. The network requires no more than 4 months data to predict the volume of the scrapped ship and 5 months data to predict the scrap prices.
- MSE and MAE are the best criteria for measuring the performance of the ANNs in both static and dynamic modes. Some other criteria are confusing and misleading e.g. percentage error.

7.2 RECOMMENDATIONS FOR FUTURE WORKS

The ANN forecasting methods demonstrated in this research might be extended in the future to predict the shipping markets and provide a means of adapting to unforeseen events and, thereby, provide stability. The development of these techniques for every section of the shipping market, especially the demolition market because of the importance of this particular market, is recommended by the author. In addition, recent progress of the computers processors is also help the abilities of the ANNs to achieve better performances and results.

There are some considerations for the continuation of this research to increase the accuracy of the obtained models and develop the employed techniques:

- The quality of the obtained ANN model is highly related to the quality of the data so the quality of the model will significantly change to a more precise model if more accurate or more specific data is available.

- Based on the sensitivity analysis of the demolition market, it is possible to implement another ANN and ignore the inputs with low sensitivities. This makes the ANN perform faster and sometimes more accurately.

- It is possible to estimate the trend of the most influential inputs to the ANN model and extend the steps of the forecasting. The accuracy of the new model depends on the number of high sensitivity inputs.

- Regarding the recent progresses in processor speed and calculation power and also the invention of new CPU generations of Duo and Quadra, it is possible to use the ANN models real time. This gives enough power to calculate the more complicated ANN algorithms in a shorter time. To implement such a real time system, it is critical to have access to the accurate data so a reliable source of data is needed.

- ANNs have been more accurate than the conventional multivariate methods in the demolition market but ANNs techniques have their own limitations and it is also important to point out that they easily can produce wrong models and subsequently wrong results. Therefore, not only the ANN model should be used and it is recommended that the user verify the network outputs with other present indicators in the market before using the network results produced by each particular model.

A number of practical considerations which may benefit others researching ANNs in the future are listed below:

- ANN modelling is a black-box modelling method so the model itself, as a mathematical relation between the inputs and outputs, is not accessible. Weight matrices are the only windows to the model and sometimes because they are significantly high dimensional matrices it is difficult to use the normal matrix mathematics to calculate the numbers that are needed.

- Maintenance of the obtained ANN models in each stage is difficult. At the beginning of the learning process all the weights are designated randomly to the inputs and as long as the ANN learns trough time they will be changed. Hence, the new settings of weights should be saved in order to continue the learning process for the next stage if a break is taken during the ANN learning process. If the weights are not saved the learning begins again from another random allocation of weights and the same point reached previously may not be achieved.

- If there has been a problem with the data it will affect the ANN and subsequently the quality of the model. Therefore, preparation of the data is needed, however choosing the best method to prepare the data is crucial and plays an important role.

This thesis can be used to aid other researchers looking to predict future shipping markets. They can use the methodology of this research to design and adjust the ANN for their interests. For example, implementing an ANN for predicting bulk carriers' freight rates is different but it is possible to use the same methodology, as it is shown in this research, to identify the unique structure of the ANN.

REFERENCES:

Abdi, H. (2003), Multivariate Analysis, *Encyclopaedia of Social Sciences Research Methods*, eds. Lewis-Beck, M., Bryman, A.and Futing, T., Thousand Oaks, CA.

Abdi, H. Valentin, D. and Edelman, B. (1999), *Neural networks,* Thousand Oaks, CA.

Abhyankar, A.,Copeland, L. S. and Wong, W. (1997), Uncovering nonlinear structure in real-time stock-market indexes: The S&P 500, the DAX, the Nikkei 225, and the FTSE-100. *Journal of Business & Economic Statistics*, 15, 1–14.

Agency for Toxic Substances and Disease Registry (2001), *Toxicological Profile for Asbestos*, US Department of Health and Human Services, Atlanta, GA.

Agency for Toxic Substances and Disease Registry (1999), *Toxicological Profile for Mercury*, US Department of Health and Human Services, Atlanta, GA.

Akaho, S. (2001), A kernel method for canonical correlation analysis, International Meeting of Psychometric Society - IMPS2001, Osaka.

Andersen, B. A. (2001), Worker safety in the ship breaking industries, Sectoral Activities Programme, WP.167, International Labour Office, February 2001, Geneva.

Arbib, M. A. (2003), *The handbook of the brain theory and Neural Networks*, 2nd Edition, The MIT Press, Cambridge, Massachusetts, London England.

Ashley, R. (2003), Statistically significant forecasting improvements: How much out-of-sample data is likely necessary? *International Journal of Forecasting*, 19, 229–239.

Bailer-Jones, D. M. (2001), Models, Metaphors and Analogies, *Blackwell Guide to Philosophy of Science,* eds. Machamer, P. and Silberstein M., Blackwell, Oxford.

Bailey, P. J. (2000), Is there a decent way to break up ships?. International Labour Organisation, Geneva.

Barron, A. R., Yang, Y. and Yu, B. (1994), Asymptotically optimal function estimation by minimum complexity criteria, Proc. International Symposium on Information Theory, Norway, pp.38.

Basel - United Nations Environment Programme (2004), Basel Basics (an overview), http://www.basel.int/pub/, 11 August 2004.

Berry, M. J. A., and Linoff, G. (1997), *Data Mining Techniques for Marketing, Sales, and Customer Support*, New York: John Wiley & Sons.

Bigus, J. P. (1996), *Data Mining with Neural networks:* Solving Business Problems from application development to decision support, McGraw Hill, 1996.

Bilodeau, M. and Brenner, D. (1999), *Theory of Multivariate Statistics*, Springer-Verlag, NewYork.

BIMCO - The Baltic and Maritime International Council (2004), The business of ship recycling, 2nd February 2004, http://www.bimco.dk/, 7 August 2004.

Boulesteix, A. L. and Strimmer, K. (2006), Partial least squares: a versatile tool for the analysis of high-dimensional genomic data, Briefings in Bioinformatics. Vol 8. No1. 32- 44.

Burr, D.J. (1993), Artificial neural networks: a decade of progress, *Artificial Neural Networks for Speech and Vision*, ed. Mammone, R. J., Chapman & Hall.

Buxton, I. L. (1991), The market for ship demolition, *Maritime Policy and Management*, Taylor & Francis Ltd., Vol. 18, No. 2, Apr-Jun 199, p 105-112. ISSN 0308-8839.

Clarkson Research (2005), Shipping Intelligence Network (SIN), www.clarksons.net/SIN, December 2005.

Castillo, E., Guijarro, B., Romero, O. and Alonso, A. (2006), A Very Fast Learning Method for Neural Networks Based on Sensitivity Analysis, *Journal of Machine Learning Research* 7 (2006) 1159–1182

Commission of the European Communities (2006), GREEN PAPER on better ship dismantling, Brussels, SEC (2006)

Cybenko, G. (1989), Approximation by superposition of sigmoidal function, *Mathematics of Control, Signals and Systems*, 2, 303-314.

Dai, H., & McBeth, C. (1997), Effects of learning parameters on learning procedure and performance of a BPNN. *Neural Networks*, 10(8), 1505-1521.

Dankel, D. and Gonzalez, A. (1993), *The Engineering of Knowledge-Based Systems*. Englewood Cliffs, New Jersey: Prentice Hall, 1993.

Demuth, H. and Beale, M. (2004), Neural Network Toolbox: for use with Matlab, Version 4, June 2004; The Math Works Inc.

DNV- Det Norske Veritas (2001), Technological and Economic Feasibility Study of Ship Scrapping in Europe, Report NO. 2000-3527, Revision NO. 01. 13th February. Oslo.

DNV - Det Norske Veritas, (1999), Decommissioning of ships, Environmental Protection and Demolition Practices. Report NO. 99-30 65, Revision NO. 03. 3rd February. Oslo.

Dorffner, G. (1996), Neural Networks for Time Series Processing, *Neural Network World*, April, pp 447-468.

Drewry Shipping Consultants (1998), Ship Scrapping – Locations, Activity, Price Trends and Problems. Briefing Report, October 1998.

Elman, J. L. (1990), Finding Structure in Time, *Cognitive Science,* 14, pp 179-211.

European Commission (2004), Oil Tanker Phase Out and the Ship Scrapping Industry. Directorate-General Energy and Transport Report P-59106-0, 31st May.

European Commission (2003), *The Impact of the new chemicals policy on health and the environment*, Environment Directorate-General final report June 2003

European Union (2003), Regulation (EC) No 1726/2003 of the European Parliament and of the Council of 22 July 2003. Official Journal of the European Union, L249/1, 1st of October 2003.

Everitt, B.S. and Dunn, G. (2001), *Applied multivariate data analysis* , 2nd Edition. Arnold/Oxford University Press, London.

Farber, J. D. and Sidorowich, J. J. (1988), Can new approaches to nonlinear modeling improve economic forecasts? In P. W. Anderson, K. J. Arrow, and D. Pines (Eds.) *The Economy as an Evolving Complex System* (pp. 99–115). Redwood City, CA: Addison-Wesley.

Farmer, D. and Sidorowich, J. (1988), Exploiting chaos to predict the future and reduce noise, Evolution, Learning, and Cognition, *World Scientific*, ed. Lee, W. C., Singapore, pp 277–330.

Fausett, L. (1994), *Fundamentals of Neural Networks: architectures, algorithms and applications*, Prentice Hall, 1994.

Fearnleys (2004), Fearnleys Review: Annual Report 2004, Fearnleys, February 2005.

Frank, R.J., Davey, N. and Hunt, S.P. (1997), *Time Series Prediction and Neural Networks,* Department of Computer Science, University of Hertfordshire, Hatfield, UK.

Fukushima, K. (1988), Neocognitron: A hierarchical neural network capable of visual pattern recognition. *Neural Networks*, pp. 57.

Gershenfeld N. A. and Weigend A.S. (1993), The future of time series: learning and understanding. In *Time Series Prediction: Forecasting the Future and Understanding the Past*, Addison-Wesley, pp. 1–70.

Gordon, K. S. (2002), *Nonlinear regression* Volume 3, John Wiley & Sons Ltd, Chichester. pp 1405–1411.

Granger, C. W. J. (1993). Strategies for modelling nonlinear time-series relationships. *The Economic Record*, 69(206), 233–238

Green peace, (2004), 'START' SHIP RECYCLING IN BRITAIN, Greenpeace media briefing, March 2004.

Grossberg, S. (1974), Classical and instrumental learning by neural networks, *Journal of Theoretical Biology*, 3, pp 51–141.

Guido. J. D. (1994), *Trading on the Edge - Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets*, Wiley, 1994.

Hall, G. (2000), The challenges of ship disposal in the 21st century, 37th annual technical symposium, Association of Scientist and Engineer, 10th May 2000.

Hardle, W. and Simar, L. (2003), *Applied Multivariate Statistical Analysis*, MD*Tech - Method and Data Technologies, Berlin and Louvain-la-Neuve.

Haykin, S. (1999), *Neural Networks: A Comprehensive Foundation, Feed-forward Neural Network*, 2nd Ed., Prentice Hall, Englewood Cliffs, NJ.

Haykin, S. (1994), *Neural Networks: A Comprehensive Foundation*, Macmillan College Publishing. , New York.

Heaton, J. (2003), Programming Neural Networks in Java, Understanding the Kohonen Neural Network, 10/19/2003. From the World Wide Web: http://www.jeffheaton.com/

Hebb, D. O. (1949), *The Organization of Behaviour*, Wiley.

Hoptroff, R. G. (1993), The principles and practice of time series forecasting and business modeling using neural networks. *Neural Computing and Applications*, 1, 59–66.

Hornik, K., Stinchcombe, M. and White, H. (1989), Multi-layer feed-forward networks are universal approximators. *Neural Networks*, 2, 359–366.

Hosmer, D.W. and Lemeshow, S. (1989), *Applied Logistic Regression*, John Wiley & Sons, New York.

Hotelling, H. (1935), The most predictable criterion, *Journal of Educational Psychology* 26, pp. 139-142.

Hsieh, W. W. (2004), Nonlinear multivariate and time series analysis by neural network methods, *Reviews of Geophysics*, Vol. 42, March.

Hu Y. H. and Hwang J. N. (2003), *Handbook of Neural Network Signal Processing*, CRC Press, London, New York.

Hush, D. R. and Horne, B. (1993), Progress in Supervised Neural Networks, *IEEE Signal Process*, pp 1–38.

IMO - International Maritime Organisation (2004), 'International Convention for the Prevention of Pollution from Ships, 1973, as modified by the Protocol of 1978 relating thereto (MARPOL 73/78)', 8 May 2004, http://www.imo.org/Conventions/

IMO - International Maritime Organisation (2004a), IMO GUIDELINES ON SHIP

ISL - Institute of Shipping Economics and Logistics (2006), World merchant fleet development, ISL Market Analysis, Issue 1/2 2006.

ISL - Institute of Shipping Economics and Logistics (2006a), Tanker fleet development, SSMR Market Analysis, March, 2006.

ISL - Institute of Shipping Economics and Logistics (2004), SSMR Market Analysis, Issue 1/2 2004.

Jankowski, N. (1999), Approximation and Classification in Medicine with IncNet Neural Networks, *Machine Learning and Applications*. Workshop on Machine Learning in Medical Applications, July, Greece, pp 53-58.

Jordan, M. (1986), Attractor dynamics and parallelism in a connectionist sequential machine, the Eight Annual Conference of the Cognitive Science Society, Erlbaum, Hillsdale, N.J.

Kasabov, N. K. (1998), *Foundations of neural networks, fuzzy systems, and knowledge*, A Bradford Book, The MIT Press, Cambridge, Massachusetts London, England.

Katijani, Y., Hipel, K. W. and McLeod, A. I. (2005), Forecasting nonlinear time series with feed-forward neural networks: A case study of Canadian lynx data. Journal of Forecasting. 24, 105-117.

Kecman, V. (2001), *Learning and soft computing: support vector machines, neural networks and fuzzy logic models*, The MIT press, London.

Kohonen, T., Mäkisara, K. and Saramäki, T. (1984), Phonotopic maps - insightful representation of phonological features for speech recognition, International

Conference on Pattern Recognition, *IEEE Computer Soc. Press*, Los Alamitos, CA, pp 182-185.

Kohonen, T. (1997), Self-Organizing Maps, *Second Extended Edition, Springer Series in Information Sciences*, Vol. 30, Springer, New York.

Kohonen, T. *(1995), The Self-Organizing Maps*, Springer Verlag, Berlin, Germany.

Kohonen, T. (1990), The self-organizing Maps, *Proceedings of the IEEE*, vol. 78, no. 9, September, pp 1464-1480.

Kohonen, T. *(1989), Self-Organization and Associative Memory*, Springer 1989.

Kohonen, T. (1982), Self-organized formation of topologically correct feature maps, *Biology Cybernetics*, pp 43 and pp 59–69.

Kohonen, T. (1977), *Associative Memory: A system theoretic Approach*, Springer-Veriag, Berlin.

Kramer, M. A. (1991), Nonlinear principal component analysis using autoassociative neural networks, *AIChE Journal*, 37 , 233-243.

Kros, B. and Van Der Smagt, P. (1996), *An introduction to Neural Network*: 8th Edition, November, The university of Amsterdam.

Kruschke, J. K. and Movellan, J. R. (1991), Benefits of gain: speeded learning and minimal hidden layers in back-propagation networks. *IEEE Trans. Systems, Man and Cybernetics* 21(1), pp 273–280.

Kulkarni, A. S. (1998), *Application of neural networks to stock market prediction*, Department of Electrical Engineering, University of Washington, Seattle.

Kung S. Y. and Hwang J. N (1998), Neural networks for intelligent multimedia processing. *Proc. IEEE* 86, pp. 1244–1272

Lachenbruch, P. A. (1975), *Discriminant Analysis: LDA Method*, Hafner, New York.

Lapedes, A., & Farber, R. (1987), *Non-linear signal processing using neural networks: Prediction and system modelling*. Tech. Rep. LA-UR-87-2662, Los Alamos National Laboratory.

Lawrence, S., Giles, C. L. and Fong, S. (1995), On the applicability of neural networks and machine learning methodologies to natural language processing, *Technical Report UMIACS-TR-95-64 and CS-TR-3479*, University of Maryland.

Lloyd's Register of Shipping (2001), World Shipbuilding Statistics, June 2001, London.

Lwin, T. and Maritz, J. S. (1989), *Empirical Bayes methods*, Chapman and Hall, London.

Madan, M., Gupta, Liang Jin, and Noriyasu H. (2003), *Static and Dynamic Neural Network, From Fundamentals to Advanced Theory.* John Wiley & Sons, Inc., Hoboken, New Jersey.

Maddala, G. S. (1997), *Econometrics*, McGraw Hill.

Magnani, L. and Nersessian, N. (2002), Modelling data: Analogies in neural networks, simulated annealing and genetic algorithms, *Model-Based Reasoning: Science, Technology, Values,* Kluwer Academic, Plenum Publishers, New York.

Mandic D. P. and Chambers J. A. (2001), *Recurrent Neural Networks for Prediction*, John Wiley & Sons LTD, New York.
McCullagh, P. and Nelder, J.A. (1989), *Generalized Linear Models*, 2nd Edition, Chapman & Hall, London.

McCulloch, W.S. and Pitts, W. (1943), A logical calculus of the ideas immanent in nervous activity. *Bull. Math. Biophys*, 5, pp 115–133.

McLachlan, G. J. (1992), *Discriminant Analysis and Statistical Pattern Recognition*, John Wiley & Sons, New York.

McNelis, P. D. (2005), *Neural Networks in Finance: Gaining Predictive Edge in the Market*, Elsevier Academic Press.

Medsker, L.R. and Jain, L.C. (2001), *Recurrent Neural Network Design and Applications*, CRC Press, London, New York, Washington, D.C.

Minsky, M. and Papert, S. A. (1969), *Perceptrons: An Introduction to Computational Geometry*. MIT Press.

Mozer, M. C. (1993), Neural net architectures for temporal sequence processing. Time Series Prediction: Forecasting the Future and Understanding the Past. *Addison-Wesley publishing*, eds. Weigend, A. S. and Gershenfeld, N.A., CA, pp. 243-256.

Myers, R. H. (1986), *Classical and Modern Regression with Applications*, Duxbury Press, Boston.

Nygren, K. (2004), *Stock Prediction – A Neural Network Approach*. Master thesis, Royal Institute of Technology, KTH, Sweden.

Ocho-Rivera, J., Garcia, R. and Andreu, J. (2003), An ANN modelling approach applied to the probabilistic management of water resources systems, Department of Hydraulic and Environmental Engineering, Polytechnic University of Valencia, Spain.

Quing, W., Xueqin, S., Quingxin Y. and Weilli,Y. (1997), Using Wavelet Neural Networks for the Optimal Design of Electro Magnetic Devices, *IEEE Trans. on Magnetics*, 33(2), March, pp 1928-19930.

Pao, Y. (1989), *Adaptive Pattern Recognition and Neural Networks,* Addison-Wesley, MA.

Personnaz, L. and Dreyfus, G. (1998), Comment on 'Discrete-time recurrent neural network architectures: a unifying review'. *Neurocomputing* 20, 325–331.

Platou, R.S. (2006), The Platou Report 2006, Platou Economic Research, February 2006, Oslo, Norway.

Platou, R.S. (2004), The Platou Report 2004, Platou Economic Research 23rd February 2004, Oslo, Norway.

Principe, J. S., Euliano, N. E. and Lefebvre, W. C. (2000), *Neural and Adaptive Systems*, John Wiley and Sons Inc., New York.

Rabunal, J. R. and Dorado, J. (2006), *Artificial Neural Network in Real-Life Application*, IDEA group publishing, London.

Ratkowsky, D.A. (1983), *Nonlinear Regression Modelling: A Unified Practical Approach*, Marcel Dekker, New York.

Rencher, A. C. (2002), *Methods of Multivariate Analysis*, 2nd Edition, John Wiley & Sons Inc, US.

Ritter, H. and Schulten, K. (1986), Kohonen's Self-Organizing Maps: Exploring their Computational Capabilities, *Biology Cybernetics* Vol 54, pp 99-106.

Rosenblatt, F. (1958), The perceptron: a probabilistic model for information storage and organization in the brain, *Psychol. Rev.* 65, pp 386–408.

Rumelhart, D. E., Hinton, G. E. and Williams, R. (1986), Learning internal representation by error propagation, *Nature 323*, pp 533–536.

Scholz, M. and Vigario, R. (2002), Nonlinear PCA: a new hierarchical approach, European Symposium on Artificial Neural Networks (ESANN), April, Belgium.

Specht, D.F. (1991), A General Regression Neural Network, *IEEE Trans. On Neural Networks*, Vol. 2, 568–576, 1991.

Stopford, M. (2004a), Oil Market Outlook, Intertanko event, 29th March 2004, Dubai.

Stopford, M. (2004b), World Shipbuilding 2004, Advance Press conference SMM 2004, 25th May, Hamburg.

Stopford M. (2003), Maritime Economics. 2nd Edition, Routledge, London and New York.

Stopford, M. (2003a), Maritime Market Overview. Maritime Cyprus Conference, Cyprus, 24th September 2003.

Stopford, M. (2003b), Prospects for the Tanker Market in 2003 and Beyond, Intertanko Maritime Conference, 17th September 2003, London.

Smyth, G. K. (2002), *Nonlinear regression* Volume 3, John Wiley & Sons, Ltd., Chichester, pp 1405–1411.

Swingler, K. M. (1994*),* Sequence categorisation using multiple recurrent layers to create many trajectories through network state space. Proc. ICANN, vol. 2, pp 1025-1028.

Tsoi, CA and Back, A. (1997), Discrete time neural network architectures: a unifying review. *Neurocomputing* 15(3), 183–223.

Tsoi, A.C. and Back, A. D. (1994), Locally recurrent globally feedforward networks: a critical review of architectures, *IEEE Transactions on Neural Networks*, 5, 229.

UNCTAD - United Nations Conference on Trade and Development (2005), Review of marine transport 2005, United Nations, New York and Geneva.

UNEP - United Nations Environment Programme (2001), Ship dismantling industry set to go green, News Centre, 19th June 2001, Geneva, http://www.unep.org/Documents.Multilingual/, 8 August 2004,

Van Eyden, R.J. (1996), *The Application of Neural Networks in the Forecasting of Share Prices*, Finance and Technology Publishing.

Vanstone, B., Finnie, F. and Clarence, T. (2004), Applying fundamental analysis and neural networks in Australian stockmarket, Bond University, Australia.

Vesanto, J. and Alhoniemi, E. (2000), Clustering of the Self-Organizing Map, *IEEE Transactions on Neural Networks*, 11(2), March, pp 586–600.

Wang, S. and Archer, N. P. (1991), Fuzzy set representation of neural network classification boundaries, *IEEE Trans. on SMC*, vol. 21, pp 735-742.

Warren, S. and Cary, S. C. (1994), Neural Networks and statistical models, 19th annual SAS users group international conference, April, USA.

Watterson, K. (1995), A Data Miner's Tools, *Byte*, October, pp. 91-96.

Wendorf, C. A. (2004), *Manuals for Univariate and Multivariate statistics*, Part IX: Multivariate Analysis of Variance, University of Wisconsin.

Werbos, P. (1990), Backpropagation through time: what it does and how to do it. *Proc. IEEE* 78, 1550–1560.

Werbos, P. (1974), Beyond Regression: New Tools for Prediction and Analysis in the Behavioural Science. PhD thesis, Harvard University, Cambridge, MA.

White, H. (1992), *Artificial Neural Networks: Approximation and Learning Theory*, Blackwell, Oxford, UK.

Widrow, B. and Hoff, M. (1960), Adaptive switching circuits, Proc. WESCON Convention, vol. IV, pp 96–104.

Widrow, B. And Hoff, M. (1988), Adaptive switching circuits, *Neurocomputing*, eds. J. Anderson and E. Rosenfeld, MIT Press, Cambridge, MA..

Wold, S., Ruhe, A., Wold, H., and Dunn, W. J. (1984), Collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses. SIAM J Sci. Comput. Stat.; 5:735–43.

Zhang, G. P. (2004), *Neural Networks in Business Forecasting*, Idea Group Publishing. London.

Zilouchian, A. and Jamshidi, M. (2001), *Intelligent control systems using soft computing methodologies*, CRC Press, London, New York, Washington, D. C.

APPENDICES

APPENDIX A: RAW DATA

Month	USA Crude Steel Production (million tonne)	EU Crude Steel Production (million tonne)	China Crude Steel Production (million tonne)	Japan Crude Steel Production (million tonne)	South Korea Crude Steel Production (million tonne)	Average Steel Price (USD/tonne-ldt)	Scrap Price – Sub Continent (USD/tonne-ldt)	Scrap Price – Far East (USD/tonne-ldt)	OPEC Oil Production (MBD)	Non OPEC Oil Production (MBD)	Oil World Trade (MBD)	Exchange Rate (Index 1995/1=100) EUR/USD	Exchange Rate (Index 1995/1=100) EUR/USD
January 95	7.7	14.0	7.4	8.6	3.0	258.7	175.0	151.6	24.8	42.7	31.7	100.0	100.0
February 95	7.7	14.2	7.4	8.6	3.0	258.7	179.7	151.6	24.8	42.6	31.8	102.5	100.0
March 95	7.9	14.3	7.6	8.7	3.0	258.7	187.5	151.6	24.9	42.4	32.0	107.4	101.8
April 95	7.9	14.3	7.6	8.6	3.1	269.6	187.5	156.3	24.9	42.2	32.1	107.1	103.4
May 95	7.9	14.4	7.6	8.7	3.1	269.6	187.5	153.1	25.0	42.3	32.5	105.2	104.0
June 95	7.9	14.6	7.6	8.7	3.1	295.7	187.5	153.1	25.0	42.4	32.8	107.4	104.0
July 95	7.8	14.4	7.7	8.7	3.1	295.7	187.5	153.1	25.1	42.6	33.4	108.0	104.0
August 95	7.8	14.4	7.7	8.7	3.1	284.8	187.5	153.1	25.2	42.7	33.2	103.1	102.2
September 95	7.8	14.3	7.7	8.6	3.1	280.4	190.6	151.6	25.2	42.7	32.9	104.6	103.1
October 95	7.8	14.3	7.7	8.5	3.1	269.6	193.8	146.9	25.2	42.8	32.6	104.9	103.4
November 95	7.8	13.9	7.7	8.4	3.1	265.2	184.4	143.8	25.4	42.9	33.0	102.8	102.5
December 95	7.8	13.4	7.8	8.4	3.1	254.3	179.7	143.8	25.7	43.2	33.5	103.1	102.2
January 96	7.9	13.3	7.9	8.3	3.2	254.3	189.1	143.8	26.0	43.4	34.0	99.4	100.3
February 96	7.9	13.3	7.9	8.1	3.2	250.0	190.6	143.8	25.9	43.4	34.0	101.2	100.6
March 96	8.0	13.3	8.0	8.1	3.2	250.0	190.6	143.8	25.7	43.4	34.0	101.5	100.6
April 96	7.9	13.3	8.0	8.1	3.3	252.2	209.4	143.8	25.5	43.4	34.0	99.1	101.5
May 96	8.0	13.3	8.0	8.1	3.4	258.7	193.8	143.8	25.7	43.4	34.0	100.0	100.0
June 96	8.0	13.6	8.1	8.1	3.4	265.2	193.8	143.8	25.8	43.4	33.9	100.6	97.5
July 96	7.9	13.5	8.3	8.2	3.4	273.9	193.8	143.8	25.9	43.4	33.9	102.8	97.2
August 96	7.9	13.5	8.4	8.3	3.3	273.9	193.8	143.8	25.9	43.6	34.1	102.8	96.3
September 96	7.8	13.3	8.4	8.3	3.3	280.4	159.4	143.8	25.9	43.9	34.6	100.3	95.7
October 96	7.9	13.3	8.5	8.3	3.3	287.0	171.9	132.8	26.0	44.2	35.2	102.2	95.7
November 96	7.8	13.4	8.5	8.4	3.3	295.7	148.4	132.8	26.2	44.2	35.1	101.2	95.4
December 96	7.8	13.3	8.6	8.4	3.3	295.7	157.8	132.8	26.4	44.2	35.1	101.2	93.8
January 97	7.8	13.6	8.6	8.5	3.3	308.7	162.5	132.8	26.7	44.2	35.0	95.7	91.7
February 97	7.8	13.7	8.5	8.5	3.3	308.7	173.4	126.6	26.5	44.2	35.0	92.6	91.7
March 97	7.9	13.8	8.6	8.6	3.4	315.2	173.4	126.6	25.9	44.2	35.3	93.8	88.6
April 97	7.9	14.1	8.7	8.7	3.5	315.2	181.3	131.3	25.3	44.1	35.4	91.1	88.9
May 97	8.0	14.2	8.7	8.7	3.5	321.7	176.6	131.3	25.5	44.1	35.5	92.3	88.6
June 97	8.0	14.5	8.7	8.8	3.5	343.5	176.6	131.3	25.8	44.2	35.7	90.8	89.5
July 97	8.0	14.6	9.0	8.8	3.5	343.5	184.4	135.9	26.0	44.2	35.8	86.8	88.9

August 97	8.1	14.6	9.1	8.9	3.6	343.5	189.1	131.3	26.3	44.3	35.8	88.0	88.0
September 97	8.0	14.7	9.0	8.8	3.6	343.5	184.4	131.3	26.8	44.7	36.1	89.8	86.8
October 97	8.1	14.9	9.1	8.8	3.7	345.7	190.6	131.3	27.5	45.0	36.4	92.3	83.1
November 97	8.1	14.8	9.2	8.7	3.7	345.7	184.4	131.3	27.8	45.2	36.7	90.8	68.3
December 97	8.3	14.7	9.2	8.7	3.7	343.5	179.7	131.3	28.2	45.2	37.1	89.2	47.1
January 98	8.5	15.0	9.1	8.6	3.9	339.1	171.9	131.3	28.6	45.3	37.4	87.1	47.4
February 98	8.5	15.1	9.0	8.5	3.7	326.1	167.2	131.3	28.5	45.4	37.6	87.7	48.3
March 98	8.6	15.3	9.1	8.4	3.7	313.0	153.1	131.3	28.5	45.2	37.6	86.8	56.9
April 98	8.7	15.4	9.2	8.1	3.6	300.0	153.1	131.3	28.4	45.0	37.7	88.6	58.2
May 98	8.7	15.5	9.3	8.0	3.6	300.0	153.1	131.3	28.3	44.8	37.7	89.2	56.9
June 98	8.7	15.6	9.3	7.9	3.5	300.0	134.4	120.3	28.2	44.6	37.4	88.6	57.2
July 98	8.6	15.4	9.3	7.8	3.5	273.9	126.6	106.3	28.0	44.3	37.1	89.5	61.2
August 98	8.7	15.0	9.5	7.9	3.4	271.7	131.3	110.9	27.9	44.1	36.8	91.7	59.7
September 98	8.3	14.8	9.5	7.8	3.3	247.8	121.9	110.9	27.8	44.5	37.1	95.1	57.2
October 98	8.1	14.6	9.6	7.8	3.3	208.7	115.6	110.9	27.7	44.7	37.3	96.0	59.4
November 98	7.9	14.3	9.8	7.8	3.2	208.7	114.1	114.1	27.6	44.8	37.5	93.8	63.7
December 98	7.7	13.9	9.8	7.7	3.2	202.2	114.1	114.1	27.6	44.7	37.5	94.8	65.8
January 99	7.6	13.7	10.0	7.6	3.2	202.2	109.4	109.4	27.7	44.7	37.5	93.5	67.4
February 99	7.5	13.7	9.8	7.5	3.3	206.5	120.3	114.1	27.8	44.7	37.5	90.2	64.6
March 99	7.5	13.7	10.0	7.5	3.4	206.5	120.3	114.1	27.2	44.5	37.4	87.7	64.6
April 99	7.5	13.6	9.9	7.4	3.4	223.9	123.4	115.6	26.6	44.2	37.3	87.1	66.8
May 99	7.7	13.9	9.9	7.5	3.4	223.9	121.9	115.6	26.0	44.1	37.1	85.2	66.8
June 99	7.8	14.2	9.9	7.5	3.4	223.9	115.6	110.9	26.1	44.2	37.1	84.0	68.6
July 99	7.8	14.2	10.0	7.5	3.4	234.8	115.6	115.6	26.2	44.5	37.2	83.1	65.8
August 99	7.9	14.2	10.2	7.7	3.5	256.5	121.9	114.1	26.1	44.6	37.2	86.2	67.1
September 99	7.9	14.2	10.3	7.9	3.5	256.5	128.1	114.1	26.1	45.0	37.2	84.6	64.9
October 99	8.0	14.4	10.4	8.0	3.6	256.5	128.1	114.1	26.0	45.3	37.1	87.1	65.8
November 99	8.1	14.4	10.5	8.1	3.6	265.2	128.1	126.6	26.0	45.7	37.1	84.3	68.3
December 99	8.2	14.2	10.8	8.4	3.7	278.3	128.1	126.6	26.3	45.8	37.3	82.2	69.5
January 00	8.5	14.4	10.7	8.5	3.7	284.8	126.6	125.0	26.6	45.9	37.5	82.2	70.2
February 00	8.6	14.6	10.6	8.6	3.6	308.7	126.6	125.0	26.8	46.0	37.7	79.4	69.8
March 00	8.7	15.0	10.5	8.7	3.7	308.7	135.9	125.0	27.2	45.9	38.2	77.5	71.1
April 00	8.7	14.8	10.4	8.7	3.7	321.7	146.9	135.9	27.4	45.8	38.6	76.3	70.8
May 00	8.7	15.0	10.3	8.7	3.7	332.6	135.9	135.9	27.7	45.7	39.1	73.5	69.8
June 00	8.8	15.1	10.0	8.7	3.7	321.7	148.4	145.3	27.9	45.8	39.2	76.9	70.5
July 00	8.8	15.0	10.2	8.8	3.7	313.0	162.5	157.8	28.2	45.9	39.3	75.7	70.5
August 00	8.7	15.0	10.4	8.9	3.7	313.0	179.7	171.9	28.4	45.9	39.5	72.3	71.1
September 00	8.6	14.8	10.5	8.9	3.7	291.3	176.6	171.9	28.7	46.0	39.7	71.4	70.8
October 00	8.5	14.9	10.6	9.0	3.7	247.8	176.6	187.5	28.9	46.1	39.8	68.0	69.5
November 00	8.2	14.7	10.8	9.0	3.7	202.2	176.6	187.5	29.1	46.2	40.0	70.2	66.2
December 00	7.9	14.5	10.9	9.0	3.7	202.2	176.6	187.5	28.9	46.3	39.8	76.6	62.5
January 01	7.8	14.6	10.9	9.0	3.6	202.2	195.3	200.0	28.6	46.3	39.7	74.5	62.8
February 01	7.7	14.7	10.9	8.9	3.7	202.2	195.3	201.6	28.4	46.4	39.5	75.1	63.1
March 01	7.7	15.0	11.0	8.9	3.7	206.5	185.9	160.9	27.9	46.3	39.2	71.1	59.4
April 01	7.7	14.9	11.0	8.8	3.7	210.9	185.9	160.9	27.5	46.2	39.0	71.4	60.3
May 01	7.8	15.0	11.1	8.8	3.7	219.6	189.1	167.2	27.0	46.1	38.8	68.6	61.8
June 01	7.9	15.1	11.2	8.7	3.8	208.7	187.5	167.2	27.1	46.2	38.9	68.3	60.6
July 01	7.9	14.8	11.3	8.7	3.9	200.0	162.5	154.7	27.2	46.4	38.9	70.2	61.2
August 01	7.9	14.7	11.5	8.7	3.9	200.0	150.0	153.1	27.3	46.6	39.0	73.5	61.8
September 01	7.8	14.5	11.5	8.7	3.9	206.5	160.9	134.4	27.1	46.8	38.9	73.5	60.3
October 01	7.6	14.5	12.0	8.7	3.9	206.5	137.5	134.4	26.6	47.2	38.7	72.9	61.2
November 01	7.5	14.2	12.2	8.6	3.9	193.5	137.5	129.7	26.2	47.4	38.5	71.7	62.5
December 01	7.2	14.0	12.3	8.5	3.9	193.5	137.5	125.0	25.9	47.5	38.4	72.3	60.0
January 02	7.1	14.0	12.7	8.5	3.9	193.5	135.9	118.8	25.5	47.6	38.2	70.5	60.3
February 02	7.0	14.1	13.0	8.4	3.9	193.5	129.7	123.4	25.3	47.7	38.0	69.8	60.0
March 02	7.0	14.4	13.3	8.4	3.9	208.7	139.1	123.4	25.0	47.8	38.2	70.8	59.4
April 02	7.0	14.3	13.7	8.4	3.9	228.3	150.0	123.4	24.9	47.9	38.3	72.9	61.2
May 02	7.3	14.4	14.0	8.6	3.9	250.0	145.3	145.3	24.6	48.0	38.4	75.7	64.3

June 02	7.5	14.7	14.3	8.7	3.9	267.4	134.4	134.4	24.8	48.0	38.7	79.4	65.2
July 02	7.5	14.9	14.5	8.8	3.9	278.3	154.7	134.4	25.0	47.9	39.1	79.1	66.5
August 02	7.7	14.7	14.8	9.0	4.0	278.3	154.7	143.8	25.3	47.8	39.6	80.0	66.2
September 02	7.8	14.6	15.0	9.2	4.0	278.3	154.7	145.3	25.7	47.9	39.8	80.0	64.0
October 02	7.9	14.7	15.2	9.3	4.0	295.7	159.4	159.4	25.9	48.2	40.0	80.0	64.3
November 02	7.9	14.7	15.4	9.3	4.0	300.0	168.8	159.4	26.2	48.5	40.0	80.0	64.9
December 02	7.9	14.5	15.6	9.3	4.0	300.0	179.7	170.3	26.4	48.5	40.1	83.7	66.2
January 03	7.9	14.6	15.6	9.2	4.0	300.0	185.9	170.3	26.5	48.5	40.4	86.8	68.0
February 03	7.8	14.7	15.7	9.1	4.0	306.5	190.6	170.3	26.6	48.5	40.7	86.8	66.5
March 03	7.8	14.8	16.2	9.1	4.0	310.9	215.6	200.0	26.6	48.4	40.9	87.7	62.8
April 03	7.7	14.9	16.4	9.1	4.0	289.1	214.1	210.9	26.5	48.3	41.1	89.2	64.6
May 03	7.6	15.0	16.8	9.1	4.0	289.1	201.6	200.0	26.5	48.0	41.4	95.4	65.2
June 03	7.8	15.1	17.0	9.1	4.0	267.4	215.6	210.9	26.5	47.9	41.6	92.6	65.8
July 03	7.7	15.0	17.6	9.1	4.0	271.7	240.6	240.6	26.6	48.2	41.7	91.1	66.8
August 03	7.7	14.8	18.0	9.3	4.0	284.8	229.7	225.0	26.8	48.5	41.8	87.4	66.8
September 03	7.5	14.6	18.4	9.3	4.0	289.1	242.2	232.8	27.1	48.7	42.0	93.8	68.6
October 03	7.5	14.6	18.8	9.3	4.0	289.1	250.0	246.9	27.4	49.0	42.3	93.5	66.8
November 03	7.5	14.6	19.1	9.3	4.0	300.0	260.9	270.3	27.7	49.3	42.6	93.5	66.8
December 03	7.5	14.6	19.4	9.3	4.0	300.0	282.8	298.4	28.0	49.6	42.9	100.6	65.5
January 04	7.7	14.7	19.6	9.3	4.0	350.0	325.0	326.6	28.2	49.6	43.5	99.7	66.8
February 04	7.7	15.0	19.7	9.3	4.0	415.2	387.5	385.9	28.3	49.6	43.9	99.4	66.8
March 04	7.8	15.3	20.0	9.3	4.0	495.7	392.2	370.3	28.4	49.6	44.3	99.1	68.9
April 04	7.9	15.3	20.2	9.3	4.0	495.7	400.0	318.8	28.4	49.7	44.3	96.6	67.1
May 04	8.0	15.5	20.4	9.3	4.0	530.4	384.4	257.8	28.5	49.9	44.5	98.5	67.7
June 04	8.1	15.7	20.5	9.3	4.0	569.6	345.3	318.8	28.6	50.0	44.5	97.8	68.0
July 04	8.1	15.9	20.9	9.3	4.0	569.6	387.5	343.8	28.8	49.9	44.7	97.2	67.4
August 04	8.4	15.7	21.5	9.4	4.0	569.6	425.0	371.9	29.1	49.8	45.0	97.2	68.3
September 04	8.4	15.7	21.9	9.4	4.0	589.1	415.6	337.5	29.6	49.7	45.1	98.8	68.3
October 04	8.5	15.7	22.6	9.5	4.0	589.1	404.7	368.8	29.7	49.9	45.3	102.2	70.2
November 04	8.5	15.7	23.4	9.5	4.0	589.1	421.9	375.0	29.9	50.0	45.4	107.4	75.4
December 04	8.5	15.6	24.2	9.5	4.0	589.1	434.4	356.3	30.0	50.2	45.5	108.9	76.0

Month	Exchange Rate (Index 1995/1=100) YEN/USD	Bunkers- 380 CST in Rotterdam (USD/ mill. tonnes)	Tanker, MR Product Building Price (mill. USD)	Tanker, Aframax DH Building Price (mill. USD)	Tanker, Suezmax DH Building Price (mill USD)	Tanker, VLCC DH Building Price (mill. USD)	Crude Carrier (105000dwt) FR Single Voyage (kUSD/day)	Crude Carrier (150000dwt) FR Single Voyage (kUSD/Day)	Crude Carrier (300000dwt) FRSingle Voyage (kUSD/Day)	Clean Carrier (70/85000dwt) FR Single Voyage (kUSD/Day)	Tanker, MR Product DS/DH 5Years Market Value (mill. USD)	Tanker, Aframax DS/DH 5Years Market Value (mill. USD)	Tanker, Suezmax SH/DH 5Years Market Value (mill. USD)
January 95	100.0	104.3	32.2	41.9	53.1	86.1	14.8	14.8	11.1	17.8	21.9	29.8	33.2
February 95	101.8	106.3	32.8	42.2	53.3	86.1	13.0	13.0	8.0	16.6	22.3	29.8	34.0
March 95	108.6	105.3	33.1	42.5	53.6	86.1	12.3	13.6	9.9	16.1	22.6	29.8	34.3
April 95	120.0	107.3	33.3	42.8	53.9	86.1	14.2	12.3	8.0	13.9	23.0	29.8	34.3
May 95	117.2	94.5	33.3	42.8	53.9	86.1	11.7	13.0	6.8	13.4	23.8	29.8	34.3
June 95	118.8	82.7	33.3	42.8	53.9	86.1	13.0	13.6	13.6	17.3	24.2	29.8	34.3
July 95	114.5	85.6	33.3	42.8	53.6	86.1	14.2	19.8	19.1	19.3	24.2	29.8	34.7
August 95	104.9	87.6	33.1	42.8	53.3	86.7	14.2	14.8	18.5	18.8	24.2	30.2	35.5
September 95	99.4	85.6	32.8	42.8	53.1	86.9	14.2	14.8	14.8	18.3	24.2	30.6	36.2
October 95	99.4	87.6	32.8	42.8	52.8	87.5	14.2	14.8	11.1	20.1	24.2	30.6	36.2
November 95	97.8	104.3	32.5	42.5	52.5	86.7	13.6	16.0	16.7	21.8	24.2	30.6	36.2
December 95	96.9	101.4	32.2	42.2	52.5	85.8	14.8	16.0	16.7	22.0	24.2	30.6	36.2
January 96	93.5	97.4	32.2	41.9	52.2	85.0	17.9	17.3	16.0	21.1	24.2	30.9	37.4
February 96	96.0	107.3	32.2	41.7	51.9	84.7	17.3	17.3	19.8	20.1	24.2	30.9	38.5
March 96	93.8	113.2	32.2	41.4	51.9	84.4	16.0	17.3	16.0	18.6	24.2	31.7	39.2
April 96	93.8	99.4	32.2	41.1	51.7	83.9	16.0	17.9	9.3	16.3	24.2	32.1	39.2
May 96	91.7	88.6	31.9	41.1	51.4	83.9	16.0	18.5	16.0	16.3	24.5	32.1	39.2
June 96	91.1	94.5	31.9	41.1	51.1	83.9	16.0	17.9	21.6	17.1	25.3	32.5	39.2
July 96	92.6	99.4	31.9	41.1	50.8	83.9	13.6	17.3	21.6	17.1	24.9	32.5	39.2
August 96	92.3	117.1	32.2	41.1	51.1	83.9	13.6	16.0	19.1	16.6	24.2	32.5	39.2
September 96	89.8	121.1	32.5	41.1	51.4	83.6	13.6	13.6	13.6	16.1	24.2	32.5	39.2
October 96	86.8	115.2	32.8	41.1	51.7	83.3	16.0	16.7	13.6	17.8	24.5	32.1	38.9
November 96	88.0	120.1	32.5	41.1	51.7	83.3	17.3	17.3	13.6	19.1	24.5	32.1	38.5
December 96	86.5	100.4	32.2	40.8	51.7	83.1	18.5	17.9	14.2	21.5	24.9	31.7	38.5
January 97	82.8	95.5	31.9	40.6	51.7	83.1	21.6	22.8	30.9	25.0	26.0	32.5	39.2
February 97	80.9	88.6	31.9	40.6	51.7	83.1	19.8	24.1	29.6	23.8	26.0	33.6	40.0
March 97	81.5	85.6	31.9	40.8	51.7	83.1	29.6	28.4	31.5	23.3	26.0	33.6	40.0
April 97	78.8	85.6	32.2	41.1	51.7	83.1	28.4	25.3	25.9	22.8	26.8	34.7	41.1
May 97	86.2	87.6	32.2	41.1	51.7	83.1	24.7	24.7	32.1	20.6	26.8	34.7	41.1
June 97	88.0	93.5	31.9	41.1	51.7	83.1	22.8	22.8	34.6	18.8	26.8	34.7	41.1
July 97	84.3	98.4	31.9	41.1	51.7	83.1	17.3	20.4	38.9	15.4	26.8	34.7	41.1
August 97	84.3	98.4	31.9	41.1	51.4	83.1	20.4	25.3	46.3	15.4	26.8	35.5	41.1
September 97	83.4	108.3	31.9	41.1	51.1	82.8	19.1	27.8	42.0	16.1	26.8	36.6	42.3
October 97	83.4	95.5	31.9	41.1	51.1	82.8	25.3	24.7	51.2	16.3	26.8	37.4	43.4
November 97	78.2	82.7	31.4	40.6	50.8	81.9	19.8	29.0	45.7	18.6	26.0	37.4	43.4
December 97	77.2	69.9	30.6	40.0	50.0	80.8	21.6	29.0	30.2	21.1	24.9	36.6	42.3
January 98	79.4	66.9	29.7	39.7	49.7	80.3	21.6	28.4	27.8	20.6	24.5	36.6	41.1
February 98	78.8	70.9	29.2	38.9	48.9	78.9	17.9	21.6	35.2	20.8	24.2	35.8	40.8
March 98	77.8	75.8	28.9	37.5	47.8	77.2	22.2	28.4	40.1	18.1	23.0	34.7	40.0
April 98	75.7	67.9	28.6	36.4	46.9	75.8	17.3	21.6	37.0	15.6	21.9	34.7	40.8

May 98	72.9	64.0	28.6	35.6	46.1	74.2	14.2	21.6	42.0	14.9	21.9	33.6	40.0
June 98	70.8	60.0	28.1	35.0	45.6	72.8	15.4	23.5	36.4	16.1	21.1	31.7	40.0
July 98	70.8	60.0	27.5	34.4	44.7	70.8	16.7	24.1	43.2	16.8	21.1	31.7	40.0
August 98	69.5	65.9	26.9	34.2	44.4	69.7	14.8	19.8	35.2	17.3	19.2	30.2	40.0
September 98	73.8	61.0	26.4	34.2	44.2	69.4	13.6	16.0	22.2	16.8	17.7	28.3	37.4
October 98	84.3	54.1	26.1	34.2	44.2	69.2	14.2	18.5	26.5	15.4	17.7	28.3	37.4
November 98	81.8	52.2	25.8	34.2	43.9	68.9	16.0	18.5	24.7	16.6	17.4	27.5	37.4
December 98	88.0	61.0	25.8	34.2	43.3	68.6	17.9	21.6	29.6	17.3	18.5	25.7	37.4
January 99	86.2	56.1	25.8	34.2	43.1	68.6	19.8	19.8	29.0	15.1	20.0	23.0	37.4
February 99	84.0	69.9	25.8	34.2	42.5	68.3	14.8	22.2	30.9	12.1	20.0	23.0	37.4
March 99	84.0	69.9	25.3	33.9	42.2	68.1	16.0	19.8	28.4	10.9	20.0	23.0	37.4
April 99	83.7	78.7	25.0	33.6	41.9	67.8	14.8	16.7	14.8	10.4	20.0	23.0	36.2
May 99	82.5	87.6	25.0	33.3	41.9	67.8	11.7	13.6	17.3	12.9	20.0	23.0	36.2
June 99	82.5	103.3	25.0	33.6	42.2	67.8	9.9	13.0	24.1	14.6	20.0	25.3	36.2
July 99	87.4	113.2	25.0	33.9	42.5	67.8	9.9	14.8	16.0	14.4	20.0	25.3	35.8
August 99	90.5	117.1	25.0	33.9	42.8	68.1	10.5	9.3	13.6	13.6	19.6	24.9	35.5
September 99	93.5	125.0	25.3	34.2	42.8	68.3	9.9	11.7	17.3	13.1	19.6	24.9	35.5
October 99	95.7	128.0	25.6	34.2	42.8	68.6	10.5	13.6	17.3	12.6	19.6	25.3	35.5
November 99	97.8	125.0	25.8	34.2	42.8	68.9	12.3	13.6	13.6	10.2	19.2	24.5	35.5
December 99	97.8	129.9	26.1	34.7	43.6	69.7	17.9	23.5	16.7	12.9	19.2	24.5	35.5
January 00	92.9	140.7	26.7	35.0	44.4	70.3	18.5	21.0	20.4	15.9	19.6	25.7	36.6
February 00	90.2	146.7	26.9	35.6	45.3	71.1	24.1	24.1	23.5	17.3	20.0	27.9	38.1
March 00	96.3	121.1	26.9	36.1	45.8	71.7	28.4	27.2	29.0	17.6	20.8	29.4	39.6
April 00	92.0	140.7	27.2	36.7	46.4	72.2	24.7	29.0	38.9	19.1	22.3	30.6	41.1
May 00	92.6	134.8	27.2	36.9	46.7	72.8	27.8	30.2	45.7	19.1	22.3	31.3	41.1
June 00	93.8	130.9	27.5	37.5	47.5	73.3	38.3	37.0	50.6	23.0	22.3	32.5	43.0
July 00	90.8	148.6	27.5	38.1	48.1	73.6	42.6	49.4	50.6	26.3	22.3	33.6	43.8
August 00	93.5	158.5	27.8	38.6	48.6	73.9	45.1	42.6	68.5	27.5	24.9	36.6	49.1
September 00	92.3	150.6	27.8	39.4	49.2	74.4	33.3	47.5	69.1	28.5	24.9	38.1	49.8
October 00	91.4	131.9	27.8	40.0	49.7	75.0	50.0	55.6	67.9	29.5	24.9	40.0	50.2
November 00	90.5	107.3	27.8	40.6	50.0	75.3	49.4	53.1	85.2	32.7	24.9	40.8	50.2
December 00	87.7	118.1	27.8	40.3	49.7	75.3	55.6	59.9	82.7	50.5	26.0	41.5	50.2
January 01	85.5	123.0	27.8	40.0	49.7	75.0	47.5	59.3	64.2	60.4	26.0	41.1	49.4
February 01	84.9	119.1	27.8	39.7	49.7	74.7	47.5	35.2	50.6	58.5	26.0	40.0	48.3
March 01	78.8	117.1	27.8	39.7	49.7	74.7	51.9	46.9	55.6	38.9	26.0	40.0	48.3
April 01	80.9	123.0	27.8	39.7	49.7	74.7	38.3	38.9	44.4	28.0	26.0	40.0	48.3
May 01	83.7	120.1	27.8	39.7	49.7	74.7	29.6	34.0	31.5	29.2	25.7	41.1	48.7
June 01	80.0	116.1	27.5	39.7	49.7	74.7	23.5	25.9	22.8	31.5	25.3	40.8	48.7
July 01	80.0	128.9	27.5	39.4	49.7	74.7	25.3	30.2	25.3	29.0	24.9	39.6	47.9
August 01	83.7	113.2	27.2	39.2	49.7	74.7	26.5	27.8	24.7	27.0	24.5	37.7	45.7
September 01	83.4	100.4	26.9	38.6	48.6	73.6	22.2	24.7	39.5	23.5	23.8	36.2	44.5
October 01	81.5	102.4	26.7	37.8	47.8	72.2	25.3	27.2	32.1	25.8	22.6	35.1	43.0
November 01	80.9	105.3	26.4	36.7	46.7	70.6	22.2	19.8	14.8	18.1	21.9	34.3	41.5
December 01	76.0	107.3	26.1	36.1	45.8	69.7	23.5	19.1	16.0	14.4	20.8	32.8	40.0
January 02	74.2	117.1	25.8	35.8	45.3	69.4	15.4	17.3	19.8	14.1	20.0	31.3	38.9
February 02	74.5	130.9	25.3	35.6	45.0	69.2	13.6	17.3	18.5	16.6	20.0	30.6	38.9
March 02	74.8	139.8	25.3	35.3	44.7	68.6	13.6	19.1	13.0	17.1	20.0	30.6	38.9
April 02	77.5	134.8	25.3	35.0	44.7	67.5	16.0	19.1	8.6	13.9	20.0	30.6	38.9
May 02	80.3	140.7	25.3	34.7	44.7	66.4	16.0	16.7	19.1	14.1	20.0	30.6	38.9
June 02	83.4	145.7	25.3	34.7	44.7	65.6	17.9	17.9	11.1	15.9	20.0	30.6	39.6
July 02	83.1	159.4	25.3	34.4	44.4	65.3	16.0	16.7	14.2	17.1	20.0	30.6	39.6
August 02	84.3	146.7	25.3	34.2	44.4	65.0	12.3	15.4	9.9	18.1	20.0	30.6	39.6
September 02	82.2	126.0	25.3	34.2	44.4	64.7	12.3	15.4	11.7	19.1	20.0	28.7	37.7
October 02	81.5	134.8	25.6	34.2	44.2	64.4	17.9	23.5	38.3	18.1	21.5	28.7	37.7
November 02	81.5	166.3	25.6	34.2	43.9	64.2	30.9	35.8	44.4	20.3	21.5	28.7	37.7
December 02	84.0	173.2	25.8	34.2	44.2	63.9	29.6	43.2	65.4	29.5	21.5	28.7	38.1
January 03	83.7	146.7	26.4	34.7	44.7	64.4	30.2	54.9	71.6	31.5	23.4	31.7	41.9
February 03	84.6	128.0	26.7	35.3	45.3	65.0	46.3	56.8	57.4	29.0	24.5	32.5	43.8

March 03	84.0	136.8	26.9	35.8	45.8	65.6	50.6	54.3	72.2	32.7	24.5	32.5	43.8
April 03	83.7	146.7	26.9	35.8	45.8	65.6	37.0	33.3	48.1	35.9	24.5	32.5	43.8
May 03	84.0	163.4	26.9	35.6	45.8	65.6	28.4	38.9	42.6	31.2	24.5	33.6	43.8
June 03	83.4	157.5	26.9	35.8	46.1	65.6	27.8	37.7	34.6	22.8	24.5	33.6	43.8
July 03	82.8	147.6	27.5	36.7	46.7	66.4	20.4	17.9	20.4	26.3	24.5	33.6	43.8
August 03	84.6	153.5	28.1	37.2	47.5	67.2	16.7	17.9	20.4	31.2	24.5	33.6	43.8
September 03	89.8	149.6	28.6	38.1	48.3	68.3	19.8	24.1	54.3	26.7	24.5	33.6	43.8
October 03	91.4	137.8	29.7	40.0	50.0	71.7	28.4	34.6	28.4	21.1	26.4	35.1	45.3
November 03	91.4	139.8	30.3	41.7	51.7	74.7	35.8	42.0	82.7	16.6	27.5	37.0	47.2
December 03	92.9	137.8	31.4	43.3	53.3	77.8	50.6	66.0	87.0	24.0	29.4	38.9	49.8
January 04	94.2	146.7	31.7	44.4	54.2	79.4	57.4	88.3	79.6	26.0	31.3	39.2	52.1
February 04	91.1	154.5	32.2	45.8	55.0	81.4	46.9	50.0	76.5	34.7	31.3	41.5	52.1
March 04	96.0	167.3	32.8	46.9	56.1	83.6	37.7	51.2	63.6	33.7	31.3	44.2	52.8
April 04	90.2	159.4	33.1	47.5	56.7	85.6	27.8	40.1	53.1	21.5	31.3	44.2	52.8
May 04	89.8	163.4	33.6	48.3	57.2	88.1	28.4	44.4	56.8	21.5	31.3	45.7	54.3
June 04	91.4	165.4	33.9	48.9	58.1	90.0	34.6	46.9	74.1	28.0	32.1	47.5	58.5
July 04	88.9	161.4	34.2	50.0	61.1	93.3	32.7	52.5	79.0	30.7	32.5	49.8	60.0
August 04	90.8	173.2	34.4	51.7	63.9	96.9	32.1	45.7	62.3	26.7	33.6	51.3	62.3
September 04	89.8	143.7	34.7	52.8	66.4	99.7	32.7	46.3	63.6	26.7	34.7	53.2	66.8
October 04	93.5	154.5	36.4	55.0	68.1	103.1	80.2	107.4	136.4	40.4	36.6	54.7	71.3
November 04	96.9	163.4	37.5	56.9	68.9	106.1	85.2	125.9	189.5	57.0	39.6	57.0	72.5
December 04	96.9	174.2	39.4	59.7	69.7	109.4	68.5	77.8	125.3	53.0	39.6	57.0	72.5

Month	Tanker, VLCC SH/DH 5Years Market Value (mill. USD)	Clean Carrier, 40/45000 dwt DB/DH 10 Years Second Hand Value (mill. USD)	Tanker Fleet, 10000 dwt+ Supply (mill. dwt)	Tanker Fleet, 10000 dwt+ Demand (mill. dwt)	Tanker Fleet, 10000 dwt + Utilisation Rate (%)	Tanker, Order Book in Percent of Existing Fleet (%)	Tankers Sold for Scrapping (mill. dwt)
January 95	52.2	16.8	260.9	218.5	83.9	9.2	4.2
February 95	51.9	17.3	259.8	216.3	83.5	9.1	4.2
March 95	51.9	17.9	258.7	215.2	83.0	9.0	4.2
April 95	51.9	18.2	258.7	213.0	82.6	8.8	4.0
May 95	51.9	18.7	258.7	215.2	83.9	8.6	4.0
June 95	51.9	19.1	258.7	218.5	85.0	8.4	4.0
July 95	52.2	19.1	258.7	220.7	85.9	8.1	1.0
August 95	53.3	19.1	258.7	220.7	85.9	7.9	1.0
September 95	53.7	19.1	258.7	220.7	85.9	7.6	1.0
October 95	53.7	19.1	258.7	220.7	86.1	7.2	1.5
November 95	53.7	19.1	258.7	221.7	86.5	7.0	1.5
December 95	53.7	19.1	258.7	222.8	87.0	6.8	1.5
January 96	54.8	19.1	258.7	223.9	87.4	6.5	0.9
February 96	55.9	19.1	259.8	230.4	87.0	6.3	0.9
March 96	56.7	19.1	260.9	237.0	86.5	6.0	0.9
April 96	56.7	19.1	262.0	242.4	86.3	6.1	3.5
May 96	56.7	19.1	263.0	237.0	86.1	6.3	3.5
June 96	57.0	19.1	264.1	231.5	86.1	6.5	3.5
July 96	57.4	19.1	264.1	226.1	86.1	6.8	1.3
August 96	57.4	19.1	264.1	227.2	86.3	7.0	1.3
September 96	57.8	19.1	264.1	228.3	86.5	7.4	1.3
October 96	58.9	19.6	264.1	229.3	86.7	7.6	1.1
November 96	59.3	19.6	265.2	229.3	86.7	7.6	1.1
December 96	60.0	20.3	266.3	229.3	86.7	7.6	1.1
January 97	61.9	21.0	267.4	229.3	86.7	7.6	1.2
February 97	62.6	21.0	267.4	230.4	87.0	9.2	1.2
March 97	62.6	21.0	266.3	230.4	87.2	10.4	1.2
April 97	61.9	21.0	266.3	231.5	87.2	11.3	1.3
May 97	61.9	21.0	265.2	231.5	87.6	12.0	1.3
June 97	61.9	21.0	265.2	232.6	88.3	12.6	1.3
July 97	61.9	21.0	265.2	233.7	88.9	13.6	0.3
August 97	61.9	21.0	265.2	234.8	89.6	14.2	0.3
September 97	63.7	21.0	265.2	237.0	90.0	14.7	0.3
October 97	64.8	20.6	265.2	239.1	90.9	15.2	0.8
November 97	65.6	20.0	265.2	239.1	90.4	15.2	0.8
December 97	64.4	19.6	266.3	238.0	90.0	15.2	0.8
January 98	63.0	18.9	267.4	237.0	89.1	15.4	0.1
February 98	61.9	18.5	268.5	238.0	89.3	15.4	0.1
March 98	60.7	17.6	268.5	239.1	89.6	15.4	0.3
April 98	60.0	17.0	269.6	240.2	89.8	15.5	0.5

May 98	58.5	16.5	269.6	239.1	89.3	15.5	0.7
June 98	57.4	16.0	270.7	238.0	88.5	15.5	0.7
July 98	55.6	15.0	270.7	237.0	87.8	15.6	0.3
August 98	52.6	13.6	271.7	235.9	87.4	15.8	0.0
September 98	48.9	13.1	271.7	235.9	87.0	16.5	0.9
October 98	48.9	13.1	272.8	235.9	86.7	16.2	1.3
November 98	48.9	12.6	272.8	235.9	86.5	15.8	0.8
December 98	51.5	12.1	273.9	235.9	86.5	15.2	1.0
January 99	54.8	13.9	275.0	237.0	86.5	14.7	0.1
February 99	54.8	13.9	276.1	235.9	86.1	14.6	1.0
March 99	54.8	13.9	277.2	234.8	85.4	14.4	1.2
April 99	53.7	13.9	277.2	231.5	84.3	14.1	1.9
May 99	53.7	13.8	277.2	230.4	83.7	13.9	0.7
June 99	53.7	13.8	277.2	229.3	83.3	13.6	0.7
July 99	53.7	13.5	277.2	228.3	82.8	13.4	0.8
August 99	53.7	13.5	277.2	227.2	82.6	13.1	0.9
September 99	53.7	13.5	278.3	228.3	82.6	12.8	1.7
October 99	53.7	13.5	278.3	229.3	82.8	12.6	0.9
November 99	52.6	13.5	279.3	231.5	83.0	12.9	3.3
December 99	52.6	13.8	279.3	234.8	84.1	13.1	2.4
January 00	54.8	14.4	279.3	237.0	85.0	13.3	2.5
February 00	56.7	14.9	280.4	240.2	86.1	13.5	3.2
March 00	60.0	15.9	280.4	244.6	87.4	13.9	1.3
April 00	61.9	17.2	281.5	250.0	88.9	14.4	1.5
May 00	62.2	17.2	281.5	252.2	90.0	15.0	0.5
June 00	63.0	17.2	281.5	253.3	90.4	15.6	0.8
July 00	64.1	17.5	281.5	254.3	90.9	16.1	1.9
August 00	70.0	18.1	281.5	255.4	91.3	16.3	0.2
September 00	70.0	18.1	282.6	258.7	92.2	16.6	0.1
October 00	70.0	18.1	282.6	262.0	92.8	16.9	1.1
November 00	70.0	18.6	282.6	264.1	93.5	17.2	0.6
December 00	70.0	19.1	283.7	260.9	92.6	17.6	0.0
January 01	70.0	19.1	283.7	258.7	91.5	17.8	0.3
February 01	67.8	19.1	284.8	255.4	90.7	18.2	0.5
March 01	67.8	19.1	284.8	253.3	89.6	18.9	1.2
April 01	67.8	19.1	284.8	251.1	88.9	19.4	0.7
May 01	67.8	19.1	284.8	250.0	88.0	20.3	1.4
June 01	67.8	19.1	284.8	248.9	87.8	21.1	1.0
July 01	66.3	18.8	284.8	248.9	87.6	22.4	1.0
August 01	64.8	18.0	284.8	247.8	87.4	22.2	0.8
September 01	63.3	17.4	284.8	246.7	87.2	21.7	1.2
October 01	62.6	16.7	284.8	246.7	87.0	21.4	1.4
November 01	61.5	16.0	284.8	246.7	87.0	21.0	4.0
December 01	60.0	15.5	284.8	242.4	85.7	20.6	1.5
January 02	58.5	15.1	283.7	235.9	83.9	20.3	1.7
February 02	58.5	15.1	282.6	231.5	82.6	20.5	2.7
March 02	57.4	14.1	281.5	231.5	82.6	20.8	1.3
April 02	56.3	14.1	281.5	232.6	82.8	21.0	2.3
May 02	54.8	14.1	281.5	232.6	83.0	20.6	3.6
June 02	54.8	14.1	281.5	233.7	83.5	20.0	0.3
July 02	54.8	14.1	280.4	234.8	84.3	19.5	0.4
August 02	54.8	14.1	280.4	235.9	85.2	19.5	1.3
September 02	51.9	14.1	280.4	240.2	86.3	19.5	1.9
October 02	51.9	14.1	280.4	244.6	87.6	19.5	0.9
November 02	51.9	14.1	281.5	248.9	88.9	19.8	0.7
December 02	51.9	15.0	281.5	252.2	90.0	20.1	1.2
January 03	57.8	16.1	282.6	254.3	90.4	20.6	1.1
February 03	58.5	16.1	283.7	255.4	91.1	21.1	0.4

March 03	58.5	16.1	284.8	257.6	90.9	21.5	0.8
April 03	58.5	16.1	285.9	257.6	90.4	21.9	1.6
May 03	60.0	16.1	288.0	257.6	90.0	22.2	4.5
June 03	60.0	16.1	289.1	256.5	89.3	22.7	1.6
July 03	60.0	16.1	290.2	255.4	88.9	23.1	1.8
August 03	60.0	16.1	290.2	254.3	88.5	23.5	3.5
September 03	60.0	16.1	291.3	255.4	88.3	23.9	0.6
October 03	62.6	16.1	291.3	259.8	89.3	24.4	1.0
November 03	68.1	17.4	292.4	264.1	90.4	24.8	0.6
December 03	72.6	19.1	292.4	265.2	91.3	25.7	0.6
January 04	75.2	19.1	292.4	266.3	91.5	26.5	0.6
February 04	75.6	21.1	293.5	267.4	91.7	26.4	0.4
March 04	75.9	21.1	293.5	267.4	92.0	26.4	1.5
April 04	75.9	21.1	296.7	266.3	91.1	26.4	1.3
May 04	81.5	21.1	298.9	266.3	90.4	26.5	0.8
June 04	84.8	21.2	300.0	268.5	89.6	26.6	1.0
July 04	87.4	23.3	301.1	271.7	90.4	26.8	0.7
August 04	90.7	24.1	302.2	275.0	91.3	26.9	0.6
September 04	93.3	25.7	303.3	278.3	92.2	27.0	0.4
October 04	101.5	26.9	304.3	280.4	93.0	27.1	0.1
November 04	105.2	28.1	305.4	282.6	93.7	26.9	0.4
December 04	105.2	28.1	306.5	284.8	94.3	26.5	0.3

List of Abbreviations

Abbreviation	Term
ADALINE	ADAptive LINear Elements
ANN	Artificial Neural Networks
ANOVA	Analysis of Variance
ART	Adaptive Resonance Theory
BIMCO	Baltic and International Maritime Council
CA	Correspondence Analysis
CAT	Category
CC	Canonical Correlation
CCA	Canonical Correlation Analysis
Cov	Covariance
CV	Cross Validation
DA	Discriminant Analysis
DF	Degrees of Freedom
DNV	Det Norske Veritas
DWT	Deadweight Tonnage
EC	European Commission
ESM	Environmentally Sound Management
ESS	Error Sum of Squares
GLM	Generalised Linear Model
GT	Gross Tonnage
ILO	International Labour Organisation
IMO	International Maritime Organisation
ISL	Institute of Shipping Economics and Logistics
LDT	Light displacement tonnes or Lightweight
LNG	Liquefied Natural Gas
MADALINE	Multiple ADAptive LINear Elements
MAE	Mean Absolute Error
MANOVA	Multivariate Analysis of Variance
MARPOL	Marine Pollution (The International Convention)
MEPC	Marine Environment Protection Committee
MCA	Multiple Correspondence Analysis
MDS	Multidimensional Scaling
MFA	Multiple Factor Analysis
MLP	Multi Layer Perceptron
MLR	Multiple linear regression
MSE	Mean Square Error
NLPCA	Non Linear PCA
OECD	Organisation for Economic Co-operation and Development
OPEC	Organisation of the Petroleum Exporting Countries
PC	Principal Componenet
PCA	Principal Component Analysis
PCBs	Polychlorinated Biphenyls
PCR	Principal Component Regression
PE	Processing Element
PLS	Partial Least Squares
PLSR	Partial Least Square Regression
PNN	Probabilistic Neural Networks
RBF	Radial Basis Functions
RNN	Recurrent Neural Networks

Abbreviation	Term
RRR	Reduced Rank Regression
RSS	Residual Sum of Squares
RMSE	Root Mean Square Error
RMSEC	Root Mean Square Error of Calibration
RMSEP	Mean Square Error of Prediction
SOM	Self-Organised Maps
SS	Sum of Squares
TBT	Tributyltin
TDNN	Time Delay Neural Network
TEU	Twenty-foot Equivalent Unit
TLRN	Time Lagged Recurrent Network
UNCTAD	United Nations Conference on Trade and Development
ULCC	Ultra Large Crude Carrier
UNEP	United Nations Environment Programme
VLCC	Very Large Crude Carrier

GLOSSARY

TERM	ABBREVIATION / ACRONYM	EXPLANATION
ACTIVATION FUNCTION		A mathematical function that a neuron uses to produce an output referring to its input value. Usually this input value has to exceed a specified threshold value that determines, if an output to other neurons should be generated.
AFRAMAX		Tankers generally 80,000-119,000 DWT
ARTIFICIAL NEURAL NETWORKS	ANN	A mathematical or computational modelling based on the biological neural networks.
BACKPROPAGATION		A learning algorithm used by artificial neural networks with supervised learning. Special form of the delta learning rule.
BACKPROPAGATION NETWORKS		A feedforward type of artificial neural networks. They have one input layer, one output layer and at least one hidden layer.
BALLAST		Seawater taken into a vessel's tanks in order to submerge the vessel to proper trim.
BALTIC AND INTERNATIONAL MARITIME COUNCIL	BIMCO	Trade organisation representing shipowners, shipbrokers and agents, and other members
BULK CARGO		Usually a homogeneous cargo stowed in bulk, and not enclosed in any container.
CANONICAL CORRELATION ANALYSIS	CCA	A tecnique for identifying and quantifying the relations between a p-dimensional random X-variable and a q-dimensional random Y-variable.
DEADWEIGHT TONNAGE	DEADWEIGHT, DWT	The lifting or carrying capacity of a ship when fully loaded. The deadweight is the difference, in tonnes, between the displacement and the lightweight. It includes cargo, bunkers, water (potable, boiler, and ballast), stores, passengers and crew.
DELTA LEARNING RULE		A learning algorithm used by artificial neural networks with supervised learning. Effects the changing of weights by multiplying a neuron's input with the difference of its output and the desired output and the learning rate.
DET NORSKE VERITAS	DNV	One of several Classification Societies - The professional organisations which class and certify the strength and seaworthiness of vessel construction. Class and certification issued to each vessel may be required for insurance purposes. DNV and Lloyds Register of Shipping are two well known classification societies in the world today.
FEED BACK		A specific connection structure of the artificial neural networks, where neurons of one layer may have connections to neurons of other layers and also to neurons of the same layer.
FEED FORWARD		A specific connection structure of the artificial neural networks where neurons of one layer may only have connections to neurons of the next layer.
GROSS TONNAGE	GT	The internal capacity of a vessel measured in units of 100 cubic feet.
HIDDEN LAYER		A type of layer that lies between the artificial neural

		networks' input and output layers. Called "hidden", because its neuron values are not visible outside the network. The usage of hidden layers extends the neural networks' abilities to learn logical operations.
INTERNATIONAL LABOUR ORGANISATION	ILO	The UN agency seeking the promotion of social justice and internationally recognized human and labour rights
INTERNATIONAL MARITIME ORGANISATION	IMO	The United Nations' agency responsible for improving maritime safety and preventing pollution from ships.
INPUT		A set of values, called "pattern", that is passed to the artificial neural networks' input layer.
INPUT LAYER		The first layer of the artificial neural networks that accepts certain input patterns and generates output values to the successing weight matrix.
LIGHT DISPLACEMENT TONNES OR LIGHTWEIGHT	LDT	The lightweight is the displacement, in t, without cargo, fuel, lubricating oil, ballast water, fresh water and feed water, consumable stores and passengers and crew and their effects, but including liquids in piping.
LEARNING ALGORITHM		A mathematical algorithm that the artificial neural networks use to learn specific problems.
LEARNING RATE		A changeable value used by several learning algorithms, which effects the changing of weight values. The greater the learning rate, the more the weight values are changed. Is usually decreased during the learning process.
MULTIVARIATE ANALYSIS OF VARIANCE	MANOVA	A classical statistical analysis method to assess the significance of effects by decomposition of a response's variance into explained parts.
	MARPOL	International Convention for the Prevention of Pollution from Ships, 1973, as modified by the Protocol of 1978 relating thereto (MARPOL 73/78).
MARINE ENVIRONMENT PROTECTION COMMITTEE	MEPC	IMO's senior technical body on marine pollution related matters.
MULTI LAYER PERCEPTRON	MLP	A feedforward type of artificial neural networks. Built of an input layer, at least one hidden layer and one output layer.
MULTIPLE LINEAR REGRESSION	MLR	A multivariate analysis method which relates the variations in a response variable (Y-variable) to the variations of several predictors (X-variables).
NEURON		An element of the artificial neural networks' layers.
ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT	OECD	An international organisation of those developed countries that accept the principles of representative democracy and a free market economy
OUTPUT		A value or a set of values (pattern), generated by the neurons of the artificial neural networks' output layer. Used to calculate the current error value of the net.
OUTPUT LAYER		The last layer of the artificial neural network that produces the output value of the net.
PANAMAX.		The maximum size ship that can fit through the Panama Canal in terms of width, length and draught generally about 80,000 DWT
PRINCIPAL COMPONENT ANALYSIS	PCA	The amount of variance in a variable that is shared by all the variables in the analysis.
POLYCHLORINATED	PCBs	A mixture of individual chemicals which are either

BIPHENYLS		oily liquids or solids that is colourless to light yellow.
PRINCIPAL COMPONENT REGRESSION	PCR	A regression analysis method including two-step procedure which first decomposes the X-matrix by PCA, then fits a MLR model.
PROCESSING ELEMENT	PE	An element of the artificial neural networks' layers.
PERCEPTRON		A feedforward type of the artificial neural networks. Built of one input layer and one output layer.
PARTIAL LEAST SQUARES	PLS	A regression analysis method which is the extension of the MLR.
PROPAGATION		The passing of values and errors through the different layers of the artificial neural networks during its learning process.
SUEZMAX		The maximum size ship that can sail through the Suez canal generally considered to be between 150-200,000 DWT depending on ships dimensions and draught.
SUPERVISED LEARNING		A specific type of a learning algorithm. The output (pattern) of the network is compared with a target output (pattern). Depending on the difference between these patterns, the network error is computed.
TRIBUTYLTIN	TBT	One of the most poisonous substances to be released to the aquatic environment.
TWENTY-FOOT EQUIVALENT UNIT	TEU	Standard unit for counting containers of various capacities and for describing the capacities of container ships or terminals. One 20 Foot ISO container equals 1 TEU.
THRESHOLD		A specific value that must be exceeded by a neuron's activation function, before this neuron generates an output.
ULTRA LARGE CRUDE CARRIER	ULCC	Tanker of 320,000 DWT & above
UNSUPERVISED LEARNING		A specific type of a learning algorithm. Unlike supervised learning, no target patterns exist.
VERY LARGE CRUDE CARRIER	VLCC	Tanker of 160,000-320,000 DWT
WEIGHT		An element of a weight matrix. A connection between two neurons with a value that is dynamically changed during the artificial neural networks' learning process.
WEIGHT MATRIX		The connection structure between two layers of the artificial neural networks. Its elements, the weights, are changed during the network's learning process.