Combining Dynamic Factor Models and Artificial Neural Networks in Time Series Forecasting with Applications

A dissertation submitted in fulfillment of the academic requirements for the degree of doctor philosophy in Statistics

By

Ali Basher Abd Allah Babikir
School of Mathematics, Statistics and Computer Science
University of KwaZulu-Natal
South Africa

April 2014

Preface

The work described in this thesis was carried out in the School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Pietermaritzburg, South Africa, under the supervision and direction of Professor Henry Mwambi (School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal).

I would like to declare that this thesis is my own work. It has not been submitted in any form for any degree or diploma to any other University. It, therefore, represents my original work. Where use has been made of the work of others, it is duly acknowledged within the text or references chapter.

....

Signature (Ali B A Babikir) Date

As the candidate's supervisor, I certify the above statement and have approved this thesis for submission.

....

Signature (Prof. Henry Mwambi) Date

Declaration 1 – Plagiarism

I, Ali B A Babikir, declare that:

- The research reported in this thesis, except where otherwise indicated, is my original research.
- This thesis has not been submitted for any degree or examination at any other university.
- This thesis does not contain other person's data, pictures, graphs, or other information, unless specifically acknowledged as being sourced from other persons.
- This thesis does not contain other person's writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
    a) Their words have been re-written, but the general information attributed to them has been referenced.
    b) Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
- This thesis does not contain text, graphics, or tables copied and pasted from the internet, unless specifically acknowledged and the source being detailed in the thesis and in the references section.

Declaration 2 - Publications and manuscripts

I, Ali B A Babikir, declare that:

- Ali Babikir and Henry Mwambi (Under review). Factor Augmented Artificial Neural Network Model. *Journal of Economics and Econometrics*
- Ali Babikir and Henry Mwambi (Under review). Artificial Neural Networks – Dynamic Factor Model (ANN-DFM). *Journal of Economics and Statistics.*
- Ali Babikir and Henry Mwambi (Accepted). A factor - artificial neural network model for time series forecasting. *IEEE Transactions on Neural Networks and Learning Systems.*
- Ali Babikir and Henry Mwambi (Under review). Evaluating the combined forecasts of the dynamic factor model and the artificial neural network model using linear and nonlinear combining methods. *Journal of Forecasting*.

Dedication

To Almighty GOD for his mercy that endures forever and for being my constant source of inspiration despite all difficulties. To my parents, thanks you for your endless prayers, support and encouragement. To my brothers, sisters and to the departed soul of my brother Mohammed, may Allah forgive him and grant him paradise. I would like to thank my parents in low for their continuous support and encouragement during the years of my study.

Finally and most important, I would like to thank my wife for her unconditional love, support and encouragement during my ups and downs. She has been a safe haven for me during troubled times and a rock solid home base for our wonderful four daughters Amal, Malak, Fatima and Nour. It goes without saying that I dedicate this thesis to her as well.

Acknowledgements

I would like to acknowledge many people for helping me during my doctoral work. First of all, I thank Allah for granting me power, support, ability and much more to make it possible to complete my PhD thesis. Special thank and deep gratitude to my thesis advisor, Prof. Henry Mwambi for his support and guidance. Without his patience, encouragement and support this research would not have been possible. This thesis would have been less readable, if he had not carefully read and commented on my early drafts. I am looking forward to our permanent collaboration. I apologize to my family for the time not spent with them during the preparation of this thesis and I am grateful for their understanding and patient. I wish to thank the anonymous referees who gave useful comments on the papers during the course of this work. My appreciation extends to the School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal for friendly environment that I enjoyed during the course of my study. I am also grateful to my friends Ahmed Abdelmageed, Faiz Awad, Ali Satty and Mohammed Omer for their consistent encouragement and help throughout my research.

Abstract

This study investigates and examines the advantages and forecasting performance of combining the dynamic factor model (DFM) and artificial neural networks (ANNs) leading to new novel models that have capabilities to produce more accurate forecasts with application to the South African financial sector data. The overall aim of the study is to provide forecasting models that accommodate all relevant variables and the presence of any nonlinearity in the data to produce more adequate forecasts and serve as an alternative to traditional and current forecasting models, particularly in the presence of a changing and interacting environment.

The thesis consists of four independent papers corresponding to four chapters. The first chapter brings together two important developments in forecasting literature; the artificial neural networks (ANNs) and factor models. The chapter introduces the Factor Augmented Artificial Neural Network (FAANN) hybrid model in order to produce a more accurate forecasting. The model is applied to forecasting three time series variables, namely, Deposit rate, Gold mining share prices and Long term interest rate. The out-of-sample root mean square error (RMSE) and Diebold-Mariano test results show that the FAANN model yields substantial improvements over the autoregressive AR benchmark model and standard dynamic factor model (DFM). The superiority of the FAANN model is due to the ANNs flexibility to account for potentially complex nonlinear relationships that are not easily captured by linear models.

In the second chapter we introduce a new model that exploits the artificial neural networks model as a data smoother to alleviate the effect of major financial crisis and nonlinearity due to

high fluctuations such as those associated with the 2008 crisis. The chapter introduces the ANN-DF model, where in the first stage the best fitted ANNs for each single series of the data set which contains 228 monthly series is used to obtain the in-sample forecasts of each series. In the second stage, the factor model is used to extract the factors from the smoothed data set, and then these factors are used as explanatory variables in forecasting. The model is applied to forecast three South Africa variables, namely, Rate on 3-month trade financing, Lending rate and Short term interest rate in the period 1992:01 to 2011:12. The results, based on the root mean square errors of three, six and twelve months ahead out-of-sample forecasts over the period 2007:01 to 2011:12 indicate that, in all of the cases, the ANN-DFM and the DFM statistically outperform the autoregressive (AR) models. In the majority of the cases the ANN-DFM outperforms the DFM. The results indicate the usefulness of smoothing and factor extraction in forecasting performance. The forecast results are confirmed by the test of the equality of forecast accuracy proposed by Diebold-Mariano (1995).

The third chapter evaluates the role of the DFM model (liner in nature) and the ANN model (with capacity to handle nonlinearity) as competing forecasting estimation methods. The chapter uses artificial neural networks (ANNs) as nonlinear method based on the fact that the relationships between input and output variables in ANNs do not need to be specified in advance. In this chapter, the same extracted factors are used as input and independent variables for ANNs and the Dynamic Factor Model. This was necessary in order to investigate the forecasting performance of the linear and the nonlinear methods under the same conditions. We refer to the new model as Factor Artificial Neural Network (FANN). The empirical results of the Root Mean Square Error (RMSE) for the out-of-sample forecasts from 2007:01 to 2011:12 indicate that the proposed FANN model is an effective way to improve forecasting accuracy over the Dynamic Factor Model (DFM), the ANN and the AR benchmark model. The results confirm the usefulness of the factors that were extracted from a large set of related variables when we compared the FANN model and the standard univariate ANN model.

Finally, combining forecasts is often considered as a successful alternative to using just an individual forecasting method. Different forecasting methods are considered especially when the forecasts are generated form the linear and the nonlinear methods. Thus, chapter four investigates the forecasting performance of combining independent forecasts of the Dynamic Factor Model and the Artificial Neural Networks models using linear and nonlinear combining procedures for the same variables of interest. The analysis was based on three financial variables namely the JSE return index, government bond return index and the Rand/Dollar exchange rate in South Africa. The out-of- sample results of three, six and twelve month horizons from 2006:01 to 2011:12 for the DFM and ANNs provided more adequate forecasts compared to benchmark auto-regressive (AR) models with reduction in the RMSE of around 2 to 12 percent for all variables and over all forecasting horizons. The ANN as a nonlinear combining method outperforms all linear combining methods and is the best individual model for all variables and over all forecasting horizons. The results suggest that the ANN combining method can be used as an alternative to linear combining methods to achieve greater forecasting accuracy. We attribute the superiority of the ANN combining method to its ability to capture any existing nonlinear relationship between the individual forecasts and the actual forecasting values.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background

Forecasting the future is an important issue in time series data analysis in order to aid in planning and the adoption of necessary policies that depend on variable of interest. Forecasting can assist policy makers with better decision making and prioritizing development agendas for the country or a specific economic sector. There are various available forecasting techniques in the academic literature; one of the most recent important developments in forecasting literature is the Dynamic Factor Model which has become popular in empirical macroeconomics in forecasting of economic variables of interest. Factor models have more advantages than other methods in various respects. They can handle a large amount of information (the number of variables can be more than the number of observations) which is to say that we can get accurate forecasting without running into scarce degrees of freedom problems which are often faced in regression based analyses. Current problems in time series are multidimensional data involving more than one time series, but luckily modern computers and software allow us to efficiently summarize the information contained in large datasets. Realistically using a large body of information related to the variable of interest may lead to reduced errors and produce more precise forecasts. A second advantage of factor models is that factor modelers do not need to rely on tight assumption as is sometimes the case in structural models. As a result of these advantages, Dynamic Factor Models have been applied successfully in a number of research papers and for different countries to forecast key macroeconomic variables. These includes, among others, Stock and Watson (2002b) for the United States, Marecllino et al. (2003) for the Euro area countries, Schneider and Spitzer (2004) for Austria, Arits et al. (2005) for England, Schumacher (2007) for Germany, Bruneau et al. (2006) for France, Matheson (2006) for New Zealand, Reijer (2005) for the Netherland and Gupta and Kabundi (2010) for South Africa. In applied forecasting literature it is difficult to find a model that can beat the autoregressive AR model specifically in out-of-sample forecasting exercises. Liu and Jansen (2007) found that a structural Dynamic Factor Model outperformed AR model. Additionally, the Dynamic Factor Model had successful applications in short sample problems with respect to forecasting key macroeconomic variables, see for example, Banerjee et al. (2008), and Matheson (2006). An inclusive summary of the superior forecasting performance of the Dynamic Factor Model has been provided by Eickmeier and Ziegler (2008) who applied Meta analyses to 46 studies which compared the Dynamic Factor Model to other forecasting models such as autoregressive (AR). They found that, on average, factor models performed significantly better than the respective benchmark models in forecasting gross domestic product (GDP) and inflation. In spite of the success of the Dynamic Factor Model and its improvement over autoregressive and other linear models, there are still some limitations. One of the major limitations is merely to depict them as linear models, also known as the model driven approach. That is the requirement that they should fit the data with prior knowledge about the relationships between the inputs and outputs before modeling.

Due to these limitations, the nonlinear time series models have been proposed in order to improve the forecasting performance of nonlinear systems. These models include bilinear models, threshold autoregressive models, smoothing transition autoregressive models, autoregressive conditional heteroscedastic models and generalized autoregressive conditional heteroscedastic models. However, limited success has been found during the last two decades using nonlinear models since most of them are developed specifically for particular problems without applicability for other situations. In addition, the formulations of these models are more complex when compared to those based on the linear models; see for example Granger and Terasvirta (1998) as well as Terasvirta et al. (2005). Consequently, a different approach has been developed and successfully used in time series forecasting. This is the neural network techniques which has become an essential tool for economic and financial forecasting. Artificial Neural networks (ANNs) have been applied in many areas in time series forecasting problems such as stock; Hamid and Iqbal (2004), interest rate; Kumar and Chaturvedi (2011), exchange rates; Pacelli et al. (2011) and Kondratenko and Kuperin (2003), electricity prices; Ganeta et al. (2006) and Abraham and Nath (2001), tourism; Palmer et al., (2006) and Claveria and Torra, (2014), breast cancer; Delen et al. (2005) and in many other areas and applications. One of the main reasons for their attractiveness is that ANNs perform better than the Autoregressive Integrated Moving Average (ARIMA) and other linear models and that ANNs are universal function approximators capable of mapping any linear or nonlinear function. ANNs also do not require any knowledge or prior assumptions about the model form during the model building process, Ma and Khorasani (2004). The ANNs have been claimed as a major application area of the forecasting studies, Hippert et al. (2001) and Zhang (2004). However we also avoid overstating the superiority of ANNs because sometimes linear models can produce satisfactory results compared to ANNs when the linear part of the time series is superior to the nonlinear part. Examples can be found in Taskaya and Ahmed (2005), Heravi et al. (2004), Caire et al. (1992) and Brace et al. (1991). But, when the nonlinear part of the time series is superior to the linear part, ANN models can give satisfactory results; see Choudhary and Haider (2012), Duzgun (2010), Binner (2005), Franses and Griensven (1998) as well as Kang (1991). Thus, in both cases, if one of these components (linearity and non-linearity) is not taken into consideration the analysis may lead to deceptive results. To overcome the deficiency of using an individual linear or nonlinear model, various hybrid approaches have been suggested in the literature, and the idea is to merge different methods in order to improve the forecasting accuracy. Hybrid models can also be defined as combined models, which can be implemented in one of the three combining forms; linear models, nonlinear models and both linear and nonlinear models. The later form of combining models is an effective way to generate more adequate forecasts based on the fact that it is difficult to determine whether the time series under study is linear or nonlinear because the real world is highly complex, so there exist some linear and nonlinear patterns in the financial time series simultaneously. It is not sufficient to use only a nonlinear model for time series as the nonlinear model might miss some linear features of time series data and vice versa. In hybridization of linear and nonlinear models, two or more models are combined together using the same data set or different data sets to produce forecasts. The majority of the studies in this category combine ARIMA and ANNs models. Examples are given by Khashei and Bijari (2012), Lee and Tong (2011), Aladag et al. (2009), Chen and C.H. Wang (2007), Jain and Kumar (2007), Pai and Lim (2005), Lu et al. (2004), Zhang (2003), as well as Tseng and Tzeng (2002).

1.2 Problem statement and motivations

In the present changing financial world, forecasting financial variables is considered to be one of the most challenging tasks. Therefore, a lot of attention has been given to forecast future values of financial time series. Financial time series are affected and mitigated by different factors such as business cycles, monetary policies, general economic conditions, expectations and political events; in addition the financial sectors and markets are also affected by external and international indicators. Hence, forecasting the correct value of financial variables has become an important area of research interest in both developed and developing worlds. However, due to the volatility of the financial time series, there are some difficulties faced in building the above-mentioned factors, Ingoo et al. (2007).

This research aims to investigate the forecasting performance of combining the Dynamic Factor Model (DFM) as a large scale linear model and the Artificial Neural Networks (ANNs) model as a nonlinear model with application to financial variables from the South African financial sector. In terms of comparing forecasting accuracy, Altavilla and Grauwe (2010) and Kilian and Taylor (2003), stated that the linear models are accurate in short horizons while the nonlinear models are accurate in long horizons. Hence, combining linear and nonlinear models can lead to accurate forecasts in short and long horizons. Considering that the DFM is a large data model it can thus accommodate all series that affect and are related to the variables of interest. When the financial crisis happened none of the macroeconomics or financial models had an effective way to predict what happened to economies. Thus this thesis will cover the years 2007 through 2011 as the out-of-sample period to investigate the forecasting accuracy of our proposed models and combined forecasts.

The motivation and importance of this study comes from the following facts:

- So far all developed factor augmented models such as those introduced by Forni et al. (2000), Bernanke et al. (2005), Banerjee and Marcellino (2008), Ng and Stevanovic (2012) as well as Dufour and Pelletier (2013) augmented factors to the linear models such as DFM, factor-augmented vector autoregressive model (FAVAR), Factor augmented Error Correction Model (FECM), factor augmented autoregressive distributed lag

(FADL) and factor-augmented vector autoregressive moving average (VARMA) respectively. Their models lack the capacity to fully capture nonlinearity.

- The existing literature of the hybrid linear and nonlinear models only considers the hybridization of ANNs to univariate linear models such as ARIMA. Examples are given by Tseng et al. (2002), Zhang (2003), BuHamra et al. (2003), Jain and Kumar (2007) as well as Khashei and Bijari (2010), where both models only used the same variable of interest, in the sense that other related variables are not considered in the models.
- There is evidence that combination forecasts that pool linear and nonlinear forecasts can outperform combination forecasts that are based only on linear forecasts, see for example, Stock and Watson (1999), Blake and Kapetanios (1999) and Teräsvirta (2006).
- Much research shows that combined forecasts outperform the best individual forecasting model with regard to possible structural breaks in the data, Hendry and Clements (2004). According to Kabundi (2002) the ANNs are better equipped to capture the structural breaks when compared to linear regression model.
- The data generating process is likely to switch its structure over the observations period between the linear and the nonlinear structures particularly if the period is long. These structures can be captured by combining the linear and nonlinear models due to the fact that the real time series are rarely pure linear or nonlinear. If the data contains both linear and nonlinear patterns, then neither the Dynamic Factor Model nor the Neural Networks can be adequate in modeling and forecasting time series since the Dynamic Factor Model cannot deal with nonlinear relationships, while the Neural Network model alone is not able to handle both linear and nonlinear patterns equally well. Therefore, combining

different models can increase the chance to capture different patterns in the data and improve forecasting performance.

- The combined forecasts and hybrid model from combining both the Dynamic Factor Model and the Artificial Neural Networks is proposed to take advantage of the unique strength of both models and to fill the gap of the need for novel models that accommodate a large dataset and consider the nonlinearity at the same time. Thus, combined or hybrid models can be an effective way to improve forecasting accuracy achieved by either of the models used separately.
- The importance and the role of the financial sector for the South African economy cannot be overstated. The financial sector in South Africa comprises over R6 trillion in assets, contributing 10.5 percent of the gross domestic product of the economy annually, employing 3.9 percent of the working class population, and contributing at least 15 per cent of corporate income tax. Since 2000 the sector has grown at an annual rate of 9.1 percent, compared to the broader economic growth of 3.6 percent. Growth in employment has also been very strong; over the same period the number of people employed in the sub-sector increased by 24.5 percent and the financial sector has become one of the fastest growing employers in South Africa. The total assets of the sector have also grown significantly, registering nominal compound average growth of 12.3 percent between 2000 and 2012, and the financial sector assets now stand at 252 percent of the gross domestic product (GDP). Thus achieving accurate forecasts for the financial variables can provide decision makers with the information they require to make precise decisions about the economy.

1.3 Research objectives and contributions

One of the major challenging issues in forecasting is the lack of researches to facilitate and introduce new effective models or formulations that consider the availability of the large amount of time series data and consider their interaction, comovement, the financial linkage between countries and nonlinearity at the same time. This study investigates the forecasting performance of three new models compared to the traditional autoregressive (AR) benchmark model and DFM. These new models account for comovement and the interaction between large data time series variables that are related to the variable of interest to be forecast through the Dynamic Factor Model, and the effect of the nonlinearity that likely occurred as a result of structural and behavioral changes that are captured through artificial neural networks. The proposed models can serve as alternative time series forecasting models. Specifically, the study seeks to:

- Evaluate the forecasting performance of the factors when augmented to nonlinear artificial neural networks.
- Examine the role of the artificial neural networks when used as a data smoother before extracting the factors.
- Assess the role of estimation method – linear or nonlinear – used for the same factors in order to produce more adequate forecasts.
- Investigate the forecasting performance of combining the forecasts of the dynamic factor model and artificial neural networks model using linear and nonlinear combining methods.

Based on both Dynamic Factor Model and Artificial Neural Networks models using Matlab and R packages, these objectives are addressed in four independent papers corresponding to four chapters. The four papers have been submitted for peer review in international journals. Of these four papers, one paper has been accepted for publication and the others are still in review. Each paper has been written as a stand-alone article that can be read separately from the rest of the thesis but draws separate conclusions that link to the overall research objectives and questions. As a result, a number of overlaps and replications occur in the sections "Dynamic Factor Model", "Artificial Neural Networks" and "determination of the number of factors" in the different chapters. This problem is negligible when one considers the critical peer review process and the fact that the different chapters are papers that can be read separately without losing the overall context. Lastly, Chapter 6 should be read after chapters 1 to 5 as it summarizes the findings. A brief outline follows:

Chapter 1: This chapter serves as an introduction to the study.

Chapter 2: This chapter brings together two important developments in forecasting literature; the Artificial Neural Networks (ANNs) and factor models. The chapter introduces the Factor Augmented Artificial Neural Network (FAANN) hybrid model in order to produce a more accurate forecasting. The model has the merit to accommodate all variables that are related and affect the variables of interest through the factor model and consider the nonlinearity that may be inherent with dependent and independent variables through the ANNs. The chapter begins with a brief review to the factor estimation and how to determine the number of the factors. An overview of the forecasting models and the formulation of the proposed model are given. An application study including three financial variables, namely, Deposit rate, Gold mining share prices and Long term interest rate over three, six and twelve month out-of-sample forecast horizons is presented. The root mean square errors (RMSEs) and Diebold-Mariano test are used as forecast comparative measures. The chapter identifies the advantages of the new proposed model over Dynamic Factor Model (DFM). In other words, the advantage of augment factors to nonlinear model over the linear model. Lastly, the chapter ends with the findings and conclusions.

Chapter 3: Employs the Artificial Neural Networks (ANNs) model as a data smoother for each single series of the dataset before extracting the factors. This is to alleviate the effects of the financial crisis to the dataset such as downturns, fluctuations and nonlinearities. After the dataset has been smoothed, factors extracted are then used as explanatory variables in regression model. Three financial variables, namely, Rate on 3-month trade financing, Lending rate and Short term interest rate are used to assess the forecasting performance of the new model "Artificial Neural Networks - Dynamic Factor Model" (ANN-DFM) compared to standard DFM and autoregressive benchmark model are presented. In-sample and out-of-sample forecasts are discussed based on the RMSEs. The RMSEs results of the out-of-sample forecast are confirmed by Diebold-Mariano test for three, six and twelve month-ahead forecast horizons. The chapter ends with concluding remarks.

Chapter 4: In this chapter we investigate the advantage of exploiting a linear or nonlinear estimation method with same inputs or independent variables in order to produce more adequate forecasts. We propose a Factor - Artificial Neural Networks (FANN) model, where the extracted factors from a dataset of 228 series are used as inputs to the nonlinear ANNs method, and the forecasting capability of the new model is compared to the forecasting results of the DFM, where the same factors are used in both models. An overview to the estimation forecasting models is given, followed by the formulation of the introduced model. A description to the dataset and the criteria to determine the number of the factors is provided. Two financial variables, namely, Johannesburg Stock Exchange (JSE) share prices and the Treasury Bill Rate are used to evaluate the forecasting performance of the introduced model and the alternatives. The results of the in-sample and out-of-sample across the models are compared. The chapter ends with a summary of the key points.

Chapter 5: This chapter deals with forecasting combination of the Dynamic Factor Model (DFM) and Artificial Neural Networks (ANNs) forecasts. The chapter describes necessary notation and concepts regarding individual forecasting models and the AR benchmark model as well. Two families of forecasting combination methods are discussed: linear and nonlinear methods. An application study is carried out based on three financial variables, namely the JSE return index, government bond return index and the Rand/Dollar exchange rate in South Africa. The forecasting results of in-sample and individual models out-of-sample of three, six and twelve month-ahead forecast horizons are presented and compared based on the AR benchmark model. Discussed results under linear and nonlinear combining methods are obtained and compared to the benchmark model as well. Conclusions and remarks are drawn from the comparison.

Chapter 6: Finally, this chapter gives a synthesis of the thesis. The findings are summarized and conclusions are derived from the preceding chapters. For future work on the applications of the proposed models for forecasting time series, relevant recommendations are made. Special focus is directed towards the comparison of the FAANN model to the Factor Augmented Vector Autoregressive (FAVAR) model and we investigate the capabilities of the proposed models for nowcasting application.

A single reference list is given at the end of the thesis.

Chapter 2

Factor Augmented Artificial Neural Network Model*

2.1 Abstract

This chapter brings together two important developments in forecasting literature; the artificial neural networks (ANNs) and factor models. The chapter introduces the Factor Augmented Artificial Neural Network (FAANN) hybrid model in order to produce a more accurate forecasting. The model is applied to forecasting three time series variables using a large South African monthly panel. The out-of-sample root mean square error (RMSE) results show that the FAANN model yields substantial improvements over the autoregressive AR benchmark model and standard Dynamic Factor Model (DFM). The Diebold-Mariano test results also further confirm the superiority of the FAANN model forecasts performance over the AR benchmark model and the DFM model forecasts. The superiority of the FAANN model is due to the ANN flexibility to account for potentially complex nonlinear relationships that are not easily captured by linear models.

JEL classification: C22, C45, C53.

---

* Ali Babikir and Henry Mwambi (in review). Factor Augmented Artificial Neural Network Model. *Journal of Economics and Econometrics.*
* The paper has been accepted for presentation at *International Symposium on Forecasting* to be held in Rotterdam from June 29-July 2, 2014.

Keywords: Artificial neural network; Dynamic factor model; Factor Augmented Artificial Neural Network Model; Forecasting.

2.2 Introduction

The use of several common factors to summarize the information from a huge set of predictor variables has been the new frontier in the forecasting literature. Forecasting financial and economic variables often needs to incorporate information from a large set of potential explanatory variables into the forecasting model, whereas most traditional prediction models are unable to deal with this, either because they are inefficient or because it is impossible to incorporate a large number of variables in a single forecasting model and fit it using standard econometric techniques. An alternative approach to this problem is to use models that are based on factors which lead to an additional advantage in the forecasting area. Factor models were introduced in macroeconomics and finance by Sargent and Sims (1977), Geweke (1977), and Chamberlain and Rothschild (1983). The literature on the large factor models starts with Forni et. al (2000) and Stock and Watson (2002a). Further theoretical advances were made, among others, by Bai and Ng (2002), Bai (2003), and Forni et al. (2004). These models can be used to forecast macroeconomic aggregates [Stock and Watson (2002b), Forni et al. (2005), Banerjee et al. (2008)], perform structural macroeconomic analysis [Bernanke et al. (2005), Favero et al. (2005)], for nowcasting and economic monitoring [Giannone et al. (2008), Aruoba et. al (2009)], to deal with weak instruments [Bai and Ng (2010), Kapetanios and Marcellino (2010)], and the estimation of dynamic stochastic general equilibrium models [Boivin and Giannoni (2006)]. Bernanke et al. (2005) propose a forecasting model which they called the Factor-Augmented Vector Autoregressive (FAVAR) model, a model which merges a factor model with a vector autoregressive component. A factor-augmented vector autoregressive moving average (VARMA) model is suggested by Dufour and Pelletier (2013). Banerjee and Marcellino (2008) introduced the Factor Augmented Error Correction Model (FECM). The FECM combines error-correction, cointegration and dynamic factor models, and has several conceptual advantages over standard ECM and FAVAR models. Ng and Stevanovic (2012) proposed a factor augmented autoregressive distributed lag (FADL) framework for analyzing the dynamic effects of common and idiosyncratic shocks.

On the other hand, artificial neural networks (ANNs) have become one of the most accurate forecasting technologies and have been widely used in many areas of forecasting purpose. Artificial neural network has many features that make it attractive and accurate for forecasting tasks. First, ANNs are universal functional approximators. Second, ANNs are data-driven self-adaptive methods in that there are few a priori assumptions to be made about the models for problems under study, thus ANN modeling is different from traditional model-based methods. Third, an ANN model is by design nonlinear in contrast to traditional time series forecasting models which assume linearity of the series under consideration. Zhang et al. (1998) demonstrated that the systems of the real world are often nonlinear. These advantages of ANNs have attracted attention in time series forecasting and have become a competitive method to traditional time series forecasting methods and the literature is very vast in this area.

The hybrid approach or combining models represent the most important developments in ANNs over the last decade; the idea is to use the unique capability of each model component to better capture different patterns in the series. Research findings have demonstrated that combining different models can produce accurate forecasts, especially when the combined models are quite different. More hybrid models of ANNs with different forecasting models, which successfully improve the forecasting performance, have been introduced in the recent times.

Yu et al. (2005) proposed the integration of the generalized linear auto regression (GLAR) with artificial neural networks in order to obtain accurate forecasts for foreign exchange markets. Tseng et al. (2002) proposed a hybrid model called SARIMABP that combines the seasonal autoregressive integrated moving average (SARIMA) model and the back-propagation neural network model to predict seasonal time series data. Khashei and Bijari (2010) introduced a hybrid model of ANNs and ARIMA models for forecasting purposes.

In this chapter, we introduce the factor augmented artificial neural network (FAANN) model. The FAANN model complements the growing literature of the factor augmented models such as the DFM, FAVAR and FECM. In our case we propose a model where the factors are augmented to a nonlinear model, while the factors in the DFM, FAVAR and FECM models are augmented to linear models. The approach has several advantages. First, as it is difficult to surely know the characteristics of the data under consideration, a hybrid approach that has both linear and nonlinear modeling capabilities can be a better approach for forecasting. Second, a hybrid model of the factors - that are extracted from a large dataset that are related to the variable of interest - with the nonlinear ANN produce an accurate forecast.

The main contributions of our chapter are, (i) to bring together two important recent improvements of forecasting literature on modeling co-movement that have a common origin but, in their implementations, have remained apart, namely, Artificial Neural Network and Dynamic Factor Model (DFM); and (ii) to evaluate the role of factors when augmented to a nonlinear model as compared to linear factor augmented model (DFM) within empirical real data from South Africa where we look at both in-sample and out-of sample performance.

The rest of the chapter is organized as follows: in the next section, we review the factor model and how to determine the number of factors. In Section 2.4, we briefly review the DFM, ANN and introduce the FAANN modeling approaches to time series forecasting. Data is reported in Section 2.5. Empirical results from three real data sets are reported in Section 2.6. Section 2.7 provides a conclusion of the main findings of the chapter and suggests directions for additional research in this area.

2.3 Methodology

In this section, the basic concepts of the estimation of factors and determination of the number of factors are briefly reviewed.

2.3.1 Estimation of the Factors

The dynamic factor (DF) model that extracts common components between the dataset series and then uses these common components to forecast represents the recent direction in forecasting literature as the availability of economic and financial time series grows in terms of both time and cross-section size. The DF model expresses individual times series as the sum of two unobserved components: a common component driven by a small number of common factors and an idiosyncratic component for each variable. The DF model extracts the few factors that explain the co-movement of the economy and financial sector. Forni et al. (2005) demonstrate that for a small number of factors relative to the number of variables and a heterogeneous panel, we can recover the factors from present and past observations.

Consider a $n \times 1$ covariance stationary process $Y_t = (y_{1t}, \ldots\ldots, y_{nt})'$. Suppose that X_t equals the standardized version of Y_t (i.e., X_t possesses a mean zero and a variance equal to one). Under DF models, we write X_t as the sum of two orthogonal components as follows:

$$X_t = \lambda F_t + \xi_t \tag{2.1}$$

where F_t equals a $r \times 1$ vector of static factors, λ equals an $n \times r$ matrix of factor loadings, and ξ_t equals a $n \times 1$ vector of idiosyncratic components. In a DF model, F_t and ξ_t are mutually orthogonal stationary process, while, $\chi_t = \lambda F_t$ equals the common component.

Since dynamic common factors are latent, we must estimate them. We note that the estimation technique used matters for factor forecasts. We adopt the Stock and Watson (2002b) method, which employs the static principal component (PC) approach[1] on X_t. The factor estimates, therefore, equal the first principal components of X_t, (i.e., $F_t = \widehat{\Lambda}' X_t$, where $\widehat{\Lambda}$ equals the $n \times r$ matrix of the eigenvectors corresponding to the r largest eigenvalues of the sample variance covariance matrix $\hat{\Sigma}$.

2.3.2 Determination of the number of factors

Recently the theory and criteria of the determination of the number of factors has been developed by, among others, Hallin and Liska (2007), Stock and Watson (2005) and Bai and Ng (2007) for the Forni, Hallin, Lippi and Reichlin class of models and by Bai and Ng (2002) for the class of dynamic factor models in the static representation. Onatski (2009) developed a statistical test to test and determine the number of dynamic factors under the null hypothesis that the number of

[1] The approach works as follow; Suppose a factor model is represented as $X_{it} = \lambda'_i F_t + e_{it}$ where X_{it} is the observed datum for the *i*th series at time *t* (*i*=1, …, *N*; *t*=1, …, *T*); F_t is a vector $(r \times 1)$ of common factors; λ_i is a vector $(r \times 1)$ of factor loadings; and e_{it} is the idiosyncratic component of X_{it}. The right hand side variables are not observed. The method of principal components minimizes $V(r) = min_{\Lambda,F}(NT)^{-1}\sum_{i=1}^{N}\sum_{t=1}^{T}(X_{it} - \lambda'_i F_t)^2$ where $\Lambda = (\lambda_1, \ldots., \lambda_N)$. Concentrating out Λ and using the normalization that $F'F/T = I_r$, where I_r is $r \times r$ identity matrix, the problem is identical to maximizing $tr(F'(XX')F)$. The estimated factor matrix, denoted by $\tilde{F}$, is $\sqrt{T}$ times the eigenvectors corresponding to the r largest eigenvalues of the $T \times T$ matrix $X'X$, and $\tilde{\Lambda}' = (\tilde{F}'\tilde{F})^{-1}\tilde{F} X = \tilde{F}X/T$ is the corresponding loading matrix.

[2] In this paper we choose iterated forecast instead of direct forecast. Marcellino et. al (2006) found that iterated forecast using AIC lag length

factors is equal to k_0 against the alternative $k_1 > k_0$ (for details see Onatski 2009). The aforementioned papers proposed various information criteria to guide us in the selection of the number of factors. In this chapter different panel information criteria developed by Bai and Ng (2002) are applied.

Principal component analysis with r factors extracted from a dataset in X allows for the calculation of the sum of squared residuals $V(r) = (NT)^{-1}\sum_{t=1}^{T}\hat{\xi}_t\hat{\xi}_t^T$ where $\hat{\xi}_t$ is a $N \times 1$ vector of the estimated idiosyncratic errors. Based on this quantity Bai and Ng (2002) suggest a number of information criteria of which some of the most popular are shown below:

$$min_r IC_{p2}(r) = ln\big(V(r)\big) + r(\frac{N+T}{NT}) lnC_{nt}^2 \tag{2.2}$$

$$min_r IC_{p3}(r) = ln\big(V(r)\big) + r(\frac{lnC_{nt}^2}{C_{nt}^2}) \tag{2.3}$$

where the sequence of constants $C_{nt}^2 = \min\{N, T\}$ represent the convergence rate for the principal component estimator. We apply the approach which proposes five static factors. The Onatski (2009) test confirms the result of the used approach, the selected number of the factors explain more than 80 percent of variation of the entire data panel.

2.4 Forecasting models

In this section, the basic concepts and modeling approaches of the dynamic factor model (DFM), autoregressive model (AR) and artificial neural networks (ANNs) models for time series forecasting are presented. The section also introduces the formulation of the proposed model.

2.4.1 Dynamic Factor model forecast

The estimated factors will be used to forecast the variables of interest. The forecasting model is specified and estimated as a linear projection of an h-step ahead transformed variable $\mathrm{y}_{\mathrm{t+h}}^{h}$ into t-dated dynamic factors. The forecasting model follows the setup in Stock and Watson (2002a) and Froni et al. (2003) which takes the form:

$$y_{t+h}^{h} = \beta(L)\hat{F}_t + \gamma(L)y_t + u_{t+h}^{h} \tag{2.4}$$

where $\hat{F}_t$ are dynamic factors estimated using the method by Stock and Watson (2002b) while $\beta(L)$ and $\gamma(L)$ are the lag polynomials, which are determined by the Schwarz Information Criterion (SIC). The u_{t+h}^{h} is an error term. The coefficient matrix for factors and autoregressive terms are estimated by ordinary least square (OLS) for each forecasting horizon h.

2.4.2 Autoregressive (AR) Forecast

The AR model is given by

$$y_t = \phi + \gamma(L)y_t + e_t \tag{2.5}$$

where y_t is the variable to forecast, ϕ is a constant, $\gamma(L)$ is the iteratively estimated lag polynomial, the lag order is chosen by SIC and e_{t} is the error term.

The h-step ahead forecast from this model is

$$y_{t+h}^{h} = \phi + \gamma^{h}(L)\, y_t + e_{t+h}^{h} \tag{2.6}$$

where y_{t+h}^{h} is the h-step ahead forecast[2] of y_t, $\gamma^{h}(L)$ are the iteratively estimated lag polynomials, e_{t+h}^{h} is the h-month ahead forecast error term.

[2] In this paper we choose iterated forecast instead of direct forecast. Marcellino et. al (2006) found that iterated forecast using AIC lag length selection performed better than direct forecasts, especially when forecast horizon increases. They argued that iterated forecast models with lag length selected based on information criterion are good estimates to the best linear predictor.

The benchmark AR forecast is individually applied to our variables of interest, namely, Deposit rate, Gold mining share prices and Long term interest rate. The optimal lag length is chosen by SIC.

2.4.3 The ANN and the formulation of the FAANN model

Artificial neural networks (ANNs) are model free dynamic, which are widely used for forecasting. One of the important advantages of the ANN models over other classes of nonlinear models is that ANNs are universal approximators that can approximate a large class of functions with a high degree of accuracy. See Chen et al. (2003) for more details. There is no need for prior assumptions about the model form during the model building process.

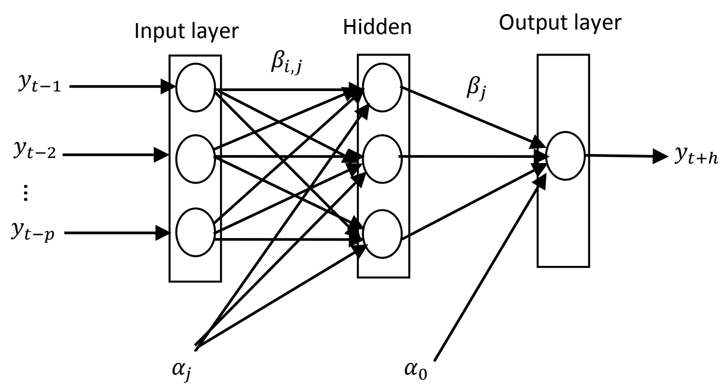

Figure 2.1 Neural network model ($\boldsymbol{N}^{(\boldsymbol{p},\mathbf{3},\mathbf{1})}$)

Fig. 2.1 shows a popular three-layer feed-forward neural network model. It consists of one input layer with p input variables, one hidden layer with q hidden nodes, and one output layer with a single output node. The hidden layers perform nonlinear transformations on the inputs from the input layer and feed the transformed values to the output layer. The connection weights and node biases are the model parameters. The model estimation process is called network training. Usually in applications of ANNs, the total available data are split into a training set and a test set. The training set is used to calibrate the network model, while the test set is used to evaluate its forecasting ability. During the training procedure, an overall error measure is minimized to get the estimates of the parameters of the models. The mathematical representation of the model in Fig. 2.1 that show the relationship between output (y_t) and the inputs ($y_{t-1}, y_{t-2}, \ldots, y_{t-p}$) is given by;

$$y_{t+h} = \alpha_0 + \sum_{j=1}^{q} \alpha_j g\left(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij}\, y_{t-i}\right) + \varepsilon_{t+h} \tag{2.7}$$

where $\alpha_j (j = 0,1, \ldots, q)$ and $\beta_{ij} (i = 0,1, \ldots, p; j = 1,2, \ldots, q)$ are the model parameters often called the connection weights. As we stated above p and q are the number of input nodes and hidden nodes respectively, ε_t is error term. The logistic function is usually used as the hidden layer transfer function, which is generally given by;

$$g(x) = \frac{1}{1+\exp(-x)} \tag{2.8}$$

There are many different approaches to find the optimal networks but these approaches are quite complicated and are difficult to implement, and in addition there is no guarantee that the optimal solution of these approaches is optimal for all real forecasting problems. Thus the procedure often used to determine (p, q) is to test numerous networks with different numbers for (p, q) to select the network that minimize the error. The minimization is done with some efficient nonlinear optimization algorithm; in our case we use Broyden, Fletcher, Goldfarb and Shanno (BFGS) algorithm, see Nocedal and Wright (2006).

2.4.4 Formulation of the FAANN model

The unique properties of ANN models motivated us to augment the factors to the ANN models to produce a more accurate forecast. The ANN models properties include; the relationships between input and output variables do not need to be specified in advance, since the method itself establishes these relationships through a training process. The ANN models do not require any assumptions on the underlying population distributions.

Time series forecasting research has demonstrated that the combined models improve forecasting performance substantially. For example see Khashei and Bijari (2010) and Zhang (2003). These combined models reduce the risk of failure compared to a single model where the underlying process cannot easily be determined or a single model may not be adequate to identify all the characteristics of the series.

In this chapter, we introduce the factor augmented artificial neural network (FAANN) model; the proposed model is a hybrid model of artificial neural network and factor model in order to produce more accurate forecasts. In the FAANN model the series is considered as nonlinear function of several past observations and the factors - that are extracted from large dataset that relate to the series under consideration – are as follows:

$$y_t = f[(y_{t-1}, y_{t-2}, \ldots, y_{t-p}), (F_1, F_2, F_3, F_4, F_5)] \tag{2.9}$$

where f is nonlinear functional form determined via ANN. In the first stage, the factor model is used to extract factors from a large related dataset. In the second stage, a neural network model is used to model the nonlinear and linear relationships existing in factors and original data. Thus,

$$y_{t+h} = \alpha_0 + \sum_{j=1}^{q} \alpha_j g\left(\beta_{0j} + \sum_{i=1}^{p} \beta_{ij}\, y_{t-i} + \sum_{i=p+1}^{p+5} \beta_{ij}\, F_{t,i}\right) + \varepsilon_{t+h} \tag{2.10}$$

As previously noted, the $\alpha_j (j = 0,1, \dots, q)$ and $\beta_{ij}(i = 0,1, \dots , p; j = 1,2, \dots , q)$ are the model parameters often called the connection weights; as we stated before p and q are the number of input and hidden nodes respectively, ε_t is the error term. Fig. 2.2 represents the FAANN model architecture.

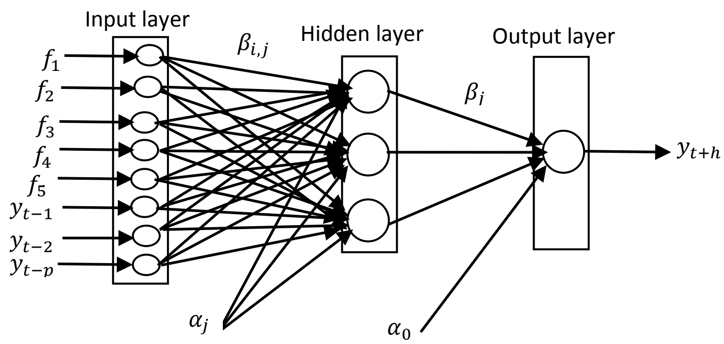

Figure 2.2: The FAANN model architecture ($\boldsymbol{N}^{(\boldsymbol{p}+\mathbf{5},\boldsymbol{q},\mathbf{1})}$)

2.5 Data

The AR benchmark model includes data on only the variable of interest, namely, deposit rate or share prices for gold mining or long term interest rate, while the FAANN and DFM models include 228 monthly time series of which 203 are from South Africa, covering the financial, real, nominal sectors and confidence indices, two global variables and 23 series of major trading partners and global financial markets. Thus besides the national variables, the chapter uses a set of global variables such as gold and crude oil prices. In addition the data also includes series from financial markets of major trading partners namely the United Kingdom, the United States, China and Japan. The in-sample period contains data that spans from 1992:01 to 2006:12, while the out-of-sample set spans from 2007:01 to 2011:12. The Augmented Dickey-Fuller (ADF) test is used to assess the degree of integration of all series. All non-stationary series are made stationary through differencing. The Schwarz information criterion (SIC) is used in selecting the appropriate lag length in such a way that no serial correlation is left in the stochastic error term. All series are standardized to have a mean of zero and a constant variance.

2.6 Evaluation of forecast accuracy

In this section, three data sets from South Africa - Deposit rate, Gold mining share prices and Long term interest rate - are used in order to demonstrate the in-sample and out-of-sample appropriateness and effectiveness of the proposed model.

2.6.1 In-sample forecast evaluation

In this subsection, we evaluate the in-sample predictive power of our proposed model and other fitted models. To do so, we estimate the forecasting models using the full sample form 1992:01 to 2011:12, giving a total of 240 observations, three data sets - Deposit rate, Gold mining share prices and Long term interest rate – in order to check the robustness of in-sample results of our proposed model and compare it to the other models. In-sample forecasting is most useful when it comes to examining the true relationship between the set of predictors and the future predictions of the variable of interest. Table 2.1 below reports the root mean square error RMSE[3] of the in-sample forecasting results. Our proposed FAANN model out performed all other models. The maximum reduction in RMSE over the AR benchmark model is around 24%, while the minimum reduction is around 14% considering all variables. Comparing the in-sample forecasting performance of our proposed model – FAANN – to the DFM, our proposed model yields lower RMSE with a reduction of between 9% and 19% for all variables. Note that the same factors are augmented to Autoregressive model to produce DFM and augmented to

[3] The RMSE statistic can be defined as $\sqrt{\frac{1}{N}\sum(Y_{t+n} - {}_t\hat{Y}_{t+n})^2}$, where Y_{t+n} denotes the actual value of a specific variable in period $t+n$ *and* ${}_t\hat{Y}_{t+n}$ is the forecast made in period t for $t+n$

artificial neural network to produce FAANN. Based on that, the in-sample forecasts results indicate significant differences between estimation methods which favour the nonlinear method over the linear one. This is potentially due to the property of the ANN models as universal approximators that can be applied to different time series to obtain accurate forecasts.

Table 2.1: The RMSE of the in-sample forecasts:

Variable	Model		
	FAANN	DFM	AR
Deposit rate	**0.1687**	0.1849	0.1954
Share prices for gold mining	**1.5922**	1.7782	1.8187
Long term interest rate	**0.1253**	0.1537	0.1640

Note: Bold entries indicate the forecasting model with the lowest RMSE

2.6.2 Out-of-sample forecast evaluation

In this subsection, we estimate the AR, DFM and our proposed FAANN model for the three variables of interest, namely, Deposit rate, Gold mining share prices and Long term interest rate, over the in-sample period 1992:01 to 2006:12 using monthly data, then compute out-of-sample for 3, 6 and 12 month-ahead forecasts for the period of 2007:01 to 2011:12. We re-estimate the models each month over the out-of-sample forecast horizon in order to update the estimate of the coefficients, before producing the 3, 6 and 12 month-ahead forecasts. We calculate the RMSE for the 60 three, 60 six and 60 twelve month-ahead forecasts for the three series across all of the different models in order to compare the forecast accuracy generated by the models. Note that the choice of the in-sample period, especially the starting date, depends on data availability of some important financial series. The out-of-sample period includes the occurrence of the financial crisis that affected economies and financial sectors in particular. None of the macroeconomics, financial or time series models predicted the crisis. Thus, we used this period as out-of-sample in order to demonstrate the appropriateness and effectiveness of our proposed model to produce accurate forecasts for such data that experienced a downturn or upturn during the financial crisis or the data that exhibits inherent nonlinearity. The result of each single variable can be summarized as follows:

- ***Deposit rate forecasting results***: in order to estimate the FAANN model. First, we use the MATLAB package to estimate the factors. Second, to obtain the optimum network architecture, based on the concepts of artificial neural networks design and using Broyden, Fletcher, Goldfarb and Shanno (BFGS) algorithm in R software. The best-fitted network is selected depending on the lowest in-sample RMSE, which is composed of eight inputs, five neurons in the hidden layer and one output (in abbreviated form $N^{(8-5-1)}$). Table 2.2 below reports the RMSEs of the 3, 6, and 12 month-ahead and the average of the 3, 6 and 12 month-ahead RMSEs. The benchmark for all forecast evaluations is the AR model forecast RMSEs. The FAANN model outperforms all other models for long and short horizons followed by the DFM, the RMSE of the FAANN model decrease as the forecasts horizon increase which in turn agreed with Greg and Sarah (1999) who found that the artificial neural network models significantly forecast better in long horizon. Compared to the AR benchmark model the FAANN performed better with large reductions in RMSE of around 25 percent to 46 percent of the RMSE of the AR, and the average RMSE reduction around 37 percent. Fig. 2.3 at the end of the chapter shows the out-of-sample forecast values for 3, 6 and 12 month horizons of the FAANN model.

Table 2.2: Out-of-sample (2007:01 – 2011:12) RMSE for Deposit rate

Model	3 month	6 month	12 month	Average
FAANN	**0.1390**	**0.1242**	**0.1240**	**0.1291**
DFM	0.1769	0.1784	0.2184	0.1912
AR	0.1862	0.1949	0.2314	0.2041

Note: Bold entries indicate the forecasting model with the lowest RMSE.

- ***Gold mining share prices***: to estimate the FAANN model, we used the same steps, software and algorism that were implemented to the previous variable. In order to obtain the optimum network, the best-fitted network is composed of eight inputs, seven neurons in the hidden layer and one output (in abbreviated form $N^{(8-7-1)}$). Table 2.3 below shows the performance measures of the FAANN, the DFM and the AR benchmark. The FAANN model stands out in forecasting both short and long horizons with a sizable reduction in RMSE relative to the AR benchmark model of 10 percent to 18 percent, the average of the RMSE reduction over the forecast horizons is 13 percent. Fig. 2.4 at the end of the chapter shows the plot of the FAANN model forecast values.

Table 2.3: Out-of-sample (2007:01 – 2011:12) RMSE for Gold mining share prices

Model	3 month	6 month	12 month	Average
FAANN	**1.6062**	**1.6349**	**1.4963**	**1.5791**
DFM	1.7131	1.7316	1.7335	1.7261
AR	1.7743	1.7924	1.8187	1.7951

Note: Bold entries indicate the forecasting model with the lowest RMSE.

- ***Long-term interest rate***: in order to estimate the FAANN model we implemented the same steps, software and algorism that we used with previous variables. Here the best-fitted network is composed of eight inputs, three neurons in the hidden layer and one output (in abbreviated form $N^{(8-3-1)}$). Table 2.4 shows the ability of the FAANN model to produce accurate forecasts that outperform other models. The FAANN model outperforms the benchmark model and the DFM on both the single level forecast horizons and the average of these horizons for each variable. The maximum reduction in the RMSE of the FAANN model compared to the AR benchmark is 45 percent, while the minimum reduction is 27 percent, the average RMSE reduction is around 38 percent. Fig. 2.5 at the end of the chapter shows the FAANN model forecast values for 3, 6 and 12 out-of-sample forecast horizons.

Table 2.4: Out-of-sample (2007:01 – 2011:12) RMSE for Long-term interest rate

Model	3 month	6 month	12 month	Average
FAANN	**0.1494**	**0.1295**	**0.1269**	**0.1353**
DFM	0.2018	0.1935	0.2212	0.2055
AR	0.2052	0.2140	0.2308	0.2167

Note: Bold entries indicate the forecasting model with the lowest RMSE.

2.6.3 Comparison of linear and nonlinear factor augmented models

Here we compare the forecasting performance of the factor augmented autoregressive model – the DFM – and factor augmented artificial neural network model – the FAANN – Table 2.5 represents the RMSE ratios of the FAANN model to the DFM over the out-of-sample period. The results indicate that the FAANN model generates accurate forecasts compared to the DFM for all variables and with all forecast horizons. The FAANN model provides improvement between 6 percent and 43 percent over the DFM for all variables and all horizons. Thus, these results indicate the superiority of nonlinear factor augmented (FAANN) over the linear factor augmented (DFM) across three different series and three different time horizons.

Table 2.5: Out-of-sample (2007:01 – 2011:12) RMSE ratio

Variables	Forecasting horizon		
	3 month	6 month	12 month
Deposit rate	0,7858	0,6962	0,5678
Gold mining share prices	0,9376	0,9442	0,0915
Long-term interest rate	0,7403	0,6693	0,5737

Note: The entries represent the ratio of the RMSE of the FAANN model to the RMSE of the DFM models.

To confirm the RMSEs results, the test of equal forecast accuracy of Diebold and Mariano (1995) is used to evaluate forecasts. The test of equal forecast accuracy employed here is given by $S = \frac{\overline{\mathrm{d}}}{\sqrt{\widehat{\mathrm{V}}(\overline{\mathrm{d}})}}$, where $\overline{\mathrm{d}} = \frac{1}{\mathrm{T}} \sum_{\mathrm{t}=1}^{\mathrm{T}} (\mathrm{e}_{1\mathrm{t}}^2 - \mathrm{e}_{2\mathrm{t}}^2)$ is the mean difference of the squared prediction error, and $\widehat{\mathrm{V}}(\overline{\mathrm{d}})$ is the estimated variance. Here $\mathrm{e}_{1\mathrm{t}}^2$ denotes the forecast errors from the FAANN model and $\mathrm{e}_{2\mathrm{t}}^2$ denotes the forecast errors from the AR benchmark model or the DFM. The S statistic follows a standard normal distribution asymptotically. Note, a negative and significant value of S indicates that the FAANN model outperforms the other model in out-of-sample forecasting. Table 2.6 provides the result of the Diebold - Mariano test between the FAANN and the AR benchmark and between the FAANN and the DFM. The test results confirm that the FAANN models provide lowest RMSEs. In summary the FAANN models provide significantly better forecasts at the 1 percent and 5 percent level compare to the AR and the DFM models.

Table 2.6: Diebold – Mariano test (2007:01 – 2011:12)

Model	Forecasting Horizons		
	3 month	6 month	12 month
Deposit Rate			
FAANN vs. AR	-2.095**	-2.108**	-3.159**
FAANN vs. DFM	-1.944**	-2.799**	-3.064**
Share Prices for Gold Mining			
FAANN vs. AR	-2.420**	-2.527**	-2.753**
FAANN vs. DFM	-2.281**	-2.337**	-2.602**
Long Term Interest Rate			
FAANN vs. AR	-2.402**	-2.339**	-2.429**
FAANN vs. DFM	-2.341**	-3.277***	-3.622***

Note: *** and ** indicate significant at the 1% and 5% levels respectively.

2.7 Conclusion

This chapter introduced the Factor Augmented Artificial Neural Network (FAANN) model which merged the factors that were extracted from a large data set – 288 series in our case – with ANN. Using the period of 1992:01 to 2006:12 as in-sample period and 2007:01 to 2011:12 as out-of-sample period, we compared the forecast performance of the FAANN with DFM and AR benchmark model for three, six and twelve month-ahead forecast horizons for three variables, namely, Deposit rate, Gold mining share prices and Long term interest rate. The study has provided evidence, using both the RMSE and Diebold-Mariano test as the comparison criteria, that FAANN models best fits the three considered variables over the 3, 6 and 12 month-ahead forecast horizons.

Tables 2.2 to 2.4 show the ability of the FAANN model to produce accurate forecasts that outperform other models, the results which seem not contradicted with in-sample model forecast performance as in Table 2.1. The FAANN model outperformed the AR benchmark model with large reduction in RMSE of around 11percent to 46 percent considering all variables and forecast horizons. Compared to the DFM the FAANN model produced more accurate forecasts that yielded a decrease in RMSE of around 6 percent to 43 percent. We attribute the superiority of the FAANN to the flexibility of the ANN to account for potentially complex nonlinear relationships that are not easily captured by linear models. On other hand, the DFM outperformed the AR benchmark with a reduction in the RMSEs of around 2 percent to 10 percent for all variables and across all forecast horizons. These results indicate that the factor augmented with linear or nonlinear models produced forecasts that are better than the AR forecasts. In other words, the models that use large data set of economic and financial variables improve the forecasting performance over models that do not use such data. We also observed that the FAANN residual decrease as the forecast horizon increase.

Further research can evaluate the FAANN forecasting performance in small and large simulated samples and compare it to FAVAR model.

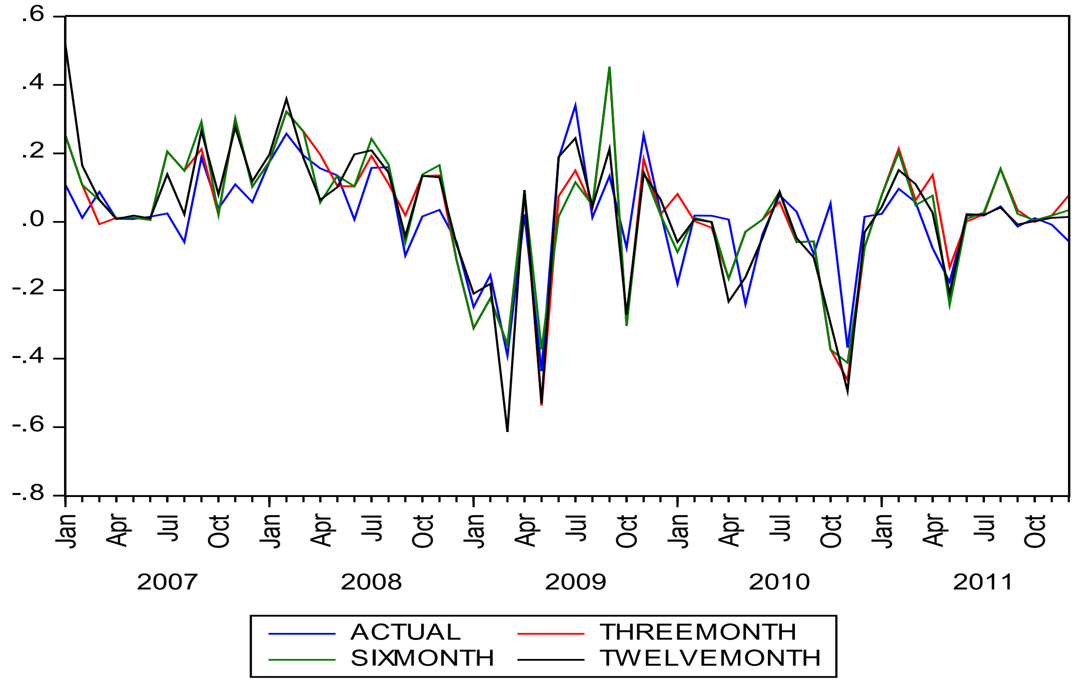

Figure 2.3: FAANN out-of-sample forecast results for three, six and twelve month-ahead: Deposit rate

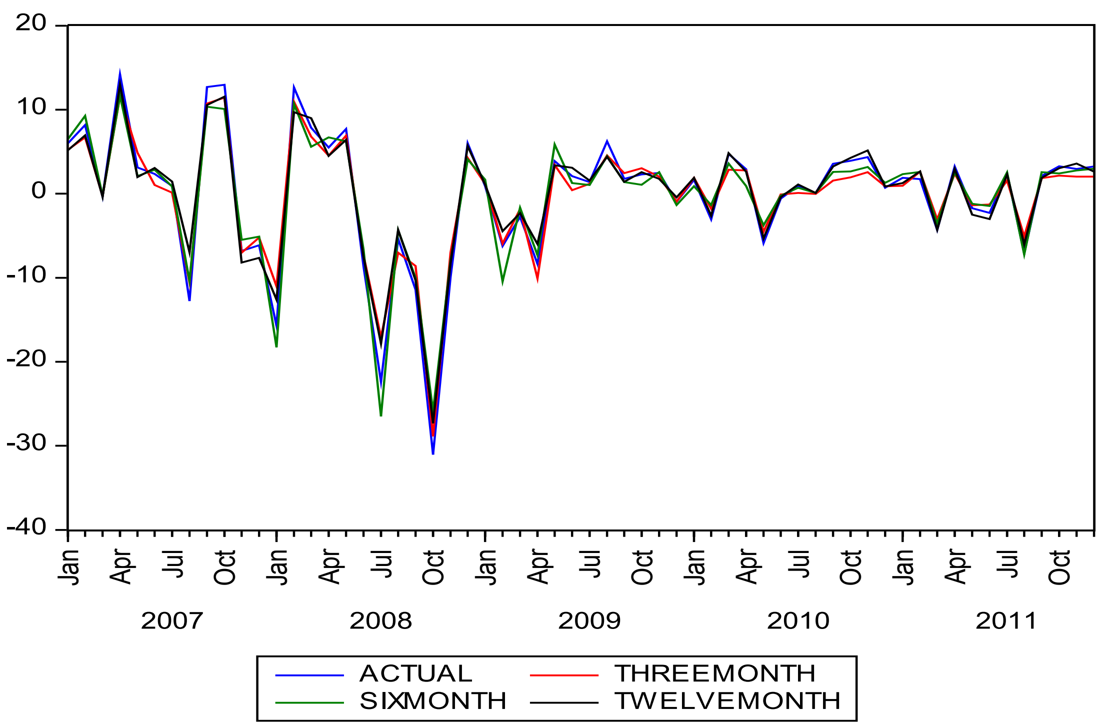

Figure 2.4: FAANN out-of-sample forecast results for three, six and twelve month-ahead: Gold mining share prices

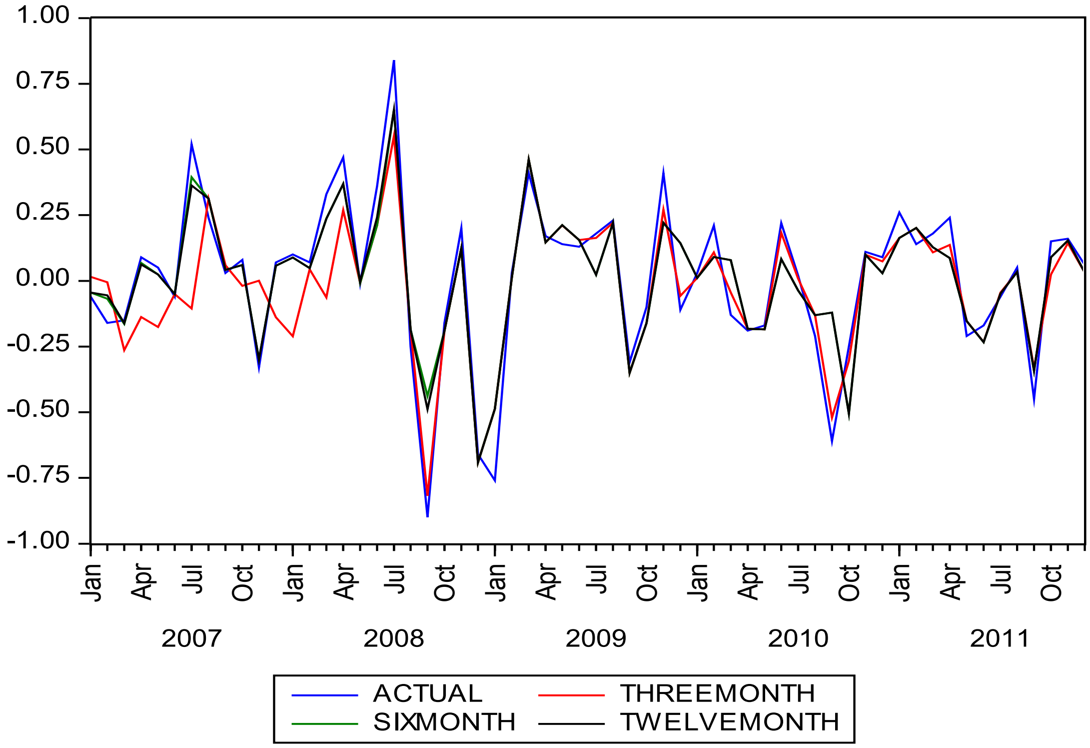

Figure 2.5: FAANN out-of-sample forecast results for three, six and twelve month-ahead: Long term interest rate

Chapter 3

Artificial Neural Networks – Dynamic Factor Model (ANN-DFM)*

3.1 Abstract

In this chapter we introduce a new model that uses the dynamic factor model (DFM) framework combined with artificial neural network (ANN) analysis, which accommodates a large cross-section of financial and macroeconomic time series for forecasting. In our new ANN-DF model we use the factor model to extract factors from ANNs in sample forecasts for each single series of the data set which contains 228 monthly series. These factors are then used as explanatory variables in order to produce more accurate forecasts. We apply this new model to forecast three South Africa variables, namely, Rate on three-month trade financing, Lending rate and Short term interest rate in the period 1992:01 to 2011:12. The model comparison results, based on the root mean square errors of three, six and twelve months ahead out-of-sample forecasts over the period 2007:01 to 2011:12 indicate that, in all of the cases, the ANN-DFM and the DFM statistically outperforms the autoregressive (AR) models. In the majority of the cases the ANN-DFM outperforms the DFM. The results indicate the usefulness of the factors in forecasting performance. The RMSE results are confirmed by the test of equality of forecast accuracy that is proposed by Diebold and Mariano.

---

* Ali Babikir and Henry Mwambi (in review). Artificial Neural Networks – Dynamic Factor Model (ANN-DFM). *Journal of Economics and Statistics*.

Keywords: Artificial neural network; Dynamic factor model; Forecast accuracy; Root mean square error.

3.2 Introduction

The need for obtaining accurate forecasts has been a strong motivation for empirical research. Efforts have been put in place in order to develop various kinds of forecasting models. Recently Dynamic Factor Models have become one of the standard econometric tools for forecasting purposes, and are increasingly being applied by forecasters, policymakers and research institutions to forecast key variables. This is partly because many time series are nowadays readily available, and modern computers and software allow us to efficiently summarize the information contained in large datasets. The use of dynamic factor models has been further improved by recent advances in estimation techniques proposed by Stock and Watson (2002a), Forni et al. (2005) as well as Kapetanios and Marcellino (2009). The two former approaches rely on static and dynamic principal component analysis (PCA) respectively and the latter one a subspace algorithm.

The goal of the techniques is to allow forecasters to easily summarize the information contained in large datasets and extract a few common factors that are useful in forecasting exercises. The small numbers of the estimated factors are then entered into simple regression models to forecast variables. Normally, exploiting information from large datasets helps to improve forecasts, and the results reported a good forecasting performance of the factor models; see among others, Angelini et al. (2011), Bańbura and Rünstler (2011), Das et al. (2011), Banerjee et al. (2005), Schumacher (2008) as well as Schumacher and Dreger (2004). The enhanced performance of the factor models has motivated researchers to extend the factor model and augment factors to other models. Bernanke et al. (2005) propose a forecasting model which they called the factor-augmented vector autoregressive (FAVAR) model, a model which merges a factor model with a vector autoregressive component. A factor-augmented vector autoregressive moving average (VARMA) model was recently suggested by Dufour and Pelletier (2013). Banerjee and Marcellino (2008) introduced the Factor augmented Error Correction Model (FECM). The FECM combines error-correction, cointegration and dynamic factor models. In another project we introduced a new model where the factors are augmented to Artificial Neural Networks (ANNs); the forecasting performance of the model has been assessed relative to the DFM and AR model. The results show that the new model – FAANN – outperforms the alternatives.

Recently artificial neural networks (ANNs) have been successfully applied to various areas including forecasting [Perez (2006) and Önder et al. (2013)], data mining [Pal (2002) as well as Craven and Shavlik (1997)] and pattern recognition [Bishop (1995)] smoothing the data or parameters [Yang and Wu (2012), Ferrari and Stengel (2005), Hill and Goring (1998) as well as Moon and Janowski (1995)]. The good results of the ANNs in all above-mentioned areas are based on the unique properties and features of the method which includes; first, ANNs are universal functional approximators, which can approximate any continuous function to any desired accuracy. Second, ANNs are data-driven self-adaptive methods in that there are few a priori assumptions to be made about the models for problems under study, thus ANN modeling is different from traditional model-based methods. Third, an ANN model is by design nonlinear in contrast to traditional time series forecasting models which assume linearity of the series under consideration. The real world is highly complex and changing; the recent financial crisis showed a clear evidence of the downturns of the economies and financial markets around the world. Beside such phenomena, there exist some linear and nonlinear patterns in the financial and economic time series. Thus, it is not sufficient to use only a linear or nonlinear model for time series because the linear/nonlinear model might miss some features of time series data. At the same time, smoothing or using high accurate approximators to the data that are affected by financial crisis and the uncertainty in the economy and financial sector can lead to more accurate forecasts.

The main purpose of this chapter is to use the dynamic factor framework for forecasting. We introduce a model that utilizes the generalized dynamic factor model proposed by Forni et al. (2000, 2003 and 2005). The factors of the new model are extracted from ANNs in sample forecasts to each single series of the dataset that contains 228[4] series. In this stage the ANNs are used as smoother or approximators as the method has proved in this regards. The chapter is organized as follows. Section 3.2 gives a brief description of the generalized dynamic factor model and how to determine the number of factors. In Sections 3.3, we briefly review the DFM, ANN and introduce the ANN-DFM modeling approaches to time series forecasting. The data set is described in Section 3.4. In Section 3.5 the forecasting performance of the factor models is assessed. The first subsection gives an overview of the in-sample forecasting performance. The next subsection presents the performance of the out-of-sample forecasting exercise. Section 3.6 draws some conclusions.

3.3 Methodology

In this section, the basic concepts of the estimation of factors and determination of the number of factors are briefly reviewed.

[4] Full details of the dataset provided in Section 3.5.

3.3.1 Estimation of the Factors

Let the panel of observations X_t be the N stationary time series variables with observations at times t = 1,......., T, where it is assumed that the series have zero mean. The idea behind the factor model is that most of the variance of the data set can be explained by a small number $q \ll N$ of factors contained in the vector f_t. In general the dynamic factor model representation is given by

$$X_t = \chi_t + \xi_t = \lambda(L)' f_t + \xi_t \tag{3.1}$$

where χ_t are the common components driven by factors f_t, and ξ_t are idiosyncratic components for each of the variables. In particular ξ_t is that part of X_t that cannot be explained by the common components. The common component is a function of the $q \times 1$ vector of dynamic factors which are common to all variables in the set $f_t = (f_{1t} \ldots\ldots f_{qt})'$, the operator $\lambda(L) = 1 + \lambda_1 L + \cdots + \lambda_s L^s$ is a lag polynomial with positive powers on the lag operator L with $Lf_t = f_{t-1}$. This way the lags of the factors are allowed to affect the current movement of the variables. The model can be re-written in static representation as;

$$X_t = \Lambda' F_t + \xi_t \tag{3.2}$$

where F_t is a vector of $r \geq q$ static factors that comprise of the dynamic factors f_t and all lags of the factors. Basically there are three methods of estimating the factors in F_t from a large data set. These methods were developed by Stock and Watson (2002a; hereafter SW), Kapetanios and Marcellino (2009) and Forni, Hallin, Lippi and Reichlin (2005, hereafter FHLR)[5]. In the current

5 For further technical details on this type of factor models, see Schumacher (2007).

chapter we employ the estimation method developed by FHLR. Below, we give a brief description of SW and FHLR methods and how they differ.

First we start with the SW model where the authors proposed estimating F_t with static principal component analysis (PCA) applied to X_t . The factor estimates are simply the first r principal components of X_t which according to SW are $F_t = \widehat{\Lambda}' X_t$, where $\widehat{\Lambda}$ is the $N \times r$ matrix of the eigenvectors corresponding to the r largest eigenvalues of the sample covariance matrix $\hat{\Sigma}$.

On the other hand, FHLR propose a weighted version of the principal components estimator suggested by SW, where the series are weighted according to their signal-to-noise ratio, which is estimated in the frequency domain. The estimation of common and idiosyncratic components is conducted using two steps. First, the covariance matrices of the common and idiosyncratic components of X_t are estimated via dynamic PCA. This involves estimating the spectral density matrix of X_t, $\Sigma(\omega)$, which has rank q. For each frequency ω, the largest q eigenvalues and the corresponding eigenvectors of $\Sigma(\omega)$ are computed, and the spectral density matrix of the common components $\Sigma_\chi(\omega)$ is estimated. Then it follows that the spectral density matrix of the idiosyncratic components is given by $\hat{\Sigma}_\xi(\omega) = \hat{\Sigma}(\omega) - \hat{\Sigma}_\chi(\omega)$. Inverse Fourier transform provides the time-domain autocovariances of the common and the idiosyncratic components given by $\hat{\Gamma}_\chi(k)$ and $\hat{\Gamma}_\xi(k)$ for lag k. Since dynamic PCA corresponds to a two-sided filter of the time series, this approach alone is not suited for forecasting. Second, a search is undertaken for the r linear combinations of X_t that maximize the contemporaneous covariance explained by the common factors $\hat{Z}_i{}' \hat{\Gamma}_\chi(0) \hat{Z}_i$, $i = 1, \dots, r$. This optimization problem is subject to the normalization $\hat{Z}_i' \hat{\Gamma}_\xi(0) \hat{Z}_j = 1$ for $i = j$ and zero otherwise. This representation can be reformulated as the generalized eigenvalue problem such that $\hat{\Gamma}_\chi(0) \hat{Z}_i = \hat{\mu}_i \hat{\Gamma}_\xi(0) \hat{Z}_i$, where $\hat{\mu}_i$ denotes the *i*-th generalized eigenvalue and $\hat{Z}_i$ its $N \times 1$ corresponding eigenvector in their non-null spaces. The factor estimates according to FHLR are then obtained as $\hat{F}_t = \hat{\mathrm{Z}}'\mathrm{X}_\mathrm{t}$ with $\hat{\mathrm{Z}} = [\hat{\mathrm{Z}}_1 \dots \hat{\mathrm{Z}}_\mathrm{r}]$.

3.3.2 Determination of the number of factors

Recently the determination of the number of the factors has been developed for both the case of the static factor model [Bai and Ng (2002) and Alessi et al. (2008)] and the dynamic factor model [Bai and Ng (2007); Amengual and Watson, (2007); Hallin and Liska (2007) and Onatski (2009, 2010)]. To specify the number of static factors, Bai and Ng (2002) and Alessi et al. (2008) use information criterion, based on AIC and BIC, to help guide the selection of the optimal number of factors r in a large data set. We apply the Bai and Ng (2002) approach which proposes five static factors. Onatski (2009) developed a statistical test to test and determine the number of dynamic factors under the null hypothesis that the number of factors is equal to k_0 against the alternative $k_1 > k_0$ (for details see Onatski (2009)).

3.4 Forecasting models

In this section, the basic concepts and modeling approaches of the dynamic factor model (DFM), autoregressive model (AR) and artificial neural networks (ANNs) models for time series forecasting are presented. The section also introduces the formulation of the proposed model.

3.4.1 Dynamic Factor model forecast

Based on subsection 3.2 five static factors and two dynamic factors are extracted from the entire data panel which both explain more than 87 percent of variation of the data. Then these estimated factors will be used to forecast the variables of interest. The forecasting model is specified and estimated as a linear projection of an h-step ahead transformed variable y_{t+h}^{h} into t-dated dynamic factors. The forecasting model follows the setup in Stock and Watson (2002a) and Froni et al. (2003) which takes the form:

$$y_{t+h}^{h} = \beta(L)\hat{F}_t + \gamma(L)y_t + u_{t+h}^{h} \tag{3.3}$$

where $\hat{F}_t$ are dynamic factors estimated using the method by Stock and Watson (2002b) while $\beta(L)$ and $\gamma(L)$ are the lag polynomials, which are determined by the Schwarz Information Criterion (SIC). Note that u_{t+h}^{h} is an error term. The coefficient matrix for factors and autoregressive terms are estimated by ordinary least square (OLS) for each forecasting horizon h.

3.4.2 Autoregressive (AR) Forecast

The AR model is given by

$$y_t = \phi + \gamma(L)y_t + e_t \tag{3.4}$$

where y_t is the variable to forecast, ϕ is a constant, $\gamma(L)$ is the iteratively estimated lag polynomial, the lag order is chosen using SIC and $\mathrm{e_t}$ is the error term.

The h-step ahead forecast from this model is

$$y_{t+h}^{h} = \phi + \gamma^{h}(L)\, y_t + e_{t+h}^{h} \tag{3.5}$$

where y_{t+h}^{h} is the h-step ahead forecast[6] of y_t, $\gamma^{h}(L)$ are the iteratively estimated lag polynomials, e_{t+h}^{h} is the h-month ahead forecast error term.

The benchmark AR forecast applied individually to our variables of interest, namely, the Lending rate, Rate on 3-month trade financing and short term interest rate. The optimal lag length is chosen by SIC.

[6] In this paper we choose iterated forecast instead of direct forecast. Marcellino et al. (2006) found that iterated forecast using AIC lag length selection performed better than direct forecasts, especially when forecast horizon increases. They argued that iterated forecast models with lag length selected based on information criterion are good estimates to the best linear predictor.

3.4.3 The Artificial Neural Network (ANN)

A neural network model can be described as a type of multiple regression in that it accepts inputs and processes them to predict some output. ANN can offer a valid approximation to the generating mechanism of a vast class of non-linear processes see for example, Hornik et al. (1989), Swanson and White (1997) and Omidi et al. (2011) for their use as forecasting tools. There are a number of properties that make the ANN model an attractive alternative to traditional forecasting models[7]. Most importantly ANN models control or are resistant to the limitations of traditional forecasting methods, including misspecification, biased outliers and assumption of linearity; Hill et al. (1996). The most significant advantage of ANN models over other classes of nonlinear models is that ANNs are universal approximators that can approximate a large class of functions with a high degree of accuracy; see Chen et al. (2003) and Zhang and Min Qi (2005). The network used in this chapter is a single hidden layer feed-forward network with n nodes in the hidden layer and linear jump connection or linear neuron activation function (see Fig 1) specified as follows:

$$y_{t+h} = \alpha_0 + \sum_{j=1}^{n} w_j\, g\left(\alpha_{0,j} + \sum_{i=1}^{p} w_{i,j}\, y_{t-i}\right) + \sum_{i=1}^{p} \beta_i\, y_{t-i} + \varepsilon_{t+h} \quad (3.6)$$

where inputs y_{t-i} represent the lagged values of the variable of interest and the output y_{t+h} is the variable being forecast, h indicates the forecast horizon, where $w_{i,j}(i = 1,2, \ldots, p, j = 1,2, \ldots, n)$ and $w_j(j = 1,2, \ldots, n)$ are the weights that connect the inputs to the hidden layer and the hidden layer to output respectively, α_0 is the bias. The function g is a logistic function given by $g(x) = \frac{1}{1 + e^{-x}}$. The ε_{t+h} is an error term. The third summation in Equation (3.6) shows the jump connection or skip-layer network that directly links the inputs y_{t-i} to the

[7] For more details about the strengths and drawbacks of ANN, see Ramlall (2010).

output y_{t+h} through β coefficients. The most important feature about this model is the combination of the pure linear model and feed-forward neural network. Therefore, if the relationship between inputs and output is pure linear, then only the skip-layer given by coefficient set β should be significant, and if the relationship is nonlinear one expects the coefficients set w and α to be highly significant, while the jump connections coefficient β will be relatively insignificant. Finally however, if the underlying relationship between input and output is mixed, then we expect all types of coefficient sets to be significant. The model is estimated by recursive least square using the Broyden, Fletcher, Goldfarb and Shanno (BFGS) algorithm; see Nocedal and Wright (2006). The selection of the lag lengths and the number of nodes in the hidden layer are chosen on the basis of the training set or the in-sample RMSE.

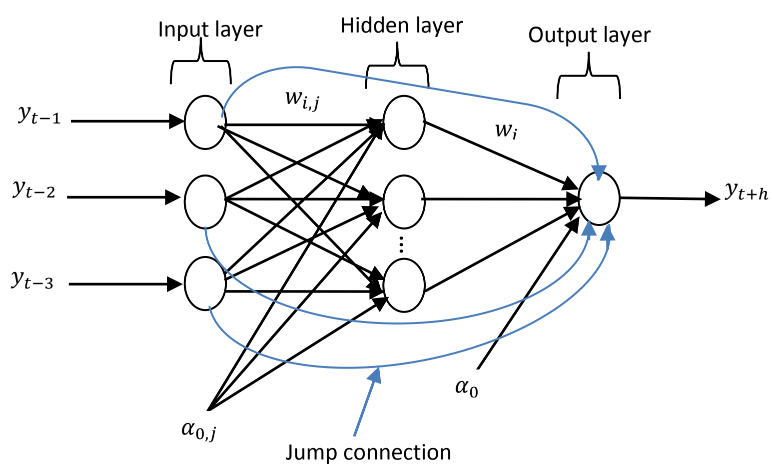

Figure 3.1: Structure of the network, $N^{(3,3,1)}$

3.4.4 Formulation of the ANN-DFM model

There are numerous time series models available but still the accuracy of time series forecasting currently is fundamental to many decision makers. Many researchers in time series forecasting have argued that predictive performance improves in combined models as these combinations reduce the risk failure that can occur as a result of difficulty to determine the pattern of the data under consideration. Our proposed model is constructed by a three-step process. In the first step, a nonlinear ANN model is estimated for each single series of our dataset. An in-sample forecast is obtained from the best fit of the model for each single series using the previous subsection. In the second step, a factor model is used to extract common components between the new estimated dataset obtained in the previous step which is then used to forecast the variable of interest. Note; in this step, three static factors and two dynamic factors are extracted - based on the subsection 3.2 – which explains more than 82 percent of variation of the entire dataset. In the third step, the extracted factors are used as explanatory variables in Equation (3.3) to produce the forecasts of the variable of interest.

3.5 Data

Our data set consists of 228 monthly time series from January 1992 to December 2011, or 239 observations for each variable. Among these 228 series are 203 series from South Africa, covering the financial, real, nominal sectors and confidence indices, two global variables and 23 series of major trading partners and global financial markets. Thus besides the national variables, the chapter uses a set of global variables such as gold and crude oil prices. In addition the data also includes series from financial markets of major trading partners namely the United Kingdom, the United States, China and Japan. We divide our sample into an estimation subsample and a subsample reserved for out-of-sample forecasting. The estimation period is from January 1992 through December 2006, and the forecasting period is from January 2007 through December 2011. We calculate forecasts for 3, 6 and 12 month forecasting horizons for three variables - Short term interest rate, Lending rate and Rate on 3-month trade financing. The

Augmented Dickey-Fuller (ADF) test is used to assess the degree of integration of all series. All non-stationary series are made stationary through differencing. The Schwarz information criterion (SIC) is used in selecting the appropriate lag length in such a way that no serial correlation is left in the stochastic error term. All series are standardized to have a mean of zero and a constant variance.

3.6 Empirical Results

3.6.1 In-sample results

To evaluate the accuracy of forecasts generated by the dynamic factor (DF) driven model, first we investigate the in-sample predictive power of the fitted models. We estimate the forecasting models using the full sample in order to check the robustness of our in-sample results. In-sample forecasting is most useful when it comes to examining the true relationship between the set of predictors and the future predictions of the variable of interest. Table 3.1 below reports the in-sample forecasting results. The first row reports the RMSE[8] for the AR benchmark model, while the remaining rows present the ratio of the RMSE of the model of interest to the RMSE of the AR benchmark model. The model with a lowest RMSE ratio is deemed to perform better than the other models. In our case the ANN-DFM out performed all other models and for all variables followed by the DFM. This result proves the superiority of DF driven models; the reason is potentially because the DF models can efficiently handle large amounts of information that include external variables that influence the South African economic and financial sector and therefore help improve the forecasting performance.

[8] The RMSE statistic can be defined as $\sqrt{\frac{1}{N}\sum(Y_{t+n} - {}_t\hat{Y}_{t+n})^2}$, where Y_{t+n} denotes the actual value of a specific variable in period $t+n$ and ${}_t\hat{Y}_{t+n}$ is the forecast made in period t for $t+n$.

Table 3.1: The RMSE of the in-sample forecasts:

Model	Rate on 3-month trade financing	Lending rate	Short term interest rate
AR	0.0394	0.5362	0.4031
ANN-DFM	**0.7949**	**0.8153**	**0.7618**
DFM	0.9231	0.9216	0.8710

Note: the first row reports the RMSE for the AR benchmark model; the remaining rows represent the ratio of the RMSE for the forecasting model to the RMSE for the AR. Bold entries indicate the forecasting model with the lowest RMSE.

3.6.2 Out-of-sample results

In this subsection we evaluate the accuracy of the forecasts generated by the ANN-DFM, and we compare its performance with the DFM and the AR benchmark model using the RMSE. We compare each of the three, six and twelve month-ahead forecasts generated by the ANN-DFM with DFM and AR over the out-of-sample horizon of 2007:01 to 2011:12. Note that the out-of-sample period includes the financial crisis which affected the South African economy in 2009. Thus, a good forecasting model can be used as an alternative to predict such crisis. Table 3.2 below reports the RMSE statistics for the AR benchmark model in the last row and the ratio of the RMSE of other models to the RMSE for the AR benchmark model. The results from the AR benchmark models show that for most cases the RMSE increases as the horizon increases. These results indicate that more accurate forecasts under the AR model are available at shorter horizons. Note; in this chapter we choose iterated forecast instead of direct forecast, on the other hand, the forecasts constructed recursively, using all available data to estimate parameters. The main observations can be summarized as follows:

- ***Rate on 3-month trade financing***: from Table 3.2 we observe that the ANN-DFM model outperforms the AR benchmark model with an average reduction of 20 percent in the RMSE for all horizons. On the other hand, the DFM outperform the AR benchmark model with an average reduction of 10 percent in the RMSE for all horizons. These results show that the ANN-DFM outperform the DFM with an average reduction of 10 percent in the RMSE.
- ***Lending rate***: as with the previous variable the ANN-DFM is the standout performer for all horizons. Compared to the AR benchmark model the reduction in the RMSE is between 12 percent to 19 percent. The DFM also beat the AR benchmark model with a reduction in RMSE of around 6 percent to 11 percent.
- ***Short term interest rate***: from Table 3.2 we see that the DFM outperforms the other models. Comparing the DFM to the ANN-DFM we find that the DFM performs slightly better than the ANN-DFM; comparing both models to the AR benchmark model there is a reduction in the RMSE of around 14 percent to 22 percent. Here we observe that the DFM forecasting errors are a bit less than the ANN-DFM forecasting errors, taking into account that the DFM model used five factors while the ANN-DFM used three factors only.

When we consider the cross model test of forecast accuracy that was proposed by Diebold and Mariano (1995)[9], Table 3.3 shows that in all cases where ANN-DFM outperform the AR benchmark model the statistics are significant at least at 5 percent level except in one case where the statistic is significant at 10 percent level. Regarding the cases where ANN-DFM outperforms DFM the statistics are significant at the maximum of 5 percent level. The above is true for two series, namely, Rate on 3-month trade financing and Lending rate. On the other hand, when the

[9] The test is given by; $S = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}}$ where $\bar{d} = \frac{1}{T}\sum_{t=1}^{T}(e_{1t}^2 - e_{2t}^2)$ is the mean difference of the squared prediction error, and $\hat{V}(\bar{d})$ is the estimated variance. Here e_{1t}^2 denotes the forecast errors from the ANN-DFM model and e_{2t}^2 denotes the forecast errors from the AR benchmark model, the DFM and ANN. The S statistic follows a standard normal distribution asymptotically. Note, a negative and significant value of S indicate that the ANN-DFM model outperforms the other model in out-of-sample forecasting.

DFM tends to outperform the ANN-DFM the test statistics are insignificant at 1 percent, 5 percent and 10 percent levels.

Table 3.2: Out-of-sample (2007:01 – 2011:12) relative RMSE

Model	Rate on 3-month trade financing		
	3 month	6 month	12 month
ANN-DFM	**0.8037**	**0.8175**	**0.7995**
DFM	0.8560	0.9097	0.9349
AR	0.0382	0.0383	0.0387
	Lending Rate		
ANN-DFM	**0.8318**	**0.8114**	**0.8826**
DFM	0.9298	0.8887	0.9399
AR	0.3617	0.3881	0.3576
	Short term interest rate		
ANN-DFM	0.8588	0.7848	0.7807
DFM	**0.8576**	**0.7827**	**0.7805**
AR	0.3519	0.3852	0.3872

Note: The last row reports the RMSE for the AR benchmark model; the remaining rows represent the ratio of the RMSE for the forecasting model to the RMSE for the AR. Bold entries indicate the forecasting model with the lowest RMSE.

Table 3.3: Diebold – Mariano test (2007:01 – 2011:12)

Model	Forecasting Horizons		
	3 month	6 month	12 month
Rate on 3-Month Trade Financing			
ANN-DFM vs. DFM	-1.642*	-1.779*	-2.327**
ANN-DFM vs. AR	-2.029**	-2.124**	-2.121**
Lending Rate			
ANN-DFM vs. DFM	-1.923*	-1.675*	-1.726*
ANN-DFM vs. AR	-2.326**	-2.017**	-2.134**
Short Term Interest Rate			
ANN-DFM vs. DFM	1.569	1.514	1.151
ANN-DFM vs. AR	-1.947*	-2.501**	-2.661***

Note: ***, ** and * indicate significant at the 1%, 5% and 10% levels respectively.

3.7 Conclusion

This chapter introduces a new model where Artificial Neural Networks (ANNs) is used to generate in-sample fit to the dataset of 228 series, then the Dynamic Factor Model (DFM) is used to extract a small number of the factors that can be used as explanatory variables in order to produce the forecasts of the three variables of interest, namely Rate on 3-month trade financing, Lending rate and Short term interest rate, using monthly data over the period 1992:01 to 2011:12. The in-sample period contains data from 1992:01 to 2006:12, and the out-of-sample forecasts are based on three, six and twelve month-ahead forecasts over a 60 month forecasting horizon covering 2007:12 to 2011:12. The forecasting performance of the new model ANN-DFM is evaluated in terms of the RMSEs by comparing it to the DFMs and the AR benchmark model. Our results indicate that the new ANN-DFM outperforms the AR benchmark model for all variables and over all forecasting horizons. On the other hand, the new model outperforms the DFM in majority of the cases; however, when the DFM outperforms the ANN-DFM the improvement is very small. In general a data-rich factor driven model is best suited in forecasting the three variables when compared to the AR benchmark model. The Diebold and Mariano (1995) test for cross model forecast accuracy confirms the superiority of the factor driven model in general and the ANN-DFM in particular.

Chapter 4

A factor - artificial neural network model for time series forecasting*

4.1 Abstract

Artificial Neural Networks (ANNs) are flexible nonlinear models that can approximate virtually any function to any desired degree of accuracy. Theoretical and empirical results support the effectiveness of the integration of different models to improve forecast performance. In this chapter the factor models (FMs) are integrated with the ANN model to produce a new hybrid method which we refer to as the Factor Artificial Neural Network (FANN) to improve the time series forecasting performance of the artificial neural networks. The FMs in this chapter use 228 monthly series over the period from 1992:01 to 2011:12. The empirical results of the Root Mean Square Error (RMSE) for the in-sample and out-of-sample forecasts from 2007:01 to 2011:12 indicate that the proposed FANN model is an effective way to improve forecasting accuracy over the dynamic factor Model (DFM), the ANN and AR benchmark model. The results confirm the usefulness of the factors that were extracted from a large set of related variables when we compare the FANN and ANN models.

---

* Ali Babikir and Henry Mwambi (Accepted). A factor - artificial neural network model for time series forecasting. *IEEE Transactions on Neural Networks and Learning Systems*.

On the other hand, as far as estimation is concerned the nonlinear FANN model is more suitable to capture nonlinearity and structural breaks compared to linear models. The Diebold-Mariano test results confirm the superiority of the FANN model forecasts performance over the AR benchmark model and the ANN model forecasts.

Keywords: Artificial neural network; Dynamic factor model; Forecast accuracy; Root mean square error.

4.2 Introduction

In recent decades considerable progress in handling large panels of time series data in forecasting using factor models has been made. The initial contributions in this area were the work of Geweke (1977) and Sargent and Sims (1977), who introduced the dynamic factor approach to macroeconomics. They exploited the dynamic interrelationships between the variables, and then reduced the number of common factors even further. However, the approach followed by Geweke (1977) and Sargent and Sims (1977) is too restrictive, in that it assumes orthogonality on the idiosyncratic components, while Chamberlain (1983) and Chamberlain and Rothschild (1983) allow for the possibility of weakly cross-sectional correlation of the idiosyncratic components. In further improvements these large factor models have been improved by accounting for serial correlation and weakly cross-sectional correlation of idiosyncratic components, through advances in estimation techniques proposed by Forni et al. (2005), Kapetanios and Marcellino (2009) and Stock and Watson (2002b). This advance, in turn, has generated an increasing amount of interest in the usage of these models in academia, international organizations, central banks, and governments, simply because they can accommodate a large panel of time series when forecasting variables. However, there is still a considerable degree of divergence in opinion as to whether or not factor models with large cross-sections of time series tend to outperform traditional econometric models with limited numbers of variables. On the one hand, studies such as those of Cristadoro et al. (2005), Forni et al. (2001), Forni et al. (2005), Giannone and Matheson (2007), Gupta and Kabundi (2011), Schneider and Spitzer (2004), Stock and Watson (1989, 1991, 1999, 2002a,b) and van Nieuwenhuyze (2005) found evidence of improvements in the forecasting performances of macroeconomic and financial variables using factor analysis, while on the other hand Angelini et al. (2001), Gosselin and Tkacz (2001), Schumacher (2007) and Schumacher and Dreger (2004) found only minor or no improvements in forecasting ability. These conflicting results have led to attracting debate as to whether or not the victory claimed by the proponents of large models was precocious. Some attribute the success of large models to different circumstances pertaining to each study. For example Banerjee et al. (2005) find that small models forecast macroeconomic variables better than factor models. In addition, they also find that the performances of factor models differ between countries. Factor models are comparatively good at forecasting real variables in the US relative to in the euro area, while the euro area nominal variables are easier to predict than the US nominal variables using factor models. Furthermore, Boivin and Ng (2006) claim that the composition of the data set and the dimensions of the cross-section are important in producing better forecasts from factor models.

Based on the success of the dynamic factor model many linear extensions were introduced such as factor augmented vector autoregressive (FVAR) and factor augmented error correction model (FECM) and their Bayesian applications. Our factor model extension brings together the factor model and the nonlinear ANN model, the mixture that we believe can accommodate the structural breaks.

Against this backdrop, this chapter exploits the information contained in the large-dimensional factor model framework developed by Forni et al. (2005) (hereafter FHLR) for forecasting Johannesburg Stock Exchange (JSE) share prices and a measure of the short-term nominal interest rate (Treasury Bill Rate) for the South Africa, over the out-of-sample period from 2007:01 to 2011:12[10], with an in-sample estimation period from 1992:01 to 2006:12. The forecasting performances of the Factor Models (FMs) estimated under linear dynamic factor model (DFM) and nonlinear Factor – Artificial Neural Network (FANN) assumptions with regard to the interaction between the factors and the variables of interest are investigated. The FMs are evaluated and compared with the performances of two other alternative models, namely Autoregressive (AR) and Artificial Neural Network (ANN) models, on the basis of the Root Mean Squared Error (RMSE) of the out-of-sample forecasts.

In this chapter we introduce the FANN model, where we model the extracted factors using ANN nonlinear method to forecast the variables of interest. The nonlinear Factor-ANN results compare to the results of the DFM and ANN models. To the best of our knowledge this is the first attempt to use the FANN model to forecast variables in South Africa in particular.

The empirical results show sizable gains in terms of the forecasting ability of the FANN compared to both the standard ANN and the DFM. Thus the FANN represents an improvement with respect to the standard ANN and the DFM.

The remainder of the chapter is organized as follows: section 4.3 describes the FMs and ANN forecasting models; section 4.4 presents the data; and the results from forecasting models are discussed in section 4.5. Finally, we close with Conclusions in section 4.6.

[10] The choice of out of sample span comes from the aim to investigate the performance of forecasting models during the period of financial crisis.

4.3 The Models

This chapter uses the FM to extract common components from a large set of variables, after which these common components are used to forecast the variables of interest using the linear DFM and the nonlinear ANN methods.

4.3.1 Estimation of the Factors and the Dynamic Factor Model

Let the panel of observations X_t be the N stationary time series variables with observations at times t = 1,......., T, where it is assumed that the series have zero mean. The idea behind the factor model is that most of the variance of the data set can be explained by a small number $q \ll N$ of factors contained in the vector f_t. In general the dynamic factor model representation is given by

$$X_t = \chi_t + \xi_t = \lambda(L)' f_t + \xi_t \qquad (4.1)$$

where χ_t are the common components driven by factors f_t, and ξ_t are idiosyncratic components for each of the variables. In particular ξ_t is that part of X_t that cannot be explained by the common components. The common component is a function of the $q \times 1$ vectors of dynamic factors which are common to all variables in the set $f_t = (f_{1t} ..... f_{qt})'$, the operator $\lambda(L) = 1 + \lambda_1 L + \cdots + \lambda_s L^s$ is a lag polynomial with positive powers on the lag operator L with $Lf_t = f_{t-1}$. In this way the lags of the factors are allowed to affect the current movement of the variables. The model can be re-written in static representation as:

$$X_t = \Lambda' F_t + \xi_t \qquad (4.2)$$

where F_t is a vector of $r \geq q$ static factors that comprise of the dynamic factors f_t and all lags of the factors. Basically there are three methods of estimating the factors in F_t from a large data set.

These methods were developed by Stock and Watson (2002a; hereafter SW), Kapetanios and Marcellino (2009) and Forni, Hallin, Lippi and Reichlin (2005, hereafter FHLR)[11]. In the current chapter we employ the estimation method developed by FHLR. Below we give a brief description of SW and FHLR methods and how they differ.

First we start with the SW model where the authors proposed estimating $\mathrm{F_t}$ with static principal component analysis (PCA) applied to $\mathrm{X_t}$. The factor estimates are simply the first r principal components of $\mathrm{X_t}$ which according to SW are $\mathrm{F_t} = \widehat{\Lambda}'\mathrm{X_t}$, where $\widehat{\Lambda}$ is the $\mathrm{N \times r}$ matrix of the eigenvectors corresponding to the r largest eigenvalues of the sample covariance matrix $\hat{\Sigma}$.

On the other hand, FHLR propose a weighted version of the principal components estimator suggested by SW, where the series are weighted according to their signal-to-noise ratio, which is estimated in the frequency domain. The estimation of common and idiosyncratic components is conducted using two steps. First, the covariance matrices of the common and idiosyncratic components of $\mathrm{X_t}$ are estimated via dynamic PCA. This involves estimating the spectral density matrix of $\mathrm{X_t}$, $\Sigma(\omega)$, which has rank q. For each frequency ω, the largest q eigenvalues and the corresponding eigenvectors of $\Sigma(\omega)$ are computed, and the spectral density matrix of the common components $\sum_{\chi}(\omega)$ is estimated. Then it follows that the spectral density matrix of the idiosyncratic components is given by $\hat{\Sigma}_{\xi}(\omega) = \hat{\Sigma}(\omega) - \hat{\Sigma}_{\chi}(\omega)$. Inverse Fourier transform provides the time-domain autocovariances of the common and the idiosyncratic components given by $\hat{\Gamma}_{\chi}(\mathrm{k})$ and $\hat{\Gamma}_{\xi}(\mathrm{k})$ for lag k. Since dynamic PCA corresponds to a two-sided filter of the time series, this approach alone is not suited for forecasting. Second, a search is undertaken for the r linear combinations of $\mathrm{X_t}$ that maximize the contemporaneous covariance explained by the common factors $\hat{\mathrm{Z}}_{\mathrm{i}}'\hat{\Gamma}_{\chi}(0)\hat{\mathrm{Z}}_{\mathrm{i}}$, $i = 1, \dots, r$. This optimization problem is subject to the

[11] For further technical details on this type of factor models, see Schumacher (2007).

normalization $\hat{Z}_i' \hat{\Gamma}_\xi(0) \hat{Z}_j = 1$ for $i = j$ and zero otherwise. This representation can be reformulated as the generalized eigenvalue problem such that $\hat{\Gamma}_\chi(0)\hat{Z}_i = \hat{\mu}_i \hat{\Gamma}_\xi(0)\hat{Z}_i$, where $\hat{\mu}_i$ denotes the *i*-th generalized eigenvalue and $\hat{Z}_i$ its $N \times 1$ corresponding eigenvector in their non-null spaces. The factor estimates according to FHLR are then obtained as $\hat{F}_t = \hat{Z}' X_t$ with $\hat{Z} = [\hat{Z}_1 \ldots \hat{Z}_r]$.

4.3.2 Dynamic Factor model

The estimated factors will be used to forecast the variables of interest. The forecasting model is specified and estimated as a linear projection of an h-step ahead transformed variable y_{t+h} into *t*-dated dynamic factors. The forecasting model follows the setup in Stock and Watson (2002a) and Froni et al. (2003) which takes the form:

$$y_{t+h} = \beta(L)\hat{f}_t + \gamma(L)y_t + u_{t+h} \tag{4.3}$$

where $\hat{f}_t$ are dynamic factors estimated using the method by FHLR while $\beta(L)$ and $\gamma(L)$ are the lag polynomials, which are determined by the Akaike Information Criterion (AIC) and the Schwarz Information Criterion (SIC). The u_{t+h} is an error term. The coefficient matrix for factors and autoregressive terms are estimated by ordinary least square (OLS) for each forecasting horizon h. To generate the estimate and forecast of the Autoregressive (AR) benchmark we impose a restriction to Eq. (4.3), where we set $\beta(L) = 0$[12].

4.3.3 The Artificial Neural Network (ANN)

A neural network model can be described as a type of multiple regression in that it accepts inputs and processes them to predict some output. ANN can offer a valid approximation to the

[12]We use the autoregressive model as our benchmark.

generating mechanism of a vast class of non-linear processes; see for example, Hornik et al. (1989), Swanson and White (1997) and Omidi et al. (2011) for their use as forecasting tools. There are a number of properties that make the ANN model an attractive alternative to traditional forecasting models[13]. Most importantly ANN models control or are resistant to the limitations of traditional forecasting methods, including misspecification, biased outliers and assumption of linearity; Hill et al. (1996). The most significant advantage of the ANN models over other classes of nonlinear models is that ANNs are universal approximators that can approximate a large class of functions with a high degree of accuracy; see Chen et al. (2003) and Zhang and Min Qi (2005). The network used in this chapter is a single hidden layer feed-forward network with n nodes in the hidden layer and linear jump connection or linear neuron activation function (see Fig 4.1) specified as follows:

$$y_{t+h} = \alpha_0 + \sum_{j=1}^{n} w_j \, g\left(\alpha_j + \sum_{i=1}^{p} w_{i,j} \, y_{t-i}\right) + \sum_{i=1}^{p} \beta_i \, y_{t-i} + \varepsilon_{t+h} \qquad (4.4)$$

where inputs y_{t-i} represent the lagged values of the variable of interest and the output y_{t+h} is the variable being forecast, h indicates the forecast horizon, where $w_{i,j}(i = 1,2, \ldots, p, j = 1,2, \ldots, n)$ and $w_j(j = 1,2, \ldots, n)$are the weights that connect the inputs to the hidden layer and the hidden layer to output respectively, α_0 is the bias. The function g is a logistic function given by $g(x) = \frac{1}{1 + e^{-x}}$. The ε_{t+h} is an error term. The third summation in Equation (4.4) shows the jump connection or skip-layer network that directly links the inputs y_{t-i} to the output y_{t+h} through β coefficients. The most important feature about this model is the combination of the pure linear model and feed-forward neural network. Therefore, if the relationship between inputs and output is pure linear, then only the skip-layer given by coefficient set β should be significant, and if the relationship is nonlinear one expects the

[13] For more details about the strengths and drawbacks of ANN, see Ramlall (2010).

coefficients set w and α to be highly significant, while the jump connections coefficient β will be relatively insignificant. Finally however, if the underlying relationship between input and output is mixed, then we expect all types of coefficient sets to be significant. The model is estimated by recursive least square using the Broyden, Fletcher, Goldfarb and Shanno (BFGS) algorithm, see Nocedal and Wright (2006). The selection of the lag lengths and the number of nodes in the hidden layer are chosen on the basis of the training set or the in-sample RMSE, where n=5.

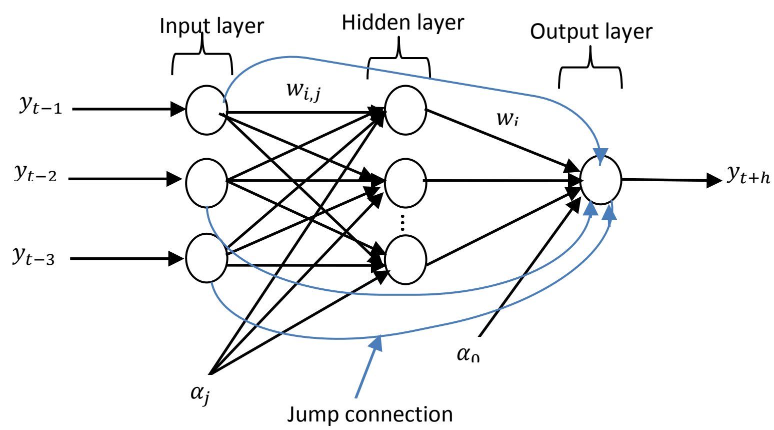

Figure 4.1: Structure of the best fitted network, $N^{(3,5,1)}$

4.3.4 Proposed Factor – Artificial Neural Network (FANN) model

Previous researchers argued that combined models improve the predictive performance of time series forecasting. The combined models reduce the risk of using an inappropriate model as the underlying process cannot easily be determined, thus the hybrid model can reduce these risk failure and obtain more accurate results. In this chapter we propose a hybrid model of artificial neural network and factor model in order to yield an enhanced predictive and forecast performance. The factor models (FM) extract components that are common between the 228 time series variables. The factor model expresses individual time series as the sum of two unobserved components, a common component which is driven by a small number of common factors, and an idiosyncratic component which is specific to each variable. The FM is able to extract a few factors that explain the co-movement of all variables. Our proposed model used the Forni et al. (2005) approach explained above to extract these factors at the first step.

In the second step, a neural network is used in order to model the nonlinear and linear relationships existing in the factors f_t and y_t the variable we need to forecast (see Fig 4.2), as follows:

$$y_{t+h} = \alpha_0 + \sum_{j=1}^{n} w_j \, g\big(\alpha_j + \sum_{i=1}^{p} w_{i,j} \, f_{t,i}\big) + \sum_{i=1}^{p} \beta_i \, f_{t,i} + \ \varepsilon_{t+h} \qquad (4.5)$$

where $w_{i,j}(i = 1,2, \dots, p, j = 1,2, \dots, n)$ and $w_j(j = 1,2, \dots, n)$ are the weights that connect the inputs to the hidden layer and the hidden layer to output respectively, p is the number of factors. In our application we arrive at $p = 5$ as determined by Bai and Ng (2002) approach and also supported by Onatski (2009) test, n is the number of nodes in the hidden layer, α_0 is the bias. The function g is a logistic function, where $g(u) = \frac{1}{1 + e^{-u}}$. The coefficients β represent the linear part of the equation (5) which directly links the inputs f_i to the output y_{t+h} . The ε_{t+h} is an error term. The number of nodes in the hidden layer are determined on the basis of the training set or in-sample RMSE, where *n*=3.

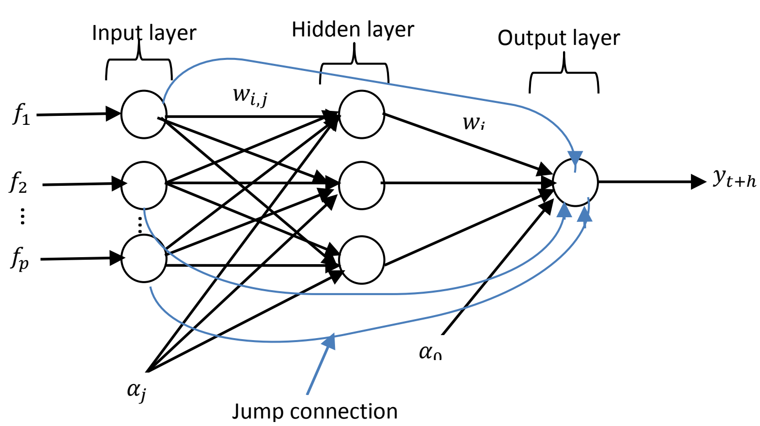

Figure 4.2: The best fitted network structure, $N^{(5,3,1)}$

4.5 Data and the number of factors

The data set contains 228 monthly series, 203 from South Africa covering the financial, real and nominal sectors, two global variables and 23 series of major trading partners and global financial markets. Thus besides the national variables, the chapter uses a set of global variables such as gold and crude oil prices. In addition the data also includes series from financial markets of major trading partners, namely the United Kingdom, the United States, China and Japan. The in-sample period contains data from 1992:01 to 2006:12, while the out-of-sample set spans from 2007:01 to 2011:12. The Augmented Dickey-Fuller (ADF) test is used to assess the degree of integration of all series. All non-stationary series are made stationary through differencing. The Schwarz information criterion (SIC) is used in selecting the appropriate lag length in such a way that no serial correlation is left in the stochastic error term. All series are standardized to have a mean of zero and a constant variance.

Recently the determination of the number of the factors has been developed for both the case of the static factor model [Bai and Ng, (2002) and Alessi et al., (2008)] and the dynamic factor model [Bai and Ng, (2007); Amengual and Watson, (2007); Hallin and Liska (2007) and Onatski (2009, 2010)]. To specify the number of static factors, Bai and Ng (2002) and Alessi et al. (2008) use information criterion, based on AIC and BIC, to help guide the selection of the optimal number of factors r in a large data set. We apply the Bai and Ng (2002) approach which proposes five static factors. Onatski (2009) developed a statistical test to test and determine the number of dynamic factors under the null hypothesis that the number of factors is equal to k_0 against the alternative $k_1 > k_0$ (for details see Onatski (2009)). In our case the test suggests two dynamic factors which both explain more than 87 percent of variation.

4.6 Results

4.6.1 In-sample results

In this subsection we evaluate the in-sample predictive power of the fitted models. We estimate the forecasting models using the full sample in order to check the robustness of our in-sample results. In-sample forecasting is most useful when it comes to examining the true relationship between the set of predictors and the future predictions of the variable of interest. Table 4.1 below reports the RMSE[14] of the in-sample forecasting results. The table reports the RMSE statistics for the AR benchmark model and the ratio of the RMSE for the other models to the RMSE for the AR benchmark model. Thus, the ratio that is higher than one indicates that the method under analysis is worse than the benchmark, so the model with a lowest RMSE ratio is deemed to perform better than the other models. Our proposed FANN model out performed all other models with a large reduction in RMSE relative to the AR benchmark model for both variables. The reason is potentially because we merge the factors that efficiently handle large amounts of information that include external variables that influence South African economy with ANN nonlinear estimation model. The ANN model also provides fairly better in-sample forecasts compared to the AR benchmark model and DFM model. In general the FANN and ANN nonlinear models perform much better than DFM and AR linear models.

[14] The RMSE statistic can be defined as $\sqrt{\frac{1}{N}\sum(Y_{t+n} -{}_t\hat{Y}_{t+n})^2}$, where Y_{t+n} denotes the actual value of a specific variable in period $t+n$ *and* ${}_t\hat{Y}_{t+n}$is the forecast made in period t for $t+n$

Table 4.1: The RMSE of the in sample forecasts:

Forecasting model	Treasury Bill Rate	JSE all Share prices
AR (benchmark model)	0.8860	0.9747
DFM	0.9369	0.9511
FANN	**0.6868**	**0.6536**
ANN	0.7731	0.8431

Note: the first row reports the RMSE for the AR benchmark model; the remaining rows represent the ratio of the RMSE for the forecasting model to the RMSE for the AR. Bold entries indicate the forecasting model with the lowest RMSE.

4.6.2 Out-of-Sample Forecasting Results

In this subsection, we evaluate out-of-sample forecasts of the Treasury Bill Rate and JSE all share prices over the period 2007:01 to 2011:12. This period includes the global financial crisis that impacted the South African economy at the end of 2008 and 2009. We consider short forecast horizon of 3 months and long forecast horizon of 12 months. Table 4.2 below reports the RMSE statistics for the AR benchmark model in the first row and the ratio of the RMSE of other models to the RMSE for the AR benchmark model. The result of the AR benchmark model shows that the RMSE increases as horizon increases, and indicates that more accurate forecasts for the AR are available at shorter horizons. Note; in this chapter we choose iterated forecast instead of direct forecast, on the other hand the forecasts constructed recursively using all available data to estimate parameters. The results of the two variables can be summarized as follows:

Treasury Bill Rate*:* the proposed FANN model outperforms all other models for short horizon producing the lowest RMSE followed by the AR benchmark model. For the long horizon, the

ANN outperforms all other models followed by the FANN model. The FANN result shows that the RMSE increases as the forecast horizon increases. Compared to the DFM[15] model the FANN model performs better, thus the estimation method used to model the factors is improtant.

***JSE all Share prices**:* the FANN model stands out in forecasting the JSE all share prices for both short and long horizons with a sizable reduction in RMSE relative to the AR benchmark model of 8 percent to 19 percent. The DFM outperforms the ANN and AR benchmark model, thus the derived factor models FANN and the DFM outperform univariate linear and nonlinear models AR and ANN respectively. These results clearly indicate the importance of the information contained in the common factors, which in turn are derived from 228 monthly series. The performance of the FANN model over the DFM model indicates the role of the estimation method that captures the nonlinearity associated to the variables of interest. Babikir et al. (2012) found evidence of structural breaks in the JSE share return index in the end of 2008 and mid-2009. These events are included in our out-of-sample period; thus it shows that the FANN model captures well the structural breaks compared to the DFM and the other models.

We attribute the forecast performance of derived factor models the FANN and the DFM for the JSE all Share prices over the Treasury Bill Rate to the data set used to extract the factors which contains more financial than macroeconomic variables.

[15] Gupta R. and Kabundi A. (2011) found that the DFM model outperforms the other models they used to forecast Treasury Bill Rate for South Africa.

Table 4.2: The RMSE of out-of-sample (2007:01 – 2011:12) for 3 and 12 month horizons:

Forecasting model	h = 3	h = 12
Treasury Bill Rate		
AR benchmark	0.5208	0.6919
DFM	1.1334	0.9829
FANN	**0.9453**	0.9364
ANN	1.0620	**0.7524**
JSE all Share prices		
AR benchmark	1.7743	1.8187
DFM	0.9655	0.9532
FANN	**0.8150**	**0.9273**
ANN	1.0325	1.0947

Note: the first row reports the RMSE for the AR benchmark model; the remaining rows represent the ratio of the RMSE for the forecasting model to the RMSE for the AR. Bold entries indicate the forecasting model with the lowest RMSE.

In order to assess the FANN model forecast accuracy, we performed the cross model test of the FANN against other models, namely AR, DFM and ANN. The cross-model test is based on the statistic proposed by Diebold and Mariano (1995), which is given by: $S = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}}$ where $\bar{d} = \frac{1}{T}\sum_{t=1}^{T}(e_{1t}^2 - e_{2t}^2)$ is the mean difference of the squared prediction error, and $\hat{V}(\bar{d})$ is the estimated variance. Here e_{1t}^2 denotes the forecast errors from the FANN model and e_{2t}^2 denotes the forecast errors from the AR benchmark model, the DFM and ANN. The S statistic follows a standard normal distribution asymptotically. Note, a negative and significant value of S indicate that the FANN model outperforms the other model in out-of-sample forecasting. Table 4.3 below shows the test results. In general the FANN model outperforms the AR and ANN in predicting the two variables of interest and for each of the short and long horizon forecasts. In other words, based on RMSE and on the Diebold and Mariano (1995) test statistics, we have relatively strong evidence that there is a significant statistical gain from using the FANN model over other models. We note that there is no significant statistical difference between the forecasts of factors derived models namely the FANN and the DFM in most cases. For the JSE all share prices variable the forecast of the FANN model outperforms linear and nonlinear univariate models for 12 month horizons with at least 5 percent level of significant, and outperforms all other models for the 3 month horizons in particular.

Table 4.3: Diebold – Mariano test (2006:01 – 2011:12):

Model	Forecasting Horizons	
	3	12
Treasury Bill Rate		
FANN vs. AR	-3.3174 ***	-2.5733**
FANN vs. DFM	-1.2825	0.0503
FANN vs. ANN	-0.5206	3.040**
JSE all share prices		
FANN vs. AR	-3.0829**	-2.0972**
FANN vs. DFM	-3.7276***	0.7960
FANN vs. ANN	-2.7126**	-2.3940**

Note: ***, ** and * indicate significant at the 1%, 5% and 10% levels respectively.

4.7 Conclusion

In this chapter, the Factor Models (FMs) are applied to introduce a new hybrid method for improving the time series forecasting performance of the artificial neural networks. The model used the factors that were extracted from 228 monthly series. Five static factors and two dynamic factors were extracted which explain more than 87 percent of the variation in the data panel. These factors are then used as independent variables or inputs to the ANN in a model we call the factor ANN model (FANN) and to estimate the common linear DFM. Besides the FANN and the DFM, we estimate standard ANN and AR benchmark models. The four models were then used to forecast the Johannesburg Stock Exchange (JSE) share prices and the Treasury Bill Rate over the estimation period 1992:01 to 2006:12. The models were evaluated based on the RMSE for 3 and 12 month-ahead forecasts over an out-of-sample horizon of 2007:01 to 2011:12.

The in-sample results showed the superiority of the FANN over the other models. The FANN outperformed the AR benchmark model with a large reduction in RMSE of around 31 percent to 35 percent. The model outperformed the standard ANN model but the ANN model outperformed the DFM, which in turn performed better than AR benchmark model.

The out-of-sample results revealed that the best performing model over all appears to be our proposed FANN model, followed by the DFM model. These results confirmed the usefulness of the factors that were extracted from large related variables. On the other hand, as far as estimation is concerned, the nonlinear FANN model was suitable to capture nonlinearity and structural breaks compared to linear models. Thus the structural breaks associated with the financial crisis that affected the economy can explain the failure of the linear DFM compared to the nonlinear FANN model. The results of the Diebold-Mariano test suggested that the FANN model produced forecasts that were significantly better than the AR benchmark model forecasts, and the standard ANN model forecasts.

Chapter 5

Evaluating the combined forecasts of the dynamic factor model and the artificial neural network model using linear and nonlinear combining methods*

5.1 Abstract

The chapter evaluates the advantages of combined forecasts from the Dynamic Factor Model (DFM) and the Artificial Neural Networks (ANN). The analysis was based on three financial variables namely the JSE return index, government bond return index and the Rand/Dollar exchange rate in South Africa. The forecasts were based on the out-of-sample period from January 2006 to December 2011. Compared to benchmark auto-regressive (AR) models both the DFM and ANN offer more accurate forecasts with reduced RMSE of around 2 to 12 percent for all variables and over all forecasting horizons. The ANN as a nonlinear combining method outperforms all linear combining methods for all variables and over all forecasting horizons. The results suggest that the ANN combining method can be used as an alternative to linear combining methods to achieve greater forecasting accuracy.

---

* Ali Babikir and Henry Mwambi (Under review). Evaluating the combined forecasts of the dynamic factor model and the artificial neural network model using linear and nonlinear combining methods. *Journal of Forecasting*.

The ANN combining method produces out-of-sample forecasts that are substantially more accurate with a sizeable reduction in RMSE of both the AR benchmark model and the best individual forecasting model. We attribute the superiority of the ANN combining method to its ability to capture any existing nonlinear relationship between the individual forecasts and the actual forecasting values.

Keywords: Dynamic factor model; Artificial neural network; Combination forecast; Forecast accuracy; Root mean square error.

5.2 Introduction

Much effort has been made in research on forecasting accuracy over the past few decades. However, no single forecasting method has been found to outperform others in all situations. The seminal work of Bates and Granger (1969) opened the horizon for research on forecast combination in various economics and financial fields. Thus, a new field in forecasting research has been to combine the forecasts produced by individual models using various combination methods. This allows the final forecast result to borrow strength from the individual forecasting methods, an attribute that cannot be achieved by a single method. The strength of individual forecasting results from the DFM and ANN models is the reason that inspired us to explore the usefulness of these models forecast combination.

The DFM is increasingly becoming popular in economics as it can handle large data sets with many variables in an effective way. The DFM has been used for various purposes. For example the DFM can be used to generate economic indicators and have also been widely used to forecast real and nominal economic variables[16]. They often provide more accurate forecasts than autoregressive and vector autoregressive models; see Eickmeier and Ziegler (2008).

On another note much work suggests that nonlinearities in financial data are well approximated by the ANN model, among others for example see the work by Hutchinson et al. (1994), Donaldson and Kamstra (1996) and Aladag et al. (2010).

The first objective of this study is to evaluate the forecasting performance of the DFM and ANN models to determine which model works best in which situation or domain. To do so, we evaluated our models using three financial variables. Our work was motivated by the more general evidence of results that combination forecasts based on both linear and nonlinear individual model can outperform combination forecasts based only on linear model forecasts; see for example, Blake and Kapetanios (1999), Stock and Watson (1999), and Teräsvirta (2006). We focus on the more specific objective of combining forecasts driven from the data rich linear model (DFM) and the univariate nonlinear model (ANN). The motivation of the selection of these models comes from their advantages as the DFM - our linear model in this combination - can accommodate a large set of variables, and consider their co-movements that assist in producing more accurate forecasts; on the other hand the ANN model has advantages that approximate complex, possibly nonlinear relationships without any prior assumptions about the underlying data generating process or mechanism. The issue of comparing the forecasting performance of the DFM and ANN models and their combination has, to the best of our knowledge, not been yet examined.

We consider forecasts of the three financial variables, namely the JSE return index, government bond return index and the Rand/Dollar exchange rate in South Africa. Recursive out-of-sample

[16] Several empirical works used DFM to forecast GDP, Inflation, per capita growth rate and etc. see, among others, Stock and Watson (2002a), Artis et al. (2005) and Schumacher (2007).

(2006:01 – 2011:12)[17] forecasts of the variables are generated using the DFM and ANN models. The empirical results show that both the DFM and ANN forecasts offer accurate forecasts that dominate the forecast from the AR benchmark model with reductions in RMSE of around 2 percent to 12 percent in all cases and over all forecasting horizons. We then generate combination forecasts based on the DFM and ANN models forecasts using several combining methods, including the mean, variance covariance, discount mean square forecast error (MSFE) as linear combination methods and the ANN as a nonlinear combination method. The combination forecast results show that the RMSE of nonlinear ANN combining method are fairly smaller that the RMSE of linear combining method which are in turn better than results of the AR benchmark model. The nonlinear ANN combining method also outperform the best individual forecasting models for all variables and at all forecasting horizons with sizable reductions in RMSE of around 8 percent to 23 percent of the RMSE of the best individual forecasts. The results of Diebold-Mariano formal statistical test confirm the superiority of the ANN combined forecasts over the AR benchmark model. This is most likely due to the nonlinear relationship between the individual forecasts and the actual forecasting values.

The remainder of the chapter is organized as follows: section 5.3 briefly describes the DFM and ANN forecasting models and the combination methods; section 5.4 presents the data; the results from forecasting models and their combinations are discussed in section 5.5; and section 5.6 gives a brief conclusion of the work and possible future extensions.

[17] The choice of out of sample period comes from the aim to investigate the performance of forecasting models and their forecasts combination during the period of financial crisis.

5.3 Individual Forecasting Models and Combination Methods

The first sub-section in this section briefly introduces the notation, formulation and estimation methods in forecasting models. In the second sub-section we introduce and discuss the different combining methods.

5.3.1 Individual Forecasting Models

5.3.1.1 The Dynamic Factor Model and the Estimation of Factors

This chapter uses the DFM to extract common components from a large set of variables, after which these common components are used to forecast the variables of interest.

Let the panel of observations X_t be the N stationary time series variables with observations at times t = 1,......., T, where it is assumed that the series have zero mean. The idea behind the factor model is that most of the variance of the data set can be explained by a small number $q \ll N$ of factors contained in the vector f_t. In general the dynamic factor model representation is given by

$$X_t = \chi_t + \xi_t = \lambda(L)' f_t + \xi_t \qquad (5.1)$$

where χ_t are the common components driven by factors f_t, and ξ_t are idiosyncratic components for each of the variables. In particular ξ_t is that part of X_t that cannot be explained by the common components. The common component is a function of the $q \times 1$ vectors of dynamic factors which are common to all variables in the set $f_t = (f_{1t} \ldots\ldots f_{qt})'$, where the operator $\lambda(L) = 1 + \lambda_1 L + \cdots + \lambda_s L^s$ is a lag polynomial with positive powers on the lag operator L with

$Lf_t = f_{t-1}$. In this way the lags of the factors are allowed to affect the current movement of the variables. The model can be re-written in static representation as:

$$X_t = \Lambda' F_t + \xi_t \tag{5.2}$$

where F_t is a vector of $r \geq q$ static factors that comprise of the dynamic factors f_t and all lags of the factors. Basically there are three methods of estimating the factors in F_t from a large data set. These methods were developed by Stock and Watson (2002a; hereafter SW), Kapetanios and Marcellino (2004) and Forni, Hallin, Lippi and Reichlin (2005, hereafter FHLR)[18]. In the current chapter we employ the estimation method developed by FHLR. Below we give a brief description of SW and FHLR methods and how they differ.

First we start with the SW model where the authors proposed estimating F_t with static principal component analysis (PCA) applied to X_t . The factor estimates are simply the first r principal components of X_t which according to SW are $F_t = \widehat{\Lambda}' X_t$, where $\widehat{\Lambda}$ is the $N \times r$ matrix of the eigenvectors corresponding to the r largest eigenvalues of the sample covariance matrix $\hat{\Sigma}$.

On the other hand, FHLR propose a weighted version of the principal components estimator suggested by SW, where the series are weighted according to their signal-to-noise ratio, which is estimated in the frequency domain. The estimation of common and idiosyncratic components is conducted using two steps. First, the covariance matrices of the common and idiosyncratic components of X_t are estimated with dynamic PCA. This involves estimating the spectral density matrix of X_t, $\Sigma(\omega)$, which has rank q. For each frequency ω, the largest q eigenvalues and the corresponding eigenvectors of $\Sigma(\omega)$ are computed, and the spectral density matrix of the common components $\sum_{\chi}(\omega)$ is estimated. Then it follows that the spectral density matrix of the idiosyncratic components is given by $\hat{\Sigma}_{\xi}(\omega) = \hat{\Sigma}(\omega) - \hat{\Sigma}_{\chi}(\omega)$. Inverse Fourier transform

18 For further technical details on this type of factor models, see Schumacher (2007).

provides the time-domain autocovariances of the common and the idiosyncratic components given by $\hat{\Gamma}_{\chi}(k)$ and $\hat{\Gamma}_{\xi}(k)$ for lag k. Since dynamic PCA corresponds to a two-sided filter of the time series, this approach alone is not suited for forecasting. Second, a search is undertaken for the r linear combinations of X_t that maximize the contemporaneous covariance explained by the common factors $\hat{Z}_i{}'\hat{\Gamma}_{\chi}(0)\hat{Z}_i$, $i = 1, \ldots, r$. This optimization problem is subject to the normalization $\hat{Z}_i'\hat{\Gamma}_{\xi}(0)\hat{Z}_j = 1$ for $i = j$ and zero otherwise. This representation can be reformulated as the generalized eigenvalue problem such that $\hat{\Gamma}_{\chi}(0)\hat{Z}_i = \hat{\mu}_i\hat{\Gamma}_{\xi}(0)\hat{Z}_i$, where $\hat{\mu}_i$ denotes the *i*-th generalized eigenvalue and $\hat{Z}_i$ its $N \times 1$ corresponding eigenvector in their non-null spaces. The factor estimates according to FHLR are then obtained as $\hat{F}_t = \hat{Z}'X_t$ with $\hat{Z} = [\hat{Z}_1 \ldots \hat{Z}_r]$.

The estimated factors will be used to forecast the variables of interest. The forecasting model is specified and estimated as a linear projection of an h-step ahead transformed variable y_{t+h} into *t*-dated dynamic factors. The forecasting model follows the setup in Stock and Watson (2002a) and Froni et al. (2004) with the form

$$y_{t+h} = \beta(L)\hat{f}_t + \gamma(L)y_t + u_{t+h} \tag{5.3}$$

where $\hat{f}_t$ are dynamic factors estimated using the method by FHLR while $\beta(L)\ and\ \gamma(L)$ are the lag polynomials. The coefficient matrix for factors and autoregressive terms are estimated by ordinary least square (OLS) for each forecasting horizon h. To generate the estimate and forecast of the AR benchmark we impose a restriction to Eq. (5.3), where, we set $\beta(L) = 0$[19].

[19]We use the autoregressive model as our benchmark.

5.3.1.2 The Artificial Neural Network

The ANN model is one of the generalized nonlinear nonparametric models (GNLNPMs). The advantage of ANNs over more traditional econometric models is that they can handle complex, possibly nonlinear relationships without any prior assumptions about the underlying data generating process; see Hornik et al., (1989, 1990) and White (1990).

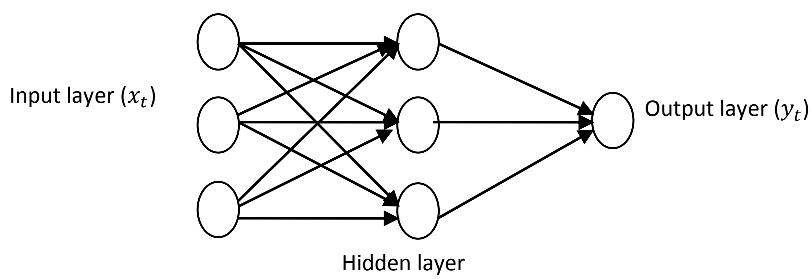

Figure 5.1: Three-layer feed-forward neural network

There are a number of properties that make ANN models an attractive alternative to traditional forecasting models. Most importantly ANN models control the limitations of traditional forecasting methods, including misspecification, biased outliers and assumption of linearity, Hill et al. (1996). Thus in particular, the network used in this chapter is a feed-forward network coupled with linear jump connection or linear neuron activation function. The network has one hidden layer that has three nodes and an input layer which also has three nodes. The input nodes are connected forward to each and every node in the hidden layer, and these hidden nodes are connected to the single node in the output layer, as shown for illustration in Fig. 5.1. The inputs - which are similar to the dependent variables used in the multiple regression model – are connected to the output node - which is similar to the dependent variable –through the hidden layer. We follow McAdam and Hallett (1999) in describing the network model. Thus, the model can be specified as follows:

$$n_{k,t} = w_0 + \sum_{i=1}^{I} w_i x_{i,t} + \sum_{j=1}^{J} \emptyset_j N_{t-1,j} \tag{5.4}$$

$$N_{k,t} = f(n_{k,t}) \tag{5.5}$$

$$y_t = \alpha_{i,0} + \sum_{k=1}^{K} \alpha_{i,k} N_{k,t} + \sum_{i=1}^{I} \beta_i x_{i,t} \tag{5.6}$$

where inputs x represent the lagged values of the variable of interest and the output y is their forecasts. The w_0 is the bias and w_i and $\alpha_{i,k}$ denote the weights that connect the inputs to the hidden layer and the hidden layer to output respectively. The $\emptyset_j$ and β_i links the input to the output via the hidden layer. The I regressors are combined linearly to form Kneurons which in turn, are combined linearly to produce the forecast or output. The Equations (5.4) to (5.6) links inputs x to outputs y through the hidden layer. The function f is a logistic function meaning that $N_{k,t} = f(n_{k,t}) = \frac{1}{1+e^{-n_{k,t}}}$. The second summation in Equation (5.6) shows that we also have a jump connection or skip-layer network that directly links the inputs x to the output y. The most important feature about this model is the combination of the pure linear model and feed-forward neural network. Therefore, if the relationship between inputs and output is pure linear, then only the skip-layer given by coefficient set β should be significant, and if the relationship is a nonlinear one then we expect the coefficients set w and α to be highly significant, while the jump connections coefficient β will be relatively insignificant. Finally however, if the underlying relationship between input and output is mixed, then we expect all coefficient sets to be significant. The selection of the lag lengths and the number of nodes in the hidden layer are chosen on the basis of the in-sample RMSE.

5.3.2 Forecast Combining Methods

We consider four methods to combine individual forecasts generated by the DFM and ANN models. The combining methods comprise of three linear combining methods (The mean, VACO and Discount MSFE based methods) and one nonlinear combining method (ANN). As some of the combining methods require a holdout period to calculate the weights used to combine individual forecasts, we use the first 24 months of the out-of-sample as holdout observations. For all combining methods, we form combination forecasts over the post holdout out of sample period. Brief details about the above combining methods are given below.

5.3.2.1 Mean combination method

The simple average serves as a useful benchmark and has been shown to perform better than some complicated methods, see for example Makridakis and Winkler (1983), Clemen and Winkler (1986), Guerard and Clemen (1989) as well as Diebold and Pauly (1990). Fang (2003) found the performance of the simple average combination method to be superior to single forecasts. The simple average combination method calculates the composite forecasts without taking the historical performance of the individual forecasts into account, as the combination weight is assigned equally to each of the individual forecasts. The simple average combination method can be expressed as:

$$\hat{y}_t^c = \sum_{i=1}^{m} w_i \, \hat{y}_t^i \tag{5.7}$$

where $\hat{y}_t^c$ is the combined forecast at time t, $\hat{y}_t^i$ is the forecast from ith individual forecasting model, $w_i = \frac{1}{m}$ is the weight of individual forecast for model i, m is the total number of individual forecasting models. Note that other forms of weights are possible as will be seen for the next two methods but generally the weights have to satisfy the condition $\sum_{i=1}^{m} w_i = 1$.

5.3.2.2 Variance Covariance (VACO) Combination Method

The VACO method calculates the weights by taking the historical performance of the individual forecasts into consideration. The method determines the weights according to the following equation:

$$w_i = \frac{\left[\sum_{j=1}^{T}\left(y_j - \hat{y}_j^i\right)^2\right]^{-1}}{\sum_{i=1}^{m}\left[\sum_{j=1}^{T}\left(y_j - \hat{y}_j^i\right)^2\right]^{-1}} \tag{5.8}$$

Then the combined forecast is given by; $\hat{y}_t^c = \sum_{i=1}^{m} w_i \hat{y}_t^i$, where y_j is the jth actual value, $\hat{y}_j^i$ is the jth forecasting value from ith individual forecasting model, T is the total number of out of sample points.

5.3.2.3 Discount Mean Square Forecast Error (DMSFE) Combination Method

The DMSFE method weights recent forecasts more heavily than distant ones. Winkler and Makridakis (1983) suggest that the weights can be calculated as:

$$w_i = \frac{\left[\sum_{j=1}^{T} \delta^{T-j+1}\left(y_j - \hat{y}_j^i\right)^2\right]^{-1}}{\sum_{i=1}^{m}\left[\sum_{j=1}^{T} \delta^{T-j+1}\left(y_j - \hat{y}_j^i\right)^2\right]^{-1}} \tag{5.9}$$

where δ is the discount factor with $0 < \delta \leq 1$, if $\delta = 1$ then Eq. (5.9) of DMSFE method becomes Eq. (5.8) meaning that the VACO is a special case of DMSFE. Note that as earlier stated the above weights satisfy the condition $\sum_{i=1}^{m} w_i = 1$.

5.3.2.4 Artificial Neural Network (ANN) Combination Method

The linear combining methods introduced above are only based on linear combinations of the individual forecasts. The problem is that linear combinations are bound to be inefficient if the individual forecasts are based on nonlinear models or if the true relationship is nonlinear. For the success of the ANN as a combination method over the linear methods, among others, see Donaldson and Kamstra (1996) and Harrald and Kamstra (1997).

Here we use the same setup used in sub-section (2.1.2), the output $\hat{\mathrm{y}}_{\mathrm{t}}^{\mathrm{c}}$ of combined forecasts can be given by

$$\hat{\mathrm{y}}_{\mathrm{t}}^{\mathrm{c}} = \alpha_{i,0} + \sum_{k=1}^{K} \alpha_{i,k} N_{k,t} + \sum_{i=1}^{m} \beta_i \hat{y}_t^i \tag{5.10}$$

where $\hat{\mathrm{y}}_{\mathrm{t}}^{\mathrm{i}}$ is the forecast from i-th individual forecasting model.

5.4 Data Presentation and Preliminary Findings

The ANN model includes data on only the variable of interest. The DFM contains 228 monthly series, 203 from South Africa, covering the financial, real and nominal sectors, two global variables and 23 series of major trading partners and global financial markets. Thus besides the national variables the chapter uses a set of global variables such as gold and crude oil prices. In addition the data also includes series from financial markets of major trading partners namely the United Kingdom, the United States, China and Japan. The in-sample period contains data from 1992:01 to 2005:12, while the out-of-sample set spans from 2006:01 to 2011:12. The Augmented Dickey-Fuller (ADF) test is used to assess the degree of integration of all series. All non-stationary series are made stationary through differencing. The Schwarz information criterion (SIC) is used in selecting the appropriate lag length in such a way that no serial correlation is left in the stochastic error term. All series are standardized to have a mean of zero and a constant variance.

Recently the determination of the number of the factors has been developed for both the case of the static factor model [Bai and Ng, (2002) and Alessi et al. (2008)] and the dynamic factor model [Bai and Ng, (2007); Amengual and Watson (2007); Hallin and Liska (2007); Onatski (2009, 2010)]. To specify the number of static factors, Bai and Ng (2002) and Alessi et al. (2008) use information criterion based on AIC and BIC to help guide the selection of the optimal number of factors r in a large data set. We apply the Bai and Ng (2002) approach which proposes five static factors. Onatski (2009) developed a statistical test to test and determine the number of dynamic factors under the null hypothesis that the number of factors is equal to k_0 against the alternative $k_1 > k_0$ (for details see Onatski (2009)). In our case the test suggests two dynamic factors, which both explain more than 87 percent of variation.

5.5 Evaluation of Forecast Accuracy

5.5.1 In-sample results

We first assess the in-sample predictive power of the fitted models. We estimate the forecasting models using the full sample, in order to check the robustness of our in-sample results. In-sample forecasting is most useful when it comes to examining the true relationship between the set of predictors and the future predictions of the variable of interest. Table 5.1 below reports the in-sample forecasting results. The first row reports the RMSE[20] for the AR benchmark model, while the remaining rows present the ratio of the RMSE of the model of interest to the RMSE of the

[20]The RMSE statistic can be defined as $\sqrt{\frac{1}{N}\sum(Y_{t+n} - {}_t\hat{Y}_{t+n})^2}$, where Y_{t+n} denotes the actual value of a specific variable in period $t+n$ *and* ${}_t\hat{Y}_{t+n}$ is the forecast made in period t for $t+n$.

AR benchmark model. The model with a lowest RMSE ratio is deemed to perform better than the other models. In our case the DFM out performed all other models including the AR for all variables. The reason is potentially because the DFM can efficiently handle large amounts of information that include external variables that influence the South African financial sector, therefore helping to improve the forecasting performance. The ANN model also provides fairly better in-sample forecasts compared to the AR benchmark model.

Table 5.1: In sample results: Relative RMSE for financial variables

Forecasting model	JSE Return Index	Government Bond Return Index	Rand/Dollar Exchange rate
AR (benchmark model)	1.000	1.000	0.9997
DFM	***0.854***	***0.859***	***0.847***
ANN	0.915	0.885	0.859

Note: The first row reports the RMSE for the AR benchmark model; the remaining rows represent the ratio of the RMSE for the forecasting model to the RMSE for the AR. Bold entries indicate the forecasting model with the lowest RMSE.

5.5.2 Performance of Individual Forecasting models

We estimate the two individual forecasting models, namely the DFM and ANN and our benchmark model AR as well, based on the data from 1992:01 to 2005:12 and then using an expanding window, we recursively estimate out-of-sample forecasts to generate 3, 6 and 12 month-ahead forecasts for the period from 2006:01 to 2011:12. In other words, we re-estimated the models by adding a month each time over the out-of-sample forecast horizon to update the estimation of the coefficients before generating the 3, 6 and 12 month-ahead forecasts. We evaluate the out-of-sample forecasts for our variables of interest - namely Johannesburg Stock Exchange Return Index, Government Bond Return Index and Rand/Dollar Exchange Rate - over the period from 2006:01 to 2011:12. This period includes the financial crisis that affected most emerging markets like South Africa, which led to very high volatility, in general, with the South African economy reaching the trough of the business cycle by the end of 2008; Venter (2011). The importance of the impact of US stock returns on South African stock returns has recently been highlighted by Gupta and Modise (2012). In light of this, as the US economy started to show mild signs of revival, its decreased uncertainty is likely to have produced lower levels of volatility in the South African stock returns. Further, as the US recession was officially called-off in the first quarter of 2009, a reduced volatility in the stock returns was observed in the early third quarter of 2009. Also, both the leading and the coincident financial indicators for South Africa had started to turn upwards in the first quarter of 2009 (Venter 2011). In addition, as indicated by Van Wyk de Vries et al. (forthcoming), during the financial crisis followed by the global uncertainty, hedging demand by South African investors for domestic stocks were much less volatile. The domestic stocks showed a positive mean value than hedging demands for US and UK stocks, with the mean value of the latter set of stocks being actually negative. Naraidoo and Raputsoane (2010) indicated that the South African Reserve Bank had systematically adjusted the financial conditions index (containing stock prices) during the recent financial crisis to minimize the forecasted volatility in the financial conditions index.

In Table 5.2 below, we compare the RMSEs of the out-of-sample forecasting results for the AR benchmark model and the other forecasting models. The table reports the RMSE statistics for the AR benchmark model and the ratio of the RMSE for the competing models to the RMSE for the AR benchmark model. A relative RMSE less than one indicates a superior forecasting performance of the model for the chosen forecast horizons h = 3, 6 and 12. In our analysis we consider 3, 6 and 12 months as short, medium and long forecast horizons respectively. The results of the three variables can be summarized as follows:

I. ***Johannesburg Stock Exchange (JSE) Return index***: among the competing models the DFM clearly outperforms the other models over all forecasting horizons under consideration. The relative RMSEs of the DFM model declines as the horizon increases, suggesting that more accurate forecasts of the return are available at longer horizon.

II. ***Government Bond Return Index*****:** similar to the variable above, the DFM clearly outperforms the other models over all forecasting horizons under investigation.

III. ***Rand/Dollar Exchange rate***: we can see that the ANN model shows better results over all forecasting horizons.

Note, we do not apply any smoothing method to the data; we let the data speak for themselves. For this variable there is evidence of high non-linearity associated with extreme value fluctuations which are best captured by models that can handle non-linearity such as the ANN.

Table 5.2: Out-of-sample (2006:01 – 2011:12) relative RMSE for financial variables (3, 6 and 12 month horizons)

Forecasting model	h = 3	h = 6	h = 12
	JSE Return Index		
AR (benchmark model)	1.571	1.588	1.591
DFM	***0.954***	***0.945***	***0.936***
ANN	0.995	0.988	0.986
	Government Bond Return Index		
AR (benchmark model)	1.448	1.571	1.541
DFM	***0.981***	***0.913***	***0.934***
ANN	0.992	0.967	0.962
	Rand/Dollar Exchange rate		
AR (benchmark model)	1.175	1.143	1.136
DFM	1.013	1.038	1.019
ANN	***0.882***	***0.898***	***0.918***

Note: The first row reports the RMSE for the AR benchmark model; the remaining rows represent the ratio of the RMSE for the forecasting model to the RMSE for the AR. Bold entries indicate the forecasting model with the lowest RMSE.

5.5.3 Combining Forecasts

Table 5.4 below shows the results of combining forecasts of the DFM and ANN models. We aim to merge the advantages of the DFM model that accommodate a large number of variables and the ANN model with its flexibility to account for potentially complex nonlinear relationships not easily captured by traditional linear models. Similarly to Table 5.2, Table 5.4 reports the RMSE for the AR benchmark model and the ratio of the RMSE for a given combining method to the RMSE for the AR benchmark model. Note that all combining methods produce forecasts that are more accurate than the AR benchmark model. Overall the nonlinear ANN combining method performs consistently well for all variables at all forecasting horizons, and hence offers a more reliable method of generating reliable forecasts of the variables of interest. The nonlinear ANN combining method consistently outperformed the AR benchmark model with large reductions in RMSE of around 10 percent to 35 percent relative to the AR over all forecasting horizons and variables. The nonlinear ANN combining method also beat the best individual forecasting models for all variables and over all forecasting horizons with sizable reductions in RMSE of around 8 percent to 23 percent of the RMSE of the best individual forecasts. We note in addition that the Discount MSFE with δ = 0.9 as a combining method performs nearly as well as the best individual model for all variables and forecasting horizons. The variance covariance (VACO) combining method performs less accurately compared to other combining methods over all forecasting horizons and variables with exception of the *Rand/Dollar Exchange rate* variable.

To determine whether the differences we observed in the forecasting performances of the ANN combined forecasts to AR benchmark model based on RMSEs are statistically significant, the test proposed by Diebold and Mariano (1995) is used[21]. Table 5.3 below shows the results of the Diebold-Mariano test; we conclude that the forecasts of the nonlinear ANN combining method are statistically more accurate than the AR benchmark forecasts for all variables and over all forecasting horizons, with the test statistic being significant at least at 10 percent level.

---

[21] The Diebold and Mariano test statistic is given by; $S = \frac{\bar{d}}{\sqrt{\hat{V}(\bar{d})}}$ where $\bar{d} = \frac{1}{T}\sum_{t=1}^{T}(e_{1t}^2 - e_{2t}^2)$ is the mean difference of the squared prediction error, $\hat{V}(\bar{d})$ is the estimated variance. Here e_{1t}^2 denote the forecast errors from the combined forecasts using ANN combination method and e_{2t}^2 denote the forecast errors from the AR benchmark model. The S statistic follows a standard normal distribution asymptotically. Note, a negative and significant value of S indicate that the ANN combination method outperforms the AR benchmark model in out of sample forecasting.

Table 5.3: Forecast Combining Results – RMSE – for financial variables (2006:01 – 2011:12)

Combination Method	h = 3	h = 6	h = 12
	JSE Return Index		
AR	1.5711	1.588	1.591
Mean	0.960	0.954	0.945
VACO	0.985	0.982	0.983
DMSFE δ = 0.95	0.959	0.953	0.943
DMSFE δ = 0.90	0.957	0.952	0.942
ANN	***0.870***	***0.759***	***0.832***
	Government Bond Return Index		
AR	1.448	1.571	1.541
Mean	0.976	0.916	0.929
VACO	0.980	0.935	0.950
DMSFE δ = 0.95	0.976	0.915	0.929
DMSFE δ = 0.90	0.976	0.916	0.928
ANN	***0.887***	***0.832***	***0.895***
	Rand/Dollar Exchange rate		
AR	1.175	1.143	1.136
Mean	0.898	0.924	0.918
VACO	0.876	0.895	0.905
DMSFE δ = 0.95	0.889	0.915	0.913
DMSFE δ = 0.90	0.894	0.919	0.917
ANN	***0.646***	***0.892***	***0.813***

Note: The first row reports the RMSE for the AR benchmark model; the remaining rows represent the ratio of the RMSE for the combining method to the RMSE for the AR. Bold entries indicate the combining method with the lowest RMSE.

Table 5.4: Diebold – Mariano test (2006:01 – 2011:12)

Model	Forecasting Horizons		
	3	6	12
JSE Return Index ANN Combined forecasts vs. AR	-2.3423**	-2.6830***	-3.1614***
Government Bond Return Index ANN combined forecasts vs. AR	-1.9383*	-2.3967**	-4.7805***
Rand/Dollar Exchange rate ANN combined forecasts vs. AR	-1.7152*	-2.5546**	-1.9525*

Note: ***, ** and * indicate significant at the 1%, 5% and 10% levels respectively.

5.6 Conclusion

This chapter evaluates the usefulness of the application of combining forecasts of a model using large panel of data - DFM - and univariate nonlinear model - ANN - using out of sample period from 2006:01 to 2011:12 to forecast financial variables, namely, JSE Return index, Government Bond Return Index and Rand/Dollar Exchange Rate in South Africa. Despite the extensive work on forecasting returns in South Africa, this is the first attempt in using the DFM and ANN combined forecasts to forecast financial variables, particularly in the South African context. This combining strategy is able to merge the advantages of the unique individual strengths of these models of accommodating a large number of related variables and flexibility to capture linear and nonlinear relationships. It is clear that forecast combinations represent a realistic approach for dealing with the misspecification biases that affect individual forecasting models. Since individual models may be biased in different directions, it is important to consider which types of forecasts to combine, i.e., forecasts from linear versus nonlinear models and forecasts from univariate versus multivariate models.

In individual forecasting models, our empirical results confirm that, compared to the AR benchmark model, both the DFM and ANN forecasts offer accurate forecasts that dominate the forecast from the AR benchmark model with reductions in RMSE of around 2 percent to 12 percent in all cases and over all forecasting horizons.

The study also used some of the recently studied linear combination methods, and used nonlinear ANN as an alternative combining method. The empirical results (Table 5.4) of combining forecasts showed that the RMSE of nonlinear ANN combining method are fairly smaller than the RMSE of linear combining methods which are in turn better than results of the AR benchmark model. The nonlinear ANN combining method also outperformed the best individual forecasting models for all variables and at all forecasting horizons with sizable reductions in RMSE of around 8 percent to 23 percent of the RMSE of the best individual forecasts. The results of the Diebold-Mariano test suggested that the ANN combining method produced forecasts that were significantly better than the AR benchmark model forecasts. This is most likely due to the nonlinear relationship between the individual forecasts and the actual forecasting values. Our empirical results confirm the usefulness of ANN in modelling and combining forecasts of financial variables. Possible future extensions include methods that can strengthen the combining power of the ANN as a combining method in the broad class of generalized nonlinear non-parametric models.

Chapter 6

Conclusions

This thesis examined the advantages of combining the dynamic factor model (DFM) and artificial neural networks (ANNs) by introducing new novel models that have capabilities to produce more accurate forecasts with application to the South African financial sector data. One of the most important advantages of the dynamic factor framework is that it can accommodate a large number of variables, and extract a few factors that explain the comovement of all South African financial sector or economy variables. On the other hand, artificial neural networks are universal approximators, nonlinear method and data-driven self-adaptive methods in which there are few a priori assumptions to be made about the models. Thus, combining models with such features in order to produce forecasts can lead to a good forecasting performance. The dataset used in this study contains 228 monthly series, 203 of which are from South Africa, covering the financial, real and nominal sectors, two global variables, namely, gold and crude oil prices and 23 series of global financial markets and major trading partners, namely, the United Kingdom, the United States, China and Japan.

In the second chapter, the Factor Augmented Artificial Neural Network (FAANN) model which merges the factors that were extracted from a large data set with ANN was introduced for forecasting. The out-of-sample forecasting performance results of the model compared to the DFM and the autoregressive (AR) benchmark model for three, six and twelve month-ahead forecast horizons for three variables, namely, Deposit rate, Gold mining share prices and Long term interest rate are reported. The FAANN models provided more accurate forecasts for the three above-mentioned variables over the out-of-sample forecast horizons. The model can provide a better alternative for time series forecasting due to its promising performance and capability in handling time series data.

Financial crisis affected financial sectors and economies around the world thus leading to downturn and fluctuations in all variables; these fluctuations are possibly inherent with nonlinearity. Based on this, the third chapter introduced a new model that uses the dynamic factor model (DFM) framework, where the ANNs were employed as smoother or data approximators before extracting the factors. After we smoothed the dataset, factors were extracted and then used as explanatory variables in order to produce more accurate forecasts. The ANN-DFM was applied to forecast three South African financial variables, namely, Rate on 3-month trade financing, lending rate and Short term interest rate. The results, based on the root mean square errors and Diebold-Mariano test of three, six and twelve months ahead out-of-sample forecasts indicated that, in all of the cases, the ANN-DFM and the DFM statistically outperformed the AR models. In the majority of the cases the ANN-DFM outperformed the DFM with a sizable reduction in RMSEs. In the minority of the cases where DFM outperformed ANN-DFM the improvement was marginal. The results demonstrated the usefulness of the factors in forecasting performance.

In the fourth chapter, we introduced a new model where the extracted factors are used as input to the nonlinear ANNs model. The chapter investigates the capability of the forecasting methods to improve forecast performance where the same extracted factors from a large dataset are used as explanatory to the regression model in the DFM and also used as inputs to ANNs to produce the

FANN model. The Johannesburg Stock Exchange (JSE) share prices and the Treasury Bill Rate were used to evaluate the in-sample and out-of-sample forecasts. In terms of the in-sample forecasts, the FANN model outperformed all alternatives. The out-of-sample empirical results revealed that our proposed FANN model statistically outperformed the DFM and the AR benchmark model for all variables and over all forecasting horizons. Furthermore, the DFM model beat the AR benchmark model in out-of-sample exercise. These results confirmed the usefulness of the factors that are extracted from large related variables. However, when we compared the forecasting performance of the FANN model that was based on nonlinear estimation method to the DFM which was based on linear estimation method, we concluded that the FANN model was more likely to do better than the DFM. Thus, the nonlinear method was more suitable to capture nonlinearity and structural breaks compared to the linear method.

The fifth chapter made use of a number of linear and nonlinear combining methods to pool the DFM and ANNs forecasts. Three financial variables, namely, the JSE return index, government bond return index and the Rand/Dollar exchange rate in South Africa were used to investigate the advantages of combining the two models over the out-of-sample forecast. The individual forecast of the ANNs and the DF models outperformed the AR benchmark model. On one hand, the results of combined forecasts showed the superiority of all combining methods compared to the AR benchmark model. On the other hand, the ANNs as a nonlinear combining method outperformed all the linear combining methods and the best individual forecasts for all variables and over all forecasting horizons. The superiority of the ANNs as a combining method can be attributed to its ability to capture any existing nonlinear relationship between the individual forecasts and the actual forecasting values.

The scope of this thesis was limited to combining forecasts of Artificial Neural Networks and Dynamic Factor Model in order to generate more accurate forecasts with applications to the South African financial sector. However, there is still a considerable amount of research that needs to be undertaken in the field of forecasting study. The following recommendations could serve as extensions and an agenda for future work:

- Evaluate the new models for forecasting performance in small and large simulated samples.
- Evaluate the introduced models to investigate their capabilities in nowcasting.
- Investigate the capabilities of the proposed models to handle the issue of regime switch.
- Compare the forecasting performance of the factor augmented artificial neural network model to the factor augmented vector autoregressive model.
- Evaluate the forecasting performance of the factor augmented artificial neural network and factor artificial neural network models using different optimization algorisms.
- Augment factors to Generalized Autoregressive Conditional Heteroskedasticity model in order to assess volatility.
- Compare the forecasting performance of combined forecasts from different models to the forecasts from the combined models.

Bibliography

Abraham A., and Nath B., (2001). A Neourofuzzy Approach for Modeling Electricity Demand In Victoria, *Applied Soft Computing Journal*, 1, 127-138.

Alessi L., Barigozzi M. and Capasso M. (2008). A robust criterion for determining the number of static factors in approximate factor models.*Working paper No 903, European Central Bank*.

Altissimo, F., A. Bassanetti, R. Cristadoro, M. Forni, M. Hallin, M. Lippi and L. Reichlin (2001). EuroCOIN: a real time coincident indicator of the euro area business cycle, *CEPR Working Paper* 3108.

Amengual D. and Watson M.W. (2007). Consistent estimation of the number of dynamic factors in a large N and T panel. *Journal of Business and Economic Statistics*, 25, 91-96.

Angelini, E., G. Camba-Mendez, D. Giannone, L. Reichlin and G. Rünstler. (2011). Short-term forecasts of euro area GDP growth, *The Econometrics Journal,* 4(1), 25-44.

Angelini, E., Henry, J. and Mestre, R. (2001). Diffusion index-based inflation forecasts for the euro area. *Working Paper No*. 61. *European Central Bank*.

Ard H.J. and den Reijer, (2005). Forecasting Dutch GDP using large scale factor models, *Working paper 28*. *Netherland Bank*.

Artis, M. J., A. Banerjee and M. Marcellino (2005). Factor forecasts for the UK, *Journal of Forecasting*. 24(4), 279-298

Aruoba, S., Diebold, F. and Scotti, C. (2009), Real-time measurement of business conditions, *Journal of Business and Economic Statistics* 27, 417–427.

Babikir A., Gupta R., Mwabutwa C. and Owusu-Sekyere E. (2012). Structural breaks and GARCH models of stock return volatility: The case of South Africa. *Economic Modelling*. 29, 2435–2443.

Bai, J. (2003), Inferential theory for factor models of large dimensions, *Econometrica* 71(1), 135– 171.

Bai, J. and Ng, S. (2002), Determining the number of factors in approximate factor models, *Econometrica,* 70(1), 191–221.

Bai, J. and Ng, S. (2007), Determining the number of primitive shocks, *Journal of Business and Economic Statistics* 25(1), 52–60.

Bai, J. and Ng, S. (2010), Instrumental variable estimation in a data-rich environment, *Econometric Theory,* 26(6), 1577–1606.

Bańbura, M. and G. Rünstler, (2011). A look into the factor model black box: publication lags and the role of hard and soft data in forecasting GDP. *International Journal of Forecasting,* 27(2), 333–346.

Banerjee, A. and Marcellino, M. (2008), Factor augmented error correction models. CEPR *Discussion Paper*, 6707.

Banerjee, A., M. Marcellino, and I. Masten, (2005). Leading indicators for euro-area inflation and GDP growth. *Oxford Bulletin of Economics and Statistics*, 67, 785-813.

Banerjee, A., Marcellino, M. and Masten, I. (2008), Forecasting macroeconomic variables using diffusion indexes in short samples with structural change, *In forecasting in the Presence of Structural Breaks and Model Uncertainty,* ed. D. Rapach and M. Wohar. Bingley, UK. Emerald Group.

Banerjee, A., Marcellino, M., and Masten, I. (2005). Leading indicators for euro-area inflation and GDP growth. *Oxford Bulletin of Economics and Statistics*, 67, 785–814.

Bates, J.M. and C.W.J. Granger, (1969). The combination of forecasts. *Operational Research Quarterly*. 20: 451-468.

Bernanke, B., Boivin, J. and Eliasz, P. (2005), Measuring the effects of monetary policy: a factor-augmented vector autoregressive (FAVAR) approach, *Quarterly Journal of Economic,* 120, 387-422.

Blake AP and Kapetanios G. (1999). Forecast combination and leading indicators: Combining artificial neural network and autoregressive forecasts. Manuscript, *National Institute of Economic and Social Research.*

Boivin, J. and Giannoni, M. (2006), DSGE models in a data-rich environment, Technical report, Columbia Business School, Columbia University.

Boivin, J., and Ng, S. (2006). Are more data always better for factor analysis? *Journal of Econometrics*, 132, 169–194.

Brace M. C., Schmidt J. and Hadlin M., (1991). Comparison of The Forecasting Accuracy of Neural Networks With Other Established Techniques. In: Proceedings of the First Forum On *Application of Neural Network to Power Systems*, Seattle, WA., 31-35.

Bruneau, C., O. de Bandt, A. Flageollet and E. Michaux (2003), "Forecasting inflation using economic indicators: the case of France", *Banque de France NER* 101.

C. M. Bishop, (1995). *Neural Networks for Pattern Recognition*, Oxford University Press, Oxford, UK.

C.H. Aladag, E. Egrioglu and C. Kadilar (2009). Forecasting nonlinear time series with a hybrid methodology, *Applied Mathematics Letters*, 22, 1467–1470.

Cagdas Hakan Aladag, Erol Egrioglu and Ufuk Yolcu (2010). Forecast Combination by Using Artificial Neural Networks. *Neural Processing Letters*. 32, 269–276.

Caire P., Hatabian G. and Muller C, (1992). Progress in Forecasting by Neural Networks, In: *Proceedings of the International Joint Conference on Neural Networks*, 2, 540-545.

Carlo Altavilla and Paul De Grauwe, (2010). Forecasting and combining competing models of exchange rate determination, *Applied Economics*, 42, 3455–3480.

Chamberlain, G. (1983). Funds, factors, and diversification in arbitrage pricing models. *Econometrica*, 51, 1281–1304.

Chamberlain, G., and Rothschild, M. (1983). Arbitrage factor structure and mean-variance analysis in large markets. *Econometrica*, 51, 1305–1324.

Chen, A., Leung, M. T., and Hazem, D. (2003). Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan Stock Index. *Computers and Operations Research*, 30, 901–923.

Chen, A., Leung, M. T., and Hazem, D. (2003). Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan Stock Index. *Computers and Operations Research*, 30, 901–923.

Choudhary M. and Haider A., (*2012*). Neural network models for inflation forecasting: an appraisal, *Applied Economic*, 44, 2631-2635.

Choudhary, Rohit and Garg Kumkum, (2008). A Hybrid Machine Learning System for Stock Market Forecasting, *World Academy of Science, Engineering and Technology,* 39*,* 315-318.

Clemen, R. T., and Winkler, R. L. (1986). Combining economic forecasts. *Journal of Business and Economic Statistics.* 4, 39–46.

Cristadoro, R., Forni, M., Reichlin, L., and Veronese, G. (2005). A core inflation indicator for the euro area. *Journal of Money, Credit and Banking*, 37, 539–560.

Dakun Yang and Wei Wu, (2012). A Smoothing Interval Neural Network. *Discrete Dynamics in Nature and Society*, 2012.

Dandan Liu and Dennis W. Jansen (2007). Macroeconomic Forecasting using Structural Factor Analysis, *Journal of Forecasting, 23, 655-677.*

Das S., Gupta R., and Kabundi A., (2011). Forecasting Regional House Price Inflation: A Comparison between Dynamic Factor Models and Vector Autoregressive Models. *Journal of Forecasting,* 30, 288–302.

Delen D., Waller G. and Kadam A., (2005). Predicting Breast Cancer Survivability: A Comparison of Three Data Mining Methods, *Artificial Intelligence in Medicine*, 34, 13-127.

Diebold, F. X., and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics.* 13,253– 263.

Diebold, F. X., and Pauly, P., (1990). The use of prior information in forecast combination. *International Journal of Forecasting.* 6, 503–508.

Dufour, J.-M. and Pelletier, D. (2013), Practical methods for modelling weak VARMA processes: Identification, estimation and specification with a macroeconomic application, *Discussion paper*.

Duzgun R. (2010). Generalized regression neural network for inflation forecasting. *International Research Journal* of *Finance* and *Economics*, 51, 59-70.

Eickmeier, S. and C. Ziegler (2008), How successful are dynamic factor models at forecasting output and inflation? A meta-analytic approach, *Journal of Forecasting*, 27(3), 237-265.

Emrah Önder, Fırat Bayır and Ali Hepşen, (2013). Forecasting Macroeconomic Variables using Artificial Neural Network and Traditional Smoothing Techniques, *Journal of Applied Finance & Banking*, 3(4), 73-104.

F.M. Tseng, H.C. Yu and G.H. Tzeng, (2002). Combining neural network model with seasonal time series ARIMA model, *Technological Forecasting and Social Change*, 69, 71–87.

Fang, Y. (2003). Forecasting combination and encompassing tests. *International Journal of Forecasting.* 19, 87–94.

Favero, C., Marcellino, M. and Neglia, F. (2005), Principal components at work: The empirical analysis of monetary policy with large datasets, *Journal of Applied Econometrics,* 20, 603-620.

Forni M., Hallin M., Lippi M. and Reichlin L. (2000). The Generalized Dynamic-Factor Model: Identification And Estimation. *The Review of Economics and Statistics*, 82(4), 540-554.

Forni, M., Hallin, M., Lippi, M., and Reichlin, L. (2001). Coincident and leading indicators for the euro area. *The Economic Journal*, 111, 62–85.

Forni M, Hallin M, Lippi M. and Reichlin L., (2003). Do financial variables help forecasting inflation and real activity in the euro area? *Journal of Monetary Economics*, 50, 1243–1255.

Forni M., Hallin M., Lippi M. and Reichlin L. (2004). The generalized dynamic factor model: Consistency and rates. *Journal of Econometrics,* 119, 231-255.

Forni M., Hallin M., Lippi M. and Reichlin L., (2005). The generalized dynamic factor model: One sided estimation and forecasting. *Journal of the American Statistical Association*. 100, 830-840.

Franses, P.H. and K. van Griensven, (1998). Forecasting exchange rates using neural networks for technical trading rules, *Studies in Nonlinear Dynamics and Econometrics,* 2, 109–116.

Ganeta R., Romeo L. M., and Gill A., (2006). Forecasting of Electricity Prices with Neural Networks. *Energy Conversion and Management*, 47, 1770-1778.

Geweke, J. (1977). The dynamic factor analysis of economic time series. In D. J. Aigner, & A. S. Goldberger (Eds.), *Latent variables in socio-economic models* (pp. 365–383). Amsterdam: North Holland.

Giannone, D., and Matheson, T. D., (2007). A new core inflation indicator for New Zealand. *International Journal of Central Banking*, 3, 145–180.

Giannone, D., Reichlin, L. and Small, D., (2008). Nowcasting: the real-time informational content of macroeconomic data, *Journal of Monetary Economics,* 55, 665–676.

Gosselin, M.-A., and Tkacz, G. (2001). Evaluating factor models: an application to forecasting inflation in Canada. *Working Paper No,* 18, *Bank of Canada*.

Granger C. W. J. and Terasvirta T., (1998). Modeling Nonlinear Economic Relationships, Oxford University Press, Oxford. In: G. Zhang et al. (Eds.), *Forecasting with Artificial Neural Networks*: The State of Art, *International Journal of Forecasting*, 14, 35-62.

Greg Tkacz and Sarah Hu, (1999). Forecasting GDP Growth Using Artificial Neural Networks, *Working Paper*, 3, *Bank of Canada*.

Gupta R, and Kabundi A., (2010). Forecasting Macroeconomic Variables in a Small Open Economy: A Comparison between Small- and Large-Scale Models*, Journal of Forecasting,* 29, 168–185.

Gupta R. and Kabundi A (2011). A large factor model for forecasting macroeconomic variables in South Africa. *International Journal of Forecasting,* 27(4), 1076–1088.

Hallin M. and Liska R. (2007). Determining the number of factors in the general dynamic factor model. *Journal of the American Statistical Association*, 102, 603-617.

Hamid S. A., and Iqbal Z., (2004).Using Neural Networks for Forecasting Volatility of S&P 500 Index Futures Prices, *Journal of Business Research*, 57, 1116-1123.

Hassan, Md. Rafiul, Nath, Baikunth and Kirley, Micheal, (2007) A fusion model of HMM, ANN and GA for Stock Market Forecasting, *Expert Systems with Applications,* 33, 171-180.

Hendry, D. F., and Clements, M. P. (2004). Pooling of forecasts. *Econometrics Journal*, **7**, 1–31.

Heravi S., Osborn D. R. and Birchenhallc R., (2004). Linear Versus Neural Network Forecasting for European Industrial Production Series. *International Journal of Forecasting*, 20, 435-446.

Hill, T., O'connor, M. and Remus, W., (1996). Neural network models for time series forecasts. *Management Science*. 42, 1082-1092

Hippert H. S., Pedreira C. E. and Souza R. C., (2001). Neural Networks for short-term load forecasting: A review and evaluation. *IEEE Transactions on Power Systems*, 16(1), 44-45.

Hornik, K., Stinchcombe, M. and White, H. (1989). Multi-layer feed forward networks are universal approximators. *Neural Network*, 2, 359-66.

Hornik, K., Stinchcombe, M. and White, H., (1990). Universal approximation of an unknown mapping and its derivatives using multilayer feed forward networks. *Neural Network.* 3, 551-60.

Hutchinson, J.M., A.W. Lo and T. Poggio, (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *Journal of Finance*. 49, 851–889.

Indranarain Ramlall (2010). Artificial Intelligence: Neural Networks Simplified. *International Research Journal of Finance and Economics*, 39.

Jain A., and Kumar A. M., (2007). Hybrid Neural Network Models for Hydrologic Time Series Forecasting, *Applied Soft Computing*, 7(2), 585-592.

Jane M. Binner, Rakesh Bissoondeeal, Thomas Elger, Alicia M. Gazely and Andrew W. Mullineux, (2005). A comparison of linear forecasting models and neural networks: an application to Euro inflation and Euro Divisia, *Applied Economic*, 37, 665-680.

Jane T, Hill and Derek G, Goring, (1998). Smoothing hydrological data: a neural network approach. *Journal of Hydrology*, 37(2), 81-94.

Jorge Nocedal and Stephen J. Wright (2006), *Numerical Optimization* (2nd ed.), Berlin, New Yor. Springer.

K.Y. Chen and C.H. Wang, (2007). A Hybrid SARIMA and support vector machines in forecasting the production values of the machinery industry in Taiwan, *Expert Systems with Applications*, 32 (1), 254–264.

Kang S., (1991). An Investigation of the Used of Feedforward Neural Networks for Forecasting. Phd Thesis. Kent State University, 1991.

Kapetanios, G. and Marcellino, M., (2010). Factor-GMM estimation with large sets of possibly weak instruments. *Computational Statistics and Data Analysis,* 54, 2655–2675.

Kapetanios, G., and Marcellino, M., (2009). A parametric estimation method for dynamic factor models of large dimensions. *Journal of Time Series Analysis*, 30(2), 208–238.

Kilian, L. and Taylor, M. P., (2003). Why is it so difficult to beat the random walk forecast of exchange rates?, *Journal of International Economics*, 60, 85–107.

Lu J.-C., Niu D.-X. and Jia Z.-L., (2004). A Study of Short Term Load Forecasting Based on ARIMA-ANN, *Proceedings of the Third International Conference on Machine Learning and Cybernetics*, Shanghai, 26-29.

M. Khashei and M. Bijari (2012). A new class of hybrid models for time series forecasting, *Expert Systems with Applications*, 39, 4344–4357.

M. Perez, (2006). Artificial neural networks and bankruptcy forecasting: a state of the art, *Neural Computing and Applications*, 15(2), 154–163.

M.W. Craven and J.W. Shavlik, (1997). Using neural networks for data mining, *Future Generation Computer Systems*, 13(2), 211–229.

Ma L., and Khorasani K., (2004). New Training Strategies for Constructive Neural Networks with Application to Regression Problem, *Neural Network*, 17, 589-609.

Marcellino, M., J. Stock and M. Watson (2003). Macroeconomic forecasting in the euro area: country-specific versus area-wide information, *European Economic Review*, 47 ,1-18.

Marcellino, M., J.H. Stock, and M.W. Watson, (2006). A Comparison of Direct and Iterated Multistep AR Methods for Forecasting Macroeconomic Series, *Journal of Econometrics*, 135, 499-526.

Matheson, T. D. (2006). Factor model forecasts for New Zealand, *International Journal of Central Banking*, 2, 169-237.

McAdam, P. and A. J. Hughes Hallett (1999). Non Linearity, Computational Complexity and Macro Economic modeling. *Journal of Economic Surveys,* 13(5), 577-618.

Mehdi Khashei and Mehdi Bijari, (2010). An artificial neural network (p, d,q) model for time series forecasting, *Expert Systems with Applications* 37, 479–489.

Naraidoo, R. and Raputsoane, L., (2010). Zone-targeting monetary policy preferences and financial market conditions: A flexible non-linear policy reaction function of the SARB monetary policy. *South African Journal of Economics,* 78, 400–417

Omidi, A., Nourani, E. and Jalili M., (2011). Forecasting stock prices using financial data mining and Neural Network. In the *Proceedings of 3rd IEEE International Conference on Computer Research and Development*, Shanghai, pp. 242-246.

Onatski A., (2009). Testing hypotheses about the number of factors in large factor models. *Econometric,* 77, 1447-1479.

Onatski A., (2010). Determining the number of factors form empirical distribution of eigenvalues. *The Review of Economics and Statistics*. 92, 1004-1016

Oscar Claveria and Salvador Torra, (2014). Forecasting tourism demand to Catalonia: Neural networks vs. time series models, *Economic Modelling*, 36, 220–228

P. G. Harrald and M. Kamstra., (1997). Evolving artificial neural network to combine financial forecasts, *IEEE Trans. Evolutionary Comput*. 1(1), 40-52.

Pai P. F. and Lim C.-S., (2005). A Hybrid ARIMA and Support Vector Machines Model in Stock Price Forecasting, *Omega*, 33, 497-505.

Palmer, A., Montaño, J.J. and Sesé, A., (2006). Designing an artificial neural network for forecasting tourism time-series, *Tourism Management,* 27, 781–790.

R. G. Donaldson and M. Kamstra., (1996). Forecast Combining with Neural Networks, *Journal of Forecasting.* 15, 49-61.

S. BuHamra, N. Smaoui and M. Gabr, (2003). The Box–Jenkins analysis and neural networks: prediction and time series modeling, *Applied Mathematical Modelling*, 27, 805–815.

S. K. Pal, V. Talwar, and P. Mitra, (2002). Web mining in soft computing framework: relevance, state of the art and future directions, *IEEE Transactions on Neural Networks*, 13(5), 1163–1177.

Sanjeev Kumar and D.K. Chaturvedi (2011). Prediction of Interest Rate using Generalized Neural Method, *International Journal of Computer Information Systems and Industrial Management Applications*, 3, 738-745.

Sargent, T. J., and Sims, C. A., (1977). Business cycle modeling without pretending to have too much a priori economic theory. In C. Sims (Ed.), *New methods in business research*. Federal Reserve Bank of Minneapolis.

Schneider, M., and Spitzer, M. (2004). Forecasting Austrian GDP using the generalized dynamic factor model. *Working Paper* No. 89. *Oesterreichische National Bank*.

Schumacher, C. (2007). Forecasting German GDP using alternative factor models based on large dataset, *Journal of Forecasting*, 26(4), 271–302.

Schumacher, C., and Dreger, C., (2004). Estimating large-scale factor models for economic activity in Germany: do they outperform simpler models? *Jahrbücher für Nationalökonomie und Statistik*, 224, 731–750.

Schumacher, C., and J. Breitung, (2008). Real-time forecasting of German GDP based on a large factor model with monthly and quarterly data, *International Journal of Forecasting*, 24, 386–398.

Serena Ng and Dalibor Stevanovic, (2012). Factor Augmented Autoregressive Distributed Lag Models, *working paper*, Columbia University, New-York.

Silvia Ferrari and Robert F. Stengel, (2005). Smooth Function Approximation Using Neural Networks. *IEEE Transactions on Neural Networks*, 16(1), 24-38.

Stock, J. H., and Watson, M. W., (1989). New indexes of coincident and leading economic indicators. *NBER Macroeconomics Annual*, 351–393.

Stock, J. H., and Watson, M. W., (1991). A probability model of the coincident indicators. In K. Lahiri, & G. Moore (Eds.), *Leading economic indicators: new approaches and forecasting record* (pp. 63–95). Cambridge: Cambridge University Press.

Stock, J. H., and Watson, M. W., (1999). Forecasting inflation. *Journal of Monetary Economics*, 44, 293–335.

Stock J, and Watson M., (1999). A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. Ch. 1 in *Cointegration, Causality and Forecasting*: A Festschrift for Clive W.J. Granger, Engle R, White H (eds); Oxford University Press: Oxford, 1-44.

Stock J. and Watson M., (2002a). Macroeconomic forecasting using diffusion indexes. *Journal of Business and Economic Statistics,* 20, 147-162.

Stock J. and Watson M., (2002b). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association,* 97, 1167-1179.

Stock, J.H., and M.W. Watson (2005). Implications of Dynamic Factor Models for VAR Analysis, Manuscript, *Princeton University*.

Swanson, Norman R. and Halbert White, (1997). A Model Selection Approach to Real-Time Macroeconomic Forecasting Using Linear Models and Artificial Neural Networks. *The Review of Economics and Statistics*, 79(4), 540–550.

Taskaya-Termizel T. and Ahmad K., (2005). Are ARIMA Neural Network Hybrids Better than Single Model? *Proceeding of International Joint Conference on NNs*, 4, Montreal, Canada.

Teräsvirta, T., (2006). Forecasting Economic Variables with Nonlinear Models. In Graham Elliott, Clive W.J. Granger and Allan Timmermann (eds.), *Handbook of Economic Forecasting*. Elsevier, 413- 457.

Timo Teräsvirta, Dick van Dijk and Marcelo C. Medeiros, (2005). Linear models, smooth transition autoregressions and neural networks for forecasting macroeconomic time series: A re-examination, *International Journal of Forecasting*, 21(4), 755–774.

Tseng F.-M. Y. H.-C. and Tzeng. G.-H., (2002). Combining Neural Network Model With Seasonal Time Series ARIMA Model, *Technological Forecasting and Social Change*, 69, 71-87.

V. Kondratenko and Yu. A Kuperin, (2003). Using Recurrent Neural Networks To Forecasting of Forex, Condensed Matter, *Statistical Finance.*

van Nieuwenhuyze, C., (2005). A generalized dynamic factor model for the Belgian economy: identification of the business cycle and GDP growth forecasts. *Journal of Business Cycle Measurement and Analysis*, 2, 213–248.

Van Wyk de Vries, E., Gupta, R., and van Eyden, R.,. Intertemporal Portfolio Allocation and Hedging Demand: An Application to South Africa, *forthcoming in Journal of Business Economics and Management*.

Venter, J.C., (2011). Business cycles in South Africa during the period 2007 to 2009. *South African Reserve Bank Quarterly Bulletins Articles and Notes*. 1–6.

Vincenzo Pacelli, Vitoantonio Bevilacqua and Michele Azzollini, (2011). An Artificial Neural Network Model to Forecast Exchange Rates, *Journal of Intelligent Learning Systems and Applications*, 3, 57-69.

White, H., (1990). Connectionist nonparametric regression multilayer feed forward networks can learn arbitrary mappings. *Neural Network.* 3, 535-49.

Winkler, R. L., and Makridakis, S., (1983). The combination of forecasts. *Journal of the Royal Statistical Society*, 146, 150–157.

Y.S. Lee and L.I. Tong, (2011). Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming, *Knowledge-Based Systems* 24, 66–72.

Young B. Moon and Rich Janowski, (1995). A neural network approach for smoothing and categorizing noisy data, *Computers in industry*, 26, 23-39.

Yu, L., Wang, S., and Lai, K. K., (2005). A novel nonlinear ensemble forecasting model incorporating GLAR and ANN for foreign exchange rates. *Computers and Operations Research*, 32, 2523–2541.

Zhang, G., Patuwo, B. E., and Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. *International Journal of Forecasting*, 14, 35–62.

Zhang, G. P., (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175.

Zhang, G. P., and Qi, G. M., (2005). Neural network forecasting for seasonal and trend time series. *European Journal of Operational Research*, 160, 501–514.

Appendix: Dataset and the Transformations

series code	Trans**	Series description
S1	Δ ln	South African Reserve Bank/Reserve Bank (SARB) Assets: Gold Coin And Bullion
S2	Δ ln	South African Reserve Bank Assets : Total Gold And Other Foreign Reserves
S3	Δ	South African Reserve Bank Assets: Advances To Banking Institutions
S4	Δ ln	South African Reserve Bank Assets : Other Advances
S5	Δ ln	South African Reserve Bank Assets: Investments In Government Stock
S6	L	South African Reserve Bank Assets : Other Investments
S7	Δ ln	South African Reserve Bank Assets : Total Discounts, Advances And Investments
S8	Δ	South African Reserve Bank Assets : Other
S9	Δ ln	Corporation For Public Deposits : Assets : Treasury Bills
S10	L	Corporation For Public Deposits (CPD): Assets: Promissory Notes
S11	L	Corporation For Public Deposits : Assets : Other Assets
S12	L	Corporation For Public Deposits : Assets : Total Assets
S13	Δ ln	Assets of Banking Institutions: Banknotes And Subsidiary Coin
S14	Δ	Assets of Banking Institutions: Deposits With The SARB
S15	L	Assets of Banking Institutions: Total Central Bank Money And Gold
S16	Δ ln	Assets of Banking Institutions: Mortgage Advances
S17	Δ2ln	Assets of Banking Institutions: Bills And Acceptances Discounted
S18	L	Assets of Banking Institutions: Interbank And Intragroup Funding, Including NCDSs/PNS
S19	Δ ln	Assets of Banking Institutions: Foreign Currency, Loans And Advances
S20	Δ ln	Assets of Banking Institutions: Redeemable Preference Shares
S21	Δ ln	Assets of Banking Institutions: Total Deposits, Loans And Advances
S22	ln	Assets of Banking Institutions: Investments Other Than Shares: Government Stock
S23	Δ ln	Assets of Banking Institutions: Other Investments
S24	Δ ln	Assets of Banking Institutions: Investments And Bills Discounted: Shares
S25	Δ ln	Assets of Banking Institutions: Specific Provisions In Re- Spect of Investments And Bills
S26	Δ	Assets of Banking Institutions: Total Investments And Bills Discounted
S27	Δ ln	Assets of Banking Institutions: Non-Financial Assets
S28	L	Assets of Banking Institutions: Other Assets
S29	Δ ln	Banking Institutions: Assets: Negotiable Certificate of Deposits / Promissory Notes (NCDSs/PNS)
S30	Δ ln	Banking Institutions: Assets: Treasury Bills Discounted
S31	ln	Banking Institutions: Assets: other Bills Discounted Including Bankers' Acceptances
S32	Δ ln	Banking Institutions: Assets: Total Loans And Advances
S33	ln	Banking Institutions: Assets: Advances to Non-Residents

S34	Δ ln	Banking Institutions: Assets: Investments By The Government Sector In Short-Term Government Stock
S35	Δ ln	Banking Institutions: Assets: Investments By The Government Sector In Long-Term Government Stock
S36	Δ ln	Banking Institutions: Assets: Other Investments By The Government Sector
S37	Δ ln	Banking Institutions: Assets: Investments In Stock of Public Enterprises/Corporations By The Private Sector
S38	L	Banking Institutions: Assets: Other Investments By The Private Sector
S39	Δ	Banking Institutions: Assets: Investments By Non-Residents
S40	Δ ln	Banking Institutions: Assets: Advances to The Provincial Governments
S41	Δ ln	Banks: Liquid Assets: Bank Notes And Subsidiary Coin
S42	Δ ln	Banks: Liquid Assets: Reserve And Clearing Account Held With SARB
S43	Δ ln	Banks: Liquid Assets: Treasury Bills
S44	Δ	Banks: Liquid Assets: Government Stock
S45	Δ ln	Banks: Liquid Assets: Total Holdings
S46	L	Banks: Liquid Assets: Required Holdings
S47	Δ ln	Assets Of Land And Agricultural Bank / Land Bank: Cash Credit Advances - Individuals
S48	Δ ln	Assets Of Land And Agricultural Bank / Land Bank: Cash Credit Advances - Co-Operatives
S49	Δ ln	Assets Of Land And Agricultural Bank / Land Bank: Cash Credit Advances - Total
S50	Δ ln	Assets Of Land And Agricultural Bank / Land Bank: Mortgage Loans Of Individuals
S51	Δ ln	Assets Of Land And Agricultural Bank / Land Bank: Mortgage Loans Of Co-Operatives
S52	Δ ln	Assets Of Land And Agricultural Bank / Land Bank: Other Loans To Individuals
S53	Δ ln	Assets Of Land And Agricultural Bank / Land Bank: Total Long-Term Loans And Advances
S54	Δ ln	Monetary Sector Assets: Long-Term Foreign Assets
S55	Δ ln	Monetary Sector Assets: Claims of The South African Reserve Bank / Reserve Bank (SARB) On The Private Sector
S56	Δ ln	Monetary Sector Assets: Claims of other Monetary Institutions on the Private Sector
S57	Δ ln	Monetary Sector Assets: Other Banks' Gold & Foreign Exchange (Excluding SARB And Government)
S58	Δ ln	Monetary Sector Assets: Claims on the Government Sector - South African Reserve Bank / Reserve Bank (SARB)
S59	Δ ln	Monetary Sector Assets: Claims on the Government Sector - Corporation For Public Deposits (CPD)
S60	L	Monetary Sector Assets: Claims on The Government Sector - Other Monetary Institutions
S61	L	Monetary Sector Assets: Claims On The Government Sector - Total Credit
S62	L	Monetary Sector Assets: Gross Claims On The Government Sector
S63	L	Monetary Sector Assets: Gross Claims On The Government Sector
S64	L	Monetary Sector Assets: Gross Claims On The Government Sector
S65	L	Monetary Counterparts : Cumulative Flow of Net Other Foreign Assets After Valuation Adjustment
S66	L	Monetary Counterparts : Cumulative Flow of Net Other Foreign Assets After Valuation Adjustment
S67	L	Monetary Counterparts : Cumulative Flow of Net Other Foreign Assets After Valuation Adjustment

S68	L	Monetary Counterparts : Cumulative Flow of Net Other Foreign Assets After Valuation Adjustment
S69	L	Monetary Counterparts : Net Other Assets And Liabilities
S70	L	Monetary Counterparts : Net Other Assets And Liabilities
S71	Δ	Banking Institutions: Assets: Advances To The Central Government
S72	ln	Monetary Sector Assets: Total Gold And Foreign Exchange (Excluding Government)
S73	Δ ln	Monetary Sector Assets: Total Foreign Assets
S74	Δ ln	Liabilities Of Banking Institutions: Total Equity And Liabilities
S75	Δ ln	Monetary Aggregates / Money Supply: M1(A)
S76	Δ ln	Monetary Aggregates / Money Supply: M1
S77	L	Monetary Aggregates / Money Supply: M2
S78	Δ ln	Monetary Aggregates / Money Supply: M3
S79	Δ ln	Secondary Market - Stock Exchange Transactions: Total Volume Of Shares Traded On The JSE
S80	Δ ln	Derivative Market (SAFEX) - Financial Futures Contracts: Open Interest
S81	Δ ln	Derivative Market (SAFEX): Financial Options On Futures Contracts - Open Interest
S82	Δ ln	Total Domestic Marketable Stock Debt Of National Government
S83	Δ2ln	National Government Revenue: Taxes on Income, Profits And Capital Gains: Income Tax
S84	Δ ln	National Government Tax Revenue: Total Other Taxes on Income, Profits And Capital Gains
S85	Δ ln	National Government Tax Revenue: Total Taxes on Income, Profits And Capital Gains
S86	Δ ln	National Government Tax Revenue: Taxes On Property: Transfer Duties
S87	Δ ln	National Government Tax Revenue: Other Taxes on Property
S88	Δ ln	Revenue: Total National Government Revenue
S89	Δ ln	National Government: Total Expenditure
S90	L	National Government Expenditure Adjusted For Cash Flows
S91	L	Gross Gold and Other Foreign Reserve
S92	Δ ln	Manufacturing: Orders And Sales: Sales
S93	Δ ln	Indicators of Real Economic Activity: Trade: Retail Sales
S94	Δ ln	Indicators of Real Economic Activity: Trade: Wholesale Sales
S95	Δ ln	Indicators of Real Economic Activity: Electric Current Generated
S96	Δ ln	Liabilities of Banking Institutions: Share Capital And Reserves
S97	Δ ln	Prime Overdraft Rate
S98	Δ ln	Brent Crude Oil Price In US Dollar
S99	Δ ln	London Gold Price In Rand
S100	Δ ln	Producer Prices Of Domestic Output: Agriculture, Forestry, Fishing And Mining (PPI)
S101	Δ ln	Producer Prices Of Domestic Output: Agriculture, Forestry And Fishing (PPI)
S102	Δ ln	Producer Prices Of Domestic Output: Mining And Quarry (PPI)
S103	L	Producer Prices Of Domestic Output: Food Manufacturing (PPI)
S104	L	Producer Prices Of Domestic Output: Paper And Paper Products Manufacturing (PPI)
S105	L	Producer Prices Of Domestic Output: Products of Petroleum And Coal Manufacturing (PPI)

S106	L	Producer Prices of Domestic Output: Chemical And Chemical Products Manufacturing (PPI)
S107	L	Producer Prices of Domestic Output: Basic Metals Manufacturing (PPI)
S108	L	Producer Prices of Domestic Output: Transport Manufacturing (PPI)
S109	L	Producer Prices of Domestic Output: Total Manufacturing (PPI)
S110	L	Producer Prices of Domestic Output: Electricity, Water, Steam And Gas (PPI)
S111	L	Total Producer Prices of Domestic Output (PPI)
S112	L	Total Producer Prices of Imported Commodities (PPI)
S113	Δ ln	Secondary Market - Stock Exchange Transactions: Total Value (Turnover) Of Shares Traded On The JSE
S114	L	Primary Market - Share Capital Raised By Companies on The JSE: Total Value of Share Capital Raised
S115	ln	Primary Market - Share Capital Raised By Companies on The JSE: Rights Issues of Ordinary Shares
S116	L	Primary Market - Share Capital Raised By Companies on the JSE: Other Share Capital
S117	ln	Net Purchases Of Shares By Non-Residents on the Johannesburg Stock Exchange (JSE)
S118	L	Purchases of Shares By Non-Residents on The JSE
S119	L	Sales of Shares by Non-Residents on the JSE
S120	Δ ln	Total Net Purchases of Shares and Bonds (Repo And Outright) By Non-Residents
S121	Δ ln	Weighted Average Rate: Fixed Deposits With Original Maturity of More Than 1 Year But Less Than 3 Years
S122	ln	Weighted Average Rate: Fixed Deposits With Original Maturity of 3 Years And More But Less Than 5 Years
S123	ln	Predominant Rate on Deposits: Postbank Investment Accounts
S124	L	Predominant Rate: Participation Bond Schemes
S125	L	Predominant Rate on New Mortgage Loans: Banks - Dwelling Units (Home Mortgage Rate)
S126	L	Predominant Rate on New Mortgage Loans: Participation Bond Schemes
S127	L	Foreign Exchange Rate : SA Cent Per Japanese Yen Middle Rates (R1 = 100 Cents)
S128	Δ ln	Foreign Exchange Rate: SA Cent Per China Yuan Middle Rate (R1 = 100 Cents)
S129	L	Foreign Exchange Rate: SA Cent Per Hong Kong Dollar Middle Rate (R1 = 100 Cents)
S130	ln	Foreign Exchange Rate: SA Cent Per India Rupee Middle Rate (R1 = 100 Cents)
S131	Δ ln	Foreign Exchange Rate : SA Cent Per UK Pound Middle Rates (R1 = 100 Cents)
S132	Δ ln	Foreign Exchange Rate : SA Cent Per USA Dollar Middle Rates (R1 = 100 Cents)
S133	Δ ln	Rate on 3-Month Trade Financing : UK
S134	Δ	Rate on 3-Month Trade Financing : US
S135	Δ	Rate on 3-Month Trade Financing : South Africa
S136	Δ ln	Nominal Effective Exchange Rate of the Rand: Average fort the Period - 15 Trading Partners
S137	Δ ln	Real Effective Exchange Rate of The Rand: Average For The Period - 15 Trading Partners - Trade In Manufactured Goods
S138	Δ ln	Yield on Loan Stock Traded on The Stock Exchange: Government Bonds - 3 To 5 Years
S139	Δ ln	Yield on Loan Stock Traded on The Stock Exchange: Government Bonds - 5 To 10 Years
S140	Δ	Yield on Loan Stock Traded on The Stock Exchange: Government Bonds - 10 Years And Over

S141	Δ ln	Yield on Loan Stock Traded on The Stock Exchange: Eskom Bonds
S142	Δ ln	Primary Market - Net Issues of Marketable Public-Sector Bonds: Government
S143	Δ ln	Primary Market - Net Issues of Marketable Public-Sector Bonds: Local Governments
S144	Δ ln	Primary Market - Net Issues of Marketable Public-Sector Bonds: Public Enterprises
S145	Δ ln	Primary Market - Net Issues of Marketable Public-Sector Bonds: Other Borrowers
S146	L	Primary Market - Total Net Issues of Marketable Public-Sec- Tor Bonds
S147	Δ	Secondary Market - Stock Exchange Transactions - Total Nominal Value of Bonds Traded On BESA
S148	Δ ln	Net Purchases of Bonds By Non-Residents on The Bond Exchange of South Africa (BESA)
S149	Δ ln	Purchases of Bonds By Non-Residents on The Bond Exchange of South Africa
S150	L	Sales of Bonds By Non-Residents on The Bond Exchange of South Africa
S151	L	National Government Financing By Government Bonds
S152	L	National Government Financing By Foreign Bonds And Loans
S153	Δ	Discount / Premium on Government Bonds
S154	Δ	Total Bonds of National Government
S155	Δ	Marketable Domestic National Government Bonds: Maturity Intervals Not Exceeding 1 Year
S156	Δ	Marketable Domestic National Government Bonds: Maturity Intervals Exceeding 1 But Not 3 Years
S157	Δ	Marketable Domestic National Government Bonds: Maturity Intervals Exceeding 3 But Not 10 Years
S158	Δ	Marketable Domestic National Government Bonds: Maturity Intervals Exceeding 10 Years
S159	Δ	Marketable Domestic National Government Bonds: Average Maturity (Months)
S160	Δ	Marketable Foreign National Government Bonds: Maturity Intervals Exceeding 1 Year But Not 3 Years
S161	Δ	Marketable Foreign National Government Bonds: Maturity Intervals Exceeding 3 Years
S162	L	Marketable Foreign National Government Bonds: Average Maturity (Months)
S163	Δ	Ownership Distribution of Domestic Marketable Bonds: Short Term: Banks
S164	Δ	Domestic Marketable Long Term National Government Bonds Held By Public Investment Corporation (PIC)
S165	Δ	Banks Total Equity
S166	Δ	United States - Federal Funds Rate
S167	Δ	South Africa - Money Market Rate
S168	Δ	South Africa - Deposit Rate
S169	Δ	South Africa - Lending Rate
S170	Δ	South Africa - Discount Rate
S171	Δ	South Africa - Principal Rate, End of Period
S172	Δ	China,P.R.: Hong Kong - Market Rate, End Of Period
S173	L	ABSA House Price Index - All Sizes - Purchase Price - Smoothed
S174	L	Share Market - Number of Shares Traded (Millions)
S175	Δ ln	United States - T-Bill Rate-3 Month
S176	L	United States - Bank Prime Loan Rate
S177	Δ ln	United States - Market Rate, End of Period

S178	Δ ln	United States – NEER From Ins
S179	Δ ln	United States - REER Based on Rel. Cp
S180	L	United Kingdom - Treasury Bill Rate Bond
S181	L	United Kingdom - Lending Rate: Clearing Banks
S182	Δ ln	United Kingdom - Market Rate, End of Period
S183	L	United Kingdom - NEER From Ins
S184	L	United Kingdom - REER Based on Rel. Cp
S185	L	S&P 500 Composite Price Index
S186	Δ	South Africa - Treasury Bill Rate
S187	Δ	South Africa - Share Prices: Indust & Comm
S188	Δ	South Africa - Share Prices: Gold Mining
S189	Δ	South Africa - Share Prices: All Shares
S190	L	South Africa Inflation Rate
S191	L	South Africa - Consumer Price Index
S192	Δ	NASDAQ Open
S193	L	Yield On Loan Stock Traded on The Stock Exchange: Government Bonds - 0 To 3 Years
S194	L	National Government Domestic Financing By Treasury Bills
S195	Δ	Total Financing of National Government
S196	L	National Government: Other Financing
S197	L	Financing of The National Government Deficit/Use of Surplus: Change In Net Indebtedness To The PIC
S198	L	Financing of National Government Deficit/Use of Surplus: Change In Debt Instruments Held By The Monetary Sector
S199	Δ	Index of Industrial Production
S200	L	Short-Term Interest Rates, Per Cent Per Annum
S201	L	Immediate Interest Rates, Call Money, Interbank Rate, Per Cent Per Annum
S202	Δ	Long-Term Interest Rates, Per Cent Per Annum
S203	Δ	Dow Jones U.S. Total Stock Market Total Return Index
S204	Δ ln	UK FTSE All-Share Index (W/GFD Extension)
S205	Δ	UK FTSE All-Share Return Index
S206	Δ	Johannesburg Se Return Index
S207	L	South Africa Business Confidence Index
S208	L	Johannesburg Se Dividend Yield
S209	L	Government Bond Yields (SA)
S210	Δ ln	FTSE/JSE All-Share Index (W/GFD Extension)
S211	Δ	South Africa 3-Month Bills Total Return Index
S212	Δ ln	Total Return Indices - Bills (USA)
S213	Δ	UK 3-Month Treasury Bill Yield
S214	Δ	Total Return Indices - Stocks
S215	L	Total Return Indices - Bonds (SARB Government Bond Return Index)
S216	Δ	Japan Nikkei 500 Index

S217	Δ	Japan Nikko Tokyo Se Performance Index (Total Return Indices - Stocks)
S218	L	South Africa Consumer Confidence Index
S219	Δ ln	Real Monetary Aggregates / Money Supply: M1
S220	Δ ln	Real Monetary Aggregates / Money Supply: M2
S221	Δ ln	Real Monetary Aggregates / Money Supply: M3
S222	Δ ln	Real ABSA House Price Index - All Sizes - Purchase Price (Smoothed Rand)
S223	Δ	Real Yield on Loan Stock Traded on The Stock Exchange: Government Bonds - 0 To 3 Years
S224	Δ	Real Yield on Loan Stock Traded on The Stock Exchange: Government Bonds - 3 To 5 Years
S225	L	Real Yield on Loan Stock Traded on The Stock Exchange: Government Bonds - 5 To 10 Years
S226	Δ	Real Yield on Loan Stock Traded on The Stock Exchange: Government Bonds - 10 Years and Over
S227	Δ ln	Real London Gold Price In Rand
S228	L	Government Bond Yield 10 Years and Over - Treasury Bill Rate

** Transformations code

L Level or no transformation

ln Logged

Δ First differenced

Δ ln Logged and first differenced

Δ^2ln Logged and second differenced

Δ^2 Second differenced