Forecasting Financial Variables using Artificial Neural Networks – Dynamic Factor Model

Ali Babikir*

University of KwaZulu-Natal

Henry Mwambi#

University of KwaZulu-Natal

Abstract

In this paper 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 dataset, 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 African variables, namely, Rate on three-month trade financing, Lending rate and Short-term interest rate in the period 1992:1 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 2007:1 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 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 proposed by Diebold-Mariano.

---

*Dr A Babikir is a postdoctoral fellow in the School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Pietermaritzburg, South Africa.

*Prof H Mwambi is a professor in the School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Pietermaritzburg, South Africa. [MwambiH@ukzn.ac.za]

1. INTRODUCTION

The need to obtain accurate forecasts has been a strong motivation for empirical research. Efforts have been made 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 summarise the information contained in large datasets. The use of dynamic factor models has been further improved by 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 on a subspace algorithm.

The goal of the techniques is to allow forecasters to easily summarise the information contained in large datasets and extract a few common factors that are useful in forecasting exercises. The small number of 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, Boivin, and Eliasz (2005) propose a forecasting model that they called the factor-augmented vector autoregressive (FAVAR) model, a model that 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) and the forecasting performance of the model is 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; Craven & Shavlik, 1997) and pattern recognition (Bishop, 1995) and smoothing the data or parameters (Moon & Janowski, 1995; Hill & Goring, 1998; Ferrari & Stengel, 2005; Yang & Wu, 2012). The good results of the ANNs in all of above-mentioned areas are based on the unique properties and features of the method, including the following: firstly, ANNs are universal functional approximators, and can approximate any continuous function to any desired accuracy. Secondly, ANNs are data-driven self-adaptive methods in that there are few a prior assumptions to be made about the models for problems under study; thus ANN modelling is different from traditional model-based methods. Thirdly, 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; for example the recent financial crisis showed clear evidence of the downturns of the economies and financial markets around the world. Beside such phenomena, there are 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 or nonlinear model might miss some features of time series data that may contain both components. 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 paper is to use the dynamic factor framework for forecasting. We introduce a model that utilises the generalised dynamic factor model proposed by Forni, Hallin, Lippi and Reichlin (2000, 2003 and 2005). The factors of the new model are extracted from ANNs in sample forecasts to each single series of the dataset, which contains 228 series (full details of the dataset are provided in Section 4).

In this stage the ANNs are used as smoothers or approximators, because their strength has been approved in this regard. The paper is organised as follows. Section 2 gives a brief description of the generalised dynamic factor model and how to determine the number of factors. In Section 3, we briefly review the DFM, ANN and introduce the ANN-DFM modelling approaches to time series forecasting. The dataset is described in Section 4. In Section 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 6 draws some conclusions.

2. METHODOLOGY

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

2.1 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 a zero mean. The idea behind the factor model is that most of the variance of the dataset 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

$$X_t = \chi_t + \xi_t = \lambda(L)'f_t + \xi_t \qquad (1)$$

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 $f_t = (f_{1t} \dots f_{qt})'$, the operator $\lambda(L) = 1 + \lambda_1 L + \dots + \lambda_s L^s$ 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:

$$X_t = \Lambda' F_t + \xi_t \qquad (2)$$

where F_t is a vector of $r \geq q$ static factors that comprise 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 dataset. These methods were developed by Stock and Watson (2002a; hereafter SW), Kapetanios and Marcellino (2009) and Forni, Hallin, Lippi and Reichlin (2005, hereafter FHLR) (for further technical details on this type of factor models, see Schumacher, 2007). In the current paper 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 $\widehat{\Sigma}$.

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 $\widehat{\Sigma}_\xi(\omega) = \widehat{\Sigma}(\omega) - \widehat{\Sigma}_\chi(\omega)$. Inverse Fourier transform provides the time-domain autocovariances of the common and the idiosyncratic components given by $\widehat{\Gamma}_\chi(k)$ and $\widehat{\Gamma}_\xi(k)$ 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 maximise the contemporaneous covariance explained by the common factors $\widehat{Z}_i'\widehat{\Gamma}_\chi(0)\widehat{Z}_i$, $i = 1, \dots, r$. This optimisation problem is subject to the normalisation $\widehat{Z}_i'\widehat{\Gamma}_\xi(0)\widehat{Z}_j = 1$ for $i = j$ and zero otherwise. This representation can be reformulated as the generalised eigenvalue problem such that $\widehat{\Gamma}_\chi(0)\widehat{Z}_i = \widehat{\mu}_i\widehat{\Gamma}_\xi(0)\widehat{Z}_i$, where $\widehat{\mu}_i$ denotes the i-th generalised eigenvalue and $\widehat{Z}_i$ its $N \times 1$ corresponding eigenvector in their non-null spaces. The factor estimates according to FHLR are then obtained as $\widehat{F}_t = \widehat{Z}'X_t$ with $\widehat{Z} = [\widehat{Z}_1 \dots \widehat{Z}_r]$.

2.2 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 & Ng, 2002; Alessi et al., 2008) and the dynamic factor model (Bai & Ng, 2007; Amengual & Watson, 2007; Hallin & Liska, 2007; Onatski, 2009, 2010). To specify the number of static factors, Bai and Ng (2002) and Alessi et al. (2008) use an information criterion, based on AIC and BIC, to help guide the selection of the optimal number of factors r in a large dataset. 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. FORECASTING MODELS

In this section, the basic concepts and modelling 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.

3.1 Dynamic factor model forecast

Based on the subsection 2.2 five static factors and two dynamic factors are extracted from the entire data panel that together explain more than 87% of variation of the data. These estimated factors are then 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 Forni et al. (2003), which takes the form:

$$y_{t+h}^h = \beta(L)\widehat{F}_t + \gamma(L)y_t + u_{t+h}^h \qquad (3)$$

where $\widehat{F}_t$ 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.

3.2 Autoregressive (AR) forecast

The AR model is given by

$$y_t = \phi + \gamma(L)y_t + e_t \qquad (4)$$

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 e_t is the error term.

The h-step ahead forecast from this model is

$$y_{t+h}^h = \phi + \gamma^h(L)y_t + e_{t+h}^h \qquad (5)$$

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

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

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.

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. The ANN model can offer a valid approximation to the generating mechanism of a vast class of non-linear processes see for example, (Hornik, Stinchcombe & White, 1989; Swanson & White, 1997; 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 (for more details about the strengths and drawbacks of ANN, see Ramlall, 2010).

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; Zhang & Min Qi, 2005). The network used in this paper 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:

$$y_{t+h} = \alpha_0 + \sum_{j=1}^{n} w_j \, g\left(\alpha_{0,j} + \sum_{i=1}^{p} w_{i,j} \, y_{t-i}\right) + \sum_{i=1}^{p} \beta_i \, y_{t-i} + \varepsilon_{t+h} \qquad \textbf{(6)}$$

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, and 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 6 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 of this model is the combination of the pure linear model and feed-forward neural network. Therefore, if the relationship between inputs and output is purely 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 & Wright, 2006). The selection of the lag lengths and the number of nodes in the hidden layer is chosen on the basis of the training set or the in-sample RMSE.
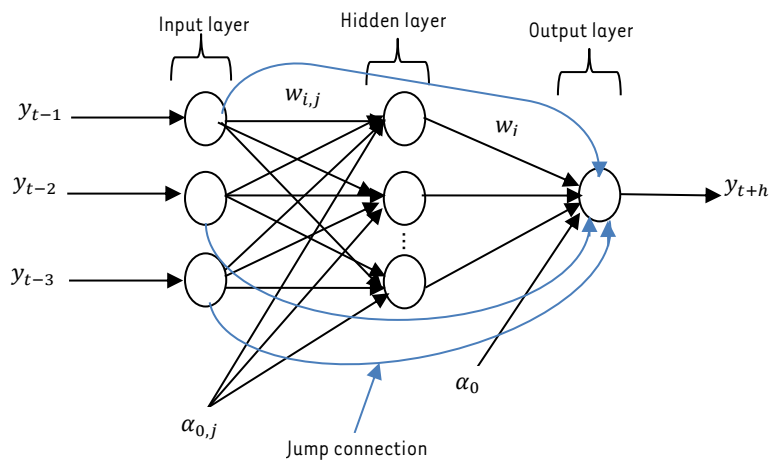

FIGURE 1:   **Structure of the network, N$^{(3,3,1)}$**

3.4 Formulation of the ANN-DFM model

There are a numerous time series models available, but still the accuracy of time series forecasting currently is fundamental to many a decision maker. Much research in time series forecasting has argued that predictive performance improves in combined models, as this combination reduces the risk of failure that can occur as a result of difficulty in determining 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. In this step, 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 are then used to forecast the variable of interest. Note that in this step three static factors and two dynamic factors are extracted - based on subsection 2.2 - which explain more than 82% of variation of the entire dataset. In the third step, the extracted factors are used as explanatory variables in equation 3 to produce the forecasts of the variable of interest.

4. DATA

Our dataset consists of 228 monthly time series from January 1992 to December 2011, or 239 observations for each variable. The data sources are the South Africa Reserve Bank, ABSA Bank, Stats South Africa, National Association of Automobile Manufacturers of South Africa (NAAMSA), South African Revenue Service (SARS), Quantec and World Bank.

Among these 228 series 203 series are from South Africa, covering the financial, real, nominal sectors and confidence indices, two are global variables and 23 series are from major trading partners and global financial markets. Thus besides the national variables, the paper uses a set of global variables such as gold and crude oil prices. In addition the data also includes series from the 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 three, six and 12- month forecasting horizons for three variables Short-term interest rate, Lending rate and Rate on three 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 standardised to have a mean of zero and a constant variance.

5. EMPIRICAL RESULTS

5.1 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 1 below reports the in-

sample forecasting results. The first row reports the RMSE 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 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 $t + n$.

The model with a lowest RMSE ratio is deemed to perform better than the other models. In our case the ANN-DFM outperformed all other models and for all variables followed by the DFM. These results prove the superiority of DF-driven models, the reason possibly being that the DF models can efficiently handle large amounts of information, including the external variables that influence the South African economic and financial sector, therefore helping improve the forecasting performance.

TABLE 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

Source: Authors' analysis

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.

5.2 Out-of-sample results

In this subsection we evaluate the accuracy of the forecasts generated by the ANN-DFM. We compare its performance with the DFM and the AR benchmark model using the RMSE. We compare each of the three, six and 12 months-ahead forecasts generated by the ANN-DFM with DFM and AR over the out-of-sample horizon of 2007:1 to 2011:12. Note that the out-of-sample period includes the financial crisis that affected the South African economy in 2009. Thus, a good forecasting model can be used as an alternative to predict such a crisis. TABLE 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 that in this paper we choose iterated forecast instead of direct forecast; in other words, the forecasts are constructed recursively, using all available data to estimate parameters. The main observations can be summarised as follows:

- *Rate on 3-month trade financing*: from TABLE 1 we observe that the ANN-DFM model outperforms the AR benchmark model with an average reduction of 20% in the RMSE for all horizons. On the other hand, the DFM outperforms the AR benchmark model with an average reduction of 10% in the RMSE for all horizons. These results show that the ANN-DFM outperforms the DFM with an average reduction of 10% 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% and 19%. The DFM also beat the AR benchmark model with a reduction in RMSE of around 6% to 11%.

- *Short-term interest rate*: from TABLE 2 we see that the DFM outperforms the other models. Comparing the DFM to the ANN-DFM model we find that the DFM performs slightly better than the ANN-DFM model. Comparing both models to the AR benchmark model the reduction in the RMSE is around 14% to 22%. 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 only three factors.

When we consider the cross-model test of forecast accuracy that was proposed by Diebold and Mariano (1995), (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 indicates that the ANN-DFM model outperforms the other model in out-of-sample forecasting), TABLE 3 shows that in all cases where the ANN-DFM outperforms the AR benchmark model the statistics are significant at least at the 5% level. Regarding the cases where ANN-DFM outperforms DFM, the statistics are significant at the maximum of the 5% level, except one case where the statistic is significant at the 1% level. This is true for two series, namely, Rate on 3-month trade financing and Lending rate. On the other hand, when the DFM tends to outperform the ANN-DFM the test statistics are insignificant at the 1%, 5% and 10% levels, where their critical values are 2.326, 1.6449 and 1.2816 respectively.

TABLE 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

Source: Authors' analysis

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: 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***

Source: Authors' analysis

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

4. CONCLUSION

This paper introduced a new model where artificial neural networks (ANNs) are 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 12 months-ahead forecasts over a 60-month forecasting horizon covering 2007:01 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 improvements are minimal. In general, a data-rich factor-driven model is best suited for 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.

LIST OF REFERENCES

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

Amengual, D. & 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*, pp. 91-96.

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

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

Bai, J. & Ng, S. (2007). Determining the number of primitive shocks in factor models. *Journal of Business and Economic Statistics, 25*, pp. 52-60.

Banbura, M. & Rünstler, G. (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)*, pp. 333-346.

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

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

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

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

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

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

Dakun, Y & Wu, W. (2012). A Smoothing Interval Neural Network. *Discrete Dynamics in Nature and Society*, 2012.

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

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

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

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

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

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

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

Hill, J.T. & Goring, D.G. (1998). Smoothing hydrological data: a neural network approach. *Journal of Hydrology, 37(2),* pp. 81-94.

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

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

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

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

Nocedal, J. & Wright, S.J. (2006), *Numerical Optimization, 2nd ed.* Berlin, New York: Springer.

Omidi, A., Nourani, E. & 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. *Econometrica. 77*, pp. 1447-1479.

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

Önder, E., Bayir, F. & Hepsen, A. (2013). Forecasting Macroeconomic Variables using Artificial Neural Network and Traditional Smoothing Techniques, *Journal of Applied Finance & Banking*, *3(4),* pp. 73-104.

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

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

Ramlall, I. (2010). Artificial Intelligence: Neural Networks Simplified. *International Research Journal of Finance and Economics, 39*, pp. 105-120.

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

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

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

Silvia, F. & Stengel, R.F. (2005). Smooth Function Approximation Using Neural Networks. *IEEE Transactions on Neural Networks, 16(1)*, pp. 24-38.

Stock, J.H. & Watson, M.W. (2002a). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association, 97,* pp. 147-162.

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

Swanson, N.R. & White, H. (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),* pp. 540-550.

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

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