Forecasting the Indonesian Government Securities Yield Curve using Neural Networks and Vector Autoregressive Model

Dedi Rosadi
Department of Mathematics, Gadjah Mada University, Indonesia
Email: dedirosadi@ugm.ac.id

Yoga Aji Nugraha
Department of Mathematics, Gadjah Mada University, Indonesia
Email: yoga.nugraha09@gmail.com

Rahmawati Kusuma Dewi
Department of Mathematics, Gadjah Mada University, Indonesia
Email: de2_rahma@yahoo.com

1. Introduction

The term structure of interest rates is a relation of the yield and the maturity of default free zero-coupon securities and provides a measure of the returns that an investor might expect for different investment periods in a fixed income market. One potential candidate for a standard model of the yield curve is the Nelson and Siegel (1987) (hereafter NS) class of yield curve models. Models of this class are already very popular with practitioners, researchers, and academics in finance and economics (see e.g. Stander, 2005; and Rosadi, Gunardi, Abdurakhman, Utami, and Wulansari, 2008; for overview for an overview of various models of yield curve). However, in emphasizing simple relationships by maturity and the fit to observed yield curve data at a given point in time, NS model (and all static yield curve models) overlook the dynamics of interest rates over time. Hence, they are not inter-temporally consistent (i.e interest rates and yield curves at different points in time cannot be mapped to each other via an underlying stochastic time-series process). As the result, they fail to model the dynamic relationship between the parameters of a term-structure model and therefore, the forecasting ability of these models is poor. This has motivated us to apply the neural networks and compare its performance with VAR for the purposes of forecasting the yield curve.

The rest of this paper can be described as follows. In the next subsection we shortly describe the yield curve and nelson siegel model for the yield curve. In section three and four, we shortly overview the neural networks model and VAR methodologies that will be used for the forecasting. Section five describes the empirical results and last section concludes.

2. Yield Curve and Estimation

Here we introduce the fundamentals of the yield curve. If $P(t,T)$ denotes the price of zero coupon bond at time t with time to maturity T, often called as *discount function* and $R(t,T)$ denotes its continuously compounded zero-coupon nominal yield to maturity (YTM), also called spot rate, then the discount curve can be obtained from the yield curve as: $P(t,T) = e^{-(T-t)R(t,T)}$. Therefore, we obtain relation between *spot rate* R(t,T) with the price of *zero coupon bond* P(t,T) as $R(t,T) = -\ln P(t,T)/(T-t)$.

Let's denote f(t,T,T+Δt) be the f*orward rate* with the contract determined at time t with the settlement date at time T and maturity at time T+Δt. The *instantaneous forward rate* $f(t,T,T+\Delta t)_{inst}$ define as *forward rate* with Δt→0, therefore, *instantaneous forward rate* can be written as f(t,T). The relation between the

price of *zero coupon* bond with the *instantaneous forward rate* can be written as $P(t,T) = \exp(-\int_t^T f(t,u)du)$ or $f(t,T) = -\frac{d}{dt}\ln P(t,T)$ and therefore the relationship between the yield to maturity and the forward rate is therefore given as:

$$R(t,T) = \int_t^T f(t,u)du / (T-t)$$

This means that the zero-coupon yield is an equally weighted mean of the forward rates. It is therefore possible to price any coupon bond as the sum of the present values of future cash flows (coupon payments and the principal payment), if the yield curve or the forward-rate curve is given.

There are various models of yield curve available in literature, see e.g. Stander (2005) and Rosadi, et.al (2008) for overview. NS models the instantaneous forward-rate curve *f(t,m)* is given as (Nelson and Siegel, 1987)

$$f(t,m) = \beta_0 + \beta_1 \exp(-\frac{m}{\tau}) + \beta_2 \frac{m}{\tau}\exp(-\frac{m}{\tau})$$

This implies that Nelson-Siegel yield curve is given as :

$$R(t,m) = \frac{1}{m}\int_{x=0}^{m} f(x)dx = \beta_0 + \beta_1\left[\left[1-\exp(-\frac{m}{\tau})\right]/\frac{m}{\tau}\right] + \beta_2\left[\left[\left[1-\exp(-\frac{m}{\tau})\right]/\frac{m}{\tau}\right] - \exp(-\frac{m}{\tau})\right]$$

The rate of exponential decay is governed by the parameter τ and $\beta_0, \beta_1, \beta_2$ are interpreted as the short-term component, long-term component, and the medium-term component, respectively, and *m* denotes the maturity time. The parameters β_0 and τ need to be positive and $\beta_0 + \beta_1 > 0$.

The Nelson Siegel model can be estimated using constrained nonlinear regression optimization procedure based on YTM function (eq. (1)) above. There are two alternative estimation procedures available in literature, namely *Full estimation algorithm* and *Partial-estimation algorithm* (e.g., Bolder and Streliski, 1999). In *Partial-estimation algorithm*, we are first fixing the τ s and β s are estimated, or the other ways, fixing the β s and τs are estimated. For the full estimation algorithm, we may apply, for example, the sequential Quadratic Programming (SQP), Nelder-Mead Simplex method, or other constrained optimization methods (see e.g., Lange, 2010; Lagarias, Reedsz, Wrightx and Wright 1998). In this paper, we use SQP optimization method, available in SPSS 15. As an alternatif implementation, we may use function *nlminb* in R (implemented PORT optimisation method, see http://netlib.bell-labs.com/netlib/port/) or function *constrOptim* (implemented Nelder-Mead and several other optimization procedures). In Nugraha and Rosadi (2011), we apply the OLS method for the starting values of the parameters, and further use these starting values for Nelder Mead estimation of the parameters using R-CLI. The R-GUI implementation of the NS model is also available in package *RcmdrPlugin.Econometrics* (Rosadi, 2010).

Although various extension of the NS model has been proposed in literature, as noted in e.g., Diebold and Li (2006), there is no guarantee of better out-of-sample forecasting performance than NS model. Moreover, that parsimonious models are often more successful for out-of sample forecasting. For this reason, in this paper we rely our forecasting technique based on the simplest NS model above.

3. Forecasting Yield Curves

A good approximation to yield-curve dynamics should not only fit well in-sample, but also forecast well out-of-sample. Because the NS yield curve depends only on four parameters τ and $\beta_0, \beta_1, \beta_2$, forecasting the yield curve is equivalent to forecasting τ and $\beta_0, \beta_1, \beta_2$. In this section we outline the method of forecasting using Neural networks and Vector autoregression.

3.1. Neural Networks

In this paper, we apply the multilayer perceptron (MLP) containing the *input layer, one hidden layer* and the *output layer* where the estimation algorithm used is the backpropagation method. Several variation of backpropagation algorithms are available in the literature, such as backpropagation variation using momentum, using Adaptive Learning Rate & Momentum and using Numerical Optimization. For forecasting the yield curve, the NN algorithm is applied for four parameters of the NS model.

3.2. Vector Autoregressive (VAR)

The VAR(p) model with k endogeneous variables $\mathbf{y}_t = (y_{1t}, \cdots, y_{kt})$ can be written as

$$\mathbf{y}_t = \mathbf{A}_1 \mathbf{y}_{t-1} + \ldots + \mathbf{A}_p \mathbf{y}_{t-p} + CD_t + \mathbf{u}_t$$

where $\mathbf{A}_i, i = 1, \ldots, p$ is the coefficient matrix with dimension $(k \times k)$, $\mathbf{u}_t$ is white noise process with dimension k and the time invariant and positive definite matriks $E\mathbf{u}_t\mathbf{u}_t^{'} = \Sigma_u$. Matrix C is the coefficient matrix of m-independent random variables D_t with dimension $k \times m$, where the matrix D_t contains all possible random variables, such as constanta, trend components, dummy variables and/or seasonal dummy variables.

The parameters of the VAR model can be estimated using ordinary least square method, where the optimal order can be found using information criteria, such as AIC (Akaike Information Criteria), HQC (Hannan Quinn Criteria), SBC (Schwarz Bayesian Criteria).

In this paper, for forecasting the NS yield curve, we apply the VAR model for the endogenous variables $\tau, \beta_0, \beta_1, \beta_2$ of the NS model. Various econometrics models as the alternatives model for forecasting yield curve are discussed in e.g. Diebold and Li (2006) including random walk model, forward curve regression, Autoregressive model, error correction model, etc. See this literature for further detail.

4. Empirical results

4.1. Data

For practical application, we use the zero coupon bond price of Indonesia Government Bond from April 2008 until July 2008, which is the data during the global financial crisis in 2008.The BI-rate and 3-month SBI are used as the data with the smallest maturity time where we use the most recent observation available in each month. The other observations are the transaction price of bond (clean price) on the respective dates. Several series of bonds are considered as the important series and will be used as the anchor for curve estimation. If no transaction series are available for the “anchor” data on a particular day, we will use the quote price of HIMDASUN (where the historical data are available from Bloomberg).

Table 1. Important bond for yield curve estimation

TTM	1-month	3-months	1-Y	3-Y	5-Y	7-Y	10-Y	15-Y	20-Y	30-Y
Series	BI rate	SBI-3	FR0002	FR0016	FR0049	FR0027	FR0048	FR0046	FR0047	FR0050

From the raw data, using Excel 2007, we calculate the time to maturity (TTM) and yield to maturity (YTM) for each observation.

4.2. Estimating Nelson Siegel Parameter

In Nelson Siegel approach, there are four parameters that will be estimated (τ, β0, β1, β2) for each daily transaction during April 2 to July 31 2008. For estimating the parameter, we apply the full estimation algorithms based on SQP method, available on SPSS version 15.

It has been known that the NS model is sensitive to the method of estimation and particularly to the starting values used for the parameters. From our studies in Rosadi et.al (2008), for our data, we find that the optimal choice of starting values (at least locally optimal) of β_0 is the yields of BI rate and for β_1, β_2 and τ are equal to 1, and for the constrains, we consider β_0 is greater or equal than the yields of BI rate and with τ≥1.

4.3. Prediction using NN

For prediction of NS parameters, we use the training-data set (τ, β0, β1, β2) from April 2 until June 30, 2008 (40 data) and the rest of the data (τ, β0, β1, β2) are used as the testing-data set (15 data). We will apply MLP model with one hidden layer, and we consider of using sigmoid bipolar (tansig) activation function ($f(x) = (1-\exp(-x))/(1+\exp(-x))$. The initial values of parameters are generated using Nguyen –Widrow method.

The computation is implemented using neural networks toolbox in Matlab 7. The following parameterizations are used for computation

- net.trainParam.mc is used to determine the momentum factor used.
- Using 10000 maximum epoch, it means that the learning process will be stopped as soon as it reach 10000 epoch, that is written as net.trainParam.epochs=10000.
- It is determined that the goal target is 0.01, it means that the iteration will be stopped when the goal value ≤ 0.01, that is written as: net.trainParam.goal=0.01.
- Before training the input and output, we need to *preprocessing* data to ensure that the input and output are on the range of [-1,1], which will be required for applying tansig function.

The number of neurons in hidden layer

The data are trained using backpropagation function traingdx (Variable Learning Rate Backpropagation) and trainlm (backpropagation with Lavenberg-Marquardt algorithm) method in MATLAB 7. The number of neurons considered in the study is between one until 10. On the other hand, when applying the traingdx, we consider the momentum factor 0.1 until 0.9 with increment 0.1. For each combination of neurons number and momentum factor, we predict the parameters value and calculate the Mean Square Error (MSE) of the prediction, and further average the MSE of all parameters to obtain the MSE Average. The minimum value of MSE Average in the training data is obtained when the neurons number is 8 and the momentum factor 0.9 with MSE Average is smaller than the result using trainglm method. When we using the training method using backpropagation with function trainglm, we obtain the minimum MSE Average is obtained when the neurons number is 8.

We further apply the method on data testing. On the data testing, we found that the above choices of neurons and momentum rate performs quite well, although the minimum MSE Average on the data testing is found for neurons number 4 and momentum rate 0.6. To save the space, detail description is omitted.

4.4. Prediction using VAR

Using the data set of (τ, β0, β1, β2) from April 2 until June 30, 2008 (40 data), we try to fit VAR(p) using our R-GUI package for time series analysis that we called as *RcmdrPlugin.Econometrics* (Rosadi, 2010). We found that the best model for the data using AIC is VAR(2) model.

4.5. Prediction of Yield Curve using NN and VAR

Using the result from NN and VAR model above, we predict the future 21 days values of the parameters (τ, $\beta 0$, $\beta 1$, $\beta 2$) (i.e. can be considered as the data testing). Based on these predicted values of parameters, we calculate the predicted yields to maturity, and calculate the mean square error (MSE) of the prediction. We found that the NN methods will give smaller MSE than VAR based method, however detail is omitted. For illustrative purpose, we show the following two figures, which compare the prediction of Yield Curve using NN and VAR on NS model and the real data for two trading days, i.e. on July 7, 2008 and July 15, 2008. The prediction intervals are calculated using the predicted values using NN model. Based on these pictures (and other similar behaviour for the other days), we found that the accuracy of prediction based on NN methods will perform better for forecasting the yield of short term bond, where for the long term bonds, the VAR method show a little bit more accurate prediction.

5. Conclusion

The above results indicate that NN and VAR method are two methods that can be used to predict the behaviour of yield curve. However, it seems that the NS model in some cases can not model the behavior of yield data accurately, and therefore resulting an inaccurate prediction of future behaviour of the yield curve. For the future research, we suggest to consider more advanced model for yield curve models, such as the extended method by incorporating the financial and macroeconomic variables into NS model (see e.g. the application of this method for a similar purpose but different type of bond in Rosadi, Qoyyimi and Wulansari, 2010).

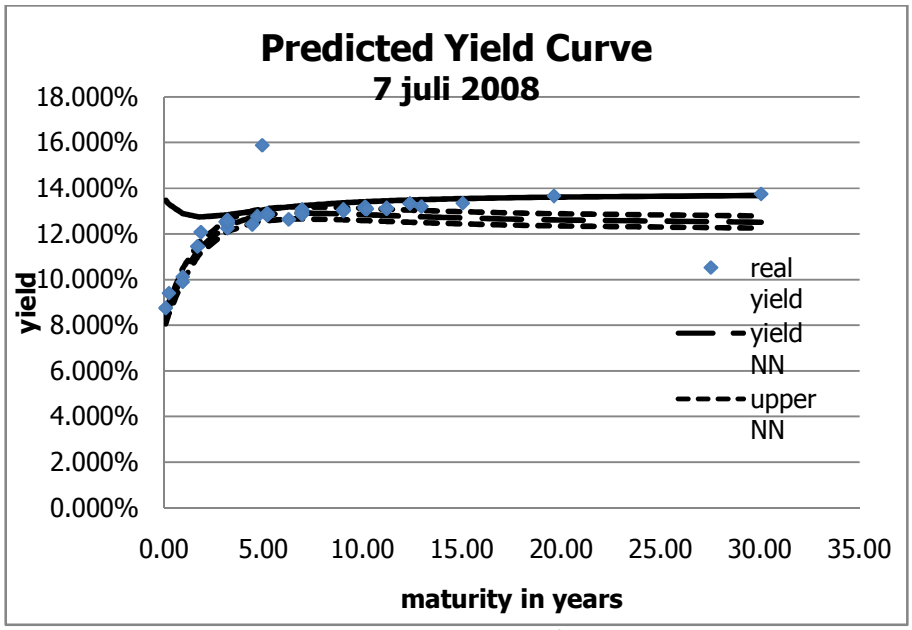

Figure 1. Predicted Yield Curve for July 7, 2008

REFERENCES

Bolder, D. and Stréliski, D., 1999, Yield Curve Modelling at the Bank of Canada, preprint

Diebold, F.X. and Li, C., 2006, Forecasting the term structure of government bond yields, Journal of Econometrics, 130, 337-364

Lagarias, J.C., Reedsz, J.A., Wrightx, M.A. and Wright, P.A., 1998, Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions, SIAM J. OPTIM. Vol. 9, No. 1, pp. 112-147

Lange, K., 2010, Numerical Analysis for Statisticians, Second Edition, Springer Verlag, New York

Nelson, C. R. and Siegel, A. F., 1987, Parsimonious modeling of yield curves, Journal of Business, 60, No. 4, 473-89

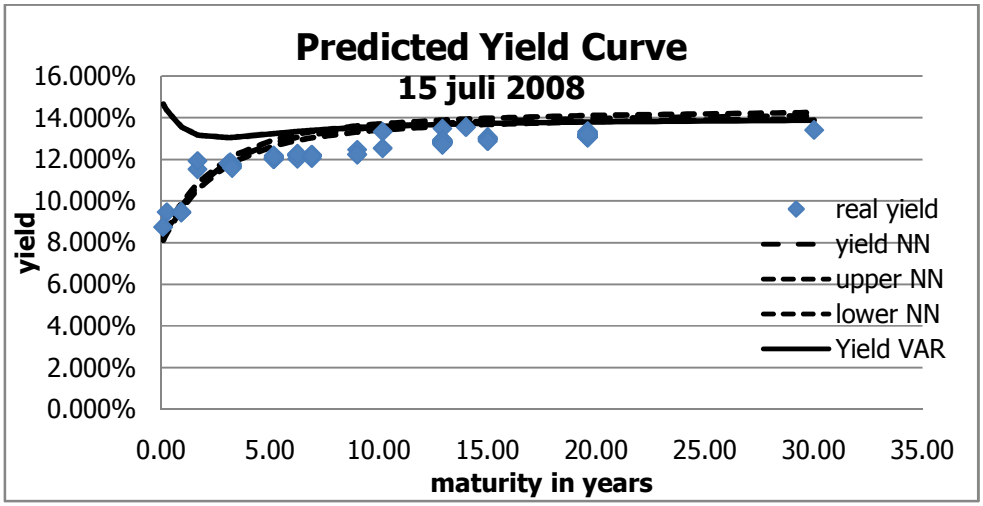

Figure 2. Predicted Yield Curve for July 15, 2008

Nugraha, Y. A. and Rosadi, 2011, The Term Structure of Nelson Siegel Modeling using Nelder Mead Algorithm and Parameter Forecasting using Vector Autoregressive model, Submitted (in Bahasa Indonesia)

Rosadi, D., 2010, Rplugin.Econometrics: R-GUI for teaching Time Series Analysis, Proceeding of CompStat 2010, 1st Edition., ISBN: 978-3-7908-2603-6, Springer Verlag, Paris

Rosadi, D., Gunardi, Abdurakhman, Utami, H. and Wulansari, Y., 2008, Validation and Goodness of Fit Test for Indonesian Government Securities Yield Curve, Research Report, Department of Statistics - Gadjah Mada University and Indonesian Bond Pricing Agency

Rosadi, D., Kumadewi, R., 2009, Forecasting the Indonesian Government Securities Yield Curve using Neural Networks, Proceedings AMC 2009, Malaysia

Rosadi, D., Qoyyimi, D. T. and Wulansari, Y. , 2010, Modeling Yield Spread for Indonesian Corporate Bond, Internal Research Report, Department of Statistics - Gadjah Mada University and Indonesian Bond Pricing Agency

Stander, S.Y., 2005, Yield Curve Modeling, Palgrave Macmilan, New York

RÉSUMÉ (ABSTRACT). In this paper, we discuss the problem of forecasting the yield curve of Indonesian Government bond using neural networks and Vector Autoregressive Model. We first model the yield curve using the Nelson-Siegel (NS) model and further estimate the parameters using a constrained optimization algorithm. Then, we forecast various parameters of the Nelson-Siegel yield curve using neural networks and Vector Autoregression (VAR). The forecasted NS parameters are then used to calculate the yield curve of the various Indonesian government bonds. We implement the methods using data of Indonesian Government Bond, obtained from Indonesia financial market. The results presented here extend studies presented in Rosadi and Kusumadewi (2009) and Nugraha and Rosadi (2011). In Nugraha and Rosadi (2011), we apply the OLS method for the starting values of the parameters, and further use these starting values for Nelder Mead estimation of the parameters, and only apply VAR forecast method in R-CLI. In this paper, we use SQP optimization method, available in SPSS 15. For VAR(p) modeling, we apply our new R-GUI package *RcmdrPlugin.Econometrics* (Rosadi, 2010). We also compare the VAR performance with neural network forecast, using neural networks estimation function in MATLAB 7.

Keywords: Yield Curve, Nelson Siegel, Neural Networks, Vector Autoregression