

Implementation of LM Algorithm and SQP for Yield Curve Optimization using Nelson-Siegel-Svensson Model

Dedy Dwi Prastyo

Department of Statistics, Institut Teknologi Sepuluh Nopember, Indonesia
E-mail: dedy-dp@statistika.its.ac.id

Suhartono

Department of Statistics, Institut Teknologi Sepuluh Nopember, Indonesia

Hizir Sofyan

Department of Mathematics, Universitas Syah Kuala, Indonesia

Abstract. To ease obligation portfolio management, a model presenting a relation between Yield to Maturity (YTM) and Time to Maturity (TTM), which is called yield curve model, is needed. Generally, yield curve can be modeled by two statistical approaches, i.e. parametric or nonparametric model. The objective of this paper is to study the adequacy of the Nelson-Siegel-Svensson (NSS) model, as one of parametric model, which is implemented to Indonesian government bonds data. NSS model is optimized by using non-linear least square optimization. Two non-linear least square methods giving good estimation are Levenberg-Marquardt (LM) algorithm and Sequential Quadratic Programming (SQP). The results of this research showed that NSS model is adequate to be used in Indonesian government bonds yield curve modeling. The optimization of NSS model showed that LM algorithm's estimation is more stable and robust than SQP's estimation. Hence, cleaning or filtering outlier data should be done to get the better result.

Key words and Phrases: *Yield curve, Nelson-Siegel-Svensson model, Levenberg-Marquardt algorithm, Sequential Quadratic Programming, Outlier*

1. INTRODUCTION

As the era development, there are many ways to save money in form of investment. In invest management, obligation is one of the capital market instruments which give fixed-income securities for the holders (Tandellin, 2001). Therefore, obligation market is interesting for many investors who don't like the existence of high risk in the decreasing economic condition (Haugen, 1997). Government bonds are ones of the risk-free obligations which are published by the government.

One of the primary attribute in obligation evaluation is yield structure (Sharpe, Gordon, and Jeffery, 1995). Yield value determination in most obligation portfolio management is approached by Yield to Maturity (YTM). To ease obligation portfolio management, a model presenting a relation between YTM

and maturity time or Time to Maturity (TTM), which is called yield curve model, is needed.

McCulloch (1975) and Fisher (1995) made a yield curve model through polynomial spline nonparametric approach by using government obligation data in England. In 1987, Nelson and Siegel introduce yield curve model by using nonlinear parametric approach. Svensson (1994) developed Nelson-Siegel (NS) model and name it Nelson-Siegel-Svensson (NSS) model. Otok et al. (2009) showed that NSS model is the best model to model yield curve of Indonesian government bonds data. There are several methods that can be used for estimating parameters of NSS model. This research focuses on the use of LM algorithm and SQP for NSS model optimization.

2. YILED CURVE MODEL

Obligation is defined as a debenture which is offered by the issuers or the borrowers to pay back the lenders in amount of the loan added by the interest for the determined years (Fabozzi, 2000). The main attributes of obligation are maturity time and yield. In obligation portfolio, investor use TTM approach more to know cash flow that will be got in the maturity time (Tandellin, 2001).

In 1987, Nelson and Siegel showed the yield curve from a model which is located in the same range form. The model uses four parameters to determine the curve. The NS model can be stated as in the following equation

$$y = f(T, \theta) + \varepsilon \quad (1)$$

with

$$f(T, \theta) = \beta_0 + \beta_1 \exp\left(-\frac{T}{\tau}\right) + \beta_2 \left(\frac{T}{\tau} \exp\left(-\frac{T}{\tau}\right)\right) \quad (2)$$

where y is yield, $y = f(T, \theta)$ is yield using forward rate approach in maturity time (T), τ is a time constant for the curve, β_0 shows the asymptotic value, β_1 and β_2 are parameters that show the direction of the curve.

NSS model is a developed yield curve model which use NS method and added by more parameters, i.e. β_3 and τ_2 , which are used to increase the curve flexibility (Amoako, Isaac, basil, 2005). The NSS model is stated as in the following equation

$$f(T, \theta) = \beta_0 + \beta_1 \exp\left(-\frac{T}{\tau_1}\right) + \beta_2 \left(\frac{T}{\tau_1} \exp\left(-\frac{T}{\tau_1}\right)\right) + \beta_3 \left(\frac{T}{\tau_2} \exp\left(-\frac{T}{\tau_2}\right)\right) \quad (3)$$

where β_3 and τ_2 are used to determine the direction and the position of the second curve. The existence of the addition parameters make NSS model smoother and follow the contour of the data.

3. OPTIMIZATION METHODS

3.1 Levenberg-Marquardt Algorithm

Levenberg-Marquardt (LM) algorithm is one of the iterative method which is used in parameter estimation of nonlinear model. Mathematically and computationally, LM algorithm bring quantitative solution for optimizing parameters in nonlinear function. LM algorithm is an interpolation between Gauss-Newton algorithm and gradient descent method, so LM algorithm is more

perfect than Gauss-Newton algorithm and be able to yield less result. Besides that, LM algorithm is very popular in curve-fitting algorithm (Wikipedia, 2008).

When the current solution is far from the correct one, the algorithm behaves like a steepest descent method: slow, but guaranteed to converge. When the current solution is close to the correct solution, it becomes a Gauss-Newton algorithm (Lourakis, 2005).

The vector of parameters $\theta' = (\theta_1, \theta_2, \dots, \theta_p)$ is estimated by using Ordinary Least Square (OLS) method by minimizing Sum Square of Error (SSE) of eq.(1). The estimated parameters are optimized by differentiating SSE to the vector parameters as in eq.(4) and (eq.5). Equation (5) cannot be solved directly, so it should be solved iteratively.

$$SSE = \sum_{i=1}^n (y_i - \hat{f}_i(\theta))^2 \quad (4)$$

$$\frac{\partial SSE}{\partial \theta} = \sum_{i=1}^n -2(y_i - \hat{f}_i(\theta)) \frac{\partial \hat{f}_i(\theta)}{\partial \theta} = 0. \quad (5)$$

Gauss-Newton algorithm uses second order of Taylor series to approach equation (5) iteratively. Estimated parameters in iteration (k+1) of Gauss-Newton algorithm is as follow:

$$\hat{\theta}_{k+1} = \hat{\theta}_k + (\mathbf{X}'_k \mathbf{X}_k)^{-1} \mathbf{X}'_k \hat{\mathbf{e}}_k \quad (6)$$

with

$$\mathbf{X}'_k = \begin{bmatrix} \frac{\partial f_1(\theta_k)}{\partial \theta_1} & \frac{\partial f_2(\theta_k)}{\partial \theta_1} & \dots & \frac{\partial f_n(\theta_k)}{\partial \theta_1} \\ \frac{\partial f_1(\theta_k)}{\partial \theta_2} & \frac{\partial f_2(\theta_k)}{\partial \theta_2} & \dots & \frac{\partial f_n(\theta_k)}{\partial \theta_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_1(\theta_k)}{\partial \theta_p} & \frac{\partial f_2(\theta_k)}{\partial \theta_p} & \dots & \frac{\partial f_n(\theta_k)}{\partial \theta_p} \end{bmatrix}; \quad \hat{\mathbf{e}}_k = \begin{bmatrix} (y_1 - \hat{f}_1(\theta_k)) \\ (y_2 - \hat{f}_2(\theta_k)) \\ \vdots \\ (y_n - \hat{f}_n(\theta_k)) \end{bmatrix}$$

where p is the number of parameters.

LM algorithm revised Gauss-Newton method by inserting a constant λ whose value change iteratively adaptive to the changing of SSE. If SSE decreases, then λ decrease and vice versa. Estimated parameters in iteration (k+1) of LM algorithm is as follow

$$\hat{\theta}_{k+1} = \hat{\theta}_k + (\mathbf{X}'_k \mathbf{X}_k + \hat{\lambda} \mathbf{I})^{-1} \mathbf{X}'_k \hat{\mathbf{e}}_k. \quad (7)$$

In this research, LM algorithm is used to estimate parameters of NSS model, as one of the yield curve model, to fit the Indonesian government bonds data. Based on eq.(3) and eq.(4), the vector of parameters and SSE of NSS model are stated as follow

$$\theta' = (\beta_0, \beta_1, \beta_2, \beta_3, \tau_1, \tau_2) \quad (8)$$

$$SSE = \sum_{i=1}^n (y_i - \hat{f}(T_i, \theta))^2. \quad (9)$$

Matrix of independent variables and vector of residual as in eq.(6) are stated as in the following forms

$$\mathbf{X}_k = \begin{bmatrix} \frac{\partial f(T_1, \boldsymbol{\theta}_k)}{\partial \beta_0} & \frac{\partial f(T_1, \boldsymbol{\theta}_k)}{\partial \beta_1} & \dots & \frac{\partial f(T_1, \boldsymbol{\theta}_k)}{\partial \tau_2} \\ \frac{\partial f(T_2, \boldsymbol{\theta}_k)}{\partial \beta_0} & \frac{\partial f(T_2, \boldsymbol{\theta}_k)}{\partial \beta_1} & \dots & \frac{\partial f(T_2, \boldsymbol{\theta}_k)}{\partial \tau_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f(T_n, \boldsymbol{\theta}_k)}{\partial \beta_0} & \frac{\partial f(T_n, \boldsymbol{\theta}_k)}{\partial \beta_1} & \dots & \frac{\partial f(T_n, \boldsymbol{\theta}_k)}{\partial \tau_2} \end{bmatrix} \text{ and } \hat{\mathbf{e}}_k = \begin{bmatrix} (y_1 - \hat{f}(T_1, \boldsymbol{\theta}_k)) \\ (y_2 - \hat{f}(T_2, \boldsymbol{\theta}_k)) \\ \vdots \\ (y_n - \hat{f}(T_n, \boldsymbol{\theta}_k)) \end{bmatrix}.$$

The LM algorithm for NSS model is as in the following

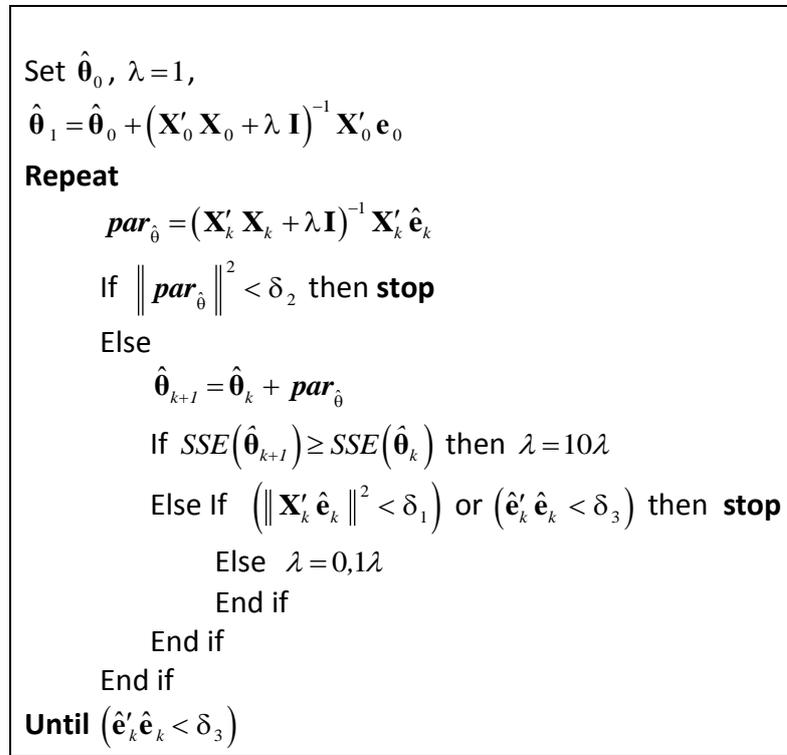


Figure 1. LM Algorithm for NSS Model

Terminating procedure is very important in an algorithm. In this research, the LM algorithm terminates when at least one of the following condition is satisfied:

1. The magnitude of the gradient of $\hat{\mathbf{e}}'_k \hat{\mathbf{e}}_k$, i.e. $\mathbf{X}'_k \hat{\mathbf{e}}_k$ in eq.(7) drops below a threshold δ_1 .
2. The relative change in the magnitude of $\hat{\boldsymbol{\theta}}_k$ drops below a threshold δ_2 ,

$$\|\hat{\boldsymbol{\theta}}_{k+1} - \hat{\boldsymbol{\theta}}_k\|^2 \leq \delta_2 \text{ or } \|(\mathbf{X}'_k \mathbf{X}_k + \lambda \mathbf{I})^{-1} \mathbf{X}'_k \hat{\mathbf{e}}_k\|^2 < \delta_2.$$
3. The error (SSE) in eq.(9) drops below a threshold $\hat{\mathbf{e}}'_k \hat{\mathbf{e}}_k < \delta_3$.
4. A maximum number of iterations are completed.

The fourth condition doesn't guarantee to converge, so it is used the first up to the third condition in this research.

3.2 Sequential Quadratic Programming

The vector of parameters and SSE of NSS model are stated as in eq.(8) is $\theta' = (\beta_0, \beta_1, \beta_2, \beta_3, \tau_1, \tau_2)$ or generally say that $\theta' = (\theta_1, \theta_2, \dots, \theta_p)$. The objective is to minimize SSE as in eq.(4), say $R(\theta)$, with constraint $C_i(\theta)$. To optimize the objective function, it should be $\frac{\partial R(\theta)}{\partial \theta} = 0$ and $C_i(\theta) = 0$ which need to calculate $H_{ij} = \frac{\partial^2 R}{\partial S_i \partial S_j}$ and $J_{ij} = \frac{\partial C_i}{\partial S_j}$. In optimization by using SQP, the main algorithm is stated in the following steps:

Step one

Pick a guess θ_0 and evaluate $\frac{\partial R(\theta)}{\partial \theta}$ which most likely $\frac{\partial R(\theta)}{\partial \theta} \neq 0$. Taylor series expansion of function $f(x)$ at point a is:

$$f(x) = f(a) + f'(a)(x-a) + O((x-a)^2)$$

Similarly we have

$$\frac{\partial R(\theta)}{\partial \theta} = \frac{\partial R(\theta)}{\partial \theta} \Big|_{\theta=\theta_0} + \frac{\partial^2 R(\theta)}{\partial \theta^2} \Big|_{\theta=\theta_0} (\theta - \theta_0) + O((\theta - \theta_0)^2) \quad (10)$$

Set $\frac{\partial R(\theta)}{\partial \theta} = 0$ and omit $O((\theta - \theta_0)^2)$, then we have

$$-\frac{\partial R(\theta)}{\partial \theta} \Big|_{\theta=\theta_0} = \frac{\partial^2 R(\theta)}{\partial \theta^2} \Big|_{\theta=\theta_0} (\theta - \theta_0), \quad (11)$$

Now, it is defined that $(\theta - \theta_0) = \theta_\alpha$ or $\theta = \theta_0 + \theta_\alpha = \theta'_1$, where θ_α is the change to θ_0 that makes derivative equal to 0.

Step two

Now, we have got θ'_1 . Evaluate $C_i(\theta'_1)$, if equal to 0, we are done but most likely it will not evaluate to 0 in the first several steps. Say that $C_i(\theta'_1) \neq 0$. Let's apply Taylor series expansion on the constraint function $C_i(\theta)$ at point θ'_1 :

$$C(\theta) = C(\theta'_1) + \frac{\partial C}{\partial \theta} \Big|_{\theta=\theta'_1} (\theta - \theta'_1) + O((\theta - \theta'_1)^2) \quad (12)$$

Set $C(\theta) = 0$ and omit $O((\theta - \theta'_1)^2)$, then we have

$$-C(\theta'_1) = J_{ij}(\theta'_1)(\theta - \theta'_1) \quad (13)$$

Now, it is defined that $(\theta - \theta'_1) = \theta_\beta$ or $\theta = \theta'_1 + \theta_\beta = \theta_0 + \theta_\alpha + \theta_\beta$, where θ_β is the change to θ_0 that makes derivative equal to 0.

In the iteration, we get the estimation $\theta_1 = \theta_0 + \theta_\alpha + \theta_\beta$. Then we will continue with step one and step two until we got a solution θ_k which minimizes our object function and also satisfies our constraints. The iteration will result these values:

$$\theta_0 \Rightarrow \theta'_1 \Rightarrow \theta_1 \Rightarrow \theta'_2 \Rightarrow \theta_2 \Rightarrow \dots \Rightarrow \theta_k \quad (14)$$

4. DATA ANALYSIS

4.1 Data and Research Method

The objective of this research were: (i) to study the adequacy of NSS model to Indonesian government bonds data, and (ii) to choose the best optimization or estimation method for NSS model. To obtain that objective, this research was done by the following steps: data preparation, cleaning or filtering data, modeling, and model evaluation.

Data preparation was focused on the data selection and cleaning or filtering data (also called outlier detection). NSS model was used to model 63 days of Indonesian government bonds data. In each day, the data consist pairs of YTM and TTM data. The first 62 days data (May 2, 2008 up to Juli 28, 2008) were used for modeling. These data were named in-sample data. Model validation was done to check wether the optimized NSS can be used to model Indonesian government data one day ahead. Model evaluation was done by using data from May 3, 2008 up to Juli 29, 2008. These data were named out-sample data. Root Mean Square Error (RMSE) is used as a criterion to choose the best model.

Preliminary identification showed that the outliers were present. Cleaning or filtering data was done by visually identification and inferentially statistics. Scatter plot was used to identified the outliers visually. Inferentially, if the values of standardized residuals of NSS model are greater than three, then the data are assumed to be an outliers.

4.2 Empirical Results

The six parameters in NSS model were estimated by two methods, i.e. LM algorithm and SQP. Parameters estimation was done after the data were cleaned. Figure 2 showed the NSS parameters estimation by using LM and SQP for the first 62 days data.

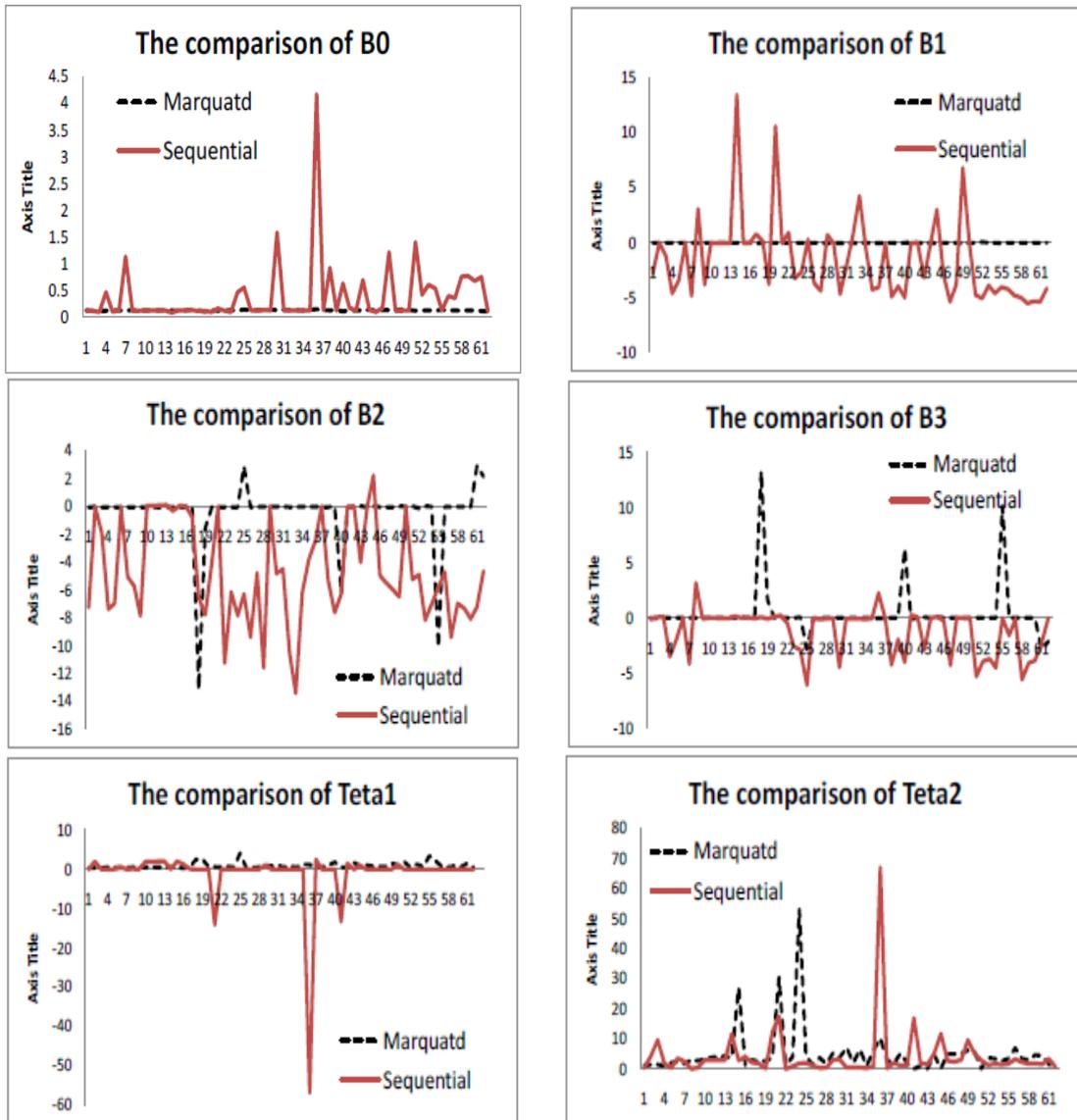


Figure 2. Estimated Parameters of NSS Model on the First 62 Days Data.

Figure 2 showed that the values of estimated NSS model parameters by using LM algorithm were relatively more stable. The estimated parameters in some certain days were very extreme, especially for parameters β_2 , β_3 , and τ_2 . Figure 2 also showed that parameters β_0 , β_1 , and τ_1 were more stable if estimated by using LM algorithm.

LM algorithm gave the better estimation for parameters of NSS model. The next, it will focus on the sensitivity of LM algorithm to the data with outlier and without outlier. Table 1 showed the summary of mean and standard deviation of estimated parameter of NSS model by using LM algorithm. There is no significant difference of parameters value before and after the outliers were omitted.

Then, the model adequacy is evaluated by using RMSE calculated from in-

sample and out-samples data. Figure3 showed the RMSE value of in-sample and out-sample data.

Table 1. NSS Model Parameters for the Data with Outlier and without Outlier

Parameters	Data with outliers		Outliers were omitted	
	Mean	Stdev	Mean	Stdev
β_0	0.1304	0.0076	0.1318	0.0080
β_1	-0.0296	0.1081	-0.0449	0.0156
β_2	-0.2414	2.0666	-0.4131	2.3195
β_3	0.1717	2.0895	0.3537	2.3406
τ_1	0.8893	1.2430	0.8820	0.7412
τ_2	3.8491	5.4809	4.8305	7.9436

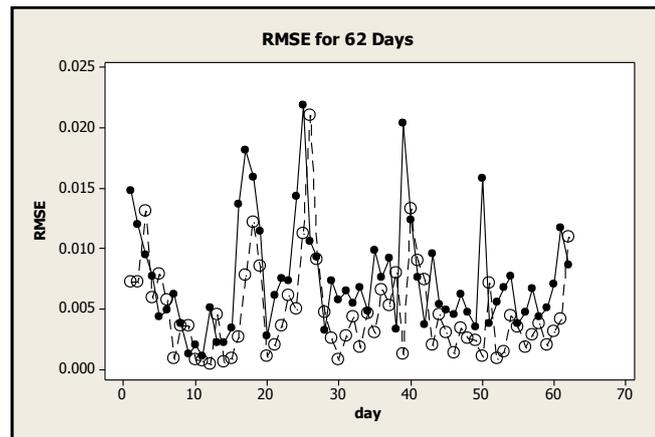


Figure 3. The Comparison of RMSE for In-Sample (circle) and Out-Sample Data (solid circle)

RMSE in-sample and RMSE out-sample is not significantly different. This fact showed that NSS model is good model to predict the YTM value one-day ahead. The mean of RMSE in-sample is 0.004771 with standard deviation is 0.003908. The mean of RMSE out-sample is 0.007480 with standard deviation is 0.004595.

Model evaluation was done to evaluate the adequacy of model, which is estimated from in-sample data, to detect the unusual observation of YTM value in out-sample data. In this paper, model evaluation was only done to NSS model optimized by using LM algorithm. Figure 4 showed the NSS model evaluation for two days data. Three lines represented the estimated NSS curve (blue line), estimated upper bond curve (red line), and estimated lower bond curve (black

line). The green line represented the observation TTM and YTM in out-sample data.

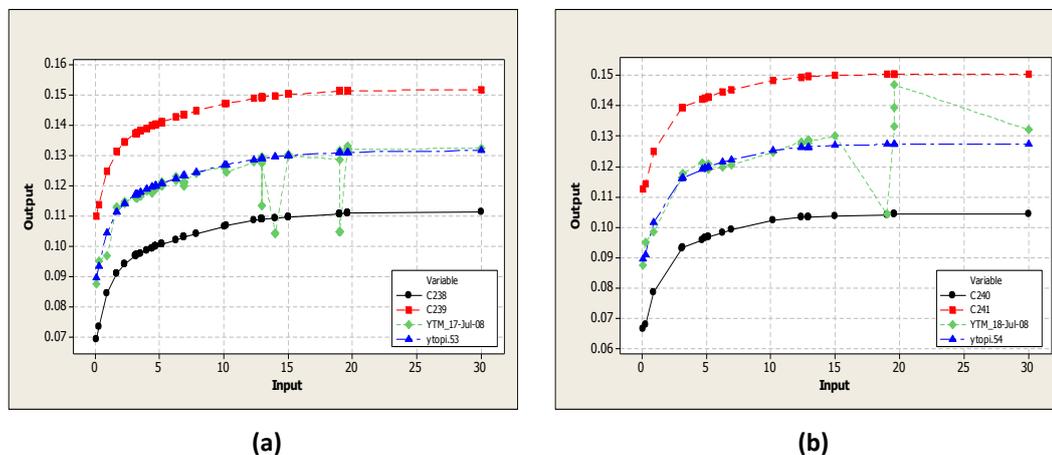


Figure 4. Plots of TTM and YTM from NSS Model: observation of out-sample data (green line) ; estimated lower bond curve (black line), estimated upper bond curve (red line) on July 17, 2008 (a), and July 18, 2008 (b).

Model evaluation process showed that average 1.99% of TTM values were detected as an unusual value. It can be said that there are average two government bonds were predicted as an unusual value from 100 government bonds in a day. Then, 95% confidence interval estimation showed that 1,486% - 2,494% government bonds were detected as an unusual value. These results showed that NSS model can be used as an early warning system (EWS) to detect the unusual values of yield.

5. CONCLUDING REMARKS

This research concludes that optimization of NSS model by using LM algorithm was better (more stable) than by using SQP. NSS model is adequate to be used as a yield curve model for Indonesian government bonds data. Then, NSS model can be used as an early warning system to detect the unusual value of government bonds data. NSS model can predict 98.01% of government bonds data as a usual (normal) value.

REFERENCES

- Amoako, B.K.O., Isaac, A. & Basil, W.H. (2005) *Svensson Model and The Nelson-Siegel Model*. Malardalens University, Mathematics and Physics.
- Fabozzi, F..J. (2000). *Investment Management*. (2 ed). Prentice-Hall, Inc.
- Fisher, M. (1995). *Fitting The term Structure of Interest Rates With Smoothing Splines*. Federal Reserve Board Finance and Economics Discussion Series1.

- Haugen, R.A. (1997). *Modern Investment Theory* (Vol. 4). New Jersey: Prentice-Hall, Inc.
- Lourakis, M.I.A., *A Brief Description of the Levenberg-Marquardt Algorithm Implemented by Levmar*. Institute of Computer Science. Foundation for Research and Technology-Helllas (FORTH). Heraklion, Crete, Greece. 2005.
- McCulloch, J.H. (1975). The Tax-Adjusted Yield Curve. *The Journal of Finance*, **30**, 811-830.
- Nelson, C.R. and Siegel, A.F. (1987). Parsimonious Modeling of Yield Curves. *The Journal of Business*, **60**(4): 473-489.
- Otok, B.W., Yaniar, R. & Suhartono. (2009). Yield Curve of Indonesian Government Bonds using Parametric and Nonparametric Methods. *Proceeding of The 4th International Conference on Mathematics and Statistics*. Bandar Lampung.
- Sharpe, F., Gordon, J.A. & Jeffery, V.B. (1995). *Investments* (2 ed). New Jersey: Prentice Hall, Inc.
- Svensson, L.E.O., (1994). *Estimating and Interpreting Forward Interest Rates: Sweden 1992 -1994*. Technical Reports 4871, National Bureau of Economic Research.
- Tandellin, E. (2001). *Analisis Investasi dan Manajemen Portofolio* (1 ed). Yogyakarta; BPFE-Yogyakarta.
- www.wikipedia.org. (2008). *The Levenberg-Marquardt Algorithm*.