International Journal on Advanced Science Engineering Information Technology

Vol.8 (2018) No. 4
ISSN: 2088-5334

Natural Cubic Spline Model for Estimating Volatility

Farida[#], Nifatamah Makaje[#], Phattrawan Tongkumchum[#], Aniruth Phon-On[#], Jetsada Laipaporn[#]

[#] *Department of Mathematics and Computer Science, Faculty of Science and Technology,*
Prince of Songkla University, Pattani Campus, 94000, Thailand
E-mail: faridatuahuns@gmail.com, mnifatam@gmail.com, phattrawan.t@psu.ac.th, aniruth.p@psu.ac.th, laipaporn.j@gmail.com

***Abstract*— Volatility measures the dispersion of returns for a market variable since a reasonable estimation of the volatility is an appropriate starting point for assessing investment risks and monetary policymaking. These risks are usually assessed by using the GARCH (1,1) model. However, the recursive term in this model makes finding the derivatives of the likelihood function mathematically intractable. In this study, the natural cubic spline model is used to estimate the volatility by fitting it to the absolute returns of the data. In estimating the parameters, the Maximum Likelihood method was applied while a simple algebra was used to find its derivatives. The damped Newton-Raphson method was then used to maximize the likelihood function with the R programming software. The proposed method was illustrated using the absolute returns of the crude oil prices data from West Texas Intermediate, and it showed similar results with the popular GARCH (1,1) model. The natural cubic spline can be an alternative for estimating the volatility of any financial time series data.**

***Keywords*— volatility; natural cubic spline model; the damped Newton-Raphson method; maximum likelihood method; GARCH (1,1) model.**

I. INTRODUCTION

Unpredictable changes in market prices can lead to losses and because of that, any risk involved needs to be estimated. Financial institutions rely on the estimation of market risk and uncertain change in market factors such as prices of securities, indices, interest rates and currency exchange rates. Underlying this market risk is the concept of volatility. The volatility of a market variable is defining as the fluctuation of the returns on prices of the variable for a period.

Over the last 2 decades, academicians and financial analyst have been paying attention to the estimation and forecasting of the volatility since information from the volatility can be used to assess the investment risk, in order to minimize losses [1]. It can also be used in monetary policymaking [2]. In order to find the volatility, returns are used instead of prices because returns show less correlation than prices. In turn, the returns are defined as the proportional increases or decreases in the prices over this period. However, because the returns fluctuate randomly due to unpredictable and unknown factors, it is necessary to have a method for assessing the nature and extent of this random variation and thus distill the information from the observed data.

The well-known model for estimating the volatility is the GARCH model proposed by Bollerslev in 1986 [3]. The most widely used is the GARCH (1,1) model given by [4]

$$\sigma_i^2 = \gamma V_L + \alpha u_{i-1}^2 + \beta \sigma_{i-1}^2, \tag{1}$$

Where γ, α and β are the parameters, u_i, σ_i^2 are the returns and variance respectively on the trading day i while V_L is the long-run variance rate. Parameters from the model are calculated using the Maximum Likelihood (ML) method, derived from the probability density function for normal distribution [5]. For m observations, the likelihood function of the normal distribution is given by

$$l_{u_i} = \sum_{i=1}^{m} \left(-\ln\left(\sigma_i\right) - \frac{u_i^2}{2\sigma_i^2} \right). \tag{2}$$

There is many applications of the GARCH (1,1) model in time series data, for estimating banking share returns, stock market returns and applications in the agricultural sector [6], [7], [8], [9]. Although the GARCH (1,1) is widely used, the recursive term σ_{i-1}^2 in this model [3] makes the derivatives of the likelihood function mathematically intractable. This leads us to find a model whose derivative of the likelihood function is easily calculated.

In this study, an alternative method was proposed to estimate the volatility of a financial time series data. The method involved fitting the natural cubic spline model to the absolute returns of the data, and then applied the ML method to find estimated parameters. In order to find such estimated parameters, it still needed to find a derivative of the likelihood function, but in this method, it will be used a simple algebra because the cubic spline was just polynomial. The damped Newton-Raphson was then used to maximize the likelihood function. The parameter estimations were carried out using R statistical software. The model was illustrated using West Texas Intermediate's data for the crude oil prices [10].

II. Material and Method

Time series of daily crude oil prices comprising of 7,735 trading days, prices from 2nd January 1986 to 29th August 2016, is shown in Fig. 1. The data shows a peak in July 2008 around 145 US$/barrel. However, the price had crashed down to 31 US$/barrel by December 2008.

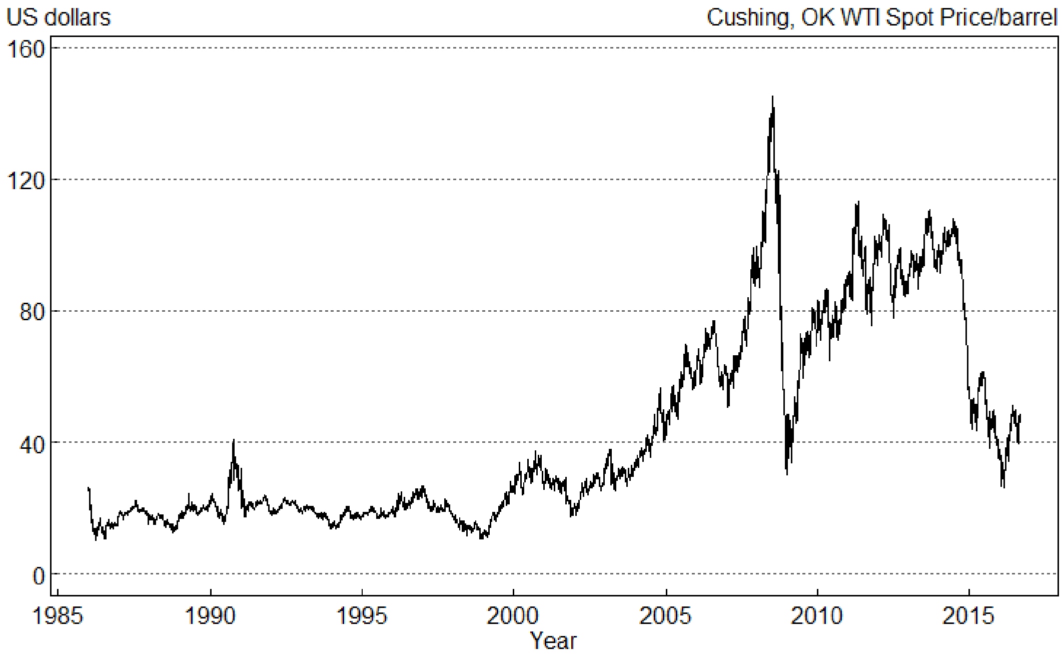
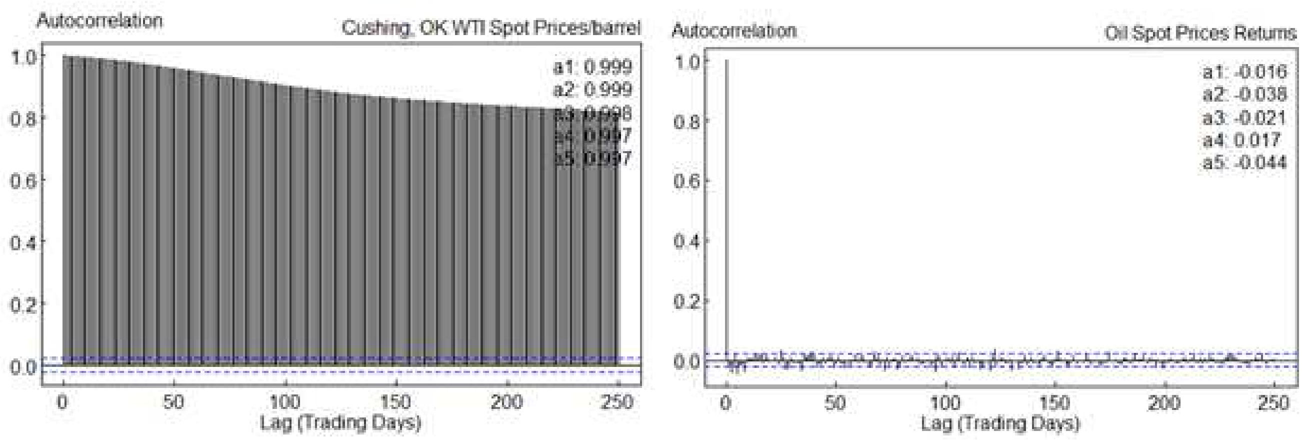


Fig. 1 The fluctuation of the crude oil prices data for 7,735 trading days

However, the high correlation between successive values of the crude oil prices data complicates fitting statistical models that assume independence of prices as shown in Fig. 2. For the oil prices, the correlation on day 1 is equal to 1, and a1 which is 0.999 is the correlation between day 1 and day 2. From Fig. 2, it shows that the autocorrelation plot for the prices on the left panel shows high correlations whereas the autocorrelation for returns shows least correlations on the right panel.



(a) (b)

Fig. 2 Autocorrelation functions for the oil prices (a) and the oil price returns (b).

For this reason, it is preferable to use the returns on the prices, rather than the actual prices. The return for i^{th} observation is denoted by u_i and can be calculated by the following formula [11].

$$u_i = \ln \frac{S_i}{S_{i-1}}; \quad i = 1,2,\ldots,m, \qquad (3)$$

Where S_i is a market price? Usually, volatility models for financial time series data use the squared returns, but empirical evidence has shown that using the absolute returns on the data gives better volatility estimation [12], [13], [14] and Ding et al. has also suggested [15].

The absolute returns of the crude oil prices are shown in Fig. 3, where the horizontal red line represents the mean of the absolute returns, which is 0.01747. The magnitudes of the crude oil price returns show an irregular rise to higher levels, as represented by thick dots. A method is needed to highlight these irregular swings by smoothing out the local fluctuations to capture the signal of the data. For this purpose, the smoothing splines are widely used [16]-[18], however, in this study the natural cubic spline is used to smooth the local fluctuations in Fig. 3.

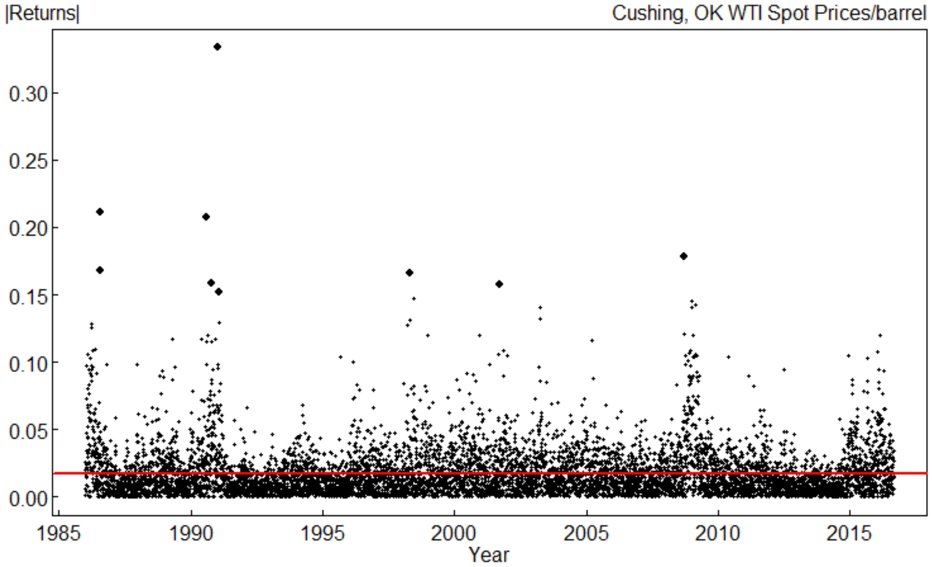


Fig. 3 The absolute returns of the crude oil prices

The natural cubic spline can be shown mathematically as,

$$s(t) = b_0 + b_1 t + \sum_{k=1}^{p-2} c_k \left[(t-t_k)^3 - \left(\frac{t_p - t_k}{t_p - t_{p-1}} \right) (t - t_{p-1})^3 + \left(\frac{t_{p-1} - t_k}{t_p - t_{p-1}} \right) (t - t_p)_+^3 \right], \qquad (4)$$

Where t denotes time, $t_1 < t_2 < \ldots < t_p$ are specified p knots and b_0, b_1 c_k are the unknown parameters. The model in Equation (4) was fitted to the absolute returns of the data, as shown by the light curve in Fig. 4.

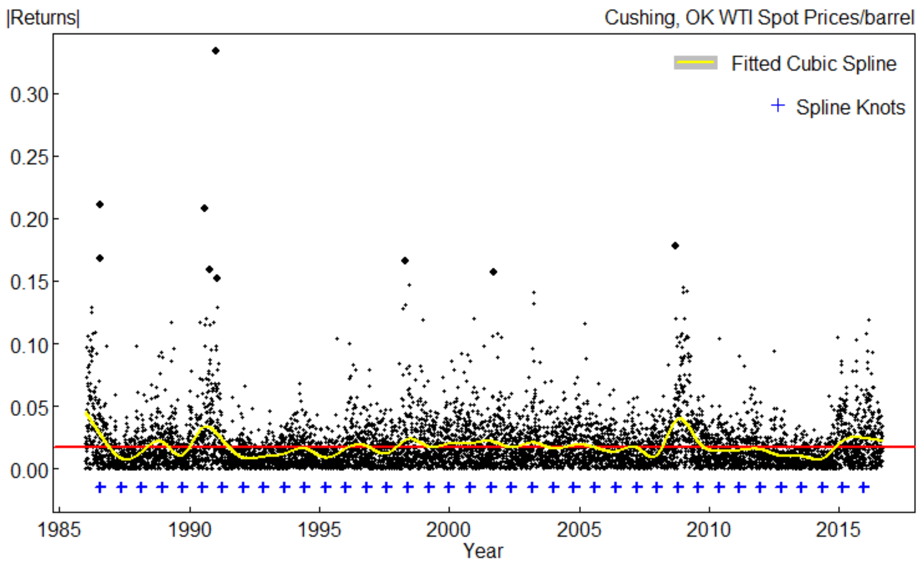


Fig. 4 Fitted natural cubic spline to the absolute returns

The parameters b_0, b_1 and c_k from the natural cubic spline was estimated by using Equation (2) where σ_i to represent $s(t)$ in Equation (4) as the volatility. The initial volatility σ_i was obtained from the fitted value of the natural cubic spline to the absolute returns in Fig. 4. The best estimate of σ_i was the value that maximizes its equation when the ML method was used, and this was obtained by applying a numerical method [19].

The damped Newton-Raphson iterative method was used since it is a powerful method for solving nonlinear equations with quadratic convergence [20], based on the formula below.

$$\mathbf{x}^{n+1} = \mathbf{x}^n - d\left[H(\mathbf{x}^n)\right]^{-1} F(\mathbf{x}^n), \quad n = 0,1,2,\ldots, \quad (5)$$

where $\mathbf{x}^{n+1}$ is the new value of $\mathbf{x} = (b_0, b_1, c_k)$ at $(n+1)th$ iteration, $\left[H(\mathbf{x}^n)\right]^{-1}$ is the inverse of the Hessian matrix (matrix of the second partial derivative) of the function l_{u_i}, $F(\mathbf{x}^n)$ is the first derivative of the function l_{u_i} and d is the damping factor such that $d \in [0,1)$. The damping factor d was used to avoid overshooting [21] and to decrease the changes at each iteration of the Newton-Raphson method.

The initial parameters for the iteration processes are typically obtained randomly or by guessing. However, in this study the initial parameters b_0, b_1 and c_k for the damped Newton-Raphson method were obtained by fitting the natural cubic spline to the absolute returns using the linear regression in Fig. 4. By using the damped Newton-Raphson iterative method, the estimated parameters b_0, b_1 and c_k were obtained.

III. Results and Discussion

Volatility estimation is frequently done by using the GARCH (1,1) model. However, in this study, the proposed method by fitting the natural cubic spline to the absolute returns of the data was presented. Fig. 5 shows the initial volatility (yellow curve), and estimated volatility (pink curve) obtained, by using 38 specific knots of the spline (blue plus sign). The estimated volatility was obtained by substituting the estimated parameters b_0, b_1 and c_k into the formula in Equation (4). The pink curve in Fig. 5 shows the estimated volatility of the absolute returns of the crude oil prices data from West Texas Intermediate.

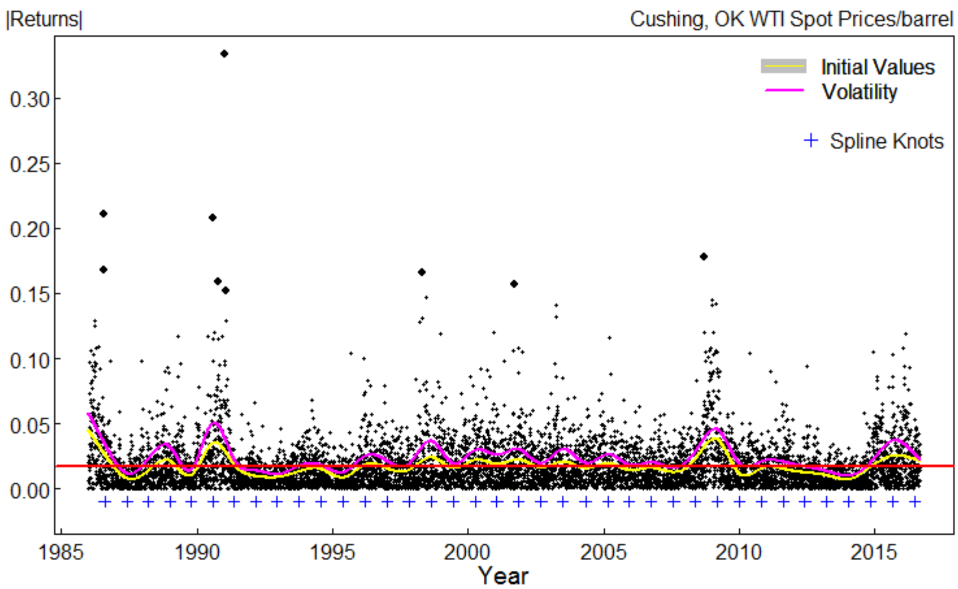


Fig. 5 The initial and estimated volatility of the absolute returns of the crude oil prices

In order to know how good the method proposed in this study is, it was compared with the Spline-GARCH (1,1) model. Note that the Spline-GARCH (1,1) model is a kind of Spline-GARCH model which was proposed by Engle and Rangel in 2005 [22]. This result is shown in Fig. 6.

The daily volatility from the GARCH (1,1) model was represented as a dotted line. It can be seen that there are some fluctuations in the volatility series from the GARCH (1,1) model around 1990 and 2009. In order to capture the signal, the natural cubic spline was applied in smoothening the daily volatility from the GARCH (1,1) model. This is represented as a blue curve, Spline-GARCH (1,1). Meanwhile, the estimated volatility from the proposed method was represented by a pink curve, Spline-fitted Model. This curve was obtained from Fig. 5 but it was multiplied by 100 in order to have the same scale with the daily volatility of the GARCH (1,1) model. The ability of the proposed method in estimating the volatility as well as the Spline-GARCH (1,1) model is illustrated in Fig. 6.

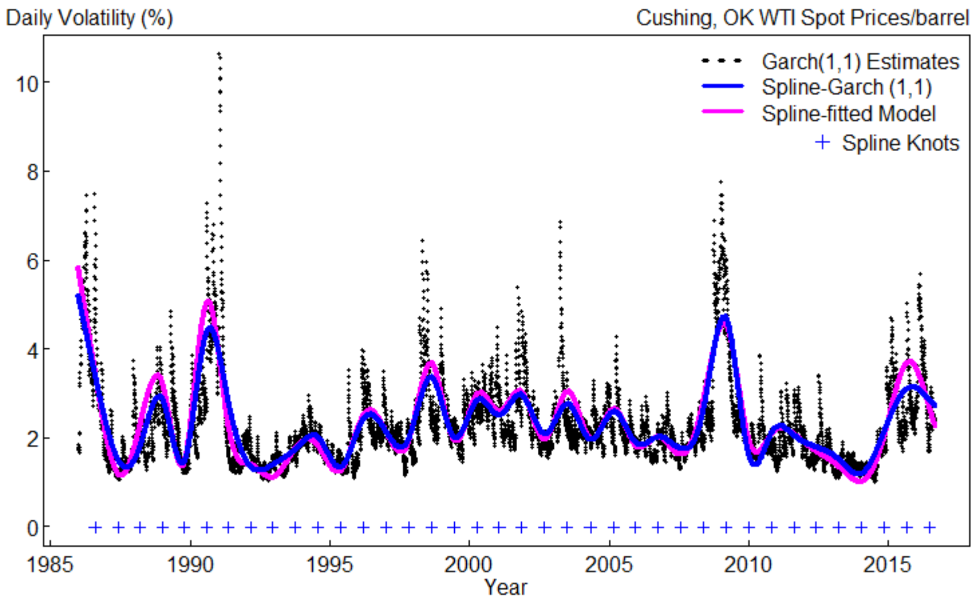


Fig. 6 Estimating the volatility of the absolute returns of the crude oil prices using the Spline-GARCH (1,1) model and Natural Cubic Spline model

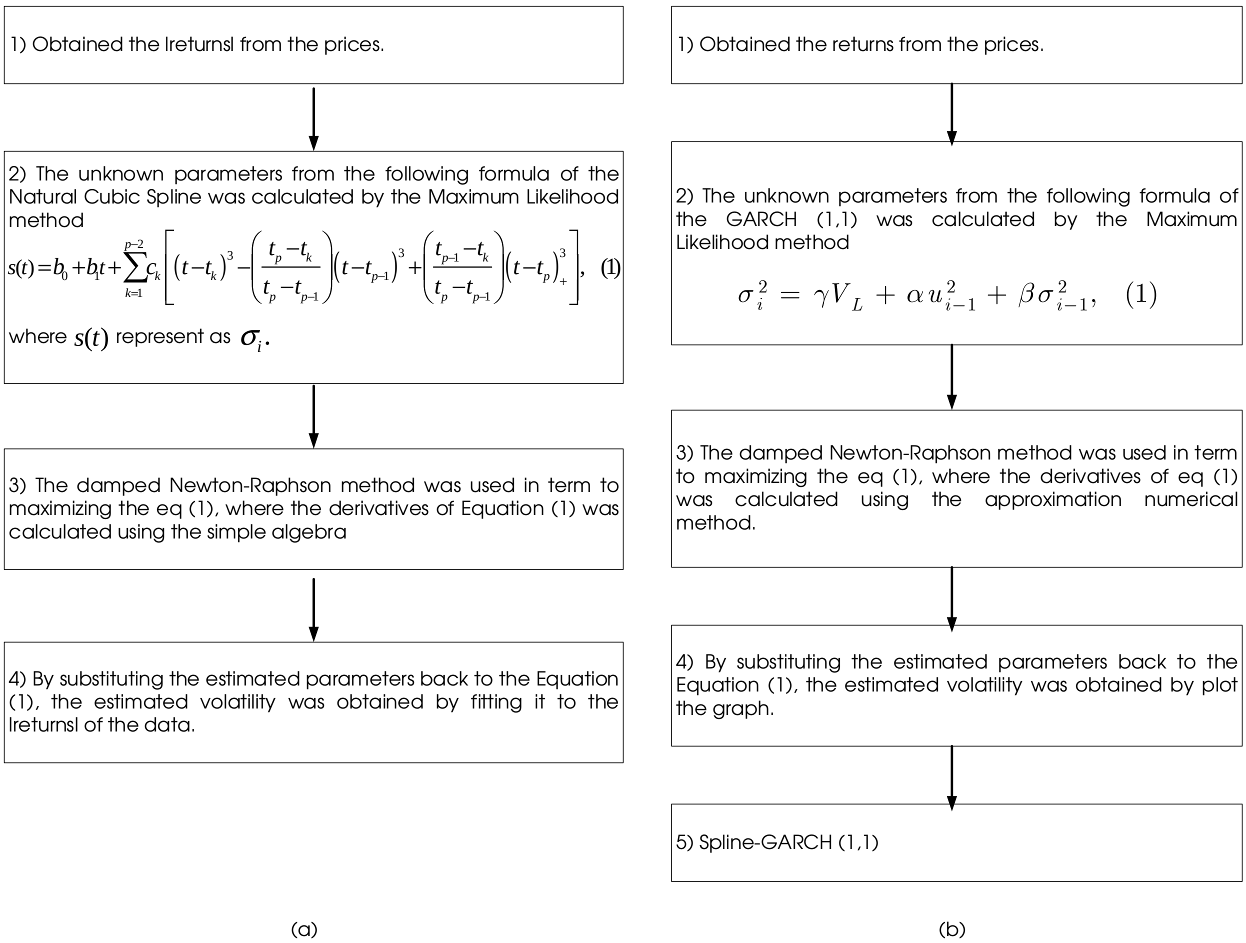


Fig. 7 The framework of the proposed method (a) and the Spline-GARCH (1,1) model (b)

Typically, to compare two models and determine which one is a better fit to the data, we could use the maximum value of their likelihood function [22] and the vague description (framework) of both methods. Table 1 shows the maximum values of the likelihood function of two models.

Based on the likelihood value in Table 1, the estimation of the volatility of the absolute returns of the crude oil prices data from West Texas Intermediate by the proposed method shows better performance. This concept based on the best estimate of the ML method is the value that maximizes its function [22].

TABLE I
THE LIKELIHOOD VALUE OF TWO METHODS

Model	The likelihood value
Spline-GARCH (1,1)	25,625.74
Proposed Method	25,681.33

As we mentioned that another diagnostic check of a model is to compare the framework. The frameworks of the two methods were shown in Fig. 7. The differences between the two methods are in the steps 3 and 5. The step 3 shows that the likelihood function in the Spline-GARCH (1,1) model was calculated using the numerical approximation method. Furthermore, the step 5 shows that the Spline-smoothed was needed in order to capture the signal of the data on a daily basis for the long period. Meanwhile, for the proposed method, the estimated volatility was obtained by substituting the estimated parameters back to the formula of the natural cubic spline then fitting it to the returns of the data.

Although the GARCH (1,1) is well-known, it does not sufficiently fit data set over an extended period, see [24]. So, a smoothing method to capture the signal from the daily volatility of the GARCH (1,1) model is needed. From the Fig. 7, the proposed method gives a more straightforward way to estimate the volatility from the financial time series data. Also, the initial parameters from the proposed method were obtained by fitting the natural cubic spline to the absolute returns using the linear regression, where generally in the GARCH (1,1) model the initial parameters were obtained by guess randomly.

Fitting the natural cubic spline to the absolute returns for estimating the volatility is a new approach for the volatility models. However, the findings from this study should depict a new approach for volatility estimation among the financial time series data. However further analyses in the future are required in order to apply this method to other financial time series data since only one dataset was used in this study.

IV. CONCLUSIONS

For over 25 years, academicians and financial analysts have been paying attention in developing better volatility estimation models for financial time series data. This is because volatility estimation plays an essential role in assessing investment risk. An alternative method has been proposed in this study to estimate the volatility of financial time series data by fitting the natural cubic spline model to the absolute returns of the data. The parameter of the model was estimated using the ML together with the damped Newton-Raphson method. The volatility that was estimated from the proposed method showed similar findings with the Spline-GARCH (1.1) model. The natural cubic spline model can be an alternative method for estimating the volatility of financial time series data. However, further analysis is required in the evaluation of the performance of the natural cubic spline model for other financial time series data.

ACKNOWLEDGMENT

The Higher Education Research Promotion and Thailand's Education Hub supported this study for Southern Region of ASEAN Countries Project Office of the Higher Education Commission. Our thanks go to Emeritus Prof. Dr. Donald McNeil for his helpful discussion. The authors declare no conflict of interest.

REFERENCES

[1] S. Poon and C. W. J. Granger, "Forecasting Volatility in Financial Markets: A Review," *Journal of Economic Literature*. Vol. 41, pp. 478-539, 2013.

[2] C. E. Alper, A. Morales, and F. Yang, "Monetary Policy Implementation and Volatility Transmission along the Yield Curve: The Case of Kenya," *Journal of Economic Literature*. Vol. E58, E43, pp. 1-29, 2016.

[3] T. Bollerslev, "Generalized Autoregressive Conditional Heteroscedasticity," *Journal of Econometrics*. Vol. 31, pp. 307-327. 1986.

[4] P. Posedel, "Properties and Estimation of GARCH (1,1) Model", *Metodoloski Zvezki.*, Vol. 2, No. 2, pp. 243-257, 2005.

[5] J. Sijbers, A. J. den Dekker, P. Scheunders and D. V. Dyck, "Maximum-likelihood Estimation of Rician Distribution Parameters", *IEEE Transactions on Medical Imaging.*, Vol. 17, No. 3, pp. 357-361, 1998.

[6] S. Saejiang, W. Pornwiriyamongkol, and D. McNeil, "Time Series Analysis of Banking Share Return in Thailand," *Songklanakarin Journal Science Technology*. Vol. 23, No. 3, pp. 444-448, 2001.

[7] V. Y. Naimy, "Parameterization of GARCH (1,1) for Paris Stock Market", *American Journal of Mathematics and Statistics.*, Vol. 3, No. 6, pp. 357-361. 2013.

[8] L. Aklimawati and T. Wahyudi, "Estimating the Volatility of Cocoa Price Return with ARCH and GARCH Models," *Pelita Perkebunan*., Vol. 29, No. 2, pp. 142-158. 2013.

[9] B. D. Kelkay and E. G. Yohannes, "The Application of GARCH Family Models to Some Agricultural Crop Products in Amhara National Regional State," *Journal of Economics and Sustainable Development.* Vol. 5, No. 11, pp. 24-35, 2014.

[10] U. S Energy Information Administration Agency. (2016). Petroleum and Other Liquids. [Online]. Available: https://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm

[11] A. E. M. Ahmed and S. Z. Suliman, "Modeling Stock Market Volatility Using Garch Models Evidence From Sudan," *International Journal of Business and Social Science*. Vol. 2, No. 23, pp. 114-128, 2011.

[12] S. J. Taylor, *Modeling Financial Time Series*. Lanchaster, U.K: Wiley, 1986.

[13] M. D. McKenzie, "Power Transformation and Forecasting the Magnitude of Exchange Rate Changes," *International Journal Forecast*. Vol. 15, pp. 49-55, 1999.

[14] M. Davidian and R. J. Carroll, "Variance Function Estimation," *Journal of the American Statistical Association*., Vol. 82, pp. 1079-1091, 1987.

[15] Z. Ding, C. W. J. Granger, and R. F. Engle, "A Long Memory Property of Stock Market Returns and A New Model," *Journal of Empirical Finance*. Vol. 1, pp. 83-106, 1993.

[16] M. A. Lukas, F. R. de Hoog, and R. S. Anderssen, "Efficient Algorithms for Robust Generalized Cross-Validation Spline Smoothing," *Journal of Computational and Applied Mathematics*., Vol. 235, pp. 102-107, 2010.

[17] R. L. Burden and J. D. Faires, *Numerical Analysis, 9th ed*. Boston, U.S.A: Brooks/Cole, 2011.

[18] S. C. Chapra, *Applied Numerical Methods with MATLAB for Engineers and Scientists, 2nd ed.* New York, U.S.A: McGraw-Hill, 2004.

[19] I. J. Myung, "Tutorial on Maximum Likelihood Estimation," *Journal of Mathematical Psychology*. Vol. 47, pp. 90-100, 2003.

[20] S. C. Chapra and R. P. Canale, *Numerical Methods for Engineers, 6th ed.* New York, U.S.A: McGraw-Hill, 2010.

[21] D. W. Marquardt, "An Algorithm for Least-squares Estimation of Nonlinear Parameters," *Journal of the Society for Industrial and Applied Mathematics*., Vol. 11, No. 2, pp. 431-441, 1963.

[22] R. F. Engle and J. G. Rangel., *The Spline GARCH Model for Unconditional Variance and Its Global Macroeconomic Causes*. Working paper. U. S. A: Stern School of Business, New York University, 2005.

[23] J. C. Hull, *Options, Futures, and Other Derivatives, 7th Ed.* New Jersey, U.S.A: Pearson Prentice Hall, 2009.

[24] T. Chuffart, "Selection Criteria in Regime-Switching Conditional Volatility Models," *Econometric*., Vol. 3, pp. 289-316, 2015.