Energy Economics 72 (2018) 341–355

Contents lists available at ScienceDirect

Energy Economics

journal homepage: www.elsevier.com/locate/eneco

The prediction of oil price turning points with log-periodic power law and multi-population genetic algorithm

Fangzheng Cheng[a], Tijun Fan[a], Dandan Fan[b], Shanling Li[c,*]

[a] School of Business, East China University of Science and Technology, Shanghai 200237, China
[b] School of Glorious Sun of Business and Management, Donghua University, Shanghai 200051, China
[c] Desautels Faculty of Management, McGill University, Quebec H3A 1G5, Canada

ARTICLE INFO

Article history:
Received 8 July 2017
Received in revised form 24 March 2018
Accepted 31 March 2018
Available online 13 April 2018

JEL classification:
C52
C53
C58
C63
G01
Q47

Keywords:
Turning point forecasting
WTI spot price
Log-periodic power law model
Multi-population genetic algorithm

ABSTRACT

The turning points in international oil price are the most significant and sudden corrections in prices in the world market. Accurate prediction of turning points can help governments and enterprises develop effective oil reserve strategies and economic decisions. Nevertheless, forecasting the turning points poses great challenges in both methodology and computational effort. Log-periodic power law (LPPL) is one state-of-the-art method to predict turning points. In this research, we propose an improved version of LPPL forecasting model by incorporating a method called multi-population genetic algorithm (MPGA) to search for optimal values of parameters in the LPPL model. By doing so, the improved LPPL model provided significantly superior performance in predicting the turning points compared to prior researches. To verify the quality of the improved LPPL model, we collected the data of WTI spot price in the period starting from April 2003 to November 2016 and used the improved LPPL model to predict the three turning points in this period based on the data prior to the turning points. In addition, we compared the improved LPPL model with three LPPL models that use other approaches to search for parameters, including simulated annealing, standard genetic and particle swarm optimization. We showed that the results from our LPPL model are superior to other three search approaches. We also concluded that the fluctuation of the WTI (West Texas Intermediate) spot price in March 2017 is a false alarm of a major turning point. The improved LPPL has great potential to predict future turning points.

1. Introduction

Crude oil is one of the most important energy resources on earth. So far, it remains the world's leading fuel, which accounts for nearly one-third of global energy consumption. Therefore oil price shock may significantly impact on economic activities and policies (e.g., Hamilton, 1983, 1996, Mork, 1989; Barsky and Kilian, 2004; Morana, 2013; Chen et al., 2016; Baumeister and Kilian, 2016a). In history, wild fluctuation in oil prices has led to recessions and even collapsing regimes so that the oil price turning point has been one of the important study topics for economists, investors and policymakers. For example, curtailed oil production (known as the "oil embargo") initiated by oil-exporting Arab nations in 1973 led to quadrupling of oil prices and contributed to economic recessions in oil-consuming, industrialized nations such as the United States and the United Kingdom (Kesicki, 2010). In late 1970s, after the Iraq-Iran war, supply shock, increased inventory demand due to the anticipation of future supply shortages and rising global demand all stimulated oil price to sky high. The price shocks had a substantial impact on US GDP and drove the US economy into a recession. In recent years, oil prices went through two cycles of highs and lows with the historic high values of oil prices of $145/barrel during 2010–2013 and the following prolonged downturn during 2014–2016 with the lowest oil price touching $26/barrel.

As discussed above, historical volatility in oil prices can be attributed to many factors, such as wars in oil production countries, strategic reserves and inventories of crude oil, U.S. dollar exchange rates, speculative trading, alternative energy sources, geopolitical conflicts and natural disasters (e.g., Kilian and Murphy, 2014; Brigida, 2014; Colgan, 2014), all of which make the accurate forecast of

* Corresponding author.
E-mail addresses: fz.cheng@mail.ecust.edu.cn (F. Cheng), tjfan@ecust.edu.cn (T. Fan), shanling.li@mail.mcgill.ca (S. Li).

https://doi.org/10.1016/j.eneco.2018.03.038
0140-9883/

crude oil prices particularly difficult. Although short-term oil price predictions has remained a constant topic of study, the results are usually less accurate than desired as in a short time span, fewer reference factors can be collected and accounted for in the forecasts. Therefore, government agencies and international organizations mainly rely on long-term forecast results to make macroeconomic policy decisions. On the other hand, accurate long-term forecast is also rather complex and difficult as it requires significant amount of comprehensive and timely information. To better predict oil price, researchers started to focus on turning points where the upward trend of oil prices suddenly turns into downward trend or vice versa (Roehner and Sornette, 1998; Zhou and Sornette, 2006; Sornette et al., 2009; Fantazzini, 2016). In recent years, the research on oil price turning point has become an emergent topic since a better understanding of when and how the turning points occur will help government decision makers develop long-term strategic oil reserve policies, i.e., buy or sell oil in advance according to the turning point prediction (Refenes et al., 1993). Specifically, after a significant turning point and if the oil price moves upwards, the Department of Energy should call on hoarding low-price oil to prevent another turning point at which oil price starts move down. If the oil price starts to decline after a downward turning point, they may reserve sufficient amount of oil for usage and sell the overage.

In this research paper, a log-periodic power law (LPPL) model based on multi-population genetic algorithm (MPGA) is proposed to evaluate the turning points of WTI (West Texas Intermediate) oil price in the time period of 1980 and 2018. The LPPL model was first proposed by Johansen, Ledoit and Sornette (Johansen and Sornette, 1999; Johansen et al., 1999, 2000; Sornette, 2002) in which the authors predicted the bubbles of financial market and found turning points that mark the beginning of the market crash. The LPPL model possesses strong predictive power as the model could forecast the bubble trend in the oil market and the time of the bubble burst as well (Yan et al., 2010). In this research, the model was further improved by optimizing the parameters using the multi-population genetic algorithm (MPGA) to improve the quality of the LPPL.

Specifically, the parameters in the LPPL model were first optimized using the MPGA, and then the forecast results were detected by Lomb periodogram analysis method. The WTI spot prices in the period between April 2003 and November 2016 were collected and the LPPL model was used to search for turning points in the historical data. The turning points generated by the LPPL were compared with the actual data, and the predicted turning points were found to be very accurate. We strongly believe that the improved LPPL model can be applied to predict turning point of the WTI spot price in the future.

In addition, the results from the improved model were compared with three other common heuristic algorithms including simulated annealing algorithm (Kirkpatrick, 1984), standard genetic algorithm (Holland, 1975) and particle swarm optimization algorithm (Eberhart and Kennedy, 1995). The numerical results show that although our new MPGA requires more computational time, its prediction results are much more accurate than the other three algorithms.

Lastly it is demonstrated that based on our LPPL model and the data from the WTI spot price, the time period of March 2017 is predicted not to be a major turning point. In addition, the turning points detected by our model were compared with oil stock market in the same time period and it was found that the turning points of crude oil price detected by our model are very consistent with those of stock price of oil. This means that the accurate forecast of turning point of oil price may also help guide the trade in the oil sector of the stock market. We believe that the high forecast quality to predict the turning points of oil price not only helps governments and oil enterprises to develop long-term oil reserve strategies, but also guides investment companies to make long-term effective investment decisions.

The rest of the paper is organized as follows: Section 2 reviews the literature on the prediction of crude oil price; Section 3 proposes the turning point forecasting model; Section 4 describes the data, and presents crude oil forecasting results and managerial implications; Section 5 concludes the paper and discusses future research.

2. Literature review

Early research works tend to study the short-term prediction of oil price. As discussed in the previous section, the government agencies tend to rely on long-term forecast of the oil price to develop oil reserve strategies. But, the long-term forecast is rather complex and difficult. Thus, recently a group of papers started to examine the turning points of the oil price. In what follows, existing literature is discussed according to the time span of study, namely short-term forecast, long-term forecast and turning point forecast.

To predict short-term trends of oil price, researchers incorporated impact factors such as oil inventory level and US dollar exchange rate. Short-term forecast models can be roughly divided into two categories: statistical and econometric models, as well as artificial intelligence (AI) models (Sehgal and Pandey, 2015). One of the most prominent and earliest statistical and econometric models was autoregressive conditional heteroskedasticity (ARCH) introduced by Engle (1982). Unlike the conventional time series or econometric models assuming constant variances, ARCH allows the conditional variance to change over time as a function of past errors leaving the unconditional variance constant. This type of model was proved to be useful in modelling different economic scenarios. As an extension of ARCH, generalized autoregressive conditional heteroskedasticity (GARCH) was developed by Bollerslev (1986), allowing for a much more flexible lag structure. Further improved models to forecast short-term oil price can be found in Morana (2001), Hansen and Lunde (2005), Narayan and Narayan (2007), Agnolucci (2009), Mohammadi and Su (2010), Hou and Suardi (2012) and Klein and Walther (2016). Lux and Kaizoji (2007) pointed out that ARCH, GARCH and related models cannot accommodate the property of multi-scaling (or multi-fractality). In other words, the models can only be applied to one-dimensional time series prediction. Since 2000s, many researchers developed multivariate time series prediction models that have been widely applied to short-term crude oil price prediction. The related research works include Calvet and Fisher (2001, 2004), Yu et al. (2008), Kilian (2009), Murat and Tokat (2009), Wang et al. (2011), Baumeister and Kilian (2012, 2016a), Zhang and Zhang (2015), Wang et al. (2016) and Cheng et al. (2018). Given that most statistical and econometric models are stationary time series models, artificial intelligence (AI) models were developed to predict short-term oil price for non-stationary time series. For example, Khashman and Nwulu (2011) developed artificial neural networks (ANN) and found that ANN prediction system achieved better prediction results. Recently, hybrid models of the statistical and AI models have been developed to capture unknown or complex structure in the oil price series. For example, VAR-SVM (vector autoregression-support vector machines) model proposed by Zhao et al. (2015) may overcome the limitations of early models and result in better forecasting accuracy.

The second stream of related literature studies the long-term prediction of crude oil price. In comparison to the short-term forecast, there is remarkably little literature on the long-term prediction of crude oil price (Haugom et al., 2016). The reason is that the information of the true global oil resources is poor or vague (Kjärstad and Johnsson, 2009), and the oil market mechanism is complex (Haugom et al., 2016). Graham and Dodd (1934) proposed an econometric model that was the first applied to the long-term prediction of oil price, but this model is difficult to implement in practice. In recent years, due to the rapid progress in computer technology and forecasting techniques for long-term oil price forecasting, a variety of forecasting models emerged such as econometric model (Haugom et al., 2016), equilibrium model (Weijermars and Sun, 2018), Hubbert model (Rehrl and Friedrich, 2006), wavelet analysis (Liang et al., 2005), stochastic processes (Gibson and Schwartz, 1990; Schwartz and Smith, 2000; Hahn et al., 2014), pooled forecast (Baumeister et al., 2014), Bayesian model (Lee and Huh, 2017) and jump and dip diffusion model (Shafiee and Topal, 2010; Hsu et al., 2016).

The last stream of research focuses on the prediction of turning points of oil price, which are the peaks and the troughs of oil price over a long time period. At present, there are mainly two methods for predicting the turning points, that is, tests for financial bubbles and the log-periodic power law (LPPL) model. Fantazzini (2016) obtained the same prediction results through these two methods. The tests for financial bubbles are based on the recursive and rolling right-tailed Augmented Dickey-Fuller unit root test, wherein the null hypothesis is of a unit root and the alternative is of a mildly explosive process that is proposed by Phillips et al. (2015) and Phillips and Shi (2014). However, this method requires the formation of a bubble as a pre-requisite for the following price crash (Fantazzini, 2016). The second method is the LPPL model and the literature on turning point forecasting of oil price by using the LPPL model are relatively few, and started to increase in very recent years. Sornette et al. (2009) and Zhang and Yao (2016) confirmed the existence of the turning point of oil bubble during 2008; Wątorek et al. (2016) and Fomin et al. (2016) detected a negative oil price bubble in 2016. It is noticed that most of the above studies that adopted the LPPL model also called the LPPL model as JLS model since it was first proposed by Johansen, Ledoit and Sornette. The LPPL model is based on the interplay between economic theory and its rational expectation postulate on one hand and statistical physics on the other hand (Johansen et al., 2000). The LPPL model is established by simulating the behavior of rational traders in the market. A rational trader can only either sell, buy or wait, but the transition from one state to the other is triggered by a threshold being exceeded and is thus a discontinuous process. The transition also involves another trader and the process is irreversible, as one trader cannot sell the commodity he bought back to the seller at the same price. In general, one trader only has information of a limited number of other traders and generally only see the cooperative response of the market as a whole in terms of an increase or a decrease in the value of the market. Thus it is rational for traders to remain their trading decisions as long as they are compensated by a higher rate of growth of the bubble for taking the risk of a crash. Based on the behavior of the rational traders which has characteristics with strong analogies to the field of statistical physics and complex systems, Johansen and Sornette (1999), Johansen et al. (1999, 2000) and Sornette (2002) proposed a rational expectation model, i.e., the LPPL model. Thus, the LPPL model is not only a statistical physics model, but also a rational economic model. During a bubble-like expansion, oil price is decoupled from the level justified by economic fundamentals (Fantazzini, 2016), and amplified by speculative behavior (Sornette et al., 2009). Therefore it cannot be explained using supply and demand alone, as highlighted by Baumeister and Kilian (2016b). Thus, the rational economic model considering speculative behavior, which is incorporated in the LPPL model, is more suitable for turning point prediction of oil price. As Pele et al. (2013) pointed out the group of LPPL models is a useful tool to recognize the behavior of market bubble and has good abilities to predict the turning points of an oil price bubble in the market. The LPPL model also has a successful application in the enterprise, such as the Research Group of BNP Paribas Fortis (Jiang et al., 2010) and leading financial service companies in China such as Guosen Securities and Guotai Junan Securities (China).

The LPPL models can be further divided into three types: simple log-periodic power law model (Sornette et al., 2009; Fantazzini, 2016; Zhang and Yao, 2016; Wątorek et al., 2016; Fomin et al., 2016), Weierstrass-type LPPL model (Sornette et al., 2009) and Landau-type LPPL model (Sornette et al., 2009). Although the latter two models are studied by many scholars, they are mainly applied to stock price prediction (Johansen and Sornette, 2000; Sornette and Zhou, 2004; Zhou and Sornette, 2005). However, in the research area of oil price turning point forecast, the simple log-periodic power law model is most commonly used, and the forecast results of these three type models are consistent for crude oil price (Sornette et al., 2009). Therefore, the simple LPPL model is chosen to predict the turning point of the WTI spot price. The key in implementing the LPPL model is to search for optimal values for parameters. This search is nonlinear in nature and it is concluded in Section 3 that the parameters optimization of simple LPPL is an NP-hard problem. Heuristic algorithms such as Tabu search were initially used for parameter optimization (Johansen et al., 2000; Sornette and Johansen, 2001; Sornette and Zhou, 2006; Jiang et al., 2010). Nevertheless, as Filimonov and Sornette (2013) pointed out, Tabu search does not guarantee the convergence in the parameter search process. Later, more heuristic algorithms were developed including the standard genetic algorithm (Jacobsson, 2009), simulated annealing (Sornette et al., 2013) and the particle swarm optimization (van Gysen, 2016).

In this paper, the multi-population genetic algorithm (MPGA) technique was employed to search for optimal values for the parameters in the LPPL model. In contrast to the heuristic approaches that do not guarantee the convergence and optimality of the parameters, the MPGA is able to find optimal parameters and our results show that the MPGA technique outperforms the current mainstream three heuristic algorithms: standard genetic algorithm, simulated annealing and particle swarm optimization. A quantitative detection method called the Lomb periodogram analysis was used to validate the optimization results.

3. The log-periodic power law (LPPL) model for the prediction of turning points

In this section, the improved LPPL model that employs the MPGA to search for optimal parameters is presented. The LPPL model has been widely used to capture the oscillating movements of bubbles in the financial market (Zhang and Yao, 2016). In this model, the ending crash of a speculative bubble is the climax of the log-periodic oscillation, and the most probable time of a crash is given by several parameters. By fitting the LPPL model to a time series, it is possible to predict the time (with the precision of up to an approximate date) of a crash. Only recently, researchers started to apply various LPPL models to predict the turning points of the oil price as the oil market is one kind of financial market with demonstrated speculative behaviors. The MPGA employed in the LPPL model and proposed in this paper not only improves the prediction accuracy, but also reduces the influence of some other factors such as crossover probability and mutation probability during the prediction process. The LPPL model is described as below:

$$y(t) = A + B(t_c - t)^{\alpha} + C(t_c - t)^{\alpha}\cos[\omega ln(t_c - t) + \phi] \tag{1}$$

where $y(t)$ is the oil price at time t, α is the exponential growth, ω is the control of the amplitude of the oscillations and A, B, C and ϕ are the parameters with no structural information.

Note that in Eq. (1), t_c is a critical time or a turning point, to be predicated. Fitting Eq. (1) to the WTI spot price data, the LPPL model may capture the speculative behavior of the oil price and thus predict any turning points. One main feature captured by Eq. (1) is the dampened yet accelerated oscillation in oil price. That is, when t approaches t_c, the oscillations occur more frequently with a decreasing amplitude. In other words, $(t_c - t)^\alpha$ is the power law term, which describes the faster than exponential change of the prices owing to positive feedback mechanisms and $(t_c - t)^\alpha \cos[\omega ln(t_c - t) + \phi]$ is the periodic term, which indicates a correction to the power law term and has the symmetry of discrete scale invariance. The most probable time of the turning point is when $t = t_c$, for $t \geq t_c$.

In the above LPPL model, there are seven parameters that need to be optimized, including four nonlinear parameters, t_c, ω, ϕ and α; and three linear parameters, A, B and C. For simplicity, the linear parameters can be directly derived from the given nonlinear parameters by using least square method (Johansen et al., 2000; Jiang et al., 2010). With the notation

$$f_j = (t_c - t_j)^\alpha \tag{2}$$

$$g_j = (t_c - t_j)^\alpha \cos[\omega ln(t_c - t_j) + \phi] \tag{3}$$

where t_j $(j = 1, 2, \ldots, J)$ with j as a time unit and J is the total time units in an interval, the linear parameters can be calculated using the following equations:

$$\begin{pmatrix} A \\ B \\ C \end{pmatrix} = \left[\left(V^T_{3*J} \cdot V_{J*3} \right)^{-1} \left(V^T_{3*J} \cdot Y_{J*1} \right) \right]_{3*1} \tag{4}$$

where $V_{J*3} = \begin{pmatrix} 1 & f_1 & g_1 \\ 1 & f_2 & g_2 \\ \vdots & \vdots & \vdots \\ 1 & f_J & g_J \end{pmatrix}_{J*3}$, which is a matrix with J rows and 3 columns; $Y_{J*1} = \begin{pmatrix} y(1) \\ y(2) \\ \vdots \\ y(J) \end{pmatrix}_{J*1}$, which is a column vector with J rows.

Therefore,

$$\begin{pmatrix} A \\ B \\ C \end{pmatrix} = \begin{pmatrix} J & \sum_{j=1}^J f_j & \sum_{j=1}^J g_j \\ \sum_{j=1}^J f_j & \sum_{j=1}^J f_j^2 & \sum_{j=1}^J f_j g_j \\ \sum_{j=1}^J g_j & \sum_{j=1}^J f_j g_j & \sum_{j=1}^J g_j^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{j=1}^J y(j) \\ \sum_{j=1}^J y(j) f_j \\ \sum_{j=1}^J y(j) g_j \end{pmatrix} \tag{5}$$

This approach is proven to be very stable and able to yield good estimation of the linear parameters A, B and C (Sornette and Zhou, 2002). However, the determination of the values of the four nonlinear parameters, t_c, ω, ϕ and α proves to be more challenging. In fact, it can be proven that searching for the optimal values of the four nonlinear parameters in the LPPL model is an NP-hard problem (see Appendix A). Therefore, it is important to find an effective and efficient heuristic algorithm to optimize the nonlinear parameters in the LPPL model. For this purpose, the multi-population genetic algorithm (MPGA) is employed to search for the optimal parameter values in the LPPL model. The MPGA is one of the most popular heuristic algorithms with the advantages of improving convergence rates and maintaining relatively low mean-square-errors (Sun et al., 2015). To the best of our knowledge, this is the first attempt to use the MPGA in the context of the LPPL model, which is to be discussed in details in Section 3.1. The framework of LPPL Model can be briefly summarized as follows:

Step1 : A sample interval is selected for the prediction of a turning point in the future time horizon. The interval is at least four-year long after the previous major turning point in history.

Step2 : The sample interval is further divided into over 100 subintervals to avoid the bias of specific sample interval and the impact of selecting a sample interval on the forecast result.

Step3 : For each subinterval, the MPGA (multi-population genetic algorithm) is employed to optimize the parameters in the LPPL model. The optimized LPPL model is then used to predict a date in the future when a turning point will occur.

Step4 : The Lomb periodogram analysis is conducted to statistically test the predicted turning points obtained by the LPPL models for all subintervals.

Step5 : The turning points that are statistically validated by the Lomb periodogram analysis are considered as predicted turning points by the LPPL model.

The specific details of the algorithms to implement the improved LPPL model are provided in Fig. B.8 (see Appendix B). In the next two sections, we discuss about the MPGA that is used to search for the optimal nonlinear parameters in the LPPL models and the Lomb periodogram analysis that is used to test the predicted turning points generated by the LPPL models.

3.1. Multi-population genetic algorithm (MPGA)

In this section, we will briefly discuss about the searching process by using the MPGA. The MPGA works differently from traditional genetic algorithms in searching for optimal parameters (particularly the nonlinear parameters in the LPPL model) in that the MPGA works on multiple populations with the objective to evaluate each subinterval. After the initial populations are produced, if the optimization criteria are not met, new populations are created and the search starts again. The above process iterates until the optimization criteria are reached.

The first step of the MPGA is to generate multiple populations. Inspired by the biology concepts of crossover and mutation, each population in the MPGA can be muted into hundreds of chromosomes and each chromosome represents a feasible solution for the four nonlinear parameters in the LPPL model. Based on Eq. (1), the MPGA measures the fitness value of each of chromosomes (i.e., the four nonlinear parameters) generated from all the populations by computing the residual sum of squares (RSS) between the historical oil price at time t or $y(t)$ and the results from the LPPL models:

$$\begin{aligned} RSS_{m,n} = \sum_{t=1}^{J} \Big\{ & y(t) - A - B\left(t_c^{m,n} - t\right)^{\alpha^{m,n}} \\ & - C\left(t_c^{m,n} - t\right)^{\alpha^{m,n}} \cos\left[\omega^{m,n} \ln\left(t_c^{m,n} - t\right) + \phi^{m,n}\right] \Big\}^2 \end{aligned} \tag{6}$$

where $RSS_{m,n}$ represents the fitness value (RSS) of the n-th chromosome in the m-th population; $t_c^{m,n}$, $\omega^{m,n}$, $\phi^{m,n}$ and $\alpha^{m,n}$ correspond to the n-th chromosome in the m-th population.

To highlight the main difference between the MPGA and other standard genetic algorithms, the operations in the MPGA are briefly described below. For each generation, the minimum fitness value and its corresponding chromosome in each population ($\min_n(RSS_{m,n})$ for each m), and the minimum fitness value in all populations ($\min_{m,n}(RSS_{m,n})$) and its corresponding chromosome are recorded. A new population will be generated by selection, crossover, mutation and re-inserting. Next, the chromosome whose fitness value is the smallest among the m-th population ($\min_n(RSS_{m,n})$ for m) will substitute the chromosome whose fitness value is the largest among the $m+1$-th population ($\max_n(RSS_{m+1,n})$ for $m+1$). This process is generally referred to as the immigration operation that combines individual populations into a unified entity. After the immigration operation, if the minimum fitness value in a new population ($\min_n(RSS_{m,n})$) is less than the corresponding record (i.e., the minimum fitness value and its corresponding chromosome in this population), the records are updated; otherwise, the original records in this population remain. The record of the minimum fitness value of all populations ($\min_{m,n}(RSS_{m,n})$) and the corresponding chromosome are processed in the same way. Finally, if the minimum fitness value of all populations does not change for a given number of consecutive iterations (set to 50 in our computation), or the total number of iterations reaches a given upper bound (set to 500 in our computation), the algorithm terminates and the latest minimum fitness value of all populations and its corresponding chromosome (t_c, ω, ϕ and α) are considered as the outputs of the MPGA. The specific procedure of the MPGA to optimize nonlinear parameters (t_c, ω, ϕ and α) is as provided in Appendix C.

After obtaining the four nonlinear parameters optimized by the MPGA, the three linear parameters can be subsequently derived from Eq. (5). The corresponding LPPL model is thus established. To identify whether the established LPPL model are statistically valid, the Lomb periodogram test is applied to detect periodic oscillations of the established LPPL model.

3.2. Validation of turning point prediction using the Lomb periodogram analysis

In order to validate the turning points predicted by the LPPL models, it is necessary to determine whether the frequency ($\omega/2\pi$) obtained from the MPGA (note that ω is one of the optimal four nonlinear parameters) and the frequency of $y(t)-A-B(t_c-t)^\alpha$ are consistent. A test method called the Lomb periodogram analysis is adopted to detect periodic oscillations for $y(t)-A-B(t_c-t)^\alpha$ and calculate its frequency.

It should be noted that the Lomb periodogram analysis method was chosen to detect periodic oscillation in preference to many existing methods (e.g., Shanks transform; (H; q)-analysis; Fourier transform) as they all have major weaknesses or limitations (Zhou and Sornette, 2002, 2003). For example, the Shanks transform method can be very subjective when identifying the critical time t_c and the Fourier transform assumes uniform time series. In comparison, the Lomb periodogram analysis method is not only able to objectively evaluate the critical time t_c or the turning point, but is also suitable for non-uniform time series.

The validation process starts with pre-setting the frequency series ($freq_i$) ($i=1,2,\ldots,M$) with M as the length of the pre-given frequency series. For a given frequency f, the power spectral density $P(f)$ can be computed by the Lomb periodogram analysis as below:

$$P(f)=\frac{1}{2\sigma^2}\left\{\frac{\left[\sum_{j=1}^{J}(x_j-\bar{x})\cos\left(2\pi f(t_j-\tau)\right)\right]^2}{\sum_{j=1}^{J}\cos^2\left(2\pi f(t_j-\tau)\right)}+\frac{\left[\sum_{j=1}^{J}(x_j-\bar{x})\sin\left(2\pi f(t_j-\tau)\right)\right]^2}{\sum_{j=1}^{J}\sin^2\left(2\pi f(t_j-\tau)\right)}\right\} \tag{7}$$

where $x_j = y_j-A-B(t_c-t_j)^\alpha$ at times t_j ($j=1,2,\ldots,J$); $\bar{x}=\frac{1}{J}\sum_{j=1}^{J}x_j$ and $\sigma^2=\frac{1}{J-1}\sum_{j=1}^{J}(x_j-\bar{x})^2$ are respectively the mean and the variance of x_j. The time offset, τ, is calculated by

$$\tau=\frac{1}{4\pi f}\arctan\frac{\sum_{j=1}^{J}\sin(4\pi f t_j)}{\sum_{j=1}^{J}\cos(4\pi f t_j)} \tag{8}$$

Invalid values are then removed from the resulting $P(freq_i)$ series ($i=1,2,\ldots,M$). These invalid values include: 1) $P(f_{mpf})$ corresponding to the most probable frequency (f_{mpf}), which is caused by the random series, and inversely proportional to the length of the given frequency series (J), $f_{mpf}\approx 1.5/J$; 2) the $P(freq_i)$ which is smaller than the critical value that is calculated by $z=-ln(1-(1-p)^{1/M})$, at the given statistical significance level of p. If there are no valid values in the $P(freq_i)$ series, the Lomb periodogram rejects the conclusion. In other words, the turning points predicted by the LPPL model are not statistically valid. Otherwise the frequency corresponding to the maximum valid values in the $P(freq_i)$ series is the result of the Lomb periodogram test.

The Lomb periodogram analysis can be briefly summarized as follows. First, an LPPL model corresponding to each subinterval is obtained. The Lomb periodogram analysis then computes the frequency value based on the periodic oscillations of the LPPL model. If the frequency value is close to the frequency ($\omega/2\pi$) optimized by the MPGA[1], the Lomb periodogram analysis concludes that the prediction by the LPPL model is effective. Otherwise, the predicted turning points are invalid and are thus deleted. Eventually only the turning points predicated by the LPPL models that pass the Lomb periodogram test are recorded.

4. Model calculation and result analysis

In this section, the data source and sample selection will be introduced; then the prediction results and comparison with other methods will be presented; and finally the management implications and applications of this model will be summarized and discussed.

4.1. Data description

The hypothesis of the LPPL model is that market crashes are caused by the slow buildup of long-range correlations between traders leading to a collapse of the market at one critical instant (Johansen et al., 1999, 2000). Thus the condition for the LPPL model is to have long enough history of data in order to forecast the turning points. Daily and weekly data of the WTI spot price at $/barrel in the period between April 1, 2003 and November 14, 2016 were collected from the U.S. Energy Information Administration (EIA) website. During this period, the WTI spot price experienced two major upturns and two major downturns. As shown in Fig. 1, the two upturns can be observed on July 3, 2008 and April 29, 2011 and the two downturns occurred approximately on December 23, 2008 and February 11, 2016. If any two adjacent turning points in history are too close, even though the LPPL model might be able to predict turning points, the Lomb periodogram test may still reject the

[1] If the absolute value of the difference between these two frequencies is <0.3, then they are close.

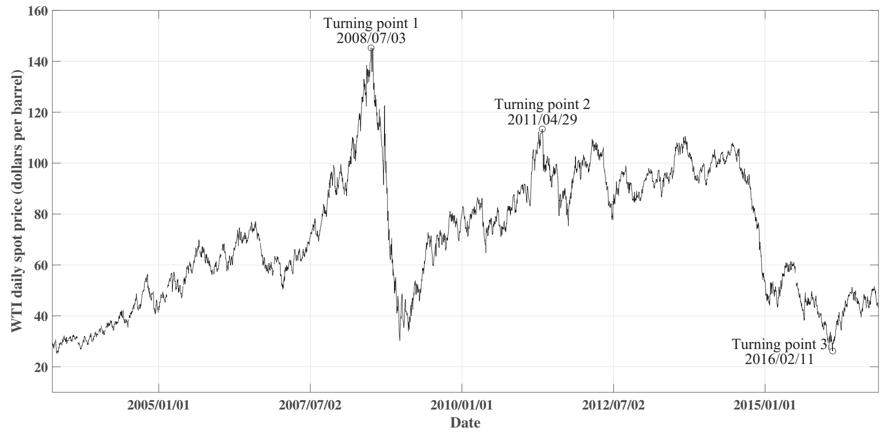
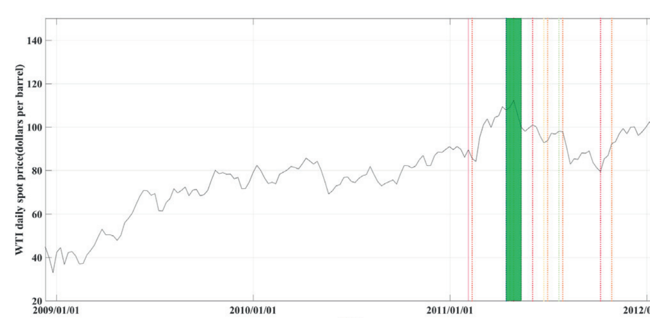


Fig. 1. The WTI spot price between April 1, 2003 and November 14, 2016.

(a) Prediction result of daily data of April 1, 2003 to January 2, 2008

(b) Prediction result of weekly data of April 1, 2003 to January 2, 2008

(c) Prediction result of daily data of February 1, 2007 to February 1, 2011



(d) Prediction result of weekly data of February 1, 2007 to February 1, 2011

(e) Prediction result of daily data of April 29, 2011 to August 1, 2015

(f) Prediction result of weekly data of April 29, 2011 to August 1, 2015

Fig. 2. Prediction result of the LPPL model based on the MPGA.

prediction due to lack of statistical significance. Therefore, our LPPL focuses on the prediction of the three major turning points on three dates: July 3 of 2008, April 29 of 2011 and February 11 of 2016, as marked in Fig. 1.

To predict future turning points by the LPPL model, 4-year length of sample data prior to the turning point were chosen, with the last data point a few months ahead of the turning points. For example, the first turning point occurred on July 3, 2008 and we select the data interval between April 1, 2003 and January 2, 2008 (four and half year data that ends 6-month ahead of the first turning point). Accordingly, similar data intervals were chosen for the prediction of the other two turning points.

4.2. Results from the LPPL model

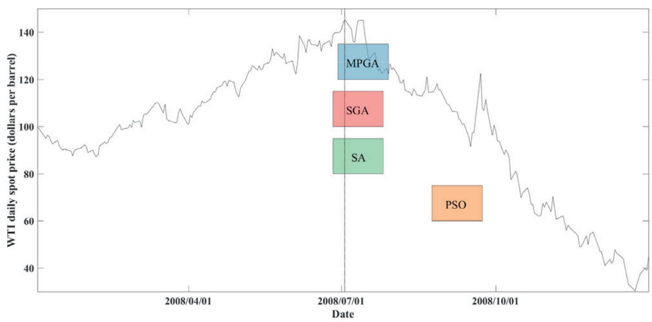
Both daily and weekly data from the WTI spot price were used to predict the three turning points mentioned in the previous section. The forecast results for the first turning point based on the data between April 1, 2003 and January 2, 2008 are shown in Fig. 2 (a) and (b); the predictions of the second turning period based on the data interval of February 1, 2007 and February 1, 2011 are shown in Fig. 2 (c) and (d); the forecast results for the third turning point based on the data interval of April 29, 2011 and August 1, 2015 are shown in Fig. 2 (e) and (f). Note that to implement the LPPL model, many subintervals in each data sample were generated (as described in Section 3) and for each subinterval, the LPPL model predicted a specific time or date for the future turning point. If the prediction algorithm is effective and robust, the predicted times for the turning points should be close to each other. As shown in Fig. 2, vertical dotted lines represent the times of future turning points predicted by the LPPL model generated by numerous subintervals. Note that only predictions that are statistically validated by the Lomb periodogram test are presented. In Fig. 2, the green areas indicate the time periods that the turning points would occur within a time window of 30 days or 4 weeks. Note that the turning point predictions within 30 days are based on the daily data and the predictions within 4 weeks are based on weekly data. In fact, the daily and weekly data provide very consistent results that prove the effectiveness of our optimized LPPL model in this paper. For example, in Fig. 2 (a), based on daily data, the LPPL model predicts that the first turning point would occur within a 30-day interval between June 29 and July 29 of 2008 and the weekly data shows that the turning point occurs during June 20 and July 18 of 2008. Since the first actual turning point occurs on July 3, 2008, the time windows predicted by our LPPL model is rather accurate. Table 1 lists the LPPL prediction results for the three turning points within 30 days and 4 weeks intervals based on both daily and weekly data, respectively.

In comparison with the historical data of the WTI spot price in Fig. 1, it was found that the prediction results are quite accurate as in all cases, the LPPL model successfully predicted the time windows when the true turning points occurred. For example, in Fig. 1, it can be seen that the exact date of the second turning point is marked as April 29, 2011. In Table 1, the prediction of the optimized LPPL model shows that the turning point would occur in the time windows of April 6, 2011–May 6, 2011 and April 15, 2011–May 13, 2011 for 30-day and 4-week intervals based on daily and weekly data respectively.

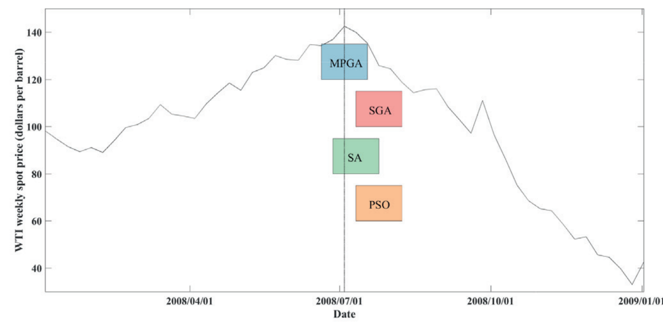
To further verify the prediction quality of the LPPL model, the results of our LPPL model that incorporates the MPGA to search for optimal parameters were compared with the results of the LPPL models that employed three other different techniques to search for optimal parameters, including standard genetic algorithm (SGA), simulated annealing (SA) and particle swarm optimization (PSO). The key characteristic of the genetic algorithm is to update solutions (or parameters in our problem) through selection, crossover and mutation. The important feature of the simulated annealing is to move one state to a neighboring state iteratively based on the probability assessment in order to find optimal solutions. The Particle swarm optimization searches for an optimal solution from a population of candidates that are dubbed as particles and the algorithm entails moving the particles around the search-space according to some simple mathematical formulas over the particles' position and velocity. The comparison results are presented in Fig. 3. In Fig. 3, the vertical black dotted lines in Fig. 3 (a) to (f) indicate the occurring dates of the historical turning points, which is obtained as follows: first of all, finding one crest (or trough) in the historical data of oil price, and then searching for the highest (or lowest) point around the crest (or trough), and the corresponding date is marked as the vertical black dotted line. For example, the vertical black dotted lines in Fig. 3 (a) and (b) indicate the first upward turning point on July 3, 2008 as shown in Fig. 1. The horizontal bars of four different colors represent the different time windows predicted by four different LPPL models that use four different techniques to search for optimal parameters. The bars represent the forecasted time windows of 30 days or 4 weeks. The blue bars indicate the forecast results of the time windows of the turning points by our LPPL model based on the MPGA; the red bars indicate the forecast results of the LPPL model based on the SGA; the green bars indicate the forecast results of the LPPL model based on the SA; the orange-color bars indicate the forecast results of the LPPL model based on the PSO. If the exact turning point dates represented by the vertical black dotted lines are within the bars, it means that the forecasting results are accurate since the bars provide the time windows of the occurrences of the turning points. From Fig. 3, it can be observed that the exact turning point dates represented by the vertical black dotted lines fall within all the blue-color bars that represent the 30-day or 4-week time windows predicted by our LPPL model. In contrast, a closer look at Fig. 3 (a) to (f) reveals that the PSO is probably the worst technique since none of orange-color bars resenting the PSO technique predict accurate time windows since all the orange-color bars are away from the vertical black dotted lines. For the SGA represented by the red-color bars, Fig. 3 (c) and (d) show that the prediction of the second turning points is away from the actual turning point for both daily and weekly data. For the SA, it is noticed that in Fig. 3 (d), this technique does not provide satisfactory results for the prediction of the second turning points based on the weekly data. The detailed description of the time windows by the four techniques is provided in Table 2. In Table 2, "True" indicates correct prediction of the time windows in which the turning points actually occurred, and "False" indicates false prediction of the time windows for the turning points to occur.

Table 1
Forecasting time range that the turning will most likely occur by the LPPL model based on the MPGA.

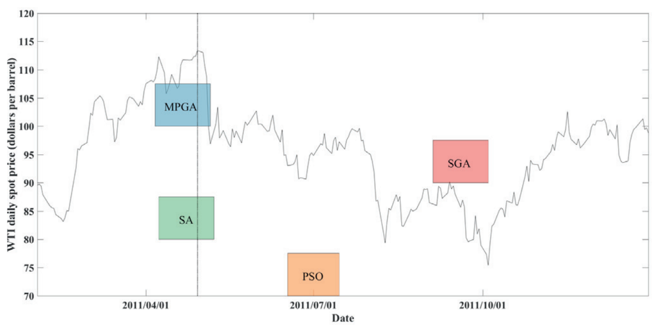
Time range	Forecast results		
	First upturn	Second upturn	Second downturn
Historical turning point	2008/07/03	2011/04/29	2016/02/11
30 days optimized by the MPGA	2008/06/29–2008/07/29	2011/04/06–2011/05/06	2016/02/10–2016/03/11
4 weeks optimized by the MPGA	2008/06/20–2008/07/18	2011/04/15–2011/05/13	2016/01/15–2016/02/12



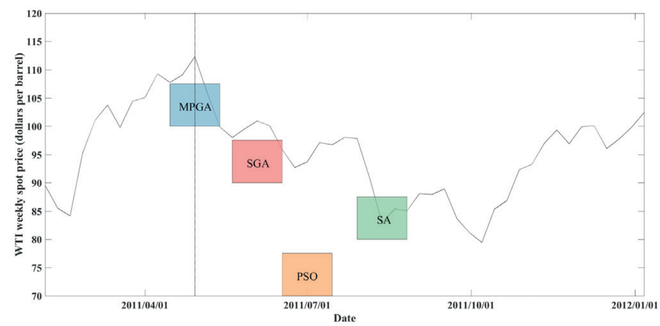
(a) Comparison of daily data of April 1, 2003 to January 2, 2008



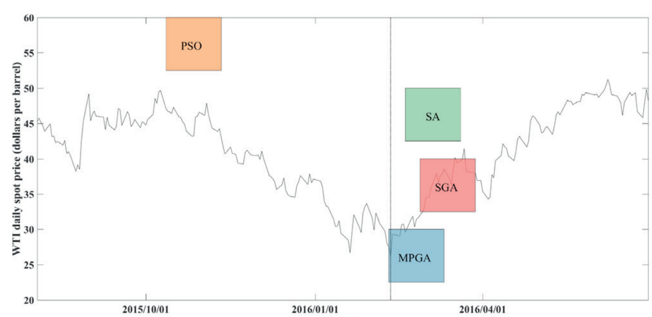
(b) Comparison of weekly data of April 1, 2003 to January 2, 2008



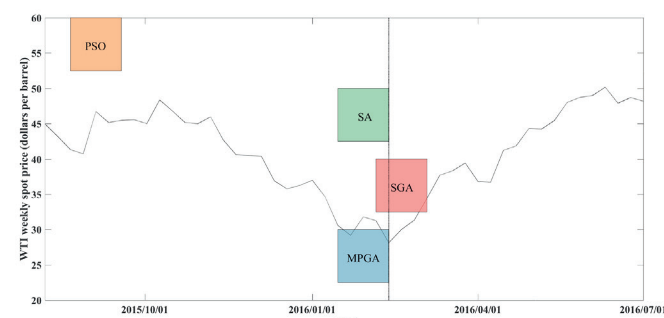
(c) Comparison of daily data of February 1, 2007 to February 1, 2011



(d) Comparison of weekly data of February 1, 2007 to February 1, 2011



(e) Comparison of daily data of April 29, 2011 to August 1, 2015



(f) Prediction result of weekly data of April 29, 2011 to August 1, 2015

Fig. 3. Comparison of the prediction results of different heuristic algorithm.

Table 2
Comparison of the prediction results of different heuristic algorithm.

Historical turning point	Turning point 1 (2008/07/03)	Turning point 2 (2011/04/29)	Turning point 3 (2016/02/11)
30 days optimized by the MPGA	True 2008/06/29–2008/07/29	True 2011/04/06–2011/05/06	True 2016/02/10–2016/03/11
4 weeks optimized by the MPGA	True 2008/06/20–2008/07/18	True 2011/04/15–2011/05/13	True 2016/01/15–2016/02/12
30 days optimized by the SGA	True 2008/06/26–2008/07/26	False 2011/09/04–2011/10/04	False 2016/02/27–2016/03/28
4 weeks optimized by the SGA	False 2008/07/11–2008/08/08	False 2011/05/20–2011/06/17	True 2016/02/05–2016/03/04
30 days optimized by the SA	True 2008/06/26–2008/07/26	True 2011/04/08–2011/05/08	False 2016/02/19–2016/03/20
4 weeks optimized by the SA	True 2008/06/27–2008/07/25	False 2011/07/29–2011/08/26	True 2016/01/15–2016/02/12
30 days optimized by the PSO	False 2008/08/24–2008/09/23	False 2011/06/17–2011/07/15	False 2015/10/12–2015/11/11
4 weeks optimized by the PSO	False 2008/07/11–2008/08/08	False 2011/06/17–2011/07/15	False 2015/08/21–2015/09/18

"True" indicates correct prediction of the time windows in which the turning points actually occurred.
"False" indicates false prediction of the time windows for the turning points to occur.

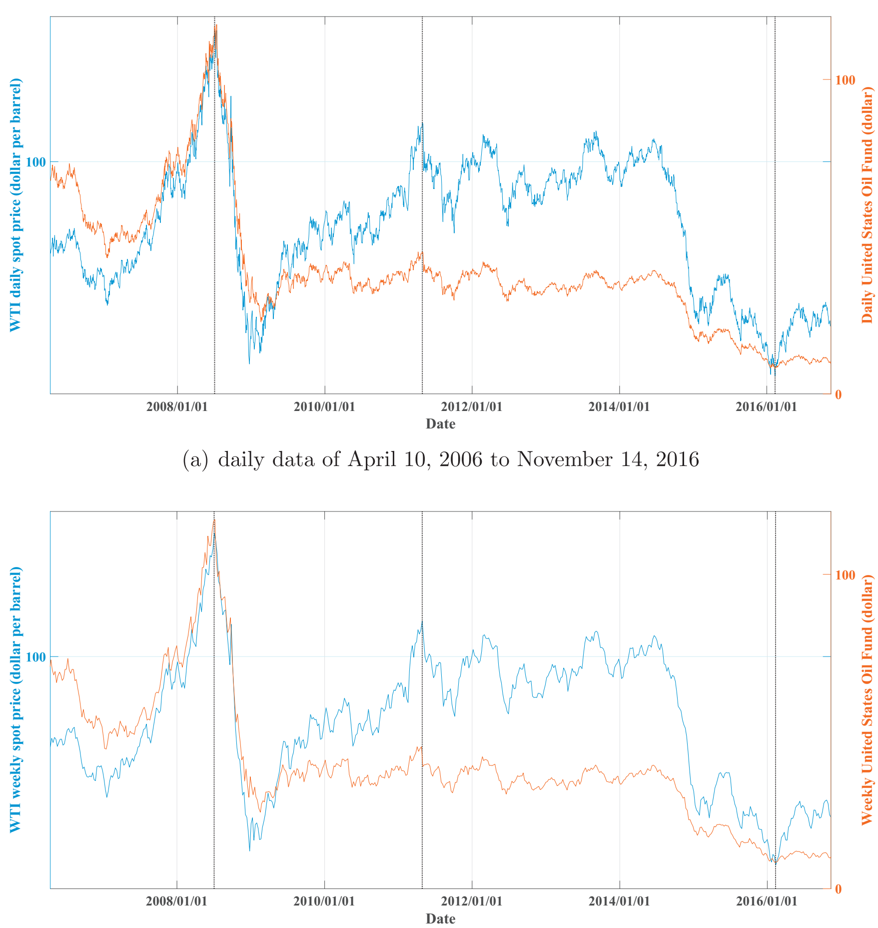
Table 2 shows that the successful predictions of the LPPL model based on the MPGA, SGA, SA and PSO is 6, 2, 4 and 0, respectively. Therefore, we conclude that the prediction results of our LPPL model that employs the MPGA are superior to those that use other techniques to search for optimal parameters in the LPPL model.

4.3. Managerial implications

In reality, organizations and companies spend millions of dollars and a lot of time trying to predict the fluctuations of the oil price. Predicting ahead of turning points of the oil price may help governments and companies come up with effective macroscopic energy reserve strategies to cope with peaks and bottoms of oil price. Clearly, the major challenges facing governments and companies arise in assessing the prospects for the price of oil in the future.

Nowadays, it is commonly accepted that crude oil prices exert a critical influence on the economic activity and, since the stock market acts as the barometer of the economy, oil prices are also likely to play a major role in the behavior of stock prices (Moya–Martínez et al., 2014). Oil market has significant volatility spillover effects on stock markets (Sadorsky, 1999; Cong et al., 2008; Arouri and Nguyen, 2010; Filis et al., 2011; Kumar et al., 2012; Awartani and Maghyereh, 2013; Gomes and Chaibi, 2014; Khalfaoui et al., 2015). Apergis and Miller (2009) also stated that oil price movements are important in explaining movements in stock returns. Furthermore, since 2006, oil future trading has been deregulated, allowing spot oil price to be increasingly determined by speculative future markets. This brings about some hidden issues which may have contributed to the outbreak of the financial crisis in 2008 (Sornette et al., 2009). Oil price movement is an important indicator for investors to make necessary investment decisions and for policy makers to adopt appropriate policies in managing stock markets (Adaramola, 2012). In that regard, the improved LPPL model can help investors and policy makers find possible future turning point and hedge against the risks for sudden increase or decrease in the oil price. Therefore, the improved LPPL model not only can help investors assess their investment opportunities, but also help the Securities and Futures Commission (SFC) to make strategies in order to stabilize stocks and other financial markets.

In particular, the prediction of turning points of oil price also has some implications of the stock price in energy sector in the financial market since the fluctuations in oil price will inevitably lead to the volatility in equity holdings of oil companies. To verify the link between the two, daily and weekly data of United States Oil (USO) Fund, an exchange-traded fund, in the period between April 10, 2006 and November 14, 2016 were retrieved from Yahoo Finance and analyzed. The turning points of the USO stock price were found to be almost identical to those of the WTI spot price. A correlation analysis revealed that the correlation coefficient between the USO fund and the WTI spot price is as high as 0.6029. Fig. 4 plots both the WTI spot price and the USO fund and shows that both prices oscillate almost synchronously. Furthermore, the LPPL model was used to predict the turning point of the United States Oil Fund. The forecast result for the United States Oil Fund between January 12, 2016 to February 11,



(a) daily data of April 10, 2006 to November 14, 2016

(b) weekly data of April 10, 2006 to November 14, 2016

Fig. 4. Comparison of the USO stock price and the WTI spot price.

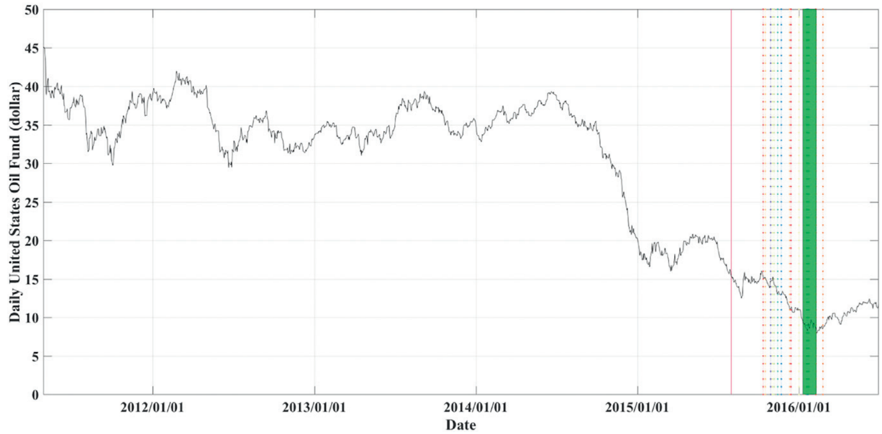


Fig. 5. Turning point forecast of the United States Oil Fund.

2016 is based on the data between April 29, 2011 and August 1, 2015 and is shown in Fig. 5. Fig. 6 presents the actual historical turning point that is marked as the vertical black dotted line, the predicted turning point for the WTI spot price that is marked as the blue bar, and the predicted turning point for the United States Oil Fund that is marked as the orange-color bar. Fig. 6 shows that 1) the predicted turning point for the United States Oil Fund verifies the validity of the predicted turning point for the WTI spot price; 2) these two predicted results are consistent with the actual historical turning points, proving the robustness of the improved LPPL method; 3) there is a close connection among the WTI spot price, the United States Oil Fund and the predicted turning point for the WTI spot price by the improved LPPL method.

Another interesting observation is derived from an online article by Nawar Alsaadi (oilprice.com) entitled "The oil market is at a major turning point". In the article, the author predicted March of 2017 to be a new turning point. Nevertheless, we conducted tests and experiments based on the data collected from the WTI spot price and we conclude that March of 2017 is not a major turning point. Two data samples were set up, both of which start from February 11, 2016 (the third major turning point in our research) and end on February 1, 2017 and March 1, 2017 respectively. By using daily data, our result did not suggest that March of 2017 is a major turning point. Fig. 7 provides the oil price in the whole month of March and it is clear that early March of 2017 is not a major turning point but a slight fluctuation.

5. Conclusions

In this research, we propose an improved log-periodic power law (LPPL) model to forecast the time windows of the turning points of oil price. The improved LPPL model is based on the multi-population genetic algorithm (MPGA) that helps search for optimal parameters in the LPPL model. The LPPL model was tested to predict the time windows of three major turning points in the period between April of 2003 and November of 2016. Then the Lomb periodogram was applied to statistically test the significance of the results generated from the LPPL model. To examine the quality of the improved LPPL model, extensive computational experiments were conducted by comparing our results with the results of other LPPL models that use the SA, SGA and PSO to search for parameters. It was found that the improved LPPL model with the MPGA is superior to other three LPPL models that employ different techniques to search optimal parameters since our LPPL model provides more accurate predictions.

As an interesting application of the optimized LPPL model, it was also predicted that the fluctuation of the oil price in March 2017 does

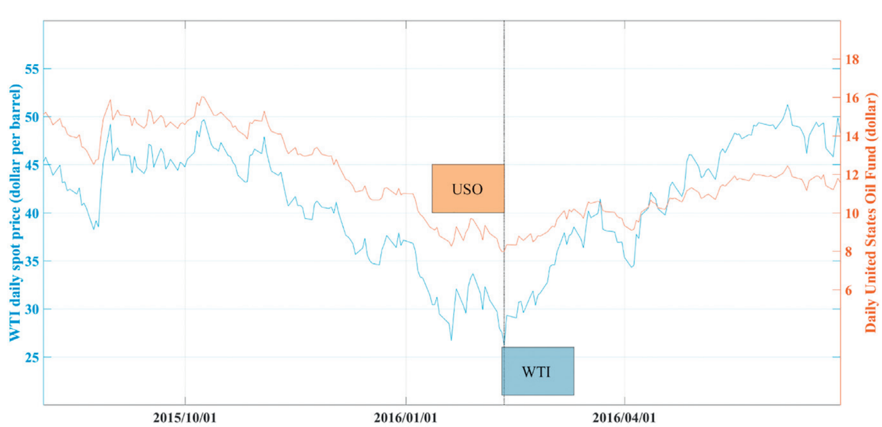


Fig. 6. Comparison of forecast results for the WTI and USO.

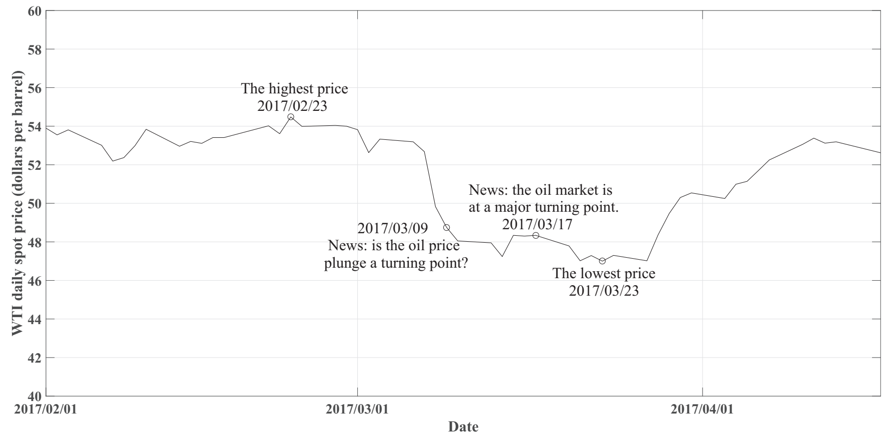


Fig. 7. The WTI spot price from February 1, 2017 to the present.

not correspond to a major turning point. Nevertheless, there may be a limitation of the application of the improved LPPL model since the model requires a relatively long period of data ahead of the turning points. Otherwise, the model may not accurately predict a turning point and may not be statistically significant. For example, between February 11, 2016 that is a major turning point and now, the length of time period is just over a year, and therefore, the one-year data is not sufficient to help predict another turning point in the future.

To successfully implement the improved LPPL model with the MPGA, a few additional remarks are noteworthy. First, more than four years of data are recommended as the sample data. This is because if the time period is too short, the Lomb periodogram test will not generate statistically significant results. Second, the starting point of the sample is preferably selected immediately after the previous major turning point. If the sample data is not sufficiently long, the starting point can move forward at most two years before the previous major turning point, otherwise it will affect the forecast results. Third, the ending point of the sample should not be chosen in the shock period of the WTI spot price because the predicted results will have large deviations. Fourth, when the predictions between daily data and weekly data are consistent, the predictions of the model are expected to be rather reliable.

Finally, we conclude that the improved LPPL has great potential to predict future turning points as long as sufficient data is available. In the future studies, we will explore more broad applications of the LPPL with the MPGA to other industries such as commodity prices and financial markets. The accurate prediction of turning points would help policy makers develop strategies to hedge against the potential risks of ups and downs of the prices and help stabilize the markets.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (71431004).

Appendix A. Proof: the optimization of the four nonlinear parameters in the LPPL model is an NP-hard problem

Definition. A class of function $\mathscr{F}$ is closed if it satisfies the following four conditions:

- $\mathscr{F}$ contains all linear functions;
- $\mathscr{F}$ is closed under addition, i.e., if $f \in \mathscr{F}$ and $g \in \mathscr{F}$, then $f + g \in \mathscr{F}$;
- $\mathscr{F}$ is closed under multiplication by a positive constant, i.e., if $f \in \mathscr{F}$ and $c > 0$, then $c \cdot f \in \mathscr{F}$;
- $\mathscr{F}$ is closed under linear substitution if whenever $f(x_1, \ldots, x_k) \in \mathscr{F}$ and c_{ik} are real numbers, we have $f(c_{10} + c_{11}x_1 + \cdots + c_{1n}x_n + \cdots + c_{k0} + c_{k1}x_1 + \cdots + c_{kn}x_n) \in \mathscr{F}$.

Proof. According to Kreinovich and Kearfott (2005), if a closed class of functions $\mathscr{F}$ contains at least one non-convex function and at least one nonlinear convex function, then for this class, the problem of finding the minimum of a given function $f \in \mathscr{F}$ on a given feasible region is NP-hard. Note that in Eq. (1), the power law term, $f_j = (t_c - t_j)^{\alpha}$, is a nonlinear convex function, and the periodic term $g_j = (t_c - t_j)^{\alpha} \cos[\omega ln(t_c - t_j) + \phi]$, is a non-convex function. Since Eq. (1) can be considered as a closed class and it has at least one nonlinear convex power law term and at least one non-convex periodic term, we thus conclude that the optimization of the four nonlinear parameters in the LPPL model is an NP-hard problem. □

Appendix B. Flowchart of the improved LPPL model

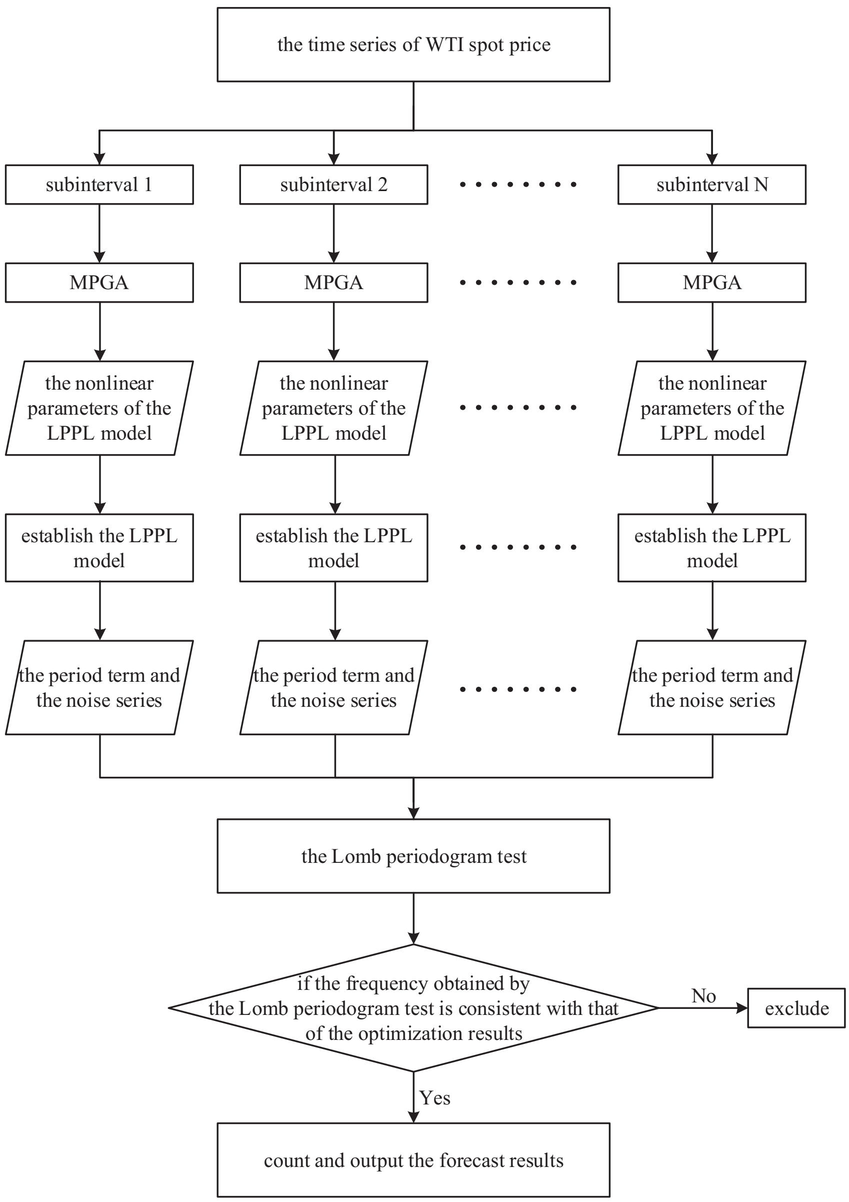


Fig. B.8. Process of the improved LPPL model.

Appendix C. Algorithm of the MPGA to optimize nonlinear parameters

This appendix describes in detail the algorithm of the MPGA to optimize nonlinear parameters (t_c, ω, ϕ and α) by the flowchart (see Fig. C.9) and pseudo-code (see Algorithm 1). The value ranges of the four nonlinear parameters, t_c, ω, ϕ and α, are predetermined; t_c is the day of a turning point to be predicted which may occur on the day after the subinterval to the future 10 years; ω is between 0 and 40; ϕ is between 0 and 2π; α is between 0.1 and 0.9. The proposed ranges of the parameter values follow the common practice for the MPGA (see Zhou, 2007). In addition, the crossover probability of each population is randomly generated between 0.7 and 0.9, and the mutation probability of each population is randomly generated between 0.001 and 0.05.

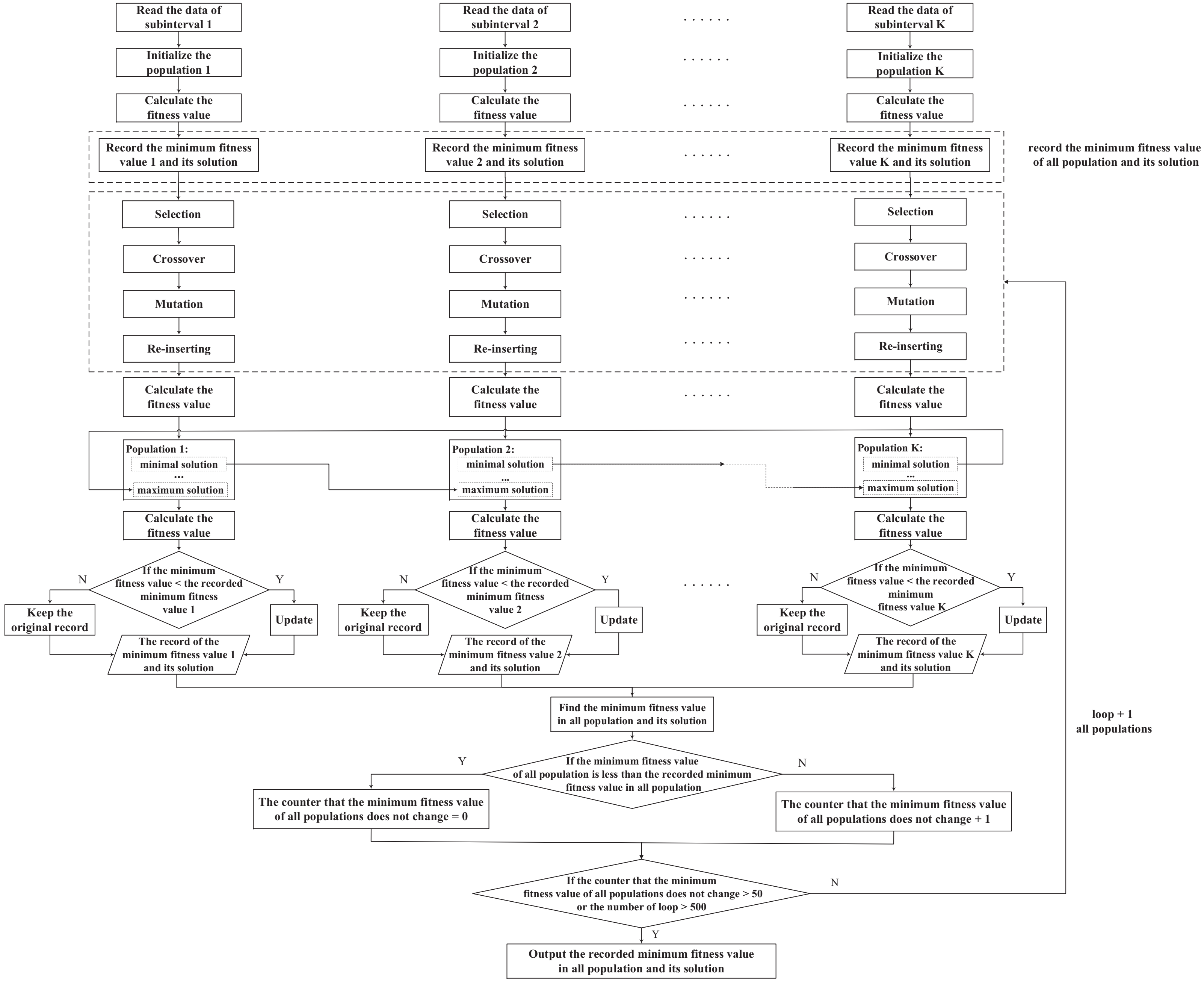


Fig. C.9. Flowchart of the MPGA to optimize nonlinear parameters.

Algorithm 1. Pseudo-code of the MPGA to optimize nonlinear parameters.

Read historical data;
Set the start time and end time of the sample, denoted as $time_{start}$ and $time_{end}$, respectively;
Predetermine the value ranges of the four nonlinear parameters, t_c, ω, ϕ and α (t_c is the day after the sample to the future 10 years; ω is between 0 and 40; ϕ is between 0 and 2π; α is between 0.1 and 0.9);
Predetermine the number of all populations, and let it equal 10;
Predetermine population size, and let it equal 100;
Predetermine the up bound of the total loop, denoted as $MaxGen$, and let it equal 500;
Predetermine the up bound that the minimum fitness value of all populations does not change, denoted as $StopGen$, and let it equal 50;
Predetermine the selection probability of each population, and let it equal 0.9;
Selecting the subintervals from the sample is as follows:
Predetermine moving step of start time of subintervals, denoted as $delta$, and let it equal the larger of $(time_{end} - time_{start}) * 0.75/$(three weeks) and three weeks;
for $subinterval_{end} = time_{end}$: - one week : $time_{end}-$ six weeks **do**
    for $subinterval_{start} = time_{start} : delta : time_{end} - (time_{end} - time_{start})/4$ **do**
        Get the subinterval $[subinterval_{start} \quad subinterval_{end}]$ from the sample $[time_{start} \quad time_{end}]$, where $subinterval_{start}$ is the start time of this subinterval; $subinterval_{end}$ is the end time of this subinterval;
        Randomly generate the crossover probability of each population between 0.7 and 0.9;
        Randomly generate the mutation probability of each population between 0.001 and 0.05;
        for each population **do**
            Initialize one population, which includes 100 chromosomes;
            Calculate the fitness value of each chromosome, according to Equation 6;
            Find the minimum fitness value of current population;
            Record this minimum fitness value, denoted as $bestObjV$, and its corresponding chromosome;
        end for
        Find and record the minimum fitness value of all populations;
        Add two counters, one is used to record the current number of loop, denoted as gen, and the other is used to record the number that the minimum fitness value of all populations does not change, denoted as $gen0$. Let gen equal 1, and let $gen0$ equal 0;
        while $gen0 < StopGen$ and $gen <= MaxGen$ **do**
            for each population **do**
                Perform the selection, crossover, and mutation and reinserting operation;
            end for
            Perform immigration operation;
            Find the minimum fitness value of each population and their corresponding chromosomes in the current loop;
            Find the minimum fitness value of all populations in the current loop, denoted as $newbestObjV$;
            if $newbestObjV < bestObjV$ **then**
                $bestObjV = newbestObjV$;
                $gen0 = 0$;
            else
                $gen0 = gen0 + 1$;
            end if
            $gen = gen + 1$;
        end while
        Save $bestObjV$ and its corresponding chromosome under the current subinterval;
    end for
end for

Appendix D. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eneco.2018.03.038.

References

Adaramola, A.O., 2012. Oil price shocks and stock market behavior: the Nigerian experience. J. Econ. 3 (1), 19–24.
Agnolucci, P., 2009. Volatility in crude oil futures: a comparison of the predictive ability of GARCH and implied volatility models. Energy Econ. 31 (2), 316–321.
Apergis, N., Miller, S.M., 2009. Do structural oil-market shocks affect stock prices? Energy Econ. 31 (4), 569–575.
Arouri, M.E.H., Nguyen, D.K., 2010. Oil prices, stock markets and portfolio investment: evidence from sector analysis in Europe over the last decade. Energy Policy 38 (8), 4528–4539.
Awartani, B., Maghyereh, A.I., 2013. Dynamic spillovers between oil and stock markets in the Gulf Cooperation Council countries. Energy Econ. 36, 28–42.
Barsky, R.B., Kilian, L., 2004. Oil and the macroeconomy since the 1970s. J. Econ. Perspect. 18 (4), 115–134.
Baumeister, C., Kilian, L., 2012. Real-time forecasts of the real price of oil. J. Bus. Econ. Stat. 30 (2), 326–336.
Baumeister, C., Kilian, L., 2016a. Forty years of oil price fluctuations: why the price of oil may still surprise us. J. Econ. Perspect. 30 (1), 139–160.
Baumeister, C., Kilian, L., 2016b. Understanding the decline in the price of oil since June 2014. J. Assoc. Environ. Resour. Econ. 3 (1), 131–158.
Baumeister, C., Kilian, L., Lee, T.K., 2014. Are there gains from pooling real-time oil price forecasts? Energy Econ. 46, S33–S43.
Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. J. Econ. 31 (3), 307–327.
Brigida, M., 2014. The switching relationship between natural gas and crude oil prices. Energy Econ. 43, 48–55.
Calvet, L., Fisher, A., 2001. Forecasting multifractal volatility. J. Econ. 105 (1), 27–58.
Calvet, L.E., Fisher, A.J., 2004. How to forecast long-run volatility: regime switching and the estimation of multifractal processes. J. Financ. Economet. 2 (1), 49–83.
Chen, H., Liao, H., Tang, B.J., Wei, Y.M., 2016. Impacts of OPEC's political risk on the international crude oil prices: an empirical analysis based on the SVAR models. Energy Econ. 57, 42–49.
Cheng, F.Z., Li, T., Wei, Y.M., Fan, T.J., 2018. The VEC-NAR model for short-term forecasting of oil prices. Energy Econ. https://doi.org/10.1016/j.eneco.2017.12.035.
Colgan, J.D., 2014. Oil, domestic politics, and international conflict. Energy Res. Soc. Sci. 1, 198–205.
Cong, R.G., Wei, Y.M., Jiao, J.L., Fan, Y., 2008. Relationships between oil price shocks and stock market: an empirical analysis from China. Energy Policy 36 (9), 3544–3553.
Eberhart, R., Kennedy, J., 1995. A new optimizer using particle swarm theory. Micro Machine and Human Science, 1995. MHS'95., Proceedings of the Sixth International Symposium on. IEEE., pp. 39–43.
Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50 (4), 987–1007.
Fantazzini, D., 2016. The oil price crash in 2014/15: was there a (negative) financial bubble? Energy Policy 96, 383–396.
Filimonov, V., Sornette, D., 2013. A stable and robust calibration scheme of the log-periodic power law model. Physica A 392 (17), 3698–3707.
Filis, G., Degiannakis, S., Floros, C., 2011. Dynamic correlation between stock market and oil prices: the case of oil-importing and oil-exporting countries. Int. Rev. Financ. Anal. 20 (3), 152–164.
Fomin, A., Korotayev, A., Zinkina, J., 2016. Negative Oil Price Bubble is Likely to Burst in March–May 2016. A Forecast on the Basis of the Law of Log-periodical Dynamics, arXiv preprint arXiv:1601.04341.
Gibson, R., Schwartz, E.S., 1990. Stochastic convenience yield and the pricing of oil contingent claims. J. Financ. 45 (3), 959–976.
Gomes, M., Chaibi, A., 2014. Volatility spillovers between oil prices and stock returns: a focus on frontier markets. J. Appl. Bus. Res. 30 (2), 509–526.
Graham, B., Dodd, D.L., 1934. Security Analysis: Principles and Technique. McGraw-Hill.
Hahn, W.J., DiLellio, J.A., Dyer, J.S., 2014. What do market-calibrated stochastic processes indicate about the long-term price of crude oil? Energy Econ. 44, 212–221.
Hamilton, J.D., 1983. Oil and the macroeconomy since World War II. J. Polit. Econ. 91 (2), 228–248.
Hamilton, J.D., 1996. This is what happened to the oil price-macroeconomy relationship. J. Monet. Econ. 38 (2), 215–220.
Hansen, P.R., Lunde, A., 2005. A forecast comparison of volatility models: does anything beat a GARCH (1, 1)? J. Appl. Econ. 20 (7), 873–889.
Haugom, E., Mydland, Ø., Pichler, A., 2016. Long term oil prices. Energy Econ. 58, 84–94.
Holland, J.H., 1975. Adaptation in Natural and Artificial Systems: An Introductory Analysis With Applications to Biology, Control, and Artificial Intelligence. University of Michigan Press, Ann Arbor, MI., pp. 439–444.
Hou, A., Suardi, S., 2012. A nonparametric GARCH model of crude oil price return volatility. Energy Econ. 34 (2), 618–626.
Hsu, T.K., Tsai, C.C., Cheng, K.L., 2016. Forecast of 2013–2025 crude oil prices: quadratic sine-curve trend model application. Energy Sources Part B 11 (3), 205–211.
Jacobsson, E., 2009. How to Predict Crashes in Financial Markets With the Log-Periodic Power Law. Master Diss. Department of Mathematical Statistics, Stockholm University.
Jiang, Z.Q., Zhou, W.X., Sornette, D., Woodard, R., Bastiaensen, K., Cauwels, P., 2010. Bubble diagnosis and prediction of the 2005–2007 and 2008–2009 Chinese stock market bubbles. J. Econ. Behav. Organ. 74 (3), 149–162.

Johansen, A., Ledoit, O., Sornette, D., 2000. Crashes as critical points. Int. J. Theor. Appl. Financ. 3 (2), 219–255.

Johansen, A., Sornette, D., 1999. Critical crashes. Risk 12 (1), 91–94.

Johansen, A., Sornette, D., 2000. Evaluation of the quantitative prediction of a trend reversal on the Japanese stock market in 1999. Int. J. Mod. Phys. C. 11 (2), 359–364.

Johansen, A., Sornette, D., Ledoit, O., 1999. Predicting financial crashes using discrete scale invariance. J. Risk 1 (4), 5–32.

Kesicki, F., 2010. The third oil price surge – what's different this time? Energy Policy 38 (3), 1596–1606.

Khalfaoui, R., Boutahar, M., Boubaker, H., 2015. Analyzing volatility spillovers and hedging between oil and stock markets: evidence from wavelet analysis. Energy Econ. 49, 540–549.

Khashman, A., Nwulu, N.I., 2011. Support vector machines versus back propagation algorithm for oil price prediction. International Symposium on Neural Networks. Springer, Berlin, Heidelberg, pp. 530–538.

Kilian, L., 2009. Not all oil price shocks are alike: disentangling demand and supply shocks in the crude oil market. Am. Econ. Rev. 99 (3), 1053–1069.

Kilian, L., Murphy, D.P., 2014. The role of inventories and speculative trading in the global market for crude oil. J. Appl. Econ. 29 (3), 454–478.

Kirkpatrick, S., 1984. Optimization by simulated annealing: quantitative studies. J. Stat. Phys. 34 (5), 975–986.

Kjärstad, J., Johnsson, F., 2009. Resources and future supply of oil. Energy Policy 37 (2), 441–464.

Klein, T., Walther, T., 2016. Oil price volatility forecast with mixture memory GARCH. Energy Econ. 58, 46–58.

Kreinovich, V., Kearfott, R.B., 2005. Beyond convex? Global optimization is feasible only for convex objective functions: a theorem. J. Glob. Optim. 33 (4), 617–624.

Kumar, S., Managi, S., Matsuda, A., 2012. Stock prices of clean energy firms, oil and carbon markets: a vector autoregressive analysis. Energy Econ. 34 (1), 215–226.

Lee, C.Y., Huh, S.Y., 2017. Forecasting long-term crude oil prices using a Bayesian model with informative priors. Sustainability 9 (2), 190. https://doi.org/10.3390/su9020190.

Liang, Q., Fan, Y., Wei, Y.M., 2005. A long-term trend forecasting approach for oil price based on wavelet analysis. Chin. J. Manag. Sci. 13 (1), 30–36.

Lux, T., Kaizoji, T., 2007. Forecasting volatility and volume in the Tokyo stock market: long memory, fractality and regime switching. J. Econ. Dyn. Control 31 (6), 1808–1843.

Mohammadi, H., Su, L., 2010. International evidence on crude oil price dynamics: applications of ARIMA-GARCH models. Energy Econ. 32 (5), 1001–1008.

Morana, C., 2001. A semiparametric approach to short-term oil price forecasting. Energy Econ. 23 (3), 325–338.

Morana, C., 2013. Oil price dynamics, macro-finance interactions and the role of financial speculation. J. Bank. Financ. 37 (1), 206–226.

Mork, K.A., 1989. Oil and the macroeconomy when prices go up and down: an extension of Hamilton's results. J. Polit. Econ. 97 (3), 740–744.

Moya-Martínez, P., Ferrer-Lapeña, R., Escribano-Sotos, F., 2014. Oil price risk in the Spanish stock market: an industry perspective. Econ. Model. 37 (574), 280–290.

Murat, A., Tokat, E., 2009. Forecasting oil price movements with crack spread futures. Energy Econ. 31 (1), 85–90.

Narayan, P.K., Narayan, S., 2007. Modelling oil price volatility. Energy Policy 35 (12), 6549–6553.

Pele, D.T., Mazurencu, M.M., Nijkamp, P., 2013. Herding Behaviour, Bubbles and Log Periodic Power Laws in Illiquid Stock Markets: A Case Study on the Bucharest Stock Exchange. https://doi.org/10.2139/ssrn.2305945.

Phillips, P., Shi, S., 2014. Financial Bubble Implosion. https://doi.org/10.2139/ssrn.2533015.

Phillips, P., Shi, S., Yu, J., 2015. Testing for multiple bubbles: historical episodes of exuberance and collapse in the S&P 500. Int. Econ. Rev. 56 (4), 1043–1078.

Refenes, A.N., Azema-Barac, M., Chen, L., Karoussos, S.A., 1993. Currency exchange rate prediction and neural network design strategies. Neural Comput. & Applic. 1 (1), 46–58.

Rehrl, T., Friedrich, R., 2006. Modelling long-term oil price and extraction with a Hubbert approach: the LOPEX model. Energy Policy 34 (15), 2413–2428.

Roehner, B.M., Sornette, D., 1998. The sharp peak-flat trough pattern and critical speculation. Eur. Phys. J. B 4 (3), 387–399.

Sadorsky, P., 1999. Oil price shocks and stock market activity. Energy Econ. 21 (5), 449–469.

Schwartz, E., Smith, J.E., 2000. Short-term variations and long-term dynamics in commodity prices. Manag. Sci. 46 (7), 893–911.

Sehgal, N., Pandey, K.K., 2015. Artificial intelligence methods for oil price forecasting: a review and evaluation. Energy Syst. 6 (4), 479–506.

Shafiee, S., Topal, E., 2010. A long-term view of worldwide fossil fuel prices. Appl. Energy 87 (3), 988–1000.

Sornette, D., 2002. Why Stock Markets Crash: Critical Events in Complex Financial Systems. Princeton University Press.

Sornette, D., Johansen, A., 2001. Significance of log-periodic precursors to financial crashes. Quant. Finan. 1 (4), 452–471.

Sornette, D., Woodard, R., Yan, W.F., Zhou, W.X., 2013. Clarifications to questions and criticisms on the Johansen-Ledoit-Sornette financial bubble model. Physica A 392 (19), 4417–4428.

Sornette, D., Woodard, R., Zhou, W.X., 2009. The 2006–2008 oil bubble: evidence of speculation, and prediction. Physica A 388 (8), 1571–1576.

Sornette, D., Zhou, W.X., 2002. The US 2000–2002 market descent: how much longer and deeper? Quant. Finan. 2 (6), 468–481.

Sornette, D., Zhou, W.X., 2004. Evidence of fueling of the 2000 new economy bubble by foreign capital inflow: implications for the future of the US economy and its stock market. Physica A 332, 412–440.

Sornette, D., Zhou, W.X., 2006. Predictability of large future changes in major financial indices. Int. J. Forecast. 22 (1), 153–168.

Sun, G., Shao, X., Jenkins, W.K., 2015. Performance of multi-parents genetic algorithms (MPGA) for IIR adaptive system identification. Circuits and Systems (MWSCAS), 2015 IEEE 58th International Midwest Symposium on. IEEE., pp. 1–4.

van Gysen, M., 2016. The Detection of Phase Transitions in the South African Market. (Doctoral dissertation, University of Cape Town).

Wang, Y., Wu, C., Wei, Y., 2011. Can GARCH-class models capture long memory in WTI crude oil markets? Econ. Model. 28 (3), 921–927.

Wang, Y.D., Wu, C.F., Yang, L., 2016. Forecasting crude oil market volatility: a Markov switching multifractal volatility approach. Int. J. Forecast. 32 (1), 1–9.

Wątorek, M., Drożdż, S., Oświęcimka, P., 2016. World financial 2014–2016 market bubbles: oil negative–US dollar positive. Acta Phys. Pol. A 129 (5), 932–936.

Weijermars, R., Sun, Z., 2018. Regression analysis of historic oil prices: a basis for future mean reversion price scenarios. Glob. Finance J. 35, 177–201.

Yan, W., Woodard, R., Sornette, D., 2010. Diagnosis and prediction of tipping points in financial markets: crashes and rebounds. Phys. Procedia. 3 (5), 1641–1657.

Yu, L.A., Wang, S.Y., Lai, K.K., 2008. Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. Energy Econ. 30 (5), 2623–2635.

Zhang, Y.J., Yao, T., 2016. Interpreting the movement of oil prices: driven by fundamentals or bubbles? Econ. Model. 55, 226–240.

Zhang, Y.J., Zhang, L., 2015. Interpreting the crude oil price movements: evidence from the Markov regime switching model. Applied Energy. 143, 96–109.

Zhao, L.T., Cheng, L., Wan, Y.T., Zhang, H., Zhang, Z.G., 2015. A VAR-SVM model for crude oil price forecasting. Int. J. Global Energy Issues 38 (1–3), 126–144.

Zhou, W.X., 2007. Jinrong Wulixue Daolun. 1st edition, Shanghai University of Finance and Economics Press.ISBN: 9787810988865, 7810988867.

Zhou, W.X., Sornette, D., 2002. Generalized *q* analysis of log-periodicity: applications to critical ruptures. Phys. Rev. E 66 (4), 046111.

Zhou, W.X., Sornette, D., 2003. Nonparametric analyses of log-periodic precursors to financial crashes. Int. J. Mod. Physics C 14 (8), 1107–1125.

Zhou, W.X., Sornette, D., 2005. Testing the stability of the 2000 US stock market "antibubble". Physica A: Statistical Mechanics and its Applications 348, 428–452.

Zhou, W.X., Sornette, D., 2006. Is there a real-estate bubble in the US? Physica A 361 (1), 297–308.

https://github.com/cfzngcfz/LPPL (retrieved on 11/17/2017).