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

### Abstract

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.

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.

### Intro

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).

The LPPL model was first proposed by Johansen, Ledoit and Sornette (Johansen and Sornette, 1999; Johansen et al., 1999, 2000; Sornette, 2002)

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 aswell(Yanetal.,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.

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 andKennedy,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.

### Literature review

Short-term forecast models can be roughly divided into two categories: statistical and econometric models, as well as artificial intelligence (AI) models. 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

Long term forecast : 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).

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)

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 explo sive process that is proposed by Phillips et al. (2015) and Phillips and Shi (2014).

The LPPL model is based on the interplay between economic theory and its rational expecta tion 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.

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

The LPPL models can be further divided into three types:

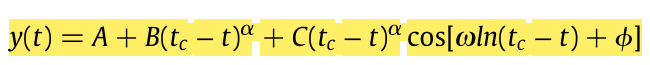
* simple log-periodic power law model
* Weierstrass-type LPPL model
* Landau-type LPPLmodel

Although the latter two models are studied by many scholars, they are mainly applied to stock price prediction.

The key in implementing the LPPL model is to search for optimal values for parameters. This search is nonlinear in nature. the parameters optimization of simple LPPL is an NP-hard problem.

### LPPL

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.



where y(t) is the oil price at time t,a is the exponential growth, y is the control of the amplitude of the oscillations and A,B, C and 0 are the parameters with no structural information. tc is a critical time or a turning point, to be predicated

One main feature captured by Eq. (1) is the dampened yet accelerated oscillation in oil prices. That is, as ttt approaches tct\_ctc​, the oscillations occur more frequently with decreasing amplitude. In other words, the term (tc−t)a(t\_c - t)^a(tc​−t)a represents the power-law behavior, which describes the faster-than-exponential change in prices due to positive feedback mechanisms. The term (tc−t)acos⁡[yln⁡(tc−t)+ϕ](t\_c - t)^a \cos[y \ln(t\_c - t) + \phi](tc​−t)acos[yln(tc​−t)+ϕ] represents the periodic component, which serves as a correction to the power-law term and exhibits the symmetry of discrete scale invariance. The most probable time for the turning point is when t=tct = t\_ct=tc​, for t≥tct \geq t\_ct≥tc​.

In the above LPPL model, seven parameters need to be optimized, including four nonlinear parameters (tc,y,ϕ,t\_c, y, \phi,tc​,y,ϕ, and aaa) and three linear parameters (A,B,A, B,A,B, and CCC). For simplicity, the linear parameters can be directly derived from the given nonlinear parameters using the least squares method.

This approach isproven to be very stable and able to yield good estimation of the linear parameters A,B and C.

However, finding the optimal values for the nonlinear parameters tc,y,ϕ,t\_c, y, \phi,tc​,y,ϕ, and aaa proves to be more challenging. In fact, it can be shown that searching for the optimal values of these four nonlinear parameters in the LPPL model is an NP-hard problem.

Framework of LLPM Model:

1. A sample interval is selected for predicting a turning point within the future time horizon. This interval is at least four years long and begins after the previous major turning point in history.
2. The selected sample interval is further divided into over 100 subintervals to avoid biases associated with specific sample intervals and to mitigate the impact of interval selection on the forecast results.
3. For each subinterval, the Multi-Population Genetic Algorithm (MPGA) is employed to optimize the parameters in the LPPL model. The optimized LPPL model is then used to predict a future date for the occurrence of a turning point.
4. Lomb periodogram analysis is conducted to statistically test the predicted turning points obtained from the LPPL models across all subintervals.
5. Turning points that are statistically validated by the Lomb periodogram analysis are considered as the predicted turning points by the LPPL model.

### Multi-populationgeneticalgorithm(MPGA)

The MPGA operates on multiple populations with the objective of evaluating each subinterval. After the initial populations are generated, if the optimization criteria are not satisfied, new populations are created, and the search process restarts.

The first step of the MPGA is to generate multiple populations.

Based on Eq. (1), the MPGA measures the fitness value of each chromosome (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 y(t) and the results from the LPPL model.

Une image contenant texte, Police, blanc, écriture manuscrite

Description générée automatiquement

### Validation using Lomb periodigram analysis

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 (y2π)\left( \frac{y}{2\pi} \right)(2πy​) optimized by the MPGA, the Lomb periodogram analysis concludes that the prediction by the LPPL model is effective. Otherwise, the predicted turning points are considered invalid and are deleted. Ultimately, only the turning points predicted by the LPPL models that pass the Lomb periodogram test are recorded.

### Model Calculation

Daily and weekly data of the WTI spot price (in $/barrel) for 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.

To predict future turning points using the LPPL model, a 4-year sample of data prior to the turning point was chosen, with the last data point a few months before the turning point.

Une image contenant texte, diagramme, Police, Plan

Description générée automatiquement