Energy Economics 49 (2015) 649–659

Contents lists available at ScienceDirect

Energy Economics

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

A novel hybrid method for crude oil price forecasting

Jin-Liang Zhang [a,d], Yue-Jun Zhang [b,c,*], Lu Zhang [d]

[a] School of Economics and Management, North China Electric Power University, Beijing 102206, PR China
[b] Business School of Hunan University, Changsha 410082, PR China
[c] Center for Resource and Environmental Management, Hunan University, Changsha 410082, PR China
[d] Center for Energy and Environmental Policy Research, Beijing Institute of Technology, Beijing 100081, PR China

ARTICLE INFO

Article history:
Received 11 September 2014
Received in revised form 25 December 2014
Accepted 17 February 2015
Available online 10 March 2015

JEL classification:
D53
G17
Q02
Q47

Keywords:
Crude oil price forecasting
EEMD
PSO
LSSVM
GARCH

ABSTRACT

Forecasting crude oil price is a challenging task. Given the nonlinear and time-varying characteristics of international crude oil prices, we propose a novel hybrid method to forecast crude oil prices. First, we use the ensemble empirical mode decomposition (EEMD) method to decompose international crude oil price into a series of independent intrinsic mode functions (IMFs) and the residual term. Then, the least square support vector machine together with the particle swarm optimization (LSSVM–PSO) method and the generalized autoregressive conditional heteroskedasticity (GARCH) model are developed to forecast the nonlinear and time-varying components of crude oil prices, respectively. Next, the forecasted crude oil prices of each component are summed as the final forecasted results of crude oil prices. The results show that, the newly proposed hybrid method has a strong forecasting capability for crude oil prices, due to its excellent performance in adaptation to the random sample selection, data frequency and structural breaks in samples. Furthermore, the comparison results also indicate that the new method proves superior in forecasting accuracy to those well-recognized methods for crude oil price forecasting.

1. Introduction

Crude oil is a kind of indispensable basic energy source, chemical material and strategic resource in socio-economic development. The changes of crude oil prices may significantly affect economic development, social stability and even national security in a country (Wu and Zhang, 2014). Therefore, it is of great significance to design scientific methods to accurately forecast crude oil price movement as much as possible, so as to address crude oil market extreme risks and find profit-making opportunities.

However, due to the confluent influence of crude oil market supply and demand, US dollar exchange rate, speculative trading, geopolitical conflicts, natural disasters etc., international crude oil prices have boomed and plummeted within 30–150 dollars per barrel in the past decade, with extreme market risks (Zhang et al., 2008b; Zhang and Wei, 2011; Zhang, 2013). Historical data shows that the complex volatility characteristics of international crude oil prices, such as nonlinearity, uncertainty and dynamics, make crude oil price forecasting difficult and the forecasting results bear high uncertainties, which may eventually cause significant uncertainties for the returns of related investors and the steady development of social economy.

Moreover, it should be noted that the past literature about crude oil price forecasting often shows that crude oil price forecasting results are sensitive to the modeling sample interval selection, sample data frequency and sample structural breaks (i.e., outliers) etc. (Yu et al., 2008; Liu and Shi, 2013; Chen et al., 2014). As a result, it is a huge challenge for crude oil price forecasting work to design a reliable forecasting method to enhance the adaptability to these factors.

Under this circumstance, in this paper, we propose a new hybrid model for crude oil price forecasting, based on the advantages of econometric models and soft-computing methods in depicting the nonlinear, dynamic features of crude oil prices. Specifically, the new method for crude oil price forecasting is a hybrid of the ensemble empirical mode decomposition (EEMD), particle swarm optimization (PSO), least squares support vector machine (LSSVM) and generalized autoregressive conditional heteroskedasticity (GARCH) model. Meanwhile, in order to validate the robustness of the newly proposed method for crude oil price forecasting, we consider not only the randomness of sample selection, the frequency of sample selection (such as the daily, weekly and monthly observations), but also the adaptability to the structural breaks. In addition, we compare the new method with previously well-known methods in terms of crude oil price forecasting accuracy. The results show that the new method has a nice robustness faced with those sensitive factors and outperforms those popular methods with respect to forecasting accuracy.

* Corresponding author. Tel./fax: +86 731 88822899.
E-mail address: zyjmis@126.com (Y.-J. Zhang).

http://dx.doi.org/10.1016/j.eneco.2015.02.018
0140-9883/

Table 1
Typical literature using the (G)ARCH type models to forecast crude oil prices or volatility.

Typical literature	Forecasting models	Main results
Sadorsky (2006)	Several GARCH type models	The TGARCH model is suitable for heating oil and natural gas while the GARCH model is appropriate for crude oil and unleaded gasoline.
Narayan and Narayan (2007)	GARCH model	Across the various sub-samples, there is inconsistent evidence of asymmetry and persistence of crude oil price shocks; and over the full sample period, the shocks have permanent effects, and asymmetric effects, on volatility.
Fan et al. (2008b)	GED-GARCH model	The GED-GARCH model has superior power in the out-of-sample forecast compared with the popular HSAF method.
Cheong (2009)	A series of ARCH models	An asymmetric long-term ARCH model has a higher forecasting ability for WTI market than Brent while in the Brent market, the GARCH model has the best forecasting performance.
Kang et al. (2009)	CGARCH and FIGARCH models	As for the WTI crude oil price volatility, CGARCH and IGARCH models are better than GARCH and FIGARCH models, respectively.
Wei et al. (2010)	GARCH models	There is not a model to win in all forecasts; however, overall the nonlinear GARCH model shows better forecasting power, especially for the long-term forecast.
Mohammadi and Su (2010)	GARCH, EGARCH, APARCH and FIGARCH models	In most cases, the APARCH model has better forecasting performance than other models.
Wang and Wu (2012)	GARCH models	The multivariate GARCH model has better forecasting performance than the univariate GARCH model overall, and the univariate GARCH model should be used if the crack spread is forecasted.
Hou and Suardi (2012)	Non-parametric GARCH model	The non-parametric GARCH model has a better forecast than the traditional GARCH model.
Arouri et al. (2012)	Three GARCH type models	Long memory is effectively present in all the series considered and a FIGARCH model seems to better fit the data, but the degree of volatility persistence diminishes significantly after adjusting for structural breaks.

The rest of the paper is organized as follows. Section 2 reviews related literature. Section 3 presents crude oil price forecasting methodology and data definitions. Section 4 puts forward crude oil forecasting results and some discussions. Finally, Section 5 concludes the paper and outlines the future work.

2. Related literature review

Crude oil price is the core of crude oil markets, and crude oil price forecasting proves an important determinant in the management of most industrial sectors across the world (Shin et al., 2013) and consequently becomes a hot spot in oil finance research. Up to now, there has been a raft of literature discussing crude oil price forecasting. Overall, previous literature mainly uses the econometric models (such as ARCH type model, VAR model), soft-computing models (such as neural networks, support vector machine) and wavelet technique. It should be noted that crude oil price forecasting methods have been increasing and forecasting performance has been improving in the past years. Here we would like to review related literature from the perspective of crude oil price forecasting methods, which may present a good foundation for the research in this paper.

First, there are a wealth of literature using the (G)ARCH type models to forecast crude oil price or its volatility and great progress has been achieved. This is mainly due to the fact the (G)ARCH type models have evident advantages in capturing the time-varying variance or volatility. Meanwhile, a body of literature compares the forecasting power of different types of ARCH or GARCH models (Fan et al., 2008b; Agnolucci, 2009). Some typical literature forecasting crude oil price or its volatility by means of the (G)ARCH type models are summarized in Table 1. It should be noted that the combined forecasting models between GARCH type models and other models, such as the implied volatility (IV) model, stochastic volatility (SV) model and support vector machine (SVM) model, tend to have superior performance. In fact, this is a promising trend for crude oil price forecasting. In other words, the hybrid forecasting models are more likely to be advocated in recent literature, which also gives some hints for our research in this paper.

Second, the neural network method has also been frequently used to forecast crude oil prices in recent years, which can be found in Table 2. It should be noted that the neural network method also has its evident disadvantages, such as over-fitting, local minima and weak generalization capability. Therefore, in the future research, the hybrid forecasting models incorporating the neural network method tend to be more appreciated. In fact, from existing related typical literature in Table 2, we have found this characteristic.

Third, the support vector machine (SVM) model has played a great role in crude oil price forecasting. Support vector machine is a new machine learning algorithm based on the statistical learning theory, and is particularly suitable for modeling with small sample size and nonlinear problems. [1]Due to the adoption of structural risk minimization (SRM) standards, SVM often has better learning performance and generalization capability compared to the traditional methods based on experience minimization. As a result, the SVM model has been extensively used for crude oil price forecasting, and some typical related literature are shown in Table 3.

Fourth, many scholars also try to forecast crude oil prices using the wavelet technique, which proves to have excellent performance in accuracy. For instance, the typical related literature shown in Table 4 basically testifies the merits of the wavelet technique in crude oil price forecasting. However, it also has some shortcomings in crude oil price forecasting. For instance, Yousefi et al. (2005) introduce a method based on the wavelet technique to forecast crude oil prices at four different time scales, i.e., 1 month, 2 months, 3 months and 4 months, and the results indicate that the wavelet technique is sensitive to the sample size. Therefore, how to avoid its sensitivity to the uncertain factors in modeling and enhance the forecasting robustness becomes a key issue for the wavelet technique.

In addition, a large number of studies use the hybrid method to forecast crude oil prices, which may combine the methods mentioned above. Some typical literature regarding the hybrid methods for crude oil price forecasting can be found in Table 5. Overall, the hybrid methods often imply the combination of interdisciplinary methods to use their strengths and can be roughly classified into three categories: (1) the combination among soft-computing methods, such as the intelligent optimization algorithms, EMD, SVM; (2) the combination among econometric methods, such as GARCH, ARIMA; and (3) the combination of soft-computing methods and econometric methods, such as the combination of GARCH type model and neural network method. These combining methods provide important hints for our study.

In summary, although international crude oil market proves a complex system, with rich multi-dimensional, nonlinear, dynamic features in crude oil price movement, the existing literature has accumulated a lot of experience in forecasting crude oil prices. In particular, the forecasting methods have been constantly emerging and the forecasting

[1] It should be noted that in industry, SVM does have been used to handle some machine learning problems in the big data environment.