Energy Economics 40 (2013) 405–415

Contents lists available at ScienceDirect

Energy Economics

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

Beyond one-step-ahead forecasting: Evaluation of alternative multi-step-ahead forecasting models for crude oil prices

Tao Xiong, Yukun Bao *, Zhongyi Hu

Department of Management Science and Information Systems, School of Management, Huazhong University of Science and Technology, Wuhan 430074, PR China

ARTICLE INFO

Article history:
Received 13 October 2012
Received in revised form 22 July 2013
Accepted 30 July 2013
Available online 16 August 2013

Keywords:
Crude oil price forecasting
Multi-step-ahead forecasting
EMD-based modeling framework
End effect
Prediction strategy

ABSTRACT

An accurate prediction of crude oil prices over long future horizons is challenging and of great interest to governments, enterprises, and investors. This paper proposes a revised hybrid model built upon empirical mode decomposition (EMD) based on the feed-forward neural network (FNN) modeling framework incorporating the slope-based method (SBM), which is capable of capturing the complex dynamic of crude oil prices. Three commonly used multi-step-ahead prediction strategies proposed in the literature, including iterated strategy, direct strategy, and MIMO (multiple-input multiple-output) strategy, are examined and compared, and practical considerations for the selection of a prediction strategy for multi-step-ahead forecasting relating to crude oil prices are identified. The weekly data from the WTI (West Texas Intermediate) crude oil spot price are used to compare the performance of the alternative models under the EMD–SBM–FNN modeling framework with selected counterparts. The quantitative and comprehensive assessments are performed on the basis of prediction accuracy and computational cost. The results obtained in this study indicate that the proposed EMD–SBM–FNN model using the MIMO strategy is the best in terms of prediction accuracy with accredited computational load.

1. Introduction

According to a report submitted by the International Energy Outlook of 2011 (IEO 2011),[1] world energy consumption will increase by 53%, from 505 quadrillion btu in 2008 to 770 quadrillion btu in 2035. World crude oil prices, one of the major determinants of energy consumption, have been exceptionally volatile over the past several years, reaching a high of $145 in July 2008 (daily spot price in nominal dollars) and a low of $30 in December 2008. More recently, growing demand from developing economies and unrest in many oil-supplying nations of the Middle East and North Africa have supported price (that is, the West Texas Intermediate crude oil spot price) increases from an average of $62 per barrel in 2009 to $79 per barrel in 2010 and then $95 per barrel in 2011 (EIA, 2012). Considering the rapid increase in international crude oil demand and the high volatility of crude oil prices over the last few decades, accurate predictions of future trends of crude oil prices are of particular importance to energy managers and analysts, both in the public and private sectors.

An extensive literature investigation reveals that it is not difficult to find that great research efforts have been expended to explore the underlying dynamics of crude oil prices and develop models suitable for forecasting crude oil prices (Azadeh et al., 2012; Ghaffari and Zare, 2009; He et al., 2012; Hou and Suardi, 2012; Knetsch, 2007; Mohammadi and Su, 2010; Narayan and Narayan, 2007; Wang et al., 2012; Xu and Ouenniche, 2011). For example, Ghaffari and Zare (2009) developed a method based on soft computing approaches to predict the daily variation of the WTI crude oil price. To improve the forecasting accuracy of crude oil prices with a deeper understanding of the market microstructure, He et al. (2012) proposed a wavelet decomposed ensemble model for crude oil price analysis and forecasting. Narayan and Narayan (2007) used a generalized autoregressive conditional heteroskedasticity (GARCH) model to forecast crude oil price volatility. However, an important point to note from past studies is their preoccupation with one-step-ahead forecasting rather than multi-step-ahead forecasting. In the one-step-ahead case, the predictor uses all or some of the observations to estimate a variable of interests for the time-step immediately following the latest observation. However, it provides no information as to the long-term future behavior of crude oil prices. For this reason, one may justifiably argue that the comparison of alternative one-step-ahead forecasting is of limited use. For an improved and more meaningful evaluation of the performances of prediction models, the long-term future behavior of crude oil prices should be taken into account explicitly.

Our focus in this study is on multi-step-ahead forecasting of crude oil prices. A multi-step-ahead forecasting extrapolates the crude oil price series by predicting many time-steps into the future without the availability of outputs in the horizon of interest. Despite the influence of many complicated factors, oil prices appear highly nonlinear and even chaotic as Panas and Ninni (2000) noted, which makes it rather difficult to forecast future oil prices, especially in the multi-step-ahead case (Fan et al., 2008b). Multi-step-ahead forecasts of crude oil prices are of greater value to decision-makers in the energy industry than one-

* Corresponding author. Tel.: +86 27 87558579; fax: +86 27 87556437.
E-mail addresses: yukunbao@hust.edu.cn, y.bao@ieee.org (Y. Bao).

[1] http://www.eia.gov/forecasts/ieo/index.cfm

0140-9883/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.eneco.2013.07.028

step-ahead ones and should be used more widely by practitioners and government agencies in their decision-making related to oil-related investments, risk management and portfolio allocation because they allow for a thorough evaluation of the future behavior of crude oil prices. Nevertheless, unlike one-step-ahead forecasting, multi-step-ahead forecasting faces a typically growing amount of uncertainties arising from various sources. For instance, an accumulation of errors and lack of information make multi-step-ahead forecasting more difficult (Weigend and Gershenfeld, 1995). As such, multi-step-ahead forecasting has been a major research topic that has significant practical implications. Generally speaking, two fundamental issues must be addressed when establishing models for multi-step-ahead forecasting. The first is prediction strategy, and the second is the selection of modeling techniques.

As for prediction strategy, currently iterated strategy, direct strategy, and multiple-input multiple-output (MIMO) strategy have been proposed, and the first two have been widely utilized in the literature. Studies regarding prediction strategies for multi-step-ahead forecasting have been investigated in the energy market (Kusiak et al., 2009; Sorjamaa et al., 2007; Tikka and Hollmén, 2008). Although past studies have clarified the need for multi-step-ahead forecasting, there has been very little, if any, effort to evaluate the performance of different strategies, particularly in the context of crude oil price prediction. As a major contribution, this study comparatively examines the performance of the three frequently used strategies (mentioned above) for multi-step-ahead forecasting in the case of crude oil prices.

After prediction strategy is determined, forecasters are faced with a choice among various modeling techniques, which has been another major research topic in the context of multi-step-ahead forecasting. According to an extensive literature investigation, some popular modeling techniques, such as autoregressive conditional heteroskedasticity (ARCH)-type models (Cheong, 2009), vector autoregressive (VAR) models (Baumeister and Kilian, 2012), support vector machines (SVMs) (El-Sebakhy, 2009), and neural networks (NNs) (Mirmirani and Li, 2004) have been successfully applied to crude oil market modeling and forecasting. However, the first two models do not conduct multi-step-ahead forecasting with the three leading prediction strategies simultaneously. (In fact, we have never found an ARCH-type model using either the direct strategy or the MIMO strategy, or SVMs using the MIMO strategy in any published work). Although NNs, because of a flexible structure, do indeed adopt all three prediction strategies, they often suffer from local minima and over-fitting. To remedy these shortcomings, an empirical mode decomposition (EMD)-based feed-forward neural network (FNN) modeling framework (EMD–FNN for short), following the philosophy of "divide and conquer", has been recently established and justified in crude oil price prediction (Yu et al., 2008). With regard to this emerging hybrid modeling technique, the present study goes further in exploring its prediction performance in the case of multi-step-ahead forecasting, especially for crude oil price forecasting. Furthermore, a technical improvement on EMD–FNN with the slope-based method (SBM) is proposed to restrain the end effect that occurred during the sifting process of EMD, which always negatively impacts the modeling quality as well as overall prediction performance when employing EMD–FNN for time series prediction. For ease of reference, EMD–SBM–FNN is hereafter used to indicate the proposed modeling framework. Therefore, for comparison purposes, two counterparts, including FNN and EMD–FNN, are selected as benchmarks. The naïve random walk approach has been widely taken as the benchmark in economic forecasting (Baumeister and Kilian, 2012; Hooper et al., 2008) and has been found to be difficult to beat in forecasting asset price (Kilian and Taylor, 2003). Thus, the naïve random walk is also selected as a benchmark in this study.

In summary, this paper proposes a revised EMD-based FNN modeling framework for multi-step-ahead forecasting of crude oil prices; it then goes a step further by investigating the performance of the three multi-step-ahead prediction strategies, including iterated strategy, direct strategy, and MIMO strategy, in the context of crude oil price forecasting. The weekly data from the West Texas Intermediate (WTI) crude oil spot price series are used as an experimental data series for the purpose of validation. The experimental results are judged on the basis of prediction accuracy and computational cost. The first contribution of this paper is that the proposed modeling framework is capable of capturing the complex dynamic of crude oil prices, resulting in higher accuracy in multi-step-ahead forecasting. The second contribution is to provide the first strong empirical evidence within the crude oil price forecasting literature on whether the superiority of an EMD-based FNN modeling framework holds consistently in the case of multi-step-ahead forecasting. Finally, the third contribution is to examine the unresolved empirical question in the literature on whether multi-step-ahead forecasting relating to the energy market is the prediction strategy that should be preferred in practice.

This paper is structured as follows. In Section 2, we provide a brief review of the literature on multi-step-ahead forecasting in the crude oil/energy market. Then, the prediction strategies, the proposed EMD–SBM–FNN modeling framework, and the benchmark models used in this study are discussed in detail in Section 3. Section 4 describes the research design, including the data source, data preprocessing, accuracy measures, input selection, and experimental procedures. In Section 5, the experimental results are discussed. Section 6 concludes.

2. Literature review

This section presents a brief review of the relevant literature related to multi-step-ahead forecasting in the crude oil price/energy market.

Crude oil price forecasting has been a challenging topic in the field of energy market research. Although multi-step-ahead forecasting can be invaluable to decision-makers in both government and industry, it is easy to see that there are few studies on multi-step-ahead crude oil price forecasting (Alquist et al., 2011; Fan et al., 2008b; He et al., 2010; Jammazi and Aloui, 2012; Ye et al., 2006; Yousefi et al., 2005; Yu et al., 2008). For example, Fan et al. (2008b) proposed a new genetic algorithm based on a generalized pattern matching approach to multi-step prediction of crude oil prices. Yu et al. (2008) presented an empirical mode decomposition based on a neural network ensemble-learning paradigm for world crude oil spot price forecasting. In contrast to most studies in energy price forecasting literatures, which typically use monthly average or close-to-close daily price data, He et al. (2010) explored the interaction between the daily high and low and the associated daily range of crude oil price forecasts. Yousefi et al. (2005) introduced a wavelet-based prediction procedure to provide forecasts of crude oil prices over different forecasting horizons. To explore the nonlinear relationship between inventory level and commodity prices, two nonlinear inventory variables were investigated by Ye et al. (2006) for their impact on forecasting ability of crude oil prices in the short term. Combining the dynamic properties of a multilayer back propagation neural network (MPNN) and the recent Haar A trous wavelet decomposition (HTW), Jammazi and Aloui (2012) proposed a hybrid HTW-MPNN model to achieve a prominent prediction of crude oil prices.

In addition, crude oil, sometimes called the blood of industries, plays an important role in the energy market. Some analyses and discussions about the existing literature on multi-step-ahead forecasting relating to the energy market will help provide some useful suggestions and implications for multi-step-ahead crude oil price forecasting. Guo et al. (2012) introduced a modified EMD–FNN model for wind speed forecasting. Based on the principle of "decomposition and ensemble", Tang et al. (2012) proposed a novel hybrid ensemble learning paradigm integrating ensemble empirical mode decomposition (EEMD) and least squares support vector machines (LSSVM) for nuclear energy consumption forecasting. To avoid excessive round-off and prediction errors, Pao (2007) developed new NNs with a single output node structure by using the direct strategy for multi-step-ahead electricity price forecasting.