Applied Energy 164 (2016) 366–379

Contents lists available at ScienceDirect

Applied Energy

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

Forecasting the term structure of crude oil futures prices with neural networks ☆

Jozef Baruník [a,b,*], Barbora Malinská [a]

[a] Institute of Economic Studies, Charles University, Opletalova 26, 110 00 Prague, Czech Republic
[b] Institute of Information Theory and Automation, The Czech Academy of Sciences, Pod Vodarenskou Vezi 4, 182 00 Prague, Czech Republic

H I G H L I G H T S

- We analyse term structure of crude oil markets.
- New model for forecasting based on neural networks is proposed.
- We show that even basic architecture of neural models performs very well against benchmarking models.

A R T I C L E I N F O

Article history:
Received 31 August 2015
Received in revised form 23 November 2015
Accepted 28 November 2015

Keywords:
Term structure
Nelson–Siegel model
Dynamic neural networks
Crude oil futures

A B S T R A C T

The paper contributes to the limited literature modelling the term structure of crude oil markets. We explain the term structure of crude oil prices using the dynamic Nelson–Siegel model and propose to forecast oil prices using a generalized regression framework based on neural networks. The newly proposed framework is empirically tested on 24 years of crude oil futures prices covering several important recessions and crisis periods. We find 1-month-, 3-month-, 6-month- and 12-month-ahead forecasts obtained from a focused time-delay neural network to be significantly more accurate than forecasts from other benchmark models. The proposed forecasting strategy produces the lowest errors across all times to maturity.

1. Introduction

Modelling and forecasting the term structures of commodity markets is attractive from an academic perspective and valuable for producers, speculators, and risk managers. Generally, the term structure illustrates expectations of the future development of the corresponding market. Notwithstanding the importance of the subject, there are almost no relevant studies forecasting commodity term structures. In this paper, we introduce a novel framework for forecasting the term structure of crude oil futures prices. We propose to couple dynamic neural networks with the Nelson–Siegel model to obtain precise forecasts of crude oil futures prices.

Crude oil is essential to the world's economies from an industrial perspective because it is a vital production input and its price is driven by distinct demand and supply shocks. Shifts in the price of oil are, to varying extents, driven by aggregate or precautionary demand related to market anxieties concerning the availability of future oil supplies. As the demand for crude oil, which is not as dependent on price as it is on income [1], continues to rise and supply is likely to decline (because crude oil is a limited resource), the literature agrees that the future development of crude oil prices will be highly volatile and hence uncertain future [2]. The main reasons that the crude oil market is one of the most volatile in the world are the growing demand for a supply that is highly dependent on the behaviour of politically and economically unstable countries, crude oil demand and production are highly correlated with the occurrence of exogenous events such as military conflicts and natural catastrophes, and the presence of speculators [3].

As the crude oil futures market is one of the most developed markets based on trading volume, understanding the behaviour of its term structure is crucial. However, studies modelling and forecasting the term structure of petroleum markets are scarce

☆ Support from the Czech Science Foundation under the P402/12/G097 DYME – "Dynamic Models in Economics" project is gratefully acknowledged. The research leading to these results received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007–2013/ under REA grant agreement number 609642.

* Corresponding author at: Institute of Economic Studies, Charles University, Opletalova 26, 110 00 Prague, Czech Republic.
E-mail address: barunik@utia.cas.cz (J. Baruník).

http://dx.doi.org/10.1016/j.apenergy.2015.11.051
0306-2619/

(see Lautier [4] for review), and most of researchers focus on directly forecasting speculative prices [5] or researching the market's efficiency [6–8]. Similar to interest rate models, there are two approaches to modelling the term structure of petroleum commodities. As a natural candidate for the state variable in a one-factor model, the spot price is modelled as geometric Brownian motion [9] or a mean-reverting process [10]. Subsequent contributions consider the convenience yield as a second state variable in a two-factor model [10]. Alternatively, Gabillon [11] employs the long-term price as the second state variable. While both of these approaches assume a constant interest rate, which implies that future spot prices and forward prices are the same, Cortazar and Schwartz [12] proposes a three-factor model.

A relatively series of contributions to the literature explaining commodity futures prices uses the approach of Diebold and Li [13], originally applied to model yield curves. Motivated by similarities in the stylized facts between commodity markets and interest rate markets, the dynamic Nelson–Siegel model is a natural candidate for this task. Among the relatively limited number of contributions on the subject, Karstanje et al. [14] examine the co-movement of factors driving commodity futures curves and their shapes by adopting the framework of the dynamic Nelson–Siegel model [13]. To study the joint dynamics of the factors driving commodity futures curves, Nomikos and Pouliasis [15] uses a multiple-regime framework. Almansour [16] model the futures term structure of crude oil and natural gas markets with switching regimes, and Heidorn et al. [17] regress futures curve factors extracted from the dynamic Nelson–Siegel model on fundamental and financial traders. While the dynamic Nelson–Siegel model explains the dynamics of factors underlying the term structure of commodity prices, the extant literature offers no suggestions for predicting future price developments, with the only exception being Grønborg and Lunde [18]. In their original work, Diebold and Li [13] propose the use of a simple autoregressive time series model to successfully forecast the dynamics of term structure factors and, hence, prices in the interest rate market. We hypothesize that factors in commodity markets may contain further nonlinear dependencies, which need to be modelled to obtain precise forecasts. Therefore, it would be sensible to apply more general methods that do not require restrictive assumptions concerning the underlying structure of factors.

A natural candidate for the forecasting task are neural networks, which can be understood as a generalized non-linear regression tool. Concisely, neural networks are semi-parametric non-linear models, which are able to approximate any reasonable function [19,20]. While the number of models using machine learning is growing rapidly in the academic literature, applications in energy markets are rather limited. Among the few works from recent increase in contributions to the literature, neural networks are applied to predict fuel consumption [21] and day-ahead electricity prices [22], model energy demand in the residential sector in the United States [23], or quantify patterns in the co-movement between futures and spot prices [24]. Several works use neural networks to forecast energy prices [25–32]. Contributing to this strand of the literature, we are the first to employ this approach for forecasting term structures.

The contribution of this work is twofold. First, we contribute to the rare literature studying the term structure of commodity prices by providing new results from the application of the dynamic Nelson–Siegel modelling strategy to crude oil futures markets for long period 1990–2014. Second, we propose the use of a time-delay neural network to forecast the term structure factors identified by the dynamic Nelson–Siegel model. Using this framework, we forecast the term structure of crude oil futures prices successfully over the 1-month, 3-month, 6-month and 12-month forecasting horizons.

2. Data

2.1. Raw data

The dataset consists of monthly closing prices of West Texas Intermediate (WTI) futures contracts,[1] traded on the New York Mercantile Exchange (NYMEX). Each contract expires three trading days prior the 25th calendar day of the month preceding the month of delivery.[2] In total, we analyse 396 monthly historical (already delivered) and to-date undelivered contracts – 12 contracts per year with delivery months in the period beginning in 1990. The undelivered contracts included in the dataset are contracts stipulating delivery in November and December 2014 and 24 contracts with delivery in the two subsequent years (2015 and 2016).

The main reason for using data beginning in 1990 is that the maximum time to maturity for contracts before this date was nine months, while later during the period considered, this duration increased to more than six years. Hence to avoid potentially large risk and inaccuracies stemming from data extrapolation, we consider only data after the year 1990. The choice of the monthly frequency is primarily driven by the fact that contracts with longer times to maturity were traded rather infrequently in the first half of the studied period. In addition, Baumeister et al. [33] find monthly data to have equal predictive power to that of daily data.

Table 1 presents an example of actual data to illustrate the structure and dimension of the dataset. To associate each futures price observation with its corresponding time to maturity, it is necessary to first determine the exact expiry date of each contract. Then, the difference between the expiry and observation dates provides us with the remaining days to maturity. Table 1 captures the end-of-month futures prices of three different (in this case consecutive) contracts with delivery in August, September and October 2003. For example, at the end of February 2001, CLQ2003 and CLU2003 contracts were traded. On February 28, 2001, it was possible to enter into a contract with delivery in August 2003 at a futures price of USD 21.72 per barrel. The time to maturity in this case (τ) was 625 trading days.

2.2. Reorganized data

After combining the days to maturity with each observed futures price quotation, the dataset should be formed into a matrix with a number of rows equal to number of days included in analysis and a number of columns equal to number of analysed maturities.

Time series captured in Table 2 are reorganized into constant-maturity futures prices. WTI crude oil futures are delivered and expire with one-month regularity; therefore, futures prices with exactly 30, 60, or 90 days to maturity are not traded every day. The literature suggests several approaches to interpolating the prices to obtain the desired data format. Diebold and Li [13] use linear interpolation for constant maturity, while Holton [34] prefers cubic splines interpolation.[3] In our work, we follow the approach of Holton [34] and use cubic spline interpolation. Fig. 1 illustrates the reorganized constant-maturity futures prices we employ, plotted against the daily evolution of the spot price.

Due to the long period we consider, which includes several turbulent periods, we present the term structure in separate periods to better highlight the rich dynamics (Fig. 2).

Fig. 2(a) illustrates the term structure dynamics in the period

[1] Available at https://www.quandl.com/c/futures/cme-wti-crude-oil-futures.

[2] Full specification of WTI futures contracts is available at http://www.cmegroup.com/trading/energy/files/en-153_wti_brochure_sr.pdf.

[3] For a detailed discussion of the interpolation methods for curve construction with application to yield curve modelling, see Hagan and West [35].