Daily Prediction of Short-Term Trends of Crude Oil Prices using Neural Networks Exploiting Multimarket Dynamics

Heping Pan, Imad Haidar*, Siddhivinayak Kulkarni*
Prediction Research Centre (PRC), University of Electronic Science and Technology of China
North Jianshe Road, Chengdu 610054, China. panhp@swingtum.com
Finance Research Centre of China, Southwest University of Finance & Economics, China
Swingtum Institute of Intelligent Finance, Australia and China
*School of ITMS, University of Ballarat, Australia
ihaidar@students.ballarat.edu.au, s.kulkarni@ballarat.edu.au

Abstract

This paper documents a systematic investigation on the predictability of short-term trends of crude oil prices on the daily basis. In a stark contrast with longer-term predictions of crude oil prices, short-term prediction with time horizons of 1-3 days posits an important problem which is quite different from what has been studied in the literature. The problem of such short-term predicability is tackled through two aspects. The first is to examine the existence of linear or nonlinear dynamic processes in the crude oil prices. This subproblem is addressed with statistical analysis involving Brock-Dechert-Scheinkman test for nonlinearity. The second aspect is to test the capability of artificial neural networks (ANN) for modeling the implicit nonlinearity for prediction. Four experimental models are designed and tested with historical data: (1) using only the lagged returns of filtered crude oil prices as input to predict the returns of the next days, this is used as the benchmark; (2) using only the information set of filtered crude oil futures price as input, (3) combining the inputs from the benchmark and second models, (4) combing the inputs from the benchmark model and the intermarket information. In order to filter out the noise in the original price data, moving averages of prices are used for all the experiments. The results provided sufficient evidence to the predictability of crude oil prices using ANN with an out-of-sample hit rate of 80%, 70%, 61% for each of the next three days trends.

Key words: Crude oil prediction, short-term trend, crude oil futures, heating oil, neural networks, intermarket analysis

1. Introduction:

Crude oil is the premier strategic commodity for the modern engine-driven economies all around the globe. As the total oil resource on the planet is limited, oil demand continues to surge ahead and production continues to decline, it is believed that oil prices will continue to rise to unprecedented levels. Unfortunately, crude oil has proven to be one of the most volatile markets in the world. In addition to daily price fluctuations, oil prices have risen substantially in last few years. As a reference, in 1999, the price of crude oil hovered around $16 a barrel; but in 2004, the crude oil price was averaging $41 per barrel. By 2008, it had crossed the $100 a barrel mark, and fluctuated between $147.96 and $69 (during June-October 2008). The unprecedented increase and the wide swings in crude oil prices have significant impacts on the well-being of the economies of both developed and developing countries. Crude oil price fluctuations and shocks have played a major source of adverse macroeconomic impacts, bringing abour economic instability in both oil exporting and oil consuming countries as well as leading to economic crises. Oil price prediction, therefore, is vital to economic agents and policy makers. Consequently there have been many efforts in developing models to predict oil prices in spot and exchange traded markets.

In a stark contrast with many published researches on long-term predictions of crude oil prices, this paper investigates the possibility of predicting short-term prices and trends of crude oil using any sensible information extracted from the historical and current price time series of the crude oil spot markets, futures markets, as well as markets which are considered to lead the spot markets. This third category of markets are called 'lead markets' to the spot markets. In our terms, the so-called 'short-term' refers to 1-5 days into the future. Practically, in this study, we limit our scope of study to 1-3 days.

Theoretically, the predictability of crude oil prices and trends on the short terms is a scientific problem with two essential aspects – two sub-problems. The first aspect is on the properties of the market prices themselves, i.e. whether the prices are solely random walks or exhibit any linear or nonlinear dynamics. The second aspect is whether there are any computational models which can capture such linear or nonlinear dynamics if existent. To address the first sub-problem, we use statistical analysis, in particular, involving using statistical tests to check whether there are nonlinear dynamics in the price time series. To tackle the second sub-problem, we use artificial neural networks (ANN) as computational models for prediction. In particular, we concentrate on using multilayer feedforward neural networks (MFNN), also called multilayer perceptrons, as general approximator to any nonlinear mapping from information extracted from historical and current price time series to the prices of the immediate futures – the next three days. The short-term trends of the prices are denoted by the short-term moving averages (MA) of the prices. In this study, we shall limit to 3-day moving average (3-MA). In this setup, we aim to predict the returns of the filtered spot prices for each of the next three days, which is what mean by prediction of short-term trends. As a historical background, this study is a further development of a methodology for financial market prediction which has been successfully applied in stock index prediction (Pan 2004; Pan et al 2005; Pan 2006). More general information is provided by (Pan et al 2006; Pan 2008).

The rest of the paper is organized as follows: Section 2 provides a brief literature review on the related studies for crude oil market analysis and prediction. Section 3 defines the problem of short-term prediction of crude oil prices precisely and presents our methodology for this investigation. Section 4 describes the data sources and their statistical properties as evidence for inherent predictability. Sections 5-8 detail four prediction models separately, of which the differences are on the information sources used for inputs, i.e. (1) the spot prices only, (2) the futures prices only, (3) the spot and futures prices, (4) the spot and lead market prices. Section 9 shows the results on multi-step prediction. Section 10 concludes the paper.

2. Crude Oil Market Analysis and Prediction in the Literature

The importance of crude oil to the economy is reflected by the number of studies in this vast area. There is a large and rich literature related to every aspect of crude oil. In particular, there are many published works dedicated to develop fundamental or econometrical models for explaining changes in crude oil prices and for predicting them correctly and accurately in spot and futures markets.

A major characteristic of crude oil market is significant price fluctuations. This volatility of oil prices could be attributed to three main factors: 1) Increase in demand and supply shortages possibly caused by economic growth or behaviours of oil producing countries; 2) Exogenous events such as wars, natural disasters, etc; 3) Endogenous factors such as speculations in the markets. Moreover, the fact that a significant amount of crude oil comes from the unstable Middle East has contributed to the oil price swings and shocks. However, the fundamental causality analysis of crude oil price fluctuations and shocks are beyond the scope of this paper. We shall limit our discussions to econometrical and computational modeling and prediction of crude oil price changes and trends.

Early and recent studies in the literature on crude oil market analysis and prediction can be grouped roughly into three clusters: 1) futures as predictors to spot prices, 2) econometrical models for explanation and prediction, 3) computational intelligence models for prediction.

2.1 Futures as Predictor to Spot Prices

Crude oil futures contract was introduced to NYMEX in 1983, however, trading in these particular contracts were relatively shallow until 1985, whereas, crude oil futures contracts are considered as an important financial instruments in the energy market (Haubrich et al 2004). The relation between futures prices and spot price has been a focus of attention for a number of scholars, and the literature is rich with many studies covering a range of aspects with respect to this relationship. Lead-lag, efficiency, prediction amongst other factors, are the most studied areas in the futures-spot literature.

It is important to note, however, that some economists believe that futures prices are not a predictor for spot prices. For example, Haubrich et al (2004) argued that crude oil futures prices is not a suitable vehicle to predict spot price, implying that futures prices do not hold any new information, not even in the short term. However, the idea of using commodity futures prices to predict spot prices is based on the assumption that the futures prices react faster to the new information entering the market than spot prices. Trading in the futures market has many advantages, such as low transaction cost, high liquidity, and low cash in up-front, amongst others. This makes it much more attractive for investors to react for new information, than taking position in the spot market. This argument applies for most of the commodities listed in the financial markets; however, it is more relevant to the energy markets. The reason for this is, when new information related to the oil market is introduced, investors have two options, either to take a position (buy or sell) in spot or in futures markets. In most of the cases, taking a position in the spot market is not the best way for reacting to the new information. Because it requires high transaction costs, storage costs, and delivery costs etc. Especially, if investors are not interested in the commodity itself rather they are hedging for another commodity, or simply just investing in the market in hope of arbitrage opportunity, i.e. speculation. In this context, the futures market is much more attractive a place for an investor to react to new information.

An early study by Bopp and Sitzer (1987) tested whether futures prices are good predictors for cash prices in the heating oil market, in attempt to answer if futures prices have the additional information to improve the predictability of econometrical models. The results showed that only futures contracts of 1 and 2 months to maturity are statistically significant for cash price prediction, in other words containing new information.

Silvapulle and Moosa (1999) examined the lead-lag relationship between crude oil futures and spot prices. Their goal was to find whether the change in crude oil futures prices causes the change in crude oil spot prices, i.e. causality, using linear and nonlinear tests. The data set was composed of daily spot prices, as well as futures contracts of 1, 3, 6 months to maturity, from 1985 to 1996. Their results showed that both futures and spot prices react to new information at the same time. Moreover, the authors also concluded that, there is some evidence in support of feedback between futures and spot prices. However, this feedback runs in one direction only; from futures to spot, and not vice versa. Finally, their results showed that the lead-lag patterns are not constant and changeable over time.

Coppola (2007) studied the relationship between crude oil spot price and futures price using the cost of carry model. The aim was to predict the out-of-sample and price movements in the futures. The author used Vector Error Correction Model (VECM) for the prediction, comparing the results to random walk model. Evidence of co-integration was recorded between spot and futures weekly prices. The author also found that futures contracts are able to reflect the information, however, these results stands only for in-sample prediction. For out-of-sample prediction the author claimed that VECM outperform random walk model in both accuracy of prediction and timing the market.

The activities of investors and hedgers and their effects on the market were also studied widely. Milunovich and Ripple (2006) presented a model to estimate the magnitude of hedging activity on crude oil futures volatility, using a combination of dynamic conditional coloration and augmented EGARCH. They found that hedging activity has a significant influence on the conditional volatility of crude oil futures returns.

Although the body of oil literature is substantial, there is still a great deal of inconsistency in the findings. This is particularly the case in the relation between spot and futures prices. While most of studies agree that futures play a important role in financial markets, however, there is no such agreement on futures as predictor of spot. Furthermore, the vast majority of the literature is based on econometrical models. A major shortfall of econometrical model is making strong assumption about the problem (Refenes 1995). This means if the assumptions are not correct; the model could generate misleading results. In these contexts artificial neural networks (ANN) are viewed as nonparametric, nonlinear, assumption-free model

(Azoff 1994). This means it does not make any assumption about the problem; rather it attempts to find the appropriate function from the data itself (Refenes 1995). Furthermore, ANN is considered as general approximator to any nonlinear mapping (Hornik *et al* 1989), and has been around for about two decades and successfully used in many applications, including crude oil price prediction.

Although, some study of crude oil prediction has used ANN models, nonetheless, to our best knowledge, not much study of futures prices as spot predictor is based on ANN modeling. Finally, reader should bear in mind that we are not testing for causality, whether futures prices cause spot prices or other variables affect the relationship. We are simply testing if futures aggregate useful information on spot price future directions. This aggregation of information could be a result of the activity of market participants, who often take a position in futures markets not only based on expectation of price raise or fall, but also to hedge from the conscience of unexpected events (Bopp and Sitzer 1987).

2.2 Econometrical Models for Crude Oil Markets

For crude oil markets, several econometrical models were developed since 1970's. The early models aimed to explain the crude oil market dynamics and the interactions amongst oil fundamental variables such as supply-demand inventory. However, these models failed to provide accurate prediction for crude oil prices (Pindyck 1999). Labonte (2004) surveyed analytical models for crude oil shock, and concluded that one of the most common cavities of econometrical models is omitted variables. This occurs when some of the variables affecting the price time series are not included in the models. This could be due to two reasons: either the economists did not know that these variable exist or the number of variables is too large to be fitted in one model. In addition to this, some models suffer structural misspecification which means a statistical or mathematical relationship does not necessarily mean that a causal relationship really exists (Labonte 2004). Time series analysis models such as linear and nonlinear regression have been around for some time and been used for a large number of studies of the econometrical modeling problems. Nevertheless, two milestones can be recorded for time series models: the Box-Jenkin's Auto Regression Intergraded Moving Average (ARIMA) in the 1970's, and the ARCH models for volatility modeling by Robert Angle, and the GARCH by Time Borslave in the 1980's. ARIMA has been applied in a large number of studies for crude oil price prediction and generated much better results than structural econometrical models and regression (Moshiri & Foroutan, 2005). However, ARIMA is a linear model, and therefore, it would be disadvantageous if the time series is nonlinear.

2.3 Computational Intelligence Models for Crude Oil Price Prediction

With the exponential lifting of computing power, computational intelligence models such as Artificial Neural Networks (ANN), Support Vector Machine (SVM), Genetic Algorithms (GA), and Case–based Reasoning (CR) have been exploited for financial prediction. These techniques were also applied to the crude oil predicting problem, however, not as intensive as time series models. As a clarification, ANN in general includes Multilayer Feedforward Neural Networks (MFNN), Recurrent Neural Networks (RNN), SVM and several other architectures. However, in this paper, whenever ANN is mentioned, it refers to MFNN, if not specified otherwise.

An early study by Kaboudan (2001) compared the ability of genetic programming and ANN to trivial predictors for predicting crude oil prices on a monthly basis. In his study Kaboudan (2001) tested different set of variables as input to the model, including lagged values of crude oil price, the world crude oil production, OECD countries consumptions, and US oil inventory amongst others. An important conclusion was reached that only the lagged values of crude oil price were useful in providing acceptable prediction. More specifically, 9 lags of crude oil monthly price provide the best inputs for one-month prediction. The author found that genetic programming was superior to ANN. However, the size of data used for this study was very small, only 60 data points were used to train both models which could significantly affect the ANN performance.

Moshiri and Foroutan (2005) began their study by testing for chaos and nonlinearity in crude oil futures prices. Performing several statistical and econometrical tests led them to conclude that futures prices time series is stochastic and non-linear. Moreover, the authors compared linear and non-linear models for predicting crude oil futures prices, including ARIMA and GARCH, to ANN, and found that ANN is superior and produces statistically significant prediction. However, in our opinion the results obtained using ANN could be limited by the use of the entire data set for training (1983 to 2000). Generally, when dealing with ANN the more data points are used, the better the network will generalize. Nonetheless this is not necessarily the case when dealing with financial or economical time series. As economic conditions change over time, non-current (old information) could affect prediction results negatively. Because training the network with irrelevant information to the current conditions could result in a poor model generalization.

Xie et al (2006) proposed a SVM model for monthly crude oil price prediction, reporting that SVM outperformed ANN (MLFF) and ARIMA for out-of-sample test. However, their results were not consistent, as ANN outperformed SVM for two of the four sub-periods tested. Nonetheless, both ANN and SVM outperformed ARIMA in all the four periods of their test. Liu et al (2007) presented a hybrid model based on neuro-fuzzy technique to predict Brent crude oil prices. Three prediction models were used, including a radial basis function network, a Markov chain-based semi-parametric model and a wavelet analysis-based model. The output of the three models was used as input to a fuzzy neural network while the target was the actual Brent crude oil price. The authors concluded that the nonlinear combination outperformed any single model in terms of root mean square errors (RMSE) as performance measure. In a related study, Yu et al (2007) proposed a decomposition method for extracting multi-scale features from price time series as input to ANN models. They applied Empirical Mode Decomposition (EMD), which was originally developed by Huang et al (1998), on daily crude oil spot prices to decompose the original price time series into a number of intrinsic mode components. Only six of these signals were selected as input to an ANN. The authors concluded that such decomposition-based ANN outperformed ANN using the original prices solely as input.

A recent research has emerged concentrating on the predicting financial markets using text embedded either in the web or directly through news provider like Reuters. This method is based on machine learning techniques, more precisely information retrieval, text classification, and categorization.

Wang et al (2005) proposed a methodology, called TEI@I, to predict crude oil prices. This methodology integrates three different techniques - text mining, econometrical modeling with ARIMA and intelligent computation (ANN). Of the crude oil price time series, the linear components are modeled with ARIMA, the nonlinear components are modeled with ANN, and the irregular event impacts are modeled using web-based text mining and rule-based expert systems. Their simulations concluded that the integrated methodology is significantly superior than individual models for monthly prediction of crude oil prices. Yu et al (2005) refined the TEI@I methodology with rough set technique. Two comments can be made regarding this methodology. First, the oil price data used for ANN training is monthly data, which limits the training and testing data significantly, as only 49 data point was reserved for out-of-sample test. Second, in the text mining model, the rule base system depends on the knowledge base which is developed by human experts; however, unlike other well-defined problem domains, expert opinions on the crude oil markets can vary wildly.

However, a survey by Mittermayer and Knolmayer (2006) compared several text mining prototypes for financial prediction, concluding that most of these prototypes have failed to produce significant results especially in terms of profitability. This could imply that even varying predictability can be achieved, the crude oil market is still highly efficient because professional traders have also developed their own capability of prediction, either by judgement or by instinct. In general, one of the disadvantage of text mining for financial prediction is that it assumes that information reaches web sites before it reaches provincial investors or at least at the same time. This assumption is not necessarily true. First and foremost the Internet is not the only medium for news and it is not the fastest one as well. Second, most professional traders (market makers) keep a track on any event via real-time news providers such as Routers. Although some of these prototypes have relied on Reuters news provider to gather the data, nonetheless, in this case, they are getting the news at the same time as professional traders but they are lagging in reaction until the system finish analysing the data and a conclusion is reached for trading decisions. Finally, the major disadvantage of these methods is the classification techniques which depend on bag of the words in other words, the content of the news is dealt with in term of frequency of occurring in the page, while human expert will deal with the meaning and possible impact (action-reaction). Furthermore, in most cases the classifiers depend on hand-crafted dictionary made by so called domain expert. This means the room of misclassification could be there.

In addition, most of the studies above have used monthly crude oil prices. This reduces the level of noise in the data but at the same time restricts the prediction horizon to monthly intervals, and limits the training and testing data significantly. Moreover, most of these studies are univariate models based on the historical prices of the crude oil while the interrelation amongst crude oil prices and other intermarket prices were ignored. In addition, when comparing a model to ANN no clear justification was provided if the ANN model used for comparison is optimal or not.

3. A Methodology and Design of Experiments for Predicting Short-Term Trends of Crude Oil Prices

From here on we shall present a specific methodology for predicting short-term trends of crude oil prices. It includes a statistical analysis of the data set and four experiments designed to find the optimal input-output mapping for the prediction. We start with a precise definition of this particular prediction problem.

Let $V(t)$ be the crude oil price at time t, in general we assume the availability of a historical price data set containing the price time series from a starting point in time $t_0 = t - N + 1$ to the current time $t > t_0$,

$$D(V, N, t) = \{V(t), t = t - N + 1, t - N + 2, \cdots, t\} \tag{1}$$

Note that in its original definition, the crude oil price prediction can be simply defined as predicting one or more future price values using all the sensible information extracted from this data set, i.e.

$$D(V, N, t) \Rightarrow ANN \Rightarrow V(t + \lambda), \qquad \lambda = 1,2,3,\cdots \tag{2}$$

Or more visually, if we use the N lagged prices to predict $\lambda = 3$ future prices, we have

$$(\cdots \quad V(t - N + 1) \quad \cdots V(t-1) \quad V(t)] \quad \Rightarrow ANN \Rightarrow \quad [V(t+1) \quad V(t+1) \quad V(t+3)] \tag{3}$$

Virtually almost all the published studies on financial time series prediction belong to this definition of prediction. However, considering the existence of innumerous unexpected factors during any trading day in the market, the prices must contain noise, as a matter of fact. Therefore, precise prediction of the raw prices is not a well-defined scientific problem. To be more practical, we choose to predict the short-term trends of the market prices. In our terms, the short-term trends of the price are denoted by short-term Moving Averages (MA), defined as

$$MA(V, q, t) = \frac{\sum_{k=t-q+1}^{t} V(k)}{q} \tag{4}$$

where q denotes the length of the time window for the moving average.

Since the prediction models we are building actually belong to statistical pattern recognition where the patterns refer to the invariances in the price process, it makes not much sense to use the absolute values of the prices. This is also because the market ecosystems must have changed during the many years in the data set period. Therefore, it is a common choice to use the relative returns of the prices. But in our case, we choose to use the relative returns of the short-term moving averages,

$$X(V,q,\tau,t) = \frac{MA(V,q,t) - MA(V,q,t-\tau)}{MA(V,q,t-\tau)} \tag{5}$$

where parameter τ expresses the scale of time, reflecting the time resolution of the input or the time horizon of the output. Note when $X(V,n,\tau,t)$ is very small, it is close to the logarithm of the two MA's ratio, i.e.

$$X(V,q,\tau,t) \approx \log \frac{MA(V,q,t)}{MA(V,q,t-\tau)} \tag{6}$$

When the variable V is clearly given, and parameters n,τ are specified, we may simply use $X(t)$ for $X(V,q,\tau,t)$. In financial econometrics literature, such $X(V,q,\tau,t)$ or $X(t)$ is called 'relative return' of the price $V(t)$. In other financial informatics literature, it is also called 'momentum'. Neuneier & Zimmermann (1998) and Gorthmann (2005) also suggested the use of the 'force' for input-output mapping,

$$Y_\tau(t) = \frac{V(t) - 2V(t-\tau) + V(t-2\tau)}{V(t-\tau)} \tag{7}$$

Their use of the force is on the original prices, but for our case, we define the force on the moving averages,

$$Y(V,q,\tau,t) = \frac{MA(V,q,t) - 2MA(V,q,t-\tau) + MA(V,q,t-2\tau)}{MA(V,q,t-\tau)} \tag{8}$$

The financial meaning of the 'force' is the direction and the strength of the short-term trend change – either trend acceleration or trend reversal.

For defining the input-output mapping in this study, there are three types of data sources:

（1） Spot prices of crude oil, denoted by 'spot' in the following formulas;
（2） Futures prices of crude oil, with different contracts with different term structures – number of months to maturity, denoted by 'futures'.
（3） Prices of lead markets, denoted by 'lead'. We will specify these intermarkets later.

With these definitions and notation given above, we now can define our problem of predicting the short-term trends of crude oil price precisely as follows

$$\begin{Bmatrix} D(X(spot,q_s,\tau_s),N_s,t) & D(X(futures,q_f,\tau_f),N_f,t) & D(X(lead,q_l,\tau_l),N_l,t) \\ D(Y(spot,q_s,\tau_s),M_s,t) & D(Y(futures,q_f,\tau_f),M_f,t) & D(Y(lead,q_l,\tau_l),M_l,t) \end{Bmatrix} \Rightarrow ANN \Rightarrow \begin{Bmatrix} X(spot,q_s,\tau_s,t+\lambda) \\ Y(spot,q_s,\tau_s,t+\lambda \end{Bmatrix} \tag{9}$$

where $\lambda = 1,2,3,\cdots$, and parameters N, M, q, τ with different subscripts imply that we may use different time resolution, time span, or time horizons for different sources of information.

Using ANN as the type of nonlinear models, finding the optimal input-output mapping is an exploration process. We actually have designed four experiments each investigating a subtype of the general model:

Model I: from spot prices to spot prices

$$\begin{Bmatrix} D(X(spot,q_s,\tau_s),N_s,t) \\ D(Y(spot,q_s,\tau_s),N_s,t) \end{Bmatrix} \Rightarrow ANN \Rightarrow \begin{Bmatrix} X(spot,q_s,\tau_s,t+\lambda) \\ Y(spot,q_s,\tau_s,t+\lambda \end{Bmatrix} \tag{10}$$

Model II: from futures prices to spot prices

$$\begin{Bmatrix} D(X(futures,q_f,\tau_f),N_f,t) \\ D(Y(futures,q_f,\tau_f),N_f,t) \end{Bmatrix} \Rightarrow ANN \Rightarrow \begin{Bmatrix} X(spot,q_s,\tau_s,t+\lambda) \\ Y(spot,q_s,\tau_s,t+\lambda \end{Bmatrix} \tag{11}$$

Model III: from spot and futures prices to spot prices

$$\begin{Bmatrix} D(X(spot,q_s,\tau_s),N_s,t) & D(X(futures,q_f,\tau_f),N_f,t) \\ D(Y(spot,q_s,\tau_s),N_s,t) & D(Y(futures,q_f,\tau_f),N_f,t) \end{Bmatrix} \Rightarrow ANN \Rightarrow \begin{Bmatrix} X(spot,q_s,\tau_s,t+\lambda) \\ Y(spot,q_s,\tau_s,t+\lambda \end{Bmatrix} \tag{12}$$

Model IV: from spot and lead market prices to spot prices

$$\begin{Bmatrix} D(X(spot,q_s,\tau_s),N_s,t) & D(X(lead,q_l,\tau_l),N_l,t) \\ D(Y(spot,q_s,\tau_s),N_s,t) & D(Y(lead,q_l,\tau_l),N_l,t) \end{Bmatrix} \Rightarrow ANN \Rightarrow \begin{Bmatrix} X(spot,q_s,\tau_s,t+\lambda) \\ Y(spot,q_s,\tau_s,t+\lambda \end{Bmatrix} \tag{13}$$

Note that although these four models cannot cover all the subtypes of input-output mapping, nonethless through the sequence of our data-based experiments, these models represent the essential different combinations. This will become clear after reading through following presentations.

Although theoretically and practically, both the momentums X and forces Y can be included in the general model of (9) and in each of the special models of (10)-(13), and although we have done a number of numerical experiments on these more general information sets, the focus of this paper is solely targeted at the prediction of the relative returns – momentums X. Therefore, in all the following descriptions, we shall only report the results on using the relative returns of the filtered prices in the input and output of all the specific models to be described throughout all the following sections.

For the actual structure of ANN as the prediction model, we use multilayer feedforward neural network (MFNN) , theoretically one hidden layer with sufficient hidden neurons of nonlinear transfer function is powerful enough to approximate any nonlinear continuous mapping from an input space $\mathbf{x} = (x_1, x_2, \cdots x_n) \in R^n$ to an output space $\mathbf{z} = (z_1, z_2, \cdots z_m) \in R^m$ (Hornik *et al* 1989). Let $\mathbf{y} = (y_1, y_2, \cdots y_h) \in R^h$ denote the *h* hidden neurons. We have the explicit representation of an ANN of the MFNN type as

$$y_j = f(x_1, x_2, \cdots, x_n) = \phi(\sum_{i=1}^{n} w_{ij} x_i + b_j), \qquad j = 1,2,\cdots,h \tag{14}$$

$$z_k = g(y_1, y_2, \cdots, y_h) = \varphi(\sum_{i=1}^{n} u_{jk} y_j + d_k), \qquad k = 1,2,\cdots,m \tag{15}$$

Generally, the transfer function ϕ for the hidden layer should be a nonlinear function. Two most commonly used transfer functions are

Sigmoid function:

$$\phi(x) = \frac{1}{1 + e^{-x}} \tag{16}$$

Hyperbolic tangent function:

$$\varphi(x) = \frac{1 - e^{-x}}{1 + e^{-x}} \tag{17}$$

The sigmoid of (16) and the hyperbolic tangent of (17) are commonly used for the neurons of the hidden layer(s) as it produces a value between [0, 1], while linear function is generally used for the neurons of the output layers as it produces a real value such as between [-1, 1]..

The three performance measures are used for these experiments: for an observed output $x(t)$, if a prediction of it is $y(t)$, and suppose there are n time points for prediction test:

(1) Hit Rate is the success ratio for market direction prediction

$$HR = \frac{1}{n}\sum_{k=1}^{n}[x(t_k)y(t_k) > 0] \tag{18}$$

where inside $[\,]$ is logical expression which returns 1 for true, 0 otherwise.

(2) Root Mean Square Errors (RMES):

$$RMSE = \sqrt{\frac{\sum_{k=1}^{n}(x(t_k) - y(t_k))^2}{n}} \tag{19}$$

(3) R-square (R^2) is used as goodness of fit measure, with R defined by

$$R = \frac{\sum_{k=1}^{n}(x(t_k) - \bar{x})(y(t_k) - \bar{y})}{\sqrt{\sum_{k=1}^{n}(x(t_k) - \bar{x})^2 \sum_{k=1}^{n}(y(t_k) - \bar{y})^2}} \tag{20}$$

In addition, Mean Square Error (MSE), Mean Absolute Error (MAE), and Sum of Square Errors (SSE) are also calculated.

4. Historical Data and Statistical Properties as Evidence for Inherent Predictability

Five time series are used in this study, West Texas Intermediate (WTI) light sweet crude oil spot prices and futures contracts traded at NYMEX. Futures data include four contracts respectively of 1, 2, 3, 4 months to maturity. The data frequency is daily closing price; from Jan 1996 to Aug 2007, so the data set includes 2912 data points for each time series. All the data were obtained from the US Department of Energy's Energy Information Administration. The data set was divided into the training and test sets, 90% of the data for training and 10% for out-of-sample test. The relative return time series for each original price time series (spot or futures) is normalized to fit between -1 and 1 as preprocessing before subsequent steps in all the experiments.

Figure 1 shows a plot of crude oil spot raw prices (WTI) in the upper section and spot price relative returns in the lower section. The plot of raw price shows that the behaviour of the crude oil time series has started to change from 2002. As the level of price has increased significantly from 2002 to 2007 compared to 1996 to 2001 level. Moreover, the plot of the returns reveals that crude oil has high volatility on daily basis with some outliers. Furthermore, it also shows evidence of volatility clustering i.e. large changes are followed by large changes and small changes are more often followed by small changes.

In order to detect whether there may be any stochastic dynamics in the spot prices, we investigated the statistical properties of the data using two tests: 1) Ljung-Box Q-test for autocorrelation, 2) Brock-Dechert-Scheinkman (BDS) test for nonlinearity.

Ljung-Box Q-test was done for autocorrelation of the input data

$$LQ = m(m+2)\sum_{i=1}^{n}\frac{\hat{\rho}_k^2}{m-i} \tag{21}$$

where m is the sample size, n is the number of lags tested for autocorrelation, $\hat{\rho}_k$ is the sample autocorrelation at lag k. Under this model the null hypothes is there is no significant correlation.

Table 1 shows the results of the Ljung-Box Q-test for the relative returns of the spot prices. Clearly, significant correlations are detected for all the lags tested (5, 10, 15, and 20) at 5% significant level. Thus, there is strong evidence that the present and past information could be useful to predict the futures direction.

Table 1: *Ljung-Box Q-test for the relative returns of spot prices*

Lag	**P-Value**	**LQ**	**Critical value**
5	0.0241	12.9276	11.0705
10	0.0090	23.5131	18.3070
15	0.0051	32.7394	24.9958
20	0.0042	40.6091	31.4104

The Brock, Dechert and Scheinkman (BDS) test for nonlinearity is by far the most used test for nonlinearity in the data. We apply the test on crude oil spot price relative returns. Linear regression was applied on the original returns to remove the linearity from the data and the BDS test was applied on the residuals. The test was used for the embedded dimension of $N = 2,3,\cdots,20$, and three different values of e were selected 1, and 1.5 of the standard deviation of the data. For $e = 1$ and apart from $N = 2$ the test rejected the null hypothesis that the data are independently identically distributed, which indicates the existence of nonlinear structure in the data. The same conclusion can be obtained for $e = 1.5$ and for all values of N tested. Moreover, the results show that the evidence of nonlinearity in the data increased with higher dimensions as the $W_{(N)e}$ increased with larger values of N and *e*. In light of these results, nonlinear models such as ANN could be used for modeling with this data.

Table 2 *BDS test for crude oil spot data*

N	e	P-value	$W_{(N)e}$	e	P-value	$W_{(N)e}$
2	1.5	0	4.7963	1	0	3.2218
3	1.5	0	6.2515	1	0	4.1548
4	1.5	0	6.9357	1	0	4.8923
5	1.5	0	7.2325	1	0	5.1992
6	1.5	0	7.7066	1	0	5.8902
7	1.5	0	8.2214	1	0	6.6529
8	1.5	0	8.5258	1	0	7.2729
9	1.5	0	8.8533	1	0	7.9657
10	1.5	0	9.0754	1	0	8.7041
11	1.5	0	9.2945	1	0	9.556
12	1.5	0	9.4367	1	0	10.296
13	1.5	0	9.6892	1	0	11.357
14	1.5	0	9.9989	1	0	12.492
15	1.5	0	10.382	1	0	13.799
16	1.5	0	10.802	1	0	15.478
17	1.5	0	11.180	1	0	17.290
18	1.5	0	11.632	1	0	19.331
19	1.5	0	12.160	1	0	21.659
20	1.5	0	12.757	1	0	24.526

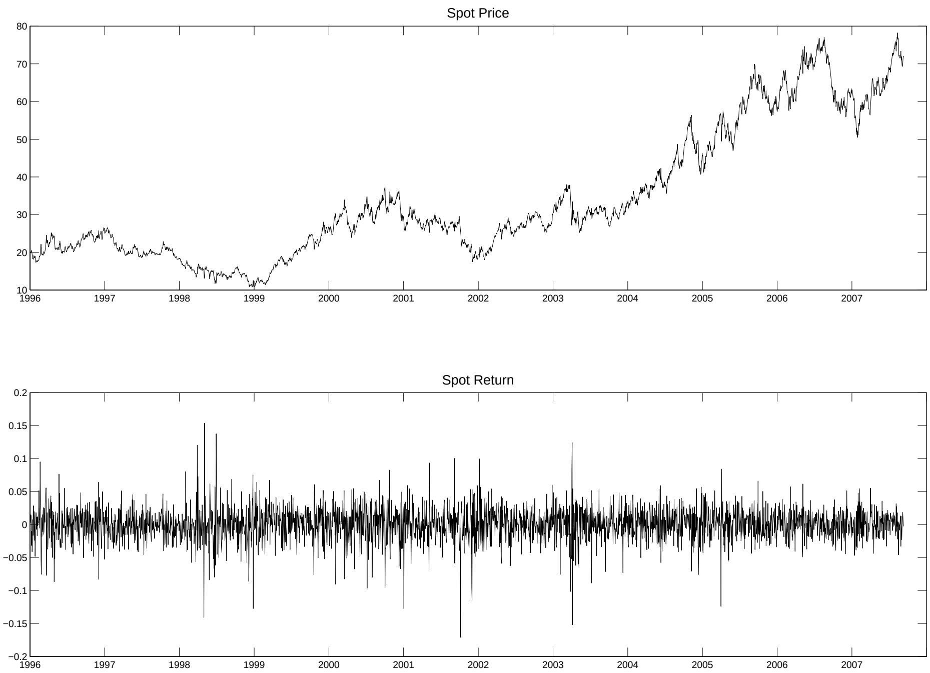

Figure 1 A plot of crude oil prices(upper) and their relative returns (lower)

5. Model I – from Spot Prices to Spot Prices

As defined by equation (10), Model I is limited to only using information extracted from spot prices to predict the momentum and/or force at the next days. This model is used as the benchmark for other variations.

A number of detailed experiments were carried out for virtually exhaustive search for the optimal input-output mapping and the optimal architecture of the ANN. As our focus is on predicting the relative returns of the filtered spot prices, a specific form of Model I for one-step prediction can be expressed as

$$D(X(spot,3,1),m,t) \Rightarrow ANN \Rightarrow X(spot,3,1,t+1) \qquad (22)$$

where m refers to the embedding dimension, i.e. the number of lagged values used in the input. Furthermore, the architecture of ANN is a MFNN with 1 hidden layer consisting of 8 hidden neurons in addition to the input and output layers. The training algorithm is Levenberg-Marquardt. Table 2 shows the performance with different embedding dimensions (different lags).

Table 2: *Performance of Model I with different lags*

Lag	Hit Rate (%)		RMSE	
	In sample	Out of sample	In sample	Out of sample
1	72.77	73.75	0.0108	0.0079
2	72.88	74.01	0.0107	0.0080
3	73.74	74.54	0.0104	0.0077
4	75.42	76.37	0.0099	0.0073
5	76.02	76.90	0.0096	0.0074
6	76.16	77.16	0.0095	0.0072
7	77.25	77.11	0.0085	0.0077
8	78.01	75.59	0.0091	0.0070
9	78.23	76.77	0.0088	0.0070
10	77.78	78.08	0.0089	0.0069
11	78.03	76.37	0.0086	0.0071
12	77.97	77.95	0.0087	0.0070
13	79.45	79.79	0.0083	0.0068
14	79.39	77.42	0.0083	0.0072
15	79.75	79.11	0.0078	0.0073
16	79.77	79	0.0081	0.0071
17	79.45	78.34	0.0080	0.0069
18	80.40	78.87	0.0076	0.0133
19	80.95	77.16	0.0076	0.0075
20	81.38	77.55	0.0074	0.0074

The result in Table 2 is also illustrated in Fig. 2 and 3.

It is important to choose the best performance with the least number of lags, and of hidden neurons to keep the model stable. 13 lags seem to produce reasonably high hit rate and the results for in- and out-of-sample are very close. Furthermore, the R^2 is 0.67, meaning the model was able to explain 67% of the variation of the data. Table 3 summarizes the performance of the benchmark with the optimal input-output mapping and ANN architecture of 6 hidden neurons. Figure 4 is a plot of out-of-sample forecast produced by Model I compared to the actual values (the vertical axis is the momentum as defined by equation (7)).

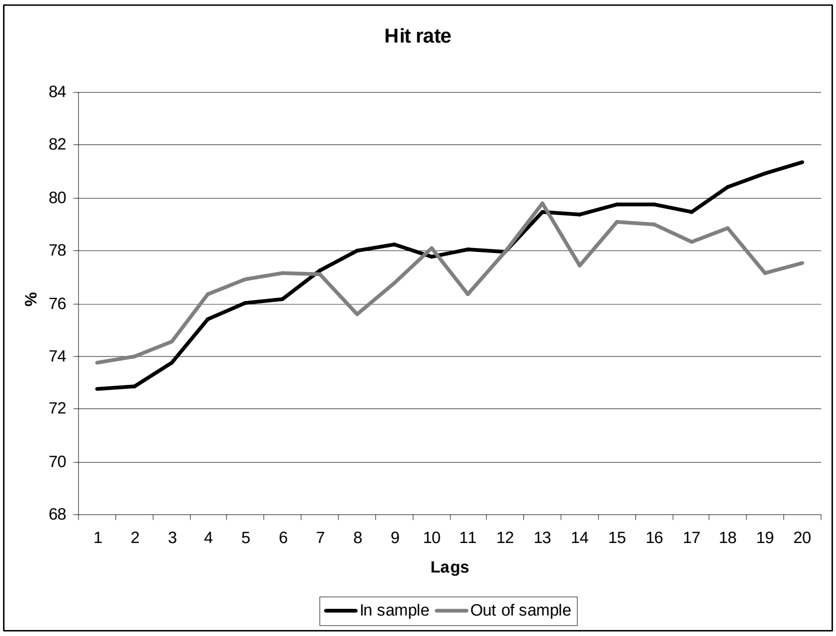

Figure 2 Hit rate for Model I at different number of lags used as input

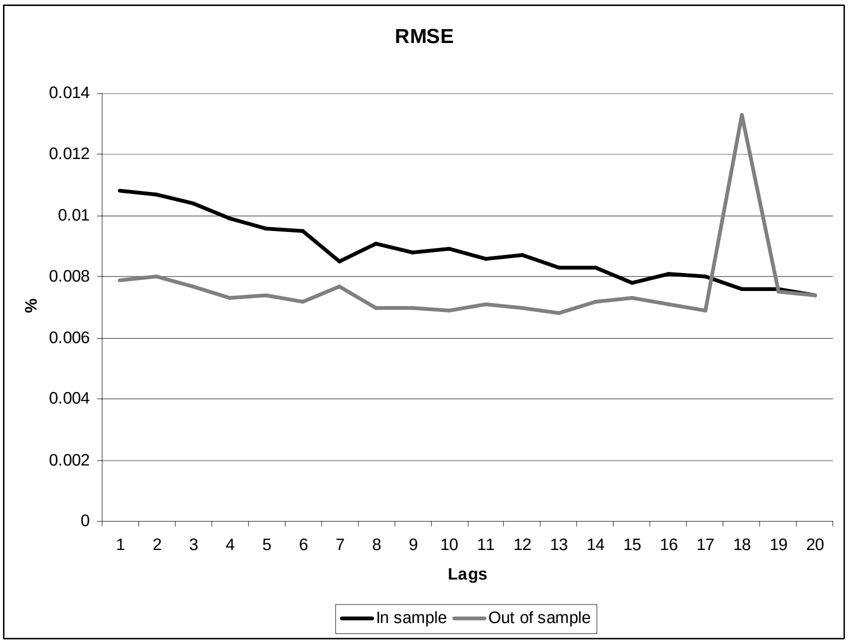

Figure 3 RMSE for Model I at different number of lags used as input

Table 3: *Summary of the benchmark performance (Model I)*

Metric	Hit Rate %	RMSE	R^2	MSE	MAE	SSE
In-sample	79.45	0.0083	0.6701	0.0001	0.0062	0.1486
Out-of-sample	79.79	0.0068	0.5762	0.0000	0.0053	0.0119

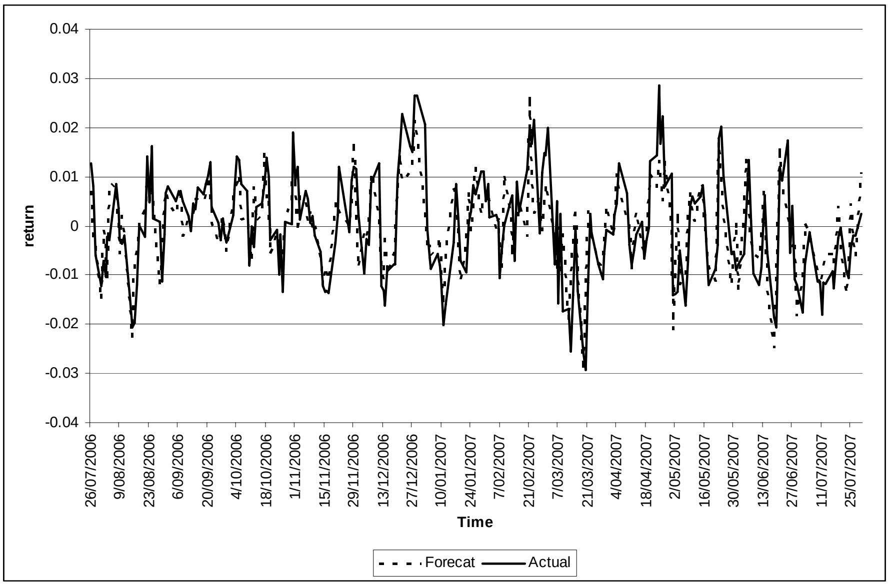

Figure 4 A plot of the Model I forecast vs. the actual values for out-of-sample

In addition to this the IC (the information coefficient) ratio was calculated

$$IC = \frac{\sqrt{\sum_{k=1}^{n}(y(t_k) - x(t_k))^2}}{\sqrt{\sum_{k=1}^{n}(x(t_k) - x(t_k - 1))^2}} \tag{23}$$

where x, y are the actual and predicted values, respectively. This ratio provides an indication of the prediction compared to the trivial predictor based on the random walk (Refenes 1995). Where $IC \geq 1$ indicate poor prediction, and $IC < 1$ means the prediction is better than the random walk. For the benchmark Model I, the IC is 0.58 for in-sample and 0.69 for out-of-sample, which means the model is significantly outperforming the trivial predictor.

6. Model II – from Futures Prices to Spot Prices

As defined in equation (11), Model II is limited to only using information extracted from one type of futures contracts for predicting the next day force of the spot prices. We use the same target variable for Model II in order to compare with the benchmark model. The four futures contracts are called ‘futures k’, for k=1, 2, 3, 4, with 1, 2, 3, 4 months to maturity respectively. A specific form of Model II for one-step prediction is defined by

$$D(X(futures,3,1),m,t) \Rightarrow ANN \Rightarrow X(spot,3,1,t+1) \tag{24}$$

The performance of the model ANN with each of the four futures contracts is shown in Tables 4-7.

Table 4: *Performance of Model II with futures 1 at different lags*

Lag	Hit Rate (%)		RMSE	
	In sample	Out of sample	In sample	Out of sample
1	71.72	73.55	0.0112	0.0079
2	71.92	73.80	0.0109	0.0081
3	71.42	71.96	0.0108	0.0079
4	73.65	75.15	0.0104	0.0075
5	74.37	74.91	0.0094	0.0085
6	74	76.51	0.0102	0.0074
7	75.23	77.61	0.0099	0.0073
8	74.99	75.26	0.0090	0.0081
9	75.40	75.52	0.0096	0.0085
10	75.64	77.49	0.0095	0.0072
11	76	76.51	0.0095	0.0073
12	75.82	77.24	0.0095	0.0070
13	76.35	76.01	0.0093	0.0073
14	76.91	75.89	0.0092	0.0074
15	76.78	75.52	0.0092	0.0074
16	76.84	78.11	0.0091	0.0072
17	77.94	75.52	0.0088	0.0076
18	78.09	76.63	0.0087	0.0079
19	77.66	77.24	0.0087	0.0078
20	78.87	76.14	0.0086	0.0078

Table 5: *Performance of Model II with futures2 at different lags*

Lag	Hit Rate (%)		RMSE	
	In sample	Out of sample	In sample	Out of sample
1	70.98	72.45	0.0114	0.0082
2	71.51	73.31	0.0111	0.0084
3	71.53	74.05	0.0110	0.0081
4	73.17	73.43	0.0106	0.0077
5	73.24	73.58	0.0096	0.0088
6	73.79	73.80	0.0104	0.0077
7	75.10	75.03	0.0102	0.0075

8	75.43	72.69	0.0101	0.0077
9	74.73	74.29	0.0099	0.0076
10	75.27	75.28	0.0097	0.0083
11	75.87	76.51	0.0098	0.0075
12	76.72	73.80	0.0094	0.0079
13	76.51	75.15	0.0094	0.0096
14	76.96	75.40	0.0094	0.0077
15	77.22	74.05	0.0094	0.0079
16	77.54	73.68	0.0091	0.0103
17	77.96	74.42	0.0090	0.0077
18	77.08	74.17	0.0089	0.0080
19	77.68	75.77	0.0088	0.0081
20	77.67	72.82	0.0088	0.0085

Table 6: *Performance of Model II with futures 3 at different lags*

Lag	Hit Rate (%)		RMSE	
	In sample	Out of sample	In sample	Out of sample
1	70.96	71.46	0.0115	0.0083
2	70.80	72.32	0.0112	0.0086
3	70.93	74.29	0.0111	0.0083
4	72.87	74.54	0.0107	0.0080
5	73.20	74.54	0.0106	0.0080
6	73.13	74.54	0.0105	0.0079
7	74.44	74.29	0.0103	0.0077
8	75.22	74.17	0.0100	0.0078
9	74.84	72.94	0.0101	0.0079
10	75.12	75.52	0.0098	0.1078
11	75.71	75.28	0.0097	0.0077
12	76.21	76.75	0.0097	0.0181
13	76.44	76.63	0.0094	0.0078
14	76.64	75.65	0.0096	0.0083
15	76.74	75.15	0.0094	0.0521
16	76.74	74.78	0.0091	0.0086
17	75.52	73.75	0.0109	0.0085
18	77.64	74.29	0.0091	0.0082
19	77.35	74.42	0.0089	0.0085
20	77.28	73.43	0.0092	0.0084

Table 7: *Performance of Model II with futures 4 at different lags*

Lag	Hit Rate (%)		RMSE	
	In sample	Out of sample	In sample	Out of sample
1	70.11	72.32	0.0116	0.0084
2	70.20	71.71	0.0113	0.0087
3	70.57	72.94	0.0112	0.0083

4	72.42	74.05	0.0108	0.0081
5	73.05	74.29	0.0107	0.0082
6	72.94	74.66	0.0106	0.0079
7	74.42	75.65	0.0104	0.0079
8	74.16	72.94	0.0103	0.0078
9	74.21	73.68	0.0102	0.0137
10	74.47	74.91	0.0100	0.0081
11	75.72	76.26	0.0099	0.0432
12	74.88	74.66	0.0100	0.0080
13	75.42	74.29	0.0097	0.0154
14	75.49	75.65	0.0098	0.0079
15	75.61	74.17	0.0097	0.0126
16	76.07	76.14	0.0097	0.0085
17	76.43	72.69	0.0095	0.0082
18	76.59	73.80	0.0093	0.0086
19	76.90	74.91	0.0090	0.0398
20	77.02	72.82	0.0092	0.0083

Tables 4-7 show that none of the futures contracts alone as input is able to outperform the benchmark. For contract 1, even with 20 lags, the prediction is less accurate than what we obtained from spot prices solely. Nonetheless, it is fair to say that the performance of futures as input is not poor either. These results remind us that the crude oil spot markets and futures markets are virtually equally efficient on the daily basis. However, this does not rule out the possibility that futures may lead the spot prices on the intraday basis.

These results are also plotted in Fig. 5.

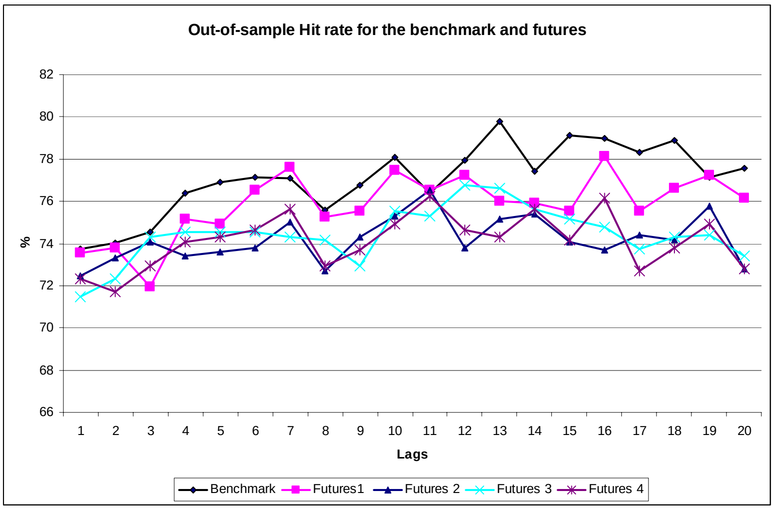

Figure 5 Out-of-sample performance of Model II at differeng number of lags

It is interesting to note that at 11 lags all the futures and spot information sets generated the same hit rate, which could indicate something in the crude oil market dynamics, worth for further study.

7. Model III – from Spot and Futures Prices to Spot Prices

As defined by equation (12), Model III is an augmentation of the benchmark model by adding any of the futures contracts to the information set of the spot prices. A specific form of Model III for one-step prediction is given by

$$\big(D(X(spot,3,1),m,t),\quad D(X(futures,3,1),1,t)\big) \Rightarrow ANN \Rightarrow X(spot,3,1,t+1) \tag{25}$$

The performance of Model III with each of the four futures contracts is shown in Tables 8-11. The result with all the four futures contracts added to the benchmarket input is given in Table 12.

Table 8: *In-sample Performance of Model III with futures 1-4 at different lags*

	Futures 1	Futures 2	Futures 3	Futures 4	All
Hit Rate	79.18	79.25	78.84	79.16	79.0398
RMSE	0.0084	0.0084	0.0084	0.0084	0.0083
R^2	0.652	0.6485	0.6461	0.6496	0.651
IC	0.5883	0.5896	0.5897	0.5876	0.8068
MSE	0.0001	0.0001	0.0001	0.0001	0.587
MAE	0.0063	0.0064	0.0064	0.0063	0.0001
SSE	0.1608	0.1624	0.1635	0.1619	0.0064

Table 9: *Out-of-sample Performance of Model III with futures 1-4 at different lags*

	Futures 1	Futures 2	Futures 3	Futures 4	All
Hit Rate	80.44	80	79.55	79.77	78.155
RMSE	0.0059	0.006	0.0063	0.0063	0.007
R^2	0.6806	0.6702	0.6383	0.6346	0.5459
IC	0.6358	0.6285	0.653	0.6539	0.7384
MSE	0	0	0	0	0.7019
MAE	0.0046	0.0047	0.0049	0.0049	0
SSE	0.0096	0.0098	0.0107	0.0108	0.0055

As can bee seen from Tables 8-9, adding futures to the benchmark do not outperform the benchmark in term of hit rate in-sample, while for out-of-sample futures 1 and futures 2 slightly outperform the benchmark. Further, there is no significant improvement for RMSE for in-sample, however, it does improve for out-of-sample for each of the futures compared to the benchmark, and futures 1 preforms the best. The R^2 is noticeable better for out-of-sample futures 1 compared to the benchmark indicating a better fit, while for in-sample it is less than the benchmark. The information coefficient ratio IC does not change for in-sample for all of the futures contracts, however, it is improved for out-of-sample for all futures contracts especially futures 1 and 2. Overall the performance is improved for out-of-sample and not much for in-sample. Finally, adding all the futures contracts 1-4 together to the benchmark has disadvantaged the model. It is safe to conclude that futures contracts of 1 and 2 months to maturity have slightly improved the out-of-sample prediction, however, this improvement is not significant enough to make further conclusions.

8. Model IV – from Spot Prices and Lead Markets to Spot Prices

As defined by equation (13), Model IV tests whether adding information from a lead market to the crude oil price information as input would improve the performance of prediction. Here we only use the latest information from the lead market. Therefore, a specific form of Model IV is given as

$$\left(D(X(spot,3,1),m,t),\quad D(X(lead,3,1),1,t)\right) \Rightarrow ANN \Rightarrow X(spot,3,1,t+1) \tag{26}$$

We have tested four markets which may be leading the crude oil prices, including S&P 500, Dollar Index, Gold, and Heating Oil. The results of these experiements are shown in Tables 10-11.

Table 10: In-sample performance of Model IV with different lead markets

	S&P 500	**Dollar Index**	**Gold**	**Heating Oil**	**All**
Hit Rate	79.18	79.20	79.14	79.37	78.96
RMSE	0.0083	0.0083	0.0082	0.0083	0.0084
R^2	0.6468	0.6456	0.6494	0.6455	0.6337
IC	0.7131	0.7142	0.7105	0.7144	0.7262
MSE	0.0001	0.0001	0.0001	0.0001	0.0001
MAE	0.0063	0.0063	0.0062	0.0063	0.0064
SSE	0.1576	0.1581	0.1565	0.1582	0.1635

Table 11: Out-of-sample performance of Model IV with different lead markets

	S&P 500	**Dollar Index**	**Gold**	**Heating Oil**	**All**
Hit Rate	78.22	78	77.48	79.11	77.25
RMSE	0.0069	0.0068	0.0069	0.0069	0.0069
R^2	0.5631	0.5749	0.573	0.5752	0.5663
IC	0.7953	0.784	0.7862	0.784	0.7922
MSE	0.000	0.000	0.000	0.000	0.000
MAE	0.00538	0.0054	0.0054	0.0054	0.0054
SSE	0.0130	0.0126	0.0126	0.0126	0.0128

From Tables 10-11, it turns out that the Heating Oil is the most informative lead market for predicting crude oil prices.

9. Multi-Step Prediction by four Models

Since moving average transformation is used, it is important to test the model capability to predict for the mid-point of the moving average (Refenes 1995). For example if the data transformed into 10 days moving average then it is important to test the model capability to predict 4 or 5 days in the future. In this study the data was transformed by 3-day moving averaging, therefore, the capability of the model to produce significant results for time $t+2$ is very important for evaluating the real performance of the model. From our first-hand many-years experience in trading futures markets (on an intraday basis), it makes tremendous difference if one can predict for time $t+2$ beyond $t+1$, simply because most professional futures traders do their predictions for $t+1$ anyway with their instinct, which has proved not less competent with advanced model-based prediction. Furthermore, to stretch the model even further, *t+3* was also included in our study. The input is still the same as the input for each of the four models presented in Sections 5-8, but now the output includes *t+1, t+2, t+3*. The ANN has the same structure of the benchmark model. Only six neurons in the hidden layer were used.

Table 12: Performance on multi-step predictions of Model I

	Hit Rate			RMSE		
	t+1	t+2	t+3	t+1	t+2	t+3
In-sample	78.37	67.97	54.57	0.0088	0.0117	0.0137
Out-of-sample	79.46	67.15	52.89	0.0088	0.0117	0.0137

Table 13: Performance on multi-step predictions of Model II

	Hit Rate			**RMSE**		
	t+1	**t+2**	**t+3**	**t+1**	**t+2**	**t+3**
Futures1	78.78	67.99	56.69	0.0088	0.0116	0.0136
Futures2	78.86	68.09	56.42	0.0088	0.0116	0.0135
Futures3	78.99	68.17	55.80	0.0088	0.0116	0.0136
Futures4	78.73	68.22	56.00	0.0088	0.0116	0.0136

Table 14: Performance on multi-step predictions of Model III

	Hit Rate			**RMSE**		
	t+1	**t+2**	**t+3**	**t+1**	**t+2**	**t+3**
Futures1	78.11	67.16	54.37	0.0088	0.0116	0.0136
Futures2	78.47	67.90	52.77	0.0088	0.0116	0.0135
Futures3	78.60	69.99	54.00	0.0088	0.0116	0.0136
Futures4	78.35	68.76	54.61	0.0088	0.0116	0.0136

By comparing table 12 to 14, it is clear that futures 1, 3, and 4 improve the out-of-sample prediction for time *t+3* while futures 2 performs as good as the benchmark.

Table 15: In-sample performance on multi-step predictions of Model IV

	Hit Rate			RMSE		
	t+1	**t+2**	**t+3**	**t+1**	**t+2**	**t+3**
S&P	78.62	67.37	55.05	0.0088	0.0117	0.0136
Dollar index	78.78	67.57	54.86	0.0089	0.0118	0.0137
Gold	78.58	68.14	55.46	0.0089	0.0117	0.0137
Heating Oil	78.39	67.72	58.75	0.0089	0.0116	0.0134

Table 16: Out-of-sample performance on multi-step predictions of Model IV

	Hit Rate			RMSE		
	t+1	**t+2**	**t+3**	**t+1**	**t+2**	**t+3**
S&P	78.47	66.54	50.80	0.0088	0.0117	0.0136
Dollar index	78.47	66.54	51.91	0.0087	0.0072	0.0791
Gold	78.11	65.93	53.14	0.0089	0.0117	0.0137
Heating Oil	78.60	69.74	60.64	0.0089	0.0116	0.0134

Azoff (1994) claimed that using multiple outputs in the output layer could affect the training of the ANN, as it could alter the weights update during the training session. Therefore, it is essential to insure that including more than one output did not affect the goodness of the results when compared to the network with only one output. Table 12 shows that the hit rate for in- and out-of-sample for network with three outputs was comparable to what obtained from the network with one output. In addition, the network performance is significant in term of hit rate for up of 2 steps ahead. While on average the performance has declined sharply for the third step prediction especially for out-of-sample. However, it is fair to recognise that on average even for time t+3 the output of the network is still better than guessing.

On the other hand, Model IV with a lead market has performed well for three-step prediction, as shown in Tables 15-16. The best result is achieved with the Heating Oil.

10. Conclusions

In this paper we have reported on a systematic study for developing ANN-based models to predict the short-term trends of the crude oil prices using the current information set from the crude oil spot prices, futrues prices, and other lead market prices. The short-term trends are represented by the 3-day moving averages of the spot prices for the next 3 days. Momentums (relative first-order differences) and forces (relative second-order differences) of the price time series are calculated to form the information set as input and output. Four models are designed and tested: from spot to spot, from futures to spot, from spot and futures to spot, and from spot and lead market to spot. For prediction at t+1, the best out-of-sample hit rate is 79.95% from the benchmark model. For t+2, and t+3, the best our-of-sample hit rates are 69.74%, and 60.64%, respectively, produced by the fourth model using the spot prices and the heating oil prices for input. The predictability of crude oil price trends for the next three steps at this level are nontrivially useful for economic agents and policy makers to make their informed decisions regarding crude oil risk management and trading.

Several transformations on the original price data were tested, it was found that 3-day moving averaging to the original data as preprocessing leads to much higher hit rate of prediction. Attentions was also paid for finding the optimal ANN model structure, as well as discovering the optimal number of lags (embedding dimensions in chaos theory). Weak evidence was found in support that futures prices of crude oil WTI contain new information about oil spot prices on the daily basis, nonetheless, futures 1 and 2 are more informative than futures 3, 4. However, this does not rule out the possibility that futures may lead the spot markets of crude oil on the intraday basis. This paper is limited to daily prediction, while the intraday prediction of crude oil prices is certainly a radically different problem which is much more dependent on real-time news analysis.

References

- Azoff, M.E. (1994). *Neural network time series forecasting of financial markets*. Chichester: John Wiley & Sons.
- Bopp, A.E., & Sitzer, S. (1978). Are petroleum futures prices good predictors of cash value? *The Journal of Futures Market, 7*(6), 705-719.
- Bowden, G.J., Dany, G.C., & Maier, H.R. (2006). An evaluation of methods for the selection of inputs for artificial neural networks based river model. In F. Recknagel (Ed.), *Ecological informatics scope, techniques and applications* (2nd Edition ed., pp. 275-292), Springer.
- Brock, W., Dechert, W., Scheinkman, J., & LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews, 15*(3), 197-235.
- Coppola, A. (2007). *Forecasting oil price movements: Exploiting the information in future market.* Working Paper (103) Centre for International Studies on Economic Growth . Retrieved June 6, 2007, from ftp://www.ceistorvergata.it/repec/rpaper/No-100.pdf
- Grothmann, R. (2005). *Multi-agent market modelling based on neural networks.* Unpublished PhD, University of Bremen, Germany.
- Haidar, I., Pan, H.P., & Kulkarni, S. (2008). Analysis of the relation between crude oil futures prices and spot price using nonlinear artificial neural networks. Proc. *37th Australian Conference of Economists* Gold Coast Australia.
- Haidar, I., Pan, H.P., Kulkarni, S., & Wu, X. (2007). A brief review of oil prices and its effect on financial markets and the economy. *Proc. Second International Workshop on Intelligent Finance (IWIF-II)*, July 2007, Chengdu, China, www.swingtum.com/institute/IWIF

- Hassoun, M.H. (1995). *Fundamentals of artificial neural networks*. London: Massachusetts Institute of Technology.
- Haubrich, J.G., Patrick, H., & Miller, J. (2004). Oil prices: Backward to the future? *Economic Commentary Federal Reserve Bank of Cleveland*, (Dec) .
- Haykin, S. (1999). *Neural networks: A comprehensive foundation* (2nd ed.). New Jersey: Prentice Hall.
- Hillard, H.G. (1994). Oil price forecasting in the 1980's: What went wrong? *Energy Journal*, *15*(2), 1-22.
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforwards are universal approximators. *Neural Networks*, *2*, 359-366.
- Huang, N.E., Shen, Z., Long S.R., Wu M.C., Shih, S.H., Zheng, Q., Yen, N-C., Tung, C.C., Liu, H.H. (1998). The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. Roy. Soc. Lond.*, 454, 903-993.
- Kaboudan, M.K. (2001). Compumetric forecasting of crude oil prices. *The 2001 Congress on Evolutionary Computation*, 283-287.
- Kodogiannis, V., & Lolis, A. (2002). Forecasting financial time series using neural networks and fuzzy system-based techniques. *Neural Computing & Applications*, *11*, 90-102.
- Labonte, M. (2004). *The effect of oil shocks on the economy: A review of the empirical evidence.* Congress report. Retrieved November 25, 2007, from http://www.ncseonline.org/NLE/CRSreports/04Jun/RL31608.pdf
- Liu, J., Bai, Y., & Li, B. (2007). A new approach to forecast crude oil price based on fuzzy neural network. *FSKD '07: Proceedings of the Fourth International Conference on Fuzzy Systems and Knowledge Discovery*, 273-277.
- McNelis, P.D. (2005). *Neural networks in finance gaining predictive edge in the market*. Massachusetts: Elsevier Academic Press.
- Milunovich, G., & Ripple, R.D. (2006). *Hedgers, investors and futures return volatility: The case of NYMEX crude oil.* Working Paper (7/2006) Department of Economics Macquarie University. Retrieved March 3, 2007, from http://ideas.repec.org/p/mac/wpaper/0607.html#provider
- Mittermayer, M., & Knolmayer, G.F. (2006). *Text mining systems for market response to news: A survey*. Working Paper (184) Institute of Information System, University of Bern. Retrieved March 15, 2007, from http://www.ie.iwi.unibe.ch/publikationen/berichte/resource/WP-184.pdf
- Moshiri, S., & Foroutan, F. (2005). Forecasting nonlinear crude oil futures prices. *The Energy Journal*, *27*(4), 81-95.
- Neuneier, R., & Zimmermann, H.G. (1998). How to train neural networks. In Genevieve B. Orr, Klaus-Robert Müller (Ed.), *Neural networks: Tricks of the trade* (pp. 273-423). Berlin: Springer.
- NYMEX. (2006). *A guide to energy hedging.* Retrieved from http://www.nymex.com
- Pan, H.P. (2004), A Swingtum Theory of Intelligent Finance for Swing Trading and Momentum Trading. Proc. First International Workshop on Intelligent Finance (IWIF 1), Dec, 2004, Melbourne.

- Pan, H.P. (2006), Multilevel Stochastic Dynamic Process Models and Possible Applications in Global Financial Market Analysis and Surveillance. *Proceedings of the 5th International Conference on Computational Intelligence in Economics and Financ*, October 2006, Taiwan.
- Pan, H.P.（2008）A Basic Theory of Intelligent Finance, *New Mathematics and Natural Computation, Special Issue on Intelligent Finance*, to appear.
- Pan, H., Tilakaratne, C., & Yearwood, J. (2005). Predicting Australian stock market index using neural networks exploiting dynamical swings and intermarket influences. *Journal of Research and Practice in Information Technology, 37*(1), 43-55.
- Pan, H.P., Sornette, D., & Kortanek, K. (2006): Intelligent Finance – An Emerging Direction. *Quantitative Finance*, Vol. 6, No. 4, August 2006, 273-277.
- Pindyck, R.S. (1999). The long-run evolution of energy prices. *The Energy Journal, 20*(2), 1-27.
- Refenes, A. (Ed.). (1995). *Neural networks in the capital markets*. New York: John Wiley & Sons.
- Silvapulle, P., & Mossa, I.A. (1999). The relation between spot and future prices: Evidence from the crude oil market. *The Journal of Futures Markets, 19*(2), 175-193.
- The International Energy Agency. (2004). *Analysis of the impact of high oil prices on the global economy* No. IEA/(2004))The International Energy Agency.
  Retrieved from http://www.iea.org/textbase/papers/2004/high_oil_prices.pdf
- Wang, S., Yu, L., & Lai, K.K.. (2005). Crude oil price forecasting with TEI@I methodology. *Journal of Systems Science and Complexity, 18*(2), 145-166.
- Xie, W., Yu, L., Xu, S., & Wang, S. (2006). A new method for crude oil price forecasting based on support vector machines. In V.N. Alexanderov et al. (Ed.), *Lecture notes in computer science* (pp. 444-451). Heidelberg: Springer.
- Yoo, P. D., Kim, M. H., & Jan, T. (2005). Machine learning techniques and use of event information for stock market prediction: A survey and evaluation. *Computational Intelligence for Modelling, Control and Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce*, 835-841.
- Yu, L., Lai, K.K., Wang, S., & He, K. (2007). Oil price forecasting with an EMD-based multiscale neural network learning paradigm. *Lecture notes in computer science* (pp. 925-932). Berlin / Heidelberg: Springer.
- Yu, L., Wang, S., & Lai, K.K. (2005). A rough-set-refined text mining approach for crude oil market tendency forecasting. *International Journal of Knowledge and System Sciences, 2*(1), 33-46.