Crude oil price forecasting techniques: a comprehensive review of literature

Niaz. Bashiri Behmiri[1], José R. Pires Manso

Department of Management and Economics, University of Beira Interior, Covilhã, Portugal.

Abstract

The goal of this article is to review the existing literature on crude oil price forecasting. We categorized the existing forecasting techniques into the two main groups of quantitative and qualitative methods; and then we performed an almost comprehensive survey on the available literature with respect to these two main forecasting techniques.

Keywords: crude oil price, forecasting, literature review.

1. Introduction

In the last several years, oil prices showed great variations, they raised and dropped down dramatically in various intervals. Since, oil is one of the strategic commodities and plays a critical role in effecting on the word's economy and macroeconomics factors such as inflation, recession, GDP, interest rates, exchange rates, and etc; therefore, the determinants of oil prices and its fluctuations have been one of the most favorable subjects for energy researchers and economists. As a result, achieving to a reliable and highly accurate forecasting and answering to the complexities of the crude oil prices would be important to policy makers. For this purpose, various techniques have been tried to forecast the movements, fluctuations or volatility of the crude oil prices. Among them, the most popular methods are econometrics methods, although more recently the computational approaches such as artificial neural networks and fuzzy expert systems gained more popularity in financial markets, because they are more flexible and accuracy than traditional methods. However, still there is not a general consensus that which methods are more reliable.

In this study, we categorize the existing oil price forecasting literature into the two main groups: quantitative and qualitative methods. Quantitative methods include econometrics and computational approaches, and qualitative methods contain computational and technological

[1] Corresponding author
Email address:bashiri.niaz@gmail.com

approaches. We divide the paper into the four sections. Section 2 categorizes the excising forecasting techniques. Section 3 assesses the existing literature, which performed forecasting on oil prices. And section 4 is the conclusion.

2. Crude oil price forecasting techniques

Among oil price forecasting literature, there are two main categories of forecasting methods: quantitative and qualitative methods. Quantitative methods represent numerical and quantitative variables that impact on oil prices; include two categories of techniques: econometrics methods and non-standard methods. Among them, econometrics models are grouped into the three classes of models: time series models, financial models, and structural models. On other hand, the main non-standard methods which are the most frequently applied in terms of oil price forecasting are artificial neural networks and support vector machines. On the other side, Qualitative methods estimate impact of infrequent events such as wars and natural events on oil prices; these approaches recently obtained more popularity among oil price forecasting literature. However, between various types of qualitative forecasting methods there are small number of them that are applied to forecast oil prices, such as Delphi method, belief networks, fuzzy logic and expert systems, and web text mining method. The classification of techniques is shown as bellow:

1. Quantitative methods:

1.1. Econometrics models: 1.1.1.Time series models,1.1.2.Financial models,1.1.3. Structural models.

1.2. Non-standard methods: Artificial Neural Networks, Support Vector Machines.

2. Qualitative methods: Delphi method, Belief Networks, Fuzzy Logic and Expert Systems, Web Text Mining method.

3. Review of the literature

1. Quantitative models

Quantitative methods are based on historical data and mathematical models and focus on short and medium term predictions. They are applied to model the quantitative variables, which impact on oil market, grouped into the two major types of models: 1. econometrics models, and 2. non-standard models.

1.1. Econometrics models

Among existing oil price forecasting literature, econometrics models are the most frequently used methods. In this study, we classify existing literature on econometrics methods based on the three main categories: time series, financial, and structural models:

1.1.1. Time series models

Time series models predict future oil prices through using historical data. In these models the future price behaviors are deduced from its own historical data. These models are mostly employed when *(a)* the data show systematic pattern, such as autocorrelation (*b)* the most number of possible explanatory variables, and their interactions provide a structural model which is very difficult to be performed *(c)* forecasting of dependant variable depends on the prediction of the explanatory variables, which may be even more complicated than forecasting the variable itself. All seem to be the cases of oil prices.

Time series models include three main categories: naïve models, exponential smoothing models, and autoregressive models such as ARIMA[2] and ARCH[3]/GARCH[4] family models which the last one applies to modeling and forecasting the volatility of crude oil prices. In this context, Pindyck (1999) examines long run behavior of crude oil, coal, and natural gas prices during 127 years covering the period from 1887 to 1996. He incorporates unobservable state variables such as marginal costs, the resource reserve base, and the demand parameters into the model and estimates the model with a Kalman filter. The author examines the forecasting ability of the model with adding mean reversion to a deterministic linear trend, under three sub-samples and full database. The results suggest that model with specification of deterministic linear trend produces more accurate forecasts. Radchenko (2005) extends the Pindyck study with performing a long term forecast of energy prices. He applies a shifting trend model with autoregressive process in error terms instead of a white noise process. This study uses 127 years data that covers the same period with Pindyck, and examines forecasting ability of the model in four sub-sample

[2] Autoregressive integrated moving average
[3] Autoregressive conditional heteroskedastisity
[4] Generalized autoregressive conditional heteroskedastisity

time horizons. The results confirm Pindyck conclusions. The author states that the shortcoming of the model is inability to consider the impact of OPEC behavior; for this reason, he combines the model with autoregressive and random walk models and concludes that the result from the combination model outperforms the single model. Lanza et al. (2005) estimate the relationship among heavy crude oil prices and product prices. The authors implement a comparison among ten heavy crude oil price series and fourteen petroleum product price series in Europe and America. The sample period is from 1994 to 2002, and they apply cointegration and error correction (ECT) tests to find out the relationships among the variables, and to forecast crude oil prices. The empirical results indicate, there is evidence that product prices are related to heavy oil prices in short and long term. Furthermore, the comparison among ECM and naïve model does not show any dominant model in America; however in the case of Europe, ECM marginally outperforms naïve model. Wang et al. (2005) carry out ARIMA approach to model the linear component of crude oil price time series. They use monthly WTI crude oil data from January 1970 to December 2003. The out-of-sample forecasting results indicate that the linear ARIMA model displays the poorest forecasting power in compare with the nonlinear artificial neural network and the nonlinear integrated fuzzy expert system approaches which will be discussed broadly in the next section. Xie et al. (2006) forecast WTI crude oil prices with applying ARIMA method. They apply WTI spot prices from January 1970 to December 2003. Then they compare the results with those of support vector machine and artificial neural networks methods. The out-of-sample forecasting results indicate that, the ARIMA model provides the poorest forecasting performance among the mentioned methods. Fernandez (2010) performs an out-of-sample forecasting for short and long term horizons with using ARIMA model. The author employs daily natural gas and Dubai crude oil prices from 1994 to 2005. The result prove that for very short term horizon forecast, the ARIMA model outperforms the artificial neural networks and the support vector machine approaches; however, for long term horizon forecast, the ARIMA model provides the poorest accurate forecasts.

Crude oil market is one of the most volatile markets in the world and shows the strong chaos, and ARIMA model is a linear model which captures time series linear characteristics; therefore, there is a general consensus that this model is not able to describe the nonlinearity components of oil price time series.

In financial institutions managing risk is an important element. For this purpose, the so called Value-at-Risk (VaR) model is a quantitative tool that its goal is to estimate the possible loss of the financial institution during a specific time period and for a given portfolio of assets. In this context, there are some studies that investigate the VaR in oil market: Goit and Laurent (2003) calculate the VaR in several commodity markets; they compare forecasting ability of ARCH-type models include RiskMetrics, Skewed student asymmetry power ARCH (APARCH) and Skewed student ARCH models on aluminum, copper, nickel, Brent and WTI crude oil and cocoa. They use daily spot prices for the whole commodities but nearby future contracts for cocoa, the data set for Brent and WTI cover the period from May, 20, 1987 to March, 18, 2002. The out-of-sample forecasts indicate that the Skewed student APARCH model provide better results and outperform the other ARCH-type models. Cabedo and Moya (2003) asses the VaR with comparing the forecasting power of three VaR calculation method, including historical simulation standard (HS) approach, historical simulation with ARMA forecasts (HSAF) approach, and variance-covariance method based on GARCH model forecasts. They use daily Brent crude oil spot prices from January 1992 to December 1992. The out-of-sample forecasting results suggest that the GARCH model overestimate the oil price changes and the HSAF methodology provides more efficient risk quantification than (HS) and GARCH approaches. Costello et al. (2008) update the research by Cabedo and Moya (2003), they calculate the VaR with comparing the forecasting ability of ARMA with historical simulation and semi-parametric GARCH model or GARCH model with historical simulation proposed by Barone-Adesi (1995) which has applied by Morana (2001). They use daily Brent crude oil spot price data from May, 20, 1987 to January, 18, 2005. They conclude that in terms of out-of-sample forecasting, the semi-parametric GARCH model is superior to the ARMA model with historical simulation. Sadeghi and Shavvalpour (2006) evaluate VaR with comparing the forecasting ability of historical simulation ARMA forecasting (HSAF) and Variance-Covariance based on GARCH forecasting models. They use weekly OPEC spot prices from January 1997 to December 2003. The out-of-sample forecasting results state that the Variance-Covariance GARCH models overestimate the oil price changes and HSAF models provide more efficient VaR, this conclusion supports the VaR estimation results by Cabedo and Moya (2003). Fan et al. (2008) calculate VaR with comparing the forecasting power of GARCH-type models based on Generalized Error Distribution (GED) including GED-GARCH, GED-TGARCH and historical simulation ARMA forecasting (HSAF) model. They apply daily WTI and Brent crude oil spot prices from May, 20, 1987 to August, 1, 2006. The results prove that the GED-GARCH type models outperform HSAF model in terms of out-of-sample forecasting. Hung et al. (2008) investigate the influence of fat-tailed innovation process in order to VaR estimation. They apply three GARCH type models, including GARCH model with normal distribution, GARCH model with student t distribution, GARCH model with heavy tailed distribution; using spot prices of five energy commodities; WTI and Brent crude oil, heating oil #2, propane and New York Harbor Conventional Gasoline regular for the period of September, 9, 1996 to August, 31, 2006. The out-of-sample VaR forecast results show that the GARCH model with heavy tailed distribution outperforms the other models in terms of accuracy and efficiency; therefore, fat-tailed in return innovation process plays a significance role in VaR estimation and should be considered in risk management investigation. Huang et al. (2009) measure the market risk with improving conditional autoregressive VaR by regression Quantiles (CAViaR) model which proposed by Engle and Manganelli (2004), the CAViaR specifications generalize the traditional GARCH-based VaR models by allowing for different stochastic processes in the tail of the return series (Huang et al 2009). They use daily WTI spot prices from November, 29, 1996 to November, 28, 2006. The results show that the model performs very well in terms of in-sample and out-of sample forecasting. Aloui and Mabrouk (2010) evaluate the VaR with comparing forecasting performance of fractionally integrated GARCH (FIGARCH), fractionally integrated asymmetry power ARCH (FIAPARCH) and hyperbolic GARCH (HYGARCH) models with normal, student and skewed student innovation's distributions. They apply daily spot prices of WTI, Brent, New York Harbor conventional regular gasoline (NYHCGR), and Rotterdam conventional regular gasoline (RCGR), covering period for WTI, NYHCGR and RCGR is January 1986 to July 2007 and for the Brent crude oil is May 1987 to July 2007. The results state that the FIAPARCH model with skewed student-t innovation's distribution is superior to the FIGARCH and HYGARCH models in terms of in and out-of sample forecasting performance but the accuracy of the FIAPARCH model in terms of out-of-sample forecasting is not as good as in-sample volatility estimation.

Furthermore, there are a vast number of studies, which assess the volatility of crude oil market via comparing the ARCH/GARCH class of models. For instance, Cheong (2009) compares the volatility forecasting ability of the GARCH type models contain GARCH, asymmetric power

ARCH (APARCH), fractionally integrated GARCH (FIGARCH) and fractionally integrated asymmetric power ARCH (FIAPARCH) all with normal and student-t distribution. The author uses daily WTI and Brent crude oil spot prices data for the period from January, 4, 1993 to December, 31, 2008. The out-of-sample forecasting accuracy is estimated for 5-20-60 and 100 day horizons. The results indicate that in the case of Brent crude oil prices the standard short memory GARCH normal and student-t models outperform for the 5 and 20 day horizons forecasts and the APARCH-normal model outperforms for the 60 and 100 day horizons forecasts; and in the case of WTI crude oil prices the FIAPARCH-student-t model has the highest accuracy for the all time horizons of the out-of-sample volatility forecasts; therefore, according to the time horizons and the type of crude oil benchmark there are mix results. Kang et al. (2009) compare the volatility prediction ability of the GARCH, component GARCH (CGARCH), integrated GARCH (IGARCH) and fractionally integrated GARCH (FIGARCH) models. They apply daily Brent, WTI and Dubai crude oil spot prices over the period from January 6, 1992 to December 29, 2006. The out-of-sample forecasting analysis considers 1, 5 and 20 days forecasting horizons. The results indicate that for the all three forecasting horizons, in the case of Brent and Dubai crude oil price the FIGARCH model outperforms the other models and in the case of WTI crude oil price the CGARCH model outperforms the other models. Wei et al. (2010) extend the work by Kang et al. (2009) with applying nine linear and nonlinear GARCH type models including RiskMetriks, GARCH, integrated GARCH (IGARCH), Glosten-Jagannathan-Runkle GARCH (GJR-GARCH), exponential GARCH (EGARCH), asymmetry power ARCH (APARCH), fractionally integrated GARCH (FIGARCH), fractionally integrated asymmetry power ARCH (FIAPARCH), and hyperbolic GARCH (HYGARCH) models. They consider one-five and twenty days out-of-sample volatility forecasts and use daily Brent and WTI crude oil spot prices covering the period from January 6, 1992 to December 31, 2009. The out-of-sample forecasts show that across the all six loss function there is no evidence that there is a single GARCH model which outperforms the other models for both Brent and WTI; The only differentiation between models is that the linear GARCH type models seem to fit better for short run (one day) volatility forecasts and the nonlinear GARCH type models seem to fit better for long run (five and twenty days) volatility forecasts. Vo (2009) compares the forecasting ability of four different models including, Markov switching stochastic volatility (MSSV) model with constant variance which is the combination of regime switching with stochastic volatility model, stochastic volatility (SV) model, GARCH model and Markov switching (MS) model. The author applies WTI weekly spot prices for the period from January, 3, 1986 to January, 4, 2008. The out-of-sample results conclude that for the in-sample forecasts the accuracy depends on the evaluation criteria and are mixed but the simple MS model seems better; in terms of the out-of -sample forecasts the MSSV outperforms the other models under the all three evaluation criteria. Mohammadi and Su (2010) compare the out-of-sample forecasting ability of the GARCH, exponential GARCH (EGARCH), asymmetry power ARCH (APARCH) and fractionally integrated GACRH (FIGARCH) models. They apply weekly data on eleven crude oil (FOB) spot prices in international markets over the period of January, 3, 1997 to February, 13, 2009. The results indicate that the forecasting accuracy of the APARCH model outperforms the other models. And Silva et al. (2010) investigate the performance of hidden Markov model (HMM) to forecast the medium term future crude oil price movements. This approach is a nonlinear time series model which use historical time series data to forecast future prices. They use daily WTI spot prices and apply wavelets to omit the high frequency of price time series, and then perform HMM to forecast oil prices. The HMM model forecast the probability distribution of accumulated return over the next days; subsequently, from this distribution, they explore future price trends. The results suggest that HMM model provides good forecasting performance.

1.1.2. Financial models

In oil price forecasting area, financial models estimate the relationship among spot and future prices, and investigate whether future contract prices are unbiased based on unbiasedness hypothesis, and efficient based on efficiency market hypothesis (EMH) predictors of future spot prices. For this goal, Bopp and Lady (1991) examine the impact of lagged future and lagged spot oil prices on future spot prices. They use monthly data on heating oil traded on NYMEX[5] covering the period from December 1980 to October 1988. For this purpose, the authors apply an autoregressive model and conclude that the predictive power of each applied series depends on the type of applied data: when deseazonalized data is applied then the performance of series are the same, but when actual prices are applied then the forecasting ability of future prices superior to actual prices. Moreover, the forecasting performance of the autoregressive financial model is

[5] New York Mercantile Exchange

compared with the random walk model and conclude that both of the models perform equally well in terms of forecasting. Serletis (1991) investigates the future market efficiency or unbiasedness. The author uses daily spot and future prices on heating oil and crude oil traded on NYMEX from July, 1, 1983 to August, 31, 1988, and daily spot and future prices on unleaded gasoline covering the period of March, 14, 1985 to August, 31, 1988. The author applies cointegration test to find out the relationships among the variables and performs Fama's interesting variance decomposition method to test the joint measurement of variation in the premium and expected future spot prices, and conclude that variation in premium worsens the forecasting ability of future market. Green and Mork (1991) examine efficiency and unbiasedness among official oil prices and ex-post spot prices. For this purpose they use generalized method of moments (GMM) estimation approach; using Middle East light and African light/North sea monthly crude oil prices data covering the period from 1978 to 1985. They conclude that ex post spot prices are not efficient or unbiased if the sub-period is 1981 to 1985, but there is evidence of improvement in efficiency during the time. Sami (1992) examines WTI crude oil future prices (three and six months) as function of WTI spot prices and interest rate. The author employs daily data from September, 20, 1991 to July, 15 1992, and monthly data from January 1984 to June 1992. The results suggest that interest rate does not play a clear role on prices; in the other hand, although spot and future prices are highly correlated but the direction of causality relationship among them is not identified. Day and Lewis (1993) compare volatility forecasting accuracy of four different models including the standard GARCH, exponential GARCH (EGARCH), implied volatility (IV) and naïve historical volatility models. They use daily nearby and more distance future contracts prices from November 1986 to March 1991. The out-of-sample forecasting result indicate that the IV model outperforms the historical volatility model and the GARCH type models. Agnolucci (2009) updates the results of Day and Lewis (1993), estimates the forecasting ability of the GARCH, asymmetric power ARCH (APARCH), exponential GARCH (EGARCH), component GARCH (CGARCH), and threshold GARCH (TGARCH) models with normal, student-t and Generalized error distribution (GED). Then compares their prediction ability with implied volatility (IV) model; using daily return WTI crude oil future prices traded on NYMEX over the period of December, 31, 1991 to May, 2, 2005. The results show that the GARCH type models outperform the IV model; furthermore, among the GARCH type models those with asymmetric effects and GED distribution provide the best accuracy forecasts. Moosa and Aloughani (1994) also investigate the efficiency and unbiasedness in crude oil future markets. They Use monthly WTI crude oil spot and future prices traded on NYMEX covering the period from January 1986 to July 1990. The results of cointegration and error correction model (ECM) tests prove that future prices (here three months and six months' future contracts) are neither unbiased nor efficient forecasts of spot prices. Moreover, they perform a GARCH in mean model to consider the time varying risk premium. Zeng and Swanson (1998) examine the forecasting ability of future prices on spot price. For this purpose, several models are applied, including random walk with drift, random walk without drift, VAR model, and VECM models. They apply daily future prices for four commodities including crude oil traded on NYMEX, gold traded on New York Commodity Exchange, Treasury bond traded on Chicago Board of Trade, and S&P 500 index traded on Chicago Mercantile Exchange for the period of April, 1, 1990 to October, 31, 1995. The results indicate that the ECM model outperforms the other models. Gulen (1998) estimates efficiency and forecasting power of posted oil prices. For this purpose, he incorporates both posted and future oil prices as explanatory variables to the model. The author uses monthly WTI crude oil spot prices, and one month, three months, and six months future prices traded on NYMEX covering the period from March 1983 to October 1995. A cointegration test is performed to estimate the efficiency hypothesis of variables, and suggests that future prices are efficient predictor and outperform the posted prices; however, the posted prices show predictive ability in very short horizon. Schwartz and Smith (2000) introduce a short-term/long-term model by developing a two-factor model which proposed by Schwartz (1997) of commodity prices which reflects two effects, first reflects mean reversion in prices, and in the same time reflects the uncertainty from the equilibrium which price revert. They use two data sets, first weekly crude oil future contracts traded on NYMEX contracts from January, 2, 1990 to February, 17, 1995 use 5 forward contracts, and second data set is crude oil future prices covering the period from January, 15, 1993 to May, 16, 1996 use 10 forward contracts. In this short-term/long-term model, the uncertainty in equilibrium price level occurs according to geometric Brownian motion process, and the short-term deviation from equilibrium which is differences between spot and equilibrium prices are expected to revert to zero according to an Ornstein-Uhlenbeck process. Morana (2001) applies a semi parametric approach suggested by Barone-Adesi et al. (1998) to forecast the volatility of Brent crude oil price, which is based on the relationship among spot and future prices. The GARCH property of oil price volatility is applied to forecast the short-term horizon of price and the one-month forward price is applied as predictor. The author uses Brent crude oil daily prices over the period from January 4, 1982 to January 21, 1999. The forecasting results prove that, Brent forward prices seem to be a biased predictor of future spot prices, but in 50 percent of the cases the sign of price changes do not be predicted accurately. Furthermore, he compares financial model with a time series random walk model and concludes that in short term horizons both specifications are unbiased. Cortazar and and Schwartz (2002) perform a financial analysis on oil prices with extension of two-factor model of Schwartz (1997) to a three-factor model and examine the relationship between spot and future prices. They use daily price of all future contracts traded on NYMEX during the period from 1991 to 2001. In this three-factor model, the long term spot price return is allowed to be stochastic and to mean revert to a long term average. The in-sample and out-of-sample forecasts specify that the three-factor model performs better than the two-factor model and fits data quite well. Furthermore, the authors suggest a minimization procedure to Kalman filter approach which performs more reliable results. Fong and See (2002) apply a Markov regime switching (MRS) model to explain the volatility of oil prices. The model is based on the standard ARCH/GARCH approach with allowing jumps in the conditional variance between regimes. For this purpose, they use WTI crude oil daily prices for the nearest future contracts covering the period from January, 2, 1992 to December, 31, 1997. The results for the three out-of-sample forecasts suggest that the regime switching model with the ARCH effects (RSARCH) outperforms the constant variance and the standard GARCH model for the all three out-of-sample forecasts. Chernenko et al. (2004) examine the efficiency and unbiasdness of broad sort of forward and future rates that among them we only consider crude oil and natural gas price futures. They use monthly data on WTI forward prices for three months, six months, and 1 year future contracts traded on NYMEX from April 1989 to December 2003. The authors found that in most of the cases the forward and future rates are not efficient or unbiased predictors for future spot prices, in the other word they are not rational expectations of actual prices. In terms of crude oil and natural gas the results are mixed and there is less evidence of risk premium. They compare their financial model to a time series random walk specification and conclude that a random walk process is a better predictor than future prices. Abosedra and Baghestani (2004) estimate the unbiasedness of 1, 3, 6, 9, and 12 months ahead crude oil future prices, and a naïve forecasting model is performed as a benchmark. They use monthly WTI spot and future prices traded on NYMEX from January 1991 to December 2002. The empirical results suggest that future prices and naïve forecasts are unbiased in all time horizons, however the 1, and 12 months future ahead prices forecasts outperform the naïve forecasts. Abosedra (2005) employs a simple univariate model to examine unbiasedness and efficiency of crude oil spot and future prices. The author uses monthly WTI crude oil spot and future prices from January 1991 to December 2001. In this study, the goal is to forecast one month a head price of crude oil for every trading day with using the previous trading day's spot prices, also uses monthly average of daily forecasts as the naïve monthly predictor, and supposes that WTI spot prices can be approximated by random walk process with no drift. He concludes that future price for one month a head contracts seems to be an unbiased and semi strongly efficient predictor. Chin et al. (2005) Examine future energy commodity prices in terms of being unbiased and producing accurate forecast of future spot price. They use monthly data on WTI, gasoline, heating oil, and natural gas spot and three, six, and twelve months future prices from January 1999 to October 2004, and suppose that spot prices follow a random walk process with drift and rational expectations. The results suggest that future prices are unbiased predictors of spot prices except in the case of natural gas in three month horizon; however, they do not perform well in producing accurate forecast of future spot prices. Moreover, they outperform time series models. Yousefi et al. (2005) apply wavelet methodology in providing out-of-sample forecasts for oil prices in 1, 2, 3, and 4 forecasting time horizons, and then investigate the efficiency of future contracts on spot prices. They use two groups of datasets, including average monthly WTI spot prices, and WTI forward prices traded on NYMEX future contracts, covering the period from January, 2, 1986 to January, 31, 2003. They conclude that, the wavelet procedure outperforms the future markets and this superiority does not decrease with extension of time horizon; furthermore, prove that future markets are not efficiently priced. Sadorsky (2006) compares different types of forecasting models, including the random walk, historical mean, moving average, exponential smoothing, linear regression model, autoregressive models, GARCH, threshold GARCH (TGARCH), GARCH in mean, state space, vector autoregressive (VAR), and bivariate GARCH (BIGARCH) models to forecast petroleum prices. He uses WTI daily future prices of crude oil, heating oil #2, unleaded gasoline covering the period from February 5, 1988 to January 31, 2003 and in the case of natural gas data covers the period of April 3, 1990 to January 31, 2003 traded on NYMEX. The results show that in the case of heating oil and natural gas the TGARCH model fits better, and in the case of crude oil and unleaded gasoline the GARCH model fits better; therefore, the GARCH type models outperform the other techniques. Coppola (2008) investigates long run relationship between spot and future oil prices; using weekly WTI spot and future prices traded on NYMEX from January 1986 to September 2006. The author performs cointegration test and VECM to examine short and long run relationships among spot and future prices. The in-sample forecasting results indicate that future prices seem to explain well the spot price movements, and the out-of-sample forecasting results suggest that the VECM outperforms the random walk model. Alizade et al. (2008) are the first to apply the markov regime switching approach to estimate the time varying minimum variance hedge ratio (Hung et al. 2011) by introducing Markov regime switching error correction model with GARCH error structure; applying weekly spot and future prices for WTI, unleaded gasoline and heating oil traded on NYMEX from January, 23, 1991 to December, 27. 2006. The in and out-of-sample forecasting results specify that the dependant hedge ratios are able to provide significant reduction in portfolio risk (Alizadeh et al 2008). Murat and Tokat (2009) examine the relationships between crude oil price with crack spread futures, and future prices. In oil market the crack spread is the difference between crude oil prices and crude oil product prices. For this purpose, they use weekly WTI spot prices and weekly prices of NYMEX future contracts from January 2000 to February 2009. They apply Johansen cointegration test and VECM approach to analyze the Granger causality relationship between the two variables and to forecast WTI oil prices. Furthermore, they apply a time series random walk model as a benchmark and conclude that the random walk model displays the poorest forecasting accuracy, and forecasting performance of VECM with crack spread futures is almost as well as ECM with crude oil futures. Hung el al. (2011) evaluate the hedge ratio of WTI crude oil future market through comparing the in and out-of-sample hedging performance of a four-regime bivariate Markov regime switching model, two-regime switching model, Constant Correlation GARCH (CC-GARCH), Time Varying Correlation GARCH (TVC-GARCH) and OLS model. They employ WTI daily spot and nearest future contract prices traded on NYMEX from January, 2, 2002 to December, 31, 2007. The results specify that the four-regime bivariate Markov regime switching model outperforms the other models in terms of the in and out-of-sample hedging performance. Nomikos et al. (2011) estimate the forecasting volatility and VaR performance of various volatility regime switching models including the MIX (distribution) GARCH and two regimes MRS-GARCH models based on the mixed conditional heteroscedastisity models proposed by Haas et al. (2004a) and Alexander and Lazar (2006) and Markov model of Haas et al. (2004.b). For the first time in a regime volatility model they append the squared lagged basis of future prices in the specification of the conditional variance GARCH-X models which proposed by Lee, 1994; NG and Pirrong, 1996; then, extend this framework with adding a conditional extreme value theory (EVT) setting. They apply daily WTI crude oil and heating oil future prices traded on NYMEX from January, 23, 1991 to December, 31, 2008, and Brent crude oil and gas traded on ICE[6] from April, 19. 1991 to December, 31, 2008. There are evidences that the MIX-GARCH and the MRS-GARCH models outperform the other models and the MIX-GARCH-X model has the best performance in terms of the out-of-sample volatility forecasting across the all markets. Furthermore, the results of VaR performance indicate that the augment GARCH-X model is the most reliable model.

1.1.3. Structural models

In the case of structural models, the oil price movement is function of a group of fundamental variables. The explanatory variables that commonly used to explain the oil price behavior are OPEC behavior, oil inventory level, oil consumption and production, and some non oil variables such as economic activity, interest rate, exchange rate, and other commodity prices. In this context, there are many studies that investigate oil prices based on fundamental variables and some of them explain the price movement fairly well; however, it does not mean that they show good forecasting performance, as there is limitation on availability for future values of the explanatory variables; therefore, due to the difficulties and complexities of structural models there is a small number of studies that performed structural analyses in order to forecast oil prices. Following, we categorize the structural models that are used to forecast oil prices based on five different models: *a)* OPEC behavior models, *b)* inventory models, c) combination of inventory and OPEC behavior models, d) supply and demand models, and e) non-oil models.

a) OPEC behavior models

According to the Huntington (1994), structural forecasting models based on supply and demand were not successful to predict oil prices in 1990s due to the two major errors. The first error was inaccurate forecast of GDP, especially for developing countries, and incorrect prediction of

[6] IntercontinentalExchange

increase in the supply of oil by non-OPEC countries. In addition to these, Tang and Hammoudeh (2002) state that another source of error was the omission of market participants' expectation on OPEC's interventions.

Tang and Hammoudeh (2002) perform an empirical investigate on OPEC attempts to control prices within a target zone model during 1988-1999. They employ the basic target zone model proposed by Krugman (1991) which have been applied in oil market by Hmammoudeh and Madan (1995). They use monthly spot the average basket price of seven types of OPEC members' crude oil products. During 1988-1999 still OPEC was not officially following a target zone policy, as the first price band was announced by OPEC in March, 2000. However, the authors state that there are evidences that OPEC supported target zone model during 1988-1999. They explain that, as oil is the major source of income for almost all OPEC member nations; hence, OPEC supports a lower limit for price; on the other side, high oil prices encourage investment by non-OPEC nations and reduce OPEC's market share; therefore, OPEC has a strong reason to put an upper limit on the price. According to this, they establish an oil price model based on production quotas, inventory levels and an expectation term, which is the expected rate of change market price based on information available at the current time. The out-of-sample forecasting results suggest that basic target zone model offers good forecasting ability, the model performs well when the price is approaching upper or lower band without any market jump and it shows a big forecasting error when the price is inside or outside of the band, this means that the model performs poor if there is a jump in the market.

b)Inventory models

Ye et al. (2002) declare a completely opposite condition with the target zone model that Tang and Hammoudeh (2002) described for OPEC during 1990s. The authors state that during 1991-2001, OPEC did a little attempt to adjust production in order to accommodate changes in demand, and if the action was taken, it was not sufficient; therefore, during this period prices showed a large volatility. In this study the authors focus on the position of OECD petroleum inventory level (crude oil and oil products) to forecast oil prices. They perform a short run forecast on nominal WTI monthly spot prices. In this model WTI spot price is function of OECD relative petroleum inventory level, which is deviation of actual inventories from the normal inventory level (they calculate normal inventory level by de-seasonalizing and de-trending historical data), the lower than normal OECD inventory level, which capture the asymmetric price changes in response to changes in inventory when the inventory level is below the normal level that when the inventory level is above normal, and the annual differences in monthly inventory. They use data from January 1992 to February 2001. The in-sample forecasting indicates that the model shows good forecasting performance. Ye et al. (2005) modify the study of Ye et al. (2002). They predict short term one month ahead nominal WTI crude oil spot price by assessment the impact of relative inventory level. In this model, the only explanatory variables is OECD industrial relative petroleum inventory level; moreover, 11 September 2001 terrorist attack and OPEC quota tightening at 1999 are dummy variables of the model. The authors exclude the lower than normal OECD inventory level variable from their new model as this variable increase the out-of-sample forecast error. They use monthly data from January 1992 to April 2003. They compare the results from the above relative stock model to the two benchmarks forecasting models: naïve autoregressive forecasting model, and modified alternative model. The in and out-of-sample evaluation criteria indicates that the relative stock model shows the best forecasting performance and the naïve model shows the poorest one. Ye et al. (2006) extend the work by Ye et al. (2005) suggest a nonlinear model to forecast monthly nominal WTI crude oil spot prices. They declare that, the short run crude oil prices are expected to behave differently when the inventory level nears its lower band than when it varies around its mid-range, and this happens because inventory has a zero lower bound or some minimum operating inventory requirement. In this model inventory is split into the low and high inventory levels and price is function of relative OECD industrial crude oil inventory level and nonlinear low and high inventory variables; moreover, 11 September 2001 terrorist attack and OPEC quota tightening at 1999 are dummy variables of the model. They use monthly data from January 1992 to October 2003 and they found that, the low inventory level variables are more significant than the high inventory level variables, this was expected because of the psychological effect of the low inventories, which leads to an asymmetric response: the price response to the low inventory is bigger that when the inventory level is high. The results indicate that the forecasting power of the new nonlinear model outperforms previous simple linear model by Ye et al. (2005) in terms of the in and-out-of sample forecasting performance, especially when the inventory level is very low or very high. Merino and Ortiz (2005) extend the inventory model that was proposed by Ye et al. (2005). They use monthly data from January 1992 to June 2004; in the first step, the authors forecast oil price with using the initial inventory model proposed by Ye et al. (2005) and they obtain the price premium of this model, which is a deviation of estimated price from actual price. In this mode, relative OECD petroleum industrial inventory level is the only explanatory variable of the model. In the next step, the authors attempt to explain the price premium through testing the Granger causality between a group of variables with price premium and investigate the systematic information that each new variable can add to the original inventory model, these new variables contain a wide range of variables including the oil market variables and financial and commodity prices. The oil market variables contain backwardation (the difference between actual prices and future prices), speculation, OPEC spare capacity, the US gasoline relative inventory level, open interest, and the US refinery capacity. Among non oil variables, they choose the US interest rate, US dollar/Euro exchange rate, the US dollar exchange rate versus, spread, and non-energy commodity prices. They perform the Granger causality test among each of the mentioned variables with price premium during three spans, 1992-2004, 1996-2004 and 1999-2004. The results indicate that, there are Granger causality from speculation, OPEC spare capacity and the US gasoline relative inventory level to price premium; however, there is not causality from any of non oil variables to price premium. Consequently, in the last step, they estimate the three extended models to forecast crude oil prices; in model A, relative OECD petroleum industrial inventory level and speculation are explanatory variables; in model B, relative OECD petroleum industrial inventory level and OPEC spare capacity are explanatory variables; and in model C, relative OECD petroleum industrial inventory level and the US gasoline relative inventory level are explanatory variables of crude oil price. They find that only speculation and oil prices are cointegrated or there is long run relationship among them; therefore, the only variable that adds systematic information to the model is speculation. As the result, they forecast the price with using model A, and compare its forecasting power with the initial inventory model of Ye et al. (2005). The result indicates that, models generate the same forecasts during 1992-2001; however, during 2001-2004 the extended model generates a better forecast, the only exception is during 2000-2001 that the extended model presents a worsen forecast than the basic model.

c) Combination of inventory and OPEC behavior models

Kaufmann (1995) proposes a Project Link model to describe world oil market during 1954 to 1989. He investigates the effect of economic, geological, and political events on oil prices. In this model, world oil price is function of market condition and strategic behavior of OPEC. The key factors are OPEC and non-OPEC capacity utilization, OPEC capacity, OPEC share from the world oil production, and OECD inventory level; moreover, OPEC quota and 1974 oil shock are included as dummy variable. The results indicate that the model has good power to describe world oil market. Kaufmann et al. (2004) investigate the impact of OPEC behavior on real oil prices. The authors examine Granger causality relationship between OPEC capacity utilization, OPEC quotas, OPEC members cheating from quotas, and days of forward consumption of OECD crude oil stocks (days) that is calculated by dividing OECD crude oil stocks by OECD crude oil demand. Furthermore, they incorporate Persian Gulf War and seasonal dummies into the model. They use quarterly data from 1986 to 2000, and perform cointegration test between variables that confirms the existence of the long run relationship between real oil price and variables of the model; subsequently, the Granger causality test by vector error correction model (VECM) indicates that there are evidence of Granger cause from OPEC behavior variables to real oil prices but not vice versa. Dees et al. (2007) examine the forecasting ability of the Kaufmann et al. (2004) model. The static and dynamic forecasting results from 1995-2000 display that forecasting performance of the model is fairly well and it is more relative to the time period and the volatility in real oil price caused by the exogenous shocks in that special time. The model performs well for the in-sample forecasting; however, it shows weak performance of the out-of-sample forecasting (2004-2006). This bias indicates that the model is suffering from the omitted variables that are responsible for the increase in oil prices among 2004-2006. In order to develop this model and solve the problem of omitted variable bias, Kaufmann et al. (2008) extend the work by Dees et al. (2007) with including the US refinery utilization rate, non-linearity in supply condition and expectations to the existing model; however, they eliminate OPEC quotas from the model and fold cheating on OPEC quotas to the capacity utilization variable. The authors explain that expectation is an expectation about the supply/demand balance which reflects in the futures market. They use a compiling observation on the price of the near month contract and the four month contract for WTI as a proxy for the expectations. Moreover, they update quarterly data from 1986-2006. The results indicate that OECD stocks, OPEC capacity utilization rates, the US refinery utilization rates and price expectations, Granger cause real oil prices very rapidly; and the one step ahead out-of-sample forecasting results show that forecasting power of the model is well and is able to account for much of the 27$ rise in crude oil price among 2004-2006. Finally the authors perform a forecast with the time series random walk model and a forecast based on future contracts as benchmarks; the results suggest that the structural econometric model produces more accurate forecasts than random walk or future markets. Chevillon and Christine (2009) assess the impact of physical market on the clearing price. In this study, the authors investigate determinant factors of real Brent crude oil spot price. They apply quarterly data from 1988-2005. In this model price is function of six explanatory variables, including OECD and non-OECD demands, OPEC quotas, OECD and non-OECD stocks, and OPEC implicit target for real price; moreover, the first and second Iraq Wars, terrorist attack of September 11, 2001 and Afghan War are included as dummy variables of the model. To the best of our knowledge this study is the first one that incorporates non-OECD inventory level to their price forecasting model. They perform a VAR analysis and concluded that worries alien to the physical market caused to increase in oil prices.

d)Supply and demand models

Yang et al. (2002) introduce a model to determine the variables that impact on the US oil price. Their model mainly focuses on OPEC production, real GDP of the US, and price and income elasticity of demand for oil in the US. They use monthly data from January 1975 to September 2000. The main purpose of this study is to investigate the world economy prosperity and recession effects on demand and consequently on oil prices. They apply GARCH model to describe volatility of oil price; then they perform a cointegration test and ECM model to investigate the short and long run relationships between oil demand with, oil price, real GDP, and natural gas and coal prices in order to find out the price and income elasticity of demand. In the next step they carry out a simulation of potential oil prices under different scenarios of OPEC production reduction, and conclude that in the case of OPEC production reduction, the oil price will increase but the magnitude and extension of this increscent depends on the harsh of recession and increase of domestic production by the US or other non-OPEC producers. Mirmirani and Li (2004) perform a structural analysis to predict crude oil prices. For this purpose they compare VAR and artificial neural network models to predict. They use monthly light sweet crude oil future prices data traded on NYMEX (lagged oil price), oil supply, petroleum consumption and money supply as explanatory variables of the model, covering the period from January 1980 to December 2002. The results prove that the ANN model performs better than VAR model.

e)Non-oil variables models

Lalonde et al. (2003) investigate the effects of the non-oil variables on real WTI crude oil spot prices. They apply quarterly data from 1974-2001. In this model, real WTI crude oil spot price is function of world output gap[7] and the real US dollar effective exchange rate gap[8]; moreover, three dummy variables are included to the model, including Iranian 1979 revolution followed by Iran-Iraq war in 1980, mid 1980s collapse of OPEC discipline, and oil price collapse at 1986 as the result of ceasing the role of swing producer by Saudi Arabia in late 1985. The result indicates that the real US dollar effective exchange rate gap is not a significant variable; therefore, they exclude this variable from the model and include petroleum inventory level. The out-of sample forecasting results show that the forecasting ability of this model outperforms the random walk and the autoregressive models benchmarks. However, the important point is that, the forecasting ability of the model excluding inventory level is worsen that the two benchmarks.

1.2. Non-standard Methods

Non-standard or computational methods are the nonlinear approaches which recently became popular in order to forecasting. Traditional approaches of forecasting such as time series methods, assume that time series under study are generated from a linear process which is totally inappropriate for a strongly nonlinear and chaos time series such as oil prices.

The main computational tool in order to oil price forecasting is Artificial Neural Networks (ANNs). Recent studies on ANNs show that ANNs have power to pattern classification and pattern recognition capabilities. ANNs are inspired by human brain biological system and have the capability to learn and generalize experiences. Currently, ANNs are being used for a wide variety of tasks in many different fields of business, industry and science (Widrow et al., 1994 , Zhang et al., 1998). Moreover, very recently Support Vector Machine (SVM) became well

[7] Difference between the world actual output and output at full capacity.
[8] Difference between the actual exchange rate and its equilibrium level.

known in order to forecasting. SVM represents a novel neural network technique, has gained ground in classification, forecasting, and regression analysis (Venables and Riplay, 2002, Chang and Lin, 2005, Dong, Cao, and Lee, 2005, Fernandez, 2010). However, one major application of ANNs is forecasting (Sharda, 1994, Zhang et al., 1998); in the case of oil price forecasting, Kaboudan (2001) performs a short term oil price forecasting with applying two compumetric forecasting methods including Genetic Programming (GP) and ANN and compares their ability with the naïve random walk model. He uses monthly closing crude oil price data from January 1993 to December 1998. The results suggest that GP forecasting performance is superior to others and ANN shows the poorest accuracy. Moshiri and Foroutan (2004) forecast crude oil prices with comparing three methods of GARCH, ARIMA, and ANNs; using daily crude oil future prices traded on NYMEX from April, 4, 1983 to January, 13, 2003. The results suggest that the feed forward multilayer neural network (FNN) method performance is superior to those of ARIMA and GARCH time series models. Mirmirani and li (2004) perform ANN model based on Genetic Algorithm (GA), and then apply Neuro Genetic Optimizer (NGO) to train and select the optimal network architecture; then, the Back Propagation Neural Network (BPNN) is chosen as network architecture. They use monthly light sweet crude oil future prices data traded on NYMEX, lagged oil prices, oil supply, petroleum consumption and money supply as explanatory variables of the model, covering the period from January 1980 to December 2002. The results indicate that BPNN-GA model performs quit better than VAR model. Wang et al. (2005) integrate the ARIMA and ANN models to forecast WTI crude oil prices. They use monthly WTI crude oil spot prices from January 1970 to December 2003. They perform ARIMA approach to model the linear components of time series, and then integrate ARIMA results with a three layer Back Propagation Neural Network (BPNN). The results show that, integrated approach outperforms the single ARIMA model. Yu et al. (2007, 2008) apply a Multi Scale Neural Network (EMD-FNN-ALNN) model instead of a Single Scale Neural Network, which is based on Empirical Mode Decomposition (EMD) approach to forecast WTI and Brent crude oil prices. In this method, the original price series decomposed into the various intrinsic modes with different scales, then with using three layers Feed Forward Neural Networks (FNN) the internal correlation structures of different components is extracted, and finally some important subseries are selected as input to an Adaptive Linear Neural Network (ALNN) for prediction. They use daily WTI crude oil price data from January, 1, 1986 to September, 30, 2006. The results indicate that multi scale neural network performance is better than the single scale neural network; therefore, this method improves prediction ability of single scale neural network. Then Yu et al. (2008) compare the forecasting performance of the EMD-FNN-ALNN model with the single ARIMA, single FNN, integrated EMD-ARIMA-averaging, integrated EMD-ARIMA-ALNN, and integrated EMD-FNN-averaging models. The results show that in both cases of WTI and Brent crude oil prices the EMD-FNN-ALNN model outperforms the other models. Zhang et al. (2008) extend the study of Yu et al (2007, 2008) with applying Ensemble Empirical Mode Decomposition (EEMD), which is an improvement of EMD. They use three dataset of WTI oil prices; the first dataset is monthly data from January 1946 to May 2006, the second dataset is WTI monthly prices from July 2000 to May 2006, and the third dataset is WTI weekly prices from July 2000 to May 2006. This method adds white noise series to the original time series and separates the scales better; the results suggest that EEMD is a reasonable technique to predict crude oil prices. Zhang et al. (2009) apply an EMD-based method to investigate the impacts of irregular and extreme events on WTI and Brent crude oil prices. The authors select Persian Gulf War, and Iraq war in 2003 effects as case studies, and conclude that this multi scale method is a promising tool to analyze the impact of irregular effects on oil prices. Shambora and Rossiter (2007) perform a financial model to predict crude oil prices. For this purpose, they apply an ANN model and use crude oil price future contracts traded on NYMEX from April, 16, 1991 to December, 1, 1997. Furthermore, they compare the results with the buy and hold strategy, the simple moving average crossover model, and the random walk model. The Sharpe ratio of each model proves that the ANN model performance is better than the other models. However, the ANN results suggest that future oil prices are not efficient predictors of spot prices. Naseri and Garache (2007) apply a hybrid ANN model to predict monthly crude oil price; using genetic algorithm (GA) to train and select network architecture, and then select Feed Forward Neural Network (FNN) as network architecture, in second stage combine k-means technique to cluster the time series data. The authors compare their hybrid model performance to those of short term energy outlook model of (EIA), Kaboudan (2001), Wang et al. (2004), and Naseri and Esfahanian (2007). The results suggest that the new proposed approach outperforms all benchmarks, unless in one case that the Wang et al. (2004) model outperforms the proposed hybrid ANN model. Fan et al. (2008) propose a new method called Generalized Pattern Matching based on Genetic Algorithm (GPMGA), which can forecast future crude oil prices based on historical observations. They use daily Brent crude oil prices from May, 5, 1987 to July, 26, 2007, and daily WTI crude oil prices from January, 2, 1986 to, July, 7, 2005. Authors compare the forecasting ability of GPMGA approach to Elman network and traditional Pattern Modeling in Recognition System approach (PMRS), and conclude that the new GPMGA approach outperforms the other models in prediction of long memory time series, and suggest that this is a useful approach in terms of forecasting. Kulkarni and Haidar (2009) use a multilayer Feed Forward Neural Network (FNN) model to perform a short term crude oil price tendency forecasting; and moreover, investigate the efficiency of future prices on spot price. For this purpose they use daily WTI crude oil spot prices from 1996 to 2007, in addition they use data on 1, 2, 3, and 4 months WTI future contracts. They conclude that future prices provide new information about spot prices especially in the case of 1 and 2 months future contracts. Xie et al. (2006) apply Support Vector Machine (SVM) method to predict crude oil prices. They use monthly WTI spot prices from January 1970 to December 2003. The authors compare the results with the ARIMA and BPNN methods; results indicate that the SVM method does not necessarily perform better than the ARIMA and BPNN methods in all sub periods. Fernandez (2010) forecasts crude oil and natural gas spot prices based on the SVM and ANN techniques, also applied the ARIMA model as a benchmark and uses daily data from 1994 to 2005. The out-of-sample forecast shows that, in short time horizons the ARIMA model outperforms the ANN and SVM; but in long run horizons, the ANN and SVM outperform the ARIMA. Therefore, time horizon impacts on forecasting ability of a model as an important element. Furthermore, the linear combination of the ANN and SVM produces more accurate forecast than the single methods.

2. Qualitative models

Apart from fundamental economic variables such as OPEC behavior, inventory level, and oil production and consumption, many irregular factors impact on oil prices, including military and political factors, natural disasters, and speculations, which are so called qualitative factors. Knowledge-based approaches are introduced to model the infrequent and irregular events which can happen in future and effect on oil market. There are very few studies that employ qualitative approaches to forecast oil prices, for example, Abramson and Finniza (1991) apply belief networks based on Mont Carlo analyses to predict OPEC and WTI crude oil prices, which is a qualitative knowledge-based technique under classification of artificial intelligence. Abramson and Finniza (1995) extend the work by Abramson and Finniza (1991), suggest a probabilistic belief network model based on Mont Carlo analyses to produce probabilistic forecast of average annual oil prices. This method combines the qualitative variables with those of algebraic formulas, conditional probabilities, and econometric relations. Nelson et al. (1994) apply Delphi approach to forecast oil prices as a qualitative computational approach. In this study, the authors investigate the impact of various variables on oil prices, such as OPEC behavior, political climate, oil demand, and etc. Wang et al. (2004) propose a new hybrid system to predict oil prices with integrating the ANN approach which has a Back Propagation Neural Network (BPNN) structure, and Rule-Based Expert Systems (RES), with Web-Based Text Mining (WTM) techniques-BPNN-WTM-RES-using monthly WTI crude oil spot price from January 1970 to December 2002. A comparison among simple BPNN and new hybrid method indicates that in an out-of-sample forecasting, the hybrid method outperforms the individual BPNN in all sub periods. Wang et al. (2005) extend the work by Wang et al. (2004) with introducing a novel nonlinear integrated approach called TEI@I to predict WTI crude oil prices. As explained in previous section, in first step they perform an ARIMA model, and in second step they integrate the ARIMA with BPNN approach to model linearity and nonlinearities of time series. Here we explain that they investigate the effects of irregular and infrequent events on oil prices by applying the Web-based Text Mining (WTM) and the Rule-based Expert Systems (RES) techniques, and integrate the ARIMA-BPNN with the WTM-RES and create the ARIMA-BPNN-WTM-RES technique. The results prove that the out-of-sample forecasting performance of the TEI@I methodology outperforms the individual ARIMA and the ARIMA-BPNN approaches. Yu et al. (2005) propose a rough set Refined Text Mining (RSTM) approach as a new knowledge-based forecasting system for crude oil price tendency forecasting. This system is a combination of two modules; the first one applies the text mining technique to produce rough knowledge, and the second one applies the rough set theory as a knowledge refiner for the rough knowledge. They use monthly crude oil data from January 1970 to October 2004. The authors compare the out-of-sample forecasting ability of the RSTM approach to the random walk, linear regression model, ARIMA model, and Back Propagation Neural Network (BPNN) model. The hit ratio of each model indicates that the new RSTM approach performance is better than the other models, and the random walk model is the poorest one. Gori et al. (2007) analyze the evolution of oil price and consumption in the last 30 years to construct a relationship between them. They forecast future trends under three scenarios of the oil price, including parabolic, linear, and chaotic behavior. Next, in the first scenario oil price prediction is obtained by parabolic curve, in the second scenario oil price prediction is obtained through linear curve, and in the third scenario a fuzzy logic is used to predict oil prices. Gaffari and Zare (2009) propose a method based on Adaptive Neuro Fuzzy Inference Systems (ANFIS) to predict daily WTI crude oil spot price tendency movements. This method is a combination of ANN and fuzzy logic. The results prove that in more than 66 percent of daily variation prediction this technique predicts the sign of oil price truly, which is predicted truly in 46.67 percent by Morana (2001), 45.76 percent by Gori et al (2007), and 54.54 percent by Fan et al (2006).

4. Conclusion

In this study we performed a review on the existing literature about crude oil price forecasting. For this purpose we distinguished forecasting methods into the two main techniques of quantitative and qualitative techniques.

Quantitative methods are grouped into: 1. Econometrics methods, including (a) time series models (b) financial models and (c) structural models, and 2. Non-standard or computational approaches. These quantitative methods are applied to model the numerical determinants of oil prices. On other hand, qualitative methods include knowledge-based techniques such as Delphi, web-based text mining method, fuzzy logic and fuzzy expert systems, and belief networks which are employed in terms of investigating the impact of irregular and infrequent events on oil prices.

There are vast numbers of studies which forecast crude oil prices with employing the mentioned methods. To the best of our knowledge, in this study we investigated almost the whole studies that performed crude oil price forecasting and are available in peer reviewed journals. Among them, the most frequently used technique is the time series econometrics in order to forecast the volatility of oil prices, the second frequently used is the financial method, the third frequently used methods are based on structural models and non standard computational models, and the least used technique is the qualitative knowledge based method.

Reference

Abosedra, S., Baghestani, H. (2004). On the predictive accuracy of crude oil futures prices. Energy Policy 32, pp. 1389–1393.

Abosedra, S. (2005). Futures versus univariate forecast of crude oil prices. OPEC Review 29, pp 231–241.

Abramson, B., Finizza, A. (1991). Using belief networks to forecast oil prices. International Journal of Forecasting 7, pp 299–315.

Abramson, B., Finizza, A. (1995). Probabilistic forecasts from probabilistic models: a case study in the oil market. International Journal of Forecasting 11, pp 63–72.

Agnolucci, P. (2009). Volatility in crude oil futures: a comparison of the predictive ability of GARCH and implied volatility models. Energy Economics 31, pp 316–321.

Alexander, C., Lazar, E. (2006). Normal mixture GARCH (1, 1): applications to exchange rate modelling. Journal of Applied Econometrics 21, pp 307–336.

Alizadeh, A.H., Nomikos, N.K., Pouliasis, P.K. (2008). A Markov regime switching approach for hedging energy commodities. Journal of Banking and Finance 32, pp 1970–1983.

Aloui, C., Mabrouk, S. (2010). Value-at-risk estimations of energy commodities via longmemory, asymmetry and fat-tailed GARCH models. Energy Policy 38, pp 2326–2339.

Baillie, R.T., Bollerslev, T., Mikkelsen, H.O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 74, pp 3–30.

Barone-Adesi, G., Bourgoin, f., Giannopoulos, K. (1998). Don’t Look Back. *Risk* 11, pp 100–104.

Bollerslev, T., (1986). Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, pp 307–327.

Bopp, A. E., Lady, G.M. (1991). A Comparison of Petroleum Futures versus Spot Prices as Predictors of Prices in the Future. Energy Economics 13, pp 274–282.

Cabedo, J.D., Moya, I. (2003). Estimating oil price ‘Value at Risk’ using the historical simulation approach. Energy Economics 25, pp 239–253.

Cai, J. (1994). A Markov model of switching-regime ARCH. Journal of Business & Economic Statistics 12, pp 309–316.

Chang, C.C., Lin, C.J. (2005). LIBSVM: a library for support vector machines. http://www.csie.ntu.edu.tw/~cjlin.

Cheong, C.W. (2009). Modeling and forecasting crude oil markets using ARCH-type models. Energy Policy 37, pp 2346–2355.

Chernenko, S., Schwarz, K., Wright, J.H. (2004). The Information Content of Forward and Futures Prices: Market Expectations and the Price of Risk. FRB International Finance Discussion Paper 808.

Chevillon, G., Rifflart, C. (2009). Physical market determinants of the price of crude oil and the market premium. Energy Economics 31. pp 537–549.

Chin, M. D., Leblanch, M., Coibion, O. (2005). The Predictive Content of Energy Futures: An Update on Petroluem, Natural Gas, Heating Oil and Gasoline, NBER Working Paper 11033.

Coppola, A. (2008) Forecasting oil price movements: exploiting the information in the futures market. Journal of Futures Market 28, pp 34–56.

Cortazar, G., Schwartz, E.S. (2005). Implementing a Stochastic Model for Oil Futures Prices. Energy Economics 25, pp 215–238.

Costello, A., Asem, E., Gardner, E. (2008). Comparison of historically simulated VaR: evidence from oil prices. Energy Economics 30, pp 2154–2166.

Darbellay, G. A., Slama, M. (2000). Forecasting the short-term demand for electricity: Do neural networks stand a better chance?. International Journal of Forecasting, 16, pp 71– 83.

Davidson, J. (2004). Moment and memory properties of linear conditional heteroscedasticity models, and a new model. Journal of Business & Economic Statistics 22, pp 16–29.

Day, T.E., Lewis, C.M. (1993). Forecasting futures market volatility. The Journal of Derivatives 1, pp 33–50.

Dees, S., Karadeloglou, p., Kaufmann, R.K., Sanchez, M. (2007). Modelling the World Oil Market: Assessment of a Quarterly Econometric Model. Energy Policy 35, pp 178–191.

Diebold, F.X., Mariano, R.S. (1995). Comparing predictive accuracy. Journal of Business & Economic Statistics 13, pp 253–263.

Ding, Z., Granger, C.W.J., Engle, R.F. (1993). A long memory property of stock market returns and a new model. Journal of Empirical Finance 1, pp 83–106.

Dong, B., Cao, C., Lee, S.E. (2005). Applying support vector machines to predict building energy consumption in tropical region.” Energy and Buildings 37, pp 545-553.

Edmundo G., Luiz F.L.L., Edmundo A. (2010). Forecasting oil price trends using wavelets and hidden Markov models, Energy Economics 32. pp 1507–1519.

Fan, Y., Y-J., Zhang, Y.J, Tsai, H.T., Wei,Y.M. (2008). Estimating value at risk of crude oil price and its spillover effect using the GED-GARCH approach. Energy Economics 30, pp 3156-3171.

Fong, W.M., See, K.H. (2002). A Markov switching model of the conditional volatility of crude oil prices. Energy Economics 24, pp 71–95.

Frey, G., Manera, M., Markanday, A., Scarpa, A. (2009). Econometrics models for oil price forecasting : A critical survey. CESifo Forum, No.1.

Gaffari, A., Zare, S. (2009). A novel algorithm for prediction of crude oil price variation based on soft computing. Energy Economics 31, pp 531–536.

Giot P., Laurent S. (2003). Market risk in commodity markets: a VaR approach. Energy Economics 25, pp 435-457.

Gori, F., Ludovisi, D., Cerritelli, P.F. (2007). Forecast of oil price and consumption in the short term under three scenarios: parabolic, linear and chaotic behavior. Energy Economics 32, pp 1291–1296.

Green, S. L., Mork, K.M. (1991). Towards Efficiency in the rude Oil Market. Journal of Applied Econometrics 6, pp 45–66.

Gulen, S. G. (1998). Efficiency in the Crude Oil Futures Markets. Journal of Energy Finance and Development 3, pp 13–21.

Haas, M., Mittnik, S., Paollela, M.S. (2004a). Mixed normal conditional heteroscedasticity. Journal of Financial Econometrics 2, pp 211–250.

Haas, M., Mittnik, S., Paollela, M.S. (2004b). A new approach to Markov-switching GARCH models. Journal of Financial Econometrics 2, pp 493–530.

Hamilton, J.D., Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. Journal of Econometrics 64, pp 307-333.

Hammoudeh, S., Madan, V. (1995). Expectations, target zones and oil price dynamics. Journal of Policy Modeling 17, pp 597–613.

Hansen, P.R. (2005). A test for superior predictive ability. Journal of Business & Economic Statistics 23, pp 365–380.

Huang, D., Yu, B., Frank, J., Fabozzi, F., Fukushima, M. (2009). CAViaR-based forecast for oil price risk, Energy Economics 31, pp 511–518.

Hung, J.C., Lee, M.C., Liu, H.C. (2008). Estimation of value-at-risk for energy commodities via fat-tailed GARCH models. Energy Economics 30, pp 1173–1191.

Hung, J.C., Wang, Y.H., Chang, M.C., Shih, K.H., Kao, H.H. (2011). Minimum variance hedging with bivariate regime-switching model for WTI crude oil. Energy 36, pp 3050–3057.

Huntington, H. (1994). Oil price forecasting in the 1980s: what went wrong?. Energy Journal 15, pp 1–22.

Kaboudan, M.A. (2001). Compumetric forecasting of crude oil prices. The Proceedings of IEEE Congress on Evolutionary Computation, pp 283–287.

Kang, S.H., Kang, S.M., Yoon, S.M. (2009). Forecasting volatility of crude oil markets. Energy Economics 31, pp 119–125.

Kaufmann, R.K. (1995). A Model of the World Oil Market for Project LINK: Integrating Economics, Geology, and Politics. Economic Modelling 12, pp 165–178.

Kaufmann, R.K. (2004). Does OPEC Matter? An Econometric Analysis of Oil Prices. The Energy Journal 25, PP 67–91.

Kaufmann, R.K., Dees, S., Gasteuil, A., Mann, M. (2008). Oil prices: the role of refinery utilization, futures markets and non-linearities. Energy Economics 30, PP 2609–2622.

Krugman, P. (1991). Target zones and exchange rate dynamics. The Quarterly Journal of Economics 106, pp 669–682.

Lalonde, R., Zhu, Z., Demers, F. (2003). Forecasting and Analyzing World Commodity Prices, Bank of Canada, Working Paper, pp 2003–2024.

Lamoureux, C.G., Lastrapes, W.D. (1990). Persistence in variance, structural change, and the GARCH model. Journal of Business & Economic Statistics 8, pp 225–234.

Lanza, A., Manera, M., Giovannini, M. (2005). Modeling and forecasting cointegrated relationships among heavy oil and product prices. Energy 27, pp 831-848.

Lee, T. (1994). Spread and volatility in spot and forward exchange rates. Journal of International Money and Finance 13, pp 375–383.

Merino, A., Ortiz, A. (2005). Explaining the So-called 'Price Premium' in Oil Markets. OPEC Review 29, pp 133–152.

Mirmirani, S., Li, H.C. (2004). A comparison of VAR and neural networks with genetic algorithm in forecasting price of oil. Advances in Econometrics 19, pp 203–223.

Mohammadi, H., Su, L. (2010). International evidence on crude oil price dynamics: applications of ARIMA–GARCH models. Energy Economics 32, pp 1001–1008.

Moosa, I. A., Al-Loughani, N.E. (1994). Unbiasedness and Time Varying Risk Premia in the Crude Oil Futures Market. Energy Economics 16, PP 99–105.

Morana, C. (2001). A semiparametric approach to short-term oil price forecasting. Energy Economics 23, PP 325–338.

Moshiri, S., Foroutan,F. (2005). Forecasting nonlinear crude oil futures prices. The Energy Journal 27, pp 81-95.

Murat, A., Tokat, E. (2009). Forecasting Oil Price Movements with Crack Spread Futures. Energy Economics 31, pp 85–90.

Naseri, M.R., Gharacheh, E.A. (2007). A hybrid artificial intelligence approach to monthly forecasting of crude oil price time series, The Proceedings of the 10th International Conference on Engineering Applications of Neural Networks, CEUR-WS284, pp. 160–167.

Nelson, Y., Stoner, S., Gemis, G., Nix, H.D. Results of Delphi VIII survey of oil price forecasts, Energy report, California Energy Commission, 1994.

Nomikos, N.K., Pouliasis, P.K. (2011). Forecasting petroleum futures markets volatility: The role of regimes and market conditions. Energy Economics 33, pp 321–337.

Pindyck, R.S. (1999). The Long-run Evolution of Energy Prices. The Energy Journal 20, pp 1–27.

Radchenko, S. (2005). The Long-run Forecasting of Energy Prices Using the Model of Shifting Trend, University of North Carolina at Charlotte,Working Paper.

Sadeghi, M., Shavvalpour, S. (2006). Energy risk management and value at risk modeling. Energy Policy 34, pp 3367-3373.

Sadorsky, P. (2006). Modeling and forecasting petroleum futures volatility. Energy Economics 28, pp 467–488.

Samii, M.V. (1992). Oil Futures and Spot Markets. OPEC Review 4, pp 409–417.

Schwartz, E., Smith, J.E. (2000). Short-term Variations and Long-term Dynamics in Commodity Prices. Management Science 46, pp 893–911.

Serletis, A. (1991). Rational Expectations, Risk and Efficiency in Energy Futures Markets. Energy Economics 13, pp 111–115.

Shambora, W., Rossiter, R. (2007). Are there exploitable inefficiencies in the futures market for oil?. Energy Economics 29, pp 18–27.

Sharda, R. (1994). Neural networks for the MS/OR analyst: An application bibliography. Interfaces 24, pp 116–130.

Stock, J.H., Watson, M.W. (1993). A simple estimator of cointegrating vectors in higher order integrated systems. Econometrica 61, pp 783–820.

Tang, L., Hammoudeh, S. (2002). An empirical exploration of the world oil price under the target zone model. Energy Economics 24, pp 577–596.

Venables, W., Ripley, B. (2002). Modern Applied Statistics with S. Forth edition. Springer-Verlag New York, Inc.

Vo, M.T. (2009). Regime-switching stochastic volatility: evidence from the crude oil market. Energy Economics 31, pp 779–788.

Wang, S.Y., Yu, L., Lai, K.K. (2004). A novel hybrid AI system framework for crude oil price forecasting. Lecture Notes in Computer Science 3327, pp 233–242.

Wei, Y., Wang, Y., Huang, D. (2010). Forecasting crude oil market volatility: Further evidence using GARCH-class models. Energy Economics 32, pp 1477–1484.

Widrow, B., Rumelhart, D.E., Lehr, M.A. (1994). Neural networks: Applications in industry, business and science. Communications of the ACM 37, pp 93–105.

Xie, W., Yu, L., Xu, S., Wang, S. (2006). A new method for crude oil price forecasting based on support vector machines. International Conference on Computational Science (Part IV), pp 444–451.

Yang, C.W., Hwang, M.J., Huang, B.N., 2002. An analysis of factors affecting price volatility of the U.S. oil market. Energy Economics 24, 107–119.

Ye, M., Zyren, J., Shore, J. (2002). Forecasting Crude Oil Spot Price Using OECD Petroleum Inventory Levels. International Advances in Economic Research 8, pp 324–334.

Ye, M., Zyren, J., Shore, J. (2005), A Monthly Crude Oil Spot Price Forecasting Model Using Relative Inventories, International Journal of Forecasting 21, pp 491–501.

Ye, M., Zyren, J., Shore, J. (2006). Forecasting Short-run Crude Oil Price Using High and Low Inventory Variables. Energy Policy 34, pp 2736–2743.

Yousefi, S.,Weinreich, I., Reinarz, D. (2005).Wavelet-based prediction of oil prices. Chaos, Solitons and Fractals 25, pp 265–275.

Yu, L., Wang, S.Y., Lai, K.K. (2005). A rough-set-refined text mining approach for crude oil market tendency forecasting. International Journal of Knowledge and Systems Sciences 2, pp 33–46.

Yu, L., Lai, K.K., Wang, S.Y., He, K.J. (2007). Oil price forecasting with an EMD-based multiscale neural network learning paradigm. Lecture Notes in Computer Science 4489, pp 925–932.

Yu, L., Wang, S.Y., Lai, K.K. (2008). Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm. Energy Economics 30, pp 2623–2635.

Zeng, T, Swanson, N.R. (1998). Predective Evaluation of Econometric Forecasting Models in Commodity Future Markets. Studies in Nonlinear Dynamics and Econometrics 2, pp 159–177.