IMA Journal of Management Mathematics Page 1 of 25
doi:10.1093/imaman/dpz011

Forecasting Crude Oil Futures Prices Using Global Macroeconomic News Sentiment

ZRYAN A. SADIK [†], PARESH M. DATE
Department of Mathematics, Brunel University London, Uxbridge, London, UB8 3PH, UK.
[†]Corresponding author.Email: zryan.sadik@alumni.brunel.ac.uk

AND

GAUTAM MITRA
OptiRisk Systems, OptiRisk R&D House, One Oxford Road, Uxbridge, UB9 4DA, UK.

We propose a method of incorporating macroeconomic news into a predictive model for forecasting prices of crude oil futures contracts. Since these futures contracts are more liquid than the underlying commodity itself, accurate forecasting of their prices is of great value to multiple categories of market participants. We utilize the Kalman filtering framework for forecasting arbitrage-free (futures) prices, and assume that the volatility of oil (futures) price is influenced by macroeconomic news. The impact of quantified news sentiment on the price volatility is modelled through a parametrized, non-linear functional map. This approach is motivated by the successful use of a similar model structure in our earlier work, for predicting individual stock volatility using stock-specific news. We claim the proposed model structure for incorporating macroeconomic news together with historical (market) data is novel and improves the accuracy of price prediction quite significantly. We report results of extensive numerical experiments which justify our claim.

Keywords: Crude oil, Macroeconomic news sentiment, Kalman filter, Forecasting.

1. Introduction and background

Commodity price and its price fluctuation both play crucial roles in affecting the global economy, even though the impacts are irregular depending on different factors such as geographic regions, specific commodity sectors, time periods etc. In particular, crude oil is an important commodity and a necessary component for the economic development and growth for industrialized and developing countries. Crude oil spot and futures contract prices can significantly impact inflation, unemployment rate, trade, poverty, and other economic conditions in many countries. Crude Oil futures are exchange-traded contracts in which the contract is an agreement to buy or sell a specific quantity of crude oil at a predetermined price, on a future delivery date. The impact of crude oil price variation reaches a large number of goods and services. Modelling the variation in price of spot and futures prices of crude oil is important for a variety of market participants, from sovereign governments to futures traders. Central banks and private sector forecasters consider the crude oil price and its futures prices as key variables in generating macroeconomic scenarios. Airlines invest in oil futures in order to hedge their exposure to aviation fuel, as fuel represents a large proportion of their cost. Given the overall importance of crude oil to economy, tools to improve the quality of prediction of oil prices and oil futures prices are clearly desirable.

The aforementioned need for accurate forecasting of oil spot and futures prices have attracted a lot of academic research. The relation between futures prices and spot price has been the centre of attention for a large number of studies, and the literature is rich with several studies covering a range of aspects with respect to this relationship. As a result, the literature on using futures prices for forecasting spot prices has grown rapidly over the past few years. Examples include approaches in McCallum (2005) and Alquist (2010). The evidence on whether using risk adjusted futures prices are useful in prediction is mixed. For instance, Moosa and Al-Loughani (1994) found evidence of a risk premium in crude oil futures markets and conclude that futures prices are not efficient forecasters of future spot prices. On the other hand, Pagano (2009) found that the use of risk premium adjustment improves forecasting ability over long time horizons beyond six months. Apart from the models based on forecasting the futures prices, one can identify the following different types of oil spot price forecasting models in the literature.

- One can use spot price history directly to build a predictive model, e.g. based on a simple autoregressive moving average (ARMA) model structure. However, Reeve (2011) shows that futures-based forecasting typically outperforms such spot price based approaches.
- Another alternative for spot prediction is to use relevant macroeconomic variables other than the spot price in a regression framework, leading to vector autoregression (VAR) models; see e.g. Baumeister (2012) shows that VAR models show good short term forecasting ability.
- Finally, structural models including a detailed structure of supply side dynamics are discussed in Kilian (2009) and Kilian and Murphy (2011, 2012), among others.

Alquist et al (2013) provide a comprehensive overview of different oil spot price forecasting methods.

There is a large literature on evaluating the forecasting performance of futures markets. Particularly, after a substantial review of the extant literature on crude oil, we found that great research efforts have been expended in two areas: first is understanding the underlying mechanisms that determine the spot price and second is the development of many models suitable for forecasting the spot price. The idea of using oil futures prices to predict the spot price is based on the assumption that the futures prices reacts faster to the new information entering the market than the spot price. For example, Kumar (1992) presents evidence to support market efficiency and finds in favour of futures prices as unbiased forecasters of crude oil prices. He investigates whether the forecasts from using futures prices can be improved by incorporating information from other forecasting techniques. Brenner and Kroner (1995) suggests that the inconsistencies observed between futures and spot prices may be as the result of carrying costs rather than a failing of the efficient market hypothesis. Girma and Paulson (1999) discovered that there are risk-arbitrage opportunities in petroleum futures spreads. Avsar and Goss (2001) observe that inefficiencies are likely to be exacerbated in relatively young and shallow futures markets such as the electricity market, where forecast errors may indicate a market still coming to terms with the true market model. Moreover, Emery and Liu (2002) examine the relationship of futures prices between electricity and natural gas, and moreover, they report that there exist opportunities to profit from trading with futures spreads. According to Silvapulle and Moosa (1999) trading in the futures market has many advantages when compared to spot trading, such as low transaction cost, high liquidity, and low cash in up-front, among others. This makes it much more attractive for investors to react for new information than taking position in the spot market. There are several studies have examined the modelling and forecasting of crude oil futures prices in the recent years. For instance, Moshiri and Foroutan (2005) examined the chaos and non-linearity 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.

Since the development of electronic media, thousands of pieces of financial news are released on different platforms every day. There is information available in terms of news sentiment which has an impact on the asset prices. There is a strong, yet complex relationship between asset price movements and the news sentiment. Traders and other market participants digest news rapidly and update their asset positions accordingly. Numerous studies have attempted to examine whether modern natural language processing (NLP) techniques and news sentiment analysis can improve financial markets prediction. For example, Joseph *et al.* (2011) examined the ability of online ticker searches (e.g. XOM for Exxon Mobil) to forecast abnormal stock returns and trading volumes. They argued specifically that online ticker searches serve as a valid proxy for investor sentiment. Oh and Sheng (2011) have attempted to discover and evaluate the predictive power of stock micro blog sentiment on future stock price directional movements. They constructed a set of robust models based on sentiment analysis and data mining algorithms. Zhang *et al.* (2016) proposed and applied a price shocks forecasting framework, which simultaneously takes the influence of social network users and their opinions about stocks into consideration. Specifically, they develop a new method to estimate social attention to stocks by influence modeling and sentiment analysis. Day and Lee (2016) investigated the influence of using different financial resources to investment and how to improve the accuracy of forecasting through deep learning. They showed that various financial resources have significantly different effects to investors and their investments, while the accuracy of news categorization could be improved through deep learning. Heston and Sinha (2016) measured sentiment with a proprietary Thomson Reuters neural network. They found that daily news predicts stock returns for only one to two days, However, weekly news predicts stock returns for one quarter. Positive news stories increase stock returns quickly, but negative stories receive a long-delayed reaction. Ren *et al.* (2018) integrated sentiment analysis into a machine learning method based on support vector machine. They took the day-of-week effect into consideration and constructed sentiment indices. Chiong *et al.* (2018) proposed a sentiment analysis-based approach for financial market prediction using news disclosures. They built a predictive model, which uses a support vector machine (SVM), that incorporate the extracted sentiment-related features from financial news. Xiang *et al.* (2018) have provided a review that clarifies the scope of natural language based financial forecasting (NLFF) research by ordering and structuring techniques and applications from related work.

The emphasis of the work presented in this paper is oil prices, oil futures prices and their volatilities. These variables play a vital role in the global economy. In the literature, several studies have found that higher oil prices have an unfavourable impact on the global economy (see for example, Morana (2013), Timilsina (2015) and Archanskaia *et al.* (2012)). In order to make appropriate decisions about the direction of economic policy, it is important to accurately forecast future oil prices with effective models [Hsu *et al.* (2016)]. All the above approaches use price histories and/or numeric data on economic variables in the predictive model. In addition, there is information available in terms of global macroeconomic news which also has an impact on oil prices. The use of news or its proxy as an input to forecasting has been increasing in the last decade. Galati and Ho (2003) investigated to what extent daily movements in the euro/dollar rate were driven by news about the macroeconomic situation in the USA and the euro zone. Kim *et al.* (2004) investigated the impact of scheduled government announcements relating to six different macroeconomic variables on the risk and return of US bond, stock and foreign exchange markets. Arshanapalli *et al.* (2006) investigated the effects of macroeconomic news on time-varying volatility as well as time-varying covariance for the US stock and bond markets; they found that stocks and bonds have higher volatility on the day of macroeconomic announcements. Nikkinen *et al.*(2006) investigated how global stock markets are integrated with respect to the U.S. macroeconomic news announcements. Hess *et al.* (2008) investigated the impact of seventeen US macroeconomic announcements on two broad and representative commodity futures indices. John *et al.* (2012) examined the intensity, direction, and speed of impact of US macroeconomic news announcements on the return, volatility and trading volume of three important commodities gold, silver and copper futures. Macroeconomic news is an important subset of all financial news; analysts qualitatively digest and apply this in financial decision making, automatic analysis of macroeconomic announcements are now finding applications, in the asset classes of Fixed income, Foreign Exchange (FX) and Commodities. For example, Erlwein-Sayer (2017) has analysed the impact of macroeconomic news in predicting the yield spreads of the European sovereign bonds. Their preliminary findings confirm that this approach improves the analytic models for monitoring spreads.

In the present work, we utilize the Kalman filtering framework to build a model for forecasting arbitrage-free (futures) prices. Kalman filter was first proposed in Kalman (1960) as a recursive Bayesian estimator in engineering. Kalman filter is a method to reduce or eliminate the noise in experimental data. It has since been widely employed in applications outside engineering, especially in econometrics. For example, Schwartz (1997) , Manoliu and Tompaidis (2002) and Lautier and Galli (2004) have applied Kalman filter to forecast spot prices with futures prices. A multi-commodity implementation is presented in Cortazar *et al.* (2008), where the futures prices of different commodities are used simultaneously to forecast the commodity prices. Inclusion of jumps to the commodity price process and a subsequent used of a particle filter for inference on commodity prices is advocated in Aiyube *et al.* (2008). In Cortazar and Schwarz (2003), a three factor model for oil futures prices is suggested which departs from Bayesian viewpoint used in filtering and infers prices using a numerical (but simple) optimisation instead. Date and Ponomareva (2011) have provided a review of applications of filtering within finance.

In addition to the Kalman filter mentioned above, we propose to use quantified macroeconomic news sentiment to enhance the predictive power of the model. Crude oil prices are affected by many different macroeconomic events, including major industrial accidents, natural disasters, wars and political upheavals. We use news sentiment as an exogenous input which can change the volatility of the spot price. Specifically we used the global macroeconomic news sentiment applied to a broad dataset of the crude oil prices. We carried out empirical experiments for forecasting the futures contract prices of crude oil using the global macroeconomic news sentiment. In Islyaev (2014) a random parameter one factor model was proposed for commodity price modelling. We extend this approach by incorporating macroeconomic news sentiment. To authors' knowledge, forecasting the futures prices of crude oil using global macroeconomic news sentiment has not been reported in the literature before. We use a linear state space model with logarithm of the spot price as the latent variable and a vector of logarithm of the futures prices as the observed variable, where we look at multiple futures with different term maturities. We use a one factor model with a constant risk premium, a random mean and a seasonality adjustment in terms of an additive sinusoid. We use logarithm of the prices of twelve futures contracts as observed variables. In addition, we assume that the volatility of spot price depends on macroeconomic news (our model structure is described and justified later in section 3). The motivation for using a filtering-based framework stems from the fact that the futures market for commodities is more liquid than the spot market. Computational results of calibration and comparison for the models as well as the out-of-sample forecasting will be presented.

The relationship of this research with the existing works on forecasting crude oil futures prices is summarised in Table 1. Unlike the other existing works, our proposed model incorporates global macroeconomic news sentiment into a Kalman filter framework to forecast spot and futures prices. The algorithms proposed in this paper provide a very useful alternative to the existing methods for futures price modelling and forecasting. Note that our method may also be employed for forecasting the futures prices of other commodities such as metals, agricultural products and raw materials using relevant price data and using an appropriately selected macroeconomic news data (see section 2.2 for more details on news data selection).

The rest of the paper is organized as follows. In section 2 we explain the two streams of time series,

Research paper	Uses futures prices?	Uses Kalman filter?	Uses macroeconomic sentiment?	Spot prices forecasted?	Futures prices forecasted?
Manoliu and Tompaidis (2002)	Yes	Yes	No	Yes	Yes
Lautier and Galli (2004)	Yes	Yes	No	No	Yes
McCallum (2005)	Yes	No	No	Yes	No
Cortazar *et al.* (2008)	Yes	Yes	No	Yes	No
Alquist (2010)	Yes	No	No	Yes	No
Islyaev (2014)	Yes	Yes	No	Yes	Yes
This paper	Yes	Yes	Yes	Yes	Yes

Table 1. Comparison of this research with recent existing works. 'Uses futures prices?' is a shorthand for using the relevant prices as measurements.

namely, the market data and macroeconomic news metadata that are used for numerical experiments. In section 3 we describe the vector autoregressive (VAR) and the one factor models. We define the linear state space representation for the one factor model. We also present our proposed model namely macroeconomic news sentiment augmented model. In section 4 we describe the Kalman filter and the issues of model calibration, model fitting and the performance measures. In section 5 we analyse the computational results of the empirical investigation of the 12 futures contracts of crude oil comparing the performances of the VAR model and the one factor model with and without the macroeconomic news data using the chosen performance measures. In section 6 we set out the conclusion.

2. Data

We have conducted extensive numerical experiments for calibration of different models and comparison of their performance when it comes to forecasting crude oil spot and futures prices. Two different types of data has been used: market data about the prices of crude oil spot and crude oil futures and global macroeconomic news data.

2.1 *Market data*

The historical market data used to estimate and validate the parameters of the models (described later in section 3) consist of daily closing prices of WTI Crude Oil traded on New York Mercantile Exchange (NYMEX) from January 2, 2014 to December 21, 2016. The time series, obtained from Thomson Reuters, includes spot prices and 12 corresponding futures contracts which expire on various dates from June 20, 2017 to May 22, 2018. Table 2 presents some statistics for the daily spot prices of WTI Crude Oil. It is easy to observe that the spot prices are skewed to the right and are leptokurtic. These two observations may be explained by seasonal characteristics.

	Count	Mean	Std	Skewness	Excess Kurtosis	Min	25%	50%	75%	Max
Spot Prices	750	61.78	24.42	0.73	1.95	26.19	44.75	49.90	91.13	107.95

Table 2. Spot Price Statistics

Figure 1 shows both spot prices and futures prices of WTI Crude Oil. From January 2, 2014 to October 31, 2014, the futures prices are below spot prices while from November 3, 2014 to December 21 2016, the futures prices are above spot prices. In futures markets, these situations are called as *backwardation* and *contango*, respectively. Normally, the spot prices are lower than the corresponding

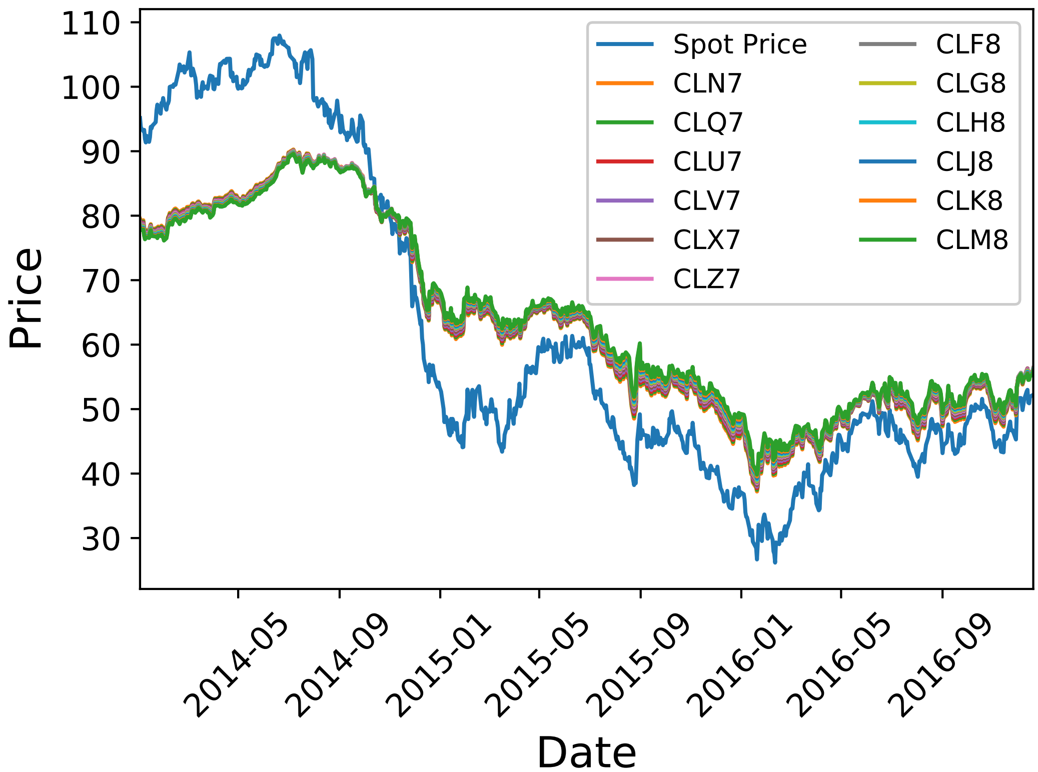

Figure 1. Spot and Futures Prices of WTI Crude Oil

futures prices. Here, the Crude Oil market experiences both backwardation and contango, which means that this market had experienced some unseen events during this time period.

Table 3 presents the corresponding relation between expiration dates and quotes for 12 futures contracts. For brevity, futures contracts including futures prices are represented with corresponding quotes.

Table 4 presents basic statistics of WTI Crude Oil futures prices. From this table, we can observe that futures prices data are slightly positively skewed and their excess kurtosis are all about 1.75, which is very similar to that for spot price data. This means that both spot price data and futures prices data have analogous statistical properties, and moreover there may be some seasonal patterned embedded in these data.

Currently, the futures contract with the farthest maturity date expires at December 2025. However, only the futures contracts with the latest 12 maturity dates are considered in this study. Having all the futures maturing beyond the last date of spot price data set also avoids the problem of very high volatility of futures price when it nears maturity.

	Quote	Expiration Date		Quote	Expiration Date
1	CLN7	2017.06.20	**7**	CLF8	2017.12.19
2	CLQ7	2017.07.20	**8**	CLG8	2018.01.22
3	CLU7	2017.08.22	**9**	CLH8	2018.02.20
4	CLV7	2017.09.20	**10**	CLJ8	2018.03.20
5	CLX7	2017.10.20	**11**	CLK8	2018.04.20
6	CLZ7	2017.11.20	**12**	CLM8	2018.05.22

Table 3. Corresponding Relation between Expiration Dates and Quotes

Quote		Count	Mean	Std	Skewness	kurtosis	Min	25%	50%	75%	Max
1	CLN7	750	63.11	15.24	0.39	1.74	37.22	50.51	60.82	79.78	90.27
2	CLQ7	750	63.2	15.11	0.39	1.74	37.47	50.69	61.06	79.79	90.16
3	CLU7	750	63.32	14.99	0.39	1.74	37.73	50.86	61.27	79.81	90.09
4	CLV7	750	63.46	14.89	0.38	1.74	37.97	51.05	61.45	79.75	90.06
5	CLX7	750	63.61	14.8	0.38	1.74	38.22	51.30	61.63	79.80	90.06
6	CLZ7	750	63.77	14.71	0.37	1.74	38.49	51.54	61.88	79.75	90.08
7	CLF8	750	63.81	14.61	0.37	1.74	38.66	51.66	61.96	79.66	89.98
8	CLG8	750	63.86	14.51	0.37	1.75	38.85	51.77	62.09	79.58	89.89
9	CLH8	750	63.92	14.41	0.37	1.75	39.06	51.91	62.27	79.47	89.81
10	CLJ8	750	64	14.31	0.37	1.75	39.30	52.04	62.45	79.41	89.73
11	CLK8	750	64.09	14.21	0.36	1.75	39.56	52.19	62.66	79.28	89.66
12	CLM8	750	64.19	14.11	0.36	1.75	39.83	52.36	62.87	79.21	89.60

Table 4. Futures Prices Statistics

2.2 *Macroeconomic news data*

2.2.1 Macroeconomic news analytic data

The RavenPack News Analytics delivers sentiment analysis and event data most likely to impact financial markets and trading around the world. The service includes analytics of more than 192,000 entities in over 130 countries and covers over 98% of the investable global market. All relevant news items about entities are classified and quantified according to their sentiment, relevance, topic, novelty, and market impact; the result is a data product that can be segmented into many distinct benchmarks and used in a variety of applications.

RavenPack classifies news items using multiple sophisticated sentiment detection algorithms. In addition, RavenPack generates a number of non-sentiment analytics including information about companies, events, relevance and market impact. Outputs are often in the form of numerical news scores that can be used as inputs in the calculation of company, sector and industry indicators.

The RavenPack News Analytics data is divided into two parts, the equity (company) related analytics and the global macro analytics. Each record in the Global Macro News Analytics database, which is used in this study, contains 34 fields including a times-tamp, entity ID, entitiy name, entity type, topic, group, type, scores for relevance, novelty and sentiment, and unique identifiers for each news story analysed. In the historical data files, each row in the database represents an entity-level record. Thus, whenever an entity such as a company or currency is mentioned in the news, RavenPack produces an entity-level record. A single news story can yield multiple records if more than one entity is mentioned.

For each entity RavenPack assigns a unique ID. The entities are classified into different types, currently RavenPack supports the following 9 entity types: Company, Organization, Currency, Commodity, Place, Nationality, People, Products and Sports Teams.

Relevant stories about entities are classified into a set of predefined event categories following the RavenPack taxonomy. There are over 2,000 types of event categories automatically detected by RavenPack. They have been divided into different groups, such that each group contains a collection of related categories. Furthermore, RavenPack classifies its data to six main taxonomy fields: Topic, Group, Type, Sub-type, Property and Category. More details about entity types and taxonomy fields can be found in the RavenPack user guide [RavenPack (2014)]. In the following subsection, we will briefly explain how the high volume of macroeconomic news data have been filtered in this research.

2.2.2 Choosing macroeconomic news

In this study, we have used macroeconomic news sentiment from a commercial data provider Ravenpack to enhance the predictive ability of the model, although the methodology is relevant for quantitative news data from other sources. First we have to decide which news item should be included into the macroeconomic news data. It is not practical to include all the available macroeconomic news items to generate positive and negative time series, which reflect the impact of macroeconomic news on the oil futures prices, one has to be careful in choosing the variables and eliminating the irrelevant news items from the macroeconomic data. The basic elements of how we filter the high volume data from Ravenpack is outlined below. The reader is referred to Ravenpack manual [RavenPack (2014)] for further information on how the data is structured.

- Ravenpack data might be categorized according to different entity types. We focus on entities related to crude oil price movements: company, organization, currency, commodity, people and product. Macroeconomic news items coming only from these entity types is considered as a part of our input.

- For the chosen entity types, the data may also be categorized by chosen events. Out of 51 event categories available, we choose the following as those related to crude oil price movement: "civil-unrest", "commodity-prices", "consumption", "domestic-product", "exploration", "foreign-exchange", "industrial-accidents", "interest-rates", "taxes", "transportation", "natural-disasters", "production", "products-services", "war-conflict".
- Finally, we restrict news items (for the above entity types and the above events) to the 15 biggest oil importing and exporting countries.

After choosing macroeconomic news data in this fashion, Ravenpack news analytics generates time-stamped news sentiment scores, novelty scores and relevance scores for each news item, for the specified time period. The quantity of interest for this study is news sentiment scores, which range from +1 (most positive news) to -1 (most negative news), with 0 as a neutral score. The novelty score ranges from 0 (least novel) to 100 (most novel) and indicates the novelty of the news item, whereas the relevance score indicates relevance and also ranges from 0 (least relevant) to 100 (most relevant). We restrict to news sentiment scores from those news items which have a novelty score of 80 or more and relevance score of 100. We also separate out positive news sentiment scores and negative news sentiment scores, as two separate time series. Finally, this gives us our time-stamped news sentiment score. This needs to be processed further, as any day might still have multiple macroeconomic news items at different frequencies and our price time series is on a day scale. The method to arrive at *macroeconomic news impact scores* from macroeconomic news sentiment scores is outlined in subsection 3.3.1.

3. Models for spot and futures prices

This section presents a detailed description of dynamic models used in this work. The main emphasis is on the one factor model as discussed in Manoliu & Tompaidis (2002). Three different models have been used in this study, namely vector autoregressive model, one factor model, and one factor model with microeconomic news. These three models are described in the following subsections.

3.1 *Vector autoregressive model*

The VAR model is one of the most successful, flexible, and easy to use models for the analysis of multivariate time series. It is a natural extension of the univariate autoregressive model for dynamic multivariate time series. Christopher Sims (1980) first proposed the VAR approach that reduced the restrictions needed by earlier econometric models and allowed the data to be modelled in an unrestricted form, where all variables in the model are considered as endogenous. The VAR model has proven to be especially useful for describing the dynamic behaviour of economic and financial time series and for forecasting. It often provides superior forecasts to those from univariate time series models and elaborate theory-based simultaneous equations models. Forecasts from VAR models are quite flexible because they can be made conditional on the potential future paths of specified variables in the model.

VAR models are used for multivariate time series. The structure is that each variable is a linear function of past lags of itself and past lags of the other variables. A VAR(p) model of the $m \times 1$ vector of time series $y_t = (y_{1t}, y_{2t}, \dots, y_{mt})$ with autoregressive order p is given by

$$y_t = Bx_t + \varepsilon_t, \tag{3.1}$$

$$x_t = \begin{bmatrix} 1 & y_{t-1}^\top & \cdots & y_{t-p}^\top \end{bmatrix}^\top, \tag{3.2}$$

and the dimensions of B and x_t are $(m \times (mp+1))$ and $((mp+1) \times 1)$, respectively. ε_t is a $m \times 1$ vector of disturbances that have mean zero and covariance matrix Σ_ε.

The VAR model has its own advantages and disadvantages. The first advantage is that the model is easy to estimate even with commercial software. Another advantage is that the VAR model treats all variables in the system symmetrically as endogenous, so there is no need to specify which variables are endogenous or exogenous. Furthermore, the VAR model can assist identification of shocks, including monetary policy shocks. Finally, VAR model has good forecasting capabilities. The main criticism of the VAR model is that the model has a very large number of parameters. This leads to problems in estimation such as over-fitting and inconsistent results with different estimation schemes. When it comes to forecasting futures prices on the same underlying asset, VAR model does not impose no-arbitrage condition on the forecasts, which is also seen as a limitation. This no arbitrage condition is automatically imposed if we consider all the futures prices driven by a single source of uncertainty affecting the spot price itself, as explained in the subsequent sections.

3.2 *One factor model*

To construct a one factor model for spot and futures prices of WTI Crude Oil, the log of spot price S_t, $x_t = \log(S_t)$, is assumed to be driven by a stochastic process in the filtered probability space (Ω , P, F_t). Here, Ω is the set that containing all the possible realisations of x_t, P is the objective probability measure and F_t is the natural filtration. For brevity, t_n, $(t_n : n = 0,\ldots,k, t_0 < t_1 < \cdots < t_k, \Delta := t_n - t_{n-1})$ is used to denote the discrete time.

Specifically, $x_t = \log(S_t)$ is assumed to follow the mean-reverting process (Ornstein-Uhlenbeck, OU):

$$dx_t = (\alpha - kx_t)dt + \sigma dW_t^P, \tag{3.3}$$

where κ is called the speed of mean reversion, α determines the long-run mean of x_t, and W_t is the Wiener process under objective probability measure P. Note that one factor model means that there is only one Wiener process in it.

Oil futures prices include risk premia like any other risky asset. This risk reflects the possibility that oil spot prices at the delivery time can be lower or higher than the contracted price. Therefore, the risk premium can be counted as a reward for holding a risky asset rather than a risk-free one. Based on this fact, equation (3.3) can be recomposed to another model that accounts for risk premium. The process x_t satisfies the following stochastic differential equation(SDE) if it is generated by the risk-neutral Wiener process W_t^R:

$$dx_t = (\tilde{\alpha} - kx_t)dt + \sigma dW_t^R. \tag{3.4}$$

These two drift items are associated with $\tilde{\alpha} \approx \alpha - \lambda_t \sigma$ for some processes λ_t i.e., $dW_t^R = dW_t^P + \lambda_t dt$: see, e.g. Duffie (1996) for the exact constraint on λ_t. For our purpose, we assume that λ_t is a constant, which is commonly assumed in the literature.

By applying Ito's lemma to the function $f(x_t,t) = e^{\kappa t}x_t$, it can be easily shown that $x_{t+\Delta}$ has the following mean and variance conditional on a past value x_t:

$$E^R(x_{t+\Delta}|F_t) = x_t e^{-k\Delta} + \frac{\tilde{\sigma}}{k}(1-e^{-k\Delta}), \tag{3.5}$$

$$Var^R(x_{t+\Delta}|F_t) = \frac{\sigma^2}{2k}(1-e^{-2k\Delta}). \tag{3.6}$$

Here, we follow the modification of the above model as suggested in Islyaev (2014) to make it more flexible at very modest increase in complexity. Specifically, we consider $\alpha \sim N(\mu_0, \theta^2)$ in equation (3.3). This allows the logarithm of spot price to converge to a random mean and potentially improves the forecasting ability of the model with only one parameter added based on the original model. We will investigate whether this added parameter improves price prediction later in the results section. The random mean is assumed to be uncorrelated with the Wiener process and has no risk premium. Hence, we can re-write the expressions for conditional mean and variance as:

$$E^R(x_{t+\Delta}|F_t) = x_t e^{-k\Delta} + \frac{\mu_0 - \lambda\sigma}{k}(1-e^{-k\Delta}), \tag{3.7}$$

$$Var^R(x_{t+\Delta}|F_t) = \frac{\theta^2}{k^2}(1-e^{-k\Delta})^2 + \frac{\sigma^2}{2k}(1-e^{-2k\Delta}). \tag{3.8}$$

Next, let $T = (T_i : i = 1,\ldots,m, 0 < T_1 < T_2 < \cdots < T_m)$ be the collection of the futures maturity dates. Then futures price for maturity T_i for a commodity with log-spot price x_t at time $t < T_i$ can be written as a conditional expectation of the commodity price at the maturity time of the futures contract: $F(t,T_i) = E^R(e^{x^i}|F_t), i = 0,\ldots,m$, where the expectation is taken under R measure and $x^i := x_{T_i}$, for brevity of notation. In case if $t > T_i$, $F(t,T_i) > 0$. The time to expiry of i^{th} futures contract is represented by $\Delta_t^i = T_i - t$. Since S_t is log-normally distributed, we have

$$F(t,T_i) = E^R(e^{x^i}|F_t) = e^{E^R(x^i|F_t)+\frac{1}{2}Var^R(x^i|F_t)}. \tag{3.9}$$

This allows us to derive an affine equation for vector of log-futures prices in terms of the log-spot price:

$$vec\{y_t^i\} = x_t e^{-k\Delta_t^i} + \frac{\mu_0 - \lambda\sigma}{k}(1-e^{-k\Delta_t^i}) + \frac{\sigma^2}{4k}(1-e^{-2k\Delta_t^i}) + \frac{\theta^2}{2k^2}(1-e^{-k\Delta_t^i})^2, \tag{3.10}$$

where $y_t^i = \log F(t,T_i)$ and the vec operator is defined by $vec(z_i) = [z_1 \quad z_2 \quad \ldots \quad z_n]^T$. Note that we do not model the convenience yield explicitly and assume that it is already reflected in the prices of futures contracts. This approach is followed in Islyaev (2014) and is consistent with the earlier framework followed in Manoliu & Tompaidis (2002). In contrast, convenience yield is explicitly modelled in Hyndman (2007).

Finally, since the futures prices of some commodities especially crop and energy commodities heavily depend on the weather conditions, it is worth taking some seasonal characters into consideration. Following Islyaev(2014), we consider a simple model for seasonality with a single sinusoid which is parametrized as follows:

$$f(t) = \exp(c_1 + c_2 \sin(c_3 t + c_4)), \tag{3.11}$$

where c_1 is a constant level, c_2, c_3 and c_4 are constants representing amplitude, the frequency and the phase of a seasonal pattern respectively. Accordingly, the prices of futures are modified as follows:

$$F(t,T_i) = f(T_i)E^Q(e^{x^i}|F_t), \tag{3.12}$$

and

$$vec\{y_t^i\} = \log(f(T_i)) + x_t e^{-k\Delta_t^i} + \frac{\mu_0 - \lambda\sigma}{k}(1 - e^{-k\Delta_t^i}) + \frac{\sigma^2}{4k}(1 - e^{-2k\Delta_t^i}) + \frac{\theta^2}{2k^2}(1 - e^{-k\Delta_t^i})^2, \quad (3.13)$$

which denotes a vector of log-futures prices, with i^{th} element of the vector denoting log-futures price for time to maturity Δ_t^i, as before. In practice, one may parametrise seasonality using multiple sinusoids. However, this complicates parameter estimation without necessary improving the quality of out of sample price forecasting.

For the one factor model described above, a linear state space representation will be used with a measurement equation based on the observed time series of futures prices and a discretized transition equation of the logarithm of spot commodity price, which is assumed to be latent. This allows us to use the Kalman filter to estimate the parameters by constructing and maximising a likelihood function, and to forecast the futures prices. The state space equations for one factor model that presented in subsection 3.2 can be written as

$$x_{n+1} = Bx_n + g + Rw_{n+1} \quad (3.14)$$
$$y_n = A_n x_n + d_n + Qz_n \quad (3.15)$$

where the state space model parameters may be expressed in terms of the original model parameters as:

$$f(t_n) = c_1 + c_2 \ \sin(c_3 t_n + c_4) \quad (3.16)$$

$$B = e^{-\kappa\Delta} \quad (3.17)$$
$$g = \frac{\mu_0}{\kappa}(1 - e^{-\kappa\Delta}) \quad (3.18)$$
$$R^2 = \frac{\sigma^2}{2\kappa}(1 - e^{-2\kappa\Delta}) + \frac{\Theta^2}{\kappa^2}(1 - e^{-\kappa\Delta})^2 \quad (3.19)$$

$$A_n = \begin{pmatrix} e^{-\kappa\Delta_n^1} \\ \vdots \\ e^{-\kappa\Delta_n^m} \end{pmatrix} \quad (3.20)$$

$$d_n = \begin{pmatrix} \frac{\mu_0 - \lambda\sigma}{\kappa}(1 - e^{-\kappa\Delta_n^1}) + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa\Delta_n^1}) + \frac{\Theta^2}{2\kappa^2}(1 - e^{-\kappa\Delta_n^1})^2 + f(T_1) \\ \vdots \\ \frac{\mu_0 - \lambda\sigma}{\kappa}(1 - e^{-\kappa\Delta_n^m}) + \frac{\sigma^2}{4\kappa}(1 - e^{-2\kappa\Delta_n^m}) + \frac{\Theta^2}{2\kappa^2}(1 - e^{-\kappa\Delta_n^m})^2 + f(T_m) \end{pmatrix} \quad (3.21)$$

here, $\Delta_n^i = \Delta_{t_n}^i$ for brevity of notation and m is the number of futures prices available at each t_n. $Q = cI_m$, where c is a scalar constant indicating the standard deviation of measurements and I_m is an $m \times m$ identity matrix. μ_0, λ, σ, Θ and κ are as defined in subsection 3.2 and c_1, c_2, c_3 and c_4 are constants.

3.3 *One factor model with macroeconomic news data*

Before proposing a method of incorporating macroeconomic news sentiment in a predictive model for forecasting prices of crude oil future contracts, in the following subsection we illustrate the process of generating two daily time series (positive and negative) that reflect the impact of macroeconomic news on the oil futures prices.

3.3.1 Generating macroeconomic impact scores

After extracting the macroeconomic news data that is related to the crude oil price movements in the first stage in subsection 2.2.2, we can select the macroeconomic news sentiment score for each news items based on their relevance and novelty scores. In this study, rather than using the macroeconomic news sentiment itself we are going to use the idea that was first proposed by Yu (2014) and used by Yu and Mitra (2016) and Sadik *et al.* (2018) to construct news impact scores. We construct macroeconomic news impact scores that can be used as proxies of global macroeconomic news impact in the new model. To achieve this, the RavenPack's macroeconomic news analytics database is used and some of its quantitative sentiment scores are employed. In the second stage, the following steps have to be done:

1. The time-stamp for each news item has to be converted from UTC to EST time, which is the timing convention of crude oil futures data from the New York Mercantile Exchange (NYMEX).

2. Then data has to be filtered such that choosing the macroeconomic news items that have relevance score of 100, which indicates how strongly related the entity is to the underlying news story, and novelty score of 80 or greater.

3. Separating the positive and negative sentiment scores so that two different time series can be obtained.

4. After separating the scores, in a similar fashion to Yu (2014), we calculate the positive and negative macroeconomic impact scores for each sentiment score.

5. Finally, we generate two daily time series that represent the positive and negative macroeconomic news impact scores.

In the last step one has to be careful in generating daily time series since we could get more than a macroeconomic news item in a day. Thus, we have to find a way to produce a single positive score and a single negative score for each day. In this study, in order to generate two daily time series that represent the positive and negative macroeconomic news impact scores, in which each day ends up with only one positive score as well as one negative score, we have attempted the following four different approaches:

(i) Aggregating all positive and negative macroeconomic news impact scores separately within a day to generate a single daily score for each time series (positive and negative).

(ii) Taking the average of the aggregated scores, which was produced by the first approach, to produce a single score for each time series for that day.

(iii) Giving weights to each macroeconomic news impact score based on how new is the news item that the sentiment score was driven from it.

(iv) The last approach is to give the highest weight to the last news item (here macroeconomic news impact score) in the day, in which a weight of 75% to the last macroeconomic news item and 25% to the rest of news items.

The following equations from (3.22) to (3.25) are representing the mathematical expressions for the aforementioned approaches, respectively.

$$\text{Aggregated} = \sum_{i=1}^{n} impact_i, \tag{3.22}$$

$$\text{Averaged} = \sum_{i=1}^{n} \frac{impact_i}{n}, \tag{3.23}$$

$$\text{Weighted} = \sum_{i=1}^{n} w_i * impact_i, \tag{3.24}$$

$$\text{Greater-Weight-To-Last-News} = 0.25 * \sum_{i=1}^{n-1} impact_i + 0.75 * impact_n, \tag{3.25}$$

where $\{impact_1, impact_2, \ldots, impact_n\}$ and $\{w_, w_2, \ldots, w_n\}$ are the macroeconomic news impact scores and their weights in the day, respectively.

Having generated the daily macroeconomic news impact scores (positive and negative) using the above approaches, we found that taking the averaged impact scores (equation 3.23) is the most appropriate approach to be used because it gives better out of sample performance in terms of the chosen error measures. Therefore, all the experiments reported in this study have used this approach to generate the two time series that represent the good and bad global macroeconomic news that impact the crude oil futures prices, and have been utilized as proxies of positive and negative macroeconomic news impact scores in the aforementioned models.

3.3.2 Macroeconomic news sentiment augmented model

In this study, we are interested in developing a forecast model for crude oil prices in which enables us to use the relative information content of the global macroeconomic news data in order to improve the predictive power of the one factor model.

To model the impact of the global macroeconomic news on the crude oil spot and futures prices, we define a scaling factor of the volatility, where it is determined by utilising the macroeconomic news sentiment score in the following way. The intuition behind this particular model structure is the same as in the earlier related work [Sadik *et al.* (2018)] and is presented below for the sake of completeness. The model structure needs to reflect the following economic realities: positive and negative news impact the volatility differently. Furthermore, positive news tends to reduce volatility whereas negative news tends to increase volatility. We also seek a functional specification which restricts the impact of news on volatility form below and above to realistic limits. Finally, we are seeking a parsimonious model with a small number of added parameters. Keeping these requirements in mind, we choose the following functional form (which was also employed in our earlier work): First, we define a function of two variables x and y:

$$f(x,y) = a + 0.5 * b\left(\frac{e^x - 1}{e^x + 1} - \frac{e^y - 1}{e^y + 1}\right), \tag{3.26}$$

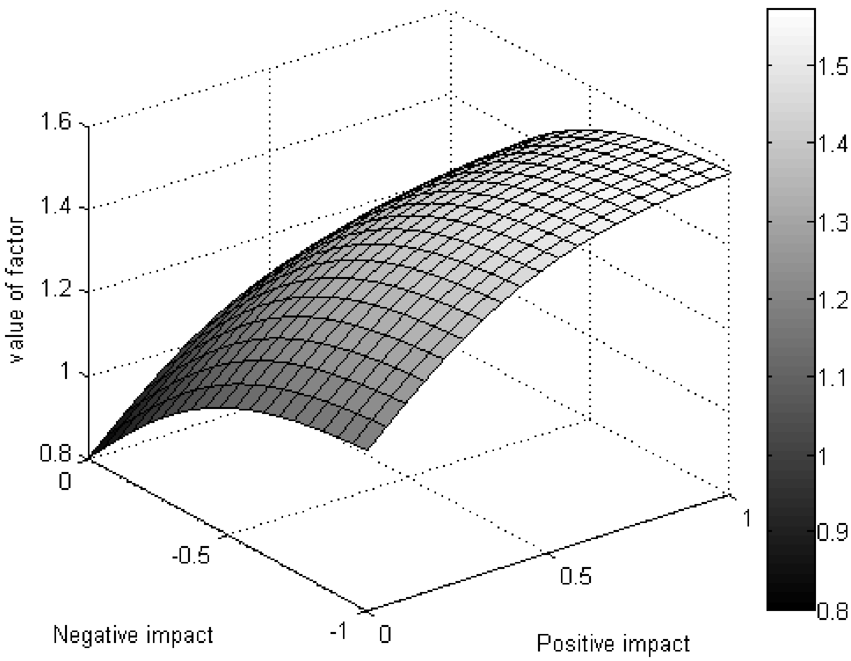

Figure 2. Plot of the scaling factor model against positive and negative impact scores

where a and b are constants. This function $f(x,y)$ lies between $(a, a+b)$ for any non-negative values of x and non-positive values of y.

Second, we create two different time series $\{\mathscr{P}_t\}$ and $\{\mathscr{N}_t\}$, as described in the previous subsection (3.3.1), that represent the positive and negative macroeconomic impact scores, respectively. Finally, we define the following function as a scaling factor of the volatility:

$$f(\mathscr{P}_t, \mathscr{N}_t) = a + 0.5 * b\left(\frac{e^{\rho \mathscr{P}_t} - 1}{e^{\rho \mathscr{P}_t} + 1} - \frac{e^{\gamma \mathscr{N}_t} - 1}{e^{\gamma \mathscr{N}_t} + 1}\right), \tag{3.27}$$

where 0.5 is a scaling factor of the function, and a, b, ρ and γ are parameters of the model. For example, if the parameters value are set as $a = 0.8$, $b = 0.8$, $\kappa = 4$ and $\gamma = 4$, the outcome will be as illustrated in figure 2. In this example, it can be clearly seen that the function reaches its highest value only when $\mathscr{P}_t = 1$ and $\mathscr{N}_t = -1$ (their highest scores). This can be interpreted in our model that the volatility increases when the value of news impact scores ($\mathscr{P}_t$ and $\mathscr{N}_t$) are increasing and vice versa. This is in keeping with the intuition behind our earlier work. For further details the reader is referred to earlier related work [Sadik *et al.* (2018)].

The new model structure of the one factor models is chosen to be one with a direct multiplicative effect of macroeconomic news on the volatility, as it is more natural to consider changes in percentage will change the volatility of price relative to its "news-neutral" level, when looking at the macroeconomic news impact. The positive and negative news may amplify or attenuate volatility level. Therefore we multiply the spot price volatility σ in equations (3.7) and (3.8) by the scaling factor in equation (3.27). In the model enhanced by macroeconomic news, equations (3.19) and (3.21) are replaced by the following two equations:

$$R^2 = \frac{(\sigma * f(\mathscr{P}_t, \mathscr{N}_t))^2}{2\kappa}(1-e^{-2\kappa\Delta}) + \frac{\Theta^2}{\kappa^2}(1-e^{-\kappa\Delta})^2 \tag{3.28}$$

$$d_n = \begin{pmatrix} \frac{\mu_0-\lambda(\sigma * f(\mathscr{P}_t,\mathscr{N}_t))}{\kappa}(1-e^{-\kappa\Delta_n^1}) + \frac{(\sigma * f(\mathscr{P}_t,\mathscr{N}_t))^2}{4\kappa}(1-e^{-2\kappa\Delta_n^1}) + \frac{\Theta^2}{2\kappa^2}(1-e^{-\kappa\Delta_n^1})^2 + f(T_1) \\ \vdots \\ \frac{\mu_0-\lambda(\sigma * f(\mathscr{P}_t,\mathscr{N}_t))}{\kappa}(1-e^{-\kappa\Delta_n^m}) + \frac{(\sigma * f(\mathscr{P}_t,\mathscr{N}_t))^2}{4\kappa}(1-e^{-2\kappa\Delta_n^m}) + \frac{\Theta^2}{2\kappa^2}(1-e^{-\kappa\Delta_n^m})^2 + f(T_m) \end{pmatrix} \tag{3.29}$$

In addition, we also need a constraint on the upper bound of the interval range that is specified by summing up the parameters $(a > 0)$ and $(b > 0)$ to keep the macroeconomic news impact related scaling in a reasonable range. As in our earlier work, the choice used here is $0 < a + b \leqslant 2$, *i.e.* the impact of macroeconomic news is assumed to change the volatility of the prices at most by a factor of 2.

The chosen model structure adds only four more model parameters and offers a reasonable compromise between increased model complexity and parsimony in terms of model parameters. The choice of model structure is essentially heuristic, and is justified through numerical experiments in the present work as well as in the earlier work on volatility prediction for stock prices, which has been reported in Sadik *et al.* (2018). One can also reduce the number of parameters by keeping the value of a and b fixed, for instance, $a = 0.5$ and $b = 1.5$. However, treating a and b as free parameters does improve results in terms of predictive power of the models.

3.4 *Summary forms for computational models*

The main part of our study, we have compared the performance of two computational models, the one factor model and the one factor model with macroeconomic news sentiment data. These two models are explained in the earlier sections and summarized below:

- **One factor model:**
  The model is given by equations (3.14)-(3.21).
- **One factor model with macroeconomic news sentiment data:**
  The model is given by equations (3.14)-(3.18), (3.20), (3.28) and (3.29).

In both the models, futures prices are treated as observable variables and the spot price is treated as a latent variable. The methodology to estimate the parameters of the models in both the cases is described in the next section. We have also compared the one factor model with VAR model described in subsection 3.1, and the results for this comparison are reported in the preliminary analysis subsection (5.1). The VAR model is given by equation (3.1), and its endogenous variables are crude oil spot price and twelve corresponding futures prices.

4. Methodology

In this study, the linear Kalman filter and the maximum likelihood estimation method are employed to estimate parameters of the three aforementioned models. The following subsections give a more detailed explanation for these two techniques.

4.1 *Kalman filter*

In 1960, R.E. Kalman published his famous paper describing a recursive solution to the discrete data linear filtering problem [Kalman (1960)]. The Kalman filter is a set of mathematical equations that can be used to estimate the state of the latent state of a process recursively using noisy observations, in a way that minimizes the trace of the covariance matrix of the estimation error. For the latent state equation (3.14) and the measurement equation (3.15), Kalman filter equations for estimating the mean and the covariance matrix of the latent variable can be summarised as follows:

Predict phase:

Predicted state estimate $$\hat{x}_{n+1|n} = B\hat{x}_{n|n-1} + g, \tag{4.1}$$

Predicted variance estimate $$P_{n+1|n} = BP_{n|n-1}B^T + RR^T - BP_{n|n-1}A_n^T S_n^{-1} A_n P_{n|n-1} B^T \tag{4.2}$$

Update phase:

Innovation or measurement residual $$V_n = y_n - (A_n\hat{x}_{n|n-1} + d_n) \tag{4.3}$$

Innovation (or residual) variance $$S_n = A_n P_{n|n-1} A_n^T + QQ^T \tag{4.4}$$

Optimal Kalman gain $$K_n = P_{n|n-1} A_n^T S_n^{-1} \tag{4.5}$$

Updated state estimate $$\hat{x}_{n|n} = \hat{x}_{n|n-1} + K_n V_n \tag{4.6}$$

Updated estimate variance $$P_{n|n} = (I - K_n A_n) P_{n|n-1} \tag{4.7}$$

$\hat{x}_{n|n-1}$ is the best estimated value of x_n at time n given observations up to time $\{n-1\}$ and $P_{n|n-1}$ is the corresponding conditional variance of $\hat{x}_{n|n-1}$ (a measure of the estimated accuracy of the state estimate). It is assumed that $\hat{x}_{0|-1}$ and $P_{0|-1}$ are known.

In the present context, Kalman filter allows us to predict the spot price x_n as a hidden state, which also gives an arbitrage-free prediction of the vector of futures prices (as $A_n\hat{x}_{n|n-1} + d_n$).

4.2 *Maximum likelihood estimation*

4.2.1 Maximum likelihood estimation for one factor model

For the given log-futures prices measurements $F = \{y_1, y_2, \ldots, y_n\}$ up to time t_n, we can apply the Kalman filter and maximum likelihood estimation algorithm to calibrate parameters of equations from(3.16) to (3.29). Now we can write joint likelihood function for F as

$$\hat{L}(F) = p(y_1) \prod_{i=2}^{n} p(y_i | F_{i-1}). \tag{4.8}$$

Since the measurements are jointly normal, the logarithm of the likelihood can be written as:

$$\log \hat{L}(F) = -\sum_{I=1}^{n} (\log |S_i| + V_i^T S_i^{-1} V_i), \tag{4.9}$$

where V_n and S_n are as defined in (4.3) and (4.4), respectively.

For a given vector-value time series $\{y_1, y_2, \dots, y_n\}$ and a vector of unknown model parameters θ, the optimisation problem can be stated as following:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}}(\log \hat{L}(F)). \tag{4.10}$$

We implement the maximisation using Matlab's in-built routine *fminsearch*, which relies on Nelder-Mead method.

4.2.2 Maximum likelihood estimation for VAR model

To derive the maximum likelihood estimation of VAR model, we first consider a VAR(p) model of the $m \times 1$ vector of time series as given in equation (3.1). Since the disturbances are assumed to normally distributed, the log likelihood in this case is:

$$l(B,\Sigma_\varepsilon) = -\left(\frac{Tm}{2}\right)\log(2\pi) + \left(\frac{T}{2}\right)\log|\Sigma_\varepsilon^{-1}| - (\frac{1}{2})\sum_{t=1}^{T}\left[(y_t - Bx_t)'\Sigma_\varepsilon^{-1}(y_t - Bx_t)\right]. \tag{4.11}$$

The value of B that maximise the log-likelihood happens to be the same as the OLS estimator:

$$\hat{B} = \left[\sum_{t=1}^{T} y_t x_t^{'}\right]\left[\sum_{t=1}^{T} x_t x_t^{'}\right]^{-1}. \tag{4.12}$$

The ML estimator for the variance is:

$$\hat{\Sigma_\varepsilon} = \left(\frac{1}{T}\right)\Sigma\hat{\varepsilon}_t\hat{\varepsilon}_t{}' \tag{4.13}$$

where

$$\hat{\varepsilon}_t \equiv y_t - \hat{B}x_t. \tag{4.14}$$

4.3 *Statistical performance measurements*

There are many statistical methods which can be used to observe the prediction accuracy of a model; for instance, Mean Absolute Error (MAE), Root Mean Square Errors (RMSE) and Mean Average Percentage Error (MAPE). For this study we have considered mean absolute error and root mean square errors as the criteria for prediction accuracy to assess the forecasting performance of the different models that have used in this study. The mean absolute error is given by

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|f_i - y_i| = \frac{1}{n}\sum_{i=1}^{n}|e_i|. \tag{4.15}$$

The root mean squared error is given by

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f_i - y_i)^2}. \tag{4.16}$$

In both equations (4.15) and (4.16), f_i is the predicted value and y_i is the observed value. Since the errors are squared before they are averaged, the RMSE gives a relatively high weight to large errors.

The AIC is an approach of selecting the best model, in terms of fit to the data, among other models. The selected model is the one which minimizes the Kullback-Leibler distance between the real data and the model [Akaike (1974)]. It is defined as:

$$\mathrm{AIC}(\hat{\Theta}) = -2(\log(likelihood)) + 2M \tag{4.17}$$

where $\hat{\Theta}$ is the set (vector) of model parameters, *likelihood* is the probability of the data given the candidate model and M is the number of estimated parameters in the candidate model (length of $\hat{\Theta}$).

The second order information criterion (AICc) takes into account the sample size by increasing the relative penalty for the model complexity with small datasets. It is defined as:

$$\mathrm{AICc}(\hat{\Theta}) = -2(\log(likelihood)) + 2M + \frac{2M(M+1)}{N-M-1} \tag{4.18}$$

where N is the sample size. When the number of observations is quite small compared to the number of parameters, such that N/M is less than 40, the use of the second-order corrected AIC (AICc) is recommended (Burnham and Anderson, 1998)..

5. Computational results

5.1 *A preliminary analysis*

In this study, the daily closing prices of twelve crude oil futures contracts has been used. For each future's contract, 500 data points are used as in sample data, which covers from 2 January 2014 to 31 December 2015, to estimate the models parameters. 250 data points are used as out of sample data, which covers from 2 January 2016 to 31 December 2016, to asses the forecasting performance of the models. From figure 1 we can see that during the few past years the crude oil market has experienced both backwardation and contango, both of which are totally different, meaning that this market had went through unpredictable events.

First, a thirteen-dimensional VAR model for the endogenous variables (crude oil spot price and twelve corresponding futures prices) has been utilized as alternative forecasting model to the one factor model. Having specified the model, the appropriate lag length of the VAR model has to be decided. In deciding the number of lags, the Akaike information criteria has been used, which is a common statistical method to use. Having specified the thirteen variables and the appropriate lag length of three, the VAR model has been estimated using the maximum likelihood estimation (MLE) method under the assumption of normality of the errors. For one factor model of the form (3.14)-(3.15), we employ Kalman filter and maximum likelihood to estimate the model parameters. The same data has been used for the one factor models that was used in the VAR model, such as using the crude oil spot price as a latent variable in the state equation (3.14) and a vector of its twelve corresponding futures prices as observed variable in the measurement equation (3.15) at each time step. After calibrating the two models, we compared them in terms of their forecast ability using the statistical measurement mentioned in subsection 4.3. Table 5 shows the average errors of the predicted values of the twelve futures prices. The results indicate that the one factor model has outperformed the VAR model by a factor of 5, on both the error criteria. Therefore, we decided to utilise the one factor model to construct a model that incorporate the global macroeconomic news data in its structure.

	Models	MAE	RMSE
1	VAR	26.18	27.22
2	One Factor	5.94	6.14

Table 5. Out of sample comparison of two forecasting models in terms of MAE and RMSE errors.

5.2 *Estimation results*

In this section, we present the empirical investigation of the estimated one factor model for the 12 contracts of oil futures prices. We analyse and compare the in-sample results obtained by the one factor model and macroeconomic news sentiment augmented model, which are described in section 3. As we mentioned earlier in subsection (2.2), we have used RavenPack's news analytics database and exploit some quantitative scores of its fields in our research.

After calculating the macroeconomic news impact scores for each sentiment score and generating two daily time series that represent the positive and negative macroeconomic news impact scores, we fit the one factor model to the observed time series of 12 futures prices via the Kalman filter and maximum likelihood method under the assumption of Gaussian error distribution. Thereafter we incorporated the macroeconomic news impact scores to the one factor model and recalibrate it again to re-estimate the model parameters including those which come from the volatility scaling factor in equation (3.27).

Table 6 shows the parameter estimates of the one factor model (with and without the macroeconomic data) for the 12 futures contracts of crude oil. We used the *fminsearch* function in MATLAB software to estimate the parameters. For the sake of brevity, we give names to the models that have been estimated in this study:

- Model 1: One factor model.
- Model 1-M: One factor model with Macroeconomic news sentiment data.

Common parameters

Models	μ	θ	κ	σ	Q	P	λ	c1	c2	c3	c4
Model 1	3.83E-01	2.35E-09	8.76E-02	1.30E-01	3.12E-03	1.71E+01	-1.35E-03	6.53E-03	-6.06E-04	1.01E-02	-4.52E-03
Model 1-M	3.87E-01	5.29E-03	9.19E-02	9.11E-02	2.83E-03	1.23E+01	3.91E-03	1.26E-02	-3.17E-02	-5.84E-03	3.70E-03

Macroeconomic parameters

Model	ρ	γ	a	b
Model 1				
Model 1-M	3.06E-02	3.85E-02	2.00E+00	2.11E-05

Table 6. Parameters estimation values of the models.

To compare the two models in terms of goodness of fit to the data, we calculated the Log-Likelihood values of the models and then applied the second order Akaike information criterion (AICc). The value of the AICc alone for a given time series is meaningless. Nevertheless, it becomes useful when it is compared to other AICc values of some other models, the model with the lowest AICc is described as the ”best” model among all other models specified for the dataset. After applying this powerful tool in comparing the one factor model with and without macroeconomic data, the results show that the one factor model with macroeconomic news sentiment data (Model 1-M) seem to be slightly better fit to the data. Table 7 presents the values of the log-likelihood and AICc for the estimated models.

Models	Log-Likelihood	AICc
Model 1	57173	-1.20E+05
Model 1-M	61458	-1.22E+05

Table 7. The log-Likelihood and AICc value of the estimated models

5.3 *Forecasting results*

This section presents the results of forecasting crude oil futures prices for twelve future contracts using the one factor model with and without macroeconomic data[1]. First we collected the prior information of global macroeconomic news data based on some filtering in the data to take the noise out. Then we created two time series that represent the positive and negative impact scores. To estimate the models we employed the Kalman filter and the maximum likelihood estimation method. The forecasting method in this study is dynamic, where the predicted future prices are used for lagged future prices instead of the real prices when forecasting the next period.

In this study, we specifically chose the futures contracts whose maturities range from 1 month to 12 months. This way, we cover short term, mid-term and long term futures prices. The prices of long term contracts have a much stronger dependence on macroeconomic news sentiment, as compared to futures with short term maturities. Incorporating a broad span maturities allows us to capture the market sentiment about the price evolution across the corresponding time spans.

As an alternative to the one factor model, the vector autoregressive (VAR) model was proposed and estimated in this study, in order to to see the predictive power of the models. As demonstrated earlier in section 5, the VAR estimation results show less forecasting accuracy than the one factor models. Thus, we carried out our experiments using the one factor model only. See Moon(1997) for further drawbacks of VAR based forecasting.

To asses the predictive power of the one factor model before and after incorporating the global macroeconomic news data, we compared the performance of the one factor model with its macroeconomic versions using some statistical measurements. The out-of-sample forecasting accuracy is mainly evaluated by two measures MAE and RMSE, which were defined earlier in subsection 4.3. The benchmarks in this study are the futures prices of crude oil. The MAE and the RMSE for each of the candidate one factor models (with and without incorporating macroeconomic data) are presented in table 8. From this tables, we can clearly see that in all cases the model with macroeconomic news data appear to outperform and give better prediction than the model which had not incorporated the macroeconomic news data in its structure. A notable result from the table is that the macroeconomic version of the one factor model have performed better for the long-term futures contracts than the short- and mid-term contracts but the overall improvement was clearly noticeable.

We start measuring how much information could be extracted from the table about the range and average of the errors for the futures contracts. For the one factor model, table 8, the MAE is ranged from 1.98 to 2.53 and the RMSE is ranged from 2.57 to 3.15, whereas the MAE and RMSE of its macroeconomic version are ranged from 1.8 to 2.07 and from 2.33 to 2.67, respectively. Both MAE and RMSE for each of the 12 contracts is better for the news augmented model as compared to the model

[1]We have investigated forecasting the crude oil spot price using the one factor model with and without macroeconomic news as well but we have not reported it. We have found that even though the one factor model with macroeconomic news performed better then the one factor model without macroeconomic news in terms of MAE and RMSE for the spot price but it was not significantly better.

which does not use news sentiment. Subsequently, we looked at the averages of the errors for the models and we observed that the MAE and RMSE for the one factor model averaged across all the 12 future contracts are 2.24 and 2.86, respectively. These error values are significantly lower after incorporating the macroeconomic news data to the model which results in 1.88 for MAE and 2.46 for RMSE.

Futures	**MAE**		**RMSE**	
	without	**with**	**without**	**with**
Futures 1	2.53	2.07	3.15	2.67
Futures 2	2.46	2.02	3.09	2.61
Futures 3	2.42	1.99	3.06	2.58
Futures 4	2.39	1.96	3.02	2.55
Futures 5	2.35	1.93	2.99	2.52
Futures 6	2.29	1.88	2.91	2.46
Futures 7	2.20	1.82	2.81	2.40
Futures 8	2.14	1.79	2.75	2.37
Futures 9	2.10	1.78	2.71	2.35
Futures 10	2.06	1.78	2.67	2.34
Futures 11	2.02	1.79	2.63	2.34
Futures 12	1.98	1.80	2.57	2.33

Table 8. MAE and RMSE for One Factor Model, with and without macroeconomic news data.

6. Summary of contributions and future research

The first and the most important contribution of this study is our model for processing macroeconomic news sentiment and its systematic incorporation into the crude oil spot price prediction model. We have demonstrated that this approach improves the price forecasting of crude oil futures. Our numerical experiments provide a strong evidence that macroeconomic news sentiment adds value to futures price forecasting. While most existing numerical studies only focus on short term futures contracts, our study includes the use of both short and long term futures contracts in model calibration as well as price forecasting.

The results of our study has widespread uses for public and private sector entities which rely on crude oil, and also for economic policy makers who need to incorporate future oil price behaviour into their economic scenarios. As an example, enhanced futures price forecasting may be used by a commodity futures trading firm as an aid for buy or sale decisions for futures, thus leading to better risk adjusted returns. Similarly, these price forecasts can feed into overall picture of the short term financial robustness of economy and can affect policy makers' decisions on setting interest rates, setting capital reserve ratios of financial institutions, scheduling required stress tests etc. While the numerical evidence provided is for oil futures, the proposed methodology is clearly applicable for other commodities where futures prices are more liquid that the underlying commodity price, and where macroeconomic news data is available. This includes other commodity futures such as metals, agricultural products and raw materials.

Acknowledgement

The Macroeconomic News Meta Data and the Market Data of crude Oil prices have been supplied by OptiRisk systems. These contributions are gratefully acknowledged. The authors also acknowledge fruitful conversations with research colleagues on Eurostar SenRisk project (http://www.senrisk.eu), whose findings on the use of analytic models with macroeconomic news sentiment for spread monitoring are consistent with those reported in this paper.

REFERENCES

AIYUBE, F., BAIDYA, T. & HUARSAYA, E. (2008). Analysis of commodity prices with the particle filter. *Energy Economics,* **30**, 597–605.

AKAIKE, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control,* **19**(6):716–723.

ALQUIST, R. & KILIAN, L. (2010), What do we learn from the price of crude oil futures?. *J. Appl. Econ.,* **25**: 539573. doi:10.1002/jae.1159

ALQUIST, RON, LUTZ KILIAN, & ROBERT J. VIGFUSSON (2013) Forecasting the price of oil. *In Handbook of Economics Forecasting (Graham Elliott and Allan Timmer-mann, eds.),* volume **2**, 427–507, North-Holland.

ARCHANSKAIA, E., CREEL, J. & HUBERT, P. (2012) The nature of oil shocks and the global economy. *Energy Policy 2012,* **42**, 509–520.

ARSHANAPALLI, B., D'OUVILLE, E., FABOZZI, F. & SWITZER, L. (2006) Macroeconomic news effects on conditional volatilities in the bond and stock markets, *Applied Financial Economics,* **16**(5), 377–384.

AVSAR, S.G., & B.A. GOSS, (2001) Forecast Errors and Efficiency in the U.S. Electricity Futures Market, *Australian Economic Papers,* Vol. **40**(4), pp. 479–499.

BAUMEISTER, C., & KILIAN, L. (2012) Real-Time Forecasts of the Real Price of Oil, *Journal of Business and Economic Statistics,* **30**, 326-336.

BRENNER, R.J., & F.K. KRONER, (1995) Arbitrage, Cointegration, and Testing the Unbiasedness Hypothesis in Financial Markets. *Journal of Financial and Quantitative Analysis,* Vol. **30**(1), pp. 23–42.

BURNHAM KP & ANDERSON DR (1998) Model Selection and Inference: A Practical Information-Theoretical Approach. 1998, Springer-Verlag, NY.

CHIONG, RAYMOND, ZONGWEN FAN, ZHONGYI HU, MARC TP ADAM, BERNHARD LUTZ, AND DIRK NEUMANN (2018) A sentiment analysis-based machine learning approach for financial market prediction via news disclosures. *In Proceedings of the Genetic and Evolutionary Computation Conference Companion,* pp. 278–279. ACM.

CORTAZAR, G. & SCHWARZ, E. (2003) Implementing a stochastic model for oil futures prices. *Energy Economics,* **25**, 215–238.

CORTAZAR, G., MILA, C. & SEVERINO, F. (2008). A multicommodity model of futures prices: Using futures prices of one commodity to estimate the stochastic process of another. *The Journal of Futures Markets,* **28**, 537–560.

DATE, P. & PONOMAREVA, K. (2011) Linear and non-linear filtering in mathematical finance: a review. *IMA Journal of Management Mathematics,* **22**, 195–211.

DAY, MIN-YUH, AND CHIA-CHOU LEE (2016) Deep learning for financial sentiment analysis on finance news providers. *In Advances in Social Networks Analysis and Mining (ASONAM), 2016 IEEE/ACM International Conference*, pp. 1127–1134.

DUFFIE, D. (1996) Dynamic Asset Pricing Theory. Princeton University Press.

EMERY, GARY W. & LIU, QINGFENG (WILSON) (2002). An analysis of the relationship between electricity and natural-gas futures prices, *Journal of Futures Markets,* **22**(2), John Wiley & Sons, Inc., 95–122.

ERLWEIN-SAYER, CHRISTINA (2017). Forecasting sovereign bond spreads with macroeconomic news sentiment. Working paper. Available online: http://senrisk.eu/wp-content/uploads/2018/10/WhitePaper_

ForecastingSovereignBondSpreads.pdf.

GALATI, G. & HO, C. (2003), Macroeconomic News and the Euro/Dollar Exchange Rate. *Economic Notes,* **32**: 371–398.

GIRMA, PAUL BERHANU & PAULSON, ALBERT S. (1999) Risk arbitrage opportunities in petroleum futures spreads, *Journal of Futures Markets*, **19**(8), John Wiley & Sons, Inc., 931–955.

HESS, D., HUANG, H. & NIESSEN, A. (2008) ow do commodity futures respond to macroeconomic news? *Financial Markets and Portfolio Management,* **22**(2), 127–146. https://doi.org/10.1007/s11408-008-0074-x

HESTON, STEVEN L., AND NITISH RANJAN SINHA (2017) News vs. sentiment: Predicting stock returns from news stories. *Financial Analysts Journal* **73**, no. 3, 67–83.

HSU, T.-K., TSAI, C.-C. & CHENG, K.-L. (2016) Forecast of 2013–2025 crude oil prices: Quadratic sine-curve trend model application. *Energy Sources Part B 2016,* **11**, 205–211.

HYNDMAN, C. (2007). Gaussian factor models - futures and forward prices. *IMA Journal of Management Mathematics,* **18**, 353–369.

ISLYAEV, SUREN (2014) Stochastic models with random parameters for financial markets. *Ph.D. Thesis,* Brunel University London, UK.

JOHN ELDER, HONG MIAO & SANJAY RAMCHANDER, (2012) Impact of macroeconomic news on metal futures, *In Journal of Banking & Finance,* Volume 36, Issue **1**, 2012, Pages 51-65, ISSN 0378-4266, https://doi.org/10.1016/j.jbankfin.2011.06.007.

JOSEPH, KISSAN, M. BABAJIDE WINTOKI, AND ZELIN ZHANG (2011) Forecasting abnormal stock returns and trading volume using investor sentiment: Evidence from online search. *International Journal of Forecasting* **27**, no. 4, 1116–1127.

KALMAN, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems, Transaction of the ASME-*Journal of Basic Engineering,* pp. 35–45 (March 1960).

KIM, SUK-JOONG; MCKENZIE, MICHAEL D. & FAFF, ROBERT W. (2004) Macroeconomic news announcements and the role of expectations: evidence for US bond, stock and foreign exchange markets, *In Journal of Multinational Financial Management,* Volume **14**, Issue 3, 2004, Pages 217-232, ISSN 1042-444X, https://doi.org/10.1016/j.mulfin.2003.02.001.

KUMAR, M.S. (1992). The Forecasting Accuracy of Crude Oil Futures Prices. *IMF Staff Papers,* Vol. **39**(2), pp. 432–461.

LAUTIER, D. & GALLI, A. (2004). Simple and extended Kalman filters: an application to term structures of commodity prices. *Applied Financial Economics,* **14**, 963–973.

LI, L & ENGLE, R. F. (1998). Macroeconomic announcements and volatility of treasury futures. Department of Economics, UCSD.

MANOLIU, M. & TOMPAIDIS, S. (2002). Energy futures prices: term structure models with Kalman filter estimation. *Applied Mathematical Finance,* **9**, 21–43.

MCCALLUM, ANDREW H. & WU, TAO, (2005) Do Oil Futures Prices Help Predict Future Oil Prices? (December 30, 2005). *FRBSF Economic Letter* No. **38**. Available at SSRN: https://ssrn.com/abstract=1967966 or http://dx.doi.org/10.2139/ssrn.1967966

MITRA, G. & MITRA, L. (2011) *The Handbook of News Analytics in Finance,* Wiley Finance.

MOON, K. S. (1997) A understanding of vector autoregressive model. *J. Korean Off. Stat. 1997,* **2**, 23–56.

MOOSA, I. A., & AL-LOUGHANI, N. E. (1994) Unbiasedness and Time-varying Risk Premia in the Crude Oil Futures Market, *Energy Economics,* Vol. **16**(2), pp. 99–105.

MORANA, K.(2013) The oil price-macroeconomy relationship since the mid-1980s: A global perspective. *Energy J. 2013,* **34**, 153–190.

MOSHIRI, S. & FOROUTAN, F. (2005) Forecasting nonlinear crude oil futures prices, *The Energy Journal* vol. **27**, pp. 81–95, 2005.

OH, CHONG, AND OLIVIA SHENG (2011) Investigating Predictive Power of Stock Micro Blog Sentiment in Forecasting Future Stock Price Directional Movement. *In ICIS,* pp. 1–19.

PAGANO, PATRIZIO & PISANI, MASSIMILIANO, (2009) Risk-Adjusted Forecasts of Oil Prices. *Bank of Italy Economic Research Paper* No. **585**; ECB Working Paper No. 999. Available at SSRN: https://ssrn.com/abstract=901951 or http://dx.doi.org/10.2139/ssrn.901951

RAVENPACK (2014) RavenPack News Analytics, Version 4.0: User Guide and Service Overview.

REEVE, TREVOR A. & VIGFUSSON, ROBERT, (2011) Evaluating the Forecasting Performance of Commodity Futures Prices (August 12, 2011). *FRB International Finance Discussion Paper* No. **1025**. Available at SSRN: https://ssrn.com/abstract=1912969orhttp://dx.doi.org/10.2139/ssrn.1912969

REN, RUI, DESHENG DASH WU, AND TIANXIANG LIU (2018) Forecasting Stock Market Movement Direction Using Sentiment Analysis and Support Vector Machine. *IEEE Systems Journal.*

SADIK, Z., DATE, P. & MITRA, G. (2018) News Augmented GARCH(1,1) Model for Volatility Prediction. *IMA Journal of Management Mathematics,* **30**(2), pp. 165–185.

SILVAPULLE, P. & MOSSA, A. (1999) The relation between spot and future prices: Evidence from the crude oil market, *The Journal of Futures Markets,* vol. **19**(2), pp. 175–193, 1999.

SCHWARTZ, E. (1997) The stochastic behavior of commodity prices: Implications for valuation and hedging. *Journal of finance New-York,* **52**, 923-974.

SIMS, C. A. (1980). Macroeconomics and Reality. *Econometrica* **48**(1), 1–48.

TIMILSINA, G.R. (2015) Oil prices and the global economy: A general equilibrium analysis. *Energy Econ. 2015,* **49**, 669–675.

XING, FRANK Z., ERIK CAMBRIA, AND ROY E. WELSCH (2018) Natural language based financial forecasting: a survey. *Artificial Intelligence Review,* **50**, no. 1: 49–73.

YU, X. (2014) Analysis of news sentiment and its applications to finance. *Ph.D. Thesis,* Brunel University London, UK.

YU, X. AND MITRA, G. (2016) An Impact Measure for News: its use in (daily) trading strategies, *The Handbook of Sentiment Analysis in Finance.* pp. 288–309.

ZHANG, LI, LIANG ZHANG, KELI XIAO, AND QI LIU (2016) Forecasting price shocks with social attention and sentiment analysis. *In Advances in Social Networks Analysis and Mining (ASONAM), 2016 IEEE/ACM International Conference,* pp. 559–566.