Oil price volatility forecast with mixture memory GARCH

Klein, T., & Walther, T. (2016). Oil price volatility forecast with mixture memory GARCH. *Energy Economics*, *58*, 46-58. https://doi.org/10.1016/j.eneco.2016.06.004

Published in:
Energy Economics

Document Version:
Peer reviewed version

Oil Price Volatility Forecast with Mixture Memory GARCH[☆]

Tony Klein[a,*], Thomas Walther[a]

[a] *Faculty of Business and Economics, Technische Universität Dresden, 01062 Dresden, Germany*

Abstract

We expand the literature of volatility and Value-at-Risk forecasting of oil price returns by comparing the recently proposed Mixture Memory GARCH (MMGARCH) model to other discrete volatility models (GARCH, RiskMetrics, EGARCH, APARCH, FIGARCH, HYGARCH, and FIAPARCH). We incorporate an Expectation-Maximization algorithm for parameter estimation of the MMGARCH and find different structures in volatility level as well as shock persistence. MMGARCH is also able to cover asymmetric and long memory effects. Furthermore, a dissimilar memory structure in variance of WTI and Brent crude oil prices is observed which is supported by additional tests. Parameter estimation and comparison of the models reveal significant long memory and asymmetry in oil price returns. In regard of variance forecasting and Value-at-Risk prediction, it is shown that MMGARCH outperforms the aforementioned models due to its dynamic approach in varying the volatility level and memory of the process. We find MMGARCH superior for application in risk management as a result of its flexibility in adjusting to variance shifts and shocks.

Keywords: GARCH-type models, long memory, asymmetry, mixture memory, oil price volatility, Value-at-Risk, volatility structure

1. Introduction

Forecasting oil price volatility is crucial for numerous industries. Research has revealed links between oil and other commodities, financial instruments, stocks, and

[☆] We are thankful for the comments of the editor Richard S. J. Tol and two anonymous reviewers which led to substantial improvement of this article. We thank Hermann Locarek-Junge, Stefan Huschens, Denise Erhardt, Christoph Lehmann, Rodrigo Lozano, Hien Pham Thu, Daniel Tillich, and the participants of the 5th INREC in Essen and the 10th Energy and Finance Conference in London for advice, remarks, and hints. We are especially grateful to Muyi Li and Guodong Li for discussing their model and providing useful information.

[*] Mail: tony.klein@tu-dresden.de, Phone: +49 351 463-34697.

bonds (e.g. Wang et al., 2014, Kang et al., 2014, Mohanty et al., 2014). Furthermore, variance predictions are of great importance to the energy and utility branch.

Typically, volatility of financial instruments and commodities in general, and of oil price returns in particular, are modeled by using Generalized Autoregressive Conditional Heteroskedasticity (GARCH) processes introduced by Engle (1982) and Bollerslev (1986). However, a complete model covering a variety of stylized facts has not been proposed so far. The so-called Long Memory property is one of these facts. Empirical results show that shocks have a much longer persistence in volatility than simple GARCH models are able to describe (see e.g. Baillie, 1996, Cont, 2001). GARCH models have a memory that declines exponentially and only an increased number of parameters provides a suitable fit. To implement this persistence of shocks, Fractional Integrated GARCH (FIGARCH) was proposed (Baillie et al., 1996). The memory process of FIGARCH models decays hyperbolically. While GARCH models have a short memory, FIGARCH models tend to depict a long memory, are still parsimonious with parameters, and show better performance in predicting future oil price volatility (Kang et al., 2009).

Researchers aim to find the most suitable and accurate model for decision making in regard to volatility forecasting. Some studies use long memory GARCH-type models (e.g. FIGARCH) to predict future oil price volatility (see e.g. Chkili et al., 2014, Wang & Wu, 2012) while other studies observe and identify different volatility regimes (Fong & See, 2002, Nomikos & Pouliasis, 2011, Chang, 2012). Motivated from recent literature, we assume that oil price volatility regimes do not only differ in level, but also in shock persistence. A model with varying memory structure would overcome drawbacks of fixed short or long memory. We expect this increased flexibility to provide better results in forecasting future oil price volatility.

We contribute to the literature by showing the presence of different memory types and confirm the existence of a high and a low volatility level in oil return series. Unlike others studies before, which utilize Markov-Regime-Switching models, we apply a mixture of two distinct GARCH models. Li et al. (2013) recently proposed the Mixture Memory GARCH (MMGARCH) which consists of a plain GARCH-part and a FIGARCH-part in volatility and provides a time-dependent and stochastic decision on whether to use short or long memory in each modeling step. With this work, we are the first to test the MMGARCH model against other GARCH-type models in predicting oil price volatility and its application for calculation of the Value-at-Risk.

Additionally to varying memory structures and volatility levels, we find structural differences between the oil blends WTI and Brent. Our results suggest that the persistence of shocks is almost infinite for WTI, while the impacts of shocks for Brent feature a shorter lifespan. This particular finding provides further evidence of a heterogeneous oil price market (Fattouh, 2010, Liu et al., 2013). Moreover, we can support the findings of studies observing asymmetric impact of good and bad news on volatility (Aloui & Mabrouk, 2010, Chkili et al., 2014).

The remainder is structured as follows: in Section 2 we introduce the models to be tested as well as estimation and forecasting methods. Section 3 is devoted to data acquisition and descriptive statistics of the utilized time series. Results are presented and discussed in Section 4. Section 5 concludes this work.

2. Methodology

2.1. Competitive GARCH models

The conditional variance h_t in a GARCH(p, q) model depends on the lagged squared residuals up to order q as well as on lagged conditional variances up to order p. Given that the conditional mean of the return series y_t is μ_t and the residuals are denoted by u_t, the GARCH$(1, 1)$ can be written as

$$\begin{aligned} y_t &= \mu_t + u_t, \\ u_t &= \varepsilon_t \sqrt{h_t}, \\ h_t = \operatorname{Var}(y_t|\mathcal{F}_{t-1}) &= \omega + \alpha u_{t-1}^2 + \beta h_{t-1}, \end{aligned}$$

where $\varepsilon_t \sim \mathcal{N}(0, 1)$ i.i.d for all $t = 1, \ldots, n$ and $\mathcal{F}_{t-1}$ denotes the sigma algebra generated by the history of the time series. We assume $\omega > 0$ and α, β to be greater or equal to zero to ensure the non-negativity of the variance process as well as $\alpha + \beta < 1$ to obtain a stationary process.

If it holds that $\alpha + \beta = 1$, every shock has an infinite persistence. The model is then called Integrated-GARCH (IGARCH) (Engle & Bollerslev, 1986). The special case with $\omega = 0, \alpha = 0.06$, and $\beta = 0.94$ is part of the "RiskMetrics" of JP Morgan. With these fixed RiskMetrics parameters it is straightforward to model the conditional variance without estimation.

Another class of GARCH models depict the stylized fact of asymmetric impact of positive and negative returns on volatility. Nelson (1991) proposed the Exponential

GARCH (EGARCH) which models the logarithm of h_t. For normally distributed u_t, the EGARCH$(1,1)$ can be written as (Engle & Ng, 1993)

$$\log\left(h_t\right) = \omega + \beta \log\left(h_{t-1}\right) + \gamma \frac{u_{t-1}}{\sqrt{h_{t-1}}} + \alpha \left(\frac{|u_{t-1}|}{\sqrt{h_{t-1}}} - \sqrt{2/\pi} \right).$$

Due to its nature, the EGARCH does not have any parameter restrictions. The parameter γ is often referred to as "leverage" parameter. A negative parameter γ indicates that negative returns have a larger impact on volatility than positive returns.

Another model to describe asymmetric effects is the Asymmetric Power ARCH by Ding et al. (1993). Furthermore, it generalizes the power of the modeled volatility. The APARCH(1,1) can be written as

$$h_t^{\delta/2} = \omega + \alpha \left(|u_{t-1}| - \gamma u_{t-1}\right)^{\delta} + \beta h_{t-1}^{\delta/2}, \tag{1}$$

with the restrictions $\omega \geq 0$, $\alpha, \beta, \delta > 0$, and $\gamma \in [-1, 1]$. For the APARCH and following FIAPARCH, a positive leverage parameter γ indicates that negative returns have a larger impact on the conditional variance than positive returns.

As mentioned before, GARCH models show only short memory when not highly parameterized. Therefore, models with fewer parameters to cover long memory are needed. FIGARCH(p, d, q) was proposed by Baillie et al. (1996) where d is the fractional integration parameter and p and q are analogues to the order of lag. A FIGARCH$(1, d, 1)$ variance process is defined as

$$h_t = \omega + \left(1 - \beta L - (1 - \varphi L)(1 - L)^d\right) u_t^2 + \beta h_{t-1},$$

where L denotes the lag-operator. We assume that $\omega > 0$, $0 \leq \beta \leq \varphi + d$, and $0 \leq d \leq 1 - 2\varphi$. The FIGARCH nests the aforementioned GARCH (if $d = 0$) as well as the IGARCH (if $d = 1$) with an infinite shock persistence.

Another long memory model we implement for comparison is the Hyperbolic GARCH (HYGARCH, Davidson, 2004). It generalizes the hyperbolic decay of memory by adding a weight b on the fractional difference to overcome the drawback that for any $0 < d < 1$ FIGARCH has no unconditional variance (Conrad, 2010). The HYGARCH$(1, d, 1)$ variance process can be formulated as

$$h_t = \omega + \left(1 - \frac{1 - \varphi L}{1 - \beta L}\left(1 + b\left((1 - L)^d - 1\right)\right)\right) u_t^2.$$

The Fractionally Integrated Asymmetric Power ARCH (FIAPARCH) proposed by Tse (1998) is a FIGARCH applied on APARCH innovations. The FIAPARCH(1,d,1) is given by

$$h_t^{\delta/2} = \omega + \left(1 - \beta L - (1 - \varphi L)(1 - L)^d\right)(|u_{t-1}| - \gamma u_{t-1})^\delta + \beta h_{t-1}^{\delta/2}$$

All parameter specifications of APARCH and FIGARCH have to hold for FIAPARCH as well.

2.2. Mixture Memory GARCH

Motivated by Li et al. (2013), we incorporate a GARCH$(1,1)$ and a FIGARCH$(1,d,0)$ component for variance modeling. Let $\varepsilon_t \sim \mathcal{N}(0,1)$ i.i.d. for all $t = 1, \ldots, n$. A random mixture between the GARCH and the FIGARCH component is incorporated by introducing a sequence of Bernoulli random variables, $(Z_t)_{t=1,\ldots,n}$, where $\mathbb{P}[Z_t = 1] = \pi_t$. The mixture proportion $\pi_t \in [0,1]$ is addressed at a later point. The Mixture Memory GARCH (MMGARCH) is then defined as

$$\begin{aligned} y_t &= \mu_t + u_t, \\ u_t &= \varepsilon_t \sqrt{h_t}, \\ h_t &= z_t h_{1,t} + (1 - z_t) h_{2,t}, \end{aligned} \tag{2}$$

where

$$\begin{aligned} h_{1,t} &= \omega_1 + \alpha_1 u_{t-1}^2 + \beta_1 h_{1,t-1}, \\ h_{2,t} &= \omega_2 + \left(1 - \beta_2 L - (1 - L)^d\right) u_t^2 + \beta_2 h_{2,t-1}, \end{aligned}$$

for $t = 1, \ldots n$, where $h_{1,t}$ denotes the GARCH, $h_{2,t}$ the FIGARCH component with the aforementioned assumptions applied, and z_t is a realization of the Bernoulli random variable Z_t at time t. Since $z_t \in \{0,1\}$, either the GARCH ($h_{1,t}$) or the FIGARCH ($h_{2,t}$) component is applied as instantaneous, conditional variance at time t. Li et al. (2013) show that under the given assumptions of non-negativity and strict stationarity for the respective components, there exists a strictly stationary solution to (2) with finite variance. In view of estimating model (2), we adopt the equivalent ARCH(∞) representation of the FIGARCH component h_2 as introduced in Bollerslev & Mikkelsen (1996):

$$
\begin{aligned}
h_{2,t} &= \omega_2 + \left((1-\beta_2 L) - (1-L)^d\right) u_t^2 + \beta_2 h_{2,t-1} \\
&= \frac{\omega_2}{1-\beta_2} + \sum_{j=1}^{\infty} \delta_j u_{t-j}^2,
\end{aligned}
\tag{3}
$$

where δ_i is calculated from the FIGARCH parameters and referred to as FIGARCH-weights. The infinite summation needs to be truncated at an index $l_T < \infty$, which is the so-called truncation lag. This truncation lag has to be handled with caution in regard to the hyperbolic decay of the FIGARCH-weights $(\delta_1, \ldots, \delta_{l_T})$ and thus, the long memory of the FIGARCH process. In literature, frequently proposed lags range from $l_T = 50$ to $l_T = 2\,000$. We set $l_T = 500$ for all long memory models.

A dynamic approach for the mixture proportion, that is dependent on a lagged realization of the time series of order one, is implemented. We adopt a logistic link function which was tested in recent research (e.g. Cheng et al., 2009). Hence, we define

$$
\begin{aligned}
\log\left(\frac{\pi_t}{1-\pi_t}\right) &= \lambda_0 + \lambda_1 \left(y_{t-1} - \mu_{t-1}\right) \\
&\Longleftrightarrow \pi_t := \frac{1}{1+\exp(-\lambda_0 - \lambda_1 \left(y_{t-1} - \mu_{t-1}\right))}.
\end{aligned}
\tag{4}
$$

While Li et al. (2013) implement the absolute (unsigned) value of the time series ($|y_{t-1} - \mu_{t-1}|$) for calculating the mixture proportion in Eq. (4), we find that incorporating the absolute value cancels some asymmetric effects; therefore, we employ the signed value. A positive parameter λ_1 indicates that negative deviations from the mean lead to a higher probability of FIGARCH being the instantaneous variance at time t.

2.3. EM Algorithm

From Li et al. (2013), we adopt the following pseudo-log-likelihood function for Model (2) with $\theta = (m, \omega_1, \alpha_1, \beta_1, \omega_2, \beta_2, d, \lambda_0, \lambda_1)$ and $t = 1, \ldots, n$:

$$
\begin{aligned}
LL(\theta) = \sum_{t=1}^{n} \log\left(\frac{\pi_t(\theta)}{\sqrt{h_{1,t}(\theta)}}\exp\left(-\frac{(y_t-\mu_t)^2}{2h_{1,t}(\theta)}\right) + \frac{(1-\pi_t(\theta))}{\sqrt{h_{2,t}(\theta)}}\exp\left(-\frac{(y_t-\mu_t)^2}{2h_{2,t}(\theta)}\right)\right) \\
- \frac{n}{2}\log(2\pi),
\end{aligned}
\tag{5}
$$

where m refers to all parameters associated with the structure of the conditional mean μ_t.

Let Ω denote the admissible parameter space originating from the parameter restrictions. The direct maximization of Eq. (5) to find the log-likelihood estimate

$$\hat{\theta} = \underset{\theta \in \Omega}{\arg\max} \, (LL(\theta)) \tag{6}$$

is not feasible. The actual decision on whether $h_{1,t}$ or $h_{2,t}$ is the instantaneous variance depends on π_t which is unobservable. Hence, the maximization problem (6) is a so-called missing data problem. We incorporate the iterative Expectation-Maximization (EM) algorithm which was introduced by Dempster et al. (1977) to replace the missing data with its conditional expectation. The initial, incomplete data log-likelihood function in Eq. (5) is replaced by the following complete data, substitute log-likelihood function:

$$\begin{aligned} LL^{\#}(\theta) = \sum_{t=1}^{n} \Big[& z_t \log\left(\pi_t(\theta)\right) + (1 - z_t) \log\left(1 - \pi_t(\theta)\right) \\ & - \frac{1}{2} z_t \left(\log\left(h_{1,t}(\theta)\right) + \frac{(y_t - \mu_t)^2}{h_{1,t}(\theta)} \right) \\ & - \frac{1}{2} (1 - z_t) \left(\log(h_{2,t}(\theta)) + \frac{(y_t - \mu_t)^2}{h_{2,t}(\theta)} \right) \Big], \end{aligned} \tag{7}$$

where z_t is a realization of the Bernoulli random variable as in Eq. (2) and represents the missing data. The maximization problem reads as follows:

$$\hat{\theta} = \underset{\theta \in \Omega}{\arg\max} \left(LL^{\#}(\theta) \right). \tag{8}$$

Dempster et al. (1977) show that a solution for the substitution problem (8) solves problem (6). Hence, the maximum-likelihood estimate for Eq. (7) is a maximum-likelihood estimate for Eq. (5). Implementing the EM algorithm, we replace the missing data z_t with its expected value τ_t, conditional on the observed $y = (y_t)_{t=1,\ldots,n}$ and estimates $\hat{\theta}$, and obtain the following algorithm structure in the k-th iteration:

E-step

Calculate the conditional expectation $\tau_t^{(k)}$ of z_t with regard to the observations y_t and the $(k-1)$-th estimate $\hat{\theta}^{(k-1)} = \left(\hat{m}^{(k-1)}, \hat{\omega}_1^{(k-1)}, \hat{\alpha}_1^{(k-1)}, \hat{\beta}_1^{(k-1)}, \hat{\omega}_2^{(k-1)}, \hat{\beta}_2^{(k-1)}, \hat{d}^{(k-1)}, \hat{\lambda}_0^{(k-1)}, \hat{\lambda}_1^{(k-1)}\right)$ of the $(k-1)$-th M-step as follows:

$$\tau_t^{(k)} = \frac{\hat{\pi}_t^{(k-1)} \phi\Big(\frac{y_t - \hat{\mu}_t^{(k-1)}}{\sqrt{\hat{h}_{1,t}^{(k-1)}}}\Big) \frac{1}{\sqrt{\hat{h}_{1,t}^{(k-1)}}}}{\hat{\pi}_t^{(k-1)} \phi\left(\frac{y_t - \hat{\mu}_t^{(k-1)}}{\sqrt{\hat{h}_{1,t}^{(k-1)}}}\right) \frac{1}{\sqrt{\hat{h}_{1,t}^{(k-1)}}} + (1 - \hat{\pi}_t^{(k-1)}) \phi\left(\frac{y_t - \hat{\mu}_t^{(k-1)}}{\sqrt{\hat{h}_{2,t}^{(k-1)}}}\right) \frac{1}{\sqrt{\hat{h}_{2,t}^{(k-1)}}}},$$

where $\phi(\cdot)$ denotes the density of a standard normal distribution and $\hat{\pi}_t^{(k-1)}$, $\hat{\mu}_t^{(k-1)}$, $\hat{h}_{1,t}^{(k-1)}$, and $\hat{h}_{2,t}^{(k-1)}$ denote the estimates of the respective processes that were calculated in the $(k-1)$-th M-step by estimating $\hat{\theta}^{(k-1)}$.

M-step

Solve problem (8) by replacing z_t with $\tau_t^{(k)}$ to obtain the estimate $\hat{\theta}^{(k)}$.[1] If there is no conditional mean structure, we decompose the maximization problem to three separate problems with $y_t = u_t$:

$$\hat{\theta}_1^{(k)} = \arg\min \sum_{t=1}^{n} \tau_t^{(k)} \left(\log(h_{1,t}(\theta)) + \frac{u_t^2}{h_{1,t}(\theta)} \right), \tag{9}$$

$$\hat{\theta}_2^{(k)} = \arg\min \sum_{t=1}^{n} (1 - \tau_t^{(k)}) \left(\log(h_{2,t}(\theta)) + \frac{u_t^2}{h_{2,t}(\theta)} \right), \tag{10}$$

$$\hat{\theta}_3^{(k)} = \arg\max \sum_{t=1}^{n} \left(\tau_t^{(k)} \log(\pi_t(\theta)) + (1 - \tau_t^{(k)}) \log(1 - \pi_t(\theta)) \right). \tag{11}$$

Consequently, we obtain $\hat{\theta}^{(k)} = \left(\hat{\theta}_1^{(k)}, \hat{\theta}_2^{(k)}, \hat{\theta}_3^{(k)}\right)$ by solving Eq. (9), (10), and (11) individually, where $\hat{\theta}_1^{(k)} = \left(\hat{\omega}_1^{(k)}, \hat{\alpha}_1^{(k)}, \hat{\beta}_1^{(k)}\right)$, $\hat{\theta}_2^{(k)} = \left(\hat{\omega}_2^{(k)}, \hat{\beta}_2^{(k)}, \hat{d}^{(k)}\right)$, and $\hat{\theta}_3^{(k)} = \left(\hat{\lambda}_0^{(k)}, \hat{\lambda}_1^{(k)}\right)$.

The main property of the EM algorithm defined above is that the log-likelihood function is monotonically increasing with every iteration,[2] we have

[1] With estimating $\hat{\theta}^{(k)}$, one obtains $\hat{\pi}_t^{(k)}$, $\hat{\mu}_t^{(k)}$, $\hat{h}_{1,t}^{(k)}$, and $\hat{h}_{2,t}^{(k)}$ which is then used in the $(k+1)$-th iteration and fed into the $(k+1)$-th E-step.

[2] While theory only ensures monotonicity, the implementation of the algorithm in `MATLAB` conveniently shows a strictly monotonically increasing log-likelihood without exception.

$$LL^{\#}\left(\hat{\theta}^{(k)}\right) \leq LL^{\#}\left(\hat{\theta}^{(k+1)}\right).$$

In our implementation, the EM algorithm is terminated after the l-th iteration if

$$\left\|\hat{\theta}^{(l)} - \hat{\theta}^{(l-1)}\right\|_2 < 10^{-5},$$

where $\|\cdot\|_2$ denotes the Euclidean norm. The estimate $\hat{\theta}^{(l)}$ is interpreted as a maximum-likelihood estimate for problem (6) and $LL(\hat{\theta}^{(l)})$ is calculated as log-likelihood of the given estimation for the initial problem.

It is noteworthy that the EM algorithm is very sensitive to starting values, e.g. one has to carefully calculate or estimate the input values $\left(\hat{\pi}_t^{(0)}, \hat{h}_{1,t}^{(0)}, \hat{h}_{2,t}^{(0)}\right)$ for the first E-step ($k = 1$). While the EM algorithm does not ensure convergence to a global maximum of the log-likelihood function, there are several techniques that increase the chance of converging to such. Deterministic and simulated annealing are considerable (Ueda & Nakano, 1998, Lavielle & Moulines, 1997), but are only tested for mixture density estimation. We implement a highly parallelized, grid-search like estimation for starting values $\hat{\theta}^{(0)}$ and $\left(\hat{\pi}_t^{(0)}, \hat{h}_{1,t}^{(0)}, \hat{h}_{2,t}^{(0)}\right)$ which is somewhat time-consuming. However, the total estimation time of the EM algorithm is vastly reduced because fewer iterations are needed if the algorithm is initiated with adequate inputs in the first (k=1) iteration.

The calculation of the standard errors of the parameter estimates is based on the missing information principle of Hartley & Hocking (1971). For technical details regarding the standard errors, we refer to Li et al. (2013) due to their complete derivation of the calculation process for this particular mixture model.

2.4. Variance Forecast

In view of the forecasting quality of the different models, the k day-ahead variance at time t, denoted by $\hat{h}_{t+k}$, is calculated for $k \in \{1, 5, 20\}$. At time t, we assume y_t, $\hat{\mu}_t$, and h_t to be known, where h_t stems from the parameter estimation of the given model.[3] An increasing window of training data as suggested in recent literature (Chkili et al., 2014, Hou & Suardi, 2012) is incorporated. For each t in the out-of-sample data a separate parameter estimation is carried out using all available data up to this time, yielding the respective parameter estimates which are used for the forecast at time t.

[3]While h_t is calculated from parameter estimates at time t, we omit the "hat" notation to distinguish from forecasted variances.

The GARCH$(1,1)$ model is then forecasted by

$$\mathbb{E}\left(h_{t+1}|\mathcal{F}_t\right) = \hat{h}_{t+1} = \hat{\omega} + \hat{\alpha}u_t^2 + \hat{\beta}h_t \tag{12}$$

for the 1 day-ahead forecast and by

$$\hat{h}_{t+k} = \frac{\hat{\omega}}{1-(\hat{\alpha}+\hat{\beta})} + (\hat{\alpha}+\hat{\beta})^{k-1}\left(\hat{h}_{t+1} - \frac{\hat{\omega}}{1-(\hat{\alpha}+\hat{\beta})}\right) \tag{13}$$

for the k day-ahead forecast. Note that we use the known variance h_t in Eq. (12). However, for Eq. (13), which yields from recursive calculation of the 1 day-ahead forecast, we set $u_{t+i}^2 = \hat{h}_{t+i}$ for all $i = 1, \ldots, k-1$. This is justified by the stationarity of the GARCH process.[4]

When predicting the variance with RiskMetrics, we use $\hat{h}_{t+1}$ for all k day-ahead predictions because the underlying IGARCH is not stationary.

The EGARCH forecast is carried out by applying the following equation recursively for $k \geq 1$:

$$\begin{aligned}\hat{h}_{t+k} = \hat{h}_{t+k-1} &\exp\left[(1-\beta)\,\omega - \alpha\sqrt{2/\pi}\right] \\ &\cdot\left(\exp\left[\frac{\gamma+\alpha^2}{2}\right]\Phi\left(\gamma+\alpha\right) + \exp\left[\frac{\gamma-\alpha^2}{2}\right]\Phi\left(\gamma-\alpha\right)\right),\end{aligned} \tag{14}$$

where $\Phi\left(\cdot\right)$ is the cumulative distribution function of the standard normal distribution (Tsay, 2013).

Forecasting the FIGARCH process, we incorporate the ARCH(∞) representation as in Eq. (3) and calculate the k day-ahead variance recursively by

$$\hat{h}_{t+k} = \frac{\hat{\omega}}{1-\hat{\beta}} + \delta(L)'\hat{h}_{(t+k-1,\ldots,1)},$$

where $\hat{h}_{(t+k-1,\ldots,1)} \coloneqq \left(\hat{h}_{t+k-1}, \ldots, \hat{h}_{t+1}, u_t^2, \ldots, u_1^2\right)$ is the data vector and $\delta(L)$ refers to the vector of the FIGARCH weights. Both vector lengths are determined by the chosen truncation lag. Predicting the HYGARCH process is done analogously to FIGARCH with respect to the additional parameter b.

[4] In literature it is common practice to already apply $u_t^2 = h_t$ in Eq. (12); we decide to use as much information as available, however.

In order to forecast the APARCH and FIAPARCH, the following additional term has to be calculated prior to calculating any $k > 1$ day-ahead forecast:

$$\mathbb{E}\left(\left(|u_{t+k-1}| - \hat{\gamma} u_{t+k-1}\right)^{\hat{\delta}} |\mathcal{F}_t\right) = \kappa h_{t+k-1}^{\frac{\hat{\delta}}{2}}, \tag{15}$$

with

$$\kappa = \frac{1}{\sqrt{2\pi}} \left((1+\hat{\gamma})^{\hat{\delta}} + (1-\hat{\gamma})^{\hat{\delta}}\right) 2^{\frac{(\hat{\delta}-1)}{2}} \Gamma\left(\frac{\hat{\delta}+1}{2}\right)$$

for normally distributed residuals (Ding et al., 1993), where $\Gamma(\cdot)$ denotes the Gamma function. Forecasts are then carried out analogously to the GARCH and FIGARCH forecasts.

The MMGARCH forecast consists of a separate GARCH and FIGARCH forecast. The mixing proportion is fixed at the value at time t. The k day-ahead forecast is calculated by

$$\hat{h}_{t+k} = \pi^*_{t+1} \hat{h}_{1,t+k} + (1 - \pi^*_{t+1}) \hat{h}_{2,t+k}, \tag{16}$$

where

$$\pi^*_{t+1} = \frac{1}{1 + \exp\left(-\hat{\lambda}_0 - \hat{\lambda}_1 (y_t - \mu_t)\right)}$$

is the last determinable mixture proportion and $\hat{h}_{1,t+k}$ and $\hat{h}_{2,t+k}$ refer to the GARCH and FIGARCH forecast, respectively.

2.5. Forecast evaluation and Value-at-Risk

In order to test and compare the prediction ability of the aforementioned models on the out-of-sample data, we apply different loss functions. We define the root mean squared error (RMSE)

$$\text{RMSE} := \sqrt{\frac{1}{M} \sum_{t=1}^{M} (\hat{h}_t - h_t)^2},$$

the mean absolute error (MAE)

$$\text{MAE} := \frac{1}{M} \sum_{t=1}^{M} |\hat{h}_t - h_t|,$$

the QLIKE which represents the loss against a normal likelihood

$$\mathrm{QLIKE} := \frac{1}{M}\sum_{t=1}^{M}\left(\log\left(\hat{h}_t\right) - \frac{h_t}{\hat{h}_t}\right),$$

the R^2LOG based on Pagan & Schwert (1990)

$$\mathrm{R}^2\mathrm{LOG} := \frac{1}{M}\sum_{t=1}^{M}\left(\log\left(\frac{h_t}{\hat{h}_t}\right)\right)^2,$$

and the mixed mean error (MME) for under-predicted values

$$\mathrm{MME(U)} := \frac{1}{M}\left(\sum_{t\in O}|\hat{h}_t - h_t| + \sum_{t\in U}\sqrt{|\hat{h}_t - h_t|}\right),$$

and over-predicted values

$$\mathrm{MME(O)} := \frac{1}{M}\left(\sum_{t\in U}|\hat{h}_t - h_t| + \sum_{t\in O}\sqrt{|\hat{h}_t - h_t|}\right),$$

where M denotes the number of observations in the out-of-sample data, $\hat{h}_t$ denotes the forecasted variance at time t and h_t denotes the actual, realized variance. The sets O and U are defined as in Brailsford & Faff (1996), where

$$O := \left\{t \in \{1,\dots,M\}\,\middle|\,\hat{h}_t > h_t\right\} \text{ and}$$
$$U := \left\{t \in \{1,\dots,M\}\,\middle|\,\hat{h}_t < h_t\right\}.$$

MME(U) and MME(O) penalize under- and over-predicted values, respectively. Therefore, both of them can be used to derive results regarding application in risk management. Following Wei et al. (2010), the realized variance h_t is set equal to the squared daily residual $u_t^2 = (y_t - \hat{\mu}_t)^2$.[5] The decision on the out-of-sample performance is made with respect to the smallest error calculated in each loss function. We apply the Superior Predictive Ability (SPA) test by Hansen (2005) to assess the outperformance of a model. The null hypothesis states that the benchmark model is not inferior to the

[5]We note that squared residuals as a measure of realized variance of a real data time series might be biased by some idiosyncratic error. For discussion on this matter, see Andersen & Bollerslev (1998).

models in its peer.

In order to compare the benefits for risk management applications we forecast the Value-at-Risk (VaR) for each $t \in 1, 2, \ldots, M$ in the out-of-sample series k day-ahead by

$$VaR_{p,t}^{(k \text{ day})} = \hat{\mu}_{t+k} + Q_p \sqrt{\hat{h}_{t+k}},$$

where Q_p denotes the p-quantile of the standard normal distribution and $\hat{\mu}_{t+k} = \mathbb{E}\left(\mu_{t+k} | \mathcal{F}_t\right)$. We examine the VaR at the 0.01 and 0.05 significance level and at the 0.99 and 0.95 significance level for long and short investment positions, respectively. For testing the VaR, we apply the popular unconditional coverage test by Kupiec (1995), the conditional coverage test by Christoffersen (1998), and the alternatives for both tests proposed by Ziggel et al. (2014). In comparison to the Kupiec and Christoffersen test, the Ziggel alternatives are one-sided and test for i.i.d. observations instead of a first-order Markov chain. We compare the performance of the models in terms of variance forecast quality by the above-mentioned tests for $k \in \{1, 5, 20\}$.

3. Data

Our data set consists of daily log return series of the U.S. West Texas Intermediate (WTI) and European Brent crude oil prices which we obtain from the U.S. Energy Information Administration (EIA). We define returns y_t of prices P_t by $y_t = \log\left(P_t/P_{t-1}\right)$ for $t = 2, \ldots, n$. The sample data is acquired from 01/01/1995 to 12/31/2014 which yields $n = 5024$ return observations for WTI and $n = 5062$ return observations for Brent. The last five years are used for out-of-sample analysis which produces an out-of-sample length of $M = 1260$ for WTI and $M = 1255$ for Brent.

This data set allows us to adjust the volatility models with two major crises (Dot-Com 2002 and Financial Crisis 2008) and test them with the decline in oil prices in 2014. The return series of each crude oil blend are displayed in Fig. 1 and 2. Both figures appear to have volatility clusters, with periods of low returns and periods of high returns.

The descriptive statistics of the return series for both blends are given in Tab. 1. The mean of the two return series is close to zero, but negative skewness, a kurtosis higher than 3, and the significant Jarque-Bera test show evidence of non-normality. This characteristic can be the result of volatility clusters and different volatility regimes. The Ljung-Box test rejects the hypothesis of i.i.d. observation in the squared returns and

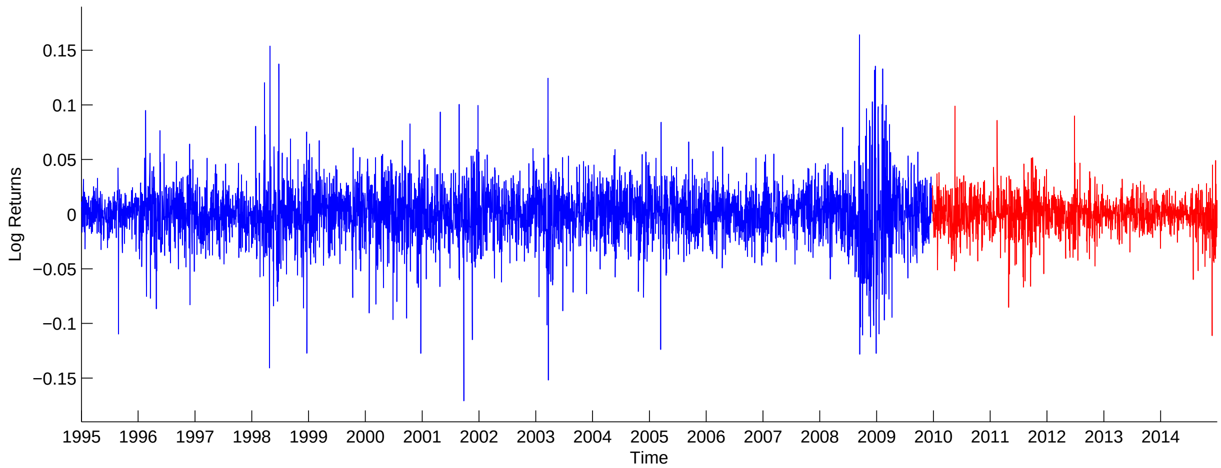

Figure 1: WTI log returns, 01/04/1995 – 12/31/2014, $n = 5024$, out-of-sample window is plotted in red.

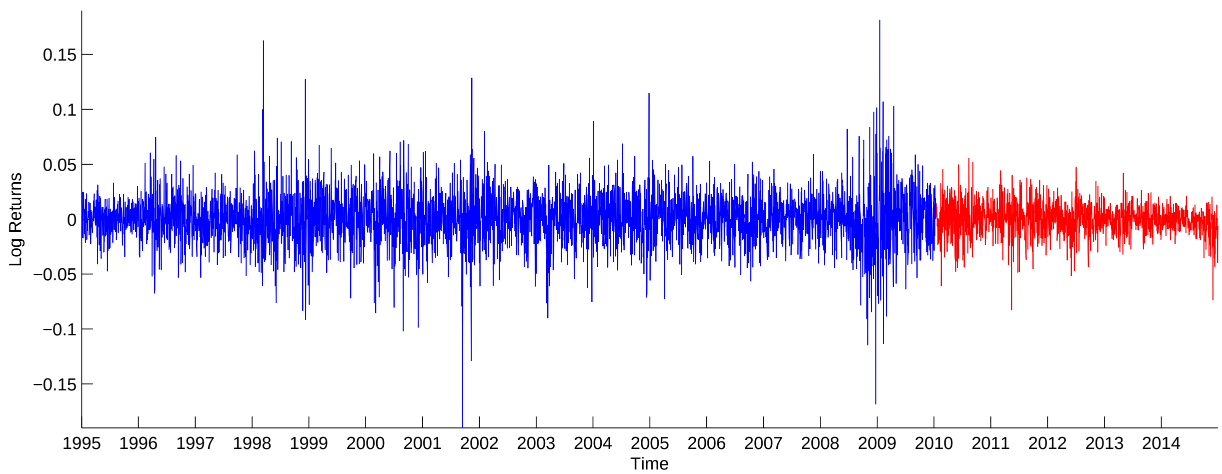

Figure 2: Brent log returns, 01/04/1995 – 12/31/2014, $n = 5062$, out-of-sample window is plotted in red.

describe highly significant autocorrelation,[6] which may imply a conditional structure in the variance of the time series. The same test finds autocorrelation in the returns of WTI and cannot reject the hypothesis of no autocorrelation in the Brent returns. The hypothesis of non-stationarity is rejected by the augmented Dickey-Fuller test and the hypothesis of stationarity is not rejected by the Kwiatkowski, Philipps, Schmidt, Shin (KPSS)-test. Therefore, we assume both series to be stationary, a requirement for most time series analysis.

[6] Ljung & Box (1978) show that their test is insensitive for deviation from the assumption of normally distributed time series.

	WTI	Brent
Descriptive Statistics		
Observations	5024	5062
Mean	0.0002	0.0002
Standard deviation	0.0241	0.0223
Minimum	−0.1709	−0.1989
Maximum	0.1641	0.1813
Skewness	−0.1983	−0.1056
Kurtosis	8.1099	8.4545
Preliminary Tests		
Jarque-Bera	5489.68***	6274.23***
Ljung-Box $Q'(8)$	31.11***	12.45
Ljung-Box $Q'^2(8)$	959.74***	466.39***
Augmented Dickey-Fuller	−16.25***	−15.94***
KPSS	0.0573	0.0788

Table 1: Descriptive statistics and preliminary tests for WTI and Brent log returns, 01/04/1995 – 12/31/2014. Rejection of the null hypothesis is displayed by *, **, *** for 10%, 5% and 1% significance level.

Prior to the analysis of the conditional volatility structure of both oil blends, the conditional mean structure is examined. We make use of the Box/Jenkins method (Box et al., 2008), which applies different autoregressive integrated moving average (ARIMA) models on the return series and compares the Bayesian Information Criterion (BIC) as a measure for the goodness-of-fit. As both time series appear to be stationary, no further differencing (degree of integration) is needed. Furthermore, we cannot identify any lag-order, neither for the autoregressive part nor for the moving average part, that yields a higher BIC than the simple unconditional mean (intercept).[7] Hence, we include an unconditional mean ($\mu_t \equiv \mu$ const.) into the parameter estimation and forecasting in the following section.

4. Results and Discussion

4.1. Estimation Results

The estimation results of the different models are presented in Tab. 2 (WTI) and Tab. 3 (Brent). Both tables show that MMGARCH clearly outperforms all other mod-

[7]These results are available upon request.

els regarding the Log-Likelihood (LL) and the BIC, indicating a better fit. All models significantly decrease the autocorrelation in the return and squared return series according to the Ljung-Box statistics of the standardized residuals. Due to the goodness-of-fit of MMGARCH, we conclude that both time series have different volatility structures, which is consistent with the findings of Fong & See (2002), Nomikos & Pouliasis (2011), and Chang (2012). All three studies find regimes of low and high volatility in oil future data by applying different Markov-Switching GARCH models. In our approach a link function is used to mix between GARCH and FIGARCH instead of a Markovian transition matrix.

Additionally, we find that these two components do not only distinguish between volatility levels but also differ in the persistence of shocks. This result expands the literature by giving evidence of different decay of shocks during the time period evaluated. Assuming no difference in the memory structure (short or long memory) of the time series, the fractional differencing parameter of MMGARCH would be $d = 0$ which is not the case for both blends. Hence, each series consist of two different variance components. The first component has a short memory, with weights declining exponentially and a low unconditional volatility. The second component has a long memory with a hyperbolic decline of weights and a high volatility level which is derived from the different dimensions of the parameters ω_1 and ω_2 in MMGARCH.

For the Brent return series, APARCH and FIAPARCH feature a relatively high and statistically significant leverage parameter γ. For both models the parameter is positive which emphasizes that downward movements have a greater impact on the conditional variance than upward movements.The EGARCH reveals the same properties which is consistent with recent literature, e.g. Chkili et al. (2014). This asymmetric news impact behavior is covered by a large mixing parameter λ_1 in the MMGARCH. The parameter λ_1 increases the proportion of the high-volatility FIGARCH in the variance mixture if large, downward movements in returns are present. Fig. 3 visualizes the estimated mixing proportion of the variance components for the Brent. In contrast, the WTI return series features a less pronounced asymmetry in its variance.

The estimation results also indicate that long memory models feature a better goodness-of-fit than short memory models for the WTI and Brent. This is supported by studies finding long memory and asymmetry in oil price return volatility (e.g. Charfeddine, 2014, Chkili et al., 2014)

Much to our surprise we discover significant structural differences regarding the MMGARCH model in the return series of WTI and Brent. Shocks and elevated variance

	GARCH	RiskMetrics	EGARCH	APARCH	FIGARCH	HYGARCH	FIAPARCH	MMGARCH
μ	0.0003	0.0003	0.0002	0.0002	0.0005	0.0005	0.0002	0.0005
	(0.0002)	(0.0000)	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0001)	(0.0001)
$\omega_1\ (\times 10^{-4})$	0.0638		−0.1024	0.0600				0.0197
	(0.0025)		(0.0487)	(0.1025)				(0.0060)
α_1	0.0529	0.0600	0.1497	0.0715				0.0305
	(0.0154)		(0.0364)	(0.0307)				(0.0036)
β_1	0.9293	0.9400	0.9857	0.9284				0.9548
	(0.0169)		(0.0066)	(0.0273)				(0.0050)
$\omega_2\ (\times 10^{-4})$					0.7169	0.2944	0.4213	0.0270
					(0.2148)	(0.1819)	(0.2192)	(0.0020)
φ					0.2675	0.2615	0.2843	
					(0.1386)	(0.1925)	(0.0910)	
β_2					0.4861	0.4542	0.4628	0.4830
					(0.0474)	(0.2091)	(0.0992)	(0.1646)
d					0.3426	0.2982	0.2962	0.9999
					(0.1440)	(0.0775)	(0.0413)	(0.0320)
$\log(b)$						0.0672		
						(0.0864)		
γ			−0.0249	0.1406			0.0949	
			(0.0164)	(0.0841)			(0.0355)	
δ				1.4358			2.1078	
				(0.3775)			(0.1034)	
λ_0								2.3533
								(0.2071)
λ_1								−2.3320
								(3.4366)
LL	12 113	12 074	12 117	12 122	12 124	12 119	12 126	12 271
BIC	−24 192	−24 157	−24 192	−24 193	−24 205	−24 187	−24 192	−24 465
LB $Q'(8)$	4.8927	3.6202	6.1364	5.9940	5.0412	4.6638	5.3189	6.3319
LB $Q'^2(8)$	19.8529**	18.9948^{h}**	28.2763***	25.1425***	11.3742	10.3755	10.9648	20.6984***

Table 2: Parameter estimates for WTI log returns, 01/04/1995 – 12/31/2014, $n = 5024$. Standard errors are given in parenthesis.

	GARCH	RiskMetrics	EGARCH	APARCH	FIGARCH	HYGARCH	FIAPARCH	MMGARCH
μ	0.0003	0.0003	0.0001	0.0002	0.0002	0.0004	0.0003	0.0004
	(0.0003)	(0.0000)	(0.0001)	(0.0002)	(0.0001)	(0.0003)	(0.0003)	(0.0002)
$\omega_1\,(\times 10^{-4})$	0.0161		−0.0540	0.0311				0.0000
	(0.0039)		(0.0219)	(0.0301)				(0.0000)
α_1	0.0532	0.060	0.1208	0.0508				0.0263
	(0.0092)		(0.0244)	(0.0068)				(0.0034)
β_1	0.9455	0.9400	0.9925	0.9490				0.9654
	(0.0075)		(0.0029)	(0.0109)				(0.0041)
$\omega_2\,(\times 10^{-4})$					0.4023	0.0818	0.0277	0.5138
					(0.0081))	(0.0601)	(0.0178)	(0.0549)
φ					0.2249	0.2138	0.2600	
					(0.0615)	(0.5307)	(0.0705)	
β_2					0.6244	0.5307	0.5682	0.6739
					(0.0827)	(0.1510)	(0.0788)	(0.0388)
d					0.4413	0.3243	0.3566	0.7069
					(0.0593)	(0.0962)	(0.0524)	(0.0120)
$\log(b)$						0.0942		
						(0.0811)		
γ			−0.0290	0.1799			0.2187	
			(0.0085)	(0.0427)			(0.0556)	
δ				1.8490			2.0639	
				(0.1883)			(0.0994)	
λ_0								1.6857
								(0.2501)
λ_1								22.0509
								(6.7507)
LL	12 586	12 563	12 585	12 596	12 589	12 589	12 606	12 670
BIC	−25 138	−25 134	−25 126	−21 141	−25 135	−25 127	−25 152	−25 263
LB $Q'(8)$	9.4290	10.0942	8.6536	8.4398	8.9478	8.9061	7.4710	7.3949
LB $Q'^2(8)$	9.0241	6.0058	14.9461*	10.5612	7.0547	5.4169	6.2585	13.5859*

Table 3: Parameter estimates for Brent log returns, 01/04/1995 – 12/31/2014, $n = 5062$. Standard errors are given in parenthesis.

levels in the WTI series appear to be significantly more persistent than in the Brent returns. We derive this result from the different fractional differencing parameters d of the FIGARCH component in MMGARCH. While we estimate $d = 0.7069$ for the Brent series, the WTI series yields $d \approx 1$.[8] The latter indicates an infinite persistence known as IGARCH effect. If $d = 1$, FIGARCH$(1, d, 0)$ decomposes to an IGARCH$(1, 1)$ which is easily shown in the following equation:

$$\begin{aligned} h_{2,t} &= \omega_2 + \left((1-\beta_2 L) - (1-L)^d\right) u_t^2 + \beta_2 h_{2,t-1} \\ &\overset{(d=1)}{=} \omega_2 + \left((1-\beta_2)L\right) u_t^2 + \beta_2 h_{2,t-1} \\ &= \omega_2 + (1-\beta_2) u_{t-1}^2 + \beta_2 h_{2,t-1}. \end{aligned}$$

The differences in the return and variance series of the WTI and Brent are of a very complex nature. Differences could be caused by variations of the USD/EUR rate and varying local supply and demand. The latter relates to the expansion of hydraulic fracturing in North America, for example. Buyuksahin et al. (2013) analyze the WTI-Brent spread and find evidence that business cycles, storage levels in the United States, and transportation infrastructure explains some of the observed spread. A connection to WTI future markets is also found. Kanamura (2015) focuses on the financialization of crude oil markets and finds significant differences in the correlation of the blends with the S&P 500 index. While these studies identify and aim to explain the price spread, our results indicate that there is also a significant difference in the conditional variance and memory structure of the WTI and Brent.

Regarding the calculated standard errors, it is noteworthy that for both blends the mixing parameter λ_1 stands out. Setting $\lambda_1 = 0$ prior to parameter estimation to obtain a constant mixing proportion worsens results in view of goodness-of-fit. We assume that the elevated errors are caused by the exponential structure of the link function given in Eq. 4. As for WTI where the parameter λ_1 is not statistically significant, the mixing proportion is somewhat stable and the lagged return has only little impact on it.

Testing the aforementioned IGARCH findings for the WTI blend, we mix two GARCH$(1, 1)$ processes in the MMGARCH model which produces the same result with

[8] Our implementation of the FIGARCH parameter estimation caps the maximum value of d at 0.9999. The interested reader will find that allowing a value of $d = 1$ causes problems in maximizing the log-likelihood function.

the same Log-Likelihood. The results are shown in Tab. 4. Inevitably, the GARCH estimates of one GARCH component are identical to the GARCH estimates in the initial MMGARCH model. The second component reveals an IGARCH process with parameter estimates $\alpha + \beta \approx 1$, which confirms our assumptions of an IGARCH effect obtained by the FIGARCH parameter estimates. Having identified an IGARCH process, we loose covariance stationarity which has major effects on variance forecasting. Depending on starting values, the GARCH/GARCH mixture yields another symmetric estimate with mixing proportion parameters $(\tilde{\lambda}_0, \tilde{\lambda}_1) = -(\lambda_0, \lambda_1)$, where the parameter estimates for the first and second GARCH component are interchanged. This behavior of the EM-algorithm is perfectly coherent and demonstrates the convergence to a maximum of the Log-likelihood function.

GARCH1		GARCH2		Mix & LL	
ω_1	1.20×10^{-6}	ω_2	1.96×10^{-6}	λ_0	-2.3553
α_1	0.5174	α_2	0.0305	λ_1	2.3440
β_1	0.4826	β_2	0.9548	LL	12 271

Table 4: Parameter estimates for WTI with a GARCH/GARCH mixture in MMGARCH, 01/04/1995 – 12/31/2014, $n = 5024$.

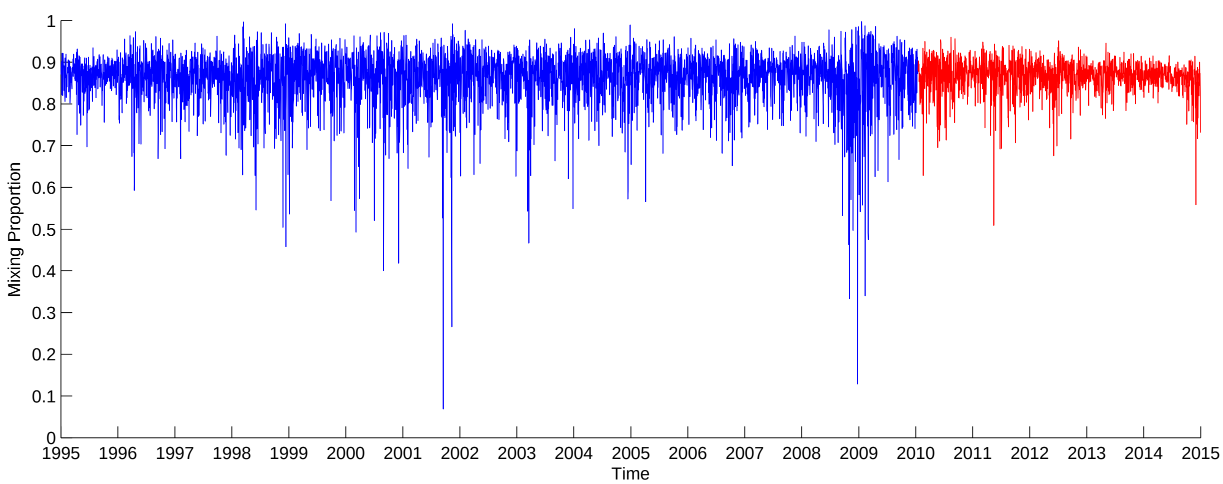

Figure 3: Brent mixing proportion for the parameter estimates from 01/04/1995 – 12/31/2014, $n = 5062$, out-of-sample window is plotted in red.

4.2. Forecasting Results

Firstly, we compare the results for variance forecasts obtained with the loss functions defined. The calculated errors are listed in Tab. 5. The smallest value for each loss function represented in the rows is given in bold type. For both WTI and Brent, we find that MMGARCH minimizes loss functions; or is close to the minimum value while the hypothesis of not being inferior to any other model cannot be rejected. This hypothesis is tested by Hansen's SPA. Regarding RMSE and MAE, we find that for both blends MMGARCH is superior to any other model tested as it produces the smallest errors. For the QLIKE and R^2LIKE loss functions, MMGARCH either features the lowest error and hence outperforms the other models or features a similar performance and is not inferior to the best model (QLIKE for 1 day-ahead forecasts). In view of the loss function MME(U), that penalizes under-predicted variances, we find that FIAPARCH is superior to any other model, while other long memory models display very good results. This might stem from their behavior of over-predicting the variance. This becomes more obvious with the loss function MME(O) which penalizes over-predicted variances. For this measure, FIAPARCH and other long memory models are inferior to any other tested model and for all forecasting horizons for WTI and Brent. We hence conclude that predicting future variance of oil prices with a FIAPARCH or FIGARCH process is not recommended due to their tendency to over-predict future variances for this time window. On one hand, this would lead to very conservative VaR-forecasts, on the other hand it could cause excessive hedging costs. The MMGARCH minimizes the MME(O) in all cases except for the 20 day-ahead WTI prediction. We do not detect any differences in forecasting quality of MMGARCH regarding the forecasting horizon. Overall, MMGARCH shows very good performance in variance prediction for different forecasting horizons and outperforms the tested models in most cases for WTI and Brent with respect to the implemented loss functions.

Lastly and more importantly, we compare the performance of the given models in regard to Value-at-Risk predictions and coverage. The results are given in Tab. 6 for the WTI and in Tab. 7 for Brent. Comparing the results of the blend WTI for the short position, we find MMGARCH to outperform most of the other models for all time horizons. Only RiskMetrics has similar coverage performance. The 5% 20 day-ahead forecast has a coverage of 3.6%, but is still closer to 5% than all other GARCH variants (except RiskMetrics, which has a slightly better coverage). However, the Kupiec and the Christoffersen tests reject the 5% 20 day-ahead prediction because of its low coverage. Another rejected test is the conditional coverage (Christoffersen) at 5% 1 day-ahead. The test detects a first order Markov chain in the Value-at-Risk violations, which is an unwanted property. Nevertheless, the statistic is still small (8.9855) compared to the other models (between 18.6987 and 42.3081), which are all

			GARCH	RiskMetrics	EGARCH	APARCH	FIGARCH	HYGARCH	FIAPARCH	MMGARCH
1 day-ahead	WTI	RMSE ($\times 10^{-4}$)	7.2969*	7.3288*	7.3062	7.2680	7.3028	7.2953	7.2948	**7.2490**
		MAE ($\times 10^{-4}$)	3.4842***	3.2046***	3.4992***	3.4749***	3.4486***	3.3636***	3.5862***	**2.9989**
		QLIKE	−7.2063**	−7.1839*	-7.2109	**-7.2170**	-7.2000	-7.2045	−7.1924***	-7.2141
		R^2LOG	8.3207***	7.3790*	8.5074***	8.3014***	8.1265***	7.8347***	8.5823***	**7.0131**
		MME(U)	0.0048	0.0057***	0.0049	0.0048	0.0047	0.0049	**0.0044**	0.0059***
		MME(O)	0.0120***	0.0099***	0.0119***	0.0120***	0.0119***	0.0114***	0.0127***	**0.0091**
	Brent	RMSE ($\times 10^{-4}$)	4.4344	4.4223	4.4331	4.4290	4.4714**	4.4455	4.5206***	**4.4114**
		MAE ($\times 10^{-4}$)	2.5842***	2.3517***	2.5993***	2.5874***	2.5785***	2.4976***	2.7878***	**2.2118**
		QLIKE	**-7.5281**	-7.5248	-7.5216	-7.5274	−7.5062**	-7.5203	−7.4833***	-7.5168
		R^2LOG	9.6161***	8.4914**	9.1295***	9.1464***	9.3648***	9.1520**	9.6998***	**8.1446**
		MME(U)	0.0043*	0.0050***	0.0043*	0.0043*	0.0043*	0.0045**	**0.0038**	0.0054***
		MME(O)	0.0103***	0.0085***	0.0104***	0.0103***	0.0102***	0.0097***	0.0115***	**0.0075**
5 day-ahead	WTI	RMSE ($\times 10^{-4}$)	7.3673	7.3481	7.3263	7.3588	7.3957*	7.3893	7.4267***	**7.3183**
		MAE ($\times 10^{-4}$)	3.6551***	3.2103***	3.5426***	3.6450***	3.6924***	3.6259***	3.9714***	**3.0841**
		QLIKE	−7.1756***	-7.1845	-7.2074	−7.1794***	−7.1546***	−7.1635**	−7.1209***	**-7.1814**
		R^2LOG	8.6595***	7.4058	8.5913***	8.6149***	8.5564***	8.3040***	9.2772***	**7.1843**
		MME(U)	0.0046	0.0057***	0.0048**	0.0046	0.0046	0.0047*	**0.0041**	0.0058***
		MME(O)	0.0128* **	0.0098***	0.0121***	0.0127***	0.0129***	0.0125***	0.0144***	**0.0095**
	Brent	RMSE ($\times 10^{-4}$)	4.4880*	4.4504	4.4488	4.4765	4.4896	4.4797	4.6346***	**4.4305**
		MAE ($\times 10^{-4}$)	2.6553***	2.3655***	2.6338***	2.6555***	2.6417***	2.5844***	3.0593***	**2.2000**
		QLIKE	-7.4991	-7.5141	**-7.5166**	−7.5037*	−7.4946**	-7.4994	−7.4178***	-7.5018
		R^2LOG	9.8667***	8.5323**	9.2229***	9.3698***	9.5862***	9.4790***	10.4065***	**8.1510**
		MME(U)	0.0042***	0.0051***	0.0042***	0.0042***	0.0042***	0.0043***	**0.0035**	0.0055***
		MME(O)	0.0107***	0.0084***	0.0106***	0.0107***	0.0106***	0.0103***	0.0130***	**0.0074**
20 day-ahead	WTI	RMSE ($\times 10^{-4}$)	7.5408***	7.4802*	7.4142	7.5491**	7.5262**	7.5137**	7.5923***	**7.4046**
		MAE ($\times 10^{-4}$)	4.0715***	3.2823	3.7258***	4.1265***	4.0148***	3.9316***	4.4024***	**3.2702**
		QLIKE	−7.0913***	−7.0778*	**-7.1898**	−7.0912***	−7.0969***	−7.1059***	−7.0523***	-7.1555
		R^2LOG	9.3988***	**7.4743**	8.8952***	9.4930***	9.1513***	8.8843***	9.9687***	7.6988
		MME(U)	0.0043	0.0059***	0.0046*	0.0042	0.0043	0.0044*	**0.0039**	0.0056***
		MME(O)	0.0145***	**0.0098**	0.0129***	0.0147***	0.0143***	0.0138***	0.0160***	0.0105***
	Brent	RMSE ($\times 10^{-4}$)	4.5994***	4.5124	4.5254	4.6022**	4.5575	4.5953**	4.7563***	**4.4884**
		MAE ($\times 10^{-4}$)	2.8838***	2.3821***	2.7805***	2.8957***	2.8377***	2.8265***	3.2828***	**2.1978**
		QLIKE	−7.4299***	-7.4500	**-7.4963**	−7.4358***	−7.4488***	−7.4351**	−7.3554***	-7.4492
		R^2LOG	10.4068***	8.5177***	9.4879***	9.9445***	9.9440***	9.9433***	10.8542***	**8.0645**
		MME(U)	0.0039**	0.0051***	0.0040**	0.0038*	0.0038*	0.0040* *	**0.0033**	0.0056***
		MME(O)	0.0119***	0.0084***	0.0113***	0.0120***	0.0118***	0.0115***	0.0139***	**0.0073**

Table 5: WTI and Brent out-of-sample forecast loss functions results for GARCH, RiskMetrics, EGARCH, APARCH, FIGARCH, HYGARCH, FIAPARCH, and MMGARCH. The results in bold face show the best result for each loss function. Rejection of the null hypothesis of the Hansen (2005) SPA test is displayed by *, **, *** for 10%, 5% and 1% significance level. The null hypothesis is rejected if the model is inferior to the other models regarding a given loss function.

rejected except for RiskMetrics.

Regarding the long position for WTI the only really concerning drawback regards the 1% long position where MMGARCH under-predicts for all time horizons. Hence, all corresponding unconditional and conditional coverage tests reject the predictions. The 5% long position coverage performance is in line with the other GARCH models. Figure 4 shows the VaR predictions for GARCH, FIGARCH, and MMGARCH over the full out-of-sample period.

The results for the short position VaR of the Brent are very much in favor to the MMGARCH. The unconditional coverage test of Kupiec does not reject any VaR prediction of MMGARCH. Additionally, all other models fail (except the RiskMetrics 1% for all time horizons and the HYGARCH 1 day-ahead 1% forecast). Comparing the conditional coverage test of Christoffersen leads to the same result. Only RiskMetrics passes the 5% VaR test for 1 day-ahead predictions along with the MMGARCH. For 5 and 20 day-ahead MMGARCH is the only model which passes the tests. Thus, we conclude that the competitive GARCH models are too conservative when forecasting short position VaR of the Brent, especially for longer forecasting horizons. The only test which rejects the MMGARCH is the conditional coverage test of Ziggel et al. (2014) at 5% and 20 day-ahead prediction. Based on the test constructions, we conclude that VaR violations do not follow a first order Markov chain (Christoffersen test) but build clusters (Ziggel test). Clustering is an unfavorable property for any VaR forecast and is caused by the lagged reaction of the 20 day-ahead forecast. Indeed, all competitive models pass this test only because they are too conservative, which has a 50% influence on the test statistic.

For the long position the MMGARCH does not perform as good as for the short position. All predictions for all time horizons under-predict the realized VaR. Hence, all tests reject the model. However, a general statement on which model is to be favored over all long position Brent forecast horizons cannot be made. For the 1 day-ahead forecast APARCH, for the 5 day-ahead forecast EGARCH, and for the 20 day-ahead forecast HYGARCH show the best coverage.

We conclude that MMGARCH has a very good variance prediction power for WTI and Brent in comparison to other models. VaR forecasts for portfolios with oil blends should also concern about different regimes of volatility level and memory. MMGARCH shows very good results for short position VaR predictions, but has a tendency to under-predict VaR for long positions, which is most obvious in the results for Brent.

4.3. Additional Remarks and Results

For additional robustness checks, estimation and forecasts were carried out for a rolling time window with fixed in-sample length of $\tilde{n} = 3800$ observations. Estimation results differed very little and there was no qualitative difference neither in the loss functions nor in the VaR forecasting results. As expected, the standard errors of the parameter estimates remained elevated.

In addition, we divided the out-of-sample data into two periods of the same length (2.5 years, approximately 630 observations per period) to cover the sharp decline of oil prices separately. VaR coverage tests and prediction ability of the models performed broadly similar to the previous analysis; MMGARCH showed a minimal better performance in the first period for the 20 day-ahead forecasts.[9]

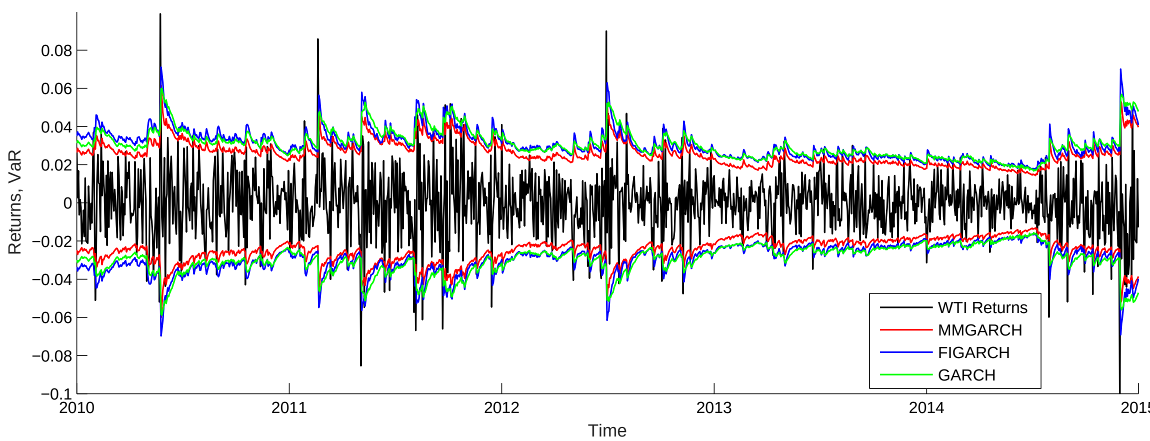

Figure 4: WTI 5% Value-at-Risk 1 day-ahead forecast of GARCH, FIGARCH, and MMGARCH from 01/01/2010 – 12/31/2014, $M = 1260$ for short (above) and long (below) position.

5. Conclusions

We are the first to test the MMGARCH on oil price returns and find it to be superior to non-regime-switching/non-mixing GARCH models with respect to the goodness-of-fit (Log-Likelihood and BIC) and to different measures for variance prediction.

In line with recent literature, the estimation results reveal significant long memory in the variance of Brent and WTI as well as asymmetric properties of the variance.

[9]The results of the loss functions and coverage tests for the split out-of-sample analysis are available upon request.

This suggests that more sophisticated GARCH-type models like APARCH and FIAPARCH should be utilized in order to capture the aforementioned stylized facts. By incorporating processes with different memory structure, the MMGARCH is able to depict long memory in the variance. With its dynamic and flexible mixture proportion, it also covers asymmetric news impact.

The MMGARCH reveals the existence of different volatility structures. This finding supports results of previous research on regime switching models, for example. In addition to different volatility levels, we detect that these different structures also feature different shock persistence, which has not yet been reported. As for WTI, we find the extreme case of infinite persistence, the so-called IGARCH effect. Comparing the two blends, Brent and WTI, we identify significant differences in structure of the variance components.

In the context of applications in risk management, we find that MMGARCH features the best over-all performance in variance forecasting for different and widely-used models tested. We also find that the long memory models for WTI and Brent tend to over-predict variances yielding very conservative VaR-forecasts. Regarding WTI the MMGARCH performs best in predicting the short-sided VaR. For the long position, it tends to slightly under-predict the VaR yielding an elevated coverage. Results for Brent differ little. The MMGARCH outperforms for short-sided VaR predictions. There is no model to be preferred for the long side, however. In general, we find that MMGARCH tends to cluster less than the models in comparison. We conclude that the differences in VaR prediction might also be caused by the structural difference of the WTI and Brent which has a significant impact on the models performance.

The MMGARCH is a suitable candidate for further research regarding oil price dynamics. Extending the model could include but is not limited to: alteration of distribution of the errors, e.g. from Normal to Student's t-distribution or generalized error distribution as well as substituting the components with asymmetric processes, e.g. a mixture of Asymmetric Power ARCH and Fractional Integrated Asymmetric Power ARCH to capture the gain/loss asymmetry in returns and therefore avoiding clustering in VaR predictions.

		pos.	VaR level p	GARCH	RiskMetrics	EGARCH	APARCH	FIGARCH	HYGARCH	FIAPARCH	MMGARCH
				WTI							
1 day-ahead	Coverage	short	0.01	0.0032	0.0103	0.0040	0.0032	0.0040	0.0040	0.0024	0.0071
			0.05	0.0214	0.0429	0.0254	0.0214	0.0238	0.0270	0.0159	0.0421
		long	0.01	0.0151	0.0262	0.0135	0.0135	0.0159	0.0183	0.0119	0.0278
			0.05	0.0413	0.0611	0.0421	0.0405	0.0389	0.0444	0.0333	0.0595
	UC_{Kup}	short	0.01	8.0799***	0.0127	6.0036**	8.0799***	6.0036**	6.0036**	10.6632***	1.1539
			0.05	27.3179***	1.4192	19.4425***	27.3179***	22.3853***	16.7563***	41.6304***	1.7620
		long	0.01	2.8411*	23.0810***	1.3991	1.3991	3.7254*	6.9696***	0.4352	27.1203***
			0.05	2.1441	3.0677*	1.7620	2.5664	3.5343*	0.8492	8.3072***	2.2737
	UC_{Zig}	short	0.01	3.9985	13.0003	5.0015	3.9982	4.9991	5.0011	3.0012	9.0001
			0.05	26.9997	54.0018	32.0012	26.9997	29.9993	34.0007	19.9995	52.9974
		long	0.01	19.0004**	32.9999***	16.9998	17.0012*	19.9997**	22.9996***	14.9992	35.0000***
			0.05	51.9992	77.0004**	52.9994	51.0001	49.0004	55.9991	41.9998	75.0011*
	CC_{Chr}	short	0.01	8.1118**	2.4381	6.0514**	8.1118**	6.0514**	6.0514**	10.6823***	1.2978
			0.05	28.5448***	1.7100	20.8123***	28.5448***	23.8983***	18.6987***	42.3081***	8.9855**
		long	0.01	3.4538	26.6200***	1.8917	1.8917	4.4031	7.8625**	0.8209	27.1774***
			0.05	5.3140*	10.4399***	9.2628***	5.9860*	4.1982	1.0477	8.6189**	3.8080
	CC_{Zig}	short	0.01	0.1835	0.5254	0.2277	0.1852	0.2773	0.2812	0.1191	0.3571
			0.05	0.2143	0.4286	0.2541	0.2143	0.2381	0.2699	0.1587	0.4310
		long	0.01	0.7540*	1.3310***	0.7229	0.6747	0.8120**	0.9749**	0.6053	1.3889***
			0.05	0.5072	0.6195*	0.4919	0.4663	0.3889	0.4484	0.3718	0.5982
5 day-ahead	Coverage	short	0.01	0.0048	0.0127	0.0040	0.0040	0.0056	0.0056	0.0024	0.0088
			0.05	0.0207	0.0430	0.0239	0.0239	0.0239	0.0271	0.0175	0.0414
		long	0.01	0.0111	0.0263	0.0135	0.0127	0.0135	0.0143	0.0088	0.0255
			0.05	0.0342	0.0629	0.0414	0.0374	0.0350	0.0366	0.0287	0.0613
	UC_{Kup}	short	0.01	4.2895**	0.8758	5.9551**	5.9551**	2.9603*	2.9603*	10.6019***	0.2043
			0.05	28.8668***	1.3599	22.1680***	22.1680***	22.1680***	16.5652***	36.8273***	2.0715
		long	0.01	0.1608	23.2132***	1.4275	0.8758	1.4275	2.0986	0.2043	21.2795***
			0.05	7.3538***	4.0808**	2.0715	4.5664**	6.5873**	5.1946**	14.1341***	3.1626*
	UC_{Zig}	short	0.01	6.0004	15.9992	5.0002	5.0016	7.0001	6.9987	3.0007	11.0007
			0.05	25.9993	53.9974	29.9997	29.9992	29.9993	33.9994	21.9996	52.0001
		long	0.01	14.0004	33.0011***	16.9999	15.9988	17.0004	18.0000*	10.9996	32.0014***
			0.05	42.9984	79.0002**	51.9998	47.0002	44.0019	45.9991	36.0010	76.9998**
	CC_{Chr}	short	0.01	4.3567	2.5710	6.0031**	6.0031**	3.0500	3.0500	10.6210***	3.2476
			0.05	34.7869***	6.0267**	23.8687***	26.6072***	26.6072***	19.8176***	40.4673***	3.5561
		long	0.01	0.4991	23.2868***	1.9217	1.3147	1.9217	2.6513	0.4164	21.3721***
			0.05	10.6798***	12.7214***	9.9561***	6.9600**	6.7926**	7.8031**	14.1934***	8.5499**
	CC_{Zig}	short	0.01	0.3423	0.6369	0.2312	0.2872	0.4478	0.4485	0.1195	0.7393
			0.05	0.2303	0.4299	0.2405	0.3046	0.3239	0.3659	0.3999	0.4140
		long	0.01	0.6720	1.3137***	0.7257	0.6910	0.7744*	0.8458**	0.6915	1.2835***
			0.05	0.4110	0.6721**	0.4875	0.4032	0.4189	0.4537	0.3476	0.6265*
20 day-ahead	Coverage	short	0.01	0.0032	0.0145	0.0040	0.0040	0.0040	0.0040	0.0024	0.0097
			0.05	0.0177	0.0451	0.0177	0.0193	0.0161	0.0185	0.0113	0.0363
		long	0.01	0.0105	0.0330	0.0121	0.0105	0.0105	0.0121	0.0073	0.0210
			0.05	0.0346	0.0629	0.0395	0.0346	0.0355	0.0387	0.0258	0.0548
	UC_{Kup}	short	0.01	7.8198***	2.2328	5.7740**	5.7740**	5.7740**	5.7740**	10.3725***	0.0138
			0.05	35.8219***	0.6410	35.8219***	31.7205***	40.2940***	33.7269***	56.3435***	5.4304**
		long	0.01	0.0279	41.4862***	0.5119	0.0279	0.0279	0.5119	1.0465	11.4298***
			0.05	6.8665***	4.0046**	3.1041*	6.8665***	6.1249**	3.6198*	18.4784***	0.5832
	UC_{Zig}	short	0.01	4.0009	17.9997*	5.0006	4.9998	4.9992	5.0006	3.0002	12.0009
			0.05	22.0022	56.0012	21.9988	23.9999	19.9998	23.0007	14.0023	45.0008
		long	0.01	12.9998	41.0015***	15.0000	12.9992	13.0019	14.9991	8.9997	25.9991***
			0.05	43.0010	78.0001**	48.9990	43.0003	43.9998	48.0004	32.0008	68.0020
	CC_{Chr}	short	0.01	7.8521**	12.3281***	5.8225*	5.8225*	5.8225*	5.8225*	10.3919***	2.7071
			0.05	39.4222***	2.8449	36.5549***	38.4532***	44.5666***	37.0243***	56.6860***	8.2052**
		long	0.01	0.3244	41.8416***	0.9036	0.3244	0.3244	0.9036	1.1926	12.5859***
			0.05	8.2268**	8.9389**	9.3664***	8.2268**	6.3195**	3.7102	18.5671***	4.9638*
	CC_{Zig}	short	0.01	0.1959	1.0010***	0.2428	0.3170	0.3067	0.3053	0.1209	0.8583**
			0.05	0.4117	0.6071	0.1773	0.3320	0.3493	0.2998	0.5687	0.5087
		long	0.01	1.2746***	1.7754***	0.6451	1.2770***	1.3697***	1.6219***	0.8492**	1.2544***
			0.05	0.5318	0.7129**	0.4591	0.5354	0.5837	0.5829	0.3728	0.8025***

Table 6: WTI out-of-sample forecast Value-at-Risk test results for GARCH, RiskMetrics, EGARCH, APARCH, FIGARCH, HYGARCH, FIAPARCH, and MMGARCH. The values given represent the test statistics of the Value-at-Risk tests by Kupiec (1995) (UC_{Kup}), Christoffersen (1998) (CC_{Chr}) and Ziggel et al. (2014) (UC_{Zig} and CC_{Zig}) at a given Value-at-Risk level (p) for short and long trading positions. Rejection of the null hypothesis is displayed by *, **, *** for 10%, 5% and 1% significance level.

		pos.	VaR level p	Brent GARCH	RiskMetrics	EGARCH	APARCH	FIGARCH	HYGARCH	FIAPARCH	MMGARCH
1 day-ahead	Coverage	short	0.01	0.0032	0.0088	0.0048	0.0040	0.0048	0.0064	0.0024	0.0120
			0.05	0.0239	0.0398	0.0223	0.0247	0.0239	0.0279	0.0167	0.0470
		long	0.01	0.0135	0.0239	0.0104	0.0112	0.0135	0.0175	0.0088	0.0263
			0.05	0.0542	0.0693	0.0430	0.0502	0.0470	0.0598	0.0327	0.0805
	UC_{Kup}	short	0.01	8.0113***	0.2018	4.2789**	5.9430**	4.2789**	1.9122	10.5866***	0.4547
			0.05	22.1138***	2.9223*	25.3137***	20.6175***	22.1138***	15.2742***	38.9705***	0.2405
		long	0.01	1.4347	17.6348***	0.0161	0.1631	1.4347	5.8702**	0.2018	23.2464***
			0.05	0.4506	8.8510***	1.3452	0.0010	0.2405	2.3757	8.9963***	20.8848***
	UC_{Zig}	short	0.01	4.0015	11.0005	5.9993	4.9987	6.0017	7.9993	2.9986	15.0008
			0.05	29.9996	49.9993	28.0000	31.0005	29.9999	35.0008	21.0008	59.000
		long	0.01	17.0000	30.0008***	13.0004	14.0001	17.0004	22.0026***	10.9989	32.9981***
			0.05	68.0001	87.0009***	54.0002	63.0022	59.0007	75.0004*	40.9998	100.9996***
	CC_{Chr}	short	0.01	8.0433**	0.4141	4.3462	5.9910*	4.3462	2.0277	10.6057***	0.8420
			0.05	23.6330***	3.4862	27.4026***	22.1406***	23.6330***	17.3409***	39.7196***	0.3563
		long	0.01	1.9293	17.7876***	0.3093	0.5017	1.9293	6.6914**	0.4141	23.3198***
			0.05	0.5910	11.5729***	4.0276	0.3312	0.3563	3.8879	9.1636**	22.9928***
	CC_{Zig}	short	0.01	0.1594	0.4383	0.3220	0.2137	0.2391	0.3187	0.1195	0.5976
			0.05	0.2391	0.4535	0.2231	0.2470	0.2391	0.2789	0.1673	0.5769
		long	0.01	1.0041***	1.2243***	0.6754	0.7104	1.0197***	1.0828***	0.7884*	1.3744***
			0.05	0.5419	0.7154***	0.4303	0.5020	0.4741	0.6306*	0.3910	0.8164***
5 day-ahead	Coverage	short	0.01	0.0032	0.0128	0.0048	0.0032	0.0032	0.0048	0.0024	0.0136
			0.05	0.0216	0.0344	0.0224	0.0232	0.0216	0.0240	0.0120	0.0464
		long	0.01	0.0112	0.0224	0.0096	0.0112	0.0136	0.0152	0.0072	0.0272
			0.05	0.0432	0.0703	0.0424	0.0392	0.0448	0.0488	0.0264	0.0839
	UC_{Kup}	short	0.01	7.9565***	0.9038	4.2369**	7.9565***	7.9565***	4.2369**	10.5253***	1.4636
			0.05	26.7860***	7.1900***	25.0842***	23.4554***	26.7860***	21.8973***	54.1401***	0.3567
		long	0.01	0.1726	14.3324***	0.0213	0.1726	1.4636	2.9347*	1.1025	25.3837***
			0.05	1.2873	9.7302***	1.6151	3.3276*	0.7472	0.0408	17.6245***	25.4134***
	UC_{Zig}	short	0.01	4.0009	16.0007	6.0008	3.9998	4.0010	5.9977	2.9985	17.0004
			0.05	26.9995	43.0027	28.0014	28.9999	27.0030	30.0007	15.0005	57.9984
		long	0.01	14.0018	28.0002***	11.9998	13.9991	16.9991	18.9996**	9.0017	33.9990***
			0.05	53.9998	88.0017***	52.9994	49.0003	55.9998	61.0008	33.0014	104.9995***
	CC_{Chr}	short	0.01	7.9886**	1.3445	4.3044	7.9886**	7.9886**	4.3044	10.5446***	1.9598
			0.05	28.0219***	7.4383**	27.1640***	25.3336***	28.0219***	23.4215***	54.5286***	3.8690
		long	0.01	0.5123	14.5754***	0.2732	0.5123	1.9598	3.5518	1.2475	25.4451***
			0.05	3.9513	13.4943***	4.5044	5.3077*	0.9385	4.6508*	17.6972***	30.0078***
	CC_{Zig}	short	0.01	0.1599	0.6395	0.3237	0.1599	0.1599	0.2397	0.1198	0.6795
			0.05	0.2158	0.3438	0.2238	0.2318	0.2159	0.2398	0.1199	0.5037
		long	0.01	0.9785***	1.2085***	0.6027	0.6397	1.0269***	1.0505***	0.5995	1.4935***
			0.05	0.4723	0.7696***	0.4237	0.4442	0.5545	0.5471	0.3705	0.9065***
20 day-ahead	Coverage	short	0.01	0.0032	0.0121	0.0024	0.0040	0.0032	0.0024	0.0008	0.0129
			0.05	0.0162	0.0388	0.0170	0.0146	0.0138	0.0210	0.0097	0.0542
		long	0.01	0.0121	0.0251	0.0089	0.0097	0.0113	0.0121	0.0065	0.0324
			0.05	0.0364	0.0736	0.0380	0.0372	0.0364	0.0421	0.0235	0.0850
	UC_{Kup}	short	0.01	7.7516***	0.5332	10.2963***	5.7139**	7.7516***	10.2963***	17.7962***	0.9908
			0.05	39.9438***	3.5021*	37.6675***	44.8067***	47.4045***	27.6586***	62.3469***	0.4488
		long	0.01	0.5332	20.0157***	0.1570	0.0107	0.2108	0.5332	1.7751	39.3022***
			0.05	5.2876**	12.7583***	4.0527**	4.6476**	5.2876**	1.7250	22.6246***	26.5212***
	UC_{Zig}	short	0.01	4.0022	15.0005	3.0004	5.0016	4.0004	3.0010	0.9996	15.9996
			0.05	20.0002	48.0009	21.0005	18.0001	16.9995	25.9995	12.0007	67.0012
		long	0.01	15.0000	30.9984***	10.9980	11.9999	14.0015	14.9981	8.0009	39.9980***
			0.05	44.9990	90.9995***	46.9995	46.0004	44.9993	51.9986	29.0003	105.0003***
	CC_{Chr}	short	0.01	7.7841**	2.4200	10.3158***	5.7627*	7.7841**	10.3158***	17.7994***	2.6602
			0.05	40.9264***	5.6189*	41.5787***	54.8795***	47.9068***	28.0175***	65.0331***	0.6002
		long	0.01	0.9265	21.6631***	0.3726	0.2657	0.5546	0.9265	1.8924	39.4458***
			0.05	5.4418*	14.5869***	4.8773*	4.7727*	5.6712*	2.1067	24.0669***	32.3411***
	CC_{Zig}	short	0.01	0.1619	0.6570	0.4158	0.2190	0.1618	0.1214	0.0404	0.6672
			0.05	0.1618	0.4441	0.1699	0.1868	0.1375	0.2104	0.2354	0.6251*
		long	0.01	0.9209**	1.4271***	0.5370	0.8828**	1.1020***	0.9322**	0.5473	1.9702***
			0.05	0.5480	0.8251***	0.4774	0.5629	0.5625	0.7340***	0.5925	0.9573***

Table 7: Brent out-of-sample forecast Value-at-Risk test results for GARCH, RiskMetrics, EGARCH, APARCH, FIGARCH, HYGARCH, FIAPARCH, and MMGARCH. The values given represent the test statistics of the Value-at-Risk tests by Kupiec (1995) (UC_{Kup}), Christoffersen (1998) (CC_{Chr}) and Ziggel et al. (2014) (UC_{Zig} and CC_{Zig}) at a given Value-at-Risk level (p) for short and long trading positions. Rejection of the null hypothesis is displayed by *, **, *** for 10%, 5% and 1% significance level.

References

Aloui, C., & Mabrouk, S. (2010). Value-at-risk estimations of energy commodities via long-memory, asymmetry and fat-tailed GARCH models. *Energy Policy*, *38*, 2326–2339. URL: http://linkinghub.elsevier.com/retrieve/pii/S0301421509009768http://www.sciencedirect.com/science/article/pii/S0301421509009768. doi:10.1016/j.enpol.2009.12.020.

Andersen, T. G., & Bollerslev, T. (1998). Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. *International Economic Review*, *39*, 885–905.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics*, *73*, 5–59. URL: http://linkinghub.elsevier.com/retrieve/pii/0304407695017321. doi:10.1016/0304-4076(95)01732-1.

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *74*, 3–30. URL: http://linkinghub.elsevier.com/retrieve/pii/S0304407695017496. doi:10.1016/S0304-4076(95)01749-6.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*, 307–327. URL: http://linkinghub.elsevier.com/retrieve/pii/0304407686900631. doi:10.1016/0304-4076(86)90063-1.

Bollerslev, T., & Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, *73*, 151–184. URL: http://linkinghub.elsevier.com/retrieve/pii/0304407695017364. doi:10.1016/0304-4076(95)01736-4.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time Series Analysis*. (4th ed.). Hoboken: Wiley.

Brailsford, T. J., & Faff, R. W. (1996). An evaluation of volatility forecasting techniques. *Journal of Banking & Finance*, *20*, 419–438. URL: http://linkinghub.elsevier.com/retrieve/pii/0378426695000151. doi:10.1016/0378-4266(95)00015-1.

Buyuksahin, B., Lee, T. K., Moser, J. T., & Robe, M. A. (2013). Physical Markets, Paper Markets and the WTI-Brent Spread. *The Energy Journal*, *34*, 129–152. doi:10.5547/01956574.34.3.7.

Chang, K.-L. (2012). Volatility regimes, asymmetric basis effects and forecasting performance: An empirical investigation of the WTI crude oil futures market. *Energy Economics*, *34*, 294–306. URL: http://linkinghub.elsevier.com/retrieve/pii/S0140988311002830. doi:10.1016/j.eneco.2011.11.009.

Charfeddine, L. (2014). True or spurious long memory in volatility: Further evidence on the energy futures markets. *Energy Policy*, *71*, 76–93. URL: http://www.sciencedirect.com/science/article/pii/S0301421514002626. doi:10.1016/j.enpol.2014.04.027.

Cheng, X., Yu, P. L. H., & Li, W. K. (2009). On a dynamic mixture GARCH model. *Journal of Forecasting*, *28*, 247–265. URL: http://doi.wiley.com/10.1002/for.1093. doi:10.1002/for.1093.

Chkili, W., Hammoudeh, S., & Nguyen, D. K. (2014). Volatility forecasting and risk management for commodity markets in the presence of asymmetry and long memory. *Energy Economics*, *41*, 1–18. URL: http://linkinghub.elsevier.com/retrieve/pii/S0140988313002363. doi:10.1016/j.eneco.2013.10.011.

Christoffersen, P. F. (1998). Evaluating Interval Forecasts. *International Economic Review*, *39*, 841–862. URL: http://www.jstor.org/stable/2527341?origin=crossref. doi:10.2307/2527341.

Conrad, C. (2010). Non-negativity conditions for the hyperbolic GARCH model. *Journal of Econometrics*, *157*, 441–457. URL: http://linkinghub.elsevier.com/retrieve/pii/S0304407610000977. doi:10.1016/j.jeconom.2010.03.045.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, *1*, 223–236. URL: http://www.tandfonline.com/doi/abs/10.1080/713665670. doi:10.1080/713665670.

Davidson, J. (2004). Moment and Memory Properties of Linear Conditional Heteroscedasticity Models, and a New Model. *Journal of Business & Economic Statistics*, *22*, 16–29. URL: http://amstat.tandfonline.com/doi/full/10.1198/073500103288619359. doi:10.1198/073500103288619359.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Soci-

ety. Series B (Methodological), *39*, 1–38. URL: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.133.4884. doi:10.1.1.133.4884.

Ding, Z., Granger, C. W., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, *1*, 83–106. URL: http://linkinghub.elsevier.com/retrieve/pii/092753989390006D. doi:10.1016/0927-5398(93)90006-D.

Engle, R. F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, *50*, 987–1007. doi:10.2307/1912773.

Engle, R. F., & Bollerslev, T. (1986). Modelling the persistence of conditional variances. *Econometric Reviews*, *5*, 1–50. URL: http://www.tandfonline.com/doi/abs/10.1080/07474938608800095. doi:10.1080/07474938608800095.

Engle, R. F., & Ng, V. K. (1993). Measuring and Testing the Impact of News on Volatility. *The Journal of Finance*, *48*, 1749. URL: http://www.jstor.org/stable/2329066?origin=crossref. doi:10.2307/2329066.

Fattouh, B. (2010). The dynamics of crude oil price differentials. *Energy Economics*, *32*, 334–342. URL: http://linkinghub.elsevier.com/retrieve/pii/S0140988309000978. doi:10.1016/j.eneco.2009.06.007.

Fong, W. M., & See, K. H. (2002). A Markov switching model of the conditional volatility of crude oil futures prices. *Energy Economics*, *24*, 71–95. URL: http://linkinghub.elsevier.com/retrieve/pii/S0140988301000871. doi:10.1016/S0140-9883(01)00087-1.

Hansen, P. R. (2005). A Test for Superior Predictive Ability. *Journal of Business & Economic Statistics*, *23*, 365–380. URL: http://www.tandfonline.com/doi/abs/10.1198/073500105000000063. doi:10.1198/073500105000000063.

Hartley, H. O., & Hocking, R. R. (1971). The Analysis of Incomplete Data. *Biometrics*, *27*, 783–823.

Hou, A., & Suardi, S. (2012). A nonparametric GARCH model of crude oil price return volatility. *Energy Economics*, *34*, 618–626. URL: http://linkinghub.elsevier.com/retrieve/pii/S0140988311001642. doi:10.1016/j.eneco.2011.08.004.

Kanamura, T. (2015). A Financialization Model of Crude Oil Markets. URL: http://ssrn.com/abstract=2712617.

Kang, S. H., Kang, S.-M., & Yoon, S.-M. (2009). Forecasting volatility of crude oil markets. *Energy Economics*, *31*, 119–125. URL: http://www.sciencedirect.com/science/article/pii/S0140988308001539. doi:10.1016/j.eneco.2008.09.006.

Kang, W., Ratti, R. A., & Yoon, K. H. (2014). The impact of oil price shocks on U.S. bond market returns. *Energy Economics*, *44*, 248–258. URL: http://linkinghub.elsevier.com/retrieve/pii/S0140988314000930. doi:10.1016/j.eneco.2014.04.009.

Kupiec, P. H. (1995). Techniques for Verifying the Accuracy of Risk Measurement Models. *The Journal of Derivatives*, *3*, 73–84. URL: http://www.iijournals.com/doi/abs/10.3905/jod.1995.407942. doi:10.3905/jod.1995.407942.

Lavielle, M., & Moulines, E. (1997). A simulated annealing version of the EM algorithm for non-Gaussian deconvolution. *Statistics and Computing*, *7*, 229–236. URL: http://link.springer.com/article/10.1023/A:1018594320699.

Li, M., Li, W. K., & Li, G. (2013). On Mixture Memory GARCH Models. *Journal of Time Series Analysis*, *34*, 606–624. URL: http://doi.wiley.com/10.1111/jtsa.12037. doi:10.1111/jtsa.12037.

Liu, L., Chen, C.-C., & Wan, J. (2013). Is world oil market "one great pool"?: An example from China's and international oil markets. *Economic Modelling*, *35*, 364–373. URL: http://www.sciencedirect.com/science/article/pii/S0264999313002988. doi:10.1016/j.econmod.2013.07.027.

Ljung, G. M., & Box, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika*, *65*, 297–303. URL: http://biomet.oxfordjournals.org/content/65/2/297.short. doi:10.1093/biomet/65.2.297.

Mohanty, S. K., Nandha, M., Habis, E., & Juhabi, E. (2014). Oil price shocks and transportation firm asset prices. *Energy Economics*, *41*, 117–124. URL: http://linkinghub.elsevier.com/retrieve/pii/S0140988313002235. doi:10.1016/j.eneco.2013.09.028.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, *59*, 347–370. URL: http://www.jstor.org/stable/2938260?origin=crossrefhttp://www.jstor.org/stable/2938260. doi:10.2307/2938260.

Nomikos, N. K., & Pouliasis, P. K. (2011). Forecasting petroleum futures markets volatility: The role of regimes and market conditions. *Energy Economics*, *33*, 321–337. URL: http://www.sciencedirect.com/science/article/pii/S014098831000201X. doi:10.1016/j.eneco.2010.11.013.

Pagan, A. R., & Schwert, G. W. (1990). Alternative models for conditional stock volatility. *Journal of Econometrics*, *45*, 267–290. doi:10.1016/0304-4076(90)90101-X.

Tsay, R. S. (2013). *An Introduction to Analysis of Financial Data with R*. Hoboken, New Jersey: Wiley.

Tse, Y. K. (1998). The conditional heteroscedasticity of the yen-dollar exchange rate. *Journal of Applied Econometrics*, *13*, 49–55. URL: http://doi.wiley.com/10.1002/(SICI)1099-1255(199801/02)13:1<49::AID-JAE459>3.0.CO;2-O. doi:10.1002/(SICI)1099-1255(199801/02)13:1<49::AID-JAE459>3.0.CO;2-O.

Ueda, N., & Nakano, R. (1998). Deterministic annealing EM algorithm. *Neural Networks*, *11*, 271–282. URL: http://linkinghub.elsevier.com/retrieve/pii/S0893608097001330. doi:10.1016/S0893-6080(97)00133-0.

Wang, Y., & Wu, C. (2012). Forecasting energy market volatility using GARCH models: Can multivariate models beat univariate models? *Energy Economics*, *34*, 2167–2181. URL: http://linkinghub.elsevier.com/retrieve/pii/S0140988312000540. doi:10.1016/j.eneco.2012.03.010.

Wang, Y., Wu, C., & Yang, L. (2014). Oil price shocks and agricultural commodity prices. *Energy Economics*, *44*, 22–35. URL: http://linkinghub.elsevier.com/retrieve/pii/S0140988314000632. doi:10.1016/j.eneco.2014.03.016.

Wei, Y., Wang, Y., & Huang, D. (2010). Forecasting crude oil market volatility: Further evidence using GARCH-class models. *Energy Economics*, *32*, 1477–1484. URL: http://linkinghub.elsevier.com/retrieve/pii/S0140988310001180. doi:10.1016/j.eneco.2010.07.009.

Ziggel, D., Berens, T., Weiß, G. N., & Wied, D. (2014). A new set of improved Value-at-Risk backtests. *Journal of Banking & Finance*, *48*, 29–41. URL: http://linkinghub.elsevier.com/retrieve/pii/S0378426614002520. doi:10.1016/j.jbankfin.2014.07.005.