Gutiérrez, Raúl de Jesús; Gutiérrez, Lidia E. Carvajal; García Salgado, Oswaldo

Article

Value at risk and expected shortfall estimation for Mexico's isthmus crude oil using long-memory GARCH-EVT combined approaches

International Journal of Energy Economics and Policy

ISSN: 2146-4553

available at http: www.econjournals.com

International Journal of Energy Economics and Policy, 2023, 13(4), 467-480.

Value at Risk and Expected Shortfall Estimation for Mexico's Isthmus Crude Oil Using Long-Memory GARCH-EVT Combined Approaches

Raúl De Jesús Gutiérrez*, Lidia E. Carvajal Gutiérrez, Oswaldo García Salgado

Facultad de Economía, Universidad Autónoma del Estado de México, Paseo Universidad, Universitaria, 50130 Toluca de Lerdo, México. *Email: rdejesusg@uaemex.mx

Received: 26 February 2023 **Accepted:** 18 June 2023 **DOI:** https://doi.org/10.32479/ijeep.14179

ABSTRACT

This paper estimates a variety of CGARCH and FIGARCH models with normal distribution to capture salient features of Mexico's Isthmus crude oil return series such as fat tails and volatility clustering as well as asymmetry and long memory; this to obtain independent and identically distributed standardized residuals series. Furthermore, extreme value theory is applied to model the tail behavior of the innovation distribution of the volatility models in estimating one-day-ahead VaR and Expected Shortfall (ES). In- and out-of-sample forecasting performance is evaluated by the unconditional coverage test of Kupiec and the Dynamic Quantile test of Engle and Manganelli. Backtesting results show strong and consistent evidence confirming that FIGARCH-EVT, ACGARCH1-EVT and CGARCH-EVT approaches yield the most accurate out-of-sample VaR and ES forecasts, for both short and long trading positions at quantiles ranging 95% to 99.9%. Findings provide useful tools for producers, consumers and portfolio investors who need sophisticated models for sound risk management and optimal hedging strategies to mitigate price risk exposure for the Isthmus crude oil.

Keywords: Crude Oil, Conditional Extreme Value Theory, Value at Risk and Expected Shortfall, Mexico's Isthmus Oil
JEL Classifications: C22, C52, G13, G32

1. INTRODUCTION

During the last three decades crude oil market has undergone important changes and innovations to such an extent that its structure has become a sophisticated financial market. However, this market has faced uncertainty and high volatility due to the presence of extreme fluctuations in crude oil prices derived from international liberalization policies of the energy sector undertaken since the mid-1980s. Upturns and reversals in oil prices have been triggered by imbalances between supply and demand originating from several exogenous events, such as geopolitical tensions, institutional OPEC's policies, military conflicts, financial and health crises, erratic economic growth, global climate change and diversification and speculation strategies as well as other factors affecting the global petroleum industry such as fracking, complex oil exploitation schemes, and oil spills (Boussena and Locatelli, 2005).

Extreme price movements and the increasing volatility do not only have negative effects on the global economy, but also imply a greater risk exposure for all crude oil market participants. That is, unexpected large losses with small probability of occurrence because of extreme market events in crude oil prices associated with left and right tails of the return distribution. Thus, risk administration processes in the energy sector have become a priority and a challenge for producers and industrial consumers of crude oil, as well as for investors seeking portfolio diversification. Consequently, in the literature, understanding and modeling extreme price fluctuations are significant importance in the development of financial models for quantifying and managing risk exposure. In this context, Value at Risk (VaR) methodology is a standard statistical measure widely used in the financial industry for quantifying market risk and estimating capital requirement sufficiency to cover market losses.

VaR is defined as the maximum possible loss that a portfolio or a security can experience with a given probability over a certain time horizon. Since VaR has been adopted as a regulation measure by the Basle Committee, several models have been developed for estimating catastrophic and extreme market risk incidents; these models include parametric methods based on GARCH volatility models and under different distributional assumptions, non-parametric historical simulation (HS) approaches based on empirical distribution of returns, and methods based on extreme value theory (EVT). An unusually large quantity of studies has shown the predictive accuracy of GARCH models and simple-filtered HS approaches in crude oil risk management. However, no consensus has been reached regarding the more appropriate model to predict the best VaR estimations about the crude oil market because phenomena such as fat tails, leptokurtosis and asymmetric effects, present in financial series, have been only partially modeled by proposed distributions in the literature (Bali and Theodossiou, 2007; Costello et al., 2008; Sowdagur and Narsoon, 2017; Riedle, 2018; Toumi et al., 2019; Echaust and Just, 2020).

In response to inconsistencies of conventional VaR approaches and their variants about modeling the magnitude and frequency of extreme returns in oil prices, EVT provides a solid statistical framework for analyzing the asymptotic behavior of extreme values and catastrophic risk related to tail areas of empirical distributions. In the pioneering work of Krehbiel and Adkins (2005), the conditional EVT is applied to energy commodity prices traded in New York Mercantile Exchange (NYMEX). Their findings reveal that the EVT-based approach provides more accurate forecasts than RiskMetrics and GARCH volatility models. Concerning WTI and Brent oil returns, Marimoutou et al. (2009) confirm the importance of filtering processes to improve both the performance of conditional EVT and filtered HS approaches, as well as capturing downside risk. In contrast, Ren and Giles (2010) suggest that unconditional EVT is sufficient to model the tails of the return distribution of the Canadian crude oil market and including more accurate VaR and ES estimations at high quantiles.

Chiu et al. (2010) offer robust evidence about poor predictive performance of the EVT approach in modeling tail risk of WTI and Brent crude oil prices. Still, Zikovic (2011) and Ghorbel and Trabelsi (2014) have found evidence favoring the EVT-based model for quantifying tail-related risk for WTI crude oil, natural gas, and heating oil futures contracts. In more recent studies, Susan and Waititu (2015) and Mi et al. (2017) reach the same results for the Brent and WTI crude oil markets by employing GARCH-EVT models based on normal and generalized error distributions. Jammazi and Nguyen (2017) estimate associated tail risks with wavelet-based extreme value theory for portfolios of crude oil prices and US dollar exchange rates. Findings show that W-EVT approaches provide an effective and powerful tool for modeling extreme movements and improving the accuracy of VaR estimates and forecasts. Weru et al. (2019) focus on modeling and forecasting the volatility and VaR of WTI crude oil and reformulated regular gasoline prices using a family of GARCH-EVT models. They find solid evidence that IGARCH-EVT and EGARCH-EVT models provide more accurate VaR estimates.

The previous literature has shown that the conditional EVT method is a robust tool for measuring the tail-related risk. However, the used GARCH processes in the filtering procedure of the EVT approach often cannot capture the long-term memory property or the high degree of volatility persistence in the crude oil market, since the autocorrelations of volatility decay at exponential rates when the lag order increases (Baillie, 1996). In fact, the recent finance literature has confirmed the presence of the long memory feature or long-range dependence in crude oil returns and conditional volatility (Elder and Serletis, 2008; Choi and Hammoudeh, 2009; Aloui and Mabrouk, 2010; Klein and Walther, 2016).

Consequently, Youssef et al. (2015) suggest using a range of long-memory GARCH models to forecast conditional volatility and EVT as a filtering process. Their results show that the FIAPARCH-EVT model performs better than the FIGARCH-EVT and HYGARCH-EVT models in forecasting the one-day-ahead VaR. Finally, Zhao et al. (2019) introduce a hybrid time-varying long memory GARCH-M-EVT model for calculation of static and dynamic VaR in the WTI crude oil market. Empirical results show that the family of FIGARCH models can capture appropriately asymmetry and long memory features in the conditional volatility. Backtesting results show that the performance of FIAPARCH-M-EVT model is superior to other GARCH-type models which only consider oil price fluctuation characteristics partially and traditional methods including Variance-Covariance and Monte Carlo in price risk measurement.

Most previous literature on energy price and returns risk measurement has focused on high quality crude oils based on their chemical and physical properties which are determined by the American Petroleum Institute (API). However, the empirical evidence is still limited for some varieties of low quality and less liquid crude oil despite the systematic evolution of the world crude oil market and its high volatility. To fill this gap in the empirical literature, this paper applies EVT to examine the conditional tail behavior of Mexico´s Isthmus crude oil returns and proposes VaR and ES measures to estimate extreme losses for long and short trading positions during the period January 02, 1995-December 31, 2020.

Isthmus is a medium sour crude oil, with an API gravity of 32-33°, and 1.8% sulfur content by weight. Its quality is like Arab light crude and Russian Urals crude. These physical properties make it appropriate for the high gasoline production and intermediate distillates. Although the Isthmus crude oil production is usually used for domestic consumption. However, exports of Isthmus crude oil have increased significantly from 11.19 to 51.11 million barrels, tripling almost its share of revenues from 722 to 1910 million dollars during the period 2018-2020. Additionally, exports of Maya heavy crude fell by 14.6% in 2021, while Isthmus export sales increased by 29.2%, implying an increase in its production.

It aims at answering the following question: the use of CGARCH-EVT and FIGARCH-EVT models could provide better information of true risk to participants in the Isthmus crude oil market?

In this respect, this paper contributes to the literature in the following two ways. First, it extends the McNeil and Frey' (2000) two-stage approach by fitting a variety of two component GARCH (CGARCH) and fractional integrated GARCH (FIGARCH) models with normal innovations to capture some common characteristics such as heteroscedasticity, long memory, fat tails, and short- and long-term asymmetry of crude oil return volatility. To the best of our knowledge, this is the first study that uses CGARCH-EVT combined approaches for forecasting the VaR and ES of benchmark and Isthmus crude oils. Second, we employ the unconditional coverage test of Kupiec (2005) and the Dynamic Quantile test of Engle and Manganelli (2004) to evaluate VaR and ES in-sample and out-of-sample forecasting performance of FIGARCH-EVT and CGARCH-EVT models for both short and long trading positions in the Isthmus crude oil market. We run the backtesting over 1000 out-of-sample observations for periods ranging from January 02, 2013 to December 30, 2016, and from January 03, 2017 to December 31, 2020. This last period includes the COVID-19 crisis.

The remainder of the paper is structured as follows. Section 2 presents a brief outline of the CGARCH and FIGARCH models and estimation approaches based on the extreme value theory. Section 3 describes the data and analyzes its basic statistics traits. Section 4 applies the models and discusses the empirical results. Finally, Section 5 presents the conclusions.

2. ECONOMETRIC METHODOLOGY

2.1. Definition of Value at Risk and Expected Shortfall

VaR statistical measure abridges portfolio risk exposure as the maximum potential loss over a given time horizon at a given confidence level. Mathematically, suppose R is a random variable denoting the loss of a given financial position or portfolio. For a given probability, p VaR is defined as

$$\mathrm{VaR}\left(p\right)_{t+1} = -\sigma_{t+1}F^{-1}\left(p\right) \tag{1}$$

where σ_{t+1} is the volatility of the cumulative distribution function that describes the loss distribution of the risky financial position and F^{-1} is the inverse of the loss distribution function, i.e., p-quantile of F.

VaR is the most widely used risk measure in the financial industry. However, VaR is not a coherent risk measure since it does not satisfy the axioms of sub-additivity and convexity. To overcome these shortcomings, Artzner et al. (1999) introduce an alternative risk measure, namely expected shortfall (ES). ES is defined as the expected value of potential loss exceeding the VaR level. Mathematically, ES for risk R and given probability p can be expressed by

$$\mathrm{ES}\left(p\right)_{t+1} = \mathrm{E}\left[R\middle|R > \mathrm{VaR}\left(p\right)_{t+1}\right] \tag{2}$$

To correctly estimate VAR and ES, it is important to model and forecast oil returns volatility. In the econometric literature, there is a family of generalized autoregressive conditional heteroscedasticity (GARCH) processes introduced by Engle (1982) and Bollerslev (1986), which are widely used to describe the time-varying structure of volatility of crude oil prices. However, they are not able to capture properly the impact of asymmetric leverage effect and the long memory.

2.2. Two Component GARCH Models

Long-term dependency, i.e., long memory observed in volatility processes, has been documented in innumerable empirical studies. However, it is well known that long memory cannot be captured by traditional GARCH modeling; for this reason, it is necessary to break down volatility into two components to describe the presence of persistence in the long and short run. An alternative specification that can capture long memory or high-persistence properties in oil return series is the CGARCH model proposed by Engle and Lee (1999). This approximation allows the decomposition of conditional volatility into two components so that the decay of the persistence in the short and long run can be analyzed properly, which cannot be achieved with the traditional GARCH models (Christoffersen et al., 2008).

The CGARCH model decomposes the volatility into its transitory and permanent components. Moreover, each component allows the variance innovations to decay at a different rate. Let $\Omega_{t-1} = \{r_{T-n+1},\ldots,r_T\}$ denote the information set of all observed returns up to time $t-1$. Assuming that the dynamics of return series can be described by an AR(1) model, the mean equation is defined as follows.

$$r_t = \mu_t + \phi r_{t-1} + \sqrt{h_t}\,z_t, \qquad z_t \sim N\left(0,1\right) \tag{3}$$

where r_t denotes return at time t, μ_t is the conditional mean, ε_t is the innovation process with mean equal to 0 and conditional variance h_t, while z_t is an independent and identically distributed process with mean equal to 0 and variance equal to 1.

The conditional variance equation of AR(1)-CGARCH model can be expressed as follows:

$$h_t = q_t + \alpha\left(\varepsilon_{t-1}^2 - q_{t-1}\right) + \beta\left(h_{t-1} - q_{t-1}\right) \tag{4}$$

$$q_t = \omega + \rho\left(\varepsilon_{t-1}^2 - h_{t-1}\right) + \varphi\left(q_{t-1} + \omega\right) \tag{5}$$

where h_t and q_t indicate the transitory and permanent components of conditional volatility, respectively. Note that the conditional volatility is mean reverting around the permanent volatility (q_t). For the permanent component, the speed of mean reversion is determined by φ, value which typically lies between $(\alpha+\beta)<\varphi<1$.

Due to the presence of positive and negative shocks of the same magnitude, but with different impact on volatility, the flexibility of the CGARCH model can be extended to capture asymmetric effects in the short and long run as follows:

$$h_t = q_t + \alpha\left(\varepsilon_{t-1}^2 - q_{t-1}\right) + \gamma\left(I\left(\varepsilon_{t-1} < 0\right)\varepsilon_{t-1}^2 - q_{t-1}\right) + \beta\left(h_{t-1} - q_{t-1}\right) \tag{6}$$

$$q_t = \omega + \rho\left(\varepsilon_{t-1}^2 - h_{t-1}\right) + \psi\left(I\left(\varepsilon_{t-1} < 0\right)\varepsilon_{t-1}^2 - h_{t-1}\right) + \varphi\left(q_{t-1} + \omega\right) \tag{7}$$

where the dummy variable governed by the Heaviside function $I(\cdot)$, is equal to 1 if $\varepsilon_{t-1}<0$ and 0 otherwise. The leverage effects are observed when $\gamma>0$ and $\psi>0$.

The second alternative to capture the asymmetric effects in the short-run and long-run is defined as follows:

$$h_t = q_t + \alpha\left(\left|\varepsilon_{t-1}\right| - q_{t-1}\right) + \gamma\left(\varepsilon_{t-1} - q_{t-1}\right) + \beta\left(h_{t-1} - q_{t-1}\right) \tag{8}$$

$$q_t = \omega + \rho\left(\left|\varepsilon_{t-1}\right| - h_{t-1}\right) + \psi\left(\varepsilon_{t-1} - h_{t-1}\right) + \varphi\left(q_{t-1} - \omega\right) \tag{9}$$

where the asymmetry parameters γ and ψ are negative, by contrast with the previous asymmetric CGARCH model (ACGARCH1 and ACGARCH2). The total effect on transitory and permanent volatility components is equal to $(\alpha-\gamma)\left|\varepsilon_{t-1}\right|$ and $(\rho-\psi)\left|\varepsilon_{t-1}\right|$ if $\varepsilon_{t-1}<0$, and $(\alpha+\gamma)\left|\varepsilon_{t-1}\right|$ and $(\rho+\psi)\left|\varepsilon_{t-1}\right|$ if $\varepsilon_{t-1}>0$.

2.3. Fractional Integrated GARCH Models

Recent empirical studies have confirmed the presence of the long-range memory (persistence) characteristics for crude oil returns volatility (Kang et al., 2009; Youssef et al., 2015; Charles and Darné, 2014; Lanouar, 2016). To capture the long memory phenomenon, Baillie et al. (1996) proposed the fractionally integrated GARCH (FIGARCH) model by combining the fractionally integrated process for the mean with the standard GARCH process for the conditional variance. This model shows that the actual autocorrelations in conditional variance decay at a slow hyperbolic rate after a volatility shock.

The conditional variance of the FIGARCH model can be specified as follows:

$$h_t = \omega + \beta h_{t-1} + \left[1 - \left(1-\beta L\right)^{-1}\left(1-\varphi L\right)\left(1-L\right)^d\right]\varepsilon_t^2 \tag{10}$$

where (L) is the lag operator and $0\leq d\leq 1$, $\omega>0$, φ, $\beta>1$. The persistence of the conditional variance or the long memory degree is measured by the parameter d. $\varphi(L)$ and $\beta(L)$ are polynomials for the lag operators of orders p and q, respectively. The roots of $\varphi(L)$ and $\beta(L)$ lie outside the unit circle.

Additionally, the FIGARCH model assumes a conditional normality process and the conditional volatility symmetrically responds to the magnitude of both positive and negative shocks, which is not a desirable characteristic of a reality. To take into account asymmetric effects and the long memory property in volatility, Bollerslev and Mikkelsen (1996) extended the EGARCH model to the so-called fractionally integrated exponential GARCH model (FIEGARCH).

The FIEGARCH model is expressed as an ARMA process in terms of the logarithm of the conditional variance h_t which is modeled by,

$$ln\left(h_t\right) = \omega + \varphi\left(L\right)^{-1}\left(1-L\right)^d\left[1+\alpha\left(L\right)\right]g\left(z_{t-1}\right) \tag{11}$$

where $g\left(z_t\right) = \theta z_t + \gamma\left[\left|z_t\right| - E\left(\left|z_t\right|\right)\right]$, the first term $\left(\theta z_t\right)$ is the sign effect, and the second term $\gamma\left[\left|z_t\right| - E\left(\left|z_t\right|\right)\right]$ is the magnitude effect. Like the EGARCH model, no parameter restrictions are imposed in this model. Then, CGARCH, FIGARCH and FIEGARCH models are fitted to the raw return series $\{r_{T-n+1},\ldots,r_T\}$ by quasi-maximum likelihood, computing standardized residuals. They are calculated as follows:

$$\left\{z_{T-n+1},\ldots,z_T\right\} = \left\{\frac{\hat{r}_{T-n+1} - \hat{\mu}_{T-n+1}}{\sqrt{\hat{h}_{T-n+1}}},\ldots,\frac{\hat{r}_T - \hat{\mu}_T}{\sqrt{\hat{h}_T}}\right\} \tag{12}$$

To check the adequacy of the CGARCH and FIGARCH modeling, the standardized residual series should be free of autocorrelation and heteroscedasticity. Furthermore, EVT requires identically and independently distributed series to model the tail behavior of the innovation distribution and measure the conditional VaR and ES.

2.4. Estimation of VaR and ES Based on the EVT

Let $R_1,R_2,\ldots,R_n$ be a sequence of independent and identically distributed random variables, i.e., losses with an unknown distribution, $F(r)=P_r(R_i\leq r)$. Since the analysis is interested in modeling the losses that exceed a threshold u, the distribution function of excess losses (EVD) of $y_i=r_i-u$ whenever r_i exceeds u for $i=1,\ldots,k$ can be defined as:

$$F_u\left(y\right) = Pr\left(R-u\leq y \middle| R>u\right) = \frac{Pr\left(u<R\leq u+y\right)}{Pr\left(R>u\right)} = \frac{F\left(u+y\right)-F\left(u\right)}{1-F\left(u\right)} \tag{13}$$

For a sufficiently high threshold u the theorems of Balkema and De Haan (1974) and Pickands (1975) show that the excess distribution converges to the GPD as follows:

$$G_\xi\left(y\right) = \begin{cases} 1-\left(1+\dfrac{\xi}{\sigma}y\right)^{-1/\xi} & if \quad \xi\neq 0 \\ 1-exp\left(-\dfrac{y}{\sigma}\right) & if \quad \xi=0 \end{cases} \tag{14}$$

where ξ is the shape parameter, and $\sigma>0$ is the scale parameter. The value of the parameter ξ may be positive, negative or zero, and serves to determine fatness of the tail. When $\xi>0$ the GPD takes the form of an ordinary Pareto distribution, which is more appropriate for crude oil series analysis since they are usually characterized by fatter tails. When $\xi=0$and $\xi<0$, the GPD has the form of the Exponential and Pareto type II distribution, respectively.

Similarly, F may be defined as $F(r)=(1-F(u))G\xi(u)+F(u))$. The non-parametric estimation of $F(u)$ is determined by $(n-k)/n$, where n is the total number of observations and k is the number of observations that exceed u. Substituting the estimated value for $F(u)$ and equation (14) into $F(r)$, the tail estimator has the following expression:

$$\hat{F}(r) = 1 - \frac{k}{n}\left(1 + \hat{\xi}\frac{(r-u)}{\hat{\sigma}}\right)^{-1/\hat{\xi}} \tag{15}$$

where $\hat{\xi}$ and $\hat{\sigma}$ are maximum likelihood estimates of ξ and σ as u gets larger.

To overcome the problems of the estimation based on the unconditional EVT, this paper utilizes the set of standardized residuals, which are closer independent and identically distributed series under the CGARCH or FIGARCH hypothesis. Let $Z_{(1)} \geq Z_{(2)} \geq, \geq Z_{(n)}$ denote the order statistics and $n=k$ is the number of observations exceeding the threshold u, then Z_{k+1} is called the random threshold.

Thus, the peak over threshold (POT) approach can be applied to the tails of the innovation distribution, i.e., $Z_{(1)}-Z_{(k+1)},\ldots,Z_{(k)}-Z_{(k+1)}$. Consequently, a new tail estimator for $F(z)$ with GPD parameters can be expressed as:

$$\hat{F}_Z(z) = 1 - \frac{k}{n}\left(1 + \hat{\xi}\frac{\left(z - z_{(k+1)}\right)}{\hat{\sigma}}\right)^{-1/\hat{\xi}} \tag{16}$$

For a determined probability p, the VaR and ES for the innovations can be estimated as:

$$\widehat{\mathrm{VaR}}(Z)_{t+1}^{p} = z_{(k+1)} + \frac{\hat{\sigma}}{\hat{\xi}}\left(\left(\frac{n}{k}p\right)^{-\xi} - 1\right) \tag{17}$$

$$\widehat{\mathrm{ES}}(Z)_{t+1}^{p} = \frac{\widehat{\mathrm{VaR}}(Z)_{t+1}^{p}}{1-\hat{\xi}} + \frac{\hat{\sigma} - \hat{\xi} z_{(k+1)}}{1-\hat{\xi}} \tag{18}$$

Finally, an estimate of the conditional VaR for a 1-day horizon is defined by,

$$\mathrm{VaR}_{t+1}^{p} = \hat{\mu}_{t+1} + \sqrt{\hat{h}_{t+1}}\,\widehat{\mathrm{VaR}}(Z)_{t+1}^{p} \tag{19}$$

Similarly for a 1-day horizon, an estimate of the conditional ES is defined by,

$$\mathrm{ES}_{t+1}^{p} = \hat{\mu}_{t+1} + \sqrt{\hat{h}_{t+1}}\,\widehat{\mathrm{ES}}(Z)_{t+1}^{p} \tag{20}$$

where $\hat{\mu}_{t+1}$ and $\hat{h}_{t+1}$ are the forecasts of the conditional mean and the conditional variance at time t+1, while $\widehat{\mathrm{VaR}}(Z)_{t+1}^{p}$ and $\widehat{\mathrm{ES}}(Z)_{t+1}^{p}$ are given by Equations (17) and (18).

2.5. Backtesting the Risk Measures

According to the Basel Committee on Banking Supervision, if risk models do not provide capital requirements sufficiency to cover the realized loss, this fact is defined as a failure. All models should be evaluated with several statistical measures comparing their forecasting ability in terms of risk measurement. The backtesting procedure consists of comparing VaR and ES estimates with actual realized returns in the next period. In the literature, there are different backtesting procedures for examining the statistical accuracy of the risk models. This study applies the likelihood ratio test proposed by Kupiec (1995). This test examines whether the failure rate is statistically equal to the expected failure rate, $\alpha=1-p$, where p is the confidence level used to estimate the VaR and the ES. If T indicates the overall number of observations, then the number of failures n follows a binomial distribution with probability α.

The likelihood ratio test statistic of Kupiec is computed as,

$$LR = 2\ln\left[\left(\frac{n}{T}\right)^{n}\left(1-\frac{n}{T}\right)^{T-n}\right] - 2\ln\left[(\alpha)^{n}(1-\alpha)^{T-n}\right] \tag{21}$$

where $LR \sim \chi^2$ with one degree of freedom under the null hypothesis: $H_0 : \frac{n}{T} = \alpha$. If the value of LR is smaller than the critical values, the null hypothesis is accepted, which means that the estimation of VaR is reliable, while the alternative hypothesis is accepted if the model generate a large or small number of failures.

The Dynamic Quantile (DQ) test is also used in this study to evaluate the absolute performance of the VaR and ES forecasts. The DQ test of Engle and Manganelli (2004) is a conditional coverage test based on a linear regression model of the hit variable on a set of explanatory variables, that is, the process of hit function defined as

$$Hit_t = I(R_t < -VaR) - (1-\alpha) = \begin{cases} \alpha, & if \quad R_t < -VaR \\ -(1-\alpha), & otherwise \end{cases} \tag{22}$$

where the null hypothesis for the DQ test is that the sequence of hits is uncorrelated with its own lagged values and with any other lagged variable (including past returns, past VaR, etc.) that belongs to the information set Ω_{t-1}, and its expected value must be equal to zero.

The dynamic quantile test is a Wald test under the null hypothesis that all parameters, β_0, β_k, γ_k are zero in the regression model.

$$Hit_t = \beta_0 + \sum_{k=1}^{K}\beta_k Hit_{t-k} + \sum_{k=1}^{K}\gamma_k z_{t-k} + \varepsilon_t \tag{23}$$

where the z_{t-k} explanatory variables belong to the information set Ω_{t-1} and ε_t is a discrete i.i.d. process.

Therefore, if we denote by $\psi=(\beta_0,\beta_1,\ldots,\beta_k, \gamma_1,\ldots,\gamma_k)$ the vector of the $2K$+1 parameters of the model and by Z the matrix of explanatory variables of model. The dynamic quantile test statistic satisfies the following relation:

$$DQ = \frac{\hat{\Psi}' Z' Z \hat{\Psi}}{\alpha(1-\alpha)} \xrightarrow{d} \chi^2_{(2K+1)} \quad as\ T \to \infty \tag{24}$$

where $\hat{\Psi}$ is the OLS estimate of Ψ.

3. DATA AND DESCRIPTIVE STATISTICS

3.1. Data Description

To examine the predictive ability of VaR and ES measures based on conditional EVT, this study uses daily Isthmus crude oil prices for the period January 02, 1995 to December 31, 2020 for a total 6717 observations. The price data has been downloaded from Bloomberg, and the sample period has been selected considering available information. We examine the performance of the VaR and ES measures over two different sample periods. The first one runs from January 02, 1995 to December 31, 2016. This period is selected because the crude oil price experienced two large shocks in the second half of 2014 and early 2016. During this period of market uncertainty, the price of Isthmus crude oil fell by 77% because to the strong supply growth. The second one includes time series data from January 02, 1995 to December 31, 2020, when the impact of the first wave of the COVID-19 pandemic together with the oil price war between Saudi Arabia and Russia hammered the global crude oil market.

3.2. Preliminary Results

Most empirical analyses focus on return series rather than price series. We compute continuously compounded daily returns by taking the difference in the logarithms of two consecutive daily prices. Table 1 shows the descriptive statistics, Jarque-Bera statistic and Ljung-Box statistic for raw and squared returns. Results indicate that Isthmus crude oil returns have positive mean but relatively small in comparison with the standard deviation. This fact indicates that oil prices tend to increase over time and evidence a pronounced unconditional volatility. Returns are significantly skewed to the left for the period under scrutiny, indicating the presence of more negative outlying returns than positive ones in Isthmus time series. The return series exhibits high excess kurtosis which implies that the tails of the distribution of crude oil returns are larger and fatter than the tails of the normal distribution, due to the occurrence of atypical observations in the crude oil market. The evidence of no-normality of the distribution is also confirmed by the high significant Jarque-Bera statistic.

Moreover, the Ljung-Box $Q^2(20)$ statistics for squared returns reveals the presence of non-linear dependence and volatility clustering in the crude oil return series; periods of high volatility are followed by further high volatility and periods of low volatility are followed by further low volatility. The dynamics of returns exhibited in Figure 1 shows that the volatility clustering phenomenon became stronger during the 1997-1998 Asian financial crisis; 2000-2001 coinciding with the terrorist attack to the World Trade Center on September 11, 2001; the 2008-2009 global financial crisis; and during the outbreak of the COVID-19 pandemic on March 11, 2020. These events led to a succession of extreme downward and upward crude oil price movements within very short time spans. Consequently, findings of this preliminary analysis support the use of asymmetric CGARCH and FIGARCH models for capturing the short- and long-term asymmetry and long memory in crude oil volatility modeling.

4. EMPIRICAL RESULTS

4.1. CGARCH Model Estimates

To remove the effect of the autocorrelation and heteroscedasticity in Isthmus crude oil return series and get i.i.d. standardized residuals, first we estimate seven GARCH, CGARCH and FIGARCH specifications under normal distribution. Table 2 displays the estimates of GARCH, CGARCH and FIGARCH models, as well as the diagnostics tests for standardized residuals and squared standardized residuals. Results of the mean equation show that the estimates of c are statistically significant at 1% and 10% levels. The coefficient of the AR(1) process is negative and statistically significant at conventional levels of significance for EGARCH, ACGARCH1 and FIEGARCH models. The optimal lag length (k) in the system (3) is determined by the Hannan-Quinn information criteria. As for the conditional variance, the ARCH and GARCH coefficients, α and β, are positive and significant at the 1% level except for FIEGARCH model, indicating that the volatility is time-varying. Moreover, the sum of coefficients (α+β) is less than 1 for all of cases, confirming a high degree of persistence in the volatility and the stationarity condition for the conditional variance in the Isthmus crude oil market, particularly for GARCH, EGARCH, ACGARCH2 and FIGARCH models.

Although this fact is not supported by the estimated coefficients φ, since the values in the CGARCH and ACGARCH1 models are 0.9625 and 0.9717 against 0.7983 and 0.7574, measured by (α+β), respectively. This finding indicates that the long-run component of conditional volatility is highly persistent and dies out a hyperbolic rate of decay, absent in the short-run component. The level of statistical significance of parameters γ and ψ indicates that short-run and long-run conditional volatility exhibits an asymmetric response to positive and negative shocks at 1% and 5% levels. This result confirms that the asymmetric impact of bad news increases the conditional volatility more than good news except for the alternative ACGARCH1 specification. The values of fractionally differencing parameters d are positive and statistically significantly different from zero, indicating that the crude oil return series exhibits the long memory volatility process. The long memory property parameter varies between 0.3749 and 0.6725 for the FIGARCH and FIEGARCH specifications, respectively.

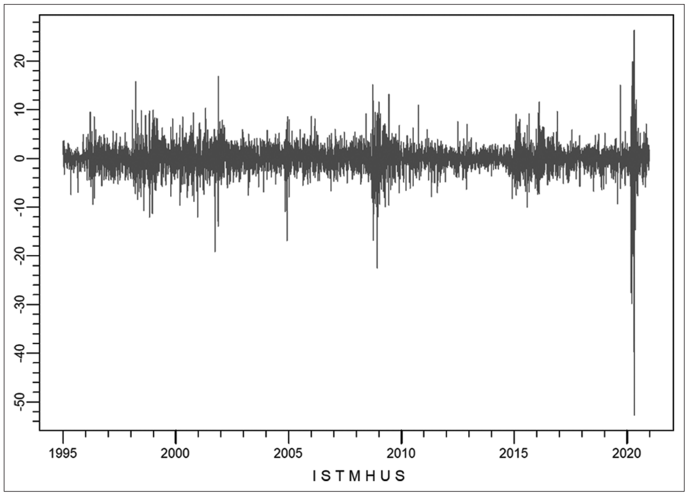
Table 1: Descriptive statistics for returns of Isthmus crude oil

Oil	Mean	Minimum	Maximum	SD	Skewness	Kurtosis	J-B test	Q^2 (20)
Isthmus	0.0171	−52.7429	26.3265	2.8546	−1.4001	38.8886	382,341*	3753*

J-B test is the Jarque-Bera normality test statistic. Q^2 (20) is the Ljung-Box statistic of the squared return series for testing serial correlation up to the 20 order. *denotes statistical significance at the 1% level

Results of the diagnosis tests for standardized residuals and squared standardized residuals are reported in the Panel B of Table 2. A comparison between the Akaike information criterion and log likelihood statistics suggests that the estimated ACGARCH2 model has the best performance to capture the usual stylized fact observed in Isthmus crude oil return series such as conditional heteroscedasticity, asymmetric effects, and persistence in the variance process. Moreover, the results of the Ljung-Box statistics up to 4^{th} and 20^{th} order in the standardized and the squared standardized residuals, confirm the absence of linear and non-linear serial correlation. Thus, the filtering procedure is effective to obtain i.i.d. residual series. However, the high value of Jarque–Bera statistics rejects the null hypothesis of a normal distribution for all standardized residual series at the 1% level. This evidence suggests using CGARCH and FIGARCH models with EVT to provide the best fit to the left and right tails of the residual distribution.

Figure 1: Evolution of Isthmus crude oil returns from January 02, 1995 to December 31, 2020



4.2. Choice of Thresholds and Estimate Parameters for the GPD

The choice of threshold is one of the main and most important tasks for identifying the tail part information before fitting the GPD to the data since this issue implies a trade-off between bias and variance. According to Coles (2001), the choice of too low thresholds might generate biased estimates because the asymptotic theory does not apply any more. Conversely, high thresholds might provide more accurate estimates with high variance due to the limited number of observations. Two tools can be used to

Table 2: Estimates of the GARCH, CGARCH and FIGARCH models for crude oil returns

	GARCH	EGARCH	CGARCH	ACGARCH1	ACGARCH2	FIGARCH	FIEGARCH
Panel A: Parameter estimation							
c	0.08876***	0.0295*	0.0725	0.0543*	0.0302	0.0838***	0.0378
	(0.0303)	(0.0307)	(0.0310)	(0.0310)	(0.0310)	(0.0303)	(0.0296)
ϕ	−0.0170	−0.0258***	−0.0249	−0.0241	−0.0256*	−0.0187	−0.0282**
	(0.0149)	(0.0139)	(0.0158)	(0.0154)	(0.0153)	(0.0156)	(0.0150)
ω	0.1433***	−0.0649***	0.0139**	0.0128***	−0.0098***	0.2534***	−0.1513***
	(0.0159)	(0.0053)	(0.0045)	(0.0040)	(0.0024)	(0.0417)	(0.0138)
α	0.0787***	0.1270***	0.0885***	0.0413***	0.0780***	0.3195***	0.2101***
	(0.0049)	(0.0084)	(0.0106)	(0.0111)	(0.0086)	(0.0368)	(0.0194)
β	0.9001***	0.9826***	0.7098***	0.7161***	0.9898***	0.5750***	0.1848**
	(0.0063)	(0.0021)	(0.0473)	(0.0441)	(0.0017)	(0.0421)	(0.0988)
ρ			0.0186***	0.0137***	0.1182***		
			(0.0041)	(0.0050)	(0.0182)		
φ			0.9625***	0.9617***	0.7791***		
			(0.0061)	(0.0062)	(0.0453)		
γ		−0.3201***		0.0759***	−0.2011**		−0.0711***
		(0.0456)		(0.0161)	(0.0835)		(0.0096)
ψ				0.01232**	−0.4396***		
				(0.0055)	(0.0772)		
δ							
d						0.3749***	0.6725***
						(0.0270)	(0.0256)
α+β	0.9788	1.1096	0.7983	0.7574	1.0678	0.8945	0.3949
Panel B: Diagnostic tests							
Log (L)	−10573	−10555	−10560	−10546	−10543	−10566	−10554
AIC	2.1156	2.1122	2.1134	2.1110	2.1105	2.1143	2.1123
SIC	2.1189	2.1160	2.1179	2.1168	2.1163	2.1182	2.1168
Q (4)	2.6716	2.3829	2.5504	2.6333	2.3586	2.4236	2.5017
	(0.6142)	(0.6657)	(0.6356)	(0.6209)	(0.6701)	(0.6584)	(0.6443)
Q (20)	9.9936	10.1636	9.9416	9.8194	9.9567	10.2877	10.6613
	(0.9683)	(0.9651)	(0.9692)	(0.9713)	(0.9690)	(0.9626)	(0.9545)
Q^2 (4)	1.5821	5.7355	1.3297	2.3370	1.2488	1.3337	1.0066
	(0.8120)	(0.2198)	(0.8563)	(0.6740)	(0.8700)	(0.8556)	(0.9088)
Q^2 (20)	16.0903	20.3941	9.0206	11.3006	11.2462	10.1849	13.0638
	(0.7110)	(0.4335)	(0.9827)	(0.9380)	(0.9396)	(0.9647)	(0.8746)
J−B test	6599 (0)	6642 (0)	6638 (0)	6634 (0)	6641 (0)	6608 (0)	6653 (0)

determine the appropriate threshold including the plot of empirical mean excess function (MEF) and the Hill plot.

In this study, the empirical MEF is employed, which is defined by

$$e_n\left(z_{(k+1)}\right) = \frac{1}{k}\sum_{i=1}^{k}\left(z_i - z_{(k+1)}\right) \tag{26}$$

where the empirical MEF is the sum of the excesses over the threshold $Z_{(k+1)}$. The MEF is a linear function of $Z_{(k+1)}$ for the GPD model, so that the empirical MEF is a straight line above the threshold $Z_{(k+1)}$. If the empirical MEF is a positively sloped straight line above certain determined threshold $Z_{(k+1)}$, it is an indication that the data follows the GPD with a positive shape parameter ξ. On the other hand, exponentially distributed data would show a horizontal MEF while short-tailed data would have a negatively sloped line.

The empirical MEF is applied directly to the positive standardized residuals z_t to choose thresholds. For negative standardized residuals, the series z_t is transformed into $-z_t$ to get the empirical MEF from those maximums. The selected threshold values, the number of exceedances and the estimated scale and shape parameters as well as their standard errors are reported in Panel A and B of Table 3 for the GARCH, EGARCH, CGARCH, ACGARCH1, ACGARCH2, FIGARCH and FIEGARCH specifications with normally distributed innovations for both negative and positive standardized residuals. According to the criterion of linearity in the empirical MEF plots, the selected thresholds are very similar for negative and positive standardized residuals under different GARCH, CGARCH and FIGARCH specifications that range from 2.00 to 2.37 for lower tail and from 2.11 to 2.33 for upper tail with several exceedances between 162 and 204 and 126 and 166 for the sample period from January 02, 1995 to December 31, 2012, respectively.

For second sample period, thresholds range from 2.00 to 2.64 for the lower tail and from 2.10 to 2.52 for the upper tail with few exceedances between 181 and 232 and 135 and 175, respectively. The number of exceedances is reasonable to model adequately the tail behavior of the distribution since it comprises between 5 and 10% of the sample, which are in line with the simulation study from McNeil and Frey (2000). The estimated scale parameter σ is positive and statistically significant at the 5% level. The σ estimates vary remarkably under different thresholds for the upper tail of standardized residuals. This finding could be associated with the magnitude and frequency of positive extreme values. For lower tail of standardized residuals, the parameter estimates take values in a narrower range, albeit they are relatively greater than the ones in upper tail except for the CGARCH-EVT, ACGARACH2-EVT and FIEGARCH-EVT models. This fact can be attributed to a greater presence of negative extreme returns and dispersion among them, particularly during the COVID-19 pandemic period.

On the other hand, the value of the shape parameter ξ is positive and statistically significant different from zero at the 5% level, suggesting that lower and upper tails of standardized residuals for Isthmus crude oil are characterized by heavy-tail distributions for

Table 3: Parameter estimates from GPD for positive and negatives standardized residuals

	GARCH-EVT		EGARCH-EVT		CGARCH-EVT		ACGARCH1-EVT		ACGARCH2-EVT		FIGARCH-EVT		FIEGARCH-EVT	
	Positives	Negatives	Positives	Negatives	Positives	Negatives	Positives	Negatives	Positives	Negatives	Positives	Negatives	Positives	Negatives
Panel A: Estimation of GPD parameters for the period from January 02, 1995 to December 31, 2012														
$z_{(k+1)}$	2.2865	2.3786	2.1646	2.1649	2.2046	2.3119	2.2248	2.1428	2.1910	2.0058	2.3376	2.2562	2.1181	2.0168
ξ	0.2302**	0.3113**	0.2787**	0.2439**	0.2587**	0.2822**	0.3808**	0.2489**	0.1748**	0.2859**	0.1738**	0.2505**	0.1504**	0.2617**
	(0.1840)	(0.1974)	(0.1806)	(0.1436)	(0.1709)	(0.1729)	(0.2007)	(0.1397)	(0.1633)	(0.1265)	(01779)	(0.1535)	(0.1431)	(0.1277)
σ	0.4952**	0.5105**	0.4519**	0.5566**	0.4670**	0.5403**	0.4004**	0.5249**	0.5283**	0.4929***	0.539**	0.5489**	0.5736**	0.5284***
	(0.1095)	(0.1149)	(0.0946)	(0.0887)	(0.0957)	(0.1101)	(0.0909)	(0.0882)	(0.1047)	(0.0747)	(0.1182)	(0.1014)	(0.1028)	(0.0815)
k	126	162	166	184	150	189	152	198	154	201	157	204	164	203
Panel B: Estimation of the GPD parameters for the period from January 2, 1995 to December 30, 2016														
$z_{(k+1)}$	2.4871	2.4504	2.2223	2.6494	2.2628	2.1739	2.2910	2.2889	2.1065	2.4522	2.5203	2.5090	2.4683	2.0053
ξ	0.2186**	0.3245**	0.2892**	0.2713**	0.2109**	0.3181**	0.2461**	0.3216**	0.1707**	0.2789**	0.2076**	0.3139**	0.2911**	0.3154**
	(0.2002)	(0.1996)	(0.1655)	(0.2343)	(0.1541)	(0.1393)	(0.1694)	(0.1654)	(0.1253)	(0.1819)	(0.1980)	(0.1931)	(0.2017)	(0.1224)
σ	0.5182**	0.5372**	0.4301**	0.6289**	0.5006**	0.4884**	0.4768**	0.5079**	0.4872**	0.5852**	0.5117**	0.5550**	0.4545**	0.4787***
	(0.1235)	(0.1217)	(0.0827)	(0.1681)	(0.0944)	(0.0806)	(0.0959)	(0.0974)	(0.0759)	(0.1256)	(0.1211)	(0.1245)	(0.1066)	(0.0695)
k	135	181	174	207	159	215	161	225	163	228	166	231	175	232

Table reports the threshold, the maximum likelihood estimates of the GPD for positives and negatives standardized residuals. Standard errors are reported in parentheses. ***, **denote significance at 1% and 5% levels two sample periods. An examination of the negative standardized residuals reveals that the shape parameter varies from 0.2439 to 0.3113 for the period 1995-2012, and from 0.2713 to 0.3245 for the period 1995-2016. Likewise, the shape parameter for the positive standardized residuals remains relatively stable and takes values from 0.1504 to 0.3808, and from 0.1707 to 0.2911. Another interesting finding is that the estimated shape parameters for negative standardized residuals are greater than the ones of positive standardized residuals under any volatility model for both sample periods except for EGARCH-EVT and ACGARCH1-EVT models. This fact implies that the lower tail is heavier and riskier than the upper tail. These findings lead to more conservative estimates of VaR and ES, particularly at extreme quantiles as shown in Panel A and B of Table 4. At quantiles of 99.5% and 99.9%, the upper tail risk of VaR and ES ranges from 2.82 to 4.21 and from 3.64 to 6.08 for the period 1995-2012, and from 2.84 to 4.00 and from 3.60 to 5.47 for the period 1995-2016, respectively. For the case of lower tail risk, estimated VaR and ES lie between 3.09 and 4.76 and between 4.15 and 6.55 for the first sample period, and between 3.03 and 4.72 and between 4.10 and 6.68 for the second sample period. These findings provide robust evidence that participants in the Isthmus crude oil market face risk of severe losses since the standardized residuals are found in the tails region of the conditional distribution, albeit results for the upper tail are not consistent with the stability of the shape parameters.

4.3. Backtesting Results of in-Sample VaR and ES Estimates

This section deals with backtesting results of the in-sample VaR and ES estimates under asymmetric CGARCH-EVT and FIGARCH-EVT approaches, which are compared to the performance of the benchmark models: GARCH-EVT approaches. To assess the accuracy of the risk measures, the in-sample VaR and ES forecasts are compared with actual returns from January 2, 2013 to December 30, 2016, and from January 3, 2017 to December 31, 2020, totaling approximately 1000 observations for each of the sample periods.

Tables 5 and 6 report the P-values of the Kupiec´s (1995) LR test and the DQ test of Engle and Manganelli (2004) for the in-sample VaR and ES estimates ranging from 95% to 99.5% and for both short and long trading positions. Forecasting performance of a model is superior to alternative models when the P > 5% significance level. For the period 1995-2016, the P-values of the LR test for the long trading position confirm that the FIGARCH-EVT approach provides the best performance to estimate in-sample VaR for Isthmus crude oil at 95% quantile. For 97.5% and 99% quantiles, backtesting results reveal that most conditional-EVT approaches perform better than the FIEGARCH-EVT model. Based on the DQ test, all conditional EVT-based approaches fail to estimate well long position risk for lower quantiles. However, the GARCH-EVT and CGARCH-EVT model families provide the most reliable VaR forecasts at the 99.5% quantile. While FIGARCH-EVT and FIEGARCH-EVT approaches perform very well at 99.9% quantile, indicating that majority of VaR violations are independently distributed.

Regarding results for the short trading position, asymmetric CGARCH-EVT models provides the better forecasting at the

Table 4: One-day-ahead in sample VaR and ES estimates for short and long trading positions

	GARCH-EVT		EGARCH-EVT		CGARCH-EVT		ACGARCH1-EVT		ACGARCH2-EVT		FIGARCH-EVT		FIEGARCH-EVT	
	Short	Long	Short	Long	Short	Long	Short	Long	Short	Long	Short	Long	Short	Long
Panel A: Estimates of the 1-day ahead VaR and ES for upper and lower tail based conditional EVT (1995-2012)														
$VaR_{0.950}$	1.7264	1.8996	1.7584	1.7474	1.7356	1.8073	1.8590	1.7745	1.6582	1.7463	1.6780	1.7858	1.5744	1.7085
$ES_{0.950}$	2.2022	2.4244	2.2280	2.3488	2.2019	2.3617	2.2807	2.3512	2.1855	2.3326	2.1916	2.3609	2.1533	2.3149
$VaR_{0.975}$	2.0016	2.1792	2.0174	2.0908	1.9980	2.1120	2.0662	2.1022	1.9788	2.0673	1.9906	2.1121	1.9338	2.0488
$ES_{0.975}$	2.5597	2.8304	2.5871	2.8030	2.5559	2.7861	2.6152	2.7875	2.5740	2.7821	2.5700	2.7962	2.5763	2.7758
$VaR_{0.990}$	2.4399	2.6547	2.4464	2.6437	2.4257	2.6180	2.4389	2.6320	2.4669	2.6020	2.4661	2.6402	2.4701	2.6046
$ES_{0.990}$	3.1290	3.5209	3.1819	3.5342	3.1329	3.4911	3.2172	3.4928	3.1656	3.5310	3.1456	3.5009	3.2075	3.5286
$VaR_{0.995}$	2.8385	3.1162	2.8520	3.1522	2.8236	3.0978	2.8212	3.1212	2.8918	3.1105	2.8796	3.1286	2.9278	3.1230
$ES_{0.995}$	3.6469	4.1910	3.7443	4.2067	3.6698	4.1596	3.8345	4.1441	3.6803	4.2431	3.6461	4.1524	3.7462	4.2308
$VaR_{0.999}$	4.0508	4.6626	4.1591	4.7238	4.0758	4.6504	4.2148	4.6421	4.1013	4.7634	4.0556	4.6497	4.1939	4.7599
$ES_{0.999}$	5.2218	6.4366	5.5564	6.2852	5.3589	6.3228	6.0851	6.1689	5.1460	6.5579	5.0695	6.1818	5.2364	6.4480
Panel B: Estimates of the 1-day ahead VaR and ES for upper and lower tail based conditional EVT (1995-2016)														
$VaR_{0.950}$	1.7845	1.8827	1.8212	1.8227	1.7101	1.7894	1.7713	1.8194	1.6901	1.7785	1.8208	1.8742	1.9182	1.7225
$ES_{0.950}$	2.2511	2.4051	2.2631	2.3779	2.1969	2.3263	2.2341	2.3454	2.1919	2.3295	2.2832	2.3927	2.3334	2.2914
$VaR_{0.975}$	2.0574	2.1570	2.0623	2.1312	1.9968	2.0733	2.0350	2.0965	1.9963	2.0823	2.0940	2.1496	2.1443	2.0242
$ES_{0.975}$	2.6004	2.8113	2.6023	2.8014	2.5601	2.7426	2.5838	2.7539	2.5612	2.7509	2.6280	2.7942	2.6523	2.7321
$VaR_{0.990}$	2.4878	2.6286	2.4650	2.6392	2.4462	2.5588	2.4603	2.5717	2.4610	2.5856	2.5210	2.6191	2.5225	2.5391
$ES_{0.990}$	3.1513	3.5094	3.1688	3.4985	3.1296	3.4546	3.1479	3.4544	3.1214	3.4488	3.1669	3.4784	3.1859	3.4842
$VaR_{0.995}$	2.8758	3.0911	2.8490	3.1167	2.8487	3.0325	2.8521	3.0367	2.8640	3.0616	2.9026	3.0756	2.8837	3.0404
$ES_{0.995}$	3.6478	4.1941	3.7089	4.1538	3.6396	4.1492	3.6677	4.1399	3.6074	4.1089	3.6484	4.1438	3.6954	4.2165
$VaR_{0.999}$	4.0394	4.6658	4.1018	4.6417	4.0447	4.6330	4.0663	4.6144	4.0059	4.5957	4.0319	4.6103	4.0650	4.7286
$ES_{0.999}$	5.1369	6.525	5.4715	6.2468	5.1553	6.4963	5.2782	6.4657	4.9844	6.2364	5.0737	6.3806	5.3617	6.6826

Table reports VaR and ES for quantiles of 95%, 97.5%, 99% and 99.9%

Table 5: Backtesting results of the in-sample VaR estimates for conditional EVT-based models

Model	Short position									
	0.950		0.975		0.990		0.995		0.999	
	LR	DQ	LR	DQ	LR	DQ	LR	DQ	LR	DQ
Panel A: In sample VaR analysis for the period from January 2, 1995 to December 31, 2012										
GARCH-EVT	0.2325**	0.0000	0.1155**	0.0000	0.0006	0.0000	0.0093	0.0000	0.3927**	0.8007**
EGARCH-EVT	0.1783**	0.0000	0.1155**	0.0000	0.0013	0.0000	0.0236	0.0000	0.3927**	0.8007**
CGARCH-EVT	0.2325**	0.0000	0.2134**	0.0000	0.0006	0.0000	0.0004	0.0000	0.3927**	0.8007**
ACGARCH1-EVT	0.1783**	0.0000	0.2134**	0.0000	0.0006	0.0000	0.0001	0.0000	0.9825**	0.9994**
ACGARCH2-EVT	0.4568**	0.0000	0.2134**	0.0000	0.0013	0.0000	0.0553	0.0000	0.9825**	0.9994**
FIGARCH-EVT	0.3318**	0.0000	0.0025	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000
FIEGARCH-EVT	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Panel B: In sample VaR analysis for the period from January 2, 1995 to December 30, 2016										
GARCH-EVT	0.0003	0.0000	0.4664**	0.0000	0.0932**	0.0000	0.0011	0.0000	0.0000	0.0000
EGARCH-EVT	0.0006	0.0000	0.9119**	0.0000	0.0932**	0.0000	0.0004	0.0000	0.0000	0.0000
CGARCH-EVT	0.0019	0.0000	0.7537**	0.0000	0.0932**	0.0000	0.0004	0.0000	0.0000	0.0000
ACGARCH1-EVT	0.0493	0.0000	0.1694**	0.0000	0.1600**	0.0000	0.0011	0.0000	0.0000	0.0000
ACGARCH2-EVT	0.2970**	0.0000	0.1528**	0.0000	0.0932**	0.0000	0.0004	0.0000	0.0000	0.0000
FIGARCH-EVT	0.2970**	0.0000	0.9119**	0.0000	0.0932**	0.0000	0.0004	0.0000	0.0000	0.0000
FIEGARCH-EVT	0.1338**	0.0000	0.2904**	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Model	**Long position**									
	0.950		0.975		0.990		0.995		0.999	
	LR	DQ	LR	DQ	LR	DQ	LR	DQ	LR	DQ
Panel A: In sample VaR analysis for the period from January 2, 1995 to December 31, 2012										
GARCH-EVT	0.0010	0.0000	0.3471**	0.0000	0.4016**	0.0000	0.3104**	0.2276**	0.0000	0.0000
EGARCH-EVT	0.0336	0.0000	0.4664**	0.0000	0.4016**	0.0000	0.3104**	0.2276**	0.0000	0.0000
CGARCH-EVT	0.0223	0.0000	0.4664**	0.0000	0.1600**	0.0000	0.9609**	0.0882**	0.0000	0.0000
ACGARCH1-EVT	0.0704**	0.0001	0.7537**	0.0000	0.1600**	0.0000	0.9609**	0.0882**	0.0000	0.0000
ACGARCH2-EVT	0.0704**	0.0001	0.4664**	0.0000	0.1600**	0.0000	0.6087**	0.1827**	0.0000	0.0000
FIGARCH-EVT	0.7862**	0.0000	0.2004**	0.0000	0.0136	0.0000	0.0236	0.0000	0.9825**	0.0915**
FIEGARCH-EVT	0.0005	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.9825**	0.0915**
Panel B: In sample VaR analysis for the period from January 2, 1995 to December 30, 2016										
GARCH-EVT	0.0001	0.0000	0.1104**	0.0001	0.4016**	0.0000	0.0236	0.0000	0.0000	0.0000
EGARCH-EVT	0.0033	0.0000	0.7537**	0.0000	0.2605**	0.0000	0.0236	0.0000	0.0000	0.0000
CGARCH-EVT	0.3719**	0.0000	0.4664**	0.0000	0.2605**	0.0000	0.0236	0.0000	0.0000	0.0000
ACGARCH1-EVT	0.8975**	0.0000	0.0400	0.0000	0.8086**	0.0000	0.0236	0.0000	0.0000	0.0000
ACGARCH2-EVT	0.3318**	0.0000	0.7537**	0.0000	0.2605**	0.0000	0.0236	0.0000	0.0000	0.0000
FIGARCH-EVT	0.0988**	0.0000	0.4664**	0.0000	0.2605**	0.0000	0.0236	0.0000	0.0000	0.0000
FIEGARCH-EVT	0.0743**	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table reports the P-values of the Kupiec test and the DQ test. Values in bold denote preferred model with the best forecasting performance. ** denotes significance at the 5% level

99.9% quantile under both LR and DQ tests. For the period 1995-2016, backtesting results reveal that none of the conditional EVT-based models does not perform well especially at extreme quantiles, where they all do not pass successfully the LR and DQ tests for both short and long trading positions since VaR estimates are too conservative. The empirical evidence is not consistent with respect to the results of Youssef et al. (2015). This finding may be explained by the study period during the COVID-19 pandemic and by the presence of extreme price fluctuations in the crude oil market.

For the ES figure, backtesting results are reported in Table 6. The evidence reveals that all conditional EVT-based models fail to estimate the tail risk of the long trading position at any quantile except FIGARCH-EVT and FIEGARCH-EVT models for the 99.5% level, where they only jointly pass the LR and the DQ tests. Nevertheless, GARCH-EVT, ACGARCH2-EVT and FIGARCH-EVT approaches improve the forecasting performance for the short trading position at the 99.9% quantile. In relation to the forecasting performance, findings evidence that conditional EVT-based models considering stylized facts such as volatility clustering, fat tails, long-run asymmetry and long-run persistence in the Isthmus crude oil series behavior improve the risk forecasts for short and long trading positions.

4.4. Backtesting for Out-of-Sample VaR and ES Forecasts

This section reports the out-of-sample forecasting performance of conditional EVT-based models over the periods from January 02, 2013 to December 30, 2016, and from January 03, 2017 to December 31, 2020.

Backtesting results of out-of-sample VaR forecasts for short and long trading positions are provided in Table 7. It is worth noting that the out-of-sample forecasts are more robust than those of the in-sample analysis in terms of P-values that are highly acceptable. For the 2013-2016 period, both P-values show that the performance of the FIGARCH-EVT model is superior to the alternative models followed by ACGARCH2-EVT approach for the long trading position at any quantile. For the short trading position, asymmetric CGARCH-EVT models perform well estimating the upper tail risk in the Isthmus crude oil return series

Table 6: Backtesting results of the in-sample ES estimates for conditional EVT-based models

Model	Short position									
	0.950		0.975		0.990		0.995		0.999	
	LR	DQ	LR	DQ	LR	DQ	LR	DQ	LR	DQ
Panel A: In sample ES analysis for the period from January 2, 1995 to December 31, 2012										
GARCH-EVT	0.0033	0.0000	0.6033**	0.0000	0.9446**	0.0000	0.2367**	0.0000	0.9825**	0.9994**
EGARCH-EVT	0.0033	0.0000	0.6033**	0.0000	0.6954**	0.0000	0.4275**	0.0000	0.0000	0.0000
CGARCH-EVT	0.0000	0.0000	0.6033**	0.0000	0.9446**	0.0000	0.2367**	0.0000	0.0000	0.0000
ACGARCH1-EVT	0.0033	0.0000	0.6033**	0.0000	0.9446**	0.0000	0.4275**	0.0000	0.0000	0.0000
ACGARCH2-EVT	0.0033	0.0000	0.6033**	0.0000	0.9446**	0.0000	0.4275**	0.0000	0.9825**	0.9994**
FIGARCH-EVT	0.1783**	0.0000	0.2134**	0.0000	0.0516	0.0000	0.0553	0.0000	0.9825**	0.9994**
FIEGARCH-EVT	0.6799**	0.0000	0.0122	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000
Panel B: In sample ES analysis for the period from January 2, 1995 to December 30, 2016										
GARCH-EVT	0.0000	0.0000	0.0221	0.0000	0.4016**	0.0000	0.0236	0.0000	0.0000	0.0000
EGARCH-EVT	0.0000	0.0000	0.0401	0.0000	0.2605**	0.0000	0.0093	0.0000	0.0000	0.0000
CGARCH-EVT	0.0003	0.0000	0.2479**	0.0000	0.1600**	0.0000	0.0011	0.0000	0.0000	0.0000
ACGARCH1-EVT	0.0019	0.0000	0.9119**	0.0000	0.0932**	0.0000	0.0003	0.0000	0.0000	0.0000
ACGARCH2-EVT	0.0019	0.0000	0.6287**	0.0000	0.0272	0.0000	0.0003	0.0000	0.0000	0.0000
FIGARCH-EVT	0.0033	0.0000	0.4984**	0.0000	0.0065	0.0000	0.0001	0.0000	0.0000	0.0000
FIEGARCH-EVT	0.0141	0.0000	0.2904**	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
Model	**Long position**									
	0.950		0.975		0.990		0.995		0.999	
	LR	DQ	LR	DQ	LR	DQ	LR	DQ	LR	DQ
Panel A: In sample ES analysis for the period from January 2, 1995 to December 31, 2012										
GARCH-EVT	0.0000	0.0000	0.0055	0.0000	0.0015	0.3675**	0.0257	0.0882**	0.0000	0.0000
EGARCH-EVT	0.0000	0.0000	0.0055	0.0000	0.0015	0.3675**	0.0257	0.5668**	0.0000	0.0000
CGARCH-EVT	0.0000	0.0000	0.0221	0.0000	0.0015	0.3675**	0.0257	0.5668**	0.0000	0.0000
ACGARCH1-EVT	0.0000	0.0000	0.0221	0.0000	0.0075	0.2261**	0.0257	0.5668**	0.0000	0.0000
ACGARCH2-EVT	0.0000	0.0001	0.0221	0.0000	0.0015	0.3675**	0.0257	0.5668**	0.0000	0.0000
FIGARCH-EVT	0.0001	0.0000	0.2479**	0.0000	0.4681**	0.0000	0.1155**	0.7990**	0.0000	0.0000
FIEGARCH-EVT	0.4568**	0.0000	0.9119**	0.0000	0.1600**	0.0000	0.9609**	0.5668**	0.0000	0.0000
Panel B: In sample ES analysis for the period from January 2, 1995 to December 30, 2016										
GARCH-EVT	0.0000	0.0000	0.0010	0.0000	0.9446**	0.0000	0.1195**	0.0000	0.0008	0.0000
EGARCH-EVT	0.0000	0.0000	0.0055	0.0000	0.8086**	0.0000	0.1195**	0.0000	0.0008	0.0000
CGARCH-EVT	0.0001	0.0000	0.0684	0.0000	0.8086**	0.0000	0.0553	0.0000	0.0001	0.0000
ACGARCH1-EVT	0.0055	0.0000	0.6033**	0.0000	0.2605**	0.0000	0.0236	0.0000	0.0000	0.0000
ACGARCH2-EVT	0.0091	0.0000	0.7537**	0.0000	0.2605**	0.0000	0.0236	0.0000	0.0000	0.0000
FIGARCH-EVT	0.0493	0.0000	0.4984**	0.0000	0.0932**	0.0000	0.0034	0.0000	0.0000	0.0000
FIEGARCH-EVT	0.8741**	0.0000	0.0198	0.0000	0.0030	0.0000	0.0011	0.0000	0.0000	0.0000

Table reports the P-values of the Kupiec test and the DQ test. Values in bold denote preferred model with the best forecasting performance. **denotes significance at the 5% level

Table 7: Backtest results of the out-of-sample VaR estimates for conditional EVT-based models

Model	Short position									
	0.950		0.975		0.990		0.995		0.999	
	LR	DQ	LR	DQ	LR	DQ	LR	DQ	LR	DQ
Panel A: Out of sample VaR analysis for the period from January 2, 2013 to December 30, 2016										
GARCH-EVT	0.7615**	0.0000	0.1528**	0.0000	0.2605**	0.0000	0.2367**	0.0000	0.9825**	0.9919**
EGARCH-EVT	0.0982**	0.0001	0.1067**	0.0000	0.2831**	0.0003	0.1155**	0.0000	0.9825**	0.9934**
CGARCH-EVT	0.6532**	0.0000	0.1067**	0.0000	0.1600**	0.0000	0.0553	0.0000	0.9825**	0.9922**
ACGARCH1-EVT	0.6532**	0.0002	0.1067**	0.0000	0.4016**	0.0000	0.6087**	0.9581**	0.9825**	0.9898**
ACGARCH2-EVT	0.9825**	0.0001	0.3853**	0.0000	0.6954**	0.0000	0.1155**	0.9612**	0.9825**	0.9948**
FIGARCH-EVT	0.0988**	0.0006	0.0002	0.0000	0.0033	0.0000	0.0553	0.0003	0.3927**	0.9896**
FIEGARCH-EVT	0.0003	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.1132**	0.2111**
Panel B: Out of sample VaR analysis for the period from January 3, 2017 to December 31, 2020										
GARCH-EVT	0.0010	0.0384	0.0114	0.5621**	0.0685**	0.8894**	0.3104**	0.9912**	0.9825**	0.9999**
EGARCH-EVT	0.0005	0.0013	0.0401	0.1868**	0.0685**	0.8776**	0.3104**	0.9890**	0.9825**	0.9999**
CGARCH-EVT	0.0010	0.0774**	0.1684**	0.7834**	0.2831**	0.9998**	0.3104**	0.9903**	0.9825**	0.9999**
ACGARCH1-EVT	0.0010	0.0635**	0.0221	0.6477**	0.0257	0.7677**	0.6087**	0.9967**	0.9825**	0.9999**
ACGARCH2-EVT	0.0036	0.0072	0.0055	0.4612**	0.4681**	0.9923**	0.6087**	0.9963**	0.9825**	0.9999**
FIGARCH-EVT	0.0010	0.0774**	0.1684**	0.7612**	0.6974**	0.9991**	0.9609**	0.9945**	0.9825**	0.9999**
FIEGARCH-EVT	0.0144	0.0000	0.0043	0.0000	0.0136	0.0000	0.0011	0.0000	0.0258	0.0000
Model	**Long position**									
	0.950		0.975		0.990		0.995		0.999	
	LR	DQ	LR	DQ	LR	DQ	LR	DQ	LR	DQ
Panel A: Out of sample VaR analysis for the period from January 2, 2013 to December 30, 2016										
GARCH-EVT	0.0982**	0.0000	0.1104**	0.0202	0.2831**	0.0013	0.1155**	0.6970**	0.9825**	0.9969**
EGARCH-EVT	0.2325**	0.0000	0.0684**	0.0012	0.1505**	0.0009	0.1155**	0.7303**	0.9825**	0.9987**

(Contd...)

Table 7: ***(Continued)***

Model	Long position									
	0.950		0.975		0.990		0.995		0.999	
	LR	DQ	LR	DQ	LR	DQ	LR	DQ	LR	DQ
CGARCH-EVT	0.1673**	0.0003	0.7537**	0.0214	0.4681**	0.0025	0.3104**	0.0000	0.9825**	0.9949**
ACGARCH1-EVT	0.3719**	0.0544	0.3471**	0.0112	0.2831**	0.0002	0.3104**	0.1274**	0.9825**	0.9949**
ACGARCH2-EVT	0.2325**	0.0645**	0.0684**	0.0014	0.2831**	0.0003	0.1155**	0.7190**	0.9825**	0.9964**
FIGARCH-EVT	0.5801**	0.1590**	0.9119**	0.4312**	0.2605**	0.2055**	0.9609**	0.1287**	0.3927**	0.1673**
FIEGARCH-EVT	0.0000	0.0000	0.0000	0.0000	0.0272	0.0000	0.4275**	0.0000	0.3927**	0.0601**
Panel B: Out of sample VaR analysis for the period from January 3, 2017 to December 31, 2020										
GARCH-EVT	0.0553	0.0000	0.2134**	0.3426**	0.0516	0.0630	0.0553	0.0000	0.1132**	0.6931**
EGARCH-EVT	0.0551	0.0034	0.0313	0.0013	0.0932**	0.0576	0.0236	0.0000	0.0048	0.0000
CGARCH-EVT	0.9895**	0.0319	0.1067**	0.6963**	0.0065	0.0310	0.0236	0.0000	0.1132**	0.7266**
ACGARCH1-EVT	0.8741**	0.9112**	0.2134**	0.7311**	0.1250**	0.1016**	0.1719**	0.7862**	0.3927**	0.9812**
ACGARCH2-EVT	0.6799**	0.5549**	0.0313	0.0841**	0.0272	0.0003	0.0236	0.0000	0.0048	0.0000
FIGARCH-EVT	0.5801**	0.5284**	0.0074	0.1993**	0.0005	0.0003	0.0011	0.0000	0.0007	0.0000
FIEGARCH-EVT	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table reports the P-values of the Kupiec test and the DQ test. Values in bold denote preferred model with the best forecasting performance. ** denotes significance at 5% level

Table 8: Backtesting results of the out-of-sample ES estimates for conditional EVT-based models

Model	Short position									
	0.950		0.975		0.990		0.995		0.999	
	LR	DQ	LR	DQ	LR	DQ	LR	DQ	LR	DQ
Panel A: Out sample ES analysis for the period from January 2, 2013 to December 30, 2016										
GARCH-EVT	0.0000	0.0000	0.0010	0.0000	0.0015	0.4624**	0.0257	0.8317**	0.0000	0.0000
EGARCH-EVT	0.0000	0.0000	0.0000	0.0002	0.0001	0.2933**	0.0257	0.8334**	0.0000	0.0000
CGARCH-EVT	0.0000	0.0000	0.0010	0.0000	0.0075	0.5082**	0.1155**	0.9600**	0.0000	0.0000
ACGARCH1-EVT	0.0001	0.0000	0.0010	0.0000	0.0015	0.4624**	0.1155*	0.9607**	0.9825**	0.9799**
ACGARCH2-EVT	0.0000	0.0000	0.0010	0.0000	0.0015	0.4624**	0.0237	0.8351**	0.0000	0.0000
FIGARCH-EVT	0.0055	0.0063	0.1694**	0.0123	0.0015	0.0081	0.6087**	0.9994**	0.0000	0.0000
FIEGARCH-EVT	0.7615**	0.0000	0.7383**	0.0000	0.5860**	0.0003	0.4275**	0.0000	0.0000	0.0000
Panel B: Out of sample ES analysis for the period from January 2, 1995 to December 30, 2020										
GARCH-EVT	0.0000	0.0000	0.0000	0.0048	0.0075	0.6254**	0.0257	0.8517**	0.9825**	0.9999**
EGARCH-EVT	0.0000	0.0000	0.0000	0.0092	0.0015	0.4545**	0.1125**	0.9522**	0.9825**	0.9999**
CGARCH-EVT	0.0000	0.0000	0.0000	0.0171	0.0075	0.6239**	0.0257	0.8525**	0.9825**	0.9999**
ACGARCH1-EVT	0.0000	0.0000	0.0000	0.0092	0.0075	0.6252**	0.0257	0.8552**	0.9825**	0.9999**
ACGARCH2-EVT	0.0000	0.0000	0.0000	0.0175	0.0015	0.4555**	0.0257	0.8530**	0.9825**	0.9999**
FIGARCH-EVT	0.0000	0.0000	0.0000	0.0322	0.0075	0.6209**	0.0257	0.8508**	0.9825**	0.9999**
FIEGARCH-EVT	0.0000	0.0000	0.0684**	0.0000	0.8086**	0.0000	0.7009**	0.0000	0.3927**	0.4655**
Model	**Long position**									
	0.950		**0.975**		**0.990**		**0.995**		**0.999**	
	LR	**DQ**	**LR**	**DQ**	**LR**	**DQ**	**LR**	**DQ**	**LR**	**DQ**
Panel A: Out sample ES analysis for the period from January 2, 2013 to December 30, 2016										
GARCH-EVT	0.0000	0.0000	0.0000	0.0003	0.0016	0.3277**	0.1155**	0.6956**	0.9825**	0.9955**
EGARCH-EVT	0.0000	0.0000	0.0000	0.0042	0.0016	0.3403**	0.0250	0.8425**	0.0000	0.0000
CGARCH-EVT	0.0000	0.0000	0.0000	0.0017	0.0016	0.3228**	0.1155**	0.6823**	0.9825**	0.9934**
ACGARCH1-EVT	0.0000	0.0000	0.0000	0.0009	0.0016	0.3341**	0.1155**	0.7139**	0.9825**	0.9948**
ACGARCH2-EVT	0.0000	0.0000	0.0000	0.0000	0.0016	0.3370**	0.0831**	0.8374**	0.9825**	0.9971**
FIGARCH-EVT	0.0000	0.0000	0.0001	0.0453	0.0016	0.3091**	0.1155**	0.6438**	0.9825**	0.9990**
FIEGARCH-EVT	0.0091	0.0004	0.0685**	0.0000	0.0685**	0.0000	0.1155**	0.7317**	0.9825**	0.9998**
Panel B: Out of sample ES analysis for the period from January 2, 1995 to December 30, 2020										
GARCH-EVT	0.0000	0.0301	0.0401	0.3773**	0.4681**	0.0025	0.9609**	0.0000	0.9825**	0.9989**
EGARCH-EVT	0.0000	0.0431	0.0221	0.2436**	0.6954**	0.0024	0.7009**	0.2254**	0.9825**	0.9990**
CGARCH-EVT	0.0000	0.0425	0.0221	0.3393**	0.2831**	0.0006	0.7009**	0.0000	0.9825**	0.9989**
ACGARCH1-EVT	0.0000	0.0205	0.0114	0.2126**	0.1505**	0.0000	0.6087**	0.8656**	0.0000	0.0000
ACGARCH2-EVT	0.0001	0.0423	0.0401	0.0983**	0.4681**	0.0007	0.9609**	0.1315**	0.9825**	0.9972**
FIGARCH-EVT	0.0011	0.1453**	0.3471**	0.7995**	0.9446**	0.1124**	0.4275**	0.0791**	0.3929**	0.9856**
FIEGARCH-EVT	0.3719**	0.6419**	0.7774**	0.1033**	0.0516	0.0000	0.0236	0.0000	0.0007	0.0000

Table reports the P-values of the Kupiec test and the DQ test. Values in bold denote preferred model with the best forecasting performance. **denotes significance at 5% level

for extreme quantiles according to the LR and DQ tests. Despite there are several conditional EVT-based models performing well at highest quantiles. However, it should mention that the FIEGARCH-EVT model provides poor forecasting to estimate the extreme risk in Isthmus crude oil market.

For out-of-sample analysis over the 2017-2020 period, the ACGARCH1-EVT model provides the most stable and highest P-values of the LR and DQ tests for the long trading position under any confidence level. For the short trading position, the predictive power of FIGARCH and CGARCH-EVT approaches is statistically superior estimating the out-of-sample VaR under any quantile. In fact, it is worth mentioning that several conditional EVT-based models are highly effective to estimate the upper tail risk at quantiles >97.5%. However, the FIEGARCH-EVT model continues to provide the worst performance for both short and long trading positions. Its performance is even worse than GARCH-EVT and EGARCH-EVT models under the LR and DQ tests. These findings no are in line with the results of Youssef et al. (2015) and Zhao et al. (2019). This result may be explained by the different analysis periods.

Table 8 reports backtesting results of out-of-sample ES forecasting for short and long trading positions. For the 2013-2016 period, results reveal that the forecasting performance of the most conditional EVT-based models improves to estimate the tail risk of short and long trading positions for quantiles over than 97.5%, especially under the DQ test. For the 2017-2020 period, the results are similar for the short trading position. However, the P-values of the LR and DQ tests show that the FIGARCH-EVT model still performs well in predicting the out of sample ES for the long trading position under any confidence level. Finally, our findings are useful for risk managers and investors to improve risk management and hedging strategies when the Isthmus oil market volatility increases in unstable economic and financial periods. Thus, we show the importance of stylized volatility features such as asymmetry, volatility clustering, long range memory and fat tails for measuring tail risk based on VaR and ES estimates, particularly during the COVID-19 crisis.

5. CONCLUSION

Crude oil returns exhibit the presence of fat tails, heteroscedasticity, asymmetry, and long memory. Accordingly, this paper adopts CGARCH and FIGARCH models under normal distribution to capture mayor stylized facts in the return volatility of Isthmus crude oil market. Additionally, McNeil and Frey (2000) approach based on extreme value theory is extended to evaluate and improve the in-sample and out-of-sample VaR and ES estimates for the short and long trading positions. Backtesting results provide robust evidence that the FIGARCH-EVT model provides the most accurate out-of-sample VaR forecasts for both short and long trading positions, followed by CGARCH-EVT and ACGARCH1-EVT models. Still the most symmetric and asymmetric CGARCH models along with conventional GARCH models only perform well at extreme quantiles. The FIGARCH-EVT model still performs well in estimating the ES. This result suggests that the ES should be considered as a sound alternative risk measure to overcome the deficiencies of VaR. Our findings have important implications for producers, consumers, portfolio investors, and policy makers. Since the conditional EVT-based models can capture the heteroscedasticity, asymmetry, and long memory, as well as explicitly modeling the tail behavior of distribution, they could improve risk assessment and management. In addition, an accurate estimation of risk level will allow to develop more effective hedging strategies for diminishing oil price risk exposure from Isthmus crude oil market participants, particularly the Mexican Federal Government which depends on the high crude oil prices to maintain sound public finance.

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(5), 2326-2339.

Artzner, P., Delbaen, F., Eber, J.M., Heath, D. (1999), Coherent measures of risk. Mathematical Finance, 9(3), 203-228.

Baillie, R.T. (1996), Long memory processes and fractional integration in econometrics. Journal of Econometrics, 73(1), 5-59.

Baillie, R.T., Bollerslev, T., Mikkelsen, H.O. (1996), Fractionally integrated generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 74(1), 3-30.

Bali, T., Theodossiou, P. (2007), A conditional-SGT-VaR approach with alternative GARCH models. Annals Operations Research, 151(1), 241-267.

Balkema, A.A., de Haan, L. (1974), Residual lifetime at great age. Annals of Probability, 2(5), 792-804.

Bollerslev, T. (1986), Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307-327.

Bollerslev, T., Mikkelsen, H.O. (1996), Modeling and pricing long memory in stock market volatility. Journal of Econometrics, 73(1), 151-184.

Boussena, S., Locatelli, C. (2005), Towards a more coherent oil policy in Russia? OPEC Energy Review, 29(2), 85-105.

Charles, A., Darné, O. (2014), Volatility persistence in crude oil markets. Energy Policy, 65, 729-742.

Chiu, Y.C., Chuang, Y., Lai, J.Y. (2010), The performance of composite forecast models of value-at-risk in the energy market. Energy Economics, 32(2), 423-431.

Choi, K., Hammoudeh, S. (2009), Long memory in oil and refined products markets. The Energy Journal, 30(2), 97-116.

Christoffersen, P., Jacobs, K., Ornthanalai, C., Wang, Y. (2008), Option valuation with long run and short-run volatility components. Journal of Financial Economics, 90(3), 272-297.

Coles, S. (2001), An Introduction to Statistical Modeling of Extreme Values. Berlin: Springer-Verlag.

Costello, A., Asem, E., Gardner, E. (2008), Comparison of historically simulated VaR: Evidence from oil prices. Energy Economics, 30(5), 2154-2166.

Echaust, K., Just, M. (2020), Value at risk estimation using the GARCH-EVT approach with optimal tail selection. Mathematics, 8(114), 1-22.

Elder, J., Serletis, A. (2008), Long memory in energy futures prices. Review of Financial Economics, 17(2), 146-155.

Engle, R.F. (1982), Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica, 50(4), 987-1007.

Engle, R.F., Lee, G.G. (1999), A permanent and transitory component model of stock return volatility. In: Engle, R., White, H., editors. Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W.J. Granger. Oxford: Oxford University Press. p475-497.

Engle, R.F., Manganelli, S. (2004), CAViaR: Conditional autoregressive value at risk by regression quantiles. Journal of Business and Economic Statistics, 22(4), 367-381.

Ghorbel, A., Trabelsi, A. (2014), Energy portfolio risk management using time-varying extreme value copula methods. Economic Modelling, 38(2), 470-485.

Jammazi, R., Nguyen D.K. (2017), Estimating and forecasting portfolio's value-at-risk with wavelet-based extreme value theory: Evidence from crude oil prices and US exchange rates. Journal of the Operational Research Society, 68(11), 1352-1362.

Kang, S.H., Kang, S.M., Yoon, S.M. (2009), Forecasting volatility of crude oil markets. Energy Economics, 31(1), 119-125.

Klein, T., Walther, T. (2016), Oil price volatility forecast with mixture memory GARCH. Energy Economics, 58, 46-58.

Krehbiel, T., Adkins, L.C. (2005), Price risk in the NYMEX energy complex: An extreme value approach. Journal of Futures Markets, 25(4), 309-337.

Kupiec, P.H. (1995), Techniques for verifying the accuracy of risk measurement models. Journal of Derivatives, 3(2), 73-84.

Lanouar, C. (2016), Breaks or long range dependence in the energy futures volatility: Out-of-sample forecasting and VaR analysis. Economic Modeling, 53, 354-374.

Marimoutou, V., Raggad, B., Trabelsi, A. (2009), Extreme value theory and value at risk: Application to oil market. Energy Economics, 31(4), 519-530.

McNeil, A.J., Frey, R. (2000), Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. Journal of Empirical Finance, 7(3), 271-300.

Mi, Z.F., Wei, Y.M., Tang, B.J., Cong, R.G., Yu, H., Cao, H., Guan, D. (2017), Risk assessment of oil price from static and dynamic modelling approaches. Applied Economics, 49(9), 929-939.

Pickands, J. (1975), Statistical inference using extreme order statistics. Annals of Statistics, 3(1), 119-131.

Ren, F., Giles, D.E. (2010), Extreme value analysis of daily Canadian crude oil prices. Applied Financial Economics, 20(12), 941-954.

Riedle, T. (2018), Using market BuVaR as countercyclical value at risk approach to account for the risks of stock market crashes. The Quarterly Review of Economics and Finance, 69, 308-321.

Sowdagur, V., Narsoo, J. (2017), Forecasting value-at-risk using GARCH and extreme-value-theory approaches for daily returns. International Journal of Statistics and Applications, 7(2), 137-151.

Susan, M., Waititu, A.G. (2015), Modelling oil price risk. American Journal of Theoretical and Applied Statistics, 4(6), 539-546.

Toumi, K., Vivani, J.L., Chayeh, Z. (2019), Measurement of the displaced commercial risk in Islamic banks. The Quarterly Review of Economics and Finance, 74, 18-31.

Weru, S.K., Waititu, A., Ngunyi, A. (2019), Modelling energy market volatility using GARCH models and estimating value at risk. Journal of Statistics and Actuarial Research, 2(1), 1-32.

Youssef, M., Belkacem, L., Mokni, K. (2015), Value-at-risk estimation of energy commodities: A long-memory GARCH-EVT approach. Energy Economics, 51, 99-110.

Zhao, L.T., Liu, K., Duan, X.L., Li, M.F. (2019), Oil price risk evaluation using a novel hybrid model based on time-varying long memory. Energy Economics, 81, 70-78.

Zikovic, S. (2011), Measuring risk of crude oil at extreme quantiles. Journal of Economics and Business, 29(1), 9-31.