Extreme energy prices forecasting: Application of alpha-recurrent neural network with generalised Pareto distribution

Check for updates

AUTHORS:
Kelebogile Bantsi[1]
Katleho Makatjane[1,*]

AFFILIATIONS:
[1]Department of Statistics, University of Botswana, Gaborone, Botswana
[*]Current: Department of Statistics and Population Studies, University of the Western Cape, Cape Town, South Africa

CORRESPONDENCE TO:
Katleho Makatjane

EMAIL:
kmakatjane@uwc.ac.za

DATES:
Received: 04 July 2025
Revised: 06 Nov. 2025
Accepted: 04 Dec. 2025
Published: 26 Mar. 2026

HOW TO CITE:
Bantsi K, Makatjane K. Extreme energy prices forecasting: Application of alpha-recurrent neural network with generalised Pareto distribution. S Afr J Sci. 2026;122(3/4), Art. #22326. https://doi.org/10.17159/sajs.2026/22326

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☒ Open data set
☐ All data included
☐ On request from author(s)
☐ Not available
☐ Not applicable

EDITOR:
Stefan Lotz

KEYWORDS:
asymmetric risk, crude oil, deep learning, financial time series, tail risk, volatility

FUNDING:
None

Tail risk assessment is crucial in financial markets, especially for commodities such as Brent crude oil, where extreme price fluctuations pose a risk for investors and policymakers. Risk models such as generalised autoregressive conditional heteroscedasticity (GARCH) often struggle to capture these extreme movements accurately, leading to potential underestimation of risk exposure. To solve this problem, we combine an alpha-recurrent neural network with a generalised Pareto distribution to better predict extreme price changes and improve tail risk estimation. Our findings demonstrate that this approach effectively captures downside risk, with backtesting results yielding high *p*-values, confirming its statistical reliability. The results from the shape parameter reveal that losses in crude oil markets are significantly riskier than gains, highlighting the asymmetric nature of price movements. Risk estimates indicate that the model provides robust assessments for both long- and short-term trading positions, making it a valuable tool for risk management. Nevertheless, these results have broader implications for financial risk modelling, particularly in commodity markets, where macroeconomic and geopolitical factors influence price volatility. Future work should focus on expanding the data set, enhancing computational efficiency and adding external risk factors such as liquidity constraints and regulatory shifts. Better calibration methods and the ability to adjust in real-time can make predictions more accurate, helping risk assessment models stay useful in changing market conditions.

Significance:

This research presents a new application of alpha-recurrent neural networks combined with the generalised Pareto distribution to assess tail risk in oil markets. The study uses a weighted nonlinear least squares likelihood moment for parameter estimation, which ensures robust and efficient modelling of extreme price movements. The findings show that losses are significantly more risky than gains and provide critical information for financial risk management. The accuracy of the model is verified by rigorous backtesting to confirm its reliability. Moreover, this approach offers methodological progress that is applicable beyond oil markets, including equity and crypto, and points to important directions for future research on market risk assessment.

Introduction

In recent years, researchers have concentrated on identifying the most suitable and accurate models for energy price data. In this pursuit, both theoretical and empirical insights play a critical role in energy forecasting and risk management.[1] When modelling the distribution of a data set, analysts typically assume that the data follow a normal distribution; however, this assumption imposes an exponentially bounded probability distribution. As Mwamba[2] has observed, the normality assumption overlooks the fat-tail characteristic, which is commonly found in high-frequency data, such as daily Brent crude oil prices. Consequently, financial time series often exhibit significant skewness and excess kurtosis, which deviate from the properties of Gaussian distributions.[3] Modelling these extreme fluctuations requires the development and application of robust methods and techniques because the assumption of independent and identically distributed (i.i.d.) market returns is no longer considered inviolable, as statistical models and extreme value theory (EVT) have challenged this assumption more effectively than the Gaussian models have. This shift has driven the demand for radical risk management systems, prompting the Basel Committee on Banking Supervision (BCBS) to introduce international standards to meet these needs.[4] Hence, from this time, this study seeks to enhance the understanding of the potential trend towards reduced vulnerability and increased resilience by estimating the extreme effects of energy price fluctuations. To ensure robust modelling of extremes in the Brent crude oil market, we introduce a novel application that combines an alpha-recurrent neural network (alpha-RNN) with the generalised Pareto distribution (GPD). To the best of the authors' knowledge, this is the first study to implement this mapping application for financial risk analysis specifically in the energy market. By assimilating EVT with deep learning methods, we aim to enhance the forecasting performance and the estimation of financial assets' risk. Furthermore, the ability of alpha-RNN to capture nonlinearity and time dependence in financial data offers substantial improvements in forecasting accuracy because the model combines exponentially smoothed weights, which makes the model essential during periods of high volatility. This combination, therefore, ensures that the model is responsive to market changes, making it suitable for forecasting extreme movements in financial markets.

Value at risk (VaR) is a widely used measure in finance that estimates the potential loss of a portfolio over a specified time horizon at a given confidence level. Despite its simplicity, VaR often underestimates the likelihood of extreme losses, highlighting the importance of accurate tail-end estimation. Models such as the GPD address this limitation by focusing on the tails of the loss distribution, enabling better assessment of rare but significant events. However, estimation uncertainty in these models can still pose challenges for risk management and forecasting.

Research highlights and key findings

This study is one of the first to apply the combination of deep learning with EVT methods, such as the GPD, to the energy market returns (i.e. Brent crude oil). Moreover, we focus on the downside risk (i.e. the losses) and the upside risk (i.e. the gains). In this way, we aim to capture the possibility of achieving losses and gains beyond the expected or average returns. For instance, if a stock has an expected return of 8%, but there is a possibility it could return 5% due to unfavourable market conditions, the difference (i.e. 8% – 5% = 3%) represents the downside risk (loss), hence, the application of our approach to these specific energy market has not been thoroughly explored in the existing literature. The highlights and key findings of this study are summarised in Table 1.

Methodology

In this section, we present the methods and procedures followed in our study. The data set used was a five-business-day selection of Brent crude oil prices from 4 January 2010 to 15 January 2025. According to Abdollahi and Ebrahimi[5], this oil is a widely used global benchmark for crude oil pricing, with approximately 80% of all crude contracts referencing it. These data were obtained from https://finance.yahoo.com/ and were accessed on 18 January 2025, using the Yahoo Finance package from Python. For a day t, the daily log return r_t is defined as the natural logarithm of the value at time t divided by the natural logarithm of the value at time t – 1. The data analysis was performed in Python 3.12.1 using Keras with a TensorFlow backend on a high-performance computer. Model training and testing were conducted on a system equipped with AMD Ryzen 7000 series 5, 16 GB of RAM and AMD Radeon Graphics. The average training time for the alpha-RNN across the three forecast horizons (12, 36 and 60 months) ranged between approximately 10 and 90 seconds per epoch. Inference (prediction) for each horizon requires less than 5 seconds, demonstrating that the computational cost, although higher than conventional econometric models, remains practical for most applications.

Alpha-RNN

A univariate alpha-RNN configured for one-step-ahead forecasting with a fixed smoothing parameter is evaluated whenever $s = t - \rho + 2, \dots, t$, which provides

$$\begin{aligned}\hat{r}_{t+m} &= \omega_y \hat{h}_t + b_r,\ \hat{h}_s \\ &= g\left(U_h \tilde{h}_{s-1} + \omega_h y_s + b_h\right),\ \tilde{h}_s \\ &= \alpha \hat{h}_s + (1+\alpha)\tilde{h}_{s-1}\end{aligned} \qquad \text{Equation 1}$$

where $\hat{r}_{t+m} = \omega_y \hat{h}_t + b_r$ is the output function, $h_s = g\left(U_h \tilde{h}_{s-1} + \omega_h y_s + b_h\right)$ is the hidden state update and $\tilde{h}_s = \alpha \hat{h}_s + (1+\alpha)\tilde{h}_{s-1}$ represents the smoothing function. This model enhances the simple RNN by replacing the hidden state $\hat{h}_{s-1}$ in the hidden layer with an exponentially smoothed hidden state $\tilde{h}_{s-1}$, and when $\alpha = 1$, the smoothing gives the appearance of infinite memory. This can be intuitively understood by simplifying the scenario and considering inactive conditions. For instance, when $b_r = b_h$ $U_h = \phi \in \mathfrak{R}$, and $\omega_y = 1$, Equation 1 reduces to Equation 2:

$$\hat{r}_{t+1} = \hat{h}_t = \phi\left(\tilde{h}_{t-1} + y_t\right) = \phi\left(\alpha \hat{h}_{t-1} + (1-\alpha)\tilde{h}_{t-2} + r_t\right) \qquad \text{Equation 2}$$

with the starting condition in each sequence $\hat{h}_{t-1\rho+1} = \phi y_{t-\rho+1}$. Now, consider $\rho = 2$ lags in the model so that $\hat{h}_{t-1} = \phi r_{t-1}$, then

$$\hat{h}_{t-1} = \phi\left(\alpha\phi y_{t-1} + (1-\alpha)\tilde{h}_{t-2} + r_t\right) \qquad \text{Equation 3}$$

and Equation 3 can be written in the following simpler form

$$\hat{r}_{t+1} = \phi_1 r_t + \phi_2 r_{t-1} + \phi(1-\alpha)\tilde{h}_{t-2} \qquad \text{Equation 4}$$

with autoregressive weights being $\phi_1 = \phi$ and $\phi_2 = \alpha\phi_2$. Upon closer inspection, we observe that Equation 4 has a third component on the right-hand side that disappears when alpha becomes unity; and this gives the model indefinite memory because h_t depends on r_1 which is the first observation in the returns series, not just the first observation in the sequence. To see this, we unroll the recursion relation in the exponential smoothing and have

$$\tilde{h}_{\alpha t+1} = \alpha \sum_{s=0}^{t-1} (1-\alpha)^s \hat{h}_{t-s} + (1-\alpha)^t y_t \qquad \text{Equation 5}$$

where $\tilde{h}_1 = y_1$ is used. It is often convenient to characterise exponential smoothing by the half-life: the number of lags needed for the coefficient $(1-\alpha)^s$ to be $s = \frac{1}{\log_2(1-\alpha)}$.

GPD and peaks over threshold

The GPD plays the role of a natural distribution for excesses over a reasonably high threshold. In other words, the peaks over threshold (POT) used in the GPD modelling plays the same role as the block maxima/minima (BMM) on the generalised extreme value (GEV)

Table 1: Performance and contributions of the alpha-recurrent neural network (alpha-RNN) generalised Pareto distribution (GPD) model

Characteristic	Highlights	Findings	Contributions
Tail risk detection	Effective in modelling extreme events	Losses are significantly riskier than gains	Captures asymmetry in crude oil price movements
Backtesting reliability	High *p*-values in the Kupiec and Christoffersen tests	Validates statistical reliability	Demonstrates robustness in tail risk estimation
Downside risk assessment	Accurate estimation of downside risks	Effective for both long- and short-term positions	Improves risk management for volatile markets
Model scalability	Computationally efficient for large data sets	Handles nonlinearity and extreme value behaviour	Applicable across diverse financial data sets
Methodological innovation	Integration of alpha-RNN with GPD	Models peaks-over-threshold effectively	Combines deep learning with extreme value theory for robust risk modelling
Shape parameter insights	Analysis reveals a higher risk of losses	Highlights asymmetry in price movement risks	Provides critical insights for tailored risk management

distribution.[6] According to Arnold[7], the parameter ξ in the GPD corresponds to the GEV distribution and serves the same purpose. For an optimal threshold selection, we use a forward-stop algorithm proposed by Troop and coworkers[8]. In this algorithm, the i.i.d. residuals from the alpha-RNN architecture are generated, and the threshold is fixed in such a way that $\boldsymbol{u}_i = \boldsymbol{x}_i$ where $i = 1, 2, 3, \ldots, m$ and $(m < n)$, $m > 0$. After this, the null hypothesis that the sample $x_1, \ldots, x_n$ follows a GPD with the shape and scale parameters denoted as ξ and β is tested, and thereafter, we compute the corresponding probability values to make the decision. The cutoff denoted as $\hat{k}$ is computed such that

$$\hat{k} = \max\left\{ k \in \{1, \ldots, m\} : -\frac{1}{k}\sum_{i=1}^{k} \log(1 - p_i) \leq \alpha \right\} \qquad \text{Equation 6}$$

where α is a significance level and p_i is the computed probability values from the tested null hypothesis. After this, the optimal threshold is selected as $u_i = u_k$.

After selecting the best (optimal) threshold, Nortey and colleagues[9] showed that the sequence of observations above the selected threshold is well estimated by a two-parameter GPD that has the following representation

$$G_{\xi,\beta}(x) = \begin{cases} 1 - \left(1 + \dfrac{\xi(x-u)}{\beta}\right)^{-1/\xi}, x \neq 0 \\ 1 - \exp\left(-\left(\dfrac{x-u}{\beta}\right)\right), x = 0 \end{cases}, \qquad \text{Equation 7}$$

where $x > 0$ when $\xi \leq 0$, otherwise, $0 \leq x \leq \dfrac{\beta}{\xi}$ for $\xi < 0$, and $\beta > 0$ to avoid undefined solutions.[10] To estimate the parameters of the GPD, we then adopt a weighted nonlinear least squares likelihood moments (WNLSLMs) estimation process[11] because it performs best for the shape parameter of the extreme value models in most cases[12].

Risk measures

The possibility of measuring risk, according to Pfaff[13], is to estimate losses that may occur when prices for the assets of the portfolio or stock returns are declining. The risk measures used to estimate risk at the tail of the distribution, according to Rockafellar and Uryasev[14], are mainly used to protect a financial market position from heavy losses. Two approaches, parametric and nonparametric, are commonly used to calculate value-at-risk (VaR) and expected shortfall (ES) for financial data. In this study, we use a parametric approach and follow a GPD as described in Equation 6; hence, the VaR is computed by following the computation presented in the work of Krężołek[15]. While the ES is computed using the mathematical computation of Rockafellar and Uryasev[14], it is defined to represent the anticipated loss magnitude, specifically measuring the average loss when it exceeds the threshold of the VaR.

The other risk measure we use in this study is the conditional tail expectation (CTE), also known as tail value-at-risk (TVaR), which measures the expected loss given that a loss exceeds a certain threshold, specifically the VaR. For a random variable that follows a GPD with parameters ξ, β and threshold u, CTE is computed. Finally, the glue-VaR risk measure is a flexible approach to risk assessment that combines aspects of VaR and ES is defined as a linear combination of VaR, which is computed with the confidence level given as α and beta risk as β and this corresponds to the restriction $(0 < \beta < \alpha < 1)$.

Backtesting risk measures

Backtesting is a statistical method used to compare actual losses with appropriate risk measures, with exceptions occurring once every 100 days with a 99% confidence level. The backtesting process evaluates whether the frequency of these exceptions aligns with the specified confidence level.[16] We therefore use the Kupiec unconditional coverage test to evaluate the effectiveness and accuracy of a model in calculating VaR, and it is favoured for confirming model adequacy.[17] We let x^p be the number of violations observed at level p, i.e. $r_t < VaR_p$ (for long positions) or $r_t > VaR_p$ (for short positions). The test procedure involves comparing the corresponding proportion of violation $[x^p/N]$ to p. Under H_0, the Kupiec likelihood ratio test follows a chi-square distribution with one degree of freedom. If the calculated probability value is less than the observed probability value, reject the null hypothesis and infer that the model is incorrect, indicating that the risk computed using the estimated risk model is not accurate.

Another test that we employ to test the adequacy of the model is the Chrisforssen test. This test is known as the Markov test, and it checks for the independence property by analysing whether the probability of a VaR violation on a given day is influenced by the outcome of the previous day. By defining p_{ij} to represent a violation probability that occurs conditionally on the state i at time $t-1$, the Chrisforssen test then follows a chi-square test of independence, and the null hypothesis is that the VaR violations at a given significance level are independent of each other. Therefore, we reject the null hypothesis if the calculated probability value is less than the observed probability value.

Finally, the accuracy of CTE and ES is jointly evaluated with VaR using the Fissler–Ziegel (FZ) scoring rule and the joint quantile (JQ)–ES regression test. These methods are meant to assess how well VaR, ES, and CTE work together, fixing the problems of the Kupiec and Christoffersen tests, which only look at VaR. The FZ scoring rule gives a reliable way to score the (VaR, ES or CTE) pair, making it easier to compare forecasts by looking at how accurate they are for both the expected losses and the size of those losses. Unlike threshold-based tests, this approach assesses tail risk forecasts comprehensively, addressing both the frequency and magnitude of extreme losses. The JQ–ES regression test offers a regression-based methodology that simultaneously models and evaluates the accuracy of VaR, ES and CTE forecasts. This helps in making statistical conclusions about both the extreme values and the average of those values, ensuring that the expected shortfall forecast is fair compared with the VaR level. These tools improve the reliability of tail risk evaluations, especially in financial contexts where precise downside risk modelling is essential.

Data analysis and results

This section deals with an empirical analysis using time series data, and the univariate data analysis for the variable of interest, the Brent crude oil prices, is shown in Figure 1. The normal density in Figure 1a shows that the distribution of the Brent crude oil closing returns is slightly skewed to the left, indicating that these prices are asymmetric, i.e. a negative skewness. Figures 1b and 1c show that the data are non-normally distributed. It is also important to note that seasonality is accompanied by both positive and negative patterns in Figure 1d, where the graph shows significant fluctuations in Brent crude oil prices over the years. Notable patterns include periods of sharp declines and recoveries, such as around 2016 and 2020, likely reflecting economic events or shocks. These sharp declines and recoveries in Brent crude oil prices are often driven by a combination of economic, geopolitical and industry-specific factors; for example, the 2015/2016 Shanghai stock market crash and the 2020 COVID-19 pandemic. In agreement with Prabheesh and coworkers[18], these events reduced oil demand significantly, causing price crashes. Conversely, periods of recovery, such as post-2020[19], are fuelled by increased industrial activity, travel and production cuts by the Organisation of the Petroleum Exporting Countries (OPEC) to stabilise prices. Long-term structural shifts, such as the US shale oil boom, renewable energy adoption and market speculation, have further amplified volatility, with the future market reacting strongly to news about supply and demand.[20,21]

The estimated average return of Brent crude oil is 0.00026, suggesting that returns vary around a value near zero. The standard deviation of 0.02252 indicates a moderate level of volatility in daily returns, as evidenced by the time series plot. The return distribution exhibits a negative skewness of –0.9571, indicating an asymmetry characterised by an extended left tail. The visual evidence from the kernel density estimate in Figure 1a indicates an asymmetric distribution. The kurtosis value of 16.91809, significantly exceeding the benchmark value of 3, indicates a leptokurtic distribution characterised by fat tails and an increased likelihood of extreme returns, while the density plot in Figure 1a and the boxplot in Figure 1c both support this observation, clearly indicating the presence of severe extremes. Formal statistical tests provided additional evidence supporting the deviation from normality, where the Shapiro–Wilk (S–W) test produced a p-value of 0.0001, whereas the Anderson–Darling (A–D) test rejected the null hypothesis of

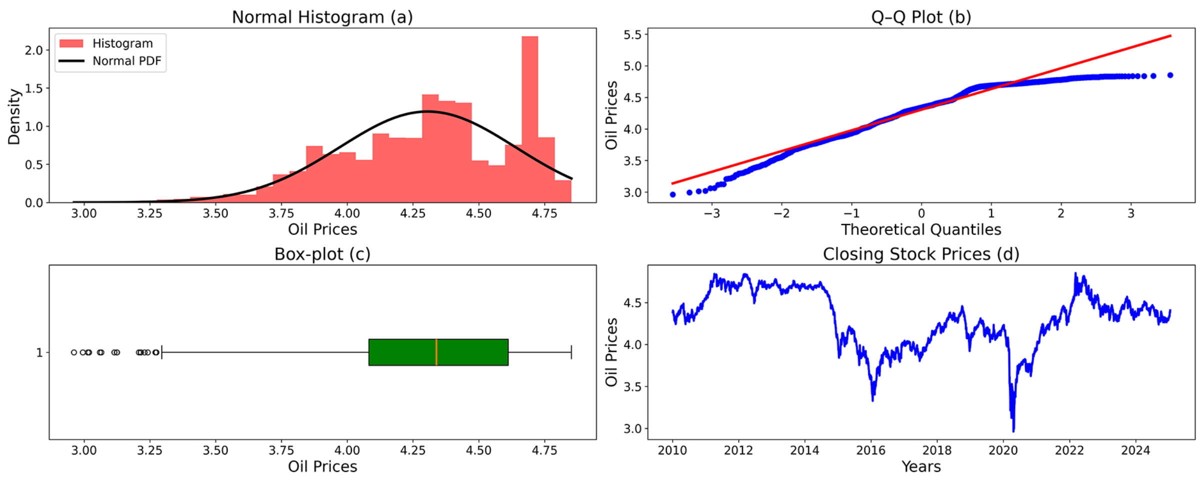

Figure 1: Plot of Brent crude oil closing.

Table 2: Descriptive statistics and test for normality

	Mean	Standard deviation	Skewness	Kurtosis	Shapiro–Wilk (*p*-value)	Anderson–Darling (*p*-value)	Modified Wald test
Returns	0.00026	0.02252	–0.9571	16.91809	0.90276 (0.0001)	56.604 (0.0001)	–3.4358 (0.0042)

normality with a test statistic of 56.604 ($p = 0.0001$). Finally, the modified Wald test showed a significant result (–3.4358; $p = 0.0042$), confirming that the Brent crude oil return distribution is not normal and has heavy tails. Table 2 presents a summary of these statistics.

Short-, medium- and long-term forecasting using alpha-RNN

The application process for the alpha-RNN involves splitting the data set into 80% training, 10% testing and 10% validation to ensure robust performance and avoid overfitting. Using the multi-input, multi-output (MIMO) forecast, the alpha-RNN is trained, tested and validated. We therefore implement a fully connected recurrent neural network using Keras, following the methodology outlined in Kumari and Sood[22]. To ensure stability, the recurrent weight matrices are initialised using an orthogonal matrix, keeping the absolute values of the eigenvalues bounded by unity as suggested by Henaff and colleagues[23]. Non-recurrent weight matrices are then initialised using the Glorot and Bengio uniform initialisation method.[24] Cross-validation is conducted over hidden neuron configurations as per the approach of Dixon.[25]

The 12-, 36- and 60-month step-ahead forecasts in Table 3 reveal that the model performs better in modelling positive returns than in modelling negative returns because the model has lower root mean squared error (RMSE) and mean squared error (MSE), and reliability is higher across all horizons. Notably, the 12-month forecasts for negative returns exhibit poor coverage (20.36%) and higher errors, suggesting challenges in short-term modelling for downward trends. Over longer horizons, both return types show improved coverage, with the 36-month forecasts achieving high reliability (above 65% for both positive and negative returns). Computational time varies, being highest for 36-month forecasts and significantly lower for 60-month predictions, reflecting potential model optimisations. These results emphasise the model's strength in capturing long-term trends while pointing to the need for better short-term calibration, particularly for negative returns.

Fitting the GPD

Nonetheless, Figure 2 shows the extreme values that are above the 95th percentile and below the 5th percentile for both negative and positive returns that are to be used to fit the GPD. Using the forwardstop algorithm, we find that an optimal threshold is −0.036297 for negative returns and 0.046467 for positive returns, and

Table 3: The 12-, 36- and 60-month step-ahead forecasts

Returns	RMSE	MSE	MAE	Coverage	Time
12 months					
Positive returns	0.3001	0.0900	0.0399	0.8791	76.55
Negative returns	0.5031	0.2531	0.0509	0.2036	20.30
36 months					
Positive returns	0.2374	0.0563	0.0818	0.687	92.24
Negative returns	0.3871	0.1498	0.1067	0.809	80.44
60 months					
Positive returns	0.1982	0.0393	0.0818	0.987	12.24
Negative returns	0.2908	0.0845	0.0786	0.951	10.44

RMSE, root mean squared error; MSE, mean squared error; MAE, mean absolute error

these resulted in 176 exceedances for the losses (i.e. negative returns) and 62 exceedances for the gains (i.e. positive returns). The fitted GPD using WNLSLMs in Table 4 indicates heavier and more dispersed negative tails. The scale parameter for losses ($\beta = 0.02273$) exceeds that for gains ($\beta = 0.01867$), whereas the shape parameter for losses ($\xi = 0.29699$) is also larger than its positive counterpart ($\xi = 0.26814$). Thus, conditional on exceeding the threshold, losses not only vary more widely but also exhibit a fatter tail, confirming a pronounced downside risk asymmetry. Standard errors for negative returns are lower, suggesting more precise estimates for losses than gains.

As both shape parameters are positive, it follows that there is no finite right endpoint for either return. This characteristic is essential in risk management, as it implies that potential extreme losses or gains are unbounded, emphasising the importance of preparing for extreme market movements. Moreover, the greater number of extreme negative returns compared with positive returns indicates a more volatile or riskier environment for losses, which is crucial information for investors and risk managers. However, this supports the idea that EVT is good at explaining the heavy-tailed behaviour of Brent crude oil prices, as shown in studies that used the peaks over threshold method to look at Brent and West Texas prices, indicating that both investors and the public face extreme price risks.[26] Finally, these findings collectively emphasise the importance of EVT in quantifying risks and preparing for extreme price fluctuations in the energy market.

The goodness of fit test

After model estimation, the goodness of fit (GoF) is assessed. The A–D and S–W tests for GoF are used, and the results are given in Table 5. Previous studies[6,27] also used these tests for testing GoF for GPD. The null hypothesis that the returns follow a GPD is not rejected because all the calculated probability values are greater than the observed

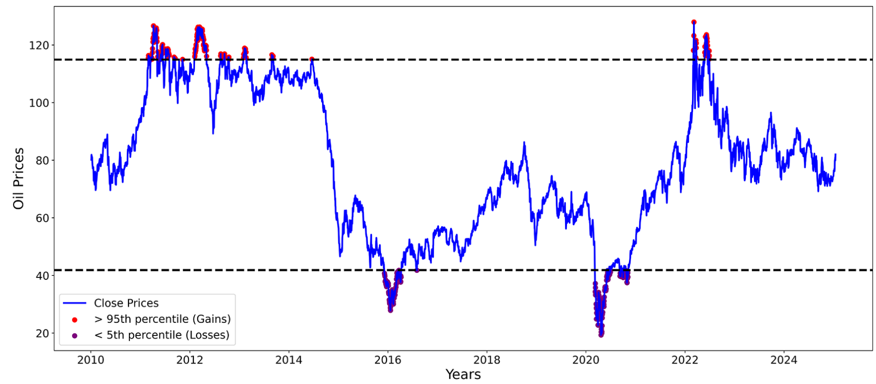

Figure 2: Threshold exceedances below 5% and above 95% levels.

Table 4: Parameter estimates of the fitted generalised Pareto distribution for positive and negative returns

	Positive returns	Negative returns
Threshold	0.046467 (62)	−0.036297 (176)
β_{ij}	0.01867	0.02273
$Se(\beta_{ij})$	0.00409	0.00155
ξ_{ij}	0.26814	0.29699
$Se(\xi_{ij})$	0.18416	0.19009

Note: Values in parentheses are the number of exceedances above and below the threshold for both negative and positive returns.

Table 5: Goodness-of-fit tests for positive and negative returns

Test		Positive returns	Negative returns
Anderson–Darling	Statistic	0.279	0.136
	p-value	0.648	0.978
Shapiro–Wilk	Statistic	0.049	0.018
	p-value	0.52	0.983

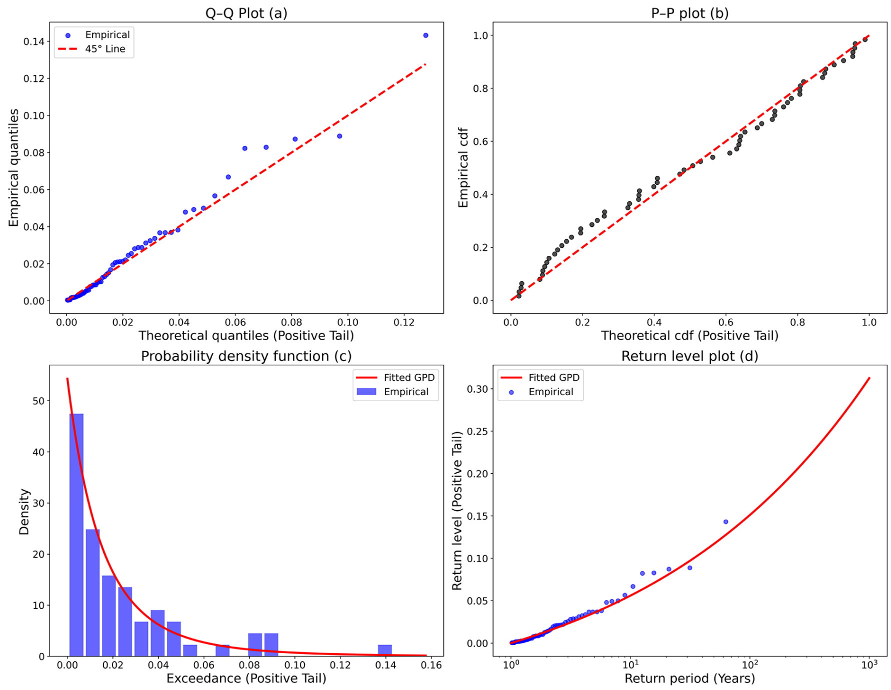

Figure 3: Diagnostics plots for the generalised Pareto distribution for both positive tails.

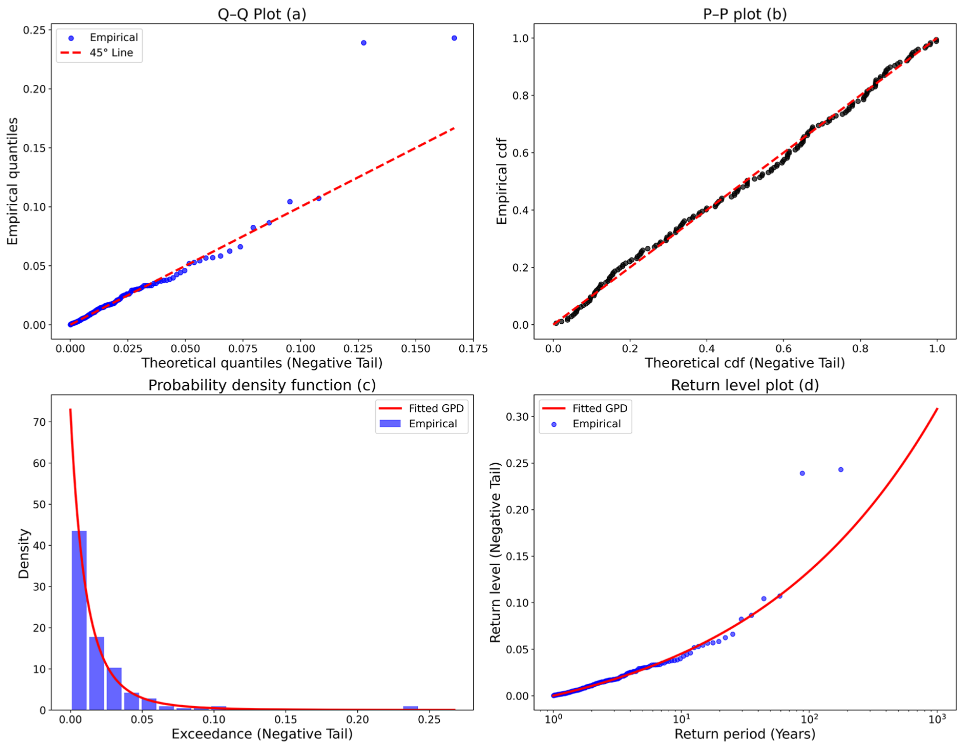

Figure 4: Diagnostics plots for the generalised Pareto distribution for the negative tail.

probability value of 5; hence, the returns are modelled very well with the specified distribution. Figures 3 and 4 also support these results.

Extreme quantiles and risk estimates from alpha-RNN-GPD

To effectively estimate the extreme quantiles, as well as downside and upside risk in the Brent crude oil market, we employ the minimum score combined algorithm, incorporating predictions from the alpha-RNN and the GPD, and the final model is alpha-RNN-GPD. The true values of extreme quantiles derived from the Alpha-RNN-GPD are presented in Table 6. These results highlight the performance of the model alongside the quantile mean method in modelling extreme positive and negative returns. For positive returns, the model yields slightly higher estimates at both the 95th (9.938) and 99th percentiles (9.937) compared with the quantile mean method (9.818 and 9.932, respectively), with a low RMSE of 0.144, indicating a good fit. For negative returns, the model also produces higher estimates at the 95th (21.138) and 99th percentiles (21.144) than the quantile mean method (21.124 and 21.144, respectively), with a RMSE of 0.185. This suggests slightly greater variability in modelling extreme negative returns, which aligns with earlier findings of heavier and more dispersed downside tail behaviour.

Four common risk measures that are known as VaR ES, CTE and glue-VaR are used to compute the risk for Brent crude oil prices. To show the results of deep learning combined with EVT as a financial risk measure, VaR, ES, CTE and glue-VaR are estimated, and the results are presented in Table 7.

For positive returns, there are relatively lower risk measures (VaR, ES and CTE) compared with negative returns, emphasising that the downside risk is more severe (Table 7). As the confidence level increases from $p = 0.95$ to $p = 0.99$, all risk measures (except VaR for positive returns) increase, indicating the growing importance of extreme tail behaviour. The glue-VaR values show a notable increase for negative returns at higher confidence levels, accentuating the disproportionate effect of extreme losses. The computed risk measures have significant implications for the energy market, particularly for Brent crude oil, which is characterised by high volatility and susceptibility to extreme price movements due to geopolitical, economic and supply–demand dynamics. Negative returns (potential losses) exhibit significantly higher risk measures (VaR, ES and CTE) than positive returns (potential gains).

Table 6: Extreme quantiles

Returns	Model	0.95	0.99	RMSE
Positive	Alpha-RNN-GPD	9.938	9.937	0.144
	Quantile mean	9.818	9.932	
Negative	Alpha-RNN-GPD	21.138	21.144	0.185
	Quantile mean	21.124	21.144	

VaR, value at risk; ES, expected shortfall; CTE, conditional tail expectation

Table 7: Computed risk measures

Returns	*p*	VaR	ES	CTE	Glue-VaR
Positive	0.95	1.14	1.37	8.479	0.842
	0.99	1.14	1.89	10.719	0.953
Negative	0.95	2.195	3.07	8.456	0.842
	0.99	1.227	2.132	10.987	3.074

VaR, value at risk; ES, expected shortfall; CTE, conditional tail expectation

This indicates that the Brent crude oil market has a heavier downside tail, meaning extreme price drops are more likely and significant than extreme price gains. Energy market participants, including producers, traders and policymakers, need to prioritise risk mitigation strategies for adverse price movements. Hedging instruments such as future contracts, options and stop-loss strategies are thus critical to protect against these severe price drops.

Table 8 summarises the backtesting results for the alpha-RNN-GPD model across four key risk measures, VaR, glue-VaR, ES and CTE, under both positive and negative crude oil returns. For positive returns, all risk measures exhibit strong statistical reliability. The Kupiec and Christoffersen tests yield

Table 8: Backtesting of risk measures for positive returns

Returns	Kupiec test			Fissler–Ziegel test		
Positive	Risk measures	95%	99%	Risk measures	95%	99%
	VaR	0.783	0.7258	ES	0.7710	0.7382
	Glue-VaR	0.7710	0.7382	CTE	0.7680	0.8433
	Christoffersen test			Joint quantile test		
	VaR	0.8880	0.8632	ES	0.687	0.6571
	Glue-VaR	0.8880	0.8632	CTE	0.8328	0.7697
Negative	Risk measures	95%	99%	Risk measures	95%	99%
	VaR	0.7904	0.8964	ES	0.8587	0.8121
	Glue-VaR	0.8917	0.7871	CTE	0.8677	0.7898
	Christoffersen test			Joint Quantile test		
	VaR	0.9221	0.8741	ES	0.2146	0.7871
	Glue-VaR	0.8391	0.7587	CTE	0.3047	0.5989

consistently high p-values, indicating accurate coverage and independence of violations for VaR and glue-VaR. Similarly, ES and CTE perform well under the FZ scoring rule and joint quantile test, with CTE showing particularly strong performance at the 95% and 99% levels (0.8328 and 0.7697), confirming well-calibrated tail forecasts. For negative returns, the model continues to perform robustly. Both VaR and glue-VaR pass both the Kupiec and Christoffersen tests with p-values above 0.75, whereas ES and CTE achieve high FZ scores (e.g. ES 0.8587 and 0.8121; CTE 0.8677 and 0.7898), highlighting the model's effectiveness in capturing downside risk. However, the JQ test shows slightly lower calibration for ES (0.2146 at 95%), suggesting some sensitivity in the left tail when jointly assessing quantile and ES forecasts. The capability of the model to provide reliable predictions for short-term scenarios strengthens its practical application in risk-management and decision-making processes.

Conclusions, limitations and recommendations

This study combines an alpha-RNN with the GPD to evaluate the risks associated with Brent crude oil price returns. The ability of the alpha-RNN to model complex time series patterns, combined with the focus on extreme tail behaviour by the GPD, provides a robust application in assessing the likelihood and magnitude of extreme price movements in financial markets, especially the Brent crude market under study. This approach offers a more precise estimations of VaR, ES, Glue-VaR and CTE, particularly at higher confidence levels. The results highlight significant tail risks, with negative returns showing a higher vulnerability to extreme price drops compared with positive returns. The combined effect of alpha-RNN and GPD improved the predictive capability of the model. The ability of alpha-RNN-GPD to capture nonlinear dependencies and stationary patterns enhanced the accuracy and reliability of tail risk estimates, making it a valuable tool for predicting rare but significant events in the Brent crude oil market. This combined approach allows for a more comprehensive understanding of the dynamics of extreme price fluctuations, equipping market participants with actionable insights to prepare for low-probability, but highly significant events. The findings of this study have strategic and regulatory implications. For market participants, the results give emphasis to the importance of adopting innovative risk management techniques, such as stress testing and hedging strategies, to mitigate extreme risks effectively. Overall, the combination of alpha-RNN and GPD represents a significant step forward in improving the assessment of tail risks and guiding the development of robust strategies to ensure market stability and sustainability.

Limitations

The alpha-RNN-GPD model demonstrates strong performance in capturing tail risk, but its effectiveness is limited to high computational costs, requiring cutting-edge hardware for training and implementation, making it less accessible for smaller institutions or individual traders. Moreover, the focus of tail risk neglects other important risk dimensions, such as liquidity risk, operational risk or macroeconomic influences that also significantly influence crude oil prices. Proper calibration of the alpha-RNN with GPD components demands expertise, posing a barrier for practitioners without substantial statistical and deep learning knowledge. Although the backtesting results validate the model's performance for the given data set, its robustness in predicting extreme events during unseen market conditions or crises is not guaranteed. Finally, the application in this study assumes that historical price movements fully encapsulate market dynamics, yet evolving geopolitical and environmental factors impacting crude oil prices are not fully captured in this study.

Recommendations

To enhance the alpha-RNN-GPD model and its applicability to broader contexts, several recommendations should be considered. First, it is crucial to expand model testing across a variety of financial markets and commodities to assess its generalisability. This will help to determine whether the model can consistently capture tail risk in diverse market conditions. By doing so, we can evaluate its adaptability to different asset classes and improve its robustness. Relaxing the assumptions of stationarity is essential for improving the performance of the model during periods of structural market changes or external shocks; moreover, exploring alternative models that can handle non-stationarity more effectively, a dynamic version of the alpha-RNN will help to address this limitation and improve the robustness during extreme events. The computational complexity of assimilating alpha-RNN and GPD models is another consideration. To make the model more accessible to smaller institutions or individual traders, it is recommended to optimise its computational efficiency. This can be achieved by simplifying the model architecture, applying dimensionality reduction techniques or utilising more efficient algorithms, and reducing the computational cost which will increase the practical applicability of the model. In addition, incorporating other risk factors beyond tail risk can provide a more holistic view of the risk environment. In this study, we have primarily focused on tail risk, neglecting other important dimensions such as liquidity risk, operational risk and macroeconomic influences. Expanding the model to include these factors would offer a more comprehensive assessment of the risk landscape, particularly for Brent crude oil prices.

Enhancing calibration methods is also crucial to ensure the model is usable by a broader audience. Developing user-friendly calibration tools or offering pre-trained models and tutorials could help bridge the knowledge gap and make the model more accessible to those without advanced statistical or machine-learning knowledge. Finally, conducting robust out-of-sample testing is essential to validate the robustness of the model. Although the backtesting results provide some assurance, it is important to test the model under various market conditions, including crises or periods of extreme volatility, and this will ensure that the model can predict extreme events reliably, even in unseen conditions. The model should be adjusted to incorporate real-time factors that influence Brent crude oil prices, such as evolving geopolitical and environmental events, which will be studied elsewhere.

Acknowledgements

We thank numerous people for their helpful comments on this manuscript.

Data availability

The entire data set supporting the results of this study has been deposited in a recognised repository and is openly accessible via https://github.com/Makatjane/Data.

Declarations

We have no competing interests to declare. We have no AI or LLM use to declare. This work is based on the mini-dissertation 'Modelling Electricity Prices in Botswana' by K.B., supervised by K.M.

Authors' contributions

K.B.: Conceptualisation, methodology, formal analysis, validation, data curation, writing – original draft; writing – review and editing. K.M.: Supervision, project leadership, project administration, investigation, writing – review and editing. Both authors read and approved the final manuscript.

References

1. Huang CS, Huang CK, Chinhamu K. Assessing the relative performance of heavy-tailed distributions: Empirical evidence from the Johannesburg Stock Exchange. J Appl Bus Res. 2014;30(4):1263–1286. https://doi.org/10.19030/jabr.v30i4.8675

2. Mwamba JM. Implementing a robust risk model for South African equity markets: A peak-over-threshold approach. S Afr J Econ. 2012;80(4):459–472. https://doi.org/10.1111/j.1813-6982.2012.01319.x

3. Sezer OB, Gudelek MU, Ozbayoglu AM. Financial time series forecasting with deep learning: A systematic literature review: 2005–2019. Appl Soft Comput. 2020;90, Art. #106181. https://doi.org/10.1016/j.asoc.2020.106181

4. Shakdwipee P, Mehta M. From Basel I to Basel II to Basel III. Int J Sci Technol Res. 2017;3(1):66–70. https://www.ijntr.org/from-basel-i-to-basel-ii-to-basel-iii

5. Abdollahi H, Ebrahimi SB. A new hybrid model for forecasting Brent crude oil price. Energy. 2020;200, Art. #117520. https://doi.org/10.1016/j.energy.2020.117520

6. Maposa D. Statistics of extremes with applications to extreme flood heights in the Lower Limpopo River Basin of Mozambique [PhD thesis]. Mankweng: University of Limpopo; 2016. https://ulspace.ul.ac.za/handle/10386/1695

7. Arnold BC. Pareto distribution. Wiley StatsRef: Statistics Reference Online. Hoboken, NJ: John Wiley & Sons; 2014. p. 1–10. https://doi.org/10.1002/9781118445112.stat01100

8. Troop D, Godin F, Yu JY. Best-arm identification using extreme value theory estimates of the CVaR. J Risk Financ Manag. 2022;15(4), Art. #172. https://doi.org/10.3390/jrfm15040172

9. Nortey EN, Asare K, Mettle FO. Extreme value modelling of Ghana Stock Exchange index. SpringerPlus. 2015;4(1), Art. #696. https://doi.org/10.1186/s40064-015-1306-y

10. Chinhamu K, Huang CK, Huang CS, Chikobvu D. Extreme risk, value-at-risk, and expected shortfall in the gold market. Int Bus Econ Res J. 2015; 14(1):107–122. https://doi.org/10.19030/iber.v14i1.9035

11. Zhao X, Zhang Z, Cheng W, Zhang P. A new parameter estimator for the generalized Pareto distribution under the peaks over threshold framework. Mathematics. 2019;7(5), Art. #406. https://doi.org/10.3390/math7050406

12. Del Castillo J, Serra I. Likelihood inference for the generalized Pareto distribution. Comput Stat Data Anal. 2015;83:116–128. https://doi.org/10.1016/j.csda.2014.10.014

13. Pfaff B. Financial risk modelling and portfolio optimization with R. Hoboken, NJ: John Wiley & Sons; 2012. https://doi.org/10.1002/9781118477144

14. Rockafellar RT, Uryasev S. Optimisation of conditional value-at-risk. J Risk. 2000;3:21–41. https://EconPapers.repec.org/RePEc:rsk:journ4:2161159

15. Krężołek D. The Gluevar risk measure and investors' attitudes to risk – an application to the non-ferrous metals market. Stat Transit New Ser. 2016;17(2):305–316.

16. Nieppola O. Backtesting value-at-risk models [master's thesis]. Helsinki: Helsinki School of Economics; 2009. Available from: https://aaltodoc.aalto.fi/bitstreams/fafacabc-c1c0-4e52-b2f3-35dd99d03e71/download

17. Basel Committee on Banking Supervision. Minimum capital requirements for market risk [webpage on the Internet]. c2019 [cited 2025 Feb 10]. Available from: https://www.bis.org/bcbs/publ/d457.htm

18. Prabheesh K, Padhan R, Garg B. COVID-19 and the oil price–stock market nexus: Evidence from net oil-importing countries. Energy Res Lett. 2020;1(2), Art. #13745. https://doi.org/10.46557/001c.13745

19. Norouzi N. Post-COVID-19 and globalisation of oil and natural gas trade: Challenges, opportunities, lessons, regulations, and strategies. Int J Energy Res. 2021;45(10):14338–14356. https://doi.org/10.1002/er.6762

20. Pandey P, Sehgal S. Dynamic currency linkages and their determinants: An empirical study for the East Asian Economic Community region. Emerg Mark Finance Trade. 2018;54(7):1538–1556. https://doi.org/10.1080/1540496X.2017.1380621

21. Kaufmann RK. The role of market fundamentals and speculation in recent oil price movements. Energy Policy. 2011;39(1):105–115. https://doi.org/10.1016/j.enpol.2010.09.018

22. Kumari PA, Sood M. Implementation of a simple RNN and LSTMs-based prediction model for coronavirus disease (COVID-19). IOP Conf Ser: Mater Sci Eng. 2021;1022, Art. #012015. https://doi.org/10.1088/1757-899X/1022/1/012015

23. Henaff M, Szlam A, LeCun Y. Recurrent orthogonal networks and long-memory tasks. Proc Int Conf Mach Learn. 2016;48:2034–2042. https://proceedings.mlr.press/v48/henaff16.html

24. Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. Proc Int Conf Artif Intell Stat. 2010;9:249–256. https://proceedings.mlr.press/v9/glorot10a

25. Dixon M. Industrial forecasting with exponentially smoothed recurrent neural networks. Technometrics. 2022;64(1):114–124. https://doi.org/10.1080/00401706.2021.1921035

26. Amachukwu NJ, Obiora-Ilouno HO, Obisue EI. Modelling of extreme crude oil price using the generalised Pareto distribution: Brent and West Texas benchmark price. Earthline J Math Sci. 2021;6(2):359–374. https://doi.org/10.34198/ejms.6221.359374

27. Huang CK, Chikobvu D, Chinhamu K. Evaluating risk in precious metal prices with generalised Lambda, generalised Pareto, and generalised extreme value distributions. S Afr Stat J. 2017;51(1):159–182. https://hdl.handle.net/10520/EJC-67c4bb7ee