Scientific Research Publishing

Journal of Financial Risk Management, 2025, 14(3), 199-223
https://www.scirp.org/journal/jfrm
ISSN Online: 2167-9541
ISSN Print: 2167-9533

Integrating Volatility Models within State Space Frameworks for Commodity Return Analysis

Kisswell Basira, Lawrence Dhliwayo, Florance Matarise

Department of Mathematics and Computational Sciences, University of Zimbabwe, Harare, Zimbabwe
Email: kbasira@science.uz.ac.zw, ldhliwayo@science.uz.ac.zw, matarise7@gmail.com

How to cite this paper: Basira, K., Dhliwayo, L., & Matarise, F. (2025). Integrating Volatility Models within State Space Frameworks for Commodity Return Analysis. *Journal of Financial Risk Management, 14,* 199-223.
https://doi.org/10.4236/jfrm.2025.143012

Received: March 3, 2025
Accepted: August 12, 2025
Published: August 15, 2025

Abstract

The study applies a Kalman filter (KF) to Generalized Autoregressive Conditional Heteroskedasticity (GARCH) models to create a hybrid model, to estimate the parameters of the GARCH model in the presence of time-varying volatility. We specify the GARCH model, represented in state space, and use the KF to estimate the volatility. Results are validated by comparing them with estimates obtained through MLE. State space models by nature of their mathematical formulation can handle both observed and latent volatility. The study used simulation data coupled with an empirical analysis of four commodity returns, Crude oil, Gold, Cotton and Lithium. Results show that the hybrid models in general outperformed their MLE counterparts. For the four commodities analyzed, the Skewed Student t-Distribution State Space-GARCH (SSTD_SS_GARCH (1, 1)) was suitable for crude oil and gold while the Process Innovations Volatility Decomposition State Space GARCH (PIVD_SS_GARCH (1, 1)) was fitted for cotton and the student t-distribution MLE GARCH (STD_MLEGARCH (1, 1)) was optimal for lithium. The hybrid model improves forecasting performance by combining the strengths of both GARCH and Kalman filter methodologies.

Keywords

Nonenegativity, Linearization, Adaptive Kalman Filter, State-Space Modeling, Volatility Persistence

1. Introduction

First introduced by (Kalman, 1960) and (Kalman & Bucy, 1961), state-space models (SSM) form a class of models that may be used to study both stationary and non-stationary financial data. These models have two components, that is a latent or hidden state which evolves with time and the observation part that is driven by the unobserved components. This framework applied to GARCH models to create a hybrid model is effective in modeling both volatility and returns with several benefits. It allows for the estimation of time-varying states (like volatility) based on new observations. In addition, the recursive nature of the SSMs enables continuous update of both volatility and parameter estimates as new data comes in, hence can respond to changes in market condition (Bulut, 2024). Further, the GARCH model in a state-space framework provides more reliable, unbiased and consistent parameter estimates. The combination of the GARCH and State Space KF is positioned to improve forecasting accuracy for both returns and volatility by leveraging the synergies abound, which in itself is an added advantage, the ability models non-linear relationships in the data, which can be difficult to capture with standard GARCH models alone (Sharma et al., 2021a).

Volatility measures the degree of variation in asset prices over a period. It is therefore a somber indicator of market hazard and uncertainty, influencing investment choices, risk aversion strategies, and the general economic policies. Understanding volatility is particularly crucial in the context of commodity prices, which play a significant role in global economic stability. The four commodities, crude oil, gold, lithium, and cotton often serve as key indicators of macroeconomic conditions (Basira et al., 2024). Variations in these prices can have implications on inflation rates, exchange rates, and overall economic growth (Pappas & Boukas, 2025). Moreover, accurate volatility estimation is vital for applications in derivatives pricing, futures and the calculation of portfolio Value at Risk (VaR) (Chevallier & Ielpo, 2013). A nuanced interpretation of volatility in these key commodities contributes to more resilient economic strategies in the volatile financial market.

Modeling volatility in commodities like crude oil, gold, lithium, and cotton offers crucial insights for investors and policymakers regarding market behavior and risk management (Engle, 2001). These models help economies hedge against price fluctuations, ensuring stable profit margins. Gold, a safe-haven asset, reveals volatility that signals economic uncertainty and monetary policy effectiveness (Swanepoel & Fliers, 2021). As global technology shifts toward lithium battery-driven vehicles, understanding lithium price dynamics is essential for effective production and supply chain management (Vega-Muratalla et al., 2024). Additionally, cotton's agricultural significance is heightened by climate change, which affects its prices and global demand (Ramos et al., 2024).

This study offers a data-driven comparative analysis of the performance of hybrid state space estimation techniques, and the KF to the traditional MLE as applied to volatility models of the GARCH family. The main objective of this study is to develop a hybrid model combining the Kalman filter with GARCH models for estimating time-varying volatility in commodity returns. This will be achieved through simulation and empirical data analysis of the hybrid model on four commodity returns: crude oil, gold, cotton, and lithium. A comparison of the forecasting performance of the hybrid model against ordinary MLE methods for GARCH models will ensue.

2. Literature Review

The literature on modeling commodity returns has advanced with the development of GARCH-type models that effectively capture volatility undercurrents. Recent research emphasizes the significance of coupling strong estimation techniques like the state space approach in volatility modeling.

The GARCH of (Bollerslev, 1986) and its derivatives, form an integral part of this breed of models. These model gained prominence due to their efficacy, usefulness and adequacy in arresting volatility clustering, leverage effects and asymmetry financial data as cited by (Basira et al., 2024).

The assumption of time-invariant probabilistic properties in classical statistical analyses of financial time series often fails to reflect real-world scenarios. Studies indicate that certain time series exhibit evolving second-moment structures and structural breaks, necessitating the conditions of non-stationarity to ensure adequate capture of the time varying in the volatility. (Likassa et al., 2025) worked a time-varying GARCH model that allows for parameter evolution over time, thereby improving modeling flexibility.

In its general form, the GARCH equation is assumed deterministic and not stochastic. This assumption is not true in any sense, even if the true equation is exact, we might suspect that it is at least subject to measurement error. Dropping this assumption complicates the situation as it leaves us faced with a stochastic GARCH process. (Hall, 1990) shows how the stochastic GARCH-M (SGARCH-M) model may be put into state space form and estimated by the Kalman Filter. This shows that the Kalman Filter provides a useful way of relaxing the Gaussian implausible assumptions in parameter estimation.

Several studies have compared the effectiveness of GARCH-family models with stochastic volatility models in various financial contexts. For instance, (Metsileng et al., 2021) analyzed BRICS foreign exchange rates and stock indices, demonstrating the superior forecasting performance of the GJR-GARCH model. Additionally, (Ewing & Malik, 2017) highlighted the benefits of EGARCH models in capturing asymmetric effects and structural breaks in oil price volatility. The persistence of leverage effects in energy and commodity markets underscores the importance of selecting appropriate volatility models (Beg & Anwar, 2014).

(Ferreira et al., 2017) proposed a state-space approach for GARCH models with time-varying parameters to deal with non-stationarity that is usually observed into the Chilean Stock Market (IPSA) and to the American Standard & Poor's 500 index (S&P500). Results show that forecasting procedures for time varying GARCH processes use the prediction equations of the *KF* with true parameters replaced by consistent estimates. In another study, (Wong et al., 2006) presented a new approach to modelling non-stationarity in EEG time series by a generalized state space approach using EEG data recorded during the onset of anaesthesia. Non-stationarity was modelled by allowing the variances of the driving noises to change with time, depending on the state prediction error within the state space model. Results show that the variances of the dynamical noises driving the components can be made time-dependent by generalizing the concept of GARCH modelling to the situation of state-space modelling.

(Omar, 2019) investigated the properties of univariate and multivariate state-space models under conditional heteroskedasticity and multiple structural breaks. The methodology allowed the extension of standard state-space models to heavy tailed data and allowed for dynamic parameters using a Gibbs sampling algorithm to carry out Bayesian inference on the parameters and the latent state vector as well as empirical analysis of ICU data. The results obtained were however not in favor of state space modes. Besides giving plausible results, they were outperformed. Further, in their book, (Kim & Nelson, 2017) discussed at length the methodology of regression coefficients that vary with time due to market dynamics and heteroscedasticity, fads and time varying volatility in financial markets.

More recent studies by (Azman et al., 2022) focused on the behaviour of volatility for the prices of cryptocurrency using a state space model framework for volatility incorporating the Kalman filter. This was applied to forecast the conditional volatility of five cryptocurrency prices (Bitcoin (BTC), Ethereum (ETH), Ripple (XRP), Litecoin (LTC) and Bitcoin Cash (BCH)) for 10,000 consecutive hours during the COVID-19 pandemic from 26 February 2020 until 18 April 2021. Among the three models used, the state space model gives the best fit and narrowest confidence interval of volatility and value-at-risk forecasts. Recent studies have also advocated for the existence of dual long memory in financial time series. In this regard, (Basira et al., 2024) studied this phenomenon in five commodity returns. They considered the following hybrid models ARFIMA-FIAPARCH, ARFIMA-FIGARCH, FIAPARCH and HYGARCH with different innovations distributions ranging from the STD, SSTD and GED. Results show that these models perform exceptionally well in the presence of dual long memory.

The application of KF to GARCH-type models facilitates real-time estimation of latent volatility processes, improving parameter precision. Recent advancements have also seen the incorporation of Bayesian inference with state space models. For instance, Integrated Nested Laplace Approximations (INLA) have been employed to estimate long memory stochastic volatility models (Lima et al., 2023). Recent studies have begun to compare these methodologies directly. (Ossandón & Bahamonde, 2013) postulated that since the dynamics of a GARCH process is nonlinear, the standard KF algorithm cannot be directly applied. The EKF generalizes the model through a linearization process to allow it to handle the nonlinearity posed by the state space formulations of GARCH models.

3. Methodology

This study is designed to estimate volatility models in commodity returns, beginning with the establishment of a state space formulation of GARCH models. We then apply state space methodology to effectively estimate volatility. In the process, the study uses both observable and latent factors of the model to analyze commodity returns data. In light of this, (Kleppe et al., 2022) introduced a factor state-space approach with stochastic volatility, effectively modeling and prediction of commodity contracts, particularly in crude oil markets.

3.1. Formulation of the GARCH Model

GARCH models and all its derivatives have been widely used as shown in the literature review section. They consist of two key components i.e. the mean equation and the volatility equation (Aduda et al., 2016) given by:

The mean equation

$$r_t = \mu + \varepsilon_t \tag{1}$$

The variance equation

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \tag{2}$$

where: σ_t^2 is the variance, ω is the constant, with the α_i and β_j coefficients of the ARCH(p) and GARCH(q) parameters on the model.

3.2. GARCH Model State Space Representation

The GARCH model, in its natural state, offers a compact and interpretable state space representation, distinguishing it as a versatile basis for modeling time-varying volatility in financial time series. This representation facilitates the incorporation of traditional econometric methods with modern state space techniques, improving the general appreciation of the dynamics driving commodity volatility.

The GARCH model is intrinsically dynamic, capturing how current volatility is influenced by past squared returns and previous volatility estimates. This characteristic aligns impeccably with the state space framework, which is designed to model systems characterized by observable outputs (returns) and latent states (volatility). Consequently, the state space representation not only simplifies the modeling process but also allows for a more nuanced analysis of volatility in financial markets (Rzayev & Ibikunle, 2018). The return equation is as in (1) while the variance equation is;

$$\begin{bmatrix} \sigma_t^2 \\ \varepsilon_t \end{bmatrix} = \begin{bmatrix} \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2 \\ r_t - \mu \end{bmatrix} \tag{3}$$

This representation allows for an efficient way to analyse the time varying nature of parameter and is predominantly useful in estimation and forecasting at the same time maintain the adaptability to various types time series data paving the way for KF recursions (Choudhry & Wu, 2008).

3.3. Estimation of Parameters

There are principally two approaches to tackling GARCH models within a state space framework. The first approach involves utilizing the nonlinear representation introduced by (Ossandón & Bahamonde, 2013) and applying the Extended KF to estimate the model parameters. The second approach involves a linearization process through Taylor series expansion, which simplifies the estimation of the GARCH model by approximating the nonlinear dynamics (Heydari et al., 2020).

3.4. Nonlinear Representation

Since the GARCH model is nonlinear as introduced by (Bollerslev, 1986). (Ossandón & Bahamonde, 2013) presented formulations for typical GARCH model with Gaussian perturbations for p, $q \geq 2$; GARCH(p, 1), $p \geq 2$ and GARCH(1, q), $q \geq 2$. Earlier on in 2011, they derived a nonlinear state space representation of the GARCH(1, 1) model.

a) Basic Structure

$$\text{GARCH}(1,1) \equiv \begin{cases} r_t = \mu_t = U_2\sqrt{U_{1,t}} & \text{Return} \\ \begin{bmatrix} U_{1,t} \\ U_{2,t} \end{bmatrix} = \begin{bmatrix} f_1\left(U_{1,t-1}, U_{2,t-1}, \theta\right) \\ 0 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \varepsilon_t & \text{volatility} \end{cases} \tag{4}$$

where $\theta = (\omega, \beta_1, \alpha_1)$ are parameters that you estimate from historical data using the GARCH model. This part updates the state variables U_1, $U_{1,t}$ and U_2, $U_{2,t}$ Here, $U_{1,t}$ represents the conditional variance (volatility) at time t, and U_2, $U_{2,t}$ is a normalized measure of the mean return.

b) Function of variables

$$f_1\left(U_{1,t-1}, U_{2,t-1}, \theta\right) = \omega + \beta_1 U_{2,t}^2 \cdot U_{1,t} + \alpha_1 U_{1,t} \tag{5}$$

is the function of the unobserved recursive state variable, $U_{1,t} = \sigma_t^2$ and $U_{2,t} = \frac{\mu_t}{\sigma_t}$ (Ossandón & Bahamonde, 2013). This function shows how the previous volatility and returns influence current volatility. If past returns were high, it can suggest that future volatility might also be high.

3.5. Formulation of EKF

To tackle the filtering problem involving nonlinear system dynamics (both state and observations), we consider the Extended Kalman Filter (EKF) as proposed by (Jategaonkar, 2015). For simplicity, we will assume that the system is free of external inputs. Given that the EKF provides approximations of the optimal parameter estimates, the nonlinearities in the system's dynamics are effectively captured by a linearized representation of the nonlinear model, centered around the most recent state estimate (Zhu et al., 2021). To achieve optimal results, the linearization process should closely approximate the nonlinear system within the bounds of the state estimates. The EKF algorithm refers to the works of citation. The prediction cycle and the filtering cycles are explained in detail. In that context, the EKF assesses the volatility of asset returns, considering volatility as a latent state variable shaped by past return data and fresh observations. The predict cycle focuses on forecasting future volatility based on established patterns, while the filtering cycle refines these predictions by incorporating actual return data, thereby enhancing accuracy. Both the EKF and the GARCH model address relationships that are not perfectly linear; volatility can exhibit unpredictable changes influenced by past returns and historical volatility. It is important to note that, if the assumptions underlying the model's linearizations are inaccurate, the EKF may produce unreliable volatility estimates (Kumar, 2018).

3.6. Linearisation

Linearization of GARCH type models refers to an approximation technique that simplifies the modeling and estimation process. To linearize GARCH models, the conditional variance equation is represented as a linear function of the model parameters and the lagged squared residuals, capturing the essential dynamics in a more tractable form. Linearization is generally accomplished by applying a first-order Taylor series expansion of the conditional variance equation around its unconditional mean (Sharma et al., 2021b). Consider a simple GARCH (1, 1) model whose conditional variance equation is given in (2), of course equating $p = q = 1$. The resultant equation will have σ_t^2 (volatility), ε_{t-1}^2 (sqaured innivations), ω (constant), β_1 and α_1 respective ARCH and GARCH parameters of the model.

$$\begin{cases} r_t = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \sigma_t^2 \\ r_{t-1}^2 \end{bmatrix} + v_t & \text{observation} \\ \begin{bmatrix} \sigma_t^2 \\ r_{t-1}^2 \end{bmatrix} = \begin{bmatrix} \alpha_1 & \beta_1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma_{t-1}^2 \\ r_{t-2}^2 \end{bmatrix} + \begin{bmatrix} \beta_1 \\ 0 \end{bmatrix} + \omega_t & \text{transition} \end{cases} \tag{6}$$

From this state space representation, we can extract and estimate the GARCH model parameters using statistical techniques, such as MLE-parameters that maximize the likelihood of the observed data given the model or the KF-algorithm iteratively updates the estimates of the state variables (volatility and returns) and can be utilized to derive estimates of the parameters.

Another method would be to use the first-order Taylor series (Servadio & Zanetti, 2021) of the conditional variance equation around its unconditional mean:

$$\sigma_t^2 \approx \omega + \beta_1 E\left[\varepsilon_{t-1}^2\right] + \alpha_1 E\left[\sigma_{t-1}^2\right] + \beta_1\left(\varepsilon_{t-1}^2 - E\left[\varepsilon_{t-1}^2\right]\right) + \alpha_1\left(\sigma_{t-1}^2 - E\left[\sigma_{t-1}^2\right]\right) \tag{7}$$

By applying the first-order Taylor series expansion, we create a linear approximation of the conditional variance (volatility) equation. This means we can break down our complex volatility model into something more manageable, which allows us to use standard statistical methods.

Yielding a linearized equation:

$$\sigma_t^2 \approx \left(+\beta_1 E\left[\varepsilon_{t-1}^2\right] + \alpha_1 E\left[\sigma_{t-1}^2\right]\right) + \beta_1 \varepsilon_{t-1}^2 + \alpha_1 \sigma_{t-1}^2 - \left(\beta_1 E\left[\varepsilon_{t-1}^2\right] + \alpha_1 E\left[\sigma_{t-1}^2\right]\right) \tag{8}$$

Once we have this simpler linear form, we can use the KF to estimate model parameters. In the context of commodity returns, it helps us continuously update our estimates of volatility as new price information comes in (Servadio & Zanetti, 2021).

3.7. MLE Based on Kalman Filter Outputs

The Kalman filter can be used to estimate unobserved states (like volatility) over time. The outputs from the Kalman filter (e.g., estimated volatility) can be used to compute the likelihood function. This approach leverages the dynamic nature of the Kalman filter, allowing for state estimation that evolves with the data. The methodology provides a recursive way to update estimates as new data arrives, which can better capture changes in volatility over time. Computing programs such as R and Python have excellent functionality to handle these algorithms.

3.8. The Theory Underpinning the Study

The tractable MLE for SS_GARCH models is supported through theoretical derivations that leverage properties of the KF and the structure of SSMs. SS_GARCH models are capable of incorporating non-Gaussian error distributions (e.g., STD, SSTD), which is crucial for modeling the fat tails and skewness observed in commodity returns (Aït-Sahalia et al., 2015). This model can be expressed in a state space framework, which allows us to describe the dynamics of both the observed variable (returns) and the unobservable state variable (volatility) in a convenient form given by Equations (1) and (3), where $\epsilon_t \mid F_{t-1} \sim N(0,1)$ with ϵ_t an independent standard normal.

Proof

To derive the MLE in a tractable form, we first notice that the joint distribution of observations can be expressed recursively. Below, we express the log-likelihood function;

$$L = \sum_{t=1}^{T} \log\left(f\left(r_t \mid \sigma_t \right) \right) \tag{9}$$

Substituting the conditional distribution of the observed returns (Gaussian) (Christoffersen et al., 2003), we get:

$$L = -\frac{T}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\left(\frac{\left(r_t - \mu\right)^2}{\sigma_t^2} \right) + \log\left(\sigma_t^2\right) \tag{10}$$

This formulation shows that the likelihood can be computed if we can derive σ_t^2 through its state space representation. By utilizing the KF's updating mechanism, we can iteratively compute σ_t^2 efficiently, facilitating the efficient maximization of the log-likelihood function to derive the necessary model parameters for the GARCH framework as required.

Given the structure of a GARCH model, we deal with latent volatility states that are not directly observable. In an SS_GARCH model, the estimation process is meticulously structured into three integral steps. Initially, the process begins with initialization, where initial estimates for both the state and variances are established to provide a foundation for subsequent calculations. Following this, the prediction phase employs the state equation to forecast the next state, particularly focusing on the anticipated volatility of the financial series.

Finally, the model enters the update step, where the most recent observation is integrated to refine the state estimates and update the likelihood, thus enhancing the precision of the model's forecasts. This methodical approach enhances the modeling of time-varying volatility, enabling a thorough representation of the dynamic traits intrinsic to financial returns.

3.9. Derivation of the Likelihood Function for Nonstationary GARCH Model

In this section, we demonstrate that SS_GARCH models can effectively accommodate nonstationary processes. This is achieved by defining the state transition matrix to encapsulate the dynamics of GARCH processes. In the scenario of a nonstationary GARCH process, the model integrates parameters that vary over time or utilizes a time index within the state equations to accurately capture the evolving nature of the underlying volatility, thereby enhancing the model's adaptability to fluctuations in the data.

The independence of ϵ_t enables inference about the entire sequence of σ_t and thus about the likelihood of the observed data to be performed despite potential nonstationarity, within the foundational data-generating mechanism. This is mathematically substantiated by employing the innovation process:

$$\epsilon_t = r_t - \mu - \hat{r}_{t|t-1} \tag{11}$$

where $\hat{r}_{t|t-1}$ represents the forecasted value derived from previous observations. and the Kalman gain adjusts the impact of ϵ_t. The tractable representation through the KF allows for efficient optimization of the likelihood function, ensuring that both stationary and nonstationary GARCH models can be estimated accurately.

To generalize the SS_GARCH model for non-Gaussian errors, we replace the Gaussian assumption with a more flexible distribution that captures skewness and kurtosis (Bianchi et al. 2011). Common candidates for non-Gaussian distributions include the STD, SSTD, GED and PIVD among many others. The extension involves allowing $\epsilon_t \sim Skew-t(\gamma, \varepsilon, \vartheta)$, γ is the degrees of freedom, ε represents the skewness and ϑ is the scale parameter.

Proposition 1: Given a non-Gaussian distribution for ϵ_t, the likelihood function needs to be articulated in relation to the probability density function (PDF) of the selected distribution. For instance, using the SSTD, the density function can be defined as:

$$f\left(r_t \mid \sigma_t\right) = \frac{2}{\sigma_t} t\left(\frac{r_t - \mu}{\sigma_t}, \gamma, \varepsilon\right) \tag{12}$$

Proof.

Since *L*, the Likelihood function given above, is;

$$L = \sum_{t=1}^{T} \log\left(f\left(r_t \mid \sigma_t\right)\right) \tag{13}$$

This equation represents the log-likelihood function, where $f\left(r_t \mid \sigma_t\right)$ is the probability density function, the pertubations of a chosen distribution (Christoffersen et al., 2003).

Substituting the skew-t density function gives:

$$L = \sum_{t=1}^{T} \left(\log(2) - \log(\sigma_t) + \log\left[t\left(\frac{r_t - \mu}{\sigma_t}, \gamma, \varepsilon \right) \right] \right) \quad (14)$$

The KF needs to be adapted to handle the non-Gaussian innovations. Instead of using the Gaussian innovation, we need to re-evaluate the Kalman gain and the state update equations to account for the PDF of the selected non-Gaussian distribution (Shumway & Stoffer, 2011). Firstly, the innovation process is modified as in (Shumway & Stoffer, 2011):

$$\epsilon_t = r_t - \hat{r}_{t|t-1}, \text{ where } \hat{r}_{t|t-1} = \mu + \hat{\sigma}_{t|t-1}.$$

Parameters are estimated using the modified log-likelihood function resulting from the chosen non-Gaussian distribution.

To address nonstationarity, the state transition equations can include time-varying parameters or local trends:

$$\sigma_t^2 = \varphi_0(t) + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (15)$$

Here, $\varphi_0(t)$ can be represented as a function of time or other variables, enabling the model to adjust to changes in the underlying data distribution over time.

The theoretical groundwork provided by the two propositions justifies the application of state space estimation techniques for GARCH models. The ability to leverage KF algorithms facilitates the efficient handling of latent states and parameter estimation processes, demonstrating the attractiveness of this approach in both stationary and nonstationary contexts (Theodossiou, 1998).

By modifying the observation equation to capture non-Gaussian behavior and incorporating suitable probability distributions, we can develop a SS_GARCH model capable of accommodating skewness and heavier tails. More recent theory around these distributions was presented by (Linton et al., 2010; Nikolaev et al., 2013) as in proposition 2 below.

Proposition 2: Given a non-Gaussian distribution for the error term ϵ_t the likelihood function $L(\theta)$ is defined as the product of the probability density functions (PDFs) of the chosen distribution evaluated at each observation, conditional on the past information (Cerqueti et al., 2020):

$$L(\theta) = \prod_{t=1}^{T} F(\epsilon_t \mid F_{t-1} : \theta) \quad (16)$$

where, θ is the vector of parameters, the state space GARCH model to be estimated, $F_{t-1} : \theta$ is the (Fan et al., 2014) conditional PDF of ϵ_t given the information set F_{t-1} and parameters θ and F_{t-1} represents the information set available up to time $t-1$, typically including past, (Tsay, 2005). Recent studies have used the skewed t-distribution to capture the skewness and heavy tails often observed in financial time series.

3.10. Data

The dataset consists of 5722 daily closing prices spanning from January 2, 2001, to October 26, 2023. This data was partitioned into an in-sample portion of 80%, while the remaining 20% was designated as out-of-sample data for model calibration. Additionally, commodity price data from 2001 to 2023 was similarly divided in a 4:1 ratio between in-sample and out-of-sample segments.

Actual return is a return which has occurred, calculated based on the historical data (Basira et al., 2024). To calculate the actual return of the stock from each sample by using daily commodity price;

$$R_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}}, \tag{17}$$

where R_{it} is the i^{th} commodity return on day t with price $P_{i,t}$, see (Basira et al., 2024).

The KFAS package in R is a powerful tool for estimating state space models, including GARCH models, which are particularly useful for modeling volatility in financial time series, such as commodity returns.

The KFAS package in R provides a basis for estimating state space representations of GARCH models, mostly useful for modeling commodity returns as a combination of underlying volatility processes and observation equations. Central to KFAS is the KF algorithm. It iteratively estimates latent state variables from observable returns and updates its estimates as new data becomes available. The estimation process begins with initializing the model using initial guesses for the state and variances, followed by making predictions based on state space equations. As new commodity return data is replicated, the model refines its parameter estimates through optimization procedures (Kantas et al., 2014).

3.11. State Space Formulation

The state space formulation of a GARCH model was presented in (1) and (2) in Section 1. Below, we present Kalman filter predictions and update formulations.

I) Prediction Step

Equation (17) predicts the conditional variance based on past observations. This equation is essential in predicting the state.

$$\hat{\sigma}^2_{t|t-1} = \omega + \alpha_1 r^2_{t-1} + \beta_1 \sigma^2_{t-1} \tag{18}$$

- **Predict the observation:**

After presicting the state, we predict the observation as given in Equation (18).

$$\hat{r}_{t|t-1} = \mu + \hat{\sigma}_{t|t-1}. \tag{19}$$

II) Update Step

Equation (19) calculates the innovation or prediction error, which is crucial for updating the state estimates.

$$\epsilon_t = r_t - \hat{r}_{t|t-1} \tag{20}$$

Equation (20) is used to update the estimation of the state variance:

$$\hat{\sigma}^2_t = \hat{\sigma}^2_{t|t-1} + K_t \epsilon_t \tag{21}$$

where K is the Kalman gain, which can be calculated based on the current estimate of the state and the error as in section 3.2.2.

3.12. Parameter Estimation

While the KF provides a mechanism for updating the state estimates, parameters ω, α_1 and β_1 are then estimated using optimization techniques that minimize the likelihood function derived from the recursive updates and observations.

4. Analysis and Results

4.1. Basic Statistics

Table 1 shows descriptive statistics, split into three categories as follows: the measures of central tendency, skewness and kurtosis, independents, normality, and stationary tests. A general increasing trend in commodity prices is evidenced by all positive return means. Skewed distributions and volatility clustering are again evident on all commodity return indices. Heavy tailed distributions are recommended.

Table 1. Summary statistics for commodity returns.

	Oil	Gld	Lith	Cot
A: Basic statistics				
Mini	−0.5100	−0.0982	−0.0597	−0.3184
Maxi	0.3200	0.0864	0.0777	0.1837
Mean	0.0002	0.0004	0.0002	0.0001
Std_Dev	0.0271	0.0108	0.0110	0.0282
Symmetry	−1.1611	−0.3376	0.6095	−1.5860
Heavy_Tail	39.1846	18.2400	9.7838	25.2066
Assumption tests (*p*-values)				
Ljung-Box test (5)	39.3 (<0.0001)	2.892 (<0.0001)	2097.7 (<0.0001)	183.5 (<0.0001)
Ljung-Box test (10)	47.3 (<0.0001)	17.84 (<0.0001)	3395.7 (<0.0001)	202.4 (<0.0001)
Jarque-Bera Test	30000 (<0.0001)	6395 (<0.0001)	31031 (<0.0001)	10000 (<0.0001)
Ljung-Box test^2 (5)	118.2 (<0.0001)	41.68 (<0.0001)	5270.4 (<0.0001)	665.4 (<0.0001)
Ljung-Box test^2 (10)	207.5 (<0.0001)	161.6 (<0.0001)	5701.8 (<0.0001)	1343 (<0.0001)
Stability tests				
Dicky-Fuller Test	−17.2 (<0.0001)	−18.3 (<0.0001)	−18.3 (<0.0001)	−31.1 (<0.0001)
Phillips-Perron Test	−5757 (<0.0001)	−5460 (<0.0001)	−5460 (<0.0001)	−3973 (<0.0001)
Kwiatkowski-Phillips-Schmidt-Shin Test	0.064 (>0.1000)	0.255 (>0.1000)	0.255 (>0.1000)	0.004 (>0.1000)

The ADF test and JB test for normality give significant *p*-values for all five commodities return series. This implies that the normality assumption is not met at all levels of significance. The Ljung-Box test, at Q(5), Q(10) and the squared values indicate that autocorrelation is insignificant for returns but highly significant in squared returns indicative of persistence in the volatility process of commodity returns. Heavy tailed asymmetric GARCH-type models are recommended under these circumstances.

The augmented Dickey-Fuller (ADF) and Phillips and Perron (PP) tests rejected the unit root hypothesis while the Kwiatkowski failed to reject the unit root hypothesis of stationarity. Hence commodity price returns are largely stationary in mean.

4.2. Exploratory Analysis of Price Trends, Return and Volatility Structure

Figures 1-4 show plots of the time series, return series, box plots, Q-Q plots and ACF of return of square returns for the five commodities. The returns of all the indices exhibit volatility clustering. In the same vein, all the ACFs and PACFs of the returns series the absence of long memory. All the QQ plots show that the tails of all the commodity indices' returns are heavier than the tails of normal distribution. They indicate the presence of heavy-tailed distributions and asymmetric dispersion of all the returns.

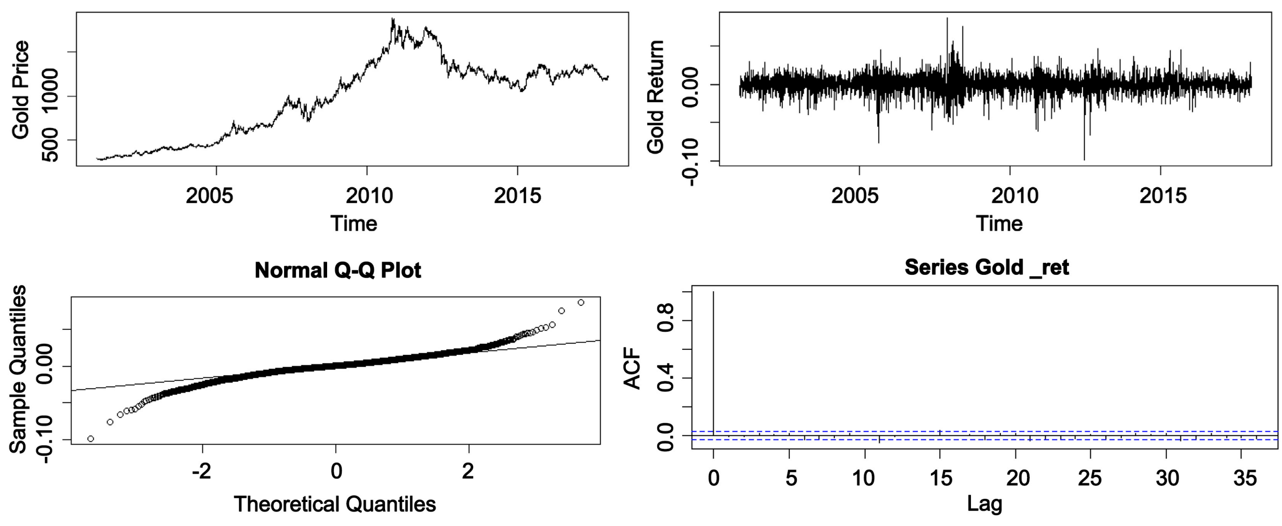
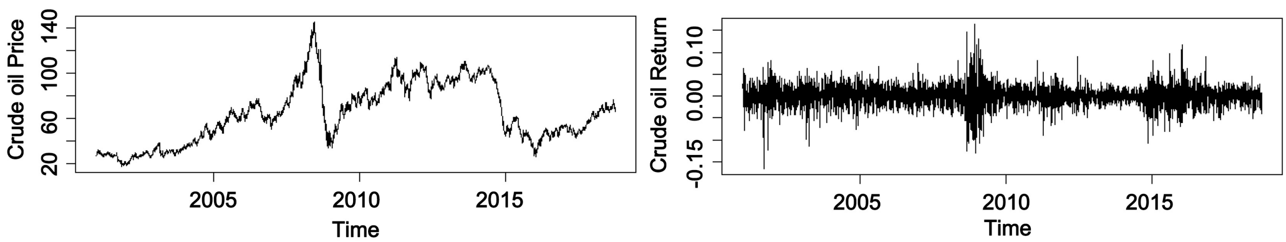


Figure 1. Composite analysis of commodity prices and returns: Left_top: Daily Price; Right_top: Gold_Returns; Left_bottom: Returns_QQ; Right-bottom: Returns_ACF for Gold.



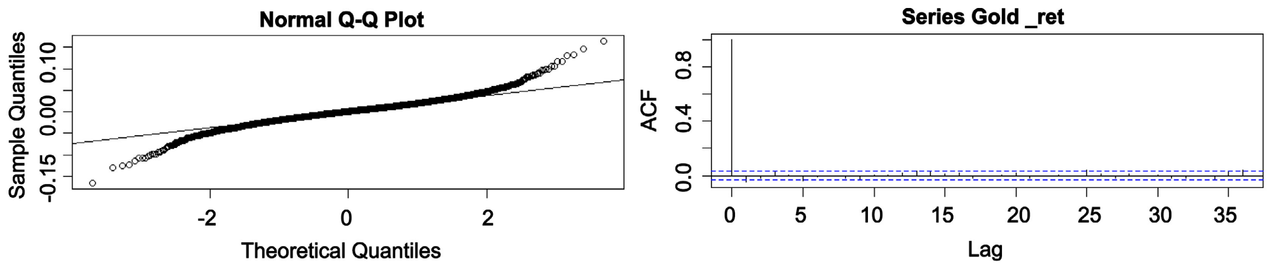


Figure 2. Composite analysis of commodity prices and returns: Left_top: Daily Price; Right_top: Crude_Returns; Left_bottom: Returns_QQ; Right_bottom: Returns_ACF for crude oil.

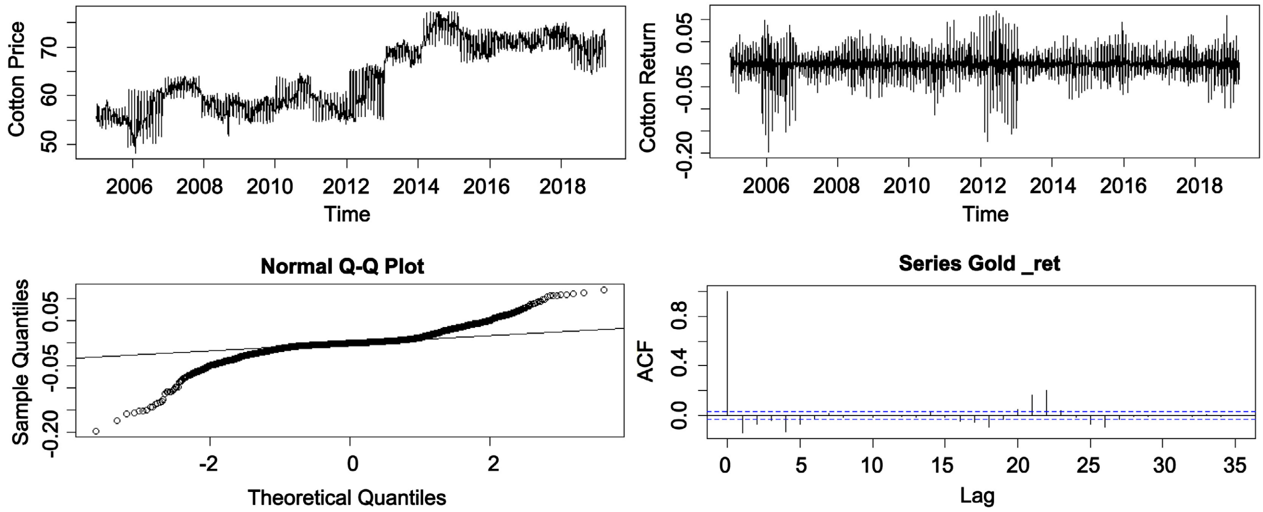


Figure 3. Composite analysis of commodity prices and returns: Left_top: Cotton_Price; Right_top: Cotton_Returns; Left_bottom: Returns_QQ; Rihgt_bottom: Returns_ACF for Cotton.

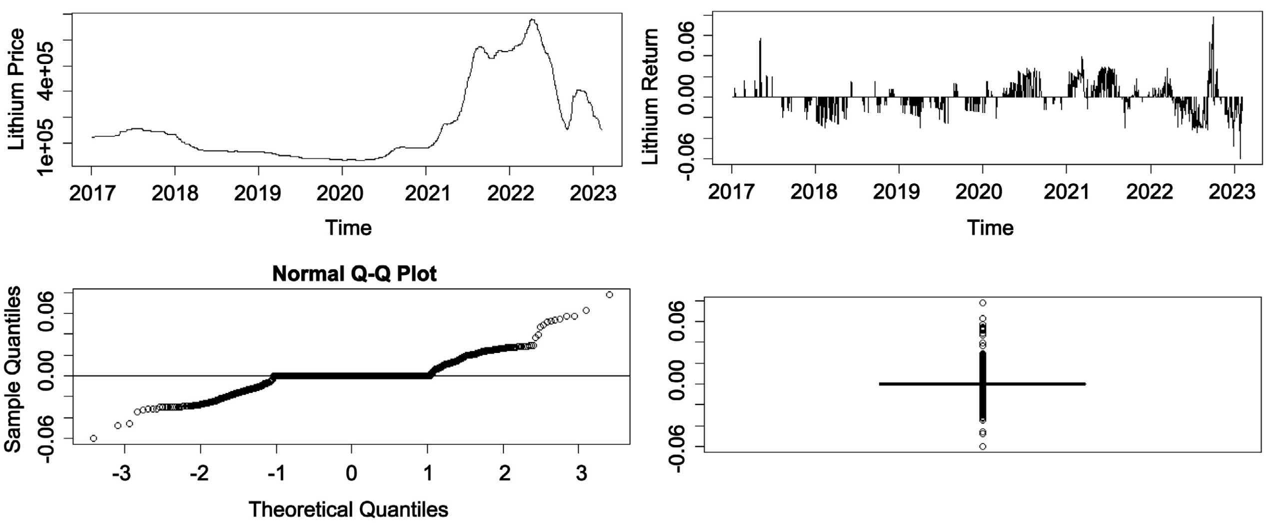


Figure 4. Composite analysis of commodity prices and returns: Top-left: Daily Price; Top-right: daily Returns; Bottom-left: Returns QQ; Bottom-right: Returns Box plot for Lithium.

4.3. Test for Long Memory

Based on the Lo Modified R/S test, there is no evidence of long memory among all the for commodity returns. We failed to reject the null hypothesis of no persistence (Table 2).

Table 2. Lo Modified R/S test.

Commodity	Mean	Volatility
Crud	1.2000	1.470
Gld	1.390	3.890
Cot	0.200	3.410
Lith	2.020	1.800

Note: t-value ranges are: 90%: [0.861, 1.747], 95%: [0.809, 1.862] and 99%: [0.721, 2.098].

4.4. Simulation Results

In financial modeling, Monte Carlo simulation is a powerful tool for evaluating complex statistical methods (Cobb & Charnes, 2004). Simulation is useful as it mimic reality and allows different analytical methods to be tested and different scenarios to be tried before using empirical data. By creating artificial datasets that mimic actual financial data, we can meticulously evaluate the efficacy and plausibility of our models.

In this section, we simulate a return series using specified simulation parameters. This return series is then used to estimate the GARCH parameters using both maximum likelihood estimation (MLE) and state space techniques. We will compare the methods to determine which approach most correctly reproduces the theoretical parameters used to simulate the series.

Table 3. Model performance with simulated GARCH (1, 1) data under normality assumptions (Sample size $n = 1000$).

Parameter	simulation parameters	MLE_GARCH (1, 1) estimates			SS_GARCH (1, 1) estimates		
		Est.	*p*-value	z-value	Est.	*p*-value	z-value
ω	0.1000	0.2097	<0.0001	4.8651	0.1476	<0.0001	4.4046
α_1	0.2000	0.1941	<0.0001	8.1253	0.3307	<0.0001	9.3796
β_1	0.700	0.5590	<0.0001	9.3742	0.6515	<0.0001	16.5533
AIC			−2.4365			−2.7582	

Table 4. Model performance with simulated GARCH (1, 1) data under STD assumptions (Sample size $n = 1000$).

Parameter	simulation parameters	MLE_GARCH (1, 1) estimates			SS_GARCH (1, 1) estimates		
		Est.	*p*-value	z-value	Est.	*p*-value	z-value
ω	0.1000	0.2096	0.0003	5.2967	0.1441	<0.0001	4.4867
α_1	0.2000	0.1933	<0.0001	3.596	0.3185	<0.0001	7.3842
β_1	0.700	0.5609	<0.0001	9.7468	0.6654	<0.0001	13.5767
AIC			−2.6218			−2.6231	

Table 5. Model performance with simulated GARCH (1, 1) data under SSTD assumptions (Sample size $n = 1000$).

Parameter	simulation parameters	MLE_GARCH (1, 1) estimates			SS_GARCH (1, 1) estimates		
		Est.	*p*-value	z-value	Est.	*p*-value	z-value
ω	0.1000	0.1965	<0.0001	5.9834	0.1215	<0.0001	5.8365
α_1	0.2000	0.1912	<0.0001	4.7361	0.2564	<0.0001	9.6537
β_1	0.700	0.5758	<0.0001	16.4854	0.6875	<0.0001	16.7843
AIC			−2.6218			−2.7926	

Table 6. Model performance with simulated GARCH (1, 1) data under PIVD assumptions (Sample size $n = 1000$).

Parameter	simulation parameters	MLE_GARCH (1, 1) estimates			SS_GARCH (1, 1) estimates		
		Est.	*p*-value	z-value	Est.	*p*-value	z-value
ω	0.1000	0.2096	0.0003	5.2967	0.1441	<0.0001	4.4867
α_1	0.2000	0.1933	<0.0001	3.596	0.3185	<0.0001	7.3842
β_1	0.700	0.5609	<0.0001	9.7468	0.6654	<0.0001	13.5767
AIC			−2.3980			−2.6834	

The simulation results presented in **Tables 3-6** highlight the parameter estimates obtained from both the MLE GARCH (1, 1) and the State Space GARCH (1, 1) methods, alongside the theoretical parameters used for simulation with different innovations assumptions among them, the STD, the SSTD and the PIVD.

Estimates for the parameters ω, α_1, and β_1 from both methods are compared to the theoretical values used in the simulation. For instance, the estimated ω from MLE is significantly higher than the theoretical value (0.1000), indicating a potential overestimation in the model's volatility baseline. In contrast, the SSG method produces an estimate of 0.1441, which is also considerably lower and can be rounded to the theoretical value. The *p*-values for both methods are uniformly low (<0.0001), indicating that the estimated parameters are statistically significant. This suggests that both models can reliably capture the underlying volatility structure. The z-values further reinforce this significance, with both methods showing strong statistical evidence for their parameter estimates.

The findings indicate that the State Space method outperforms Maximum Likelihood Estimation (MLE) for parameter estimation in GARCH (1, 1) models at different error distributions. Notably, heavy-tailed distributions, such as the STD, SSTD and PIVD, demonstrated superior performance in terms of the AIC (−2.7582, −2.7926 and −2.6834, respectively) and the accuracy of parameter estimates relative to the theoretical values used during data simulation for the SSG. These results underscore the State Space approach as a more robust choice for time series analysis in financial contexts, where precise volatility estimation is essential, especially in the presence of heavy-tailed data characteristics. Overall, the simulation results validate the efficacy of using both MLE and state space methodologies for estimating GARCH model parameters.

4.5. Empirical Results

After analyzing the simulated data, focus is now on empirical results. This section deals with validation procedures applied to real-life commodity data to assess the practicality and robustness of the coupled skewed GARCH and state space GARCH models across various error distributions. By examining actual market data, we aim to gain insights into how effectively our models capture the complexities of financial returns, including volatility clustering and asymmetry, thus enhancing our understanding of their performance in real-world scenarios.

Table 7. Parameter estimates for the GARCH model, Crude oil returns.

		MLE_GARCH (1, 1)			SS_GARCH (1, 1)		
Innovations	Parameters	Estimate	*p*-value	t-values	Estimate	*p*-value	t-values
NORM	ω	0.0000	0.4091	0.8254	0.0000	0.6738	1.2653
	α_1	0.0584	0.0037	2.8987	0.1432	<0.0001	3..4637
	β_1	0.9359	<0.0001	40.1582	0.7474	<0.0001	31.3851
	AIC	−4.7568			−4.9360		
STD	ω	0.0000	0.7110	0.8254	0.0000	0.6535	1.2653
	α_1	0.05093	0.0198	2.8987	0.1035	<0.0001	3..4637
	β_1	0.9447	<0.0001	40.1582	0.8203	<0.0001	31.3851
	AIC	−4.8701				−4.98562	
SSTD	ω	0.0000	0.2546	0.1,645	0.0000	0.0523	2.0895
	α_1	0.08437	<0.0001	6.7438	0.1548	<0.0001	8.7341
	β_1	0.8453	<0.0001	23.6547	0.8002	<0.0001	19.7252
	AIC		−4.9794			−4.9887	

Table 8. Parameter estimates for the GARCH model, Gold returns.

		MLE_GARCH (1, 1)			SS_GARCH (1, 1)		
Innovations	Parameters	Estimate	*p*-value	t-values	Estimate	*p*-value	t-values
NORM	ω	0.0000	0.0309	2.1583	0.0000	0.0214	2.795
	α_1	0.0372	<0.0001	14.4088	0.0678	<0.0001	11.6543
	β_1	0.9558	<0.0001	38.6249	0.8153	<0.0001	26.9476
	AIC	−6.2756			−6.5340		
STD	ω	0.0000	0.2534	1.8452	0.0000	0.0207	3.8534
	α_1	0.0339	<0.0001	12.6431	0.0678	<0.0001	19.6381
	β_1	0.9618	<0.0001	9.8756	0.8153	<0.0001	33.2351
	AIC	−6.2787				−6.7806	
SSTD	ω	0.0000	0.2058	1.9845	0.0000	0.0213	2.9134
	α_1	0.0311	<0.0001	2.6614	0.0815	<0.0001	11.7923
	β_1	0.9274	<0.0001	12.9760	0.8016	<0.0001	27.7615
	AIC		−6.5392			−6.7349	

From Table 7, on Crude oil, the SSTD_SS_GARCH (1, 1) model exhibits all parameter estimates as statistically significant and achieves the lowest Akaike Information Criterion (AIC) value, indicating its superior performance compared to the other models.

Table 9. Parameter estimates for the GARCH model, Lithium returns.

		MLE_GARCH (1, 1)			SS_GARCH (1, 1)		
Innovations	Parameters	Estimate	*p*-value	t-values	Estimate	*p*-value	t-values
NORM	ω	0.0000	0.718472	0.36054	0.0000	0.2453	0.3367
	α_1	0.1490	<0.0001	4.15568	0.1765	<0.0001	5.4896
	β_1	0.8246	<0.0001	10.40838	0.7954	<0.0001	14.7352
	AIC	−6.7859			−6.8101		
STD	ω	0.0000	0.8184	0.6214	0.0000	0.718472	0.36054
	α_1	0.15918	0.2894	0. 8453	0.15918	<0.0001	4.15568
	β_1	0.8336	<0.0001	6.8568	0.8336	<0.0001	10.40838
	AIC	−6.7867				−6.7867	
SSTD	ω	0.0000	0.4634	0.3258	0.0000	0.5723	0.3548
	α_1	0.1076	0.2341	0. 7546	0.1054	<0.0001	7.9546
	β_1	0.7983	<0.0001	4.8237	0.8462	<0.0001	12.6381
	AIC		−6.7859			−6.7823	

Table 10. Parameter estimates for the GARCH model, Cotton returns.

		MLE_GARCH(1, 1)			SS_GARCH (1, 1)		
Innovations	Parameters	Estimate	*p*-value	t-values	Estimate	*p*-value	t-values
NORM	ω	0.0000	0.27757	1.08580	0.0000	0.0009	1.4538
	α_1	0.0243	<0.0001	15.2553	0.0243	<0.0001	11.8467
	β_1	0.9703	<0.0001	49.0348	0.9703	<0.0001	44.9568
	AIC	−4.8066			−4.7456		
STD	ω	0.0000	<0.0562	2.241	0.1232	<0.0001	1.08580
	α_1	0.0215	<0.0001	8.573	0.2143	<0.0001	12.7865
	β_1	0.9184	<0.0001	5.6231	0.6987	<0.0001	43.3562
	AIC	−4.8942			−4.8674		
SSTD	ω	0.0000	<0.0001	1.6389	0.1257	<0.0001	4..9237
	α_1	0.0243	<0.0001	17.7452	0.2068	<0.0001	11.6982
	β_1	0.9703	<0.0001	35.6231	0.6885	<0.0001	23.4276
	AIC		−4.7756			**−4.8132**	
PIVD	ω	0.0000	<0.0001	4.7254	0.000	<0.0001	5.8265
	α_1	0.03862	<0.0001	11.8754	0.25433	<0.0001	12.0032
	β_1	0.9023	<0.0001	19.6753	0.6439	<0.0001	21.7643
	AIC		−4.7813			**−4.8945**	

Table 8 exhibits parameter estimates for Gold. The STD_SS_GARCH (1, 1) model demonstrated notable effectiveness for Gold, outperforming all other models assessed. Its superior performance is evidenced by its statistical significance across all parameter estimates and its favorable performance metrics, which collectively highlight its reliability and robustness compared to the alternatives.

Table 9 shows parameter estimates for Lithium. The STD_MLE_GARCH (1, 1) model exhibited strong performance, surpassing the other models evaluated. Its effectiveness is reflected in the statistical significance of its parameter estimates and favorable performance metrics, which collectively underscore its reliability and robustness.

As depicted in Table 10, the PIVD_SS_GARCH (1, 1) model outperformed all other models for Cotton returns. Its superior performance is highlighted by the statistical significance of its parameter estimates and advantageous performance metrics (Ferreira et al., 2017).

4.6. Discussion of Results

Estimating GARCH (1, 1) models for commodity returns through both state space and maximum likelihood estimation (MLE) techniques offer significant insights into the dynamics of these markets. MLE relies on several critical assumptions concerning the disturbances distribution and the model's functional structure. If these assumptions do not hold, the resulting estimates can be biased or inconsistent. Additionally, MLE is sensitive to the choice of initial parameter values; poorly chosen initial values may cause the optimization algorithm to converge to local minima, leading to suboptimal estimates (Audet et al., 2024). The results obtained for the four commodity returns analyzed in this study acme the significance of understanding the characteristics of the data underlying data when selecting an estimation method. The adaptability and robustness of the state space method make it particularly suitable for modeling commodities with long memory properties (Karanasos et al., 2021). In contrast, MLE may be more effective in simpler, stationary contexts. As commodity markets continue to evolve, ongoing research into advanced modeling techniques that can accommodate various characteristics will be essential for accurate risk assessment, estimation, and forecasting.

To address the limitations of state space GARCH models, it's essential to consider several key factors. These models often exhibit computational complexity, requiring sophisticated estimation techniques that can lead to longer processing times, especially with large datasets (Bollerslev, 1986). They are also sensitive to initial conditions, where inadequate starting values may cause convergence issues and affect reliability. Additionally, the risk of model specification errors can result in biased estimates, highlighting the need for thorough diagnostics (Engle, 2001). While state space models provide flexibility, their interpretability may be hindered by complexity, limiting practical application (Bollerslev, 1986). Finally, they generally require larger datasets for reliable estimation, which can compromise performance in data-limited scenarios. Addressing these considerations offers a more comprehensive understanding of state space GARCH models and their trade-offs.

4.7. Diagnosis of Models

4.7.1. Evaluation of Estimates

In **Table 11**, the best-performing model's performance and reliability were assessed using out of sample data to ensure that they accurately capture the commodity volatility patterns and can generalize well to new data. The comparable estimates across these models indicate that they did not suffer from overfitting, a common issue where a model captures noise instead of underlying patterns, leading to poor generalization to new data (Dufays et al., 2022; Dufays & Rombouts, 2020).

Table 11. Model validation using out of sample data.

Commodity	Distribution	ω	*p*-value	t-statistic	α_1	*p*-value	t-statistic	β_1	*p*-value	t-statistic
Crude	SSTD	0.0000	0.3546	1.6545	0.1432	<0.0001	3.5741	0.7538	<0.0001	25.8769
Gold	SSTD	0.0000	0.9217	2.1834	0.1674	<0.0001	14.6429	0.8012	<0.0001	12.7634
Lithium	STD	0.0000	0.7140	0.3670	0.1480	<0.0001	4.4863	0.7835	<0.0001	9.7256
Cotton	PIVD	0.0000	0.0091	10.8013	0.0257	<0.0001	16.8236	0.8638	<0.0001	33.4863

These models maintained consistent estimates in out-of-sample applications, demonstrating an effective balance between complexity and generalization. Models that are overly complex may excel in training but underperform on new data, characterized by high variance and low bias (Mukhamediev et al., 2022). In contrast, overly simplistic models risk underfitting by failing to capture essential trends.

The findings suggest that these models are reliable for practical applications, showcasing robustness against overfitting while effectively generalizing to new datasets. This aligns with best practices in model validation and mitigates the risk of overfitting (Ying, 2019).

4.7.2. Forecasts Evaluation

The forecasting evaluation metrics presented in **Table 12** compare the performance of the State Space GARCH (SS_GARCH) and Maximum Likelihood Estimation GARCH (MLE_GARCH) models across different commodities: Crude oil, Gold, Lithium, and Cotton.

The SS_GARCH model consistently shows lower MSE values for most commodities compared to the MLE_GARCH model, particularly for Crude oil (0.0100 vs. 0.0102) and Lithium (0.0015 vs. 0.0012). Lower MSE indicates that the SS_GARCH model generally provides more accurate forecasts, which is significant for risk management and investment decisions (Pourkhanali et al., 2020). The same pattern was observed in MAE, particularly in Cotton, the SS_GARCH MAE (0.0108) is slightly lower than that of MLE_GARCH (0.0115). Regarding the DM Test, the *p*-values below 0.05 (e.g., for Crude oil, Gold and Lithium) suggest that the SS_GARCH model significantly outperforms the MLE_GARCH model in forecast accuracy for those commodities, reinforcing the validity of the SS_GARCH approach. The R^2 values indicate the proportion of variance explained by the model. While both models show similar R^2 values, the SS_GARCH model's slightly lower values may reflect its more dynamic nature in capturing time-varying volatility (Dufays & Rombouts, 2020). The other metrics in Table 12 tell the same story.

Table 12. Forecasting evaluation metrics of the SS_GARCH and MLE_GARCH (1, 1) models.

Measure	Method	Crud		Gld		Lith		Cot	
		STD	SSTD	STD	SSTD	STD	SSTD	STD	SSTD
MSE	SS	0.0100	0.0020	0.0028	0.0103	0.0015	0.0028	0.0028	0.0101
	MLE	0.0102	0.0033	0.0030	0.0102	0.0012	0.0030	0.0031	0.0102
MAE	SS	0.0086	0.0186	0.0086	0.0086	0.0076	0.0086	0.0108	0.0086
	MLE	0.0093	0.0167	0.0082	0.0093	0.0074	0.0082	0.0115	0.0093
DM Test (p-value)	SS	0.0631	0.0324	0.0465	0.0235	0.0753	0.0630	0.0603	0.0304
	MLE	0.0823	0.0472	0.501	0.0423	0.0543	0.0468	0.0437	0.0219
Directional Accuracy (%)	SS	56	59	57	61	54	58	59	
	MLE	53	56	57	58	55	56	55	
MCS	SS MLE	Included		Included		Included		Included	
TIC	SS	0.5776	0.6623	0.6027	0.5776	0.6817	0.6027	0.5853	0.5776
	MLE	0.5657	0.6497	0.6458	0.5657	0.6962	0.6458	0.5719	0.5657
R^2	SS	0.0816	0.08132	0.0892	0.0816	0.07884	0.0892	0.07965	0.0816
	MLE	0.0986	0.0872	0.0968	0.0986	0.0801	0.0968	0.0846	0.0986

4.7.3. Model Selection

In Table 13, the SSTD_SS_GARCH (1, 1) model is particularly effective for crude oil and gold, as evidenced by its performance across the metrics. The PIVD_SS_GARCH (1, 1) model performs well for cotton. Although the STD_MLE_GARCH (1, 1) model is well-suited for lithium, all models exhibit different levels of effectiveness in capturing the intricate volatility patterns of commodity returns.

4.7.4. Sensitivity Analysis

Below we present sensitivity analysis of the best fitted model.

Table 13. Top performing models out of sample test.

Parameter	Gld-SSTD_ SS_GARCH (1, 1)			Crud-SSTD_SS_ GARCH (1, 1)			Cot-PIVD_SSP_ GARCH (1, 1)		
	Est	p-value	z-value	Est	p-value	z-value	Est	p-value	z-value
ω	0.000	<0.0001	5.8654	0.000	<0.0001	4.5683	0.000	<0.0001	4.6784
α_1	0.1201	<0.0001	11.1584	0.1834	<0.0001	10.0042	0.2558	<0.0001	14.8645
β_1	0.7832	<0.0001	14.0723	0.8012	<0.0001	17.5437	0.6439	<0.0001	22.2564
AIC		−4.8889			−5.003			−4.8348	

5. Conclusion

In conclusion, this study demonstrates the effectiveness of the state space approach and the KF in estimating GARCH-type models for commodity returns, particularly in scenarios with limited data. The superior performance of models such as the SSTD_SS-GARCH (1, 1) for crude oil and gold, and the PIVD_SS_GARCH (1, 1) for cotton, underscores the potential of these methods to enhance forecasting accuracy.

From a policy perspective, the findings suggest that regulators and policymakers can leverage these advanced modeling techniques to better understand and predict commodity price volatility. This understanding is crucial for implementing effective policies that stabilize markets and mitigate risks associated with price fluctuations.

For practitioners in risk management and investment decision-making, the study highlights the importance of adopting SSMs in their analytical toolkit. By utilizing models that account for complex volatility dynamics, investors can make more informed decisions, optimize their portfolios, and develop robust hedging strategies.

Future research should consider alternative state space methodologies to enhance GARCH-type model estimation for commodity returns. For instance, particle filters could effectively address non-linearities and non-Gaussian noise, providing a robust complement to the EKF employed in this study. Additionally, the Unscented KF (UKF) may be explored for its capacity to handle abrupt price movements in commodities such as crude oil. Bayesian state space models could also offer improved parameter estimation in data-limited contexts, while Hidden Markov Models (HMM) might be useful for capturing regime shifts in commodity prices, particularly for gold.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Aduda, J., Weke, P., Ngare, P., & Mwaniki, J. (2016). Financial Time Series Modelling of Trends and Patterns in the Energy Markets. *Journal of Mathematical Finance, 6,* 324-337. https://doi.org/10.4236/jmf.2016.62027

Aït-Sahalia, Y., Cacho-Diaz, J., & Laeven, R. J. A. (2015). Modeling Financial Contagion Using Mutually Exciting Jump Processes. *Journal of Financial Economics, 117,* 585-606. https://doi.org/10.1016/j.jfineco.2015.03.002

Audet, C., Bouchet, P.-Y., & Bourdin, L. (2024). *A Derivative-Free Approach to Partitioned Optimization (Version 2).* https://doi.org/10.48550/ARXIV.2407.05046

Azman, S., Pathmanathan, D., & Thavaneswaran, A. (2022). Forecasting the Volatility of Cryptocurrencies in the Presence of COVID-19 with the State Space Model and Kalman Filter. *Mathematics, 10,* Article No. 3190. https://doi.org/10.3390/math10173190

Basira, K., Dhliwayo, L., Chinhamu, K., Chifurira, R., & Matarise, F. (2024). Estimation and Prediction of Commodity Returns Using Long Memory Volatility Models. *Risks, 12,* Article No. 73. https://doi.org/10.3390/risks12050073

Beg, A. B. M. R. A., & Anwar, S. (2014). Detecting Volatility Persistence in GARCH Models in the Presence of the Leverage Effect. *Quantitative Finance, 14,* 2205-2213. https://doi.org/10.1080/14697688.2012.716162

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics, 31,* 307-327. https://doi.org/10.1016/0304-4076(86)90063-1

Bulut, E. (2024). Market Volatility and Models for Forecasting Volatility. In J. C. Rouco, & P. C. N. Figueiredo (Eds.), *Business Continuity Management and Resilience: Theories, Models, and Processes* (pp. 220-248). IGI Global. https://doi.org/10.4018/979-8-3693-1658-0.ch010

Cerqueti, R., Giacalone, M., & Mattera, R. (2020). Skewed Non-Gaussian GARCH Models for Cryptocurrencies Volatility Modelling. *Information Sciences, 527,* 1-26. https://doi.org/10.1016/j.ins.2020.03.075

Chevallier, J., & Ielpo, F. (2013). *The Economics of Commodity Markets*. Wiley. https://doi.org/10.1002/9781118710098

Choudhry, T., & Wu, H. (2008). Forecasting Ability of GARCH vs Kalman Filter Method: Evidence from Daily UK Time-Varying Beta. *Journal of Forecasting, 27,* 670-689. https://doi.org/10.1002/for.1096

Christoffersen, C., Bollano, E., Lindegaard, M. L. S., Bartels, E. D., Goetze, J. P., Andersen, C. B., & Nielsen, L. B. (2003). Cardiac Lipid Accumulation Associated with Diastolic Dysfunction in Obese Mice. *Endocrinology, 144,* 3483–3490. https://doi.org/10.1210/en.2003-0242

Cobb, B. R., & Charnes, J. M. (2004). Real Options Volatility Estimation with Correlated Inputs. *The Engineering Economist, 49,* 119-137. https://doi.org/10.1080/00137910490453392

Dufays, A., & Rombouts, J. V. K. (2020). Relevant Parameter Changes in Structural Break Models. *Journal of Econometrics, 217,* 46-78. https://doi.org/10.1016/j.jeconom.2019.10.008

Dufays, A., Houndetoungan, E. A., & Coën, A. (2022). Selective Linear Segmentation for Detecting Relevant Parameter Changes. *Journal of Financial Econometrics, 20,* 762-805. https://doi.org/10.1093/jjfinec/nbaa032

Engle, R. (2001). GARCH 101: The Use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives, 15,* 157-168. https://doi.org/10.1257/jep.15.4.157

Ewing, B. T., & Malik, F. (2017). Modelling Asymmetric Volatility in Oil Prices under Structural Breaks. *Energy Economics, 63,* 227-233. https://doi.org/10.1016/j.eneco.2017.03.001

Fan, J., Han, F., & Liu, H. (2014). Challenges of Big Data analysis. *National Science Review, 1,* 293-314. https://doi.org/10.1093/nsr/nwt032

Ferreira, G., Navarrete, J. P., Rodríguez-Cortés, F. J., & Mateu, J. (2017). Estimation and Prediction of Time-Varying GARCH Models through a State-Space Representation: A Computational Approach. *Journal of Statistical Computation and Simulation, 87,* 2430-2449. https://doi.org/10.1080/00949655.2017.1334778

Hall, S. G. (1990). *A Note on the Estimation of GARCH-M Models Using the Kalman Filter*. Bank of England.

Heydari, M. H., Avazzadeh, Z., & Cattani, C. (2020). Taylor's Series Expansion Method for Nonlinear Variable-Order Fractional 2D Optimal Control Problems. *Alexandria Engineering Journal, 59,* 4737-4743. https://doi.org/10.1016/j.aej.2020.08.035

Jategaonkar, R. V. (2015). *Flight Vehicle System Identification: A Time-Domain Method-*

ology (2nd ed.). American Institute of Aeronautics and Astronautics, Inc. https://doi.org/10.2514/4.102790

Kalman, R. E. (1960). A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering, 82,* 35-45. https://doi.org/10.1115/1.3662552

Kalman, R. E., & Bucy, R. S. (1961). New Results in Linear Filtering and Prediction Theory. *Journal of Basic Engineering, 83,* 95-108. https://doi.org/10.1115/1.3658902

Kantas, N., Doucet, A., Singh, S. S., Maciejowski, J., & Chopin, N. (2014). On Particle Methods for Parameter Estimation in State-Space Models. *Statistical Science, 30,* 328-351. https://doi.org/10.1214/14-sts511

Karanasos, M., Yfanti, S., & Christopoulos, A. (2021). The Long Memory HEAVY Process: Modeling and Forecasting Financial Volatility. *Annals of Operations Research, 306,* 111-130. https://doi.org/10.1007/s10479-019-03493-8

Kim, C.-J., & Nelson, C. R. (2017). *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. MIT Press. https://books.google.com/books?hl=en&lr=&id=Grt-NEAAAQBAJ&oi=fnd&pg=PA1&dq=estimatin+and+prediction+of+volatility+models+using+state+space+methodology&ots=ocXJi7fQ9J&sig=gVDHN7B5Jr29_Te3Keq8EriTBLo

Kleppe, T. S., Liesenfeld, R., Moura, G. V., & Oglend, A. (2022). Analyzing Commodity Futures Using Factor State-Space Models with Wishart Stochastic Volatility. *Econometrics and Statistics, 23,* 105-127. https://doi.org/10.1016/j.ecosta.2021.03.008

Kumar, D. (2018). Volatility Prediction: A Study with Structural Breaks. *Theoretical Economics Letters, 08,* 1218-1231. https://doi.org/10.4236/tel.2018.86080

Likassa, H. T., Chen, D., Nadarajah, S., Sema, M., Chen, J. K., Temesgen, S. et al. (2025). A Multivariate GARCH Model with Time-Varying Correlations: What Do Inflation Data Show in Ethiopia? *Computational Economics*. https://doi.org/10.1007/s10614-025-10951-y

Lima, J. F., Pereira, F. C., Gonçalves, A. M., & Costa, M. (2023). Bootstrapping State-Space Models: Distribution-Free Estimation in View of Prediction and Forecasting. *Forecasting, 6,* 36-54. https://doi.org/10.3390/forecast6010003

Linton, O., Pan, J., & Wang, H. (2010). Estimation for a Nonstationary Semi-Strong Garch(1, 1) Model with Heavy-Tailed Errors. *Econometric Theory, 26,* 1-28. https://doi.org/10.1017/s0266466609090598

Metsileng, L. D., Moroke, N. D., & Tsoku, J. T. (2021). Modeling the Exchange Rate Volatility Using the BRICS Garch-Type Models. *International Journal of Financial Research, 12,* 166-179. https://doi.org/10.5430/ijfr.v12n5p166

Mukhamediev, R. I., Popova, Y., Kuchin, Y., Zaitseva, E., Kalimoldayev, A., Symagulov, A. et al. (2022). Review of Artificial Intelligence and Machine Learning Technologies: Classification, Restrictions, Opportunities and Challenges. *Mathematics, 10,* Article No. 2552. https://doi.org/10.3390/math10152552

Nikolaev, N. Y., Boshnakov, G. N., & Zimmer, R. (2013). Heavy-Tailed Mixture GARCH Volatility Modeling and Value-at-Risk Estimation. *Expert Systems with Applications, 40,* 2233-2243. https://doi.org/10.1016/j.eswa.2012.10.038

Omar, Z. (2019). *State-Space Models with GARCH Errors: Application to Health Data*. Master's Thesis, McGill University (Canada). https://search.proquest.com/openview/3e30a2e200a78a4b536e54ac6b69ace0/1?pq-origsite=gscholar&cbl=18750&diss=y

Ossandón, S., & Bahamonde, N. (2013). A New Nonlinear Formulation for GARCH Models. *Comptes Rendus. Mathématique, 351,* 235-239. https://doi.org/10.1016/j.crma.2013.02.014

Pappas, A., & Boukas, N. (2025). Examining Impact of Inflation and Inflation Volatility on Economic Growth: Evidence from European Union Economies. *Economies, 13,* Article No. 31. https://doi.org/10.3390/economies13020031

Pourkhanali, A., Keith, J., & Zhang, X. (2020). *Conditional Heteroscedasticity Models with Time-Varying Parameters: Estimation and Asymptotics*. https://www.monash.edu/business/ebs/research/publications/ebs/wp15-2021.pdf

Ramos, A., Valladão, D., & Street, A. (2024). *Time Series Analysis by State Space Learning (Version 1)*. https://doi.org/10.48550/ARXIV.2408.09120

Rzayev, K., & Ibikunle, G. (2018). The Information Content of High Frequency Trading Volume. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3106599

Servadio, S., & Zanetti, R. (2021). Differential Algebra-Based Multiple Gaussian Particle Filter for Orbit Determination. *Journal of Optimization Theory and Applications, 191,* 459-485. https://doi.org/10.1007/s10957-021-01934-8

Sharma, S., Aggarwal, V., & Yadav, M. P. (2021a). Comparison of Linear and Non-Linear GARCH Models for Forecasting Volatility of Select Emerging Countries. *Journal of Advances in Management Research, 18,* 526-547. https://doi.org/10.1108/jamr-07-2020-0152

Sharma, S., Aggarwal, V., & Yadav, M. P. (2021b). Comparison of Linear and Non-Linear GARCH Models for Forecasting Volatility of Select Emerging Countries. *Journal of Advances in Management Research, 18,* 526-547. https://doi.org/10.1108/jamr-07-2020-0152

Shumway, R. H., & Stoffer, D. S. (2011). *Time Series Analysis and Its Applications*. Springer. https://doi.org/10.1007/978-1-4419-7865-3

Swanepoel, C., & Fliers, P. T. (2021). The Fuel of Unparalleled Recovery: Monetary Policy in South Africa between 1925 and 1936. *Economic History of Developing Regions, 36,* 213-244. https://doi.org/10.1080/20780389.2021.1945436

Theodossiou, P. (1998). Financial Data and the Skewed Generalized T Distribution. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.65037

Tsay, R. S. (2005). *Analysis of Financial Time Series*. Wiley. https://doi.org/10.1002/0471746193

Vega-Muratalla, V. O., Ramírez-Márquez, C., Lira-Barragán, L. F., & Ponce-Ortega, J. M. (2024). Review of Lithium as a Strategic Resource for Electric Vehicle Battery Production: Availability, Extraction, and Future Prospects. *Resources, 13,* Article No. 148. https://doi.org/10.3390/resources13110148

Wong, K. F. K., Galka, A., Yamashita, O., & Ozaki, T. (2006). Modelling Non-Stationary Variance in EEG Time Series by State Space GARCH Model. *Computers in Biology and Medicine, 36,* 1327-1335. https://doi.org/10.1016/j.compbiomed.2005.10.001

Ying, X. (2019). An Overview of Overfitting and Its Solutions. *Journal of Physics: Conference Series, 1168,* Article ID: 022022. https://doi.org/10.1088/1742-6596/1168/2/022022

Zhu, Y., Lang, Z. Q., & Guo, Y. (2021). Nonlinear Model Standardization for the Analysis and Design of Nonlinear Systems with Multiple Equilibria. *Nonlinear Dynamics, 104,* 2553-2571. https://doi.org/10.1007/s11071-021-06429-9