Nonlinear Analysis: Modelling and Control, Online First, 1–24
https://doi.org/10.15388/namc.2026.31.47080

Vilnius University Press

News-Augmented GARCH model: Theoretical properties and simulation-based hyperparameter analysis*

Diana Vereškaitė, Jurgita Markevičiūtė, Andrius Buteikis

Institute of Applied Mathematics, Vilnius University
Naugarduko str. 24, Vilnius
veresk.diana@gmail.com

Received: December 19, 2025 / **Revised:** April 10, 2026 / **Published online:** May 26, 2026

Abstract. Volatility is a key measure of financial risk, and GARCH models are widely used to describe its dynamics. However, they do not account for the influence of news sentiment, which can significantly shape market volatility. Recently proposed News-Augmented GARCH model addresses it by incorporating sentiment signals in a nonlinear, asymmetric, and multiplicative form. This paper examines its theoretical properties and performs a simulation-based hyperparameter study. The analysis establishes the existence of a unique and causal solution, derives a stability condition, and evaluates model sensitivity and parameter recovery under controlled scenarios. Results demonstrate robust performance across various settings and provide guidance for informed hyperparameter selection and evaluation, enhancing model's reliability for empirical applications.

Keywords: volatility prediction, GARCH, news sentiment, asymmetry, stability condition, simulation study.

1 Introduction

In the realm of stock market, analyzing the risk component is a primary objective. The volatile part of returns is undefined and therefore considered the risky portion. Modeling volatility allows investors to quantify risk, supporting more informed decision-making. Among the various methods available, GARCH (Generalized AutoRegressive Conditional Heteroskedasticity) model, proposed by Bollerslev back in 1993 [9], stands out as one of the most widely accepted and commonly used approaches. It effectively captures volatility by analyzing stock's own historical data, providing a structured way to understand and forecast risk. However, stock market volatility is not shaped by historical price movements alone. It is strongly influenced by traders' sentiments, which are often

*Publication/Research is funded by Research Council of Lithuania under the Programme "University Excellence Initiatives" of the Ministry of Education, Science and Sports of the Republic of Lithuania (measure No. 12-001-01-01-01 "Improving the Research and Study Environment"), project No. S-A-UEI-23-11.

not directly observable. News serve as a viable proxy for these sentiments, as it reflects collective expectations and market reactions. Furthermore, news sentiment often exhibits asymmetric effects, with negative news typically having different influence on volatility than positive news. Increasing volume of information available through social and traditional media has significantly impacted the stock market, amplifying the influence of sentiment on market dynamics. Fortunately, advancements in machine learning have made it possible to better quantify the sentiment of these sources, differentiate between positive and negative effect, and therefore allow integration of news data into traditional volatility modeling.

While existing research has explored integrating sentiment into volatility models, most approaches adopt additive exogenous variables, assuming a linear relationship. However, this overlooks potential nonlinear interactions between sentiment and volatility, which a multiplicative approach captures more effectively. Therefore, the study focuses on a recently proposed News-Augmented GARCH model [26], which addresses this gap by incorporating asymmetric news sentiment in a multiplicative way. Together with nonlinearity and asymmetry involved, it enhances the model's flexibility in capturing the complex interactions between sentiment and market fluctuations and also increases predictive power under selected empirical setting. Nevertheless, it also introduces additional complexity due to the indirect estimation of parameters, increased computational burden, and sensitivity to sentiment data quality. While this model offers a promising approach, its mathematical properties, such as stability, uniqueness of the solution, and parameter sensitivity, remain underexplored. This research aims to bridge this gap by providing formal analysis for model stability and uniqueness, while also conducting a simulation-based hyperparameter analysis to better understand the capabilities and limitations of the model.

The rest of the paper is organized as follows. Section 2 reviews the literature. In Section 3, News-Augmented GARCH model is presented together with novel theoretical framework. Section 4 describes simulation study, providing both methodology and results. Finally, discussion and conclusion are provided in Section 5.

2 Literature review

Integrating news information into volatility analysis represents a distinct and significant area of research. First attempts to treat news information were through asymmetry without modeling news impact independently. The first to mention asymmetry was Black in 1976 [8]. He observed that stock price declines tend to lead to higher volatility than stock price increases of the same magnitude. This work kicked-off further works that tried to model the so called leverage effect (different impact of bad and good news). The first attempt to incorporate asymmetry idea as information of news into volatility modeling was by Nelson in 1991 [23], where the logarithm of conditional variance responds asymmetrically to positive and negative shocks. This approach laid the foundation for later asymmetric models and became a standard in asymmetric volatility modeling. It was followed by the widely used GJR-GARCH model of Glosten, Jagannathan, and Runkle (1993) [18], which introduces asymmetry through an indicator capturing the sign of past shocks.

Although these models capture different effect of news in a certain way, they rely on historical price and return data only. To have a better understanding of news, different types of proxies were incorporated in volatility modeling. One of the pioneers of using an external factor as a proxy for news was Lamoureux with his work from 1990 [21]. The paper examines the relationship between stock volatility and trading volume, showing that including volume as an explanatory variable often eliminates GARCH effects, suggesting that it captures the flow of new information. The definition of a news proxy was further improving: [16] explored macroeconomic announcements as binary or categorical exogenous variables. Kalev (2002) [20] introduced a more detailed approach by incorporating the timing and type of information events, finding that different categories exhibit different persistence and that overnight information spills into the opening period. In summary, early works typically relied on news volume [6, 28] or social-media volume [29], but more sophisticated methods soon appeared. Mitra and Mitra in their book [22] summarize sentiment construction techniques such as lexicon-based scoring [34] and NLP-based methods [25, 33]. With advancements in NLP, companies such as RavenPack and Thomson Reuters began providing sentiment extraction services for financial news and social media, which made news incorporation more easy and accurate. The reader is referred to the book for further details on sentiment construction, while this work focuses solely on news augmentation within GARCH models.

The rapid development of GARCH models incorporating news sentiment likely began gaining momentum around 2020, coinciding with advancements in machine learning and natural language processing. This enabled construction of news indices from various sources, such as traditional media [27], social media [1,11], scheduled versus unscheduled news [24], and macroeconomic [26] or firm-specific information [25]. In most works, news indices are included additively, either in the mean equation [30] or in the variance equation [19], and sometimes via regression with fitted GARCH volatility [13]. Studies using these methods consistently found a significant relationship between GARCH modeling and news information across crypto, oil, regional markets, sectors, and shock periods such as COVID-19. The pandemic further accelerated research, as markets became especially sensitive to news and shifts in sentiment. For instance, Chen (2022) [10] investigates how Twitter sentiment related to COVID-19 and vaccines affects pharmaceutical stock volatility using a TGARCH-M model, showing that negative sentiment has a notably stronger influence. Trabelsi (2024) [31] studies the Tunisian market with a GJR-GARCH model, finding that both COVID-19 case numbers and government interventions significantly increase market volatility. Similar evidence arises in other regions' financial markets like China [12] or Taiwan [19] both finding dominant negative sentiment effects.

Despite methodological variety, most literature incorporates news additively, imposing a linear structure that may not represent real market dynamics. Alternative approaches are far less explored. One such direction is time-varying GARCH models beginning with paper [3], where volatility follows a multiplicative structure. Multiplicative component is an endogenous smooth-transition function based on past data. It allows the model to adapt to changing market conditions but without external inputs, where transitions usually have log-linear or logistic form. A more flexible, nonlinear approach was proposed by Sadik [25], who integrated external asymmetric news information into the GARCH

model. It is incorporated in a multiplicative form, allowing the impacts of news shocks to interact with volatility dynamics in a more unified framework without assuming linear relationship. Model suggests a good trade between complexity and simplicity, it sparking the interest and being a great candidate for further research. Follow-up work adapted this news augmentation to a one-factor crude-oil model [26] and applied it to microblog sentiment [7]. This paper builds upon the News-Augmented GARCH model and fills the gap of mathematical and numerical understanding of it.

3 Model overview

3.1 News-Augmented GARCH

Before delving deeper into the model, let us review the existing work proposed by Sadik [25]. The model is unique in the way that news sentiment is formed and included into GARCH framework. Firstly, news sentiment index is included in a multiplicative way, which is novel and important because it allows the model to dynamically scale volatility based on the intensity of market sentiment and has a practical interpretation. Secondly, news sentiment is formed in a nonlinear and asymmetric way, which covers the complexity that occurs in news information arrival. Due to the following reasons, the model is of main interest in this paper as it has potential to cover the complexity of the question raised with the adequate amount of parameters.

The model is formulated as follows: Let $\{r_t\}_{t\in\mathbb{Z}}$ denote a discrete-time return process, where t represents the time index. Let $\mathscr{P}_t \in [0,1]$ and $\mathscr{N}_t \in [-1,0]$ denote the positive and negative news indices at time t, respectively. The news function is defined as

$$f(\mathscr{P}_t, \mathscr{N}_t) = a + \frac{b}{2}\left(\frac{\mathrm{e}^{\kappa\mathscr{P}_t}-1}{\mathrm{e}^{\kappa\mathscr{P}_t}+1} - \frac{\mathrm{e}^{\gamma\mathscr{N}_t}-1}{\mathrm{e}^{\gamma\mathscr{N}_t}+1}\right),$$
$$f_t := f(\mathscr{P}_t, \mathscr{N}_t)$$

with hyperparameters $a, b > 0$, $\kappa, \gamma \geqslant 0$, where a and b control the upper and lower bounds of the function, and κ, γ correspond to positive and negative news, respectively. Further, the return process is specified as

$$r_t = \mu + \epsilon_t, \quad \epsilon_t = z_t\sigma_t, \quad z_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1),$$

and the conditional variance equation is

$$\sigma_t^2 = f_{t-1}\big(\omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2\big),$$

where $\omega > 0$, $\alpha, \beta \geqslant 0$.

Parameters ω, α, β serve as regular GARCH parameters, which are scaled by a multiplicative function. This function is bounded by $[a,\ a+b]$, determining how much volatility is reduced or amplified. Parameters κ and γ control sensitivity to positive and negative news accordingly and define the curvature. An example is shown in Fig. 1, where it is seen that function value increases as positive and negative impact is increasing, stating

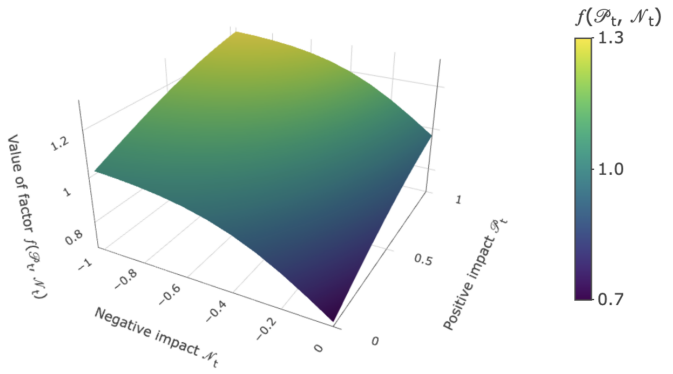

Figure 1. Surface plot of the news function $f(\mathscr{P}_t, \mathscr{N}_t)$ of positive $(\mathscr{P}_t)$ and negative $(\mathscr{N}_t)$ news indices, showing the nonlinear response to news shocks for parameter values $a = 0.7$, $b = 0.7$, $\kappa = 2$, and $\gamma = 4$.

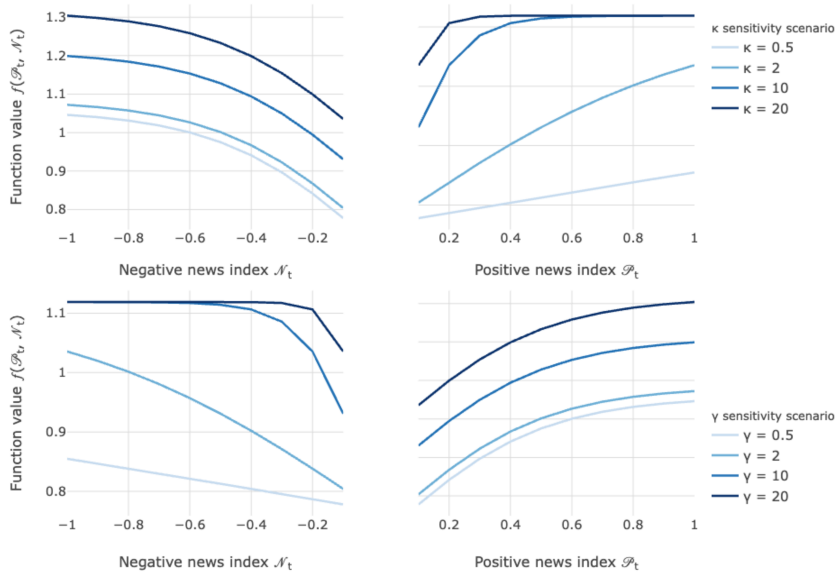

Figure 2. Sensitivity analysis of the news function $f(\mathscr{P}_t, \mathscr{N}_t)$ with respect to κ and γ. The upper panels show the effect of varying κ, while the lower panels illustrate changes as γ varies, with other parameters fixed. $\mathscr{N}_t = -0.1$ when plotting against $\mathscr{P}_t$, and $\mathscr{P}_t = 0.1$ when plotting against $\mathscr{N}_t$.

that increasing news impact increase the volatility. It is also observed in Fig. 2 that as the news sensitivity parameter increases, the function value rises, but its rate of change slows down at higher news impact levels. This reflects the market becoming accustomed to strong news, and the slowdown is governed by κ and γ.

Having the form of the function, the next step is the construction of the news index. In the original paper, the index is derived from RavenPack data, where sentiment scores

from news and social media are processed using NLP techniques [35]. Relevant events are extracted, assigned sentiment values, and transformed using an exponential decay to reflect diminishing impact over time. Positive and negative news are treated separately, and the resulting scores are aggregated and rescaled to the range $[-1, 1]$ to form the final daily news index for each asset.

The main advantage of the News-Augmented GARCH model is that it enhances volatility modeling by incorporating news sentiment, allowing a more comprehensive view of market volatility beyond historical price movements. One of its strengths is capturing asymmetric news impact, as positive and negative news often affect volatility differently. Additionally, the flexibility of the model is enhanced by the nonlinear form of $f(\mathscr{N}_t, \mathscr{P}_t)$, which can be adjusted through four hyperparameters, providing adaptability to different market conditions and datasets. Moreover, the multiplicative form adds nonlinear news effects with clear interpretation, since a and b define the ceiling and floor of volatility responses.

Despite its advantages, the model comes with certain limitations or incompleteness. Firstly, mathematical properties have not been explored. Another gap is lack of clear view of model capabilities under different parameter settings and general model sensitivity view. One of the main challenges is that the parameters in the function $f(\mathscr{N}_t, \mathscr{P}_t)$ are not analytically estimated, requiring additional computations to optimize the model. This increases complexity in determining the best parameter values. Additionally, parameters κ, γ, which govern good and bad news sentiments, do not have straight-forward interpretation. Furthermore, the News-Augmented GARCH model is limited by the classic GARCH framework, which may restrict its application in certain scenarios. The model is also sensitive to the quality of news sentiment quantification, inaccurate sentiment analysis can distort predictions, leading to potentially misleading conclusions. These limitations motivated further work specifically in providing mathematical properties and performing simulations for thorough hyperparameter analysis.

3.2 Theoretical properties

To ensure the robustness and applicability of the proposed model, the existence of a unique and causal solution is established under certain conditions.

Theorem 1. *Let $\{r_t\}_{t\in\mathbb{Z}}$ be a discrete-time stochastic process satisfying the News-Augmented GARCH*(1,1) *equations*

$$\begin{aligned} r_t &= \mu + \epsilon_t, \qquad \epsilon_t = z_t\sigma_t, \\ \sigma_t^2 &= f_{t-1}\big(\omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2\big) \end{aligned} \tag{1}$$

with

$$z_t \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$$

and

$$f_t := f(\mathscr{N}_t, \mathscr{P}_t) = a + 0.5b\left(\frac{\mathrm{e}^{\kappa\mathscr{P}_t} - 1}{\mathrm{e}^{\kappa\mathscr{P}_t} + 1} - \frac{\mathrm{e}^{\gamma\mathscr{N}_t} - 1}{\mathrm{e}^{\gamma\mathscr{N}_t} + 1}\right),$$

where the sequences $\{\mathscr{P}_t\}_{t\in\mathbb{Z}}$ *and* $\{\mathscr{N}_t\}_{t\in\mathbb{Z}}$ *are assumed to be bounded exogenous sequences with* $-1 \leqslant \mathscr{N}_t \leqslant 0$ *and* $0 \leqslant \mathscr{P}_t \leqslant 1$ *for all* $t \in \mathbb{Z}$ *and* $a, b > 0$, $\kappa, \gamma \geqslant 0$, $\omega > 0$, $\alpha, \beta \geqslant 0$. *Let* σ_t^{*2} *be defined by the following equation*:

$$\sigma_t^{*2} = \omega f_{t-1} \sum_{i=0}^{\infty} \prod_{j=1}^{i} f_{t-j-1} A_{t-j}, \tag{2}$$

where $A_t = \alpha z_t^2 + \beta$ *and* $\prod_{j=1}^{0} f_{t-j-1} A_{t-j} = 1$. *Then Eq.* (2) *is a unique and causal solution of Eq.* (1) *if* $(\alpha + \beta)(a + b) < 1$.

Proof. The procedure begins with repeated substitution of σ_{t-j}^2 and ϵ_{t-j}^2, $j = 1, \ldots, k$, into Eq. (1). The first substitution results in the following equation:

$$\begin{aligned}
\sigma_t^2 &= f_{t-1}\big(\omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2\big) = f_{t-1}\big(\omega + \big[\alpha z_{t-1}^2 + \beta\big]\sigma_{t-1}^2\big) \\
&= f_{t-1}\big(\omega + \big[\alpha z_{t-1}^2 + \beta\big] f_{t-2}\big[\omega + \alpha\epsilon_{t-2}^2 + \beta\sigma_{t-2}^2\big]\big) \\
&= f_{t-1}\big(\omega + \omega f_{t-2}\big[\alpha z_{t-1}^2 + \beta\big] + f_{t-2}\big[\alpha z_{t-1}^2 + \beta\big]\big[\alpha\epsilon_{t-2}^2 + \beta\sigma_{t-2}^2\big]\big)
\end{aligned}$$

Substituting once more yields the following:

$$\begin{aligned}
\sigma_t^2 &= f_{t-1}\big(\omega + \omega f_{t-2}\big[\alpha z_{t-1}^2 + \beta\big] \\
&\quad + f_{t-2}\big[\alpha z_{t-1}^2 + \beta\big]\big[\alpha z_{t-2}^2 + \beta\big] f_{t-3}\big[\omega + \alpha\epsilon_{t-3}^2 + \beta\sigma_{t-3}^2\big]\big) \\
&= f_{t-1}\big(\omega + \omega f_{t-2}\big[\alpha z_{t-1}^2 + \beta\big] + \omega f_{t-2} f_{t-3}\big[\alpha z_{t-1}^2 + \beta\big]\big[\alpha z_{t-2}^2 + \beta\big] \\
&\quad + f_{t-2} f_{t-3}\big[\alpha z_{t-1}^2 + \beta\big]\big[\alpha z_{t-2}^2 + \beta\big]\big[\alpha z_{t-3}^2 + \beta\big]\sigma_{t-3}^2\big).
\end{aligned}$$

Continuing this process results in the following expression:

$$\begin{aligned}
\sigma_t^2 &= f_{t-1}\left(\omega + \omega \sum_{i=1}^{k} \prod_{j=1}^{i} f_{t-j-1}\big[\alpha z_{t-j}^2 + \beta\big]\right) + \sigma_{t-k-1}^2 \prod_{j=1}^{k+1} f_{t-j}\big[\alpha z_{t-j}^2 + \beta\big] \\
&= \omega f_{t-1} \sum_{i=0}^{k} \prod_{j=1}^{i} f_{t-j-1} A_{t-j} + \sigma_{t-k-1}^2 \prod_{j=1}^{k+1} f_{t-j} A_{t-j}.
\end{aligned}$$

Then the difference between σ_t^{*2} and σ_t^2 is

$$\sigma_t^{*2} - \sigma_t^2 = \omega f_{t-1} \sum_{i=k+1}^{\infty} \prod_{j=1}^{i} f_{t-j-1} A_{t-j} - \sigma_{t-k-1}^2 \prod_{j=1}^{k+1} f_{t-j} A_{t-j}.$$

Since $a, b > 0$, $\alpha, \beta \geqslant 0$, and $\mathbf{E}(A_t) = \alpha + \beta > 0$, it follows that

$$\begin{aligned}
&\mathbf{E}\big|\sigma_t^{*2} - \sigma_t^2\big| \\
&\quad \leqslant \omega f_{t-1} \sum_{i=k+1}^{\infty} \prod_{j=1}^{i} f_{t-j-1}(\alpha + \beta) + \mathbf{E}\big(\sigma_{t-k-1}^2\big) \prod_{j=1}^{k+1} f_{t-j}(\alpha + \beta) \\
&\quad = \omega f_{t-1} \sum_{i=k+1}^{\infty} (\alpha + \beta)^i \prod_{j=1}^{i} f_{t-j-1} + (\alpha + \beta)^{k+1} \mathbf{E}\big(\sigma_{t-k-1}^2\big) \prod_{j=1}^{k+1} f_{t-j}.
\end{aligned}$$

Given the functional form of f_t, its boundedness is first established. Observe that

$$\frac{\mathrm{e}^x - 1}{\mathrm{e}^x + 1} \in (-1, 1) \quad \forall x \in \mathbb{R}.$$

Using $\mathscr{P}_t \in [0, 1]$, $\mathscr{N}_t \in [-1, 0]$, and $\kappa, \gamma \geqslant 0$, it follows that

$$0 \leqslant \left(\frac{\mathrm{e}^{\kappa \mathscr{P}_t} - 1}{\mathrm{e}^{\kappa \mathscr{P}_t} + 1} - \frac{\mathrm{e}^{\gamma \mathscr{N}_t} - 1}{\mathrm{e}^{\gamma \mathscr{N}_t} + 1} \right) \leqslant 2,$$

which implies

$$a \leqslant f_t \leqslant a + b.$$

Since $f_{t-j-1} \leqslant a + b$, it follows that

$$\prod_{j=1}^{i} f_{t-j-1} \leqslant (a + b)^i.$$

Hence,

$$\begin{aligned}
&\mathbf{E}\big|\sigma_t^{*2} - \sigma_t^2\big| \\
&\quad \leqslant \omega(a + b) \sum_{i=k+1}^{\infty} (\alpha + \beta)^i \prod_{j=1}^{i} (a + b) + (\alpha + \beta)^{k+1} \mathbf{E}\big(\sigma_{t-k-1}^2\big) \prod_{j=1}^{k+1} (a + b) \\
&\quad = \omega(a + b) \sum_{i=k+1}^{\infty} \big[(\alpha + \beta)(a + b)\big]^i + \big[(\alpha + \beta)(a + b)\big]^{k+1} \mathbf{E}\big(\sigma_{t-k-1}^2\big).
\end{aligned}$$

If $(\alpha + \beta)(a + b) < 1$, then

$$\begin{aligned}
&\mathbf{E}\big|\sigma_t^{*2} - \sigma_t^2\big| \\
&\quad \leqslant \omega(a + b) \frac{\big[(\alpha + \beta)(a + b)\big]^{k+1}}{1 - (\alpha + \beta)(a + b)} + \big[(\alpha + \beta)(a + b)\big]^{k+1} \mathbf{E}\big(\sigma_{t-k-1}^2\big).
\end{aligned}$$

In this case, $\big[(\alpha + \beta)(a + b)\big]^{k+1} \to 0$ as $k \to \infty$, which means that

$$\mathbf{E}\big|\sigma_t^{*2} - \sigma_t^2\big| \to 0 \quad \text{as } k \to \infty.$$

Consequently, the parameters α, β, a, and b are required to satisfy one of the following restrictions:

- $\alpha + \beta \geqslant 1$ and $a + b < 1$ such that $(\alpha + \beta)(a + b) < 1$.
- $\alpha + \beta < 1$ and $a + b \geqslant 1$ such that $(\alpha + \beta)(a + b) < 1$.
- $\alpha + \beta < 1$ and $a + b < 1$. □

Corollary 1. *Let $\{r_t\}_{t \in \mathbb{Z}}$ be an NA-GARCH$(1, 1)$ process, as defined in Theorem* 1*, and let $(\alpha + \beta)(a + b) < 1$. Then*

$$\mathbf{E}\big(\sigma_t^2\big) \leqslant \frac{\omega(a + b)}{1 - (\alpha + \beta)(a + b)}.$$

Proof. If $(\alpha+\beta)(a+b)<1$, then

$$\begin{aligned}\mathbf{E}\big(\sigma_t^2\big) &= \mathbf{E}\left(\omega f_{t-1}\sum_{i=0}^{\infty}\prod_{j=1}^{i} f_{t-j-1}A_{t-j}\right) = \omega f_{t-1}\sum_{i=0}^{\infty}\prod_{j=1}^{i} f_{t-j-1}\mathbf{E}\big(\alpha z_{t-j}^2+\beta\big)\\ &= \omega f_{t-1}\sum_{i=0}^{\infty}(\alpha+\beta)^i\prod_{j=1}^{i} f_{t-j-1} \leqslant \omega(a+b)\sum_{i=0}^{\infty}\big[(\alpha+\beta)(a+b)\big]^i\\ &= \frac{\omega(a+b)}{1-(\alpha+\beta)(a+b)}.\end{aligned}$$

□

4 Simulation study

This section presents a simulation study examining the behavior of the News-Augmented GARCH model. The main objective is to assess how variations in hyperparameters affect estimation accuracy and forecasting performance. To address this, a simulation-based methodology is employed, which allows model behavior to be analyzed in a controlled setting where true parameter values are known. Synthetic time series are generated from predefined parameter sets, the model is fitted to the simulated data, and performance is evaluated using standard metrics. Results are then aggregated across parameter combinations to reveal general patterns and sensitivity characteristics. The section is structured into two parts: the simulation methodology and the resulting analysis, which examines parameter–metric relationships, parameter recovery, and individual parameter effects. Together, these dimensions provide a comprehensive view of the model's behavior, sensitivity, and overall reliability. All simulations are implemented using the R programming language.

4.1 Methodology

This subsection describes selected simulation design used to analyze the hyperparameters of the News-Augmented GARCH model.

4.1.1 Simulation pseudo-code

Step 1: Simulate news sentiment. For simulation purposes, positive and negative news impact indices are generated from half-normal distributions for each time point $t = 1,\dots,T$ as

$$\mathcal{P}_t \sim \text{HalfNormal}^{+}, \qquad \mathcal{N}_t \sim \text{HalfNormal}^{-},$$

and rescaled using min-max normalization. $\mathcal{P}_t$ and $\mathcal{N}_t$ are treated as exogenous variables. The half-normal distribution is used solely to generate simulated sentiment paths. Letting $X_t = |Z_t|$ and $Y_t = -|Z_t|$ with $Z_t \sim \mathcal{N}(0,\sigma^2)$,

$$\mathcal{P}_t = 0.01 + 0.99\,\frac{X_t-\min(X)}{\max(X)-\min(X)}, \qquad \mathcal{N}_t = -1 + 0.99\,\frac{Y_t-\min(Y)}{\max(Y)-\min(Y)},$$

so that $\mathcal{P}_t \in (0, 1]$ and $\mathcal{N}_t \in [-1, 0)$. The news impact function is then defined by

$$f_t = a + 0.5b\left(\frac{\mathrm{e}^{\kappa\mathcal{P}_t} - 1}{\mathrm{e}^{\kappa\mathcal{P}_t} + 1} - \frac{\mathrm{e}^{\gamma\mathcal{N}_t} - 1}{\mathrm{e}^{\gamma\mathcal{N}_t} + 1}\right).$$

The selection of the distribution shape follows the intuition that high-impact news events are usually relatively rare, while the majority of news items carry lower informational impact. Some empirical analyses of financial text also suggest that sentiment signals are often weak in magnitude. For example, large-scale analyses of financial news datasets show that most news content is objective or neutral, with only a small portion expressing strong sentiment [14]. This suggests that aggregated sentiment scores tend to concentrate around weak or near-neutral values, whereas strongly polarized observations occur less frequently. At the same time, the exact distribution of sentiment indices is not universal and may depend on the data source [15], the natural language processing method, the sentiment construction procedure, or the sector analyzed [2]. Nevertheless, for simulation purposes, the half-normal specification serves as a simple and transparent representation of a possible sentiment impact pattern.

Step 2: Simulate returns. The simulation spans $T = 2500$ trading days, each divided into 78 equal intraday intervals, which translates to 5-minute intervals required for realized volatility calculation. The 5-minute frequency captures sufficient intraday volatility and limits computational burden. This corresponds to approximately 10 years of data. For the data generating process, conditional variance is initialized using the following expression:

$$\sigma_1^2 = \frac{\omega\,(a+b)}{1-(\alpha+\beta)(a+b)},$$

to ensure positivity, and i.i.d. standard normal innovations $z_t \sim \mathcal{N}(0, 1)$ are generated. Returns are then obtained as

$$\epsilon_t = z_t\sigma_t, \qquad r_t = \mu + \epsilon_t$$

with the conditional variance evolving according to

$$\sigma_t^2 = f_{t-1}\big(\omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2\big).$$

GARCH parameters α and β are defined conditionally to satisfy stability. Returns are later aggregated from intraday intervals to daily frequency for further modeling.

Step 3: Evaluate News-Augmented GARCH across parameter combinations. For each combination of tuning parameters (a, b, κ, γ), the News-Augmented GARCH model is fitted to the simulated series. The hyperparameters are drawn from a predefined grid $\mathcal{G}$:

$$a, b \in \{0.1, 0.4, 0.7, 1, 2, 3\}, \qquad \kappa, \gamma \in \{0.1, 0.5, 2, 4, 5, 10, 20\},$$
$$\mathcal{G} = \big\{(a, b, \kappa, \gamma)\big\}.$$

Each element of the grid $\mathcal{G}$ corresponds to a parameter combination indexed by $j = 1, \ldots, N_{\mathrm{comb}}$ with $j \in \mathcal{G}$. Parameters ω, α, β are evaluated using Maximum Likelihood

Estimation method. For each parameter combination j, the simulation is repeated $R = 100$ times to account for stochastic variation in returns. Then $\hat{\sigma}_t^{2\,(j,r)}$ is estimated for each parameter combination j and simulation repetition r, and performance metrics are calculated. In particular, the RMSE with respect to realized volatility is computed as

$$\mathrm{RMSE}_{\mathrm{RV}}^{(j,r)} = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\left(\mathrm{RV}_t^{(r)} - \hat{\sigma}_t^{(j,r)}\right)^2},$$

where

$$\mathrm{RV}_t^{(r)} = \sqrt{\sum_{k=1}^{n} \ln^2 \frac{P_{t,k}^{(r)}}{P_{t,k-1}^{(r)}}},$$

where $P_{t,k}^{(r)}$ denotes the simulated price at intraday interval k of day t in simulation repetition r, and $n = 78$ represents the number of intraday observations within a trading day. The realized volatility $\mathrm{RV}_t^{(r)}$ is therefore computed separately for each simulation repetition r. Additional metrics include $\mathrm{RMSE}_{\mathrm{HV}}$, $\mathrm{MAE}_{\mathrm{RV}}$, $\mathrm{MAE}_{\mathrm{HV}}$, and AIC. Further discussion on metric selection is provided in Section 4.1.2. This procedure is repeated for each repetition $r = 1, \ldots, 100$, providing a robust evaluation of model performance across the entire parameter grid.

Step 4: Aggregate metrics. For each parameter combination j, the simulation is repeated $R = 100$ times. For each repetition r, a new dataset is generated by redrawing the news indices $\mathcal{P}_t, \mathcal{N}_t$ and the innovation sequence z_t, producing a new return series.

Metrics are then averaged over repetitions to obtain

$$\overline{\mathrm{RMSE}}_{\mathrm{RV}}^{(j)} = \frac{1}{R}\sum_{r=1}^{R} \mathrm{RMSE}_{\mathrm{RV}}^{(j,r)}$$

with analogous expressions for $\overline{\mathrm{RMSE}}_{HV}^{(j)}$, $\overline{\mathrm{MAE}}_{\mathrm{RV}}^{(j)}$, $\overline{\mathrm{MAE}}_{HV}^{(j)}$, $\overline{\mathrm{AIC}}^{(j)}$. The aggregation is needed to smooth out the stochastic variability of individual simulations.

Step 5: Calculate standard deviation. For a given parameter $H \in \{a, b, \kappa, \gamma\}$, define the subset of the grid $\mathcal{G}$ corresponding to a specific value h of that parameter as

$$\mathcal{G}_{H=h} = \{\mathbf{x} \in \mathcal{G} \colon \mathbf{x}_H = h\}.$$

Then the mean standard deviation of the metric across all parameter combinations in $\mathcal{G}_{H=h}$ is

$$\overline{\mathrm{StDev}}_H = \frac{1}{M_H}\sum_{j=1}^{M_H} \mathrm{StDev}_j,$$

where j indexes the parameter combinations in $\mathcal{G}_{H=h}$ and M_H denotes the total number of such combinations. Specifically, $M_H = 294$ when $H \in \{a, b\}$, and $M_H = 252$ when $H \in \{\kappa, \gamma\}$.

Step 6: Select best parameter set. The optimal parameter combination j^* is identified as the one minimizing the aggregated RMSE:

$$j^* = \arg\min_j \overline{\mathrm{RMSE}}^{(j)}_{\mathrm{RV}}.$$

Step 7: Sensitivity analysis. To assess the impact of the true model parameters on forecast performance, a one-at-a-time sensitivity analysis is performed. True parameters of the data-generating process are denoted by $(a_{\mathrm{true}}, b_{\mathrm{true}}, \kappa_{\mathrm{true}}, \gamma_{\mathrm{true}})$. Starting from a baseline parameter configuration, each hyperparameter $H \in \{a_{\mathrm{true}}, b_{\mathrm{true}}, \kappa_{\mathrm{true}}, \gamma_{\mathrm{true}}\}$ is varied independently, while the remaining parameters are kept fixed. For each parameter H, five values are considered (including the baseline), namely

$$a_{\mathrm{true}}, b_{\mathrm{true}} \in \{0.2, 0.7, 1, 2, 3\},$$
$$\kappa_{\mathrm{true}}, \gamma_{\mathrm{true}} \in \{0.5, 2, 4, 10, 20\}.$$

These values correspond to the true parameters of the data-generating process and are therefore independent of the estimation grid considered in Steps 3–6. In total, this yields $4 \times 5 = 20$ true parameter configurations, for each of which Steps 1–6 are repeated.

Additionally, as true parameter value is increased, simulated data gets higher dispersion, and therefore error values increase. To remove this effect, mean normalization was performed both for RMSE and MAE metrics

$$\mathrm{RMSE}^{(s)}_{\mathrm{RV,\,norm}} = \frac{\overline{\mathrm{RMSE}}^{(j^*)}_{\mathrm{RV}}}{\overline{\mathrm{RV}}^{(j^*)}},$$

where s indexes the true parameter configuration, and $\overline{\mathrm{RMSE}}^{(j^*)}_{\mathrm{RV}}$ denotes the RMSE obtained from the optimal parameter combination identified for that scenario.

4.1.2 Forecast and performance metrics

One of the main objectives in financial time-series modeling is to obtain reliable forecasts. Once the simulation and model estimation are completed, the natural next step is to evaluate the forecasting ability of the News-Augmented GARCH model. Volatility forecast is created using out-of-sample method, taking 20% of the data. To measure performance, some of the most common metrics are selected: forecasting accuracy measurement, using RMSE and MAE, and goodness-of-fit measurement, using AIC:

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}, \qquad \mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|,$$
$$\mathrm{AIC} = 2k - 2\ln\hat{L},$$

where y_i is the observed value, $\hat{y}_i$ is the predicted value, n is the number of observations, k is the number of estimated parameters, $\hat{L}$ is the maximum likelihood of the model.

While the calculation of information criteria is straightforward, the evaluation of true volatility is not trivial because it is not directly observed. In general terms, volatility can be defined as an index of variation of a stock price over time, but that can be measured in different ways. In the literature, commonly observed proxies for actual volatility can be summarized into realized volatility (further – RV), historical volatility (further – HV). Historical volatility is defined as standard deviation of past returns over a fixed window:

$$\mathrm{HV} = \frac{1}{n-1}\sum_{t=1}^{n}(r_t - \bar{r})^2.$$

It is the most simple approach and therefore widely used, but it does receive some criticism. HV is often considered a less efficient estimator of true volatility because it does not utilize all available intraday information and might be undervaluing the volatility due to its inability to respond to short time volatility shifts and jumps. General information on HV disadvantages can be found in book [32], with further details in paper [4]. Realized volatility is a measure that covers historical volatility limitations. It is defined as the sum of squared intraday returns over a given period:

$$\mathrm{RV}_t = \sqrt{\sum_{i=1}^{n} \ln^2 \frac{P_{t,i}}{P_{t,i-1}}},$$

where $P_{t,i}$ is the price of the asset at time i of the day t, and n is the number of observations within the day t. Further discussion about realized volatility can be found in aforementioned book [32] with explanations and source to the papers going into further discussion [5] and original paper proposing the idea [17]. As high-frequency data became more accessible, realized volatility has been used more frequently in the literature as actual volatility for error measure calculations. Its advantages are that it captures volatility with finer granularity using high-frequency data, making it more responsive to market shocks and sudden changes, while avoiding the smoothing effects seen in historical volatility. An important note for using realized volatility is that the asset has to be highly traded. The author of original paper, that is being investigated, used historical volatility for error calculation. However, since realized volatility is widely used and offers advantages, this paper includes both RV and HV, along with their corresponding RMSE and MAE metrics, for a more comprehensive model evaluation.

4.1.3 Statistical analysis tools

Once the simulation is done together with forecast and metric calculation, further statistical tools are used in order to answer the main questions about the model and its dynamics. To understand the solo influence of hyperparameters, the following statistical methods are introduced: partial dependence, Sobol indices, and SHAP values. These methods cover different aspects of the analysis, allowing both visual and quantitative examination of solo hyperparameter effects, considering mean and variance contributions.

Partial dependence (PD) is used to visualize how the model error changes on average when varying a single hyperparameter, while averaging over the remaining ones. Let $\mathbf{x}$ denote the vector of hyperparameters, and let $g(\mathbf{x})$ represent the resulting RMSE value. For a selected parameter x_p, the one-dimensional partial dependence function is defined as

$$\mathrm{PD}_p(z) = \mathbf{E}_{\mathbf{x}_{-p}}\big[g(z, \mathbf{x}_{-p})\big],$$

where $\mathbf{x}_{-p}$ denotes the vector of all parameters except x_p. The expectation is approximated by averaging RMSE values over all observed parameter combinations. The resulting curve shows the average direction and magnitude of the influence of the selected hyperparameter on RMSE.

Sobol indices quantify how much each hyperparameter contributes to the variability of RMSE. Let $Y = g(\mathbf{x})$ denote the RMSE produced by parameter vector $\mathbf{x}$. The first-order Sobol index for parameter x_p is defined as

$$S_p = \frac{\mathrm{Var}_{x_p}(\mathbf{E}_{\mathbf{x}_{-p}}[Y \mid x_p])}{\mathrm{Var}(Y)},$$

which measures the variance contribution of parameter x_p acting independently. The total-order index is defined as

$$S_{T_p} = 1 - \frac{\mathrm{Var}_{\mathbf{x}_{-p}}(\mathbf{E}_{x_p}[Y \mid \mathbf{x}_{-p}])}{\mathrm{Var}(Y)}$$

and captures the overall contribution of x_p, including all interaction effects with other parameters. Conditional expectations appearing in the numerators are approximated by averaging RMSE values across subsets of the evaluated parameter combinations, and the variances are computed from these empirical averages. Consequently, Sobol indices represent the proportion of total RMSE variability explained by each hyperparameter.

SHAP values are used to decompose the RMSE obtained for a given parameter configuration into contributions of individual hyperparameters, revealing their local effects. Let $h(\mathbf{x})$ denote the RMSE associated with parameter vector $\mathbf{x}$. For a parameter vector $\mathbf{x}$, the additive decomposition is

$$h(\mathbf{x}) = \phi_0 + \sum_{p=1}^{d} \phi_p.$$

Here ϕ_0 represents the baseline RMSE defined as the average RMSE across all evaluated parameter configurations, while ϕ_p denotes the marginal contribution of parameter x_p relative to this baseline. For each configuration, SHAP values distribute the difference between the observed RMSE and the baseline RMSE among the parameters. Aggregating the absolute values $|\phi_p|$ across all configurations, using the mean absolute value, provides a global measure of parameter importance. This method complements PD and Sobol analysis by highlighting both local effects for individual parameter combinations and the overall influence of each hyperparameter.

4.2 Results

This subsection presents results from simulations used to investigate behavior of the News-Augmented GARCH model. Analysis is structured into three parts: parameter–metric dependence, accuracy in recovering true parameter values, and parameter solo effects. By systematically varying hyperparameters and evaluating key performance metrics, results reveal patterns in estimation accuracy and forecasting performance. Findings give direction to which parameters most strongly influence model behavior, which exhibit weaker interactions, and under what conditions the model yields stable and reliable estimates.

4.2.1 Hyperparameter sensitivity

Hyperparameter sensitivity refers to how performance metrics change as the true parameter values are varied. It helps assess whether the model is accurate and adequate across a broad range of settings. Each parameter is changed while holding other parameter values constant to clear out the effect. As indicated in simulation pseudo-code – each parameter is changed 4 times from the selected baseline. Six metrics are observed: AIC, normalized RMSE, MAE, and their counterparts for realized or historical volatility. Broad selection of metrics is chosen in order to inspect if the outcome is still stable and to compare the predictive accuracy of historical volatility versus realized volatility.

Table 1 presents results from the aforementioned true parameter simulation. Firstly, it is observed that errors increase as true parameter value increases. This pattern is visible in most metrics. It indicates that as the true parameter value increases, it becomes harder

Table 1. Simulation metrics and their standard deviations under different true parameter scenarios: RV – realized volatility, HV – historical volatility, SD – standard deviation.

	RMSE				MAE				AIC	
	RV	SD	HV	SD	RV	SD	HV	SD	Min	SD
Baseline	0.15	0.57	0.79	0.22	0.12	0.59	0.62	0.21	10598	328
$a = 0.2$	0.13	0.22	0.77	0.09	0.10	0.23	0.61	0.07	8464	135
$a = 1$	0.23	0.70	0.81	0.34	0.20	0.74	0.64	0.25	11614	1854
$a = 2$	0.21	1.13	0.81	0.45	0.18	1.17	0.64	0.35	12303	1305
$a = 3$	0.26	1.30	0.82	0.65	0.22	1.38	0.64	0.45	13602	1583
$b = 0.2$	0.17	0.49	0.79	0.19	0.14	0.51	0.63	0.12	10073	2098
$b = 1$	0.16	0.60	0.80	0.26	0.13	0.63	0.63	0.17	10953	2667
$b = 2$	0.15	0.64	0.78	0.25	0.12	0.66	0.62	0.19	10875	398
$b = 3$	0.20	0.65	0.80	0.31	0.17	0.69	0.63	0.23	11453	1942
$\kappa = 0.5$	0.15	0.51	0.78	0.20	0.12	0.52	0.62	0.14	10250	273
$\kappa = 2$	0.15	0.53	0.78	0.20	0.12	0.55	0.63	0.15	10410	309
$\kappa = 10$	0.16	0.65	0.79	0.26	0.13	0.67	0.63	0.19	10929	428
$\kappa = 20$	0.16	0.70	0.79	0.29	0.13	0.72	0.63	0.21	11155	456
$\gamma = 0.5$	0.14	0.50	0.78	0.20	0.11	0.51	0.61	0.18	10252	284
$\gamma = 2$	0.15	0.53	0.78	0.21	0.12	0.55	0.63	0.15	10410	311
$\gamma = 10$	0.15	0.65	0.79	0.26	0.12	0.67	0.62	0.19	10930	412
$\gamma = 20$	0.16	0.70	0.79	0.29	0.13	0.72	0.62	0.19	11148	483

for the model to identify its exact position, and the range of plausible values expands. This is followed by guideline that if evaluated parameter value is within a higher range, it should be interpreted with more caution. Further looking at standard deviation, it seems to be relatively high, on average around 1.5 times higher than the error, especially for realized volatility based measures. This might come from relatively low simulation size of 100, which was selected due to resource limitations. Nevertheless, the way standard deviation is calculated in this simulation gives a specific interpretation. Calculation of standard deviation is provided in pseudo-code Section 4.1.1, in short – it is calculated as the mean variation of the metric across the parameter grid. The interpretation therefore is that a lower standard deviation indicates that metric values across the grid are relatively consistent, suggesting that the choice of grid point is less critical, as alternative combinations yield similar metric values. On the contrary, a higher standard deviation reflects a wider spread of metric values across the grid, implying that selecting the optimal grid point is more important, since other combinations may result in substantially different performance. Given the aforementioned interpretation, resulting outcome is that standard deviation also increases as true parameter values increase, translating a higher sensitivity in true value exploration.

Comparing metrics it is observed that RMSE values are higher than MAE values, which is expected since RMSE squares the deviations before averaging and then takes the square root. This mathematical form gives more weight to larger deviations, making RMSE particularly sensitive to the tails of the distribution. Although Gaussian errors were used in the simulation, the model still captures tail behavior that naturally occurs in financial data, demonstrating its robustness. Therefore, RMSE will be used as the primary performance metric in the remainder of this paper.

Another topic in question is realized and historical volatility comparison. The empirical findings of this study show that using realized volatility produces over 4 times lower errors than historical volatility, confirming the earlier theoretical discussion that realized volatility is a more accurate and effective proxy for actual volatility. Therefore, for the remainder of this paper, only realized volatility is considered. However, the general patterns and observations remain similar to those obtained with historical volatility, confirming the stability of the conclusions.

4.2.2 Hyperparameter recovery power

One of the most important questions to ask while developing a model – how well is it able to recover true parameter values? In this study model, a successful finding of the parameter is defined as either around minimum RMSE value point or around the second derivative, where RMSE value starts to stabilize, and model is invariant to further parameter changes. As briefly presented in Section 4.1.1, true parameter values are changed, and it is observed under which conditions model is able to find true parameter value.

Findings summarized in Table 2 show that using RMSE RV as a metric, the model is able to closely recover 60% of selected true parameter values. The number is generally on an acceptable side as the model itself is quite complex – includes nonlinearity, asymmetry, and multiplication. The general tendency is that model was not able to correctly find any

Table 2. Hyperparameter recovery success rates.

Metric	a	b	κ	γ	Mean
RMSE RV	60%	40%	60%	80%	60%
MAE RV	60%	20%	60%	80%	55%
RMSE HV	40%	40%	60%	40%	45%
MAE HV	60%	40%	40%	0%	35%
AIC	60%	20%	20%	0%	25%
Mean	71%	38%	63%	58%	

of the low parameter values < 1 and generally preferred to select higher parameters. This suggests that recovering low parameter values is relatively more difficult, and estimated values near the lower boundary should be interpreted with caution. It was also observed that across various metrics, model was able to find parameter a most times (71% on average) and parameter b least times (38% on average). This suggests that estimates of a can generally be interpreted with higher confidence, while b requires more careful consideration. In practice, simulations involving b may benefit from a denser grid to capture its behavior more reliably.

Additionally, $\hat{\alpha}$ and $\hat{\beta}$ were observed. News impact share defined as $\hat{\alpha}/(\hat{\alpha}+\hat{\beta})$ averages to 0.04 for different true parameter combinations, which is a regular level of how much shocks persist in the model. News-Augmented GARCH was able to find correct levels even though time series were simulated with initial average news impact share of 0.27. The model was returning to lower share values as such a high ratio is not logical in a few perspectives. Intuitively, the shock term $\alpha\varepsilon_{t-1}^2$ captures only the immediate reaction of volatility to a new innovation, so it is relatively small, while the persistence term $\beta\sigma_{t-1}^2$ carries forward most of the previous day's variance. Since financial volatility is highly persistent ($\alpha+\beta \approx 1$), the model naturally assigns more weight to β, making the ratio $\alpha/(\alpha+\beta)$ small. Moreover, the innovations in this simulation are assumed to be Gaussian, so extreme returns are less frequent than under heavy-tailed distributions. As a result, the estimated contribution of shocks is reduced, leading to expected smaller values of α.

Delving deeper into the pattern of true parameter search (see Fig. 3), a nonlinear response is observed for all parameters together with a stabilizing tendency once a certain threshold is reached. Increasing parameter values beyond the threshold typically yields no further improvements, even a decline in performance or a very slow decrease in metric value. It indicates that the model has a relatively well-defined region of optimal parameter values. The presence of a stabilizing pattern also suggests robustness, as extreme parameter values do not drastically impact results. Furthermore, considering the stability condition derived in Subsection 3.2, a logical interpretation arises: if $a+b$ increases, $\alpha+\beta$ decrease toward zero, in which case the News-Augmented GARCH model effectively looses its GARCH component and behaves more like a purely news-driven process. However, there are notable exceptions. For information-criterion metrics, parameter b follows a linear relationship with metric values increasing as b increases. Similarly, κ in RMSE HV and RV exhibits a continuous decline in metric values, suggesting a preference for higher κ values. Nonetheless, the actual value tends to be around the point of the highest rate of change, supporting the general pattern.

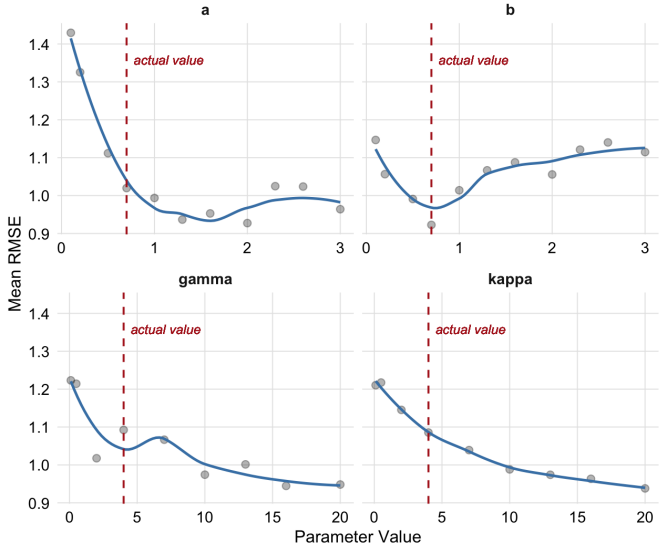

Figure 3. RMSE changes in response to varying parameter values under the baseline condition.

4.2.3 Solo hyperparameter influence

To understand the relative importance of each hyperparameter in the News-Augmented GARCH model, it is essential to analyze individual effects. Isolating them in a nonlinear setting is not trivial, but it is necessary to obtain a complete picture of the model's behavior. The goal of this subsection is to quantify and visualize solo effect of each hyperparameter a, b, κ, γ on model performance, measured by RMSE. The main questions to be answered are therefore: which hyperparameter has the strongest independent effect on RMSE, and which hyperparameter exhibits the lowest interaction with other parameters. These questions are addressed using statistical tools introduced in Subection 4.1.3, which goes by combining partial dependence, Sobol indices, and SHAP values. Together they capture mean trends, variance contributions, and local or global influences, allowing to identify dominant parameter and assess parameter entanglement. For the analysis, a denser parameter grid is used focusing on the baseline hyperparameters. Thirteen combinations are selected for each hyperparameter. While the general trends observed in this denser grid are similar to those in the previous section, the higher resolution provides more stable and reliable estimates, particularly when examining detailed parameter interactions.

Figure 4 gives general overview of each hyperparameter importance using Sobol indices and SHAP values. It is seen that parameter a consistently dominates across all methods. It has the highest first-order index, indicating a strong solo influence on RMSE. SHAP analysis confirms this result, changes in a move predictions of RMSE on average by approximately 0.14, which corresponds to about $25-30\%$ of the typical RMSE range $(0.44-0.60)$. Partial dependence results in Fig. 5 also highlight the largest variation in RMSE when a is varied, showing model's sensitivity to this parameter and therefore

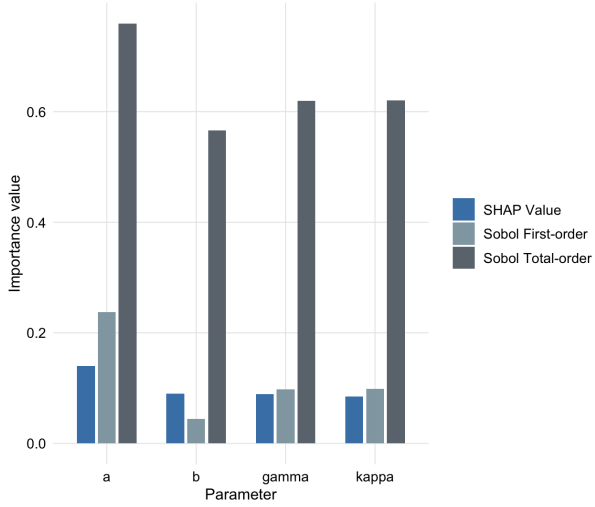

Figure 4. Comparison of solo effects across parameters. Sobol indices and SHAP values are compared.

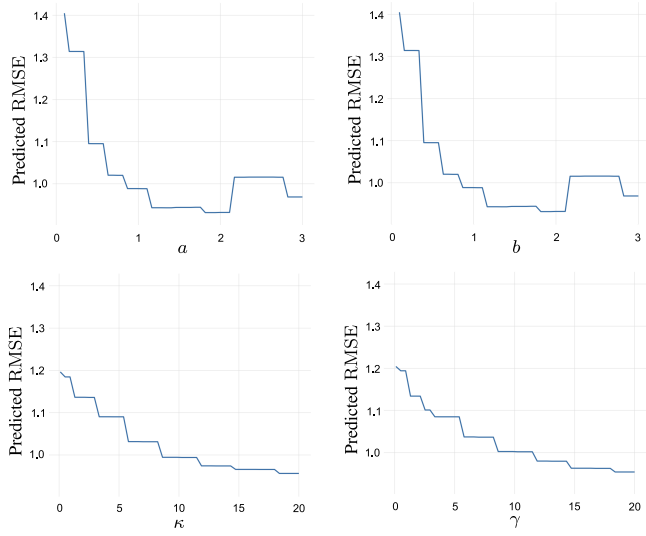

Figure 5. Partial dependence plots for parameters a, b, κ, and γ.

further highlighting its central role. Although the total-order Sobol index indicates that a has the highest impact including interaction terms, the first-order to total-order ratio shows that parameter a has the lowest interaction weight. Meaning that adding other parameters gives less of additional impact comparing to other parameters. This therefore again complements the statement that parameter a is central in this model with both high solo effect and relatively lower interaction strength.

Parameter b shows lowest solo importance and highest interaction effects in terms of solo-to-total ratio. Partial dependence curve has the lowest scale, which confirms lower impact levels. A noteworthy observation is that partial dependence exhibits a U-shaped effect with an optimal region around $b = 0.7$, which is true parameter value for the selected data set. Drawing from one example, this suggests that parameter b might be relatively sensitive in correctly identifying the true value and is not invariant to other parameter value regions. Additionally, the absence of a stabilizing pattern is advantageous, as it indicates that under current conditions, b finds lowest RMSE exactly at its true value rather than becoming indifferent at higher values.

Parameters κ and γ have almost identical and moderate level influence. Their Sobol indices and SHAP values are smaller than those of a, but still nonnegligible. Partial dependence curves decrease steadily and have mid-sized scale, suggesting that higher values of κ or γ lead to modest reductions in RMSE. Additionally, all parameters in general show high level difference between first order and total order indices, indicating that there exists strong interactions. This suggests that parameters are entangled, and their joint influence on the model is substantially higher than their individual effects.

In summary, the simulation study shows that the model generally recovers near-true parameter values and exhibits stabilizing, robust behavior even under extreme settings. Some sensitivity persists at the upper and lower ranges, so estimates there should be interpreted cautiously. Parameter a consistently dominates the dynamics, requiring particular care in hyperparameter grid construction.

5 Discussion and conclusions

Over the last few decades, there has been a shift towards enriching volatility models by incorporating various proxies of news and information flow. A consistent evidence is found in the literature that news sources have a significant effect on volatility. However, most existing approaches introduce sentiment indices as additive exogenous regressor, which shows that the relation in question is by default linear. Nevertheless, such complex phenomena as public reaction to the news and its relation to volatility are likely to exhibit nonlinear behavior. A novel approach provided by Sadik [25] address this feature and provide a new form of news inclusion into volatility. Multiplicativity together with nonlinearity and asymmetry were introduced, which although targeted those limitations and covered the complexity of the question, also increased model complexity. Therefore, additional questions had to be raised to fully understand the mathematical and numerical background of the model. This paper bridges this gap and aims to evaluate mathematical properties and numerical behavior of the News-Augmented GARCH model. Results demonstrate that under the derived condition, the model is stable and uniquely solvable. It is also capable of accurately recovering actual parameters in simulated settings. Simulation-based sensitivity analysis further revealed nuanced relationships between model parameters and performance metrics.

The main findings of the paper provide confidence in its reliability and suitability for empirical application. Firstly, under a newly found condition $(a + b)(\alpha + \beta) < 1$,

the model is stable and has a unique causal solution. The proof and the condition give an important base for reliability of the model and a new approach to how α and β are evaluated and interpreted as a conjunction to time-varying sentiment-related function rather than a constant parameter by itself. This interpretation differs to the original News-Augmented GARCH paper, which was working under plain GARCH conditions, and also to other most commonly used models. Secondly, through a simulation study, it was found that model is able to closely recover actual values under a broad set of hyper-parameters. Although there are some exceptions or conditions to this finding, the general trend remains, which is a novel result and provides a basis for relying on the model in further empirical analysis with unknown parameters. Further simulation study findings also provided the reader with newly found guidelines on how to interpret the estimates. Firstly, more careful interpretation is needed for estimates in low and high ranges, as the model tends to be less accurate in those regions. Additionally, hyperparameter a stands out as the most influential, having the strongest impact on forecasting accuracy. Thus, more attention should be given to this parameter when shaping the grid and interpreting results. Lastly, a comparison between realized volatility and historical volatility has been established, an aspect that is not commonly emphasized in studies on GARCH model accuracy. It was found that RV yields lower RMSE values and better approximates actual volatility. Those findings point to News-Augmented GARCH model being a strong candidate among the wide GARCH family, as it captures the complex dynamics of volatility and news effects and also proves to lean towards robustness, reliability, and accuracy in its performance.

From a practical standpoint, the validated uniqueness and stability, together with findings from the simulation study, extend confidence in the model's use for further empirical applications, offering practical value for investors and risk managers. For instance, the model with the new condition can be used in diversifying portfolio, separating stocks under news-sensitivity appetite. It also enables strategies that adjust exposure based on the market's sensitivity to positive or negative information. Moreover, it can enhance stress testing and scenario analysis by simulating the impact of sudden news shocks on market volatility, helping institutions better prepare for extreme conditions.

While the study provides valuable insights into the model, there are certain limitations that should be taken into consideration when interpreting its results. Firstly, the model is subject to distributional limitations. News are modeled with a half-normal distribution, while returns are assumed to be normally distributed, even though heavy tails are also observed in stock return data. These assumptions can be limiting, therefore should be further examined, and additional calculations may be required to assess their impact. The simulation design also introduces a few limitations related to sample and grid size. Each parameter combination was simulated 100 times, which may not be sufficient to achieve fully stable results. The grid of tuning parameters currently contains 1764 combinations, but this could be extended to explore the parameter space more thoroughly. Additionally, 4 parameter changes were selected for observing true values. Current sizes were chosen due to resource constraints but still produced sufficiently robust patterns to support meaningful conclusions. Lastly, analysis is limited to simulation, but empirical findings would provide additional validation. Therefore, future work may concentrate on addressing these

limitations, while the authors will focus on extending the model to multivariate settings together with an empirical evaluation.

Conflicts of interest. The authors declare no conflicts of interest.

Acknowledgment. The authors are thankful for the high performance computing resources provided by the Information Technology Research Center of Vilnius University.

References

1. H. Abdollahi, Oil price volatility and new evidence from news and Twitter, *Energy Econ.*, **122**:106711, 2023, https://doi.org/10.1016/j.eneco.2023.106711.
2. J. Ahmed, M. Ahmed, A framework for sentiment analysis of online news articles, *Int. J. Emerg. Technol.*, **11**(3):267–274, 2020, ISSN 0975-8364.
3. C. Amado, T. Teräsvirta, Modelling conditional and unconditional heteroskedasticity with smoothly time-varying structure, SSE/EFI Working Paper Series in Economics and Finance, No. 691, Stockholm School of Economics, Economic Research Institute, 2008, https://hdl.handle.net/10419/56173.
4. T.G. Andersen, T. Bollerslev, Answering the skeptics: Yes, standard volatility models do provide accurate forecasts, *Int. Econ. Rev.*, **39**(4):885–905, 1998, https://doi.org/10.2307/2527343.
5. T.G. Andersen, T. Bollerslev, F.X. Diebold, H. Ebens, The distribution of realized stock return volatility, *J. Financ. Econ.*, **61**(1):43–76, 2001, https://doi.org/10.1016/S0304-405X(01)00055-1.
6. L. Bauwens, W. Ben Omrane, P. Giot, News announcements, market activity and volatility in the euro/dollar foreign exchange market, *J. Int. Money Finance*, **24**(7):1108–1125, 2005, https://doi.org/10.1016/j.jimonfin.2005.08.008.
7. S. Berry, G. Mitra, Z. Sadik, Improved volatility prediction and trading using StockTwits sentiment data, *SSRN Electron. J.*, 2019, https://doi.org/10.2139/ssrn.3527557.
8. F. Black, Studies of stock price volatility changes, in *Proceedings of the 1976 Meeting of the Business and Economic Statistics Section, Chicago, IL*, American Statistical Association, Washington DC, 1976, pp. 177–181.
9. T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, *J. Econom.*, **31**:307–327, 1986, https://doi.org/10.1016/0304-4076(86)90063-1.
10. J.Q. Chen, H. Li, Y. Lv, Using TGARCH-M to model the impact of good news and bad news on COVID-19 related stocks' volatilities, *J. Financ. Risk Manag.*, **11**(2):441–480, 2022, https://doi.org/10.4236/jfrm.2022.112023.
11. J.Q. Chen, H. Li, Y. Lv, Using TGARCH-M to model the impact of good news and bad news on COVID-19 related stocks' volatilities, *J. Financ. Risk Manag.*, **11**(2):441–480, 2022, https://doi.org/10.4236/jfrm.2022.112023.
12. S.T. Chen, K.Y.A. Haga, Using E-GARCH to analyze the impact of investor sentiment on stock returns near stock market crashes, *Front. Psychol.*, **12**:664849, 2021, https://doi.org/10.3389/fpsyg.2021.664849.

13. S.T. Chen, N.D. Phuoc, T.T.N. Bui, V.H. Nguyen, Using E-GARCH to analyze the impact of investor sentiment on stock returns near stock market crashes, *Front. Psychol.*, **12**:699861, 2021, https://doi.org/10.3389/fpsyg.2021.699861.

14. M. Davidović, J. McClear, News sentiment and stock market dynamics: A machine learning approach, *J. Risk Financ. Manag.*, **18**(8):412, 2025, https://doi.org/10.3390/jrfm18080412.

15. K. Du, F. Xing, R. Mao, E. Cambria, Financial sentiment analysis: Techniques and applications, *ACM Comput. Surv.*, **56**(9):220, 2024, https://doi.org/10.1145/3649451.

16. L.H. Ederington, J.H. Lee, The economic determinants of the volatility of stock returns, *J. Financ. Econ.*, **36**(3):355–384, 1993.

17. K.R. French, G.W. Schwert, R.F. Stambaugh, Expected stock returns and volatility, *J. Financ. Econ.*, **19**(1):3–29, 1987, https://doi.org/10.1016/0304-405X(87)90026-2.

18. L. Glosten, R. Jagannathan, D. Runkle, On the relation between the expected value and the volatility of the nominal excess return on stocks, *J. Finance*, **48**:1779–1801, 1993, https://doi.org/10.1111/J.1540-6261.1993.TB05128.X.

19. Y.-J. Hsu, Y.-P. Wang, S.-C. Chen, News sentiment and stock market volatility, *Rev. Quant. Finance Account.*, **55**(2):525–548, 2020, https://doi.org/10.1007/s11156-019-00801-4.

20. P. Kalev, Public information arrival and volatility of intraday stock returns, *J. Empir. Finance*, **9**(1):55–75, 2002.

21. C. Lamoureux, W. D. Lastrapes, Heteroskedasticity in stock return data: Volume versus GARCH effects, *J. Finance*, **45**:221–229, 1990, https://doi.org/10.1111/J.1540-6261.1990.TB05088.X.

22. G. Mitra, L. Mitra (Eds.), *The Handbook of News Analytics in Finance*, John Wiley & Sons, Chichester, 2011, https://doi.org/10.1002/9781118467411.

23. D.B. Nelson, Conditional heteroskedasticity in asset returns: A new approach, *Econometrica*, **59**:347–370, 1991, https://doi.org/10.2307/2938260.

24. L. Rognone, S. Hyde, S.S. Zhang, News sentiment in the cryptocurrency market: An empirical comparison with Forex, *Int. Rev. Financ. Anal.*, **69**:101462, 2020, https://doi.org/10.1016/j.irfa.2020.101462.

25. Z. Sadik, P. Date, G. Mitra, News augmented GARCH(1,1) model for volatility prediction, *IMA J. Manag. Math.*, **30**, 2018, https://doi.org/10.1093/imaman/dpy004.

26. Z.A. Sadik, G. Mitra, K. Ahmad, Forecasting crude oil futures prices using global macroeconomic news sentiment, *IMA J. Manag. Math.*, **31**(1):1–26, 2019, https://doi.org/10.1093/imaman/dpy016.

27. J.-M. Sahut, P. Hajek, V. Olej, L. Hikkerova, The role of news-based sentiment in forecasting crude oil price during the COVID-19 pandemic, *Ann. Oper. Res.*, **345**:861–884, 2025, https://doi.org/10.1007/s10479-024-05821-z.

28. S. Sidorov, P. Date, V. Balash, Using news analytics data in GARCH models, *Appl. Econom.*, **29**(1):82–96, 2013.

29. T.T.P. Souza, O. Kolchyna, P.C. Treleaven, T. Aste, Twitter sentiment analysis applied to finance: A case study in the retail industry, 2015, arXiv:1507.00955.

30. E. Trabelsi, COVID-19 and uncertainty effects on Tunisian stock market volatility: Insights from GJR-GARCH, wavelet coherence, and ARDL, *J. Risk Financ. Manag.*, **17**(9):403, 2024, https://doi.org/10.3390/jrfm17090403.

31. E. Trabelsi, Covid-19 and uncertainty effects on tunisian stock market volatility: Insights from GJR-GARCH, wavelet coherence, and ARDL, *J. Risk Financ. Manag.*, **17**(9):403, 2024, https://doi.org/10.3390/jrfm17090403.

32. R.S. Tsay, *An Introduction to Analysis of Financial Data with R*, John Wiley & Sons, Chichester, 2013.

33. X. Yu, G. Mitra, C. Arbex-Valle, T. Sayer, An impact measure for news: Its use in (daily) trading strategies, in G. Mitra, L. Mitra (Eds.), *The Handbook of Sentiment Analysis in Finance*, John Wiley & Sons, Chichester, 2016, https://doi.org/10.2139/ssrn.3706827.

34. Y. Zhai, A. Hsu, S. Halgamuge, Combining news and technical indicators in daily stock price trends prediction, in *Data Mining and Knowledge Discovery Handbook*, Springer, Berlin, Heidelberg, 2007, pp. 1087–1096, https://doi.org/10.1007/978-3-540-72395-0_132.

35. *RavenPack News Analytics, Version 4.0: User Guide and Service Overview,* RavenPack, 2014.