LVI SBPO SIMPÓSIO BRASILEIRO DE PESQUISA OPERACIONAL
DE 04 ATÉ 07 NOVEMBRO 2024
FORTALEZA | CEARÁ

Non-gaussian score-driven models with non-linear unobserved components combinations

ABSTRACT

A commom technique in time series modeling is to decompose the time series into it's trend and seasonal components. Inside the Score-Driven Models class, this decomposition is usually made in an additive form, so that the series is expressed as the sum of its trend and seasonal components. However, it is not unusual that, even with the seasonal component being considered into the model, the residuals still show signs of seasonal dependency that were not captured by the model. With that said, the main objective of this article is to study if a non-linear combination of those components is able to improve forecast accuracy in score driven models.

KEYWORDS. Time series. Score-driven models. Unobserved components.

EST&MP - Statistics and Probabilistic Models, BDA - Big Data e Analytics, PM - Mathematical Programming

1. Introduction

Score-driven models are a general and flexible class of time series models developed simultaneously by (Harvey [2013]) and (Creal et al. [2013]). Known by different names - generalized autoregressive score (GAS), dynamic conditional score models, or simply score-driven models, which is the term chosen in this article - they offer several advantages for time series modeling, as described below.

Firstly, these models can handle non-Gaussian time series, where the conditional probability distribution is not normal. This capability is crucial for cases where the normal distribution is inadequate, such as wind speed series (which only take positive values) and financial return series (which exhibit excess kurtosis). Count data, such as the number of items sold in a store, also benefit from this flexibility, as they require discrete distributions. The ability to model non-normal data makes GAS models highly applicable across various domains.

In finance, they effectively model asset volatility. For instance, (Xu and Lien [2022]) found that score-driven models outperform GARCH and EGARCH for oil asset volatility and are superior for 1- and 5-step-ahead predictions for natural gas assets. (Fuentes et al. [2023]) used score-driven models for predicting extreme financial return events, improving accuracy by incorporating realized volatility measures. In retail, (Hoeltgebaum et al. [2021]) applied GAS models with a lognormal distribution for daily demand forecasts in distribution centers, proving competitive with usual benchmarks and offering closed-form predictive density. In climate and energy, (Blazsek and Escribano [2023]) used GAS models to measure long-term nonlinear climate effects on variables like global ice volume and atmospheric CO2 levels. (Kushwah and Wadhvani [2019]) employed GAS models for wind power series forecasts, monitoring turbine performance, and integrating neural network activation functions in nonlinear GAS models (NLGASX). Lastly, to demonstrate the versatility of score-driven models, (Koopman and Lit [2019]) used score-based models to predict football match outcomes in national championships and concluded that this class of models outperformed various benchmark models in both forecast accuracy and minimizing betting losses.

Several packages and software for score-driven models are available, facilitating their computational implementation. Examples include the GAS package developed in R (Ardia et al. [2016]), the ScoreDrivenModels package implemented in Julia (Bodin et al. [2020]), and Time Series Lab (Lit et al. [2023]), a free time series modeling software.

Score-driven models can be implemented primarily in two ways: through an ARMA (Auto Regressive Moving Average) process and through an unobserved components specification, which is the framework adopted in this article. The ARMA approach leverages the familiar structure of autoregressive and moving average terms to capture the dynamics of the time series, making it straightforward to apply standard ARMA techniques within the score-driven framework (Creal et al. [2013]). On the other hand, the unobserved components specification involves decomposing the time series into various latent components, such as trend, seasonality, and irregular components. This approach, developed by (Harvey [2013]), is strongly connected with structural time series models (Harvey [1990]) and (Durbin and Koopman [2012]), which commonly decompose the time series into a linear combination of its trend and seasonal components.

However, even with a correctly specified seasonal component, there are some time series for which this linear combination is insufficient to capture the complete seasonal effect, often diagnosed by a significant seasonal lag in the autocorrelation function of the linear model's residuals. Given this shortfall, and inspired by the ETS models from (Hyndman et al. [2008]) that allow for non-linear components combinations and by the non-linear interactions for unobserved components structural models proposed in (Koopman and Lee [2009]), this article aims to study whether non-linear combinations of trend and seasonal components in a score-driven model can fully capture the

seasonal dynamics of a time series where the linear model falls short.

2. Theoretical Background

The first step in modeling a time series through a score-driven process is to define the probability distribution to be followed by the series to be modeled. In other words, given a series $\boldsymbol{y_t}$ and its past $\mathbf{Y_{t-1}} = (y_{t-1}, y_{t-2}, ..., y_2, y_1)$, one should define a conditional distribution such that

$$y_t \sim p(y_t | f_{t|t-1}; \mathbf{Y_{t-1}}, \theta) \tag{1}$$

where $f_{t|t-1}$ is the time-varying parameter estimated at time t and θ is a vector containing fixed hyperparameters to be estimated.

Once the conditional probability distribution is defined, the next step is to define the updating mechanism for the time-varying parameters. There are two possibilities: the first is for the updating function to follow an ARMA(p,q) process, and the second is to develop this function based on the framework of unobserved components models (UCM), which is the specification used in this article.

Before going into the dynamics' specification, it is important to define some essential concepts for score-driven models, the first of which is the score. This concept drives the dynamics of the time series model at each time point and is defined as:

$$\nabla_t = \frac{\partial \ln p(y_t | f_{t|t-1}; \mathbf{Y_{t-1}}, \theta)}{\partial f_{t|t-1}} \tag{2}$$

When combined with the Fisher information matrix, $I_{t|t-1}$, we obtain the so-called standardized score s_t, which will guide the model development. Their definitions are:

$$I_{t|t-1} = E_{t-1}[\nabla_t \nabla_t'] = -E_{t-1}\left[\frac{\partial^2 \ln p(y_t | f_{t|t-1}; \theta, \mathbf{Y_{t-1}})}{\partial f_{t|t-1} f_{t|t-1}'}\right] \tag{3}$$

$$s_t = I_{t|t-1}^{-d} \nabla_t, \quad d = \{0, 1, 0.5\} \tag{4}$$

It is worth noting that, when $d = 0.5$, the definition of s_t is slightly different, as shown by (Sarlo et al. [2023]).

A convenient property of observation-driven models is that parameter estimation can be performed via maximum likelihood estimator (MLE) (Creal et al. [2013]). The likelihood function, denoted by $L(\theta)$, is simply the joint conditional probability distribution written as a function of θ, which is the vector that encompasses all the fixed parameters of the model. It is common to use the log-likelihood function, represented by $l(\theta) = \ln L(\theta)$. The maximum likelihood maximization problem can be written as:

$$\hat{\theta} = \arg\max_{\theta} \sum_{t=1}^{T} l_t \tag{5}$$

where $l_t = \ln p(y_t | f_{t|t-1}, \mathbf{Y_{t-1}}; \theta)$ is the log-likelihood function for a realization of y_t. According to (Creal et al. [2013]), estimating via MLE for score-driven models is particularly straightforward, provided an efficient nonlinear optimization algorithm is used, requiring only the implementation of the updating function and the evaluation of the l_t function at a particular value θ^* of θ.

As is common for time series models, the quality of the model is assessed based on diagnostics of the residuals from the estimated model. In the case of score-driven models, there are two possible classes of residuals: Pearson residuals and quantile residuals.

Pearson residuals are defined as:

$$r_t^P = \frac{y_t - E(y_t|f_{t|t-1}, \mathbf{Y_{t-1}}, \theta)}{\sqrt{Var(y_t|f_{t|t-1}, \mathbf{Y_{t-1}}, \theta)}} \quad \forall t = 1, ..., T \tag{6}$$

where $E(y_t|f_{t|t-1}, \mathbf{Y_{t-1}}, \theta)$ and $Var(y_t|f_{t|t-1}, \mathbf{Y_{t-1}}, \theta)$ are, respectively, the mean and variance of the probabilistic model $p(y_t|f_{t|t-1}, \mathbf{Y_{t-1}}, \theta)$.

Quantile residuals, on the other hand, are defined as:

$$r_t^q = \Phi^{-1}(F(y_t|f_{t|t-1}, \mathbf{Y_{t-1}}, \hat{\theta})) \quad \forall t = 1, ..., T \tag{7}$$

where $\Phi^{-1}(.)$ is the quantile of a standard normal distribution Z and $F(y_t|f_{t|t-1}, \mathbf{Y_{t-1}}, \hat{\theta})$ is the cumulative distribution function associated with $p(y_t|f_{t|t-1}, \mathbf{Y_{t-1}}, \theta)$.

It is important to note that, as shown by (Kalliovirta [2012]), if the GAS model is well specified, the quantile residuals are asymptotically normal and independently distributed, even if the probabilistic model is not Gaussian. Thus, normality tests for the residuals can still be used as a diagnostic tool. Since the Pearson residuals are appropriate only for linear and Gaussian models, one should choose to use the quantile residuals for the model's diagnostics.

This study focuses on the effect of non-linear combinations in capturing the seasonal component, the residual diagnostics will primarily emphasize the autocorrelation function (ACF) of the residuals. This emphasis is crucial because the presence of significant autocorrelation at seasonal lags in the residuals indicates that the model has not adequately captured the seasonal dynamics. By carefully analyzing the ACF of the residuals, it is possible to assess whether the non-linear combination of trend and seasonal components has effectively addressed the seasonal patterns in the time series. A well-specified model should exhibit residuals with no significant autocorrelation at seasonal lags (Hyndman and Athanasopoulos [2021]), thereby confirming that the non-linear approach has successfully captured the seasonal component.

Just as correlation measures the extent of a linear relationship between two variables, autocorrelation measures the linear relationship between lagged values of a time series (Hyndman and Athanasopoulos [2021]). The autocorrelation between y_t and y_{t-k} is denoted as r_k and can be expressed as:

$$r_k = \frac{\sum_{t=k+1}^{T}(y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{T}(y_t - \bar{y})^2} \tag{8}$$

where T is the length of the time series, $\bar{y}$ is the sample mean of the time series y_t, y_t is the value of the time series at time t, and y_{t-k} is the value of the time series at lag k.

In addition to analyzing the autocorrelation function (ACF) of the residuals, this study will also compute the Mean Absolute Percentage Error (MAPE) and Mean Absolute Scaled Error (MASE) metrics for both in-sample and forecasted values. These metrics provide a comprehensive evaluation of the model's performance. MAPE offers insight into the average magnitude of errors in percentage terms, making it easily interpretable. MASE, on the other hand, scales the errors based on the errors from a naive forecasting method, providing a robust measure that is less sensitive to outliers. By including these metrics, the study aims to evaluate not only the effectiveness of the model in capturing the seasonal component but also its out-of-sample prediction quality.

Both MAPE and MASE metrics are defined as follows:

$$MAPE = 100 \times \frac{1}{T}\sum_{t=1}^{T}\left|\frac{y_t - \hat{y}_{t|t-1}}{y_t}\right|, \quad y_t \neq 0 \tag{9}$$

where T is the length of the series, y_t is the time series, and $\hat{y}_t$ is the value estimated or forecasted by the model.

$$MASE = \frac{\frac{1}{T}\sum_{t=1}^{T}\left|y_t - \hat{y}_{t|t-1}\right|}{\frac{1}{T-1}\sum_{t=2}^{T}|y_t - y_{t-1}|} \tag{10}$$

where T is the length of the series, the denominator represents a naive forecast, and the numerator can be understood as the Mean Absolute Error (MAE). MASE is interpreted as follows: if it is greater than 1, it means the model is performing worse than the naive forecast, while if it is less than 1, the model is performing better than the naive forecast.

Finally, in order to define an unobserved components score-driven model, as presented in (Harvey [2013]) and (Caivano et al. [2016]), one should refer to an unobserved components structural model, such as the basic structural model (BSM), in which the time series y_t is decomposed into a trend component μ_t, modeled as a random walk with stochastic slope b_t and a stochastic seasonal component γ_t, modeled with trigonometric terms:

$$\begin{aligned}
&y_t = m_t + \gamma_t + \epsilon_t, \epsilon_t \sim N(0, \sigma^2)\\
&m_{t+1} = m_t + b_t + \eta_t, \eta_t \sim N(0, \sigma_\eta^2)\\
&b_{t+1} = b_t + \xi_t, \xi_t \sim N(0, \sigma_\xi^2)\\
&\gamma_{t+1|t} = \sum_{i=1}^{S/2} \gamma_{i,t}\\
&\gamma_{i,t} = \gamma_{i,t-1}\cos(\omega_{i,t}) + \gamma^*_{i,t-1}\sin(\omega_{i,t}) + \nu_t, \nu_t \sim N(0, \sigma_\nu^2)\\
&\gamma^*_{i,t} = -\gamma_{i,t-1}\sin(\omega_{i,t}) + \gamma^*_{i,t-1}\cos(\omega_{i,t}) + \nu^*_t, \nu^*_t \sim N(0, \sigma_\nu^2)
\end{aligned} \tag{11}$$

The main difference between a structural model and a score-driven model is that the first is a parameter-driven model, which means the dynamics of the time-varying components is driven by unobserved shocks $\epsilon_t, \eta_t, \xi_t, \nu_t$, while the second, as already explained, has its parameters driven by the past values of the series through the score.

Suppose the unobserved components score-driven model to be defined follows a Gaussian distribution $y_t \sim N(\mu_{t|t-1}, \sigma^2)$, where $E(y_t|f_{t|t-1}, \mathbf{Y_{t-1}}, \theta) = \mu_{t|t-1}$ and $V(y_t|f_{t|t-1}, \mathbf{Y_{t-1}}, \theta) = \sigma^2$. Notice that only the mean parameter is being defined as a time-varying parameter. With that being said, one could define unobserved components dynamics such as the one from the BSM for these parameters.

$$\begin{aligned}
\mu_{t+1|t} &= m_{t+1|t} + \gamma_{t+1|t}\\
m_{t+1|t} &= m_{t|t-1} + b_{t|t-1} + \kappa_m s_t\\
b_{t+1|t} &= b_{t|t-1} + \kappa_b s_t\\
\gamma_{t+1|t} &= \sum_{i=1}^{S/2} \gamma_{i,t}\\
\gamma_{i,t} &= \gamma_{i,t-1}\cos(\omega_{i,t}) + \gamma^*_{i,t-1}\sin(\omega_{i,t}) + \kappa_\gamma s_t\\
\gamma^*_{i,t} &= -\gamma_{i,t-1}\sin(\omega_{i,t}) + \gamma^*_{i,t-1}\cos(\omega_{i,t}) + \kappa_\gamma s_t
\end{aligned} \tag{12}$$

The main difference between the formulation of a state-space model described in equation (11) and a score-driven model described in equation (12) is the fact that the former has random shocks in its component equations, while the latter replaces these shocks with a function of the

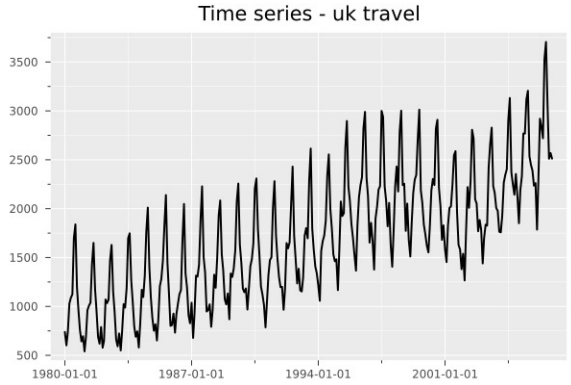

Figura 1: UK travel monthly time series.

score, making the parameter update mechanism dependent on past observations of the time series. Given this simplicity, the steps for defining the score-driven model from the BSM were omitted and the random shocks were simply replaced by the standardized score. As a result, we now have the formulation of a score-driven model whose time-varying parameters follow an unobserved components process, where the components' stochasticity is being modeled through κ_m, κ_b and κ_γ, that multiply the standardized score s_t. If a κ is estimated to zero, this means that its component is deterministic, not stochastic.

3. Proposed Model

To choose the appropriate score-driven model specification, it is useful to examine the time series to be modeled. Inspired by (Koopman and Lee [2009]), a time series from the same article was chosen. It consists of the monthly number of international trips made by residents of the United Kingdom between January 1980 and December 2006, summing up to 324 observations. This series accounts for all means of transportation used for entering and leaving the United Kingdom. The data was retrieved from `https://www.ons.gov.uk/peoplepopulationandcommunity/leisureandtourism/datasets/internationalpassengersurveytimeseriesspreadsheet`. This series was chosen because it presents, in addition to a growing trend and well-defined seasonality, an increase in the variation of seasonality. This increase, as noted by the author, is not resolved by using the logarithmic transformation of the time series, making it an excellent example for this work. Figure 1 shows its plot. From this point on, this series will always be referred to as the UK travel series.

Since the series is not defined for all real numbers, but only for positive ones, the Gaussian distribution could not be appropriate for it. Because of that, both the lognormal distribution and gamma distribution were chosen.

The gamma distribution parameterization chosen was $y_t \sim Gamma(\alpha, \lambda_{t|t-1})$, such that $E[y_t] = \lambda_{t|t-1}$ and $V[y_t] = \lambda^2_{t|t-1}/\alpha$. This parameterization was chosen to make the mean parameter explicit, so that one could easily define an appropriate unobserved component dynamic.

From a visual inspection of Figure 1, one could specify the appropriate dynamics for the mean parameters of both models. For simplicity, it was chosen to implement an unobserved components score-driven model with only the mean parameter being time-varying. Furthermore, given preliminary results, it was chosen to specify just the deterministic seasonal component, by fixing $\kappa_\gamma = 0$. It was also defined that the scale parameter d would be 1.

With that being said, the benchmark linear model was the BSM equivalent in the score-driven framework, but with a damped slope. Suppose $\mu_{t|t-1}$ is the mean parameter for both given

distributions. The model is:

$$\begin{aligned}
\mu_{t+1|t} &= m_{t+1|t} + \gamma_{t+1|t} \\
m_{t+1|t} &= m_{t|t-1} + b_{t|t-1} + \kappa_m s_t \\
b_{t+1|t} &= \phi b_{t|t-1} + \kappa_b s_t \\
\gamma_{t+1|t} &= \sum_{i=1}^{[S/2]} [\gamma_i \cos(\omega_i t) + \gamma_i^* \sin(\omega_i t)]
\end{aligned} \tag{13}$$

where S is the seasonal period of the series (12 for monthly series), $m_{t+1|t}$ is the trend component, $\gamma_{t+1|t}$ is the seasonal component, $\omega_{i,t} = 2\pi i/S$, and $s_t = I_{t|t-1}^{-d} \nabla_t$ is the standardized score.

Following the non-linear structural model presented in (Koopman and Lee [2009]), the non-linear unobserved components score-driven model was specified as:

$$\begin{aligned}
\mu_{t+1|t} &= m_{t+1|t} + e^{\beta m_{t+1|t}} \times \gamma_{t+1|t} \\
m_{t+1|t} &= m_{t|t-1} + b_{t|t-1} + \kappa_m s_t \\
b_{t+1|t} &= \phi b_{t|t-1} + \kappa_b s_t \\
\gamma_{t+1|t} &= \sum_{i=1}^{[S/2]} [\gamma_i \cos(\omega_i t) + \gamma_i^* \sin(\omega_i t)]
\end{aligned} \tag{15}$$

where S is the seasonal period of the series (12 for monthly series), $m_{t+1|t}$ is the trend component, $\gamma_{t+1|t}$ is the seasonal component, $\omega_{i,t} = 2\pi i/S$, and $s_t = I_{t|t-1}^{-d} \nabla_t$ is the standardized score.

It is important to notice that the benchmark additive model is a particular case of the non-linear model when $\beta = 0$, which is the reason this specific components interaction was chosen from (Koopman and Lee [2009]). Moreover, equation (14) does not change each unobserved component equation, only their interaction.

All models were implemented in the Julia programming language, using the non-linear optimization solver Ipopt (Dunning and Lubin [2013]). All necessary positivity constraints were specified as literal constraints in the optimization problem, ensuring the robustness and reliability of the estimation process. The non linear optimization problem written to estimate the unobserved components score-driven model with linear components interaction through MLE is:

$$\begin{aligned}
\min_{\Theta} \; & -\sum_{t=1}^{T} \ln p(y_t | f_{t|t-1}; \mathbf{Y_{t-1}}, \Theta) \\
s.t. \quad & f_{t+1|t} = m_{t+1|t} + \gamma_{t+1|t} \quad \forall t = 1, ..., T \\
& m_{t+1|t} = m_{t|t-1} + b_{t|t-1} + \kappa_m s_t \quad \forall t = 1, ..., T \\
& b_{t+1|t} = \phi b_{t|t-1} + \kappa_b s_t \quad \forall t = 1, ..., T \\
& \gamma_{t+1|t} = \sum_{i=1}^{[S/2]} [\gamma_i \cos(\omega_i t) + \gamma_i^* \sin(\omega_i t)] \quad \forall t = 1, ..., T \\
& -1 \leq \phi \leq 1
\end{aligned} \tag{16}$$

where $\Theta = \{\phi, \kappa_m, \kappa_b, \gamma, \gamma^*\}$, $f_{t|t-1}$ is the time varying parameter of a given distribution and $\omega_{i,t} = 2\pi i/S$.

4. Results

To validate the models' prediction accuracy, the UK travel time series was split into in-sample and out-of-sample portions. The in-sample portion consists of the first 212 months, from January 1980 to December 2005. The out-of-sample portion comprises the last 12 months, from January 2006 to December 2006.

To evaluate the forecasting accuracy for the additive and non-linear models, Figure 2 presents the ACF plot of the quantile residuals, which is the main focus of this section. Additionally, Figure 3 shows the 12-step-ahead point forecast out-of-sample plot. The MAPEs and MASEs for both in-sample and out-of-sample periods are available in Tables 2 and 3, respectively, along with the estimated hyperparameters in Table 1. Furthermore, Tables 4 shows the pvalues for some quantile residuals diagnostics, namely the jarque-bera normalitity test, the ljung-bux autocorrelation test and also the ljung-box test for the squared residuals in order to study heterocedasticity. Finally, Tables 5 and 6 reveal some descriptive statistics for the quantile residuals from both the lognormal and gamma fitted models.

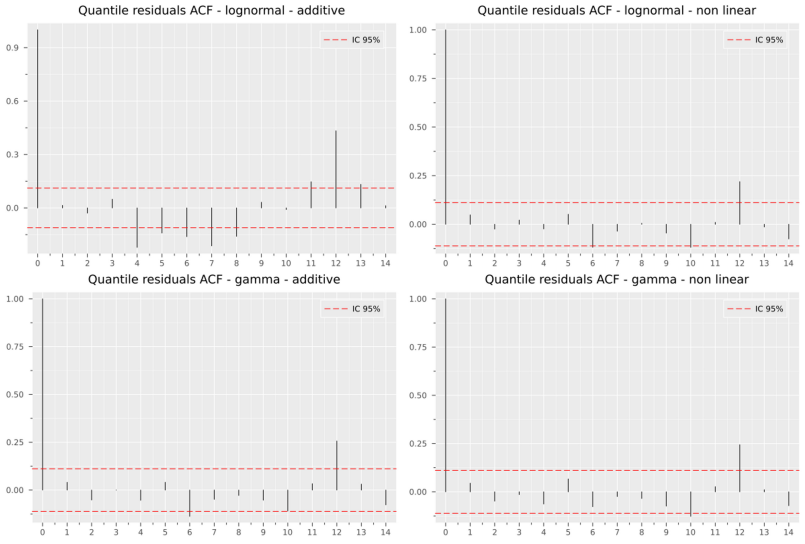

Figura 2: ACF for both gamma and lognormal, additive and non-linear models' quantile residuals.

Figure 2 illustrates the core objective of this work: demonstrating that the additive model fails to fully capture the seasonal dependency of the time series. This is evidenced by a significant spike in the ACF of the quantile residuals at the seasonal lag. The main result of this article is also highlighted in this figure. Notice that the lognormal non-linear model significantly reduces the seasonal lag spike, nearly reaching the confidence interval, indicating that the ACF is statistically indistinguishable from zero.

To corroborate this result, note that the estimated β for the lognormal non-linear model in Table 1 was significantly greater than 0, unlike the β for the gamma model. This indicates that the interaction of the trend and seasonal components is indeed non-linear for the lognormal model, while for the gamma model, the estimated β is practically 0, making it essentially an additive model. Consequently, little change is seen in the ACF in Figure 2 for the Gamma model.

Moreover, Tables 2 and 3 indicate that for both distributions, the estimated in-sample values and the 12-step-ahead out-of-sample forecasts were more accurate for the multiplicative models, even for the gamma model. The increase in forecast accuracy in the lognormal non-linear model can even be seen through visual inspection of Figure 3, in which the forecasted values follow

Estimated hyperparameters								
	Lognormal				Gamma			
	κ_m	κ_b	ϕ	β	κ_m	κ_b	ϕ	β
additive	0.66801	0.001	1.0	-	0.36407	0.001	1.0	-
non linear	0.29769	0.001	1.0	-0.7794	0.34748	0.001	1.0	0.00011

Tabela 1: Estimated hyperparameters for each fitted model.

	In sample MAPE(%)		Out of sample MAPE(%)	
	Lognormal	Gamma	Lognormal	Gamma
additive	6.8	5.36	8.51	7.09
non-linear	**5.14**	**5.2**	**6.06**	**6.76**

Tabela 2: MAPE's for in sample and out of sample predicted values for each fitted model.

the out of sample time series significantly closer than the additive model.

Some auxiliary and interesting results that can also be mentioned are the fact that the estimated κ_b for all models are practically equal to 0, indicating that this component was estimated deterministically. Additionally, the slope components were always estimated without damping, as all the ϕ values were estimated as 1.

Furthermore, the quantile residuals diagnostics through hypothesis tests also reveals some interesting results in Table 4. Notice that, for the lognormal fitted model, the non-linear model increased the jarque-bera normality test pvalue, meaning that this intercation was able to produce normaly distributed residuals. Even more, the ljung-box autocorrelation test on the squared residuals also show an increase in its pvalue, which means that the non-linear model improved the residuals homocedasticity, as expected. Those results, however, were not obtained with the gamma model.

5. Conclusion

The main goal of this article was achieved: demonstrating that non-linear combinations of trend and seasonal components in score-driven models can effectively capture the seasonal dynamics of a time series where linear models fall short. Our results are consistent with those obtained by (Koopman and Lee [2009]) using non-linear structural models. Not only did the non-linear model better capture the seasonal effect, but it also generated better in-sample adherence metrics and better out-of-sample forecast values for both lognormal and gamma distributions.

Even though the model was simplified with only the mean parameter being time-varying and the seasonal component being deterministic, the results were already satisfactory. This simpli-

	In sample MASE		Out of sample MASE	
	Lognormal	Gamma	Lognormal	Gamma
additive	0.42928	0.32925	0.68679	0.51979
non-linear	**0.31857**	**0.32132**	**0.44438**	**0.49015**

Tabela 3: MASE's for in sample and out of sample predicted values for each fitted model.

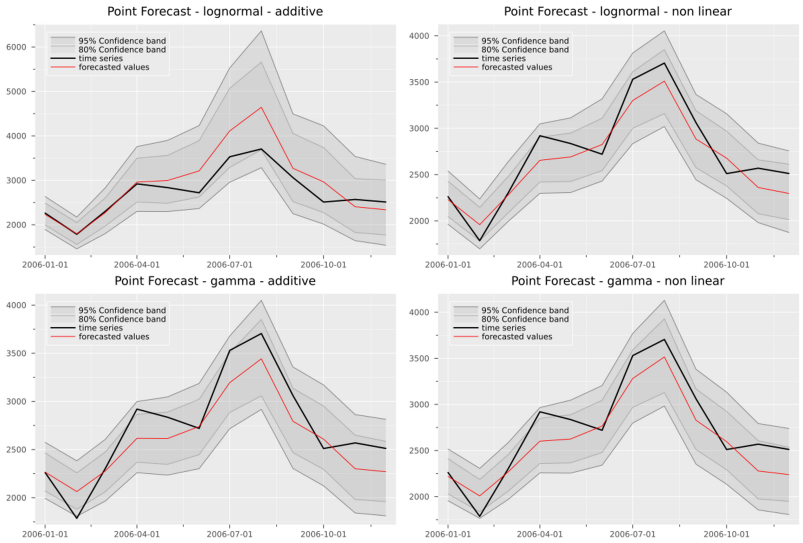

Figura 3: Out of sample forecast for both gamma and lognormal, additive and non-linear models.

Hypothesis tests pvalues						
	Lognormal			Gamma		
	jarque bera	ljung box	ljung box (resid^2)	jarque bera	ljung box	ljung box (resid^2)
additive	0.5418	0.0	0.0597	0.6212	0.0003	0.0248
non linear	0.6738	0.0057	0.1109	0.309	0.0009	0.0129

Tabela 4: Pvalues for residuals diagnostics hypothesis tests (jarque-bera, ljung-box and lkung-box for squared residuals).

residuals descriptive statistics - lognormal								
	mean	median	min	max	q_{05}	q_{95}	skewness	kurtosis
additive	0.0008	-0.0148	-2.5132	2.8389	-1.6609	1.6859	0.0999	-0.2339
non linear	0.0079	0.0611	-2.5725	2.805	-1.6332	1.6791	0.061	0.2146

Tabela 5: Quantile residuals descriptive statistics for each lognormal fitted model.

residuals descriptive statistics - gamma								
	mean	median	min	max	q_{05}	q_{95}	skewness	kurtosis
additive	-0.0121	0.006	-2.6514	2.8573	-1.6867	1.766	0.0608	0.2422
non linear	0.016	0.0127	-2.6121	2.8899	-1.6583	1.7717	0.1523	0.2975

Tabela 6: Quantile residuals descriptive statistics for each gamma fitted model.

fication did not hinder the model's ability to significantly reduce the seasonal lag spike in the ACF of the quantile residuals, as evidenced by the improved performance of the non-linear lognormal model.

Possible future expansions of this work include testing this methodology on time series with different patterns and generalizing the model to its full flexibility: allowing all parameters to be time-varying and incorporating a stochastic seasonal component. This would provide a more comprehensive evaluation of the model's capabilities and potentially enhance its applicability across various domains.

Referências

Ardia, D., Boudt, K., and Catania, L. (2016). Generalized autoregressive score models in r: The gas package. *arXiv preprint arXiv:1609.02354*.

Blazsek, S. and Escribano, A. (2023). Score-driven threshold ice-age models: Benchmark models for long-run climate forecasts. *Energy Economics*, 118:106522.

Bodin, G., Saavedra, R., Fernandes, C., and Street, A. (2020). Scoredrivenmodels. jl: A julia package for generalized autoregressive score models. *arXiv preprint arXiv:2008.05506*.

Caivano, M., Harvey, A., and Luati, A. (2016). Robust time series models with trend and seasonal components. *SERIEs*, 7:99–120.

Creal, D., Koopman, S. J., and Lucas, A. (2013). Generalized autoregressive score models with applications. *Journal of Applied Econometrics*, 28(5):777–795.

Dunning, I. and Lubin, M. (2013). Ipopt.jl github. `https://github.com/jump-dev/Ipopt.jl`. Access: 2023-01-26.

Durbin, J. and Koopman, S. J. (2012). Time series analysis by state space methods oxford university press.

Fuentes, F., Herrera, R., and Clements, A. (2023). Forecasting extreme financial risk: A score-driven approach. *International Journal of Forecasting*, 39(2):720–735.

Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*. Cambridge university press.

Harvey, A. C. (2013). *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*, volume 52. Cambridge University Press.

Hoeltgebaum, H., Borenstein, D., Fernandes, C., and Veiga, Á. (2021). A score-driven model of short-term demand forecasting for retail distribution centers. *Journal of Retailing*, 97(4):715–725.

Hyndman, R. and Athanasopoulos, G. (2021). Forecasting: principles and practice.

Hyndman, R., Koehler, A. B., Ord, J. K., and Snyder, R. D. (2008). *Forecasting with exponential smoothing: the state space approach*. Springer Science & Business Media.

Kalliovirta, L. (2012). Misspecification tests based on quantile residuals. *The Econometrics Journal*, 15(2):358–393.

Koopman, S. J. and Lee, K. M. (2009). Seasonality with trend and cycle interactions in unobserved components models. *Journal of the Royal Statistical Society Series C: Applied Statistics*, 58(4): 427–448.

Koopman, S. J. and Lit, R. (2019). Forecasting football match results in national league competitions using score-driven time series models. *International Journal of Forecasting*, 35(2):797–809.

Kushwah, A. K. and Wadhvani, R. (2019). Performance monitoring of wind turbines using advanced statistical methods. *Sādhanā*, 44(7):163.

Lit, R., Koopman, S., and Harvey, A. (2023). Time series lab. `https://timeserieslab.com/downloads-main`. Access: 2024-01-26.

Sarlo, R., Fernandes, C., and Borenstein, D. (2023). Lumpy and intermittent retail demand forecasts with score-driven models. *European Journal of Operational Research*, 307(3):1146–1160.

Xu, Y. and Lien, D. (2022). Forecasting volatilities of oil and gas assets: A comparison of gas, garch, and egarch models. *Journal of Forecasting*, 41(2):259–278.