

Introdução

O café arábica (*Coffea Arabica*) representa mais de 60% do café produzido mundialmente. É reconhecido por sua complexidade aromática e sabor adocicado, o que o define como uma das marcas mais valorizadas do mercado. Nesse sentido, estudos que capturam mudanças nos preços dessa *commodity* são de grande importância. O produto em sua grande maioria é negociado em sacas de 60kg e tem seus preços diários atualizados pela bolsa de Nova York. Em Minas Gerais existem alguns mercados que se baseiam no de Nova York e atuam realizando a compra interna do produto, esses mercados em geral representam os que compram dos pequenos produtores dos estados, dentre estes tem a cooperativa regional de cafeicultores localizada na cidade de Guaxupé.

Objetivos

O principal objetivo deste trabalho foi modelar a volatilidade condicional da série de preços do café arábica, e ajustar um modelo para estimar as variações no preço deste produto utilizando os dados disponibilizados pela cooperativa cooxupé.

Material e Método

A série temporal avaliada possui $n = 348$ observações, sendo que cada uma dessas observações corresponde ao preço médio mensal de uma saca de café de 60 kg, tipo 6 bebida dura bica corrida, preço médio avaliado ao final de cada mês pela Cooxupé. Os dados referidos são de janeiro de 1995 até dezembro de 2023. A análise inicial ocorreu fazendo a série de log retornos, que é definida por Morettin e Toloi (2006) da seguinte forma:

Definição 1: Seja P_t o preço de um ativo em um instante t . Por conveniência supõe-se que não são pagos dividendos no intervalo $(t, t - 1)$, defini-se então o retorno simples como:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}} = \frac{\Delta P_t}{P_{t-1}},$$

o log retorno de um ativo é então definido como sendo:

$$r_t = \log\left(\frac{P_t}{P_{t-1}}\right) = \log(1 + R_t).$$

Em seguida verificou-se o ACF e PACF das séries log retorno dos preços do café. O próximo passo foi realizar o teste de Cox Stuart para verificar a existência de alguma tendência na série log de retornos. O periodograma também foi utilizado para verificar a existência de alguma periodicidade na série, em conjunto o teste G de Fisher. Posteriormente, ajustou-se um modelo AR(p) com uma componente cosseno para modelar a periodicidade encontrada e a autocorrelação residual remanescente na série.

Definição 2: O processo autorregressivo de ordem p ou $AR(p)$ é representado pelo modelo $Y_t = \mu + \varphi_1 Y_{t-1} + \dots + \varphi_p Y_{t-p} + \varepsilon_t$, para todo $t = 1, \dots, n$ (MORETTIN E TOLOI, 2006). Os valores Y_{t-1}, \dots, Y_{t-p} são os p termos anteriores, a Y_t nos tempos t_1, \dots, t_{t-p} , os termos $\varphi_1, \dots, \varphi_p$, representam os parâmetros reais, μ simboliza uma média geral, ε_t indica o erro aleatório a qual se supõe ser um ruído branco com média zero e variância σ^2 independentemente do tempo t ele é indicado por $\varepsilon_t \sim RB(0, \sigma^2)$. Na análise residual foram utilizados os testes de Shapiro Wilk para verificar a normalidade, o teste de Box-Pierce para verificar a existência de autocorrelação e heterocedasticidade nos resíduos, a estatística do teste é definida por (LJUNG e BOX, 1978) como segue:

$$Q(K) = n(n+2) \sum_{j=1}^m \frac{\hat{r}_j^2}{n-j},$$

sendo \hat{r}_j a autocorrelação estimada da série no lag j ; n é a quantidade de observações e; m é o número de lags que estão sendo testados. Com o intuito de modelar a volatilidade existente nos resíduos, ajustou-se um modelo autorregressivo de heteroscedasticidade condicional generalizada, conhecido como modelo GARCH(r,s), (MORETTIN E TOLOI, 2006) o define da seguinte forma:

Definição 3:

$$X_t = \sqrt{h_t} \varepsilon_t$$

$$h_t = \alpha_0 + \sum_{i=1}^r \alpha_i X_{t-i}^2 + \sum_{j=1}^s \beta_j h_{t-j},$$

em que ε_t i.i.d. (0,1), $\alpha_0 > 0$, $\alpha_i \geq 0$, $\beta_j \geq 0$, $\sum_{i=1}^q (\alpha_i + \beta_j) < 1$, $q = \max(r, s)$. Devido a sua simplicidade e eficiência para capturar a volatilidade presente nos dados, o modelo GARCH(1,1) é comumente utilizado no processo de modelagem da volatilidade de séries de retorno, (Brooks e Burke 2003). O modelo pode então ser expresso por:

$$h_t = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 h_{t-1},$$

em que $0 \leq \alpha_1$, $\beta_1 < 1$, $\alpha_1 + \beta_1 < 1$.

Resultados e Discussão

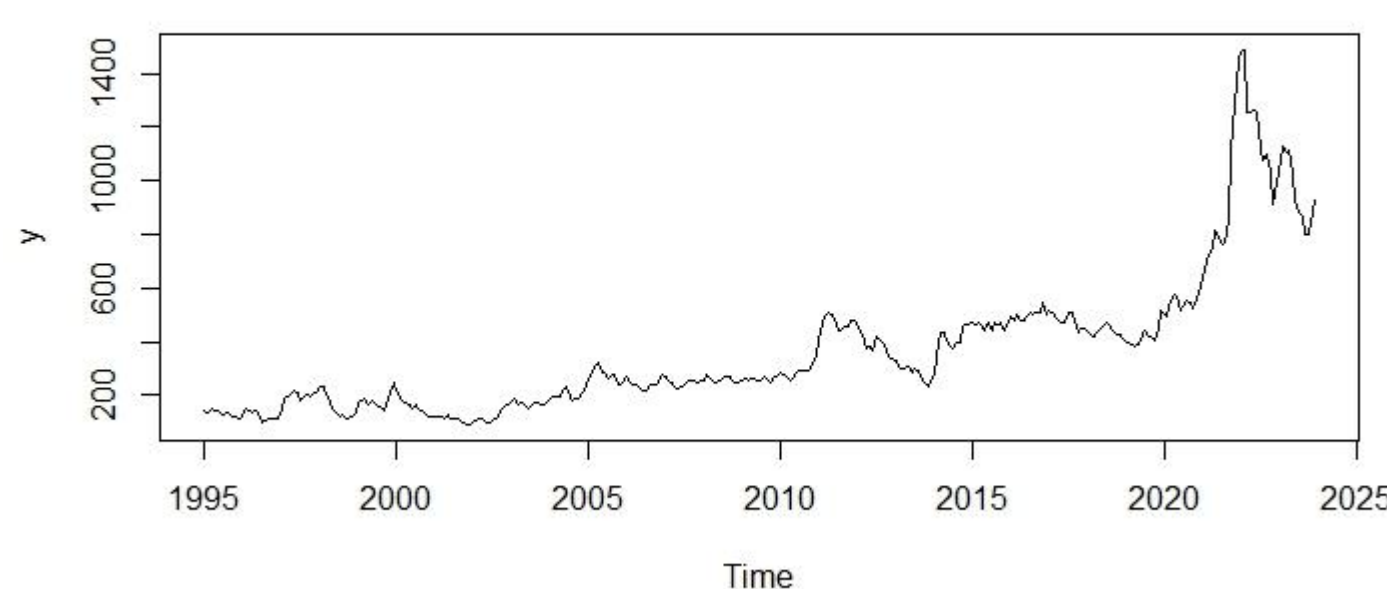


Figura 1: Série histórica de preços

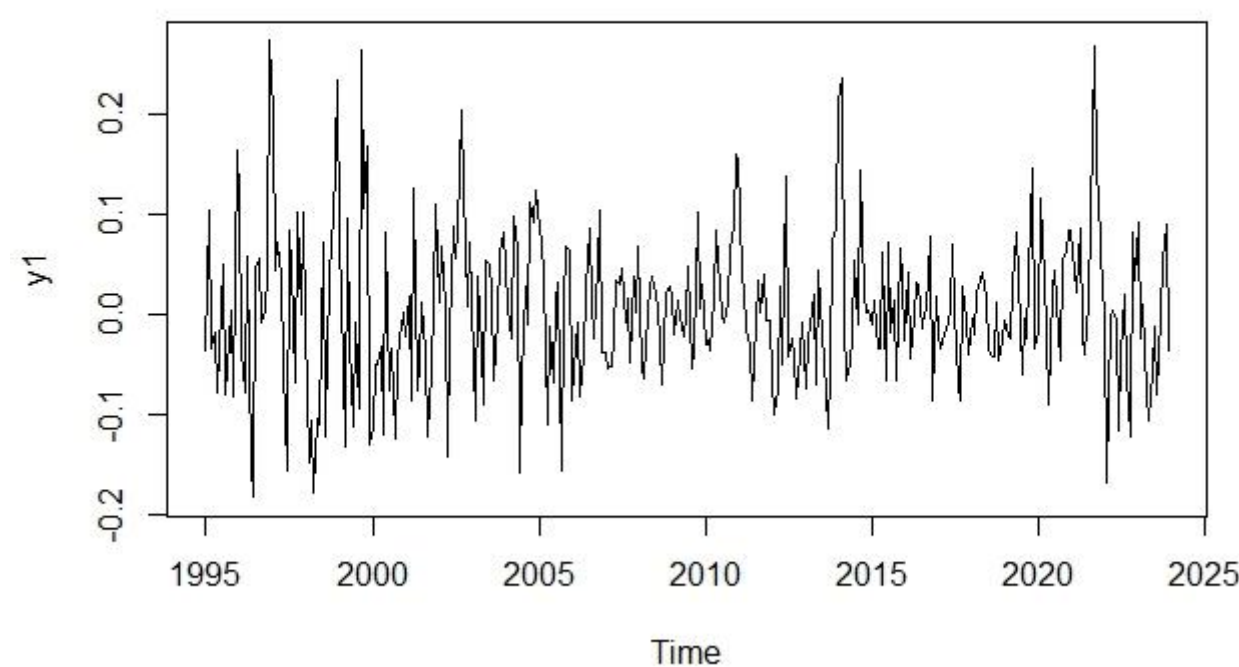


Figura 2: Série log retorno dos preços

O teste de Cox - Stuart indicou que não existem indícios de tendência e rejeitou a hipótese alternativa a 5% de probabilidade (valor de p de 49,52%). Por outro lado, o teste G de Fisher juntamente com o periodograma rejeitaram a hipótese de nulidade de que a periodicidade é não significativa a 5% de probabilidade, (valor p de 0,0853%) e indicaram a frequência angular $w = 0,52$ (valor de $p = 0,0853\%$), o que equivale a uma periodicidade de 12 meses. Para modelar a volatilidade e contornar a periodicidade existente, ajustamos um modelo AR(1) com uma componente cosseno.

O ajuste com a componente de cossenos foi escolhido no lugar de um componente seno pois o primeiro teve a estimativa significativa, diferentemente do seno, que não apresentou significância a 5% de probabilidade. Os valores das estimativas, significativas foram; $AR(1)(\hat{\phi}_1 = 0.2472)$, $\cos(\hat{\eta} = 0.0254)$. A figura 3 mostra que o quadrado do resíduo apresenta fortes indícios de autocorrelação, o teste de Box-Pierce confirma tal fato com um (valor p de 0.1519%).

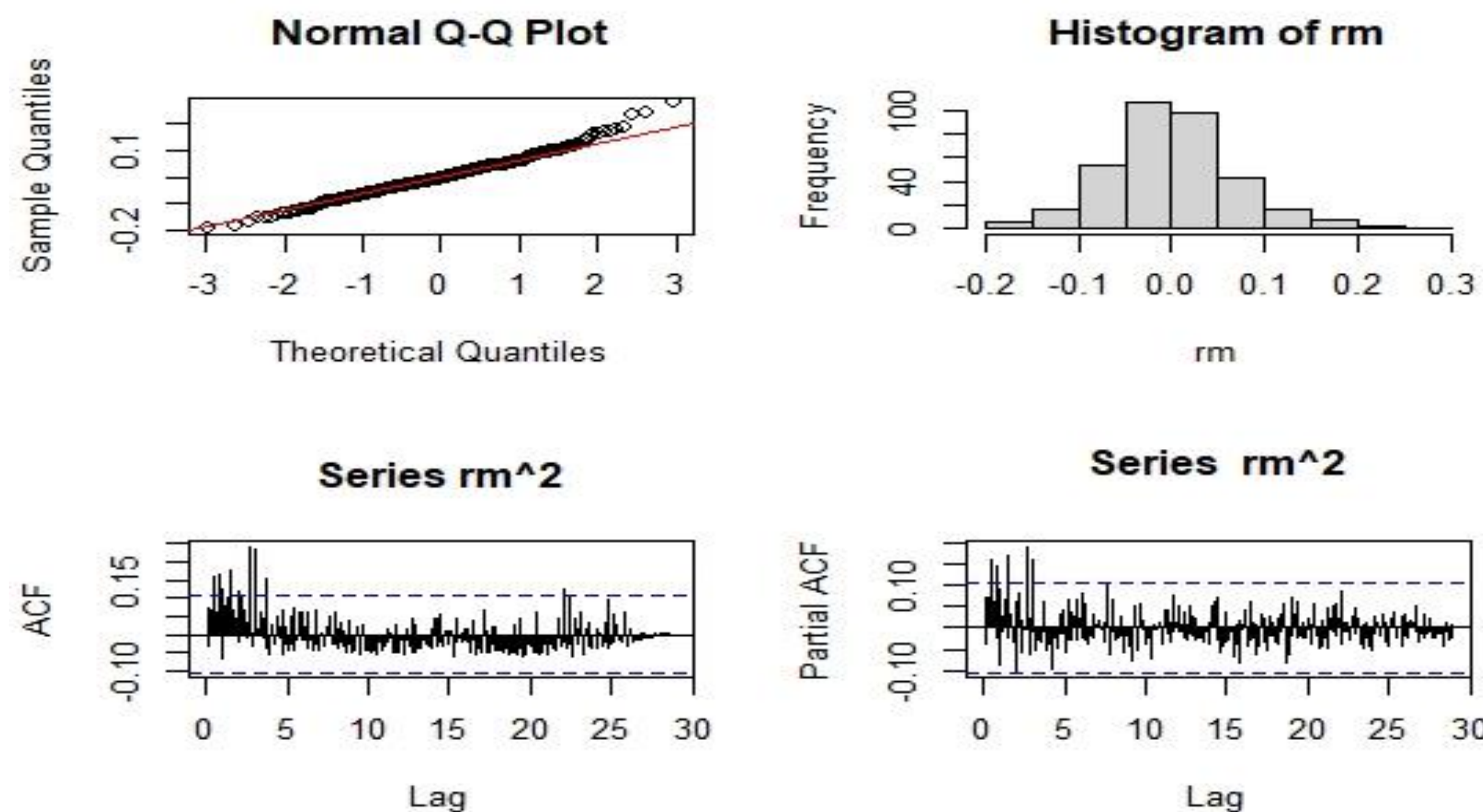


Figura 3: Gráfico da análise de resíduos do modelo AR(1)

Para contornar essa situação e capturar a assimetria e a leptocurtose existente nas séries de retornos, ajustamos um modelo GARCH(1,1) juntamente com uma distribuição skew t-Student, para modelar a autocorrelação nos resíduos. Outras distribuições também foram testadas, a escolha do modelo se deu em função do maior valor para a função Log de verossimilhança. O modelo escolhido apresentou a estimativa para a função de verossimilhança de $\hat{l} = 442.5308$, os parâmetros $\hat{\alpha}_1 = 0,04523$, $\hat{\beta}_1 = 0,9393$, $\hat{\nu} = 10$, $\hat{\lambda} = 1,159$, (todos eles significativos a 5% de probabilidade). A especificação final do modelo é dada por:

$$Y_t = 0.2472137 \cdot Y_{t-1} + 0.2774813 \cdot \cos(0.52t) + X_t,$$

em que X_t segue um modelo GARCH(1,1) do tipo $X_t = \sqrt{h_t} \varepsilon_t$. No qual $h_t = 0.04523 \cdot X_{t-1}^2 + 0.9393 \cdot h_{t-1}$, com $\varepsilon_t \sim \text{skew-t}(10, 1.159)$.

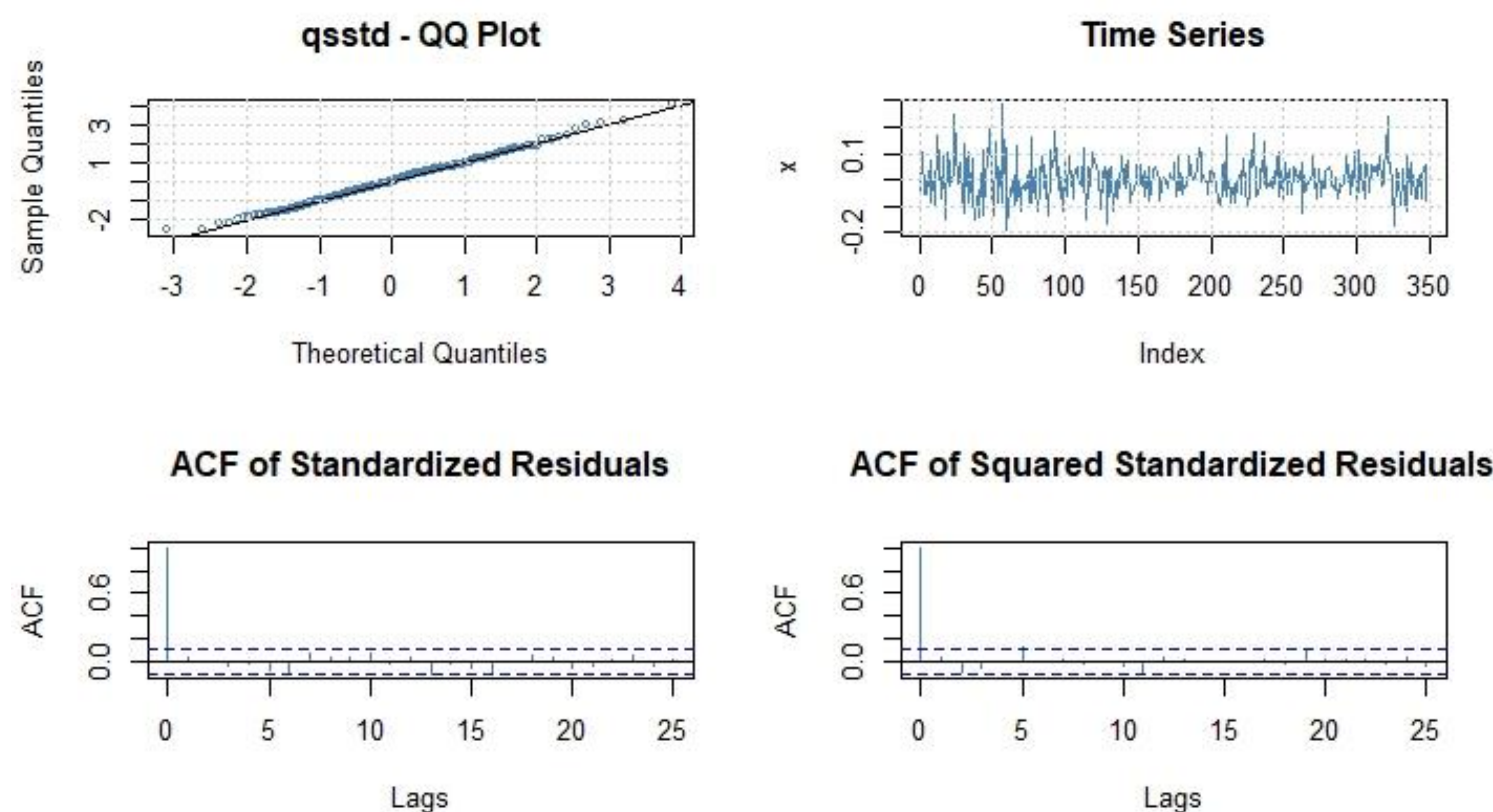


Figura 4: Gráfico da análise de resíduos do modelo GARCH(1,1)

Para verificar a adequação do modelo foram aplicados os testes de Ljung-Box aos resíduos padronizados, $Q(10) = 10.892(0.366)$, $Q(15) = 17.174(0.308)$, $Q(30) = 23.112(0.283)$, e aos quadrados do resíduo padronizado, $Q(10) = 11.928(0.289)$, $Q(15) = 16.854(0.327)$, $Q(30) = 22.248(0.327)$. Os resultados indicam a não existência de autocorrelação nos resíduos.

Conclusões

Constatou-se que os resíduos do modelo GARCH(1,1) podem ser considerados ruídos branco ($p = 30,89\%$), além disso a projeção intervalar realizada para a volatilidade condicional indicou que os preços do café para os próximos 6 meses podem cair em até 12% ou subir em até 14%, aproximadamente.

Bibliografia

BAILLIE, R. T.; BOLLERSLEV, T. The message in daily exchange rates: a conditional-variance tale. **Journal of Business and Economic Statistics**, v. 7, n. 3, p. 297-305, 1986.

BROOKS, C.; BURKE, S. P. Information criteria for GARCH model selection. **The European journal of finance**, v. 9, n. 6, p. 557-580, 2003.

LJUNG, G. M.; BOX, G. E. On a measure of lack of fit in time series models. **Biometrika**, 174 v. 65, n. 2, p. 297-303, 1978.

MORETTIN, P. A.; TOLOI, C. M. C. **Análise de séries temporais**. 2. ed. São Paulo: E. Blücher, 538 p. 2006.

BROOKS, C.; BURKE, S. P. Information criteria for GARCH model selection. **The European journal of finance**, v. 9, n. 6, p. 557-580, 2003.