



Séries Temporais e Previsão

Mestrado em Ciência de Dados

[OBJ]

Leonardo Nogueira Pucetti - 2230394

Daniel Filipe Ferreira Leal - 2220123

Leiria, Janeiro de 2025

Índice

Índice.....	4
Índice Imagens.....	5
1. Introdução.....	6
2. Análise do Dataset Hor.....	7
2.1. Descrição e Análise Preliminar.....	7
2.2. Médias Móveis.....	9
2.3. Decomposição e Modelagem Holt.....	10
3. Análise do Dataset Hycarb.....	12
3.1. Descrição e Análise Preliminar.....	13
3.2. Modelagem Holt-Winters.....	14
3.3. Modelagem SARIMA.....	15
3.4. Comparação de Modelos e Conclusões.....	20
4. Análise dos Retornos Diários de Fechamento das Ações da Ford.....	21
4.1. Retornos Diários Ajustados.....	21
4.1.1. Inicialização dos Dados e Visualização Preliminar.....	21
4.1.2. Especificação do Modelo GARCH.....	22
4.1.3. Ajuste e Comparação de Modelos.....	22
4.1.4. Interpretação dos Resultados e Previsões.....	24
4.2. Retornos Diários Fechamento.....	25
4.2.1. Inicialização dos Dados e Visualização Preliminar.....	25
4.2.2. Especificação de Modelos EGARCH e GJR-GARCH.....	26
4.2.3. Ajuste e Comparação de Modelos.....	26
4.2.4. Interpretação dos Resultados e Previsões.....	28

Índice Imagens

Imagem 1: Taxa de Ocupação de Hotéis no Havaí.....	7
Imagem 2: Função de Autocorrelação (ACF).....	8
Imagem 3: Função de Autocorrelação (ACF).....	9
Imagem 4: Gráfico com as 6 séries em separado.....	10
Imagem 5: Previsões pelo método Holt.....	11
Imagem 6: Decomposição aditiva.....	12
Imagem 7: Decomposição multiplicativa.....	12
Imagem 8: Concentração de hidrocarbonetos entre 1970 e 1979.....	13
Imagem 9: Comparativo ACF e PACF ts Hycarb.....	14
Imagem 10: Comparativo entre modelos aditivo e multiplicativo.....	15
Imagem 11: Boxplot Série temporal Hycarb.....	16
Imagem 12: ACF e PACF série temporal Hycarb.....	18
Imagem 13: Modelo SARIMA com previsão para 12 períodos de tempo.....	19
Imagem 14: Comparativo previsões HoltWinters e SARIMA.....	20
Imagem 15: Retornos diários ajustados ações Ford.....	21
Imagem 16: Histograma dos Resíduos Padronizados.....	23
Imagem 17: QQ-Plot dos Resíduos Padronizados.....	23
Imagem 18: Gráfico das Previsões do Modelo GARCH.....	24
Imagem 19: Fechamentos diários ações Ford.....	25
Imagem 20 e 21: Comparativo entre os 6 modelos.....	27
Imagem 21: Modelo eGARCH(2,1).....	28

1. Introdução

Este trabalho acadêmico explora a aplicação de técnicas avançadas de análise de séries temporais em três datasets distintos, cada um representando variáveis econômicas e ambientais específicas. Os datasets analisados incluem: a taxa de ocupação de hotéis no Havaí, as concentrações de hidrocarbonetos semanais, e os retornos diários de fechamento das ações da Ford, com uma análise adicional sobre os retornos diários ajustados. Através destas análises, buscou-se entender padrões, tendências, e comportamentos sazonais que possam influenciar as decisões de negócios, políticas ambientais, ou estratégias de investimento.

Para cada série temporal, foi empregado uma série de modelos estatísticos e métodos de decomposição para decompor as séries em seus componentes de tendência, sazonalidade, e ruído, além de aplicar modelos preditivos para avaliar o comportamento futuro das séries. Os modelos e técnicas utilizados incluem:

- **Médias Móveis:** Usadas para suavizar as séries temporais e identificar tendências mais claras, removendo flutuações de curto prazo.
- **Decomposição:** Técnica utilizada para separar uma série temporal em seus componentes básicos. Aplicando tanto decomposições aditivas quanto multiplicativas, dependendo da natureza e comportamento dos dados.
- **Modelos Holt-Winters e Holt:** Modelos de suavização exponencial que ajustam a série temporal para tendência e sazonalidade, úteis para fazer previsões de curto prazo.
- **Modelagem SARIMA (Seasonal Autoregressive Integrated Moving Average):** Um modelo avançado que ajusta não apenas as autocorrelações, mas também as não estacionariedades e sazonalidades dentro dos dados, fornecendo uma ferramenta poderosa para análise e previsão de séries temporais complexas.
- **Modelos GARCH (Generalized Autoregressive Conditional Heteroskedasticity):** Para o estudo dos retornos diários de fechamento e ajustados da Ford, foi analisada a volatilidade e previsões de mercado através de modelos GARCH, crucial para a gestão de risco e estratégias de trading.

O primeiro dataset aborda a taxa de ocupação de hotéis no Havaí, uma métrica vital para entender a saúde econômica do setor turístico nesta região. O segundo dataset, as concentrações de hidrocarbonetos, é essencial para monitoramento ambiental e políticas de saúde pública. O terceiro dataset compreende os retornos diários das ações da Ford, proporcionando insights sobre a volatilidade do mercado de ações e o comportamento econômico que pode influenciar decisões de investidores e stakeholders. Adicionalmente, o estudo dos retornos diários ajustados busca identificar períodos de maior volatilidade e modelar a natureza dinâmica dos dados financeiros para previsões futuras.

Através desta análise, procurou-se oferecer uma visão holística e multidimensional dos padrões das séries temporais, utilizando os dados históricos para fazer previsões informadas e apoiar a tomada de decisões baseada em evidências. Esta introdução serve como a base para a exploração detalhada de cada dataset nos capítulos subsequentes, onde cada análise específica é contextualizada e discutida em profundidade.

2. Análise do Dataset Hor

2.1. Descrição e Análise Preliminar

O dataset 'Hor' compreende dados sobre a taxa trimestral de ocupação de hotéis no Havaí, cobrindo o período do primeiro trimestre de 1982 até o quarto trimestre de 2015. A série temporal é explorada inicialmente através de uma visualização dos seus primeiros valores, que revela as taxas de ocupação no início do período estudado. A partir do resumo estatístico gerado têm-se alguns insights sobre a distribuição dos dados, incluindo média, mediana, valores mínimo e máximo, facilitando uma primeira compreensão das características gerais da série.

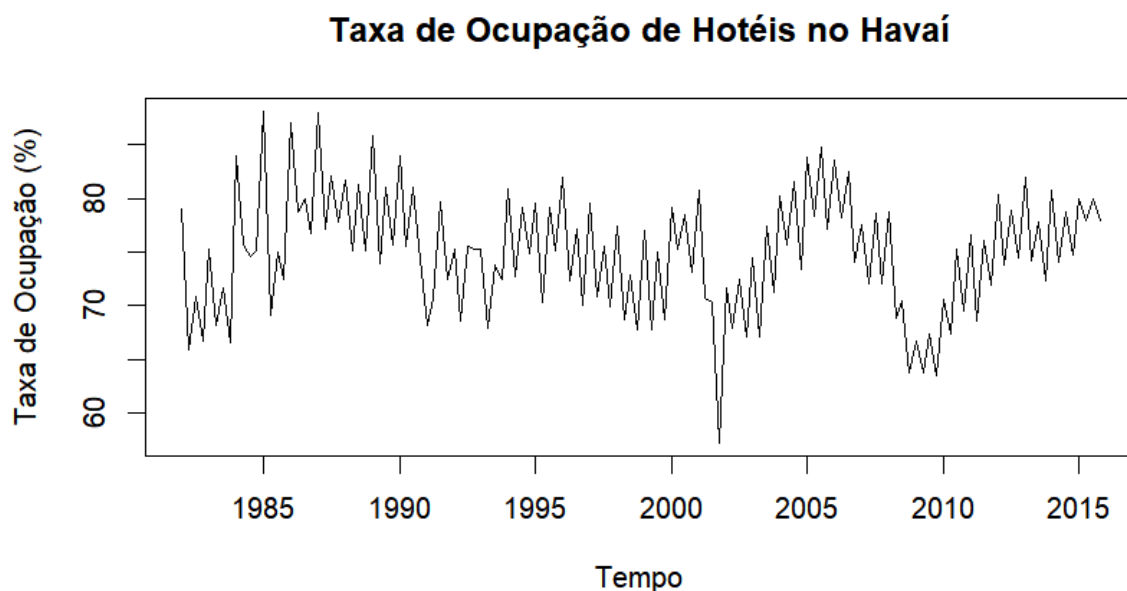


Imagem 1: Taxa de Ocupação de Hotéis no Havaí

A partir da Imagem 1, foi possível observar que há tendências ao longo da série temporal. A taxa de ocupação dos hotéis flutua, mas há períodos de aumento ou diminuição gradual. Por exemplo, há um aumento nos primeiros anos (1982 a 1990), uma queda entre

1990 e 2002, seguido de um crescimento até perto de 2008, depois veio a crise de 2008 houve uma queda e de seguida uma recuperação gradual após 2010.

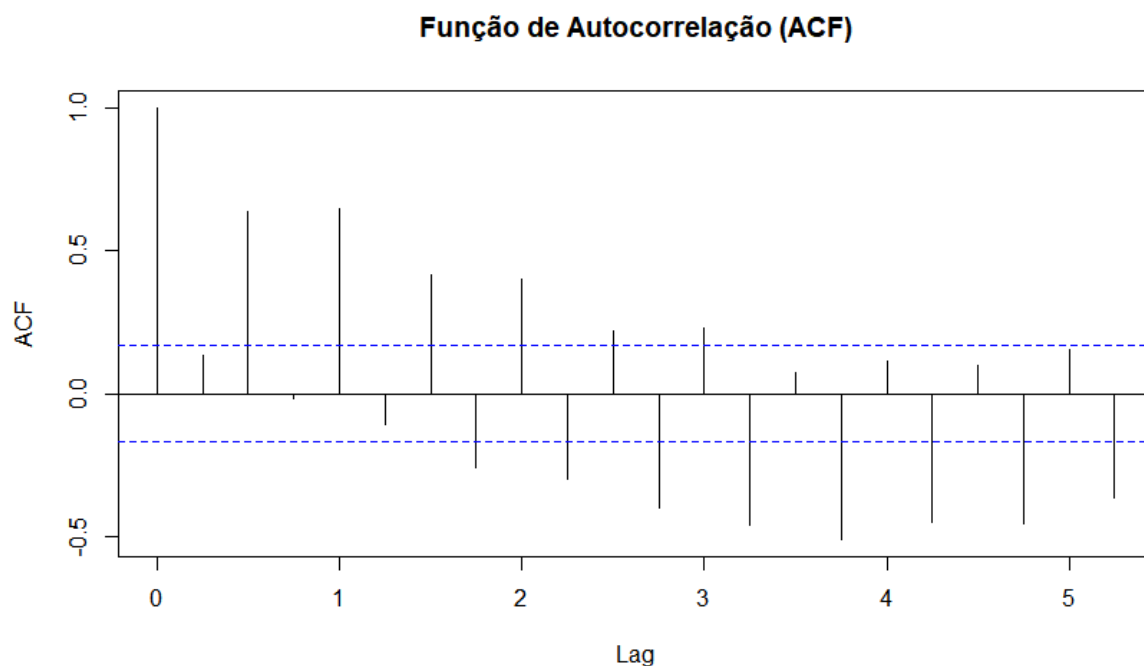


Imagem 2: Função de Autocorrelação (ACF)

O gráfico da ACF (Imagem 2) mostra picos significativos nos primeiros lags, mas não há um padrão claro de repetição regular em múltiplos da frequência. Isso sugere que, até este ponto, não há uma sazonalidade evidente que se repita de forma consistente.

Como a frequência da série é trimestral (4 observações por ano), é esperado observar picos significativos nos lags múltiplos de 4. No entanto, o gráfico da ACF não apresenta picos claros nesses lags, o que indica a ausência de uma sazonalidade trimestral clara.

A análise da função de autocorrelação não sugere a presença de sazonalidade na série temporal. Se a série possuir sazonalidade, ela não é suficientemente forte para se refletir na função de autocorrelação de forma significativa.

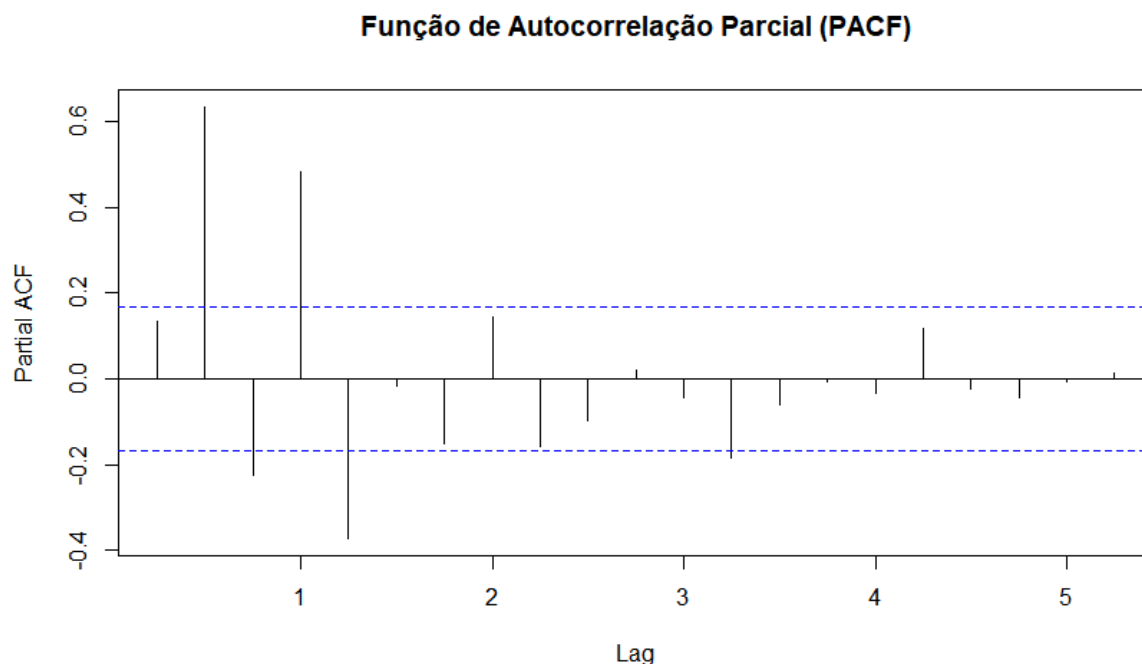


Imagem 3: Função de Autocorrelação (ACF)

O gráfico da PACF mostra picos significativos nos primeiros lags. Contudo, os lags posteriores apresentam valores muito próximos de zero e não há outros picos significativos.

Se existisse uma sazonalidade clara, espera-se que seja possível identificar picos significativos no lag correspondente à frequência da série (neste caso, no lag 5, devido à periodicidade trimestral). No entanto, o lag 5 não apresenta um pico significativo, reforçando a ideia de que não há uma componente sazonal evidente na série.

Concluindo a análise da função de autocorrelação parcial (PACF) também não sugere a existência de sazonalidade na série. Os picos nos primeiros lags refletem correlações de curto prazo, mas não há padrões sazonais claros em múltiplos da frequência da série (trimestral).

2.2. Médias Móveis

Para entender melhor as tendências subjacentes e suavizar possíveis flutuações de curto prazo na taxa de ocupação, aplicou-se médias móveis de diferentes ordens (2, 3, 4, 5, e 6). As médias móveis são implementadas utilizando a função `stats::filter`, que ajusta a série com base na média dos valores em uma janela especificada, resultando em várias séries suavizadas. Cada uma das médias móveis foram analisadas em gráficos separados, que demonstram como as oscilações são atenuadas com o aumento da ordem da média móvel.

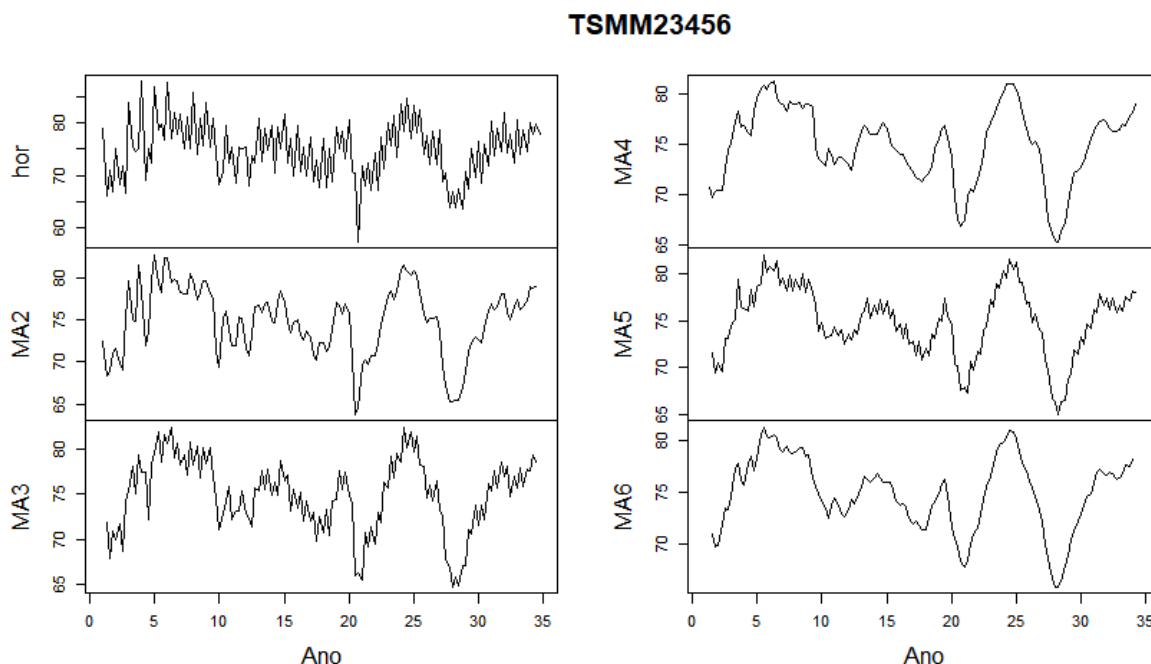


Imagem 4: Gráfico com as 6 séries em separado

Ao analisar os gráficos das médias móveis, concluiu-se que as médias móveis de números ímpares (como MA3 e MA5) não são as mais adequadas para esta série temporal. Isso ocorre porque as linhas resultantes apresentam muitas oscilações abruptas, comprometendo a suavização dos dados. Essa característica é evidente nos gráficos, onde a variabilidade permanece elevada, dificultando a identificação de tendências mais claras.

Além disso, considerando que a frequência da série temporal é trimestral (ou seja, 4 observações por ano), optou-se por adotar a média móvel de número par, especificamente a de valor 4 (MA4). Essa escolha é justificada porque uma média móvel de número igual à frequência da série temporal oferece uma suavização mais equilibrada, respeitando a estrutura cíclica dos dados e proporcionando uma melhor representação da tendência subjacente.

Assim, pode-se concluir que a média móvel de ordem 4 é a mais apropriada para esta análise.

2.3. Decomposição e Modelagem Holt

Foi utilizado então o modelo de Holt para prever futuras ocupações, ajustando o modelo aos dados com parâmetros de suavização que minimizem o erro de previsão. Os parâmetros de suavização, como α (0.2631) e β (1×10^{-4}), são calibrados para refletir a contribuição da última observação e da tendência anterior, respectivamente.

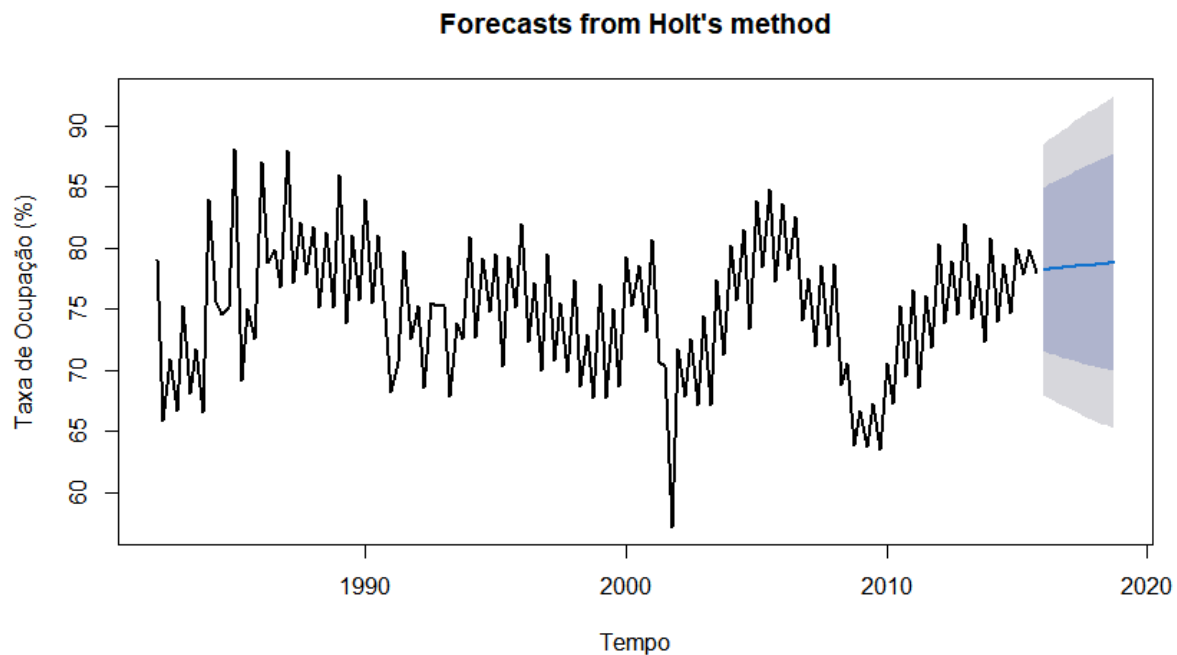


Imagem 5: Previsões pelo método Holt

A partir do estudo foi possível encontrar as seguintes equações de suavização para uso:

Equação de nível:

$$l_t = 0,2631 * x_t + (1 - 0,2631)(l_{t-1} + T_{t-1})$$

Equação de tendência:

$$T_t = 0,0001(l_t - l_{t-1}) + (1 - 0,0001)T_{t-1}$$

Foi aplicado a técnica de decomposição, que separa a série temporal em componentes de tendência, sazonalidade e ruído. Este processo é visualizado em um gráfico que mostra claramente como a série pode ser decomposta.

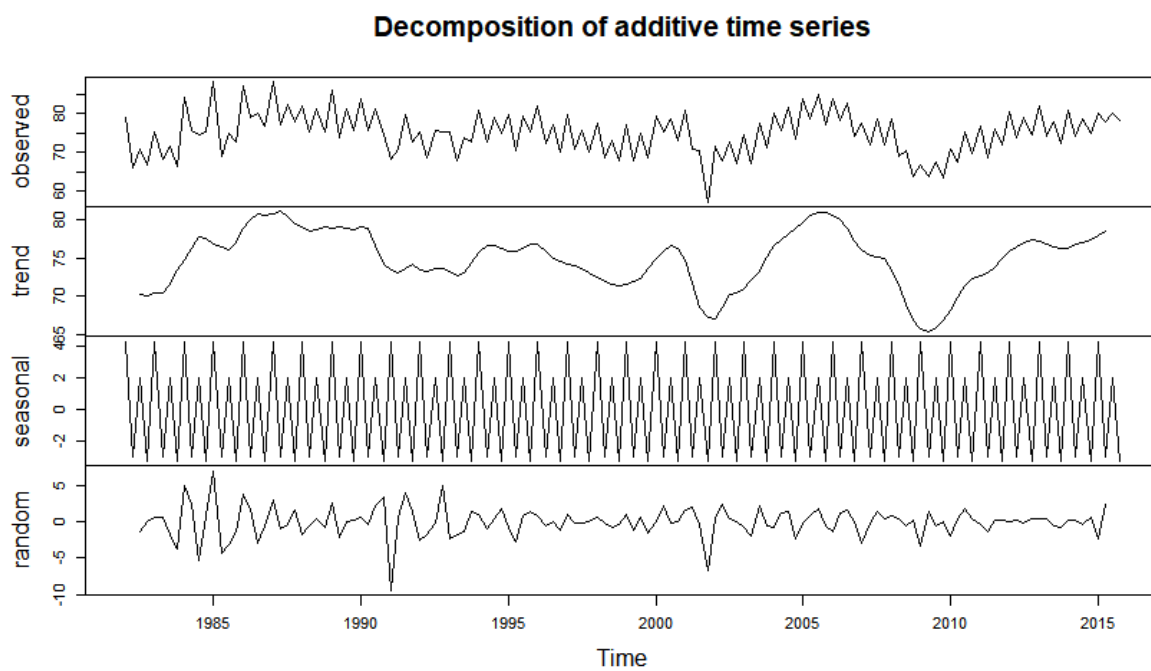


Imagem 6: Decomposição aditiva

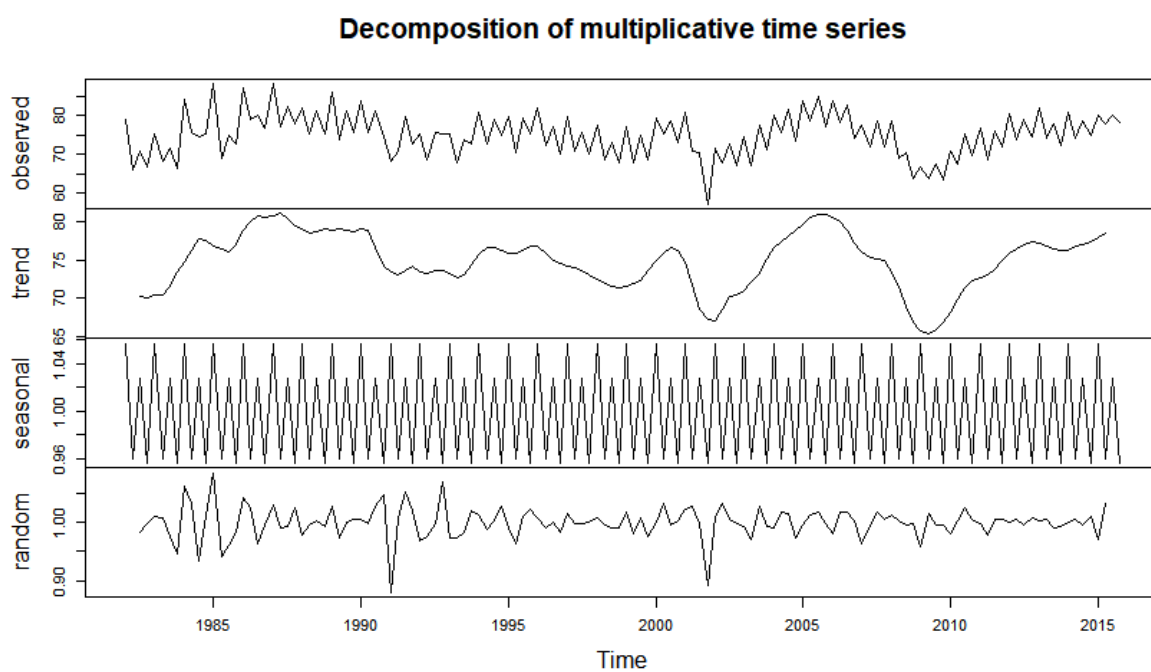


Imagem 7: Decomposição multiplicativa

3. Análise do Dataset Hycarb

3.1. Descrição e Análise Preliminar

O dataset 'Hycarb' registra as concentrações semanais de hidrocarbonetos, captando as variações dessas concentrações de 1970 até a semana 40 de 1979. Os dados são inicialmente explorados através de visualizações básicas e estatísticas descritivas para fornecer uma compreensão inicial da distribuição e das características fundamentais da série.

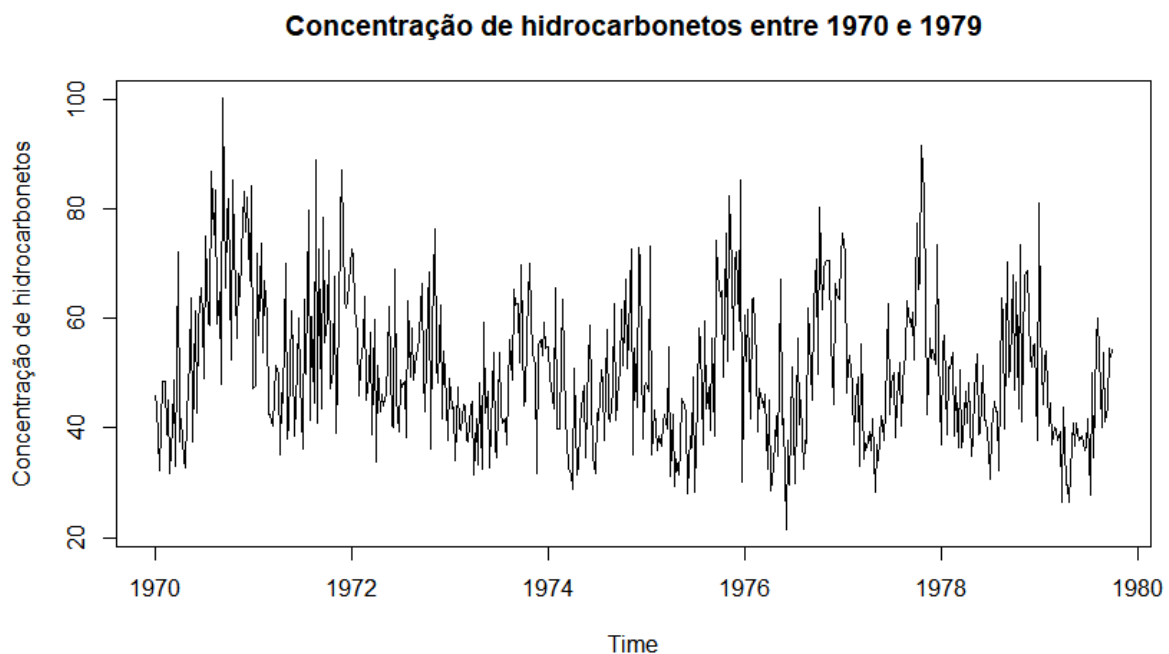


Imagem 8: Concentração de hidrocarbonetos entre 1970 e 1979

É possível observar uma componente sazonal na série temporal, caracterizada pela repetição de padrões aparentemente ano a ano. Dado que a série possui uma frequência semanal, essa sazonalidade apresenta um período próximo de 52 semanas.

Esta série não apresenta uma tendência clara ao longo do tempo, contudo nota-se a presença de um possível ciclo entre os anos de 1970-1971 e entre 1977-1978, com picos significativos de concentração em comparação com os demais períodos da série. Além disso, observa-se uma variação fora do padrão, sugerindo possíveis fatores externos ou fenômenos atípicos influenciando a série nesses intervalos específicos.

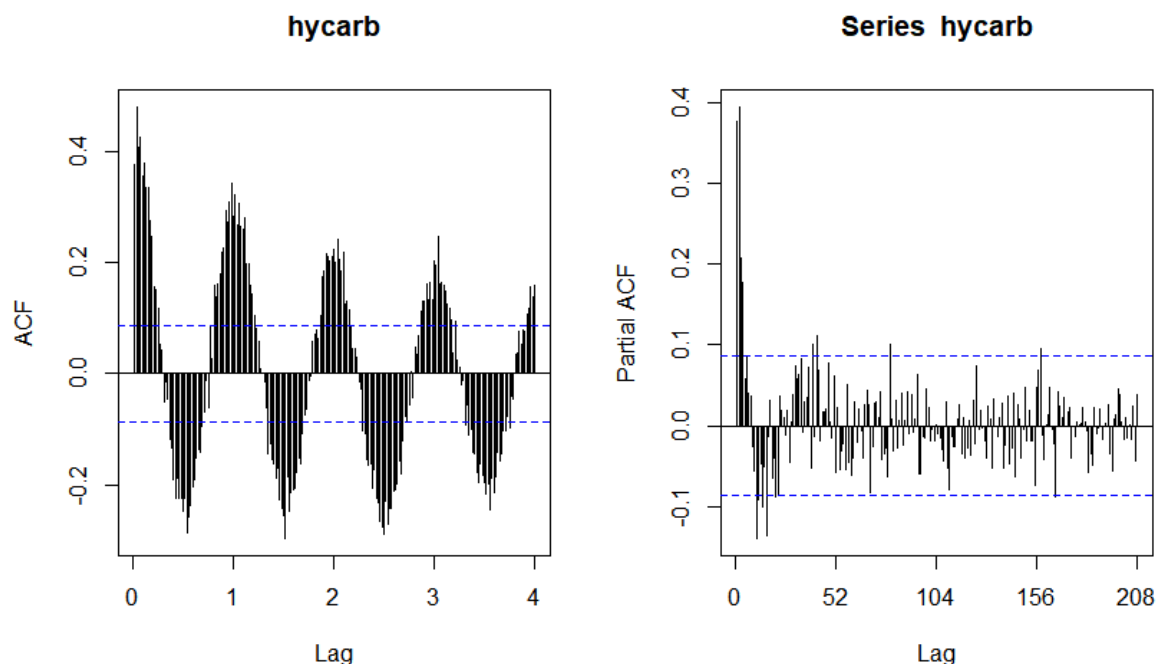


Imagem 9: Comparativo ACF e PACF ts Hycarb

Apartir da avaliação dos gráficos PACF e ACF aparentemente a série não é estacionária, logo possivelmente essa componente deverá ser trabalhada para a composição de um modelo SARIMA. A avaliação nos leva a crer também que a série temporal não possui uma tendência clara ou minimamente pode possuir uma tendência de longo prazo. Contudo fica clara a existência de uma sazonalidade de lag 52, logo uma sazonalidade ano a ano.

3.2. Modelagem Holt-Winters

Foi utilizado o modelo Holt-Winters para aplicar uma suavização exponencial tripla, ajustando a série temporal para componentes de tendência, sazonalidade e erro. O modelo é ajustado com opções de sazonalidade aditiva e multiplicativa, e comparou-se os resultados para determinar qual configuração se adapta melhor aos dados. As equações do modelo Holt-Winters ajustam os parâmetros para otimizar a previsão de curto prazo, e os resultados são visualizados para avaliar a eficácia do modelo em capturar a dinâmica da série temporal.

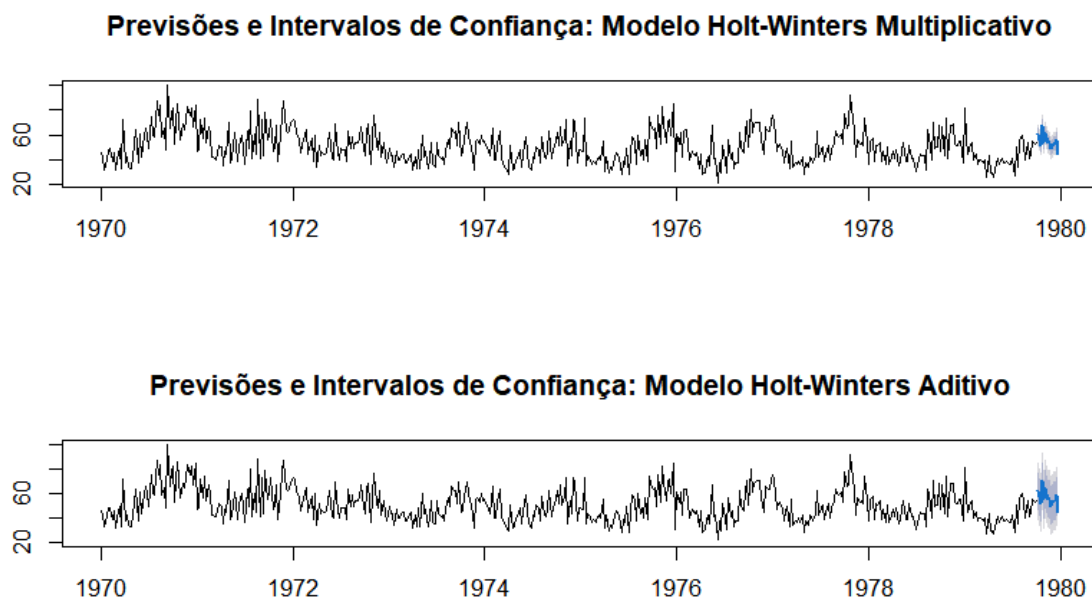


Imagem 10: Comparativo entre modelos aditivo e multiplicativo

Após avaliação dos métodos aditivo e multiplicativo do modelo HoltWinters vê-se que o modelo multiplicativo seria o mais adequado. Essa decisão, apesar de não ser clara com base nos gráficos, fica mais evidente uma vez que é possível comparar as medidas de erro de cada um dos modelos em que temos que os valores RMSE, MAE e MAPE são menores para o modelo Multiplicativo em comparação com o aditivo (Tabela 1).

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF
Multiplicativo	0.011832	11.98449	9.072858	-4.00488	18.8055	0.757568	-0.11887
Aditivo	0.542074	12.28347	9.246977	-2.73997	19.04204	0.772106	-0.12116

Tabela 1: Comparativo de erros modelo Aditivo x Multiplicativo

3.3. Modelagem SARIMA

A análise inicial identificou a presença de dois outliers na série temporal (100.12 e 91.47) com base no boxplot.

Estudo de possíveis outliers

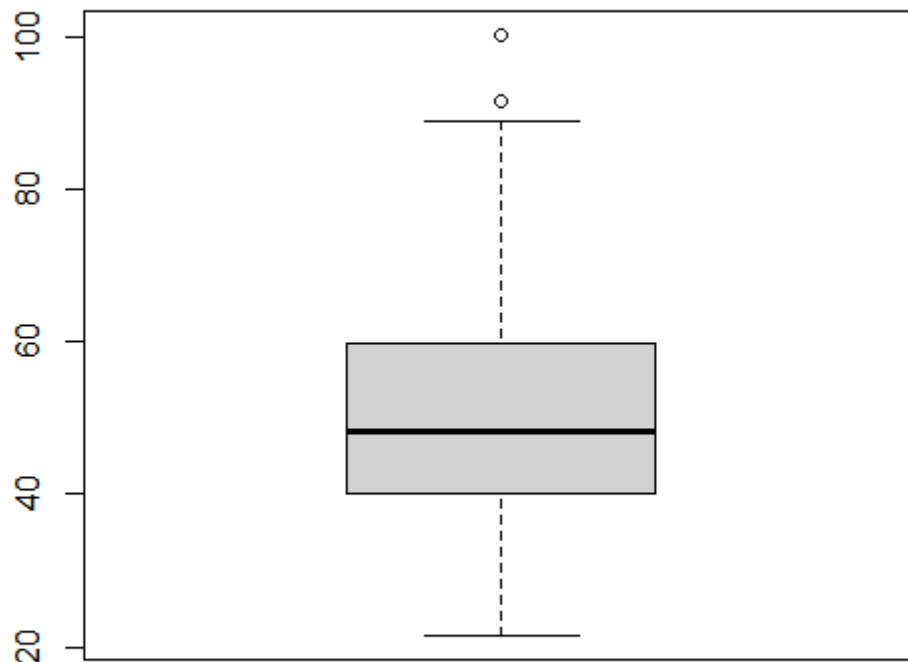


Imagem 11: Boxplot Série temporal Hycarb

No entanto, como esses valores representam apenas 0,4% dos dados totais (2 em 508 observações), concluiu-se que os outliers não representam um problema significativo para a modelagem. Assim, eles foram mantidos na série.

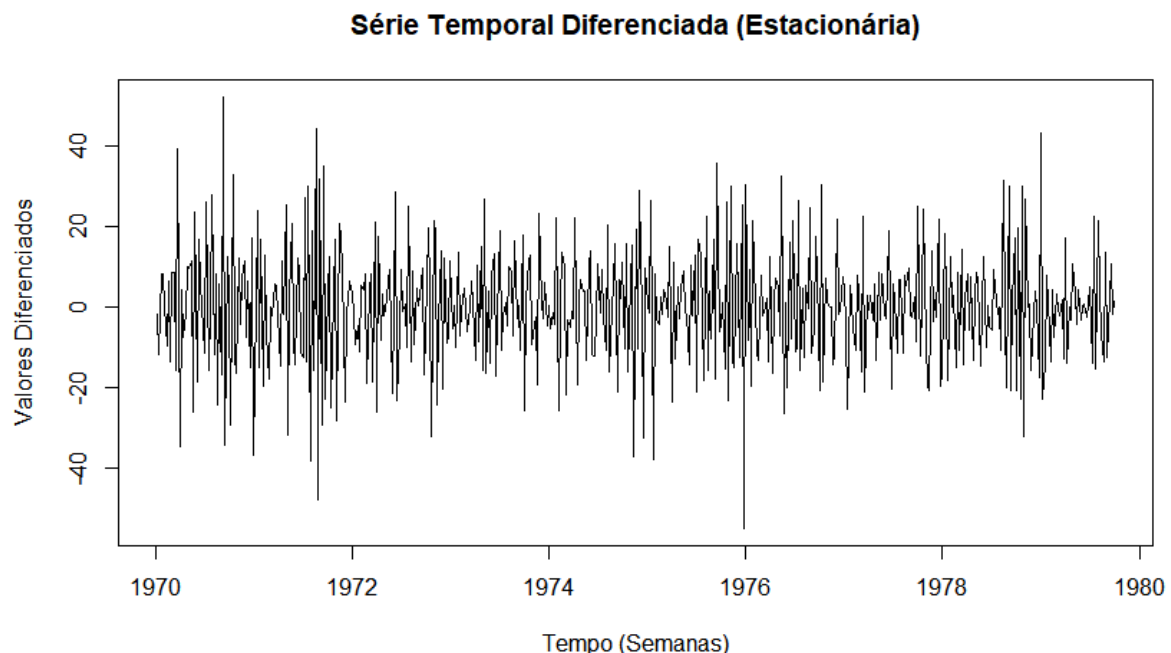
Foi realizada uma análise visual para avaliar se a variância da série era constante. Apesar de ligeiramente alta, a variância apresentou-se constante ao longo do tempo, indicando que não seria necessário aplicar uma transformação logarítmica.

A estacionaridade da série foi avaliada utilizando os testes de ADF e KPSS, com os seguintes resultados:

- ADF:
 - p-value = 0.01097 (<0.05).
 - Conclusão: Rejeita-se a hipótese nula pois há evidência estatística que indica que a série seja estacionária.
- KPSS:
 - p-value = 0.02172 (<0.05).
 - Conclusão: Rejeita-se a hipótese nula pois há evidência estatística que indica que a série não seja estacionária.

Os testes de hipóteses fornecem evidências conflitantes sobre a estacionariedade da série, isso pode ser devido a existência de uma tendência a longo prazo ou devido a sazonalidade existente. Logo foi aplicada uma diferenciação, ordinária e sazonal, para avaliar o melhor modelo SARIMA. Dado que a função `nsdiffs()` não indicou a necessidade de uma diferenciação sazonal, e a `ndiffs()` indicou a necessidade de uma diferenciação ordinária, a mesma foi aplicada e refeitos os testes ADF e KPSS para confirmar a estacionariedade da série temporal.

- ADF após a diferenciação:
 - pp-valor < 0.01 .
 - Conclusão: Série estacionária.
- KPSS após a diferenciação:
 - pp-valor > 0.1 .
 - Conclusão: Série estacionária.



Após a diferenciação, as funções ACF e PACF foram utilizadas para identificar os possíveis valores de pp , qq , PP , e QQ :

- ACF (para MA e SMA):
 - Lags significativos: 1, 2, 3, 17, 40, 41, 42, 43, 80, 81, 110, 146, 159, 167, 168, 1731, 2, 3, 17, 40, 41, 42, 43, 80, 81, 110, 146, 159, 167, 168, 173.
- PACF (para AR e SAR):
 - Lags significativos: 1, 2, 3, 4, 5, 10, 17, 22, 30, 33, 39, 41, 79, 1571, 2, 3, 4, 5, 10, 17, 22, 30, 33, 39, 41, 79, 157.

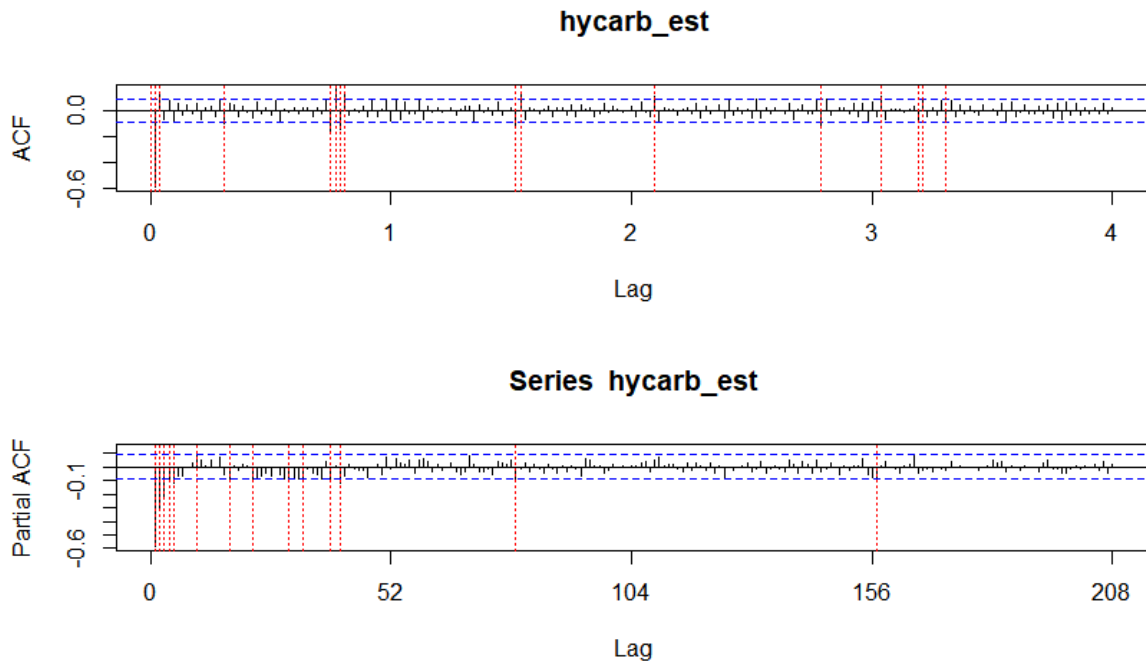


Imagem 12: ACF e PACF série temporal Hycarb

Possíveis Componentes $AR(p) = 0,1,2,3,4,5$.

Dada a avaliação dos lags significativos temos que os valores mencionados acima sobre a avaliação da Pacf representam possíveis valores para se chegar a um modelo SARIMA, demais lags significativos apresentam apenas uma repetição sazonal referente ao modelo.

Possíveis Componentes $MA(q) = 0,1,2$.

Dada a avaliação dos lags significativos temos que os valores mencionados acima sobre a avaliação da Acf representam possíveis valores para se chegar a um modelo SARIMA, demais lags significativos apresentam apenas uma repetição sazonal referente ao modelo.

Possíveis Componentes $SAR(P) = 1,2$.

Aparentemente há possíveis padrões de repetição ao longo dos lags pela avaliação da Pacf, possivelmente a cada 38 lags, o que seria aproximadamente a cada 3 trimestres.

Possíveis Componentes $MA(Q) = 0$.

Aparentemente não há padrões claros de repetição ao longo dos lags pela avaliação da Acf.

Um total de 36 modelos SARIMA foram ajustados considerando todas as combinações possíveis de parâmetros $p, d, q, P, D, Q, p, d, q, P, D, Q$. Para cada modelo, os critérios de informação AIC, BIC, e AICc foram calculados. Os três melhores modelos com base nesses critérios foram:

	AIC	BIC	AIC _c
ARIMA(5,1,2)(2,0,0)[52]	3897.46	3935.17	3897.46
ARIMA(5,1,2)(1,0,0)[52]	3898.55	3932.11	3898.55
ARIMA(0,1,2)(2,0,0)[52]	3903.22	3928.43	3903.22

Tabela 2: Modelos SARIMA comparativos

O modelo ARIMA(5,1,2)(2,0,0)[52] foi selecionado como o melhor devido aos menores valores de AIC, AICc e BIC.

O modelo escolhido foi validado por seus resíduos, no qual, o gráfico da ACF dos resíduos não apresentou autocorrelação significativa, sugerindo que o modelo capturou adequadamente as dependências na série, e validado pelo teste Ljung-box, que apresentou p-value de 0.9996 ($>\alpha=0.05$), indicando que não há evidência de autocorrelação nos resíduos.

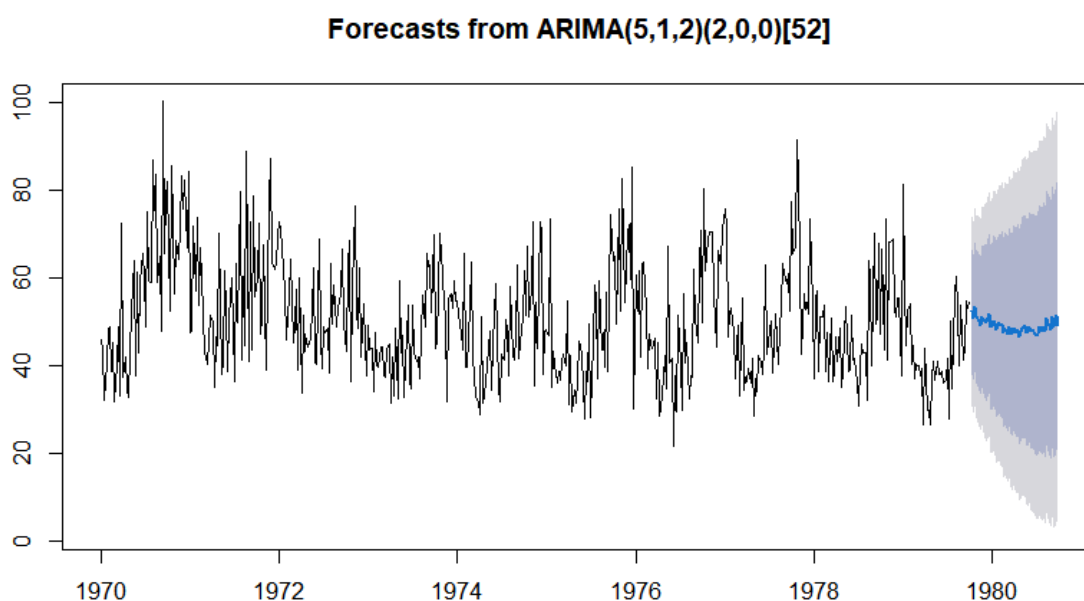


Imagem 13: Modelo SARIMA com previsão para 12 períodos de tempo

3.4. Comparação de Modelos e Conclusões

Finalizado a análise comparando os modelos Holt-Winters e SARIMA em termos de sua capacidade de previsão e ajuste aos dados uma vez que os critérios de seleção incluiu análise de resíduos, testes de autocorrelação residual, e as métricas de erro como RMSE e MAPE foi possível chegar a algumas conclusões.

A comparação é ilustrada com gráficos que destacam as previsões de cada modelo, permitindo uma visualização clara de qual modelo proporciona uma melhor adaptação e previsões mais precisas.

Conclui-se qual modelo oferece a melhor interpretação e previsão para os dados de concentração de hidrocarbonetos, fornecendo recomendações baseadas em nossa análise detalhada.

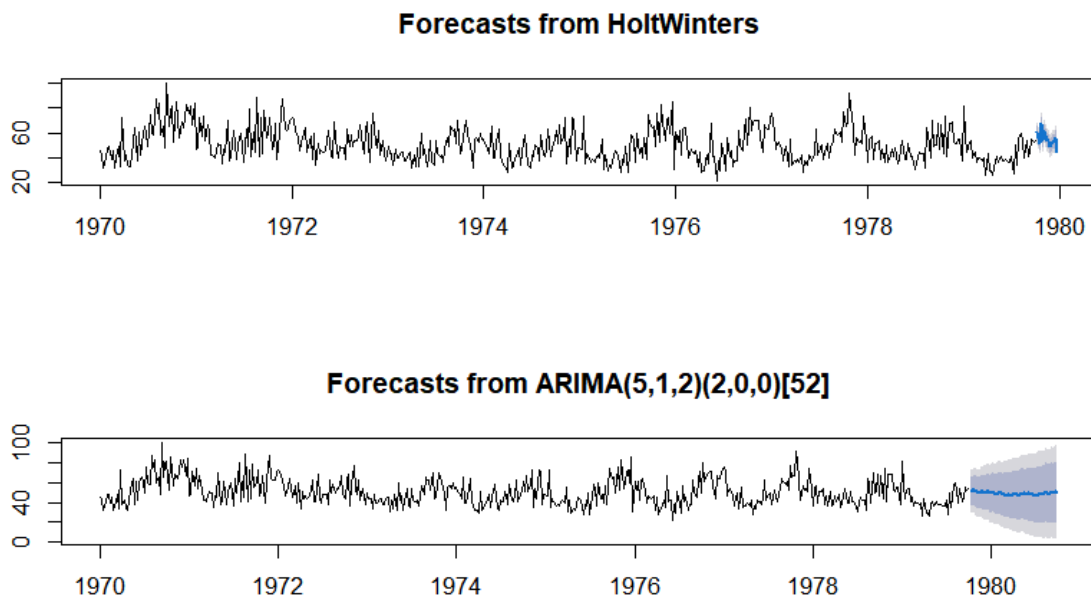


Imagem 14: Comparativo previsões HoltWinters e SARIMA

4. Análise dos Retornos Diários de Fechamento das Ações da Ford

4.1. Retornos Diários Ajustados

4.1.1. Inicialização dos Dados e Visualização Preliminar

A análise foi realizada utilizando os dados históricos das ações da Ford (*ticker: F*), coletados no período entre 3 de janeiro de 2007 e 31 de dezembro de 2023. Para a análise de volatilidade, foi utilizado o preço ajustado de fechamento (*Adjusted Close*), que considera dividendos e splits. A série de retornos diários foi calculada como a variação percentual dos valores ajustados entre dois períodos consecutivos. Principais características dos dados:

- A série apresenta flutuações esperadas nos preços ao longo do tempo, refletindo os movimentos diários do mercado.
- As estatísticas descritivas indicam a amplitude dos valores ajustados, com o menor preço registrado de 0.6961 e o maior de 20.6325.

Para avaliar a série temporal, foi plotado o gráfico dos retornos diários. Essa visualização permitiu identificar picos de volatilidade em determinados períodos, evidenciando variações anômalas nos preços.

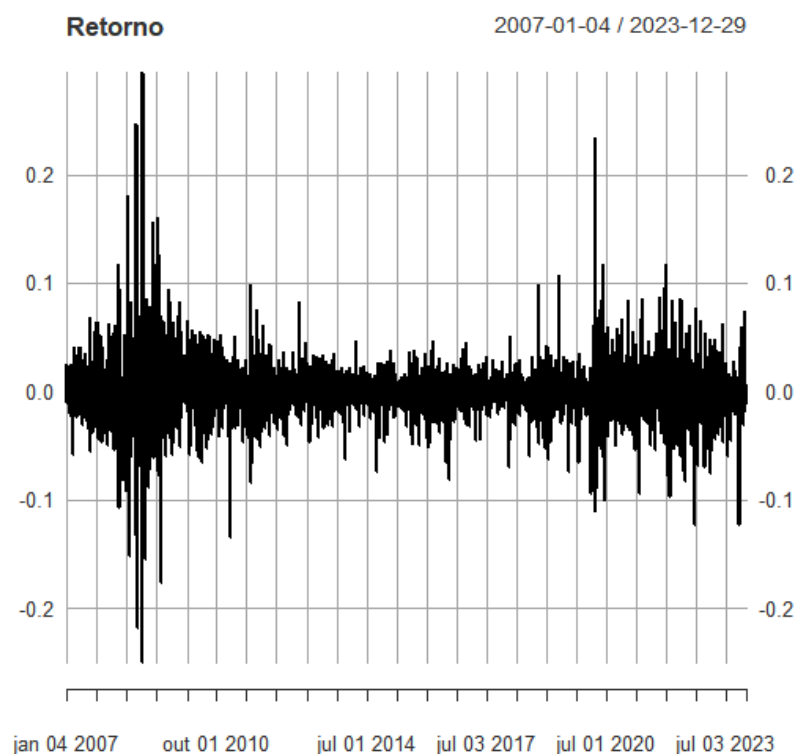


Imagem 15: Retornos diários ajustados ações Ford

4.1.2. Especificação do Modelo GARCH

Foi aplicado o teste ARCH LM para verificar a presença de heterocedasticidade, ou seja, se a variância condicional dos retornos muda ao longo do tempo.

- Resultados do teste ARCH LM:
 - Estatística Qui-quadrado: 4210.5
 - Graus de liberdade: 1
 - p-value: $< 2.2e-16$

O p-value extremamente pequeno levou à rejeição da hipótese nula de ausência de efeitos ARCH. Isso confirma a presença de heterocedasticidade nos retornos, justificando o uso de modelos GARCH para capturar a dinâmica da volatilidade condicional.

O primeiro modelo ajustado foi um GARCH(1,1), que captura a volatilidade condicional com base em choques recentes (α_1) e na persistência da volatilidade (β_1). Este modelo assume distribuição normal para os resíduos.

- Parâmetros estimados:
 - Variância incondicional (ω): 0.000006
 - Choques recentes (α_1): 0.0588
 - Persistência da volatilidade (β_1): 0.9321
- Log-verossimilhança: 10278.89
- Critérios de informação:
 - AIC: -4.8273
 - BIC: -4.8213

O valor elevado de β_1 indica que a volatilidade é altamente persistente, ou seja, choques de volatilidade têm efeito prolongado ao longo do tempo.

4.1.3. Ajuste e Comparação de Modelos

Para avaliar a adequação do modelo, foram analisados os resíduos padronizados:

- Histograma dos Resíduos Padronizados: O gráfico mostrou uma distribuição aproximadamente simétrica, com pequenas discrepâncias em relação à normalidade.
- QQ-Plot dos Resíduos Padronizados: Evidenciou alinhamento com a distribuição teórica normal, exceto por pequenos desvios nas caudas.

Os resíduos não apresentaram padrões significativos de autocorrelação, validando a capacidade do modelo de capturar a dinâmica da volatilidade.

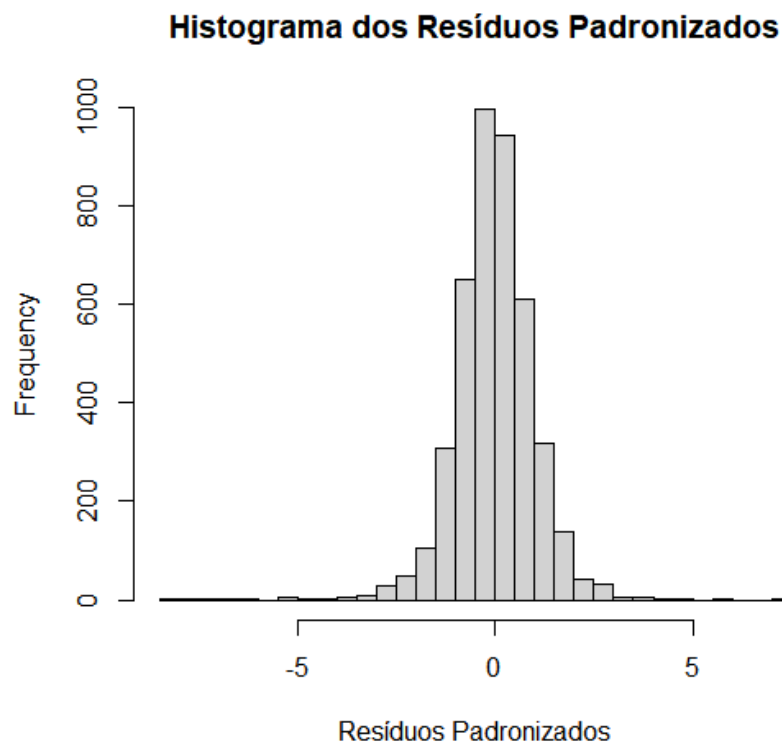


Imagem 16: Histograma dos Resíduos Padronizados

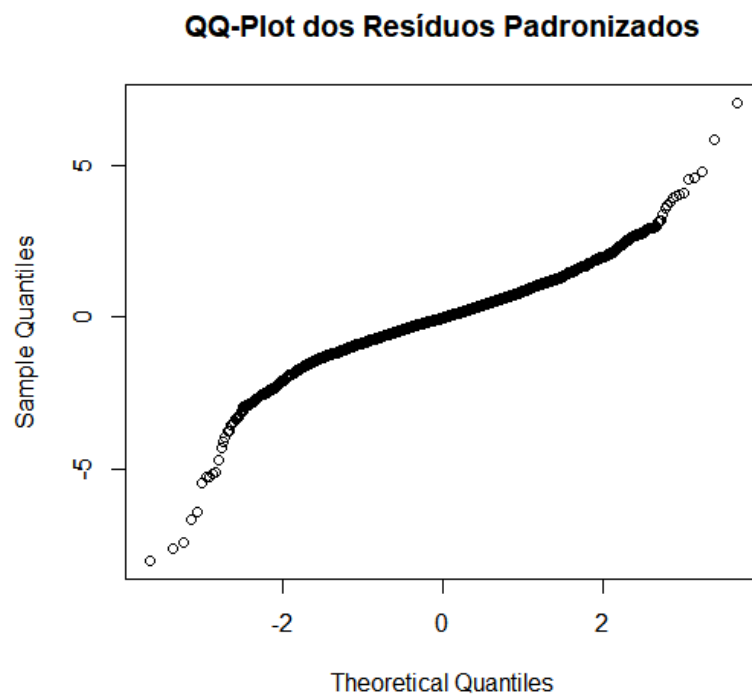


Imagem 17: QQ-Plot dos Resíduos Padronizados

O modelo foi utilizado para prever a volatilidade condicional em um horizonte de 20 períodos (equivalente a 20 dias úteis). As previsões geradas incluem:

Série ajustada média (forecast): Representa o comportamento esperado dos retornos ajustados.

Volatilidade condicional (σ): Indicadores de incerteza para cada período futuro.

No primeiro dia (T+1), a previsão média é de 0.0004766, com volatilidade condicional de 0.02583. Ao final do horizonte (T+20), a previsão média se mantém estável, com leve aumento na volatilidade condicional para 0.02622.

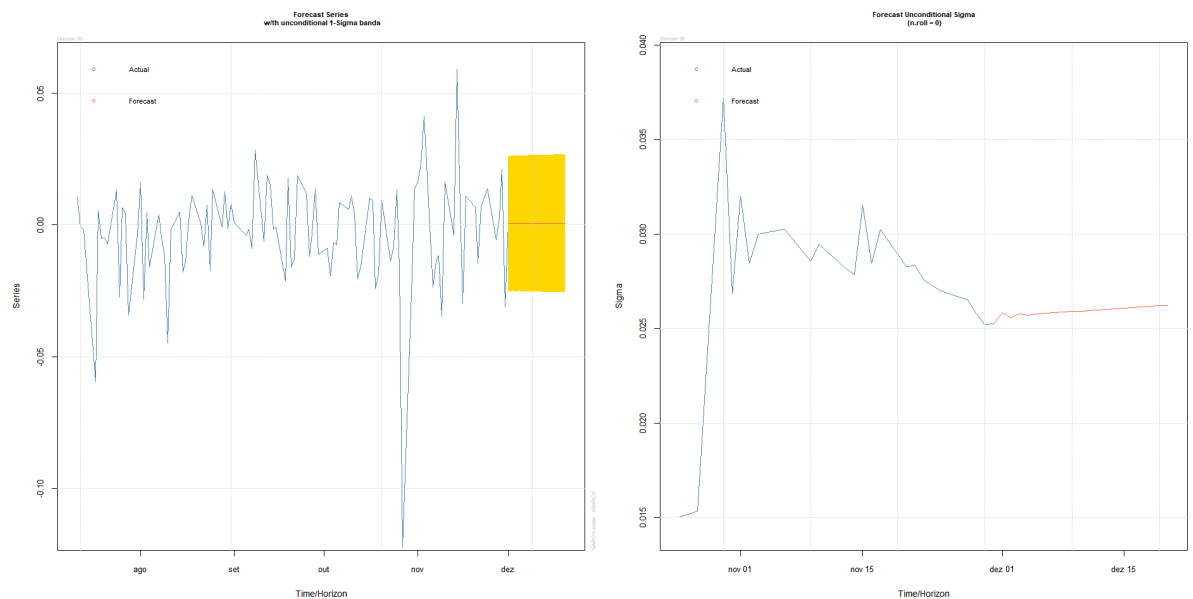


Imagem 18: Gráfico das Previsões do Modelo GARCH

4.1.4. Interpretação dos Resultados e Previsões

O modelo GARCH(1,1) demonstrou ser adequado para modelar a volatilidade condicional dos retornos diários ajustados das ações da Ford. As análises de resíduos indicaram que o modelo capturou bem a dinâmica dos dados, e as previsões forneceram estimativas confiáveis para um horizonte de 20 períodos. Este modelo será utilizado como base para comparações com outras especificações GARCH.

4.2.Retornos Diários Fechamento

4.2.1. Inicialização dos Dados e Visualização Preliminar

Iniciou-se o processo de análise obtendo os dados das ações da Ford, especificamente os preços de fechamento diários ajustados, de 1 de janeiro de 2007 até 31 de dezembro de 2023, usando a função `getSymbols`. Após carregar os dados, foi extraída a série de preços de fechamento (`F$F.Close`) e visualizada a série temporal para obter uma percepção inicial da variação dos preços ao longo do tempo. A série é então transformada em retornos logarítmicos diários para normalizar os dados e facilitar a análise subsequente. Estes retornos são limpos para remover quaisquer valores faltantes, preparando-os para a modelagem de volatilidade.

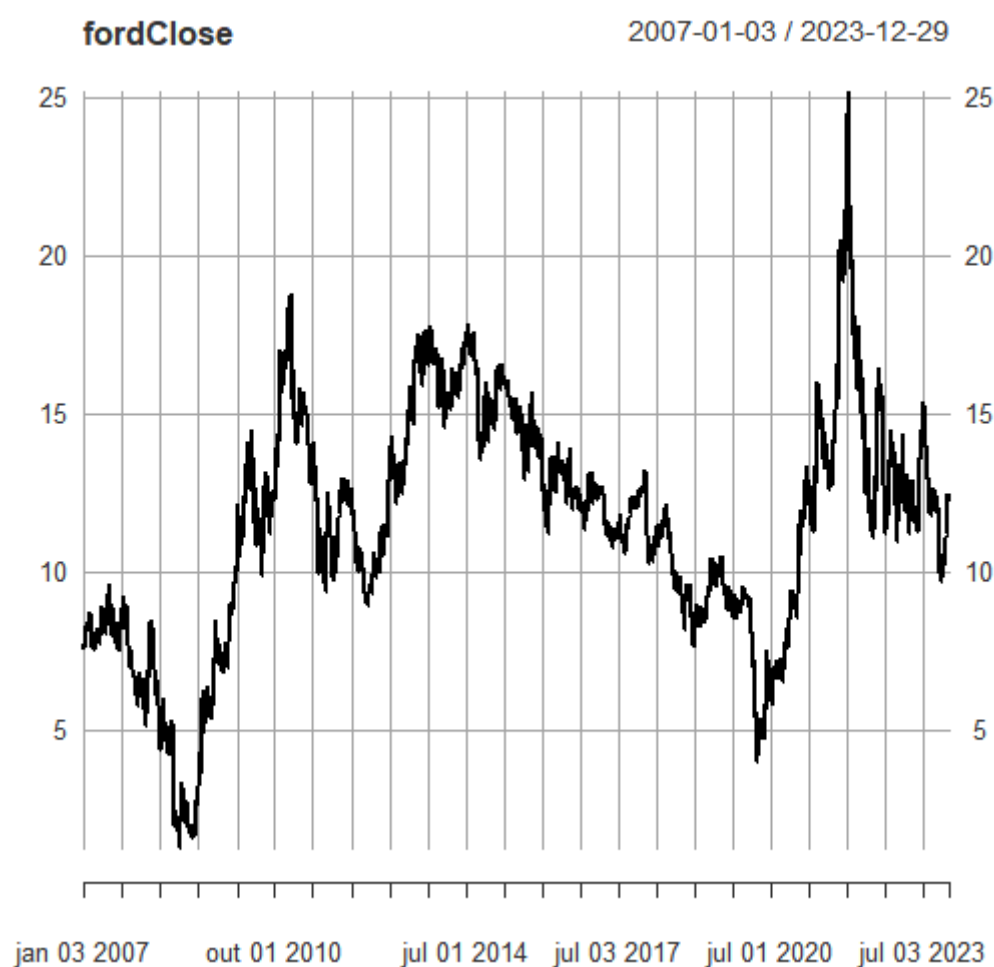


Imagem 19: Fechamentos diários ações Ford

4.2.2. Especificação de Modelos EGARCH e GJR-GARCH

Diversos modelos GARCH são especificados para capturar a volatilidade dos retornos diários. Modelos EGARCH e GJR-GARCH de várias ordens são configurados para testar diferentes estruturas de dependência na variação temporal da volatilidade. Cada especificação inclui modelos com diferentes ordens de reação à volatilidade passada e aos choques de informação, utilizando distribuições normais e t-Student para acomodar a presença de caudas pesadas nos resíduos financeiros.

4.2.3. Ajuste e Comparação de Modelos

Os modelos especificados são ajustados aos retornos logarítmicos diários limpos, e os resultados são avaliados com base nos critérios de informação de Akaike (AIC) e Bayesiano (BIC). Uma tabela é criada para comparar estes critérios entre os diferentes modelos, facilitando a seleção do modelo que melhor se ajusta aos dados em termos de equilíbrio entre ajuste e complexidade. O melhor modelo é selecionado com base no menor AIC, indicando um bom ajuste com penalidade mínima por excesso de parâmetros.

	AIC	BIC
EGARCH(1,1)	-4.949.887	-4.940.963
EGARCH(1,2)	-4.950.575	-4.940.164
EGARCH(2,1)	-4.952.237	-4.940.339
GJR-GARCH(1,1)	-4.943.180	-4.934.257
GJR-GARCH(1,2)	-4.943.972	-4.933.562
GJR-GARCH(2,1)	-4.942.918	-4.931.020

Tabela 3: Comparativos AIC e BIC modelos

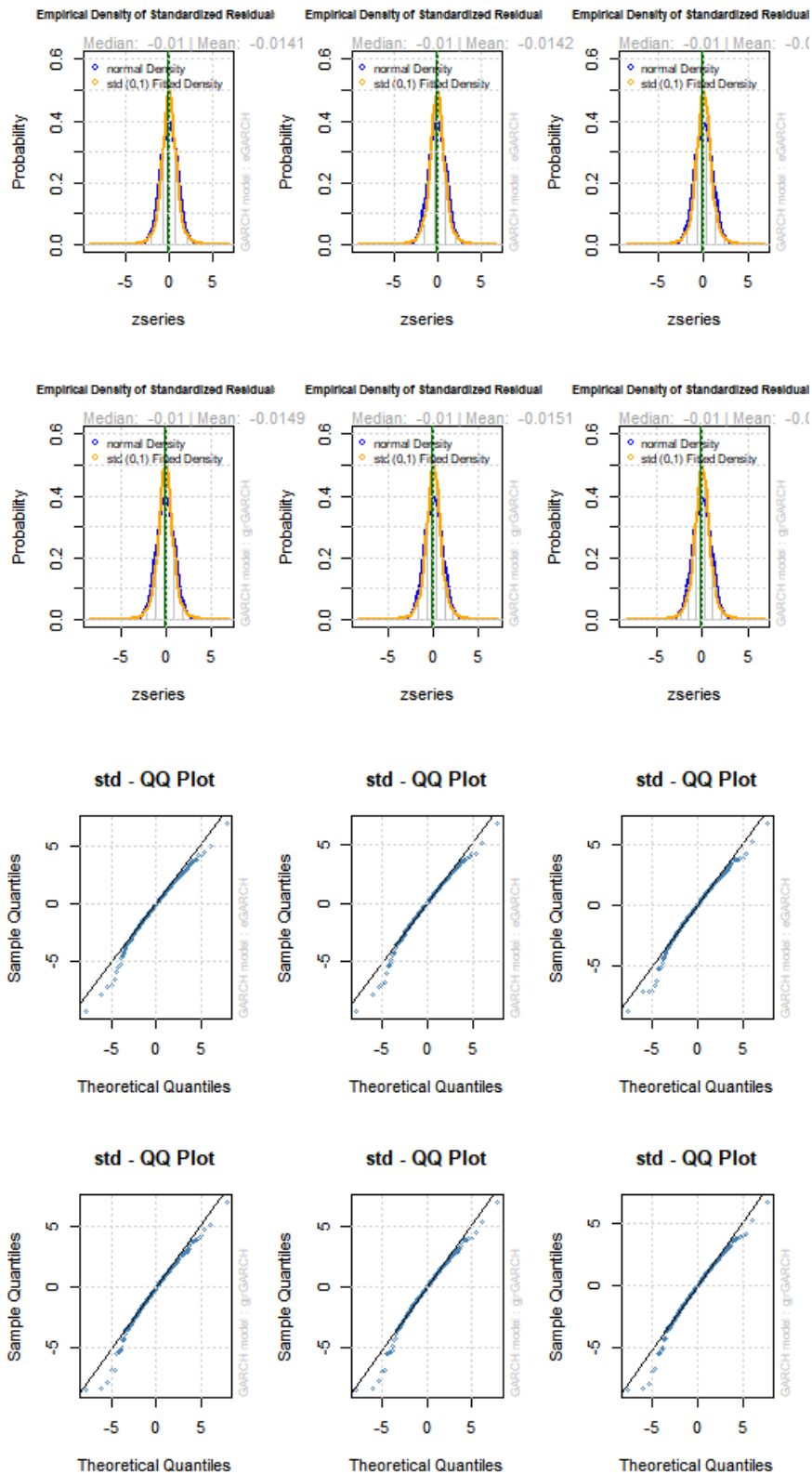


Imagem 20 e 21: Comparativo entre os 6 modelos

4.2.4. Interpretação dos Resultados e Previsões

Após a seleção do modelo, foi feito um resumo estatístico detalhado para entender melhor os parâmetros e a performance do modelo escolhido. Os resíduos do modelo são examinados através de histogramas e QQ-plots para avaliar a normalidade e identificar qualquer comportamento atípico.

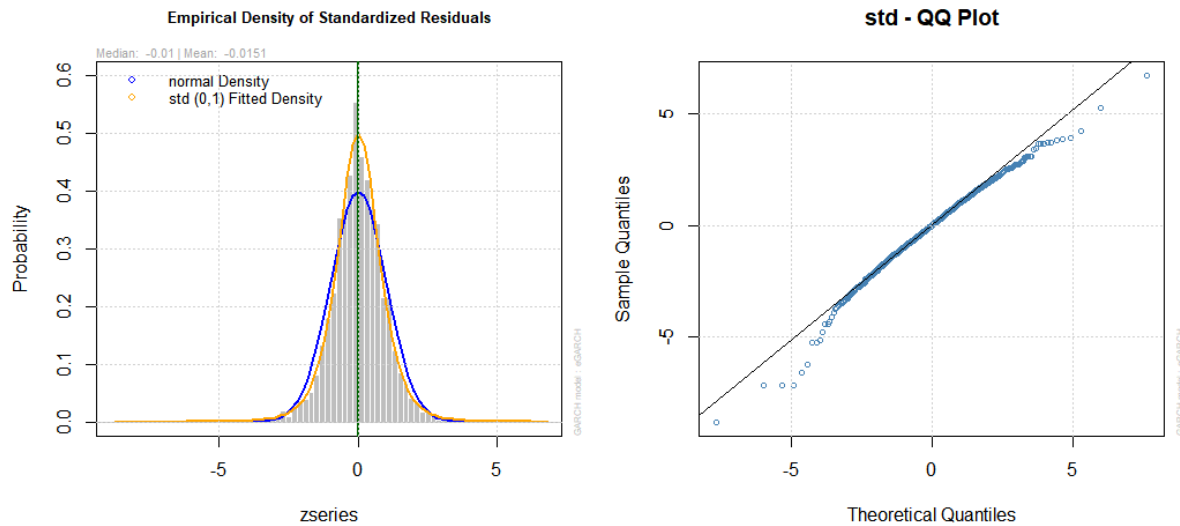


Imagem 21: Modelo eGARCH(2,1)