

Uma abordagem para previsão da taxa de câmbio por meio de modelos de séries temporais

Fernando Muller Banzato
Dissertação de Mestrado apresentada à
Faculdade de Ciências da Universidade do Porto em Matemática
2021

MSc

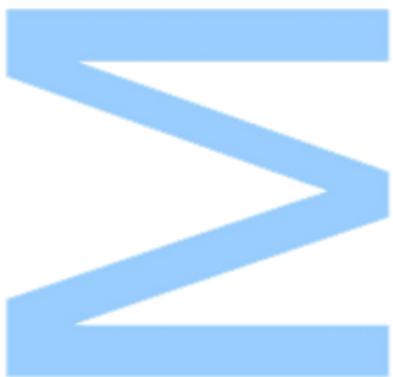
2.º
CICLO

FCUP
ANO



Uma abordagem para previsão da taxa de câmbio por meio de modelos de séries temporais

Fernando Muller Banzato



Uma abordagem para previsão da taxa de câmbio por meio de modelos de séries temporais

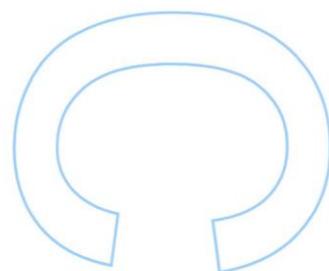
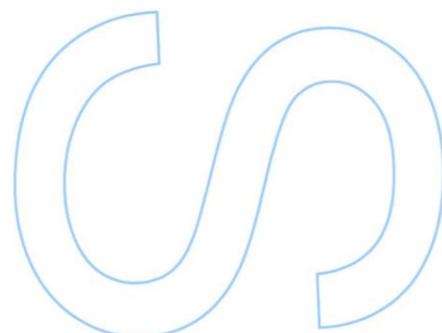
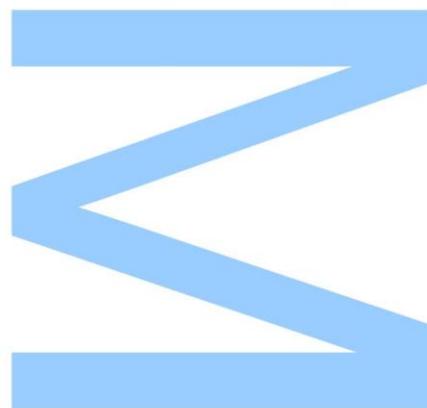
Fernando Muller Banzato

Mestrado em Matemática

Departamento de Matemática, Faculdade de Ciências, Universidade do
Porto
2021

Orientador

Alberto Adrego Pinto, Professor Catedrático, Faculdade de Ciências,
Universidade do Porto

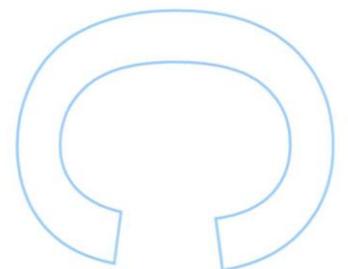
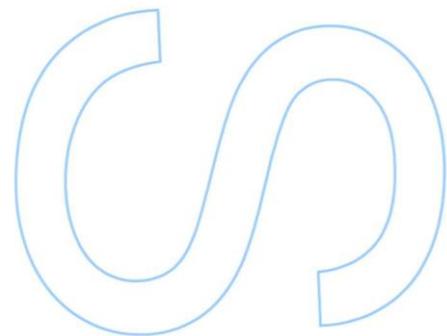
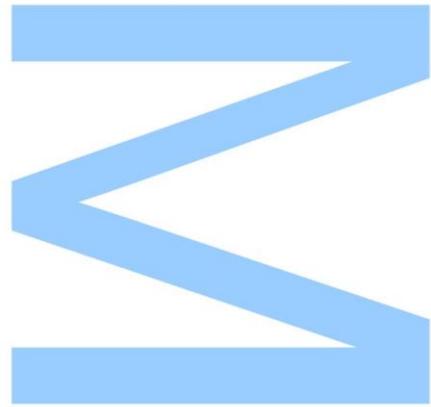




Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, ____ / ____ / ____



Agradecimentos

¹ Primeiramente, queria agradecer ao professor Alberto Adrego Pinto pela orientação desta dissertação, por sua disponibilidade e também por ter compartilhado o seu conhecimento e a sua experiência, possibilitando-me concluir este trabalho com o rigor exigido.

Agradeço aos meus colegas de Mestrado, e em especial: Carla, Jorge, Cláudio e Pedro, pela solidariedade e apoio constantes.

Expresso a minha gratidão aos meus queridos pais, Marlene e Cláudio, que desempenharam um papel determinante por sempre me apoiarem e por priorizarem a educação, ensinando-me a encará-la como um dos pilares fundamentais da vida. Outrossim, agradeço aos meus irmãos: Cláudio, Renata e João, pelo constante incentivo e ajuda.

Registro também um agradecimento especial à minha amada esposa, Xênia, que além de sempre estar ao meu lado, encorajando-me a aceitar este desafio, também ajudou-me a encontrar a motivação adequada. Por último e da mesma forma, agradeço ao meu querido filho, Lucas, pela paciência, contínuas revisões de texto deste relatório e pela companhia nesta longa jornada em Portugal.

¹**Nota sobre a ortografia adotada:** relatório de dissertação escrito na variante brasileira da língua portuguesa.

Resumo

Em tempos atuais de economia global, as taxas de câmbio desempenham um relevante papel no cenário econômico, pois impactam diretamente os negócios, dos mais variados setores da economia. Além disso, tratam-se de informações dinâmicas com características de alta instabilidade, influenciadas por fatores econômicos, políticos e sociais; esse conjunto de fatores é responsável por uma complexidade adicional ao exercício de construção de modelos preditivos. Neste tema, é enorme o desafio e concomitantemente, existem muitas oportunidades a explorar-se em termos de abordagem de modelação, nesta vasta área de conhecimento.

Este estudo tem como objetivo investigar as relações de dependência entre as taxas de câmbio envolvendo as moedas Euro (EUR), Dólar americano (USD), a moeda brasileira Real (BRL) e alguns índices do mercado acionário, de modo a explorar alguns modelos preditivos em séries temporais, que consigam explicar as variações entre as cotações diárias das taxas de câmbio relacionadas com estas moedas.

Palavras-chave: Séries Temporais, Taxa de câmbio, *Forecasting*, Previsão, GARCH, ARIMA, Regressão Dinâmica, Volatilidade

Abstract

Nowadays of global economy, the exchange rates play an important role in the economic scenario because businesses, from the most sectors of economy, is directly impacted. In addition, it is a dynamic information scenario with characteristics of high volatility, influenced by economic, political and social factors, which is why the complexity to build forecasting models is considerably increased. In this context, the challenges are big and, at the same time, there are a lot of existing opportunities in terms of the modeling approach in this wide area of knowledge.

This study has a proposal to investigate the dependency relationships between exchange rates involving the currencies Euro (EUR), US dollar (USD) and the Brazilian real (BRL) and also some stock market indexes, in order to explore some forecasting models in time series that can explain the variations over time between the daily quotations of the exchange rates involving these currencies.

Keywords: Time Series, Exchange Rate, Forecasting, GARCH, ARIMA, Dynamic Regression, Volatility

Conteúdo

1	Introdução	1
2	Conceitos Teóricos	3
2.1	Séries Temporais	3
2.1.1	Correlação	4
2.1.1.1	Função de Autocorrelação	4
2.1.1.2	Função de Autocorrelação Parcial	5
2.1.1.3	Função de Correlação Cruzada	6
2.1.2	Componentes de uma Série Temporal	6
2.1.3	Estacionariedade	7
2.1.4	Testes de hipótese	9
2.1.4.1	Ljung-Box	10
2.1.4.2	Multiplicadores de Lagrange (LM)	10
2.1.4.3	<i>Augmented Dickey Fuller</i> (ADF)	11
2.1.4.4	Kwiatkowski-Phillips-Schmidt-Shin (KPSS)	11
2.2	Conceitos Financeiros	12
2.2.1	Retorno	12
2.2.2	Volatilidade	13
2.3	Modelos	14
2.3.1	AR	14
2.3.2	MA	16
2.3.3	ARMA	16
2.3.4	ARIMA	17
2.3.5	Regressão Dinâmica	19
2.3.6	ARCH	20

2.3.7	GARCH	22
2.3.8	APARCH	24
2.4	Coeficiente de determinação	25
2.5	Critérios de Informação	26
2.5.1	AIC	26
2.5.2	AICc	27
2.5.3	BIC	27
2.6	Indicadores de erro	28
2.6.1	Medidas dependentes de escala	29
2.6.2	Medidas relativas	30
3	Aplicação e Resultados	32
3.1	Descrição dos dados	32
3.2	Leitura e preparação dos dados	33
3.3	Análise exploratória dos dados	34
3.3.1	Taxa de câmbio EUR/BRL	34
3.3.2	Taxa de câmbio EUR/USD	45
3.4	Modelação	48
3.4.1	Taxa de câmbio EUR/BRL	49
3.4.1.1	Modelo homocedástico	49
3.4.1.2	Modelo heterocedástico	57
3.4.2	Taxa de câmbio EUR/USD	69
4	Conclusões	76
4.1	Conclusões	76
4.2	Trabalho futuro	77
	Bibliografia	78
	Apêndices	81
A	Métodos de estimação	82

A.1 Mínimos quadrados	82
A.2 Máxima verossimilhança	83
B Aplicação e Resultados	88
B.1 Taxa de câmbio EUR/BRL	89
B.2 Taxa de câmbio EUR/USD	90
C Instruções do <i>R Software</i>	92
C.1 Bibliotecas	92
C.2 Funções criadas	94
C.3 Principais funções utilizadas	98
D Estrutura das tabelas de dados	100

Lista de Figuras

2.1	Exemplo de um gráfico da função de autocorrelação (ACF)	5
2.2	Exemplo de um gráfico da Função de autocorrelação parcial (PACF)	6
3.1	Taxa de câmbio EUR/BRL	34
3.2	Estatísticas descritivas da taxa de câmbio EUR/BRL	35
3.3	<i>Boxplots</i> da taxa de câmbio EUR/BRL por período.	36
3.4	ACF/PACF da taxa de câmbio EUR/BRL	36
3.5	Retorno da taxa de câmbio EUR/BRL	37
3.6	Estatísticas descritivas do retorno da taxa de câmbio EUR/BRL	38
3.7	Comparativo entre funções de densidade de probabilidade (teórica <i>versus</i> retorno da taxa de câmbio EUR/BRL).	39
3.8	ACF/PACF do retorno da taxa de câmbio EUR/BRL	40
3.9	Decomposição da taxa de câmbio EUR/BRL.	42
3.10	Decomposição do retorno da taxa de câmbio EUR/BRL.	43
3.11	Taxa de câmbio EUR/USD	45
3.12	Estatísticas descritivas do retorno da taxa de câmbio EUR/USD.	46
3.13	ACF/PACF do retorno da taxa de câmbio EUR/USD	47
3.14	Comparativo entre funções de densidade de probabilidade (teórica <i>versus</i> retorno da taxa de câmbio EUR/USD).	47
3.15	Fluxograma do processo de modelação adotado.	48
3.16	Comparativo dos resíduos entre os modelos de regressão e regressão dinâmica ajustado por ARIMA para a taxa de câmbio EUR/BRL.	53
3.17	Resíduos do modelo de regressão dinâmica para a taxa de câmbio EUR/BRL.	53

3.18 Comparação dos quantis do modelo de Regressão Dinâmica para a taxa de câmbio EUR/BRL.	54
3.19 <i>Forecast</i> da taxa de câmbio EUR/BRL obtido por modelo de Regressão Dinâmica.	55
3.20 Indicador MAPE do <i>forecast</i> , obtido por modelo de Regressão Dinâmica para a taxa de câmbio EUR/BRL.	56
3.21 ACF/PACF do retorno quadrado da taxa de câmbio EUR/BRL.	58
3.22 Retorno quadrado da taxa de câmbio EUR/BRL	58
3.23 Resultados da modelação <i>ARMA + GARCH</i> para o retorno da taxa de câmbio EUR/BRL.	61
3.24 Densidade e quantis dos resíduos padronizados da taxa de câmbio EUR/BRL, obtidos por modelo <i>ARMA(1, 0) + GARCH(1, 1)</i>	63
3.25 ACF dos resíduos padronizados e dos resíduos padronizados quadrados da taxa de câmbio EUR/BRL, obtidos por modelo <i>ARMA(1, 0) + GARCH(1, 1)</i>	64
3.26 Série de retornos e volatilidade condicional da taxa de câmbio EUR/BRL, obtidos por modelo <i>ARMA(1, 0) + GARCH(1, 1)</i>	64
3.27 <i>Rolling forecast</i> da volatilidade e retorno da taxa de câmbio EUR/BRL obtidos por modelo <i>ARMA(1, 0) + GARCH(1, 1)</i>	66
3.28 Densidade de probabilidade da volatilidade e retorno da taxa de câmbio EUR/BRL, obtido por meio de simulações do modelo <i>ARMA(1, 0) + GARCH(1, 1)</i>	67
3.29 Indicador MAPE do <i>forecast</i> , obtido por modelo <i>ARMA(1, 0) + GARCH(1, 1)</i> para o retorno da taxa de câmbio EUR/BRL.	68
3.30 ACF/PACF do retorno quadrado da taxa de câmbio EUR/USD.	69
3.31 Resultados da modelação <i>ARMA + GARCH</i> para o retorno da taxa de câmbio EUR/USD.	71
3.32 Densidade e quantis dos resíduos padronizados da taxa de câmbio EUR/USD, obtidos por modelo <i>ARMA(0, 0) + GARCH(1, 1)</i>	73
3.33 ACF dos resíduos padronizados e dos resíduos padronizados quadrados da taxa de câmbio EUR/USD, obtidos por modelo <i>ARMA(0, 0) + GARCH(1, 1)</i>	73

3.34	Série de retornos e volatilidade condicional da taxa de câmbio EUR/USD, obtidos por modelo $ARMA(0, 0) + GARCH(1, 1)$	74
3.35	Densidade estimada pelo método <i>rolling density forecast</i> , baseado em modelo $ARMA(0, 0) + GARCH(1, 1)$	75
B.1	Retorno da taxa de câmbio EUR/BRL, obtido por meio de simulações do modelo $ARMA(1, 0) + GARCH(1, 1)$	89
B.2	Volatilidade do retorno da taxa de câmbio EUR/BRL, obtido por meio de simulações do modelo $ARMA(1, 0) + GARCH(1, 1)$	89
B.3	Densidade e quantis dos resíduos padronizados da taxa de câmbio EUR/USD, obtidos por modelo $ARMA(0, 0) + GARCH(1, 1)$, sem os <i>outliers</i>	90
B.4	Volatilidade e retorno da taxa de câmbio EUR/USD, obtido pelo método <i>rolling density forecast</i>	91

Lista de Tabelas

3.1	Séries financeiras utilizadas no componente prático.	32
3.2	Correlação entre a taxa de câmbio EUR/BRL e índices de ações.	40
3.3	Correlação entre o retorno da taxa de câmbio EUR/BRL e outras séries.	41
3.4	Correlação cruzada (CCF) entre o retorno da taxa de câmbio EUR/BRL e o retorno dos índices de ações.	42
3.5	Indicador MAPE do <i>forecast</i> , obtido por modelo de Regressão Dinâmica para a taxa de câmbio EUR/BRL.	56
3.6	Parâmetros dos modelos <i>ARMA</i> + <i>GARCH</i> testados para o retorno da taxa de câmbio EUR/BRL.	59
3.7	Coefficientes estimados pelo modelo <i>ARMA</i> (1, 0) + <i>GARCH</i> (1, 1) para o retorno da taxa de câmbio EUR/BRL.	62
3.8	Relação das variáveis externas testadas como preditoras em modelo <i>GARCH</i> . . .	65
3.9	Indicador MAPE do <i>forecast</i> , obtido por modelo <i>ARMA</i> (1, 0) + <i>GARCH</i> (1, 1) para o retorno da taxa de câmbio EUR/BRL.	69
3.10	Parâmetros dos modelos <i>ARMA</i> + <i>GARCH</i> testados para o retorno da taxa de câmbio EUR/USD.	70
3.11	Coefficientes estimados pelo modelo <i>ARMA</i> (0, 0) + <i>GARCH</i> (1, 1) para o retorno da taxa de câmbio EUR/USD.	72
C.1	Funções do <i>R Software</i> utilizadas para testes de hipótese	98
C.2	Funções do <i>R Software</i> utilizadas para <i>Forecasting</i>	99
C.3	Funções do <i>R Software</i> utilizadas na modelação e <i>forecast</i> <i>GARCH</i>	99
D.1	Relação dos <i>datasets</i> utilizados.	100
D.2	Conteúdo do <i>dataset</i> (estrutura padrão).	101

Lista de Abreviaturas

<i>ACF</i>	<i>Autocorrelation Function</i>
<i>ADF</i>	<i>Augmented Dickey Fuller</i>
<i>AIC</i>	<i>Akaike Information Criterion</i>
<i>AICc</i>	<i>Akaike Information Criterion corrected</i>
<i>APARCH</i>	<i>Assimetric Power Autoregressive Conditionally Heteroscedastic</i>
<i>AR</i>	<i>Autoregressive</i>
<i>ARCH</i>	<i>Autoregressive Conditionally Heteroscedastic</i>
<i>ARIMA</i>	<i>Autoregressive Integrated Moving Average</i>
<i>ARMA</i>	<i>Autoregressive Moving Average</i>
<i>BIC</i>	<i>Bayesian Information Criterion</i>
<i>BRL</i>	Real brasileiro
<i>CCF</i>	<i>Cross Correlation Function</i>
<i>CV</i>	<i>Coeficiente de Variação</i>
<i>EGARCH</i>	<i>Exponential Generalized Autoregressive Conditionally Heteroscedastic</i>
<i>EUR</i>	Euro
<i>EURBRL</i>	Taxa de câmbio Euro (EUR) para Real brasileiro (BRL)
<i>EURUSD</i>	Taxa de câmbio Euro (EUR) para Dólar americano (USD)

<i>GARCH – M</i>	<i>Generalized Autoregressive Conditionally Heteroscedastic in mean</i>
<i>GARCH</i>	<i>Generalized Autoregressive Conditionally Heteroscedastic</i>
<i>GJR – GARCH</i>	<i>Glosten-Jagannathan-Runkle Generalized Autoregressive Conditionally Heteroscedastic</i>
<i>GLS</i>	<i>Generalized Least Squares</i>
<i>IBOVESPA</i>	Índice da Bolsa de Valores, Mercadorias e Futuros de São Paulo
<i>iid</i>	<i>independent and identically distributed</i>
<i>IQR</i>	<i>Interquartile range</i>
<i>KPSS</i>	Kwiatkowski-Phillips-Schmidt-Shin
<i>LM</i>	<i>Lagrange Multiplier</i>
<i>MA</i>	<i>Moving Average</i>
<i>MAE</i>	<i>Mean Absolute Error</i>
<i>MAPE</i>	<i>Mean Absolute Percentage Error</i>
<i>ME</i>	<i>Mean Error</i>
<i>MLE</i>	<i>Maximum Likelihood Estimation</i>
<i>MPE</i>	<i>Mean Percentage Error</i>
<i>MSE</i>	<i>Mean Squared Error</i>
<i>NASDAQ</i>	<i>National Association of Securities Dealers Automated Quotations</i>
<i>norm</i>	<i>Normal distribution</i>
<i>OLS</i>	<i>Ordinary Least Squares</i>
<i>PACF</i>	<i>Partial Autocorrelation Function</i>
<i>RSME</i>	<i>Root Mean Squared Error</i>

<i>S&P</i>	<i>Standard & Poor's</i>
<i>snorm</i>	<i>Skewed Normal distribution</i>
<i>sstd</i>	<i>Skewed Student's t distribution</i>
<i>std</i>	<i>Student's t distribution</i>
<i>STL</i>	<i>Seasonal and Trend decomposition using Loess</i>
<i>TS – GARCH</i>	<i>Taylor-Schwert Generalized Autoregressive Conditionally Heteroscedastic</i>
<i>USD</i>	Dólar americano
<i>VaR</i>	<i>Value-at-Risk</i>
<i>VIF</i>	<i>Variance Inflation Factor</i>

Capítulo 1

Introdução

Desde o período final do século XX, temos vivido a Era da Informação, também chamada de Era Digital, caracterizada por um período de acelerados avanços tecnológicos e que vem transformando fortemente a nossa sociedade nos âmbitos social, político e econômico. Os impactos sentidos na nossa vida cotidiana são muito evidentes pois agora estamos conectados à rede de computadores internet cem por cento de nosso tempo, cercado de informações por todos os lados e atualizadas de forma quase instantâneas. Essa demanda frenética por informações pode ser observada pelo recente fenômeno *Big Data* que descreve um conceito de grande volume de dados, estruturados ou não, gerados a cada fração mínima de tempo e que foi adotada como prática corporativa. Muitas empresas passaram a armazenar dados de sua operação, das mais variadas naturezas e às vezes, sem ainda sequer conhecer o motivo para armazená-las, apostando-se na ideia de que esses dados oferecem importante potencial para serem futuramente tratados, analisados e convertidos em informações, produzindo-se relevante conhecimento para o Negócio. Embora seja uma prática recente e não reste dúvidas quanto à sua viabilidade, é fato que pode-se agregar valor ao seu Negócio a partir desta iniciativa de análise de dados, o que abre um futuro bastante promissor em termos de oportunidade profissional para matemáticos, estatísticos, cientistas de dados, de computadores e afins.

Neste contexto de modelação de dados, *Forecasting* foi o tema escolhido desta dissertação por motivação pessoal e profissional, tanto pelo interesse pela fundamentação matemática dos modelos quanto pelas aplicações práticas reais, uma vez que já vivenciei trabalho com esta responsabilidade, nomeadamente em planejamento de demanda (em *Supply Chain*) na indústria e empresa de consultoria por vários anos. Por estas razões, tenho a clara percepção de sua

importância e da complexidade envolvida, além das amplas possibilidades de aplicação em diversos segmentos de mercado e também pelo fato de se apresentar como um grande desafio de Negócio, no qual uma parte das empresas dependem para a sua sobrevivência.

Além da abordagem percorrida com os principais fundamentos matemáticos englobados de alguns modelos preditivos de séries temporais, tais como ARIMA, Regressão Dinâmica, GARCH, entre outros, escolheu-se como dados, as séries financeiras contendo a taxa de câmbio entre as moedas Euro e Real do Brasil e a taxa de câmbio entre Euro e Dólar americano, além dos índices de ações *Dow Jones Industrial Average*, *Standard and Poor's 500*, *NASDAQ Composite*, *ESTX 50 PR.EUR*, *EURONEXT 100* e *IBOVESPA*, como um experimento de *Forecasting*, constituindo o componente prático deste trabalho. Tradicionalmente, os dados de ativos financeiros possuem um padrão de alta volatilidade, apresentam um comportamento instável ao longo do tempo, são sensíveis às mudanças dos cenários políticos e econômicos e sociais, e estas características elevam a complexidade da modelação. Explora-se nessa aplicação, alguns modelos mais simples que em princípio não são apropriados para este contexto se utilizado em sua forma individual, a exemplo do modelo ARIMA cuja premissa é de variância constante ao longo do tempo, por meio de modelo híbrido, combinando-se diferentes modelos, de modo que se reúna as suas forças individuais resultando um modelo final mais robusto, até modelos mais sofisticados como o GARCH que realiza uma estimativa da variância condicional e que de acordo com Shumway e Stoffer (2017) [20], pertencente à classe dos modelos condicionalmente não estocásticos, uma vez que sua variância é estimada de forma determinística.

Nos modelos determinísticos consegue-se determinar exatamente o resultado que o modelo produz diante das condições sob as quais um experimento foi realizado, ou seja, as estimativas obtidas baseiam-se exclusivamente nos dados observados no passado. Já os componentes estocásticos, estão vinculados com uma determinada distribuição de probabilidade.

Este estudo trata especificamente de ativos financeiros, com enfoque em taxas de câmbio, que tradicionalmente, não exhibe um determinado padrão de comportamento. O seu objetivo central consiste em investigar e desvendar esta difícil e não aparente relação de dependência estatística.

Capítulo 2

Conceitos Teóricos

2.1 Séries Temporais

Uma coleção de variáveis aleatórias, X_t , indexados por uma variável t que denota o instante, caracteriza um processo estocástico, também conhecido por *random process*. O termo série temporal refere-se à dados observados em um determinado espaço de tempo após a realização de um processo estocástico. Usualmente, esta série é representada por x_t , onde t é o parâmetro temporal dos instantes observados. Ao longo desta dissertação, ambos os termos serão empregados conceitualmente como sinônimos.

As séries temporais x_t , onde $t \in T$, podem ser discretas, quando as observações são realizadas em um instante t pertencente a um conjunto discreto, a exemplo de $T = \{1, 2, 3, 4, 5\}$, que é exatamente o caso deste estudo; ou contínuas, quando as observações são obtidas continuamente a partir de instantes t delimitados em um intervalo de tempo, ou seja, quando T é um intervalo, tal qual no exemplo de $T = [0, 1]$. As séries temporais também são designadas regulares quando o espaçamento de tempo entre as observações é fixo e irregulares, em situação contrária.

De acordo com Shumway e Stoffer (2017) [20], a aplicação de técnicas convencionais de análise estatística a esse conjunto de dados é limitada pela introdução de uma relação de dependência pelos dados da amostra com o tempo, objeto fundamental de qualquer abordagem de Análise de Séries Temporais.

2.1.1 Correlação

2.1.1.1 Função de Autocorrelação

Para medir o grau de dependência linear entre duas variáveis aleatórias, X e Y , utiliza-se uma medida estatística, designada coeficiente de correlação, definida por:

$$\rho_{x,y} = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}} = \frac{E[(X - \mu_x)(Y - \mu_y)]}{\sqrt{E[(X - \mu_x)^2(Y - \mu_y)^2]}}$$

assumindo que as variâncias de X e Y , referidas respectivamente por $Var(X)$ e $Var(Y)$, existam e que μ_x e μ_y correspondem às suas respectivas médias. Este coeficiente de correlação assume valores no intervalo $-1 \leq \rho_{x,y} \leq 1$, sendo que se $\rho_{x,y} = 0$, então diz-se que as variáveis X e Y são não correlacionadas. Assim, quanto mais próximo de 1 (em valor absoluto) o coeficiente de correlação estiver, mais forte será o grau de correlação.

Estende-se este mesmo conceito para o universo das séries temporais através de uma função, denominada função de autocorrelação (ACF), de uma única série temporal x_t e com um argumento δ , que representa a defasagem (*lag*) no tempo, e que determina o coeficiente de correlação entre os seus valores no instante t e $t + \delta$, ou seja, mede-se o grau de dependência entre os pares de informação que distam δ unidades temporais. Assim, nas mesmas condições supracitadas para a correlação, define-se a autocorrelação *lag* δ de x_t por

$$\rho(\delta) = \rho(t, t + \delta) = \frac{Cov(x_t, x_{t+\delta})}{\sqrt{Var(x_t)Var(x_{t+\delta})}}$$

onde $Cov(x_t, x_{t+\delta})$ representa a covariância entre x_t e $x_{t+\delta}$; $Var(x_t)$, refere-se à variância de x_t e $Var(x_{t+\delta})$, a variância de $x_{t+\delta}$.

A análise da autocorrelação desempenha papel relevante na avaliação de modelos preditivos sobretudo na análise de resíduos do modelo com o objetivo de inspecionar se os mesmos exibem um padrão de independência, isto é, se as relações de dependência são explicadas pelo modelo sem a colaboração dos resíduos (cenário ideal). Desta forma, os resíduos representam a porção não explicada pelo modelo (componente aleatório).

Ilustra-se com a figura 2.1 um exemplo de um gráfico da função de autocorrelação, também conhecido por correlograma, para um conjunto de dados ¹, sendo que a linha tracejada, desta-

¹temperatura média anual, medida em graus Fahrenheit, registrada na cidade americana de *New Haven, Connecticut*, no período

cada em azul, representa os limites do intervalo de confiança (ao nível de 95%), isto é, quando o valor da autocorrelação, para uma determinada defasagem (*lag*), está fora da faixa de tolerância (intervalo de confiança), diz-se que este apresenta autocorrelação estatisticamente significativa.

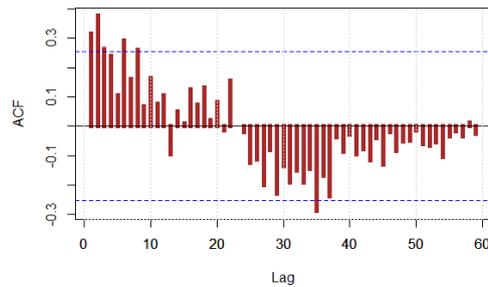


Figura 2.1: Exemplo de um gráfico da função de autocorrelação (ACF)

2.1.1.2 Função de Autocorrelação Parcial

A função de autocorrelação parcial (PACF) é definida com uma função que determina o coeficiente de correlação entre x_t e $x_{t-\delta}$ depois de remover o efeito de $x_{t-1}, x_{t-2}, x_{t-3}, \dots, x_{t-\delta+1}$, ou seja, uma função que indica a medida de correlação após o ajuste para a presença de todos os outros termos de menor defasagem (uma parcela da ACF). Como um exemplo, pode-se dizer que autocorrelação parcial para a defasagem $\delta = 4$ refere-se à parte da medida de autocorrelação sem computar os efeitos provocados pelas defasagens $\delta = 1, \delta = 2$ e $\delta = 3$.

Usualmente analisada em conjunto com a autocorrelação, a autocorrelação parcial é uma importante ferramenta que ajuda a direcionar a escolha da ordem de um modelo autoregressivo (AR) e que possui uma representação gráfica análoga à ACF. Também inclui um intervalo de confiança que determina para quais defasagens PACF assume valores estatisticamente significativos, e é exemplificado pela figura 2.2, cujos dados referem-se à mesma série temporal utilizada anteriormente no exemplo ACF.

de 1912 a 1971 e disponível através do *dataset nhtemp* no *R Software*.

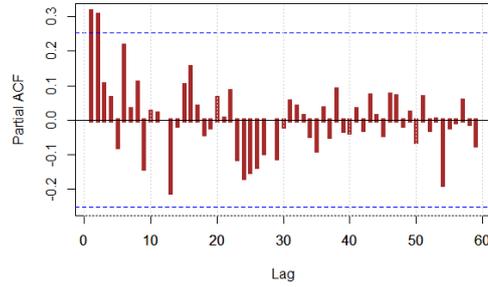


Figura 2.2: Exemplo de um gráfico da Função de autocorrelação parcial (PACF)

2.1.1.3 Função de Correlação Cruzada

A função de correlação cruzada (CCF) determina uma medida de correlação entre duas séries, porém não sincronizada em seus instantes. De acordo com Shumway e Stoffer (2017) [20], a CCF entre duas séries x_t e y_t é definida por:

$$\rho_{xy}(s, t) = \frac{Cov_{xy}(s, t)}{\sqrt{Var_x(s)Var_y(t)}},$$

onde $Cov_{xy}(s, t)$ denota a covariância entre x_t e y_t , Var_x e Var_y representam, respectivamente, a variância de x_t e y_t ; e s e t , os instantes de referência para a medição. De forma similar à ACF, $-1 \leq \rho_{xy}(s, t) \leq 1, \forall s, t$.

2.1.2 Componentes de uma Série Temporal

Em geral, uma série temporal exibe uma variedade de padrões resultante da combinação de quatro principais componentes, na qual é usualmente caracterizada pelos mesmos. Esses componentes podem ser identificados ao decompô-la através dos dados observados, e são descritos por:

- **Tendência (T_t):**

refere-se a um padrão de elevação ou declínio, não necessariamente linear, observado por um longo período de tempo;

- **Sazonalidade (S_t):**

padrão observado quando a série temporal é afetada por fatores sazonais, que

repete-se a uma periodicidade fixa e conhecida;

- **Ciclo (C_t):**

caracterizada por padrões de elevação e declínio (flutuações) que se repetem ao longo de tempo em uma frequência não fixa;

- **Resíduo (R_t):**

é o componente residual, ou seja, refere-se à parte de que não se adéqua a nenhum padrão particular. Pode ser referido por padrão aleatório, e não é explorado por nenhuma técnica estatística de análise de séries temporais.

É frequente agregar os componentes tendência e ciclo em um único (tendência-ciclo) e referenciá-lo por apenas tendência por simplicidade [13].

Considerando x_t uma série temporal, pode-se descrever os dois tipos de modelos mais empregados em séries temporais, que consideram o efeitos desses componentes da seguinte forma:

- **Modelo aditivo:** $x_t = T_t + S_t + C_t + R_t$,

onde assume-se a independência entre quaisquer componentes.

- **Modelo multiplicativo:** $x_t = T_t \times S_t \times C_t \times R_t$,

onde um componente pode afetar um outro componente, ou seja, pode existir interação entre os componentes.

Dessa forma, a grande diferença entre eles reside na premissa assumida pelos modelos aditivos de independência entre os componentes.

2.1.3 Estacionariedade

Em análise de séries temporais, utiliza-se frequentemente técnicas de inferência estatística baseada nos dados observados, e como condição para aplicá-las é necessário a existência de características de estacionariedade, ou seja, é desejável que a série apresente algumas propriedades estatísticas, a exemplo dos dois primeiros momentos (média e variância), manterem-se invariantes no tempo. Assim, uma série temporal estacionária é aquela cujas propriedades não dependem do tempo em qual a série é observada. Com mais rigor, pode-se dizer que se x_t é uma série temporal estacionária, se a distribuição de $(x_t, x_{t+1}, \dots, x_{t+s})$ não depende de t para todos os valores de s .

Em situações de não estacionariedade, que ocorrem quando se tem a raiz unitária da equação característica (processo estocástico linear com raiz unitária é não estacionário), recorre-se normalmente à transformação dos dados. Um dos exemplos de transformação de dados mais utilizadas em séries temporais é a diferenciação da série. Em linhas gerais, trata-se de um método, presente nos modelos da classe ARIMA, onde define-se uma nova série a partir das diferenças entre os instantes subsequentes. Este processo pode ser repetido k (parâmetro chamado de ordem da diferenciação) vezes, isto é, quando há uma diferenciação sobre uma série já diferenciada, até obter-se uma outra série y_t estacionária. Uma série diferenciada em ordem k , obviamente, tem o seu tamanho reduzido em k observações, comparado à série original.

No exemplo da diferenciação de primeira ordem (uma única diferenciação), tem-se a série transformada y_t , definida por

$$y_t = \nabla x_t = x_t - x_{t-1}.$$

No caso da série conter múltiplas raízes unitárias, a mesma precisará ser diferenciada algumas vezes até atingir o comportamento de estacionariedade.

Como a série temporal é a realização de um processo estocástico, pode-se classificá-la de acordo com conceitos relacionados à uma certa noção de regularidade e descritos por dois tipos de processos estocásticos estacionários:

- **Processo estocástico estacionário em sentido estrito:**

Um processo estocástico Y_t é designado estacionário em sentido estrito (*strictly stationary*) se a distribuição conjunta de $Y_{t_1}, Y_{t_2}, Y_{t_3}, \dots, Y_{t_n}$ for igual à distribuição conjunta de $Y_{t_1-\delta}, Y_{t_2-\delta}, Y_{t_3-\delta}, \dots, Y_{t_n-k}$ para todos os instantes $t_1, t_2, t_3, \dots, t_n$ e para quaisquer escolhas da defasagem δ , conforme Cryer e Chan (2008) [7]. Dessa maneira, se trasladarmos as variáveis variáveis no tempo, a distribuição conjunta de qualquer conjunto de variáveis permanece inalterada.

- **Processo estocástico estacionário em sentido lato (fraco ou 2º ordem):**

Um processo estocástico Y_t é dito estacionário em sentido lato (*weakly stationary*) se:

- (1) A função média apresenta um valor constante ao longo do tempo, ou seja, $\mu_t = E(y_t) = \mu$;
- (2) A função de autocovariância $\gamma(s, t)$ depende apenas da diferença $|s - t|$. Assim,

$cov(y_t, y_{t+\delta}) = \gamma(t, t + \delta) = \gamma(\delta)$ depende apenas de δ , para qualquer instante t e defasagem δ .

O conceito de estacionariedade estrita é mais forte, porém pouco usual no campo nas aplicações pois dificilmente encontram-se conjuntos de dados com essas propriedades. Por consequência, o termo estacionário refere-se mais de forma usual à condição de estacionariedade no sentido lato, cujas funções de média e autocovariância (caracterização anteriormente descrita) são determinadas pelo primeiro e segundo momento de sua distribuição de probabilidade.

2.1.4 Testes de hipótese

O teste de hipótese consiste em um procedimento de diagnóstico que permite a partir de uma amostra de dados, confirmar ou rejeitar algumas suposições (hipóteses) para a população, com um determinado grau de significância estatística. Para cada teste, há duas hipóteses: a hipótese nula, representada por H_0 , que é a hipótese testada, enquanto que a hipótese alternativa, representada por H_1 , reflete a condição complementar (caso contrário). A regra de decisão para a rejeição de H_0 envolve uma comparação da estatística de teste observada na amostra com o valor teórico da distribuição (chamado valor crítico e que delimita a zona de rejeição), calculado de forma que a probabilidade de rejeitar H_0 , sendo H_0 verdadeira, coincida com o nível de significância α .

De acordo com Rice (2007) [19], a evidência do teste pode ser resumida pelo *p-value*, que neste trabalho, é designado por `valor_p`, que é definido como o menor nível de significância no qual a hipótese nula seria rejeitada. O `valor_p` pode ser compreendido como a probabilidade de se obter um efeito ao menos tão extremo quanto aquele em seus dados amostrais e permite avaliar o resultado de um teste, comparando-o diretamente com α , ao invés da comparação entre a estatística de teste observada e o valor crítico. Todos os resultados de testes de hipótese realizados no componente prático deste estudo, são comunicados dessa forma, referenciando o `valor_p`.

Os testes de hipótese mais utilizados neste trabalho, estão detalhados nesta seção [2.1.4](#).

2.1.4.1 Ljung-Box

O teste de hipótese Ljung-Box é um teste de independência e é realizado sob a hipótese nula da autocorrelação (ρ) conjunta para as m defasagens ($i = \{1, 2, \dots, m\}$) ser nula. Assim, é baseada na formulação:

$$H_0 : \rho_1, \rho_2, \dots, \rho_m = 0 \text{ versus } H_i : \rho_i \neq 0 \text{ para algum } i = \{1, 2, \dots, m\}.$$

De acordo com Tsay (2010) [23], sua estatística de teste, aprimorada do *Portmanteau Test*, é definida por:

$$Q_m = T(T+2) \sum_{k=1}^m \frac{\hat{\rho}_k^2}{T-k},$$

onde T representa o tamanho amostral, e $\hat{\rho}_k$ é o estimador enviesado de ρ_k para amostras finitas, e segue a distribuição de probabilidade χ_m^2 , com m graus de liberdade. A decisão de rejeitar H_0 é tomada quando a condição $Q_m > \chi_\alpha^2$ é satisfeita, sendo α , a significância estatística do teste.

2.1.4.2 Multiplicadores de Lagrange (LM)

O teste do Multiplicadores de Lagrange, também conhecido por teste de efeitos ARCH, é utilizado para verificar a existência de heterocedasticidade condicional, e portanto, a sua prática é recomendável para sustentar a iniciativa de modelação ARCH/GARCH. Este teste é equivalente à estatística F para testar a nulidade os coeficientes de uma regressão linear. Assim, para uma série de retornos que tem a sua média regida pela equação $r_t = \mu_t + z_t$, onde μ_t é um processo médio e z_t , representa os resíduos, tem-se equação dos resíduos quadrados dada por:

$z_t^2 = \alpha_0 + \alpha_1 z_{t-1}^2 + \dots + \alpha_m z_{t-m}^2 + \varepsilon_t$, $i = \{m+1, m+2, \dots, T\}$, tem-se as hipóteses formuladas por:

$$H_0 : \alpha_i = 0, \forall i, i = \{1, 2, \dots, m\} \text{ versus } \alpha_i \neq 0 \text{ para algum } i, i = \{1, 2, \dots, m\}.$$

É baseado em uma estatística de teste F distribuída por χ_m^2 , com m graus de liberdade e que segundo Tsay (2010) [23] é calculada por:

$$F = \frac{(RSS_0 - RSS_1)/m}{RSS_1/(T - 2m - 1)},$$

onde $RSS_0 = \sum_{t=m+1}^T (z_t^2 - \bar{\omega})$, $\bar{\omega} = \frac{\sum_{t=1}^T z_t^2}{T}$, representa a média dos resíduos quadrados, enquanto que RSS_1 , definido por $RSS_1 = \sum_{t=m+1}^T \hat{\varepsilon}_t^2$, refere-se à soma dos resíduos quadrados obtidos na regressão linear pelo método dos mínimos quadrados. A regra de decisão para o teste LM sugere que se a condição $F > \chi_\alpha^2$ for verificada, sendo α , o nível de significância do teste, recomenda-se a rejeição de H_0 , isto é, há a presença de heterocedasticidade nos dados testados.

2.1.4.3 Augmented Dickey Fuller (ADF)

Para examinar se a série tem um comportamento estacionário, costuma-se empregar testes de hipótese de raiz unitária (*unit root tests*), e o teste *Augmented Dickey Fuller* faz parte deste grupo de testes. Denotando a série temporal por x_t , este teste é baseado em uma estatística de teste τ (estatística-t) sob a hipótese nula testada (H_0), com a seguinte formulação:

$$H_0 : x_t \text{ possui raiz unitária versus } H_1 : x_t \text{ é estacionária.}$$

Com $\nabla x_j = x_t - x_{t-1}$ representando a diferenciação de primeira ordem de x_t , de acordo com Tsay (2010) [23], para verificar a presença de uma raiz unitária em um processo AR(p), onde tem-se a seguinte equação:

$$\nabla x_t = c_t + \phi x_{t-1} + \sum_{i=1}^{p-1} b_i \nabla x_{t-i} + \varepsilon_t ,$$

onde c_t é uma função determinística indexado ao tempo t , que na prática pode assumir valor nulo, constante ou $c_t = \omega_0 + \omega_1 t$, o teste é equivalente às hipóteses:

$$H_0 : \phi = 0 \text{ versus } H_1 : \phi < 0 .$$

Assim, se a série original x_t precisa ser diferenciada, então $\hat{\phi} \approx 0$ (não rejeição de H_0). Por outro lado, se x_t já é estacionário, $\hat{\phi} < 0$ (rejeição de H_0).

2.1.4.4 Kwiatkowski-Phillips-Schmidt-Shin (KPSS)

Este teste de raiz unitária é similar ao teste ADF, indicado para o mesmo diagnóstico de estacionariedade, porém suas hipóteses H_0 e H_1 são invertidas (em relação ao ADF), e descritas

por:

$H_0 : x_t$ é estacionária *versus* $H_1 : x_t$ possui raiz unitária.

Este teste é baseado em um modelo regressão linear e se a série for estacionária, o modelo terá um elemento fixo de intercepto ou a mesma ficará estacionária acerca de um nível. A estatística de teste foi derivada do teste unilateral Multiplicadores de Lagrange (LM) pelos autores Kwiatkowski, Phillips, Schmidt e Shin (1992) [15]. Também possui um parâmetro k , que depende do número de observações, referente ao número de *lags* para o cálculo da estatística de teste e no *R Software*, há duas versões em termos das defasagens, curto e longo prazo, com uma diferença na fórmula de cálculo do valor padrão. Sendo n o tamanho da série, tem-se: $k_{short} = trunc(4 \sqrt[4]{\frac{n}{100}})$ e $k_{long} = trunc(12 \sqrt[4]{\frac{n}{100}})$. Obviamente, pode ser utilizado em conjunto com o teste ADF como uma estratégia de reforço de diagnóstico.

2.2 Conceitos Financeiros

2.2.1 Retorno

Como exemplo de um ativo financeiro tem-se a cotação do valor de uma ação de um determinada empresa, que mensuradas em um determinado intervalo de tempo, forma uma série temporal financeira. A análise estatística direta desses preços de ativo é difícil porque em geral as séries de preços apresentam forte correlação e não apresentam estacionariedade, e portanto sua variância não se mantém constante ao longo do tempo, o que gera dificuldades na aplicação da inferência estatística.

A maior parte dos estudos que envolvem ativos financeiros utiliza em substituição ao preço do ativo, o seu retorno financeiro. Campbell, Lo e MacKinlay (1997) [4], destacam os principais motivos em se utilizar os retornos:

- Para os investidores, o retorno é uma medida completa e livre de escala da oportunidade;
- O retorno de um ativo financeiro que possui propriedades estatísticas mais fáceis de tratar.

Estes mesmos argumentos favoráveis em se utilizar os retornos foram também corroborados por Tsay (2010) [23]. No universo dos pesquisadores na análise de séries temporais financeiras

ras, é usual a utilização do **retorno composto continuamente** (*Continuously Compounded Return*), também denominado **log-retorno** ou por apenas **retorno**.

Assumindo-se a premissa de que a série temporal dos preços seja regular (espaçamento constante entre as observações consecutivas) e que não haja pagamento de dividendos neste intervalo de tempo, pode-se definir que se P_t é o preço de um ativo no instante t , a razão $\frac{P_t}{P_{t-1}}$ é denominada retorno bruto simples e o logaritmo do retorno bruto simples, chamado de retorno composto continuamente, e representado por r_t , é definido por:

$$r_t = \log \frac{P_t}{P_{t-1}} = \log_{P_t} - \log_{P_{t-1}} = \nabla \log_{P_t} . \quad (2.1)$$

Também existe o conceito do retorno líquido simples que é calculado por $\frac{P_t - P_{t-1}}{P_{t-1}}$, que denota o ganho líquido da operação “vender hoje o ativo adquirido ontem”, também denominado variação percentual, que quando é representado por pequenos valores, aproxima-se do retorno composto continuamente [20], isto é, $\frac{P_t - P_{t-1}}{P_{t-1}} \approx \nabla \log_{P_t}$.

O escopo deste trabalho é estudar taxas de câmbio, que expressa a relação entre dois ativos financeiros, que dito em outras palavras significa a taxa à qual se pode trocar a moeda de um país pela moeda de outro país, através de operações de compra e venda, além de alguns outros ativos financeiros (índices de ações). Adotou-se neste trabalho como retorno, especificamente o retorno composto continuamente (r_t), definido anteriormente, e que a partir daqui, por questão de simplificação, será referenciado apenas pelo termo retorno.

2.2.2 Volatilidade

A volatilidade, que é muito utilizada para a mensuração do risco de instrumentos financeiros, pode ser compreendida como uma medida de dispersão dos preços (ou retornos) de um ativo financeiro, e que portanto, avalia o grau de variabilidade. Assim, uma série de preços que comporta-se de forma instável, é considerada uma série volátil.

A volatilidade está relacionada ao risco, mas possui um conceito distinto. O risco está associado somente a resultados indesejáveis, como por exemplo, medindo-se o grau de incerteza de se obter um retorno esperado em um investimento financeiro enquanto que a volatilidade é uma medida mais rigorosa porque capta também as incertezas diante de um resultado positivo,

conforme Poon (2005) [18].

Apesar de tratar-se de uma medida não diretamente observável, há várias maneiras de estimá-la, sendo o desvio padrão amostral, denotado por $\hat{\sigma}$, uma das formas mais simples e muito frequentemente empregada nos estudos, definida por:

$$\hat{\sigma} = \sqrt{\frac{1}{T-1} \sum_{i=1}^T (r_i - \mu)^2}, \quad (2.2)$$

onde T é o número de observações, r_t é o retorno financeiro no instante t , e μ , representa a média dos retornos observados. A variância, definida pelo quadrado do desvio padrão, também é utilizada neste mesmo sentido.

Em uma série financeira, a volatilidade tende a se apresentar em grupos (*Volatility Clustering*), motivo pelo qual intui-se que existe algum grau de dependência de seus valores com o instante t . É comum observar nas séries temporais dessa natureza, padrões de períodos estáveis sendo sucedidos por períodos estáveis e períodos turbulentos (instáveis) sucedidos por períodos turbulentos. De acordo com Poon (2005) [18], dentre outras características marcantes nas séries de retorno do mercado financeiro documentadas pela literatura, aparece o fenômeno da alavancagem, mais conhecido pelo termo expresso na língua inglesa *leverage effect*, no qual a volatilidade comporta-se de maneira assimétrica e tende a aumentar depois de se observar um retorno negativo, sendo esse aumento mais acentuado no período de grandes baixas. Isso se deve à elevação do risco financeiro da empresa na qual suas ações sofreram queda de preço.

2.3 Modelos

2.3.1 AR

Os modelos autoregressivos, mais comumente conhecidos pela sigla AR, foram concebidos a partir da ideia central de que os valores atuais em uma série temporal podem ser explicados por uma função dos valores observados no passado, ou seja, existe uma relação de dependência temporal entre as observações da mesma série.

Se y_t representa uma série temporal, estacionária por premissa para esta classe de modelos, onde o índice t tem significado temporal, é possível construir um modelo autoregressivo de

ordem p , abrevia-se $AR(p)$, de modo que se possa determinar y_t por uma função de seus p valores passados, referenciados por $y_{t-1}, y_{t-2}, \dots, y_{t-p}$ e que desempenham o papel de variáveis explicativas no referido modelo.

Definição: Um modelo autoregressivo de ordem p é descrito pela expressão:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t, \quad (2.3)$$

onde a série y_t é estacionária.

Assim, diante de um modelo $AR(1)$, como um exemplo, a estimação realizada de y_t , em termos temporais, depende exclusivamente de seu valor passado no instante anterior (y_{t-1}), e este modelo pode ser representado matematicamente por:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \varepsilon_t,$$

onde ε_t é um ruído branco ² de média 0 e variância σ^2 , idealmente não autocorrelacionados, assumindo-se a hipótese de que média e variâncias são contantes.

Desta forma, $E(y_t) = E(y_{t-1}) = \mu$ e $E(\varepsilon_t) = 0$, o que implica dizer que $E(y_t) = \phi_0 + \phi_1 E(y_{t-1})$. Substituindo os termos $E(y_t)$ e $E(y_{t-1})$ por μ e isolando o termo μ , chega-se à expressão $\mu = \frac{\phi_0}{1-\phi_1}$, onde observa-se, que o seu valor esperado existe desde que $\phi_1 \neq 1$ e que este valor, não depende do instante t , mas exclusivamente dos coeficientes ϕ_0 e ϕ_1 .

Considerando a hipótese de variância constante, tem-se $var(y_t) = var(y_{t-1})$. Assim, tem-se a expressão $var(y_t) = var(\phi_0 + \phi_1 y_{t-1} + \varepsilon_t) = \phi_1^2 var(y_t) + \sigma^2$, que desenvolvida chega-se à equação da variância:

$$var(y_t) = \frac{\sigma^2}{1 - \phi_1^2},$$

onde estabelece-se a restrição de que $|\phi_1| < 1$, por conta da estacionariedade.

²ruído branco ou *white noise* é o nome dado às séries temporais geradas por variáveis não correlacionadas e utilizadas como um modelo de ruído em aplicações de engenharia [20].

2.3.2 MA

Um modelo de média móvel de ordem q , usualmente abreviado por $MA(q)$ por ser um termo proveniente do inglês *moving average*, é definido por uma expressão típica de um modelo de regressão, onde os erros ocorridos em instantes passados $t - q, t - q + 1, \dots, t - 2, t - 1$ exercem papel de preditores (variáveis explicativas). Assim, um modelo $MA(q)$, média móvel de ordem q , é descrito matematicamente por:

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} , \quad (2.4)$$

onde c é uma constante e ε_t é um ruído branco. Pode ser compreendido como uma média ponderada dos erros passados embora não se trata de um regressão no sentido mais usual, pois os termos referentes aos resíduos $(\varepsilon_t, \varepsilon_{t-1}, \dots)$ não são valores observados.

2.3.3 ARMA

De forma análoga ao modelo AR, no modelo ARMA, estima-se a média condicional aos valores passados, com a premissa de que a variância condicional mantém-se contante. Basicamente, este modelo refere-se a um modelo AR agregado a um novo componente de médias móveis (MA), e que desempenha o papel de capturar a tendência da série temporal. O modelo $ARMA(p, q)$ possui dois parâmetros:

- (1) p : ordem do modelo autoregressivo (AR). Exemplo: se $p = 2$, então y_t é função de y_{t-1} e y_{t-2} ;
- (2) q : ordem do modelo de médias móveis (MA). Exemplo: se $q = 1$, então o modelo de y_t , em termos de resíduos, depende de ε_t e ε_{t-1} .

Preservando as mesmas representações de um modelo AR descrito anteriormente (2.3) e acrescentando $\theta_1, \theta_2, \dots, \theta_q$ como os coeficientes do modelo de médias móveis, pode-se escrever a equação geral de um modelo $ARMA(p, q)$ por:

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t , \quad (2.5)$$

sendo os resíduos $\varepsilon_t \sim iid(0, \sigma^2)$.

2.3.4 ARIMA

Um modelo autoregressivo integrado de médias móveis, interpretação literal da sigla ARIMA, tem a sua ordem descrita por uma tripla de parâmetros:

- (1) p : ordem do modelo autoregressivo (AR);
- (2) d : parâmetro que denota o número de vezes que a série y_t é diferenciada;
- (3) q : ordem do modelo de médias móveis (MA).

O modelo pode ser escrito em sua forma polinomial, utilizando um operador de defasagem conhecido por B . Este operador de defasagem (B), apresenta o seguinte conceito: diante de uma série y_t , sua primeira diferença é definida pela operação $\nabla y_t = y_t - y_{t-1} = y_t - B y_t = (1 - B)y_t$. Em termos gerais, se esta operação for de ordem d , isto é, realiza-se d sucessivas diferenças na série, a mesma é denotada por $(1 - B)^d y_t$.

Assim, de acordo com Hyndman e Athanasopoulos (2018) [13], um modelo ARIMA pode ser escrito por:

$$(1 - \phi_1 B - \dots - \phi_p B^p) (1 - B)^d y_t = c + (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t, \quad (2.6)$$

onde o termo “ $(1 - \phi_1 B - \dots - \phi_p B^p)$ ” corresponde ao modelo $AR(p)$, “ $(1 - B)^d y_t$ ” denota as d diferenças, enquanto que o lado direito da equação “ $c + (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t$ ” representa o modelo $MA(q)$.

Assumindo $c = \mu (1 - \theta_1 - \dots - \theta_p)$, $y'_t = (1 - B)^d y_t$ e μ , como a média de y_t , a expressão 2.6 pode ser reparametrizada por:

$$(1 - \phi_1 B - \dots - \phi_p B^p) (y'_t - \mu) = (1 + \theta_1 B + \dots + \theta_q B^q) \varepsilon_t. \quad (2.7)$$

Embora possa haver necessidade da série ser diferenciada mais que uma vez ($d \geq 1$), a título ilustrativo, um modelo ARIMA com a série diferenciada uma única vez (parâmetro $d = 1$), ou seja, $y'_t = \nabla y_t = y_t - y_{t-1}$, também pode ser escrito por:

$$y'_t = \phi_0 + \phi_1 y'_{t-1} + \phi_2 y'_{t-2} + \dots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} + \varepsilon_t. \quad (2.8)$$

O modelo ARIMA também é chamado de raiz unitária não estacionário porque o AR polinomial possui raiz unitária.

Existe uma correspondência entre o modelo ARIMA(p,d,q) e os modelos ARMA(p,q), AR(p)

e $MA(q)$, dependendo dos parâmetros utilizados. Assim, tem-se:

$$ARIMA(p, 0, q) \equiv ARMA(p, q) , ARIMA(p, 0, 0) \equiv AR(p) , ARIMA(0, 0, q) \equiv MA(q) .$$

Desta forma, com o intuito de tornar mais concreta a comunicação pode-se descrever um processo y_t regido por um modelo $ARIMA(p, d, q)$ exemplificando a expressão matemática de um modelo $ARIMA(1, 1, 1)$ tem-se:

$$\nabla y_t = y_t - y_{t-1} = \phi_1 \nabla y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} , \quad (2.9)$$

onde ε_t representa os resíduos, idealmente como ruído branco, ϕ_1 representa o coeficiente do modelo $AR(1)$ enquanto que θ_1 exerce papel de coeficiente do modelo $MA(1)$. Como o parâmetro de diferenciação $d = 1$, o termo $\nabla y_t = y_t - y_{t-1}$ está presente na equação do modelo, indicando que y_t é diferenciada uma única vez.

Assim, o mesmo modelo $ARIMA(1, 1, 1)$ pode ser reescrito por:

$$(1 - B)y_t = \phi_1 (1 - B) y_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} . \quad (2.10)$$

Como $y_{t-1} = B y_t$ e $\varepsilon_{t-1} = B \varepsilon_t$, pode-se substituí-los na expressão 2.10 transformando-a em 2.11:

$$(1 - B)y_t = \phi_1(1 - B)B y_t + \varepsilon_t + \theta_1 B \varepsilon_t , \quad (2.11)$$

que desenvolvida, tem a sua expressão final:

$$(1 - \phi_1 B)(1 - B)y_t = (1 + \theta_1 B)\varepsilon_t , \quad (2.12)$$

onde ε_t é ruído branco (sem correlação significativa) e B é o operador de defasagem (*lag operator*).

Informações complementares e mais detalhadas sobre o modelo ARIMA, e seus modelos componentes (AR, MA e ARMA), podem ser obtidas através das referências Shumway e Stoffer (2017) [20], Hyndman e Athanasopoulos (2018) [13], Box e Jenkins (1970) [2] e Brockwell e Davis (2016) [3].

2.3.5 Regressão Dinâmica

Diferentemente do que ocorre com os modelos da classes ARIMA, onde a estimação dos valores atuais de uma série temporal baseia-se unicamente nos valores observados no passado, no modelo de regressão dinâmica há também a presença de um outro componente obtido por um modelo de regressão múltipla com outras séries temporais relevantes (variáveis externas), que exercem alguma influência na série a ser estudada, desempenhando o papel de variáveis explicativas. Pode-se resumir o modelo de regressão dinâmica como um modelo híbrido composto por um modelo de regressão e um outro modelo, que trata de modelar os resíduos da regressão. No componente prático deste estudo, o modelo de regressão dinâmica utilizado foi na configuração regressão linear múltipla combinado com o modelo ARIMA.

Pode-se compreender a ideia de como os modelos de regressão dinâmica são construídos, baseado em Hyndman e Athanasopoulos (2018) [13], a seguinte suposição: deseja-se realizar a modelação de uma série temporal y_t (variável resposta ou variável dependente) contando também com a disponibilidade de outras k séries temporais relevantes, representadas por $(x_{1,t}, x_{2,t}, x_{3,t}, \dots, x_{k,t})$. Matematicamente, este modelo de regressão linear múltipla, sendo $(x_{1,t}, x_{2,t}, x_{3,t}, \dots, x_{k,t})$ as variáveis explicativas (ou independentes), é descrito por:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \dots + \beta_k x_{k,t} + \eta_t . \quad (2.13)$$

Os resíduos η_t de um modelo de regressão clássico devem, idealmente ser não correlacionados, e descritos como um processo de ruído branco (*white noise*), de modo que as variáveis predictoras $(x_{1,t}, x_{2,t}, x_{3,t}, \dots, x_{k,t})$ consigam explicar totalmente o padrão de comportamento de y_t .

O processo de modelação por regressão dinâmica inicia-se a partir de uma regressão clássica. No entanto, nesta fase da modelação aceita-se autocorrelação significativa nos resíduos (η_t), uma vez que os mesmos serão ajustados posteriormente por um outro modelo, que neste exemplo, será um modelo $ARIMA(1, 1, 1)$.

Assim se η_t segue um modelo $ARIMA(1, 1, 1)$, pode-se escrever:

$$\nabla \eta_t = \eta_t - \eta_{t-1} = \phi_1 \nabla \eta_{t-1} + \varepsilon_t + \theta_1 \varepsilon_{t-1} , \quad (2.14)$$

$$(1 - \phi_1 B)(1 - B)\eta_t = (1 + \theta_1 B) \varepsilon_t , \quad (2.15)$$

onde ε_t é ruído branco (sem correlação significativa) e B é o operador de defasagem (*lag operator*), neste caso η_t é diferenciada uma única vez, ϕ_1 é o coeficiente do modelo $AR(1)$ enquanto que θ_1 refere-se ao coeficiente do modelo $MA(1)$. Os resíduos finais (ε_t) idealmente precisam ser não autocorrelacionados.

Informações complementares sobre o modelo de regressão dinâmica podem ser encontradas no livro de Hyndman e Athanasopoulos (2018) [13], especificamente no capítulo 9, intitulado *Dynamic regression models*, e também na referência Pankratz (1991) [16].

2.3.6 ARCH

Os valores dos retornos passados utilizados nos modelos da classe ARMA são igualmente importantes nos modelos ARCH, embora no primeiro modelo, estima-se o valor esperado condicional de um processo, cuja variância mantém-se constante ao longo do tempo, enquanto que nos modelos da classe ARCH (*Autoregressive Conditionally Heteroscedastic*), a variabilidade altera-se sistematicamente ao longo do tempo, resultando assim em uma alternância entre períodos de comportamento mais estável e períodos mais turbulentos, com padrão de expressiva variação. A designação heteroscedasticidade remete exatamente a essa característica de variação da volatilidade. O objetivo do modelo é capturar essa volatilidade persistente através de uma autoregressão de valores passados dos retornos quadrados.

Essa variabilidade estimada por este modelo (variância condicional), mais conhecida por volatilidade, é o seu grande diferencial. Este modelo é mais adequado à séries temporais não lineares com fortes características de volatilidade, razão pelo qual este modelo é bastante empregado nas séries financeiras.

Estes modelos foram introduzidos em 1982 pelo economista norte-americano Robert Engle [9], foram concebidos com a ideia de caracterizar e prever os movimentos voláteis nos preços do mercado financeiro, propiciando uma melhor gestão dos riscos. A prova da relevância desta contribuição de Engle foi que este trabalho levou-o a ganhar o Prêmio Nobel de Economia em 2003.

O grande interesse deste modelo é estimar a variância condicional dos retornos, ou seja, conhecendo-se previamente os valores passados do retorno $(r_{t-1}, r_{t-2}, \dots)$, conseguir calcular o valor de sua variância no instante atual (instante t). Assim, em uma situação de valor esperado condicional nulo, isto é $E(r_t | r_{t-1}, r_{t-2}, \dots) = 0$, que é verdadeira em muitos casos, a variância condicional é dada por:

$$\text{var}(r_t | r_{t-1}, r_{t-2}, \dots) = E(r_t^2 | r_{t-1}, r_{t-2}, \dots) .$$

Genericamente, estabelece-se que se o retorno r_t é definido por um processo médio representado por μ_t (ou simplesmente referenciado por μ , no caso de uma média constante ou invariante no tempo), agregada a um processo z_t , isto é, $r_t = \mu_t + z_t$, então z_t assume o papel de desvio, ou erro, em relação à sua média. Segundo Brockwell e Davis (2016) [3], diz-se que se z_t é um processo $ARCH(p)$, lê-se ARCH de ordem p , então o mesmo é descrito como uma solução estacionária das equações:

$$z_t = \sqrt{h_t} \varepsilon_t = \sigma_t \varepsilon_t , \tag{2.16}$$

onde os resíduos $\varepsilon_t \sim iid N(0, 1)$.

$$h_t = \omega + \alpha_1 z_{t-1}^2 + \alpha_2 z_{t-2}^2 + \dots + \alpha_p z_{t-p}^2 , \tag{2.17}$$

onde h_t representa a variância condicional de z_t dado $(z_{t-1}, z_{t-2}, \dots, z_{t-p})$ e σ_t , designa o desvio padrão de z_t . Como condição para que a variância h_t seja estritamente positiva, há restrições aos coeficientes do modelo: $\omega > 0$ e $(\alpha_{t-1} \geq 0, \alpha_{t-2} \geq 0, \dots, \alpha_{t-p} \geq 0)$, que serão estimados e avaliados em relação às suas significâncias estatísticas.

A variância incondicional de r_t é dada por $\sigma^2 = \frac{\omega}{1 - \sum_{i=1}^p \alpha_i}$.

Assim, diz-se que z_t segue um processo $ARCH(p)$ quando os valores quadrados z_t^2 seguem um modelo $AR(p)$.

Em uma situação hipotética de média constante e nula ($\mu = 0$), tem-se $r_t = z_t$, pode-se caracterizar o modelo mais simples desta classe, com o modelo $ARCH(1)$, de acordo com as equações do retorno e da variância condicional dos retornos definidas por $r_t = \sigma_t \varepsilon_t$ e $\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2$, respectivamente, sendo os resíduos $\varepsilon_t \sim iid N(0, 1)$. Assim, conclui-se que a

variância condicional dos retornos (também é interpretada como uma medida de volatilidade), pode ser explicada por uma função dos retorno quadrado observado no instante anterior $t-1$.

2.3.7 GARCH

Os modelos ARCH foram generalizados por Bollerslev [1] em 1986, sob o acrônimo GARCH, com a designação de *Generalized Autoregressive Conditionally Heteroscedastic*. Esta generalização dos modelos GARCH consiste numa extensão de um modelo ARCH, incorporando um novo componente que é função também dos valores passados da própria volatilidade. Embora o modelo ARCH seja mais simples, frequentemente demanda muitos parâmetros para descrever apropriadamente a volatilidade (Tsay, 2002) [22], o que estabelece uma vantagem ao modelo GARCH em termos de sua aplicabilidade. O modelo GARCH padrão (*standard*) de z_t é matematicamente descrito pelas equações de média condicional e variância condicional, respectivamente, por:

$$z_t = \sqrt{h_t} \varepsilon_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim iid(0, 1), \quad (2.18)$$

$$h_t = \sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i z_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2, \quad (2.19)$$

onde z_t descreve o desvio dos retornos em relação a um processo médio, isto é, $z_t = r_t - \mu_t$, h_t representa a variância condicional de z_t , conhecendo-se previamente os retornos passados e σ_t é o desvio padrão condicional de z_t . Embora seja usual assumir que a distribuição condicional de z_t é Normal de média nula, $z_t \sim iid N(0, \sigma_t^2)$, é frequente, na análise empírica de séries financeiras, não fixá-la como premissa e ajustar o modelo com distribuições com outras características, a exemplo da distribuição *t-Student*, que pelo fato de possuir caudas mais pesadas, geralmente produz melhores ajustes, segundo os autores Brockwell e Davis (2016) [3].

O termo cauda de uma distribuição designa a parte da distribuição correspondente aos valores extremos de uma variável aleatória. Conforme Klugman, Panjer e Willmot (2019) [14], variáveis aleatórias que tendem a atribuir probabilidades mais elevadas a valores extremos são consideradas terem caudas pesadas. Este conceito do “peso” de uma distribuição pode ser abordado tanto de forma relativa (distribuição A possui cauda mais pesada que distribuição B),

analisando-se a razão entre as funções de distribuição ou a razão entre as funções de densidade de probabilidade de duas variáveis aleatórias, quanto pode ser abordada por conceito absoluto (distribuições com certas propriedades são classificadas como caudas pesadas).

Este modelo $GARCH(p, q)$ apresenta, além dos coeficientes do modelo ARCH, ω ($\omega > 0$) e $(\alpha_{t-1} \geq 0, \alpha_{t-2} \geq 0, \dots, \alpha_{t-p} \geq 0)$, também contempla os q coeficientes β , associados aos termos do modelo relacionados à volatilidade de z_t , isto é, $(\beta_{t-1} \geq 0, \beta_{t-2} \geq 0, \dots, \beta_{t-q} \geq 0)$.

Caso a soma dos parâmetros α e β seja inferior a 1, tem-se um modelo GARCH estacionário, isto é, quando $\sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$ se verifica.

Partindo-se da premissa de que $\mu = 0$, o implica que $r_t = z_t$, pode-se ilustrar um modelo $GARCH(1, 2)$, para fins didáticos, pelas equações $r_t = \sigma_t \varepsilon_t$, $\sigma_t^2 = \omega + \alpha_1 r_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + \beta_2 \sigma_{t-2}^2$. A interpretação é que a variância condicional do retorno, ou o quadrado da volatilidade do retorno no instante t , pode ser explicado por uma função do retorno quadrado do instante anterior $t-1$ e do quadrado da volatilidade observadas nos dois instantes anteriores $(t-1, t-2)$.

Embora não seja foco deste trabalho, vale registrar que existem diversas variantes de modelos GARCH entre as quais é mostrado uma pequena relação, acompanhado de uma breve descrição de sua essência:

- TS-GARCH (Taylor-Schwert GARCH):

modelo que tem a variância condicional substituída pelo desvio padrão condicional, que por sua vez, são modelados como uma função linear de resíduos absolutos defasados.

- EGARCH (Exponential GARCH):

diferencia-se do GARCH padrão pela estrutura da variância (substitui-se σ_t^2 por $\log \sigma_t^2$).
Permite os efeitos assimétricos dependendo do sinal do termo erro;

- GJR-GARCH (Glosten-Jagannathan-Runkle GARCH):

Permite efeitos assimétricos de erros passados de forma simples (utiliza uma função indicadora na equação da variância);

- GARCH-M (GARCH-in-mean):

incorpora-se o efeito da volatilidade da série na equação da média.

Para o processo de modelação GARCH, vários *softwares* estatísticos disponibilizam funções de ajustamento aos dados que modelam simultaneamente a média condicional μ_t com um

modelo ARMA e a variância condicional h_t com um modelo GARCH, a exemplo das funções *garchFit* do *Package 'fGarch'* [25] e *ugarchfit* do *Package 'rugarch'* [11], pertencentes ao *R Software* e que foram utilizados no componente prático deste projeto, especificamente no capítulo Aplicação e Resultados (3).

Informações mais detalhadas sobre os modelos de heterocedasticidade condicional ARCH e GARCH, utilizados para modelação da volatilidade, podem ser encontradas nas referências Brockwell e Davis (2016) [3], Enders (2015) [8], Shumway e Stoffer (2017) [20] e Tsay (2010) [23].

2.3.8 APARCH

Em algumas séries, tipicamente nas séries financeiras, aparece os *leverage effects* em que os altos retornos negativos provocam um aumento substancialmente maior na volatilidade comparado aos retornos positivos de mesma magnitude. O modelo GARCH padrão, descrito anteriormente, não contempla esses efeitos uma vez que a volatilidade é função dos retornos quadrados, e portanto, não importa se é positivo ou negativo, geram impactos idênticos (efeitos simétricos) se tiverem a mesma grandeza. O modelo denominado APARCH (Assimetric Power ARCH) posiciona-se exatamente neste contexto e pode ser uma alternativa para a modelação de volatilidade com efeitos assimétricos.

Assim, se $z_t \sim APARCH(p, q)$, o modelo pode ser escrito pelas equações da média e variância de z_t , respectivamente, da seguinte forma:

$$z_t = \sqrt{h_t} \varepsilon_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim iid(0, 1), \quad (2.20)$$

$$\sigma_t^\delta = \omega + \sum_{i=1}^p \alpha_i (|z_{t-i}| - \gamma_i z_{t-i})^\delta + \sum_{j=1}^q \beta_j \sigma_{t-j}^\delta, \quad (2.21)$$

em que $\omega > 0$, $\delta > 0$, $\alpha_i \geq 0$, $-1 < \gamma_i < 1$ para $i = 1, 2, \dots, p$ onde γ_i desempenha o papel do coeficiente de alavancagem (*leverage*) e $\beta_j \geq 0$ para $j = 1, 2, \dots, q$. Comparado ao GARCH padrão, o modelo APARCH mostra-se mais robusto e flexível porque possui um parâmetro adicional δ como expoente, além de contar com coeficiente de alavancagem em sua equação.

A família APARCH inclui os modelos ARCH e GARCH e alguns modelos podem ser derivados deste modelo. Wurtz, Chalabi e Luksan (2006) [26] relaciona alguns deles:

- ARCH: quando $\delta = 2$, $\gamma_1 = \gamma_2 = \dots = \gamma_p = 0$ e $\beta_1 = \beta_2 = \dots = \beta_q = 0$;
- GARCH: quando $\delta = 2$ e $\gamma_1 = \gamma_2 = \dots = \gamma_p = 0$;
- TS-GARCH: quando $\delta = 1$ e $\gamma_1 = \gamma_2 = \dots = \gamma_p = 0$;
- GJR-GARCH: quando $\delta = 2$.

2.4 Coeficiente de determinação

Em um modelo de regressão linear descrito por $Y = X\beta + \varepsilon$, sendo ε os resíduos com $\varepsilon \sim iid(0, \sigma^2)$, não necessariamente com distribuição Normal, Y é a variável resposta (variável dependente) e X , o vetor das variáveis explicativas $X = (x_1, x_2, x_3, \dots)$, o coeficiente de determinação, designado por R^2 , tem a interpretação do percentual de variância de Y (variável resposta) explicada pelo modelo.

O R^2 corresponde ao quadrado do coeficiente de correlação linear de Pearson amostral entre os n valores observados Y e os valores ajustados \hat{Y} , isto é,

$$R^2 = \rho_{Y, \hat{Y}}^2 = \sum_{i=1}^n \frac{Cov(Y, \hat{Y})}{\sigma_Y \sigma_{\hat{Y}}} = \frac{(\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}}))^2}{\sum_{i=1}^n (y_i - \bar{y}) \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})}$$

Desta forma, sendo RSS a soma dos quadrados dos resíduos e TSS, a soma dos quadrados total, é definido por: por $R^2 = 1 - \frac{RSS}{TSS}$ onde, $RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ e $TSS = \sum_{i=1}^n (y_i - \bar{y})^2$. Também pode ser expresso por uma relação entre os estimadores de máxima verossimilhança da variância residual ($\hat{\sigma}^2$) e da variância de Y ($\hat{\sigma}_Y^2$), definidos por $\hat{\sigma}^2 = \frac{RSS}{n}$ e $\hat{\sigma}_Y^2 = \frac{TSS}{n}$, através da equação $R^2 = 1 - \frac{\hat{\sigma}^2}{\hat{\sigma}_Y^2}$.

O coeficiente de determinação possui algumas propriedades importantes tais como: não tem unidade de medida e o seu valor está compreendido no intervalo entre zero e um ($0 \leq R^2 \leq 1$), com o significado de quanto mais próximo do valor um, melhor é o ajustamento do modelo. Além disso, apresenta uma importante limitação, sempre que adiciona-se uma nova variável explicativa ao modelo, o novo R^2 calculado será sempre superior ou igual ao do modelo anterior, uma vez que a soma dos quadrados dos resíduos não pode aumentar e a soma total dos quadrados dos desvios de Y relativamente à sua média se mantém constante, podendo-se gerar uma falsa

percepção de que modelo o novo modelo é mais eficiente que o anterior.

O coeficiente de determinação ajustado, representado por \bar{R}^2 , corrige essa limitação relatada, adequando o cálculo tendo em conta o número de graus de liberdade envolvidos no cálculo da variância amostral de Y e de \hat{Y} . Considerando p o número de parâmetros do modelo, e os estimadores não enviesados da variância residual, representado por $\bar{\sigma}^2$ e da variância de Y , representado por $\overline{\sigma_Y^2}$, onde $\bar{\sigma}^2 = \frac{RSS}{n-(p+1)}$ e $\overline{\sigma_Y^2} = \frac{TSS}{n-1}$, tem-se a sua definição descrita pela seguinte equação:

$$\bar{R}^2 = 1 - \frac{\bar{\sigma}^2}{\overline{\sigma_Y^2}} = 1 - \frac{RSS/(n - (p + 1))}{TSS/(n - 1)} = 1 - \frac{n - 1}{n - (p + 1)}(1 - R^2) .$$

O coeficiente de determinação ajustado \bar{R}^2 pode ser utilizado para avaliar a qualidade de ajustamento do modelo (é uma métrica do tipo *goodness-of-fit* e segue a mesma lógica do R^2 de quanto mais próximo do valor 1, melhor é o ajustamento alcançado) e tem interpretação da percentagem dos desvios de Y em relação à média que é explicada pelo modelo. Pode assumir valor negativo, que é uma inconveniência, e é sempre limitado pelo valor de R^2 ($\bar{R}^2 \leq R^2 \leq 1$).

2.5 Critérios de Informação

Uma técnica utilizada como diagnóstico de um modelo estatístico para avaliar o grau de ajustamento aos dados da amostra é a adoção de critérios de informação. Trata-se de estatísticas que medem a quantidade de informação que o modelo recolhe dos dados. Vale lembrar que os erros (ruídos) referem-se à porção que o modelo não consegue explicar. Esses critérios de informação são muito úteis pois viabiliza a comparação entre diferentes modelos, de modo que ajuda a justificar a decisão da escolha do melhor modelo.

2.5.1 AIC

Um dos critérios de informação bastante usual, também utilizado para identificar a ordem de um modelo AR, é conhecido pelo acrônimo AIC (*Akaike Information Criterion*), que de acordo com Cowpworthwait e Metcalfe (2011) [6] é definido por:

$$AIC = -2 LL + 2 k ,$$

onde LL refere-se ao logaritmo da função de verossimilhança, onde k é o número total de parâmetros do modelo. Assim, uma medida da quantidade de informação é o logaritmo da verossimilhança do modelo com a interpretação de quanto menor o seu valor, maior é a aderência do modelo aos dados (modelo mais favorável) e portanto, um elevado número de parâmetros k do modelo, penaliza o critério de informação, uma vez que isto proporciona uma elevação de seu valor.

2.5.2 AICc

Um outro critério de informação relevante é o AICc (*Akaike Information Criterion corrected*) que praticamente elimina o viés (*bias*) do AIC, e conforme Cryer e Chan (2008) [7] é definido por:

$$AICc = AIC + \frac{2(k+1)(k+2)}{(n-k-2)},$$

sendo k , o número total de parâmetros do modelo e n , o número total de observações (dimensão da amostra).

2.5.3 BIC

Outra abordagem possível para determinar-se a ordem de um modelo ARMA é selecionar o modelo com o menor valor do critério de informação BIC (*Bayesian Information Criterion*), que Cryer e Chan (2008) [7] define por:

$$BIC = -2 LL + k \log(n),$$

onde k representa o número total de parâmetros do modelo, enquanto que n refere-se ao tamanho amostral. Esta expressão do BIC equivale ao critério AIC para um tamanho amostral $n = 100$, sendo que à medida que n aumenta, penaliza-se o critério BIC (valores mais altos são desfavoráveis) com a ideia de que um modelo obtido por uma quantidade mais elevada de observações tivesse a “obrigação” de extrair mais informações, isto é, ter um maior poder de explicação.

2.6 Indicadores de erro

Em um processo de modelação de *Forecasting* requer-se instrumentos para avaliar o nível de acurácia dos modelos testados, ou seja, precisa-se medir de alguma forma se a qualidade do ajustamento é ou não satisfatória. Neste sentido, recorre-se aos indicadores (ou métricas) que medem o nível de erro das previsões geradas pelo modelo. Neste tema, pode-se também abordar também a acurácia, um conceito complementar relacionado ao seu grau de precisão, ou seja, mede-se o quanto o modelo acertou, ao invés do quanto o modelo errou. Assim, utilizar a acurácia, que de forma indireta tem o mesmo significado que utilizar o erro, pode ser mais interessante no aspecto psicológico, dependendo das circunstâncias, porque comunica uma qualidade positiva do modelo. No entanto, existem muitos indicadores de erro, com particularidades e indicação de uso próprias, sendo todos convergentes na indicação do nível de erro do modelo.

Em termos práticos, para mensurá-los, recorre-se usualmente à estratégia de dividir o conjunto de dados em duas partes chamadas de treino, que concentra a maior quantidade de observações ($\approx 80\%$ apenas com valor típico de referência) e é utilizado para a calibração do modelo (*fit*); e teste, que contém os dados complementares e que desempenha o papel de propiciar a avaliação do modelo obtido, com dados não utilizados no ajuste do modelo.

Conforme Hyndman e Athanasopoulos (2018) [13], o termo “erro” de previsão (*Forecasting error*), aqui designado por e_t , deve ser compreendido como um desvio, calculado por simples operação aritmética da diferença entre o valor observado (valor realizado) e o valor previsto (valor estimado pelo modelo). Não há o significado literal de erro, a métrica refere-se à porção que o modelo não capturou da observação, isto é, a parte não prevista pelo modelo, que pode inclusive ser aleatória e portanto, não previsível.

Matematicamente, o erro é descrito por: $e_{T+h} = y_{T+h} - \hat{y}_{T+h|T}$, onde y_t refere-se ao valor observado no instante $T + h$ enquanto $\hat{y}_{T+h|T}$ representa o valor estimado pelo modelo para este mesmo instante $T + h$, $\{y_1, y_2, \dots, y_T\}$ representa as observações do conjunto de treino enquanto que $\{y_{T+1}, y_{T+2}, \dots\}$, refere-se ao conjunto de teste, no qual o modelo será avaliado. Desta forma, a diferença básica entre os resíduos e os erros de previsão é que usam diferentes conjuntos de dados, treino e teste, respectivamente, além de o cálculo do *forecast* envolver várias etapas, a

exemplo do *rolling forecast*, enquanto nos resíduos, é sempre uma única etapa.

Esses indicadores possuem um conceito que também se aplica para mensurar o grau de aderência do modelo aos dados, e portanto, fora da perspectiva futura, de previsão. O conceito e a descrição de alguns indicadores de erro de previsão, que tem uso frequente em modelação, aparecem na sequência deste trabalho, divididos em duas categorias: medidas dependentes de escala e medidas relativas.

2.6.1 Medidas dependentes de escala

As medidas dependentes de escala tratam de valores que apenas são passíveis de interpretação conhecendo-se previamente a ordem de grandeza dos valores. Nessas métricas, os erros e os dados possuem a mesma escala de valores.

- **ME:**

O erro médio (ME) é calculado pela simples média aritmética dos erros medidos em cada um dos instantes. Sua fórmula de cálculo é definida por:

$$ME = \frac{1}{h} \sum_{i=1}^h e_i ,$$

onde h representa o número de observações utilizadas na construção do modelo se o propósito for mensurar o erro dos valores ajustados pelo modelo ou, o comprimento do horizonte de previsão (período futuro), quando deseja-se avaliar os erros de previsão (*forecasting errors*).

- **MAE:**

O indicador erro absoluto médio, conhecido pelo acrônimo MAE (*Mean Absolute Error*), é definido por:

$$MAE = \frac{1}{h} \sum_{i=1}^h |e_i| .$$

É um indicador de erro popular, com fórmula de cálculo simples podendo ser aplicado em uma ou em múltiplas séries, com o propósito comparativo desde que se respeite a mesma escala de valores. De acordo com Hyndman e Athanasopoulos (2018) [13], o método de minimizar o erro de previsão usando o MAE, direciona o modelo à sua mediana. Sua fórmula expressa cálculo similar ao ME com a diferença de que no MAE, há um maior

rigor, uma vez que considera somente valores absolutos impedindo que no cômputo geral que erros, produzidos em diferentes instantes e cujos valores apresentam sinal trocado, se anulem e assim, mitiga-se o risco de decisões equivocadas por “erros aceitáveis”.

- **MSE:**

O erro quadrático médio, comumente designado por MSE (*Mean Squared Error*), definido pela fórmula de cálculo:

$$MSE = \frac{1}{h} \sum_{i=1}^h e_i^2,$$

tem os seus componentes quadráticos e portanto, a exemplo do MAE, também goza da propriedade do indicador nunca tem o seu valor reduzido, considerando a medição em um período.

- **RMSE:**

O erro RMSE (*Root Mean Squared Error*) é definido pela raiz quadrada do erro quadrático médio, ou seja, é calculado a partir da expressão matemática:

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{h} \sum_{i=1}^h e_i^2}.$$

Embora tenha interpretação mais difícil que o MAE, o seu uso é comumente empregado e utilizá-lo como método de minimização do erro de previsão, conduz o modelo ao valor de sua média [13].

2.6.2 Medidas relativas

Essas métricas não têm caráter absoluto, porque tratam-se de medidas relativizadas e apresentam a vantagem de não estarem vinculadas à nenhuma unidade de medida (*unit-free measure*) e assim, permite-se comparações entre diferentes conjuntos de dados independentemente de suas escalas de valores. Por outro lado, existe o aspecto negativo de que seus cálculos nem sempre sejam possíveis de se realizar, no caso do valor de referência, na qual são relativizadas, assumir o valor nulo (denominador da equação). Adicionalmente, esses indicadores estão sujeitos à valores extremos, quando o valor de referência toma valor próximo a zero.

- **MPE:**

O MPE, *Mean Percentage Error* relativiza os erros pela grandeza das observações y_t e é

obtido pelo cálculo:

$$MPE = \frac{100}{h} \sum_{i=1}^h \frac{e_i}{y_i} .$$

- **APE:**

O indicador de erro absoluto percentual (APE) é medido apenas em cada instante t , e não sob a perspectiva de um período de tempo. Sua forma de cálculo é dada por:

$$APE = 100 \left| \frac{e_i}{y_i} \right| .$$

- **MAPE:**

O MAPE, *Mean Absolute Percentage Error*, refere-se à média do indicador APE e portanto, é calculado por:

$$MAPE = \frac{1}{h} APE = \frac{100}{h} \sum_{i=1}^h \left| \frac{e_i}{y_i} \right| .$$

Essa métrica é bastante relevante em modelação, considerado uma boa métrica de erro para avaliar a acurácia (ou grau de precisão) de modelos preditivos, mais conhecido pelo termo *forecast accuracy*, e foi adotado neste estudo. Também existem algumas outras variantes do MAPE a exemplo do WMAPE (*Weighted Mean Absolute Percentage Error*) que ao invés de empregar a tradicional média aritmética do MAPE, baseia-se no cálculo de um MAPE ponderado pelos valores observados y_t (média ponderada dos APE's considerando os valores observados como pesos), tornando o indicador evidentemente mais sensível aos erros ocorridos em instantes cujos valores observados são mais elevados.

Capítulo 3

Aplicação e Resultados

Neste capítulo, apresenta-se o componente prático do estudo, viabilizado com o apoio do *R Software*. As informações mais detalhadas em relação à tecnologia empregada são apresentadas no Apêndice (C).

3.1 Descrição dos dados

Neste trabalho foi utilizado um conjunto de oito séries temporais de ativos financeiros, cujos dados foram extraídos do portal Yahoo Finance e que podem ser acessados por <https://finance.yahoo.com>, referem-se à cotações diárias correspondentes a um período de cinco anos, compreendido entre as datas 01/09/2015 e 31/08/2020 ¹ (ver tabela 3.1).

Categoria	Ativo financeiro	Código (<i>ticker symbol</i>)	Unidade de medida	Bolsa de valores
taxa de câmbio	EUR/BRL	EURBRL	Moeda BRL (preço com atraso)	-
taxa de câmbio	EUR/USD	EURUSD	Moeda USD (preço com atraso)	-
Índice de ações	Dow Jones Industrial Average	DJI	Moeda USD (preço em tempo real)	Nova Iorque, Estados Unidos
Índice de ações	S & P 500	GSPC	Moeda USD (preço em tempo real)	Nova Iorque, Estados Unidos
Índice de ações	NASDAQ Composite	IXIC	Moeda USD (preço em tempo real)	Nova Iorque, Estados Unidos
Índice de ações	ESTX 50 PR.EUR	STOXX50E	Moeda EUR (preço com atraso)	Paris, França
Índice de ações	EURONEXT 100	N100	Moeda EUR (preço com atraso)	Zurique, Suíça
Índice de ações	IBOVESPA	BVSP	Moeda BRL (preço com atraso)	São Paulo, Brasil

Tabela 3.1: Séries financeiras utilizadas no componente prático.

Como o mercado financeiro é dinâmico, operações de compra e venda de ações das empresas ocorrem com muita frequência ao longo de cada dia útil, dentro de horário pré-estabelecido para

¹período válido para todas as séries, com a exceção da taxa de câmbio EUR/BRL, cujos dados são referentes ao período de 01/09/2015 a 31/12/2020 (por estratégia de avaliação do *forecast*).

cada bolsa de valores, e conseqüentemente, existem muitas cotações, a cada instante, dentro do mesmo dia. Adotou-se, para todos os ativos supracitados, a cotação do fechamento do dia, isto é, capturou-se os seus respectivos valores no encerramento do pregão diário.

Este estudo tem ênfase na taxa de câmbio EUR/BRL (EURBRL), isto é, o preço da compra de 1 EUR expresso em unidades da moeda BRL, e na taxa de câmbio EUR/USD (EURUSD), que representa o custo de aquisição de 1 EUR medido em USD, uma vez que estas serão as variáveis protagonistas da modelação, ou seja, a essência do trabalho é explicar o comportamento destas variáveis. Além das séries temporais com estas taxas de câmbio, também foram selecionadas algumas séries do mercado acionário (categoria “Índice de ações”), com o propósito de contribuir na explicação do comportamento das taxas de câmbio. Para este trabalho de investigação, partiu-se de uma premissa inicial de que esses índices ², por representarem composições de empresas atuantes em diferentes segmentos de mercados, e possuírem potencial para relevar alguma importante informação que possa ajudar a descrever a volatilidade das taxas de câmbio.

3.2 Leitura e preparação dos dados

Uma vez coletado os dados, iniciou-se a etapa de *data cleansing*, isto é, o processo pelo qual avalia-se a integridade e a consistência dos dados, e se necessário, realiza-se ações corretivas de modo a se obter uma base de dados devidamente validada e apta para as análises exploratórias. Esta etapa basicamente consistiu na análise de dados faltantes (*missing data*) e na sincronização de todas as séries em relação às datas de referência às observações coletadas. Como os papéis (ativos) abrangem diferentes mercados (europeu, norte-americano e brasileiro) e as bolsas de valores operam em dias úteis, há situações onde não existe cotação de algum destes ativos, normalmente por conta de feriados locais, dia em que a bolsa de valor local não opera. Como este estudo busca investigar as relações de dependência entre as séries, é importante que os dados observados refram-se exatamente aos mesmos instantes de modo que se consiga retratar cada instante de forma integrada, sob diferentes perspectivas. Neste contexto, para a modelação

²A título ilustrativo, o índice ESTX 50 PR.EUR é formado por empresas como Airbus SE, L'Oréal SA., Allianz SE, LVMH Moët Hennessy entre outras, enquanto que no índice EURONEXT 100, há a participação de algumas empresas tais como BNP Paribas SA, Accor SA, EDP Renováveis SA.

da taxa de câmbio EUR/BRL foram desprezados todos os dados coletados nas datas em que havia ao menos uma incidência de dado faltante, finalizando a fase de preparação dos dados. Esse procedimento resultou em um total “líquido” de 1152 observações em cada uma das oito séries, exatamente no mesmo horizonte temporal. Apesar deste procedimento de saneamento dos dados alterar o espaçamento temporal entre as observações e que a rigor, pode ser corrigido por interpolações, optou-se em assumir a premissa que este conjunto conjunto de dados que será objeto de análise e modelação da taxa de câmbio EUR/BRL possui características de séries temporais regulares, isto é, igual espaçamento temporal entre as observações.

Especificamente na modelação da taxa de câmbio EUR/USD, como as demais séries não integraram o escopo do estudo, não faria sentido utilizar os dados em sua versão “reduzida” (com dados desprezados de acordo com o procedimento descrito anteriormente). Dessa forma, trabalhou-se com uma série de maior tamanho, com 1302 observações.

3.3 Análise exploratória dos dados

Como primeiro passo da investigação, é importante uma minuciosa inspeção visual do gráfico da série temporal, de modo a extrair algumas percepções preliminares que auxiliam na identificação das análises mais pertinentes ao problema proposto, e que demanda um exame mais criterioso.

3.3.1 Taxa de câmbio EUR/BRL



Figura 3.1: Taxa de câmbio EUR/BRL

Em relação à série taxa de câmbio EUR/BRL, conforme já esperado, nota-se graficamente que os seus valores exibem um padrão de tendência crescente, a série comporta-se de forma fortemente instável e evidencia-se uma brusca elevação no período final da série (últimos seis meses), relacionada ao surgimento da doença causada pelo coronavírus (COVID-19) que depois espalhou-se globalmente de forma pandêmica. Nota-se também que a barreira do valor cinco (5) (desembolsa-se o valor de cinco unidades de reais (BRL) para aquisição de uma unidade de euro), foi ultrapassada em 04/03/2020, quase em simultâneo com o primeiro caso de infecção COVID-19 registrado no Brasil, ocorrido em 26/02/2020 e com uma certa defasagem ao início da pandemia nos continentes asiático e europeu. É evidente que trata-se de dados não estacionários e as estatísticas descritivas gerais tem pouco significado uma vez que sofrem profundas alterações, de acordo com o período temporal em que se situa, como por exemplo, ao comparar-se o período completo da série (Set/2015 a Ago/2020) com os últimos seis meses (Mar/2020 a Ago/2020), nota-se expressivas variações na mediana (4.23 e 5.98, respectivamente) e no desvio padrão (0.70 e 0.38, respectivamente), conforme mostram as figuras 3.2 e 3.3, referentes, respectivamente, às estatísticas descritivas e aos gráficos *boxplot*:

Descriptive Statistics		Descriptive Statistics	
ts.m.eubr		x	
N: 1152		N: 121	
	ts.m.eubr		x
Mean	4.26379	Mean	5.92808
Std.Dev	0.69676	Std.Dev	0.37637
Min	3.22850	Min	4.92880
Q1	3.73690	Q1	5.68210
Median	4.22570	Median	5.97870
Q3	4.46320	Q3	6.14940
Max	6.63540	Max	6.63540
MAD	0.50920	MAD	0.40090
IQR	0.72580	IQR	0.46730
CV	0.16341	CV	0.06349
Skewness	1.33333	Skewness	-0.33797
SE.Skewness	0.07208	SE.Skewness	0.21999
Kurtosis	1.81855	Kurtosis	-0.23469
N.Valid	1152.00000	N.Valid	121.00000
Pct.Valid	100.00000	Pct.Valid	100.00000

Figura 3.2: Estatísticas descritivas da taxa de câmbio EUR/BRL, referente ao período completo e aos últimos seis meses, respectivamente.

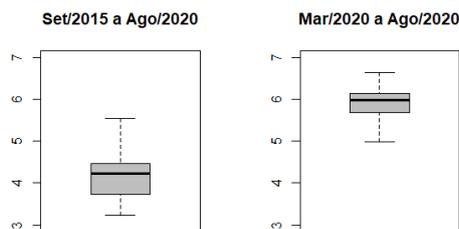


Figura 3.3: *Boxplots* da taxa de câmbio EUR/BRL por período.

Como tradicional na análise de séries temporais, convém sempre examinar o padrão de autocorrelação da série em questão, de modo a avaliar o grau de correlação entre as informações separadas por uma certa defasagem δ , ou seja, determina-se o valor da correlação entre x_t e $x_{t-\delta}$ para diferentes defasagens δ , e exibe-os graficamente por meio de um correlograma, onde também visualiza-se os limites de significância estatística. Esta análise propicia uma ideia prévia de quanto os valores passados influenciam o valor atual e assim, pode-se extrair algumas impressões ou suposições úteis para a etapa futura, no qual elaborar-se-á os modelos preditivos. Observa-se o padrão de uma autocorrelação bastante pronunciada e estatisticamente significativa até aproximadamente à defasagem (*lag*) 300, com decaimento dos valores de forma lenta.

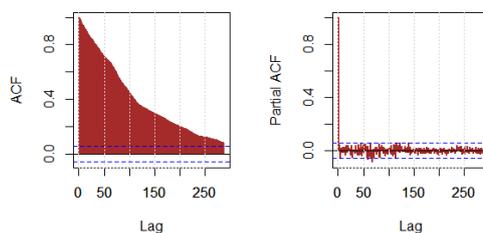


Figura 3.4: ACF/PACF da taxa de câmbio EUR/BRL

Ainda analisando-se a taxa de câmbio EUR/BRL, mas agora na série de retornos³, observa-se visualmente pelo gráfico que há um padrão de comportamento completamente diferente da taxa bruta, de acordo com as expectativas teóricas, com estacionariedade aparente, valor médio

³com os dados transformados com a adoção do conceito do retorno composto continuamente, definido por: $r_t = \log \frac{P_t}{P_{t-1}} = \log P_t - \log P_{t-1} = \nabla \log P_t$.

próximo do zero e aspecto de aleatoriedade presente. Além disso, percebe-se discretamente a presença de períodos mais instáveis caracterizando a apresentação da volatilidade em grupos (ou *clusters*) (ver figura 3.5).

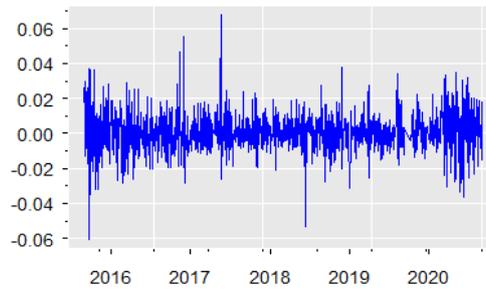


Figura 3.5: Retorno da taxa de câmbio EUR/BRL

Essas impressões iniciais podem ser avaliadas por meio de testes de hipótese. Assim, avaliou-se se a média dos retornos da taxa de câmbio EUR/BRL (o valor da sua média é 0.0004) pode ser assumida como nula por meio da realização do teste de hipótese t (*t-Student*) para uma amostra para confirmar a sua nulidade, assumindo-se a premissa de que os retornos possuem distribuição Normal. A um nível de confiança de 95%, obteve-se um $valor_p = 0.2132$, sugerindo a não rejeição da hipótese testada. Quanto à questão da estacionariedade, procedeu-se o teste de hipótese *Augmented Dickey-Fuller* (ADF), cuja hipótese nula considera a existência de raiz unitária (não estacionária). Aos 95% de nível de confiança, obteve-se uma estatística de teste *Dickey-Fuller* = -9.6827 , testado com o parâmetro $lag = 10$ e um $valor_p < 0.01$ e portanto $valor_p \leq \alpha$). Desta forma, recomenda-se a rejeição da hipótese nula, e por conseguinte, é equivalente a considerá-la como estacionária. Este resultado de estacionariedade foi também ratificado através da realização do teste KPSS dado que a estatística de teste produzida foi $KPSS = 0.46096$ com $lag = 7$ e $valor_p = 0.05088$, cuja conclusão é a não rejeição de H_0 , ou seja, o retorno da taxa de câmbio EUR/BRL é estacionário.

Na análise das estatísticas descritivas do retorno da taxa de câmbio EUR/BRL extrai-se algumas informações interessantes como a constatação que a sua média tem valor próximo de zero ($\bar{x} = 0.0004$) e superior ao quantil 50% que corresponde à sua mediana ($med_x = -0.0001$).

Descriptive Statistics	
m.tb	
N: 1151	
	m.tb
Mean	0.00042
Std. Dev	0.01142
Min	-0.06043
Q1	-0.00621
Median	-0.00014
Q3	0.00664
Max	0.06782
MAD	0.00951
IGR	0.01283
CV	27.23680
Skewness	0.25565
SE. Skewness	0.07211
KURTOSIS	2.98131
N.Valid	1151.00000
Pct.Valid	100.00000

Figura 3.6: Estatísticas descritivas do retorno da taxa de câmbio EUR/BRL

Outra indicação relevante é a presença de uma leve assimetria, representado pelo coeficiente de assimetria k_3 , *Skewness coefficient* na língua inglesa ($k_3 = 0.2556$). A rigor, o coeficiente *Skewness* mede a ausência de simetria, cuja fórmula de cálculo denota a razão entre o momento centrado de terceira ordem (m_3) e o desvio padrão (σ) elevado à terceira potência. Considerando uma variável aleatória X , sua representação matemática é dada por:

$$k_3 = \frac{m_3}{\sigma^3} = \frac{E[(X - \mu)^3]}{\sigma^3}.$$

Também percebe-se que sua distribuição de probabilidade é mais “achatada” que a distribuição Normal uma vez que o coeficiente de curtose ($k_4 = 2.9813$) tem valor superior ao zero (valor do coeficiente de curtose da distribuição Normal). O coeficiente de curtose mede o grau de “achatamento” de uma distribuição e é determinado usando-se o momento de quarta ordem (m_4), de acordo com a equação:

$$k_4 = \frac{m_4}{\sigma^4} = \frac{E[(X - \mu)^4]}{\sigma^4}.$$

Com estes resultados do retorno da taxa de câmbio EUR/BRL, imagina-se que sua distribuição de probabilidade tende a se ajustar melhor com a distribuição *t-Student* comparada à distribuição Normal, em virtude de sua característica de caudas mais pesadas comparada à distribuição Normal, conforme ilustra a figura 3.7, que exibe uma comparação entre a sua distribuição e as funções de densidade de probabilidade da distribuição Normal, de média nula e variância unitária e a distribuição *t-Student* padronizada com três graus de liberdade.

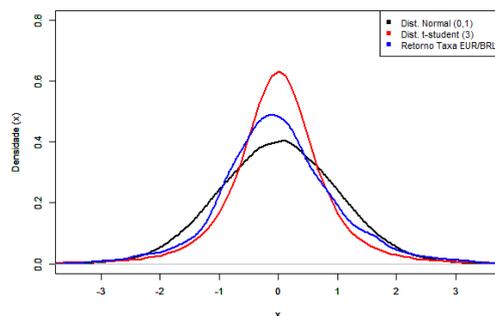


Figura 3.7: Comparativo entre funções de densidade de probabilidade teórica (Normal e *t-Student* padronizada com 3 graus de liberdade) e a densidade do retorno da taxa de câmbio EUR/BRL.

Além desta percepção gráfica, gerou-se também algumas informações básicas para ratificar esta impressão. Assim, se X segue uma distribuição Normal $X \sim N(0, 1)$ então x se concentra de tal modo que $\approx 99\%$ das observações pertencem ao intervalo $x \pm 2.576$, ou dito em outras palavras, a probabilidade de ocorrência de valores de x fora deste intervalo é de $\approx 1\%$. Se X tem distribuição *t-Student* padronizada (média nula e variância unitária) com três graus de liberdade, a probabilidade de ocorrência de valores de x fora deste mesmo intervalo é $\approx 2.20\%$, e portanto, mais provável, comparado à Normal. Ao avaliar a série de retornos da taxa de câmbio EUR/BRL, também padronizada para permitir uma comparação direta, obtém-se uma probabilidade de $\approx 2.09\%$, considerando o evento de ocorrência de valores extremos na mesma situação descrita anteriormente, que traduz-se em um indício de cauda mais pesada, comparada à distribuição Normal. Vale destacar que as observações para estes valores extremos, correspondentes a $\approx 1\%$ da distribuição da retornos da taxa de câmbio EUR/BRL, aparecem de forma mais concentrada na cauda direita $\approx 62\%$ corroborando a ideia de assimetria.

Seu correlograma apresenta correlações significativas em algumas defasagens, embora predominantemente os valores de correlação apresentem valores baixos e dentro da faixa de tolerância. Essas correlações estatisticamente significativas, que mostram a existência de alguma dependência, precisam ser transferidas e explicadas pelos modelos que serão ajustados.

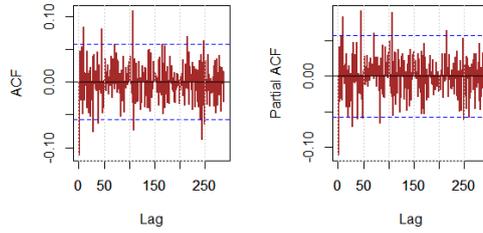


Figura 3.8: ACF/PACF do retorno da taxa de câmbio EUR/BRL

Para a avaliação do grau de relação entre as diferentes séries (taxa de câmbio *versus* índices de ações), utilizou-se o coeficiente de correlação ρ , que de acordo com Feller (1980) [10] tem o seguinte enunciado: considerando X e Y duas variáveis aleatórias com médias μ_x e μ_y e variâncias positivas σ_x^2 e σ_y^2 , o coeficiente de correlação é definido por $\rho_{x,y} = \frac{Cov(X,Y)}{\sigma_x \sigma_y}$ onde $Cov(X, Y) = E[(X - \mu_x)(Y - \mu_y)] = E(XY) - \mu_x \mu_y$.

Embora a correlação capte apenas o grau de relação linear entre duas variáveis, esta medida pode sinalizar alguma associação interessante, motivo pelo qual foi adotada neste estudo, com o intuito de avaliar a “força preditiva” de algumas séries temporais e podem contribuir na explicação do comportamento das séries de taxa de câmbio, que desempenham o papel de variável resposta, e coletar indicações para determinar quais os modelos serão testados. Neste contexto, observou-se que a taxa de câmbio EUR/BRL apresenta uma associação mais forte com os índices de ações da bolsa norte-americana, tais como NASDAQ ($\approx 73\%$) e S & P 500 ($\approx 59\%$).

Série	Coefficiente de correlação
dowjones	0.4655
sp500	0.5918
nasdaq	0.7336
estx50	-0.1439
euronext100	0.0773
ibovespa	0.4558

Tabela 3.2: Correlação entre a taxa de câmbio EUR/BRL e índices de ações.

Como na parte de modelação avaliar-se-á também algumas variáveis externas, todas refe-

rentes a índices de ações, decidiu-se investigar também se os retornos negativos dessas séries exercem maior influência na série de retornos da taxa de câmbio EUR/BRL (*leverage effect*), ou seja, analisar-se-á se há relação mais forte com suas porções negativas e por conseguinte, potencializando a força preditora. Diante desta lógica, construiu-se uma nova série para cada índice de ação, que serão referenciadas neste estudo como “retorno negativo“ e que foram nomeadas no *R Software* pelo sufixo “ret.neg.“ acrescido pelo nome do ativo financeiro, definida por $n_t = \min(r_t, 0)$, onde n_t denota a série de retornos negativos enquanto r_t representa a série de retornos. Assim, foram obtidos os coeficientes de correlação entre a série de retornos da taxa de câmbio EUR/BRL e as séries de retornos e retornos negativos dos índices de ações (ver tabela 3.3), onde percebe-se, que em valores absolutos, o grau de associação com os retornos negativos são sempre mais elevados comparado aos retornos, sendo que essa superioridade de valor é mais acentuada e percebida nos índices de ações Dow Jones e IBOVESPA.

Série	Coefficiente de correlação
ret.dowjones	0.0090
ret.sp500	0.0094
ret.nasdaq	0.0236
ret.estx50	-0.1218
ret.euronext100	-0.1342
ret.ibovespa	0.0001
ret.neg.dowjones	-0.0513
ret.neg.sp500	-0.0506
ret.neg.nasdaq	-0.0310
ret.neg.estx50	-0.1290
ret.neg.euronext100	-0.1368
ret.neg.ibovespa	-0.0448

Tabela 3.3: Correlação entre o retorno da taxa de câmbio EUR/BRL e outras séries.

Investigou-se também, se existia alguma uma relação entre as séries de forma mais pronunciada ao dessincronizá-las no horizonte temporal em algumas defasagens. Este exercício foi realizado para com até três unidades temporais de defasagem (do *lag* -3 ao *lag* 3), na qual mediu-se a correlação cruzada (CCF). O resultado (ver tabela 3.4) mostrou que o “*lag* -1“ apresenta-se mais relevante para todos os índices de ações, a exemplo do IBOVESPA, cuja medição da CCF apontou um valor mais elevado ($\approx 42\%$). A interpretação é de que o re-

torno EUR/BRL no instante t apresenta uma correlação de valor $\approx 42\%$ com o IBOVESPA no instante $t - 1$. Assim, de forma análoga ao procedimento adotado para os retornos negativos, construiu-se também novas séries, designadas pelo prefixo “lag1”, de modo a ampliar as possibilidades, em termos de variáveis explicativas, para a modelação.

Série	Lag -3	Lag -2	Lag -1	Lag 1	Lag 2	Lag 3
ret.dowjones	-0.0502	0.0532	-0.2629	0.0402	-0.0612	0.0125
ret.sp500	-0.0463	0.0491	-0.2715	0.0536	-0.0546	0.0188
ret.nasdaq	-0.0290	0.0336	-0.2515	0.0532	-0.0375	0.0068
ret.estx50	-0.0050	0.0086	-0.2334	0.0533	-0.0324	0.0259
ret.euronext100	0.0018	0.0132	-0.2422	0.0705	-0.0224	0.0197
ret.ibovespa	-0.0805	0.0677	-0.4181	0.0398	-0.0682	0.0265

Tabela 3.4: Correlação cruzada (CCF) entre o retorno da taxa de câmbio EUR/BRL e o retorno dos índices de ações até lag 3.

Além disso, ambas as séries taxa de câmbio EUR/BRL e retorno da taxa de câmbio EUR/BRL foram decompostas pelo método STL, cujo acrônimo tem significado de *Seasonal and Trend decomposition using Loess*⁴ por tratar-se de método robusto para *outliers* [13] e por meio de inspeção de seus gráficos (ver figuras 3.9 e 3.10), onde observou-se um padrão razoavelmente semelhante, com a ausência do componente sazonalidade e com a tendência revelando um aspecto ascendente, com notória elevação no período final da série (últimos seis meses).

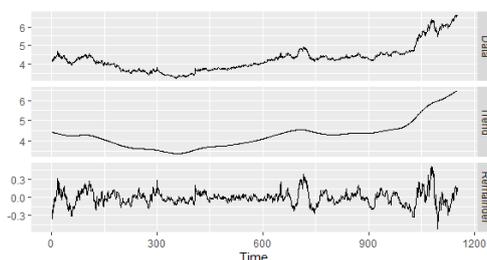


Figura 3.9: Decomposição da taxa de câmbio EUR/BRL pelo método STL (“ $x_t = T_t + R_t$ ” assumindo independência entre os seus componentes).

⁴Loess resume-se em um método de estimação de relações não lineares.

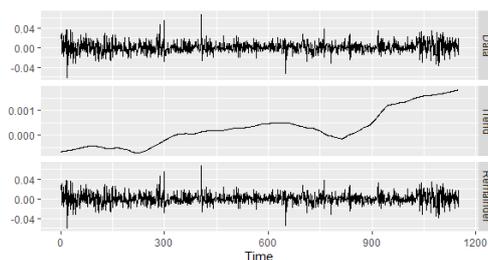


Figura 3.10: Decomposição do retorno da taxa de câmbio EUR/BRL pelo método STL. (“ $r_t = T_t + R_t$ ” assumindo independência entre os seus componentes).

Um outro ponto da análise exploratória que é sensível e relevante em muitos estudos estatísticos, é a análise dos *outliers*, mesmo que nem sempre seja abordado em profundidade. O *outlier* pode ser compreendido pela sua definição intuitiva dada por Hawkins (1980) [12], onde é referido como uma observação que se desvia tanto das outras observações e que levanta suspeitas de que tenha sido gerado por um mecanismo diferente. Embora inicialmente essas informações sejam pouco prováveis e possam parecer erros, levando o investigador a ficar tentado a removê-las, convém ser cuidadoso pois podem ser informações reais por se tratar de observações influenciadoras e que podem representar informações críticas, perante o qual se recomenda uma investigação mais criteriosa.

Não há um método unanimemente aceito para identificação dos *outliers*. Neste estudo, a escolha foi identificá-los pelo método de Tukey (1977) [24]⁵, cuja essência pode ser descrita em quatro etapas:

- (1) Calculam-se os quartis Q_1 (quantil 25%) e Q_3 (quantil 75%);
- (2) Calcula-se a métrica IQR , definido por $IQR = Q_3 - Q_1$, cujo significado é a medida de distância entre os quartis;
- (3) Constrói-se um intervalo I definido por:

$$I = \left[Q_1 - k IQR, Q_3 + k IQR \right], \quad (3.1)$$

sendo k uma constante⁶ onde $k \in \mathbb{R} \mid c > 0$, que caracteriza a observação como “normal”, desde que o seu valor não ultrapasse os limites deste intervalo;

- (4) Classifica-se como *outliers* todas as observações da série x_t , onde t é o índice temporal

⁵método famoso e conhecido por ser ter sido empregado no gráfico *Boxplot*.

⁶no caso do gráfico *Boxplot*, esta contante k assume o valor 1.5.

$t = 1, 2, 3, \dots$, cujos valores aparecem fora do intervalo I , isto é, se $x_t \notin I$, então x_t é considerado um *outlier*.

No entanto, optou-se por um nível de severidade mais elevado, com $k = 2.5$, de modo a restringir ainda mais a incidência dos *outliers*. A partir de um intervalo I mais amplo, para ser classificada como *outlier*, a observação precisa mostrar um valor ainda mais extremo, comparada à situação usual, onde $k = 1.5$.

Dessa maneira, obteve-se para a série taxa de câmbio EUR/BRL,

$$IQR = 0.7258 ; I = [1.9226, 6.2774],$$

sendo IQR a distância entre quartis e I , o intervalo que caracteriza as observações “normais” (não extremas). Este procedimento resultou na identificação de um total de 23 *outliers*, que representa $\approx 2\%$ das 1152 observações. Nota-se que dentre essas informações apontadas como “anômalas”, há uma forte concentração ($\frac{17}{23} \approx 74\%$) de observações no mês de agosto/2020, que é o período final da série, e as restantes, referem-se à maio/2020, que também integra o período dos últimos seis meses.

A estratégia foi considerá-los na modelação e verificar adicionalmente se em um outro cenário, considerando a remoção dos *outliers*, se haveria conclusões diferentes, ou seja, se os *outliers* exercem alguma influência nos resultados.

Para a identificação dos *outliers* na série de retornos da taxa de câmbio EUR/BRL, repetiu-se o mesmo procedimentos (método de Tukey e parâmetro $k = 2.5$), calculou-se o IQR e o intervalo das observações “normais” I , e obteve-se:

$$IQR = 0.01283 ; I = [-0.03827, 0.03872].$$

Por tratar-se de uma série estacionária e menos volátil, o número de ocorrência de *outliers* foi drasticamente reduzido. Em um total de 1151 observações ⁷, foram identificados apenas 5 *outliers*, que representa $\approx 0.42\%$, não localizados em algum período de tempo específicos (espalhados no horizonte temporal). Salienta-se que em todos estes casos de *outliers*, coincidentemente tratam-se de valores picos de retornos positivos.

⁷número reduzido em uma observação se comparada à série original em função da transformação em retorno financeiro, dado que envolve uma única diferenciação da série.

3.3.2 Taxa de câmbio EUR/USD

A série da taxa de câmbio EUR/USD apresenta características típicas de séries financeiras como um comportamento não estacionário e períodos de instabilidade (ver figura 3.11).

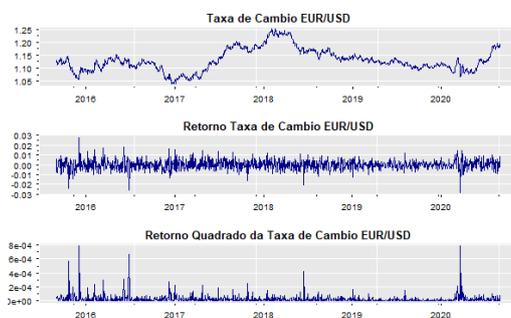


Figura 3.11: Taxa de câmbio EUR/USD

Transformando-a em série de retornos, usando o conceito do retorno composto continuamente, calculado de acordo com a equação 2.1, nota-se um aparente padrão de estacionariedade (ver figura 3.11). Esta impressão inicial foi investigada por meio da realização do teste de hipótese ADF, com nível de confiança $(1 - \alpha)$ de 95% e parâmetro $lag = 10$, que resultou em uma estatística de teste *Dickey – Fuller* = -11.922 e um *valor_p* < 0.01 . Desta forma, como o *valor_p* $\leq \alpha$, sugere-se a rejeição da hipótese nula (H_0 : não estacionária), que dito em outras palavras, ratifica a sua condição de estacionariedade.

Com a perspectiva de realizar a modelação desta série de retornos com modelo GARCH, pertencente à classe de heterocedasticidade condicional, é interessante também avaliar os seus retornos quadrados, uma vez que eles podem indicar a presença de heterocedasticidade, chamado de efeitos ARCH. Esta percepção pode ser extraída ao se observar que há algum tipo de relação entre os maiores valores do retorno quadrado (picos) e os períodos de maior variabilidade, isto é, se a ocorrência dos picos é nos períodos mais instáveis, e portanto relacionado à alta volatilidade (ver figura 3.11, gráfico do "retorno quadrado").

Examinando-se algumas estatísticas descritivas (ver figura 3.12), nota-se uma média praticamente nula ($\bar{x} = 0.00001$) com valor muito próximo à sua mediana ($med_x = -0.00001$), revelando assim medidas de tendência central quase alinhadas. Em termos de assimetria, percebe-se que a medida do *Skewness coefficient* é negativa, seu valor é pouco expressivo, próximo do zero

($k_3 = -0.07748$), fornecendo indícios de uma distribuição simétrica. Também percebe-se que sua distribuição de probabilidade é mais “achatada” que a distribuição Normal uma vez que o coeficiente de curtose ($k_4 = 3.39241$).

```

Descriptive Statistics
m.eu2
N: 1301

-----+-----
                m.eu2
-----+-----
      Mean      0.00005
    Std.Dev    0.00470
       Min    -0.02814
        Q1    -0.00281
     Median    0.00001
        Q3     0.00276
        Max     0.02815
         MAD     0.00414
         IQR     0.00557
          CV    102.21603
    Skewness   -0.07748
SE.Skewness    0.06783
    Kurtosis    3.39241
     N.Valid    1301.00000
    Pct.Valid   100.00000
    
```

Figura 3.12: Estatísticas descritivas do retorno da taxa de câmbio EUR/USD.

Também investigou-se a presença de *outliers* na série, pelo método de Tukey parametrizado com uma constante $k = 2.5$, que resultou:

$$IQR = 0.00557 ; I = [- 0.01673, 0.01667],$$

onde *IQR* representa a distância entre os quartis enquanto *I*, o intervalo das observações “normais”, calculado conforme a expressão 3.1. Em um total de 1301 observações, foram detectados 7 *outliers*, o que representa $\approx 0.5\%$, sendo o mais recente deles ocorrido em 19/03/2020 na forma de retorno negativo ($ret \approx -0.028144$). Este foi um momento em que a pandemia do COVID-19 ainda estava em sua fase inicial mas já ganhava força gerando muitas incertezas para o mercado financeiro. Nesta data, a cotação da taxa de câmbio EUR/USD era ≈ 1.0962 e encontrava-se em período de baixa, ou seja, o EUR havia se desvalorizado em relação à moeda norte-americana USD.

A série apresenta aspecto de não autocorrelacionada (ver figura 3.13). Esta hipótese foi testada através da execução do teste de hipótese Ljung-Box cujo resultado aparente se confirmou. A um nível de confiança $(1 - \alpha)$ de 95%, obteve-se uma estatística de teste $\chi_1^2 = 0.13913$ a um grau de liberdade e o *valor-p* = 0.7091 > $\alpha = 5\%$, indicando o diagnóstico de não rejeitar a hipótese nula ($H_0:iid$) testada.

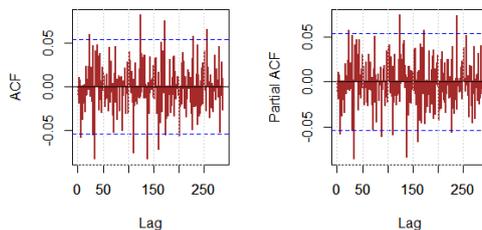


Figura 3.13: ACF/PACF do retorno da taxa de câmbio EUR/USD

Com o objetivo de coletar um *insight* da distribuição de probabilidade que tende a se ajustar melhor ao retorno da taxa de câmbio EUR/USD, repetiu-se o mesmo exercício realizado para o retorno da taxa de câmbio EUR/BRL, onde a sua função de densidade, obtida pelo método não paramétrico de Kernel, é comparada às funções de densidade da distribuição Normal, de média nula e variância unitária e *t-Student* padronizada com três graus de liberdade (ver figura 3.14). Embora tenha uma curtose mais elevada que a Normal, a inspeção visual sugere uma melhor aderência à distribuição Normal. Esta impressão foi investigada através da realização do teste de hipótese Shapiro-Wilk, que é baseado no cálculo da correlação entre os dados testados e os quantis correspondentes na distribuição Normal, onde não se verificou a normalidade do retorno da taxa de câmbio EUR/USD (hipótese nula de normalidade rejeitada, considerando um nível de significância $\alpha = 5\%$, pois o teste produziu uma estatística de teste $W = 0.97142$ e o *valor-p* $= 2.31 \times 10^{-15} \leq \alpha$).

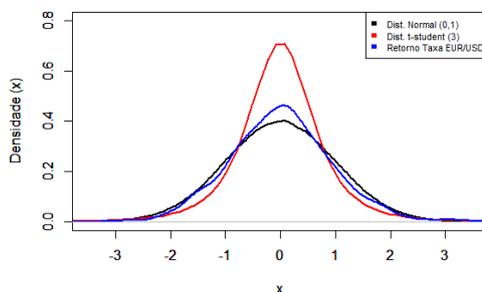


Figura 3.14: Comparativo entre funções de densidade de probabilidade teórica (Normal e *t-Student* padronizada com 3 graus de liberdade) e a densidade do retorno da taxa de câmbio EUR/USD.

3.4 Modelação

Depois do conhecimento preliminar adquirido na etapa de análise exploratória dos dados, para as duas séries estudadas: taxa de câmbio EUR/BRL e taxa de câmbio EUR/USD, inicia-se a fase de modelação, propriamente dita, pautado na busca de um modelo fidedigno, isto é, um modelo capaz de reproduzir um padrão dos dados observados de forma aderente. Para este processo do estudo foi adotado uma metodologia que contempla a abordagem de Box e Jenkins (1970) [2], proposta originalmente elaborada para a construção de um modelo ARIMA, no entanto, o mesmo foi adaptado e estendido para outros outros modelos aplicados neste estudo.

Esta metodologia está ilustrada por um fluxograma, construído para este estudo, de modo a elucidar as atividades de cada etapa do processo de modelação, além fornecer uma ideia sequencial do caminho a percorrer, incluindo as decisões, que podem determinar novas direções no fluxo de atividades. (ver figura 3.15).

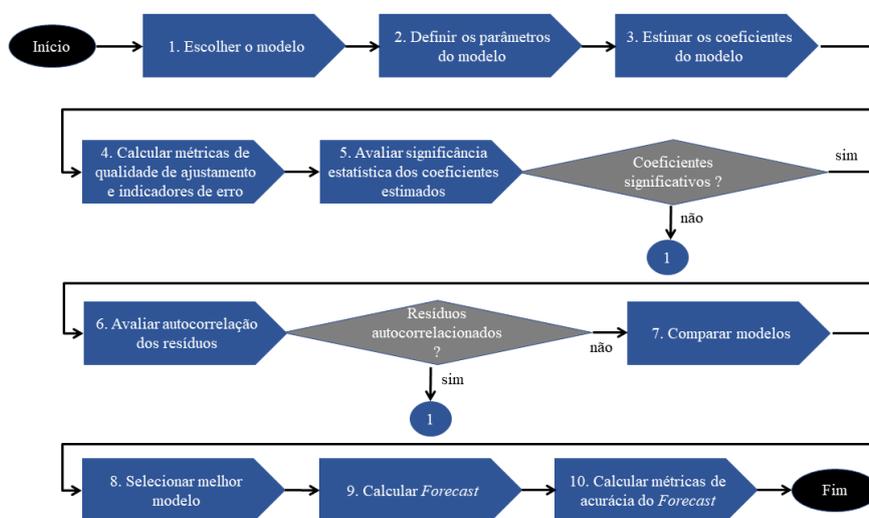


Figura 3.15: Fluxograma do processo de modelação adotado.

Para fins didáticos, esse fluxograma pode ser explicado por um exemplo hipotético: escolha-se um modelo GARCH (etapa 1); de ordem $(p, q) = (1, 1)$ com distribuição Normal dos erros (etapa 2); estima-se os coeficientes do modelo: ω , α_1 e β_1 pelo método de máxima verossimilhança, (etapa 3); calculam-se as métricas de qualidade de ajustamento (critérios de informação e coeficiente de determinação ajustado)(etapa 4); examina-se, por meio de testes de hipótese,

a significância estatística dos coeficientes ω , α_1 e β_1 estimados, a um nível α pré-definido de significância (etapa 5); avalia-se a independência dos resíduos (etapa 6); dentre os modelos candidatos (com coeficientes estatisticamente significativos e com resíduos não autocorrelacionados), compara-se os modelos pelos critérios de informação AIC/BIC (etapa 7); seleciona-se o modelo mais favorável (etapa 8); estima-se os valores previstos de retorno para um período futuro (etapa 9) e calcula-se o nível de acurácia do *forecast*, dividindo-se os dados em dois conjuntos: treino, para a elaboração do modelo e teste, para avaliar o nível de erro, ou o nível de acurácia, alcançado pelo modelo (etapa 10).

3.4.1 Taxa de câmbio EUR/BRL

No processo de modelação da taxa de câmbio EUR/BRL, foram exploradas variadas abordagens em duas classes: modelos homocedásticos e modelos heterocedásticos.

3.4.1.1 Modelo homocedástico

Resgatando-se algumas descobertas relevantes extraídas da análise exploratória (3.3.1), sabe-se que a taxa de câmbio EUR/BRL exibe uma autocorrelação significativa em várias defasagens além de apresentar um padrão de alta instabilidade. Embora haja evidências de características adversas, o processo de modelação iniciou-se com a suposição de variância constante, ou seja, de não haver mudanças na variabilidade das observações ao longo de todo o horizonte temporal da série. Nesta perspectiva, buscou-se identificar um modelo de regressão dinâmica, que neste estudo teve a configuração de um modelo ARIMA aplicado aos resíduos de um modelo de regressão linear múltipla. Para o modelo de regressão, os diferentes índices de ações e interações ⁸ entre eles assumiram o papel de potenciais variáveis explicativas, testados em diferentes modelos obedecendo uma sequência decrescente dos valores absolutos de seus coeficientes de correlação relacionados à série taxa de câmbio EUR/BRL.

Como tratou-se de um cenário de múltiplas variáveis, foi necessário definir uma estratégia de

⁸No modelo de regressão linear $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon$, os coeficientes β_i são os efeitos principais associados às variáveis explicativas X_i e representam a influência de cada variável explicativa sobre a variável resposta Y quando o modelo está controlado pelas demais variáveis explicativas. Quando existe uma interação, por exemplo entre as variáveis X_1 e X_2 e representada por $X_1 * X_2$, o efeito de qualquer uma das variáveis, X_1 ou X_2 , passa agora a depender dos valores da outra variável.

seleção de variáveis, onde optou-se pelo procedimento *stepwise selection*, que envolve inclusão e exclusão de variáveis e a cada passo, avalia-se a contribuição proporcionada pela adição de uma nova variável. Em caso afirmativo, retém-se essa nova variável e testa-se individualmente a significância estatística das demais variáveis explicativas e o ganho de qualidade do modelo. Este método, idealmente, identifica o menor conjunto de variáveis explicativas a ser considerado, além de ser mais sensível ao fenômeno da multicolinearidade, que consiste na existência de associações lineares entre as variáveis explicativas e conseqüentemente, redundância de informação fornecida pelas variáveis, acarretando problemas como instabilidade na inferência estatística. Neste estudo, a verificação da presença de multicolinearidade no modelo de regressão foi baseada na métrica VIF (*Variance Inflation Factor*), cuja definição é $VIF(X_i) = \frac{1}{1-R_i^2}$, onde X_i representa as variáveis explicativas X_1, X_2, \dots , R_i^2 é o coeficiente de determinação associado à regressão de X_i sobre as restantes variáveis explicativas. Cada variável X_i tem um valor de VIF associado, sendo que fortes relações lineares (R^2 com valor próximo de 1) implica em altos valores do VIF. A regra aplicada foi $VIF(X_i) < 10, \forall i$, ou seja, uma vez verificada esta condição, assume-se que não há efeitos de multicolinearidade.

Dizer que um coeficiente é estatisticamente significativo é equivalente a dizer que a variável correspondente produz efeitos significativos ao modelo, porque as alterações no valor da variável preditora estão relacionadas à alterações na variável resposta. Este diagnóstico ocorre por meio da realização de teste de hipótese, cuja hipótese nula H_0 , a ser testada, é definida para um coeficiente nulo, isto é, não produz efeitos.

A ideia central consiste em buscar o modelo mais simples porém com o maior poder de explicação, ou seja, novas variáveis só serão incorporadas ao modelo, desde que consigam revelar alguma informação adicional relevante.

O modelo de regressão linear foi obtido após a comparação entre os modelos testados, cujas variáveis explicativas mostraram-se estatisticamente significativas, usando o coeficiente de determinação ajustado (\bar{R}^2), como decisor do melhor ajuste alcançado, sendo que mais valores mais elevados tem a interpretação de que o modelo tem maior aderência aos dados. Esta modelação por **regressão linear múltipla**, onde a taxa de câmbio EUR/BRL é representada por Y_t (variável resposta) produziu o modelo descrito descrito por:

$$\begin{aligned}
 Y_t = & 9.39154392 + 0.00065966(nasdaq_t) - 0.00239145(estx50_t) \\
 & - 0.00008007(ibovespa_t) + 0.00000002 (estx50_t)(ibovespa_t) + \eta_t \quad (3.2)
 \end{aligned}$$

Este modelo resultou nos valores 345.2289 e 375.4984 para os critérios de informação AIC e BIC, respectivamente, além de um coeficiente de determinação ajustado de $\bar{R}^2 \approx 85\%$, cujos resíduos mostraram-se fortemente autocorrelacionados, confirmado pelo teste de hipótese Ljung-Box que indicou a rejeição da hipótese nula de resíduos iid a um nível de confiança $(1 - \alpha)$ de 95%, dado que resultou em estatística de teste $\chi^2 = 1064.4$ e $valor_p = 2.21 \times 10^{-16} < \alpha = 5\%$.

Este resultado, em termos de resíduos, não é um problema no caso de uma modelação usando regressão dinâmica, uma vez que o processo tem prosseguimento na identificação de seu segundo componente, onde tentar-se-á explicar o comportamento desses resíduos gerados pela regressão, através de um modelo ARIMA. Assim, preservando-se todos os preditores (variáveis explicativas) do modelo de regressão, testou-se diversos modelos da classe ARIMA variando-se apenas os parâmetros “p” (ordem do modelo autoregressivo), orientado também pelo PACF, e “q” (ordem do modelo de médias móveis) da tripla (p,d,q) referente à ordem do modelo ARIMA, onde “d” é o parâmetro que indica o número de vezes que a série é diferenciada. Testou-se modelos de diferentes ordens além do modelo ARIMA automático ⁹ no *R Software*, com o parâmetro “d” fixo ($d = 1$). Para seleção do modelo de melhor ajuste, comparou-se os critérios de informação e o modelo ARIMA (3,1,2), apresentou melhor desempenho (valores dos critérios de informação mais baixos), e produziu $AIC = -3601.81$, $AICc = -3601.62$ e $BIC = -3551.33$. Embora este modelo tenha apresentado resíduos iid, comprovado por meio da realização de um teste de hipótese Ljung-Box que indicou não rejeição de H_0 pois produziu uma estatística de teste de $\chi^2_1 = 0.0056727$, com um grau de liberdade e $valor_p = 0.94$ a um nível de significância $\alpha = 5\%$, a variável resposta, representada neste caso pela taxa de câmbio EUR/BRL, exhibe um comportamento bastante instável em termos de variabilidade, motivo pelo qual decidiu-se testar uma outra versão do modelo, utilizando a transformação Box-Cox.

Nestas situações onde os dados exibem diferentes níveis de variabilidade pode-se realizar uma transformação matemática dos dados, com o objetivo de estabilizar a sua variância, a exemplo

⁹instrução `auto.arima` do pacote `package forecast`

da transformação Box-Cox, pertencente à classe das transformações de potência, conforme destaca Hyndman e Athanasopoulos (2018) [13]. Matematicamente, a série com as observações transformadas por Box-Cox, representada por w_t , é definida por:

$$w_t = \begin{cases} \log(y_t), & \text{se } \lambda = 0; \\ \frac{y_t^\lambda - 1}{\lambda}, & \text{se } \lambda \neq 0. \end{cases}, \text{ onde } y_t \text{ refere-se à série com as observações originais.}$$

Calculando-se λ de maneira a minimizar o coeficiente de variação (CV), que é uma medida de dispersão definida pelo desvio padrão relativizada pela média, isto é, $CV = \frac{\sigma}{\mu}$, da variável resposta taxa de câmbio EUR/BRL, obteve-se um valor $\lambda \approx -0.5761$. A partir desse λ definido, elaborou-se um novo modelo usando a transformação Box-Cox e preservando-se todos os preditores do modelo anterior, resultando em um modelo $ARIMA(1, 1, 4)$. Esse novo modelo apresentou resultados bastante mais favoráveis comparado ao modelo anterior (aplicado à série original sem transformação dos dados).

Na perspectiva dos critérios de informação, que quantificam a informação do modelo, os resultados alcançados foram: $AIC = -9025.28$ (*versus* -3601.81 do modelo anterior), $AICc = -9025.08$ (-3601.62) e $BIC = -8974.79$ (-3551.33). A mesma condição foi também observada no indicador MAPE, que apresentou uma melhora discreta, e exibe um baixíssimo nível de erro e portanto, um excelente resultado: $MAPE = 0.008192791 \approx 0.819\%$ (*versus* $\approx 0.820\%$ do modelo anterior).

Desta forma, a taxa de câmbio EUR/BRL, representada por Y_t , é descrito pelo **modelo de regressão dinâmica**, com transformação Box-Cox de parâmetro $\lambda \approx -0.5761$, em duas partes (I e II), pelas seguintes equações:

(I) modelo de regressão linear múltipla, para explicar Y_t , com resíduos η_t :

$$Y_t = 9.39154392 + 0.00065966(nasdaq_t) - 0.00239145(estx50_t) - 0.00008007(ibovespa_t) + 0.00000002 (estx50_t)(ibovespa_t) + \eta_t \quad (3.3)$$

(II) modelo $ARIMA(1, 1, 4)$, para explicar η_t , com resíduos ε_t :

$$\eta_t = 0.2001455736 \eta_{t-1} - 0.3517124098 \varepsilon_{t-1} + 0.0162259142 \varepsilon_{t-2} + 0.0429152304 \varepsilon_{t-3} - 0.0724210480 \varepsilon_{t-4} + \varepsilon_t \quad (3.4)$$

As figuras 3.16 e 3.17 evidenciam a diferença entre o comportamento dos resíduos do modelo de regressão (η_t), que carrega um padrão de não aleatoriedade (resíduos autocorrelacionados) não capturado e aceito pelo modelo, de acordo com a sua premissa inicial e os resíduos do modelo ARIMA (ε_t), que comportam-se como iid. Para fins de validação deste modelo, a ausência de autocorrelação significativa foi ratificado através do teste de hipótese Ljung-Box, com $\alpha = 5\%$, cujo resultado gerou uma estatística de teste de $\chi_1^2 = 0.00051228$, com um grau de liberdade e o *valor_p* = 0.9819 (*valor_p* $\not\leq 5\%$), reforçando a recomendação da não rejeição da hipótese nula de resíduos independentes.

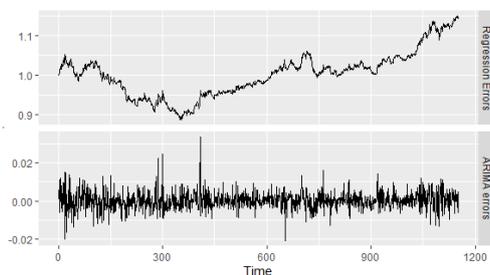


Figura 3.16: Comparativo dos resíduos entre os modelos de regressão (η_t) e regressão dinâmica ajustado por ARIMA (ε_t) para a taxa de câmbio EUR/BRL.

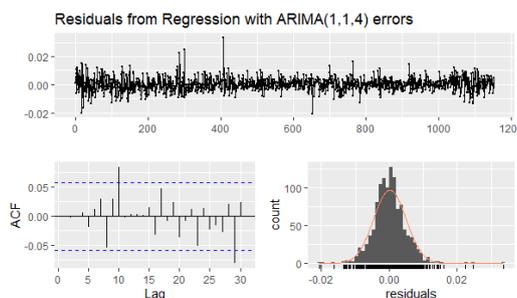


Figura 3.17: Resíduos do modelo de regressão dinâmica para a taxa de câmbio EUR/BRL.

Através da inspeção do gráfico quantil-quantil (*QQ Plot*), que compara a distribuição dos resíduos teóricos (com distribuição Normal) com a distribuição dos resíduos obtidos pelo modelo (resíduos observados), consegue-se avaliar o grau de ajustamento da distribuição de probabilidade atingido. Nota-se que existe um desajuste nas caudas (nos valores extremos, os pontos estão mais distantes da reta de referência), e de forma mais pronunciada, na cauda superior, o que corresponde a dizer que este modelo não consegue capturar apropriadamente os movimen-

tos de elevados valores da taxa de câmbio EUR/BRL. Ou seja, o modelo falha principalmente nos instantes de alta instabilidade, dado que baseia-se em premissa de variância constante.

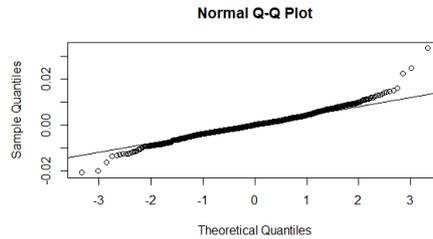


Figura 3.18: Comparação dos quantis, empírico *versus* teórico, obtido pelo modelo de Regressão Dinâmica para a taxa de câmbio EUR/BRL (gráfico quantil-quantil \equiv *QQ Plot*).

Antes do cálculo efetivo do *forecast*, foi idealizada a estratégia de como o mesmo seria avaliado. O indicador MAPE, que tinha sido utilizado para este modelo apenas no sentido de avaliar a qualidade de ajuste alcançada (mensura o nível de erro entre os valores observados e ajustados), teve a sua participação estendida para esta perspectiva “futura”, com o propósito de avaliar também a qualidade das previsões proporcionadas pelo modelo. Nesta direção, optou-se por coletar novos dados da série de taxa de câmbio EUR/BRL, de modo a conhecer os reais valores observados “futuros”. Estes novos dados não foram utilizados na construção do modelo, e portanto, este procedimento reproduz exatamente a ideia do método de avaliação de modelos, baseado na divisão dos dados em treino, utilizado para a modelação, e teste, usado unicamente para avaliar o modelo.

Este novo conjunto de dados, extraído da mesma fonte ¹⁰, refere-se às novas observações referente ao período de 01/09/2020 a 31/12/2020 e que obedecem todas as mesmas propriedades, já citados anteriormente.

Calculou-se também o *forecast*, medido em pontos para os primeiros instantes futuros, considerando um horizonte de previsão de curto prazo, pelas características da série financeira utilizada, especificamente de 6 dias (base de teste). Este processo de cálculo do *forecast* a partir da equação do modelo é um processo iterativo, onde em cada iteração determina-se as estimativas para um único instante “futuro”, respeitando-se a lógica de que as observações futuras serão substituídas por suas respectivas previsões (*forecast*), os erros futuros serão con-

¹⁰url: <https://finance.yahoo.com>

siderados nulos e os erros passados corresponderão aos seus resíduos.

Este *forecast* foi calculado em seu ponto médio, bem como nos pontos referente aos limites inferior e superior dos intervalos de confiança a que foram associados, com níveis de confiança fixados em 80% e 95%, de modo a clarificar a região de valores mais prováveis.

A figura 3.19 fornece uma ideia de como este modelo de *Forecasting* comporta-se na linha temporal, uma vez que reúne graficamente os valores observados e valores previstos, incluindo os limites de seus intervalos de confiança, sendo que o intervalo de confiança de 80% está representado em azul escuro, enquanto que o de 95%, em azul claro. É bastante evidente que o período que finda os valores observados (agosto/2020) mostrou-se bastante turbulento, com forte desvalorização da moeda BRL frente ao EUR, e o modelo, que possui característica conservadora e não é sensível à momentos de instabilidade, aparentemente reflete um panorama mais factível, em termos de *forecast*, apenas na zona próxima aos limites superiores de seus dois intervalos de confiança e ainda assim, a barreira do intervalo de 80% é ultrapassada em muitos instantes, devido à forte alta da taxa de câmbio EUR/BRL no período que o *forecast* foi avaliado.

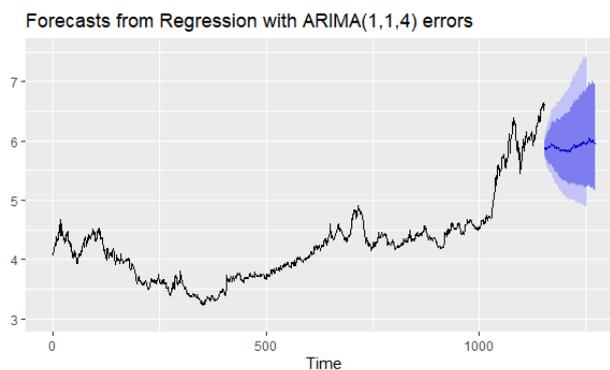


Figura 3.19: *Forecast*, para um período de 6 dias, obtido por modelo de Regressão Dinâmica com $ARIMA(1, 1, 4)$ (valor médio e intervalos de confiança a 80% e 95% de probabilidade).

Com esses novos dados coletados, foi viabilizada a avaliação da acurácia do *forecast* através do cálculo do indicador MAPE. Este procedimento foi realizado para períodos acumulados, especificamente, período de 1 dia, de 1 a 2 dias, ..., até o período de 1 a 6 dias, cujos resultados são mostrados na figura 3.20 e tabela 3.5. Percebe-se que o nível de erro pelo *forecast* médio, reside na faixa de 8% a 10%, e é mais acentuado comparado com as previsões dos limites superiores do

intervalos de confianças, pois sua avaliação foi realizada em período de alta. Também nota-se que, embora os níveis de erros sejam menores para as previsões do limite superior do intervalo de confiança de 95%, ainda assim, o modelo falha neste período de avaliação por corresponder à cauda superior da distribuição da taxa de câmbio de EUR/BRL, ou seja, quando ocorre valores extremos e pouco prováveis. A taxa de câmbio real excedeu o limite superior do intervalo de confiança de 95% em todos os 6 dias avaliados (ver figura 3.20).

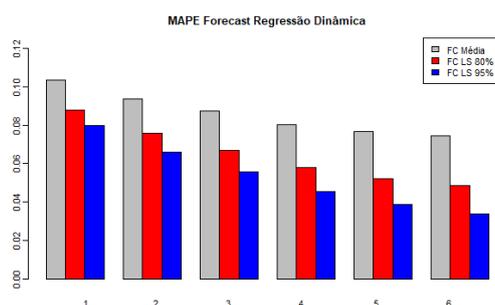


Figura 3.20: Indicador MAPE do *forecast*, obtido por modelo de Regressão Dinâmica para a taxa de câmbio EUR/BRL, considerando os valores médio e os limites superior (LS) dos intervalos de confiança de 80% e 95% para período acumulado de 1 a 6 dias.

Período (t) em dias	FC Médio	FC LS 80%	FC LS 95%
1	0.1036	0.0881	0.0797
[1, 2]	0.0940	0.0758	0.0659
[1, 3]	0.0874	0.0670	0.0558
[1, 4]	0.0805	0.0578	0.0454
[1, 5]	0.0768	0.0522	0.0388
[1, 6]	0.0747	0.0484	0.0340

Tabela 3.5: Indicador MAPE do *forecast*, obtido por modelo de Regressão Dinâmica para a taxa de câmbio EUR/BRL, considerando os valores médio e os limites superior (LS) dos intervalos de confiança de 80% e 95% para período acumulado de 1 a 6 dias.

Como comentário adicional, a maior taxa de câmbio EUR/BRL registrada no período em que os dados foram modelados era de 6.635 em 27/08/2020, justamente no período final da série (30/09/2020). Considerando o complemento de mais quatro meses (01/09/2020 a 31/12/2020), esse valor máximo foi ultrapassado em 04/11/2020, onde a taxa bateu um novo recorde com um valor de 6.7662, evidenciando o grande desafio de se construir um bom modelo preditivo para esta série.

3.4.1.2 Modelo heterocedástico

Após realizada a modelação da taxa de câmbio EUR/BRL através de modelos fundamentados na premissa de variância constante, explorou-se os modelos da classe de heterocedasticidade condicional, essencialmente usando-se modelos do tipo GARCH, de modo a compreender o comportamento da volatilidade desta taxa de câmbio, que é um cenário típico quando trata-se de dados de um ativo financeiro. Zivot e Wang (2003) [27] considera um procedimento usual e recomendado como “boa prática“ a investigação inicial da presença dos efeitos ARCH (heterocedasticidade), isto é, aplicar técnicas de inferência estatística aos dados, mais especificamente testes de hipótese, para avaliar a presença autocorrelação significativa, em diferentes defasagens de tempo (*lags*), nos resíduos ¹¹ quadrados. Este procedimento pode ser realizado tanto através pelo teste de hipótese LM (*Lagrange Multiplier*) quanto pelo teste Ljung-Box, considerando o quadrado da série testada.

Como é usual na modelação de séries financeiras, escolheu-se modelar o retorno financeiro da série da taxa de câmbio EUR/BRL, que é uma série estacionária e que possui média próxima de zero. Se a sua média pode ser assumida como nula, modelar os resíduos, em relação à sua média, é o mesmo que modelar os próprios retornos. De fato, esta premissa de média nula pode ser adotada a partir do resultado do teste de hipótese *t-Student*, assumindo a premissa de que os retornos tem distribuição Normal, para a hipótese testada de média nula ($H_0 : \mu = 0$), que indicou, considerando $\alpha = 5\%$, uma estatística de teste $t = 1.2456$ e o *valor_p* = 0.2132, superior ao nível do α utilizado ($0.2132 > 5\%$), sugerindo a não rejeição de H_0 .

Como a intenção é de construir um modelo GARCH de ordem baixa (parâmetros P e Q assumem baixos valores), o exame da autocorrelação dos retornos quadrados foi considerado até 20 defasagens, intui-se visualmente, por análise gráfica, a presença de efeitos ARCH (ver figura 3.21), dado que para várias defasagens, a exemplo dos seis primeiros *lags*, a autocorrelação do retorno quadrado apresenta-se como significativa (fora da faixa de tolerância).

¹¹resíduos calculados em relação a um processo médio.

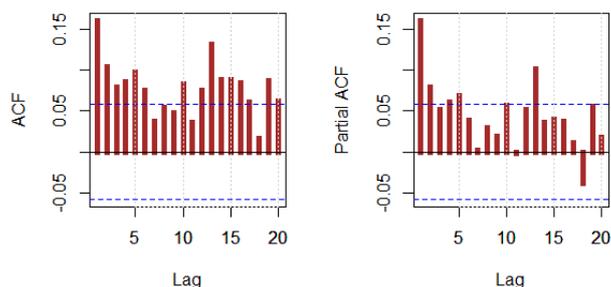


Figura 3.21: ACF/PACF do retorno quadrado da taxa de câmbio EUR/BRL (até lag 20).

Além disso, a figura 3.22 fornece informação complementar às evidências mostradas 3.21 sugerindo a existência de efeitos ARCH. Nota-se que os picos apresentados apresentam alguma relação com os períodos de alta volatilidade.

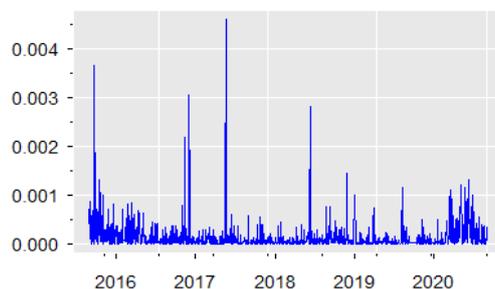


Figura 3.22: Retorno quadrado da taxa de câmbio EUR/BRL

Este diagnóstico foi confirmado através da aplicação do teste LM para 20 defasagens, assumindo um nível de significância $\alpha = 5\%$, uma vez que obteve-se uma estatística de teste $\chi^2_{20} = 71.101$ com 20 graus de liberdade e produziu o *valor-p* $= 1.203 \times 10^{-7}$, valor inferior ao α pré-estabelecido e que, por consequência, sugere a rejeição da hipótese nula testada de que não existem efeitos ARCH. Em termos práticos, este resultado ratifica a viabilidade na busca de um modelo heterocedástico, independente da autocorrelação da série de retornos em si, que não influencia esta decisão.

Os efeitos ARCH também foram confirmados pelo do teste de hipótese Ljung-Box, que aplicado à série dos retornos quadrados da taxa de câmbio EUR/BRL, que resultou na indicação de rejeição da hipótese nula de independência testada. O teste foi conduzido a um significância

$\alpha = 5\%$, produziu uma estatística de teste $\chi_1^2 = 29.727$ e $valor_p = 4.974 \times 10^{-8} < 5\%$.

Na busca de um modelo apropriado, da classe dos heterocedásticos, adotou-se uma estratégia fundamentada nos procedimentos seguidos no artigo “A GARCH Tutorial with R”, de Perlin *et al.*(2020) [17], onde estabelece-se *a priori* um conjunto de valores que cada parâmetro do modelo pode assumir e processa-se a estimação de múltiplos modelos em lote, reduzindo-se significativamente o esforço computacional. Particularmente, o processo de estimação dos modelos GARCH desta dissertação foi elaborado a partir do programa desenvolvido em [17] cujos autores basearam-se no *Package ‘rugarch’* [11] do *R Software*, todavia, o mesmo foi adaptado, tanto por conta da troca de contexto estudado, onde assume-se diferentes valores de parâmetros, distribuições de probabilidade e tipos de modelo, quanto pela questão de ter utilizado um outro pacote do *R Software: Package ‘fGarch’* [25].

O objetivo foi obter um único modelo, que conseguisse descrever de forma apropriada a média condicional e a variância condicional dos retornos da taxa de câmbio EUR/BRL, isto é, encontrar um modelo da classe *ARMA + GARCH*, que produz essas duas equações, sendo o seu primeiro componente (ARMA), com o papel de mapear a média condicional dos retornos enquanto o componente GARCH exerce a função de modelar a variância condicional dos retornos (volatilidade).

Embora a escolha dos parâmetros ideais do modelo (argumentos da função) seja complexa, demanda muitas análises e ainda assim, não haja garantias de conseguir um resultado satisfatório, iniciou-se o procedimento definindo os possíveis valores (candidatos) para cada um dos seis parâmetros, conforme ilustrado na tabela 3.6.

Parâmetro	Descrição	Possíveis valores
tipo de modelo	classe padrão ou assimétrica	“garch“, “aparch“
distribuição de probabilidade	distribuição de probabilidade dos erros	“norm“, “snorm“, “std“, “sstd“
p	ordem do modelo $AR(p)$	“1“
q	ordem do modelo $MA(q)$	“0“
P	ordem do modelo $ARCH(P)$	“1“, “2“
Q	ordem do modelo $GARCH(Q)$	“0“, “1“, “2“

Tabela 3.6: Parâmetros dos modelos *ARMA + GARCH* testados para o retorno da taxa de câmbio EUR/BRL.

Como a quantidade de modelos diferentes é obtida pela combinação dos possíveis valores de cada um dos seis parâmetros e com a atenção de produzir uma quantidade viável (número não tão elevado) de modelos diferentes para análises, decidiu-se limitar o escopo de cada um dos parâmetros com o seguinte racional:

- (1) **tipo de modelo:** foi explorado além do modelo tipo padrão (“garch”), onde há simetria no comportamento dos retornos positivos e negativos, o aspecto assimétrico (“aparch”), ou seja, onde há evidências de comportamentos distintos entre retornos positivos e retornos negativos;
- (2) **distribuição de probabilidade:** pelo padrão observado na análise exploratória de cauda mais pesada, comparada à distribuição Normal (“norm”), optou-se também em considerar a distribuição *t-Student* (“std”) além de suas vertentes assimétricas, “snorm” e “sstd”, respectivamente;
- (3) **p:** parâmetro da ordem do modelo autoregressivo fixado em 1, sugerido pela função de autocorrelação parcial (PACF) dado que o seu valor já ultrapassa a faixa de tolerância na primeira defasagem;
- (4) **q:** parâmetro da ordem do modelo de médias móveis fixado em 0 pelo padrão de estacionariedade (ausência de tendência);
- (5) **P:** Pela existência dos efeitos ARCH ratificada, definiu-se os possíveis valores desse parâmetro como 1 ou 2 para ampliar as possibilidades de teste, embora seja mais comum esse parâmetro assumir o valor 1. Além disso, quando a volatilidade passada também integra o modelo (GARCH), esse parâmetro P costuma assumir valores mais baixos (uma das vantagens do modelo GARCH em relação ao modelo ARCH);
- (6) **Q:** Esse parâmetro referente à ordem do modelo GARCH ficou livre para assumir qualquer valor inteiro entre 0 e 2, sendo o valor 2, definido por questões de viabilidade, conforme já mencionado.

O método de estimação adotado foi por máxima verossimilhança, termo que designa a qualidade do que é verossímil, mais conhecido pela expressão em inglês *maximum likelihood estimation*, cuja ideia central resume-se em determinar as estimativas para os parâmetros que maximizam a probabilidade de ocorrência dos dados, no modelo estatístico assumido para esses

dados. O detalhamento deste método de estimação aparece no apêndice A.2. Este procedimento produziu um total 48 modelos $ARMA + GARCH$ diferentes e assumindo $\alpha = 5\%$, apresenta-se os resultados dos valores dos critérios de informação (AIC e BIC) e a significância estatística dos parâmetros codificada, com os seguintes significados: “1“: significativo; “0“: não significativo, “-“: não integra o modelo e “na“: não calculado, pela figura 3.23:

#Ref.	MODELO	AIC	BIC	PARÂMETROS ESTADISTICAMENTE SIGNIFICATIVOS COM $\alpha=5\%$?											
				ar1	omega	alpha1	alpha2	beta1	beta2	gamma1	gamma2	delta	skew	shape	Todos par.
1	arma(1,0)+garch(1,0)distr.norm	-6,1707	-6,1576	1	1	-	-	-	-	-	-	-	-	-	1
2	arma(1,0)+garch(1,0)distr.std	-6,2334	-6,2128	1	1	1	-	-	-	-	-	-	-	1	1
3	arma(1,0)+garch(1,0)distr.snorm	-6,1777	-6,1602	1	1	1	-	-	-	-	-	-	-	1	1
4	arma(1,0)+garch(1,0)distr.sstd	-6,2354	-6,2134	1	1	1	-	-	-	-	-	-	-	1	1
5	arma(1,0)+garch(1,1)distr.norm	-6,2386	-6,2211	1	1	1	-	1	-	-	-	-	-	-	1
6	arma(1,0)+garch(1,1)distr.std	-6,2917	-6,2698	1	1	1	-	1	-	-	-	-	-	1	1
7	arma(1,0)+garch(1,1)distr.snorm	-6,2520	-6,2300	1	1	1	-	1	-	-	-	-	-	1	1
★ 8	arma(1,0)+garch(1,1)distr.sstd	-6,2998	-6,2735	1	1	1	-	1	-	-	-	-	-	1	1
9	arma(1,0)+garch(1,2)distr.norm	-6,2422	-6,2202	1	1	1	-	1	1	-	-	-	-	-	1
10	arma(1,0)+garch(1,2)distr.std	-6,2915	-6,2652	1	1	1	-	1	1	-	-	-	-	1	1
11	arma(1,0)+garch(1,2)distr.snorm	-6,2564	-6,2301	1	1	1	-	1	1	-	-	-	-	1	1
12	arma(1,0)+garch(1,2)distr.sstd	-6,3011	-6,2694	1	1	1	-	1	1	-	-	-	-	1	1
13	arma(1,0)+garch(2,0)distr.norm	-6,1959	-6,1790	1	1	1	1	-	-	-	-	-	-	1	1
14	arma(1,0)+garch(2,0)distr.std	-6,2437	-6,2218	1	1	1	1	-	-	-	-	-	-	1	1
15	arma(1,0)+garch(2,0)distr.snorm	-6,2093	-6,1813	1	1	1	1	-	-	-	-	-	-	1	1
16	arma(1,0)+garch(2,0)distr.sstd	-6,2484	-6,2221	1	1	1	1	-	-	-	-	-	-	1	1
17	arma(1,0)+garch(2,1)distr.norm	-6,2374	-6,2154	1	1	1	1	0	1	-	-	-	-	-	1
18	arma(1,0)+garch(2,1)distr.std	-6,2903	-6,2640	1	1	1	1	0	1	-	-	-	-	1	1
19	arma(1,0)+garch(2,1)distr.snorm	-6,2506	-6,2243	1	1	1	1	0	1	-	-	-	-	1	1
20	arma(1,0)+garch(2,1)distr.sstd	-6,2984	-6,2677	1	1	1	1	0	1	-	-	-	-	1	1
21	arma(1,0)+garch(2,2)distr.norm	-6,2404	-6,2141	1	1	1	1	0	1	0	-	-	-	-	1
22	arma(1,0)+garch(2,2)distr.std	-6,2898	-6,2591	1	1	1	1	0	1	0	-	-	-	1	1
23	arma(1,0)+garch(2,2)distr.snorm	-6,2347	-6,2240	1	1	1	1	0	1	0	-	-	-	1	1
24	arma(1,0)+garch(2,2)distr.sstd	-6,2984	-6,2633	1	1	1	1	0	1	0	-	-	-	1	1
25	arma(1,0)+aparch(1,0)distr.norm	-6,1723	-6,1503	1	1	1	-	-	-	0	-	1	-	-	1
26	arma(1,0)+aparch(1,0)distr.std	-6,2289	-6,2026	1	1	1	-	-	-	0	-	1	-	1	1
27	arma(1,0)+aparch(1,0)distr.snorm	-6,1788	-6,1525	1	1	1	-	-	-	0	-	1	-	1	1
28	arma(1,0)+aparch(1,0)distr.sstd	-6,2336	-6,2029	1	1	1	-	-	-	0	-	1	-	1	1
29	arma(1,0)+aparch(1,1)distr.norm	-6,0899	-6,0636	1	1	1	-	-	-	1	-	1	-	-	1
30	arma(1,0)+aparch(1,1)distr.std	-6,2040	-6,1733	1	1	1	-	-	-	1	-	1	-	1	1
31	arma(1,0)+aparch(1,1)distr.snorm	-5,9480	-5,9173	1	1	1	-	-	-	1	-	1	-	1	1
32	arma(1,0)+aparch(1,1)distr.sstd	-6,2056	-6,1705	1	1	1	-	-	-	1	-	1	-	1	1
33	arma(1,0)+aparch(1,2)distr.norm	-6,0167	-5,9860	1	1	1	-	-	-	1	1	1	-	1	1
34	arma(1,0)+aparch(1,2)distr.std	-6,1910	-6,1559	1	1	1	-	-	-	1	1	1	-	1	1
35	arma(1,0)+aparch(1,2)distr.snorm	-5,8132	-5,7781	1	1	1	-	-	-	1	1	1	-	1	1
36	arma(1,0)+aparch(1,2)distr.sstd	-6,1965	-6,1570	1	1	1	-	-	-	1	1	1	-	1	1
37	arma(1,0)+aparch(2,0)distr.norm	-6,2015	-6,1708	1	1	1	1	-	-	0	0	1	-	-	1
38	arma(1,0)+aparch(2,0)distr.std	-6,2427	-6,2076	1	1	1	1	-	-	0	0	1	-	1	1
39	arma(1,0)+aparch(2,0)distr.snorm	-6,2077	-6,1726	1	1	1	1	-	-	0	0	1	-	1	1
40	arma(1,0)+aparch(2,0)distr.sstd	-6,2471	-6,2076	1	1	1	1	-	-	0	0	1	-	1	1
41	arma(1,0)+aparch(2,1)distr.norm	-5,7401	-5,7050	1	1	1	1	na	1	-	1	na	1	-	1
42	arma(1,0)+aparch(2,1)distr.std	-6,1767	-6,1372	1	1	1	1	na	1	-	1	na	1	-	1
43	arma(1,0)+aparch(2,1)distr.snorm	-5,1991	-5,1596	1	1	1	1	na	1	-	1	na	1	-	1
44	arma(1,0)+aparch(2,1)distr.sstd	-6,1825	-6,1367	1	1	1	1	na	1	-	1	na	1	-	1
45	arma(1,0)+aparch(2,2)distr.norm	-6,0132	-5,9737	1	1	1	1	na	1	-	1	na	1	-	1
46	arma(1,0)+aparch(2,2)distr.std	-6,1875	-6,1496	1	1	1	1	na	1	-	1	na	1	-	1
47	arma(1,0)+aparch(2,2)distr.snorm	-5,8140	-5,7701	1	1	1	1	na	1	-	1	na	1	-	1
48	arma(1,0)+aparch(2,2)distr.sstd	-6,1990	-6,1448	1	1	1	1	na	1	-	1	na	1	-	1

Figura 3.23: Resultados da modelação $ARMA + GARCH$ para o retorno da taxa de câmbio EUR/BRL.

Diante de todos os modelos testados, procedeu-se a seleção somente daqueles que apresentaram todos os parâmetros satisfatórios, isto é, aqueles cujos coeficientes estimados, mostraram-se estatisticamente significativos (linhas destacadas em verde da figura 3.23). Após realizado esse filtro inicial, elegeu-se o modelo com melhor qualidade de ajustamento, ou seja, aquele que apresentou menores valores dos critérios de informação, especificamente para os critérios AIC e BIC. Neste cenário, o modelo escolhido foi o de referência 8 (marcado na figura com uma estrela), que o teve o desempenho mais satisfatório, com a seguinte configuração: tipo “garch“, $ARMA(1,0) + GARCH(1,1)$ com distribuição dos erros t -Student assimétrica (“sstd“). Esses resultados são exibidos na 3.23 assim como os coeficientes estimados do melhor modelo aparecem na tabela 3.7.

Coefficiente	Valor do coeficiente estimado
ar1	-0.113424
ω	0.000003
α_1	0.079631
β_1	0.898949
<i>skew</i>	1.146784
<i>shape</i>	6.640549

Tabela 3.7: Coeficientes estimados pelo modelo $ARMA(1, 0) + GARCH(1, 1)$ para o retorno da taxa de câmbio EUR/BRL, com distribuição dos erros *t-Student* assimétrica (“sstd”).

Dessa forma, o modelo $ARMA(1, 0) + GARCH(1, 1)$ identificado para os retornos da taxa de câmbio EUR/BRL (r_t), é descrito por:

$$r_t = -0.113424 r_{t-1} + z_t, \tag{3.5}$$

representando a sua média condicional (modelo $AR(1)$), onde

$$z_t = \sigma_t \varepsilon_t, \tag{3.6}$$

e os erros ε_t são iid e possuem distribuição *t-Student* assimétrica ($\varepsilon_t \sim sstd(0, 1)$), de modo que a sua variância condicional (σ_t^2) é expressa por:

$$\sigma_t^2 = 0.000003 + 0.079631 z_{t-1}^2 + 0.898949 \sigma_{t-1}^2. \tag{3.7}$$

A equação deste modelo determinado evidencia que a volatilidade (variância) no instante t é mais fortemente influenciada pela volatilidade do instante anterior ($t - 1$) comparado aos retornos passados, dado que seu coeficiente β_1 é cerca de 11 vezes superior ao coeficiente α_1 , referente ao termo dos erros quadráticos do retorno no mesmo instante $t - 1$.

Os resultados desse modelo podem ser conferidos visualmente (ver figura 3.24), sendo que especificamente pelo gráfico *QQ Plot*, percebe-se uma boa qualidade de ajustamento, ou seja, os quantis empíricos e teóricos se aproximam e os pontos se distribuem ao longo de uma reta, salvo algumas exceções de desajuste.

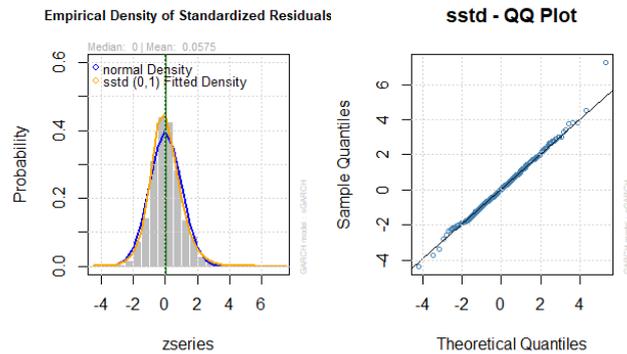


Figura 3.24: Densidade e quantis dos resíduos padronizados da taxa de câmbio EUR/BRL, obtidos por modelo $ARMA(1,0) + GARCH(1,1)$ em comparação aos quantis teóricos.

Na sequência do processo de modelação, etapa 6 do fluxograma do processo de modelação 3.15), é necessário e fundamental analisar os resíduos, ou os resíduos padronizados pelo modelo. Os resíduos padronizados, de acordo com Shumway e Stoffer (2017) [20], são definidos por

$$e_t = \frac{(x_t - \hat{x}_t^{t-1})}{\sqrt{\hat{P}_t^{t-1}}},$$

onde \hat{x}_t^{t-1} é a previsão calculada um passo adiante (*one-step-ahead prediction*) de x_t baseado no modelo ajustado onde \hat{P}_t^{t-1} representa a variância condicional estimada, também um passo adiante. Nos casos de bom ajustamento do modelo os resíduos padronizados devem comportar-se com um sequência iid de média zero e variância unitária.

Nesta análise, nota-se que os resíduos padronizados tem aparente aspecto favorável, isto é, apresentam um padrão de aleatoriedade, pois seus valores de autocorrelação posicionam-se dentro da faixa de tolerância (ver figura 3.25), o que sugere que não são estatisticamente significativos. Obviamente esta impressão foi verificada, especificamente por meio da realização do teste de hipótese Ljung-Box, que produziu uma estatística de teste de $\chi_1^2 = 0.85842$, com um grau de liberdade e $valor_p = 0.3542 > \alpha = 5\%$, cuja conclusão foi a não rejeição da hipótese H_0 que os resíduos padronizados são iid. O mesmo teste de hipótese Ljung-Box, a um nível de significância $\alpha = 5\%$, também foi aplicado aos resíduos padronizados quadrados e o resultado foi similar: não se rejeita a hipótese H_0 que os resíduos padronizados quadrados são iid (estatística de teste de $\chi_1^2 = 3.8143$, com um grau de liberdade e $valor_p = 0.05082 > \alpha = 5\%$), reforçando a ideia de robustez do modelo identificado. Este fato de os resíduos não serem

autocorrelacionados, descreve o cenário ideal de que todo o padrão de comportamento dos dados é revelado somente pelas variáveis explicativas e consequentemente, os resíduos denotam apenas a característica da aleatoriedade.

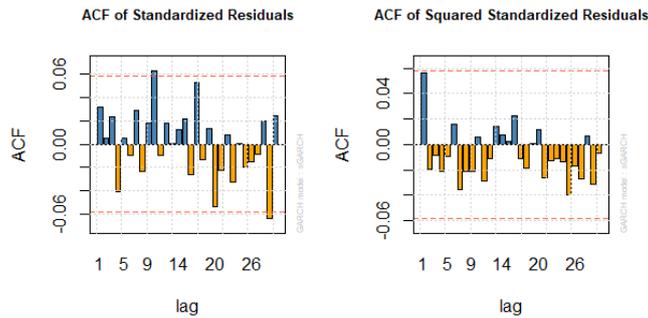


Figura 3.25: ACF dos resíduos padronizados e dos resíduos padronizados quadrados da taxa de câmbio EUR/BRL, obtidos por modelo $ARMA(1, 0) + GARCH(1, 1)$.

O modelo $ARMA + GARCH$ obtido consegue captar razoavelmente os movimentos de instabilidade do retorno. Considerando um intervalo de $r_t \pm 2 \sigma_t$, onde σ_t representa a volatilidade condicional (ou desvio padrão condicional), o modelo falha em $\approx 14\%$, isto é, em $\approx 14\%$ dos instantes, o retorno observado está fora destes limites (ver figura 3.26). Também nota-se uma sensibilidade do modelo a esses períodos de turbulência (maior volatilidade), principalmente quando os valores observados não são muito extremos (picos e vales), ou seja, a volatilidade é captada corretamente, porém, de forma mais branda e não, em sua real intensidade.

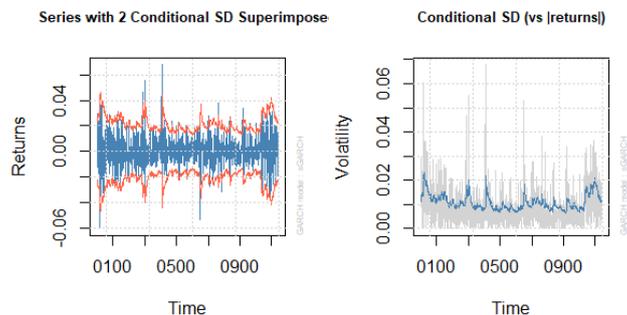


Figura 3.26: Série de retornos da taxa de câmbio EUR/BRL delimitada $\pm 2 \sigma$ e volatilidade condicional da taxa de câmbio EUR/BRL obtidos por modelo $ARMA(1, 0) + GARCH(1, 1)$.

Após a identificação de um modelo GARCH com razoável qualidade de ajustamento, testou-

se também mais alguns cenários de modelação onde manteve-se todos os seus parâmetros e incorporou-se variáveis externas como predictoras, analogamente à estratégia adotada para a regressão dinâmica. Estas variáveis externas, candidatas à integrarem um novo modelo, referem-se aos retornos dos índices de ações e suas variantes, em termos da parte negativa dos retornos e também, considerando uma defasagem (*lag* 1) aparecem, referenciadas por respectivamente por retorno, retorno negativo e retorno com defasagem, aparecem na tabela 3.8.

retorno	retorno negativo	retorno com defasagem
ret.dowjones	ret.neg.dowjones	lag1.ret.dowjones
ret.sp500	ret.neg.sp500	lag1.ret.sp500
ret.nasdaq	ret.neg.nasdaq	lag1.ret.nasdaq
ret.estx50	ret.neg.estx50	lag1.ret.estx50
ret.euronext100	ret.neg.euronext100	lag1.ret.euronext100
ret.ibovespa	ret.neg.ibovespa	lag1.ret.ibovespa

Tabela 3.8: Relação das variáveis externas testadas como predictoras em modelo GARCH.

Diante destas 18 variáveis relacionadas ao mercado acionário, foram testados os seguintes cenários:

- todas as variáveis: um novo modelo com a incorporação completa destas variáveis externas;
- uma única variável: um novo modelo a cada variável adicionada, incorporada de forma individual;
- um par de variáveis: um novo modelo para os pares de variáveis externas desde que a variável “ret.neg.estx50” ou “ret.neg.euronext100” estivessem presentes, escolhidas por apresentarem maior correlação, em valor absoluto, com o retorno da taxa de câmbio EUR/BRL.

Neste exercício, não foram detectados modelos cuja todos os coeficientes estimados permaneciam estatisticamente significativos, testados com $\alpha = 5\%$, ou seja, a incorporação destas novas variáveis, em todos os cenários testados, “piora” o que já se tinha ajustado e portanto, não há melhorias no modelo. Mesmo que esse aspecto da significância estatística pudesse ser desprezado, esta iniciativa não resultaria, dado que os critérios de informação, usados neste estudo para comparar a qualidade de ajustamento dos modelos, mostraram-se menos favoráveis (valores mais altos). Dessa maneira, o melhor modelo GARCH identificado para a retorno da taxa de câmbio EUR/BRL permaneceu inalterado. Finalizada a etapa da escolha do modelo

$ARMA + GARCH$ com melhor ajustamento, optou-se por replicá-lo, com as mesmas parametrizações, utilizando o *Package ‘rugarch’* do *R Software* [11], por conta de alguns recursos e funcionalidades adicionais relacionados aos temas *Forecasting* e simulação. Dessa forma, procedeu-se a estimação das previsões pelo do método de *rolling forecast*¹² para o período de um dia ($h = 1$) e com cinco repetições ($n = 5$). Estas parametrizações refletem um horizonte total de 6 unidades temporais, neste caso, 6 dias. Este procedimento gerou previsões do retorno e da volatilidade para 6 dias, correspondente ao período final (6 instantes) na linha temporal da série de retornos (ver figura 3.27). Em termos de acurácia, o *forecast* dos retornos teve um desempenho ruim, o indicador MAPE medido no período integral (6 dias), considerando os erros acumulados, foi elevado ($MAPE = 1.1134$).

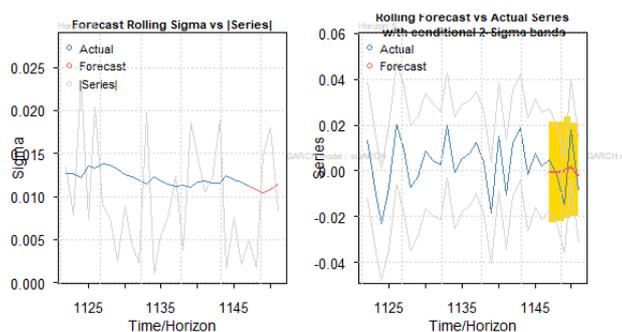


Figura 3.27: *Rolling forecast* da volatilidade (σ) e retorno da taxa de câmbio EUR/BRL obtidos por modelo $ARMA(1,0) + GARCH(1,1)$, para um horizonte de 6 dias.

Nota-se que os pontos médios da previsão do retorno não acompanha os movimentos reais dos valores observados, que mostraram-se muito oscilantes nesse período, no entanto, para o intervalo de confiança delimitado por $\pm 2 \sigma_t$, onde σ_t representa a volatilidade condicional (zona amarela do gráfico anterior), consegue englobar integralmente os valores observados no período.

Além do *rolling forecast*, este mesmo modelo $ARMA(1,0) + GARCH(1,1)$ foi também base para explorar-se um outro método, cuja abrangência de aplicações contempla inclusive o tema das previsões: as simulações. Com a especificação do modelo já conhecida e os seus parâmetros estimados, a partir dos dados, a ideia da simulação consiste em gerar um número elevado de séries temporais, de qualquer tamanho, e em número elevado mas viável, pois este procedi-

¹²método de previsão que desconsidera as últimas m observações e realiza n sucessivas rodadas de estimação, onde $n \leq m$, para um horizonte de previsão h , de modo a avançar um instante na linha do tempo, a cada iteração.

mento demanda alto seu custo computacional, de modo que as séries simuladas, que possuem as mesmas propriedades da série de interesse, possam ser utilizadas para a inferência estatística, propiciando a realização de análises probabilísticas e previsão de períodos futuros. Dessa maneira, partindo-se última observação conhecida do retorno da taxa de câmbio EUR/BRL, referente à data “2020-08-28”, realizou-se um conjunto de 5000 simulações para um período de 6 dias como horizonte de previsão, onde estimou-se a função de densidade de probabilidade. Este processo de estimação da função da densidade ocorreu pelo método não paramétrico de Kernel, que tem forte relação com os histogramas e que pode ser descrito superficialmente como um método que ajusta sobre cada ponto, uma superfície suavemente curva, com o emprego de diferentes funções Kernel, tais como gaussiana, cosseno, triangular, parabólica entre outras, e que possuem diferentes aspectos. O resultado com as funções de densidade dos retornos quanto para a volatilidade está ilustrado na figura 3.28, que mostra uma comparação com os dados reais (observados).

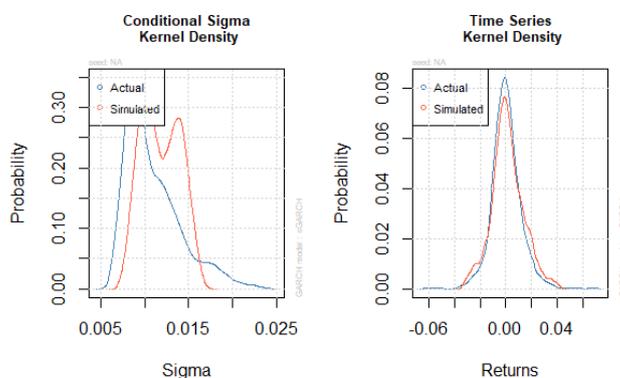


Figura 3.28: Densidade de probabilidade da volatilidade (Sigma) e retorno da taxa de câmbio EUR/BRL, obtido por realização de 5000 simulações do modelo $ARMA(1, 0) + GARCH(1, 1)$.

É evidente que o método da simulação se mostrou eficiente para captar de forma mais precisa o comportamento dos retornos. Por inspeção visual, o ajuste parece bastante favorável. Do lado da volatilidade condicional (σ_t), a impressão não é a mesma. O resultado da volatilidade mostra-se aderente apenas em níveis de volatilidade baixa, algo como $\sigma_t \leq \approx 0.010$ (ver figura 3.28).

Para a avaliação das previsões do retorno, obtidas por ambos os métodos, tanto na perspec-

tiva do *rolling forecast* quanto das simulações, utilizou-se o indicador de erro MAPE. Embora em ambos os casos tenha sido previsto um horizonte de 6 dias futuros, o cálculo da métrica de erro não é realizado nos mesmos instantes, porque o método *rolling forecast* desconsidera-se as últimas observações e a suas estimações correspondem à linha temporal do período desprezado, enquanto que na simulação, a primeira previsão refere-se a um período ainda não observado (período futuro).

Assim, representando-se as previsões por \hat{y}_t , tem-se $(\hat{y}_T, \hat{y}_{T+1}, \hat{y}_{T+2}, \dots, \hat{y}_{T+h-1})$ para previsões obtidas por meio de simulação, onde T refere-se ao primeiro instante não observado e h , o horizonte da previsão. Já no cenário de *rolling forecast*, adotando-se o mesmo horizonte h , as previsões referem à $(\hat{y}_{T-h}, \hat{y}_{T-h+1}, \hat{y}_{T-h+2}, \dots, \hat{y}_{T-2}, \hat{y}_{T-1})$, o que implica em uma distância de h unidades temporais entre os métodos. Apesar desta particularidade, procedeu-se o cálculo do MAPE, para todos os períodos acumulados de 1 a 6 dias para ambos os métodos. Nas simulações, procedeu-se um cálculo prévio do valor médio previsto, dentre as 5000 simulações a cada instante previsto (6 dias). Neste critério de avaliação, as previsões geradas por diferentes métodos mostraram-se muito pouco acuradas, para todos os períodos apurados (ver figura 3.29 e tabela 3.9). Mesmo no melhor desempenho obtido ($MAPE \approx 88\%$ para *rolling forecast* do período de 1 a 5 dias), o nível do erro é muito elevado, e significa que o desvio médio tem quase ($\approx 88\%$) da mesma grandeza do valor observado, e o modelo não consegue reproduzir as reais oscilações do retorno ao longo do tempo.

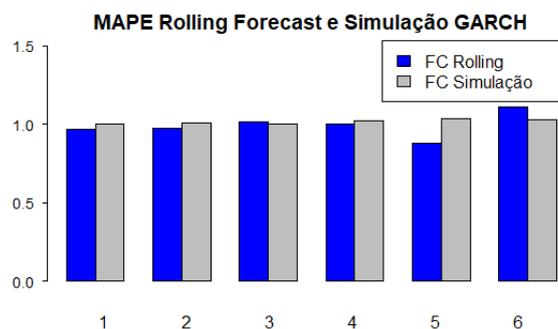


Figura 3.29: Indicador MAPE do *forecast*, obtido por modelo $ARMA(1, 0) + GARCH(1, 1)$ para o retorno da taxa de câmbio EUR/BRL, calculado para pelos métodos *rolling forecast* e simulação, considerando o período acumulado de 1 a 6 dias.

Período (t) em dias	FC rolling	FC simulação
1	0.9706	1.0043
[1, 2]	0.9729	1.0063
[1, 3]	1.0127	1.0018
[1, 4]	1.0019	1.0205
[1, 5]	0.8783	1.0353
[1, 6]	1.1134	1.0316

Tabela 3.9: Indicador MAPE do *forecast*, obtido por modelo $ARMA(1, 0) + GARCH(1, 1)$ para o retorno da taxa de câmbio EUR/BRL, calculado para pelos métodos *rolling forecast* e simulação, considerando o período acumulado de 1 a 6 dias.

3.4.2 Taxa de câmbio EUR/USD

Conforme já visto anteriormente (3.3.2), a série de retornos da taxa de câmbio EUR/USD é estacionária e sugere a presença de uma heterocedasticidade condicional (efeitos ARCH). Como o objetivo é explicá-la por um modelo GARCH, que apresenta termos em sua equação de variância condicional (2.19) que são função dos retornos quadrados, convém examinar a autocorrelação dos retornos quadrados. Adotando como parâmetro até 20 *lags*, pela questão da modelação abordar somente de modelos GARCH de baixa ordem, o correlograma dos retornos quadrados (ver figura 3.30) exibe uma autocorrelação estatisticamente significativa em vários pontos ($lag = 2, 4, 5, 9, \dots$), o que reforça o suspeita da existência de efeitos ARCH. Esses efeitos foram investigados pela aplicação do teste LM, para 20 *lags* e com o nível de significância $\alpha = 5\%$. Produziu-se uma estatística $\chi^2_{20} = 47.702$ com 20 graus de liberdade e o *valor_p* = 0.0004682, que é um valor inferior ao nível α . Desta forma, há evidências para a rejeição de H_0 : Sem efeitos ARCH, e a garantia de que a modelação GARCH pode ser prosseguida.

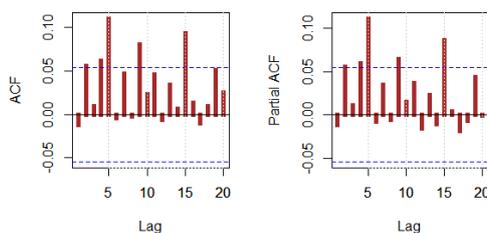


Figura 3.30: ACF/PACF do retorno quadrado da taxa de câmbio EUR/USD (até *lag* 20).

Os mesmos procedimentos realizados para a identificação do “melhor” modelo $ARMA + GARCH$ para os retornos da taxa de câmbio EUR/BRL (descrito na seção 3.4.1.2 com base no artigo “A GARCH Tutorial with R” [17]) foram repetidos para o retorno da taxa de câmbio EUR/USD, todavia, com uma parametrização distinta (ver tabela 3.10). Vale ressaltar que neste processo de estimação, conduzido através do método de máxima verossimilhança (MLE), que aparece mais detalhado no apêndice A.2, houve algumas ocorrências de matriz hessiana ¹³ singular, predominantemente em modelos de ordem GARCH mais elevada ($Q > 1$). Isto impedia a execução do procedimento de estimação e forçou uma redução no conjunto de “possíveis valores” dos parâmetros.

Parâmetro	Descrição	Possíveis valores
tipo de modelo	classe padrão ou assimétrica	“garch”
distribuição de probabilidade	distribuição de probabilidade dos erros	“norm”, “snorm”, “std”, “sstd”
p	ordem do modelo $AR(p)$	“0”, “1”
q	ordem do modelo $MA(q)$	“0”
P	ordem do modelo $ARCH(P)$	“1”, “2”
Q	ordem do modelo $GARCH(Q)$	“0”, “1”

Tabela 3.10: Parâmetros dos modelos $ARMA + GARCH$ testados para o retorno da taxa de câmbio EUR/USD.

Basicamente, as diferenças referem-se ao conjunto de “possíveis valores” dos parâmetros testados localizadas no “tipo de modelo”, que nesta modelação, ficou restrita ao “garch” devido à baixa assimetria apresentada (seção 3.3.2) e nos parâmetros “p” (ordem do modelo AR), que teve a sua lista de “possíveis valores” ampliada e “Q” (ordem do modelo GARCH), com abrangência mais limitada. Com o apoio das funções do *Package ‘fGarch’* [25] do *R Software*, procedeu-se a estimação dos modelos, pelo método de máxima verossimilhança, para todas as combinações de parâmetros pré-definidos resultando 32 diferentes modelos $ARMA + GARCH$. Para todos os modelos gerados, foram identificados os coeficientes estatisticamente significativos ($\alpha = 5\%$), ou seja, aqueles para os quais recomenda-se a rejeição da hipótese nula (H_0) de coeficiente nulo (não produz efeito ao modelo) e calculados os critérios de informação AIC e BIC. Estes resultados são exibidos pela figura 3.31, onde a significância estatística dos parâmetros aparece codificada, com a seguinte interpretação: “1”: significativo; “0”: não significativo e “-”: não integra o modelo. As linhas da figura 3.31, destacadas em verde, referem-se aos modelos

¹³matriz quadrada formada pelas derivadas parciais de segunda ordem de uma determinada função.

onde se verificou significância estatística para todos os seus parâmetros, sendo que esta condição se verificou apenas para os modelos de referência 5 e 7. Dentre esses dois modelos, adotou-se inicialmente como “melhor” modelo o de referência 5 (marcado na figura com uma estrela), $ARMA(0,0) + GARCH(1,1)$ com distribuição Normal dos erros, por apresentar valores do AIC e BIC mais favoráveis (valores mais baixos) em ambos critérios.

#Ref.	MODELO	AIC	BIC	PARÂMETROS ESTADÍSTICAMENTE SIGNIFICATIVOS COM $\alpha=5\%$?							
				ar1	omega	alpha1	alpha2	beta1	skew	shape	Todos par.
1	arma(0,0)+garch(1,0)distr. norm	-7,8792	-7,8712	1	0						0
2	arma(0,0)+garch(1,0)distr. std	-7,9490	-7,9371	1	0					1	0
3	arma(0,0)+garch(1,0)distr. snorm	-7,8776	-7,8657	1	0				1		0
4	arma(0,0)+garch(1,0)distr. sstd	-7,9477	-7,9318	1	0				1	1	0
★ 5	arma(0,0)+garch(1,1)distr. norm	-7,9269	-7,9150	1	1			1			1
6	arma(0,0)+garch(1,1)distr. std	-7,9802	-7,9643	0	1			1		1	0
7	arma(0,0)+garch(1,1)distr. snorm	-7,9254	-7,9095	1	1			1	1		1
8	arma(0,0)+garch(1,1)distr. sstd	-7,9788	-7,9589	0	1			1	1	1	0
9	arma(0,0)+garch(2,0)distr. norm	-7,8885	-7,8766	1	0	0	1				0
10	arma(0,0)+garch(2,0)distr. std	-7,9512	-7,9353	1	0	0	0			1	0
11	arma(0,0)+garch(2,0)distr. snorm	-7,8872	-7,8713	1	0	0	1			1	0
12	arma(0,0)+garch(2,0)distr. sstd	-7,9500	-7,9302	1	0	0	0			1	0
13	arma(0,0)+garch(2,1)distr. norm	-7,9302	-7,9143	1	0	0	0	1			0
14	arma(0,0)+garch(2,1)distr. std	-7,9807	-7,9609	0	0	0	0	1		1	0
15	arma(0,0)+garch(2,1)distr. snorm	-7,9288	-7,9089	1	0	0	0	1	1		0
16	arma(0,0)+garch(2,1)distr. sstd	-7,9794	-7,9556	0	0	0	0	1	1	1	0
17	arma(1,0)+garch(1,0)distr. norm	-7,8790	-7,8670	0	1	0					0
18	arma(1,0)+garch(1,0)distr. std	-7,9496	-7,9337	0	1	0				1	0
19	arma(1,0)+garch(1,0)distr. snorm	-7,8774	-7,8615	0	1	0				1	0
20	arma(1,0)+garch(1,0)distr. sstd	-7,9484	-7,9285	0	1	0				1	0
21	arma(1,0)+garch(1,1)distr. norm	-7,9264	-7,9105	0	1	1		1			0
22	arma(1,0)+garch(1,1)distr. std	-7,9803	-7,9604	0	0	1		1		1	0
23	arma(1,0)+garch(1,1)distr. snorm	-7,9249	-7,9050	0	1	1		1	1		0
24	arma(1,0)+garch(1,1)distr. sstd	-7,9789	-7,9551	0	0	1		1	1	1	0
25	arma(1,0)+garch(2,0)distr. norm	-7,8879	-7,8720	0	1	0	0	1			0
26	arma(1,0)+garch(2,0)distr. std	-7,9516	-7,9317	0	1	0	0	0		1	0
27	arma(1,0)+garch(2,0)distr. snorm	-7,8866	-7,8667	0	1	0	0	1		1	0
28	arma(1,0)+garch(2,0)distr. sstd	-7,9504	-7,9266	0	1	0	0	0		1	0
29	arma(1,0)+garch(2,1)distr. norm	-7,9296	-7,9098	0	1	0	0	0	1		0
30	arma(1,0)+garch(2,1)distr. std	-7,9808	-7,9570	0	0	0	0	0	1	1	0
31	arma(1,0)+garch(2,1)distr. snorm	-7,9283	-7,9044	0	0	0	0	0	1	1	0
32	arma(1,0)+garch(2,1)distr. sstd	-7,9796	-7,9518	0	0	0	0	0	1	1	0

Figura 3.31: Resultados da modelação $ARMA + GARCH$ para o retorno da taxa de câmbio EUR/USD.

Antes de proceder uma análise minuciosa do modelo, foi aplicado um teste de normalidade aos resíduos deste modelo, com o intuito de validar a veracidade da premissa assumida de erros com distribuição Normal, caso contrário, perde-se eficiência das estimações produzidas. O teste de hipótese realizado foi Shapiro-Wilk que atestou “quebra” de premissa, dado que o teste produziu uma estatística de teste $W = 0.97142$ e o $valor-p = 2.31 \times 10^{-15} \leq \alpha$, considerando um nível de significância $\alpha = 5\%$, recomendando a rejeição da hipótese nula de normalidade. Diante deste cenário, trocou-se a escolha do “melhor” modelo para o de referência 6, com distribuição t -Student, que apresentou os valores dos critérios de informação AIC e BIC mais baixos (favoráveis), embora o coeficiente estimado para a constante da equação da variância condicional ω não tenha exibido significância estatística. O fato deste coeficiente estimado ter valor próximo de zero, reduz a sua contribuição na explicação da volatilidade e consequentemente, a sua importância.

Os valores dos coeficientes estimados para este modelo $ARMA(0, 0) + GARCH(1, 1)$, com distribuição t -Student para os erros, são exibidos pela tabela 3.11.

Coeficiente	Valor do coeficiente estimado
ω	0.00000007
α_1	0.02060155
β_1	0.97639434
$shape$	6.71487471

Tabela 3.11: Coeficientes estimados pelo modelo $ARMA(0, 0) + GARCH(1, 1)$ para o retorno da taxa de câmbio EUR/USD, com distribuição dos erros t -Student (“std”).

Dessa forma, a média dos retornos da taxa de câmbio EUR/USD é constante, pois foi modelada por um $ARMA(0, 0)$, que denota um processo de ruído branco, e nula (esse parâmetro foi desprezado no processo de estimação). Assim, o modelo determinado, tem a representação da média condicional dos retornos da taxa de câmbio EUR/USD, descrita por:

$$r_t = z_t, \tag{3.8}$$

$$r_t = \sigma_t \varepsilon_t, \tag{3.9}$$

onde os erros ε_t são iid e têm distribuição t -Student, ou seja, $\varepsilon_t \sim iid\ std(0, 1)$, enquanto que a variância condicional dos retornos da taxa de câmbio EUR/USD (σ_t^2), é expressa pela equação:

$$\sigma_t^2 = 0.00000007 + 0.02060155 r_{t-1}^2 + 0.97639434 \sigma_{t-1}^2. \tag{3.10}$$

O modelo elaborado é estacionário ($\alpha_1 + \beta_1 \approx 0.997 < 1$) e a volatilidade é bastante mais sensível pela volatilidade do instante anterior, comparativamente ao retornos quadrados passados. Observa-se um bom ajustamento dos resíduos padronizados à distribuição t -Student, todavia, o ajuste exibe algumas discrepâncias aparentes na caudas, e de forma mais presente, na cauda esquerda da distribuição dos retornos da taxa de câmbio EUR/USD (ver figura 3.32), uma vez que há desalinhamento de quantis em alguns pontos.

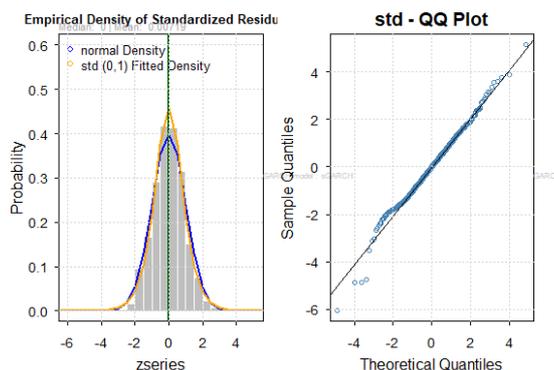


Figura 3.32: Densidade e quantis dos resíduos padronizados da taxa de câmbio EUR/USD, obtidos por modelo $ARMA(0, 0) + GARCH(1, 1)$ em comparação aos quantis teóricos.

Os resíduos e os resíduos quadrados deste modelo foram avaliados, sob a perspectiva de autocorrelação. Para ambos, nota-se claramente um aspecto visual de não autocorrelacionados (ver figura 3.33), o que contribui na avaliação positiva do modelo uma vez que é a parte não explicada pelo modelo e que idealmente devem se mostrar iid.

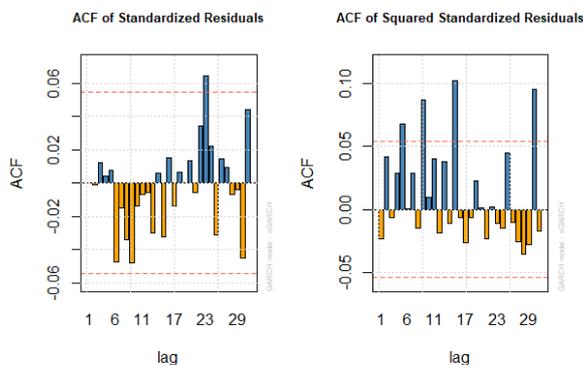


Figura 3.33: ACF dos resíduos padronizados e dos resíduos padronizados quadrados da taxa de câmbio EUR/USD, obtidos por modelo $ARMA(0, 0) + GARCH(1, 1)$.

Estas impressões foram verificadas pelo teste hipótese Ljung-Box, com significância $\alpha = 5\%$ e um grau de liberdade, para ambos os casos. Para os resíduos do retorno da taxa de câmbio EUR/USD, o resultado do teste indicou uma estatística $\chi_1^2 = 0.13913$ e o $valor_p = 0.7091 > \alpha = 5\%$, cuja recomendação foi a não rejeição da hipótese nula (H_0 : iid). Para os resíduos quadrados do retorno da taxa de câmbio EUR/USD, a estatística de teste de foi $\chi_1^2 = 0.18187$ e o $valor_p = 0.6698$. Como $valor_p > \alpha$, há evidências de que os resíduos quadrados são

independentes (não autocorrelacionados).

As oscilações do retorno da taxa de câmbio EUR/USD capturadas por este modelo descrevem o mesmo padrão observado ao longo do tempo, porém em menor amplitude, de modo que o modelo não é sensível à fortes turbulências. Considerando um intervalo de $0 \pm 2 \sigma_t$, onde σ_t representa o desvio padrão condicional, o modelo falha (fora do intervalo) em $\approx 19\%$ dos instantes medidos, especificamente para valores extremos, independentemente se o retorno é negativo ou positivo. Neste caso, a maior incidência de falhas ocorreu para altos valores dos retornos positivos, cerca de $\approx 60\%$, (ver figura 3.34).

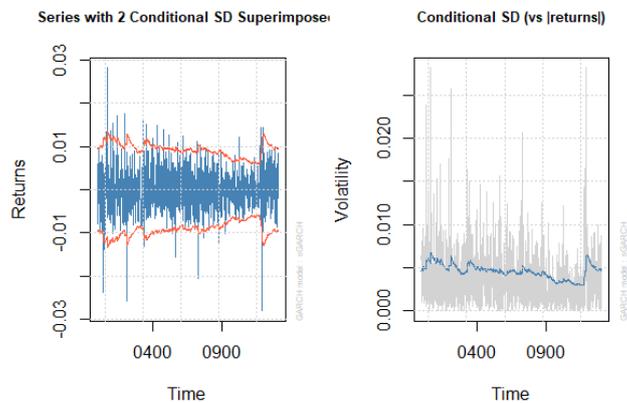


Figura 3.34: Série de retornos da taxa de câmbio EUR/USD delimitada $\pm 2 \sigma$ e volatilidade condicional da taxa de câmbio EUR/USD obtidos por modelo $ARMA(0,0) + GARCH(1,1)$, com distribuição *t-Student* para os erros.

Para a elaboração do *forecast* para o retorno da taxa de câmbio EUR/USD, foi utilizada uma outra abordagem, onde estima-se a densidade a partir de um modelo $ARMA + GARCH$, por um método chamado *rolling density forecast*. Este método calcula a média condicional e o desvio padrão condicional para o período do *forecast*, de acordo com uma parametrização previamente estabelecida.

Esta funcionalidade está disponível através da função “*ugarchroll*“ do *R Software*, pertencente ao *Package ‘rugarch’* [11].

Sua lista de parâmetros é extensa, sendo que os principais deles, estão relacionados abaixo, acompanhado pela parametrização escolhida para a execução da estimação da densidade:

- (1) dados: retorno da taxa de câmbio EUR/USD;

- (2) especificação do modelo: $ARMA(0, 0) + GARCH(1, 1)$, conforme 3.11;
- (3) distribuição de probabilidade dos erros: distribuição t -Student;
- (4) número de períodos de estimação: 1 (1 dia);
- (5) horizonte total de previsão: 6 (6 dias);
- (6) frequência de reestimação: 1 (reestima-se a cada iteração).

Em termos de desempenho, este modelo gerou as seguintes métricas:

- *Mean Squared Error*: $MSE = 1.702 \times 10^{-5}$;
- *Mean Absolute Error*: $MAE = 2.927 \times 10^{-3}$.

A partir do resultado da densidade estimada para o retorno da taxa de câmbio EUR/USD para cada um dos 6 dias previstos, foi possível também elaborar um gráfico que retrata a distribuição de probabilidade, para cada um dos dias previstos, destacando-se cinco quantis, definidos por: $(Q_1, Q_2, Q_3, Q_4, Q_5) \equiv (Q_{0\%}, Q_{25\%}, Q_{50\%}, Q_{75\%}, Q_{100\%})$, conforme figura 3.35.

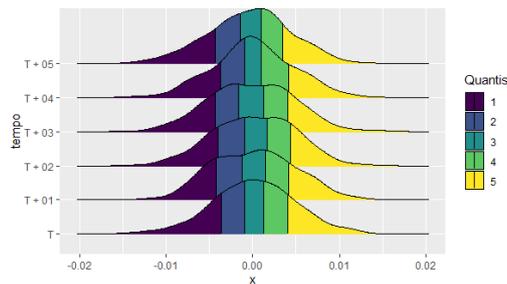


Figura 3.35: Densidade estimada pelo método *rolling density forecast*, baseado em modelo $ARMA(0, 0) + GARCH(1, 1)$, com distribuição t -Student para os erros e horizonte de 6 dias.

Capítulo 4

Conclusões

4.1 Conclusões

Embora as séries financeiras sejam normalmente caracterizadas pela ausência de estacionariedade e alternância de momentos de baixa e alta volatilidade, que são elementos dificultadores para a investigação, “sempre” há possibilidade de extrair informações relevantes a partir de seus dados históricos. É preciso ter o conhecimento e a capacidade de discernir entre o sinal e o ruído. De acordo com o estatístico Nate Silver (2012) [21] “O sinal é a verdade. O ruído é o que nos distrai, afastando-se da verdade.” Assim, na perspectiva deste estatístico, o sinal é interpretado como “... uma indicação da verdade subjacente a um problema estatístico ou preditivo.” enquanto que “...ruídos são padrões aleatórios que poderiam ser facilmente confundidos com sinais...”.

No momento presente, onde a disponibilidade de informações é gigantesca, um dos objetivos deste trabalho é apresentar alguns dos modelos mais utilizados em séries temporais, considerando os universos acadêmico e corporativo, tanto pelos seus fundamentos teóricos quanto por aplicações empíricas ao mercado financeiro, especificamente realizadas com as taxas de câmbio EUR/BRL e EUR/USD, onde transitou-se entre os modelos de regressão dinâmica e os modelos de volatilidade GARCH, pertencente à classe dos modelos heterocedásticos, no qual mereceu maior atenção.

Explorou-se um cenário multivariado com a participação de alguns índices de ações no modelo de regressão dinâmica bem como realizou-se diferentes abordagens do *forecast* por meio dos modelos GARCH, incluindo simulação e a estimação da densidade (*rolling density forecast*)

além da modelação tradicional, onde explorou-se diferentes distribuições de probabilidade para os erros. Além disso, mostrou-se a aplicação de alguns testes de hipótese para a detecção de volatilidade, assim como o emprego de critérios de informação, como medida de qualidade de ajustamento dos modelos e indicadores de erro, para a avaliação da acurácia do *forecast*.

O modelo **GARCH** mostrou-se mais apropriado para o problema investigado pois consegue acompanhar o movimento da volatilidade das séries de retorno das taxas de câmbio estudadas, embora de forma mais branda que a variação efetivamente observada. Apesar de não se mostrar devidamente acurado no período avaliado, o modelo GARCH pode ser utilizado como uma ferramenta que viabiliza a realização de ensaios probabilísticos e calcula medidas de risco, conseqüentemente, torna-se um importante instrumento de tomada de decisões.

No processo de estimação, nomeadamente no caso dos modelos *ARMA + GARCH*, conduzido através do método de máxima verossimilhança (MLE), enfrentou-se dificuldades computacionais no processo de estimação em função de algumas ocorrências de matriz hessiana singular enquanto que o algoritmo de otimização exige não singularidade para as mesmas, para o cômputo da matriz de covariância dos estimadores MLE. Este problema foi contornado limitando a abrangência do conjunto de “possíveis valores” dos parâmetros.

4.2 Trabalho futuro

Como trabalho futuro, pode-se aprimorar este estudo implementando outros modelos e explorando variantes do modelo GARCH e/ou outras parametrizações, tais como:

- Redes Neurais (*Neural Network*);
- Volatilidade Estocástica (SV);
- Vetor Autoregressivo (VAR);
- EGARCH (*Exponential GARCH*);
- GJR-GARCH (Glosten-Jagannathan-Runkle GARCH);
- testar outras distribuições de probabilidade para os erros, a exemplo da distribuição GED (distribuição do erro generalizada) e outros algoritmos, no processo de estimação.

Bibliografia

- [1] T. Bollerslev. “Generalized autoregressive conditional heteroscedasticity”. Em: *J. Econ.* 31 (1986), pp. 307–327.
- [2] G. Box e G. Jenkins. *Time Series Analysis, Forecasting and Control*. Holden-Day, 1970.
- [3] P. Brockwell e R. Davis. *Introduction to Time Series and Forecasting*. Third edition. Springer Texts in Statistics. Springer, 2016.
- [4] J. Campbell, A. Lo e C. Mackinlay. *The Econometrics of Financial Markets*. Princeton University Press, 1997.
- [5] G. Casella e R. L. Berger. *Statistical Inference*. Second edition. Duxbury Thomson Learning, 2002.
- [6] P. S. P. Cowpertwait e A. V. Metcalfe. *Introductory Time Series with R*. Springer, 2011.
- [7] J. D. Cryer e K.-S. Chan. *Time Series Analysis with Applications in R*. Second edition. Springer, 2008.
- [8] W. Enders. *Applied Econometric Time Series*. Fourth edition. Wiley, 2015.
- [9] R. F. Engle. “Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation”. Em: *Econometrica* 50 (1982), pp. 987–1007.
- [10] W. Feller. *An Introduction to Probability Theory and Its Applications*. Third edition. Vol. I. John Wiley e Sons, Inc., 1950.
- [11] A. Ghalanos. *Package rugarch*. Jul. de 2020. URL: <https://CRAN.R-project.org/package=rugarch>.
- [12] D. M. Hawkins. *Identification of Outliers*. Chapman e Hall, 1980.

- [13] R. J. Hyndman e G. Athanasopoulos. *Forecasting: Principles and Practice*. Second edition. Monash University, Australia, 2018.
- [14] S. A. Klugman, H. H. Panjer e G. E. Willmot. *Loss models : from data to decisions*. Fifth edition. John Wiley e Sons, Inc., 2019.
- [15] Denis Kwiatkowski et al. “Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?” Em: *Journal of Econometrics* 54 (1992), pp. 159,178. DOI: [https://doi.org/10.1016/0304-4076\(92\)90104-Y](https://doi.org/10.1016/0304-4076(92)90104-Y).
- [16] Alan Pankratz. *Forecasting with Dynamic Regression Models*. John Wiley e Sons, Inc., 1991.
- [17] M. S. Perlin et al. “A GARCH Tutorial with R”. Em: *Revista de Administração Contemporânea* 25.1 (2020). DOI: <https://doi.org/10.1590/1982-7849rac2021200088>.
- [18] S.-H. Poon. *A Practical Guide to Forecasting Financial Market Volatility*. John Wiley e Sons, Inc., 2005.
- [19] J. A. Rice. *Mathematical Statistics and Data Analysis*. Third edition. Thomson Brooks/Cole, 2007.
- [20] R. H. Shumway e D. S. Stoffer. *Time Series Analysis and Its Applications*. Fourth edition. Springer, 2017.
- [21] N. Silver. *O sinal e o ruído: por que tantas previsões falham e outras não*. Editora Intrínseca Ltda, 2012.
- [22] R. S. Tsay. *Analysis of Financial Time Series*. John Wiley e Sons, Inc., 2002.
- [23] R. S. Tsay. *Analysis of Financial Time Series*. Third edition. John Wiley e Sons, Inc., 2010.
- [24] J. W. Tukey. *Exploratory Data Analysis*. Addison-Wesley, 1977.
- [25] D. Wuertz. *Package fGarch*. Mar. de 2020. URL: <https://CRAN.R-project.org/package=fGarch>.

- [26] Diethlem Wurtz, Yohan Chalabi e Ladislav Luksan. “Parameter Estimation of ARMA Models with GARCH APARCH Errors An R and SPlus Software Implementation”. Em: *Journal of Statistical Software* (2006).
- [27] E. Zivot e J. Wang. *Modeling Financial Time Series with S-Plus*. Springer, 2003.

Apêndices

Apêndice A

Métodos de estimação

Neste capítulo, estão descritos dois métodos de estimação amplamente utilizados: **Mínimos quadrados**, que é o método padrão utilizado nos *softwares* de modelação estatística para modelos de regressão e ARIMA e **Máxima verossimilhança**, adotado tipicamente, segundo Shumway e Stoffer (2017) [20], na estimação dos parâmetros dos modelos ARCH e GARCH (máxima verossimilhança condicional).

A.1 Mínimos quadrados

O método de estimação por mínimos quadrados, conhecido por *Least squares estimation*, baseia-se no princípio de estimar o valor dos coeficientes de um determinado modelo de modo que a soma dos quadrados dos erros (ou resíduos) assumam valor mínimo. Assim, diante de um modelo de regressão onde pretende-se explicar a variável dependente y_t através das variáveis independentes $x_{1,t}, x_{2,t}, x_{3,t}, \dots, x_{k,t}$, tem-se:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \dots + \beta_k x_{k,t} + \varepsilon_t ,$$

e portanto, os resíduos ε_t , são descritos por:

$$\varepsilon_t = y_t - \beta_0 - \beta_1 x_{1,t} - \beta_2 x_{2,t} - \dots - \beta_k x_{k,t} .$$

O método calcula o valor dos estimadores $\hat{\beta}_0, \hat{\beta}_1, \dots, \hat{\beta}_k$, referentes, respectivamente, aos coeficientes $\beta_0, \beta_1, \dots, \beta_k$, de modo que $\sum_{t=1}^T \varepsilon_t^2$ tenha valor mínimo, isto é, para

$$\min \sum_{t=1}^T \varepsilon_t^2 = \min \sum_{t=1}^T (y_t - \beta_0 - \beta_1 x_{1,t} - \beta_2 x_{2,t} - \dots - \beta_k x_{k,t})^2 .$$

De acordo com Shumway e Stoffer (2017) [20], os estimadores de mínimos quadrados são

não enviesados, isto é, $E(\hat{\beta}) = \beta$ e possuem a menor variância dentre os estimadores da classe linear não enviesado. Caso os resíduos ε_t tenham distribuição normal, o estimador $\hat{\beta}$ obtido por mínimos quadrados coincide com o estimador produzido pelo método de máxima verossimilhança, que é apresentado na sequência.

A.2 Máxima verossimilhança

O método de estimação por máxima verossimilhança, comumente referido por MLE (*Maximum likelihood estimation*), determina as estimativas baseado na maximização de uma função L , denominada função de verossimilhança. Segundo Casella e Berger (2002) [5], sua definição pode ser compreendida considerando uma amostra $X = (X_1, X_2, \dots, X_n)$, dado que $X = x$ é observado, a função de verossimilhança de θ é definida por:

$$L(\theta|x) = f(x|\theta) ,$$

onde θ , representa os parâmetros e $f(x|\theta)$, a função de densidade de probabilidade conjunta de X . Se X é um vetor aleatório discreto, tem-se $L(\theta|x) = P_\theta(X = x)$. Assim, o método MLE atesta que $\theta = \theta_1$ é mais plausível que $\theta = \theta_2$ caso $L(\theta_1|x) > L(\theta_2|x)$, que é equivalente a comparar as suas probabilidades de ocorrência, ou seja, diante dos possíveis valores θ_1 e θ_2 para θ , dá-se preferência na escolha de θ_1 , desde que se verifique a condição

$$P_{\theta_1}(X = x) > P_{\theta_2}(X = x) .$$

Para comunicar a ideia central deste método de forma mais concreta, segue um exemplo extraído do livro de Tsay (2010) [23] onde considera-se uma série temporal dos retornos de um certo ativo financeiro representada por r_t , com n observações, isto é, r_1, r_2, \dots, r_n . Se a distribuição condicional $f(r_1, r_2, \dots, r_n|\theta)$, é Normal de média μ_t e variância σ_t^2 então θ consiste nos parâmetros μ_t e σ_t^2 e a função de verossimilhança dos dados é definida por:

$$L(\theta|r_1, r_2, \dots, r_n) = f(r_1, r_2, \dots, r_n|\theta) = f(r_1|\theta) \prod_{t=2}^n \frac{1}{\sqrt{2\pi} \sigma_t} \exp \left[\frac{-(r_t - \mu_t)^2}{2\sigma_t^2} \right],$$

onde $f(r_1|\theta)$ é a função densidade marginal da primeira observação. O método MLE calcula θ , de modo que a maximizar a função $L(\theta|r_1, r_2, \dots, r_n)$. Como o logaritmo é uma função crescente, $\max L \equiv \max \ln L$, sendo a função $\ln L$ mais conveniente e fácil de tratar. Desta maneira, o

valor de θ é determinado pela maximização de:

$$\ln f(r_1, r_2, \dots, r_n | \theta) = \ln f(r_1 | \theta) - \frac{1}{2} \sum_{t=2}^n \left[\ln(2\pi) + \ln(\sigma_t^2) + \frac{(r_t - \mu_t)^2}{\sigma_t^2} \right].$$

Alguns algoritmos presentes nos *softwares* estatísticos trocam este procedimento de maximização pela minimização da função logarítmica de verossimilhança negativa, isto é, θ é determinado para $\min [-\ln f(r_1, r_2, \dots, r_n | \theta)]$.

Para a obtenção da função de máxima verossimilhança para $f(r_1, r_2, \dots, r_n | \theta)$ para outra distribuição de probabilidade condicional, procede-se de forma similar, adequando-a à função de densidade correspondente.

Os estimadores MLE, que estão denotados por $\hat{\theta}_{ML}$, são atrativos pelo fato de apresentarem algumas propriedades assintóticas, que sob certas condições de regularidade (algo como a existência das derivadas parciais finitas até a 3ª ordem de L). As derivadas parciais de 1ª ordem, em relação a θ , da função logarítmica de verossimilhança $\ln L(\theta|x)$ define o vetor gradiente $g(\theta|x)$, de dimensão $q \times 1$, onde q representa o número de parâmetros de θ , isto é, $\theta = \theta_1, \theta_2, \dots, \theta_q$. Assim,

$$g(\theta|x) = \frac{\partial \ln L(\theta|x)}{\partial \theta}.$$

Os possíveis candidatos para MLE são solução da equação $g(\theta|x) = 0$, que é uma condição necessária para a obtenção do máximo, mas não uma condição suficiente. Através das derivadas parciais da função logarítmica de verossimilhança, também em relação a θ , forem de 2ª ordem, obtém-se a matriz hessiana $H(\theta|x)$, de dimensão $q \times q$, definida por:

$$H(\theta|x) = \frac{\partial}{\partial \theta} \left(\frac{\partial \ln L(\theta|x)}{\partial \theta} \right)' = \frac{\partial^2 \ln L(\theta|x)}{\partial \theta \partial \theta'},$$

que é base para a matriz de informação de Fisher, definida por: $I(\theta|x) = -E[H(\theta|x)]$. Esta matriz hessiana tem relevância na classificação dos pontos críticos, pois busca-se o máximo global, e também no cálculo da matriz de covariância dos parâmetros. $H(\theta|x)$ exerce papel importante nos algoritmos numéricos de otimização e pode-se apresentar como uma restrição na execução do algoritmo, a exemplo dos casos onde não é uma matriz inversível.

Suas principais propriedades podem ser descritas por:

- **Consistência:** $\hat{\theta}_{ML}$ é consistente pois converge em probabilidade para θ , se n , que representa

a dimensão da amostra, tende a infinito. Assim, tem-se:

$$\lim_{n \rightarrow \infty} P(\hat{\theta}_{ML}) = \theta ;$$

- **Normalidade assintótica:** Quando n tende a infinito, a distribuição de $\hat{\theta}_{ML}$ aproxima-se da Normal de média θ e variância $I(\theta|x)^{-1}$, isto é,

$$\lim_{n \rightarrow \infty} \hat{\theta}_{ML} \sim N(\theta, I(\theta|x)^{-1}) .$$

- **Eficiência assintótica:** Um estimador é eficiente se não é enviesado e tem variância mínima para todos os parâmetros. $\hat{\theta}_{ML}$ atinge o limite inferior de Cramér-Rao para os estimadores, que estabelece que $Var(\hat{\theta}) \geq I(\theta|x)^{-1}$.
- **Invariância:** MLE é invariante para certas transformações nos dados.

As funções de probabilidade costumam ser complicadas e não há a possibilidade de obter a solução ótima dos estimadores analiticamente. Nestas condições, recorre-se a algoritmos para maximizar numericamente a função de verossimilhança, que estão incorporados nos *softwares* estatísticos *softwares* estatísticos. Para ilustrar a dificuldade em realizar o processo de estimação pelo método MLE em modelos ARCH ou GARCH, que utiliza relações não lineares, foi extraído um exemplo do livro de Enders (2015) [8], conforme descrito como continuidade desta seção. Suponha que os valores da sequência $\{\varepsilon_t\}$ tenha distribuição Normal de média nula e variância constante σ^2 , assim, $\varepsilon_t \sim N(0, \sigma^2)$. A função de verossimilhança para qualquer realização de ε_t é definida por:

$$L_t = \frac{1}{\sqrt{2\pi} \sigma} \exp \left[\frac{-\varepsilon_t^2}{2\sigma^2} \right] .$$

Como as realizações de $\{\varepsilon_t\}$ são independentes, a função de verossimilhança das realizações conjuntas de $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$, é definida pelo produto das verossimilhanças individuais, e portanto, descrita por:

$$L = \prod_{t=1}^n \frac{1}{\sqrt{2\pi} \sigma} \exp \left[\frac{-\varepsilon_t^2}{2\sigma^2} \right] ,$$

que transformada em função logarítmica e manejada de forma mais tratável, é escrita por:

$$\ln L = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=1}^n [\varepsilon_t]^2 . \tag{A.1}$$

Neste exemplo, suponha que $\{\varepsilon_t\}$ represente os resíduos de um modelo de regressão, onde assume-se média nula, variância constante σ^2 e que as várias realizações de $\{\varepsilon_t\}$ são indepen-

dentos. Como $y_t = \beta x_t + \varepsilon_t$, o termo ε_t pode ser isolado e representado por:

$$\varepsilon_t = y_t - \beta x_t, \quad (\text{A.2})$$

Substituindo ε_t dado por A.2 na equação A.1, tem-se a equação da função logarítmica de verossimilhança, dado uma amostra de n observações, que é descrita por:

$$\ln L = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma^2 - \frac{1}{2\sigma^2} \sum_{t=1}^n [y_t - \beta x_t]^2, \quad (\text{A.3})$$

que precisa ser maximizada para a obtenção dos estimadores MLE. A primeira derivada parcial desta função (A.3) em relação a σ^2 resulta em:

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2\sigma^2} \ln(2\pi) + \frac{1}{2\sigma^4} \sum_{t=1}^n [y_t - \beta x_t]^2, \quad (\text{A.4})$$

enquanto que, se sua primeira derivada for em relação a β , o resultado obtido é expresso por:

$$\frac{\partial \ln L}{\partial \beta} = \frac{1}{\sigma^2} \sum_{t=1}^n [x_t y_t - \beta x_t^2], \quad (\text{A.5})$$

Resolvendo as equações $\frac{\partial \ln L}{\partial \sigma^2} = 0$ e $\frac{\partial \ln L}{\partial \beta} = 0$, obtém-se, respectivamente, os estimadores da variância ($\hat{\sigma}^2$) e de β ($\hat{\beta}$) obtidos pelo conhecido método *ordinary least squares*, mais comumente mencionado pela sua sigla OLS e que produz estimadores consistentes e assintoticamente eficientes. Estes cálculos determinam os seguintes resultados:

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n \varepsilon_t^2, \quad (\text{A.6})$$

$$\hat{\beta} = \frac{\sum_{t=1}^n x_t y_t}{\sum_{t=1}^n x_t^2}. \quad (\text{A.7})$$

A complexidade da estimação MLE aumenta nos modelos ARCH e GARCH pela introdução de componentes não lineares, cujas soluções dependem da aplicação computacional de métodos numéricos. Ao introduzir um erro $ARCH(1)$ v_t ao modelo de regressão A.2, que ainda continua válido, tem-se que $\varepsilon_t = v_t \sigma_t$, onde σ_t^2 é a variância condicional de ε_t e que não é constante, a função de verossimilhança das realizações conjuntas de $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$ sofre natural modificações e é expressa por:

$$L = \prod_{t=1}^n \frac{1}{\sqrt{2\pi} \sigma_t} \exp \left[\frac{-\varepsilon_t^2}{2\sigma_t^2} \right], \quad (\text{A.8})$$

que transformada em função logarítmica representa-se por:

$$\ln L = -\frac{n}{2} \ln(2\pi) - \frac{n}{2} \ln \sigma_t^2 - \frac{1}{2} \sum_{t=1}^n \left[\frac{\varepsilon_t^2}{\sigma_t^2} \right]. \quad (\text{A.9})$$

Pelo modelo $ARCH(1)$, como a variância condicional de ε_t é descrita por $\sigma_t^2 = \omega + \alpha_1 \varepsilon_{t-1}^2$, pode-se substituí-la na equação A.9, resultando na expressão:

$$\ln L = -\frac{n-1}{2} \ln(2\pi) - \frac{1}{2} \sum_{t=2}^n \ln(\omega + \alpha_1 \varepsilon_{t-1}^2) - \frac{1}{2} \sum_{t=2}^n \left[\frac{(y_t - \beta x_t)^2}{(\omega + \alpha_1 \varepsilon_{t-1}^2)} \right]. \quad (\text{A.10})$$

A primeira observação é perdida pois ε_0 não integra a amostra. Realizando a substituição de ε_{t-1}^2 por $\varepsilon_{t-1}^2 = (y_{t-1} - \beta x_{t-1})^2$ em A.10, consegue-se maximizar $\ln L$ em relação aos parâmetros ω , α_1 e β somente com a ajuda dos algoritmos computacionais, uma vez que não existem soluções analíticas para as condições de primeira ordem (a primeira derivada da função de verossimilhança deve ser zero) para determinação de um máximo.

Apêndice B

Aplicação e Resultados

Este apêndice contém algumas análises complementares, realizado no componente prático do estudo, com menor relevância.

Considerando o “melhor” modelo $ARMA + GARCH$ conseguido para explicar o retorno da taxa de câmbio EUR/BRL, obteve-se o modelo de $ARMA(1,0) + GARCH(1,1)$ com a distribuição de erros *t-Student* assimétrica e descrito pelas equações 3.5, 3.6 e 3.7. Apropriando-se deste modelo, com todas as suas características, foram geradas 5000 simulações diferentes para um horizonte de 1000 dias e para cada uma delas, um nova série temporal de tamanho 1000 períodos é gerada, que colocadas em um único gráfico (ver figura B.1), exhibe o padrão mais frequente para o retorno da taxa de câmbio EUR/BRL, observado pelas regiões mais escuras, dado que há vários pontos sobrepostos. As cores em si não expressam nenhum significado adicional à figura e servem exclusivamente para explicitar as diferentes séries temporais.

B.1 Taxa de câmbio EUR/BRL

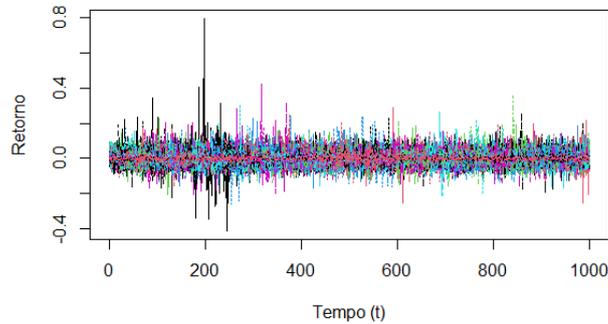


Figura B.1: Retorno da taxa de câmbio EUR/BRL, obtido por realização de 5000 simulações do modelo $ARMA(1, 0) + GARCH(1, 1)$, para um horizonte de 1000 dias.

A partir deste procedimento de simulações realizado, pode-se observar também o padrão gerado para a volatilidade do retorno da taxa de câmbio EUR/BRL (ver figura B.2).

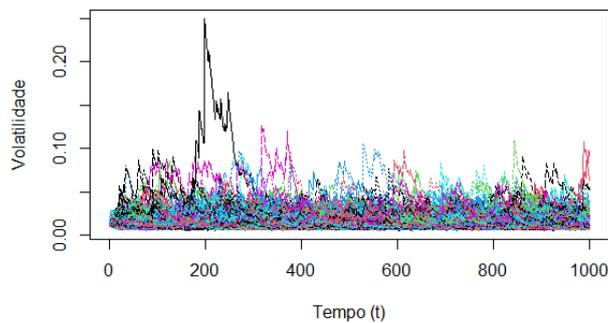


Figura B.2: Volatilidade do retorno da taxa de câmbio EUR/BRL, obtida por realização de 5000 simulações do modelo $ARMA(1, 0) + GARCH(1, 1)$, para um horizonte de 1000 dias.

Nota-se visualmente que há uma forte concentração da volatilidade nos valores mais baixos (zona escura da figura B.2). A partir dos valores gerados, pode-se realizar exercícios probabilísticos, como por exemplo, calcular que $Prob(\sigma_t > \approx 0.025) = 1\%$, considerando todos os períodos.

B.2 Taxa de câmbio EUR/USD

As imprecisões no ajustamento do “melhor“ modelo às caudas foram também investigadas para elucidar se podem ser explicadas pela presença dos *outliers*. Desta forma, procedeu-se uma nova modelação, mantendo-se todos os parâmetros e removendo-se os *outliers* identificados da amostra (7 observações, conforme descrito na seção 3.3.2). Nesta perspectiva, novos gráficos de densidade e quantis dos resíduos foram gerados (ver figura B.3), onde, por inspeção visual, nota-se uma melhoria no ajustamento na cauda esquerda (valores extremos de retornos negativos), em contrapartida, neste novo cenário considerando a remoção dos *outliers*, a imprecisão no ajuste da cauda direita foi intensificada. Este exercício mostrou que o modelo tem limitações para responder à ocorrência de valores extremos, embora possa ser tratado como aceitável.

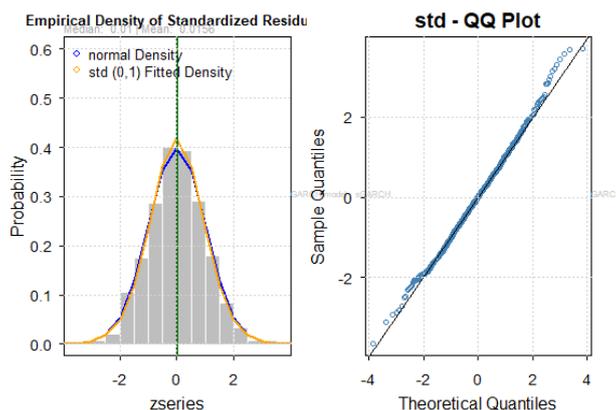


Figura B.3: Densidade e quantis dos resíduos padronizados da taxa de câmbio EUR/USD, obtidos por modelo $ARMA(0,0) + GARCH(1,1)$ em comparação aos quantis teóricos (*QQ Plot*), sem os *outliers*.

Como “melhor“ modelo $ARMA + GARCH$ identificado para explicar o retorno da taxa de câmbio EUR/USD, obteve-se o modelo de $ARMA(0,0) + GARCH(1,1)$ com distribuição *t-Student* para os erros ($\varepsilon_t \sim std(0,1)$) e descrito pelas equações 3.8, 3.9 e 3.10.

Na abordagem do método *rolling density forecast* a partir deste modelo, foi possível também abordar a medida de risco VaR (*Value-at-Risk*).

O VaR pode ser interpretado como o valor de perda máximo associado à probabilidade de ocorrência de um acontecimento. Para além disso, uma vez que se trata de uma estimativa,

está associado a um certo nível de confiança. Por outras palavras, pode então dizer-se que, ao nível de confiança $1 - \alpha$, o VaR é o valor da perda máxima com probabilidade α . Desta forma, a medida VaR pode ainda ser vista como a inversa da função de distribuição (ou função de quantil), $F_S^{-1}(1 - \alpha)$, onde S representa um determinado risco.

Nota-se uma discreta variação da volatilidade estimada assim como a não ocorrência de retornos abaixo da medida VaR (*Value-at-Risk*), a um $\alpha = 1\%$, considerando o horizonte de 6 dias (ver figura B.4).

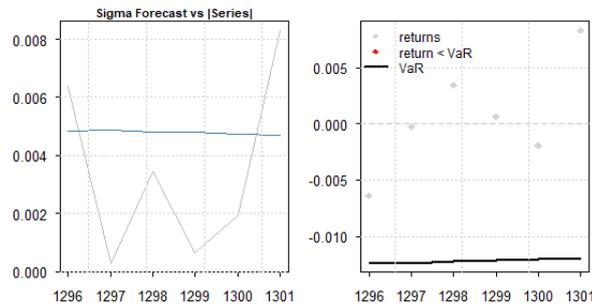


Figura B.4: Volatilidade e retorno da taxa de câmbio EUR/USD, estimado pelo método *rolling density forecast*, obtida a partir de modelo $ARMA(0, 0) + GARCH(1, 1)$, distribuição *t-Student* dos erros para um horizonte de 6 dias.

Apêndice C

Instruções do *R Software*

Este relatório foi escrito em linguagem de programação LaTeX, por meio do uso do editor *TeXstudio*, versão 4.0.4, sendo que a execução do trabalho, contou com um apoio pontual do *software Microsoft Office 365* e fundamentalmente dos *softwares RStudio*, versão 1.4.1106 (11/02/2021) e **R** versão 4.0.4 (15/02/2021), usados em plataforma Windows, sistema operacional de 64 bits e na edição *Windows 10 Home*. Este apêndice, possui foco exclusivo no *R Software*, e tem o propósito de mencionar as bibliotecas usadas, exibir as principais utilizadas, englobando as funções desenvolvidas (criadas e/ou adaptadas pelo autor) e as funções já disponíveis, que aparecem citadas por tema; além de fazer menção a alguns aspectos relevantes à linguagem de programação R que surgiram no transcorrer do trabalho, nomeadamente na parte da aplicação empírica.

C.1 Bibliotecas

Nesta seção, aparecem todas as bibliotecas do *do R Software* utilizadas na realização das aplicações empíricas, considerando as modelações da taxa de câmbio EUR/BRL e taxa de câmbio EUR/USD. Essas bibliotecas foram demandadas em algum estágio do estudo, mesmo que não haja uma relação direta com o material apresentado (relatório da dissertação).

```
# bibliotecas requeridas
library ( astsa )
library ( car )
library ( corrplot )
library ( diagram )
```

```
library (dplyr)
library (dynlm)
library (expm)
library (fGarch)
library (FinTS)
library (forecast)
library (ggplot2)
library (ggridges)
library (knitr)
library (lmtest)
library (markovchain)
library (MASS)
library (MLmetrics)
library (MIS)
library (nlme)
library (PerformanceAnalytics)
library (pracma)
library (psych)
library (QuantPsyc)
library (quantmod)
library (readr)
library (rmgarch)
library (Rsolnp)
library (rugarch)
library (stats)
library (summarytools)
library (texreg)
library (tsm)
library (tidyquant)
library (tidyverse)
library (tseries)
library (xts)
library (zoo)
```

Listing C.1: Bibliotecas

C.2 Funções criadas

Nesta parte do trabalho, estão exibidas as instruções do *R Software* (código em linguagem de programação) das principais funções desenvolvidas (criadas e/ou adaptadas pelo autor) especificamente para atender as necessidades requeridas por este estudo.

(1) Identificação de *outliers*

```

# funcao para a identificacao de outliers
# argumentos da funcao: x: dados, c: constante (sera multiplicada pelo IQR)
oliers = function (x,c) {

  na.rm = TRUE
  qnt <- quantile (x , probs=c ( .25 , .75) , na.rm = na.rm)

  # amplitude do intervalo
  H <- c*IQR( x , na.rm = na.rm)

  y <- x
  y[x < ( qnt [1] - H) ] <- NA
  y[x > ( qnt [2] + H) ] <- NA
  y

  # definicao do intervalo
  lim <- c((qnt [1] - H) ,(qnt [2] + H))

  # identificacao dos outliers (valores fora do intervalo)
  pos <- which(is.na(y))
  olx <- x[pos]

  # % de outliers detectados
  tam <- length(pos)
  perc <- round(100*tam/length(y),1)

  return(list(lim = lim , pos = pos , olx = olx , tam=tam, perc=perc))

}

```

Listing C.2: outliers

(2) Gráfico de ACF/PACF

```
# grafico das funcoes acf e pacf customizadas
func_acf<-function(x,lagm,na.action=na.pass,lags=1:288,...){
  #lags:horizonte dos lags
  plot(acf(x,lag.max=lagm,na.action=na.action,plot=F)[lags],lwd=5,lend=2,col="brown"
  ,...)
  axis(1,tck=1,col="lightgrey",lty="dotted")
  plot(pacf(x,lag.max=lagm,na.action=na.action,plot=F)[lags],lwd=5,lend=2,col="brown"
  ,...)
  axis(1,tck=1,col="lightgrey",lty="dotted")
}
```

Listing C.3: Gráfico ACF/PACF personalizado.

(3) Transformação da série (inversa do retorno)

```
# funcao que gera serie a escala da serie original (x) partindo da serie de retornos
# a partir de um valor de referencia da serie original (y)
ret2orig = function (ret,y) {
  logx <- diffinv(ret,xi=log(y))
  x <- exp(logx)
  return(x)
}
```

Listing C.4: Função usada para transformar os valores modelados em retorno para a escala da série original.

(4) Estimação de modelo GARCH (função adaptada de [17]).

```
# estimacao dos modelos GARCH (uso do package fGarch)
estimate_garch <- function(ar_lag,
ma_lag,
arch_lag,
garch_lag,
models_to_estimate,
distribution_to_estimate) {
```

```

message('Estimating ARMA(', ar_lag, ', ', ma_lag, ')', '- ',
'(', arch_lag, ', ', garch_lag, ') ',
'dist = ', distribution_to_estimate)

if (models_to_estimate=="garch") {
  f=as.formula(paste("m.eb ~ arma(", ar_lag, ", ", ma_lag, ") + garch(", arch_lag, ", ",
garch_lag, ")"))
} else {
  f=as.formula(paste("m.eb ~ arma(", ar_lag, ", ", ma_lag, ") + aparch(", arch_lag, ", ",
garch_lag, ")"))
}

my_garch <- garchFit(f, data = m.eb, cond.dist = distribution_to_estimate,
trace = F, include.mean=F)
return(my_garch)
}

```

Listing C.5: Função utilizada para a estimação de modelos GARCH (*package fGarch*).

(5) Nomeação de modelo GARCH

```

#nomeacao dos modelos
ref_garch <- function(ar_lag,
ma_lag,
arch_lag,
garch_lag,
models_to_estimate,
distribution_to_estimate) {
  model_ref <- paste (models_to_estimate, ar_lag, ma_lag, arch_lag,
garch_lag, distribution_to_estimate)
  model_name <- paste ("arma(", ar_lag, ", ", ma_lag, ")+" , models_to_estimate, "(", arch_lag,
", ",
garch_lag, ") distr. ", distribution_to_estimate)
  return(model_name)
}

```

Listing C.6: Função de nomeia os modelos GARCH gerados de acordo com a parametrização escolhida.

(6) Geração de base de dados para construção de gráfico

```
#funcao que gera base para grafico de densidade do forecast em cada instante
func_gera_df_plot<-function(mu,sigma,n){
  N=length(sigma) #tam do horizonte do forecast
  df = as.data.frame(cbind(rep("T",n),rnorm(n,mu[1],sigma[1])))
  for (i in seq(2,N)){
    hor=paste("T +",str_pad(i-1,2,pad = "0")) #label do instante
    aux <- as.data.frame(cbind(rep(hor,n),rnorm(n,mu[i],sigma[i])))
    df <- bind_rows(df,aux)
  }
  colnames(df)<- c("tempo","x")
  df$x<-as.numeric(df$x)
  return(df)
}
```

Listing C.7: Função que realiza a geração da base de dados para o gráfico das densidades estimadas (*GARCH rolling density forecast*) a partir de modelo ARMA+GARCH (figura 3.35).

(7) Identificação dos coeficientes significativos

```
#funcao que gera vetor que indica, com valor 1, os coef. significativos (valor 1)
who_signif <- function(coef) {
  n = NROW(coef)
  pos <- as.numeric(which(coef[,5]<=0.05)) #par. significativos (5%)
  v = rep(0,n) # vetor nulo
  for (i in pos) {
    v[i]=1
  }
  return(v)
}
```

Listing C.8: Função que indica os coeficientes significativos a partir de uma lista de resultados de múltiplos modelos

C.3 Principais funções utilizadas

Nesta seção, estão discriminadas as principais funções do *R Software*, acessíveis a qualquer usuário, utilizadas para viabilizar o componente prático deste estudo. Estas funções aparecem em diferentes tabelas, relacionadas às suas respectivas bibliotecas, e estão organizadas pelos seguintes temas:

- Testes de hipótese (ver tabela C.1);
- *Forecasting* (ver tabela C.2);
- Modelação e *forecast* GARCH (ver tabela C.3).

instrução	biblioteca
t.test	stats
adf.test	tseries
kpss.test	tseries
Box.test	stats
ArchTest	FinTS
shapiro.test	stats

Tabela C.1: Funções do *R Software* utilizadas para testes de hipótese

instrução	biblioteca
tslm	forecast
autoarima	forecast
Arima	forecast
forecast	forecast
accuracy	forecast
MAPE	MLmetrics

Tabela C.2: Funções do *R Software* utilizadas para *Forecasting*.

instrução	biblioteca
garchFit	fGarch
ugarchspec	rugarch
ugarchfit	rugarch
ugarchforecast	rugarch
ugarchsim	rugarch
ugarchroll	rugarch

Tabela C.3: Funções do *R Software* utilizadas na modelação e *forecast* GARCH.

A função “stl“, que utiliza o método STL (*Seasonal and Trend decomposition using Loess*), pertence à biblioteca “stats“, utilizada para a realização da decomposição das séries temporais em seus componentes de tendência e sazonalidade, não aparece nas tabelas anteriores (C.1, C.2 e C.3), embora também mereça ser mencionada.

Apêndice D

Estrutura das tabelas de dados

Todos os conjuntos de dados (*datasets*) utilizados neste estudo possuem a mesma estrutura “bruta”, foram extraídos da mesma fonte: portal Yahoo Finance e encontram-se disponíveis em <https://finance.yahoo.com>. A relação de todos estes documentos (cada documento refere-se à uma única série temporal) aparece na tabela D.1 enquanto que a estrutura padrão (conteúdo integral disponível, embora o estudo tenha demandado parte), está representada pela tabela D.2.

Nome do documento	Descrição	Categoria	Formato
EURBRL=X	EUR/BRL	taxa de câmbio	csv
EURUSD=X	EUR/USD	taxa de câmbio	csv
DJI	Dow Jones Industrial Average	Índice de ações	csv
GSPC	S & P 500	Índice de ações	csv
IXIC	NASDAQ Composite	Índice de ações	csv
STOXX50E	Euro STOXX50E Index	Índice de ações	csv
N100	EURONEXT 100	Índice de ações	csv
BVSP	IBOVESPA	Índice de ações	csv

Tabela D.1: Relação dos *datasets* utilizados.

Nome do campo	Descrição	Formato
Date	Data de realização do pregão	chr
Open	Cotação do ativo na abertura do pregão	num
High	Cotação máxima do ativo no dia	num
Low	Cotação mínima do ativo no dia	num
Close	Cotação do ativo no fechamento do pregão	num
Adj.Close	Cotação ajustada do ativo no fechamento do pregão	num
Volume	Volume do ativo negociado no dia, em quantidade	num

Tabela D.2: Conteúdo do *dataset* (estrutura padrão).