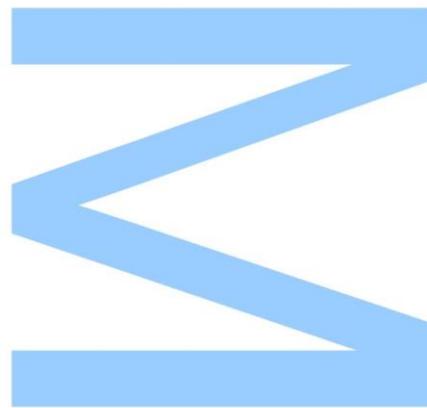


Modelação e Previsão de Preços de Eletricidade

Marina Mendes Pereira
Mestrado em Engenharia Matemática
Departamento de Matemática
2017

Orientador
Laura Cavalcante, INESC TEC

Coorientador
Professora Doutora Margarida Brito, FCUP

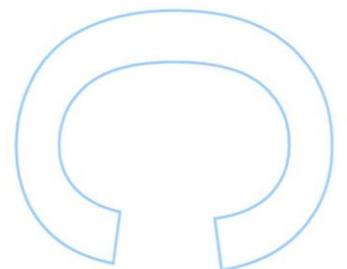
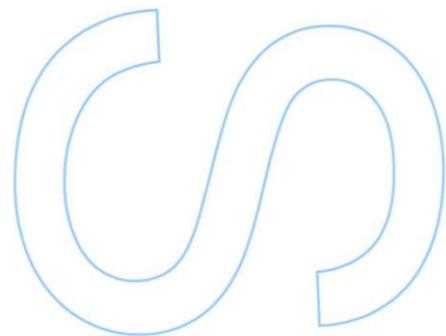
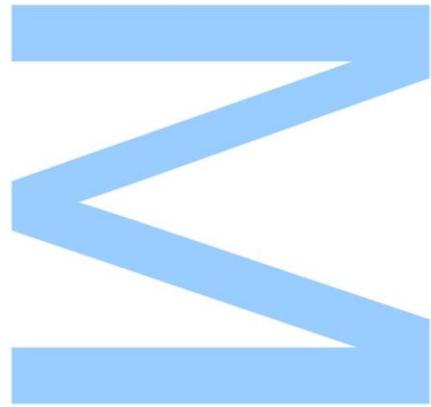




Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, ____ / ____ / ____



Agradecimentos

Ao pilar da minha vida, Tiago, pelo apoio incansável ao longo destes anos. Sem ele nada disto seria possível. Obrigado por nunca duvidares de mim, mesmo quando eu achei que era impossível.

Aos meus pais, pelo apoio incondicional.

Um obrigado muito especial à minha irmã Marlene pelas longas conversas “terapêuticas” e à minha irmã Adriana por todo o amor genuíno que só uma criança pode demonstrar.

A toda a minha restante família pelo apoio e por perdoar as ausências e as “visitas de médico”.

Aos meus amigos “Pombalenses” por estarem sempre lá para um café, um jantar, um ombro amigo... Em especial as minhas queridas Ângela, Filipa e Cindy, apesar de todas as mudanças nas nossas vidas, a amizade permaneceu.

Aos meus amigos “de Coimbra” a quem eu chamo de segunda família: Inês, Fernanda, Peixoto, Romica, Paulo, Bruno e Silvie. Vocês sabem o que significam!

Agradeço também à minha professora de Matemática do Secundário, M^a Helena Oliveira, por ter despertado em mim o gosto pela Matemática.

Agradeço igualmente ao INESC TEC pela disponibilização dos dados e às minhas Orientadoras pelo apoio prestado.

“Sempre chega a hora em que descobrimos que sabíamos muito mais do que antes julgávamos” – *José Saramago*

Resumo

Os preços da eletricidade sofrem picos repentinos e grandes quedas nos seus valores e é importante modelar estas flutuações, implementando uma ferramenta eficaz para gerir o risco de preço da energia. O comportamento dos preços de eletricidade ao longo do tempo pode ser descrito por modelos de séries temporais, que têm em consideração o tempo decorrido entre as suas alterações. A Teoria de Valores Extremos (TVE) providencia-nos uma forma de captar o comportamento das alterações mais bruscas.

O objetivo principal desta dissertação é usar a análise de séries temporais e a TVE para encontrar modelos que possam ser úteis na previsão de acontecimentos extremos no mercado de eletricidade ibérico, abordando a questão da dependência estatística de forma adequada.

Palavras-chave: Preços de Eletricidade, Séries Temporais, Teoria de Valores Extremos, Distribuição Generalizada de Pareto

Abstract

Electricity prices suffer sudden spikes and large declines in their values and it is important to model these fluctuations by implementing an effective tool for manage energy price risk. The behavior of electricity prices over time can be described by time series models, which take into account the time elapsed between their changes. The Extreme Value Theory (EVT) provides a way to capture the behavior of the most sudden changes.

The main goal of this dissertation is to use time series analysis and EVT to find models that can be useful in predicting extreme events in the Iberian electricity market, approaching the issue of statistical dependency in an appropriate way.

Keywords: Electricity Prices, Time Series, Extreme Value Theory, Generalized Pareto Distribution

Conteúdo

1	Introdução	11
2	Metodologia	14
2.1	Introdução	14
2.2	Modelos de Séries Temporais	14
2.3	Modelação de Valores Extremos	17
2.4	Descrição da Metodologia	22
3	Desenvolvimento	24
3.1	Análise descritiva dos dados	24
3.2	Modelação dos retornos	27
3.3	Modelação dos resíduos do Modelo AR-GARCH	30
3.4	Estimação de quantis condicionais e discussão	41
3.5	Avaliação dos Modelos	42
4	Comentários finais	48
	Apêndices	53
A	Ferramentas de Programação	54
B	Modelos ARMA	55

Lista de Tabelas

3.1	Testes de Estacionariedade aplicados nos retornos dos preços de eletricidade	25
3.2	Resumo descritivo dos retornos dos preços da eletricidade	26
3.3	Estimativas para os coeficientes do modelo AR-GARCH com lags 1, 24 e 168	29
3.4	Estimativas para os coeficientes do modelo AR-GARCH com lag 1 e lag 24	29
3.5	Comparação dos Modelos AR-GARCH	30
3.6	N.º Excedências vs Threshold - resíduos normalizados do modelo AR-GARCH com inovações t-Student	31
3.7	N.º Excedências vs Threshold - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas	37
3.8	Estimação do n.º de excedências para diferentes probabilidades	41
3.9	Estimativas para os coeficientes do modelo AR-GARCH (Ano 2014): lag 1, 24 e 168	42

Lista de Figuras

2.1	Método Máximo por Blocos	20
2.2	Método <i>Peaks-Over-Threshold</i>	20
3.1	Preços da eletricidade	25
3.2	QQ-Plot Retornos dos preços da eletricidade	26
3.3	ACF Retornos (esquerda) e PACF dos retornos (direita)	27
3.4	ACF resíduos normalizados (esquerda) e do quadrado dos resíduos normalizados (direita) - Modelo AR-GARCH com inovações t-Student	31
3.5	Gráfico dos Resíduos (esquerda) e gráfico da sua volatilidade (direita) - Modelo AR-GARCH com inovações t-Student	31
3.6	Hill Plot dos resíduos normalizados - Modelo AR-GARCH com inovações t-Student	32
3.7	Método de Eyeball - resíduos normalizados do Modelo AR-GARCH com inovações t-Student	32
3.8	Mean Excess Plot dos resíduos normalizados do Modelo AR-GARCH com inovações t-Student	33
3.9	GPD para $N_u = 8585$ (25% das observações) - resíduos normalizados do modelo AR-GARCH com inovações t-Student	33
3.10	GPD para $N_u = 10320$ (30% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações t-Student	34
3.11	GPD para $N_u = 12019$ (35% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações t-Student	34
3.12	GPD para $N_u = 13736$ (40% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações t-Student	34
3.13	GPD para $N_u = 15454$ (45% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações t-Student	35
3.14	GPD para $N_u = 17171$ (50% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações t-Student	35
3.15	ACF resíduos normalizados - Modelo AR-GARCH com inovações gaussianas	36
3.16	Resíduos - Modelo AR-GARCH com inovações gaussianas	36
3.17	Hill Plot resíduos normalizados - Modelo AR-GARCH com inovações gaussianas	37
3.18	Método de Eyeball - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas	37

3.19 Mean Excess Plot - resíduos normalizados do modelo AR-GARCH com inovações gaussianas	38
3.20 GPD para $N_u = 8585$ (25% das observações) - resíduos normalizados do modelo AR-GARCH com inovações gaussianas	38
3.21 GPD para $N_u = 10320$ (30% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas	39
3.22 GPD para $N_u = 12019$ (35% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas	39
3.23 GPD para $N_u = 13736$ (40% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas	40
3.24 GPD para $N_u = 15454$ (45% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas	40
3.25 GPD para $N_u = 17171$ (50% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas	40
3.26 Hill Plot dos resíduos normalizados - Modelo AR-GARCH com inovações gaussianas (Ano 2014)	43
3.27 Mean Excess Plot - resíduos normalizados do modelo AR-GARCH com inovações gaussianas (Ano 2014)	43
3.28 Método de Eyeball - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas (Ano 2014)	44
3.29 GPD para $N_u = 2538$ (30% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas (Ano 2014)	44
3.30 Diagrama de <i>reliability</i>	45
3.31 Diagrama de <i>reliability</i> - Zoom	46
3.32 Diagrama de <i>Sharpness</i>	47

Lista de Abreviaturas

ACF – Autocorrelation Function

ADF - Augmented Dickey Fuller Teste

AIC - Akaike's Information Criterion

AR – Autoregressive

ARCH - Autoregressive Conditional Heteroskedasticity

ARIMA – Autoregressive Integrated Moving Average

ARMA - Autoregressive Moving Average

BIC - Bayesian Information Criterion

CH - Condicionalmente Heterocedásticos

GARCH – Generalized Autoregressive Conditional Heteroskedasticity

GEV - Generalized Extreme Value

GPD - Generalized Pareto Distribution

LB - Ljung - Box Test

KPSS - Kwiatkowski–Phillips–Schmidt–Shin Test

MB – Máximo por Blocos

MEF – Mean Excess Plot

PACF – Partial Autocorrelation Function

POT - Peaks- Over-Threshold

PP – Phillips-Perron Test

TVE - Teoria dos Valores Extremos

Capítulo 1

Introdução

O mercado da eletricidade ibérico tem sofrido grandes mudanças ao longo do tempo, sendo a liberalização do mercado aquela que mais impacto teve no quotidiano da população [1]. A eletricidade pode ser considerada uma "mercadoria" e transacionada como tal, dando origem aos diversos mercados da eletricidade (como acontece por exemplo com os mercados da bolsa). Usualmente, os preços de eletricidade são determinados pela relação oferta-procura e idealmente deve verificar-se o princípio da conservação da energia, isto é, a quantidade de eletricidade produzida deve ser igual à quantidade consumida em todos os instantes. Este equilíbrio nem sempre se verifica, principalmente devido à incapacidade de armazenamento de eletricidade [2], e alterações nesta relação contribuem para oscilações nos preços e variações extremas ao longo do tempo, evidenciando o seu comportamento volátil. Outra característica destes preços é a sazonalidade. Por exemplo, é habitual o consumo elétrico aumentar durante o verão e inverno, devido ao aumento do uso de aparelhos de ar condicionado e aquecedores [3]. Por outro lado, a oferta pode diminuir em alturas de menor precipitação ou velocidade do vento já que as centrais hidroelétricas e os aerogeradores são uma das mais importantes fontes de produção de energia elétrica [4]. Estas e outras propriedades dos preços de eletricidade são descritas em [5].

Encontrar modelos confiáveis e realísticos para a previsão de preços de eletricidade é um grande desafio no mundo empresarial que envolve este mercado. A previsão de preços de eletricidade é extremamente importante para a diminuição do risco tanto para os produtores, que tentam vender a preços mais elevados, como para os comercializadores, que tentam obter a energia a preços mais atrativos que permitam obter maiores lucros na venda posterior ao consumidor final. Várias metodologias podem ser consideradas na investigação de modelos apropriados para descrever o comportamento dos preços (ver por exemplo [6]). Apesar de na estatística clássica as observações serem consideradas independentes e identicamente distribuídas (i.i.d.), a evolução dos valores dos preços ao longo do tempo estão relacionados com os seus valores passados. Assim, a ordem dos acontecimentos é um atributo importante já que, na análise de séries financeiras, o valor observado em determinado instante depende dos valores observados nos instantes anteriores. Desta forma, de entre as possíveis abordagens, destacam-se a análise de séries temporais, que, derivando da teoria de processos estocásticos, permite estudar as características das variáveis indexadas pelo tempo.

Estes modelos têm vindo a desempenhar um papel de interesse nas diversas áreas científicas (Economia, Meteorologia, Ciências Exatas,...) devido às suas características de modelação estatística [7]. A análise de séries temporais foca-se em dois principais objetivos: a modelação e a previsão. Na modelação procura-se encontrar um modelo que permita descrever o comportamento da série, enquanto a previsão é a avaliação dos valores futuros, considerando o modelo encontrado na fase de modelação.

Neste trabalho, dá-se particular atenção aos modelos auto-regressivos (AR) que desde o seu desenvolvimento [8] tiveram um importante desempenho na compreensão do comportamento de séries temporais. Estes modelos lineares são casos particulares da classe dos modelos auto-regressivos médias móveis (ARMA) introduzidos por Box e Jenkins [9]. Enquanto os modelos AR permitem descrever o valor presente da série em função dos seus valores passados, os modelos ARMA incluem também os valores de um termo estocástico interpretado como a inovação da série. A linearidade pressuposta nestes modelos constitui não só a sua principal característica, mas também limitação, levando a que não se adequem à modelação de alguns problemas.

Com o crescente interesse em analisar dados financeiros, como por exemplo o índice da bolsa, surgiu também a necessidade de encontrar modelos para descrever, com a maior exatidão possível, comportamentos voláteis. As séries financeiras, em particular as que representam preços de electricidade, experienciam alterações bruscas nos seus valores e apresentam grandes flutuações ao longo do tempo. Estatisticamente, a volatilidade pode ser confirmada através do desvio-padrão ou da variância da série e a necessidade de modelar esta propriedade levou ao aparecimento de diversos modelos que admitem aspectos de não-linearidade. De entre os modelos existentes para a modelação deste tipo de séries, destaca-se a classe dos modelos condicionalmente heterocedásticos. Estes modelos, introduzidos por Engle [10] e Bollerslev [11] são caracterizados pela presença de variância condicional aleatória e através do seu estudo é possível estimar e fazer previsões sobre a volatilidade das séries. As propriedades destes e de outros modelos de séries temporais são descritas em [7], onde também se encontram exemplos práticos e reais relativos à modelação de preços.

Os preços de electricidade são caracterizados também pelo elevado número de ocorrência de preços "extremos", e a modelação destas observações tem tido um interesse particular. A Teoria dos Valores Extremos (TVE) fornece uma ferramenta sólida para a modelação destes eventos. É uma abordagem capaz de captar as alterações bruscas nos preços de electricidade e está relacionada com os extremos de uma distribuição (máximos ou mínimos), permitindo o seu estudo e modelação. A TVE surgiu no início do século XX e é desde então uma das principais ferramentas utilizadas em diversas áreas: hidrologia, companhias de seguros, mercados financeiros, etc. Exemplos típicos são, por exemplo, a estimação da probabilidade de ocorrer um evento de inundação de proporções invulgarmente grandes, como foi o caso das cheias de 1967 em Lisboa [12], ou a previsão da ocorrência de sismos de grande escala, como o de 1755 também ocorrido em Lisboa [13]. Diversos exemplos de aplicação desta teoria podem ser consultados na literatura (ver, por exemplo [14], [15] e [16]). De acordo com [17], a análise de extremos remonta à década de 1920 no trabalho desenvolvido em [18] para a obtenção da distribuição exata do máximo (ou mínimo) de um conjunto de n observações

independentes e identicamente distribuídas (i.i.d.). Posteriormente, [19] descreveu a distribuição assintótica do máximo. Um relevante resultado desta área foi introduzido por [20], provado por [21] e complementado por [22], obtendo-se três possibilidades para os tipos de distribuição limite da distribuição do máximo: Gumbel, Fréchet ou Weibull. [23] apresentou uma forma de aplicar esta teoria na área dos estudos atuariais, envolvendo a análise da força de mortalidade. Mais recentemente, foi desenvolvida uma abordagem diferente para a modelação de extremos, focando-se nas observações acima de um determinado limiar (*threshold*). [24] e [25] concluíram que a distribuição assintótica dos excessos acima de um determinado *threshold* pode ser aproximada pela Distribuição Generalizada de Pareto (GPD - *Generalized Pareto Distribution*). A modelação de preços extremos através desta abordagem pode ser consultada, por exemplo, em [26] e [27].

Este trabalho centra-se no estudo e modelação das oscilações significativas que ocorrem nos preços de eletricidade, tendo em conta a volatilidade, a ocorrência de picos e a dependência significativa que existe nas séries de preços da eletricidade ao longo de determinado período de tempo. Para tal, serão combinados conhecimentos de séries temporais com a aplicação da TVE. Em 2000, [28] propôs uma estratégia que combina a aplicação do modelo de séries temporais com a TVE, anteriormente apresentada por [29], que consiste em modelar a cauda dos resíduos resultantes do modelo de séries temporais utilizando a TVE, uma vez que os resíduos têm uma propensão maior a serem i.i.d. do que a série original, e posteriormente modelar os quantis condicionais dos dados. A modelação de preços de eletricidade utilizando esta estratégia foi igualmente estudada por [30], [5] e [31]. [30] aplicou-a aos preços do mercado *Nord Pool*¹, e concluiu que esta abordagem tem uma melhor precisão na estimação dos quantis, quando comparada com os tradicionais modelos de séries temporais. Em [5] e [31] podemos ver esta metodologia aplicada ao mercado de eletricidade alemão. [31] dedica especial atenção à determinação do valor do *threshold* e evidencia a importância dessa escolha. É esta abordagem que irá ser seguida neste estudo, avaliando a sua capacidade de descrever o comportamento dos quantis dos retornos associados aos preços da eletricidade no mercado ibérico.

Este trabalho está subdividido em três principais capítulos: Metodologia (capítulo 2), Desenvolvimento (capítulo 3) e Conclusão (capítulo 4) e está estruturado da seguinte forma: o Capítulo 2 apresenta os principais conceitos referentes aos modelos de séries temporais que irão ser estudados e os conceitos fundamentais da Teoria de Valores Extremos. É também aqui feita uma breve descrição da metodologia que vai ser posta em prática. No Capítulo 3 apresentam-se os resultados obtidos pela aplicação da abordagem estudada aos preços da eletricidade do mercado português. No Capítulo 4 apresentam-se os comentários finais relativamente a todo o trabalho desenvolvido.

¹Nord Pool é uma empresa multinacional responsável pelas transações do mercado elétrico, ou parte dele, de alguns países, como, por exemplo, a Suécia.

Capítulo 2

Metodologia

2.1 Introdução

A modelação adequada de preços extremos pode ser efetuada recorrendo à Teoria de Valores Extremos (TVE), que possui resultados bem estabelecidos para esse fim. No entanto, uma vez que a abordagem clássica da TVE pressupõe observações independentes e identicamente distribuídas (i.i.d.) e que essa propriedade não é verificada nas séries de preços de mercado, é necessário seguir uma abordagem que permita lidar adequadamente com a dependência. Uma solução proposta na literatura [28] consiste em aplicar a TVE não à série original, mas aos resíduos resultantes da aplicação de modelos de séries temporais, uma vez que estes estão mais perto de serem i.i.d.. Usualmente, as inovações associadas aos resíduos destes modelos seguem distribuições gaussianas centradas e reduzidas, no entanto será também considerada a possibilidade de seguirem uma distribuição t-Student. Esta escolha é suportada pelo facto de a distribuição dos preços de eletricidade ser caracterizada por ter caudas mais pesadas do que a distribuição Normal [30]. Para essa etapa da análise, considera-se pertinente a combinação de um modelo auto-regressivo (AR) e um modelo condicionalmente heterocedástico (CH), de modo a permitir lidar apropriadamente com propriedades como a sazonalidade e volatilidade, típicas do mercado de eletricidade.

Neste capítulo, apresentar-se-ão com mais detalhe os modelos escolhidos: auto-regressivo integrado de médias móveis (ARIMA) e modelo com heterocedasticidade condicional auto-regressiva generalizado (GARCH). Os modelos ARIMA permitem descrever eventos que evoluem no tempo como função linear do seu passado, enquanto que os modelos GARCH permitem modelar a sua volatilidade. Os resultados aqui apresentados podem ser consultados em [7].

A descrição da metodologia adotada neste trabalho, relativa à análise de séries temporais e à TVE (incluindo as diferentes abordagens que podem ser utilizadas), irá ser realizada ao longo deste capítulo.

2.2 Modelos de Séries Temporais

Começa-se por introduzir a classe dos modelos **ARIMA- *autoregressive integrated moving average***. Estes modelos são uma generalização dos modelos auto-regressivos médias móveis

(ARMA), descritos em detalhe em [7]. Deste modo, a representação ARIMA(p, d, q) refere-se, respectivamente, às ordens de auto-regressão, de integração e de média móvel: p é o número de termos auto-regressivos, d é o número de diferenças e q é o número de termos da média móvel.

Diz-se que uma série temporal X_t segue um modelo ARIMA (p, d, q) se a diferença de ordem $d \geq 1$, $Y_t = (1 - B)^d X_t$, é um processo estacionário ARMA(p, q), em que B é o operador de atraso, ou seja, se

$$\Phi(B)(1 - B)^d X_t = \Theta(B)\varepsilon_t, \quad t \in \mathbb{Z},$$

onde $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ é um ruído branco, $\Phi(B) = 1 - \varphi_1 B - \dots - \varphi_p B^p, \varphi_p \neq 0$, $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q, \theta_q \neq 0$, são polinómios cujas raízes são de módulo superior a 1 e $(X_{-1}, X_{-2}, \dots, X_{-p-d}, \varepsilon_{-1}, \dots, \varepsilon_{-q})$ são não correlacionadas com $\varepsilon_0, \dots, \varepsilon_t$, e $d \in \mathbb{N}$.

Para verificar a estacionariedade¹ da série em estudo, podemos aplicar testes baseados na hipótese de existência de raiz unitária². Um dos testes de hipóteses utilizados é o teste de **Dickey-Fuller aumentado (ADF - Augmented Dickey Fuller Test)** que é uma versão do teste de Dickey-Fuller [32] para aplicar em modelos de séries temporais complexos e com um número elevado de observações e testa a hipótese de a série ter pelo menos uma raiz unitária - hipótese nula. Outro teste utilizado é o teste de **Phillips-Perron** [33]. É uma generalização do teste aumentado de Dickey-Fuller para os casos em que os resíduos ε_t são correlacionados e possivelmente heterocedásticos. Em 1992, Denis Kwiatkowski, Peter Phillips, Peter Schmidt e Yongcheol Shin [34] propuseram um outro teste, denominado **Kwiatkowski-Phillips-Schmidt-Shin (KPSS)**, para a hipótese nula de que a série é estacionária contra a hipótese alternativa de que a série tem uma raiz unitária.

Para além da estacionariedade, é igualmente importante analisar a correlação entre as observações. Por exemplo, na modelação de preços, é útil perceber se alterações atuais estão correlacionadas de alguma forma com acontecimentos anteriores. No presente estudo, irá ser analisada essa correlação através do teste de Ljung - Box (LB), um teste "portmanteu"³ que avalia a hipótese nula de que uma série não apresenta autocorrelação para um número fixo de lags (defasagens) L , contra a hipótese alternativa de que algum coeficiente de autocorrelação $\rho(k), k = 1, \dots, L$, é diferente de zero [36]. A estatística de teste é dada por

$$Q = T(T + 2) \sum_{k=1}^L \frac{\rho(k)^2}{T - k},$$

em que T é o tamanho da amostra, L é o número de lags e $\rho(k)$ é o coeficiente de autocorrelação na lag k . Sob a hipótese nula, a distribuição assintótica de Q é a Qui-quadrado com L graus de liberdade.

¹Uma série diz-se estacionária se a esperança matemática for constante e se a covariância entre duas observações dependem apenas da distância temporal que as separa.

²Diz-se que um processo estocástico tem uma raiz unitária se 1 for solução da equação homogénea $\Phi(B) = 0$, em que $\Phi(B)$ é o polinómio característico do modelo em estudo.

³Teste "portmanteau" é um teste em que a hipótese nula está bem definida, mas a hipótese alternativa é menos específica.

Os modelos ARIMA foram desenvolvidos com o intuito de descrever fenómenos que evoluem com o tempo, exprimindo os valores da série como função linear dos seus valores passados e considerando transformações adequadas para tornar a série estacionária. No entanto, séries estudadas em diversas áreas científicas apresentam características de não-linearidade, sofrendo grandes oscilações nos seus valores, como por exemplo os preços de eletricidade, que sofrem grandes picos ou descidas repentinos. Assim, surgiu a necessidade de ter em conta a volatilidade das séries na sua modelação. Existem várias classes destes modelos, dais quais aqui se destacam a dos modelos *condicionalmente heterocedásticos (CH)*.

Seja $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ um processo estocástico real e $\underline{\varepsilon}_t$ a σ -álgebra gerada por $(\varepsilon_t, \varepsilon_{t-1}, \dots)$. Diz-se que o processo estocástico real $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ segue um modelo condicionalmente heterocedástico de ordens p e q , $CH(p, q)$, se

$$\begin{cases} E(\varepsilon_t | \underline{\varepsilon}_{t-1}) = 0 \\ V(\varepsilon_t | \underline{\varepsilon}_{t-1}) = h_t = H(\varepsilon_{t-1}, \dots, \varepsilon_{t-q}, h_{t-1}, \dots, h_{t-q}) \end{cases}, \forall t \in \mathbb{Z},$$

onde H é uma função real definida sobre \mathbb{R}^{p+q} , estritamente positiva.

O processo ε diz-se condicionalmente gaussiano se a lei condicional de ε_t relativamente à σ -álgebra $\underline{\varepsilon}_{t-1}$, $\mathcal{L}(\varepsilon_t | \underline{\varepsilon}_{t-1})$, é a lei normal centrada de variância h_t , $N(0, h_t)$.

Características dos Modelos condicionalmente heterocedásticos (CH)

1. A condição $E(\varepsilon_t | \underline{\varepsilon}_{t-1}) = 0$ pode ser interpretada como uma condição de ortogonalidade ao passado do processo e permite-nos concluir a ortogonalidade e inexistência de correlações condicionais a qualquer passado do processo. Nomeadamente:
 - (a) $\forall h \in \mathbb{N}, \quad E(\varepsilon_t | \underline{\varepsilon}_{t-h}) = 0;$
 - (b) $\forall h \in \mathbb{N}, \forall k \in \mathbb{N}, \quad Cov(\varepsilon_t, \varepsilon_{t+h} | \underline{\varepsilon}_{t-k}) = 0.$
2. Relativamente à lei condicional, estes processos são centrados e não correlacionados. No entanto não são, em geral, homocedásticos, pelo que não são ruídos brancos no sentido clássico e chamam-se processos de erro.
3. Esta família de processos permite modelar processos de erro não gaussianos. Em particular: um processo de 2^a ordem condicionalmente Gaussiano não é necessariamente marginalmente Gaussiano.
4. A curtose associada a estes processos é, em geral, superior à da lei normal centrada e reduzida.

Dentro da classe de modelos condicionalmente heterocedásticos, destacam-se os modelos com heterocedasticidade condicional auto-regressiva (ARCH) [10] e a sua generalização GARCH [11]. Estes modelos permitem estimar e fazer previsões sobre a volatilidade devido à variância condicional aleatória que os caracteriza e, por isso, salienta-se a sua aplicação na análise de séries com alterações bruscas, nomeadamente as séries financeiras.

Diz-se que ε segue um modelo com heterocedasticidade condicional auto-regressiva de ordem q , ARCH(q), se

$$\forall t \in \mathbb{Z}, V(\varepsilon_t | \underline{\varepsilon}_{t-1}) = h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2,$$

com $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, q$.

Com esta formulação quadrática, garante-se a sua positividade e tal formulação é equivalente a uma representação auto-regressiva de ordem q para o processo $\varepsilon^2 = (\varepsilon_t^2, t \in \mathbb{Z})$.

No teorema seguinte apresenta-se um resultado importante sobre a estacionariedade de um processo ARCH(q).

Teorema 2.1. 1. Se ε é um processo estocástico real estacionário satisfazendo um modelo ARCH(q) então

$$\sum_{i=1}^q \alpha_i < 1.$$

2. Se ε é um processo satisfazendo um modelo ARCH(q) tal que $\sum_{i=1}^q \alpha_i < 1$, então ε é assintoticamente estacionário de 2ª ordem.

Muitas vezes, a variância condicional não depende apenas do passado do processo mas também do seu próprio passado, como se verá mais à frente na modelação dos preços de eletricidade. Bollerslev [11] propôs uma generalização dos modelos ARCH tendo em conta esta situação. Diz-se que ε segue um modelo ARCH generalizado de ordem p e q , GARCH(p, q), se a sua variância condicional é tal que

$$\forall t \in \mathbb{Z}, V(\varepsilon_t | \underline{\varepsilon}_{t-1}) = h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j},$$

com $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, q, \beta_j \geq 0, j = 1, \dots, p$.

Este modelo é assintoticamente estacionário ⁴ se os parâmetros envolvidos na definição satisfazem:

$$\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1.$$

2.3 Modelação de Valores Extremos

Como o interesse é o estudo de valores extremos, focar-se-á em particular na distribuição do valor máximo de n observações - M_n ⁵. Considere-se uma sequência de variáveis aleatórias independentes e identicamente distribuídas (i.i.d.), $\{X_1, X_2, \dots, X_n\}$, com função de distribuição F e seja $M_n = \max\{X_1, X_2, \dots, X_n\}$, então a distribuição de M_n , $P(M_n \leq z)$ é dada por $F^n(z)$. Na prática a função F não é conhecida e, mesmo sendo possível aplicar técnicas estatísticas para estimá-la,

⁴No sentido fraco

⁵A teoria aqui apresentada é análoga para o estudo de valores mínimos.

pequenas discrepâncias nessa estimação podem conduzir a desvios significativos no cálculo de F^n . A alternativa é encontrar boas aproximações para F^n através da análise do seu comportamento à medida que n aumenta. Contudo, o limite de $F^n(z)$ para $n \rightarrow \infty$ e $z < z_+$ ⁶ é uma função degenerada⁷, logo é necessário aplicar uma transformação adequada à variável que descreve o máximo, para ultrapassar esta barreira, por exemplo através da normalização da variável M_n : $M_n^* = \frac{M_n - b_n}{a_n}$ para sucessões de constantes $\{a_n > 0\}$ e $\{b_n\}$ escolhidas apropriadamente.

Neste capítulo apresentam-se duas distribuições que caracterizam o comportamento de eventos extremos: a distribuição generalizada dos valores extremos, conhecida como a distribuição GEV - *Generalized Extreme Value* e a distribuição generalizada de Pareto, GPD - *Generalized Pareto Distribution*. Os resultados aqui apresentados são baseados em [15], [16] e [14].

Introduz-se de seguida o Teorema de *Fisher-Tippett-Gnedenko* [20] [21], conhecido como *Teorema dos tipos*, relativo à distribuição assintótica de M_n^* . No caso em que existem sucessões $\{a_n > 0\}$ e $\{b_n\}$ tal que a distribuição do máximo normalizado, M_n^* , tenha como distribuição limite uma distribuição não degenerada, G , então G é de um dos três tipos apresentados no teorema:

Teorema 2.2. (*Teorema de Fisher-Tippett-Gnedenko*)

Se existem sucessões $\{a_n > 0\}$ e $\{b_n\}$ tal que

$P(\frac{M_n - b_n}{a_n} \leq z) \rightarrow G(z)$, $n \rightarrow \infty$, em que G é uma função de distribuição não degenerada, então G é uma função de distribuição pertencente a uma das seguintes famílias:

$$\text{Gumbel : } G(z) = \exp\{-\exp[-(\frac{z-b}{a})]\}, \quad -\infty < z < \infty$$

$$\text{Frechet : } G(z) = \begin{cases} 0, & z \leq b \\ \exp\{-(\frac{z-b}{a})^{-\alpha}\}, & z > b \end{cases}$$

$$\text{Weibull : } G(z) = \begin{cases} \exp\{-\{-(\frac{z-b}{a})^\alpha\}\}, & z < b \\ 1, & z \geq b \end{cases},$$

em que $a > 0$ parâmetro de escala, $b > 0$ parâmetro de localização e $\alpha > 0$ parâmetro de forma.

De referir que este teorema envolve a distribuição do máximo normalizado, mas também pode ser aplicado na aproximação da distribuição da sucessão de máximos⁸.

Diz-se se que F pertence ao domínio de atração de G , $F \in DA(G)$, se para alguma sucessão $\{a_n > 0\}$ e $\{b_n\}$ se verifica $F^n(a_n z + b_n) \rightarrow G(z)$, $n \rightarrow \infty$. Portanto existem três domínios de atração distintos, um para cada uma das distribuições consideradas no teorema anterior. Estas distribuições distinguem-se também por serem "max-stable" - cada uma pertence ao seu próprio domínio de atração.

Estas três famílias de funções de distribuição podem ser combinadas numa única função, a distribuição GEV:

$$G(z) = \exp\{-[1 + \xi(\frac{z-\mu}{\sigma})]^{-\frac{1}{\xi}}\},$$

⁶ z_+ é o menor valor de z tal que $F(z) = 1$

⁷Note-se que sendo F uma função de distribuição, $0 \leq F(z) \leq 1$ e portanto $F^n(z) \rightarrow 0, z \rightarrow \infty, z < z_+$

⁸Para n suficientemente grande, $P(\frac{M_n - b_n}{a_n} \leq z) \approx G(z) \Leftrightarrow P(M_n \leq z) \approx G(\frac{z - b_n}{a_n}) = G^*(z)$, em que G^* pertence a uma das três famílias apresentadas.

definida em $\{z : 1 + \xi(\frac{z-\mu}{\sigma}) > 0\}$, com $-\infty < \mu < \infty$ parâmetro de localização, $\xi \neq 0$ parâmetro de forma e $\sigma > 0$ parâmetro de escala. O valor de ξ determina o tipo de distribuição correspondente: caso ξ seja positivo temos a distribuição de Fréchet e se ξ for negativo estamos perante a distribuição de Weibull. No caso em que $\xi = 0$ considera-se o limite $\xi \rightarrow 0$, obtendo a distribuição de Gumbel: $G(z) = \exp\{-\exp[-(\frac{z-\mu}{\sigma})]\}$, $-\infty < z < \infty$.

Apesar de a distribuição GEV ser extremamente útil na análise de valores extremos, em muitas situações é mais adequado considerar apenas as observações que ultrapassam um determinado *threshold* u . É neste sentido que surge a necessidade de determinar a probabilidade de, dado que o *threshold* é ultrapassado, uma observação exceder o *threshold* no máximo em y unidades. A função de distribuição condicional dos excessos $Y = X - u$, dado que X ultrapassa o *threshold* u , é dada por: $F_u(y) = P(X - u \leq y | X > u)$, $y > 0$ e é possível escrevê-la a partir de F :

$$\begin{aligned} F_u(y) &= \frac{P(u < X \leq u + y)}{P(X > u)} \\ &= \frac{F(u + y) - F(u)}{1 - F(u)}. \end{aligned}$$

Mais uma vez surge como dificuldade o facto de não conhecermos F , pelo que é necessário uma boa aproximação para F_u . O Teorema de *Pickands-Balkema-de Hann* [24] [25], muitas vezes referido como o 2.º Teorema da TVE, especifica a distribuição limite de F_u :

Teorema 2.3. (*Teorema de Pickands-Balkema-de Hann*)

Sejam X_1, X_2, \dots, X_n variáveis *i.i.d.* com função de distribuição F , tal que $F \in DA(G)$, G não degenerada, nomeadamente tendo-se que $P(M_n \leq z) \approx G(z)$, em que

$$G(z) = \exp\{-[1 + \xi(\frac{z-\mu}{\sigma})]^{-\frac{1}{\xi}}\},$$

para algum μ , $\sigma > 0$ e $\xi \neq 0$.

Então, para u suficientemente grande:

$$P(X_1 - u \leq y | X_1 > u) \approx H_{\xi, \tilde{\sigma}}(y) = 1 - (1 + \xi \frac{y}{\tilde{\sigma}})^{-\frac{1}{\xi}},$$

definida em $\{y : y > 0 \text{ e } (1 + \xi \frac{y}{\tilde{\sigma}}) > 0, \xi \neq 0\}$ e onde $\tilde{\sigma} = \sigma + \xi(u + \mu)$.

Note-se que, no caso de $\xi = 0$ considera-se o limite da distribuição quando ξ se aproxima de zero, obtendo-se a função $H(y) = 1 - \exp(-\frac{y}{\sigma})$.

Também aqui se consideram três famílias de distribuição distintas: se $\xi = 0$, tem-se uma distribuição exponencial e se ξ for diferente de zero, têm-se distribuições do tipo Pareto. O parâmetro de forma ξ determinado na distribuição GPD é o mesmo parâmetro de forma determinado na distribuição GEV e há uma correspondência entre as famílias da distribuição GEV e da distribuição GPD, isto é, as distribuições da GPD para $\xi > 0$, $\xi < 0$ e $\xi = 0$ pertencem aos domínios de atração das funções de Fréchet, Weibull e Gumbel, respectivamente.

Para aplicar a TVE a um conjunto de dados, é importante selecionar corretamente as observações extremas e perceber qual o tipo de distribuição que deve ser considerada. Destacam-se dois métodos para definir tais observações extremas: o método *Máximo por Blocos (MB)* e o método *Peaks-Over-Threshold (POT)*. O método MB consiste em dividir os dados em blocos de igual tamanho e amplitude e considerar o valor máximo em cada bloco (Figura 2.1). O método POT baseia-se em extrair e estudar as observações com valores acima de um determinado *threshold* u (Figura 2.2). De acordo com os dados em análise, é necessário escolher qual a abordagem adequada a adotar. O método MB apresenta uma grande desvantagem que se prende com o facto de se escolher apenas uma observação em cada bloco, o que pode conduzir a uma amostra de pequena dimensão. A abordagem POT contorna esta dificuldade e é uma ferramenta mais poderosa para a modelação de eventos extremos [31] e será a abordagem a seguir neste estudo. No entanto, esta abordagem também tem obstáculos a contornar na sua aplicação. Um dos principais obstáculos é o facto de partir da suposição de independência das observações e na maioria dos casos estudados existe uma dependência temporal. Por outro lado, a seleção de um *threshold* adequado é outra questão importante a ter em conta, pois uma escolha inadequada pode comprometer a modelação da cauda da distribuição.

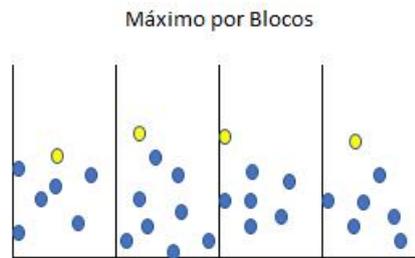


Figura 2.1: Método Máximo por Blocos

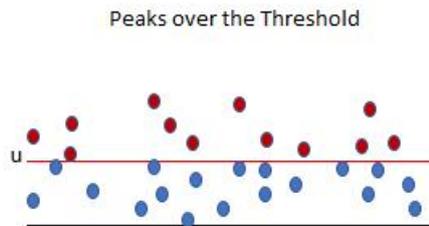


Figura 2.2: Método *Peaks-Over-Threshold*

Definir estimadores adequados para os parâmetros ξ e $\tilde{\sigma}$ é um dos tópicos mais importantes da TVE. Aqui o foco será a estimação dos parâmetros da GPD já que será adotada a abordagem POT, no entanto os mesmos métodos podem ser usados no caso da distribuição GEV. As abordagens para encontrar estes estimadores dividem-se em paramétricas e não-paramétricas. Da classe dos estimadores paramétricos fazem parte alguns dos mais conhecidos, como é o caso do estimador pelo método dos momentos, do estimador da máxima verosimilhança e o método dos momentos ponderados. Para mais detalhes ver por exemplo [37].

Neste trabalho, após o valor do *threshold* determinado, os parâmetros da GPD serão estimados pelo método da máxima verossimilhança. Sejam y_1, \dots, y_{N_u} as N_u excedências do threshold u . Para $\xi \neq 0$, a função de log-verossimilhança é expressa da seguinte forma:

$$\begin{aligned} \ell(\xi, \tilde{\sigma}) &= \sum_{i=1}^{N_u} \log [h(y_i, \xi, \tilde{\sigma})] \\ &= \sum_{i=1}^{N_u} \log \left[\frac{1}{\tilde{\sigma}} \left(1 + \frac{\xi}{\tilde{\sigma}} y_i \right)^{-\frac{1}{\xi}-1} \right] \\ &= -N_u \log(\tilde{\sigma}) - \left(\frac{1}{\xi} + 1 \right) \sum_{i=1}^{N_u} \log \left(1 + \frac{\xi}{\tilde{\sigma}} y_i \right), \end{aligned}$$

para $\left(1 + \frac{\xi}{\tilde{\sigma}} y_i \right) > 0$, caso contrário temos uma log-verossimilhança infinita.

Para $\xi = 0$, a função de log-verossimilhança é dada por $\ell(\tilde{\sigma}) = -N_u \log(\tilde{\sigma}) - \frac{1}{\tilde{\sigma}} \sum_{i=1}^{N_u} y_i$.

Maximizando a função de log-verossimilhança obtida para os parâmetros considerados, obtêm-se as estimativas pretendidas.

O parâmetro de forma ξ tem recebido especial atenção por retratar o comportamento da cauda. Para $\xi \in \mathbb{R}$, realçam-se os estimadores propostos por Pickands [25] e Dekkers [38] que são generalizações do clássico estimador de Hill, proposto por [39] para $\xi > 0$, e que fará parte deste estudo. Para uma amostra de n variáveis aleatórias i.i.d., $\{X_1, \dots, X_n\}$, considere-se a amostra ordenada $X_{(1,n)} \leq X_{(2,n)} \leq \dots \leq X_{(n,n)}$, em que $X_{(i,n)}$ denota a i -ésima estatística de ordem. O estimador de Hill é dado por:

$$\hat{H}(k) = \frac{1}{k} \sum_{i=1}^k \log X_{(n-i+1,n)} - \log X_{(n-k,n)}.$$

A partir da estimação dos parâmetros da GPD, é possível estimar os quantis desta distribuição. Denotando-se por \bar{F} função da cauda de uma dada distribuição F , isto é, $\bar{F}(y) = 1 - F(y)$, pode escrever-se:

$$\bar{F}_u(y) = \frac{\bar{F}(y+u)}{\bar{F}(u)} \Leftrightarrow \bar{F}_u(y) \bar{F}(u) = \bar{F}(y+u).$$

Pelo Teorema 2.3, tem-se $\bar{F}_u(y) \approx \bar{H}_{\xi, \tilde{\sigma}}(y)$, logo

$$\bar{F}(y+u) \approx \bar{F}(u) \bar{H}_{\xi, \tilde{\sigma}}(y) = \bar{F}(u) \left(1 + \xi \frac{y}{\tilde{\sigma}} \right)^{-\frac{1}{\xi}}.$$

Considerando $x = y + u$, tem-se

$$\bar{F}(x) = \bar{F}(u) \left(1 + \xi \frac{x-u}{\tilde{\sigma}} \right)^{-\frac{1}{\xi}}.$$

Então, para $0 < p < 1$, o quantil de ordem p , q_p é dado por:

$$q_p = u + \frac{\tilde{\sigma}}{\tilde{\xi}} \left[\left((1-p) \frac{1}{\bar{F}(u)} \right)^{-\tilde{\xi}} - 1 \right].$$

Um estimador para $\bar{F}(u)$ é dado pela distribuição empírica: $\hat{\bar{F}}(u) = \frac{N_u}{N}$, em que N_u é o número de observações que ultrapassam u e N é o número total de observações. Então, um estimador para q_p é dado por:

$$\hat{q}_p = u + \frac{\hat{\sigma}}{\hat{\xi}} \left[\left((1-p) \frac{N}{N_u} \right)^{-\hat{\xi}} - 1 \right], \quad (2.1)$$

onde $\hat{\sigma}$ e $\hat{\xi}$ designam estimativas para os parâmetros $\tilde{\sigma}$ e ξ , respetivamente.

A escolha do *threshold* u tem uma grande influência na qualidade do ajustamento da GPD aos dados. A escolha de um valor muito elevado pode resultar num número baixo de excedências. Em contrapartida, um *threshold* muito baixo pode conduzir a estimadores enviesados e impedir o bom ajustamento da GPD [27]. Serão aplicados alguns métodos para facilitar essa escolha:

1. Hill Plot

Através do estimador de Hill para o índice de cauda, é possível obter uma aproximação para o número de observações a considerar como "extremas", estimando o índice de cauda ξ para diferentes números de observações na cauda N_u e identificando graficamente a região onde o coeficiente se torna estável.

2. Mean Excess Function (MEF)

Como o próprio nome indica, esta função define o valor esperado dos excessos e é dada por $E(X - u | X > u)$. Se a distribuição subjacente é a $GDP(\xi, \tilde{\sigma})$, então a MEF é dada por: [15]

$$e(u) = E(X - u | X > u) = \frac{\tilde{\sigma} + \xi u}{1 - \xi}.$$

O *threshold* pode ser determinado através do gráfico da MEF estimada: dado que é uma função linear em u , u deverá ser tal que o gráfico mostra um comportamento linear.

3. Método de Eyeball

Este método, apresentado em [31], consiste na estimação do índice de cauda e do parâmetro de escala simultaneamente. Para diferentes números de observações "extremas", estimam-se os parâmetros da correspondente GPD e identificam-se graficamente as regiões em que esses valores apresentam um comportamento estável.

2.4 Descrição da Metodologia

O trabalho desenvolvido segue a abordagem apresentada por [28] que consiste em modelar a volatilidade da série dos retornos (ganhos ou perdas) associados aos preços de eletricidade através de um modelo GARCH e aplicar o método POT aos resíduos resultantes desse modelo para depois estimar os quantis condicionais dos retornos. Como já referido anteriormente, está-se perante uma situação em que os retornos não são i.i.d., portanto aplicar o POT diretamente não é uma solução

viável. Daí surge a combinação deste método com o uso de modelos de séries temporais. Embora a suposição de independência nos retornos não seja válida, pode ser sustentável quando se trata dos resíduos destes modelos. Inicialmente, os retornos são modelados através de uma combinação de um modelo auto-regressivo (AR) com um modelo GARCH, para que se capte a dependência do passado e a volatilidade dos dados. Será considerado que os resíduos do modelo resultante podem ter inovações gaussianas ou distribuição t-Student. Inclui-se a distribuição t-Student pelo facto de o comportamento dos preços usualmente evidenciar que estes seguirão uma distribuição com caudas mais pesadas do que a distribuição normal [30]. Após modelar os resíduos através da abordagem POT, estimam-se os quantis condicionais da série original dos retornos. O quantil condicional é calculado tendo em conta as estimativas do desvio-padrão condicional e os quantis das distribuições Normal, t-Student e GPD.

Capítulo 3

Desenvolvimento

Neste capítulo procede-se à aplicação dos conceitos introduzidos no capítulo anterior na modelação de séries temporais relativas a observações dos preços de eletricidade no mercado Ibérico, por hora, entre a primeira hora do dia 25-12-2011 e a última hora do dia 01-01-2016, perfazendo um total de 35256 observações. Após uma breve análise descritiva dos preços correspondentes a Portugal, a fase seguinte consiste na modelação dos retornos associados a estes preços (que vão de encontro às expectativas dos investidores) e, para tal, utilizar-se-á um modelo auto-regressivo e um modelo condicionalmente heterocedástico para modelar a variância.

Seguidamente, aplicar-se-á a TVE na modelação da cauda dos resíduos associados aos modelos anteriores e na estimação dos quantis condicionais dos retornos.

3.1 Análise descritiva dos dados

A série temporal relativa aos preços de mercado pode ser observada na Figura 3.1. Esta evidencia as alterações bruscas nos preços (aumentos ou quedas repentinas), tomando valores positivos até um máximo de 112€/MWh e com um preço médio de 46€/MWh. O valor da curtose (cerca de 4) sugere uma distribuição com caudas pesadas.

Perante tais extremas alterações nos preços, torna-se complicado encontrar modelos adequados à sua versatilidade. Neste estudo irão ser considerados os retornos dos preços que fornecem uma indicação da variabilidade dos mesmos ao longo do tempo.

Consideram-se usualmente dois tipos de retornos:

- retornos simples: $\frac{P_t - P_{t-1}}{P_{t-1}}$,
- retornos logarítmicos: $\ln(P_t) - \ln(P_{t-1})$,

em que P_t é o preço no instante t .

Neste trabalho consideram-se os retornos simples. Os retornos logarítmicos tendem a desvalorizar as alterações dos preços. Por exemplo, se se considerar o aumento do preço de 100€/MWh, para 150€/MWh, o retorno simples será de 50% enquanto o retorno logarítmico é apenas 40,55%. Uma desvantagem na aplicação dos retornos simples é que os preços são limitados inferiormente e

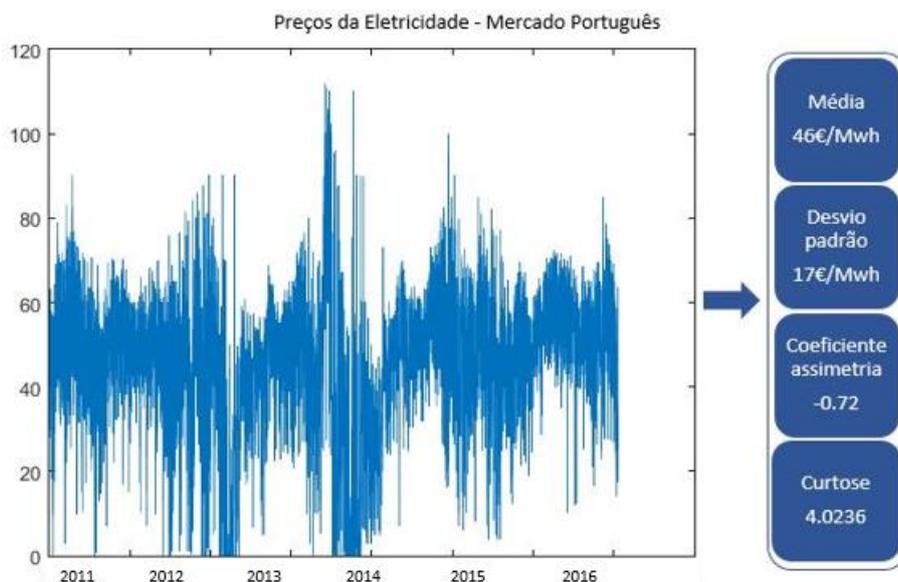


Figura 3.1: Preços da eletricidade

portanto se se pretender analisar as quedas nos preços, a maior perda possível será 100% (retorno de -100%), tornando a distribuição assimétrica positiva: uma descida nos preços de 100€/MWh, para 10€/MWh, representa uma perda de 90%, enquanto que um aumento do preço de 10€/MWh, para 100€/MWh, corresponde a um ganho de 900%. No entanto, como o nosso interesse é estudar a cauda positiva, esta dificuldade é ultrapassada. De modo a evitar retornos muito próximos de zero, consideramos preços superiores 1€, totalizando 34342 observações.

Para testar a estacionariedade foram utilizados os testes Dickey-Fuller aumentado (ADF), Phillips-Perron (PP) e KPSS. Os seus resultados estão exibidos na Tabela 3.1. Para um nível de 1%, e para ambos os testes ADF e PP, a hipótese nula de que a série tem uma raiz no círculo unitário é de rejeitar, o que sugere a estacionariedade da série. O teste KPSS dá um resultado contraditório, rejeitando-se a hipótese de que a série é estacionária. Seguir-se-á o estudo assumindo a estacionariedade dos retornos, uma vez que este teste tende a rejeitar a hipótese nula muitas vezes [40].

Tabela 3.1: Testes de Estacionariedade aplicados nos retornos dos preços de eletricidade

Teste	Hipótese nula	Versão do teste	Estatística de teste	Valor crítico a 5 %
ADF	A série temporal tem uma raiz unitária	Com constante	-26.5147	-2.8610
		Com constante e tendência	-26.6287	-3.4123
		Sem constante ou tendência	-25.2570	-1.9416
PP	A série temporal tem uma raiz unitária	Com constante	-163.8215	-2.8610
		Com constante e tendência	-163.8427	-3.4123
		Sem constante ou tendência	-163.6937	-1.9416
KPSS	A série temporal é estacionária	Com constante	2.9507	0.4630
		Com constante e tendência	1.5967	0.1460

Para além das usuais medidas de localização e dispersão, média e desvio-padrão, serão também considerados os coeficientes da curtose e assimetria para avaliar o comportamento dos dados. O resumo descritivo dos dados está presente na Tabela 3.2. O valor da curtose (438.9347) indica uma distribuição leptocúrtica, sendo a sua curva muito mais afunilada do que a curva da distribuição Normal¹. O valor elevado do coeficiente de assimetria, aproximadamente igual a 17, sugere que a distribuição tem caudas pesadas, o que se pode também concluir através da visualização do gráfico *QQ-Plot* presente na Figura 3.2.

Tabela 3.2: Resumo descritivo dos retornos dos preços da eletricidade

Mínimo	1.ºQ	Mediana	Média	3.ºQ	Máximo	Desvio-padrão
-0.94222	-0.04457	0.00000	0.02017	0.04045	12.5	0.3042
Curtose	Assimetria	$LB(24)$	$LB^2(24)$	$LB(168)$	$LB^2(168)$	
438.9347	16.5702	2955.3	461.9863	8012.6	1974.2	

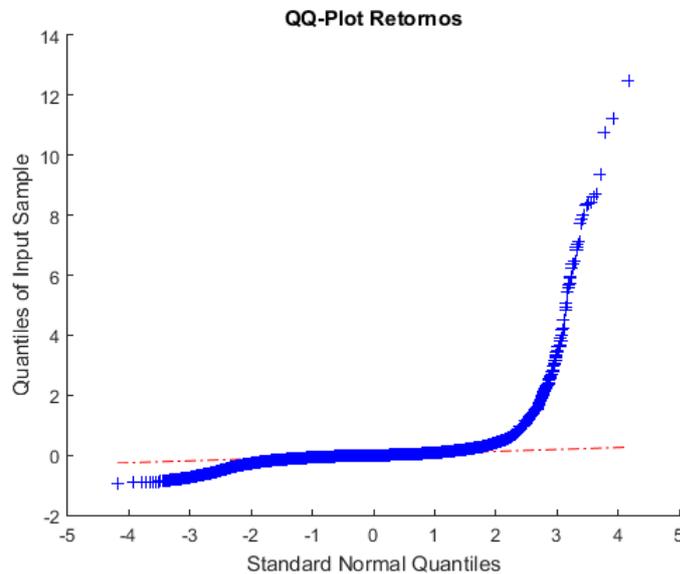


Figura 3.2: QQ-Plot Retornos dos preços da eletricidade

¹Para a distribuição Normal, o valor da curtose é igual a 3.

Os valores obtidos pelo teste Ljung-Box (LB)² para os retornos e para o quadrado dos retornos indicam forte autocorrelação da série dos retornos. Na Figura 3.3 pode observar-se o comportamento das funções de autocorrelação (ACF) e autocorrelação parcial (PACF) dos retornos que também evidencia dependência existente.

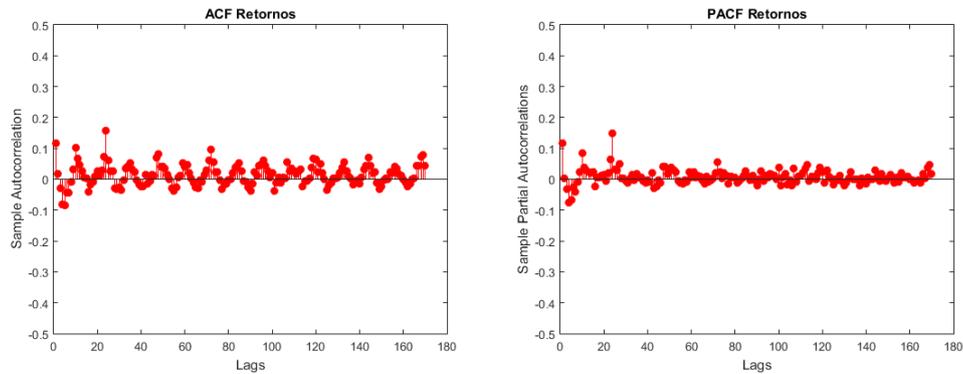


Figura 3.3: ACF Retornos (esquerda) e PACF dos retornos (direita)

3.2 Modelação dos retornos

A combinação de um modelo auto-regressivo com um modelo condicionalmente heterocedástico parece ser um bom ajustamento para os dados em estudo. Para além de se considerar a dependência do passado (modelo AR), tem-se também em conta a volatilidade dos dados (modelo GARCH). A sazonalidade da série é notória nos gráficos da ACF e PACF (Figura 3.3), especialmente nas primeira e vigésima-quarta horas, pelo que o modelo AR deverá incluir a lag 1 e a lag 24. O padrão de sazonalidade está diretamente relacionado com as mudanças na procura de eletricidade durante as 24 horas do dia, bem como os 7 dias da semana: a procura durante o dia é maior do que durante a noite e também há uma queda na procura quando é fim de semana. Será incluída também a lag 168 no modelo de modo a considerar a sazonalidade semanal. Combinando o modelo AR com o modelo GARCH mais simples possível, GARCH(1,1), que modela a volatilidade condicional em função dos seus valores e valores dos resíduos passados, obtém-se a seguinte expressão:

$$r_t = \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-24} + \alpha_3 r_{t-168} + \epsilon_t,$$

$$\sigma_t^2 = \phi_0 + \phi_1 \epsilon_{t-1}^2 + \phi_2 \sigma_{t-1}^2,$$

em que σ_t^2 é a variância condicional de ϵ_t , $\epsilon_t = \sigma_t Z_t$ e Z_t segue uma distribuição Normal centrada e reduzida ou uma distribuição t-Student com ν graus de liberdade (uma vez que a dinâmica dos preços da eletricidade é caracterizada por ter caudas mais pesadas do que a distribuição normal). Para encontrar um bom modelo, estimaram-se os coeficientes para quatro modelos diferentes: dois

² $LB(j)$ e $LB^2(j)$ são as estatísticas do teste de Ljung-Box para a autocorrelação na lag j , para $j = 24, 168$ na série dos retornos e do quadrado dos retornos, respetivamente.

considerando as lags 1, 24 e 168 e dois considerando apenas as lags 1 e 24, com inovações gaussianas e inovações t-Student. As estimativas para os coeficientes dos modelos, determinadas pelo método da máxima verossimilhança, estão apresentadas nas Tabelas 3.3 e 3.4.

Tabela 3.3: Estimativas para os coeficientes do modelo AR-GARCH com lags 1, 24 e 168

$Z_t \sim N(0, 1)$	Coef. estimado	Desvio-padrão
α_0	0.001658	0.00030608
α_1	0.189915	0.00214902
α_2	0.33661	0.00139224
α_3	0.152715	0.000626793
ϕ_0	0.000178892	2.59734×10^{-6}
ϕ_1	0.166056	0.000607936
ϕ_2	0.833944	0.000433678
Log V.	28844	
$Z_t \sim t_\nu$	Coef. estimado	Desvio-padrão
α_0	-0.00280013	0.000260392
α_1	0.203332	0.00400691
α_2	0.259433	0.00238141
α_3	0.200534	0.00219412
ϕ_0	0.0013326	5.85847×10^{-5}
ϕ_1	0.501815	0.0199022
ϕ_2	0.498185	0.00671315
Log V.	38758	
Graus de liberdade	2.69121	0.0366038

Tabela 3.4: Estimativas para os coeficientes do modelo AR-GARCH com lag 1 e lag 24

$Z_t \sim N(0, 1)$	Coef. estimado	Desvio-padrão
α_0	0.00254008	0.000336991
α_1	0.2243	0.00206842
α_2	0.402787	0.00125258
ϕ_0	0.000194403	2.94351×10^{-6}
ϕ_1	0.164262	0.000601094
ϕ_2	0.835737	0.0004162
Log V.	27696	
$Z_t \sim t_\nu$	Coef. estimado	Desvio-padrão
α_0	-0.00325529	0.000280001
α_1	0.255421	0.0042159
α_2	0.348503	0.00242934
ϕ_0	0.00185448	7.86406×10^{-5}
ϕ_1	0.582429	0.0228125
ϕ_2	0.417571	0.00723263
Log V.	37415	
Graus de liberdade	2.69866	0.0366098

Os parâmetros estimados para a equação da variância, ϕ_i , são positivos em ambos os modelos, no entanto a sua soma não é significativamente menor do que 1, o que não permite excluir a hipótese de ter uma variância condicional infinita.

Para comparar os modelos obtidos, determinam-se os valores dos critérios de informação Akaike (AIC) e Bayesian (BIC). Estes critérios são baseados no valor ótimo para a maximização do logaritmo da função de verosimilhança e são caracterizados pela penalização de modelos com mais parâmetros [42].

- Critério de Informação de **Akaike**: $AIC = -2 \times \log L + 2 \times Np$
- Critério de Informação **Bayesian**: $BIC = -2 \times \log L + Np \times N$

em que $\log L$ é o valor ótimo para a maximização do logaritmo da função de verosimilhança, Np é o número de parâmetros e N é o número de observações.

Os valores obtidos para cada critério e para cada modelo podem ser observados na Tabela 3.5. Em primeira análise, parece adequado adicionar o termo AR(168). Os critérios AIC e BIC têm menores valores para os modelos que consideram este termo e portanto irá prosseguir-se com estes modelos. Os resultados para estes critérios confirmam também a suspeita inicial de que um modelo com inovações que seguem uma distribuição t-Student seria o mais adequado para modelar os retornos.

Tabela 3.5: Comparação dos Modelos AR-GARCH

Modelo	AIC	BIC
Inovações Gaussianas (lag 1 e lag 24)	-55380	-55329
Inovações t-Student (lag 1 e lag 24)	-74816	-74757
Inovações Gaussianas (lag 1, lag 24 e lag 168)	-57674	-57615
Inovações t-Student (lag 1, lag 24 e lag 168)	-77500	-77432

3.3 Modelação dos resíduos do Modelo AR-GARCH

Modelação dos resíduos do modelo AR-GARCH com inovações t - Student

Na Figura 3.4 pode observar-se a autocorrelação dos resíduos normalizados e a autocorrelação do quadrado dos mesmos. Os gráficos indicam que nem toda a autocorrelação foi eliminada. Na Figura 3.5 é possível ver a que a maioria da heteroscedasticidade dos dados originais é refletida pelo nosso modelo de variância condicional GARCH.

Assumindo que os resíduos normalizados do modelo AR-GARCH são i.i.d., serão modelados através da abordagem POT da Teoria de Valores Extremos. Para estimar o valor do *threshold*, combinar-se-ão os métodos introduzidos no capítulo anterior: Hill Plot, Mean Excess Plot e Eyeball.

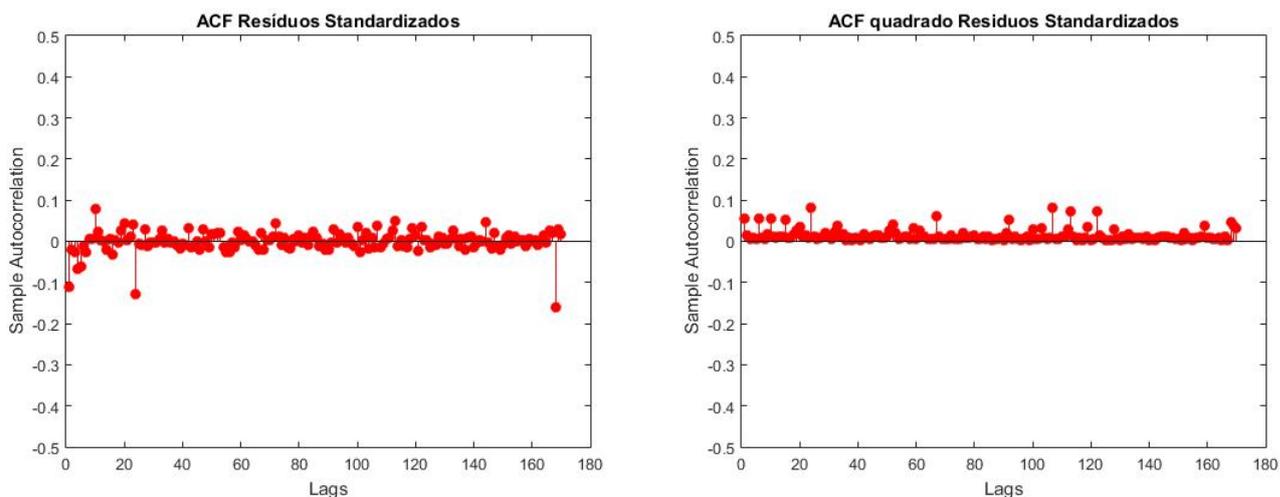


Figura 3.4: ACF resíduos normalizados (esquerda) e do quadrado dos resíduos normalizados (direita) - Modelo AR-GARCH com inovações t-Student

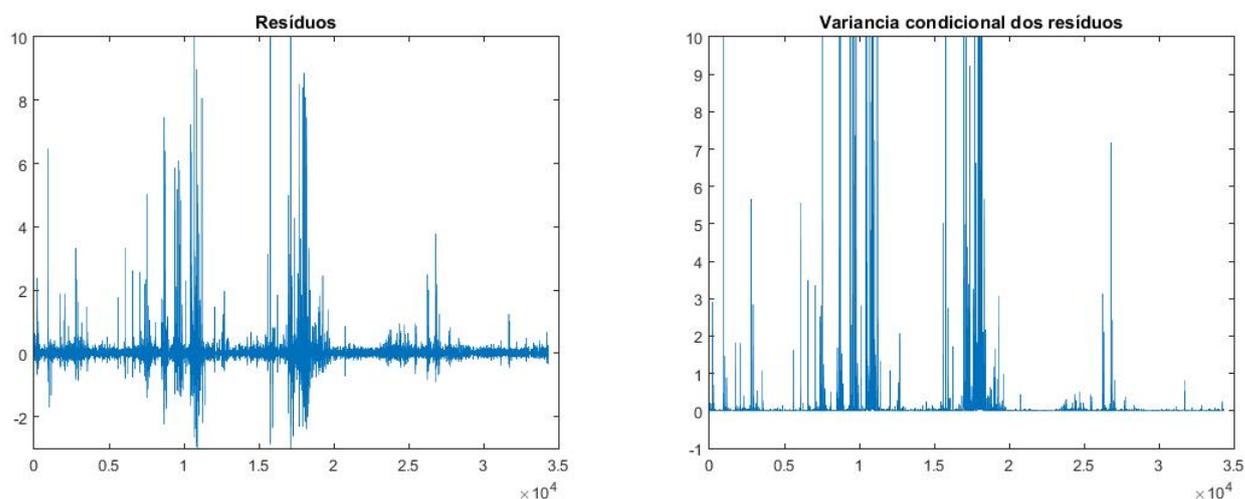


Figura 3.5: Gráfico dos Resíduos (esquerda) e gráfico da sua volatilidade (direita) - Modelo AR-GARCH com inovações t-Student

Para a análise através do método Hill Plot, estimou-se o índice de cauda, utilizando o estimador de Hill para diferentes números de observações na cauda até 10000 e o gráfico resultante está presente na Figura 3.6. Podemos ver que o estimador de Hill para o índice de cauda decresce consideravelmente rápido quando é considerado um baixo número de excedências, mantendo-se estável à medida que aumenta o número de observações na cauda.

Tabela 3.6: N.º Excedências vs Threshold - resíduos normalizados do modelo AR-GARCH com inovações t-Student

%	10%	20%	25%	30%	35%	40%	45%	50%
N.º Excedências	3434	6868	8585	10320	12019	13736	15454	17171
threshold (u)	0.2801	0.1109	0.0679	0.0397	0.0172	-0.0013	-0.0170	-0.0322

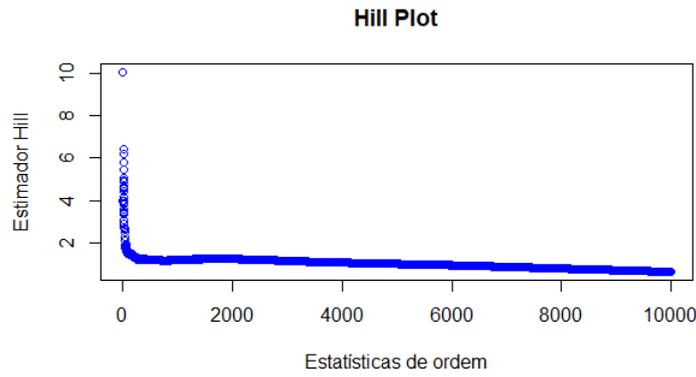


Figura 3.6: Hill Plot dos resíduos normalizados - Modelo AR-GARCH com inovações t-Student

Os gráficos provenientes da aplicação do método de Eyeball apresentam-se na Figura 3.7. O estimador para o índice de cauda é mais estável quando se consideram entre 25% (8588 observações, $u=0.0679$) a 40% (13736 observações, $u=-0.0013$) das observações na cauda, enquanto que o estimador para o parâmetro de escala estabiliza na região em que mais de 20% dos resíduos normalizados pertencem à cauda. Os gráficos da MEF (Figura 3.8) mostram que o valor do *threshold* deverá ser maior do que -0.04.

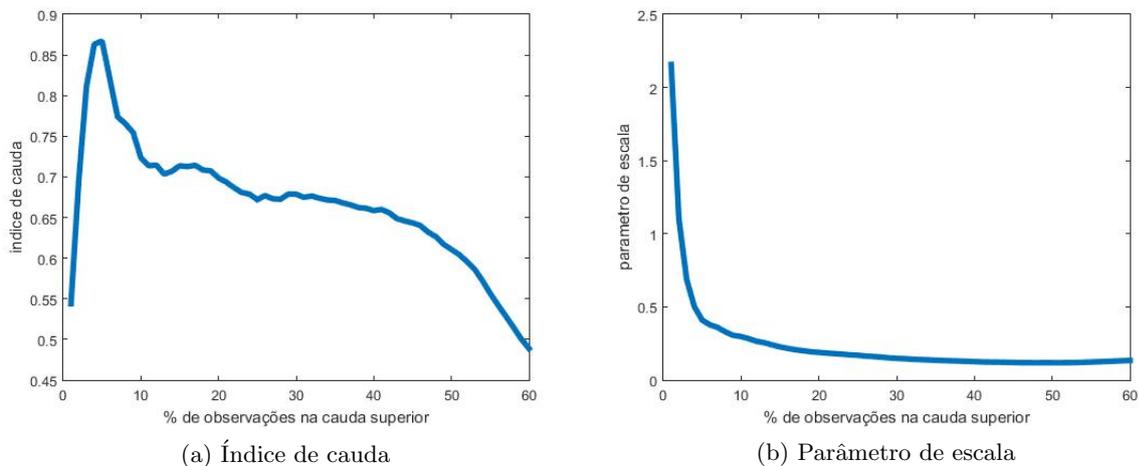


Figura 3.7: Método de Eyeball - resíduos normalizados do Modelo AR-GARCH com inovações t-Student

Para verificar a influência que a escolha do *threshold* tem no ajustamento da distribuição, fez-se o ajustamento para diferentes valores de N_u , conforme Figuras 3.9 ($N_u = 8585$), 3.10 ($N_u = 10320$), 3.11 ($N_u = 12019$), 3.12 ($N_u = 13736$), 3.13 ($N_u = 15454$) e 3.14 ($N_u = 17171$), em que se consideram desde 25% a 50% observações extremas.

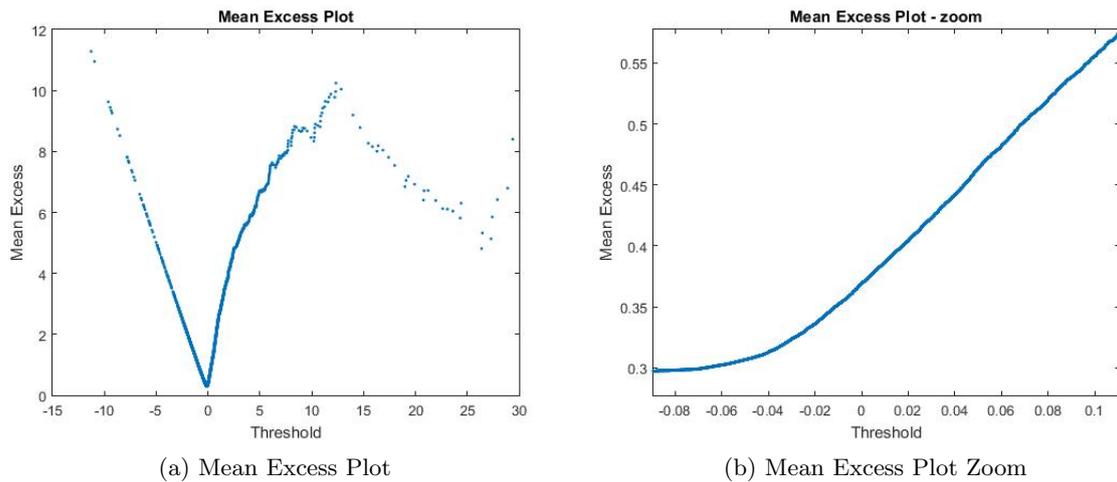


Figura 3.8: Mean Excess Plot dos resíduos normalizados do Modelo AR-GARCH com inovações t-Student

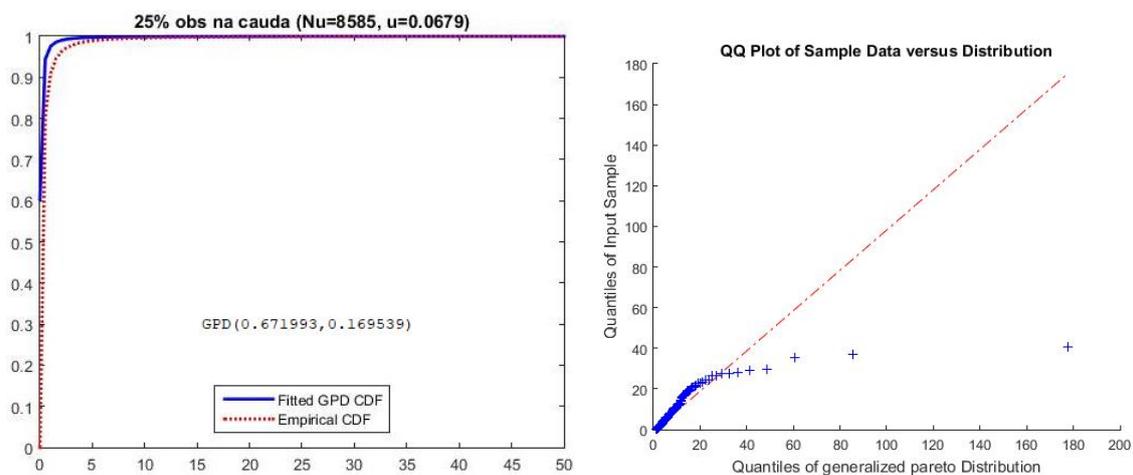


Figura 3.9: GPD para $N_u = 8585$ (25% das observações) - resíduos normalizados do modelo AR-GARCH com inovações t-Student

Os *QQ-Plots* das Figuras 3.9 - 3.14, são construídos de modo a comparar os quantis de cada uma das distribuições consideradas com os quantis da amostra, cujo tamanho varia de acordo com a percentagem de observações extremas. Analisando o comportamento destes gráficos, estes refletem a informação providenciada pelo gráfico da Figura 3.7a onde é notória a estabilidade do estimador para o parâmetro de forma ξ quando a cauda é constituída por 30% das observações. Podemos então escolher a GPD com parâmetros $\hat{\xi} = 0.68$ e $\hat{\sigma} = 0.15$ para modelar a cauda dos resíduos normalizados com inovações t-Student. Neste caso, temos 10320 observações "extremas" e um *threshold* igual a 0.0397.

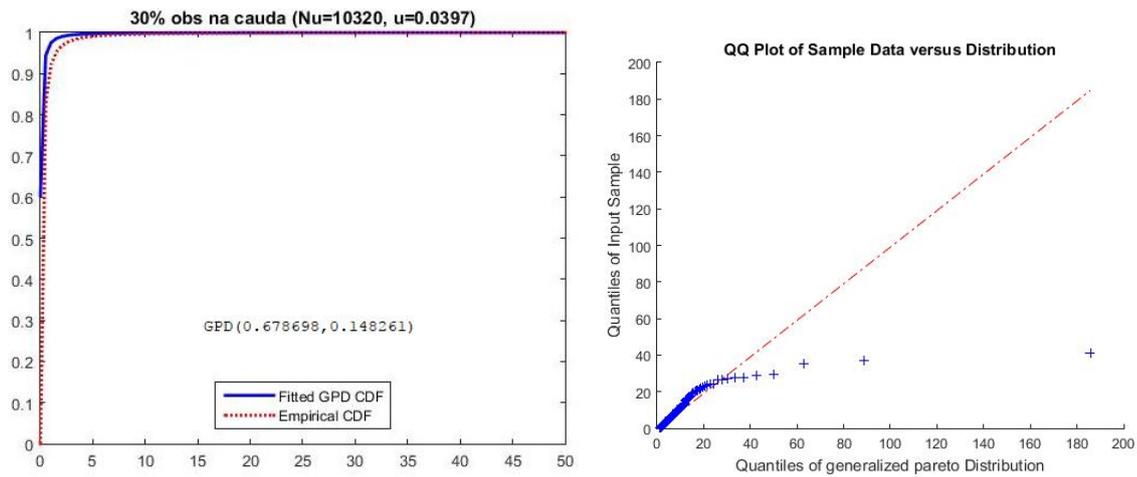


Figura 3.10: GPD para $N_u = 10320$ (30% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações t-Student

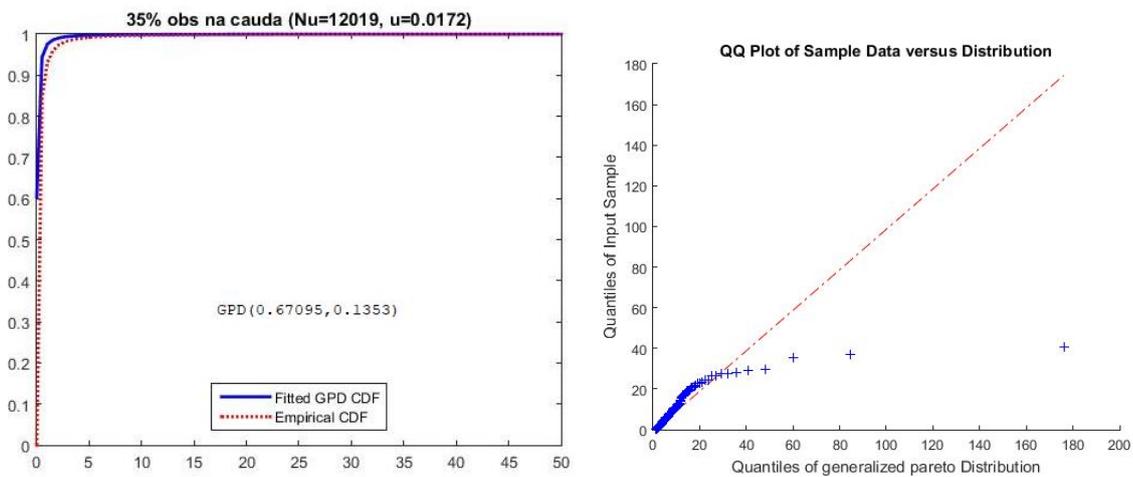


Figura 3.11: GPD para $N_u = 12019$ (35% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações t-Student

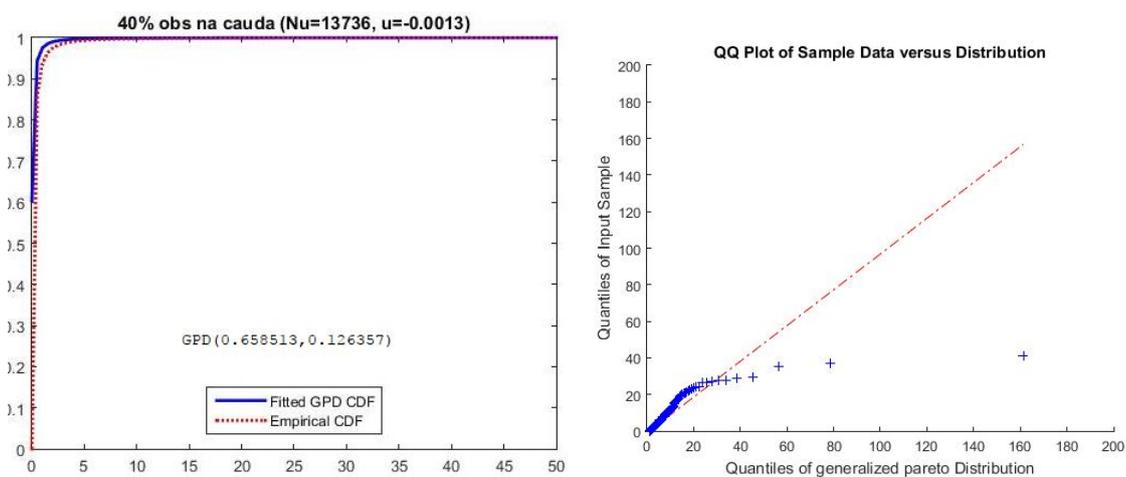


Figura 3.12: GPD para $N_u = 13736$ (40% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações t-Student

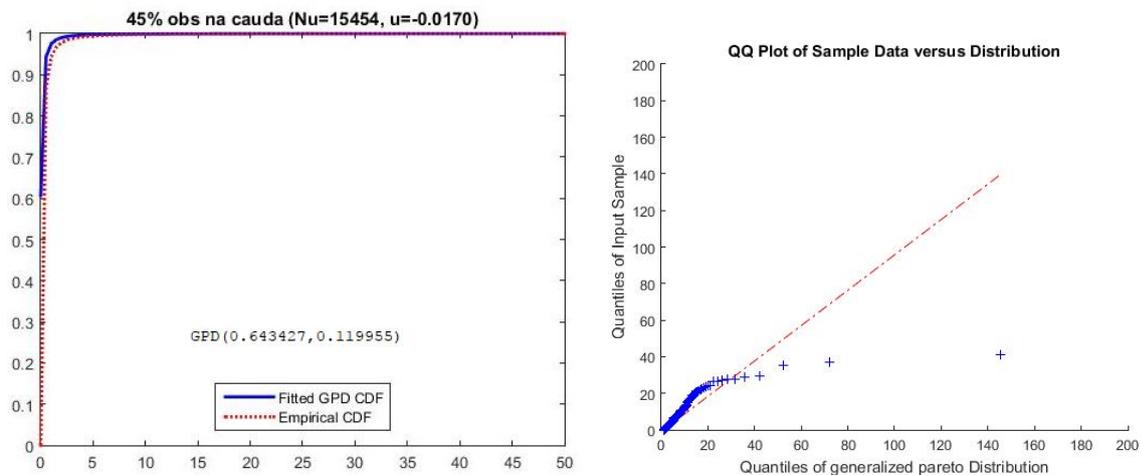


Figura 3.13: GPD para $N_u = 15454$ (45% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações t-Student

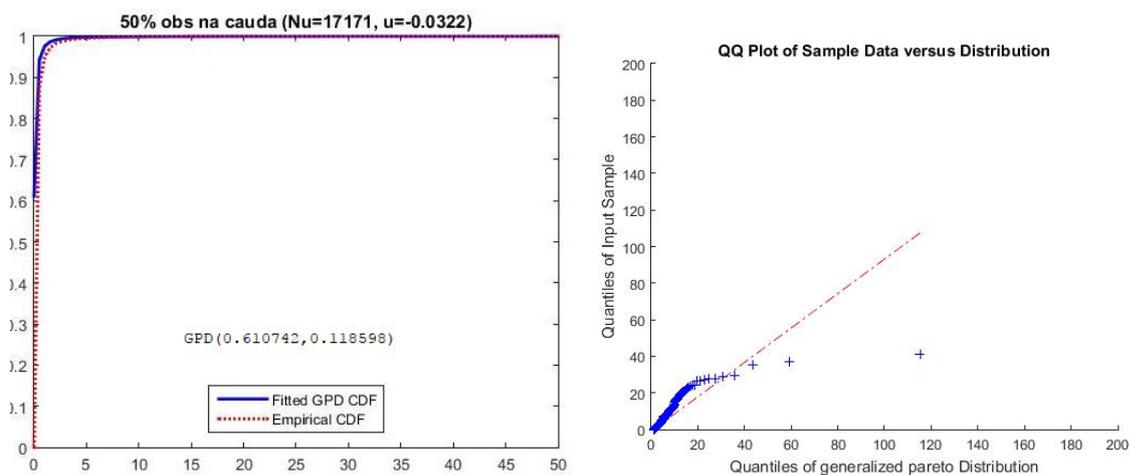


Figura 3.14: GPD para $N_u = 17171$ (50% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações t-Student

Modelação dos resíduos do modelo AR-GARCH com inovações gaussianas

A Figura 3.15 mostra os gráficos da função de autocorrelação dos resíduos e do seu quadrado, que indicam a presença de alguma autocorrelação nos resíduos normalizados com inovações gaussianas. Na Figura 3.16 podem observar-se os gráficos dos resíduos e da volatilidade destes, que sugerem que o modelo com inovações gaussianas também capta parte da heterocedasticidade.

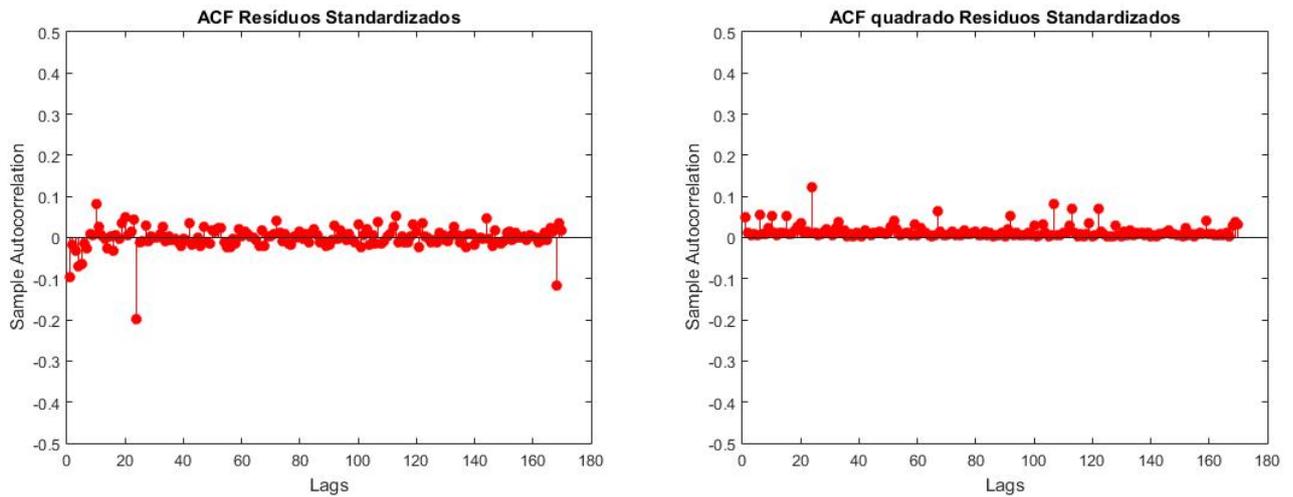


Figura 3.15: ACF resíduos normalizados - Modelo AR-GARCH com inovações gaussianas

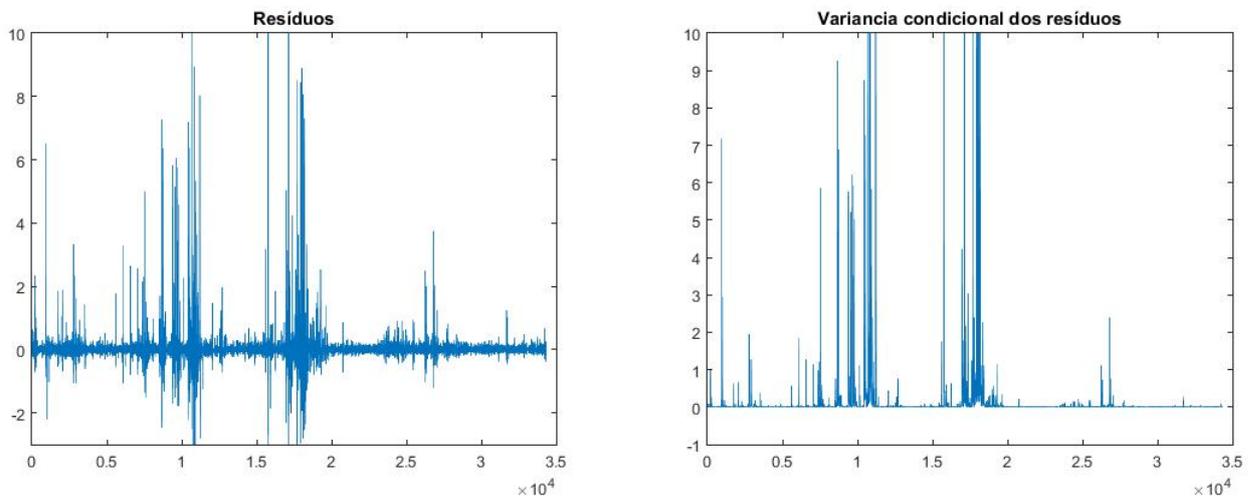


Figura 3.16: Resíduos - Modelo AR-GARCH com inovações gaussianas

Para escolher o *threshold*, e à semelhança do que se fez para a modelação dos resíduos usando o AR-GARCH com inovações t-Student, serão combinados os gráficos do estimador de Hill (Figura 3.17) para o índice de cauda, o gráfico da média dos excessos (Figura 3.19) e os gráficos para os estimadores do índice de cauda e de escala simultaneamente (Figura 3.18).

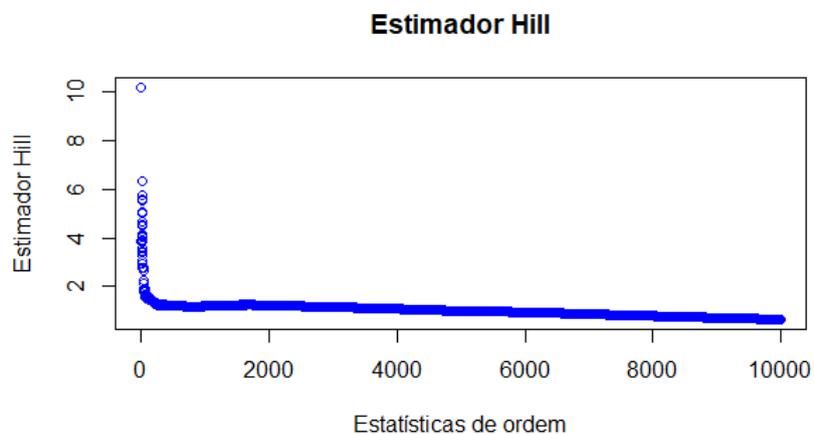
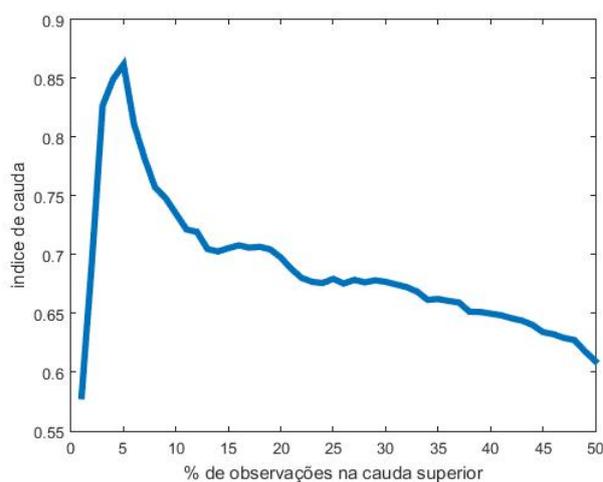


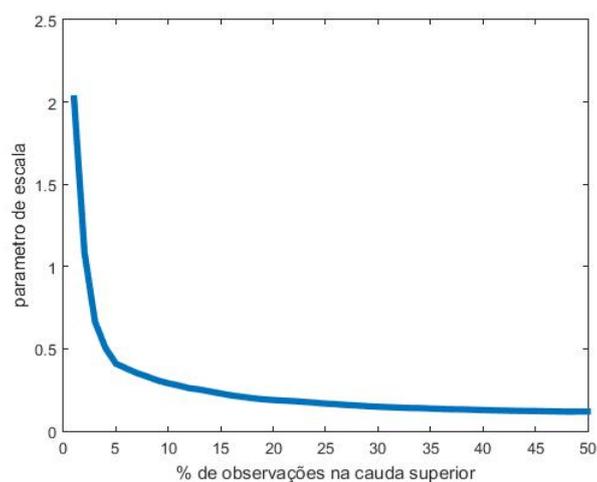
Figura 3.17: Hill Plot resíduos normalizados - Modelo AR-GARCH com inovações gaussianas

Tabela 3.7: N.º Excedências vs Threshold - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas

%	10%	20%	25%	30%	35%	40%	45%	50%
N.º Excedências	3434	6868	8585	10320	12019	13736	15454	17171
threshold (u)	0.2832	0.1126	0.0705	0.0415	0.0183	-0.0004	-0.0163	-0.0309



(a) Índice de Cauda



(b) Parâmetro de escala

Figura 3.18: Método de Eyeball - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas

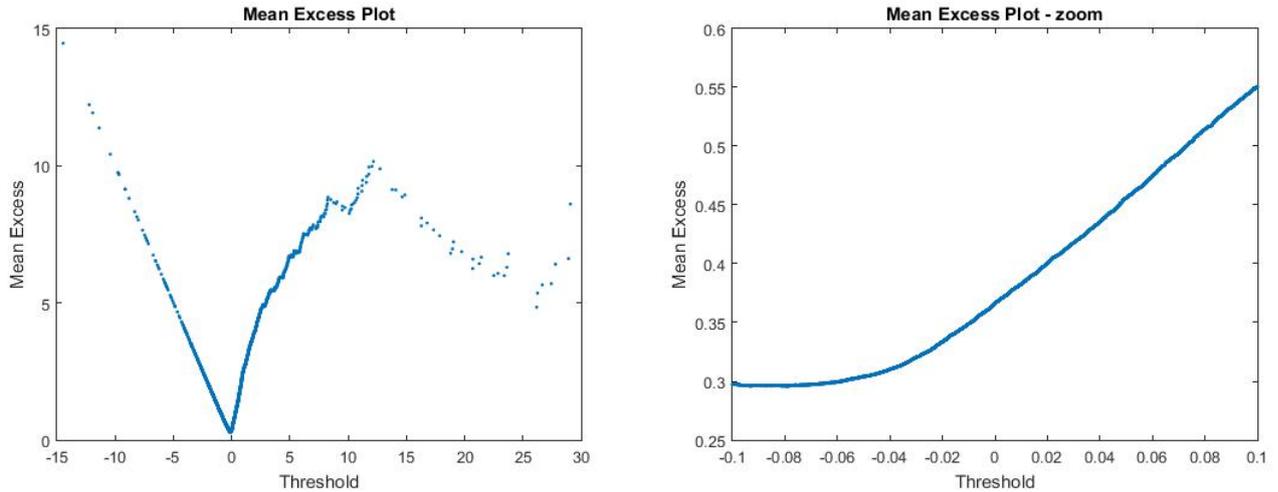


Figura 3.19: Mean Excess Plot - resíduos normalizados do modelo AR-GARCH com inovações gaussianas

O gráfico do estimador de Hill mantém-se estável (quase linear), sendo apenas decrescente de forma acentuada para valores próximos de zero. Por outro lado, o estimador do índice de cauda pelo método de eyeball (Figura 3.18) estabiliza por volta dos 25%, ou seja considerando cerca de 8585 observações na cauda e um *threshold* de 0.0705, mantendo esse comportamento até cerca dos 35% (12019 excedências - $threshold = 0.0183$). O gráfico da média dos excessos (Figura 3.19) é linear a partir de $u = -0.04$. Nas Figuras 3.20 - 3.25 está presente o ajustamento da GPD para diferentes valores de N_u ($N_u = 8585, 10320, 12019, 13736, 15454$ e 17171), considerando a percentagem de observações excedentes entre 25% e 50%.

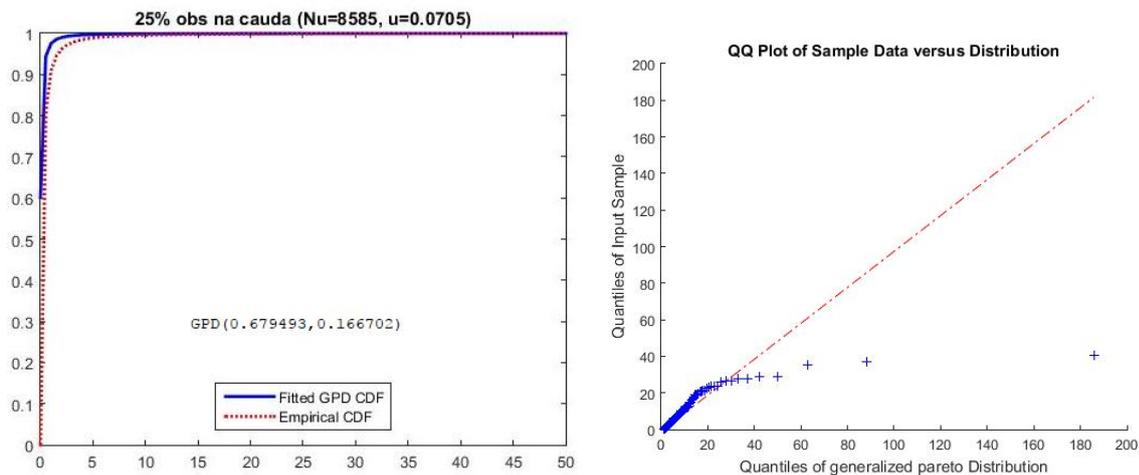


Figura 3.20: GPD para $N_u = 8585$ (25% das observações) - resíduos normalizados do modelo AR-GARCH com inovações gaussianas

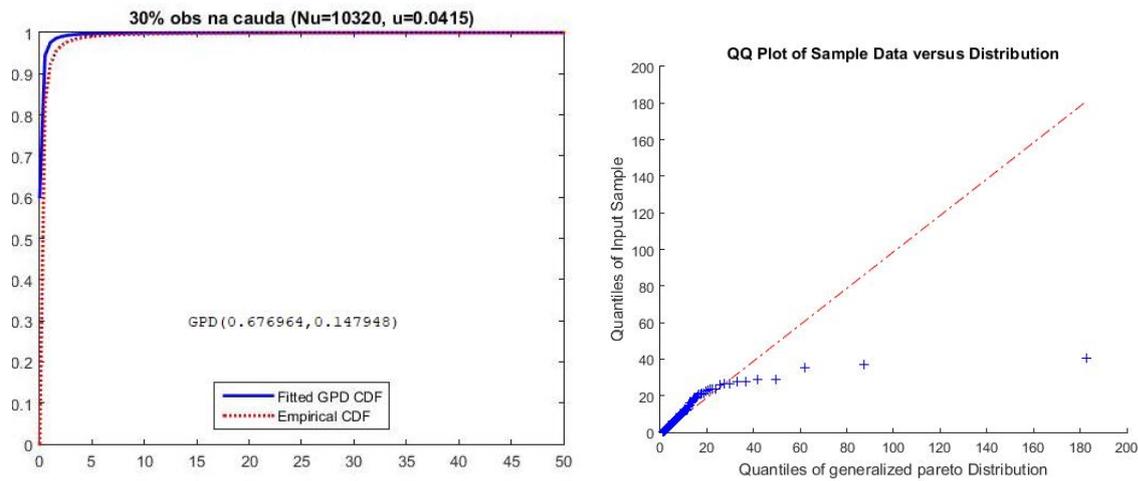


Figura 3.21: GPD para $N_u = 10320$ (30% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas

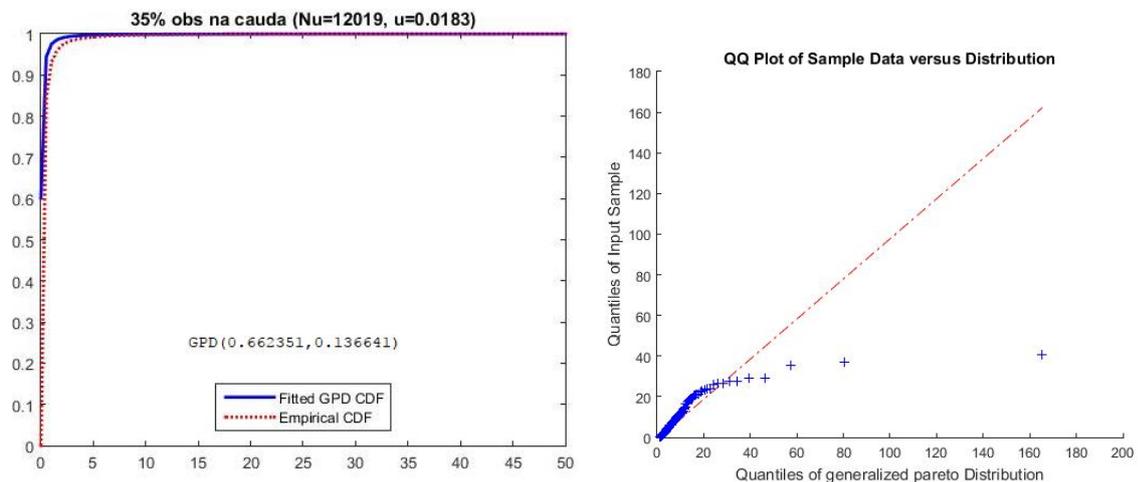


Figura 3.22: GPD para $N_u = 12019$ (35% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas

Analisando o comportamento dos *QQ-Plots* presentes nas Figuras 3.20 - 3.25, pode excluir-se as distribuições representadas nas duas últimas figuras. Esta exclusão é suportada pelo comportamento instável do estimador de ξ no gráfico da Figura 3.18a a partir de 40% observações extremas. Pode também excluir-se a distribuição no caso em que se consideram 25% dos resíduos na cauda (Figura 3.20), pois o gráfico do ajustamento (gráfico do lado esquerdo da Figura 3.20) evidencia uma grande diferença entre a distribuição ajustada e a empírica. De entre os restantes casos, será de escolher a GPD com parâmetros $\hat{\xi} = 0.68$ e $\hat{\sigma} = 0.15$ (Figura 3.21). Neste caso, consideram-se 10320 observações "extremas" e um *threshold* igual a 0.0415.

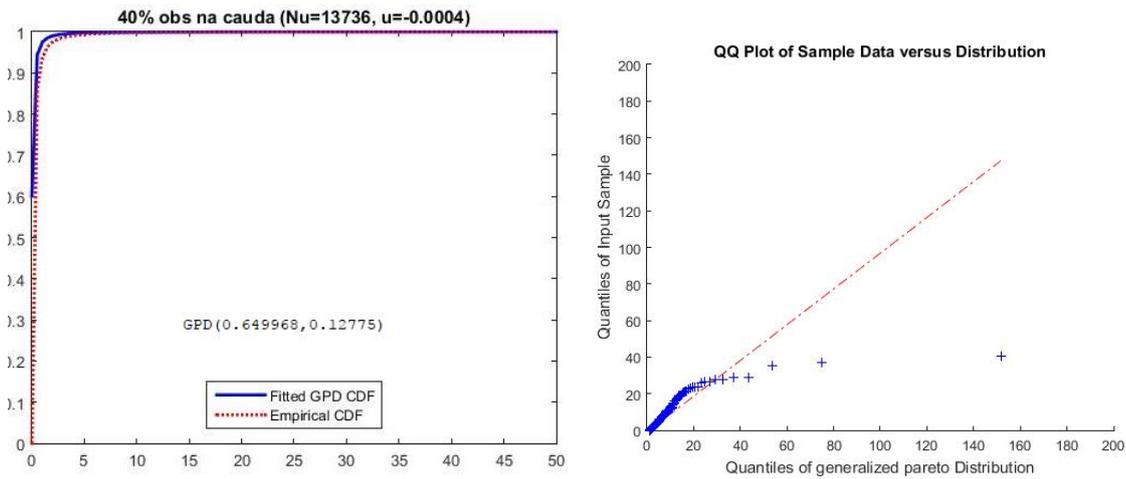


Figura 3.23: GPD para $N_u = 13736$ (40% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas

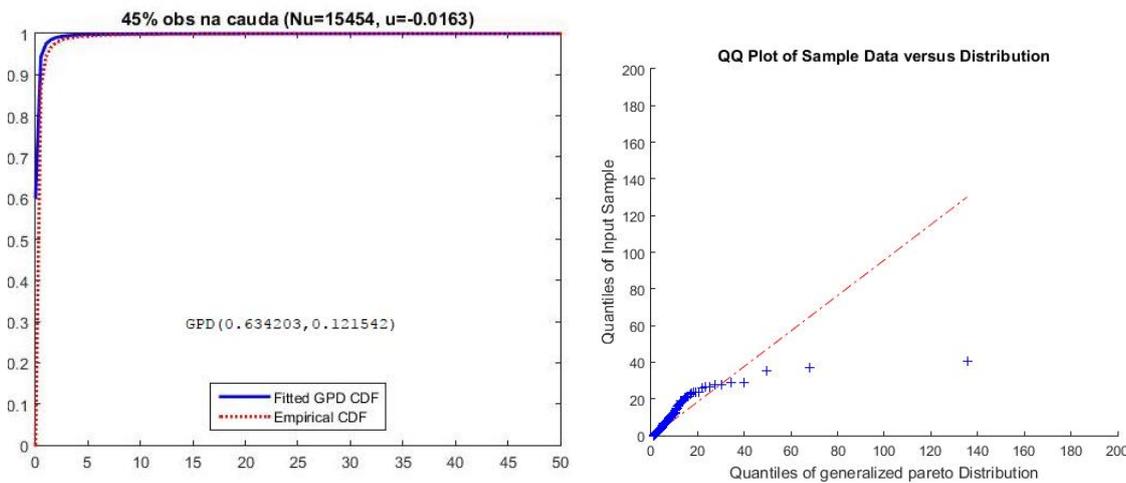


Figura 3.24: GPD para $N_u = 15454$ (45% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas

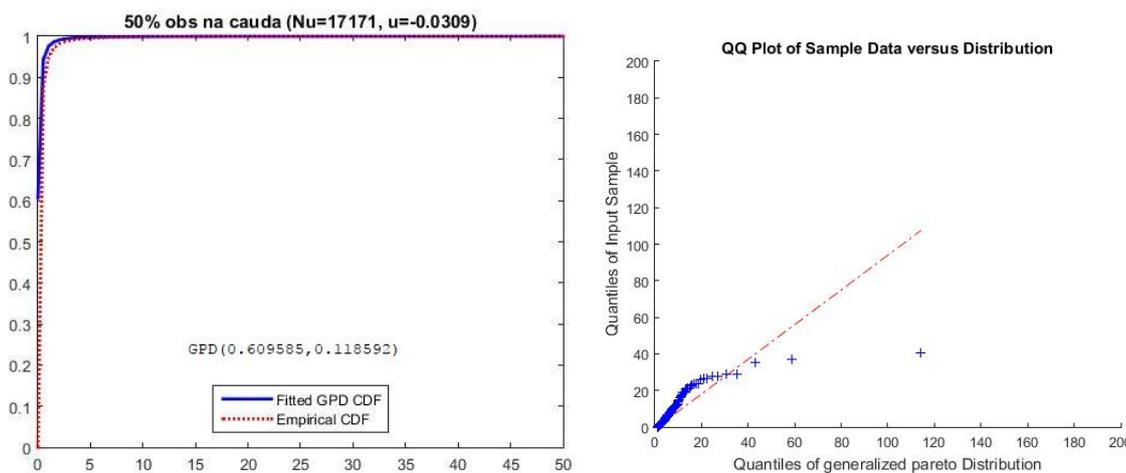


Figura 3.25: GPD para $N_u = 17171$ (50% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas

3.4 Estimação de quantis condicionais e discussão

Tendo as estimativas para os parâmetros para os modelos AR-GARCH e para a GPD ajustada aos resíduos normalizados, têm-se condições para estimar os quantis condicionais:

$$\alpha_{t,p} = \alpha_0 + \alpha_1 r_{t-1} + \alpha_2 r_{t-24} + \alpha_3 r_{t-168} + \sigma_t q_p$$

Usando as estimativas para o quantil condicional dos diferentes modelos, avaliar-se-á a sua precisão, comparando o número de excedências estimadas em cada modelo com o número de observações (retornos) que "teoricamente" são consideradas extrema. Por exemplo, para o quantil 0.95, tem-se 1717 retornos considerados valores extremos se o threshold for tal que 5% das observações pertencem à cauda. Este teste "in-sample" foi feito, considerando a amostra toda, ou seja, 34341 observações.

Assim, determinou-se o quantil condicional $\alpha_{34341,p}$, para $p = \{0.95, 0.99, 0.995, 0.999, 0.9995, 0.9999\}$, considerando o quantil q_p de cada distribuição associada a cada modelo.

Tabela 3.8: Estimação do n.º de excedências para diferentes probabilidades

Probabilidade	N_u	AR-GARCH-N(0,1)	AR-GARCH- $t_{2.7}$	GPD - $t_{2.7}$	GPD-N(0,1)
0,95	1717	2929	2034	14051	11379
0,99	343	1682	652	2864	2116
0,995	172	1413	418	1250	915
0,999	34	1046	188	229	188
0,9995	17	931	125	120	99
0,9999	3	777	56	26	20

Dos resultados da Tabela 3.8, pode verificar-se que, entre os modelos AR-GARCH, aquele que tem inovações t-Student se ajusta melhor aos dados e tem um melhor desempenho na estimação dos quantis. Também se pode observar que este modelo melhora à medida que diminui o número de excedências, ao contrário do modelo com inovações gaussianas. Dos modelos com a abordagem TVE, aquele que parte da modelação dos resíduos com inovações gaussianas tem estimativas mais próximas do valor empírico. Quando comparados com os modelos AR-GARCH, ambos os modelos resultantes da aplicação da TVE parecem apresentar melhores estimativas para o número de observações extremas. Estes resultados vão de encontro à opinião de que os modelos de séries temporais são mais adequados para a modelação da distribuição cujo comportamento é mais estável, e portanto não captam tão bem os eventos extremos. Surpreendentemente, o modelo AR-GARCH cujas inovações seguem uma distribuição Normal apresenta um melhor desempenho em relação àquele com inovações t-Student.

3.5 Avaliação dos Modelos

Para avaliação dos modelos apresentados neste trabalho, consideraram-se apenas os dados correspondentes aos anos civis 2014 e 2015, sendo que os dados do ano de 2014 serão utilizados para estimar os coeficientes dos modelos (conjunto de dados de treino) e os dados de 2015 servirão para testar os modelos (conjunto de dados de teste). Os coeficientes encontrados para os modelos AR-GARCH, tendo em contas as lags 1, 24 e 168, estão apresentados na Tabela 3.9.

Tabela 3.9: Estimativas para os coeficientes do modelo AR-GARCH (Ano 2014): lag 1, 24 e 168

$Z_t \sim N(0, 1)$	Coef. estimado	Desvio-padrão
α_0	0.00156981	0.000698654
α_1	0.190064	0.00586646
α_2	0.377172	0.00526797
α_3	0.214955	0.00124667
ϕ_0	3.72056×10^{-5}	1.72944×10^{-6}
ϕ_1	0.059348	0.000521097
ϕ_2	0.940652	0.000424939
Log V.	5710.3	
$Z_t \sim t_\nu$	Coef. estimado	Desvio-padrão
α_0	-0.00261208	0.000527485
α_1	0.193104	0.00601409
α_2	0.270256	0.0054526
α_3	0.184742	0.00455807
ϕ_0	0.00022234	3.39092×10^{-5}
ϕ_1	0.0771722	0.0105133
ϕ_2	0.922828	0.00386754
Log V.	8562.2	
Graus de liberdade	2.39031	0.066453

À semelhança do que foi feito na secção anterior, os resíduos do modelo AR-GARCH serão modelados aplicando a TVE. De acordo com os resultados da Tabela 3.8, a abordagem TVE através da modelação dos resíduos cujas inovações seguem uma distribuição Normal pela GPD tem um melhor desempenho, pelo que para avaliação desta abordagem será considerado apenas esse caso. Para escolher o valor do *threshold* serão considerados os métodos gráficos já introduzidos no capítulo 2, ou seja, Hill Plot, Mean Excess Plot e Eyeball, que aqui estão representados nas Figuras 3.26, 3.27 e 3.28, respetivamente.

Combinando a análise dos 3 métodos, é aceitável considerar 30% das maiores observações (2538 observações), escolhendo assim um *threshold* igual a 0.0256. Na Figura 3.29 está presente o ajustamento da GPD para o número de observações considerado, com os parâmetros estimados $\hat{\xi} = 0.8$ e $\hat{\sigma} = 0.13$.

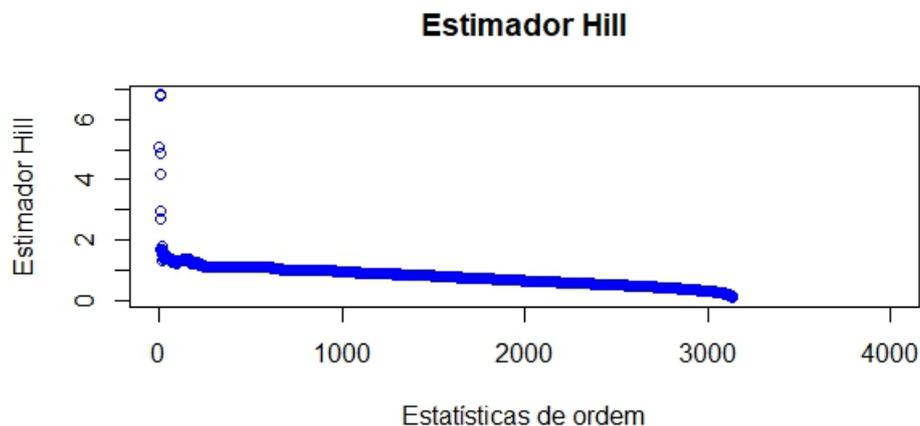


Figura 3.26: Hill Plot dos resíduos normalizados - Modelo AR-GARCH com inovações gaussianas (Ano 2014)

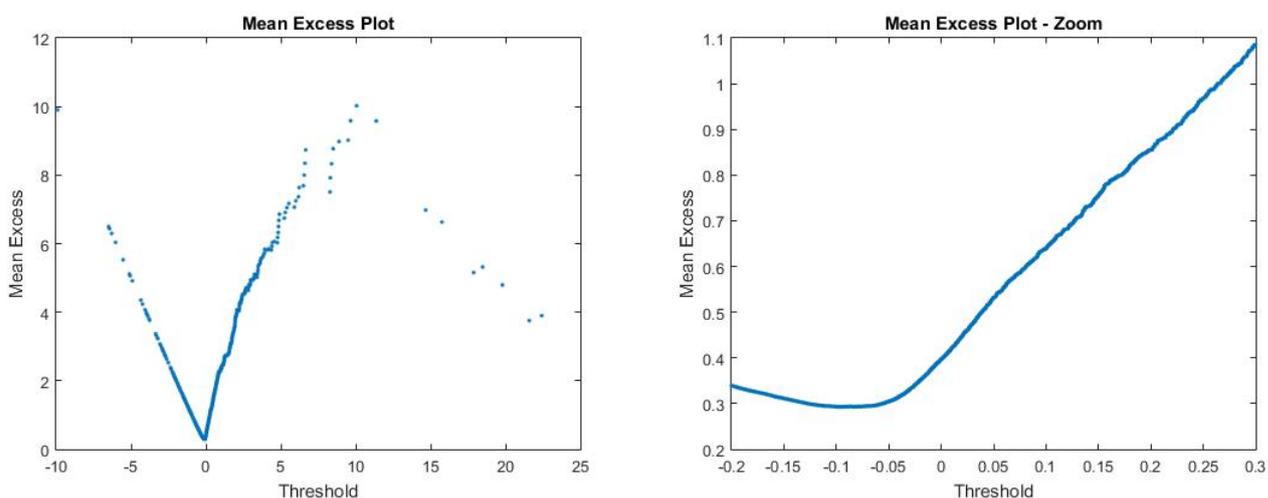


Figura 3.27: Mean Excess Plot - resíduos normalizados do modelo AR-GARCH com inovações gaussianas (Ano 2014)

Os dados relativos ao ano de 2015 serão então utilizados para avaliação dos modelos através dos coeficientes encontrados, comparando os quantis condicionais obtidos em 3 casos distintos: modelo AR-GARCH com inovações gaussianas, modelo AR-GARCH com inovações t-Student e modelo com a abordagem TVE (considerando os resíduos do modelo AR-GARCH com inovações gaussianas). A avaliação será feita através das medidas de *reliability* e *sharpness* [2], a seguir apresentadas.

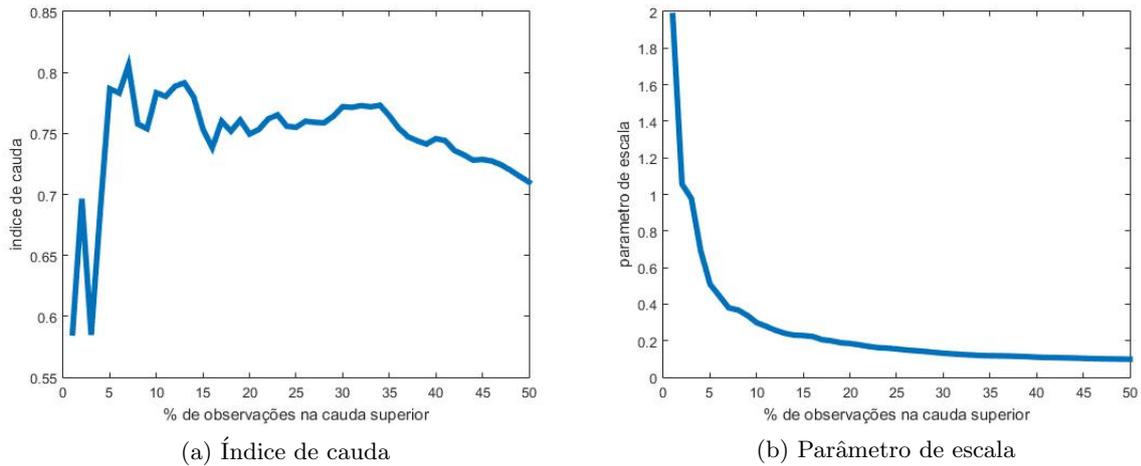


Figura 3.28: Método de Eyeball - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas (Ano 2014)

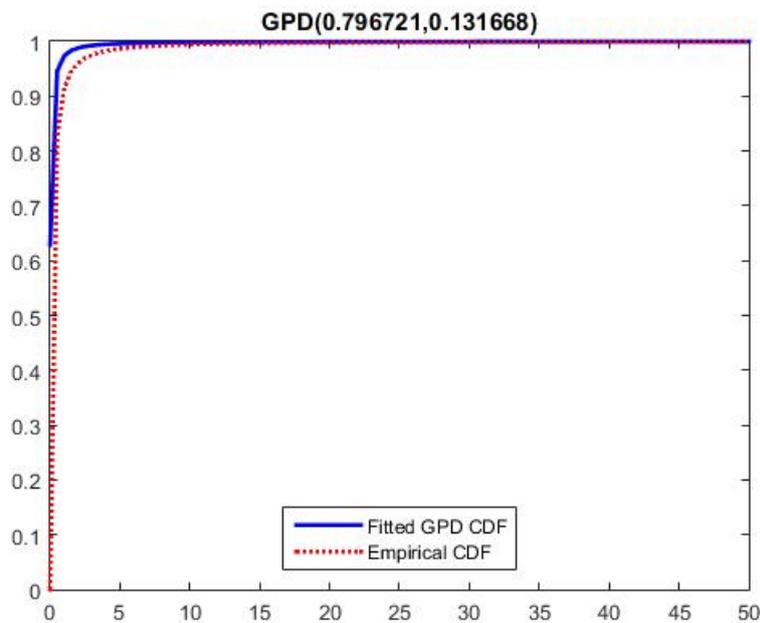


Figura 3.29: GPD para $N_u = 2538$ (30% das observações) - resíduos normalizados do Modelo AR-GARCH com inovações gaussianas (Ano 2014)

Reliability

A *Reliability* mede o desvio entre o quantil estimado e o correspondente quantil nominal. Para tal, é necessário introduzir a variável indicatriz $\zeta_{t,k}^{(\alpha)}$. Dado o quantil estimado no instante t para o instante $t+k$, $\hat{q}_{t+k|t}$, e o valor real observado p_{t+k} , $\zeta_{t,k}^{(\alpha)}$ é dado por:

$$\zeta_{t,k}^{(\alpha)} = \begin{cases} 1, & \text{se } p_{t+k} < \hat{q}_{t+k|t} \\ 0, & \text{caso contrário} \end{cases}$$

Um estimador $\hat{a}_k^{(\alpha)}$ para a proporção $a_k^{(\alpha)}$ é dado por $\hat{a}_k^{(\alpha)} = \hat{E}[\zeta_{t,k}^{(\alpha)}] = \frac{1}{N} \sum_{t=1}^N \zeta_{t,k}^{(\alpha)}$ e o desvio entre o quantil nominal α e o estimado é dado por:

$$b_k^{(\alpha)} = \alpha - \hat{a}_k^{(\alpha)}$$

A Figura 3.30 mostra a comparação da *reliability* para os diferentes modelos. Apesar de se notar que o modelo com inovações gaussianas apresenta maior estabilidade nos quantis aqui apresentados, o modelo utilizando a GPD vai melhorando à medida que se aumenta o quantil, sendo que este modelo tem melhor desempenho para quantis mais extremos, tal como era esperado. Na Figura 3.31 está presente a *reliability* para quantis superiores a 0.988 e é evidente o desvio que se obtém quando se estimam quantis considerando inovações gaussianas. Também se pode verificar a proximidade entre o modelo com inovações t-Student e o modelo que inclui a GPD quando estimados quantis elevados.

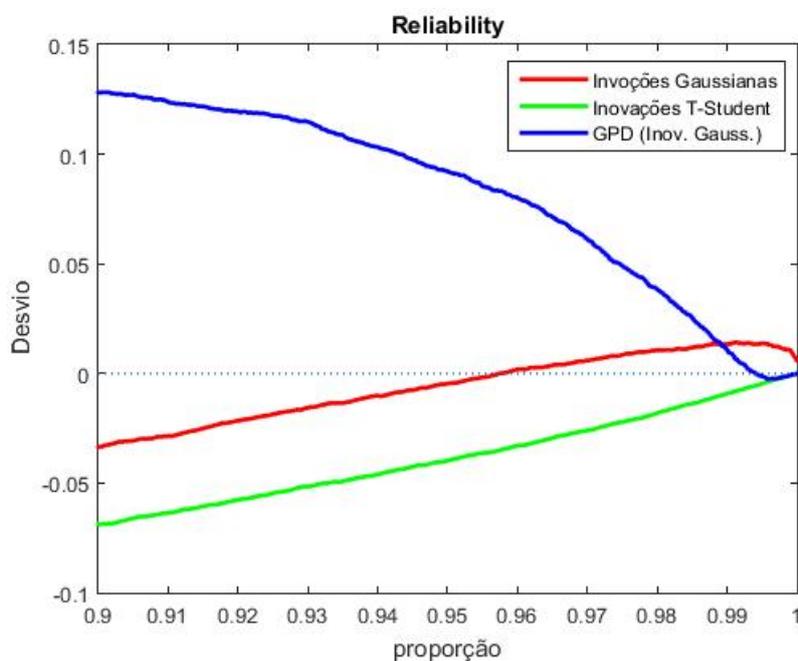


Figura 3.30: Diagrama de *reliability*

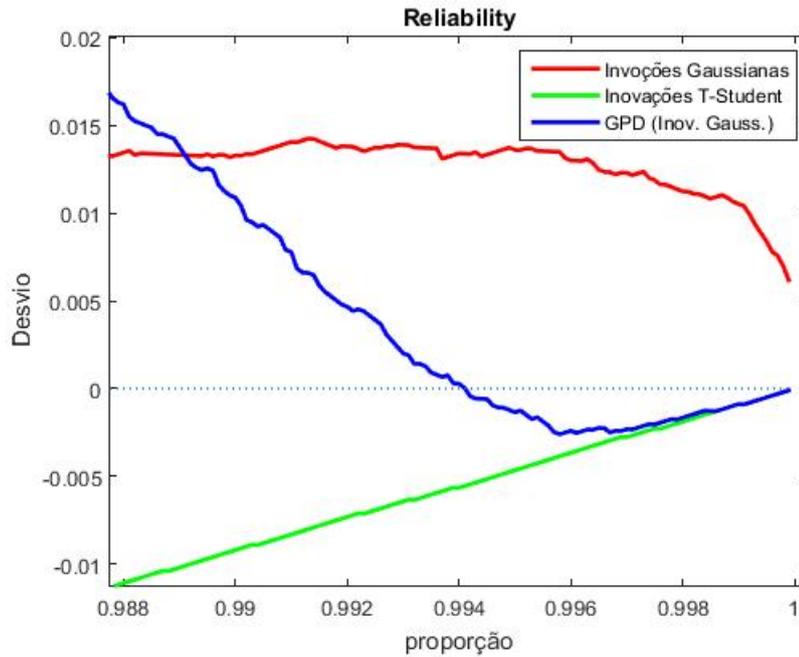


Figura 3.31: Diagrama de *reliability* - Zoom

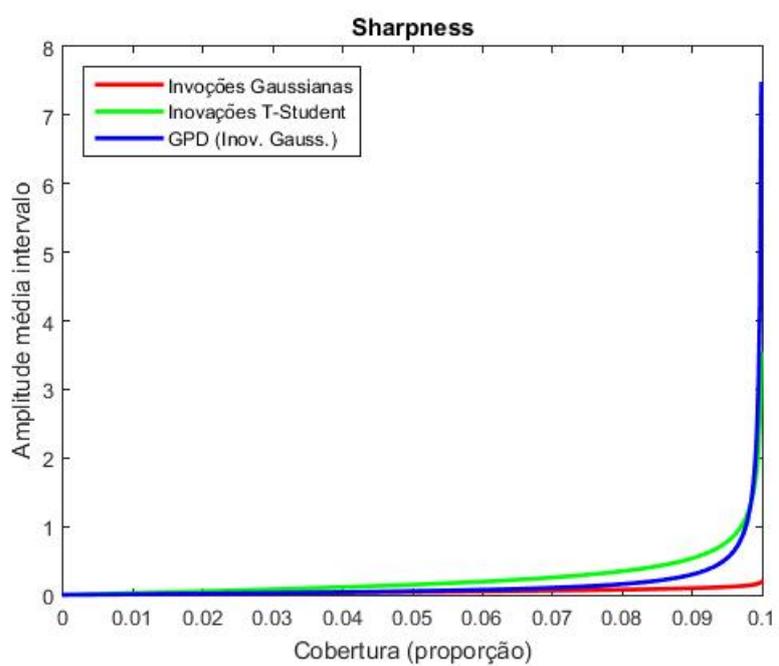
Sharpness

A *sharpness* mede a distância entre dois quantis, permitindo avaliar a forma da distribuição, a fim de obter um conjunto de intervalos centrais estimados com diferentes taxas de cobertura nominal. O diagrama de *sharpness* mostra o tamanho médio dos intervalos em função da taxa de cobertura nominal.

Seja $\delta_{t,k}^{(\beta)} = \hat{q}_{t+k|t}^{(1-\beta/2)} - \hat{q}_{t+k|t}^{(\beta/2)}$, o tamanho do intervalo central estimado no instante t para o instante $t+k$, com taxa de cobertura nominal $(1-\beta)$, a *sharpness* para estes intervalos e para o horizonte k é dada pela média dos seus tamanhos:

$$\bar{\delta}_k^{(\beta)} = \frac{1}{N} \sum_{t=1}^N (\hat{q}_{t+k|t}^{(1-\beta/2)} - \hat{q}_{t+k|t}^{(\beta/2)})$$

A *sharpness* para os modelos em análise está representada na Figura 3.32. De modo geral, pode observar-se que o modelo com inovações gaussianas tem melhor comportamento em termos de *sharpness*.

Figura 3.32: Diagrama de *Sharpness*

Capítulo 4

Comentários finais

O trabalho desenvolvido teve como objetivo principal a modelação de eventos extremos nos preços da eletricidade usando a TVE. Para tal houve necessidade de lidar adequadamente com a dependência existente nos dados. Motivados pelo estudo desenvolvido em [31] aplicou-se um modelo AR-GARCH aos retornos associados aos preços e posteriormente usou-se a TVE clássica para modelar os resíduos. Para encontrar o melhor modelo que descreve os retornos, testaram-se quatro modelos AR-GARCH, dois com inovações gaussianas e dois com inovações t-Student. Em cada um destes casos incluiu-se para além das lags 1 e 24, a lag 168 correspondente ao número de horas de uma semana. Os resultados dos critérios AIC e BIC permitiram concluir que os modelos que incluem o termo AR(168) são mais apropriados. Constatou-se igualmente que o modelo cujas inovações seguem uma distribuição t-Student descrevem melhor a série dos retornos. A análise feita aos resíduos de ambos os modelos corrobora a ideia inicial de que estes modelos conseguem captar bem a volatilidade dos retornos. Também se verificou que a maioria da autocorrelação foi eliminada por estes modelos.

Na modelação dos resíduos através da TVE deu-se especial atenção à escolha do *threshold*. Este valor tem uma grande influência na estimação dos parâmetros da Distribuição Generalizada de Pareto (GPD). Para uma adequada escolha, recorreu-se a métodos gráficos, como o gráfico de Hill e o gráfico da média dos excessos. Depois de estimados os quantis condicionais dos retornos, foi feita uma avaliação para comparar os modelos considerados, sugerindo que a TVE consegue captar de melhor forma o comportamento das caudas dos retornos.

Para analisar o desempenho de cada modelo, foram utilizadas as medidas de avaliação *reliability* e *sharpness*. Os valores obtidos para estas medidas favorecem a aplicação da TVE para a estimação de valores extremos, dado que o modelo que inclui a GPD apresentou melhores resultados para quantis elevados.

Bibliografia

- [1] <http://www.mibel.com>. [Online; última visita: 28-Set-2017].
- [2] A. Eydeland and K. Wolyniec, *Energy and Power Risk Management: New Developments in Modeling, Pricing, and Hedging*, vol. 206. John Wiley & Sons, 2003.
- [3] D. Pilipovic, *Energy Risk: Valuing and Managing Energy Derivatives*. McGraw Hill Professional, 2007.
- [4] R. Weron, “Market price of risk implied by asian-style electricity options,” tech. rep., EconWPA, 2005.
- [5] F. Paraschiv, “Price dynamics in electricity markets,” in *Handbook of Risk Management in Energy Production and Trading*, pp. 47–69, Springer, 2013.
- [6] R. Weron, “Electricity price forecasting: A review of the state-of-the-art with a look into the future,” *International Journal of Forecasting*, vol. 30, no. 4, pp. 1030–1081, 2014.
- [7] E. Gonçalves and N. M. Lopes, *Séries Temporais. Modelações lineares e não lineares*. Sociedade Portuguesa de Estatística, 2^a edição, 2008.
- [8] G. U. Yule, “On a method of investigating periodicities in disturbed series, with special reference to wolfer’s sunspot numbers,” *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 226, pp. 267–298, 1927.
- [9] G. E. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control, revised ed.* Holden-Day, 1976.
- [10] R. F. Engle, “Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation,” *Econometrica: Journal of the Econometric Society*, pp. 987–1007, 1982.
- [11] T. Bollerslev, “Generalized autoregressive conditional heteroskedasticity,” *Journal of Econometrics*, vol. 31, no. 3, pp. 307–327, 1986.
- [12] F. Rebelo, “Um novo olhar sobre os riscos? o exemplo das cheias rápidas (flash floods) em domínio mediterrâneo,” in *Territorium*, pp. 7–14, RISCOS-Associação Portuguesa de Riscos, Prevenção e Segurança, 2008.

- [13] M. A. Baptista, S. Heitor, J. M. Miranda, P. Miranda, and L. M. Victor, “The 1755 lisbon tsunami; evaluation of the tsunami parameters,” *Journal of Geodynamics*, vol. 25, no. 1-2, pp. 143–157, 1998.
- [14] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*, vol. 208. Springer, 2001.
- [15] P. Embrechts, C. Klüppelberg, and T. Mikosch, *Modelling Extremal Events: for Insurance and Finance*. Springer-Verlag, 1997.
- [16] M. Leadbetter, H. Rootzén, and G. Lindgren, *Extremes and Related Properties of Random Sequences and Processes*. Springer, 1983.
- [17] J. T. De Oliveira, “Perspectivas sobre a estatística de extremos, resultados básicos e problemas em aberto,” *Revista de la Real Academia de Ciencias Exactas Físicas y Naturales*, vol. 84, no. 3, pp. 453–481, 1990.
- [18] E. L. Dodd, “The greatest and the least variate under general laws of error,” *Transactions of the American Mathematical Society*, vol. 25, no. 4, pp. 525–539, 1923.
- [19] M. Fréchet, “Sur la loi de probabilité de l’écart maximum,” in *Annales de la societe Polonaise de Mathematique*, [sn], 1928.
- [20] R. A. Fisher and L. H. C. Tippett, “Limiting forms of the frequency distribution of the largest or smallest member of a sample,” in *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 24, pp. 180–190, Cambridge Univ Press, 1928.
- [21] B. Gnedenko, “Sur la distribution limite du terme maximum d’une serie aleatoire,” *Annals of mathematics*, pp. 423–453, 1943.
- [22] R. Von Mises, “La distribution de la plus grande de n valeurs,” *Rev. math. Union Interbalcanique*, vol. 1, no. 1, 1936.
- [23] E. J. Gumbel, “The return period of flood flows,” *The Annals of Mathematical Statistics*, vol. 12, no. 2, pp. 163–190, 1941.
- [24] A. A. Balkema and L. De Haan, “Residual life time at great age,” *The Annals of probability*, pp. 792–804, 1974.
- [25] J. Pickands III, “Statistical inference using extreme order statistics,” *The Annals of Statistics*, pp. 119–131, 1975.
- [26] A. J. McNeil, “Extreme value theory for risk managers,” *Departement Mathematik ETH Zentrum*, 1999.
- [27] V. O. Andreeva, S. E. Tinykov, O. P. Ovchinnikovaa, and G. P. Parahinc, “Extreme value theory and peaks over threshold model in the russian stock market,” *Journal of Siberian Federal University, Engineering and Technologies*, vol. 1, no. 5, pp. 111–121, 2012.

- [28] A. J. McNeil and R. Frey, “Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach,” *Journal of Empirical Finance*, vol. 7, no. 3, pp. 271–300, 2000.
- [29] F. X. Diebold, T. Schuermann, and J. D. Stroughair, “Pitfalls and opportunities in the use of extreme value theory in risk management,” *The Journal of Risk Finance*, vol. 1, no. 2, pp. 30–35, 2000.
- [30] H. N. Byström, “Extreme value theory and extremely large electricity price changes,” *International Review of Economics & Finance*, vol. 14, no. 1, pp. 41–55, 2005.
- [31] F. Paraschiv, R. Hadzi-Mishev, and D. Keles, “Extreme value theory for heavy-tails in electricity prices,” *The Journal of Energy Markets*, no. tba, 2015.
- [32] D. A. Dickey and W. A. Fuller, “Distribution of the estimators for autoregressive time series with a unit root,” *Journal of the American statistical association*, vol. 74, no. 366a, pp. 427–431, 1979.
- [33] P. C. Phillips and P. Perron, “Testing for a unit root in time series regression,” *Biometrika*, vol. 75, no. 2, pp. 335–346, 1988.
- [34] D. Kwiatkowski, P. C. Phillips, P. Schmidt, and Y. Shin, “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?,” *Journal of econometrics*, vol. 54, no. 1-3, pp. 159–178, 1992.
- [35] J. Fan and Q. Yao, “Nonlinear time series: Nonparametric and parametric methods,” *New York*, 2003.
- [36] G. M. Ljung and G. E. Box, “On a measure of lack of fit in time series models,” *Biometrika*, vol. 65, no. 2, pp. 297–303, 1978.
- [37] R. Deidda and M. Puliga, “Performances of some parameter estimators of the generalized pareto distribution over rounded-off samples,” *Physics and Chemistry of the Earth, Parts A/B/C*, vol. 34, no. 10, pp. 626–634, 2009.
- [38] A. L. Dekkers, J. H. Einmahl, and L. De Haan, “A moment estimator for the index of an extreme-value distribution,” *The Annals of Statistics*, pp. 1833–1855, 1989.
- [39] B. M. Hill *et al.*, “A simple general approach to inference about the tail of a distribution,” *The Annals of Statistics*, vol. 3, no. 5, pp. 1163–1174, 1975.
- [40] N. Cappuccio and D. Lubian, “The fragility of the kpss stationarity test,” *Statistical Methods & Applications*, vol. 19, no. 2, pp. 237–253, 2010.
- [41] E. Gonçalves and N. M. Lopes, *Probabilidades: Princípios Teóricos*. Escolar Editora, 2000.
- [42] Y. Aragon, “R pour les séries temporelles,” *Séries Temporelles avec R*, pp. 44–45, 2011.

- [43] C. Moler *et al.*, *MATLAB users' guide*. University of New Mexico, 1982.
- [44] R. Gençay, F. Selçuk, and A. Ulugülyagci, “Evim: A software package for extreme value analysis in matlab,” *Studies in Nonlinear Dynamics and Econometrics*, pp. 213–239, 2001.
- [45] R Core Team, *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, 2014.
- [46] M. Brito, L. Cavalcante, and A. C. M. Freitas, “Bias-corrected geometric-type estimators of the tail index,” *Journal of Physics A: Mathematical and Theoretical*, vol. 49, no. 21, p. 214003, 2016.
- [47] G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, 2015.

Anexos

Anexo A

Ferramentas de Programação

Para a implementação dos modelos referidos, foi utilizado o software MATLAB (MATrix LABoratory) criado no fim da década de 70 por Clever Moler [43]. Este software destina-se ao cálculo envolvendo matrizes e foi utilizado para reproduzir os resultados apresentados nesta tese, nomeadamente a estimação dos parâmetros dos modelos de séries temporais através da *toolbox Econometrics* e a modelação da distribuição generalizada de Pareto para os valores extremos com as funções existentes na *toolbox Probability Distributions* e no *package Evim* [44].

Foi também utilizado o software *R* [45] para auxílio na determinação do *threshold* a considerar na modelação dos valores extremos pela GPD [46], nomeadamente a construção do gráfico de Hill.

Anexo B

Modelos ARMA

Os modelos ARMA - *Auto Regressive Moving Average* foram introduzidos por Box e Jenkins na década de 70 [47], sendo nos dias de hoje os modelos que mais se aplicam devido à sua versatilidade: podemos obter uma série auto-regressiva pura (AR), uma média móvel (MA) ou conjugar os dois (ARMA).

Modelos Auto-Regressivos (AR)

Um modelo Auto-regressivo (AR) caracteriza-se pela possibilidade de escrever a série como função linear dos seus valores passados. Formalmente,

Definição 1. *Um processo estocástico centrado $X = (X_t, t \in \mathbb{Z})$ de segunda ordem admite uma representação auto-regressiva de ordem p , $X_t \sim AR(p)$, se existem números reais $\varphi_1, \dots, \varphi_p$ e um ruído branco $\varepsilon_t (t \in \mathbb{Z})$ de variância $\sigma^2 > 0$, tais que*

$$X_t = \varphi_1 X_{t-1} + \dots + \varphi_p X_{t-p} + \varepsilon_t, \quad \varphi_p \neq 0$$

ε_t diz-se ruído branco se tiver média nula e variância constante para cada realização da série e apresentar correlação nula:

$$E(\varepsilon_t) = 0 \quad \forall t \quad V(\varepsilon_t) = \sigma^2 \quad \forall t, \quad cov(\varepsilon_i, \varepsilon_j) = 0, \forall i \neq j$$

Diz-se que esta representação é canónica se ε_t é não correlacionado com o passado de X à data t .

Seja B o operador de atraso associado a X tal que $BX_t = X_{t-1}$ e $\Phi(B) = 1 - \varphi_1 B - \dots - \varphi_p B^p$, $\Phi(B)$ polinómio característico, então temos a relação de recorrência

$$\Phi(B)X_t = \varepsilon_t$$

Propriedade Entre as soluções X de $\Phi(B)X_t = \varepsilon_t, \quad \forall t \in \mathbb{Z}$, há uma estacionária se e só se todas as raízes do polinómio $\Phi(B)$ forem de módulo diferente de 1. Além disso uma tal solução estacionária é única.

Sendo estes modelos caracterizados pela forte correlação com o passado, é natural assumir que a identificação de processos AR se relacione com a sua função de autocorrelação.

Efetivamente, tais condições baseiam-se no estudo da sucessão das autocorrelações $\rho(h)$, $h \in \mathbb{Z}$, que verifica uma equação de recorrência linear homogénea e é possível escrevê-la como combinações de funções exponenciais de h e na propriedade da sucessão das autocorrelações parciais $r(h)$, $h \in \mathbb{Z}$ que se anula a partir de uma certa ordem.

Em particular, a sucessão das autocorrelações parciais de um processo $AR(p)$ anula-se a partir da ordem $p + 1$. Esta última propriedade é extramente útil na prática, quando se pretende estimar a ordem do modelo auto-regressivo que se ajusta aos dados em estudo.

Modelos Médias Móveis (MA)

Quando a série temporal pode ser escrita a partir dos valores passados de um termo estocástico, designamos o modelo resultante por *Média Móvel*. Por exemplo, a evolução do preço de um ativo financeiro é uma média móvel. Formalmente tem-se a seguinte definição:

Definição 2. *Um processo estocástico $X = (X_t, t \in \mathbb{Z})$ de segunda ordem admite uma representação média móvel de ordem q - Moving Average, $X_t \sim MA(q)$, se*

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}, \quad \forall t \in \mathbb{Z}$$

onde $\theta_1, \dots, \theta_q \in \mathbb{R}$, $\theta_q \neq 0$, e $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ é um ruído branco de variância $\sigma^2 > 0$.

Considerando novamente o operador de atraso B , agora associado ao processo ε , podemos escrever a igualdade anterior na forma

$$X_t = \Theta(B)\varepsilon_t$$

$$\text{com } \Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$$

À semelhança dos processos auto-regressivos, também estes podem ser identificados através das propriedades de correlação com o passado. O Teorema seguinte descreve uma condição necessária e suficiente para que um processo estacionário verifique um processo Média Móvel:

Teorema B.1. *Um processo estacionário regular admite uma representação MA de ordem finita se e só se a sucessão das autocovariâncias se anula a partir de uma certa ordem. Se $q + 1$ designa esta ordem mínima, q é a ordem da representação MA.*

Modelar simultaneamente processos auto-regressivos e médias móveis conduz-nos, na maioria das vezes, a uma melhor aproximação das séries reais com um número de parâmetros inferior.

Um processo estocástico $X = (X_t, t \in \mathbb{Z})$ de segunda ordem admite uma representação *Auto-regressiva média móvel*, $X_t \sim ARMA(p, q)$, se verifica a equação:

$$X_t = \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}$$

ou de forma equivalente

$$\Phi(B)X_t = \Theta(B)\varepsilon_t$$

com $\Phi(B) = 1 - \varphi_1 B - \dots - \varphi_p B^p$, $\Theta(B) = 1 - \theta_1 B - \dots - \theta_q B^q$, $\varepsilon = (\varepsilon_t, t \in \mathbb{Z})$ um ruído branco de variância não nula.

Os polinómios Φ e Θ não têm raízes em comum e as suas raízes são de módulos estritamente superiores a 1.

Uma forma de identificar modelos ARMA de ordens mínimas baseia-se no cálculo de determinantes envolvendo a função de autocovariância γ do processo. Consideremos as matrizes seguintes:

$$A(i, j) = \begin{pmatrix} \gamma(j+1) & \gamma(j+2) & \dots & \gamma(j+i+1) \\ \gamma(j) & \gamma(j+1) & \dots & \gamma(j+i) \\ \dots & \dots & \dots & \dots \\ \gamma(j+1-i) & \gamma(j+2-i) & \dots & \gamma(j+1) \end{pmatrix}$$

e seja $\Delta(i, j) = \det A(i, j)$.

A propriedade seguinte dá-nos uma condição necessária e suficiente para que um processo regular admita uma representação auto-regressiva média móvel:

Propriedade Seja X um processo regular. Este processo admite uma representação $ARMA(p, q)$ mínima se e só se

$$\begin{cases} \Delta(i, j) = 0, \forall i \geq p, \forall j \geq q \\ \Delta(p, q-1) \neq 0 \\ \Delta(p-1, q) \neq 0 \end{cases}$$

Além disso,

$$\Delta(i, q-1) \neq 0, \forall i \geq p-1 \quad e \quad \Delta(p-1, q) \neq 0, \forall j \geq q-1$$