



**Universidade do Minho**

Escola de Engenharia

Carla Maria Teniz Queirós

**Previsões da procura numa empresa do  
setor têxtil e do vestuário**

setembro de 2022



**Universidade do Minho**

Escola de Engenharia

Carla Maria Teniz Queirós

**Previsões da procura numa empresa do  
setor têxtil e do vestuário**

Dissertação de Mestrado em Engenharia Industrial –  
Logística e Distribuição

Trabalho efetuado sob a orientação do

**Professor Doutor Manuel Carlos Barbosa Figueiredo**

setembro de 2022

## **DIREITOS DE AUTOR E CONDIÇÕES DE UTILIZAÇÃO DO TRABALHO POR TERCEIROS**

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

### ***Licença concedida aos utilizadores deste trabalho***



**Atribuição**

**CC BY**

<https://creativecommons.org/licenses/by/4.0/>

## **AGRADECIMENTOS**

Finalizado o mestrado em Engenharia Industrial, com especialização em Logística e Distribuição, resta-me agradecer a todos aqueles que contribuíram para a sua concretização.

Em primeiro lugar, ao Professor Doutor Manuel Figueiredo pela orientação, acompanhamento, disponibilidade e flexibilidade concebidas durante todo o desenvolvimento do presente projeto.

Um especial agradecimento à empresa Linhas, S.A. pela oportunidade concebida em desenvolver o projeto desejado através do usufruto dos seus dados, concebendo-me a minha primeira experiência profissional na minha área de especialização. À Engenheira orientadora da empresa, agradeço a orientação facultada ao longo de todo o projeto e conhecimentos partilhados. Agradeço ainda aos demais Engenheiros por toda a disponibilidade oferecida durante o meu acolhimento.

Agradeço o acompanhamento e motivação de todos os meus colegas e amigos que estiveram ao meu lado nesta jornada. Um especial agradecimento: a todos os meus amigos a que me afeiçoei durante a minha licenciatura, à Maria Vieira, ao Cláudio Silva e ao meu colega Luís Araújo.

Finalmente, um agradecimento aos meus pais e avó Pimpa, por todo o apoio durante todo o meu percurso académico.

## **DECLARAÇÃO DE INTEGRIDADE**

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

## RESUMO

A indústria têxtil e do vestuário é caracterizada, maioritariamente, pela volatilidade da sua procura, os curtos ciclos de vida dos produtos, a grande variedade de produtos com dados históricos de vendas limitados e por um processo de transformação até ao produto final composto por diferentes estágios, envolvendo diversas empresas distanciadas geograficamente. Destas características advêm a complexidade e a grande importância associadas às previsões da procura para empresas que se encontram inseridas neste tipo de indústria, estando o seu desempenho organizacional fortemente dependente da precisão destas previsões.

O presente projeto de dissertação foi desenvolvido numa empresa inserida na têxtil-vestuário, Linhas, S.A., cuja atividade principal passa pela fabricação de linhas de costura. A previsão da procura tem vindo a ser baseada em modelos de previsão desadequados, o que se tem traduzindo em ineficiências para a empresa. Neste sentido, definiu-se como principal objetivo a estimação de modelos de previsão da procura mais indicados para os diferentes artigos produzidos pela Linhas, S.A..

Ao longo do projeto foram aplicados diferentes modelos de previsão da procura às dez séries alvo de análise. Pelos resultados obtidos na aplicação destes modelos, não foi possível definir apenas um modelo que melhor se comportasse para todos os casos. Em particular, dependendo da família e subfamília em estudo, obtiveram-se em alguns casos melhores resultados com modelos estatísticos e em outros com modelos de inteligência artificial. Dos modelos estatísticos aplicados, destaca-se o comportamento do modelo de alisamento exponencial *Holt-Winters* aditivo para 50% das séries. Por outro lado, salienta-se o comportamento do modelo de alisamento exponencial simples para 10% das séries, assim como do *Naïve*. Além disso, a aplicação do modelo *Extreme Learning Machines*, modelo de inteligência artificial, obteve um melhor comportamento para 30% das séries em análise. Neste seguimento, conclui-se que em 70% das séries os modelos de previsão estatísticos são os mais bem-comportados, podendo este comportamento ser justificado, maioritariamente, pela consideração apenas das variáveis vendas e tempo para o estudo.

## PALAVRAS-CHAVE

Alisamento Exponencial, *Extreme Learning Machines*, *Holt-Winters* aditivo, Previsões da procura, Têxtil-Vestuário

## **ABSTRACT**

The textile and clothing industry is mainly characterized by the volatility of its demand, the short product life cycles, the wide variety of products with limited historical sales data and by a multiple stage transformation process, involving several geographically distant companies, till the final product is obtained. From these characteristics, come the complexity and the great importance associated with demand forecasts for companies that are inserted in this type of industry, as their organizational performance is strongly dependent on the accuracy of these forecasts.

The present dissertation project was developed in a company inserted in the textile-clothing industry, Linhas, S.A., whose main activity is the manufacture of sewing threads. Demand forecasting has been based on inadequate forecasting models, which has resulted in inefficiencies for the company. In this sense, the main objective was defined to estimate the most suitable demand forecast models for the different articles produced by Linhas, S.A..

Throughout the project, different demand forecast models were applied to the ten series targeted for analysis. Due to the results obtained in the application of these models, it was not possible to define only one model that would behave best for all cases. In particular, depending on the family and subfamily under study, better results were obtained in some cases with statistical models and in others with artificial intelligence models. Of the applied statistical models, the behaviour of the additive Holt-Winters exponential smoothing model for 50% of the series stands out. On the other hand, the behaviour of the simple exponential smoothing model for 10% of the series, as well as sNaïve, is highlighted. In addition, the application of the Extreme Learning Machines model, an artificial intelligence model, obtained a better behaviour for 30% of the series under analysis. In this follow-up, it is concluded that in 70% of the series the statistical forecast models are the best behaved, and this behaviour can be justified, mainly, by the consideration of only the sales and time variables for the study.

## **KEYWORDS**

Demand Forecasting, Exponential Smoothing, Extreme Learning Machines, Holt-Winters Additive, Textile-Clothing

## ÍNDICE

Agradecimentos.....	iii
Resumo.....	v
Abstract.....	vi
Índice.....	vii
Índice de Figuras.....	xi
Índice de Tabelas .....	xiii
Lista de Abreviaturas, Siglas e Acrónimos .....	xvi
1. Introdução .....	1
1.1 Enquadramento .....	1
1.2 Descrição do Problema e Objetivos da Investigação .....	2
1.3 Metodologia de Investigação.....	2
1.4 Estrutura da Dissertação .....	3
2. Apresentação da empresa.....	4
2.1 Linhas, S.A.....	4
2.2 Produtos comercializados.....	4
2.3 Estrutura Organizacional .....	7
2.4 Cadeia de Abastecimento .....	7
3. Revisão de literatura .....	10
3.1 Classificação de artigos pela Análise ABC .....	10
3.2 Indústria Têxtil e do Vestuário.....	10
3.2.1 Caracterização da Indústria Têxtil e do Vestuário e Importância de Previsões .....	11
3.2.2 Métodos de Previsão da Procura tipicamente implementados na Têxtil-Vestuário .....	12
3.3 Métodos de Previsão da Procura.....	14
3.3.1 Métodos de Previsão Estatísticos.....	15
3.3.2 Métodos de Inteligência Artificial .....	27
3.3.3 Métodos Híbridos .....	29
3.4 Medidas de Avaliação de Desempenho das Previsões .....	30
3.4.1 Erro de Previsão .....	30

3.4.2	Erro Quadrático Médio .....	30
3.4.3	Raiz do Erro Quadrático Médio.....	31
3.4.4	Desvio Absoluto Médio.....	31
3.4.5	Erro Absoluto Percentual Médio .....	31
3.4.6	Cr�terio de Informa�o <i>Bayesiano</i> .....	32
3.4.7	Coeficiente de Determina�o e Coeficiente de Determina�o Ajustado.....	32
3.4.8	Estat�stica de <i>Ljung-Box</i> .....	33
4.	Caso de estudo.....	35
4.1	An�lise por Fam�lias de Produtos.....	35
4.1.1	An�lise da Evolu�o das Vendas e das Quotas anuais por Fam�lia de produtos .....	35
4.1.2	An�lises ABC por Fam�lia de produtos.....	39
4.2	An�lise por Subfam�lias de Produtos.....	40
4.3	Metodologia de Previs�o.....	41
5.	Previs�o da Procura atrav�s de M�todos Estat�sticos .....	43
5.1	Fam�lia A.....	43
5.1.1	Subfam�lia AK.....	44
5.1.2	Subfam�lia AI .....	49
5.1.3	Subfam�lia AM .....	51
5.1.4	Subfam�lia AA.....	53
5.1.5	Conclus�es da Fam�lia A.....	54
5.2	Fam�lia Z.....	56
5.2.1	Subfam�lia ZI .....	56
5.2.2	Subfam�lia ZG.....	58
5.2.3	Subfam�lia ZM .....	61
5.2.4	Conjunto das restantes subfam�lias Z.....	63
5.2.5	Conclus�es da Fam�lia Z.....	65
5.3	Fam�lia D .....	66
5.4	Fam�lia S.....	68
6.	Previs�o da Procura atrav�s de M�todos de Intelig�ncia Artificial e Compara�o com M�todos Estat�sticos.....	72

6.1	Família A.....	73
6.1.1	Subfamília AK.....	73
6.1.2	Subfamília AI.....	75
6.1.3	Subfamília AM.....	77
6.1.4	Subfamília AA.....	80
6.2	Família Z.....	82
6.2.1	Subfamília ZI.....	82
6.2.2	Subfamília ZG.....	84
6.2.3	Subfamília ZM.....	86
6.2.4	Conjunto das restantes subfamílias Z.....	88
6.3	Família D.....	90
6.4	Família S.....	93
7.	Conclusões e Trabalho Futuro.....	95
	Referências Bibliográficas.....	97
	APÊNDICES.....	99
	Apêndice A.....	99
	Apêndice B.....	100
	Apêndice C.....	101
	Apêndice D.....	103
	Apêndice E.....	104
	Apêndice F.....	105
	Apêndice G.....	109
	Apêndice H.....	110
	Apêndice I.....	111
	Apêndice J.....	112
	Apêndice K.....	113
	Apêndice L.....	114
	Apêndice M.....	115
	Apêndice N.....	116
	Apêndice O.....	117
	Apêndice P.....	118

Apêndice Q.....	119
Apêndice R.....	120
Apêndice S.....	121
Apêndice T.....	122
ANEXO.....	125

## ÍNDICE DE FIGURAS

<b>Figura 1</b>	Estrutura organizacional da sede da Linhas, S.A.....	7
<b>Figura 2</b>	Cadeia de abastecimento da Linhas, S.A. e respetivos fluxos físicos.....	8
<b>Figura 3</b>	Etapas de previsão .....	14
<b>Figura 4</b>	Etapas para a aplicação do modelo ARIMA.....	23
<b>Figura 5</b>	Evolução das vendas anuais (em kg) nos últimos 5 anos para as oito famílias de produtos produzidas pela Linhas, S.A. ....	36
<b>Figura 6</b>	Evolução das quotas anuais nos últimos 5 anos para as oito famílias de produtos produzidas pela Linhas, S.A. ....	36
<b>Figura 7</b>	Evolução das vendas anuais (em kg) nos últimos 5 anos para as famílias de produtos produzidas pela Linhas, S.A. (sem a família A) .....	37
<b>Figura 8</b>	Evolução das quotas anuais nos últimos 5 anos para as famílias de produtos produzidas pela Linhas, S.A. (sem a família A) .....	38
<b>Figura 9</b>	Análise ABC, por família de produtos, em 2020 .....	39
<b>Figura 10</b>	Análise ABC, por família de produtos, em 2021 .....	40
<b>Figura 11</b>	Vendas quadrimestrais de AK (em kg) .....	45
<b>Figura 12</b>	Vendas quadrimestrais de AI (em kg).....	50
<b>Figura 13</b>	Vendas quadrimestrais de AM (em kg).....	51
<b>Figura 14</b>	Vendas quadrimestrais de AA (em kg).....	53
<b>Figura 15</b>	Vendas quadrimestrais de ZI (em kg) .....	57
<b>Figura 16</b>	Vendas quadrimestrais de ZG (em kg) .....	59
<b>Figura 17</b>	Vendas quadrimestrais de ZM (em kg) .....	61
<b>Figura 18</b>	Vendas quadrimestrais do conjunto das restantes subfamílias de Z (em kg) .....	63
<b>Figura 19</b>	Quotas anuais de 2021 das restantes subfamílias Z a aplicar em previsões 2022 .....	65
<b>Figura 20</b>	Vendas quadrimestrais do conjunto da família D (em kg) .....	66
<b>Figura 21</b>	Quotas anuais de 2021 das subfamílias D a aplicar em previsões 2022.....	68
<b>Figura 22</b>	Vendas quadrimestrais do conjunto da família S (em kg).....	69
<b>Figura 23</b>	Quotas anuais de 2021 das subfamílias S a aplicar em previsões 2022 .....	71
<b>Figura 24</b>	Vendas reais quadrimestrais e respetiva procura prevista para sNaïve, HW aditivo e ELM para a série AK.....	74

<b>Figura 25</b> Vendas reais quadrimestrais e respetiva procura prevista para sNaïve, HW aditivo e ELM para a série AI .....	76
<b>Figura 26</b> Vendas reais quadrimestrais e respetiva procura prevista para sNaïve, HW aditivo e ELM para a série AM .....	79
<b>Figura 27</b> Vendas reais quadrimestrais e respetiva procura prevista para sNaïve, HW aditivo e ELM para a série AA .....	81
<b>Figura 28</b> Vendas reais quadrimestrais e respetiva procura prevista para sNaïve, HW aditivo e ELM para a série ZI .....	83
<b>Figura 29</b> Vendas reais quadrimestrais e respetiva procura prevista para sNaïve, HW aditivo e ELM para a série ZG .....	85
<b>Figura 30</b> Vendas reais quadrimestrais e respetiva procura prevista para sNaïve, HW aditivo e ELM para a série ZM .....	87
<b>Figura 31</b> Vendas reais quadrimestrais e respetiva procura prevista para sNaïve e ELM para a série do conjunto das restantes subfamílias de Z .....	90
<b>Figura 32</b> Vendas reais quadrimestrais e respetiva procura prevista para sNaïve, AES e ELM para a série da família D .....	92
<b>Figura 33</b> Vendas reais quadrimestrais e respetiva procura prevista para sNaïve, HW aditivo e ELM para a série da família S .....	94

## ÍNDICE DE TABELAS

<b>Tabela 1</b>	Famílias de produtos acabados comercializados pela Linhas, S.A. ....	5
<b>Tabela 2</b>	Aplicação da metodologia de previsão .....	42
<b>Tabela 3</b>	Componentes das séries temporais das subfamílias A obtidas através do método de decomposição multiplicativo .....	44
<b>Tabela 4</b>	Medidas de avaliação de desempenho dos métodos de previsão de AE e sNaïve aplicados à série AK .....	46
<b>Tabela 5</b>	Medidas de avaliação de desempenho dentro da amostra do modelo ARIMA (0,1,0) com e sem transformações aplicados à série AK .....	47
<b>Tabela 6</b>	Medidas de avaliação de desempenho dos modelos HW aditivo e ARIMA (0,1,0) aplicados à série AK .....	49
<b>Tabela 7</b>	Medidas de avaliação de desempenho dos métodos de previsão aplicados à série AI .....	50
<b>Tabela 8</b>	Medidas de avaliação de desempenho dos métodos de previsão aplicados à série AM .....	52
<b>Tabela 9</b>	Medidas de avaliação de desempenho dos métodos de previsão aplicados à série AA .....	54
<b>Tabela 10</b>	Constantes de alisamento otimizadas segundo série temporal .....	55
<b>Tabela 11</b>	Definição de constantes de alisamento para as séries de família A.....	55
<b>Tabela 12</b>	Componentes das séries temporais das subfamílias Z obtidas através do método de decomposição multiplicativo .....	56
<b>Tabela 13</b>	Medidas de avaliação de desempenho dos métodos de previsão aplicados à série ZI .....	57
<b>Tabela 14</b>	Medidas de avaliação de desempenho dos métodos de previsão aplicados à série ZG....	59
<b>Tabela 15</b>	Medidas de avaliação de desempenho dos métodos de previsão aplicados à série ZM ...	62
<b>Tabela 16</b>	Medidas de avaliação de desempenho dos métodos de previsão aplicados à série do conjunto das restantes subfamílias de Z.....	64
<b>Tabela 17</b>	Métodos de previsão estatísticos selecionados para a família Z.....	65
<b>Tabela 18</b>	Definição de constantes de alisamento para as séries das subfamílias ZI, ZG e ZM .....	66
<b>Tabela 19</b>	Medidas de avaliação de desempenho dos métodos de previsão aplicados à série da família D.....	67
<b>Tabela 20</b>	Medidas de avaliação de desempenho dos métodos de previsão aplicados à série da família S.....	69
<b>Tabela 21</b>	Definição de constantes de alisamento para a série da família S.....	71

<b>Tabela 22</b> Medidas de avaliação de desempenho dos métodos de previsão sNaïve, HW aditivo e ELM aplicados à série AK .....	73
<b>Tabela 23</b> Taxas de variação associadas à implementação de modelos de HW aditivo ou ELM em detrimento do modelo sNaïve para a série AK .....	75
<b>Tabela 24</b> Medidas de avaliação de desempenho dos métodos de previsão sNaïve, HW aditivo e ELM aplicados à série AI .....	75
<b>Tabela 25</b> Taxas de variação associadas à implementação de modelos de HW aditivo ou ELM em detrimento do modelo sNaïve para a série AI .....	77
<b>Tabela 26</b> Medidas de avaliação de desempenho dos métodos de previsão sNaïve, HW aditivo e ELM aplicados à série AM.....	78
<b>Tabela 27</b> Taxas de variação associadas à implementação de modelos de HW aditivo ou ELM em detrimento do modelo sNaïve para a série AM .....	79
<b>Tabela 28</b> Medidas de avaliação de desempenho dos métodos de previsão sNaïve, HW aditivo e ELM aplicados à série AA .....	80
<b>Tabela 29</b> Taxas de variação associadas à implementação de modelos de HW aditivo ou ELM em detrimento do modelo sNaïve para a série AA .....	81
<b>Tabela 30</b> Medidas de avaliação de desempenho dos métodos de previsão sNaïve, HW aditivo e ELM aplicados à série ZI.....	82
<b>Tabela 31</b> Taxas de variação associadas à implementação de modelos de HW aditivo ou ELM em detrimento do modelo sNaïve para a série ZI.....	83
<b>Tabela 32</b> Medidas de avaliação de desempenho dos métodos de previsão sNaïve, HW aditivo e ELM aplicados à série ZG .....	84
<b>Tabela 33</b> Taxas de variação associadas à implementação de modelos de HW aditivo ou ELM em detrimento do modelo sNaïve para a série ZG .....	86
<b>Tabela 34</b> Medidas de avaliação de desempenho dos métodos de previsão sNaïve, HW aditivo e ELM aplicados à série ZM.....	87
<b>Tabela 35</b> Taxas de variação associadas à implementação de modelos de HW aditivo ou ELM em detrimento do modelo sNaïve para a série ZM.....	88
<b>Tabela 36</b> Medidas de avaliação de desempenho dos métodos de previsão sNaïve e ELM aplicados à série do conjunto das restantes subfamílias Z .....	89
<b>Tabela 37</b> Medidas de avaliação de desempenho dos métodos de previsão sNaïve, AES e ELM aplicados à série da família D .....	91

<b>Tabela 38</b> Taxas de variação associadas à implementação de modelos de AES ou ELM em detrimento do modelo sNaïve para a série da família D.....	92
<b>Tabela 39</b> Medidas de avaliação de desempenho dos métodos de previsão sNaïve, HW aditivo e ELM aplicados à série da família S.....	93
<b>Tabela 40</b> Taxas de variação associadas à implementação de modelos de HW aditivo ou ELM em detrimento do modelo sNaïve para a série da família S .....	94

## LISTA DE ABREVIATURAS, SIGLAS E ACRÓNIMOS

*ACF – Autocorrelation Function*

AES – Alisamento Exponencial Simples

*ANN – Artificial Neural Networks*

*ARIMA – Autoregressive Integrated Moving Average*

*BIC – Bayesian Information Criterion*

CAE – Classificação Portuguesa das Atividades Económicas

*EELM – Extended Extreme Learning Machines*

*ELM-FF – Extended Extreme Learning Machine Fast Forecasting*

*ELM – Extreme Learning Machines*

*ENN – Evolutionary Neural Networks*

*HW – Holt-Winters*

IA – Inteligência Artificial

*IFSFM – Intelligent Fast Sales Forecasting Model*

*MAD – Mean Absolute Deviation*

*MAPE – Mean Absolute Percentage Error*

MMP – Média Móvel Ponderada

MMS – Média Móvel Simples

*MSE – Mean Square Error*

*PACF – Partial Autocorrelation Function*

*RMSE – Root Mean Square Error*

*SARIMA – Seasonal Autoregressive Integrated Moving Average*

*SLFN – Single Hidden Layer Feedforward Neural Network*

sNaïve – Naïve Sazonal

# 1. INTRODUÇÃO

O corrente documento descreve o projeto de dissertação desenvolvido na empresa Linhas, S.A., no âmbito do Mestrado em Engenharia Industrial – Logística e Distribuição. O presente capítulo destina-se à apresentação do enquadramento do projeto de dissertação, descrição do problema e objetivos de investigação, apresentação da metodologia de investigação e descrição da estrutura da dissertação.

## 1.1 Enquadramento

A previsão da procura tem um papel de grande importância na gestão de operações, sendo um dos maiores desafios para retalhistas, grossistas e fabricantes. Previsões imprecisas podem levar a diversas consequências, nomeadamente, a quebras ou excessos de *stock*, baixo nível de serviço aos clientes, utilização ineficiente de recursos, pedidos urgentes e propagação do efeito chicote ao longo da cadeia de abastecimento. Consequentemente, a eficácia da otimização da cadeia de abastecimento depende da previsão da procura do produto acabado.

Ferramentas de previsão mais precisas podem auxiliar as empresas a determinar, de forma mais rigorosa, a quantidade de cada produto que deve ser pedido aos fornecedores, contribuindo positivamente para o aprimoramento do processo de compra. Este processo acabará por se refletir num aumento da lucratividade. (Loureiro et al., 2018, p. 89)

De forma a aumentar a vantagem competitiva num ambiente como o da indústria têxtil e do vestuário, onde a procura é volátil, o ciclo de vida dos produtos é curto e há uma grande variedade de produtos com dados históricos de vendas limitados, as empresas inseridas na indústria têxtil e do vestuário devem possuir sistemas de previsão da procura adequados ao seu ambiente incerto e em constante flutuação. Posto isto, o processo de previsão de procura nesta indústria é conhecido como sendo complexo, tendo vindo a ser um tópico muito estudado nas últimas décadas e concluindo-se que não existe um método analítico comumente usado para a previsão da procura na indústria têxtil e do vestuário. Perante isto, destaca-se a importância do estudo de métodos de previsões da procura adaptados às diferentes realidades empresariais.

## **1.2 Descrição do Problema e Objetivos da Investigação**

O processo de compras da Linhas, S.A. encontra-se pouco automatizado, não sendo considerado para o mesmo a previsão da procura baseada em modelos de previsão adequados, o que se tem vindo a traduzir em ineficiências causadas por *stocks* desadequados face à procura, isto é, com excessos ou ruturas. Nestas condições, que se vêm agravando dado o cenário de expansão da empresa, torna-se fulcral melhorar o processo de programação e processamento de compras através da definição de modelos de previsão da procura de forma a possibilitar, posteriormente, o controlo de *stock*.

Definiu-se como objetivo fulcral do presente projeto de dissertação, o estudo e definição de modelos de previsão da procura mais indicados para os diferentes artigos produzidos pela empresa. A definição destes modelos tem como objetivo de curto prazo a diminuição dos erros de previsão associados aos modelos de previsão da procura atualmente adotados pela empresa e, como objetivos de longo prazo para a empresa, o desenvolvimento de modelos de gestão de *stock*, que permitam manter um nível de serviço ao cliente de 95% e diminuir potenciais riscos de rutura e excessos de *stock*.

## **1.3 Metodologia de Investigação**

O projeto de dissertação decorreu num horizonte temporal longitudinal, onde se estudou o mesmo fenómeno em diferentes períodos de tempo. No decorrer do projeto foi adotada como estratégia de investigação a investigação-ação, de forma a permitir um envolvimento no projeto por parte do investigador e dos trabalhadores e a possibilitar um ambiente colaborativo entre ambos. A investigação-ação é desenvolvida em ciclos repetitivos de quatro fases: diagnóstico, planeamento, ação e avaliação (Saunders et. al, 2009), tendo o presente projeto de dissertação seguido essas etapas na implementação de métodos de previsão.

Na primeira etapa, de diagnóstico, efetuou-se uma descrição, análise e diagnóstico da empresa. Em concreto, procedeu-se à descrição dos produtos comercializados pela empresa, da sua estrutura organizacional e da sua cadeia de abastecimento e, posteriormente, realizou-se uma análise e diagnóstico da empresa baseados nas vendas passadas dos produtos produzidos pela mesma.

Numa segunda fase, fase de planeamento, foi desenvolvido um planeamento de soluções possíveis para o problema de previsões e das ações a desenvolver. Assim, nesta etapa, procedeu-se a uma revisão de literatura acerca da indústria em que a empresa se encontra inserida e de diferentes modelos de previsão de procura. Esta revisão tinha como objetivo a caracterização da indústria e identificação dos modelos de previsão que melhor se poderiam aplicar ao caso em estudo.

Na etapa três, etapa de ação, aplicaram-se os modelos de previsão da procura considerados mais indicados na etapa anterior para os artigos considerados relevantes para o caso em estudo. E, por fim, na etapa quatro, de avaliação, procedeu-se à avaliação dos resultados obtidos pelos modelos implementados com recurso a medidas de avaliação de desempenho e comparação das mesmas com os modelos anteriormente adotados.

Para o desenvolvimento do projeto contou-se, principalmente, com fontes primárias (como relatórios e dissertações) e secundárias (como livros e revistas científicas). Estas dizem respeito a contribuições de estudos realizados entre os anos 1998 e 2021, obtidas, principalmente, através das bases de dados referenciais *B-on* e *SCOPUS*.

#### **1.4 Estrutura da Dissertação**

O presente documento está dividido em sete capítulos e ainda três secções compostas por Referências Bibliográficas, Apêndices e Anexo. Inicia-se com a Introdução, na qual é realizado um enquadramento do projeto de dissertação, elaborada uma descrição do problema e objetivos da investigação, apresentada a metodologia de investigação e descrita a estrutura da dissertação.

Seguidamente, é realizada uma apresentação da empresa na qual o projeto de dissertação se debruçou, em particular, são apresentados os produtos comercializados pela mesma, a sua estrutura organizacional e cadeia de abastecimento.

Em terceiro lugar, procede-se a uma revisão de literatura, iniciando-se pela apresentação de um método associado à classificação de artigos. Seguidamente, a literatura revista dá a conhecer a indústria na qual a empresa se insere, bem como métodos de previsão da procura e medidas de avaliação de desempenho associadas à mesma.

No quarto capítulo, é apresentado o caso de estudo em investigação, em particular, é elaborada uma análise dos produtos produzidos pela empresa por famílias e subfamílias e, posteriormente, definida uma metodologia de previsão a aplicar aquando das previsões da procura.

Posteriormente, o quinto capítulo é destinado à previsão da procura através de métodos estatísticos e o sexto capítulo à previsão da procura através de inteligência artificial, bem como a comparação dos mesmos com os resultados obtidos pelos métodos estatísticos.

Por último, elaboram-se as conclusões e é proposto trabalho futuro a considerar pela empresa.

É de realçar que, de forma a garantir a confidencialidade da empresa, todos os valores apresentados ao longo da dissertação estão corrigidos por um fator. Além disso, a denominação da empresa, famílias e subfamílias dos produtos foram alteradas, assim como outros dados que permitiam a sua identificação.

## **2. APRESENTAÇÃO DA EMPRESA**

O presente capítulo destina-se à apresentação da empresa na qual se desenvolveu o projeto de investigação que suportou o presente projeto de dissertação, Linhas, S.A.. Neste seguimento, primeiramente procede-se a uma descrição geral da empresa, seguindo-se uma apresentação dos produtos comercializados pela mesma, passando-se à exposição da sua estrutura organizacional e, por fim, a uma descrição da sua cadeia de abastecimento. Ao longo desta apresentação foram omitidos ou codificados dados confidenciais que levassem à identificação da empresa.

### **2.1 Linhas, S.A.**

A Linhas, S.A., fundada em Portugal, nos anos 80, encontra-se inserida na indústria transformadora, pela fabricação de têxteis. Esta empresa possui a Classificação Portuguesa das Atividades Económicas (CAE) número 13104, tendo como atividade principal a fabricação de linhas de costura de diversas naturezas e composições, metragens, espessuras e tonalidades (INE, 2007).

Desde o seu nascimento, a empresa está sediada na região do Minho, sendo esta unidade destinada à produção e venda de produtos acabados, realizando ainda a sua armazenagem. Na década de 2010, a Linhas, S.A. passa a possuir uma segunda unidade também na região do Minho, destinada apenas à venda e armazenagem de produtos acabados e de matérias-primas.

### **2.2 Produtos comercializados**

A Linhas, S.A., tem como produtos acabados linhas de costura, que podem ser agrupados em dez diferentes famílias, distinguindo-se, entre si, pela natureza e composição do fio. Estas famílias de produtos podem ser denominadas por dez letras do abecedário, sendo elas, A, B, G, D, H, K, M, O, S e Z. Cada família pode ainda ser fracionada em diferentes subfamílias segundo a espessura do fio. Além da natureza, composição do fio e espessura, os produtos acabados distinguem-se pelas suas metragens e tonalidades (preto, branco e cru, sendo o cru ainda tingido em diversas cores de cartaz), permitindo, assim, à empresa a comercialização de milhares de artigos.

A **Tabela 1** resume os tipos de fios e algumas especificidades e principais usos das diferentes famílias de produtos acabados comercializados pela empresa Linhas, S.A..

**Tabela 1**

Famílias de produtos acabados comercializados pela Linhas, S.A.

<b>Família</b>	<b>Tipo de Fio</b>	<b>Especificidades e Principais usos</b>
<b>A</b>	100% Poliéster	<p>Fio 100% poliéster fiado com grampo, de alta tenacidade, lubrificado, de construção uniforme e resistente à luz solar, à degradação química e à maioria dos ácidos minerais.</p> <p>É utilizado, principalmente, na produção de camisas, blusas, roupa interior, roupa de criança, calças, camisolas, roupa de trabalho e vestidos.</p>
<b>B</b>	100% Monofilamento de nylon	<p>Fio 100% monofilamento de nylon (nylon 6.6), com grande resistência e diâmetro externo estreito.</p> <p>Empregue, principalmente, em roupas femininas e masculinas, pela costura de, por exemplo, punhos, cós, bolsos e bainhas.</p>
<b>D</b>	100% Algodão	<p>Fio 100% algodão mercerizado, produzido a partir de algodão de fibra longa de alta qualidade e resistência. O lubrificante de costura é de baixo atrito, oferecendo um bom desempenho de costura.</p> <p>É usado, principalmente, para costurar peças de roupa de 100% algodão, não tratadas que devem ser pós-tingidas.</p>
<b>G</b>	100% Liocel	<p>Fio 100% liocel torcido, com boa resistência da costura.</p> <p>É uma boa escolha para, por exemplo, costurar tecidos elásticos, roupas de malha fina e roupas de liocel com tingimentos excessivos, sendo utilizado, principalmente, em camisas, tops e calças casuais, roupa para criança e vestidos.</p>
<b>H</b>	100% Nylon de alta tenacidade	<p>Fio 100% nylon de alta tenacidade, lubrificado, feito de filamentos contínuos pré-estabilizados de nylon 6.6. Fio resistente à abrasão, com características de elasticidade, um acabamento macio e uma superior lubrificação de baixo atrito.</p> <p>É 30% mais forte que o fio de nylon comum.</p> <p>Ideal para produtos de couro, calçado, artigos de desporto, cintos de segurança e <i>airbags</i>, filtros industriais, interiores de automóveis, colchões e têxteis técnicos e geotêxtis.</p>
<b>K</b>	Poliéster- algodão	<p>Fio corepun que combina filamentos contínuos de poliéster de alta tenacidade revistos por algodão. Por um lado, os filamentos de poliéster permitem o alcance de uma excelente resistência e durabilidade e, por outro lado, o</p>

		<p>algodão permite excelentes desempenhos de costura e resistência ao calor. É ainda dotado de um lubrificante especial.</p> <p>Utilizado, especialmente, em camisas, fatos, calças, lingerie premium, produtos de couro, toalhas, artigos de decoração e móveis estofados.</p>
<b>M</b>	100% Poliéster texturizado a granel	<p>Fio 100% poliéster texturizado de filamento torcido contínuo, de alta elasticidade, extensibilidade, flexibilidade, resistência e excelente cobertura de costura. É resistente à maioria dos ácidos minerais.</p> <p>Usado, principalmente, na produção de roupa interior, de banho, de desporto, de bebés, malhas e têxtis lar.</p>
<b>O</b>	100% Poliéster texturizado	<p>Fio 100% poliéster texturizado de alta elasticidade, caracterizado pela sua alta protuberância, pelo seu lubrificante de costura de baixo atrito, alto valor de elasticidade, baixo encolhimento e elevado grau de solidez da cor.</p> <p>Utilizado, especialmente, para roupa interior, de desporto e de banho.</p>
<b>S</b>	100% Poliéster (Poliéster-poliéster)	<p>Fio 100% poliéster corepun, que combina filamentos contínuos de poliéster revestidos a poliéster. Fio lubrificado e dotado de versatilidade, produtividade e excelente resistência à abrasão e tração. Resistente à maioria dos ácidos minerais.</p> <p>Fio adequado para roupas delicadas, casacos e calças de ganga, blusas, camisas, lingerie, roupas de banho, calças, artigos de couro, uniformes e roupas de trabalho.</p>
<b>Z</b>	100% Poliéster de alta tenacidade	<p>Fio 100% poliéster lubrificado, constituído por filamentos contínuos de poliéster de alta tenacidade pré-estabilizados. De excelente costura e desempenho de costura, características de elasticidade controladas, com ótimas características de estiramento, excelente resistência à rutura e à abrasão, acabamento macio com lubrificação superior de baixo atrito.</p> <p>Ideal para artigos de couro, desportivos, filtros industriais, cintos de segurança e <i>airbags</i>, malas, têxteis técnicos e geotêxtis, interiores de automóveis, colchões e móveis estofados.</p>

*Nota.* Elaboração própria baseada em informação retirada do catálogo Linhas, S.A. (2022)

Das dez famílias de produtos comercializados pela empresa, duas delas, B e O, ao contrário das restantes, não passam pelo processo produtivo internamente, destinando-se apenas a revenda.

Os produtos acabados são produzidos, na sua generalidade, para encomenda. No entanto, é de destacar que algumas famílias de artigos são produzidas também para *stock* nas cores primárias, isto é, branco, cru e preto, nas metragens consideradas padrão. Além disso, para os artigos da família A, há uma produção para *stock* para algumas cores de cartaz, além das cores primárias, dada a sua procura.

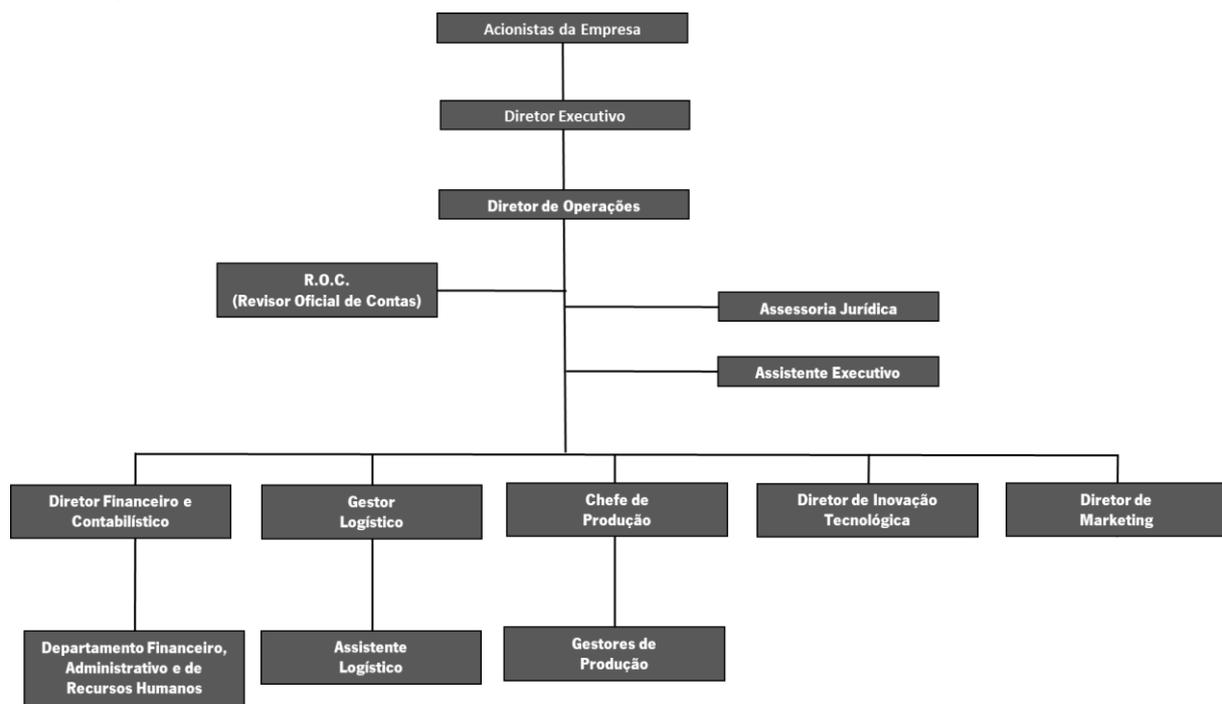
## 2.3 Estrutura Organizacional

Na sede da Linhas, S.A. a equipa é constituída pela direção geral e por cinco departamentos: contabilístico e financeiro, logístico, da produção, de inovação tecnológica, e de *marketing*. A estrutura organizacional desta unidade é esquematizada no organograma representado na **Figura 1**.

Já a segunda unidade é composta apenas por um departamento comercial com uma equipa de vendas que mantém o contacto com os clientes.

**Figura 1**

*Estrutura organizacional da sede da Linhas, S.A.*



*Nota.* Elaboração própria

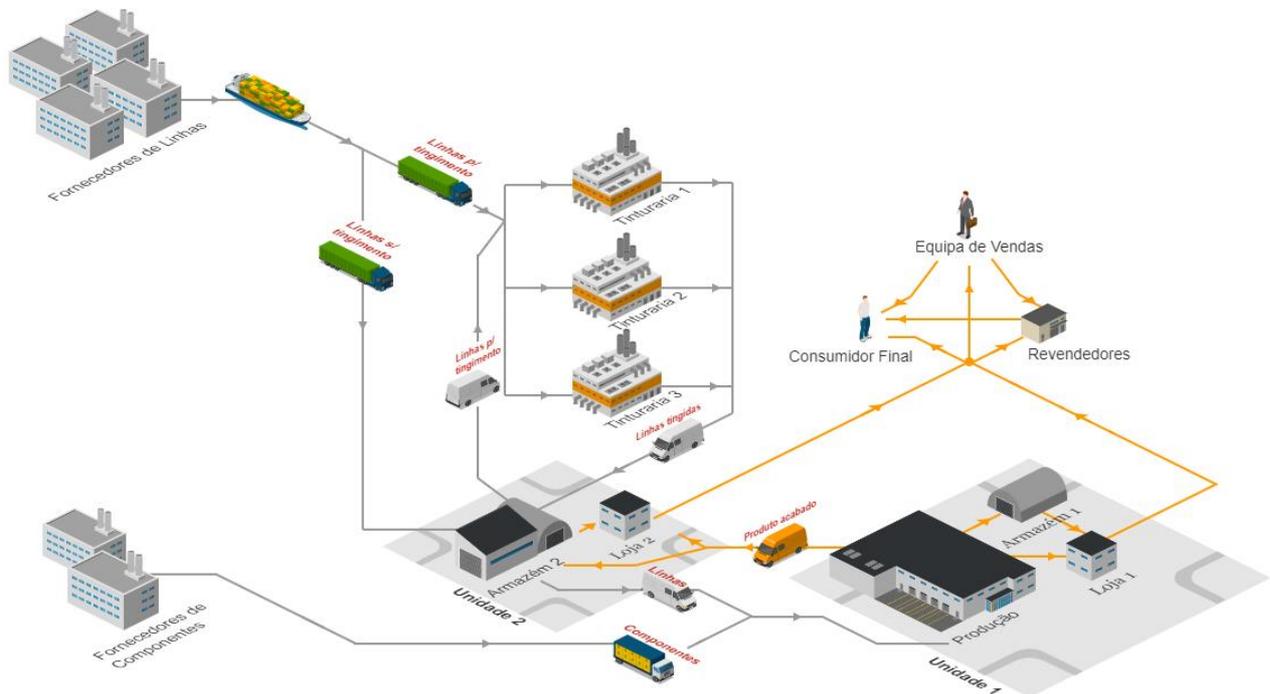
## 2.4 Cadeia de Abastecimento

A Linhas, S.A. possui mais de 50 distribuidores oficiais e trabalha diretamente com muitas fábricas que produzem vestuário, *lingerie*, calçado, roupa para cama, têxtis-lar, entre outros.

A **Figura 2** representa a cadeia de abastecimento da Linhas, S.A. e os fluxos físicos ao longo da mesma.

**Figura 2**

*Cadeia de abastecimento da Linhas, S.A. e respetivos fluxos físicos*



*Nota.* As setas a cor de laranja destacam os fluxos físicos relativos aos produtos acabados. Elaboração própria.

Os elementos mais a montante da cadeia de abastecimento dizem respeito aos fornecedores de linhas e aos fornecedores de componentes (cones, caixas de embalagem, etiquetas, sacos e outros).

A Linhas, S.A. trabalha com diversos fornecedores segundo os diversos fios adquiridos, que diferem pelas tonalidades (branco, cru ou preto) e pelas espessuras. Estes fornecedores são, na sua maioria, asiáticos, sendo o transporte das matérias-primas efetuado por via marítima até aos portos em Portugal (de Leixões ou de Sines). Após chegarem aos portos, dependendo das necessidades, as linhas podem seguir diferentes destinos via terrestre, através de camiões. Por um lado, as linhas podem seguir para as três tinturarias parceiras subcontratadas pela Linhas, S.A. para serem tingidas; por outro lado, podem ir diretamente para a segunda unidade para serem armazenadas, podendo, posteriormente, ser ou não, tingidas. Independentemente desse seu destino, isto é, independentemente de a linha sofrer ou não tingimento e do local onde ocorre esse tingimento, as linhas seguem para a sede da empresa (Unidade 1), para sofrer transformação pelo processo produtivo.

De forma similar, os componentes, que advêm de fornecedores ibéricos, seguem também para a sede da empresa (Unidade 1), via rodoviária, para serem incorporados no processo produtivo.

Após o processo produtivo, segundo as necessidades da procura e as políticas de gestão de *stock*, o produto acabado pode seguir diferentes destinos. Pode, por um lado, ficar em *stock* nos armazéns das

unidades (Unidade 1 e 2). Por outro lado, pode ser vendido nas respetivas lojas abertas ao público (Loja 1 e 2), diretamente ao consumidor final (menos comum) ou a revendedores que depois vendem ao consumidor final (mais comum). Além disso, a Linhas, S.A. detém uma equipa de vendas que faz a entrega dos produtos finais a revendedores e aos consumidores finais.

### **3. REVISÃO DE LITERATURA**

O atual capítulo destina-se à revisão de literatura na qual o presente estudo se baseou, começando por abordar a análise ABC, método tipicamente utilizado para a classificação de artigos. Posteriormente, é realizada uma caracterização da indústria têxtil e do vestuário e conhecida a importância da aplicação de métodos de previsão às empresas que nela operam, bem como os métodos de previsão tipicamente empregues nesta indústria. Por fim, o restante capítulo é reservado à revisão de literatura acerca dos métodos de previsão tipicamente utilizados na têxtil-vestuário, bem como à compreensão de diversas medidas de avaliação do desempenho de previsões.

#### **3.1 Classificação de artigos pela Análise ABC**

A análise ABC, também denominada por método de Pareto ou Regra 80/20, permite classificar artigos segundo o seu grau de importância, possibilitando, assim, um foco nos mais importantes. Neste seguimento, de forma a proceder a uma classificação de artigos, estes são divididos em três classes: o grupo “A” onde os artigos são classificados como “muito importantes”, o grupo “B” classificados como “moderadamente importantes” e o grupo “C” como “menos importantes”.

Para a classificação ABC podem ser utilizados diferentes critérios, desde o mais clássico, o volume de vendas (obtido pelo produto entre o valor monetário unitário do item e sua taxa de uso), aos menos clássicos, como a procura, o *lead-time*, a taxa de consumo ou a criticidade dos itens (Bhattacharya et al., 2001 *as cited in* Bhattacharya et al., 2007).

A divisão entre os três diferentes grupos nem sempre é precisa, isto é, as percentagens não são absolutas, sendo apenas usadas como diretrizes para determinar a classificação ABC de um item, podendo os seus valores diferir segundo o caso em estudo. Contudo, pode definir-se o grupo “A” com 80-60% do valor total das vendas, o grupo “B” com 35-15% e o grupo “C” com 15-5%, correspondendo, geralmente, o grupo “A” a um total de 10-20% dos artigos em análise, o “B” cerca de 30% e o “C” de 50-60%. Perante isto, artigos pertencentes ao grupo “A” devem ser priorizados de forma a ser possível usufruírem de melhores previsões de vendas, seguindo-se como prioridade os do grupo “B” e, por fim, do grupo “C”. (Heizer et al., 2017; Reid & Sanders, 2012).

#### **3.2 Indústria Têxtil e do Vestuário**

A Linhas, S.A., encontra-se inserida na indústria transformadora pela fabricação de têxtis, em particular, a produção de linhas de costura. Segundo Thomassey (2010), a indústria associada ao têxtil pode ser

dividida, geralmente, em três mercados, são eles, têxteis-vestuário, têxteis-lar e têxteis-técnicos. Tendo por base os principais usos das famílias de produtos comercializados pela empresa apresentados no capítulo anterior, repara-se que, grande parte as linhas produzidas pela empresa têm como principal finalidade o vestuário. Desta forma, torna-se essencial aprofundar o conhecimento no mercado têxtil-vestuário, tendo sido considerado para o estudo o sinónimo entre têxtil de vestuário e de moda, dado que, grande parte da literatura revista não faz a distinção entre os termos, além de que, segundo Liu et al. (2013), o setor moda incluiu vestuário, calçado e produtos de moda e beleza.

### 3.2.1 Caracterização da Indústria Têxtil e do Vestuário e Importância de Previsões

Diferentes literaturas referem-se à indústria têxtil e do vestuário como uma indústria caracterizada, predominantemente, pela sua procura volátil e sazonal, pelos curtos ciclos de vida dos produtos, pela grande variedade de produtos e pelos dados históricos de vendas limitados. Estas características advêm de muitos fatores, como da alta variação nos gostos dos consumidores, nas tendências da moda e nos comportamentos do consumo, bem como de fatores aleatórios, como feriados, promoções sobre as vendas, situações económicas e climáticas, entre outros (Liu et al., 2013; Loureiro et al., 2018; Ren et al., 2020; Thomassey, 2010).

Por outro lado, o processo de transformação até ao produto final é composto por diferentes estágios, passando por diferentes empresas, desde a fição à distribuição, estando, normalmente, estas empresas distanciadas geograficamente, com as fábricas, geralmente, situadas em países asiáticos. Este processo longo e complexo faz com que a cadeia de abastecimento seja mais sensível ao efeito chicote<sup>1</sup> e tenha um *lead time* associado maior (Beheshti-Kashi et al., 2015; Thomassey, 2010).

Como resultado das características predominantes neste tipo de indústria, as empresas que lhe pertencem têm de realizar planos de produção e tomar decisões a nível de *stock* com base em previsões de vendas realizadas quase em tempo real e com uma quantidade muito limitada de dados (Beheshti-Kashi et al., 2015; Choi et al., 2014).

Estes fatores tornam as previsões da procura com precisão na têxtil-vestuário uma tarefa complexa, contudo, de grande relevância, dado o sucesso e desempenho organizacional estarem fortemente dependentes de boas previsões da procura (Loureiro et al., 2018; Ren et al., 2020).

---

<sup>1</sup> O efeito chicote, em inglês *bullwhip effect*, diz respeito ao efeito de distorção da procura amplificado ao longo da cadeia de abastecimento, do consumidor final até ao fornecedor, que leva a disparidades entre a procura real e a prevista (Reiff, 2022).

A previsão precisa de vendas pode trazer benefícios uma vez que possibilitará um controlo dos níveis de *stock*, o que permitirá o aumento dos níveis de satisfação e fidelização dos clientes (Caglayan et al., 2020; Loureiro et al., 2018). Além disso, pode ser vista como uma ferramenta utilizada para reduzir o efeito chicote dentro de uma cadeia de abastecimento (Liu et al., 2013). Em contrapartida, previsões de vendas imprecisas podem levar a ruturas de *stock* (que se podem traduzir em insatisfação de clientes e perdas de receitas para os concorrentes), *stock* em excesso, obsolescência, pedidos urgentes, ou à utilização ineficiente de recursos, resultando em perdas para a empresa (Beheshti-Kashi et al., 2015; Kharfan et al., 2021; Nenni et al., 2013).

### 3.2.2 Métodos de Previsão da Procura tipicamente implementados na Têxtil-Vestuário

Existem diferentes *softwares* comerciais que permitem a aplicação de métodos de previsão, contudo, estes, raramente, são implementados nas empresas pertencentes à indústria têxtil e do vestuário, dado ser necessário um investimento muito grande na sua aquisição, a sua implementação ser muito exigente e as empresas pretenderem ter controlo total sobre as suas previsões. Neste seguimento, as empresas da têxtil-vestuário, usualmente, implementam os seus próprios sistemas de previsão. Para isso, geralmente, utilizam as vendas do ano transato como base e, após isso, as previsões são ajustadas segundo variáveis explicativas consideradas relevantes consoante a experiência de um trabalhador da empresa especializado na área. Contudo, este sistema apresenta diversas desvantagens, como o facto de serem consideradas para a previsão um número reduzido de variáveis e as previsões estarem sujeitas ao julgamento do trabalhador. Deste modo, de forma a aumentar a precisão das previsões, devem ser implementados métodos de previsão mais sofisticados (Thomassey, 2014).

Diferentes métodos de previsão têm sido privilegiados no estudo das previsões da procura na indústria têxtil e do vestuário, desde os métodos tradicionais estatísticos, aos métodos de inteligência artificial (IA) e híbridos. Neste sentido, dever-se-á conhecer os diferentes modelos de previsão tipicamente utilizados na indústria têxtil e do vestuário, bem como proceder à sua comparação.

Para comparar os modelos de previsão entre si, dois fatores devem ser considerados: Por um lado, os modelos utilizados devem atingir os menores erros de previsão possíveis. Por outro lado, o esforço de implementação dos métodos é relevante no que diz respeito à sua aplicabilidade na prática. (Moroff et al., 2021, p. 47)

Tradicionalmente, na têxtil-vestuário, têm vindo a ser empregues métodos estatísticos, nomeadamente, modelos causais e modelos não causais. Dos modelos causais tipicamente empregues, enfatizam-se os

modelos de regressão. Já dos modelos não causais, podem-se destacar, a média móvel (em particular, modelos de média móvel simples e ponderada), o alisamento exponencial (nomeadamente, os modelos de alisamento exponencial simples, de *Holt* e de *Holt-Winters*), o *Autoregressive Integrated Moving Average (ARIMA)* e a extensão *ARIMA* sazonal (*SARIMA*) (Liu et al., 2013; Loureiro et al., 2018; Ni & Fan, 2011; Ren et al., 2020).

Os métodos estatísticos apresentam diversas vantagens. Além de serem métodos simples, permitindo uma fácil compreensão e implementação, os métodos estatísticos são reconhecidos pela sua rapidez no cálculo dos resultados. Por outro lado, estes métodos, por serem empregues há décadas, apresentam uma grande maturidade.

Apesar das vantagens caracterizantes dos métodos estatísticos, segundo Liu et al. (2013), estes métodos dependem muito das características dos dados da série temporal, afetando, assim, a precisão da previsão. Diferentes autores defendem que os métodos estatísticos são limitados, dado que apenas se adequam a situações em que há uma grande quantidade de dados, poucos fatores a afetar as vendas reais e fatores quase lineares, o que não acontece na têxtil-vestuário, que tem vários fatores não lineares a afetar as vendas reais (Caglayan et al., 2020; Ni & Fan, 2011; Ren et al., 2020; Yu et al., 2011). Perante isto, dadas as características predominantes na indústria têxtil e do vestuário, estes métodos tradicionais tornam-se insuficientes para a previsão de vendas (Beheshti-Kashi et al., 2015; Choi et al., 2011; Nenni et al., 2013), acrescentando-se que estes métodos não têm vindo a ser estudados nos últimos anos (Liu et al., 2013).

Perante as limitações dos métodos estatísticos tradicionais, aliadas ao avanço das tecnologias de computação, métodos de IA também têm vindo a ser empregues neste tipo de indústria, fornecendo, segundo diferentes autores, resultados de previsão mais precisos que os métodos tradicionais (Au et al., 2008; Choi et al., 2011; Liu et al., 2013; Yu et al., 2011). Destes, segundo a literatura revista, podem-se destacar diferentes métodos tipicamente implementados na indústria têxtil e do vestuário: *Evolutionary Neural Networks (ENN)*, *Artificial Neural Networks (ANN)*, *Extreme Learning Machines (ELM)* e *Extended Extreme Learning Machines (EELM)* (Choi et al., 2011; Liu et al., 2013; Ren et al., 2020).

Embora alguns autores defendam que os métodos de IA possam fornecer um desempenho de precisão melhor comparativamente com os métodos estatísticos, estes métodos requerem um elevado tempo de previsão e de um requisito computacional elevado, o que se torna uma grande barreira para a sua aplicação em algumas indústrias, em particular, na têxtil, dadas as características particulares da mesma (Choi et al., 2011; Liu et al., 2013; Sun et al., 2008; Yu et al., 2011).

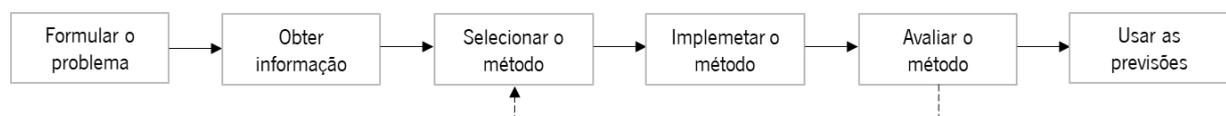
Perante as limitações tanto dos métodos estatísticos tradicionais como dos de IA, surgem os métodos de previsão híbridos para a têxtil-vestuário. Estes combinam as demais classes, utilizando, para isso, métodos estatísticos para prever as componentes lineares e métodos de IA para prever as não lineares. Desta forma, os métodos híbridos permitem que as previsões usufruam das vantagens de ambos os métodos, ao mesmo tempo que superam as desvantagens dos mesmos, tornando-se métodos ideais para a indústria têxtil e do vestuário ao presarem pela eficiência e pela eficácia (Choi et al., 2011; Liu et al., 2013; Ren et al., 2020).

### 3.3 Métodos de Previsão da Procura

Segundo Armstrong (2002), uma ordem lógica de etapas deve ser seguida na implementação de métodos de previsão da procura, sendo a mesma apresentada na **Figura 3**. Segundo esta ordem de etapas de previsão, primeiramente, deve haver uma formulação do problema, seguindo-se a obtenção de informação e, posteriormente, a seleção de métodos. Após a seleção de métodos de previsão, dever-se-á proceder a uma implementação dos mesmos e, seguidamente, a uma avaliação. Finalizada a avaliação, poderão ser seguidos dois diferentes caminhos. Se, por um lado, houver satisfação com os resultados obtidos consoante a avaliação dos métodos, dever-se-á seguir com as previsões efetuadas, utilizando-as. Se, em contrapartida, ainda não houver satisfação com as previsões efetuadas, deve-se voltar à etapa de seleção dos métodos de previsão e repetir o processo até haver uma satisfação com os resultados obtidos. Estas etapas de previsão devem ser seguidas, independentemente do método de previsão em questão.

**Figura 3**

*Etapas de previsão*



*Nota.* Elaboração própria baseado em Armstrong (2002, p. 9).

Para a previsão da procura na indústria têxtil e do vestuário, diferentes métodos de previsão da procura têm vindo a ser propostos, destacando-se, alguns métodos de previsão estatísticos, métodos de IA e métodos híbridos. Perante isto, torna-se crucial estudar, no presente projeto, estes diferentes métodos implementados.

### 3.3.1 Métodos de Previsão Estatísticos

Os métodos estatísticos, ou métodos quantitativos, recorrem à matemática, baseando-se em dados históricos da variável a prever e em variáveis associadas à variável estudada. Estes métodos podem ser divididos em duas categorias: Modelos não causais e Modelos causais (Heizer et al., 2017; Reid & Sanders, 2012).

Os modelos não causais, também denominados por modelos de séries temporais ou modelos extrapolativos, “assumem que todas as informações necessárias para gerar uma previsão estão contidas na série temporal de dados” (Reid & Sanders, 2012, p. 286), considerando uma série temporal “uma série de observações feitas em intervalos regulares durante um período de tempo especificado.” (Reid & Sanders, 2012, p. 286). Assim, estes modelos obtêm previsões baseadas apenas no histórico dos dados da variável estudada. Já os modelos causais, ou modelos explicativos, pressupõem que a variável a prever está relacionada de alguma forma com outras variáveis ou fatores, sendo necessário perceber essa relação em termos matemáticos de forma a obter previsões sobre a variável em questão.

Comparando os modelos não causais com os causais, verifica-se que os primeiros são mais fáceis de aplicar e podem gerar previsões mais rapidamente, uma vez que não exigem a construção de um modelo como os causais (Heizer et al., 2017; Reid & Sanders, 2012).

#### **A. Modelos Não Causais**

##### **Componentes das Séries Temporais**

Com o objetivo de obter previsões através de modelos não causais, numa fase inicial, deve-se proceder a uma análise das séries temporais em estudo, através da identificação das suas componentes de forma a projetá-las nas previsões futuras.

Uma série temporal é composta por uma componente sistemática, que faz parte do padrão que se pretende prever (pode conter nível, tendência, sazonalidade e ciclo), e por uma componente aleatória, que deve ser estimada, mas não pode ser prevista (Heizer et al., 2017; Reid & Sanders, 2012):

1. *Nível*: Padrão em que os dados flutuam em torno de uma média constante. Sendo um valor médio dos valores da série temporal, indica-nos a escala de ordem de grandeza da mesma;
2. *Tendência*: padrão crescente ou decrescente dos dados ao longo do tempo;
3. *Sazonalidade*: qualquer padrão que se repete regularmente e tem uma duração constante de curto prazo, por exemplo, dias, semanas, meses, trimestres, quadrimestres ou semestres;

4. *Ciclos*: oscilações, positivas ou negativas, de longo prazo, em relação ao nível da série. Distingue-se da sazonalidade pela sua extensão e magnitude, tendo uma duração irregular e superior a um ano. O padrão cíclico é, assim, mais difícil de prever do que os outros padrões e está, geralmente, associado a ciclos de negócios;
5. *Variações aleatórias*: variações irregulares dos dados causadas pelo acaso ou por situações involuntárias, sem qualquer padrão identificável e, por isso, não podem ser previstas. Assim, quanto mais variação aleatória a série tiver, mais difícil será prever com precisão.

Tanto a tendência como a sazonalidade podem ser de natureza aditiva ou multiplicativa, sendo as suas características abaixo explicadas.

### **Decomposição**

Os métodos de decomposição tornam possível a identificação das componentes das séries temporais que, mais tarde, serão cruciais para a identificação do melhor método de previsão a adotar. Podem ser identificados diferentes modelos de decomposição, de entre os quais, o modelo de decomposição aditiva e multiplicativa, representados pelas **equações (1) e (2)**.

$$\text{Modelo de Decomposição Aditiva:} \quad Y_t = T_t + S_t + R_t \quad (1)$$

$$\text{Modelo de Decomposição Multiplicativa:} \quad Y_t = T_t * S_t * R_t \quad (2)$$

onde

$Y_t$  é o valor observado no período  $t$  da série temporal,

$T_t$  é a componente do ciclo-tendência no período  $t$ ,

$S_t$  é a componente sazonal no período  $t$ ,

$R_t$  é a componente aleatória no período  $t$ .

A decomposição aditiva é mais adequada se a magnitude das flutuações sazonais ou a variação em torno do ciclo-tendência não variarem com o nível da série, isto é, apresentarem flutuações constantes independentemente do nível da série temporal. Caso as flutuações sazonais ou as flutuações em torno do ciclo-tendência sejam proporcionais ao nível da série, isto é, aumentam e diminuem proporcionalmente com aumentos e diminuições no nível da série, dependendo de um determinado fator, então um modelo multiplicativo é o mais apropriado (Hyndman & Athanasopoulos, 2018; Makridakis et al., 1998).

## Naïve

O método Naïve, utiliza a observação mais recente como previsão, isto é, considera que o valor previsto de  $t$  ( $\hat{Y}_t$ ), é igual ao valor da última observação da série temporal ( $Y_{t-1}$ ), como é ilustrado na **equação (3)**.

$$\hat{Y}_t = Y_{t-1} \quad (3)$$

Este método também é conhecido por método das previsões de caminho aleatório, uma vez que tem um bom comportamento na previsão de séries com dados que detêm de uma forte componente aleatória (Hyndman & Athanasopoulos, 2018). Por outro lado, não é aconselhável utilizar este método para casos em que há uma grande variação dos dados de período para período.

Dadas as suas vantagens de simplicidade e facilidade de aplicação, o método Naïve, normalmente, é usado para avaliar o desempenho das previsões de outros modelos de previsão que se está a aplicar, funcionando, assim, como método de comparação (Reid & Sanders, 2012).

### *Modelo Naïve Sazonal*

O modelo Naïve Sazonal (sNaïve) é uma extensão do método Naïve que consegue capturar possíveis variações sazonais. Para isso, sNaïve assume que o valor previsto de  $t$  ( $\hat{Y}_t$ ) é igual ao valor observado conhecido do mesmo período sazonal (por exemplo, dia da semana, mês, trimestre, quadrimestre, semestre) ( $Y_{t-s}$ ), onde  $s$  é o período sazonal (por exemplo, sete para dados semanais, doze para dados mensais, quatro para trimestrais, três quadrimestres e dois para semestrais). A **equação (4)** ilustra a aplicação do modelo explicado (Hyndman & Athanasopoulos, 2018; Spiliotis et al., 2020).

$$\hat{Y}_t = Y_{t-s} \quad (4)$$

## Médias Móveis

### *Método de Médias Móveis Simples*

O método de Médias Móveis Simples (MMS) é semelhante à média simples, contudo, nela é incluído apenas dados dos  $n$  períodos mais recentes, ao invés de todos os períodos como no caso da média simples. Assim, para o modelo de MMS, é definido o número de períodos a incluir para a média ( $t - n$ ) e, à medida que vai havendo novos dados, há uma substituição da observação mais antiga pela mais recente, mantendo-se o número de observações constante. A **equação (5)** representa a aplicação do modelo.

$$\hat{Y}_t = \frac{Y_{t-1} + Y_{t-2} + \dots + Y_{t-n}}{n} = \frac{\sum_i^t Y_i}{n} \quad (5)$$

onde  $i = t - n$

#### *Modelo de Médias Móveis Ponderadas*

O modelo de Médias Móveis Ponderadas (MMP) é uma extensão à MMS, contudo, enquanto a MMS dá um peso igual aos  $t - n$  períodos de dados para obter as suas previsões (ignorando todos os dados mais antigos anteriores ao primeiro período considerado), a MMP admite pesos distintos a diferentes períodos (permitindo que as observações passadas sejam ponderadas de forma diferente, em vez de ignoradas). Assim, o valor previsto do período  $t$  ( $\hat{Y}_t$ ) é obtido pela soma dos produtos entre os valores observados nos períodos anteriores a  $t$  ( $Y_i$ ) e o valor os respetivos pesos ( $\alpha_i$ ), como apresentado pela **equação (6)**.

$$\hat{Y}_t = \sum_i^t \alpha_i Y_i \quad \text{com } 0 \leq \alpha \leq 1 \quad (6)$$

onde  $i = t - n$

Neste modelo cada período pode ser ponderado de forma diferente, desde que a soma dos pesos totais seja 1, não existindo uma fórmula definida para determinar o peso para cada período. Desta forma, os pesos devem ser atribuídos com suporte à experiência, tendo por base diferentes fatores como o conhecimento do setor (Reid & Sanders, 2012). No caso de haver tendência, por exemplo, é possível atribuir pesos maiores às observações mais recentes, tornando as previsões mais responsivas às mudanças na procura comparativamente com a MMS.

Tanto as médias móveis simples, como as ponderadas, tendem a suavizar irregularidades na procura de curto prazo, fornecendo, assim, estimativas estáveis. Perante isto, estas devem ser utilizadas para séries com dados estáveis, isto é, séries em que os dados flutuem em torno de um determinado nível e sem tendência e sazonalidade observáveis (Chopra & Meindl, 2013; Heizer et al., 2017).

Por outro lado, destaca-se que quanto menor for o número de períodos a considerados ( $n$ ), maior será o ajustamento feito às observações anteriores, havendo uma previsão mais responsiva às mudanças na procura, mas também à aleatoriedade. Em contrapartida, o aumento do número de períodos leva a que as previsões sejam mais estáveis, mas menos sensível a alterações nos dados. Todavia, não existe um número fixo de períodos a considerar, devendo a sua seleção ser baseada nas características dos dados.

## Alisamento Exponencial

### *Método Alisamento Exponencial Simples*

O método de Alisamento Exponencial Simples (AES), utiliza um procedimento mais sofisticado da Média Móvel Ponderada, nomeadamente, onde os pesos atribuídos aos dados vão diminuindo ou aumentando exponencialmente, segundo uma constante de alisamento ( $\alpha$ ), à medida que os dados se vão tornando mais antigos. Assim, neste modelo, o valor previsto para o período  $t$  ( $\hat{Y}_t$ ) é obtido atribuindo um determinado peso  $\alpha$  ao valor observado no período anterior ( $Y_{t-1}$ ) e um peso  $(1 - \alpha)$  à previsão para o período anterior ( $\hat{Y}_{t-1}$ ), como é ilustrado pela **equação (7)** abaixo apresentada (Reid & Sanders, 2012).

$$\hat{Y}_t = \alpha Y_{t-1} + (1 - \alpha) \hat{Y}_{t-1} \quad (7)$$

onde  $\alpha$  é a constante de alisamento do nível, estando  $0 \leq \alpha \leq 1$

Substituindo o valor previsto do período anterior pela expressão desenvolvida, verificar-se-á a diminuição exponencial referida<sup>2</sup>.

Para obter o valor previsto para o período  $t$  ( $\hat{Y}_t$ ) é necessário possuir o valor observado no período anterior ( $Y_{t-1}$ ). Neste sentido, de forma a iniciar a previsão, pode-se utilizar como método de inicialização o pressuposto de que o valor observado do período um é igual ao valor previsto do mesmo, isto é,  $Y_1 = \hat{Y}_1$ .

A escolha do valor apropriado para a constante de alisamento do nível não é uma tarefa fácil e pode ditar a precisão das previsões. Esta constante pode assumir valores mais altos quando se pretende dar mais peso aos dados mais recentes ou, em contrapartida, pode assumir valores mais baixos, quando se pretende que a previsão seja menos responsiva a observações mais recentes. No caso de ser provável que o nível da série mude, vulgarmente, atribuem-se valores mais altos à constante, estando assim, as previsões, também mais sujeitas a variações aleatórias dos dados. Em contrapartida, são usados valores mais baixos quando a média da série é estável, levando a previsões igualmente mais estáveis (Heizer et al., 2017; Reid & Sanders, 2012). Segundo, Heizer et al. (2017), em ambiente de negócios, a constante de alisamento, geralmente, assume valores entre 0,05 e 0,50. Contudo, existem diferentes *softwares* que permitem a otimização das constantes de alisamento para cada caso em estudo.

O método de AES é apropriado para o caso em que a série não tem tendência nem sazonalidade observáveis (Chopra & Meindl, 2013) e fornece excelentes resultados de previsão para previsões de curta e média duração (Reid & Sanders, 2012).

---

<sup>2</sup>  $\hat{Y}_t = \alpha Y_{t-1} + (1 - \alpha)[\alpha Y_{t-2} + (1 - \alpha)\hat{Y}_{t-2}]$

Ao longo dos anos foram surgindo extensões ao método de Alisamento Exponencial Simples, nomeadamente, o modelo *Holt* e o modelo de *Holt*Winters.

#### *Modelo de Holt*

O modelo de *Holt*, ou modelo alisamento exponencial corrigido pela tendência, surgiu como extensão ao método AES de forma a permitir a previsão de séries temporais com tendência. Para isso, o modelo utiliza duas equações de suavização e outra equação para obter uma previsão com tendência. Primeiramente, o nível da série é suavizado pela **equação (8)**, posteriormente, através da **equação (9)**, a tendência é suavizada e, por fim, na **equação (10)** é gerada uma previsão que inclui tendência através da soma dos resultados das equações relativas ao nível e tendência.

*Nível:*

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (8)$$

*onde*

$L_t$  é a estimativa para o nível da série no período  $t$ ;

$\alpha$  é a constante de alisamento do nível, estando  $0 \leq \alpha \leq 1$ .

*Tendência:*

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_t \quad (9)$$

*onde*

$T_t$  é a estimativa para a tendência da série no período  $t$ ;

$\beta$  é a constante de alisamento da tendência, estando  $0 \leq \beta \leq 1$ .

*Previsão:*

$$\hat{Y}_{t+m} = L_t + T_t m \quad (10)$$

*onde  $m$  = número de períodos à frente para qual é realizada a previsão*

No mesmo sentido que a constante de alisamento do nível, a constante de alisamento da tendência quando toma valores mais altos leva a que a previsão seja mais responsiva a mudanças recentes na tendência, em contraste com um  $\beta$  baixo que tende a suavizar a tendência atual.

No seguimento do método de AES, no modelo simples de *Holt* é necessário recorrer a um método de inicialização. Para este é necessário recorrer a mais duas estimativas, uma para  $L_1$  e outra para  $T_1$ .

Assim, tem-se que,  $L_1 = Y_1$  e  $T_1 = Y_2 - Y_1 = \frac{(Y_4 - Y_1)}{3}$ .

O modelo de *Holt* é apropriado para o caso em que a série tem um nível e uma tendência na componente sistemática, mas não tem sazonalidade observável (Heizer et al., 2017; Reid & Sanders, 2012). Neste seguimento, mais tarde, surge o modelo de *Holt-Winters*, que vem suavizar a sazonalidade.

#### *Modelo de Holt-Winters*

O modelo de *Holt-Winters* (*HW*), ou modelo alisamento exponencial corrigido pela tendência e sazonalidade, surgiu como extensão ao modelo de *Holt* de forma a permitir a previsão de séries temporais com tendência e sazonalidade. Para isso, o modelo utiliza três equações de suavização, para o nível, tendência e sazonalidade. Por fim, usa outra equação que gera uma previsão que considera estas três componentes da série.

Existem dois modelos distintos de *Holt-Winters* segundo a sazonalidade seja aditiva ou multiplicativa estando as **equações (11), (12), (13) e (14)** abaixo apresentadas a ilustrar a obtenção de previsões para o caso de sazonalidade aditiva e as **equações (15), (16), (17) e (18)** para o caso de sazonalidade multiplicativa.

#### *HW Sazonalidade Aditiva:*

*Nível:*

$$L_t = \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (11)$$

*Tendência:*

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (12)$$

*Sazonalidade:*

$$S_t = \gamma(Y_t - L_t) + (1 - \gamma)S_{t-s} \quad (13)$$

*Previsão:*

$$\hat{Y}_{t+m} = L_t + T_t m + S_{t-s+m} \quad (14)$$

#### *HW Sazonalidade Multiplicativa:*

*Nível:*

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1}) \quad (15)$$

*Tendência:*

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1} \quad (16)$$

*Sazonalidade:*

$$S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s} \quad (17)$$

*Previsão:*

$$\hat{Y}_{t+m} = (L_t + T_t m)S_{t-s+m} \quad (18)$$

Onde  $s$  é o número de períodos de cada ciclo sazonal,  $L_t$ ,  $T_t$  e  $S_t$  representam o nível, tendência e sazonalidade da série, respetivamente,  $\alpha$ ,  $\beta$  e  $\gamma$  estão entre 0 e 1 (inclusive) e dizem respeito às constantes de alisamento do nível, tendência e sazonalidade, respetivamente,  $m$  diz respeito ao número de períodos à frente para qual é realizada a previsão e  $\hat{Y}_{t+m}$  representa a previsão para o período  $t + m$  (Makridakis et al., 1998).

Também o modelo de *HW* recorre ao método de inicialização para  $L_s$ ,  $T_s$  e  $S_1$ , apresentados pelas **equações (19), (20), (21) e (22)**, respetivamente.

$$L_s = \frac{1}{s}(Y_1 + \dots + Y_s) \quad (19)$$

$$T_s = \frac{1}{s} \left( \frac{Y_{s+1} - Y_1}{s} + \dots + \frac{Y_{s+s} - Y_s}{s} \right) \quad (20)$$

$$S_1 = Y_1 - L_s, \dots, S_s = Y_s - L_s \quad (\text{sazonalidade aditiva}) \quad (21)$$

$$S_1 = \frac{Y_1}{L_s}, \dots, S_s = \frac{Y_s}{L_s} \quad (\text{sazonalidade multiplicativa}) \quad (22)$$

Neste sentido, o modelo de *HW* multiplicativo é apropriado para o caso em que a série tem tendência e o *HW* aditivo para o caso em que tem tendência e sazonalidade observáveis (Chopra & Meindl, 2013).

## Média Móvel Integrada Autorregressiva

### *ARIMA*

Os modelos de média móvel integrada autorregressiva, em inglês *Autoregressive Integrated Moving Average (ARIMA)*, resultam da combinação de três componentes, nomeadamente, a autorregressiva (AR), a de diferenciação (I) e a das médias móveis (MA). Do modelo geral não sazonal, *ARIMA* ( $p, d, q$ ), onde  $p$  representa a ordem do processo autorregressivo (AR),  $d$  o grau de diferenciação envolvido (I), e o  $q$  a ordem do processo de média móvel (MA), surgem casos específicos de modelos *ARIMA*, em particular,

- *ARIMA* (0, 0, 0): Se  $p = d = q = 0$ , é um modelo de ruído branco, dado que não há parte AR (ou seja,  $Y_t$  não depende de  $Y_{t-1}$ ), não há diferenciação envolvida (I), nem há parte MA (isto é,  $Y_t$  não depende de  $E_{t-1}$ );
- *ARIMA* (0, 1, 0): Se  $p = q = 0$  e  $d = 1$ , é um modelo de passeio aleatório, uma vez que não possui partes AR e MA e envolve uma diferença.

Se  $p$ ,  $d$  ou  $q$  forem igual a zero, o modelo pode ser escrito com uma notação mais pequena, usando as partes do modelo que estão no modelo, veja-se o exemplo do caso *ARIMA* (0, 1, 0) poder ser denominado como  $I(1)$ .

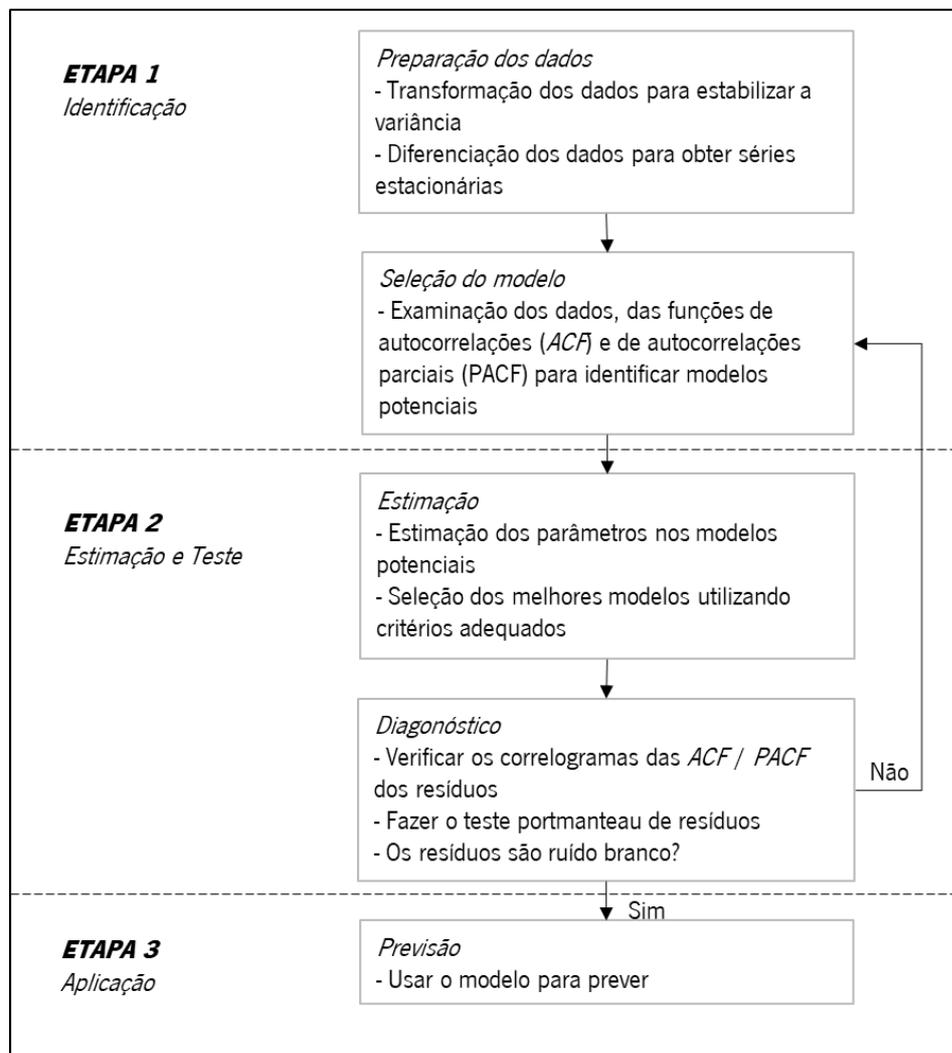
A aplicação dos modelos *ARIMA* deverá seguir-se uma abordagem composta por três etapas principais, nomeadamente:

- Etapa 1 – Identificação do modelo *ARIMA* apropriado;
- Etapa 2 – Estimativa dos parâmetros do modelo;
- Etapa 3 – Teste e Aplicação do modelo.

Abaixo, pela **Figura 4**, é apresentado um esquema que sintetiza a abordagem a seguir.

**Figura 4**

*Etapas para a aplicação do modelo ARIMA*



*Nota.* Elaboração própria baseado em Makridakis et al. (1998, p. 314).

Salienta-se que a aplicação dos modelos *ARIMA* requer que a série temporal seja estacionária, isto é, que, ao longo do tempo, os dados da série flutuam em torno de uma média constante e a variação da flutuação permaneça constante. Desta forma, caso a mesma não o seja, dever-se-á proceder a uma transformação prévia dos dados de forma a estabilizar a variância da série. Uma maneira de remover a não-estacionariedade é através do método de diferenciação que, no caso das primeiras diferenças, consiste em subtrair observações consecutivas para eliminar tendências.

Através das primeiras diferenças, obtém-se uma série diferenciada de primeira ordem pela **equação (23)**.

$$Y'_t = Y_t - Y_{t-1} \quad (23)$$

De forma a simplificar o método das diferenças, pode utilizar-se um operador de retrocesso,  $B$ . Este operador tem o efeito de retroceder os dados um período, como é ilustrado pela **equação (24)**.

$$BY_t = Y_{t-1} \quad (24)$$

Assim, a série diferenciada de primeira ordem pode ser obtida através da **equação (25)**.

$$Y'_t = Y_t - BY_t = (1 - B)Y_t \quad (25)$$

A diferenciação dos dados ( $I$ ) deve ser realizada tantas vezes quantas for necessário, até que a série se torne estacionária, sendo esse grau de diferenciação definido por  $d$ . Contudo, na prática, geralmente não é necessário ir além das diferenças de segunda ordem. A **equação (26)** ilustra uma série diferenciada de ordem  $d$ .

$$Y_t^d = (1 - B)^d Y_t \quad (26)$$

Após a série se tornar estacionária, poder-se-á aplicar um modelo *ARMA* ( $p, q$ ). Assim, o valor previsto para o período  $t$  utilizando o modelo *ARIMA* ( $p, d, q$ ) é obtido através da **equação (27)**.

$$(1 - \phi_1 B - \dots - \phi_p B^p) (1 - B)^d Y_t = c + (1 - \theta_1 B - \dots - \theta_q B^q) e_t \quad (27)$$

onde

$\phi_p$  é o  $p^o$  parâmetro autorregressivo, sendo que para

$p = 1$ , então  $-1 < \phi_1 < 1$

$p = 2$ , então  $-1 < \phi_1 < 1$ ,  $-1 < \phi_2 < 1$ ,  $\phi_2 + \phi_1 < 1$ ;  $\phi_2 - \phi_1 < 1$ ;

$c$  é um termo constante;

$\theta_q$  é o  $q^o$  parâmetro da média móvel, sendo que para

$p = 1$ , então  $-1 < \theta_1 < 1$

$p = 2$ , então  $-1 < \theta_1 < 1$ ,  $-1 < \theta_2 < 1$ ,  $\theta_2 + \theta_1 < 1$ ;  $\theta_2 - \theta_1 < 1$ ;

$e_t$  é o termo de erro no período  $t$ , estando normalmente distribuído com média 0 e variância 1.

## SARIMA

Os modelos sazonais de média móvel integrada autorregressiva, em inglês, *Seasonal Autoregressive Integrated Moving Average (SARIMA)*, são uma extensão dos modelos *ARIMA* com a adição das componentes relacionadas com a sazonalidade. A notação abreviada do modelo geral sazonal corresponde a  $ARIMA(p, d, q)(P, D, Q)^S$ , onde  $(p, d, q)$  diz respeito à parte não sazonal e  $(P, D, Q)^S$  à parte sazonal. As componentes  $P$ ,  $D$  e  $Q$  tomam os mesmos significados que as componentes  $p$ ,  $d$  e  $q$ , mas para a parte sazonal. Já  $S$  corresponde ao número de períodos de um ciclo sazonal.

De forma similar aos modelos *ARIMA*, a aplicação dos modelos *SARIMA* requerem que a série temporal seja estacionária. Assim, para remover a não-estacionariedade, além do método de diferenciação para eliminar tendências, dever-se-á proceder à diferença sazonal, isto é, subtrair observações separadas por um ciclo sazonal de forma a eliminar sazonalidades.

Neste sentido, uma série sazonalmente diferenciada obtém-se pela **equação (28)**, sendo que, tal como nas primeiras diferenças, a diferenciação pode ser repetida para obter uma diferenciação sazonal de segunda ordem, contudo isso é raramente necessário.

$$Y_t^{SD} = (1 - B^S)^D Y_t \quad (28)$$

$$\text{com } B^S Y_t = Y_{t-S}$$

O valor previsto para o período  $t$  utilizando o modelo  $ARIMA(p, d, q)(P, D, Q)^S$  é obtido através da **equação (29)**.

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - \Phi_1 B^S - \dots - \Phi_P B^{PS}) (1 - B)^d (1 - B^S)^D Y_t = \quad (29) \\ = (1 - \theta_1 B - \dots - \theta_q B^q)(1 - \Theta_1 B^S - \dots - \Theta_Q B^{QS}) e_t$$

onde

$\phi_p$ ,  $\theta_q$  e  $e_t$  correspondem ao  $p^{\circ}$  parâmetro autorregressivo da parte não sazonal, ao  $q^{\circ}$  parâmetro da média móvel da parte não sazonal e ao termo de erro no período  $t$ ,  
respetivamente, e tomam as mesmas restrições que o modelo não sazonal;

$\Phi_P$  é o  $P^{\circ}$  parâmetro autorregressivo da parte sazonal e  $\Theta_Q$  é o  $Q^{\circ}$  parâmetro da média móvel da parte sazonal, ambos com as mesmas restrições que os parâmetros não sazonais.

Realça-se que, apesar de não haver limite para a variedade dos modelos *ARIMA* e *SARIMA*, geralmente, na prática, raramente se recorre a valores para as componentes  $p$ ,  $d$ ,  $q$ ,  $P$ ,  $D$  e  $Q$  superiores a 2 (Makridakis et al., 1998).

## A. Modelos Causais

### Regressão

#### *Regressão Linear Simples*

Através da Regressão Linear Simples é possível construir um modelo que estabeleça uma relação linear entre duas variáveis, a variável dependente (que está a ser prevista) e a variável independente. Neste sentido, a variável dependente estará linearmente relacionada com a variável independente, sendo a sua relação dada pela **equação (30)**.

$$\hat{Y}_t = a + b X_t \quad (30)$$

onde

$\hat{Y}_t$  = Variável dependente, i.e., o valor previsto do período  $t$ ;

$X_t$  = Variável independente (Número de períodos de tempo quando  $X_t = 0$ );

$a$  = Constante representativa da interceção do eixo  $y$  (Valor de  $\hat{Y}_t$  quando  $X_t = 0$ );

$b$  = Constante representativa da inclinação da linha de regressão.

Os parâmetros  $a$  e  $b$  são definidos pela minimização do quadrado dos erros ou dos desvios das observações relativamente à reta que define a relação das duas variáveis.

#### *Regressão Múltipla*

A Regressão Múltipla é uma extensão da Regressão Linear Simples, permitindo, não só a incorporação de uma variável independente, mas sim de várias, como ilustra a **equação (31)**.

$$\hat{Y}_t = a + b_1 X_{1,t} + b_2 X_{2,t} + \dots + b_k X_{k,t} \quad (31)$$

onde

$\hat{Y}_t$  = Variável dependente, i.e., o valor previsto do período  $t$ ;

$X_{1,t}, X_{2,t}$  e  $X_{k,t}$  = Variáveis independentes;

$a$  = Constante representativa da interceção do eixo  $y$ ;

$b_1, b_2$  e  $b_k$  = Coeficientes que representam a influência das variáveis independentes sobre a variável dependente;

$k$  = o número de variáveis independentes consideradas no modelo.

A linha de regressão múltipla estabelece, assim, a relação entre uma variável dependente e várias variáveis independentes (Heizer et al., 2017; Reid & Sanders, 2012).

### 3.3.2 Métodos de Inteligência Artificial

A implementação de diferentes métodos de IA tem vindo a ser estudada para a previsão da procura na indústria têxtil e do vestuário, recorrendo, estes métodos, a técnicas de IA, tal como redes neurais artificiais (*ANN*).

Diferentes autores desenvolveram estudos onde utilizaram métodos de previsão baseados em *ANN* para prever as vendas na têxtil-vestuário. A título de exemplo, Caglayan et al. (2020) e Frank et al. (2003) estudaram a aplicação de *ANN* básicas neste tipo de indústria. Caglayan et al. (2020) aplicaram métodos de previsão baseados em *ANN* básicas e compararam os seus resultados com métodos de regressão, obtendo melhor desempenho em termos de precisão da previsão por parte das *ANN*. Por outro lado, Frank et al. (2003), compararam o desempenho de métodos estatísticos não causais, em particular, modelos de alisamento exponencial simples e de *Holt-Winters*, com as *ANN* básicas, tendo obtido melhor desempenho utilizando como critério apenas uma medida de avaliação de desempenho, contudo, sem diferenças significativas e pior desempenho no que diz respeito às correlações entre as vendas previstas e as vendas observadas.

Alternativamente às *ANN* básicas, outros autores estudaram a implementação de métodos de previsão baseados em métodos avançados de *ANN*, as *ENN*. Estas dizem respeito a uma combinação híbrida entre as *ANN* e procedimentos de procura evolutiva e têm vindo a apresentar resultados promissores na aplicação de previsões. Veja-se o caso de Au et al. (2008), que compararam o desempenho da aplicação das *ENN* na indústria da moda, com as *ANN* básicas e com o modelo estatístico *SARIMA*. Neste estudo os autores concluíram que as *ENN* apresentam melhores resultados comparativamente com o *SARIMA* para produtos em que há uma baixa incerteza e a componente sazonal é baixa, acrescentando-se o facto de que, a aplicação de *ENN* não exige o conhecimento humano que os métodos tradicionais exigem. Por outro lado, comparando as *ANN* básicas com as evolutivas, as básicas levam menos tempo a completar previsões (Au et al., 2008 e Yu et al., 2011).

Conclui-se que tanto as *ANN*, como as *ENN*, podem aumentar substancialmente a precisão da previsão em comparação com vários métodos tradicionais, contudo requerem um elevado tempo computacional para concluir as previsões (Liu et al., 2013). Perante isto, Yu et al. (2011), referem que, apesar de os modelos de previsão propostos baseados em *ANN* permitirem previsões precisas, estes são inadequados para a indústria têxtil e do vestuário. Segundo os autores, as *ANN* têm problemas de *overfitting*<sup>2</sup> e um

---

<sup>2</sup> O *overfitting* ocorre quando um modelo está muito ajustado ao conjunto de dados da amostra usada para o desenvolver, gerando bom desempenho para dentro dessa amostra e mau desempenho para fora da mesma (Twin, 2021).

maior tempo computacional, contudo, para este tipo de indústria, dada a variedade de produtos caracterizante, é necessária uma previsão imediata.

Perante as desvantagens das *ANN* básicas e evolutivas, surgem outros métodos baseados noutra algoritmo aplicados na têxtil-vestuário: as *ELM*. As *ELM* vêm permitir reduzir, em muito, o tempo de aprendizagem das *ANN* e, por isso, uma aplicação em tempo real. Desta forma, os algoritmos de *ELM* apresentam-se menos lentos que as *ANN* básicas e as *ENN* (Liu et al., 2013; Yu et al., 2011).

### Extreme Learning Machines

As *ELM*, desenvolvidas por Huang et al. (2004), dizem respeito a um algoritmo de aprendizagem para uma única rede neural *feedforward* de camada oculta, denominada, em inglês, por *single hidden layer feedforward neural network (SLFN)*, “que escolhe, aleatoriamente, nós ocultos e determina analiticamente os pesos de saída da *SLFN*” (Huang et al., 2006, p. 489). No desenvolvimento deste algoritmo, Huang et al. (2004, 2006), começam por comprovar que os pesos dos *inputs* e os *bias* da camada oculta de *SLFN* podem ser atribuídos de forma aleatória se as funções de ativação na camada oculta forem infinitamente diferenciáveis. Posteriormente, os autores propõem a *ELM* para *SLFN*, que se encontra abaixo apresentada de forma resumida, através da **equação (32)** e do processo adjacente.

*Algoritmo ELM:*

Dado um conjunto de treino  $\mathfrak{X} = \{(x_i, t_i) | x_i \in R^n, t_i \in R^m, i = 1, \dots, N\}$ , uma função de ativação  $g(\mathcal{X})$  e um número de nó oculto  $\tilde{N}$ ,

*Etapa 1:* atribuir, aleatoriamente, um peso ( $w_i$ ) ao *input* e ao *bias* ( $b_i$ ), com  $i = 1, \dots, \tilde{N}$ ;

*Etapa 2:* calcular a matriz do *output* da camada oculta,  $H$ ;

*Etapa 3:* calcular o peso do *output*,  $\beta$ , onde

$$\beta = H^\dagger T \text{ com } T = [t_1, \dots, t_N]^T \quad (32)$$

A definição deste algoritmo pode ser vista de forma mais detalhada em Huang et al. (2004, 2006), assim como as observações e teoremas subjacentes. Também em Wang et al. (2021), são apresentados os fundamentos teóricos por detrás deste algoritmo.

Apesar da vantagem de rapidez dos métodos baseados em *ELM*, estes apresentam instabilidade quando comparados com as *ANN* básicas, as *ENN* e os métodos estatísticos tradicionais (Liu et al., 2013; Yu et al., 2011). Neste seguimento, surge uma extensão às *ELM*, as *ELM* estendidas (*EELM*), que consistem na execução das *ELM* várias vezes, permitindo, assim, alcançar resultados mais estáveis, comparativamente às *ELM*, e um melhor desempenho quando o número de vezes a realizar a execução das *ELM* é determinado apropriadamente (Choi et al., 2011). As *ELM* estendidas (*EELM*) vêm melhorar

as *ELM* em termos de precisão da previsão, contudo vêm aumentar o tempo de previsão (Ren et al., 2020; Yu et al., 2011).

Dos diferentes autores que estudaram a implementação das *ELM* na previsão de vendas na têxtil-vestuário, destacam-se Sun et al. (2008), que aplicaram as *ELM* e *EELM*, tendo obtido resultados promissores.

Apesar dos métodos baseados em *ELM* e *EELM*, terem sido “desenvolvidos para lidar com problemas de previsão de vendas de moda” (Ren et al., 2020, p. 764), e as *ELM* e *EELM* serem mais rápidas que as *ANM* básicas e *ENN*, estas ainda requerem uma quantidade substancial de tempo para a realização de previsões (Liu et al., 2013; Ren et al., 2020).

### 3.3.3 Métodos Híbridos

De forma a enfrentar as limitações encontradas nos métodos estatísticos e métodos de IA, foram implementados métodos de previsão híbridos na indústria têxtil e do vestuário, com o objetivo de atingir um nível razoável de precisão num tempo curto de previsão, com a restrição de que possam existir poucos dados históricos disponíveis (Choi et al., 2014). Neste sentido, os métodos híbridos combinam os métodos estatísticos e os de IA na previsão da procura, utilizando, geralmente, métodos estatísticos para prever as componentes lineares dos dados e métodos de IA para as componentes não lineares. Ao fazê-lo, os métodos híbridos conseguem usufruir das vantagens de precisão dos métodos de IA e da rapidez dos estatísticos, presando tanto pela eficácia como pela eficiência (Choi et al., 2011; Liu et al., 2013; Ren et al., 2020).

Diversos modelos híbridos foram propostos para aplicação na têxtil-vestuário, desde modelos híbridos baseados em *ANM* básicas, modelos híbridos baseados em *ELM*, a outros modelos híbridos (Liu et al., 2013). Veja-se o sucesso dos estudos de Choi et al. (2012) e Ni & Fan (2011), que desenvolveram modelos híbridos baseados em *ANM* básicas ou Choi et al. (2014) e Wong & Guo (2010) que combinavam *ELM* com outros modelos, apresentando, todos eles, resultados promissores.

A título de exemplo, destaca-se o caso de Yu et al. (2011), que desenvolveram um modelo inteligente de previsão de vendas rápida, em inglês, *intelligent fast sales forecasting model (IFSFM)*, nomeadamente, *ELM* de previsão rápida (*ELM-FF*), para os produtos de moda. O modelo *ELM-FF* proposto combinava o algoritmo *ELM* com métodos tradicionais, podendo utilizar diferentes modelos estatísticos, influenciando a escolha do modelo estatístico na precisão da previsão final. Após o estudo, o modelo mostrou-se eficiente e eficaz para a realização de previsões rápidas de produtos de moda, isto é, as previsões puderam ser concluídas no tempo restrito e a precisão da previsão foi otimizada.

### 3.4 Medidas de Avaliação de Desempenho das Previsões

Segundo Makridakis et al. (1998) “na maioria das situações de previsão, a precisão é tratada como o critério primordial para a seleção de um método de previsão” (p. 41). Esta precisão refere-se ao “quão bem o modelo de previsão é capaz de reproduzir os dados que já são conhecidos” (Makridakis et al., 1998, pp. 41-42).

De forma a avaliar a precisão, são conhecidas diversas medidas de avaliação de desempenho das previsões obtidas através dos resultados dos métodos de previsão, sendo algumas delas apresentadas neste subcapítulo.

#### 3.4.1 Erro de Previsão

Considera-se um bom método de previsão aquele que consegue capturar a componente sistemática da procura em detrimento da componente aleatória, sendo esta última manifestada na forma de erro de previsão.

O erro de previsão da procura para o período  $t$  pode ser dado por  $E_t$ , sendo obtido pela diferença entre o valor da procura real ( $Y_t$ ) e a procura prevista para o período  $t$  no mesmo período ( $\hat{Y}_t$ ). A equação representativa do erro de previsão é apresentada na **equação (33)**.

$$E_t = Y_t - \hat{Y}_t \quad (33)$$

Salienta-se a importância de cálculo e análise desta medida de avaliação de previsões. Pode-se destacar o facto de, caso se verifique que o método de previsão em análise tem consistentemente erros positivos ou erros negativos, poderá estar-se a cometer um erro sistemático, que deve ser corrigido. De igual forma, se estiverem a ser obtidos erros muito elevados, comparativamente a erros observados no passado, poder-se-á estar a cometer também erros sistemáticos, devendo-se ponderar a alteração do método de previsão (Chopra & Meindl, 2013; Makridakis et al., 1998).

#### 3.4.2 Erro Quadrático Médio

O Erro Quadrático Médio, em inglês, *Mean Square Error (MSE)*, é obtido pelo quociente entre a soma dos quadrados dos erros de previsão e o número de períodos utilizados no cálculo do erro de previsão ( $n$ ), tal como é observável pela **equação (34)**.

$$MSE_n = \frac{\sum_{t=1}^n E_t^2}{n} \quad (34)$$

Dado que no *MSE* os erros são elevados ao quadrado, verifica-se que o mesmo penaliza mais os erros mais elevados do que os erros menores. Concluiu-se ser indicado utilizar o *MSE* como medida de

avaliação quando o erro de previsão tem uma distribuição simétrica em torno de zero (Chopra & Meindl, 2013; Makridakis et al., 1998, Reid & Sanders, 2012).

### 3.4.3 Raiz do Erro Quadrático Médio

A Raiz do Erro Quadrático Médio, em inglês, *Root Mean Square Error (RMSE)*, como o próprio nome indica, é obtido pela raiz quadrada do *MSE*, ilustrada pela **equação (35)**.

$$RMSE_n = \sqrt{\frac{\sum_{t=1}^n E_t^2}{n}} \quad (35)$$

A *RMSE* é idêntica ao *MSE*, contudo, ao contrário do *MSE*, esta é expressa na mesma unidade de medida que a série dos dados e que as previsões. Assim, esta medida de avaliação de desempenho também é chamada de desvio padrão dos erros de previsão (Armstrong, 2002).

### 3.4.4 Desvio Absoluto Médio

O Desvio Absoluto Médio, em inglês, *Mean Absolute Deviation (MAD)*, como o próprio nome indica, é definido como a média do desvio absoluto de todos os períodos, obtido, assim, pelo quociente entre o somatório dos valores absolutos dos erros de previsão e o número de períodos utilizados no cálculo do erro de previsão (*n*), tal como é representado na **equação (36)**.

$$MAD_n = \frac{\sum_{t=1}^n |E_t|}{n} \quad (36)$$

Caso o erro de previsão não tenha uma distribuição simétrica, considera-se o *MAD* uma medida de avaliação de previsão melhor do que o *MSE* (Chopra & Meindl, 2013; Makridakis et al., 1998, Reid & Sanders, 2012).

### 3.4.5 Erro Absoluto Percentual Médio

O Erro Absoluto Percentual Médio, em inglês, *Mean Absolute Percentage Error (MAPE)* é obtido pelo quociente entre a soma dos valores absolutos do erro percentual (PE) e o número de períodos utilizados no cálculo do erro de previsão (*n*). A **equação (37)** apresenta a obtenção do *MAPE*.

$$MAPE_n = \frac{\sum_{t=1}^n |PE_t|}{n} \quad (37)$$

$$\text{onde } PE_t = \frac{E_t}{Y_t} \times 100$$

O *MAPE* é considerado um bom avaliador de previsões para séries com sazonalidade significativa e séries em que a procura varia, consideravelmente, de um período para o outro. Contudo, quando a série temporal tem dados com zero ou valores muito próximos de zero, o *MAPE* pode não ser um bom

indicativo uma vez que o cálculo do PE pode não fazer sentido (Chopra & Meindl, 2013, Makridakis et al., 1998).

#### 3.4.6 Critério de Informação *Bayesiano*

O Critério de Informação *Bayesiano*, em inglês, *Bayesian Information Criterion (BIC)* é uma medida de avaliação de desempenho utilizada para comparar modelos alternativos pertencentes à mesma classe. Desta forma, pertencendo os modelos à mesma classe, tendo como critério de avaliação o BIC, o modelo que minimiza o seu valor deverá ser escolhido, uma vez que deverá ser aquele que apresentará melhor desempenho em termos de previsão. O *BIC* pode ser obtido pela **equação (38)**.

$$BIC = T \log\left(\frac{\sum_{t=1}^T E_t^2}{T}\right) + (k + 2) \log(T) \quad (38)$$

onde

$T$  é o número de observações usadas para a estimativa;

$(k + 2)$  é o número de parâmetros no modelo, com  $k$  igual ao número de preditores no modelo

Esta estatística é usualmente utilizada para modelos *ARIMA* (Hyndman & Athanasopoulos, 2018; Makridakis et al., 1998).

#### 3.4.7 Coeficiente de Determinação e Coeficiente de Determinação Ajustado

O coeficiente de determinação, ou  $R^2$ , permite medir o ajustamento dos modelos de regressão. Este pode ser obtido pelo coeficiente entre a variação explicada e a variação total, conforme ilustrado através da **equação (39)**.

$$R^2 = \frac{\sum(\hat{Y}_t - \bar{Y})^2}{\sum(Y_t - \bar{Y})^2} \quad (39)$$

Uma forma a selecionar o melhor modelo tendo por base o coeficiente de determinação seria selecionar o modelo que maximizasse o valor de  $R^2$ . Contudo, através deste critério, o modelo selecionado seria aquele que contivesse todas as variáveis explicativas, pois o aumento das mesmas permite um aumento do  $R^2$ . Contudo, nem todas as variáveis explicativas devem ser incluídas no modelo de previsão. Perante isto, maximizar o valor de  $R^2$  não deve ser o objetivo. Assim, de forma a ter em conta os graus de liberdade, superando o problema apresentado de  $R^2$ , surge um ajustamento ao mesmo.

O coeficiente de determinação ajustado, ou  $\overline{R^2}$ , é, portanto, uma melhoria do coeficiente de determinação, na medida em que é ajustado ao número de parâmetros do modelo. Assim, este é obtido pela **equação (40)**.

$$\overline{R^2} = 1 - (1 - R^2) \frac{n - 1}{n - k - 1} \quad (40)$$

onde

$n$  é o número de observações;

$k$  é o número de variáveis explicativas (preditores) no modelo.

Tendo como base esta medida para avaliação de desempenho, o melhor modelo é aquele que maximiza o coeficiente de determinação ajustado (Heizer et al., 2017; Makridakis et al., 1998).

### 3.4.8 Estatística de *Ljung-Box*

A estatística de *Ljung-Box*, obtida pela **equação (41)**, testa a existência de correlação nos erros de ajustamento do modelo.

$$Q = n(n + 2) \sum_{k=1}^h (n - k)^{-1} r_k^2 \quad (41)$$

onde

$n$  é o número de observações na série;

$h$  é a defasagem máxima considerada (com, normalmente,  $h \approx 20$ );

$k$  é o número de variáveis explicativas no modelo;

$r_k$  é o coeficiente de correlação de  $k$ .

A *Ljung-Box* é, normalmente, utilizada para modelos *ARIMA* ou modelos de regressão. Esta estatística tem uma distribuição aproximada à distribuição qui-quadrado com  $h - m$  graus de liberdade (onde  $m$  é o número de parâmetros estimado no modelo). Pelo que, após a obtenção o valor da estatística (Q), poder-se-á recorrer à tabela dos valores críticos para a distribuição qui-quadrado de forma a avaliar a significância do modelo. Neste seguimento, a hipótese nula para o teste *Ljung-Box* refere a não existência de correlação nos erros de ajustamento do modelo. Esta hipótese é rejeitada se o valor de Q for superior ao valor crítico da distribuição qui-quadrado para um determinado número de graus de liberdade  $h - m$  e um nível de significância padrão de 0,05, concluindo-se que os erros não são aleatórios e independentes ao longo do tempo (Hyndman & Athanasopoulos, 2018; Makridakis et al., 1998).

Deve-se ter em conta que as medidas de avaliação de desempenho das previsões descritas medem a qualidade do ajuste do modelo aos dados históricos, sendo que, esse ajuste pode não ser, necessariamente, indicativo de uma boa previsão. Ao haver um excessivo ajuste dos dados, o método de previsão pode estar a incluir componentes aleatórias nas previsões.

Além disso, métodos de previsão diferentes utilizam procedimentos diferentes na fase de ajuste. Havendo, por exemplo, métodos que, com o seu procedimento, minimizam o *MSE*, como é o caso dos métodos de regressão ou o *ARIMA*.

De forma a evitar estes problemas, é importante distinguir para esta avaliação, dois conjuntos de dados. Um conjunto de dados de “iniciação” e outro de dados “teste”, onde serão utilizados os dados de “iniciação” para prever os dados “teste” que não são utilizados no ajuste do modelo, mas sim para calcular as medidas de avaliação de desempenho das previsões. Por outro lado, torna-se crucial a comparação dos diferentes métodos de previsão baseado em várias medidas de avaliação de desempenho das previsões para uma abordagem mais completa desta avaliação (Makridakis et al., 1998).

## 4. CASO DE ESTUDO

A Linhas, S.A. comercializa milhares de artigos diferentes com comportamentos muito variáveis em termos de vendas e diversos graus de importância para a empresa. Com o objetivo de selecionar os produtos comercializados pela empresa a considerar no caso de estudo, numa perspetiva de diagnóstico da empresa, surge a necessidade de classificar os produtos segundo o seu grau importância, de forma a hierarquizar as suas análises e previsões. Perante isto, no presente capítulo realizar-se-á uma análise por famílias de produtos tendo como critério de análise as vendas anuais em quilogramas e, posteriormente, para as famílias consideradas de maior importância, passar-se-á para uma análise por subfamílias. Por fim, tendo por base as análises anteriormente realizadas, definir-se-á uma metodologia de previsão, que conta com dois procedimentos de previsão a aplicar às diferentes subfamílias.

### 4.1 Análise por Famílias de Produtos

Numa análise inicial, procedeu-se a um estudo da evolução das quantidades de quilogramas (kg) anuais vendidos, por família de produtos, nos últimos cinco anos, bem como das suas quotas anuais no total dos kg anualmente vendidos. Posteriormente, selecionou-se como método de classificação das famílias de produtos comercializados pela empresa, o método de análise ABC, tendo como critério de análise as quantidades vendidas anuais em quilogramas.

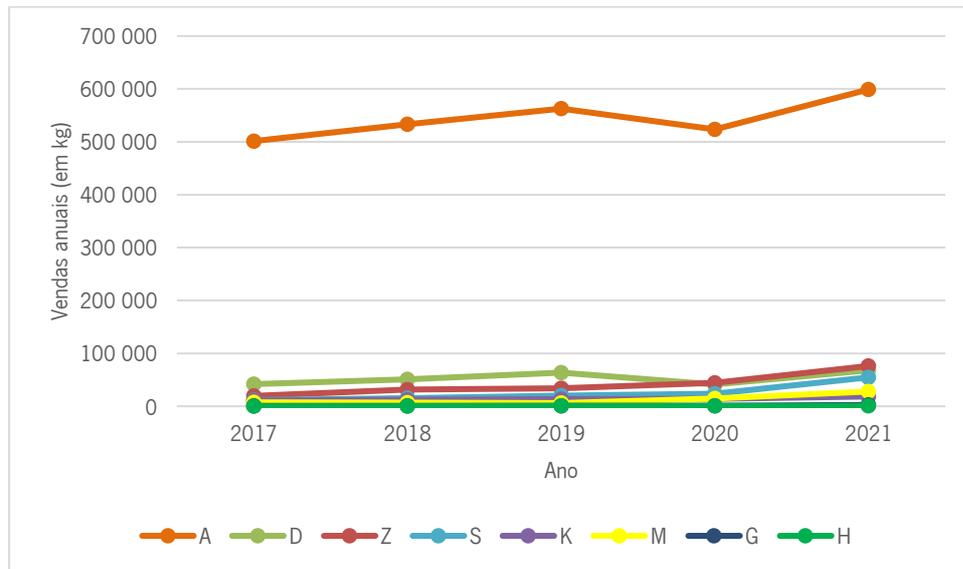
A pedido da Linhas, S.A., para as análises referidas consideram-se apenas os artigos que sofrem transformação internamente pela empresa, libertando-se, assim, para análise, os artigos das famílias B e O.

#### 4.1.1 Análise da Evolução das Vendas e das Quotas anuais por Família de produtos

Realizando-se uma análise da evolução das quantidades anuais vendidas (em kg) nos últimos cinco anos, isto é, de 2017 até 2021 (**Figura 5**), bem como das respetivas quotas anuais (**Figura 6**), é logo de destacar que, a família A é aquela que apresenta maior quota desde 2017.

**Figura 5**

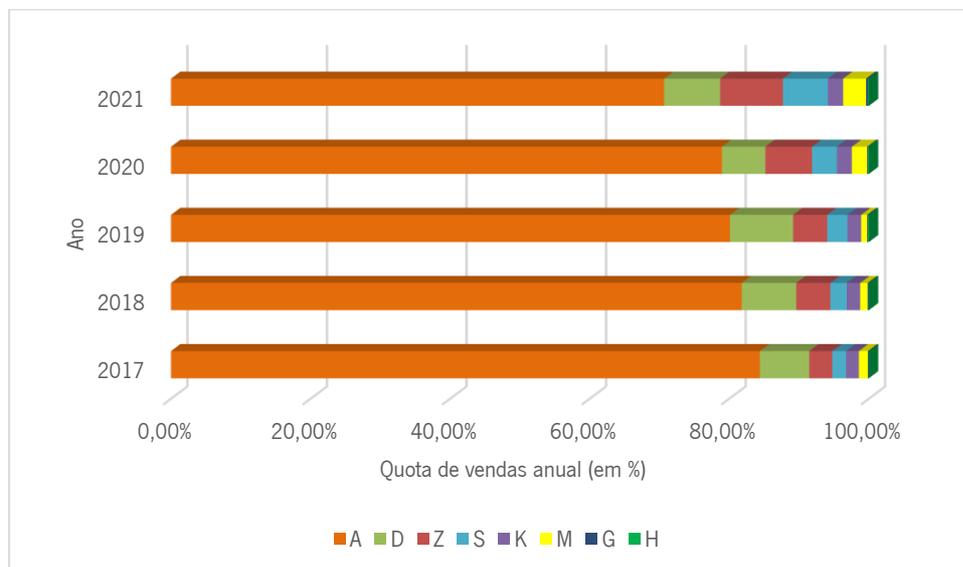
*Evolução das vendas anuais (em kg) nos últimos 5 anos para as oito famílias de produtos produzidas pela Linhas, S.A.*



*Nota.* Elaboração própria

**Figura 6**

*Evolução das quotas anuais nos últimos 5 anos para as oito famílias de produtos produzidas pela Linhas, S.A.*



*Nota.* Elaboração própria

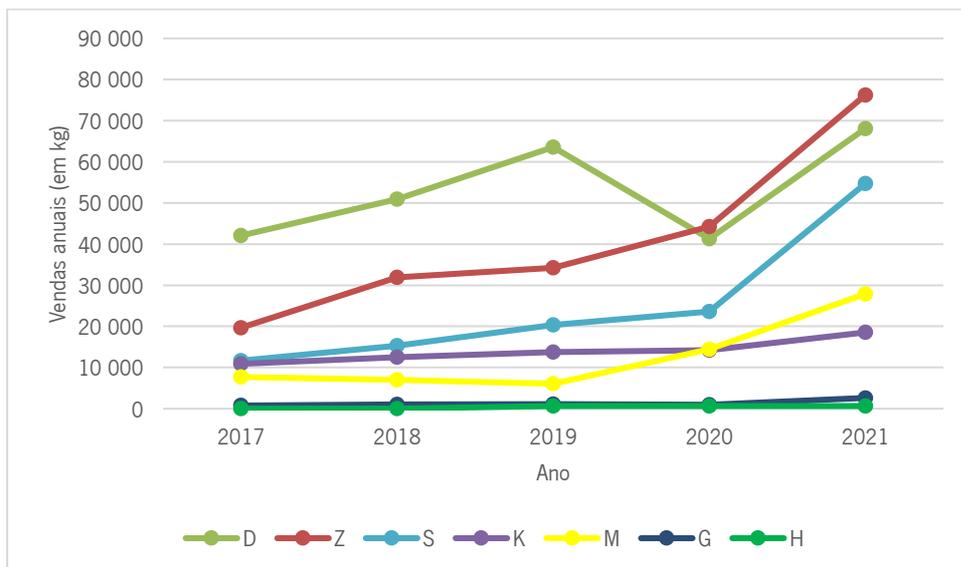
A família A apresenta, ao longo dos anos, na sua generalidade, uma evolução positiva das quantidades vendidas, à exceção de 2020 face a 2019, onde houve uma diminuição dos kg vendidos, que pode ser explicada pelo início da pandemia *Covid-19* que o mundo atravessara. Apesar desta diminuição, em 2021 volta-se a verificar um aumento dos kg vendidos, não só em relação ao ano anterior, mas também face a 2019. Embora se tenha verificado uma evolução positiva dos kg vendidos de, aproximadamente,

19,5% de 2021 face a 2017, verificou-se uma evolução negativa da quota anual da família A face aos kg vendidos para as oito famílias, passando de, aproximadamente, 84% dos quilos vendidos totais em 2017 para, aproximadamente, 71% em 2021. Esta evolução negativa da quota anual da família A, acompanhada pela evolução positiva das quantidades vendidas da mesma, pode ser explicada pela evolução positiva das quotas das restantes famílias de produtos.

Devido aos elevados valores apresentados pela família A, de forma a ser visualmente mais perceptível o comportamento das restantes sete famílias, aos gráficos apresentados nas **Figuras 5 e 6** foram retirados, visualmente, os dados relativos à família A, sendo desenvolvida a **Figura 7** onde é apresentada a evolução das quantidades vendidas anuais (em kg) nos últimos cinco anos das restantes sete famílias e a **Figura 8** representante da evolução das respetivas quotas anuais.

**Figura 7**

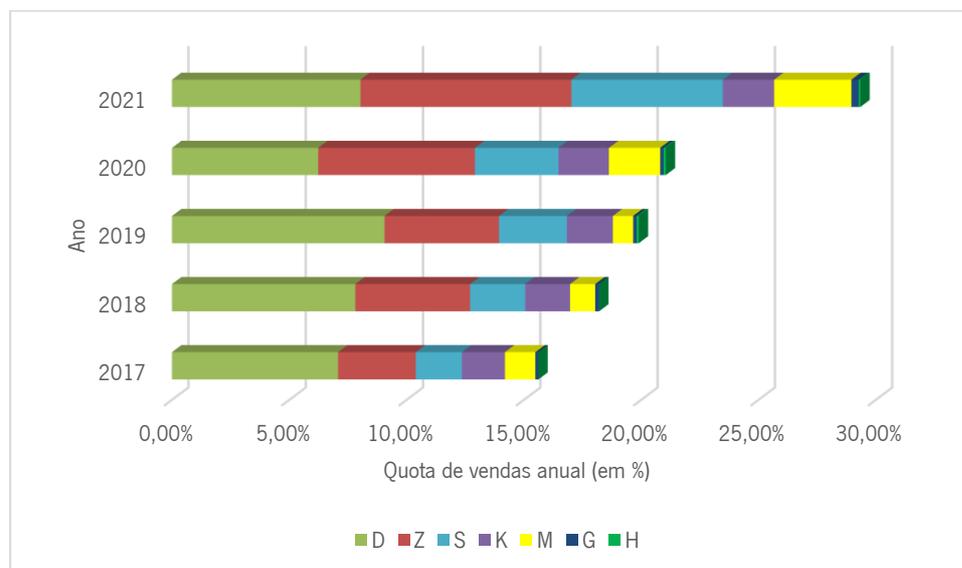
*Evolução das vendas anuais (em kg) nos últimos 5 anos para as famílias de produtos produzidas pela Linhas, S.A. (sem a família A)*



*Nota.* Elaboração própria

**Figura 8**

*Evolução das quotas anuais nos últimos 5 anos para as famílias de produtos produzidas pela Linhas, S.A. (sem a família A)*



*Nota.* O total dos kg vendidos incluem também os kg vendidos da família A, apesar de as mesmas não estarem ilustradas no gráfico, variando, por isso, a soma das quotas das sete famílias entre, aproximadamente, 15% e 30%. Elaboração própria.

Além da família A, é de destacar o comportamento das famílias D, Z e S. A família D, à exceção do ano 2020, apresenta uma evolução positiva generalizada das quantidades de kg anuais vendidas (com uma taxa de variação de cerca de 62% de 2021 face a 2017), bem como uma evolução positiva das quotas anuais (de, aproximadamente, 7% para 8%). Apesar disso, em 2020, a família D passa de segunda maior quota para terceira, trocando, assim, de posição com a família Z. Esta alteração de posição entre D e Z é justificada pelo crescimento exuberante nas quantidades vendidas que a família Z obteve (de uma taxa de variação de, aproximadamente, 287%, de 2021 face a 2017). Esta família de produtos desde 2017 até 2021 quase que triplicou a sua quota, passando a representar, aproximadamente, 9% do total dos kg produzidos e vendidos pela empresa em 2021, em contrapartida com os cerca de 3% de 2017. Já a família S, apesar de se ter mantido sempre na quarta posição de quota, apresentou ao longo dos anos, não só uma evolução positiva das quantidades vendidas (a uma taxa de variação de, aproximadamente, 369%, de 2021 face a 2017), mas também da sua quota anual (passando de, aproximadamente, 2% para 6,5% de 2017 para 2021), tendo-se verificado este crescimento, principalmente, no último ano. As restantes famílias, M, K, G e H, dadas as suas quotas baixas nos últimos cinco anos comparativamente com as famílias já analisadas, não serão alvo de estudo. No entanto, é também de destacar que a família M nos últimos dois anos tem vindo a crescer, podendo, no futuro, vir a ser considerada uma família para análise, caso mantenha essa tendência no comportamento. Para o caso em estudo, uma vez que os últimos dois anos são considerados atípicos devido à Covid-19 e dado que a família apresentou taxas de

variação negativas nos três anos anteriores, não se considera pertinente proceder ao estudo desta família.

#### 4.1.2 Análises ABC por Família de produtos

Após a análise do comportamento das oito famílias de produtos produzidos pela empresa Linhas, S.A., de forma a confirmar a importância das famílias selecionadas, completou-se a análise pela elaboração de classificações ABC por famílias de produtos.

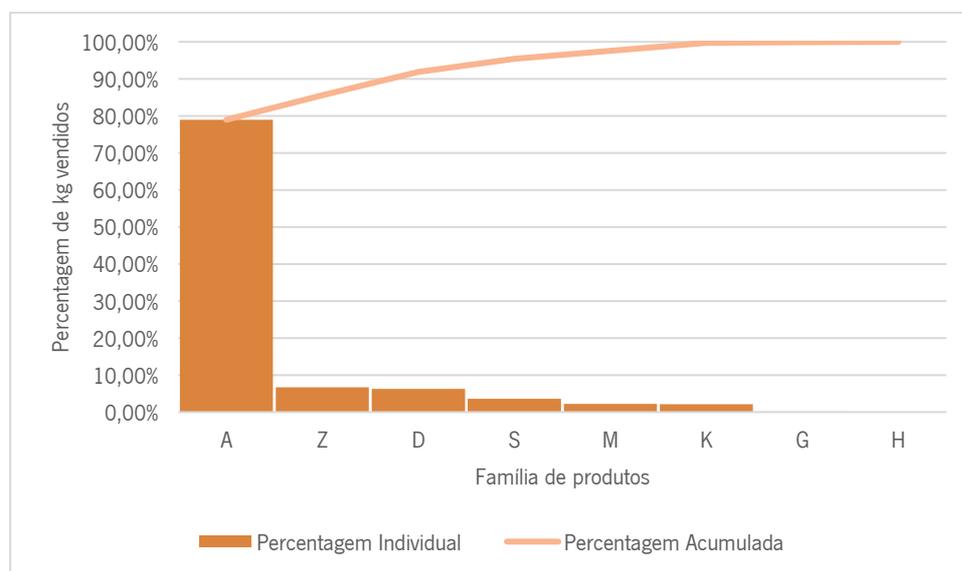
Por uma questão de confidencialidade, para o presente caso em estudo utilizou-se como critério para a classificação ABC, as quantidades de kg vendidos anuais corrigidos por um fator, em detrimento do clássico critério volume das vendas anuais.

Considerando os anos de 2020 e 2021 como anos atípicos, devido à pandemia Covid-19 que o mundo atravessava, considerou-se pertinente realizar duas análises ABC, uma para cada ano, de forma a perceber que famílias de produtos deveriam ser consideradas para o caso em estudo.

Após a realização da análise ABC para o ano 2020 (**Figura 9**), concluiu-se que apenas a família A pertencia ao grupo “A” (com, aproximadamente, 79% do total dos quilos vendidos em 2020), seguindo-se as famílias Z e D a pertencerem ao grupo “B” e as restantes ao “C”.

**Figura 9**

*Análise ABC, por família de produtos, em 2020*

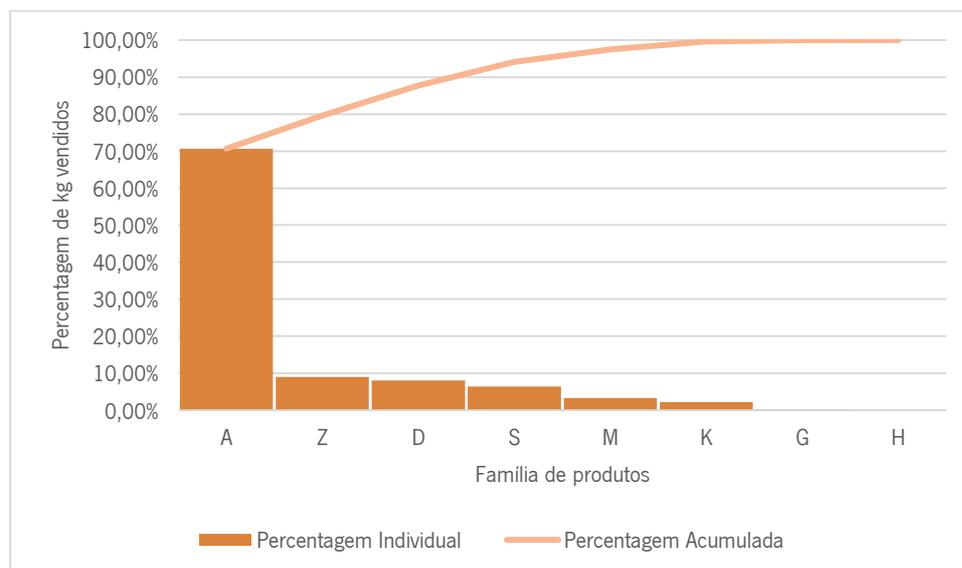


*Nota.* Elaboração própria

Em 2021, na análise ABC elaborada (**Figura 10**), a família Z passa a pertencer também ao grupo “A” após a A, e a família S passa a pertencer ao grupo “B” após a D.

**Figura 10**

Análise ABC, por família de produtos, em 2021



Nota. Elaboração própria

Finalizadas as análises às oito famílias de produtos produzidos pela empresa, é possível concluir quais as famílias de produtos merecem maior foco e quais as suas prioridades no presente estudo. A família A é aquela que maior prioridade deverá ter dada a sua quota elevada, a evolução positiva das quantidades vendidas e a sua classificação como “A” tanto em 2020, como em 2021. Seguidamente aos produtos da família A, pode-se apontar como prioridade para estudo os da família Z, dadas as suas classificações como “A” e “B” e a sua evolução positiva dos últimos anos. À família Z, seguem-se, com menos prioridade, a análise dos produtos da família D, posteriormente, da S, considerando-se as restantes famílias de importância reduzida, devido à sua evolução e classificações em “C”. Perante isto, o presente projeto debruçar-se-á sobre estas quatro famílias de produtos.

## 4.2 Análise por Subfamílias de Produtos

Os produtos comercializados pela Linhas, S.A. são divididos em famílias segundo a natureza e composição do fio, havendo em cada família uma subdivisão em subfamílias segundo as suas diferentes espessuras. Uma vez que a empresa faz as suas encomendas por subfamílias de produtos, torna-se vantajoso proceder a uma análise por subfamílias.

Procedeu-se a análises ABC para as subfamílias pertencentes às famílias A e Z, dada a importância destas famílias de produtos.

Após a análise ABC à família A para os anos 2020 e 2021 (**Apêndice A**), família anteriormente classificada como “A”, importa salientar que, aproximadamente, 93% dos quilos vendidos da família A

em 2020 e 2021 corresponderam à subfamília AK, classificando-se esta subfamília como “A”. Já as subfamílias AI, AM e AA (classificadas como “B”) representam, em conjunto, cerca de 5% do total dos kg vendidos da família A em 2020 e mais de 7% em 2021. De notar que, a família A foi classificada como “A” e apresenta valores de quota entre 84% e 71% sobre o total dos kg anuais vendidos desde 2017 até 2021. Desta forma, dever-se-á dar importância a essas três subfamílias, dado que estas têm um peso considerável no total dos kg anuais vendidos considerando todas as famílias de produtos.

Para a família Z, família classificada como “A” em 2021 e “B” em 2020, elaborou-se também análises ABC para 2020 e 2021 (**Apêndice B**). Tendo por base esta análise, é possível definir como subfamílias prioritárias para estudo as ZI e ZG uma vez que foram classificadas como “A” em ambos os anos, acrescentando-se da subfamília ZM que passa de “B” a “A” em 2021.

### **4.3 Metodologia de Previsão**

Tendo por base as análises ABC, por famílias e por subfamílias, anteriormente realizadas, definiu-se uma metodologia de previsão a ter em conta para as diferentes famílias e subfamílias. A metodologia de previsão foi baseada nas importâncias atribuídas pelas respetivas análises, definindo-se, assim, que para as subfamílias que merecem maior prioridade dada a sua importância para a empresa, dever-se-ia aplicar procedimentos de previsão individuais por subfamília. Por outro lado, para as subfamílias consideradas com importância moderada, dever-se-ia aplicar uma previsão agrupada por família tendo em consideração as quotas por subfamília, ou seja, numa primeira instância aplicar métodos de previsão aos valores totais por família de produtos e, após a escolha do método de previsão, definir os valores previstos de cada subfamília baseados nas quotas anuais por subfamília observadas no último ano civil (dado que se havia verificado estabilidades ou variações progressivas das mesmas, a verificar em **Apêndice C**). Por fim, para as famílias consideradas menos importantes não se aplicaria qualquer método de previsão no caso em estudo. A **Tabela 2** resume a aplicação da metodologia de previsão definida e, em participar, os procedimentos a aplicar para cada situação: previsão individual por subfamília ou previsão agrupada por família baseada em quotas.

**Tabela 2***Aplicação da metodologia de previsão*

<b>ABC por família 2021, 2020</b>	<b>ABC por subfamília 2021, 2020</b>	<b>Procedimento a aplicar</b>	<b>Família/ Subfamília</b>
A, A	A, A	Previsão individual por subfamília	AK
	B, B	Previsão individual por subfamília	AI, AM, AA
	C, C	Sem aplicação	Restantes subfamílias da família A
A, B	A, A	Previsão individual por subfamília	ZI, ZG
	A, B	Previsão individual por subfamília	ZM
	B, B	Previsão agrupada por família baseada em quotas	Restantes subfamílias de família Z
	B, C		
C, C			
B, B	-	Previsão agrupada por família baseada em quotas	Família D
B, C	-	Previsão agrupada por família baseada em quotas	Família S
C, C	-	Sem aplicação	Famílias M, K, G e H

*Nota.* Elaboração própria

É de destacar que se definiu para as restantes subfamílias pertencentes à família A a não aplicabilidade de métodos de previsão dado que, as vendas em conjunto destas subfamílias não só têm vindo apresentar grande irregularidade, como também têm vindo a ser muito reduzidas, chegando, em 2017, o conjunto destas subfamílias a apresentar vendas anuais de menos de cinco kg (**Apêndice D**). Dadas estas circunstâncias, considerou-se não ser justificável o esforço do estudo de aplicação de métodos de previsão elaborados. Caso contrário, pela classificação atribuída à família em questão, teria sido relevante considerar para as mesmas uma previsão agrupada por família baseada em quotas.

## 5. PREVISÃO DA PROCURA ATRAVÉS DE MÉTODOS ESTATÍSTICOS

O presente capítulo é destinado à previsão da procura com recurso a métodos estatísticos, em particular, modelos não causais, uma vez que não se detém dados suficientes para a aplicação de modelos causais, isto é, apenas temos ao nosso dispor para estudo as variáveis vendas (em kg) e tempo (quadrimestres). Para isso, o capítulo encontra-se dividido em quatro subcapítulos, sendo cada um destinado a uma família de produtos em estudo. De forma a chegar à escolha do método de previsão a implementar, para cada situação serão analisadas as componentes das séries temporais, aplicados diferentes modelos não causais em conformidade com as características da série e, por fim, será realizada uma avaliação de desempenho dos modelos de previsão estatísticos aplicados. Para esta avaliação, serão também considerados os valores obtidos pelo método de previsão utilizado atualmente pela Linhas, S.A..

Importa saber que a Linhas, S.A. aplica o mesmo método de previsão, independentemente do produto em questão. Neste seguimento, inicialmente, a empresa assume que o valor previsto mensal é igual ao valor observado do mês homólogo do ano anterior. Após esta previsão por mês, a empresa considera que o valor previsto para cada quadrimestre utilizado para fazer encomendas de matéria-prima, será igual à soma das previsões dos respetivos quatro meses que compõem o quadrimestre em questão. Neste sentido, o método de previsão aplicado pela Linhas, S.A. pode ser denominado por *sNaïve* com período sazonal quadrimestral.

Salienta-se ainda que, para a aplicação dos diferentes modelos de previsão, os dados diários das vendas por produto fornecidos pela empresa foram compilados por subfamílias de produtos e por quadrimestres. Esta compilação é realizada uma vez que, segundo a informação facultada pela Linhas, S.A., as encomendas das matérias-primas são efetuadas por subfamílias com uma periodicidade de três vezes ao ano. Além disso, acrescenta-se o facto de que, a previsão agregada dos produtos pode trazer vantagens no que diz respeito à sua precisão.

Para os modelos de previsão implementados, definiu-se um *holdout* de três, isto é, utilizaram-se os dados relativos ao último ano (últimos três quadrimestres) para teste. Posteriormente, de forma a aumentar a representatividade, aplicou-se um *holdout* de seis com o objetivo de confirmar as escolhas anteriormente realizadas.

### 5.1 Família A

Da família A, com o objetivo final de obtenção de previsões, selecionou-se para análise individual as séries temporais relativas às vendas, em kg, das subfamílias AK, AI, AM e AA. Para as restantes

subfamílias A, tal como anteriormente indicado, não foram testados métodos de previsão dada a ordem de grandeza das vendas e a sua irregularidade não justificarem os mesmos, podendo, assim, a empresa continuar a aplicar o método de previsão atualmente empregue, sNaïve.

Com recurso ao *Software Forecast Pro*, às diferentes séries temporais implementou-se o método de decomposição multiplicativo de forma a identificar as componentes das séries (estando os resultados apresentados na **Tabela 3**). Constatou-se que ambas as séries temporais apresentam fortemente uma componente ciclo-tendência, na ordem dos 80 e 90%, além de que, é também de notória a grande irregularidade observável nas séries, não havendo grandes sazonalidades observáveis. Dadas as semelhanças nas componentes das séries temporais, é de esperar que a escolha pelo melhor método de previsão a adotar, realizada posteriormente, seja igual para ambas as séries temporais.

**Tabela 3**

*Componentes das séries temporais das subfamílias A obtidas através do método de decomposição multiplicativo*

	<b>Componentes das séries temporais</b>		
<b>Subfamília</b>	Ciclo-tendência	Sazonalidade	Variações aleatórias
AK	83,70%	0,54%	15,76%
AI	93,90%	0,32%	5,78%
AM	87,17%	1,23%	11,61%
AA	85,19%	1,50%	13,31%

*Nota.* Elaboração própria (dados obtidos através do *Software Forecast Pro*)

Perante as componentes caracterizantes das séries, procedeu-se à implementação de diferentes modelos de alisamento exponencial (AE) para as diferentes séries temporais. Além disso, de forma a proceder a uma comparação na avaliação dos métodos, aplicou-se o método de previsão implementando atualmente pela Linhas, S.A., sNaïve.

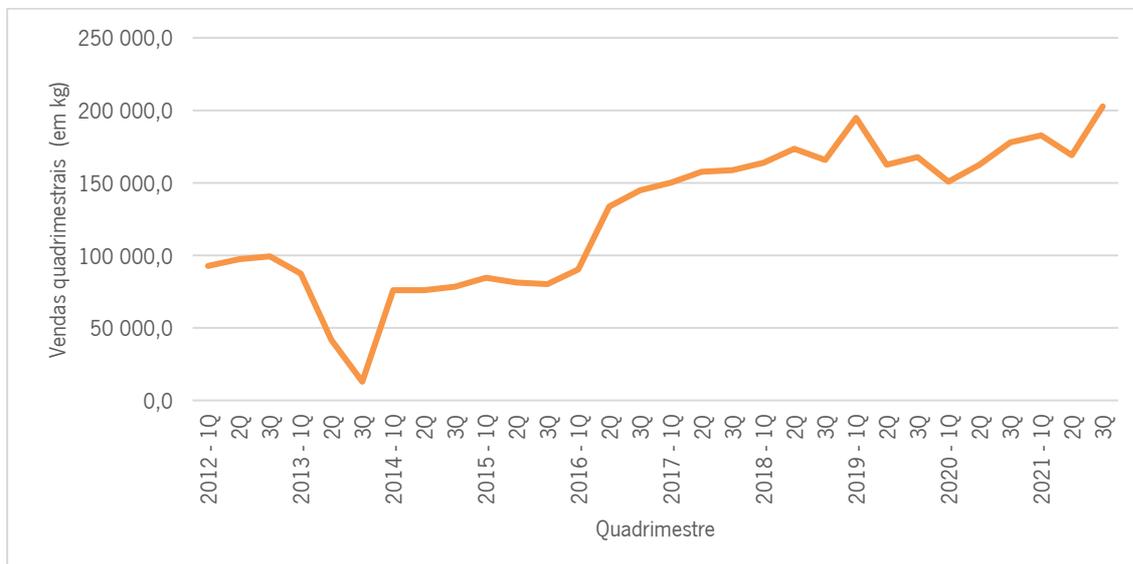
#### 5.1.1 Subfamília AK

A subfamília AK é aquela que maior importância tem para a empresa Linhas, S.A., dado que representa, em todos os anos em análise, cerca de 90% ou mais do total dos kg vendidos da família A. Para o estudo desta subfamília compilaram-se os dados da mesma com a subfamília AK(E), tratando o seu conjunto como uma só subfamília, a qual foi denominada por AK. Esta compilação deve-se ao facto destas subfamílias deterem da mesma matéria-prima, distinguindo-se entre si pelos seus processos de transformação.

A **Figura 11** representa a série temporal das vendas de AK por quadrimestre, em quilogramas, de 2012 a 2021.

**Figura 11**

*Vendas quadrimestrais de AK (em kg)*



*Nota.* Elaboração própria

Aplicaram-se diferentes modelos de AE para a série AK. Analisando as medidas de avaliação de desempenho dentro e fora da amostra, para um *holdout* de três e seis (**Tabela 4**), é de salientar que os modelos de AE pouco diferem entre si, tendo-se obtido medidas de avaliação de desempenho equivalentes. Já comparando estes modelos com o aplicado pela Linhas, S.A., existem grandes melhorias nas medidas de avaliação de desempenho ao aplicar modelos de AE, na ordem dos 50 e 60% numa perspetiva dentro da amostra e na ordem dos 20 e 50% numa perspetiva fora da amostra (**Apêndice E**). Apesar da proximidade dos resultados para os métodos de AE, destaca-se o comportamento do modelo *HW* aditivo dentro da amostra para ambos os *holdouts* e fora da amostra para um *holdout* de três. Já o modelo de *Holt* apresentou melhores resultados fora da amostra para um *holdout* de seis.

**Tabela 4**

Medidas de avaliação de desempenho dos métodos de previsão de AE e sNaive aplicados à série AK

	<b>Dentro da amostra</b>							
	<b>Holdout = 3</b>				<b>Holdout = 6</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	46 833	34 041	0,5589	-	48 817	36 042	0,6143	-
<b>AES</b>	21 140	14 270	0,2262	22 470	21 810	14 210	0,2422	23 300
<b>Holt</b>	21 100	14 030	0,2279	23 840	21 790	14 040	0,2435	24 870
<b>HW multiplicativo</b>	21 100	14 020	0,2263	25 330	21 560	13 740	0,2368	26 290
<b>HW aditivo</b>	21 030	13 890	0,2204	25 250	21 310	13 620	0,2271	25 990
	<b>Fora da amostra</b>							
	<b>Holdout = 3</b>				<b>Holdout = 6</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	23 570	21 015	0,1115	-	23 036	17 094	0,1030	-
<b>AES</b>	15 766	13 057	0,0681	-	17 664	13 948	0,0773	-
<b>Holt</b>	14 306	12 312	0,0498	-	15 487	12 753	0,0717	-
<b>HW multiplicativo</b>	14 388	12 220	0,0646	-	16 780	14 249	0,0823	-
<b>HW aditivo</b>	14 287	11 726	0,0621	-	16 398	13 591	0,0786	-

Nota. Elaboração própria (dados obtidos através do *Microsoft Excel* e *Software Forecast Pro*)

Perante estes resultados, pode-se apurar, dentro da classe de AE, o modelo *HW* aditivo como método a selecionar para prever os valores da série AK, dado este apresentar melhores resultados na generalidade. Paralelamente à aplicação dos modelos da classe de AE, dada a importância da subfamília em análise, procedeu-se ainda à implementação de modelos pertencentes à classe *ARIMA*.

De forma a aplicar o método *ARIMA*, primeiramente, para um *holdout* de três, selecionou-se o método automático do *Software Forecast Pro* para esta classe de modelos, o qual nos recomendou um *ARIMA* (0,1,0) (**Apêndice Figura F 1**). Repara-se, desde já, que o modelo reconhece a necessidade de remover a tendência da série quando procede a uma transformação simples,  $I(1)$ , adotando, assim, um grau de diferenciação 1.

Além da transformação simples realizada à série para eliminar a tendência, de forma a avaliar a estacionariedade da série na variância, definiu-se a hipótese nula de que a variância da segunda metade da série já com as transformações relativas à tendência era igual à da segunda metade da mesma série. Realizado o teste F para um nível de significância de 5%, constatou-se que o  $F = 4,77 > F \text{ crítico} = 1,95$ ,

pelo que se conclui que pode haver diferenças significativas entre as variâncias estudadas, rejeitando-se a hipótese nula de que as variâncias são iguais (**Apêndice Figura F 2**). Perante isto, partiu-se do modelo *ARIMA* (0,1,0) e procedeu-se a transformações adicionais à série para estabilizar a variância. Comparando este modelo sem transformações adicionais com os de transformações logarítmica e raiz quadrada, verificou-se que o modelo sem transformações é aquele que minimiza o *BIC*, adotando-se, assim, este modelo. Verificou-se ainda que, as medidas *MAPE*, *MAD*, *RMSE* se mantiveram iguais. Além disso, a estatística *Ljung-Box* (Q) vem afirmar a escolha do modelo, dado que, ao definir a hipótese nula de que não existe correlação entre os erros de ajustamento do modelo, se verificou que o valor calculado de  $Q=11,97$ , é inferior ao valor crítico da distribuição qui-quadrado para o nível de significância de 0,05 e 18 graus de liberdade, de 28,86. Neste seguimento, aceitou-se a hipótese nula definida, concluindo-se que os erros do modelo eram aleatórios e independentes entre si. Os resultados aqui descritos são apresentados na **Tabela 5**.

**Tabela 5**

*Medidas de avaliação de desempenho dentro da amostra do modelo ARIMA (0,1,0) com e sem transformações aplicados à série AK*

Modelo	Transformação	Medida de avaliação de desempenho dentro da amostra				
		<i>BIC</i>	<i>MAPE</i>	<i>MAD</i>	<i>RMSE</i>	<i>Ljung-Box</i> (18)
<i>ARIMA</i> (0,1,0)	Sem transformação	21 160	0,2191	14 120	21 160	11,97
	Transformação Logarítmica	48 080	0,2191	14 120	21 160	4,526
	Transformação Raiz Quadrada	28 930	0,2191	14 120	21 160	7,771

*Nota. Elaboração própria (dados obtidos através do Software Forecast Pro)*

De forma a descobrir qual o valor que deveriam tomar os restantes parâmetros, isto é, a ordem de processo autorregressivo e a de média móvel para as componentes simples, partiu-se do modelo *ARIMA* (0,1,0) sem transformação, procedendo-se à análise das funções de autocorrelações (*ACF*) e de autocorrelações parciais (*PACF*), apresentados nos **Apêndices Figura F 3** e **Figura F 4**, respetivamente. Numa perspetiva de análise da componente simples, pela leitura das *ACF* e das *PACF*, verificou-se que ambas estavam a diminuir exponencialmente, sendo esta análise compatível com o modelo *ARMA* (1,1), o que significa que se tomará como ordem de processo autorregressivo 1 e ordem de processo de média móvel 1, para a componente simples. Assim, o modelo foi modificado para *ARIMA* (1,1,1) sem transformação.

Como o *software* estimou os parâmetros, falta agora proceder à fase de teste do modelo de forma a validá-lo. O teste do modelo é dividido em três partes, são elas: análise da significância dos parâmetros, verificação dos erros do modelo e saturação do modelo.

a) Análise da significância dos parâmetros

O modelo identificado, *ARIMA* (1,1,1), tem 2 parâmetros: para a ordem do processo autorregressivo da componente simples ( $p$ ) e para a ordem do processo de média móvel da parte simples ( $q$ ). Verificou-se que para os parâmetros  $p$  e  $q$ , o valor absoluto da estatística  $t$  (de 0,38 e 0,47, respetivamente) eram claramente inferiores a 2,04 (para 30 graus de liberdade e 2,5% de significância), pelo que os coeficientes estimados para esses parâmetros do modelo não podem ser considerados significativamente diferentes de zero (**Apêndice Figura F 5**). Perante isto, o primeiro teste do modelo não foi superado. Contudo, o *software* sugere o modelo *ARIMA* (0,1,1), sendo este também testado.

O modelo sugerido, *ARIMA* (0,1,1), tem um parâmetro:  $q$ . Realizada a análise da sua significância, verificou-se que o valor absoluto da estatística  $t$  (de 0,2) é inferior a 2,04, pelo que, o coeficiente estimado não pode ser considerado significativamente diferente de zero, não superando o primeiro teste do modelo (**Apêndice Figura F 6**).

Perante isto, continuar-se-á a testar o modelo *ARIMA* (0,1,0), que não tem parâmetros.

b) Verificação dos erros do modelo

Para validar o modelo, pretende-se que os erros do modelo tenham duas características: sejam baixos e independentes entre si. Para confirmar estas características, procedeu-se à análise dos erros do modelo e da função de autocorrelação dos erros.

Pela análise dos erros cometidos para todos os períodos (**Apêndice Figura F 7**), verifica-se que, de uma maneira geral, os erros são relativamente baixos dada a ordem de grandeza da variável prevista, ultrapassando os limites admissíveis apenas em dois períodos.

Por outro lado, a função de autocorrelação dos erros (**Apêndice Figura F 8**), indica que as autocorrelações não são significativamente diferentes de zero, podendo, portanto, concluir-se que o modelo passa no teste de autocorrelação da independência dos erros. Assim, conclui-se que os erros do modelo são aleatórios, situando-se tanto acima como abaixo do valor observado de forma aleatória.

c) Saturação do modelo

Para obter a comprovação de que o modelo *ARIMA* (0,1,0) foi bem identificado, e que não excluiu nenhuma componente importante, foi necessário saturá-lo. Ao utilizar termos adicionais que o modelo não contemplava, obtiveram-se alguns modelos com termos adicionais que não passavam no primeiro

teste, o de análise da significância dos seus parâmetros, selecionando-se, assim, o modelo *ARIMA* (0,1,0) como modelo eleito da classe dos modelos *ARIMA*.

Após a seleção do modelo *ARIMA* (0,1,0), importa agora proceder a uma comparação entre os modelos estatísticos selecionados para cada classe analisada: Alisamento Exponencial e *ARIMA*. A **Tabela 6** compila as medidas de avaliação de desempenho dentro e fora da amostra para os modelos *HW* aditivo e *ARIMA* (0,1,0) aplicados à série temporal AK.

**Tabela 6**

*Medidas de avaliação de desempenho dos modelos HW aditivo e ARIMA (0,1,0) aplicados à série AK*

	Dentro da amostra							
	<i>Holdout = 3</i>				<i>Holdout = 6</i>			
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>
<b><i>HW</i> aditivo</b>	21 030	13 890	0,2204	-	21 310	13 620	0,2271	-
<b><i>ARIMA</i> (0,1,0)</b>	21 160	14 120	0,2191	-	21 820	14 060	0,2353	-
	Fora da amostra							
	<i>Holdout = 3</i>				<i>Holdout = 6</i>			
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>
<b><i>HW</i> aditivo</b>	14 287	11 726	0,0621	-	16 398	13 591	0,0786	-
<b><i>ARIMA</i> (0,1,0)</b>	15 445	12 825	0,0671	-	17 605	13 895	0,0708	-

*Nota.* Elaboração própria (dados obtidos através do *Microsoft Excel* e *Software Forecast Pro*)

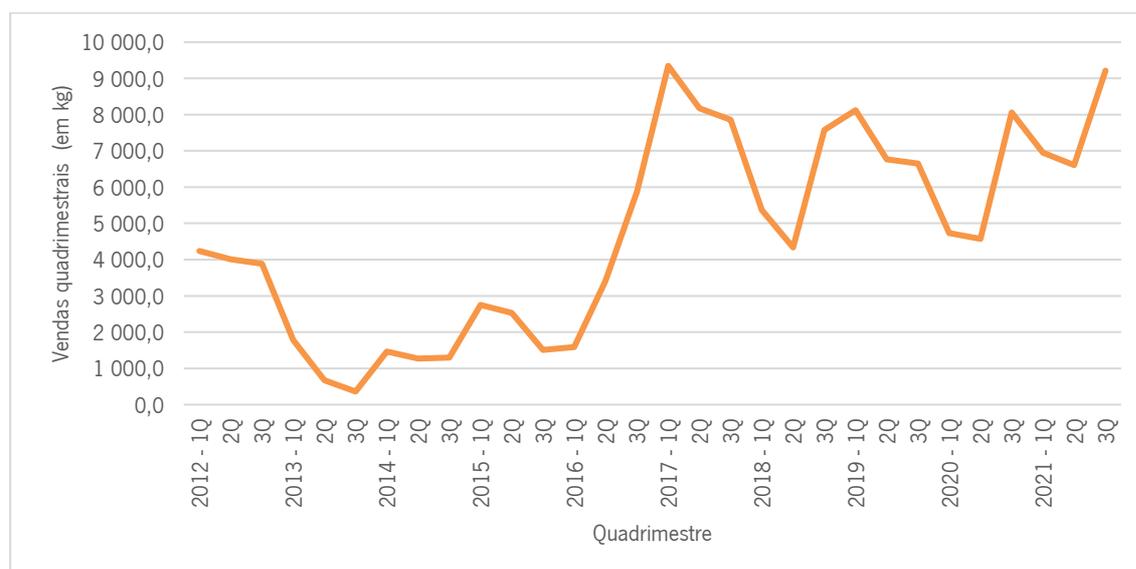
Pela análise dos valores apresentados na **Tabela 6**, foi possível constatar que, de uma forma geral, o desempenho do *ARIMA* (0,1,0) piorou os resultados obtidos em comparação ao modelo *HW* aditivo. Além disso, a classe de modelos *ARIMA* exigem procedimentos muito mais rebuscados em comparação com os modelos de AE, não sendo estes esforços justificados dado os resultados obtidos. Perante isto, e acrescentando-se o facto de o histórico detido ser reduzido, considerou-se dispensável a aplicação desta classe às restantes subfamílias futuramente estudadas.

### 5.1.2 Subfamília AI

A subfamília AI é a segunda subfamília com maior importância na família A. A série temporal correspondente à subfamília AI, como visto anteriormente no início do presente capítulo, apresenta uma forte tendência, sendo a mesma agora visível pela observação da **Figura 12**.

**Figura 12**

Vendas quadrimestrais de AI (em kg)



Nota. Elaboração própria

Para a série AI, aplicaram-se diversos modelos de AE, bem como o método adotado pela Linhas, S.A., tendo-se obtido as medidas de desempenho apresentadas pela **Tabela 7**.

**Tabela 7**

Medidas de avaliação de desempenho dos métodos de previsão aplicados à série AI

	Dentro da amostra							
	Holdout = 3				Holdout = 6			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	3 070	2 523	1,0694	-	3 137	2 548	1,1461	-
<b>AES</b>	1 598	1 169	0,3649	1 698	1 487	1 083	0,3742	1 589
<b>Holt</b>	1 600	1 172	0,3668	1 808	1 490	1 085	0,3756	1 700
<b>HW multiplicativo</b>	1 599	1 170	0,3658	1 920	1 489	1 087	0,3752	1 817
<b>HW aditivo</b>	1 562	1 151	0,3787	1 876	1 487	1 083	0,3742	1 814
	Fora da amostra							
	Holdout = 3				Holdout = 6			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	1 859	1 801	0,2505	-	2 208	2 106	0,3615	-
<b>AES</b>	1 249	1 238	0,1682	-	1 664	1 385	0,2269	-
<b>Holt</b>	1 240	1 230	0,1670	-	1 663	1 384	0,2269	-
<b>HW multiplicativo</b>	1 244	1 235	0,1675	-	1 650	1 376	0,2266	-
<b>HW aditivo</b>	1 047	1 044	0,1391	-	1 664	1 384	0,2268	-

Nota. Elaboração própria (dados obtidos através do Microsoft Excel e Software Forecast Pro)

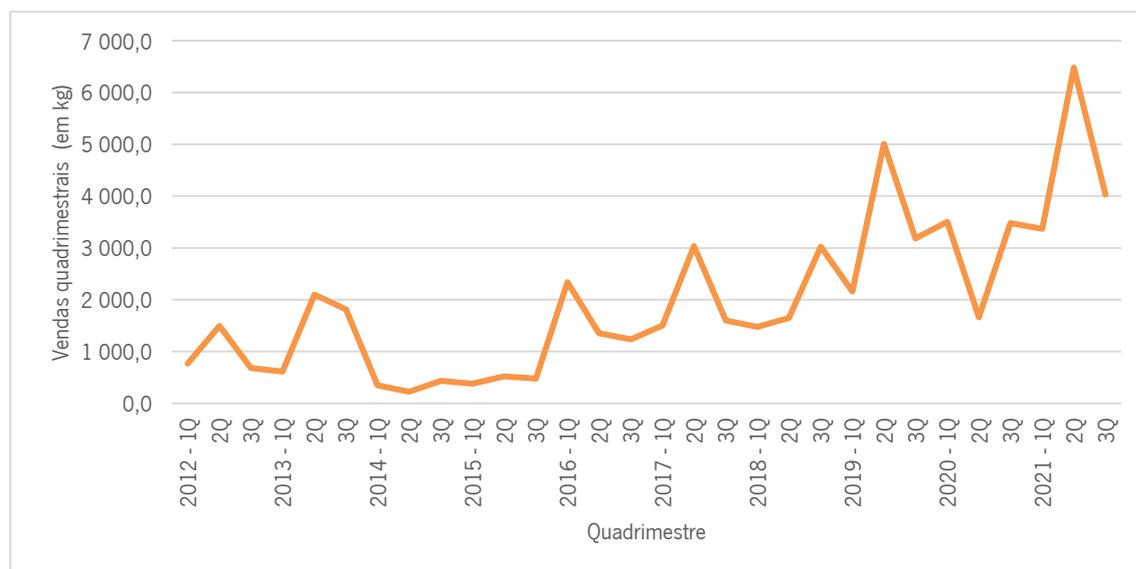
Primeiramente, é possível destacar a melhoria nas medidas de avaliação de desempenho quando aplicados modelos de AE, comparando com o desempenho do modelo sNaïve (em específico, com taxas de variação negativas das medidas entre os 25 e 67%, apresentadas em **Apêndice G**). Dos modelos de AE aplicados, à semelhança da série AK, pode-se destacar o desempenho do modelo *HW* aditivo para ambos os *holdouts*. Apesar de o modelo *HW* multiplicativo ter apresentado melhores resultados para um *holdout* de seis para fora da amostra, prefere-se o modelo aditivo, uma vez que, neste caso, este não se distancia do desempenho do multiplicativo, além de que apresentou melhores resultados que o modelo multiplicativo para as restantes situações. Por outro lado, por uma questão de estabilidade de resultados, considera-se que modelos de cariz aditivo são preferíveis aos multiplicativos, uma vez que os multiplicativos podem ser instáveis devido aos quocientes com denominadores de valor muito baixo. Perante os resultados obtidos, definiu-se como escolha de modelo de AE a aplicar, o *HW* aditivo.

### 5.1.3 Subfamília AM

A subfamília AM é a terceira subfamília com maior importância na família A. Observando a **Figura 13**, é possível confirmar de forma visual a tendência da série AM a que os dados têm vindo a ser sujeitos desde 2012 até 2021 e que no início do capítulo já fora analisada.

**Figura 13**

*Vendas quadrimestrais de AM (em kg)*



*Nota.* Elaboração própria

Pela aplicação dos modelos de AE, foi possível obter as medidas de avaliação de desempenho apresentadas pela **Tabela 8**.

**Tabela 8**

Medidas de avaliação de desempenho dos métodos de previsão aplicados à série AM

	<b>Dentro da amostra</b>							
	<b>Holdout = 3</b>				<b>Holdout = 6</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	1 335,9	1 007,1	0,9443	-	1 209	925,04	0,9586	-
<b>AES</b>	954,9	658,6	0,5940	1 015	938,6	631,0	0,6041	1 003
<b>Holt</b>	928,2	708,8	0,8017	1 049	915,9	654,3	0,7221	1 046
<b>HW multiplicativo</b>	928,2	708,5	0,8014	1 115	862,8	619,2	0,6922	1 052
<b>HW aditivo</b>	928,1	710,7	0,8067	1 115	891,6	621,5	0,6964	1 088
	<b>Fora da amostra</b>							
	<b>Holdout = 3</b>				<b>Holdout = 6</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	2 804	1 838	0,3073	-	2 466	1 708	0,5599	-
<b>AES</b>	2 122	1 646	0,3051	-	1 502	1 013	0,3007	-
<b>Holt</b>	1 928	1 401	0,2496	-	1 337	866	0,2819	-
<b>HW multiplicativo</b>	1 925	1 396	0,2486	-	1 299	868	0,3316	-
<b>HW aditivo</b>	1 919	1 397	0,2493	-	1 317	868	0,3049	-

Nota. Elaboração própria (dados obtidos através do Microsoft Excel e Software Forecast Pro)

Comparando as medidas de avaliação de desempenho dos modelos de AE aplicados, volta-se a constatar que os resultados relativos aos modelos de AE pouco diferem entre si, principalmente quando analisados os resultados fora da amostra. Além disso, verificam-se melhorias dos mesmos quando comparados com o método implementado atualmente pela Linhas, S.A. (as respectivas taxas de variação apresentam-se no **Apêndice H**). Apesar de o modelo AES ser aquele que apresenta, na generalidade, melhores resultados dentro da amostra, o mesmo não se verifica numa perspectiva fora da amostra, no qual é o modelo que apresenta piores resultados em todas as medidas de avaliação de desempenho comparativamente aos restantes modelos de AE.

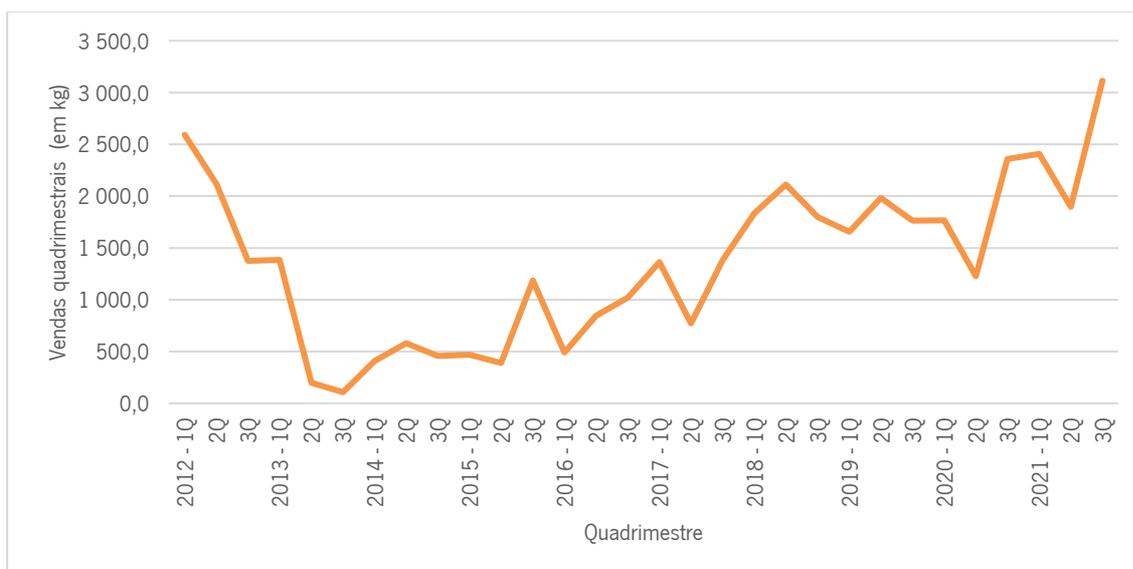
Dada a semelhança entre os resultados obtidos para os diferentes modelos, e não havendo nenhum modelo a destacar-se positivamente dos demais, parece prudente escolher, novamente, como modelo de AE a aplicar, o *HW* aditivo, mantendo-se assim, a escolha do modelo anteriormente eleito para as subfamílias anteriormente estudadas.

#### 5.1.4 Subfamília AA

A subfamília AA é a última subfamília da família A a considerar para análise. Pela análise da série temporal representante das vendas quadrimestrais em quilogramas de AA desde 2012, é notória a tendência representativa da série a partir de 2014, tal como observável na **Figura 14**.

**Figura 14**

*Vendas quadrimestrais de AA (em kg)*



*Nota.* Elaboração própria

Aplicado os modelos de AE e obtidas as medidas de avaliação de desempenho para os mesmos, bem como para o método aplicado pela Linhas, S.A. (**Tabela 9**), à semelhança das restantes subfamílias, é notória a semelhança entre os modelos de AE, bem como a melhoria das medidas de avaliação de desempenho dos modelos de AE em comparação com o método da Linhas, S.A. (Taxas de variação apresentadas no **Apêndice I**).

**Tabela 9**

Medidas de avaliação de desempenho dos métodos de previsão aplicados à série AA

	<b>Dentro da amostra</b>							
	<b>Holdout = 3</b>				<b>Holdout = 6</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	980,7	709,7	1,2754	-	1 021,3	737,6	1,3961	-
<b>AES</b>	484,2	383,1	0,6289	514,7	453,6	359,5	0,6335	484,7
<b>Holt</b>	484,2	383,0	0,6275	547,1	453,6	359,4	0,6327	517,8
<b>HW multiplicativo</b>	485,9	386,7	0,5662	583,5	455,4	396,0	0,6389	555,5
<b>HW aditivo</b>	484,3	386,4	0,5769	581,6	453,4	360,9	0,6288	553
	<b>Fora da amostra</b>							
	<b>Holdout = 3</b>				<b>Holdout = 6</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	690	688	0,2872	-	599	561	0,3004	-
<b>AES</b>	591	489	0,1830	-	685	536	0,2422	-
<b>Holt</b>	589	488	0,1827	-	685	536	0,2422	-
<b>HW multiplicativo</b>	483	375	0,1435	-	559	462	0,2282	-
<b>HW aditivo</b>	497	399	0,1547	-	683	537	0,2423	-

Nota. Elaboração própria (dados obtidos através do *Microsoft Excel* e *Software Forecast Pro*)

Perante as medidas de avaliação obtidas, uma vez mais, seleciona-se como modelo a considerar, o *HW* aditivo, com a ênfase de que, apesar de o multiplicativo apresentar valores um pouco melhores para fora de amostra, considera-se mais seguro, por uma questão de estabilidade e semelhança de resultados, selecionar o modelo do tipo aditivo.

#### 5.1.5 Conclusões da Família A

Considerando os resultados obtidos para cada uma das subfamílias da família A analisadas, em particular, AK, AI, AM e AA, conclui-se que, tal como esperado pela análise inicialmente realizada através da decomposição das séries, a escolha pelo modelo de AE foi a mesma, o *HW* aditivo. Estas escolhas deveram-se, sobretudo, à semelhança de resultados entre os diferentes modelos de AE, bem como a melhoria dos mesmos em comparação com o modelo adotado pela Linhas, S.A., sNaïve.

Importa agora definir os valores a considerar para as contantes de alisamento na aplicação do modelo *HW* aditivo. Perante isto, começou-se por compilar os valores ótimos definidos pelo *Software Forecast*

Pro para as constantes de alisamento do nível ( $\alpha$ ), da tendência ( $\beta$ ) e da sazonalidade ( $\gamma$ ), quando aplicado o modelo de *HW* aditivo com *holdout* de três e seis para cada série temporal. Os mesmos encontram-se na **Tabela 10**.

**Tabela 10**

*Constantes de alisamento otimizadas segundo série temporal*

Subfamília	Nível ( $\alpha$ )		Tendência ( $\beta$ )		Sazonalidade ( $\gamma$ )	
	Holdout = 3	Holdout = 6	Holdout = 3	Holdout = 6	Holdout = 3	Holdout = 6
<b>AK</b>	0,95425	0,96146	0,00930	0,00927	0,36756	0,99894
<b>AI</b>	0,96444	0,99926	0,00005	0,00005	0,98956	0,25476
<b>AM</b>	0,13828	0,32764	0,05935	0,04425	0,00381	0,08272
<b>AA</b>	0,94485	0,89592	0,00004	0,00004	0,58047	0,06599

*Nota.* Elaboração própria (dados obtidos através do *Software Forecast Pro*).

Após a compilação das diferentes constantes de alisamento ótimas obtidas pelo *software*, sugere-se uma escolha comum das mesmas para todas as séries dado o impacto das afinações destas constantes ser, normalmente, reduzido no que diz respeito aos resultados das medidas de avaliação de desempenho. Para esta definição, calculou-se uma média dos oito valores obtidos para cada contante de alisamento, sendo definido o valor da mesma baseada nos seguintes critérios:

- Valor definido = 0,1, para um valor médio baixo da constante, abaixo de 0,2;
- Valor definido = 0,3, para um valor médio intermédio da constante, entre 0,2 e 0,4;
- Valor definido = 0,7, para um valor médio alto da constante, acima de 0,4.

Calculados os valores médios, determinou-se os valores para as constantes de alisamento a aplicar ao modelo de alisamento exponencial *HW* aditivo aplicados às séries da família A. Os referidos valores encontram-se apresentados na **Tabela 11**.

**Tabela 11**

*Definição de constantes de alisamento para as séries de família A*

	Nível ( $\alpha$ )	Tendência ( $\beta$ )	Sazonalidade ( $\gamma$ )
<b>Valor médio</b>	0,8	0,02	0,4
<b>Valor definido</b>	0,7	0,1	0,3

*Nota.* Elaboração própria

## 5.2 Família Z

Para a família Z, selecionou-se para análise individual as séries temporais relativas às vendas em kg das subfamílias ZI, ZG e ZM, sendo que, para as restantes subfamílias Z, tal como mencionado anteriormente, considerou-se ser adequada uma análise agrupada por família baseada em quotas.

Numa primeira instância, de forma a identificar as componentes caracterizantes das diferentes séries temporais consideradas, implementou-se o método de decomposição multiplicativo através do *Software Forecast Pro*, estando os resultados do mesmo apresentados na **Tabela 12**.

**Tabela 12**

*Componentes das séries temporais das subfamílias Z obtidas através do método de decomposição multiplicativo*

Subfamília	Componentes das séries temporais		
	Ciclo-tendência	Sazonalidade	Variações aleatórias
ZI	72,99%	9,14%	17,88%
ZG	28,50%	45,34%	26,16%
ZM	60,34%	2,93%	36,73%
Conjunto das restantes subfamílias Z	84,65%	1,28%	14,07%

*Nota. Elaboração própria (dados obtidos através do Software Forecast Pro)*

Contrariamente ao que se havia verificado para as séries da família A, as componentes caracterizantes das séries da família Z apresentam algumas diferenças entre si, o que se poderá refletir aquando da escolha do método de previsão adequado para cada série temporal, esperando-se que a escolha venha a diferir segundo a série em questão. Ao passo que, por exemplo, a subfamília ZI apresenta uma forte componente ciclo-tendência, acima dos 70%, e uma componente sazonal abaixo de 10%, a subfamília ZG já apresenta valores mais reduzidos no que diz respeito à componente ciclo-tendência, na ordem dos 30% e uma forte componente sazonal acima dos 45%. Contudo, é de destacar a componente irregular em ambas as séries que apresenta valores entre cerca de 14% e 37%.

As características das séries das subfamílias de Z levam a que seja selecionada como classe de modelos não causais a aplicar, a de AE.

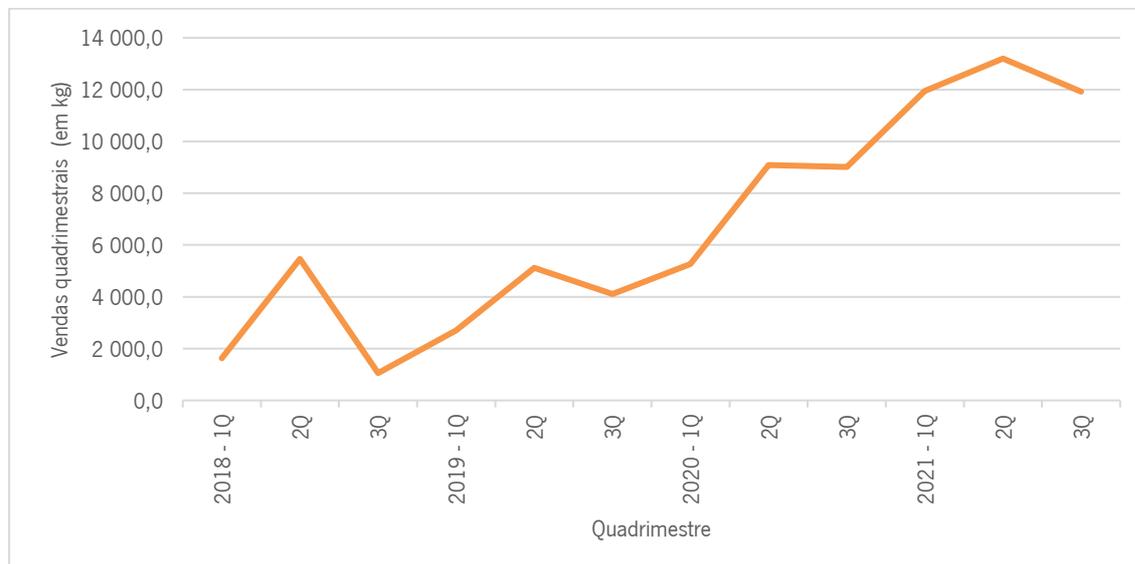
### 5.2.1 Subfamília ZI

Para a subfamília ZI considerou-se os dados a partir de 2018, sendo desta forma a série ZI composta pelas vendas dos kg vendidos desde o primeiro quadrimestre de 2018 até ao terceiro de 2021 (**Figura**

**15).** Esta consideração deveu-se ao facto da Linhas, S.A. ter entrado, em 2018, num novo mercado, o mercado de acolchoados, onde a empresa já se havia preparado para expandir esta subfamília de artigos.

**Figura 15**

*Vendas quadrimestrais de ZI (em kg)*



*Nota.* Elaboração própria

Pela análise da série ZI, tal como conhecido anteriormente, é possível destacar a forte tendência positiva das vendas de kg quadrimestrais da subfamília ZI desde 2018.

Aplicaram-se os modelos de AE à série ZI, considerando apenas um *holdout* de três dado o tamanho da série em questão. Pela aplicação dos modelos de AE e do método aplicado pela Linhas, S.A., obtiveram-se as medidas de avaliação de desempenho, dentro e fora da amostra, apresentadas na

**Tabela 13.**

**Tabela 13**

*Medidas de avaliação de desempenho dos métodos de previsão aplicados à série ZI*

<b>Dentro da amostra</b>				
	<b>Holdout = 3</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	3 173	2 673	0,6305	-
<b>AES</b>	2 387	1 951	0,6332	2 697
<b>Holt</b>	2 048	1 722	0,6787	2 614
<b>HW multiplicativo</b>	1 640	1 243	0,4125	2 366
<b>HW aditivo</b>	1 109	777	0,1998	1 600

<b>Fora da amostra</b>				
	<b>Holdout = 3</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	4 832	4 569	0,3717	-
<b>AES</b>	3 970	3 925	0,3161	-
<b>Holt</b>	2 991	2 852	0,2290	-
<b>HW multiplicativo</b>	2 872	2 328	0,1836	-
<b>HW aditivo</b>	1 808	1 739	0,1427	-

*Nota.* Os modelos implementados consideram apenas dados a partir de 2018. Elaboração própria (dados obtidos através do *Microsoft Excel* e *Software Forecast Pro*)

Pela análise dos resultados obtidos parece evidente a escolha do melhor modelo de AE, o *HW* aditivo, uma vez que é aquele que apresenta menores valores nas medidas de avaliação de desempenho tanto para dentro, como para fora da amostra, destacando-se pelos seus valores em relação aos restantes. Além disso, comparando os resultados do modelo *HW* aditivo com o método implementado atualmente pela Linhas, S.A., verifica-se que o modelo eleito apresenta uma diminuição em todas as medidas de desempenho calculadas, a taxas de variação entre 65 e 71% dentro da amostra e 62 e 63% fora de amostra (**Apêndice J**).

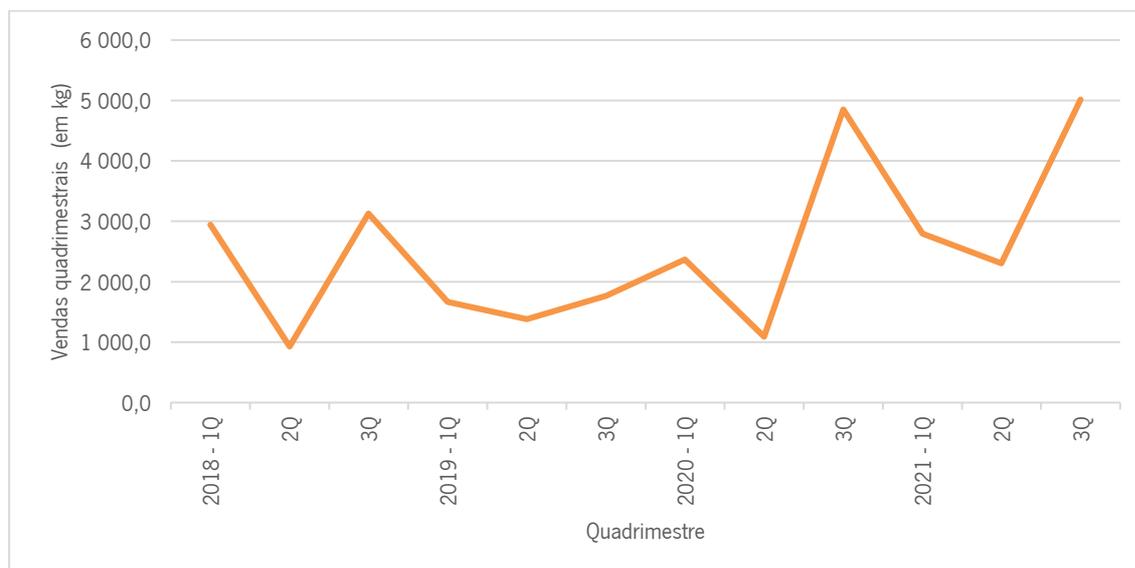
### 5.2.2 Subfamília ZG

Para o estudo da série ZG compilaram-se os dados das subfamílias ZG com os da ZG(E), tratando o seu conjunto como uma só subfamília, a qual foi denominada por ZG, pelo facto de ambos deterem da mesma matéria-prima.

À semelhança da subfamília ZI, a subfamília ZG entrou em 2018 no mercado de acolchoados. Perante isto, apesar da subfamília ZG apresentar vendas desde há muitos anos, considerou-se, igualmente, para a série ZG, os dados a partir de 2018 até 2021, inclusive. A série ZG, apresentada na **Figura 16**, apresenta uma tendência positiva menos evidente em comparação com as séries anteriormente analisadas, verificando-se uma componente ciclo-tendência na ordem dos 28,50%. Já a sazonalidade, com valores em mais de 45%, é notória pela observação da série, com valores a diminuir no segundo quadrimestre (valores de componentes obtidas através do *software Forecast Pro*).

**Figura 16**

Vendas quadrimestrais de ZG (em kg)



Nota. Elaboração própria

Aplicaram-se os modelos de AE apenas para um *holdout* de três dada a pequena dimensão da série uma vez que apenas foram considerados para estudo os dados a partir do ano de 2018. Para estes modelos e para o método implementado pela Linhas, S.A., foram obtidas medidas de avaliação de desempenho apresentadas na **Tabela 14**.

**Tabela 14**

Medidas de avaliação de desempenho dos métodos de previsão aplicados à série ZG

<b>Dentro da amostra</b>				
	<b>Holdout = 3</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	1 918	1 574	0,6737	-
<b>AES</b>	1 295	1 030	0,5856	1 463
<b>Holt</b>	1 284	1 062	0,6194	1 640
<b>HW multiplicativo</b>	1 005	710,2	0,3545	1 449
<b>HW aditivo</b>	994,2	707,8	0,3552	1 434

<b>Fora da amostra</b>				
	<b>Holdout = 3</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	749	602	0,2375	-
<b>AES</b>	1 476	1 008	0,2315	-
<b>Holt</b>	1 282	907	0,2225	-
<b>HW multiplicativo</b>	932	806	0,2399	-
<b>HW aditivo</b>	887	720	0,2019	-

*Nota.* Os modelos implementados consideram apenas dados a partir de 2018. Elaboração própria (dados obtidos através do *Microsoft Excel* e *Software Forecast Pro*)

As medidas de avaliação de desempenho para dentro da amostra, mostram que o modelo de AE com melhor desempenho é o *HW* aditivo (acompanhando-se, logo a seguir, pelo *HW* multiplicativo), apresentando melhorias nas medidas de avaliação de desempenho com taxas de variação entre os 47% e 55%, comparativamente ao método implementado pela Linhas, S.A..

Numa perspectiva fora da amostra, o método da Linhas, S.A. apresenta melhores resultados quando considerado como critério o *RMSE* ou o *MAD* em comparação com qualquer modelo de AE aplicado. Ainda assim, o modelo *HW* aditivo é o modelo de AE que melhor se comporta depois do modelo aplicado pela Linhas, S.A. quando analisadas *RMSE* e *MAD* (apresentando taxas de variação na ordem dos 18 e 20%, respetivamente, em comparação com o método da Linhas, S.A., em contraste com as taxas entre 24 e 97% dos restantes modelos de AE). Além disso, quando considerado o *MAPE*, o *HW* aditivo é o modelo que melhor desempenho apresenta em relação a qualquer outro modelo, comportando-se também melhor que o *sNaive* (com taxa de variação de -15% em comparação com o mesmo).

Neste sentido, uma vez que apenas é possível considerar um *holdout* reduzido para análise dada a falta de dados históricos, fará sentido proceder também a uma análise das componentes anteriormente determinadas, para melhor escolha do modelo a eleger. A componente sazonal é aquela que mais caracteriza a série em questão, componente essa que ambos os modelos (*sNaive* e *HW* aditivo) têm em consideração para a previsão. Além disso, a série apresenta valores tanto da componente ciclo-tendência, como da componente irregular, entre 25 e 30%.

Perante a análise dos resultados obtidos, bem como das componentes caracterizantes da série, parece prudente manter a eleição do modelo inalterável em relação às restantes escolhas, mantendo-se como escolha o modelo *HW* aditivo.

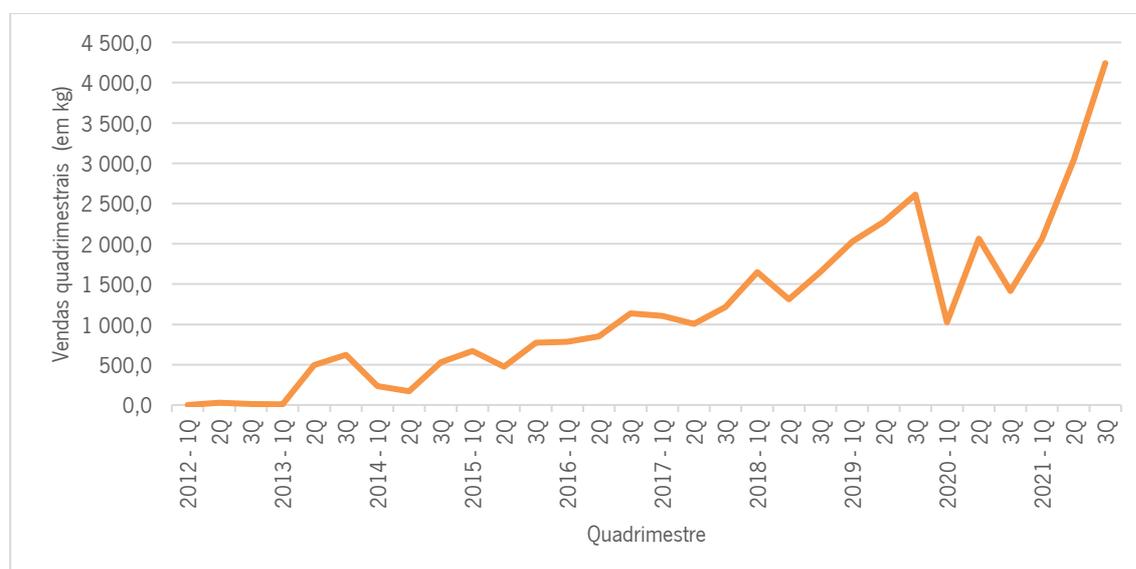
De notar que as taxas de variação referidas, relativas às medidas de avaliação de desempenho que comparam modelos de AE com o método sNaïve, encontram-se apresentadas no **Apêndice K**.

### 5.2.3 Subfamília ZM

A subfamília ZM é a terceira subfamília mais importante para a família Z. Analisado, através da **Figura 17**, as vendas quadrimestrais dos quilogramas vendidos desde 2012 até 2021, é notória a componente ciclo-tendência, de cerca de 60%, referida aquando da decomposição multiplicativa da série.

**Figura 17**

*Vendas quadrimestrais de ZM (em kg)*



*Nota.* Elaboração própria

Aplicando os modelos de AE à série das vendas de ZM desde 2012, obtiveram-se valores de *MAPE* dentro da amostra elevados, superiores a 100% (**Apêndice L**). Estes valores deveram-se ao facto de, nos primeiros anos, as vendas serem de zero ou próximas de zero, pelo que, considerou-se benéfico proceder a uma implementação e avaliação dos modelos considerado dados apenas a partir do ano 2014, onde as vendas quadrimestrais já apresentavam valores acima de 100 kg.

Após a implementação dos modelos de AE, bem como ao método implementado pela Linhas, S.A., para a série ZM com dados desde 2014, foi possível obter as medidas de avaliação de desempenho apresentadas na **Tabela 15**.

**Tabela 15**

Medidas de avaliação de desempenho dos métodos de previsão aplicados à série ZM

	Dentro da amostra							
	Holdout = 3				Holdout = 6			
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>
<b>Método Empresa</b>	541,9	443,8	0,4645	-	452,63	384,0	0,4350	-
<b>AES</b>	425,1	322,7	0,3053	457,0	252,4	212,4	0,2146	273,5
<b>Holt</b>	396,6	261,4	0,3216	458,5	194,8	164,5	0,2318	228,7
<b>HW multiplicativo</b>	396,7	261,3	0,3220	493,0	179,0	138,8	0,2030	227,8
<b>HW aditivo</b>	396,6	261,2	0,3211	493,0	168,6	130,7	0,139	214,6
	Fora da amostra							
	Holdout = 3				Holdout = 6			
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>
<b>Método Empresa</b>	1 828	1 614	0,4970	-	995	937	0,5992	-
<b>AES</b>	1 725	1 476	0,4263	-	1 110	991	0,5753	-
<b>Holt</b>	1 256	968	0,2585	-	1 542	1 428	0,8710	-
<b>HW multiplicativo</b>	1 254	967	0,2584	-	1 587	1 438	0,8949	-
<b>HW aditivo</b>	1 256	968	0,2585	-	1 201	1 059	0,6899	-

Nota. Os modelos implementados consideram apenas dados a partir de 2014. Elaboração própria (dados obtidos através do *Microsoft Excel* e *Software Forecast Pro*)

Pela análise das medidas de avaliação de desempenho, dentro da amostra, verifica-se que os modelos apresentam comportamentos muito semelhantes, principalmente o *Holt* e os modelos *HW*, apesar de o modelo *HW* aditivo ser aquele que apresenta melhor comportamento. Além disso, constata-se que os modelos de AE apresentam todos melhores resultados das medidas de avaliação de desempenho em comparação com o método aplicado atualmente pela Linhas, S.A., com taxas de variação entre os -22% e -68% (**Apêndice M**). Embora estes comportamentos se verificarem também para fora da amostra quando o *holdout* é de três, o mesmo não se verifica para uma análise fora da amostra com *holdout* de seis.

Para fora da amostra com um *holdout* de seis, o método implementado pela empresa apresenta, na generalidade, melhores resultados comparativamente aos modelos de AE considerando o *RMSE*, *MAD* e *MAPE*, seguindo-se o método AES com os melhores resultados e, posteriormente, o *HW* aditivo (que apresenta taxas de variação, em comparação com o método da Linhas, S.A., entre os 13 e 21%).

Face a estas análises para dentro e fora da amostra, e acrescentando-se o facto de a série apresentar, conforme verificado pela aplicação do método de decomposição multiplicativo, uma componente ciclo-tendência na ordem dos 60%, quando considerados valores desde 2012, e mais de 90%, quando considerados valores apenas desde 2014 (**Apêndice N**), seleciona-se como modelo a implementar o *HW* aditivo.

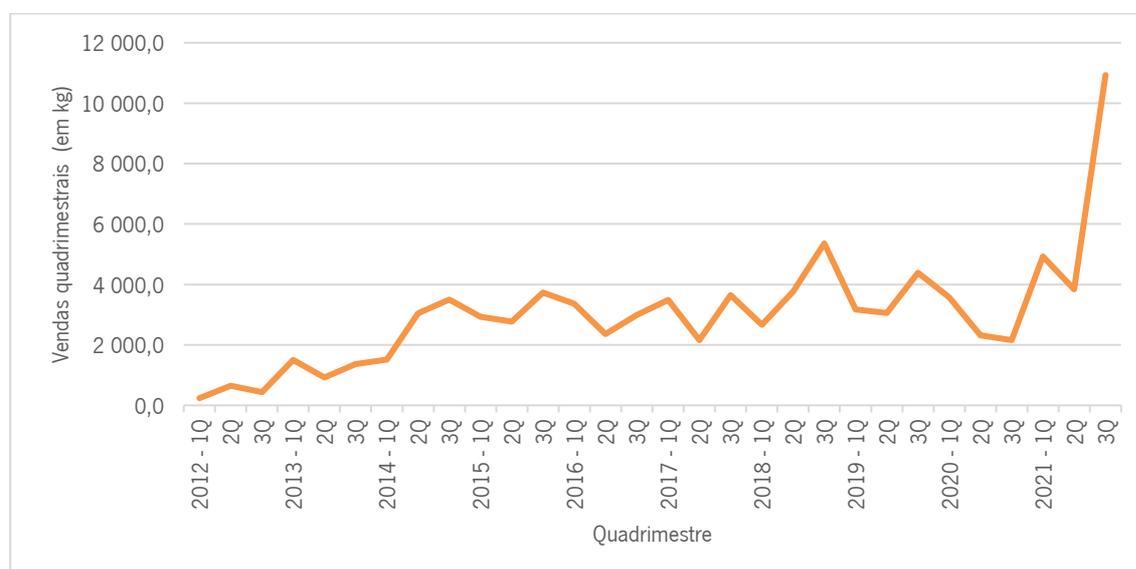
#### 5.2.4 Conjunto das restantes subfamílias Z

Finalizadas as análises das séries das subfamílias Z mais importantes para a Linhas, S.A. e eleitos os modelos quantitativos não causais para cada uma das situações individuais, importa proceder à análise da série temporal do conjunto de dados das vendas (em quilogramas) das restantes subfamílias Z sem os dados referentes às subfamílias ZI, ZG e ZM.

Tal como referido anteriormente, no início do subcapítulo, a série que agrupa as restantes subfamílias de Z apresenta uma componente ciclo-tendência de cerca de 85%, bem como uma irregularidade na ordem dos 14%, visualmente perceptíveis aquando da observação da **Figura 18**.

**Figura 18**

*Vendas quadrimestrais do conjunto das restantes subfamílias de Z (em kg)*



*Nota.* Elaboração própria

Após a implementação de modelos de AE e do método aplicado atualmente pela Linhas, S.A., foi possível obter as medidas de avaliação de desempenho apresentadas pela **Tabela 16**.

**Tabela 16**

Medidas de avaliação de desempenho dos métodos de previsão aplicados à série do conjunto das restantes subfamílias de Z

	<b>Dentro da amostra</b>							
	<b>Holdout = 3</b>				<b>Holdout = 6</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	1 042	824,2	0,3987	-	993,2	786,6	0,3878	-
<b>AES</b>	881,5	710,5	0,3067	937,0	869,9	691,5	0,3035	929,5
<b>Holt</b>	880,7	723,7	0,3446	995,0	816,1	638,2	0,3384	931,6
<b>HW multiplicativo</b>	825,6	668,1	0,3038	991,5	761,2	598,4	0,3155	928,5
<b>HW aditivo</b>	830,0	665,4	0,3147	996,7	763,7	595,7	0,3203	931,5
	<b>Fora da amostra</b>							
	<b>Holdout = 3</b>				<b>Holdout = 6</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	5 197	3 880	0,4905	-	1 566	1 410	0,5256	-
<b>AES</b>	5 094	4 028	0,5304	-	3 059	1 958	0,4059	-
<b>Holt</b>	5 086	3 972	0,5166	-	2 815	2 032	0,4892	-
<b>HW multiplicativo</b>	4 987	4 032	0,5438	-	2 696	1 890	0,4525	-
<b>HW aditivo</b>	5 114	4 232	0,5855	-	2 710	1 905	0,4562	-

Nota. Elaboração própria (dados obtidos através do *Microsoft Excel e Software Forecast Pro*)

Numa perspetiva dentro da amostra, verificou-se que os modelos de AE apresentaram comportamentos idênticos entre eles e melhores que o modelo sNaïve, com taxas de variação entre os 13 e os 24%. Já numa perspetiva fora da amostra, constatou-se que, na generalidade, o modelo sNaïve, modelo aplicado atualmente pela Linhas, S.A., foi aquele que apresentou melhores resultados em comparação com os modelos de AE. Uma vez que se deve preferir modelos melhores comportados fora de amostra, em detrimento de dentro da amostra, acrescendo-se o facto de dentro da amostra o modelo sNaïve não se afastar em muito dos resultados dos modelos de AE, considera-se não ser necessário alterar o modelo de previsão atualmente empregue pela Linhas, S.A..

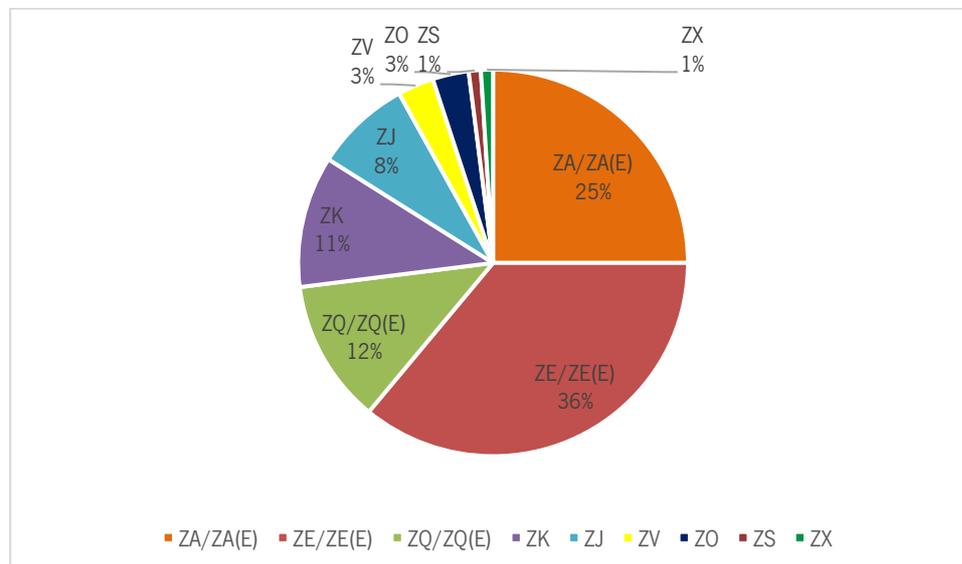
As taxas de variação relativas às medidas de avaliação de desempenho, que comparam os modelos de AE com o modelo sNaïve, encontram-se apresentadas no **Apêndice O**.

Definido o método de previsão a aplicar para a série do conjunto das restantes subfamílias de Z, importa considerar as quotas para cada subfamília Z de forma a proceder a uma previsão individual das mesmas. Para isso, consideram-se as quotas anuais apresentadas no ano civil anterior ao ano previsto. Para as

previsões do ano de 2022, considera-se as quotas anuais de 2021, apresentadas na **Figura 19**, considerando-se ser necessária as suas atualizações a cada fim de ano civil.

**Figura 19**

*Quotas anuais de 2021 das restantes subfamílias Z a aplicar em previsões 2022*



*Nota.* Elaboração própria.

### 5.2.5 Conclusões da Família Z

Tendo como fundamento os resultados obtidos para cada uma das subfamílias Z analisadas, nomeadamente, ZI, ZG, ZM e o conjunto das restantes subfamílias, selecionaram-se dois modelos de previsão diferentes, *HW* aditivo e *sNaïve*, sendo esta predileção apresentada na **Tabela 17**.

**Tabela 17**

*Métodos de previsão estatísticos selecionados para a família Z*

Subfamília	Método de Previsão Estatístico
ZI	<i>HW</i> Aditivo
ZG	<i>HW</i> Aditivo
ZM	<i>HW</i> Aditivo
Restantes subfamílias	Previsão agrupada através <i>sNaïve</i> e posterior previsão individual através de ponderação de quotas

*Nota.* Elaboração própria.

Torna-se necessário definir os valores a considerar para as constantes de alisamento na aplicação do modelo *HW* aditivo para as séries das subfamílias ZI, ZG e ZM. Para esta definição, utilizou-se o mesmo procedimento anteriormente utilizado na família A, estando na **Tabela 18** apresentados os valores

médios calculados e os valores definidos para as constantes de alisamento do nível ( $\alpha$ ), da tendência ( $\beta$ ) e da sazonalidade ( $\gamma$ ).

**Tabela 18**

*Definição de constantes de alisamento para as séries das subfamílias ZI, ZG e ZM*

	<b>Nível (<math>\alpha</math>)</b>	<b>Tendência (<math>\beta</math>)</b>	<b>Sazonalidade (<math>\gamma</math>)</b>
<b>Valor médio</b>	0,4	0,3	0,5
<b>Valor definido</b>	0,3	0,3	0,7

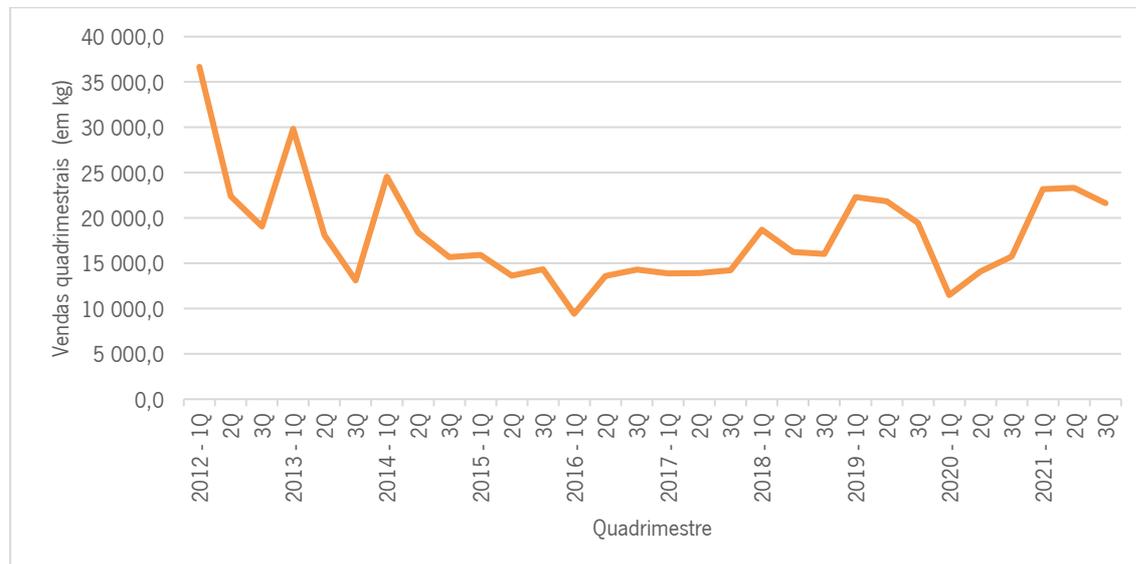
*Nota.* Elaboração própria.

### 5.3 Família D

A família D, tal como podemos verificar pela análise da **Figura 20**, tem vindo a apresentar uma tendência negativa dos seus quilogramas vendidos. Apesar disso, esta família continua a ter importância sobre as vendas da Linhas, S.A., sendo a terceira família a apresentar maiores quilogramas vendidos, importando, assim, proceder à análise das duas previsões.

**Figura 20**

*Vendas quadrimestrais do conjunto da família D (em kg)*



*Nota.* Elaboração própria.

Dadas as componentes caracterizantes da série D (**Apêndice P**), foram aplicados modelos de AE em paralelo com o modelo atualmente empregue pela Linhas, S.A., tendo-se obtido os resultados apresentados na **Tabela 19** para as medidas de avaliação de desempenho.

**Tabela 19**

Medidas de avaliação de desempenho dos métodos de previsão aplicados à série da família D

	Dentro da amostra							
	<i>Holdout = 3</i>				<i>Holdout = 6</i>			
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>
<b>Método Empresa</b>	10 156	6 419	0,3399	-	10 396	6 293	0,3103	-
<b>AES</b>	5 086	3 906	0,2272	5 406	5 099	3 939	0,2174	5 448
<b>Holt</b>	5 022	4 041	0,2314	5 674	4 363	3 526	0,1868	4 981
<b>HW multiplicativo</b>	4 102	3 269	0,2151	4 926	2 975	2 411	0,1518	3 629
<b>HW aditivo</b>	4 252	3 312	0,2146	5 106	3 281	2 526	0,1574	4 001
	Fora da amostra							
	<i>Holdout = 3</i>				<i>Holdout = 6</i>			
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>
<b>Método Empresa</b>	9 237	8 925	0,3906	-	8 429	7 820	0,4777	-
<b>AES</b>	7 238	7 197	0,3163	-	4 938	4 463	0,2955	-
<b>Holt</b>	8 351	8 332	0,3668	-	10 101	9 795	0,6110	-
<b>HW multiplicativo</b>	8 620	8 578	0,3772	-	9 453	8 765	0,5691	-
<b>HW aditivo</b>	10 623	10 269	0,4497	-	9 252	8 640	0,5614	-

Nota. Elaboração própria (dados obtidos através do *Microsoft Excel* e *Software Forecast Pro*).

Pelo exame dos resultados obtidos para as medidas de avaliação de desempenho, apesar de todos os modelos de AE apresentarem melhores resultados que o modelo atualmente empregue pela empresa numa perspectiva dentro da amostra, é possível destacar o comportamento do modelo *HW* multiplicativo. Por outro lado, fora da amostra, nem todos os modelos de AE apresentam melhores resultados que o modelo *sNaive*, sendo que apenas o modelo de AES apresenta melhores resultados fora da amostra tanto para o *holdout* de três como de seis. Utilizando como critério de desempate a escolha por modelos melhor comportados fora da amostra, selecionou-se como critério a aplicar o de AES. Este modelo, para um *holdout* de seis, consegue melhorias para cada medida de desempenho em cerca de 40%, em relação ao método atualmente empregue pela empresa (**Apêndice Q**).

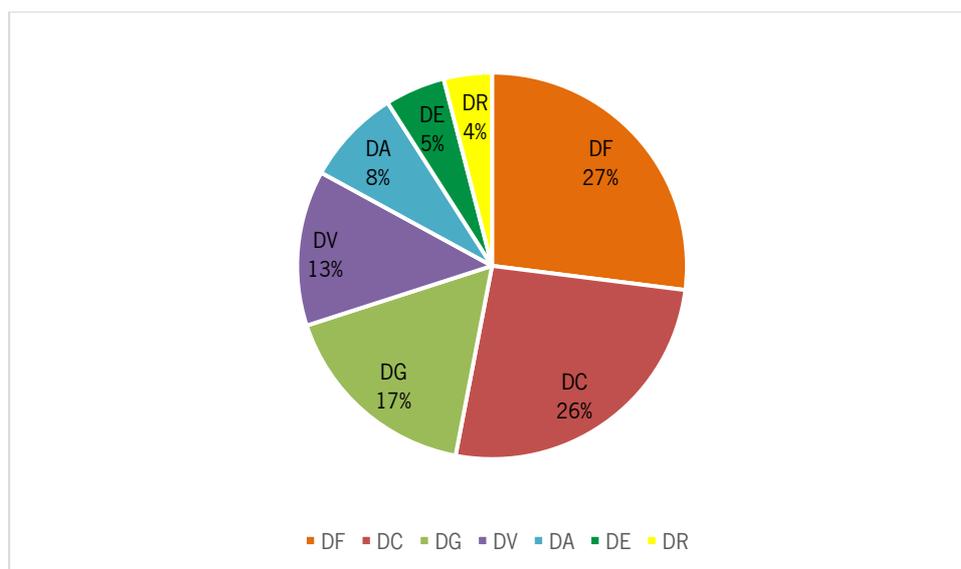
Para a aplicação do modelo de AES dever-se-á considerar uma constante de alisamento do nível ( $\alpha$ ) de 0,4, dado ser esse o valor ótimo aproximado que o *software* calculou para ambos os *holdouts*.

Após a seleção do melhor modelo quantitativo a considerar para a série temporal do conjunto da família D, um modelo de AES com  $\alpha=0,4$ , importa agora realizar uma previsão individual para cada subfamília

em questão. De forma a obter previsões para o ano de 2022, aos valores previstos obtidos pelos dados da série, dever-se-á aplicar as percentagens correspondentes às quotas anuais do último ano, 2021, para subfamília D, tendo como total a soma das vendas em kg de todas as subfamílias D. Assim dever-se-á aplicar, em 2022, as quotas apresentadas na **Figura 21**.

**Figura 21**

*Quotas anuais de 2021 das subfamílias D a aplicar em previsões 2022*



*Nota.* Elaboração própria.

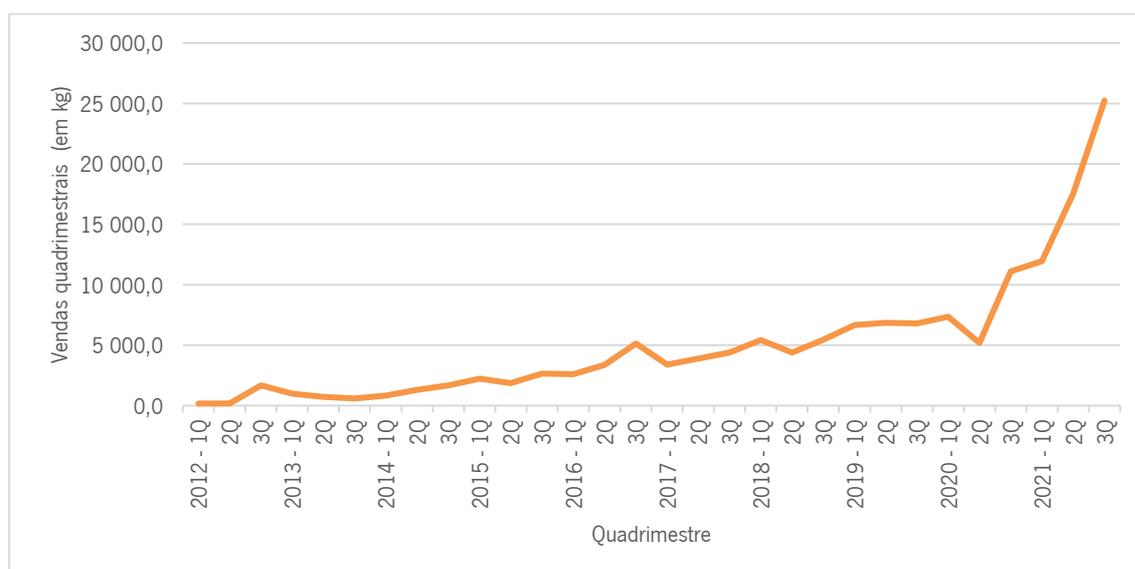
A empresa deverá, em cada ano civil, redefinir estas percentagens de forma a obter previsões mais próximas dos valores observados reais.

#### 5.4 Família S

A quarta e última família a ser estudada será a família S. Esta família, como é visível na **Figura 22**, tem vindo a apresentar uma tendência positiva.

**Figura 22**

Vendas quadrimestrais do conjunto da família S (em kg)



Nota. Elaboração própria.

A partir da aplicação do método de decomposição multiplicativo, é possível confirmar esta tendência, identificando-se uma componente ciclo-tendência de mais de 90% e irregular de quase 8% (**Apêndice R**).

Devido às características da série da família S, procedeu-se à aplicação de modelos de AE, assim como o método atualmente implementado pela empresa, tendo-se obtido as medidas de avaliação de desempenho apresentadas na **Tabela 20**.

**Tabela 20**

Medidas de avaliação de desempenho dos métodos de previsão aplicados à série da família S

	Dentro da amostra							
	Holdout = 3				Holdout = 6			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	1 445	1 145,6	0,4762	-	1 202,5	1 011,9	0,5023	-
<b>AES</b>	1 313	861,0	0,3077	1 396	803,8	624,5	0,2588	858,8
<b>Holt</b>	1 045	662,6	0,3108	1 180	657,1	463,6	0,2589	750,1
<b>HW multiplicativo</b>	945,5	678,2	0,3662	1 135	645,4	468,0	0,2256	787,2
<b>HW aditivo</b>	983,0	673,4	0,2975	1 181	648,1	480,2	0,2650	790,5

	Fora da amostra							
	<i>Holdout = 3</i>				<i>Holdout = 6</i>			
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>BIC</i>
<b>Método Empresa</b>	11 154	10 356	0,5495	-	5 963	4 645	0,3793	-
<b>AES</b>	10 924	9 464	0,4724	-	9 161	6 798	0,4242	-
<b>Holt</b>	10 419	9 105	0,4578	-	7 673	5 578	0,3610	-
<b>HW multiplicativo</b>	10 852	9 714	0,4965	-	7 648	5 507	0,3493	-
<b>HW aditivo</b>	10 357	9 141	0,4624	-	7 655	5 545	0,3557	-

Nota. Elaboração própria (dados obtidos através do *Microsoft Excel e Software Forecast Pro*)

Após a análise das medidas de avaliação de desempenho obtidas pela implementação dos modelos de previsão, constatou-se que os modelos de AE apresentavam grandes semelhanças nos resultados dentro da amostra para um *holdout* de três e seis, além de que apresentam melhorias em relação ao modelo aplicado pela Linhas, S.A..

Para fora da amostra, com um *holdout* de três, apesar das semelhanças dos resultados dos modelos, podem-se destacar os modelos de *Holt* e *HW* aditivo, a apresentarem melhorias nos resultados das medidas de desempenho em relação ao método sNaïve, entre os 7 e 16%. Já, para um *holdout* de seis, ainda numa perspectiva fora de amostra, o método implementado pela Linhas, S.A. foi aquele que apresentou melhores resultados de *RMSE* e *MAD*, contudo, em relação ao *MAPE*, o *HW* multiplicativo foi o melhor-comportado, seguindo-se do *HW* aditivo e depois do *Holt*.

Dada a componente ciclo-tendência que caracteriza a série em análise, a representar mais de 90%, deve-se implementar um modelo que tenha em consideração esta componente. De entre os modelos considerados, os modelos que o fazem são o modelo de *Holt* e os modelos de *HW*. Analisando o comportamento dos mesmos, para dentro e fora da amostra, a escolha deverá recair no modelo que melhor se comportou, apesar das grandes semelhanças entre os mesmos, sendo ele o *HW* aditivo, mantendo o modelo maioritariamente eleito para as restantes famílias.

As taxas de variação referidas, relativas às medidas de avaliação de desempenho que comparam modelos de AE com o método sNaïve, encontram-se apresentadas no **Apêndice S**.

Após a definição do modelo de AE a aplicar, *HW* aditivo, torna-se essencial definir os valores a considerar para as constantes de alisamento na aplicação do modelo para a família S. Para isso, utilizou-se o mesmo procedimento anteriormente utilizado nas famílias A e Z, estando na **Tabela 21** apresentados os valores médios calculados e os valores definidos para as constantes de alisamento do nível ( $\alpha$ ), da tendência ( $\beta$ ) e da sazonalidade ( $\gamma$ ).

**Tabela 21**

*Definição de constantes de alisamento para a série da família S*

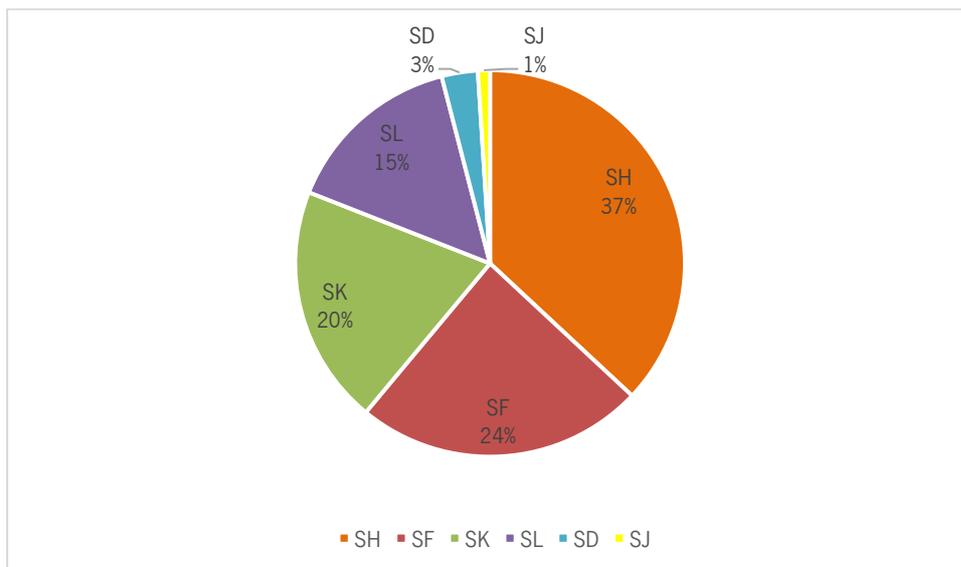
	<b>Nível (<math>\alpha</math>)</b>	<b>Tendência (<math>\beta</math>)</b>	<b>Sazonalidade (<math>\gamma</math>)</b>
<b>Valor médio</b>	0,1	0,4	0,1
<b>Valor definido</b>	0,1	0,3	0,1

*Nota.* Elaboração própria.

Finalizada a aplicação do modelo de previsão para a série da família S, dever-se-á considerar as quotas para cada subfamília de S de forma a proceder a uma previsão individual das mesmas. Para isso, tal como realizado na subfamília D, consideram-se as quotas anuais apresentadas no ano civil anterior ao ano previsto. Para as previsões do ano de 2022, considera-se as quotas anuais de 2021, apresentadas na **Figura 23**, considerando-se ser necessária as suas atualizações a cada fim de ano civil.

**Figura 23**

*Quotas anuais de 2021 das subfamílias S a aplicar em previsões 2022*



*Nota.* Elaboração própria.

## 6. PREVISÃO DA PROCURA ATRAVÉS DE MÉTODOS DE INTELIGÊNCIA ARTIFICIAL E COMPARAÇÃO COM MÉTODOS ESTATÍSTICOS

O presente capítulo destina-se à previsão da procura com recurso à inteligência artificial para as quatro diferentes famílias de produtos e subfamílias subjacentes selecionadas como cruciais para o estudo, estando, por isso, subdividido em quatro partes respeitantes cada uma, a uma família: A, Z, D e S. Considerou-se ser pertinente, para este estudo, proceder à aplicação de apenas um método de previsão baseado em inteligência artificial tendo em conta os quatro métodos considerados pela literatura revista no capítulo três como os métodos mais utilizados na têxtil-vestuário: *ANN* básicas, *ENN*, *ELM* e *EELM*. Neste sentido, selecionou-se para o estudo a aplicação do método baseado em *ELM*, em detrimento dos demais, uma vez que se pretendia obter resultados a nível de precisão das previsões superiores aos métodos estatísticos aplicados no capítulo anterior, presando, ao mesmo tempo, pela rapidez da sua aplicação.

À semelhança do capítulo anterior, definiram-se dois cenários de previsão, nomeadamente, um com *holdout* de três, utilizando o ano de 2021 como teste, e outro com *holdout* de seis, utilizando os anos 2020 e 2021 para teste.

Para a aplicação da *ELM* recorreu-se a um código em linguagem de *Python* apresentado em Anexo.

De forma a treinar a *ELM* utilizando os dados dos kg vendidos fornecidos pela empresa, é essencial adequar os dados ao modelo, sendo por isso necessário proceder a um tratamento dos mesmos antes de treinar a *ELM*. Para isso, foi realizada uma normalização da base dos dados, centrando os dados de treino e os dados de teste à média dos dados de treino. Posteriormente, as previsões foram agrupadas por janelas de adjacência, isto é, para que a *ELM* seja capaz de fazer uma previsão, esta recebe uma janela de tamanho  $l$  (usualmente chamado de lag) com dados de procura de dias adjacentes  $\{p_1, p_2, \dots, p_l\}$  e realiza uma previsão para o dia consecutivo ao último da janela, ou seja prevê  $\{p_{l+1}\}$ , de onde resulta que o número de neurónios de *input* da rede seja exatamente  $l$ . Este parâmetro  $l$  pode ser otimizado e faz sentido que se ajuste ao tamanho da base de dados. No caso particular desta aplicação foi escolhida um  $l = 10$  dias para todas as famílias, exceto para o caso particular das séries ZG e ZI onde  $l$  foi reduzida para 9, uma vez que apenas existiam dados desde 2018.

Após o tratamento dos dados, seguiram-se as etapas apresentadas na revisão de literatura relativas ao treino do algoritmo *ELM*.

As previsões da procura obtidas através do algoritmo *ELM* foram diárias, contudo, para posterior cálculo das medidas de avaliação de desempenho e comparação com os resultados dos métodos estatísticos

selecionados no capítulo anterior e do atual método da empresa, somaram-se as previsões da procura diárias, tendo sido obtidas previsões de procura quadrimestrais.

Para a comparação entre os métodos de previsão recorreu-se apenas aos resultados para fora da amostra, dado serem os que melhor transparecem os comportamentos dos métodos de previsão.

## 6.1 Família A

### 6.1.1 Subfamília AK

Procedeu-se à previsão da procura quadrimestral para a subfamília AK com recurso ao algoritmo *ELM*, seguindo-se com o cálculo das medidas de avaliação de desempenho *RMSE*, *MAD* e *MAPE* para fora da amostra para um *holdout* de três e de seis, de forma a estabelecer uma comparação dos resultados tendo em conta os métodos previamente aplicados, em particular, o método aplicado pela Linhas, S.A. e o método estatístico anteriormente selecionado para a série em questão: *HW* aditivo. Estes resultados encontram-se compilados pela **Tabela 22**. Além disso, procedendo-se ao cálculo das taxas de variação associadas à escolha do método *ELM* em detrimento dos demais modelos de forma facilitar a comparação entre os métodos aplicados (**Apêndice Tabela T 1**).

**Tabela 22**

*Medidas de avaliação de desempenho dos métodos de previsão sNaive, HW aditivo e ELM aplicados à série AK*

	Fora da amostra					
	<i>Holdout = 3</i>			<i>Holdout = 6</i>		
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>
<b>Método Empresa</b>	23 570	21 015	0,1115	23 036	17 094	0,1030
<b><i>HW</i> aditivo</b>	14 287	11 726	0,0621	16 398	13 591	0,0786
<b><i>ELM</i></b>	10 763	7 171	0,0323	19 464	13 790	0,1039

*Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)*

Pela análise dos resultados obtidos verificou-se que a implementação do método baseado em *ELM*, quando considerando um *holdout* de três, apresenta melhores resultados comparativamente aos restantes métodos considerados. Contudo, para um *holdout* de seis, apesar de apresentar, na generalidade, melhores resultados em relação ao método implementado atualmente pela Linhas, S.A., o mesmo não se verifica em relação ao modelo de AE anteriormente selecionado, piorando os seus resultados comparativamente ao *HW* aditivo a taxas de variação de 19, 1 e 32%, tendo como critério o *RMSE*, *MAD* e *MAPE*, respetivamente.

Tendo como critério de desempate a escolha pelo método mais bem-comportado para um *holdout* de seis, pode-se eleger como método a aplicar à série AK o modelo *HW* aditivo. A isto acresce-se a vantagem de a aplicação do *HW* aditivo ser mais simples e de fácil compreensão comparativamente à *ELM*.

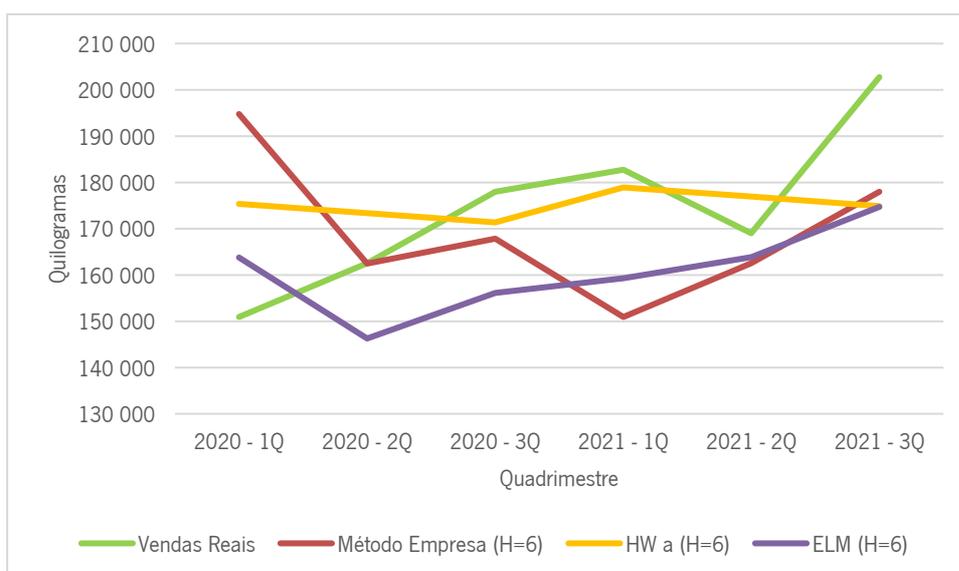
### Escolha do método a implementar à série AK

De forma enfatizar a escolha do modelo *HW* aditivo para a série AK em detrimento dos demais modelos analisados, compilaram-se os resultados obtidos para um *holdout* de seis para os métodos comparados, isto é, o aplicado atualmente pela Linhas, S.A. (*sNaive*), o de AE anteriormente eleito (*HW* aditivo) e o de IA (*ELM*).

Abaixo encontram-se ilustradas, através da **Figura 24**, as vendas reais quadrimestrais em kg para os anos 2020 e 2021, bem como as suas previsões obtidas através dos três métodos em questão (considerado um *holdout* de seis).

**Figura 24**

*Vendas reais quadrimestrais e respetiva procura prevista para sNaive, HW aditivo e ELM para a série AK*



*Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)*

Procedeu-se ainda ao cálculo das taxas de variação associadas à utilização dos métodos *HW* aditivo ou *ELM* em detrimento do método aplicado atualmente pela Linhas, S.A., para um *holdout* de seis, sendo as mesmas apresentadas na **Tabela 23**. Para o cálculo destas taxas de variação utilizaram-se como dados as medidas de avaliação de desempenho, anteriormente obtidas, para um *holdout* de seis.

**Tabela 23**

Taxas de variação associadas à implementação de modelos de HW aditivo ou ELM em detrimento do modelo sNaive para a série AK

	<b>Fora da amostra</b>		
	<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>HW aditivo</b>	-29%	-20%	-24%
<b>ELM</b>	-16%	-19%	1%

Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

A análise do gráfico e das taxas de variação vêm confirmar o já analisado, isto é, a escolha pelo modelo de HW aditivo. Este método de AE permitiu uma redução de todas as medidas de avaliação de desempenho acima dos 20% para um *holdout* de seis, quando comparado com o método atualmente empregue pela Linhas, S.A..

#### 6.1.2 Subfamília AI

Aplicou-se o algoritmo *ELM* à subfamília AI de forma a obter previsões da procura e, posteriormente, com o objetivo de proceder a comparações quantitativas, calcularam-se medidas de avaliação de desempenho fora da amostra, para os três métodos considerados com *holdout* de três e seis, estando as mesmas apresentadas na **Tabela 24** e as taxas de variação associadas às mesmas que estabelecem comparação entre o método baseado em *ELM* e os restantes métodos no **Apêndice Tabela T 2**.

**Tabela 24**

Medidas de avaliação de desempenho dos métodos de previsão sNaive, HW aditivo e ELM aplicados à série AI

	<b>Fora da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>Método Empresa</b>	1 859	1 801	0,2505	2 208	2 106	0,3615
<b>HW aditivo</b>	1 047	1 044	0,1391	1 664	1 384	0,2268
<b>ELM</b>	1 427	1 115	0,1215	1 858	988	0,1039

Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

Estabelecendo uma comparação entre as medidas de avaliação de desempenho obtidas para os três métodos considerados, constatou-se que o método que utiliza *ELM* conseguiu melhorias nos resultados comparativamente ao método aplicado pela Linhas, S.A., para os cenários de *holdout* de três e seis.

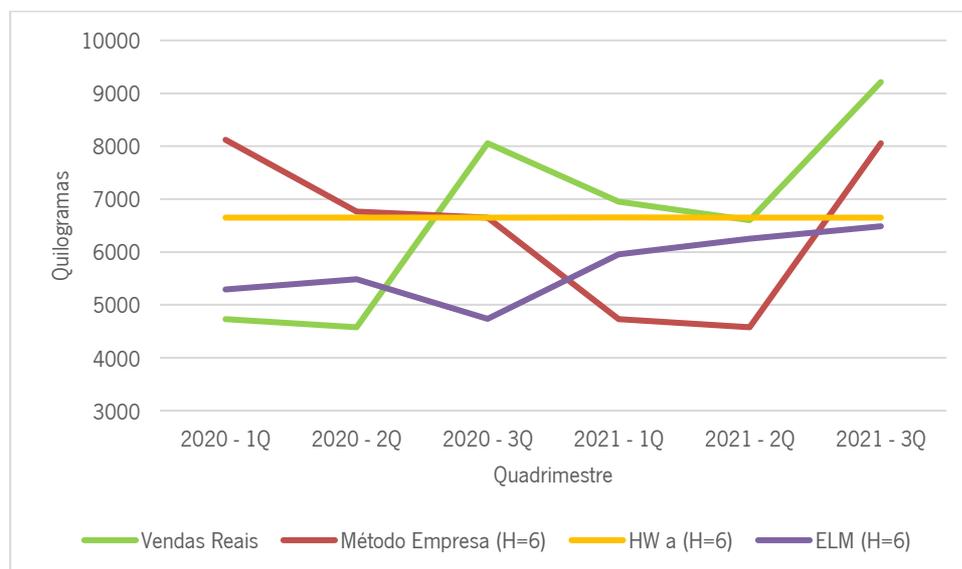
Já em relação ao modelo de AE anteriormente selecionado, *HW* aditivo, o modelo de AE comportou-se, de forma geral, melhor para um *holdout* de três, mas o algoritmo *ELM* apresentou melhores resultados, na generalidade, para um *holdout* de seis. Em particular, para um *holdout* de três, o *HW* aditivo apresenta um melhor comportamento comparativamente ao *ELM*, com taxas de variação na ordem dos 36 e 7% tendo como critério o *RMSE* e *MAD*, respetivamente, mas apresenta piores resultados em 13% relativamente ao *MAPE*. Para um *holdout* de seis, o *ELM* apresenta melhores resultados relativamente ao *MAD* e *MAPE* em 29 e 54%, respetivamente, e pior desempenho em 12% tendo como critério o *RMSE*. Neste seguimento, utilizando como critério de desempate a preferência por modelos melhor comportados para um *holdout* de seis, acrescentando-se o facto de as melhorias calculadas apresentarem taxas mais altas, elege-se-ia como método a aplicar à série AI o método baseado em *ELM*.

### Escolha do método a implementar à série AI

Foi desenvolvido o gráfico apresentado na **Figura 25**, que contempla as vendas reais e a previsão da procura para os anos testes, 2020 e 2021 com um *holdout* de seis, obtidas através dos métodos *sNaive*, *HW* aditivo e *ELM*. Este gráfico permite estabelecer uma comparação visual do comportamento dos métodos de previsão aplicados à série AI.

**Figura 25**

*Vendas reais quadrimestrais e respetiva procura prevista para sNaive, HW aditivo e ELM para a série AI*



*Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)*

Na **Tabela 25** encontram-se as taxas de variação, relativas às medidas de avaliação de desempenho para um *holdout* de seis, tendo como seleção os métodos *HW* aditivo ou *ELM* em detrimento do método aplicado atualmente pela Linhas, S.A..

**Tabela 25**

Taxas de variação associadas à implementação de modelos de *HW* aditivo ou *ELM* em detrimento do modelo *sNaive* para a série *AI*

	<b>Fora da amostra</b>		
	<b><i>Holdout</i> = 6</b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b><i>HW</i> aditivo</b>	-25%	-34%	-37%
<b><i>ELM</i></b>	-16%	-53%	-71%

*Nota.* Elaboração própria (dados obtidos com recurso a *Python* e *Software Forecast Pro*)

Considerando os resultados obtidos, verificou-se que, apesar de ambos os métodos, de *AE* e de *IA*, terem permitido melhorias em relação ao método aplicado atualmente pela Linhas, S.A., o método baseado em *ELM* permitiu, na generalidade, melhorias maiores nas medidas de avaliação de desempenho, permitindo uma diminuição em mais de metade do *MAD* e em mais de dois terços do *MAPE*.

Contudo, deve ter-se em consideração que o *ELM* apresenta a desvantagem de instabilidade de resultados comparativamente ao *HW* aditivo. Além disso, tendo em consideração os dados anteriormente apresentados para um *holdout* de três, o modelo mais bem-comportado seria o *HW* aditivo. Acresce ainda que, o *HW* aditivo apresenta as vantagens de simplicidade, maturidade e fácil compreensão e implementação. Perante isto, pode considerar-se que o melhor modelo a aplicar para este caso será o *HW* aditivo, dado os benefícios associados ao melhor comportamento das medidas de avaliação de desempenho não compensarem o risco de instabilidade dos resultados, uma vez que as medidas de avaliação de desempenho obtidas com recurso ao modelo *HW* aditivo já apresentam bons resultados, na ordem dos mil kg para *RMSE* e *MAD* e na ordem dos 20% para o *MAPE* (valores com um *holdout* de seis).

### 6.1.3 Subfamília AM

Obtidas as previsões com *holdout* de três e seis para a procura quadrimestral através do método baseado em *ELM*, calcularam-se as medidas de avaliação de desempenho associadas para fora da amostra, apresentadas na **Tabela 26**. À tabela foram acrescentadas as medidas de avaliação de desempenho anteriormente obtidas para os modelos *sNaive* e *HW* aditivo. Foram ainda calculadas as taxas de variação associadas às medidas de avaliação de desempenho respeitantes à implementação do *ELM* em detrimento dos demais métodos, apresentadas no **Apêndice Tabela T 3**.

**Tabela 26**

Medidas de avaliação de desempenho dos métodos de previsão sNaïve, HW aditivo e ELM aplicados à série AM

	Fora da amostra					
	Holdout = 3			Holdout = 6		
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>
<b>Método Empresa</b>	2 804	1 838	0,3073	2 466	1 708	0,5599
<b>HW aditivo</b>	1 919	1 397	0,2493	1 317	868	0,3049
<b>ELM</b>	3 216	2 009	0,5372	3 126	1 556	0,8785

Nota. Elaboração própria (dados obtidos com recurso a *Python* e *Software Forecast Pro*)

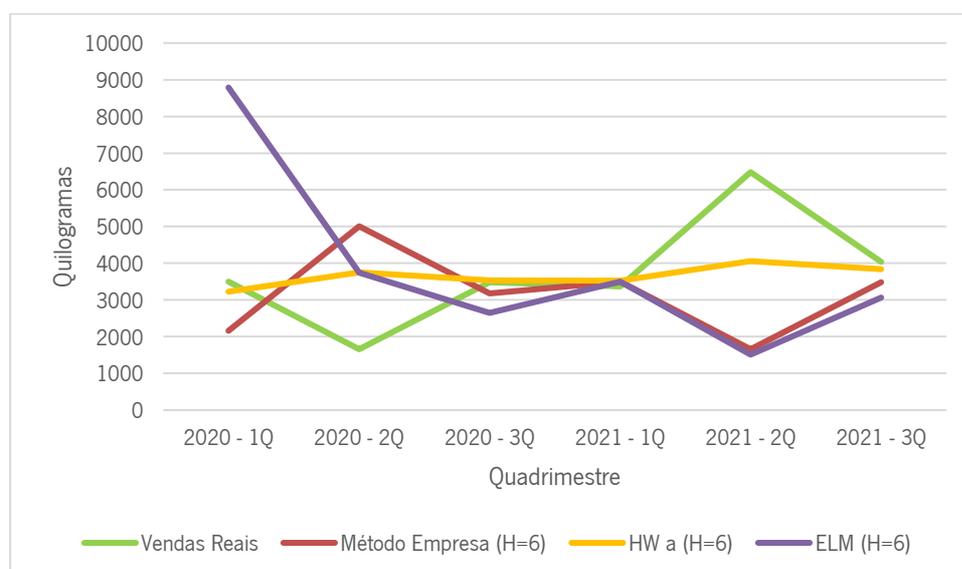
Pela comparação dos métodos de previsão considerados, verificou-se que a implementação de *ELM* não permite, na generalidade, melhoria dos resultados. Pelo contrário, a implementação de *ELM* vai piorar os resultados das medidas de avaliação de desempenho com taxas entre os 9 e 75% quando comparado com o método atualmente empregue pela Linhas, S.A. (salvo a exceção do *MAD* para *holdout* de seis, que apresentou uma melhoria de 9%) e entre os 44 e 188% quando comparado com o modelo *HW* aditivo. Neste sentido, descarta-se a implementação do *ELM* para a série AM.

### **Escolha do método a implementar à série AM**

Desenvolveu-se um gráfico com o objetivo de examinar as vendas reais quadrimestrais em kg para os anos de 2020 e 2021, e respetivas previsões da procura com um *holdout* de seis obtidas através dos métodos sNaïve, *HW* aditivo e *ELM* (**Figura 26**).

**Figura 26**

Vendas reais quadrimestrais e respetiva procura prevista para sNaive, HW aditivo e ELM para a série AM



Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

Pela análise do gráfico obtido, constata-se desde logo, de forma visual, o melhor comportamento do modelo *HW* aditivo em relação aos restantes.

Calculou-se ainda as taxas de variação associadas às medidas de avaliação de desempenho para um *holdout* de seis, para o caso em que se elege os métodos *HW* aditivo ou *ELM* em detrimento do método atualmente empregue pela Linhas, S.A.. Estas encontram-se apresentadas na **Tabela 27**.

**Tabela 27**

Taxas de variação associadas à implementação de modelos de *HW* aditivo ou *ELM* em detrimento do modelo *sNaive* para a série AM

	Fora da amostra		
	<i>Holdout = 6</i>		
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>
<i>HW</i> aditivo	-47%	-49%	-46%
<i>ELM</i>	27%	-9%	57%

Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

Confirma-se, assim, a escolha pelo modelo *HW* aditivo para a previsão da procura da série AM, constatando-se que, para um *holdout* de seis, ao contrário do *ELM*, o *HW* aditivo permite melhorias em todas as medidas de avaliação de desempenho em relação ao método atualmente empregue pela Linhas, S.A., nomeadamente, em quase metade.

#### 6.1.4 Subfamília AA

Obtiveram-se previsões da procura para a subfamília AA através do método baseado em *ELM* com um *holdout* de três e de seis. Para estas previsões calcularam-se as medidas de avaliação de desempenho para fora da amostra, tendo-se obtido os resultados apresentados na **Tabela 28**. À tabela acrescentaram-se as medidas de avaliação de desempenho associadas aos modelos *sNaïve* e *HW* aditivo, já anteriormente apresentadas. Além disso, procedeu-se ao cálculo de taxas de variação associadas às medidas de avaliação de desempenho respeitantes à implementação do *ELM* em detrimento dos demais métodos, apresentadas no **Apêndice Tabela T 4**.

**Tabela 28**

*Medidas de avaliação de desempenho dos métodos de previsão sNaïve, HW aditivo e ELM aplicados à série AA*

	Fora da amostra					
	Holdout = 3			Holdout = 6		
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>
<b>Método Empresa</b>	690	688	0,2872	599	561	0,3004
<b>HW aditivo</b>	497	399	0,1547	683	537	0,2423
<b>ELM</b>	695	385	0,1769	722	506	0,3095

*Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)*

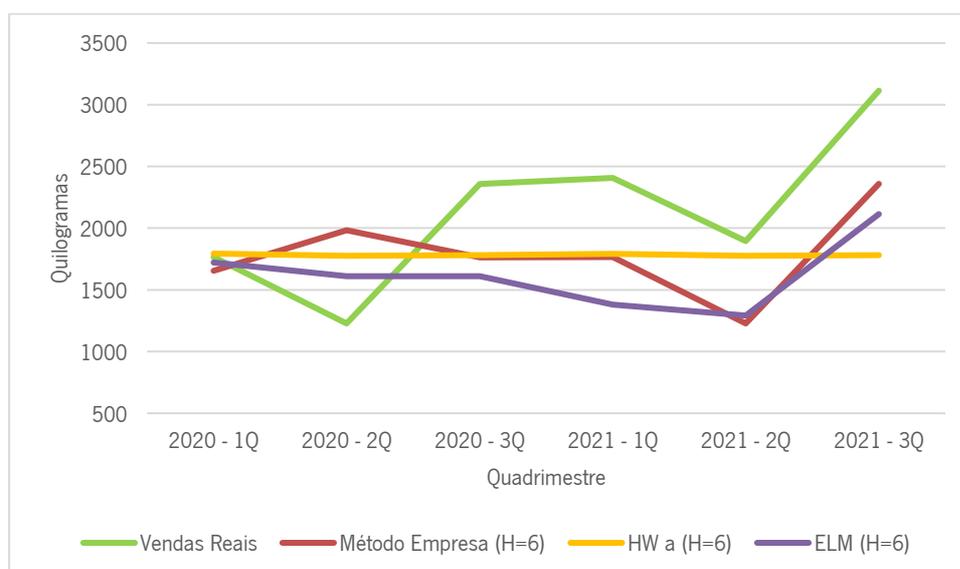
Pela análise dos resultados das medidas de avaliação de desempenho obtidas para a série AA, constatou-se que a aplicação de *ELM* apresenta, na generalidade, para um *holdout* de três, melhorias nas medidas de avaliação de desempenho comparativamente com o método implementado pela Linhas, S.A.. Contudo, o mesmo não se verifica para um *holdout* de seis. Comparando o método baseado em *ELM* com o modelo *HW* aditivo, as melhorias só se verificam para o *MAD*, sendo que as restantes medidas apresentaram piores resultados que o *HW* aditivo, tanto para um *holdout* de três como de seis. Neste sentido, parece ser benéfico manter a aplicação do modelo *HW* aditivo, em detrimento dos demais.

#### **Escolha do método a implementar à série AA**

De forma a confirmar a escolha do método de previsão a aplicar para a série AA, compilaram-se, através da **Figura 27**, os dados das vendas quadrimestrais reais em kg dos anos 2020 e 2021, bem como as respetivas previsões da procura dos métodos *sNaïve*, *HW* aditivo e *ELM*, tendo como teste esses anos.

**Figura 27**

Vendas reais quadrimestrais e respetiva procura prevista para sNaive, HW aditivo e ELM para a série AA



Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

Pela análise da **Figura 27**, verifica-se que nos dois primeiros quadrimestres de 2020 todos os modelos de previsão aplicados apresentam previsões superiores às vendas reais, já a partir do terceiro quadrimestre de 2020, todos os modelos apresentam previsões inferiores às vendas reais, sendo o *HW* aditivo aquele que mais se aproxima, de forma geral, às vendas reais.

A esta análise, acrescentou-se o cálculo de taxas de variação associadas à escolha dos modelos *HW* aditivo ou *ELM* em detrimento do *sNaive*, apresentadas pela **Tabela 29**, utilizando como critério as medidas de avaliação de desempenho para um *holdout* de seis.

**Tabela 29**

Taxas de variação associadas à implementação de modelos de *HW* aditivo ou *ELM* em detrimento do modelo *sNaive* para a série AA

	Fora da amostra		
	<i>Holdout = 6</i>		
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>
<i>HW</i> aditivo	14%	-4%	-19%
<i>ELM</i>	21%	-10%	3%

Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

Constata-se que a aplicação do modelo *HW* aditivo consegue melhorias nas previsões, comparativamente ao *sNaive*, em 4 e 19% em relação ao *MAD* e *MAPE* para um *holdout* de seis, apesar de se verificar piores resultados para o *RMSE*. Já o *ELM* apenas consegue melhorias no *MAD*. Dada a análise geral, elege-se o modelo *HW* aditivo para prever a procura da série AA.

## 6.2 Família Z

### 6.2.1 Subfamília ZI

Para a subfamília ZI, tal como já realizado no capítulo antecedente, considerou-se apenas dados das vendas desde o primeiro quadrimestre de 2018. Procedeu-se, assim, à previsão da procura para a série ZI através de *ELM* utilizando somente como ano teste o ano de 2021. Perante as previsões obtidas através do *ELM*, calcularam-se as medidas de avaliação de desempenho associadas, as quais se fazem acompanhar, na **Tabela 30**, pelas medidas de avaliação de desempenho obtidas através do modelo *HW* aditivo e *sNaive*. Para uma análise mais minuciosa, calcularam-se as taxas de variação associadas às medidas de avaliação de desempenho respeitantes à implementação do *ELM* em detrimento dos demais métodos, apresentadas no **Apêndice Tabela T 5**.

**Tabela 30**

*Medidas de avaliação de desempenho dos métodos de previsão sNaive, HW aditivo e ELM aplicados à série ZI*

	<b>Fora da amostra</b>		
	<b>Holdout = 3</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>Método Empresa</b>	4 832	4 569	0,3717
<b>HW aditivo</b>	1 808	1 739	0,1427
<b>ELM</b>	1 034	460	0,1653

*Nota.* Elaboração própria (dados obtidos com recurso a *Python* e *Software Forecast Pro*)

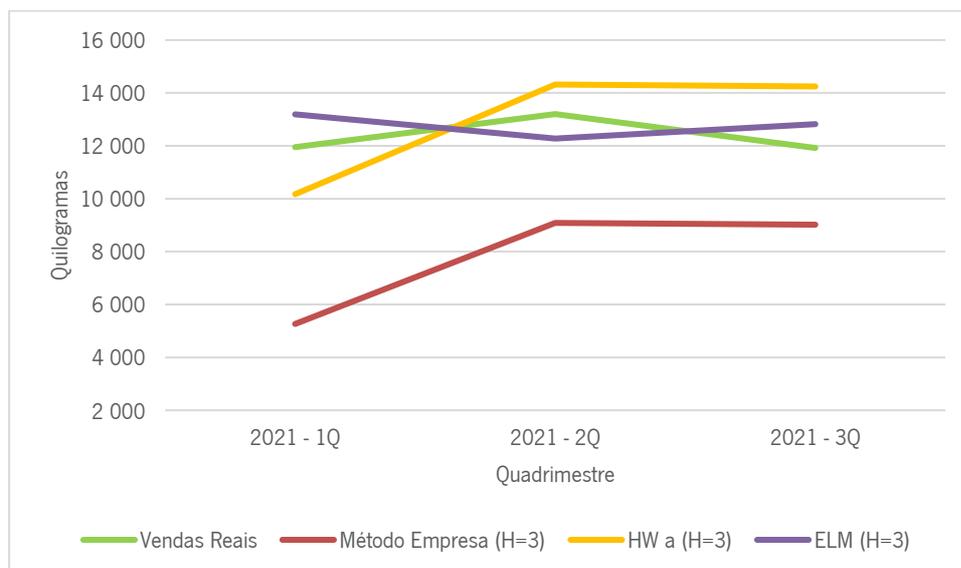
Através da implementação do método baseado em *ELM*, conseguem-se fortes melhorias nas medidas de avaliação de desempenho, nomeadamente, melhorias entre 56 e 90% quando comparando com o *sNaive* e entre 43 e 74% quando comparando com o *HW* aditivo (excetuando-se o *MAPE* que piora o desempenho em 16% comparativamente ao *HW* aditivo). Neste sentido, parece ser preferível adotar como método o *ELM*, dadas as fortes melhorias alcançadas, contudo, é necessário ter em consideração que o período considerado foi reduzido, tendo sido realizada apenas previsões para um *holdout* de três.

### **Escolha do método a implementar à série ZI**

Para a série ZI desenvolveu-se o gráfico apresentado na **Figura 28** que contempla as vendas quadrimestrais reais verificadas em 2021. A este gráfico acrescentou-se a procura prevista segundo o método aplicado pela Linhas, S.A. e o modelo de AE anteriormente selecionado, *HW* aditivo, de forma a estabelecer uma comparação entre os três métodos.

**Figura 28**

Vendas reais quadrimestrais e respetiva procura prevista para sNaive, HW aditivo e ELM para a série ZI



Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

Salienta-se logo o facto que o método aplicado pela Linhas, S.A. ser aquele que pior se comporta no que diz respeito às previsões para o ano de 2021, sendo isso visualmente perceptível pela análise do gráfico apresentado pela **Figura 28**. Já os demais modelos parecem apresentar comportamentos idênticos no que diz respeito aos períodos analisados.

Com o objetivo de confirmar a escolha do método de previsão a aplicar à série ZI, desenvolveu-se a **Tabela 31**, que contempla as taxas de variação associadas às medidas de avaliação de desempenho relativas à escolha dos modelos *HW* aditivo e *ELM* em detrimento do método aplicado pela Linhas, S.A..

**Tabela 31**

Taxas de variação associadas à implementação de modelos de *HW* aditivo ou *ELM* em detrimento do modelo *sNaive* para a série ZI

	Fora da amostra		
	<i>Holdout = 3</i>		
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>
<i>HW</i> aditivo	-63%	-62%	-62%
<i>ELM</i>	-79%	-90%	-56%

Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

Tal como observado anteriormente, tanto o *HW* aditivo como o *ELM* conseguem melhorias nas medidas de avaliação de desempenho comparativamente com o método atualmente empregue na Linhas, S.A.. Apesar de o modelo *HW* aditivo conseguir melhorias em mais de três quintos para as medidas de avaliação de desempenho comparativamente ao método aplicado atualmente pela Linhas, S.A., com os

dados considerados, a escolha pelo método a aplicar à série ZI deveria recair sobre a *ELM*, uma vez que esta consegue melhorias entre 56 e 90% nas medidas de avaliação de desempenho.

Contudo, é necessário ser prudente e recordar que a aplicação de *ELM* tem a desvantagem de instabilidade de resultados comparativamente ao *HW* aditivo, além de que a dimensão dos dados históricos considerada foi reduzida, não havendo termos de comparação para um período maior. Perante isto, e uma vez que o *HW* aditivo já consegue um bom comportamento (apresentando valores de *RMSE* e *MAD* na ordem dos mil kg e menos de 15% de *MAPE*, conseguindo, assim, melhorias em mais de 60% para todas as medidas de avaliação de desempenho consideradas, comparativamente ao método atualmente empregue pela Linhas, S.A.) elege-se como modelo de previsão da procura a implementar para este caso o *HW* aditivo.

### 6.2.2 Subfamília ZG

À semelhança da subfamília anteriormente analisada, para a previsão da procura quadrimestral para a série ZG, consideram-se apenas dados de vendas desde o primeiro quadrimestre de 2018. Após a aplicação do algoritmo *ELM* para a previsão quadrimestral da procura para a série da subfamília ZG, calcularam-se as medidas de avaliação de desempenho para fora da amostra apenas para um *holdout* de três, dada a dimensão da série. As medidas de avaliação de desempenho para os métodos considerados, bem como para o método atualmente empregue pela Linhas, S.A., encontram-se apresentadas na **Tabela 32**. Além disso, procedeu-se ao cálculo das taxas de variação associadas à escolha do método *ELM* em detrimento dos demais, apresentadas no **Apêndice Tabela T 6**.

**Tabela 32**

*Medidas de avaliação de desempenho dos métodos de previsão sNaive, HW aditivo e ELM aplicados à série ZG*

	<b>Fora da amostra</b>		
	<b><i>Holdout</i> = 3</b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>Método Empresa</b>	749	602	0,2375
<b><i>HW</i> aditivo</b>	887	720	0,2019
<b><i>ELM</i></b>	765	1 856	0,4555

*Nota.* Elaboração própria (dados obtidos com recurso a *Python* e *Software Forecast Pro*)

Tendo em conta o período considerado verifica-se que a aplicação do algoritmo *ELM*, comparativamente ao *HW* aditivo, não conseguiu melhorias de resultados em nenhuma das medidas de avaliação de

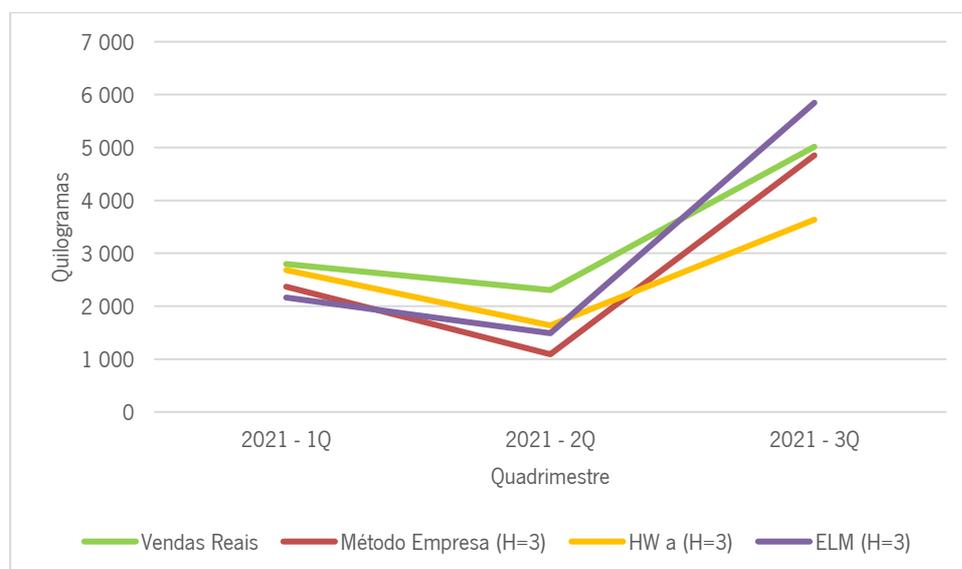
desempenho calculadas, à exceção do *RMSE*. Desta forma, considera-se inalterada a escolha do modelo de previsão para a série ZG anteriormente selecionado.

### Escolha do método a implementar à série ZG

Desenvolveu-se para a série ZG um gráfico representativo das vendas reais quadrimestrais em kg em 2021 e respetivas procuras previstas considerando os modelos sNaïve, *HW* aditivo e *ELM* (**Figura 29**).

**Figura 29**

*Vendas reais quadrimestrais e respetiva procura prevista para sNaïve, HW aditivo e ELM para a série ZG*



*Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)*

Pela análise do gráfico, o modelo *HW* aditivo parece ser aquele que melhor se comporta considerando os primeiros quadrimestres, contudo no último quadrimestre é aquele que pior comportamento apresenta, sendo o método da Linhas, S.A. aquele que mais se aproxima das vendas reais.

Considerando as medidas de avaliação de desempenho anteriormente calculadas, de forma a confirmar a escolha pelo modelo de previsão selecionado para a série ZG, procedeu-se ao cálculo das taxas de variação associadas às mesmas com a implementação dos modelos *HW* aditivo ou *ELM* em detrimento do modelo aplicado atualmente pela empresa. Estas taxas encontram-se apresentas na **Tabela 33**.

**Tabela 33**

Taxas de variação associadas à implementação de modelos de *HW* aditivo ou *ELM* em detrimento do modelo *sNaive* para a série *ZG*

	<b>Fora da amostra</b>		
	<b><i>Holdout</i> = 3</b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b><i>HW</i> aditivo</b>	18%	20%	-15%
<b><i>ELM</i></b>	2%	208%	92%

*Nota.* Elaboração própria (dados obtidos com recurso a *Python* e *Software Forecast Pro*)

Pela análise das taxas de variação das medidas de avaliação de desempenho, descarta-se logo a implementação do método *ELM*, dado que as medidas de avaliação de desempenho chegam a apresentar aumentos em mais do dobro comparativamente ao método aplicado atualmente pela empresa. Por outro lado, verifica-se que a aplicação do modelo *HW* aditivo vai permitir melhorias no *MAPE* em 15% comparativamente ao método *sNaive*, apesar de o mesmo não se verificar para as restantes medidas de avaliação de desempenho.

Tal como verificado anteriormente, a dimensão dos dados para a série em questão é reduzida, o que leva a que a decisão entre o modelo *sNaive* e *HW* aditivo levante dúvidas. Desta forma, matem-se a escolha sobre o modelo *HW* aditivo, que permite um *RMSE* e *MAD* abaixo dos mil kg e *MAPE* na ordem dos 20%.

### 6.2.3 Subfamília ZM

Obtiveram-se previsões da procura quadrimestral em kg para a série da subfamília ZM através da aplicação do algoritmo *ELM* considerando um *holdout* de três e de seis, tendo sido posteriormente calculadas as medidas de avaliação de desempenho associadas. Estas últimas encontram-se exibidas na **Tabela 34** em conjunto com as anteriormente calculadas para o método quantitativo selecionado (*HW* aditivo) e o método da empresa. Além disso, no **Apêndice Tabela T 7** são apresentadas as taxas de variação das medidas de avaliação de desempenho associadas à escolha do método *ELM* em detrimento dos demais modelos.

**Tabela 34**

Medidas de avaliação de desempenho dos métodos de previsão sNaive, HW aditivo e ELM aplicados à série ZM

	Fora da amostra					
	Holdout = 3			Holdout = 6		
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>
<b>Método Empresa</b>	1 828	1 614	0,4970	995	937	0,5992
<b>HW aditivo</b>	1 256	968	0,2585	1 201	1 059	0,6899
<b>ELM</b>	1 270	884	0,3255	1 007	603	0,4655

Nota. Elaboração própria (dados obtidos com recurso a *Python* e *Software Forecast Pro*)

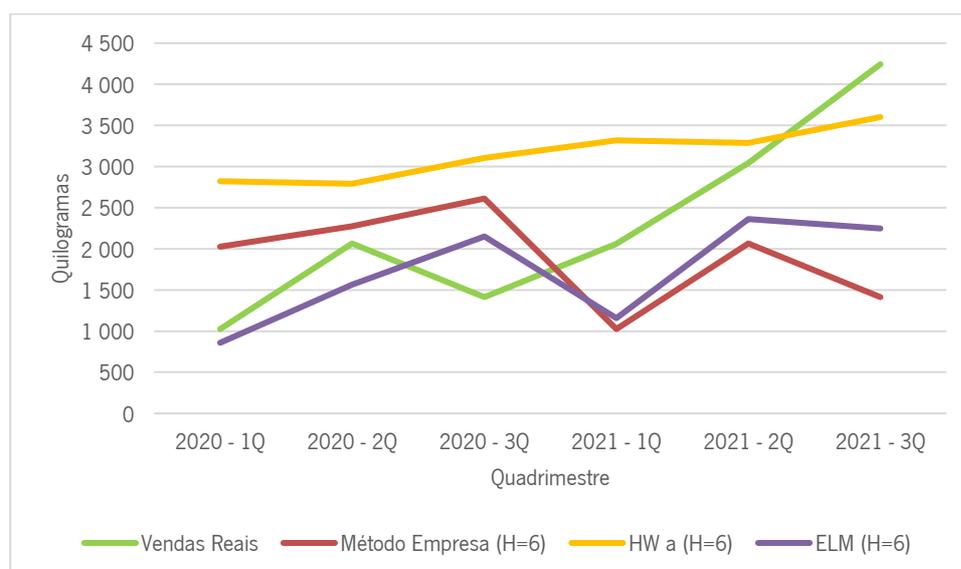
Através da comparação das medidas de avaliação de desempenho obtidas para os diferentes modelos e cenários considerados, verifica-se que o HW aditivo foi aquele que melhor se comportou para um *holdout* de três, seguindo-se ao mesmo, como mais bem-comportada, a ELM. Já para um *holdout* de seis, a ELM foi quem apresentou o melhor *MAD* e *MAPE* e o segundo melhor *RMSE*.

### Escolha do método a implementar à série ZM

Após a aplicação dos vários modelos de previsão, foi possível desenvolver um gráfico que compilasse as vendas reais quadrimestrais em kg de 2020 e 2021 e a procura prevista obtida para um *holdout* de seis considerado os métodos sNaive, HW aditivo e ELM para a série ZM (**Figura 30**).

**Figura 30**

Vendas reais quadrimestrais e respetiva procura prevista para sNaive, HW aditivo e ELM para a série ZM



Nota. Elaboração própria (dados obtidos com recurso a *Python* e *Software Forecast Pro*)

Pela análise visual do comportamento dos modelos aplicados, constata-se que o modelo *HW* aditivo apresenta maioritariamente erros positivos, isto é, previsões superiores às vendas reais, ao contrário do método *ELM* que apresenta geralmente erros negativos. Já o método aplicado atualmente pela Linhas, S.A. apresenta tanto erros positivos como negativos.

Calcularam-se as taxas de variação associadas à implementação de modelos de *HW* aditivo ou *ELM* em detrimento do modelo *sNaïve* para a série ZM, para um *holdout* de seis, apresentadas na **Tabela 35**.

**Tabela 35**

*Taxas de variação associadas à implementação de modelos de HW aditivo ou ELM em detrimento do modelo sNaïve para a série ZM*

	<b>Fora da amostra</b>		
	<b><i>Holdout = 6</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b><i>HW</i> aditivo</b>	21%	13%	15%
<b><i>ELM</i></b>	1%	-36%	-22%

*Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)*

Confirma-se como mais bem-comportado o modelo *ELM*, apresentando melhorias no *MAD* e no *MAPE* em 36% e 22%, respetivamente, tendo como comparação o modelo atualmente empregue pela Linhas, S.A.. Apesar disso, considerando os dados anteriormente apresentados para um *holdout* de três, o modelo mais bem-comportado seria o *HW* aditivo. Acresce-se ainda que, como já anteriormente referido, a *ELM* apresenta a desvantagem de instabilidade de resultados. Perante isto, poderá considerar-se que os benefícios de associados ao melhor comportamento das medidas de avaliação de desempenho para um *holdout* de seis, não compensarão a instabilidade associada à aplicação de *ELM*, sendo mais cautelosa a aplicação do modelo *HW* aditivo mais estável.

#### 6.2.4 Conjunto das restantes subfamílias Z

Compilaram-se os dados das vendas para as restantes subfamílias da família Z, isto é, todas as subfamílias Z à exceção de ZI, ZG e ZM, e aplicou-se o método baseado em *ELM* para a obtenção de previsões da procura. Posteriormente, obtiveram-se as medidas de avaliação de desempenho associadas às previsões alcançadas e compilando-se as mesmas com as obtidas pelo método *sNaïve* na **Tabela 36**.

**Tabela 36**

Medidas de avaliação de desempenho dos métodos de previsão sNaive e ELM aplicados à série do conjunto das restantes subfamílias Z

	<b>Fora da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>Método Empresa</b>	5 197	3 880	0,4905	1 566	1 410	0,5256
<b>ELM</b>	4 207	3 374	0,5964	3 152	1 414	0,6260

Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

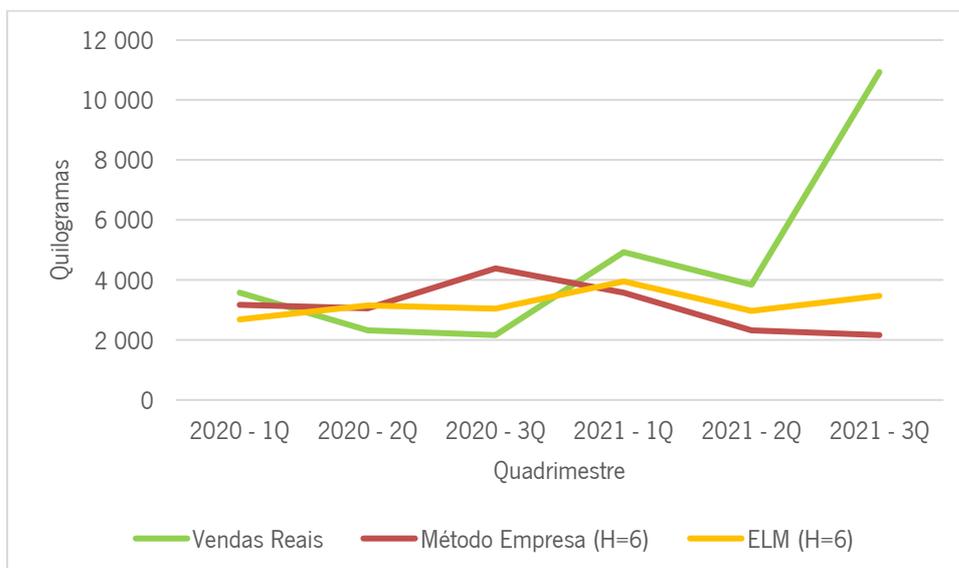
Pela análise das medidas de avaliação de desempenho e taxas de variação associadas, verifica-se que a implementação de ELM não traz, na generalidade, benefícios às previsões da procura. Apesar do RMSE e MAD apresentarem melhores valores em ELM para um holdout de três, o mesmo não acontece para as restantes medidas de avaliação de desempenho. Perante isto, parece ser benéfico manter o método atualmente empregue pela empresa na previsão da procura para o conjunto das restantes subfamílias de Z: sNaive.

### **Escolha do método a implementar à série do conjunto das restantes subfamílias de Z**

Para a série que compila os dados relativos ao conjunto das restantes subfamílias de Z, desenvolveu-se o gráfico apresentado na **Figura 31** onde são ilustradas as vendas reais quadrimestrais para os anos 2020 e 2021, bem como a respetiva procura prevista através dos métodos sNaive e ELM para um holdout de seis.

**Figura 31**

Vendas reais quadrimestrais e respetiva procura prevista para sNaïve e ELM para a série do conjunto das restantes subfamílias de Z



Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

Apesar de não ser visualmente perceptível o melhor comportamento do sNaïve, verifica-se que o sNaïve apresenta melhorias em mais de 100% no *RMSE* e 19% no *MAPE* para um *holdout* de seis comparativamente com o método baseado em *ELM*. Estas taxas de variação associadas à escolha do método *ELM* em detrimento do sNaïve tendo como critério as medidas de avaliação de desempenho para fora da amostra, encontram-se apresentadas no **Apêndice Tabela T 8**. Perante isto, elege-se como modelo de previsão a aplicar à série representativa do conjunto das restantes subfamílias Z, o sNaïve.

### 6.3 Família D

Para a família D, tal como aconteceu no capítulo anterior, compilaram-se os dados relativos às vendas em kg de todas as subfamílias da família D. Após a obtenção da procura prevista quadrimestral em kg para a família D para um *holdout* de três e de seis, obtiveram-se as medidas de avaliação de desempenho apresentadas na **Tabela 37**. Para estas calcularam-se ainda taxas de variação associadas à escolha do método *ELM* em detrimento do sNaïve e do AES, tendo como critério as medidas de avaliação de desempenho para fora da amostra (**Apêndice Tabela T 9**).

**Tabela 37**

Medidas de avaliação de desempenho dos métodos de previsão sNaive, AES e ELM aplicados à série da família D

	Fora da amostra					
	Holdout = 3			Holdout = 6		
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>
<b>Método Empresa</b>	9 237	8 925	0,3906	8 429	7 820	0,4777
<b>AES</b>	7 238	7 197	0,3163	4 938	4 463	0,2955
<b>ELM</b>	7 042	6 993	0,2803	5 964	1 906	0,3587

Nota. Elaboração própria (dados obtidos com recurso a *Python* e *Software Forecast Pro*)

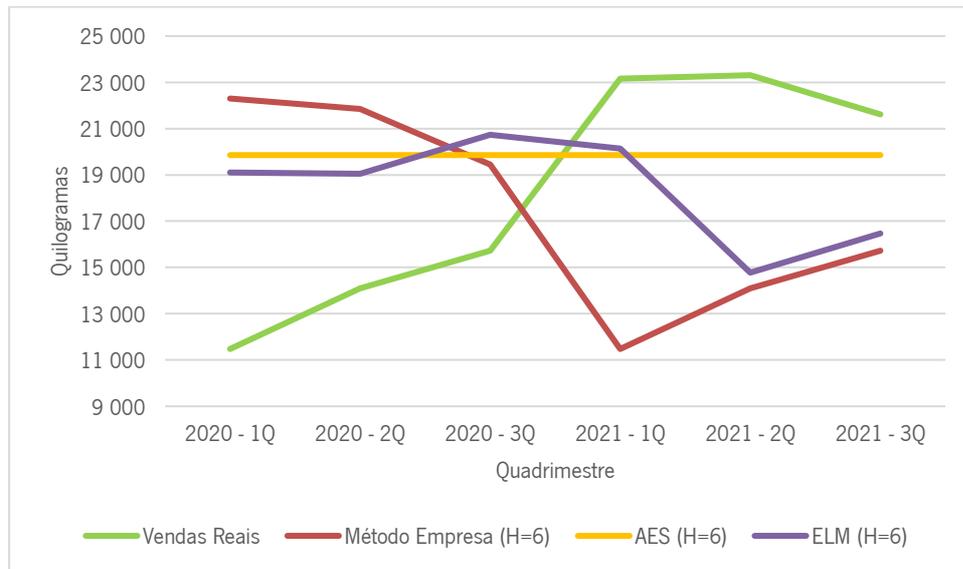
Através dos resultados obtidos salienta-se o comportamento do algoritmo *ELM* para um cenário de *holdout* três, uma vez que o modelo apresentou melhores resultados para todas as medidas de avaliação de desempenho comparativamente aos demais modelos, contudo, destaca-se a proximidade dos resultados de *ELM* e *AES*. Já para um *holdout* de seis, o modelo *ELM* apresentou melhores resultados que o método atualmente empregue pela empresa para todas as medidas de avaliação de desempenho, com taxas de variação entre -25% e -76%. Contudo, comparando com o modelo *AES*, o modelo *ELM* melhorou os resultados do *MAD* em 57%, mas piorou o seu comportamento em 21% no *RMSE* e *MAPE*.

### Escolha do método a implementar à série da Família D

Para uma escolha mais ponderada do método a implementar à série representativa da família D, compilaram-se, pela **Figura 32**, as vendas reais quadrimestrais em kg nos anos 2020 e 2021 e a respetiva procura prevista através dos modelos sNaive, AES e *ELM*, para um *holdout* de seis.

**Figura 32**

Vendas reais quadrimestrais e respetiva procura prevista para sNaive, AES e ELM para a série da família D



Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

Pela análise do gráfico destaca-se primeiramente que as vendas reais apresentaram valores inferiores às previstas para ambos os modelos em 2020 e superiores em 2021. Além disso, é visualmente perceptível o pior comportamento do sNaive em comparação aos demais modelos.

Procedeu-se ao cálculo das taxas de variação associadas à implementação de modelos de AES ou ELM em detrimento do modelo sNaive para a série da família D para um *holdout* de seis, estando as mesmas apresentadas na **Tabela 38**.

**Tabela 38**

Taxas de variação associadas à implementação de modelos de AES ou ELM em detrimento do modelo sNaive para a série da família D

	Fora da amostra		
	Holdout = 6		
	RMSE	MAD	MAPE
AES	-41%	-43%	-38%
ELM	-29%	-76%	-25%

Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

Tal como verificado anteriormente, constata-se que ambos os modelos, AES e ELM, comparativamente ao modelo atualmente empregue pela Linhas, S.A., vão permitir melhorias nas medidas de avaliação de desempenho, apresentando estas melhoria uma média de 40%. Perante estes resultados, tendo como critério para a escolha do método a implementar as medidas de avaliação de desempenho, considera-se indiferente a preferência pelo modelo de AES ou pela ELM. Desta forma, por uma questão de simplicidade

de aplicação do modelo AES e de instabilidade de resultados obtidos pelo modelo baseado em *ELM*, seleciona-se como modelo a implementar à série da família D o modelo de AES.

## 6.4 Família S

Procedeu-se à previsão da procura quadrimestral para a família S através do algoritmo *ELM*, seguindo-se com o cálculo das medidas de avaliação de desempenho *RMSE*, *MAD* e *MAPE* para fora da amostra. A **Tabela 39** apresenta estas medidas de avaliação de desempenho, juntamente com as anteriormente calculadas para o método aplicado pela Linhas, S.A. e o método estatístico anteriormente selecionado: *HW* aditivo. Além disso, procedendo-se ao cálculo das taxas de variação associadas à escolha do método *ELM* em detrimento dos demais (**Apêndice Tabela T 10**).

**Tabela 39**

*Medidas de avaliação de desempenho dos métodos de previsão sNaive, HW aditivo e ELM aplicados à série da família S*

	Fora da amostra					
	<i>Holdout = 3</i>			<i>Holdout = 6</i>		
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>
<b>Método Empresa</b>	11 154	10 356	0,5495	5 963	4 645	0,3793
<b><i>HW</i> aditivo</b>	10 357	9 141	0,4624	7 655	5 545	0,3557
<b><i>ELM</i></b>	11 913	10 695	0,4979	7 959	6 016	0,4289

*Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)*

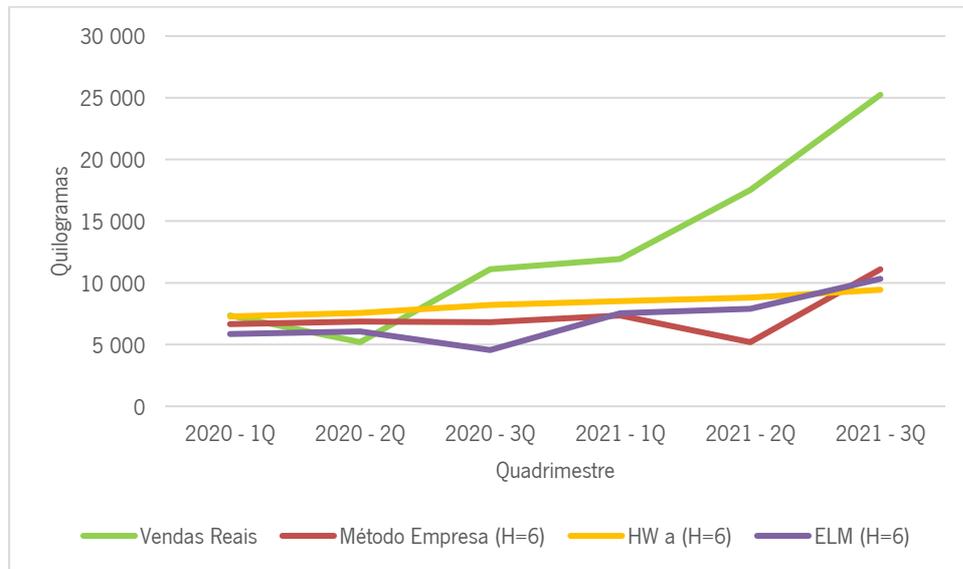
Verifica-se que a aplicação do algoritmo *ELM* na previsão da procura da série da família S não apresentou vantagens em relação aos métodos anteriormente aplicados considerando as medidas de avaliação de desempenho *RMSE*, *MAD* e *MAPE*.

### Escolha do método a implementar à série da Família S

Desenvolveu-se um gráfico com o objetivo de examinar as vendas reais quadrimestrais em kg, para os anos de 2020 e 2021, e respetivas previsões com um *holdout* de seis obtidas através dos métodos sNaive, *HW* aditivo e *ELM* (**Figura 33**).

**Figura 33**

Vendas reais quadrimestrais e respetiva procura prevista para sNaive, HW aditivo e ELM para a série da família S



Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

Pela análise do gráfico apresentado na **Figura 33** verifica-se que, apesar de no primeiro ano a procura prevista quadrimestral para ambos os modelos se ter aproximado das vendas quadrimestrais reais, no ano de 2021 as vendas reais quadrimestrais têm vindo a afastar-se cada vez mais das previstas, estando sempre acima do previsto.

Calcularam-se as taxas de variação associadas à implementação de modelos de *HW* aditivo ou *ELM* em detrimento do *sNaive* para a série da família S com *holdout* de seis (**Tabela 40**).

**Tabela 40**

Taxas de variação associadas à implementação de modelos de *HW* aditivo ou *ELM* em detrimento do modelo *sNaive* para a série da família S

	Fora da amostra		
	<i>Holdout = 6</i>		
	<i>RMSE</i>	<i>MAD</i>	<i>MAPE</i>
<i>HW</i> aditivo	28%	19%	-6%
<i>ELM</i>	33%	30%	13%

Nota. Elaboração própria (dados obtidos com recurso a Python e Software Forecast Pro)

Apesar do modelo atualmente empregue pela Linhas, S.A. ter sido aquele que, na generalidade, apresentou melhores resultados comparativamente aos restantes modelos, como já referido no capítulo anterior, a série apresenta uma componente ciclo-tendência de mais de 90%, mantendo-se, por isso, como escolha de modelo a implementar, o *HW* aditivo.

## 7. CONCLUSÕES E TRABALHO FUTURO

Numa primeira instância do projeto de dissertação, recorreu-se ao método de classificação ABC por famílias de produtos produzidos pela empresa de forma a obter uma classificação destas segundo o grau de importância, hierarquizando as previsões a realizar. Neste seguimento, tendo como critério de análise das vendas anuais em kg, selecionaram-se para o estudo de modelos de previsão da procura, apenas quatro famílias com a seguinte hierarquia de análise: A, Z, D e S. Dada a importância atribuída pelas análises ABC às famílias A e Z, posteriormente, recorreu-se ao método ABC para uma classificação por subfamílias para cada respetiva família, hierarquizando a aplicação de modelos de previsão a estas subfamílias. Ainda numa fase inicial, desenvolveu-se uma metodologia que permitiu definir procedimentos de previsão a aplicar às diferentes famílias e subfamílias segundo os graus de importância atribuídos pelas análises ABC. Consequentemente, foram definidas dez séries temporais para aplicação de modelos de previsão da procura no desenvolvimento do presente projeto de dissertação, representantes das vendas trimestrais em kg das subfamílias AK, AI, AM, AA, ZI, ZG, ZM, do conjunto das restantes subfamílias Z, do conjunto da família D e do conjunto da família S. Estas séries foram definidas por trimestre e subfamília dado que a empresa realiza encomendas trimestralmente por subfamílias e a previsão conjunta de produtos consegue trazer benefícios para a precisão das previsões.

O estudo dos modelos de previsão da procura associados ao caso Linhas, S.A. foi dividido em três fases: (1) seleção, implementação e avaliação dos modelos estatísticos de previsão da procura, em particular, modelos não causais; (2) seleção, implementação e avaliação do modelo baseado em inteligência artificial *ELM*; (3) comparação de resultados do modelo não causal eleito e do modelo *ELM*.

Na aplicação de modelos estatísticos de previsão da procura, constatou-se que todas as séries temporais apresentavam uma componente ciclo-tendência forte (à exceção da série representativa da subfamília ZG que apresentou uma componente ciclo-tendência em quase 30%, contudo a componente sazonalidade era superior a 45%). Dadas estas características, priorizou-se a implementação de modelos de alisamento exponencial para o caso em estudo.

Com a aplicação dos modelos de alisamento exponencial conseguiu-se, na generalidade, melhorias significativas em todas as medidas de avaliação de desempenho consideradas no estudo comparativamente ao modelo atualmente empregue pela Linhas, S.A., com exceção apenas da série temporal representativa do conjunto das restantes subfamílias Z, tendo-se considerado como melhor modelo a aplicar o sNaïve. Já para a série do conjunto da família D selecionou-se como melhor modelo estatístico o AES e para as restantes séries o *HW* aditivo.

Para a aplicação de modelos baseados em IA, selecionou-se como modelo a implementar o *ELM* dado que, entre os modelos tipicamente empregues na indústria têxtil-vestuário, este era aquele que, segundo a literatura revista, melhor se comportaria no que diz respeito à rapidez comparativamente aos demais modelos de IA e à precisão comparativamente aos modelos estatísticos. Contudo, desta aplicação conseguiu-se apenas melhorias nas medidas de avaliação de desempenho em três dos casos.

Salienta-se que, segundo a literatura revista, os modelos estatísticos têm a vantagem de rapidez comparativamente aos modelos baseados em IA e, estes últimos, a vantagem da precisão das previsões comparativamente aos primeiros. Contudo, considerando os resultados obtidos pelo estudo realizado, correr o algoritmo *ELM* foi conseguido em segundos, similarmente aos modelos estatísticos aplicados, além de que em 70% dos casos, o *ELM* não apresentou melhorias nas medidas de avaliação de desempenho e nos 30% dos casos em que apresentou, a aplicação do modelo *HW* aditivo já apresentava melhorias comparativamente ao modelo aplicado pela empresa. Acresce-se ainda que para um destes últimos casos considerou-se um período reduzido e, para os restantes, se se tivesse considerado o *holdout* três como referência em vez de seis, o modelo mais bem-comportado seria o *HW* aditivo.

Estes comportamentos podem dever-se, principalmente, à reduzida quantidade de variáveis consideradas para as previsões, dado que foram apenas consideradas as variáveis vendas e tempo. Assim, para casos em que não estão disponíveis variáveis além da variável em estudo e do tempo, pode-se preferir a aplicação de modelos não causais dada a sua simplicidade, maturidade e fácil compreensão e implementação comparativamente aos modelos de IA. Por outro lado, sempre que surgem novos dados, as *ELM* precisam de ser treinadas e os dados precisam de ser tratados. Além disso, realça-se que estas apresentam a desvantagem de instabilidade de resultados comparativamente modelos aos estatísticos.

De forma a dar continuidade ao trabalho realizado de definição de modelos de previsão da procura, sugere-se para trabalho futuro a consideração de outras variáveis independentes além do tempo para implementação de modelos de previsão, veja-se o caso de variáveis que reflitam indicadores económicos e sociais relevantes, comportamento dos concorrentes, fornecedores e clientes, entre outras. Para isso, sugere-se a implementação de modelos híbridos que combinem modelos estatísticos, não causais ou causais, com modelos de IA de forma a alcançar resultados com maior precisão, presando, ao mesmo tempo, pela rapidez de implementação.

Finalizado o estudo de modelos de previsão da procura, sugere-se a utilização dos resultados obtidos para o desenvolvimento de modelos de gestão de *stock* que permitam controlo dos mesmos, reduzindo os riscos de rutura e excesso de *stock* e aumentando o nível de serviço ao cliente.

## REFERÊNCIAS BIBLIOGRÁFICAS

- Armstrong, J. S. (2002). *Principles of Forecasting: A Handbook for Researchers and Practitioners*. New York: Frederick S. Hillier, Series Editor.
- Au, K. F., Choi, T. M., & Yu, Y. (2008). Fashion retail forecasting by evolutionary neural networks. *International Journal of Production Economics*, 114(2), 615-630. <https://doi.org/10.1016/j.ijpe.2007.06.013>
- Beheshti-Kashi, S., Karimi, H. R., Thoben, K. D., Lütjen, M., & Teucke, M. (2015). A survey on retail sales forecasting and prediction in fashion markets. *Systems Science and Control Engineering*, 3(1), 154-161. <https://doi.org/10.1080/21642583.2014.999389>
- Bhattacharya, A., Sarkar, B., & Mukherjee, S. K. (2007). Distance-based consensus method for ABC analysis. *International Journal of Production Research*, 45(15), 3405–3420. <https://doi.org/10.1080/00207540600847145>
- Caglayan, N., Satoglu, S. I., & Kapukaya, E. N. (2020). Sales forecasting by artificial neural networks for the apparel retail chain stores-An application. *Journal of Intelligent and Fuzzy Systems*, 39(5), 6517-6528. <https://doi.org/10.3233/JIFS-189115>
- Chaudhuri, K. D. (2021, dezembro 13). *Time Series Forecasting with Extreme Learning Machines*. Analytics Vidhya. <https://www.analyticsvidhya.com/blog/2021/12/time-series-forecasting-with-extreme-learning-machines/>
- Choi, T. M., Hui, C. L., & Yu, Y. (2011). Intelligent time series fast forecasting for fashion sales: A research agenda. *2011 International Conference on Machine Learning and Cybernetics*, 3(1), 1010-1014. <https://doi.org/10.1109/ICMLC.2011.6016870>
- Choi, T. M., Hui, C. L., Liu, N., Ng, S. F., & Yu, Y. (2014). Fast fashion sales forecasting with limited data and time. *Decision Support Systems*, 59(1), 84-92. <https://doi.org/10.1016/j.dss.2013.10.008>
- Choi, T. M., Hui, C. L., Ng, S. F., & Yu, Y. (2012). Color trend forecasting of fashionable products with very few historical data. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, 42(6), 1003-1010. <https://doi.org/10.1109/TSMCC.2011.2176725>
- Chopra, S., & Meindl, P. (2013). *Supply Chain Management: Strategy, Planning, and Operation* (5th ed.) Pearson Education.
- Frank, C., Garg, A., Raheja, A., Sztandera, L., & Raheja, A. (2003). Forecasting women's apparel sales using mathematical modeling. *International Journal of Clothing Science and Technology*, 15(2), 107-125. <https://doi.org/10.1108/09556220310470097>
- Heizer, J., Render, B., & Munson, C. (2017). *Principles of Operations Management: Sustainability and Supply Chain Management* (10th ed.). England: Pearson Education Limited.
- Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2004). Extreme learning machine: A new learning scheme of feedforward neural networks. *2004 IEEE International Conference on Neural Networks*, 2(1), 985–990. <https://doi.org/10.1109/IJCNN.2004.1380068>
- Huang, G. B., Zhu, Q. Y., & Siew, C. K. (2006). Extreme learning machine: Theory and applications. *Neurocomputing*, 70(1-3), 489–501. <https://doi.org/10.1016/j.neucom.2005.12.126>
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice* (2nd ed.). Melbourne: OTexts.
- INE. (2007). *Classificação Portuguesa das Actividades Económicas, Revisão 3*. [https://www.ine.pt/ine\\_novidades/semin/cae/CAE\\_REV\\_3.pdf](https://www.ine.pt/ine_novidades/semin/cae/CAE_REV_3.pdf)
- Kharfan, M., Chan, V. W. K., & Efendigil, T. F. (2021). A data-driven forecasting approach for newly launched seasonal products by leveraging machine-learning approaches. *Annals of Operations Research*, 303(1), 159-174. <https://doi.org/10.1007/s10479-020-03666-w>
- Liu, N., Ren, S., Choi, T. M., Hui, C. L., & Ng, S. F. (2013). Sales forecasting for fashion retailing service industry: A review. *Mathematical Problems in Engineering*, 2013(1).

- <https://doi.org/10.1155/2013/738675>
- Loureiro, A. L. D., Miguéis, V. L., & Silva, L. F. M. (2018). Exploring the use of deep neural networks for sales forecasting in fashion retail. *Decision Support Systems*, 114(), 81-93. <https://doi.org/10.1016/j.dss.2018.08.010>
- Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and Applications* (3rd ed.). New York: John Wiley & Sons.
- Moroff, N. U., Kurt, E., & Kamphues, J. (2021). Machine Learning and Statistics: A Study for assessing innovative Demand Forecasting Models. *Procedia Computer Science*, 180(), 40–49. <https://doi.org/10.1016/j.procs.2021.01.127>
- Nenni, M. E., Giustiniano, L., & Pirolo, L. (2013). Demand forecasting in the fashion industry: A review. *International Journal of Engineering Business Management*, 5(). <https://doi.org/10.5772/56840>
- Ni, Y., & Fan, F. (2011). A two-stage dynamic sales forecasting model for the fashion retail. *Expert Systems with Applications*, 38(3), 1529-1536. <https://doi.org/10.1016/j.eswa.2010.07.065>
- Reid, D. R., & Sanders, N. R. (2012). *Operations Management: An Integrated Approach* (5th ed.). USA: John Wiley & Sons, Inc.
- Reiff, N. (2022, julho 2). *Bullwhip Effect*. Infopedia. <https://www.investopedia.com/bullwhip-effect-definition-5499228>
- Ren, S., Chan, H. L., & Siqin, T. (2020). Demand forecasting in retail operations for fashionable products: methods, practices, and real case study. *Annals of Operations Research*, 291(), 761-777. <https://doi.org/10.1007/s10479-019-03148-8>
- Saunders M., Lewis P., & Thornhill A. (2009). *Research methods for business students*, 5th. Edition. Harlow: Prentice Hall.
- Spiliotis, E., Makridakis, S., Semenoglou, A. A., & Assimakopoulos, V. (2020). Comparison of statistical and machine learning methods for daily SKU demand forecasting. *Operational Research*, 22(), 3037-3061. <https://doi.org/10.1007/s12351-020-00605-2>
- Sun, Z. L., Choi, T. M., Au, K. F., & Yu, Y. (2008). Sales forecasting using extreme learning machine with applications in fashion retailing. *Decision Support Systems*, 46(1), 411-419. <https://doi.org/10.1016/j.dss.2008.07.009>
- Thomassey, S. (2010). Sales forecasts in clothing industry: The key success factor of the supply chain management. *International Journal of Production Economics*, 128(2), 470-483. <https://doi.org/10.1016/j.ijpe.2010.07.018>
- Thomassey, S. (2014). Sales forecasting in apparel and fashion industry: A review. *Intelligent Fashion Forecasting Systems: Models and Applications*, 9-27. [https://doi.org/10.1007/978-3-642-39869-8\\_2](https://doi.org/10.1007/978-3-642-39869-8_2)
- Twin, A. (2021, outubro 22). *Overfitting*. Infopedia. <https://www.investopedia.com/terms/o/overfitting.asp>
- Wang, J., Lu, S., Wang, S. H., & Zhang, Y. D. (2021). A review on extreme learning machine. *Multimedia Tools and Applications*. <https://doi.org/10.1007/s11042-021-11007-7>
- Wong, W. K., & Guo, Z. X. (2010). A hybrid intelligent model for medium-term sales forecasting in fashion retail supply chains using extreme learning machine and harmony search algorithm. *International Journal of Production Economics*, 128(2), 614–624. <https://doi.org/10.1016/j.ijpe.2010.07.008>
- Yu, Y., Choi, T. M., & Hui, C. L. (2011). An intelligent fast sales forecasting model for fashion products. *Expert Systems with Applications*, 38(6), 7373-7379. <https://doi.org/10.1016/j.eswa.2010.12.089>

## APÊNDICES

### APÊNDICE A

#### Análise ABC da Família A

**Tabela A 1**

*Análise ABC, por subfamília de A, em 2020*

<b>Subfamília</b>	<b>Total de vendas anuais (em kg)</b>	<b>Percentagem Individual</b>	<b>Percentagem Acumulada</b>	<b>Classificação ABC</b>
<b>AK</b>	491 439	93,93%	93,93%	A
<b>AI</b>	17 367	3,32%	97,25%	B
<b>AM</b>	8 639	1,65%	98,90%	B
<b>AA</b>	5 352	1,02%	99,92%	B
<b>AQ</b>	251	0,05%	99,97%	C
<b>AF</b>	170	0,03%	100%	C
<b>Total Geral</b>	523 218			

**Tabela A 2**

*Análise ABC, por subfamília de A, em 2021*

<b>Subfamília</b>	<b>Total de vendas anuais (em kg)</b>	<b>Percentagem Individual</b>	<b>Percentagem Acumulada</b>	<b>Classificação ABC</b>
<b>AK</b>	554 483	92,58%	92,58%	A
<b>AI</b>	22 769	3,80%	96,38%	B
<b>AM</b>	13 881	2,32%	98,69%	B
<b>AA</b>	7 417	1,24%	99,93%	B
<b>AQ</b>	373	0,06%	100%	C
<b>AF</b>	27	0,00%	100%	C
<b>Total Geral</b>	598 950			

## APÊNDICE B

### Análise ABC da Família Z

**Tabela B 1**

*Análise ABC, por subfamília de Z, em 2020*

Subfamília	Total de vendas anuais (em kg)	Percentagem Individual	Percentagem Acumulada	Classificação ABC
ZI	23 362	52,80%	52,80%	A
ZG+ZG(E)	8 310	18,78%	71,59%	A
ZM	4 505	10,18%	81,77%	B
ZE+ZE(E)	3 555	8,04%	89,81%	B
ZA+ZA(E)	1 987	4,49%	94,30%	B
ZQ+ZQ(E)	1 154	2,61%	96,91%	C
ZK	987	2,23%	99,14%	C
ZV	182	0,41%	99,55%	C
ZO	135	0,31%	99,85%	C
ZY	47	0,11%	99,96%	C
ZX	10	0,02%	99,98%	C
ZS	5	0,01%	99,99%	C
ZJ	3	0,01%	100,00%	C
<b>Total Geral</b>	76 235			

**Tabela B 2**

*Análise ABC, por subfamília de Z, em 2021*

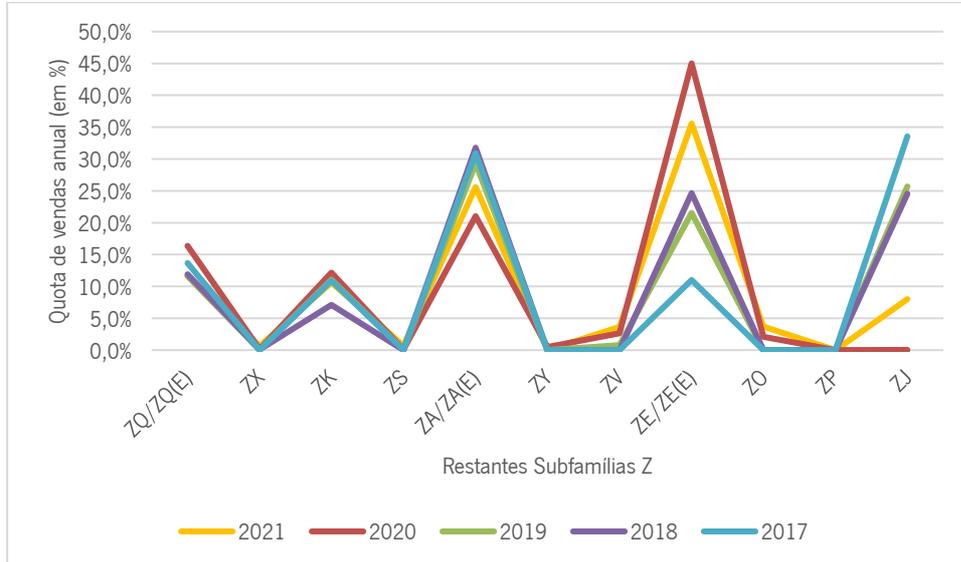
Subfamília	Total de vendas anuais (em kg)	Percentagem Individual	Percentagem Acumulada	Classificação ABC
ZI	37 068	48,62%	48,62%	A
ZG+ZG(E)	10 116	13,27%	61,89%	A
ZM	9 348	12,26%	74,15%	A
ZE+ZE(E)	6 947	9,11%	83,27%	B
ZA+ZA(E)	4 345	5,70%	88,97%	B
ZK	2 957	3,88%	92,85%	B
ZQ+ZQ(E)	2 379	3,12%	95,97%	C
ZJ	1 510	1,98%	97,95%	C
ZV	670	0,88%	98,83%	C
ZO	598	0,78%	99,61%	C
ZX	202	0,26%	99,87%	C
ZS	95	0,13%	100,00%	C
<b>Total Geral</b>	76 235			

## APÊNDICE C

### Evolução das quotas anuais por subfamília nos últimos 5 anos

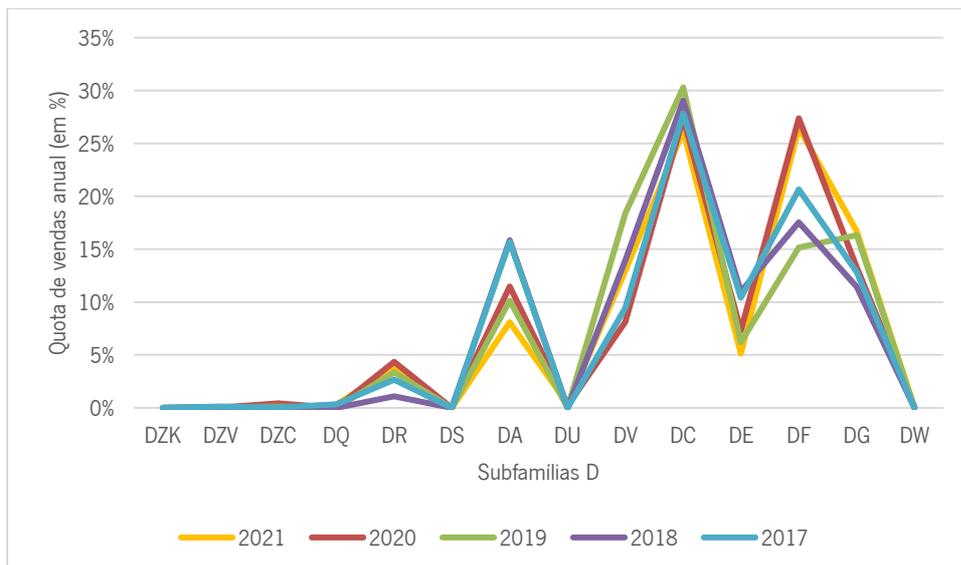
**Figura C 1**

*Evolução das quotas anuais nos últimos 5 anos das restantes subfamílias Z (excluindo-se ZI, ZG e ZM)*



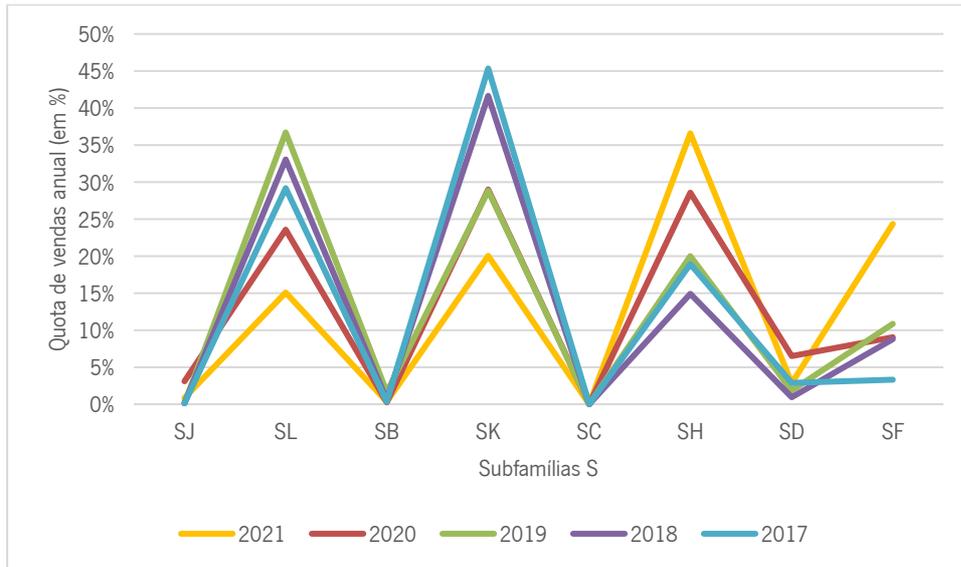
**Figura C 2**

*Evolução das quotas anuais nos últimos 5 anos de todas as subfamílias D*



**Figura C 3**

*Evolução das quotas anuais nos últimos 5 anos de todas as subfamílias S*

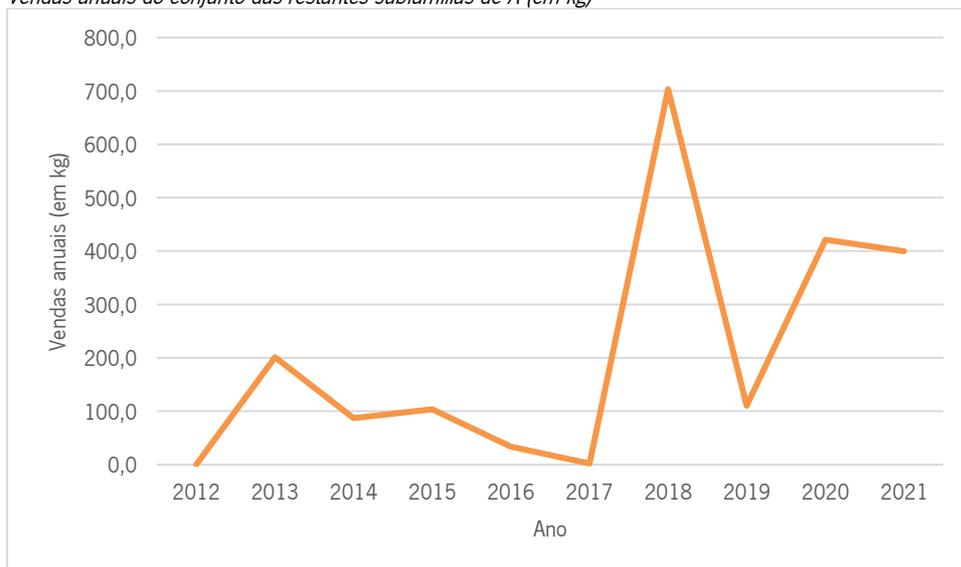


## APÊNDICE D

Série representativa das vendas em kg do conjunto das restantes subfamílias de A

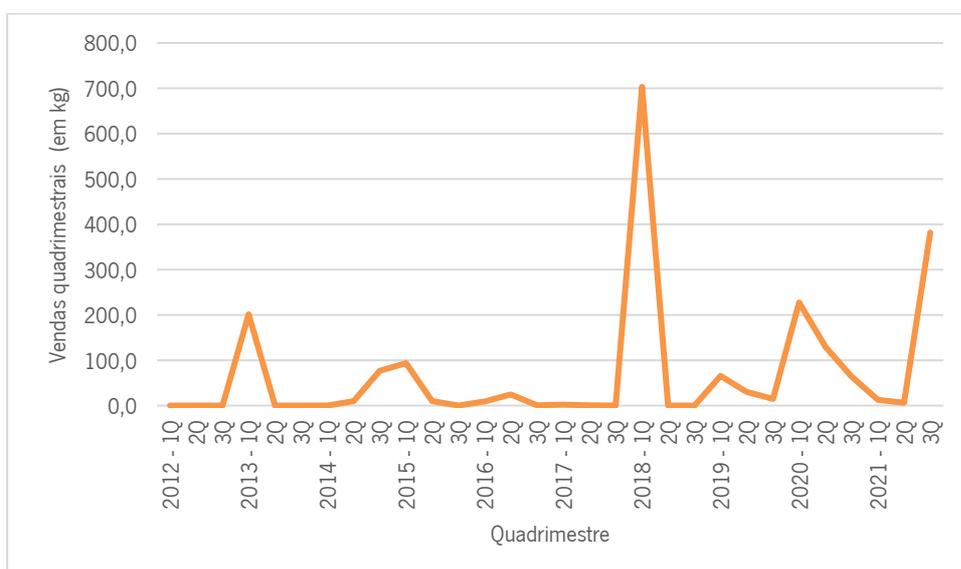
**Figura D 1**

*Vendas anuais do conjunto das restantes subfamílias de A (em kg)*



**Figura D 2**

*Vendas quadrimestrais do conjunto das restantes subfamílias de A (em kg)*



## APÊNDICE E

Taxas de variação associadas à escolha de modelos de AE em detrimento do modelo sNaïve, tendo como critério as medidas de avaliação de desempenho, para a série AK

	<b>Dentro da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>AES</b>	-55%	-58%	-60%	-55%	-61%	-61%
<b>Holt</b>	-55%	-59%	-59%	-55%	-61%	-60%
<b>HW multiplicativo</b>	-55%	-59%	-60%	-56%	-62%	-61%
<b>HW aditivo</b>	-55%	-59%	-61%	-56%	-62%	-63%
	<b>Fora da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>AES</b>	-33%	-38%	-39%	-23%	-18%	-25%
<b>Holt</b>	-39%	-41%	-55%	-33%	-25%	-30%
<b>HW multiplicativo</b>	-39%	-42%	-42%	-27%	-17%	-20%
<b>HW aditivo</b>	-39%	-44%	-44%	-29%	-20%	-24%

## APÊNDICE F

Resultados da aplicação de modelos *ARIMA* à subfamília AK

**Figura F 1**

*Resultados do modelo ARIMA recomendado pela software para a série AK: ARIMA (0,1,0)*

**Forecast Model for AK\_UendasKG**  
**ARIMA(0,1,0)**

### Within-Sample Statistics

Sample size 27	Number of parameters 0
Mean 1.208e+005	Standard deviation 4.744e+004
R-square 0.7934	Adjusted R-square 0.8011
Durbin-Watson 2.047	Ljung-Box(18)=11.97 P=0.151
Forecast error 2.116e+004	BIC 2.116e+004
MAPE 0.2191	RMSE 2.116e+004
MAD 1.412e+004	

### Out-of-Sample Rolling Evaluation

H	N	MAD	Cumulative Average	MAPE	Cumulative Average	GMRAE	Cumulative Average
1	3	17401	17401	0.091	0.091	1.000	1.000
2	2	14494	16238	0.076	0.085	1.000	1.000
3	1	24746	17656	0.122	0.091	1.000	1.000

*Nota.* Resultados obtidos através do *Software Forecast Pro*

**Figura F 2**

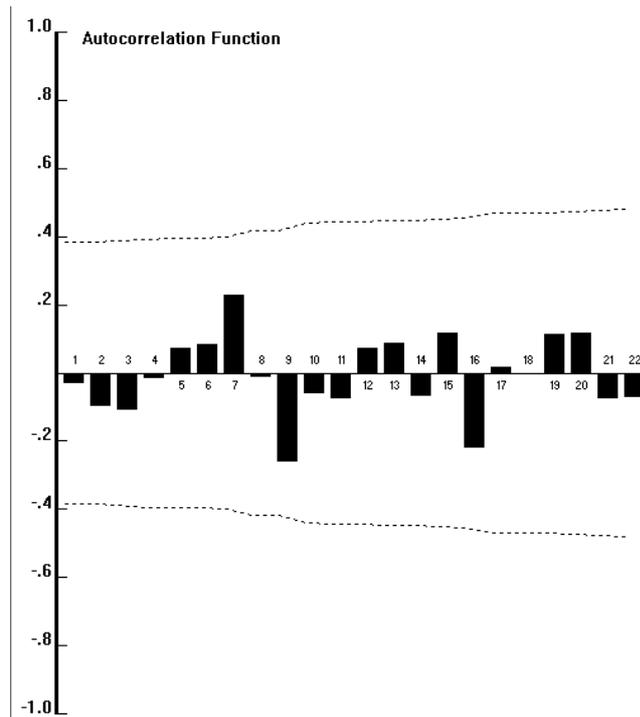
*Teste F para igualdade de variâncias*

Teste F: duas amostras para variâncias		
	Variável 1	Variável 2
Média	120833,249	3280,006156
Variância	2250105328	472173482,8
Observações	27	26
gl	26	25
F	4,76542078	
P(F<=f) uni	0,00010334	
F crítico uni	1,94719913	

*Nota.* Resultados obtidos através do *Microsoft Excel*

**Figura F 3**

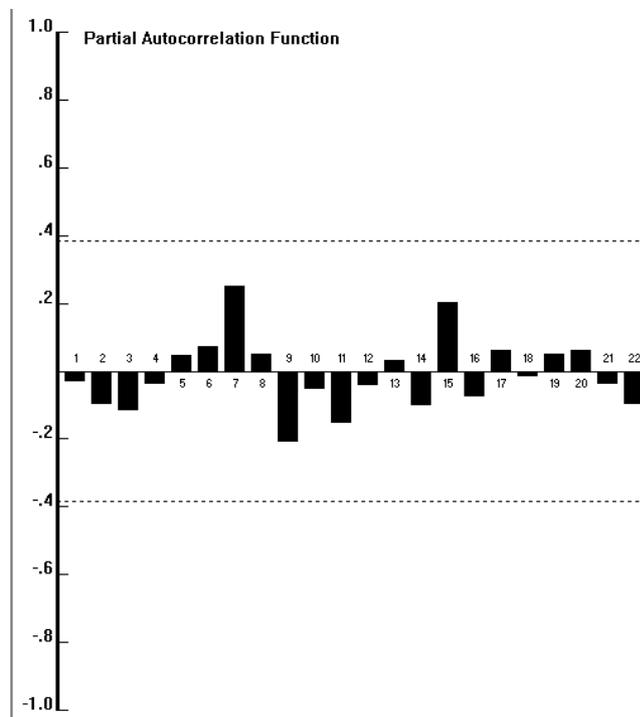
Função de autocorrelações para o modelo ARIMA (0,1,0) sem transformações aplicados à série AK



Nota. Gráfico obtido através do *Software Forecast Pro*

**Figura F 4**

Função de autocorrelações parciais para o modelo ARIMA (0,1,0) sem transformações aplicados à série AK



Nota. Gráfico obtido através do *Software Forecast Pro*

**Figura F 5***Resultados do modelo ARIMA (1,1,1) aplicado à série AK***Forecast Model for AK\_UendasKG  
ARIMA(1,1,1)**

Term	Coefficient	Std. Error	t-Statistic	Significance
a[1]	0.5419	1.4118	0.3838	0.2955 <-
b[1]	0.6266	1.3234	0.4735	0.3599 <-

Try alternative model ARIMA(0,1,1)

**Within-Sample Statistics**

Sample size 27	Number of parameters 2
Mean 1.208e+005	Standard deviation 4.744e+004
R-square 0.7955	Adjusted R-square 0.7873
Durbin-Watson 1.904	Ljung-Box(18)=13.92 P=0.2658
Forecast error 2.188e+004	BIC 2.378e+004
MAPE 0.2338	RMSE 2.105e+004
MAD 1.462e+004	

**Out-of-Sample Rolling Evaluation**

H	N	MAD	Cumulative Average	MAPE	Cumulative Average	GMRAE	Cumulative Average
1	3	17300	17300	0.091	0.091	1.048	1.048
2	2	14425	16150	0.075	0.084	0.920	0.995
3	1	27163	17985	0.134	0.092	1.098	1.011

*Nota. Resultados obtidos através do Software Forecast Pro***Figura F 6***Resultados do modelo ARIMA (0,1,1) aplicado à série AK***Forecast Model for AK\_UendasKG  
ARIMA(0,1,1)**

Term	Coefficient	Std. Error	t-Statistic	Significance
b[1]	0.0394	0.1980	0.1988	0.1560 <-

Insignificant MA terms are harmless.

**Within-Sample Statistics**

Sample size 27	Number of parameters 1
Mean 1.208e+005	Standard deviation 4.744e+004
R-square 0.7937	Adjusted R-square 0.7937
Durbin-Watson 1.976	Ljung-Box(18)=13.34 P=0.2291
Forecast error 2.154e+004	BIC 2.247e+004
MAPE 0.2254	RMSE 2.114e+004
MAD 1.425e+004	

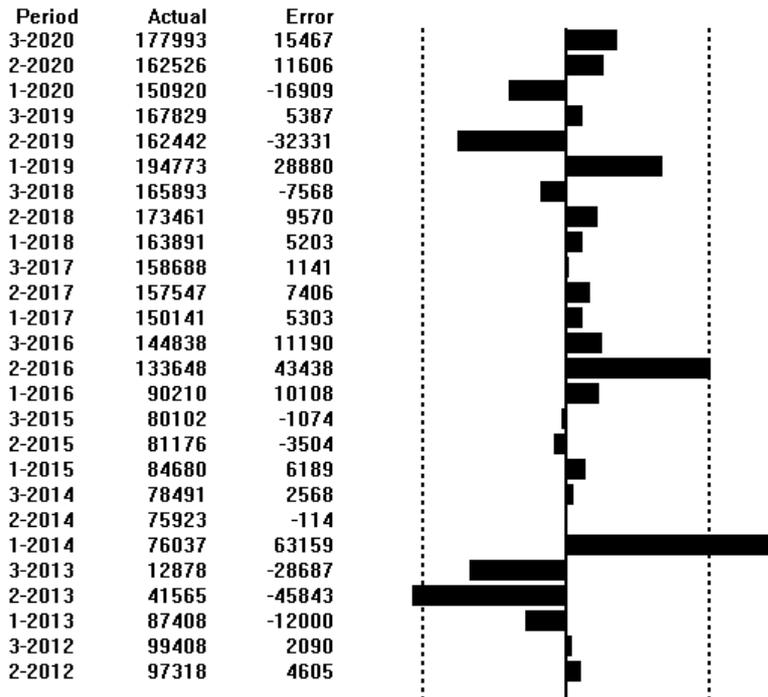
**Out-of-Sample Rolling Evaluation**

H	N	MAD	Cumulative Average	MAPE	Cumulative Average	GMRAE	Cumulative Average
1	3	17362	17362	0.091	0.091	1.031	1.031
2	2	14286	16132	0.075	0.084	0.970	1.006
3	1	25372	17672	0.125	0.091	1.025	1.009

*Nota. Resultados obtidos através do Software Forecast Pro*

**Figura F 7**

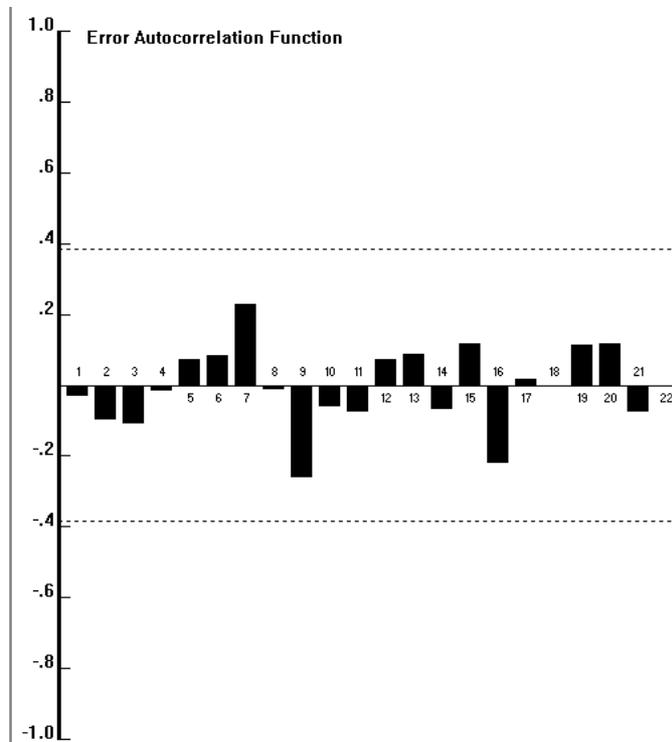
Análise dos erros do modelo ARIMA (0,1,0) aplicado à série AK



Nota. Resultados obtidos através do Software Forecast Pro

**Figura F 8**

Função de autocorrelação dos erros do modelo ARIMA (0,1,0) aplicado à série AK



Nota. Resultados obtidos através do Software Forecast Pro

## APÊNDICE G

Taxas de variação associadas à escolha de modelos de AE em detrimento do modelo sNaïve, tendo como critério as medidas de avaliação de desempenho, para a série AI

	<b>Dentro da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>AES</b>	-48%	-54%	-66%	-53%	-57%	-67%
<b>Holt</b>	-48%	-54%	-66%	-53%	-57%	-67%
<b>HW multiplicativo</b>	-48%	-54%	-66%	-53%	-57%	-67%
<b>HW aditivo</b>	-49%	-54%	-65%	-53%	-57%	-67%
	<b>Fora da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>AES</b>	-33%	-31%	-33%	-25%	-34%	-37%
<b>Holt</b>	-33%	-32%	-33%	-25%	-34%	-37%
<b>HW multiplicativo</b>	-33%	-31%	-33%	-25%	-35%	-37%
<b>HW aditivo</b>	-44%	-42%	-44%	-25%	-34%	-37%

## APÊNDICE H

Taxas de variação associadas à escolha de modelos de AE em detrimento do modelo sNaïve, tendo como critério as medidas de avaliação de desempenho, para a série AM

	<b>Dentro da amostra</b>					
	<b><i>Holdout = 3</i></b>			<b><i>Holdout = 6</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>AES</b>	-29%	-35%	-37%	-22%	-32%	-37%
<b>Holt</b>	-31%	-30%	-15%	-24%	-29%	-25%
<b>HW multiplicativo</b>	-31%	-30%	-15%	-29%	-33%	-28%
<b>HW aditivo</b>	-31%	-29%	-15%	-26%	-33%	-27%
	<b>Fora da amostra</b>					
	<b><i>Holdout = 3</i></b>			<b><i>Holdout = 6</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>AES</b>	-24%	-10%	-1%	-39%	-41%	-46%
<b>Holt</b>	-31%	-24%	-19%	-46%	-49%	-50%
<b>HW multiplicativo</b>	-31%	-24%	-19%	-47%	-49%	-41%
<b>HW aditivo</b>	-32%	-24%	-19%	-47%	-49%	-46%

## APÊNDICE I

Taxas de variação associadas à escolha de modelos de AE em detrimento do modelo sNaïve, tendo como critério as medidas de avaliação de desempenho, para a série AA

	<b>Dentro da amostra</b>					
	<b><i>Holdout = 3</i></b>			<b><i>Holdout = 6</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>AES</b>	-51%	-46%	-51%	-56%	-51%	-55%
<b>Holt</b>	-51%	-46%	-51%	-56%	-51%	-55%
<b>HW multiplicativo</b>	-50%	-46%	-56%	-55%	-46%	-54%
<b>HW aditivo</b>	-51%	-46%	-55%	-56%	-51%	-55%
	<b>Fora da amostra</b>					
	<b><i>Holdout = 3</i></b>			<b><i>Holdout = 6</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>AES</b>	-14%	-29%	-36%	14%	-4%	-19%
<b>Holt</b>	-15%	-29%	-36%	14%	-4%	-19%
<b>HW multiplicativo</b>	-30%	-45%	-50%	-7%	-18%	-24%
<b>HW aditivo</b>	-28%	-42%	-46%	14%	-4%	-19%

## APÊNDICE J

Taxas de variação associadas à escolha de modelos de AE em detrimento do modelo sNaïve, tendo como critério as medidas de avaliação de desempenho, para a série ZI

	<b>Dentro da amostra</b>		
	<i>Holdout = 3</i>		
	<i><b>RMSE</b></i>	<i><b>MAD</b></i>	<i><b>MAPE</b></i>
<b>AES</b>	-25%	-27%	0%
<b>Holt</b>	-35%	-36%	8%
<b>HW multiplicativo</b>	-48%	-53%	-35%
<b>HW aditivo</b>	-65%	-71%	-68%
	<b>Fora da amostra</b>		
	<i>Holdout = 3</i>		
	<i><b>RMSE</b></i>	<i><b>MAD</b></i>	<i><b>MAPE</b></i>
<b>AES</b>	-18%	-14%	-15%
<b>Holt</b>	-38%	-38%	-38%
<b>HW multiplicativo</b>	-41%	-49%	-51%
<b>HW aditivo</b>	-63%	-62%	-62%

## APÊNDICE K

Taxas de variação associadas à escolha de modelos de AE em detrimento do modelo sNaïve, tendo como critério as medidas de avaliação de desempenho, para a série ZG

	<b>Dentro da amostra</b>		
	<i>Holdout = 3</i>		
	<i><b>RMSE</b></i>	<i><b>MAD</b></i>	<i><b>MAPE</b></i>
<b>AES</b>	-32%	-35%	-13%
<i><b>Holt</b></i>	-33%	-33%	-8%
<i><b>HW multiplicativo</b></i>	-48%	-55%	-47%
<i><b>HW aditivo</b></i>	-48%	-55%	-47%
	<b>Fora da amostra</b>		
	<i>Holdout = 3</i>		
	<i><b>RMSE</b></i>	<i><b>MAD</b></i>	<i><b>MAPE</b></i>
<b>AES</b>	97%	67%	-3%
<i><b>Holt</b></i>	71%	51%	-6%
<i><b>HW multiplicativo</b></i>	24%	34%	1%
<i><b>HW aditivo</b></i>	18%	20%	-15%

## APÊNDICE L

Medidas de avaliação de desempenho dos modelos de previsão aplicados à série ZM com dados desde 2012

	<b>Dentro da amostra</b>							
	<b>Holdout = 3</b>				<b>Holdout = 6</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	-	-	-	-	-	-	-	-
<b>AES</b>	393,6	290,2	0,4122	418,4	253,8	202,7	0,3947	271,2
<b>Holt</b>	360,7	250,5	1,9300	407,5	217,2	186,4	1,782	247,9
<b>HW multiplicativo</b>	nd	nd	nd	nd	nd	nd	nd	nd
<b>HW aditivo</b>	360,6	250	1,9170	433,1	214,5	182,9	1,812	261,7
	<b>Fora da amostra</b>							
	<b>Holdout = 3</b>				<b>Holdout = 6</b>			
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>BIC</b>
<b>Método Empresa</b>	1 828	1 614	0,4970	-	995	937	0,5992	-
<b>AES</b>	1 725	1 476	0,4263	-	1 110	991	0,5753	-
<b>Holt</b>	1 252	964	0,2573	-	1 168	1 048	0,6689	-
<b>HW multiplicativo</b>	nd	nd	nd	-	nd	nd	nd	-
<b>HW aditivo</b>	1 251	963	0,2566	-	1 169	1 039	0,6679	-

*Nota.* Não foi possível obter valores para o modelo HW multiplicativo. Resultados obtidos através do *Software Forecast Pro*

## APÊNDICE M

Taxas de variação associadas à escolha de modelos de AE em detrimento do modelo sNaïve, tendo como critério as medidas de avaliação de desempenho, para a série ZM

	<b>Dentro da amostra</b>					
	<b><i>Holdout = 3</i></b>			<b><i>Holdout = 6</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>AES</b>	-22%	-27%	-34%	-44%	-45%	-51%
<b>Holt</b>	-27%	-41%	-31%	-57%	-57%	-47%
<b>HW multiplicativo</b>	-27%	-41%	-31%	-60%	-64%	-53%
<b>HW aditivo</b>	-27%	-41%	-31%	-63%	-66%	-68%
	<b>Fora da amostra</b>					
	<b><i>Holdout = 3</i></b>			<b><i>Holdout = 6</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>AES</b>	-6%	-9%	-14%	12%	6%	-4%
<b>Holt</b>	-31%	-40%	-48%	55%	52%	45%
<b>HW multiplicativo</b>	-31%	-40%	-48%	59%	53%	49%
<b>HW aditivo</b>	-31%	-40%	-48%	21%	13%	15%

## APÊNDICE N

### Aplicação do método de decomposição multiplicativa à série temporal da subfamília ZM

#### Figura N 1

Aplicação do método de decomposição multiplicativa à série temporal da subfamília ZM com dados desde 2012

```
Expert data exploration of dependent variable ZM_UendasKG
-----
Length 30 Minimum 0 Maximum 4243
Mean 1183 Standard deviation 978

Classical decomposition (additive)
Trend-cycle: 60.34% Seasonal: 2.93% Irregular: 36.73%

Nonpositive series. Multiplicative seasonality ruled out.
Series is too short to consider Box-Jenkins.

Series is trended and seasonal.

Recommended model: Exponential Smoothing

Forecast Model for ZM_UendasKG
Simple exponential smoothing: No trend, No seasonality

Component          Smoothing      Final
                   Weight         Value
-----
Level              0.84375        4029.5

Within-Sample Statistics
-----
Sample size 30          Number of parameters 1
Mean 1183              Standard deviation 995
R-square 0.7149        Adjusted R-square 0.7149
Durbin-Watson 1.809   Ljung-Box(18)=6.59 P=0.006849
Forecast error 531.3   BIC 552.8
MAPE 0.4279           RMSE 522.3
MAD 363.4
```

Nota. Resultados obtidos através do Software Forecast Pro

#### Figura N 2

Aplicação do método de decomposição multiplicativa à série temporal da subfamília ZM com dados desde 2014

```
Expert data exploration of dependent variable ZM_UendasKG
-----
Length 24 Minimum 169 Maximum 4243
Mean 1431 Standard deviation 935

Classical decomposition (multiplicative)
Trend-cycle: 92.40% Seasonal: 0.94% Irregular: 6.67%

Log transform recommended for Box-Jenkins.

Series is too short to consider Box-Jenkins.

Series is trended and nonseasonal.

Recommended model: Exponential Smoothing

Forecast Model for ZM_UendasKG
Holt exponential smoothing: Linear trend, No seasonality
Confidence limits proportional to level

Component          Smoothing      Final
                   Weight         Value
-----
Level              0.14224        2773.1
Trend              0.06027         115.97

Within-Sample Statistics
-----
Sample size 24          Number of parameters 2
Mean 1431              Standard deviation 954.9
R-square 0.6773        Adjusted R-square 0.6626
Durbin-Watson 1.202   Ljung-Box(17)=10.15 P=0.1029
Forecast error 554.7   BIC 606.3
MAPE 0.2945           RMSE 531.1
MAD 342.1
```

Nota. Resultados obtidos através do Software Forecast Pro

## APÊNDICE O

Taxas de variação associadas à escolha de modelos de AE em detrimento do modelo sNaive, tendo como critério as medidas de avaliação de desempenho, para a série do conjunto das restantes subfamílias de Z

	<b>Dentro da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>AES</b>	-15%	-14%	-23%	-12%	-12%	-22%
<b>Holt</b>	-15%	-12%	-14%	-18%	-19%	-13%
<b>HW multiplicativo</b>	-21%	-19%	-24%	-23%	-24%	-19%
<b>HW aditivo</b>	-20%	-19%	-21%	-23%	-24%	-17%
	<b>Fora da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>AES</b>	-2%	4%	8%	95%	39%	-23%
<b>Holt</b>	-2%	2%	5%	80%	44%	-7%
<b>HW multiplicativo</b>	-4%	4%	11%	72%	34%	-14%
<b>HW aditivo</b>	-2%	9%	19%	73%	35%	-13%

## APÊNDICE P

Aplicação do método de decomposição multiplicativa à série temporal da família D

-----  
Length 30 Minimum 9423 Maximum 36669  
Mean 18165 Standard deviation 5616

Classical decomposition (multiplicative)  
Trend-cycle: 72.55% Seasonal: 5.02% Irregular: 22.43%

Log transform recommended for Box-Jenkins.

Stationary series requires Box-Jenkins.

Series is trended and seasonal.

Recommended model: Box-Jenkins

Forecast Model for Deltas\_UendaskG  
ARIMA(0,0,1) with log transform

Term	Coefficient	Std. Error	t-Statistic	Significance
b[1]	-0.5637	0.1518	-3.7143	0.9991
_CONST	9.7649			

-----  
Within-Sample Statistics

Sample size 30	Number of parameters 1
Mean 9.765	Standard deviation 0.29
R-square 0.2473	Adjusted R-square 0.2473
Durbin-Watson 2.01	Ljung-Box(18)=9.241 P=0.0461
Forecast error 0.2516	BIC 4559
MAPE 0.1932	RMSE 4964
MAD 3582	

Nota: Dados obtidos através do *Software Forecast Pro*

## APÊNDICE Q

Taxas de variação associadas à escolha de modelos de AE em detrimento do modelo sNaïve, tendo como critério as medidas de avaliação de desempenho, para a série da família D

	<b>Dentro da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>AES</b>	-50%	-39%	-33%	-51%	-37%	-30%
<b>Holt</b>	-51%	-37%	-32%	-58%	-44%	-40%
<b>HW multiplicativo</b>	-60%	-49%	-37%	-71%	-62%	-51%
<b>HW aditivo</b>	-58%	-48%	-37%	-68%	-60%	-49%
	<b>Fora da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>AES</b>	-22%	-19%	-19%	-41%	-43%	-38%
<b>Holt</b>	-10%	-7%	-6%	20%	25%	28%
<b>HW multiplicativo</b>	-7%	-4%	-3%	12%	12%	19%
<b>HW aditivo</b>	15%	15%	15%	10%	10%	18%

## APÊNDICE R

Aplicação do método de decomposição multiplicativa à série temporal da família S

-----  
Length 30 Minimum 176 Maximum 25240  
Mean 5052 Standard deviation 5345

Classical decomposition (multiplicative)  
Trend-cycle: 91.12% Seasonal: 0.99% Irregular: 7.90%

Square root transform recommended for Box-Jenkins.

Series is too short to consider Box-Jenkins.

Series is trended and nonseasonal.

Recommended model: Exponential Smoothing

Forecast Model for Sigmas\_UendaskG  
Additive Winters: Linear trend, Additive seasonality

Component	Smoothing Weight	Final Value
Level	0.65662	23904.
Trend	0.99771	6905.7
Seasonal	0.15192	

Seasonal Indexes

-----  
Periods 1-3 -95.041 -406.49 501.53

Nota: Dados obtidos através do *Software Forecast Pro*

## APÊNDICE S

Taxas de variação associadas à escolha de modelos de AE em detrimento do modelo sNaïve, tendo como critério as medidas de avaliação de desempenho, para a série da família S

	<b>Dentro da amostra</b>					
	<b><i>Holdout = 3</i></b>			<b><i>Holdout = 6</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>AES</b>	-9%	-25%	-35%	-33%	-38%	-48%
<b>Holt</b>	-28%	-42%	-35%	-45%	-54%	-48%
<b>HW multiplicativo</b>	-35%	-41%	-23%	-46%	-54%	-55%
<b>HW aditivo</b>	-32%	-41%	-38%	-46%	-53%	-47%
	<b>Fora da amostra</b>					
	<b><i>Holdout = 3</i></b>			<b><i>Holdout = 6</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>AES</b>	-2%	-9%	-14%	54%	46%	12%
<b>Holt</b>	-7%	-12%	-17%	29%	20%	-5%
<b>HW multiplicativo</b>	-3%	-6%	-10%	28%	19%	-8%
<b>HW aditivo</b>	-7%	-12%	-16%	28%	19%	-6%

## APÊNDICE T

Taxas de variação associadas à escolha do método *ELM* em detrimento do sNaïve (Método Empresa) e modelo estatístico anteriormente selecionado, tendo como critério as medidas de avaliação de desempenho para fora da amostra

**Tabela T 1**

*Taxas de variação associadas à série AK*

	<b>Fora da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>Método Empresa</b>	-54%	-66%	-71%	-16%	-19%	1%
<b>HW aditivo</b>	-25%	-39%	-48%	19%	1%	32%

**Tabela T 2**

*Taxas de variação associadas à série AI*

	<b>Fora da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>Método Empresa</b>	-23%	-38%	-51%	-16%	-53%	-71%
<b>HW aditivo</b>	36%	7%	-13%	12%	-29%	-54%

**Tabela T 3**

*Taxas de variação associadas à série AM*

	<b>Fora da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>Método Empresa</b>	15%	9%	75%	27%	-9%	57%
<b>HW aditivo</b>	68%	44%	115%	137%	79%	188%

**Tabela T 4**

*Taxas de variação associadas à série AA*

	<b>Fora da amostra</b>					
	<b>Holdout = 3</b>			<b>Holdout = 6</b>		
	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>	<b>RMSE</b>	<b>MAD</b>	<b>MAPE</b>
<b>Método Empresa</b>	1%	-44%	-38%	21%	-10%	3%
<b>HW aditivo</b>	40%	-4%	14%	6%	-6%	28%

**Tabela T 5***Taxas de variação associadas à série ZI*

	<b>Fora da amostra</b>		
	<b><i>Holdout = 3</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>Método Empresa</b>	-79%	-90%	-56%
<b><i>HW</i> aditivo</b>	-43%	-74%	16%

**Tabela T 6***Taxas de variação associadas à série ZG*

	<b>Fora da amostra</b>		
	<b><i>Holdout = 3</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>Método Empresa</b>	2%	208%	92%
<b><i>HW</i> aditivo</b>	-14%	158%	126%

**Tabela T 7***Taxas de variação associadas à série ZM*

	<b>Fora da amostra</b>					
	<b><i>Holdout = 3</i></b>			<b><i>Holdout = 6</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>Método Empresa</b>	-31%	-45%	-35%	1%	-36%	-22%
<b><i>HW</i> aditivo</b>	1%	-9%	26%	-16%	-43%	-33%

**Tabela T 8***Taxas de variação associadas à série do conjunto das restantes subfamílias Z*

	<b>Fora da amostra</b>					
	<b><i>Holdout = 3</i></b>			<b><i>Holdout = 6</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>Método Empresa</b>	-19%	-13%	22%	101%	0%	19%

**Tabela T 9***Taxas de variação associadas à série da família D*

	<b>Fora da amostra</b>					
	<b><i>Holdout = 3</i></b>			<b><i>Holdout = 6</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>Método Empresa</b>	-24%	-22%	-28%	-29%	-76%	-25%

<b>AES</b>	-3%	-3%	-11%	21%	-57%	21%
------------	-----	-----	------	-----	------	-----

**Tabela T 10**

*Taxas de variação associadas à série da família S*

	<b>Fora da amostra</b>					
	<b><i>Holdout = 3</i></b>			<b><i>Holdout = 6</i></b>		
	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>	<b><i>RMSE</i></b>	<b><i>MAD</i></b>	<b><i>MAPE</i></b>
<b>Método Empresa</b>	7%	3%	-9%	33%	30%	13%
<b><i>HW</i> aditivo</b>	15%	17%	8%	4%	8%	21%

## ANEXO

### Código de aplicação de ELM

```
▶ def init_ELM(input_size, num_neurons=50):
    hidden_size = num_neurons # no. of hidden neurons
    mu, sigma = 0, 1
    w_lo = -1
    w_hi = 1
    b_lo = -1
    b_hi = 1
    np.random.seed(seed=2022)
    #initialising input weights and biases randomly drawn from a truncated normal distribution
    input_weights = stats.truncnorm.rvs(
        (w_lo - mu) / sigma, (w_hi - mu) / sigma, loc=mu, scale=sigma, size=[input_size, hidden_size])
    biases = stats.truncnorm.rvs(
        (b_lo - mu) / sigma, (b_hi - mu) / sigma, loc=mu, scale=sigma, size=[hidden_size])
    return input_weights, biases

def relu(x): # hidden layer activation function
    return np.maximum(x, 0, x)

def sigmoid(x):
    return 1/(1 + np.exp(-x))

[ ] def hidden_nodes(X, input_weights, biases):
    G = np.dot(X, input_weights)
    G = G + biases
    H = relu(G)
    return H

[ ] def train(X_train, y_train, neurons, window_size):
    input_weights, biases=init_ELM(window_size,neurons)
    H = hidden_nodes(X_train, input_weights, biases)
    return np.dot(pinv(H), y_train), input_weights, biases

[ ] def predict(X_test, output_weights, input_weights, biases):
    out = hidden_nodes(X_test, input_weights, biases)
    out = np.dot(out, output_weights)
    return out

[ ] def calculate_metrics(y_test, prediction, total):
    rmse = math.sqrt(mean_squared_error(y_test, prediction))
    mape_sum = 0
    #map
    for i, j in zip(y_test, prediction):
        mape_sum = mape_sum + (abs((i-j)/i))
    mape = (mape_sum/total)*100
    #mad
    mad = np.mean(np.absolute(prediction - np.mean(y_test)))
    return rmse,mape,mad
```

```

[ ] def groupby_sum_quarter(prediction,split,family,holdout):
    #group by Qi

    df = pd.DataFrame(prediction)
    df.join(split[split.Familia == family].Data.reset_index())
    df = df.join(split[split.Familia ==
                    family].Data.reset_index())

    #sum prediction by Qi
    bymonth = np.array(df.groupby(
        [df.Data.dt.year, df.Data.dt.month])[0].sum())
    list(map(np.sum, np.array(bymonth).reshape((holdout,4))))
    #sum real by Qi

    real = split[split.Familia == family]
    bymonth_real = np.array(real.groupby(
        [real.Data.dt.year, real.Data.dt.month]).QuilosVendidosFinal.sum())
    list(map(np.sum, np.array(bymonth_real).reshape((holdout,4))))

    #prediction grouped
    y = list(map(np.sum, np.array(bymonth_real).reshape((holdout,4))))
    prediction = list(map(np.sum, np.array(bymonth).reshape((holdout, 4))))

    return y, prediction

[ ] def train_and_predict_by_family(test_split, train_split, holdout, window_size, neurons):
    to_keep = train_split.Familia.unique()
    families = []
    for family in to_keep:
        #train with data just after 2018 for both families
        if holdout == 6 and (family in ["ZI", "ZG"]):
            continue

        if family == "ZI":
            window_size = 9
            neurons = 5000
        elif family == "ZG":
            window_size = 9

        #split train
        X_train, mean_train, std_train = filter_normalize_train(
            train_split, family)
        X_train_windows, y_train = test_or_train_split(X_train, window_size)
        #split test
        X_test = filter_normalize_test(
            test_split, family, mean_train, std_train)
        X_test = np.append(X_train_windows[-window_size:], X_test)
        X_test, y_test = test_or_train_split(X_test, window_size)

        if family not in ["ZI"]:
            neurons = (len(y_test)//3)
        #predict OUT sample
        output_weights, input_weights, biases = train(
            X_train_windows, y_train, neurons, window_size)

        prediction = predict(X_test, output_weights, input_weights, biases)
        y_test = (y_test*std_train) + mean_train
        prediction = (prediction*std_train) + mean_train
        out_sample_y_byQ, out_sample_pred_byQ = groupby_sum_quarter(prediction, test_split, family,holdout)
        # out SAMPLE metrics
        out_sample_rmse, out_sample_mape, out_sample_mad = calculate_metrics(
            out_sample_y_byQ, out_sample_pred_byQ, X_test.shape[0])

        families.append([family, int(out_sample_rmse),
                        int(out_sample_mad), out_sample_mape]+out_sample_pred_byQ)
    return families

```

Holdout = 3

```
[ ] train_split_holdout3 = grouped_df[grouped_df.Serie.isin(
    [year for year in range(2012, 2021)]]]
test_split_holdout3 = grouped_df[grouped_df.Serie.isin([2021])]
```

```
[ ] results_holdout_3= train_and_predict_by_family(
    test_split_holdout3, train_split_holdout3, holdout=3, window_size=10, neurons=50)
```

```
[ ] df_2021 = pd.DataFrame(results_holdout_3, columns=[
    "Familia", "RMSE", "MAD", "MAPE", "2021 - 1Q", "2021 - 2Q", "2021 - 3Q"])
df_2021
```

Holdout = 6

```
[ ] train_split_holdout6 = grouped_df[grouped_df.Serie.isin(
    [year for year in range(2012, 2020)]]]
test_split_holdout6 = grouped_df[grouped_df.Serie.isin([2020, 2021])]
```

```
[ ] results_holdout_6=train_and_predict_by_family(test_split_holdout6, train_split_holdout6, holdout=6,window_size= 10,neurons=100)
```

```
[ ] df_2020_2021 = pd.DataFrame(results_holdout_6, columns=[
    "Familia", "RMS", "MAP", "MAD", "2020 - 1Q", "2020 - 2Q", "2020 - 3Q", "2021 - 1Q", "2021 - 2Q", "2021 - 3Q"])
df_2020_2021
```

*Nota.* Retirado de Chaudhuri (2021)