

Otimização do método de previsão de vendas no setor agroalimentar

Inês Isabel Ferreira Viegas

Dissertação de Mestrado

Orientador na FEUP: Prof. Maria Teresa Peixoto Braga Bianchi de Aguiar



Mestrado Integrado em Engenharia Mecânica

2020-06-29

A ti, avó Lina.

Resumo

O projeto desenvolvido no âmbito desta dissertação de mestrado foi realizado na Cerealis, na Maia, inserindo-se no Departamento Logístico, na área de Planeamento Tático e Aprovisionamentos, tendo como objetivo a otimização dos métodos de previsão de amortecimento exponencial através da reparametrização das constantes de amortecimento e da agregação dos artigos em grupos adequados, para posterior implementação na ferramenta de planeamento de produção SO99+.

Para tal, foram realizadas análises detalhadas aos históricos de vendas de cada artigo da empresa, de forma a ser realizada uma caracterização adequada dos mesmos. Esta caracterização permitiu uma modelação correta dos dados .

Posteriormente, foram desenvolvidos programas em linguagem *Python*, de forma a se realizarem simulações de vários cenários de previsões, onde foram testados três métodos de previsão de vendas por amortecimento exponencial. Para cada um dos métodos foram testados também diferentes valores das constantes de amortecimento, por forma a selecionar a combinação que levaria a previsões mais assertivas, e diferentes agregações de artigos, aplicando-se inclusive técnicas de *clustering* para agregação de artigos com perfis similares.

A comparação dos métodos de previsão foi realizada através de erros de previsão, calculados para o mesmo nível de agregação, bem como para o mesmo período. De forma a aferir o grau de assertividade das previsões obtidas pelos métodos selecionados, foi efetuada a sua comparação com um método *naïve* de previsão, método esse correspondente a médias móveis calculadas com uma janela de dois meses.

Foi realizada uma análise final comparativa entre as previsões obtidas através das simulações realizadas neste projeto com as obtidas pelo SO99+ no início do projeto. Aferiu-se que houve uma diminuição de 8,68 p.p. no erro VWMAPE na família dos CPA e de 7,56 p.p. na família das Massas. Quanto à família das Bolachas, a diminuição foi de 2,56 p.p..

As conclusões retiradas deste projeto permitiram aferir a importância da parametrização correta dos métodos de previsão através de valores de constantes ótimos, bem como a necessidade de agregação em níveis que permitam obter séries temporais mais estáveis. Constatou-se que a implementação de valores adequados das constantes e a agregação dos artigos em níveis adaptados leva a uma diminuição dos erros de previsão, objetivo fulcral desta dissertação.

Optimization of the forecasting methods in food industry

Abstract

The project developed within the scope of this master thesis was carried out at Cerealis, in Maia, integrated in the Logistics Department, in the Tactical Planning and Procurement sector, with the objective of optimizing exponential smoothing forecasting methods through the reparametrization of the smoothing constants and the aggregation of the SKU in suitable levels, with the perspective of further implementation in the production planning tool SO99+.

To this end, it was carried out a detailed analysis to the sales history of each SKU of the company, to make an appropriate characterization of them. This characterization allowed a correct data modelling.

Then, programs have been developed using *Python*, to carry out simulations of different scenarios, where three exponential smoothing methods were tested. Different values of the smoothing constants were also tested for each forecasting method, to select the combination that would lead to more accurate predictions, and different aggregation levels of the SKU, inclusive the application of clustering methods to aggregate SKU with similar properties.

The comparison of the forecasting methods was performed using forecasting errors, calculated for the same aggregation level, as well as for the same period. To evaluate the accuracy of the predictions obtained by the selected methods, a comparison was with a naïve forecasting method, the moving average method with a two-month window.

It was performed a final analysis, comparing the predictions obtained through simulations carried out in this project with those obtained by SO99+ at the beginning of the project. It was verified that there was a decrease of 8,68 p.p. in the VWMAPE of the CPA's family and a decrease of 7,56 p.p. in the VWMAPE of the Massas' family. As for the Bolachas' family, the decreased verified in the VWMAPE was 2,56 p.p..

The conclusions drawn from this project allowed to assess the importance of a correct parameterization of forecasting methods through optimal values of the smoothing constants, as well as the need of an appropriate aggregation level for more stable time series. It was found that the implementation of adequate values of the smoothing constants and the aggregation of the SKU in certain levels leads to a decrease in forecasting errors, the main objective of this dissertation.

Agradecimentos

A realização desta dissertação não seria possível sem o apoio de várias pessoas e instituições, que me deram os meios para alcançar os objetivos.

O meu sentido agradecimento à Professora M.^a Teresa de Aguiar, pelo constante acompanhamento e apoio ao longo da realização desta dissertação, pelo incentivo e pela confiança mostrada.

À Cerealis, pela oportunidade dada na inserção de um projeto inovador, pelo apoio financeiro e pela confiança depositada.

À Eng.^a Bárbara Fraga e à Eng.^a Mafalda Marques pelo apoio e pela disponibilidade prestada ao longo da realização desta dissertação.

A todos os elementos do Departamento Logístico, em particular aos colegas da área de Planeamento Tático e Aprovisionamentos, por toda a boa disposição e pela forma como me receberam e integraram na equipa, motivando-me sempre para ir mais longe.

À Faculdade de Engenharia da Universidade do Porto, professores e colegas que me acompanharam neste percurso académico e que me proporcionaram ferramentas de desenvolvimento profissional e pessoal.

Ao GASPORTO, por me fazer acreditar que nada é impossível se batalharmos para alcançar as nossas metas e por me ajudar a manter o foco naquilo que realmente importa. Não descurando todos aqueles que cruzaram o meu caminho, deixo um agradecimento particular à Filipa Rodrigues, à Ana Fernandes, à Eva Sousa e à Cláudia Miguel, por todas as palavras de motivação, pela energia e pelo foco que me deram.

Agradeço também ao Alexandre, por toda a amizade e pela força dada durante a realização deste projeto de dissertação. A todos os meus amigos que sempre me apoiaram, não só durante esta etapa, mas em todo o meu percurso.

Aos meus pais e à minha irmã Lúcia, por tudo o que me ensinaram a ser, por tudo o que deram; Por serem o meu porto seguro. A toda a minha família, um obrigada por estarem presentes.

E um último especial agradecimento aos meus sobrinhos, que me trouxeram sempre alegria e luz mesmo nos períodos mais difíceis.

A todos, um muito Obrigada.

Índice de Conteúdos

1	Introdução	1
1.1	Grupo Cerealis	1
1.2	Organigrama da Empresa	2
1.3	Enquadramento e objetivos do projeto	3
1.4	Metodologia do projeto	4
1.5	Estrutura da dissertação	5
2	Enquadramento Teórico	6
2.1	Função Logística	6
2.2	Análise ABC/XYZ	6
2.3	Análise de Sazonalidade e de Tendência	8
2.4	Decomposição de Séries Temporais	9
2.5	Métodos de Previsão	10
2.5.1	Amortecimento Exponencial Simples	11
2.5.2	Amortecimento Exponencial Linear de Holt	11
2.5.3	Amortecimento Exponencial de Holt-Winters	12
2.6	Erros de Previsão	14
2.6.1	MAE	15
2.6.2	MAPE	15
2.6.3	VWMAPE	16
2.6.4	BIAS	16
2.7	Níveis de Agregação	16
3	Situação inicial de planeamento	18
3.1	O departamento logístico	18
3.2	Caracterização dos artigos	19
3.2.1	Análise ABC/XYZ	19
3.2.2	Análise de sazonalidade e tendência	22
3.3	Previsão de vendas	24
3.4	Estrutura das necessidades para planeamento de produção	25
3.5	Apresentação do Caso de Estudo	25
4	Métodos de previsão de vendas propostos	27
4.1	Métodos de Amortecimento Exponencial	27
4.1.1	Níveis de Agregação	28
4.1.2	Modelos Aditivos e Modelos Multiplicativos	31
4.1.3	Constantes de Amortecimento	32
4.1.4	Métodos de Previsão	32
4.1.5	Análise de Resultados	33
4.1.6	Apresentação dos resultados dos métodos de previsão de vendas propostos	36
4.2	Implementação de métodos de <i>Machine Learning</i>	44
5	Conclusões e perspectivas de trabalho futuro	46
	Referências	49
	ANEXO A: Gráficos do <i>Elbow Method</i> para obtenção do valor ótimo de <i>clusters</i>	52
	ANEXO B: Constantes de Amortecimento Seleccionadas	54

Siglas

ACF - *Autocorrelation function*

BIAS - Enviesamento do Erro

CPA - Cereais de pequeno-almoço

CV - Coeficiente de variabilidade

DL - Direção Logística

ERP - *Enterprise Resource Planning*

MAE - *Mean Absolute Error*

MAPE - *Mean Absolute Percentage Error*

ML - *Machine Learning*

P.P. - Pontos Percentuais

SC - *Supply Chain*

SKU - *Stock Keeping Unit*

SO99+ - *Service Optimizer 99+*

VWMAPE - *Volume Weighted Mean Absolute Percentage Error*

Índice de Figuras

Figura 1 - Logótipo da Cerealis.....	1
Figura 2 - Organigrama da Cerealis.	2
Figura 3 - Logótipos das marcas próprias da Cerealis Produtos Alimentares.....	2
Figura 4 - Logótipos de marcas próprias da Cerealis Moagens.	3
Figura 5 - Organigrama do Departamento Logístico da Cerealis.....	3
Figura 6 - Cronograma do projeto de dissertação.....	4
Figura 7 - Exemplo de gráfico ACF (Lizardo, 2015).	9
Figura 8 - Exemplo de vista do SO99+ (ToolsGroup, 2018).	18
Figura 9 - Gráfico ABC global.	20
Figura 10 - Gráfico ACF de sazonalidade da família das Massas.....	22
Figura 11 - Gráfico de médias móveis da família das Massas.	22
Figura 12 - Gráfico ACF de sazonalidade da família das Bolachas.	23
Figura 13 - Gráfico de médias móveis da família das Bolachas.	23
Figura 14 - Gráfico ACF de sazonalidade da família dos CPA.....	24
Figura 15 - Gráfico de médias móveis da família dos CPA.	24
Figura 16 - Horizonte temporal das previsões realizadas.....	28
Figura 17 - Exemplo de coordenadas utilizadas para implementação do método de <i>clustering K-Means</i>	30
Figura 18 - Agregação dos artigos da família de Massas, por métodos de <i>clustering K-Means</i>	30
Figura 19 - Agregação dos artigos da família de Bolachas, por métodos de <i>clustering K-Means</i>	31
Figura 20 - Agregação dos artigos da família de CPA, por métodos de <i>clustering K-Means</i> ..	31
Figura 21 - Tabela dos métodos de previsão selecionados.....	34
Figura 22 - Previsão de vendas do grupo <i>Cluster 1</i> Massas, através do método de amortecimento exponencial Linear de Holt.....	37
Figura 23 - Previsão de vendas do grupo <i>Cluster 2</i> Massas, através do método de amortecimento exponencial linear de Holt.	38
Figura 24 - Previsão de vendas do grupo <i>Cluster 3</i> Massas, através do método de amortecimento exponencial linear de Holt.	39
Figura 25 - Previsão de vendas do grupo <i>Cluster 4</i> Massas, através do método de amortecimento exponencial linear de Holt-Winters.....	40
Figura 26 - Previsão de vendas do grupo <i>Cluster 5</i> Massas, através do método de amortecimento exponencial linear de Holt-Winters.....	41
Figura 27 - Previsão de vendas da família das Bolachas, através do método de amortecimento exponencial de Holt-Winters.	42
Figura 28 - Previsão de vendas da família dos CPA, através do método de amortecimento exponencial de Holt-Winters.	43

Figura 29 - Gráfico de <i>Elbow Method</i> para obtenção de valor ótimo k de <i>clusters</i> , para a família das Massas.....	52
Figura 30 - Gráfico de <i>Elbow Method</i> para obtenção de valor ótimo k de <i>clusters</i> , para a família das Bolachas.....	53
Figura 31 - Gráfico de <i>Elbow Method</i> para obtenção de valor ótimo k de <i>clusters</i> , para a família dos CPA.	53
Figura 32 - Constantes de amortecimento ótimas para os três métodos de previsão em análise.	54

Índice de Tabelas

Tabela 1 - Critério de classificação ABC.	7
Tabela 2 - Critérios de classificação XYZ.....	8
Tabela 3 - Matriz ABC/XYZ.....	8
Tabela 4 - Percentagem de cada família na totalidade dos artigos.....	19
Tabela 5 - Peso de cada família nas classes A, B e C.....	20
Tabela 6 - Matriz ABC/XYZ global.....	21
Tabela 7 - Matriz ABC/XYZ das Massas.....	21
Tabela 8 - Matriz ABC/XYZ das Bolachas.....	21
Tabela 9 - Matriz ABC/XYZ dos CPA.	21
Tabela 10 - Erros de previsão do SO99+, na situação inicial.....	25
Tabela 11 - Erros VWMAPE para seleção de níveis de agregação.....	33
Tabela 12 - Somatórios dos quadrados dos coeficientes para seleção dos modelos de decomposição.	34
Tabela 13 - Comparação dos métodos de amortecimento exponencial selecionados com método <i>naïve</i>	35
Tabela 14 - Comparação dos métodos de amortecimento exponencial selecionados com previsões obtidas pelo SO99+.....	36
Tabela 15 - Comparação entre previsões por métodos tradicionais e métodos de <i>machine learning</i> (adaptado) (Demand Forecasting Methods: Using Machine Learning and Predictive Analytics to See the Future of Sales, 2019).....	44

1 Introdução

Este capítulo tem como objetivo apresentar o Grupo Cerealis, fazendo-se um resumo da sua história e do seu organigrama. É também feito o enquadramento do projeto de dissertação e apresentados os objetivos do mesmo. Por fim, expõem-se a metodologia seguida neste projeto e a sua estrutura.

1.1 Grupo Cerealis

O grupo Cerealis (Figura 1) desenvolve a sua atividade industrial e comercial na área agroalimentar, com foco no fabrico de massas, bolachas, cereais de pequeno-almoço e farinhas industriais e de uso culinário. Atualmente, com uma faturação da ordem dos 200 milhões de euros, produz mais de 160 artigos e conta com cerca de 700 colaboradores.



Figura 1 - Logótipo da Cerealis.

Este grupo surgiu em 1919 com a fundação da Amorim, Lage & Soares, Lda., detida pelos fundadores José Amorim, Manuel Lage e Aníbal Soares, passando, em 1930, a Amorim & Lage, Lda., com a saída de Aníbal Soares. Três anos mais tarde é lançada a marca Milaneza e, com ela, é inaugurada a fábrica de massas alimentícias.

Ao longo do tempo, o grupo foi ampliando o seu parque moageiro e construindo a primeira bateria de silos de trigo. Contudo, foi na década de 60, que iniciaram a montagem de novas linhas de fabrico de esparguete e massas cortadas e a instalação de linhas de embalagem semiautomáticas, às quais se seguiu a construção da primeira semolaria portuguesa de trigo duro e a de uma nova moagem de trigo em Coimbra, o que permitiu aumentar a cobertura comercial do setor das farinhas industriais na zona sul.

Nos anos 80, iniciou-se uma operação de modernização tecnológica, com a renovação das infraestruturas industriais e montagem de novas linhas de fabrico de massas alimentícias (Cerealis, 2018).

No final da década de 90, o grupo adquiriu a totalidade do capital da Moacir e da Companhia de Moagens Harmonia, S.A., onde concentrou toda a produção de farinhas industriais. Contudo, este período ficou marcado pela obtenção da certificação de qualidade da Milaneza para o fabrico de massas e farinhas de usos culinários, segundo a norma ISO 9001, e pela aquisição da Companhia Industrial de Transformação de Cereais e Derivados, S.A., detentora da marca Nacional, o que fez da Cerealis o maior grupo português no setor agroalimentar e na área da moagem de cereais e fabrico de massas alimentícias. Com esta

aquisição, procedeu-se ao desenvolvimento de um plano estratégico de investimentos que contemplou a construção da nova fábrica de massas na Maia, que passou a ser um dos centros de produção mais modernos da Península Ibérica.

Em 2004, o grupo reestrutura-se e assume a designação Cerealis, tendo como resultado o aparecimento das empresas Cerealis SGPS, S.A., empresa *holding*, Cerealis Produtos Alimentares - S.A., Cerealis Moagens - S.A., Cerealis Internacional – Comércio de Cereais e Derivados, S.A. e a Sociedade Imobiliária Paradense (Cerealis, 2018).

Em 2014 é estabelecida a parceria com o Instituto Kaizen no desenvolvimento de metodologias de gestão Lean e práticas 5S nas áreas das Operações da Cerealis Produtos Alimentares (Cerealis, 2018).

Até à data, o grupo Cerealis investe na constante inovação e melhoria dos seus produtos, promovendo o *empowering* dos seus recursos humanos alcançando, desta forma, o nível de excelência e posições de liderança nos mercados em que se encontra (Cerealis, 2018).

1.2 Organigrama da Empresa

A atual organização do grupo conta com a Cerealis SGPS, S.A. (empresa *holding*) como coordenadora dos três ramos do grupo (Figura 2):



Figura 2 - Organigrama da Cerealis.

- **Cerealis Produtos Alimentares, S.A.**

Empresa direcionada para a produção e comercialização de produtos alimentares, nomeadamente de massas alimentícias, de cereais de pequeno-almoço, de bolachas e de farinhas de uso culinário. Nela produz-se para as suas próprias marcas (Figura 3), bem como para clientes nacionais.



Figura 3 - Logótipos das marcas próprias da Cerealis Produtos Alimentares.

De referir, ainda, que esta empresa é detentora de 33,33% do capital social da Europasta, fabricante de massas alimentícias com origem na República Checa, o que coloca o Grupo Cerealis numa posição privilegiada para a sua expansão e crescimento.

- **Cerealis Moagens, S.A.**

Empresa de produção e comercialização de farinhas, quer de uso culinário quer industriais, detentora de três centros industriais de moagem de elevado padrão tecnológico, produzindo para quatro marcas, a Nacional, a Harmonia, a Concordia e a BIG (Figura 4).



Figura 4 - Logótipos de marcas próprias da Cerealis Moagens.

- **Cerealis Internacional – Comércio de Cereais e Derivados, S.A.**

Empresa responsável pela exportação de produtos das marcas Cerealis, presentes em mais de 40 países, e pela importação das matérias-primas para transformação, como o trigo duro, trigo mole, centeio e *hominy gritz* (usado no fabrico de corn flakes) (Cerealis, 2018).

1.3 Enquadramento e objetivos do projeto

- **Departamento Supply Chain**

Logística é a parte da *Supply Chain* (SC) que se dedica ao planeamento, implementação e controlo eficiente do fluxo e armazenamento de bens, serviços e informação entre o ponto de origem e o de consumo, de forma a alcançar os requisitos dos clientes (Professionals, 2013).

A Direção Logística (DL) tem como responsabilidade o planeamento da produção, a gestão da cadeia de distribuição, dos stocks de produtos acabados, de armazenamento, receção e acompanhamento de encomendas, de gestão de aprovisionamentos de matérias primas, subsidiárias e material de embalagem, de colocação de encomendas a fornecedores, bem como o serviço de apoio ao cliente. Na Figura 5 é apresentada o organigrama do Departamento Logístico da Cerealis.

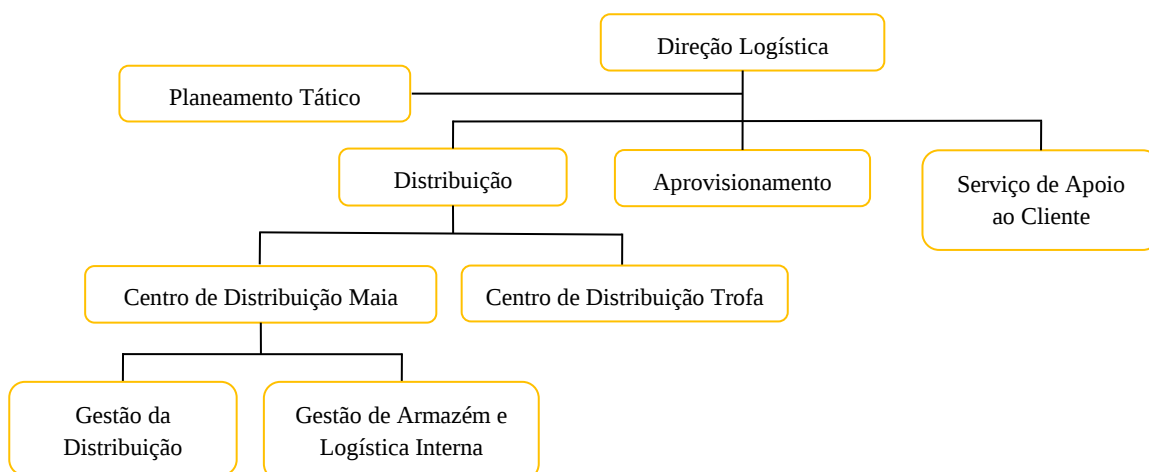


Figura 5 - Organigrama do Departamento Logístico da Cerealis.

- **Objetivos**

No âmbito da realização desta dissertação, a DL da Cerealis, com o objetivo de maximizar a eficiência das linhas de produção e os níveis de serviço, pretende implementar uma ferramenta de planeamento de produção ao artigo – *Service Optimizer 99+ (SO99+)* – na família de Massas, tendo por base a previsão de vendas e a atividade promocional planeada.

Para tal, pretendeu-se estudar aspetos inerentes aos artigos, tais como a sua sazonalidade e a tendência, de forma a compreender o seu comportamento ao longo do tempo.

Este projeto encontra-se focado na previsão de vendas dos artigos para fins de planeamento de produção, sendo o objetivo final fornecer à DL propostas de melhoria das

suas previsões de vendas, de forma a minimizar os erros de previsão e diminuir os níveis de stock, sem comprometer os níveis de serviço.

No que diz respeito à previsão de vendas, foi estudado qual o nível de agregação adequado dos artigos e qual o melhor método de previsão a ser usado, bem como os valores dos respetivos parâmetros.

1.4 Metodologia do projeto

A metodologia seguida neste projeto pode ser dividida em duas etapas: o estudo abrangente do contexto atual da empresa e o estudo mais detalhado do problema.

Na primeira etapa, de forma a ter-se uma visão geral do grupo, foi feito o acolhimento nos vários setores e departamentos da Cerealis Produtos Alimentares, com o objetivo de se perceber quais as rotinas diárias e a forma como cada um comunica entre si. Em simultâneo, foi dado a conhecer os projetos em curso na DL, com foco particular no projeto de implementação da ferramenta de planeamento de produção no departamento de Planeamento Tático.

A segunda etapa consistiu no estudo detalhado do problema deste projeto. Realizou-se a análise dos dados necessários para o estudo, nomeadamente históricos de vendas, stocks disponíveis e propostas de produção sugeridas pela ferramenta de planeamento em uso. Esta fase pode ser dividida em subfases:

- Análise da situação atual e identificação do problema;
- Potenciamento da ferramenta de planeamento, através da sua reparametrização;
- Programação de vários métodos de previsão de vendas, de forma a obter a melhor combinação de método preditivo, nível de agregação e constantes de amortecimento;
- Recolha de dados das propostas de melhorias, recorrendo-se para isso a simulações de várias metodologias de previsões de vendas para um período passado;
- Análise de dados, realizando-se a comparação entre a situação atual da empresa e a situação proposta de melhoria;
- Propostas de melhorias adicionais, com recurso a métodos diferentes dos utilizados pela ferramenta.

Na Figura 6 encontra-se o cronograma final de execução do projeto de dissertação.

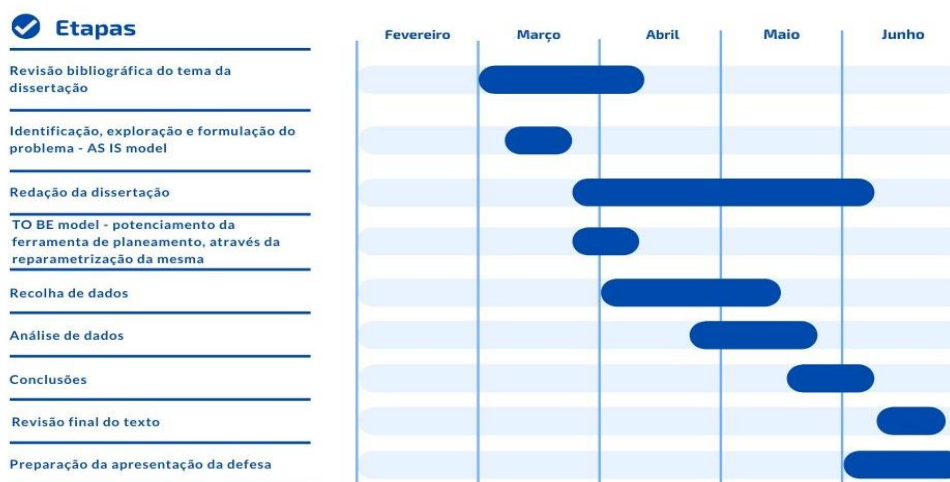


Figura 6 - Cronograma do projeto de dissertação.

1.5 Estrutura da dissertação

Para garantir coerência ao longo da realização desta dissertação, o presente documento encontra-se estruturado da seguinte forma:

No primeiro capítulo é apresentada a empresa de forma breve, bem como o projeto desenvolvido. É também definida a metodologia a seguir.

No segundo capítulo são abordados os conceitos relevantes que dizem respeito à análise de dados de séries temporais, previsão de vendas e seus erros.

O terceiro capítulo foca a situação atual de planeamento da empresa, apresentando-se o departamento logístico e os métodos de planeamento de produção e de previsão de vendas utilizados. É analisado o caso de estudo, tendo-se em vista o levantamento dos parâmetros sujeitos a melhoria.

O quarto capítulo refere-se ao desenvolvimento de melhorias dos parâmetros identificados, bem como a proposta dessas melhorias. É realizada também a avaliação das melhorias propostas, através da análise de indicadores definidos.

No quinto capítulo são apresentadas as conclusões e trabalhos futuros.

2 Enquadramento Teórico

No presente capítulo, serão abordadas as principais temáticas para o desenvolvimento deste trabalho, nomeadamente a função logística, análise ABC/XYZ para classificação dos artigos quanto ao volume e variabilidade das vendas e os métodos de previsão de vendas.

2.1 Função Logística

O termo Logística resulta de um conceito antigo, a gestão coordenada. O uso deste conceito remonta a 1844, onde a ideia de se trocar um tipo de custo por outro (custo de transporte e custos de inventário), em prol de uma vantagem maior, já era discutida (Ballou, 1999).

Foi durante a Segunda Guerra Mundial, em contexto militar, que o conceito Logística começou a ser mais utilizado, definindo-se como *“ramo da ciência militar ligada à procura, manutenção e transporte de material, pessoal e instalações”* (Ballou, 1999).

Atualmente, a Logística define-se como *“um processo que faz parte da cadeia de abastecimento, tendo por objetivo o planeamento, implementação e controlo de eficiência, fluxos e inventários de bens materiais, serviços e informações relacionadas, desde o ponto de origem até ao ponto de consumo, de forma a cumprir com os requisitos dos clientes”* (Professionals, 2013).

Tendo como principal objetivo a criação de valor, para os clientes, para os fornecedores e para a própria empresa, a Logística apresenta um conjunto de atividades principais, genéricas para os vários setores, tais como: estabelecimento dos níveis de serviço ao cliente standard, transporte de mercadorias, gestão de stocks e seu armazenamento, previsão de vendas, planeamento de compras, análise de dados e controlo dos processos da cadeia de abastecimento.

Estas atividades, inseridas na cadeia de abastecimento, devem acrescentar valor ao processo, do ponto de vista da gestão logística, seja ele em termos de lucro ou cumprimento dos requisitos dos clientes (Ballou, 1999).

2.2 Análise ABC/XYZ

A análise ABC, baseada na regra de Pareto onde se afirma que 80% das vendas de uma empresa são geradas por 20% dos seus artigos (Kourentzes, 2016), tem como objetivo identificar quais os artigos de maior importância, distinguindo-os em três níveis:

- **Classe A** – artigos de alta rotação, com grande grau de importância. Pelo seu elevado valor, deve ter-se uma maior atenção na sua gestão;
- **Classe B** – artigos de rotação intermédia. O seu grau de importância é mediano. Contudo, tendo em conta que representam algum volume das vendas, deve ser dada alguma atenção à gestão de artigos desta classe;

- **Classe C** – artigos de baixa rotação, cujo grau de importância é mais reduzido, devido ao seu diminuto valor. Não requerem tanta atenção na sua gestão.

A distinção dos artigos nos níveis referidos é feita através de parâmetros pré-definidos pela empresa, como o custo médio unitário dos artigos, o volume de vendas anuais ou o consumo anual, dependendo do foco que se pretender.

Para se classificar cada artigo, deve-se primeiramente obter a contribuição de cada um relativamente ao critério selecionado, depois ordenar por ordem crescente de valor cumulativo de contribuições e, finalmente, agrupar nas respetivas classes.

Neste projeto, o critério definido foi o volume de vendas, apresentando-se o mesmo na Tabela 1.

Tabela 1 - Critério de classificação ABC.

Classe	Critério	Volume de Vendas (%)
A	Volume de Vendas Anual	80
B		15
C		5

Além da categorização de um artigo quanto à sua importância, torna-se também fundamental a sua caracterização relativamente à variabilidade no volume de vendas ao longo do horizonte temporal, de forma a haver uma indicação do grau de dificuldade na obtenção das previsões de vendas de um determinado artigo. Assim, como complemento da análise ABC surge a análise XYZ, que, mais uma vez, agrupa os artigos em três níveis:

- **Classe X** – artigos com baixa variabilidade no volume de vendas e, por isso com baixa dificuldade de previsão.
- **Classe Y** – artigos com variabilidade moderada. As suas previsões são, ainda, com algum grau de certeza;
- **Classe Z** – artigos com grande variabilidade no volume de vendas, implicando que as suas previsões sejam mais difíceis de obter com precisão.

O critério usado para classificar os artigos em X, Y ou Z, é o coeficiente de variabilidade, obtido através do quociente entre o desvio padrão do volume de vendas anual e a sua média. O cálculo deste coeficiente é dado pela Equação 2.1.:

$$CV = \frac{\sigma}{\bar{V}} \quad (2.1.)$$

Onde:

- CV – coeficiente de variabilidade;
- σ – desvio padrão do volume de vendas anual;
- $\bar{V}$ – média do volume de vendas anual.

Na Tabela 2 apresentam-se os valores de CV correspondentes a cada uma das classes da análise XYZ usados neste projeto. Estes valores dependem da situação da empresa e devem ser adaptados à realidade de cada caso.

Tabela 2 - Critérios de classificação XYZ.

Classe	Valor de CV
X	$CV < 0,5$
Y	$0,5 < CV < 1,5$
Z	$CV > 1,5$

A conjugação destas duas classificações permite uma gestão logística dos artigos adequada, bem como a agregação dos mesmos em subfamílias para posterior previsão de vendas das mesmas, ou para diferenciação do grau de atenção na validação e ajuste das previsões: artigos AZ devem ser aqueles com uma validação mais cuidada e artigos CX aqueles com menor necessidade de validação (Tabela 3).

Tabela 3 - Matriz ABC/XYZ.

	X	Y	Z
A	AX	AY	AZ
B	BX	BY	BZ
C	CX	CY	CZ

2.3 Análise de Sazonalidade e de Tendência

A sazonalidade de uma série temporal é identificada através de padrões que se repetem ao longo de intervalos de tempo periódicos (Makridakis, Wheelwright, & Hyndman, 1998).

Uma forma de detetar a existência de sazonalidade numa série temporal é através de coeficientes de autocorrelação, com um determinado grau de desfasamento k . Os coeficientes de autocorrelação são medidas estatísticas que indicam o grau de relação de uma variável em dois períodos diferentes (Makridakis, Wheelwright, & Hyndman, 1998). Estes podem ser calculados através da Equação 2.2.:

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad (2.2.)$$

Onde:

- r_k – coeficiente de autocorrelação, com grau de desfasamento k ;
- k – grau de desfasamento;
- Y_t – variável no período atual t ;
- Y_{t-k} – variável no período $t-k$;
- $\bar{Y}$ – média de todos os valores da variável Y .

O conjunto de coeficientes de autocorrelação constitui a função de autocorrelação (ACF), sendo que o seu gráfico é uma ferramenta que ajuda na melhor perceção de padrões em séries temporais, tais como a sazonalidade, antes de se iniciar a previsão de vendas. A ACF permite identificar se existe relação entre um valor de uma observação, no período atual, e outros valores do passado. Desta forma, a sazonalidade pode ser detetada com recurso a gráficos representativos dos coeficientes de autocorrelação, ou gráficos ACF, onde picos desses coeficientes, num determinado período k , indica que existirá sazonalidade com um período sazonal k (Figura 7).

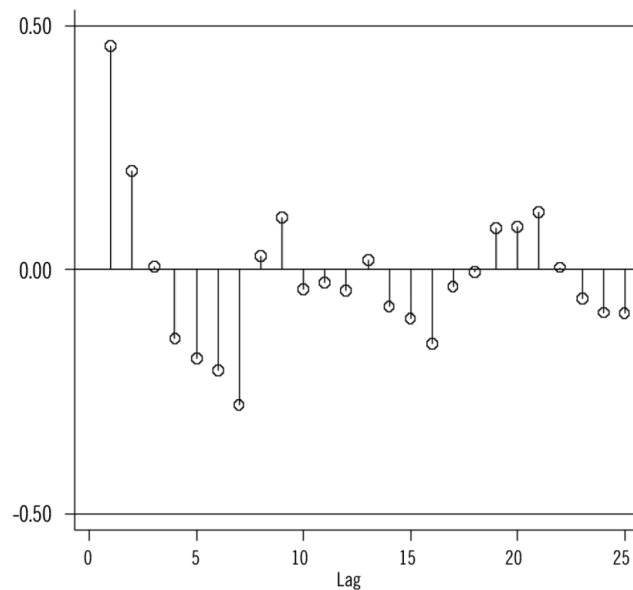


Figura 7 - Exemplo de gráfico ACF (Lizardo, 2015).

Quanto à tendência de uma série temporal, esta pode ser definida como uma alteração nos dados a longo prazo de sentido crescente ou decrescente (Hyndman & Athanasopoulos, 2013).

Para se captar a tendência de uma série temporal, um método eficaz é o de médias móveis, traçando-se uma nova série temporal, em que cada ponto é uma média móvel rolante com um número de período igual à sazonalidade. Havendo tendência, a regressão realizada com os pontos da série temporal deve apresentar declive positivo ou negativo, consoante apresenta tendência crescente ou decrescente, respetivamente.

2.4 Decomposição de Séries Temporais

A decomposição de séries temporais é um método útil para o conhecimento dos três parâmetros que afetam o desenvolvimento das mesmas: sazonalidade, tendência e erros. Esta torna-se uma análise anterior ao início das previsões de vendas, uma vez que ajuda na seleção do método preditivo mais adequado. Existem dois tipos de decomposição: decomposição aditiva ou multiplicativa.

Decomposição aditiva – a série temporal surge através da soma dos três parâmetros acima mencionados, como se demonstra na Equação 2.3.:

$$Y_t = S_t + T_t + E_t \quad (2.3.)$$

Onde:

- Y_t – valor da série temporal, no período atual t ;
- T_t – componente de tendência, no período atual t ;
- S_t – componente de sazonalidade, no período atual t ;
- E_t – componente de erros, no período atual t .

É um modelo adequado para séries temporais em que a amplitude das variações de sazonalidade não muda ao longo do tempo.

A decomposição aditiva inicia-se com a determinação da tendência T_t pelo método de médias móveis (Secção 2.3). Após o seu cálculo, subtraímos este fator da série temporal para que esta apenas contenha sazonalidade e erros, S_t e E_t . Retirando-se a componente tendência

da série, calculam-se os índices de sazonalidade de cada período, através da média das vendas do mesmo nos vários anos. Finalmente, os erros E_t obtêm-se subtraindo a sazonalidade S_t (Makridakis, Wheelwright, & Hyndman, 1998).

Decomposição multiplicativa – as séries temporais são obtidas através do produto da tendência, da sazonalidade e dos erros (Equação 2.4.):

$$Y_t = S_t \times T_t \times E_t \quad (2.4.)$$

Onde:

- Y_t – valor da série temporal, no período atual t ;
- T_t – componente de tendência, no período atual t ;
- S_t – componente de sazonalidade, no período atual t ;
- E_t – componente de erros, no período atual t .

O modelo multiplicativo é utilizado quando as séries temporais apresentam uma variação crescente ou decrescente nas amplitudes das sazonalidades ao longo do tempo.

À semelhança do modelo aditivo, começa-se por determinar a tendência T_t da série temporal através do método de médias móveis. No entanto, para se retirar este parâmetro da série deve realizar-se o quociente entre os valores reais de vendas Y_t e os valores das tendências. De seguida, calculam-se os índices de sazonalidade pelo processo referido na decomposição aditiva, obtendo-se a sazonalidade S_t da série. Por fim, os erros E_t obtêm-se através da divisão dos valores reais das vendas Y_t pelo produto dos parâmetros de tendência T_t e sazonalidade S_t (Makridakis, Wheelwright, & Hyndman, 1998).

2.5 Métodos de Previsão

Os métodos de previsão definem-se como “*algoritmos que indicam valores previsionais num determinado período futuro t* ” (Hyndman, Koehler, Ord, & Snyder, 2008), de forma a diminuir a incerteza dos acontecimentos nesse período.

A necessidade de se realizarem previsões ocorre devido à existência de um desfaseamento entre o momento de planeamento de produção relativamente ao momento da compra do artigo (Makridakis, Wheelwright, & Hyndman, 1998).

A realização de previsões tem um papel relevante numa organização pois contribuem para o alcance dos seus objetivos. Pela sua importância, são várias as áreas onde a podemos realizar, nomeadamente no planeamento e na aquisição de mercadorias.

Existem dois tipos de métodos de previsão:

- **Métodos Qualitativos** – são aplicados quando não há informação quantitativa sobre o passado ou quando o passado não permite inferir sobre o futuro. Dependem de *inputs* de intuição e experiência previamente adquirida.
- **Métodos Quantitativos** – requerem três condições: 1. Histórico de dados sobre as variáveis em estudo; 2. Capacidade de quantificar numericamente esses dados; 3. Possibilidade de identificar padrões nesses históricos e inferir sobre a sua repetição no futuro – existência de continuidade (Makridakis, Wheelwright, & Hyndman, 1998).

Foram utilizados métodos quantitativos na realização deste projeto, adequados para previsões de vendas para horizontes de tempo próximos, como os que são necessários para o planeamento da produção.

Como já foi anteriormente referido, uma série temporal divide-se em três componentes: tendência, sazonalidade e erros. Os métodos de previsão focados para a análise das séries temporais deste projeto de dissertação foram os de amortecimento exponencial simples, amortecimento exponencial linear de Holt e de amortecimento exponencial de Holt-Winters, uma vez que são tidos como métodos de séries temporais bastante assertivos e com capacidade de se adaptarem a grandes mudanças nos dados (Ballou, 1999).

2.5.1 Amortecimento Exponencial Simples

O método de amortecimento exponencial simples é usualmente utilizado em séries temporais que não apresentam tendência nem sazonalidade. É um tipo de modelo baseado em médias móveis ponderadas, onde as observações mais recentes têm pesos superiores às mais antigas. O problema de armazenamento de informação é reduzido, pois não é necessário guardar um grande número de dados já que o efeito de todo o histórico está presente na previsão P_t . Assim, a previsão do período a seguir ao atual, $t+1$, é obtida pela Equação 2.5.:

$$P_{t+1} = P_t + \alpha(Y_t - P_t) \quad (2.5.)$$

Onde:

- t – período de tempo atual;
- α – constante de amortecimento exponencial ($0 \leq \alpha \leq 1$);
- Y_t – procura no período t ;
- P_t – previsão para o período t ;
- P_{t+1} – previsão para o período seguinte a t .

Constata-se que o erro de previsão para o período t , $(Y_t - P_t)$, é usado na correção da previsão do período seguinte, na direção oposta ao mesmo. Assim, quando α é próximo de 1, existe um grande ajuste na previsão para $t+1$ através do erro. No caso de α ser praticamente nulo, a previsão não apresenta correção (Hyndman R. J., Koehler, Ord, & Snyder, 2008).

A escolha do valor de α é importante, na medida em que quanto mais elevado for o seu valor, maior é o peso atribuído às procuras mais recentes, Y_t . Por este motivo, deve-se selecionar um α que permita ao modelo de previsão detetar variações significativas nos dados e minimizar flutuações aleatórias, para que o erro de previsão seja o menor possível.

Para se iniciar as previsões, e sendo o valor de P_t desconhecido, assume-se que a previsão do período 1 é igual à procura no mesmo período, $P_1 = Y_1$. Uma outra forma de iniciar esta tarefa é usar a média de cerca de cinco valores dos dados recentes como previsão do período 1 (Makridakis, Wheelwright, & Hyndman, 1998).

2.5.2 Amortecimento Exponencial Linear de Holt

Quando os dados da série temporal apresentam tendência, o método de previsão a usar será o de amortecimento exponencial linear de Holt. Neste, são utilizadas duas constantes de amortecimento, α e β , obtendo-se a previsão final, para o período pretendido, através das Equações 2.6., 2.7. e 2.8.:

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + b_{t-1}) \quad (2.6.)$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1} \quad (2.7.)$$

$$P_{t+n} = L_t + b_t n \quad (2.8.)$$

Onde:

- L_t – nível da série no período t ;
- b_t – declive da série no período t ;
- α – constante de amortecimento exponencial ($0 \leq \alpha \leq 1$);
- β – constante de amortecimento do declive ($0 \leq \beta \leq 1$);
- P_{t+n} – previsão para o período $t+n$, sendo n o número de períodos à frente do atual.

A Equação 2.6. permite calcular uma estimativa do nível da série temporal no período t , ajustando os valores atuais de procura Y_t relativamente à tendência verificada no período anterior. Já a Equação 2.7. atualiza a tendência ao longo dos vários períodos, sendo este processo semelhante ao que se verifica no método de previsão por amortecimento exponencial simples quando se realizam as previsões. No entanto, neste caso, esta equação permite estabilizar a tendência e respetivas flutuações aleatórias. Por fim, a Equação 2.8. reflete o valor da previsão final para o período $t+1$ (Swamidass, 2000).

Para se iniciar o processo de previsão admite-se que $L_1 = Y_1$ e que o valor da tendência b_1 é o declive entre Y_1 e Y_2 , isto é, $b_1 = Y_2 - Y_1$.

2.5.3 Amortecimento Exponencial de Holt-Winters

Quando os dados da série temporal apresentam tendência e padrões de sazonalidade, o método adequado de previsão é o de amortecimento exponencial de Holt-Winters.

Mediante o tipo de sazonalidade, existem dois tipos de métodos de Holt-Winters. Caso esta seja aditiva, isto é, o padrão de sazonalidade é constante ao longo do tempo e independente da série temporal, as equações utilizadas para o cálculo da previsão final para o período pretendido são as seguintes:

$$L_t = \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + b_{t-1}) \quad (2.9.)$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1} \quad (2.10.)$$

$$S_t = \gamma(Y_t - L_t) + (1 - \gamma)S_{t-s} \quad (2.11.)$$

$$P_{t+n} = L_t + b_t n + S_{t-s+n} \quad (2.12.)$$

Onde:

- L_t – nível da série no período t ;
- b_t – declive da série no período t ;
- S_t – componente de sazonalidade do período t ;
- α – constante de amortecimento exponencial ($0 \leq \alpha \leq 1$);
- β – constante de amortecimento do declive ($0 \leq \beta \leq 1$);
- γ – constante de amortecimento de sazonalidade ($0 \leq \gamma \leq 1$);
- Y_t – procura no período t ;
- s – comprimento do período de sazonalidade;
- P_{t+n} – previsão para o período $t+n$, sendo n o número de períodos à frente do atual.

No caso de a sazonalidade ser de carácter multiplicativo, isto é, com variações crescentes ou decrescentes na amplitude da sazonalidade e dependente dos valores da série temporal ao longo do tempo, as equações utilizadas para o cálculo da previsão final para o período em análise são:

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + b_{t-1}) \quad (2.13.)$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1} \quad (2.14.)$$

$$S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s} \quad (2.15.)$$

$$P_{t+n} = (L_t + b_t n) \times S_{t-s+n} \quad (2.16.)$$

Onde:

- L_t – nível da série no período t ;
- b_t – declive da série no período t ;
- S_t – componente de sazonalidade do período t ;
- α – constante de amortecimento exponencial ($0 \leq \alpha \leq 1$);
- β – constante de amortecimento do declive ($0 \leq \beta \leq 1$);
- γ – constante de amortecimento de sazonalidade ($0 \leq \gamma \leq 1$);
- Y_t – procura no período t ;
- s – comprimento do período de sazonalidade;
- P_{t+n} – previsão para o período $t+n$, sendo n o número de períodos à frente do atual.

Note-se que a diferença existente entre os dois métodos reside no facto de os termos onde se faz o amortecimento do valor da procura atual serem divididos por uma componente, quando a sazonalidade é de carácter multiplicativo, e não subtraindo-se essa componente como se verifica quando a sazonalidade é de carácter aditivo. Ou seja, na Equação 2.13. o valor da procura Y_t é dividido pela componente de sazonalidade S_{t-s} , ao passo que na Equação 2.9. a componente de sazonalidade é subtraída ao valor da procura atual. A mesma diferença é verificada entre a Equação 2.11. e Equação 2.15., onde na primeira se subtrai L_t a Y_t e na seguinte Y_t é dividido pelo nível da série L_t . Por último, no cálculo da previsão final, a diferença entre os dois métodos reside no facto de que no aditivo a componente de sazonalidade é adicionada às componentes nível e declive da série, enquanto que no método multiplicativo a mesma é multiplicada pela soma destas duas componentes.

Comparativamente ao método linear de Holt, o método de Holt-Winters, para além de acrescentar mais uma equação nos cálculos para se obter a previsão final, apresenta mais uma constante de amortecimento exponencial, γ , servindo esta para amenizar flutuações que ocorram na componente de sazonalidade.

Em ambos os métodos de Holt-Winters, a inicialização do nível e do declive da série é feita da mesma forma, assumindo-se que $t = s$.

A inicialização do nível faz-se pela média das procuras verificadas no primeiro período de sazonalidade s (Equação 2.17.):

$$L_s = \frac{1}{s} \sum_{i=1}^s Y_i \quad (2.17.)$$

Onde:

- L_s – nível da série no período s ;
- Y_t – procura no período i ;
- s – comprimento do período de sazonalidade.

O declive é inicializado através da média das procuras verificadas nos dois primeiros períodos de sazonalidade, ou seja, $2s$ (Equação 2.18.):

$$b_s = \frac{1}{s} \sum_{i=1}^s \frac{Y_{s+i} - Y_i}{s} \quad (2.18.)$$

Onde:

- b_s – declive da série no período s ;
- Y_{s+i} – procura no período $s+i$;
- Y_i – procura no período i ;
- s – comprimento do período de sazonalidade.

Relativamente à inicialização da componente de sazonalidade, caso esta apresente padrão aditivo, realiza-se a subtração entre procuras e o nível no período s (Equação 2.19.). Se a sazonalidade apresentar padrão multiplicativo, os seus índices são obtidos através do quociente entre as procuras e o nível da série no período s (Equação 2.20.).

$$S_i = Y_i - L_s, i = 1, 2, \dots, n \quad (2.19.)$$

$$S_i = \frac{Y_i}{L_s}, i = 1, 2, \dots, n \quad (2.20.)$$

Onde:

- S_i – sazonalidade da série no período i ;
- Y_i – procura no período i ;
- L_s – nível da série no período s ;
- s – comprimento do período de sazonalidade.

2.6 Erros de Previsão

Com o objetivo de avaliar métodos de previsão, quer para os comparar entre si, quer para se medir a precisão do método em uso, recorre-se a medidas estatísticas de erros de previsão. Desta forma, obtém-se facilmente a precisão com que foi calculada a previsão. Neste contexto, foram calculados os Erros Absolutos Médios (MAE – *Mean Absolute Error*), os Erros Absolutos Percentuais Médios (MAPE – *Mean Absolute Percentage Error*), o Erro Absoluto Percentual Médio Ponderado por Volume (VWMAPE – *Volume Weighted Mean Absolute Percentage Error*) e o Enviesamento do Erro (BIAS).

Para além da avaliação dos métodos de previsão, a minimização destes erros permite obter, também, os valores ótimos das constantes de amortecimento α , β e γ . No presente

projeto de dissertação, foi utilizado como critério de seleção destas constantes a minimização do VWMAPE.

2.6.1 MAE

O MAE traduz o valor médio absoluto dos desvios entre os valores observados Y_t e as previsões realizadas P_t , no respetivo período t . É um erro habitualmente utilizado na comparação de previsões efetuadas em conjuntos de dados iguais, apesar de ser complicado comparar métodos de previsão de séries temporais diferentes através do MAE. O erro é calculado pela Equação 2.21.:

$$MAE = \frac{1}{n} \sum_{t=1}^n |Y_t - P_t| \quad (2.21.)$$

Onde:

- MAE – erro médio absoluto;
- n – número de observações;
- Y_t – procura no período t ;
- P_t – previsão para o período atual t .

Os valores obtidos através do cálculo do MAE são visualmente mais fáceis de interpretar, comparativamente com os erros quadráticos médios, por exemplo, uma vez que indicam objetivamente o desvio existente entre a procura real e a previsão. No entanto, quando se pretende comparar métodos de previsão aplicados a dados em escalas temporais distintas, o MAE torna-se ineficiente (Makridakis, Wheelwright, & Hyndman, 1998).

2.6.2 MAPE

Segundo Armstrong et al. (1992) e Makridakis (1998), para colmatar as falhas do MAE, quando se pretende comparar métodos de previsão, deve utilizar-se o MAPE, o qual traduz o valor percentual médio dos desvios absolutos entre as observações Y_t e as previsões P_t , no período t :

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|Y_t - P_t|}{Y_t} \times 100\% \quad (2.22.)$$

Onde:

- $MAPE$ – erro percentual médio absoluto;
- n – número de observações;
- Y_t – procura no período t ;
- P_t – previsão para o período atual t .

Este critério tem a vantagem de permitir calcular o erro médio percentual da previsão, em ordem à respetiva observação. Isto é, não depende da escala dos dados, permitindo a comparação das previsões realizadas através de diferentes métodos. No entanto, apresenta a desvantagem do erro calculado se tornar indefinido quando a observação Y_t é zero.

2.6.3 VWMAPE

Sendo comum nas empresas existir agregação dos vários artigos, o cálculo dos erros de previsão através de médias simples dos critérios acima mencionados torna-se pouco representativo, uma vez que artigos com poucas vendas podem afetar o resultado de forma significativa. Desta forma, recorre-se a um critério que traduz o valor percentual médio dos desvios absolutos entre as observações Y_t e as previsões P_t , ponderados relativamente ao volume respetivo (normalmente o volume de vendas), no período t (Singh, 2015). Neste critério o erro é calculado através da Equação 2.23.:

$$VWMAPE = \frac{\sum_{t=1}^n |Y_t - P_t|}{\sum_{t=1}^n |Y_t|} \times 100\% \quad (2.23.)$$

Onde:

- $VWMAPE$ – erro percentual médio absoluto, ponderado relativamente ao volume;
- n – número de observações;
- Y_t – procura no período t ;
- P_t – previsão para o período atual t .

Através da determinação do VWMAPE, é possível ter uma visão mais objetiva e correta dos erros de previsão quando queremos comparar a eficácia de métodos preditivos diferentes e aplicados a níveis de agregação diferentes.

2.6.4 BIAS

Não interessando apenas o conhecimento do valor dos erros de previsão, mas também a sua direção (positiva ou negativa), um critério complementar aos acima mencionados é o BIAS, que se traduz no cálculo do VWMAPE, mas com desvios relativos e não absolutos:

$$BIAS = \frac{\sum_{t=1}^n Y_t - P_t}{\sum_{t=1}^n Y_t} \times 100\% \quad (2.24.)$$

Onde:

- $BIAS$ – enviesamento do erro;
- n – número de observações;
- Y_t – procura no período t ;
- P_t – previsão para o período atual t .

Através do BIAS conseguimos ter uma perceção mais clara para onde pende o erro de previsão, isto é, se as previsões efetuadas estão acima ou abaixo das observações, o que permite tomar medidas adequadas aquando do planeamento, por exemplo.

2.7 Níveis de Agregação

Tal como referido por Zotteri et al. (2005), uma das principais problemáticas na implementação de métodos de previsão é a escolha de níveis de agregação adequados para os dados em análise.

Dependendo da tipologia dos dados e dos objetivos das previsões, os níveis de agregação servem como forma de estabilização das séries temporais com a diminuição da sua variabilidade. Os níveis mais baixos de agregação, como SKU, apresentam maior variabilidade e levam a maiores erros de previsão, ao passo que níveis mais abrangentes têm

menores erros e menor variabilidade. No entanto, os níveis mais agregados, apesar de menor ruído, podem perder alguma qualidade nas previsões uma vez que determinados padrões de níveis mais baixos podem não ser detetados (Wilson, 2019).

Os tipos de níveis de agregação mais comuns são as agregações naturais, como por exemplo:

- Famílias de artigos;
- Clientes;
- Marcas.

É necessário perceber-se também que podem ser atribuídos diferentes níveis de agregação ao conjunto de dados em análise. Deve, então, avaliar-se caso a caso e aferir qual o nível que leva à obtenção de resultados de previsão mais assertivos.

Uma outra forma de se proceder à agregação de artigos é através de métodos de *clustering*, que envolvem a agregação de artigos em grupos específicos, em que os dados das séries temporais apresentam características semelhantes, tal como sazonalidade e tendência.

O *clustering*, sendo um método de reconhecimento e modelação de dados não supervisionado, para além da vantagem inerente à agregação de padrões semelhantes, permite também uma maior autonomia no estudo das séries temporais.

Um dos algoritmos mais comuns para aplicação de *Clustering* é o *K-Means*. Este tem por base o cálculo das distâncias entre pontos dos dados e os centros dos grupos de *clustering* – *cluster centers*. Estes centros não fazem necessariamente parte dos dados, podendo ser ponto imaginários ou reais. O método de *K-Means* define a semelhança dos dados como a proximidade a que estes se encontram do *cluster center*. Isto é, um determinado artigo é incluído num *cluster* se a sua distância ao centro do mesmo for próxima.

Uma das dificuldades da aplicação deste algoritmo deve-se à escolha do número de *clusters* a serem utilizados. Para tal, é usado o método do cotovelo (*Elbow Method*) para encontrar o número ótimo de *clusters* a usar. Este método heurístico analisa a variância explanatória em função do número de *clusters*. Com o aumento do seu número, a variância vai sendo cada vez menor, até ao ponto em que se torna injustificável continuar a aumentar o número de *clusters*. Este ponto é caracterizado como uma saliência no gráfico, semelhante ao formato de um cotovelo.

3 Situação inicial de planeamento

No presente capítulo é feito o levantamento do contexto atual no departamento logístico, especificamente na divisão de planeamento, bem como a análise de dados e dos métodos de previsão usados até ao momento pela Cerealis. Finalmente, é apresentado o caso de estudo analisado neste projeto de dissertação.

Salienta-se que, por motivos de confidencialidade, todos os valores apresentados neste projeto são mostrados numa proporção relativa aos valores absolutos verificados.

3.1 O departamento logístico

Atualmente, a Cerealis utiliza a ferramenta *Service Optimizer 99+* (SO99+) no planeamento de produção de artigos. Esta é composta por um conjunto de ferramentas de suporte à tomada de decisões logísticas integrado num ambiente singular (Figura 8), aplicando-se no planeamento tático e estratégico de materiais e requisitos essenciais para as cadeias de distribuição das empresas (ToolsGroup, 2018).

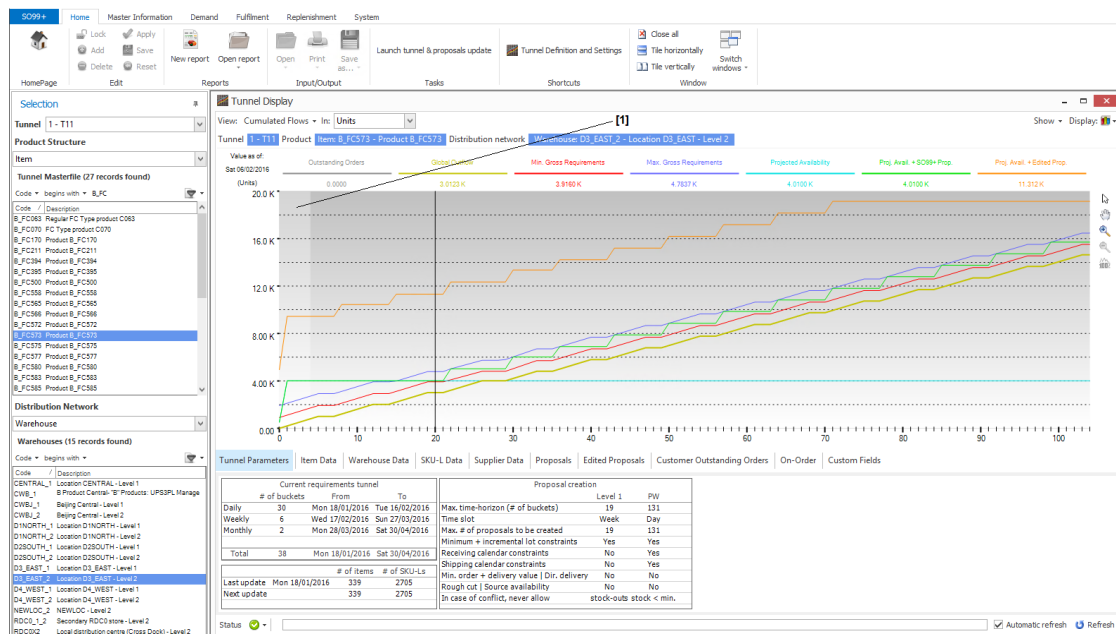


Figura 8 - Exemplo de vista do SO99+ (ToolsGroup, 2018).

No departamento logístico da Cerealis, esta ferramenta é crucial na realização de tarefas como:

- Previsão de vendas;
- Planeamento de produção;
- Gestão de stocks.

Além destas, o SO99+ permite criar quadros de *reporting* de forma a monitorizar os erros de previsão, possíveis ruturas de stock ou *overstock*, bem como a criação de alertas diários de situações que podem influenciar o planeamento de produção de artigos:

- *High demand peak alert*: ajuda a monitorizar variações significativas da procura em contraste com a previsão definida, em específico nos casos em que a procura é acima do normal;
- *Overstock alert*: permite aos planeadores identificar e monitorizar situações de excedente de produção e melhorar as produções futuras;
- *Stockout alert*: ajuda a identificar situações de rotura de stock e antecipar eventuais produções, evitando problemas nos prazos de entregas.

É também possível a exportação de ficheiros, em formato *Excel*, de vários tipos de dados, tais como histórico de vendas, previsões, classificação de artigos e erros de previsão, o que facilita a análise mais detalhada dos processos de previsão e desenvolvimento de possíveis melhorias.

No âmbito deste projeto de dissertação, o recurso ao SO99+ foi fundamental na medida em que permitiu obter os dados necessários ao desenvolvimento das várias análises e atingir os objetivos definidos.

3.2 Caracterização dos artigos

Nesta secção são apresentadas duas análises realizadas aos artigos da Cerealis, sendo estas a análise ABC/XYZ e análise de sazonalidade e tendência. Estas permitiram caracterizar os artigos, fundamental para uma melhor parametrização do método de previsão de vendas.

De referir que a Cerealis classifica os seus artigos mediante a sua tipologia, isto é, mediante a família de produtos em que se inserem. Neste projeto de dissertação foram abordadas as famílias de Massas, Bolachas e Cereais de Pequeno Almoço (CPA).

3.2.1 Análise ABC/XYZ

Neste momento, é possível afirmar que a família das Massas, que corresponde a 53,43% da totalidade dos artigos, apresenta um grande peso no volume de vendas, bem como na variabilidade. A esta segue-se a família dos CPA, a qual representa 34,80% dos artigos totais, e, por fim, a família das Bolachas, com 11,77% da totalidade dos artigos (Tabela 4).

Tabela 4 - Percentagem de cada família na totalidade dos artigos.

Família	
Massas	53,43%
Bolachas	11,77%
CPA	34,80%

Considerou-se como parâmetro de classificação ABC o volume de vendas, em unidade de Kg. Concluiu-se que no panorama global existem:

- 18,84% de artigo A;
- 18,84% de artigos B;
- 62,31% de artigos C (Figura 9).

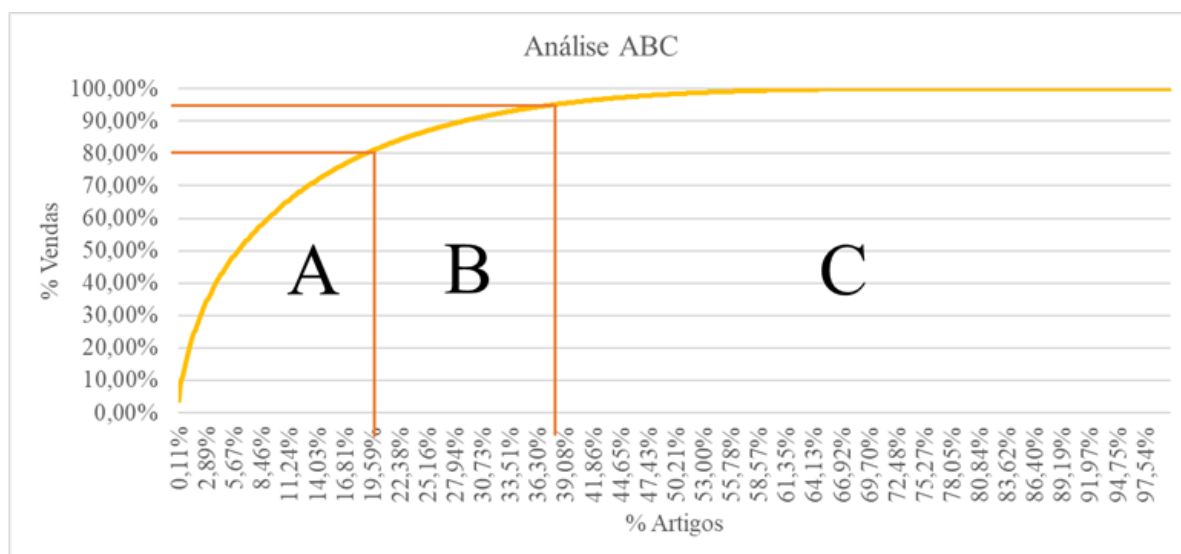


Figura 9 - Gráfico ABC global.

Tal significa que 80% do volume de vendas global corresponde a 18,84% dos artigos (Classe A) e que 5% das vendas corresponde a 62,31% (Classe C).

Fazendo uma análise do ponto de vista das várias famílias de artigos dentro de cada classe (Tabela 5), verifica-se que:

- relativamente aos artigos de **Classe A**, dentro dos 18,84%, 78,98% correspondem a artigos de Massas, 2,27% a artigos de Bolachas e 18,75% a artigos de CPA;
- dos 18,84% dos artigos de **Classe B**, 59,66% são artigos de Massas, 10,80% são artigos de Bolachas e 29,54% são CPA;
- quanto aos artigos de **Classe C**, dos 62,31%, 43,81% correspondem a artigos de Massas, 14,95% a artigos de Bolachas e 41,24% são artigos de CPA.

Tabela 5 - Peso de cada família nas classes A, B e C.

Classe A		Classe B		Classe C	
Global	18,84%	Global	18,84%	Global	62,31%
Massas	78,98%	Massas	59,66%	Massas	43,81%
Bolachas	2,27%	Bolachas	10,80%	Bolachas	14,95%
CPA	18,75%	CPA	29,54%	CPA	41,24%

Denota-se, assim, uma grande prevalência dos artigos incluídos na família das Massas, principalmente nas classes A e B (artigos de maior rotação).

Quanto à análise ABC/XYZ (Tabela 6), verifica-se que, no panorama global, os artigos apresentam uma grande variabilidade no volume de vendas, correspondendo à classe Z ($CV > 1,5$) 43,68% dos artigos.

Tabela 6 - Matriz ABC/XYZ global.

	X	Y	Z
A	15,10%	2,78%	0,96%
B	11,56%	5,14%	2,14%
C	8,99%	12,74%	40,58%

Olhando para estes dados numa perspetiva mais focada em cada uma das três grandes famílias da Cerealis, temos:

Massas

Dentro dos artigos da família das Massas, verifica-se que a maioria apresenta uma baixa variabilidade nas vendas (37,28% dos artigos de Massas) (Tabela 7).

Tabela 7 - Matriz ABC/XYZ das Massas.

	X	Y	Z
A	17,43%	3,81%	1,60%
B	12,42%	5,01%	3,01%
C	7,82%	16,23%	32,67%

Bolachas

Na família das Bolachas, verifica-se que a variabilidade no volume de vendas é grande, dado que 50,91% dos artigos são de Classe Z (Tabela 8).

Tabela 8 - Matriz ABC/XYZ das Bolachas.

	X	Y	Z
A	18,18%	3,64%	0,91%
B	11,82%	6,36%	5,45%
C	1,82%	7,27%	44,55%

CPA

Por último, observando os dados da família dos CPA verifica-se de novo uma grande quantidade de artigos que apresentam elevada variabilidade no volume de vendas, classificados como Classe Z (51,08%) (Tabela 9).

Tabela 9 - Matriz ABC/XYZ dos CPA.

	X	Y	Z
A	14,15%	2,77%	0,31%
B	9,85%	3,38%	2,77%
C	9,85%	8,92%	48,00%

3.2.2 Análise de sazonalidade e tendência

A análise de sazonalidade e tendência foi realizada com agregação dos artigos por famílias.

Para a determinação da sazonalidade, foi utilizado o gráfico ACF com o objetivo de detetar os seus padrões. Já na tendência, foi traçada uma nova série temporal, onde cada ponto é uma média móvel rolante com um período igual à sazonalidade.

Neste contexto, em cada uma das famílias de artigos da Cerealis, pôde apurar-se:

Massas

Na família das Massas, por análise do gráfico ACF (Figura 10) denota-se a existência de picos com um padrão de repetição ao longo dos *lags*. É nítido que existe um padrão de sazonalidade, com o aumento de vendas nos meses 1, 11 e 12 de cada ano. Desta forma, pode aferir-se que o período de sazonalidade das Massas é de doze meses.

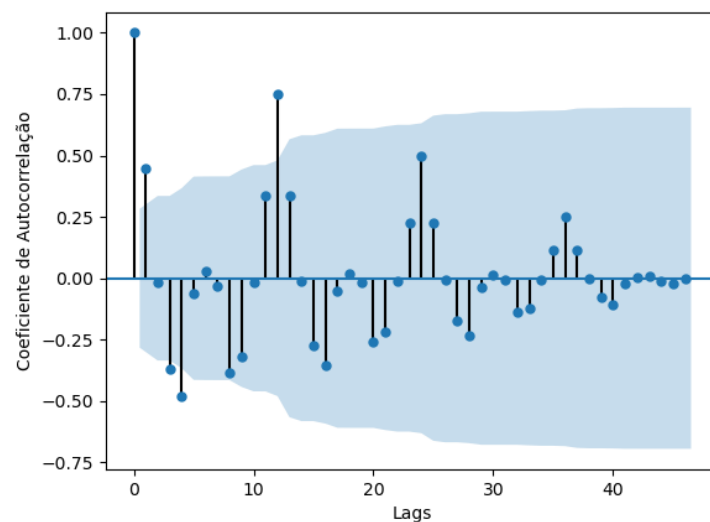


Figura 10 - Gráfico ACF de sazonalidade da família das Massas.

A partir do gráfico da Figura 11, que retrata as médias móveis para a família das Massas, pode ver-se que o declive da linha de tendência desta família é inexistente. Daqui se conclui que a família das Massas apresenta uma tendência constante (nem crescente, nem decrescente).

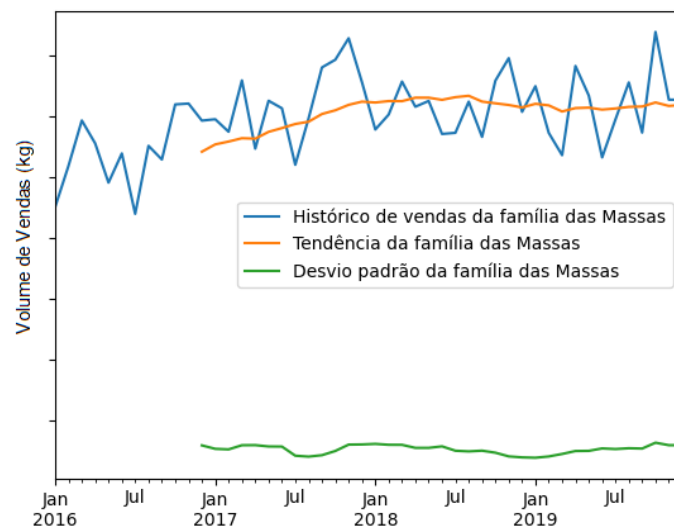


Figura 11 - Gráfico de médias móveis da família das Massas.

Bolachas

Relativamente à família das Bolachas, através do gráfico ACF (Figura 12) verificamos a existência de picos, salientando-se o acentuado aumento no mês de dezembro, o que permite inferir que há sazonalidade com um período de doze meses.

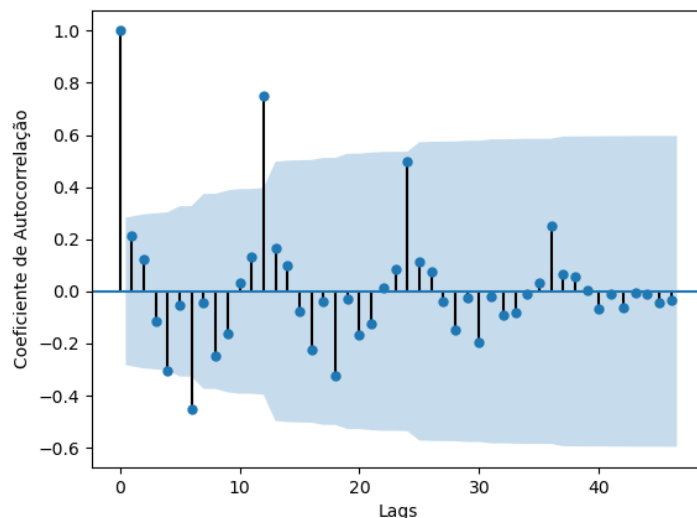


Figura 12 - Gráfico ACF de sazonalidade da família das Bolachas.

Quanto à tendência, o gráfico da Figura 13 mostra que a família das Bolachas apresenta uma tendência crescente de vendas, uma vez que a sua linha tem um declive positivo. No entanto, podemos verificar também uma estabilização do volume de vendas no ano de 2019.

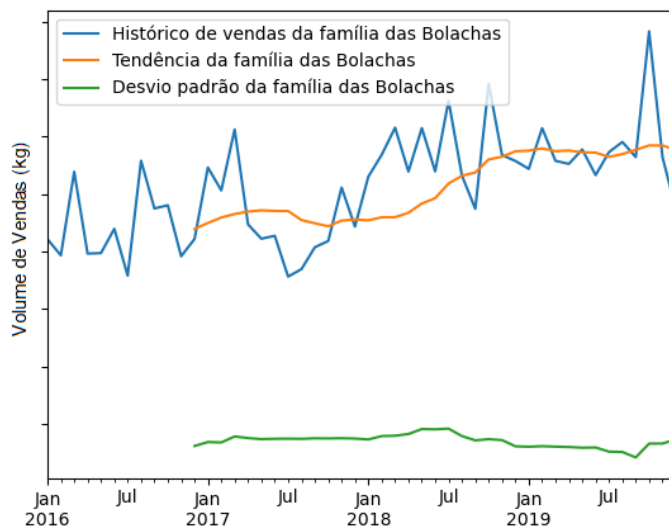


Figura 13 - Gráfico de médias móveis da família das Bolachas.

Ao observar-se a Figura 13, constata-se, ainda, que o maior crescimento na tendência ocorre entre os anos de 2018 e 2019.

CPA

Por análise do gráfico ACF da família dos CPA (Figura 14), verificamos um padrão de sazonalidade onde se encontram picos em determinados *lags*. Tal permite inferir que existe sazonalidade, sendo o seu período de doze meses. Tal como na família das Bolachas, verifica-se um pico no mês de dezembro.

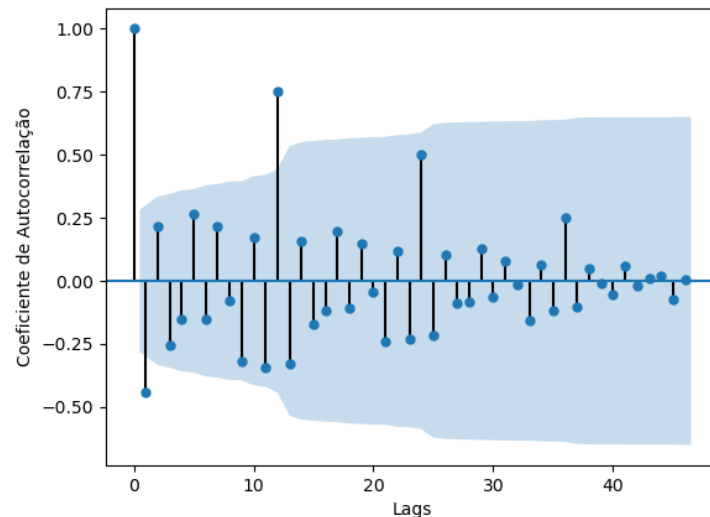


Figura 14 - Gráfico ACF de sazonalidade da família dos CPA.

Através do gráfico de médias móveis (Figura 15) apura-se, também, que o declive da linha de tendência da família dos CPA é positivo, ou seja, uma tendência crescente. Constatase, ainda, que o declive se torna mais acentuado em meados do ano 2018, concluindo-se que o volume de vendas tem vindo a aumentar cada vez mais.

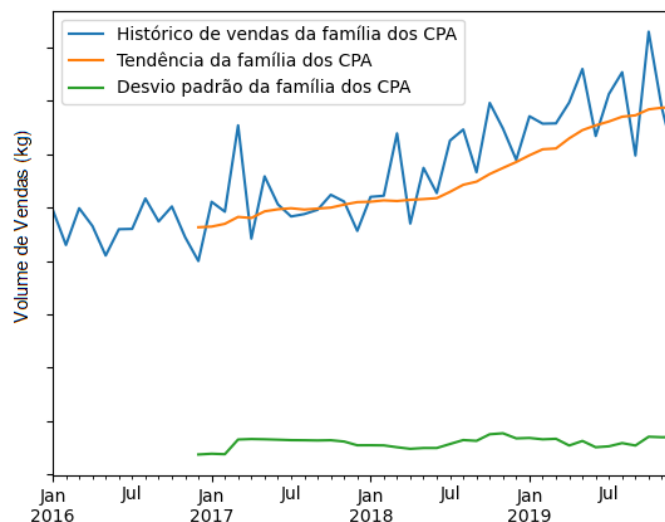


Figura 15 - Gráfico de médias móveis da família dos CPA.

3.3 Previsão de vendas

Na Cerealis, o processo de planeamento de produção passa primeiramente pela previsão de vendas mensais de cada artigo. Essa previsão é realizada através da ferramenta SO99+, que utiliza métodos de amortecimento exponencial para obter os seus valores.

Através da importação diária de dados do *Infor M3*, um sistema *enterprise resource planning* (ERP) onde se encontram todas as informações das várias áreas de negócio como a produção, a área comercial, a distribuição e os recursos humanos, o SO99+ é capaz de gerar propostas de produção para os vários artigos.

Analisando o histórico de vendas de cada artigo, esta ferramenta de planeamento reúne-os em grupos que apresentam padrões de venda semelhantes. Em cada um desses grupos, permite identificar se existe sazonalidade e, caso esta se verifique, quais os seus índices, e se estes apresentam algum tipo de tendência.

Após esta análise, calcula a previsão de vendas dos vários grupos.

Sendo um dos objetivos deste projeto de dissertação a minimização dos erros de previsão, foi necessário avaliar a situação inicial dos mesmos. Para isso, foram calculados para o mês de fevereiro de 2020, período mais recente de dados disponíveis, os erros alusivos às previsões fornecidas pelo SO99+. Esses valores, apresentados na Tabela 10, foram obtidos para um nível de agregação por famílias de artigos.

Tabela 10 - Erros de previsão do SO99+, na situação inicial.

Família	MAE (kg)	MAPE (%)	VWMAPE (%)	BIAS (%)
Massas	1576760	27,63	27,63	-0,88
Bolachas	60050	20,68	20,68	2,08
CPA	400872	35,30	35,30	4,91

3.4 Estrutura das necessidades para planeamento de produção

Tendo como base todo este processo, o SO99+ calcula os diversos parâmetros que irão influenciar as estruturas das necessidades e, posteriormente, os planos de produção, nomeadamente:

Reorder Interval

O intervalo de reabastecimento traduz a frequência mínima, em dias, que uma produção deve ser realizada, calculando-se com base no stock relativo e no nível de serviço a atingir. É baseado em condições internas e impostas pela gestão da cadeia de aprovisionamentos, dando a noção de qual a quantidade necessária a produzir para assegurar que o nível de stocks nunca é excedido.

Este parâmetro permite obter o *reorder level*, nível de stock que normalmente gera a nova produção para cada artigo.

Minimum Lead Time

O *lead time* mínimo indica o intervalo, em dias, entre a data em que se processam os cálculos das propostas de produção e a data de entrega das encomendas, previamente estabelecida com os fornecedores. O SO99+ atribui valores de 1, 3, 5 e 7 dias a cada um dos grupos de produção, influenciando, desta forma o *reorder level*, o objetivo de nível de serviço e as propostas de produção geradas por este.

Safety Stock

Os stocks de segurança refletem o grau de variação nas vendas e nos *lead times*. Para evitar excedentes ou ruturas de stock, estes são criados como forma de amortecer o impacto dessas variações. O SO99+ calcula este parâmetro por recurso à Equação 3.1.:

$$\text{Safety Stock} = \text{Reorder Level} - \text{Minimum Lead Time} \quad (3.1.)$$

Através destes parâmetros, o SO99+ fornece propostas de produção, que cabe ao planeador validá-las posteriormente.

3.5 Apresentação do Caso de Estudo

O trabalho proposto tem como objetivo analisar e otimizar os métodos de previsão utilizados pela ferramenta SO99+, de forma a minimizar os erros de previsão. Posto isto, foi detetado que as propostas de produção fornecidas por esta ferramenta conduziam a erros

elevados. Neste contexto, esses erros tornaram-se o principal foco de melhoria deste projeto de dissertação. Após análise dos grupos de agregação que o *SO99+* gerava, verificou-se que a grande maioria dos artigos é classificada como “indefinido”. Isto leva a que não lhes seja atribuído qualquer tipo de índices de sazonalidade ou tendência e, conseqüentemente, as previsões obtidas estejam desajustadas. Em paralelo, tal classificação permitiu constatar que os valores das constantes de amortecimento exponencial α , β e γ não seriam os adequados.

Com o trabalho desenvolvido, pretende-se escolher o melhor método de previsão, melhorar os níveis de agregação dados aos artigos e obter níveis ótimos das constantes, para que, em simultâneo, os erros de previsão sejam minimizados e exista uma maior praticidade no processo de planeamento de produção. Para tal, com recurso à linguagem de programação *Python*, replicaram-se os métodos utilizados pelo *SO99+* e realizaram-se diversas simulações a partir dos históricos de vendas dos artigos da Cerealis, o que permitiu aferir qual o nível de agregação mais eficiente, o melhor método de previsão e quais os valores ótimos para as constantes de amortecimento exponencial. Com a combinação destas melhorias, pretende-se obter uma maior eficácia no processo de planeamento de produção, minimizando-se os erros previsionais.

4 Métodos de previsão de vendas propostos

No presente capítulo será apresentada a metodologia seguida no projeto de dissertação, após análise inicial dos dados e identificação de possíveis melhorias a implementar. Esta metodologia teve como base um estudo aprofundado das propostas de melhoria, nomeadamente na seleção dos valores ótimos das constantes de amortecimento, do nível de agregação favorável e do método de previsão mais adequado.

Os dados analisados e os métodos foram implementados recorrendo à linguagem de programação *Python*, tendo-se realizado diversas simulações para a obtenção dos resultados pretendidos.

4.1 Métodos de Amortecimento Exponencial

Com este projeto estudaram-se os métodos de previsão de vendas de amortecimento exponencial, uma vez que a ferramenta de planeamento utilizada pela Cerealis, o *SO99+*, realiza as suas previsões através de métodos preditivos deste tipo.

Para se iniciar o estudo, foi necessário realizar-se, em primeira instância, uma limpeza de artigos obsoletos ao histórico de vendas. Para tal, utilizou-se como critério de eliminação a inexistência de vendas há mais de dois anos. Isto é, artigos que não apresentassem uma única venda desde 2018 foram retirados da análise.

Com o objetivo de avaliar a possibilidade de melhorar as previsões da Cerealis, e uma vez que se pretendia comparar os níveis de precisão dos métodos de previsão, os dados utilizados dos históricos de vendas mensais de 2016 a 2019 foram divididos em dois *sets*:

- *Set* de Estudo ou de Treino – neste grupo utilizaram-se os históricos de todos os artigos dos anos de 2016 a 2018 para deteção do nível, tendência e sazonalidade, necessários para a realização das previsões.
- *Set* de Teste – através dos modelos obtidos, calcularam-se as previsões mensais para 2019, que foram comparadas com os valores reais de vendas desse mesmo ano.

Estes dois *sets* de dados foram os considerados para a realização das análises que se seguem, onde se retiraram as conclusões necessárias através da comparação das previsões realizadas para o *set* de teste com as vendas do mesmo período.

Deve notar-se que as previsões obtidas por intermédio das simulações desenvolvidas foram previsões rolantes. Isto é, foram obtidas para uma janela temporal mensal e considerando sempre o histórico mais recente, de forma a minimizar os erros previsionais, uma vez que quanto maior for o intervalo da previsão, tendencialmente maior será o desvio entre os valores calculados e os valores verificados realmente (Figura 16).

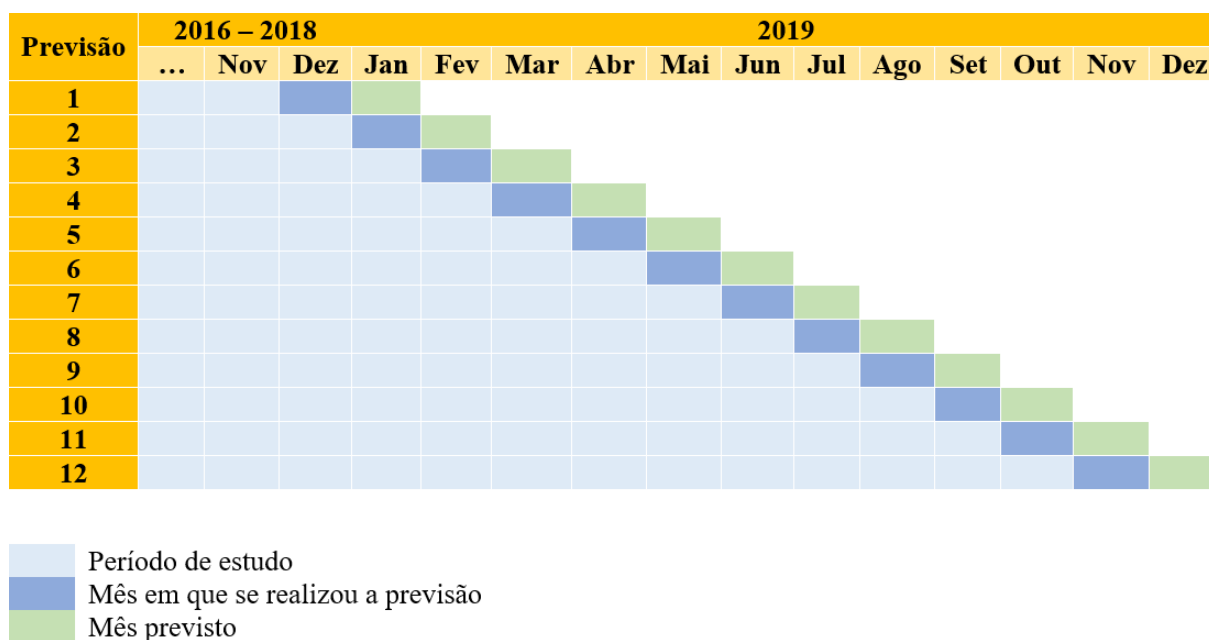


Figura 16 - Horizonte temporal das previsões realizadas.

Para se prosseguir com o estudo, foi necessário tomar uma sequência de decisões, antes da seleção final do método de previsão de vendas adequado. Estas seguiram a ordem das subsecções seguintes:

- Nível de Agregação;
- Modelo de Decomposição;
- Constantes de Amortecimento;
- Método de Previsão.

Para se tomar esta sequência de decisões foi usado o ano de 2018, para posterior avaliação final do método no período do *set* de teste, o ano de 2019.

4.1.1 Níveis de Agregação

Com o objetivo de estudar o nível de agregação que levaria à obtenção de menores erros de previsão, foram estabelecidos cinco níveis:

- SKU – as previsões geradas aplicavam-se a cada artigo individualmente, não havendo, desta forma, qualquer tipo de agregação de dados;
- Grupos ABC/XYZ – os artigos agruparam-se mediante a sua classificação na matriz ABC/XYZ. Foram criados nove grupos neste nível;
- Famílias de produtos – consoante a família de cada artigo (Massas, Bolachas ou CPA), estes foram agregados em três grupos;
- ABC/XYZ à Família – os artigos foram agrupados primeiramente à família e, dentro dessa família, foi realizada a análise ABC/XYZ. Dependendo da sua classificação, eram agregados no respetivo grupo, tendo-se criado vinte e sete grupos para este nível.
- *Clustering* à Família – geraram-se doze grupos de agregação, sendo cada artigo primeiramente agrupado à família e, dentro dessa, foram posteriormente agrupados mediante as suas semelhanças nas propriedades dos históricos de vendas.

Para cada um dos níveis, foram geradas previsões de vendas mensais para o ano de 2019. Como forma de os comparar, foi necessário calcular os erros VWMAPE, mencionados na Secção 2.6, sendo que estes teriam de ser determinados para o mesmo nível, para poderem ser comparáveis. Para tal, foi feita a desagregação dos dados para o nível de agregação mais baixo, o de SKU, e depois novamente agregados ao nível de Famílias, com o objetivo de se obterem os erros de previsão.

Para a desagregação dos dados, foram criados rácios através da média das vendas de cada artigo, nos últimos dois meses, obtendo-se o peso de cada um no total de vendas do seu grupo. Depois de se terem os rácios, foram obtidas as previsões de cada artigo através da Equação 4.1.:

$$P_t = P_t^{grupo} \times \text{rácio}_{artigo} \quad (4.1.)$$

Onde:

- P_t – previsão do artigo para o período atual t ;
- P_t^{grupo} – previsão do grupo para o período atual t ;
- rácio_{artigo} – peso das vendas do artigo no volume de vendas total do grupo.

Para agregar os dados para o cálculo do VWMAPE, foram determinados os desvios absolutos das previsões do ano de 2019 para cada um dos grupos do respetivo nível, recorrendo-se à Equação 4.2.:

$$\text{Desvio Absoluto} = |Y_t - P_t| \quad (4.2.)$$

Onde:

- Y_t – procura do artigo para o período atual t ;
- P_t – previsão do artigo para o período atual t ;

De seguida, mediante a Família a que corresponde cada grupo dos níveis a agregar, somaram-se os desvios absolutos por família. Por fim, dividiu-se o somatório pelo respetivo valor total de vendas da família, obtendo-se os vários valores dos erros VWMAPE.

Estes erros foram calculados para os três métodos de previsão, com o objetivo de se analisar, de forma global, qual o melhor nível de agregação.

Clustering à Família

A agregação dos artigos neste nível foi realizada com recurso ao algoritmo *clustering K-Means*.

O primeiro passo na agregação é a definição das coordenadas dos pontos. Estas foram definidas com base em treze dimensões:

- Volume de vendas, por SKU na família (normalizado);
- Índices de sazonalidade mensais.

O volume de vendas normalizado foi obtido realizando-se uma média dos valores de vendas normalizados dos anos 2016 a 2018. Já os índices de sazonalidade foram obtidos fazendo-se uma média dos índices mensais de cada ano. Isto é, relativamente a janeiro, calcularam-se os índices sazonais dos anos de 2016 a 2018 relativos a esse mês, fazendo-se posteriormente uma média desses três valores. Na Figura 17 encontra-se um exemplo das coordenadas utilizadas para um determinado artigo de cada uma das famílias. Por motivos de confidencialidade, a cada artigo foi atribuída uma combinação de letras idênticas, de forma a substituir o código do mesmo da empresa.

SKU	Volume de Vendas (normalizado)	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez
AAA Massas	0,58	0,57	0,78	0,64	0,31	0,29	0,28	0,30	0,42	0,47	2,42	3,31	2,20
BBB Bolachas	0,57	0,91	0,57	1,06	1,66	1,52	0,67	1,01	1,30	0,74	1,26	0,50	0,80
CCC CPA	0,41	2,11	2,07	0,42	0,24	0,47	0,36	0,70	0,21	0,40	1,35	2,06	1,61

Figura 17 - Exemplo de coordenadas utilizadas para implementação do método de *clustering K-Means*.

Além da definição das coordenadas dos pontos, é importante a determinação do número ótimo de *clusters*. Com recurso ao método do cotovelo, foi possível determinar o ponto que corresponderia ao valor ótimo k de *clusters*. Através das figuras do Anexo A, concluiu-se que o número ótimo seria de cinco *clusters* em cada família.

Nas Figura 18, Figura 19 e Figura 20 encontram-se os gráficos dos resultados da análise *clustering K-Means*, tendo estes sido traçados com recurso às coordenadas dos centroides dos *clusters* de cada uma das famílias de artigos.

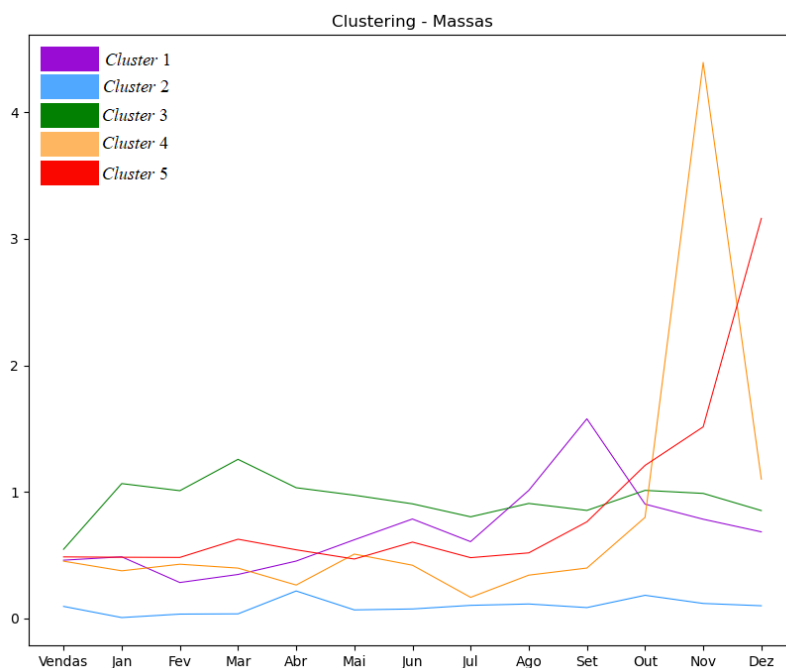


Figura 18 - Agregação dos artigos da família de Massas, por métodos de *clustering K-Means*.

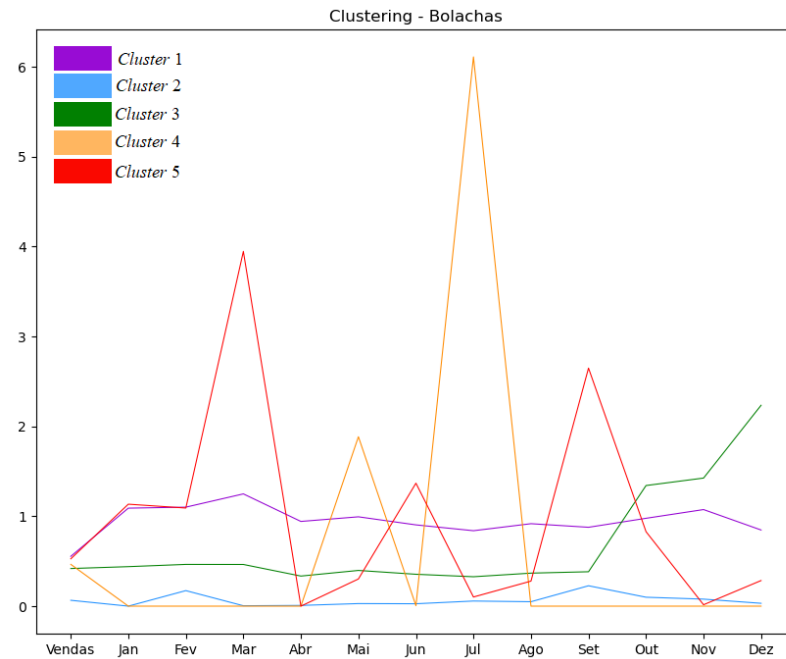


Figura 19 - Agregação dos artigos da família de Bolachas, por métodos de *clustering K-Means*.

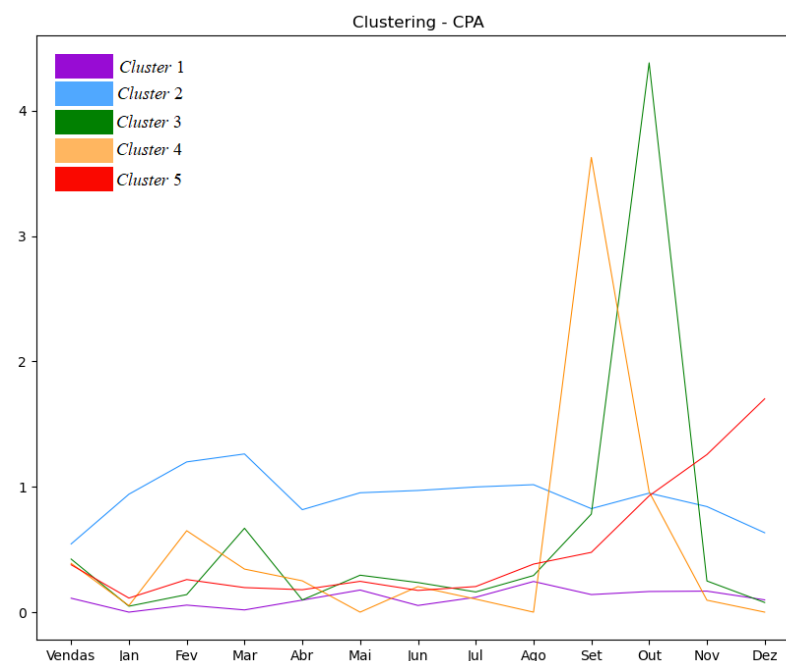


Figura 20 - Agregação dos artigos da família de CPA, por métodos de *clustering K-Means*.

4.1.2 Modelos Aditivos e Modelos Multiplicativos

Para se perceber qual o melhor tipo de modelo para cada grupo, foram realizadas simulações de forma a ser possível selecionar-se aquele modelo que levaria ao menor erro de previsão.

Para tal, realizou-se a decomposição das séries temporais de cada grupo, quer por decomposição aditiva quer por decomposição multiplicativa. Pegando-se na componente do Erro, e aplicando a função de autocorrelação (referida na Secção 2.3), foram traçados os gráficos ACF para cada um dos modelos de decomposição. De salientar que, quando a componente Erro das séries temporais apresentava valores negativos, não era possível aplicar-

se a decomposição multiplicativa, pelo que se assumiu, nesses casos, que o modelo a usar seria o aditivo.

De seguida, e uma vez que alguns coeficientes poderiam tomar valores negativos, realizou-se a soma dos quadrados dos respetivos coeficientes de autocorrelação da componente Erro de cada modelo através da Equação 4.3.:

$$\text{Somatório dos Quadrados de Coeficientes} = \sum_{t=1}^n i_t^2 \quad (4.3.)$$

Onde:

- i_t – coeficiente de autocorrelação da componente Erro, para o período t ;

O modelo mais adequado seria aquele que apresentasse o menor valor de soma, sendo que nos casos em que a componente Erro da série apresentava valores negativos, não sendo possível aplicar a decomposição multiplicativa, os valores atribuídos a estes somatórios foram NA (Não Aplicável).

Salienta-se que, de forma a tornar as simulações mais práticas, a seleção do modelo foi feita de forma automática para cada grupo em análise e em simultâneo com a escolha dos níveis de agregação.

4.1.3 Constantes de Amortecimento

Nesta altura, tornou-se necessário determinar quais os valores ótimos para cada uma das constantes de amortecimento dos três métodos de previsão, α , β e γ .

Para a obtenção destes valores, realizaram-se simulações, tendo-se dividido o *set* de treino em duas partes. A primeira, correspondente aos anos de 2016 e 2017, usada para a modelação da série e a segunda, o ano de 2018, como forma de comparar os erros de previsão gerados pelas combinações dos diferentes valores das constantes e seleccionar aqueles que levariam ao menor valor de erro. Foram, então, atribuídos valores entre 0,1 e 0,9 a cada uma das constantes e realizaram-se previsões com as várias combinações possíveis entre as três constantes. Para evitar valores extremos, no qual apenas se considera a observação mais antiga ou mais recente, os valores de 0 e 1 não foram testados nas simulações.

A decisão sobre o valor ótimo para as constantes foi tomada tendo por base a minimização do erro de previsão VWMAPE.

Desta forma, foi possível obter-se o melhor valor de cada constante para cada um dos métodos de previsão de vendas e para cada grupo previsional, uma vez que um valor pode ser ótimo num dos métodos e não noutra.

4.1.4 Métodos de Previsão

A última decisão a ser tomada foi relativa a qual o método de previsão mais benéfico a implementar em cada um dos grupos de previsão, mediante os valores ótimos de constantes de amortecimento obtidas. Os métodos equacionados foram o método de amortecimento exponencial simples, o método de amortecimento exponencial Linear de Holt e o método de amortecimento exponencial de Holt-Winters.

Para tal, foram realizadas simulações de previsão para o ano de 2019, utilizando as constantes de amortecimento, para cada um dos métodos, seleccionadas anteriormente.

O método de previsão foi escolhido tendo como critério a minimização do erro VWMAPE.

4.1.5 Análise de Resultados

O critério de comparação de resultados dos métodos de previsão é o erro VWMAPE, calculado com os desvios absolutos obtidos ao nível do artigo e por mês e, posteriormente, agregados ao nível de agregação selecionado e ao ano de 2019.

As primeiras decisões tomadas foram relativas aos níveis de agregação e aos modelos de decomposição das séries temporais.

Na Tabela 11 encontram-se os erros obtidos para cada nível de agregação, após desagregação e agregação para o nível de Famílias.

Tabela 11 - Erros VWMAPE para seleção de níveis de agregação.

Família	Método	VWMAPE (%)				
		Nível SKU	Nível Grupos ABC/XYZ	Nível Famílias	Nível ABC/XYZ à Família	Nível Clustering à Família
Massas	Simples	35,68	41,43	42,67	41,19	35,59
	Holt	39,32	45,70	42,89	41,74	43,21
	Holt-Winters	43,87	43,21	43,31	42,44	44,52
Bolachas	Simples	41,63	42,36	41,20	46,44	47,66
	Holt	43,64	74,56	41,02	46,01	48,56
	Holt-Winters	53,17	45,75	41,58	49,23	47,66
CPA	Simples	42,41	41,75	42,71	44,05	45,21
	Holt	46,94	54,34	43,74	46,19	46,51
	Holt-Winters	53,43	43,88	43,03	48,91	46,54

O primeiro nível a ser eliminado foi o de Grupos ABC/XYZ, pois, como se constata, é o que apresenta maiores erros de previsão. De seguida, eliminou-se o nível de SKU, uma vez que, apesar de na família das Massas e dos CPA apresentar valores bastante competitivos, tendo em conta a dimensão da empresa e a quantidade de artigos existente, tornava-se pouco prático a implementação de previsões a este nível.

Restam, então, os níveis de Famílias, o de ABC/XYZ à Família e o de Clustering à Família.

Quanto à família de Massa, o nível selecionado foi o de Clustering à Família, pois é o que de entre os dois restantes níveis apresenta menores valores de erro VWMAPE. Relativamente às famílias de Bolachas e de CPA, observando a Tabela 11, constata-se que para o nível de Famílias os erros de previsão são os mais baixos. Por este motivo, para as Bolachas e os CPA o nível de agregação escolhido foi o de Famílias.

Na Tabela 12 são apresentados os valores dos somatórios dos quadrados dos coeficientes obtidos pela aplicação da função de autocorrelação na componente de erros das séries temporais de cada grupo de previsão.

Tabela 12 - Somatórios dos quadrados dos coeficientes para seleção dos modelos de decomposição.

Nível de Agregação	Grupo	Somatório dos Quadrados dos Coeficientes	
		Modelo Aditivo	Modelo Multiplicativo
Clustering à Família	<i>Cluster 1</i> Massas	1,7135	1,4650
	<i>Cluster 2</i> Massas	NA	NA
	<i>Cluster 3</i> Massas	1,5447	1,5537
	<i>Cluster 4</i> Massas	1,6115	1,5005
	<i>Cluster 5</i> Massas	1,6690	1,6633
Famílias	Bolachas	1,3194	1,3406
Famílias	CPA	1,5823	1,4870

Observando-se os resultados da Tabela 12, foram tomadas as seguintes decisões relativamente ao modelo a usar para cada grupo:

- Modelo Aditivo – *Cluster 2* Massas, *Cluster 3* Massas e Bolachas;
- Modelo Multiplicativo – *Cluster 1* Massas, *Cluster 4* Massas, *Cluster 5* Massas e CPA.

Após a escolha dos níveis de agregação e da utilização de modelos aditivos/multiplicativos, seguiu-se a análise dos valores ótimos para as três constantes de amortecimento. Na Figura 32 do Anexo B constam os valores dos erros de previsão obtidos, mediante o método de previsão, e os valores selecionados das respetivas constantes.

Foi também através da Figura 32 do Anexo B que se retiraram conclusões quanto ao método de previsão de vendas a usar, tendo-se escolhido aquele que teria o menor erro VWMAPE dentro do seu grupo.

Na Figura 21 encontram-se os métodos de previsão selecionados, bem como as respetivas constantes de amortecimento e os erros MAE, MAPE e o BIAS.

Grupo	Nível	Método	Modelo	α	β	γ	MAE (kg)	MAPE (%)	VWMAPE (%)	BIAS (%)
<i>Cluster 1</i> Massas	<i>Clustering à Família</i>	Holt	Multiplicativo	0,3	0,6	-	4660,46	47,02	36,43	-2,30
<i>Cluster 2</i> Massas		Holt	Aditivo	0,1	0,1	-	4505,07	245,85	79,11	26,60
<i>Cluster 3</i> Massas		Holt	Aditivo	0,1	0,3	-	5405,51	33,63	29,29	2,59
<i>Cluster 4</i> Massas		Holt-Winters	Multiplicativo	0,1	0,1	0,4	5034,28	65,53	64,26	-2,63
<i>Cluster 5</i> Massas		Holt-Winters	Multiplicativo	0,1	0,1	0,4	6049,72	129,49	50,63	-4,66
Bolachas	Famílias	Holt-Winters	Aditivo	0,1	0,1	0,7	1071,13	66,42	40,08	-1,13
CPA		Holt-Winters	Multiplicativo	0,1	0,1	0,1	1776,27	57,34	42,00	9,30

Figura 21 - Tabela dos métodos de previsão selecionados.

Constata-se que o método de previsão de vendas adequado para a maioria dos grupos é o de Holt-Winters, sendo que para os grupos *Cluster 1*, *Cluster 2* e *Cluster 3* da família das Massas o método adequado é o de Amortecimento Exponencial Linear de Holt. Analisando-se também o BIAS de cada um dos grupos, verifica-se que a grande maioria apresenta enviesamento negativo do erro, querendo isto dizer que as previsões realizadas têm um desvio negativo relativamente ao valor de vendas que de facto ocorreu nesse período. No caso dos grupos *Cluster 2* e *Cluster 3* da família das Massas, uma vez que o enviesamento do erro é

positivo, conclui-se que as vendas observadas ficaram acima do valor previsto para esse período.

Pode constatar-se que o grupo *Cluster 2 Massas* é o que apresenta o maior desvio nas previsões, o que implica que existe uma tendência de os valores previsionais de vendas desse grupo estarem abaixo da média dos valores reais. Este desvio deve-se ao facto de no período utilizado para a modelação da série não haver registo de vendas, não se tendo conseguido captar as componentes de sazonalidade, tendência e erro de forma adequada neste grupo. Nos restantes grupos, os desvios nas previsões apresentam uma diferença menor que 10%.

De forma a aferir a *performance* dos métodos seleccionados para a realização das previsões de vendas, foi estabelecido um critério com recurso a um método de previsão *naïve*, o método de médias móveis com uma janela de observações passadas de três meses. Comparando-se os valores dos erros VWMAPE, para cada grupo, entre o método *naïve* com os do método preditivo seleccionado, se os erros VWMAPE do método *naïve* forem superiores ao do método seleccionado, então pode afirmar-se que o método de amortecimento exponencial escolhido apresenta uma *performance* favorável.

Na Tabela 13 constam os erros de previsão VWMAPE do método *naïve* e do método seleccionado de cada grupo.

Tabela 13 - Comparação dos métodos de amortecimento exponencial seleccionados com método *naïve*.

Grupo	VWMAPE (%)	
	Método <i>Naïve</i>	Método de Amortecimento Exponencial Seleccionado
Cluster 1 Massas	45,55	36,43
Cluster 2 Massas	82,11	79,11
Cluster 3 Massas	36,73	29,29
Cluster 4 Massas	68,99	64,26
Cluster 5 Massas	53,61	50,63
Bolachas	45,62	40,08
CPA	46,63	42,00

Verifica-se que as previsões obtidas pelos métodos de amortecimento exponencial apresentam transversalmente erros VWMAPE menores que os obtidos pelo método *naïve*, sendo essa diferença na ordem dos 3-10 pontos percentuais (p.p.), permitindo inferir que estes levarão a resultados mais assertivos.

Realizou-se também o estudo comparativo com os resultados de previsões de vendas obtidos pelo SO99+ para o mês de fevereiro de 2020, mencionados na Secção 3.3. Foram, então, obtidas as previsões para este período específico, utilizando-se os métodos e respectivas constantes de amortecimento seleccionadas até aqui, para cada uma das famílias de artigos e, posteriormente, os valores de VWMAPE respectivos, de forma a se compararem ambos os cenários. Na Tabela 14 encontram-se sumariados esses valores.

Deve referir-se que, ao contrário do que foi feito até agora, estes erros foram calculados ao nível da família, de forma a poderem comparar-se de forma direta os resultados.

Tabela 14 - Comparação dos métodos de amortecimento exponencial selecionados com previsões obtidas pelo SO99+.

Grupo	VWMAPE (%)	
	SO99+	Método de Amortecimento Exponencial Selecionado
Massas	27,63	19,70
Bolachas	20,68	18,12
CPA	35,30	26,62

Como se verifica pela Tabela 14, a otimização dos valores das constantes de amortecimento levou à minimização dos erros de previsão nas três famílias de artigos. Deve salientar-se que a maior diferença se verificou na família dos CPA, onde houve uma diminuição do erro VWMAPE em 8,68 p.p..

Na família das Massas houve uma diminuição de 7,93 p.p. no erro VWMAPE relativamente ao valor obtido no erro das previsões do SO99+, enquanto que na família das Bolachas essa diferença foi de 2,56p.p..

Pode então inferir-se que a otimização dos valores das constantes, a seleção adequada do nível de agregação e a escolha do método adequado de previsão são fatores essenciais para se conseguir obter os valores mínimos de erros VWMAPE.

4.1.6 Apresentação dos resultados dos métodos de previsão de vendas propostos

A presente subsecção tem como objetivo mostrar a evolução das previsões de cada grupo ao longo do tempo, comparativamente com as observações do período correspondente. De referir que os valores do eixo vertical foram ocultados por motivos de confidencialidade, correspondendo estes ao volume de vendas verificado em cada período correspondente.

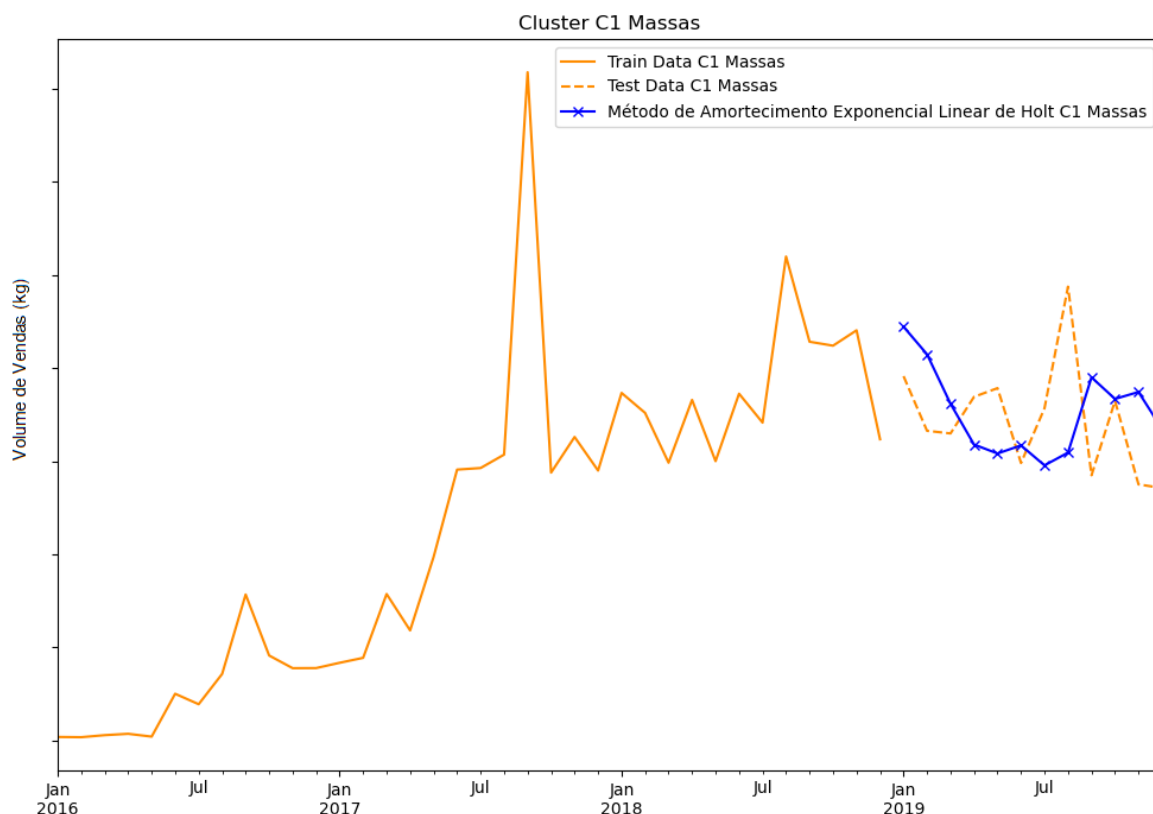


Figura 22 - Previsão de vendas do grupo *Cluster 1 Massas*, através do método de amortecimento exponencial Linear de Holt.

Na Figura 22 apresentam-se os resultados obtidos para o grupo *Cluster 1* da família das Massas, obtidos através do método de amortecimento exponencial Linear de Holt, com constante de amortecimento α igual a 0,3 e β igual a 0,6.

Observando-se o gráfico, consegue verificar-se que não há um padrão sazonal nítido. No entanto, pode aferir-se a existência de um pico de vendas por volta do mês de julho, repetindo-se tal em todos os anos. A impercetibilidade de um padrão sazonal levou a que o método de previsão selecionado fosse o de amortecimento exponencial Linear de Holt, onde apenas se considera a existência de tendência.

Ao longo do período de *set* de estudo utilizado para modelação da série antes de se simularem os valores das constantes de amortecimento, denota-se que a tendência do volume de vendas sofre um abrupto aumento do declive em meados de 2017, estabilizando posteriormente. Devido à tendência crescente das vendas verificada neste período, o valor obtido para a constante de amortecimento β foi alto, levando à consideração de uma tendência variável nas previsões do ano de 2019, visto que se atribuiu maior peso às observações mais recente.

As previsões obtidas apresentam uma correção moderada relativamente ao erro verificado nas observações próximas, o que é explicado pelo valor da constante de amortecimento α , que dita o grau de correção a dar ao nível da série.

Esta combinação de constantes de amortecimento levou aos resultados preditivos observados. O erro de previsão obtido, de 36,43%, pode justificar-se pelos valores das constantes de α e β , que atribuíram um peso maior às observações recentes.

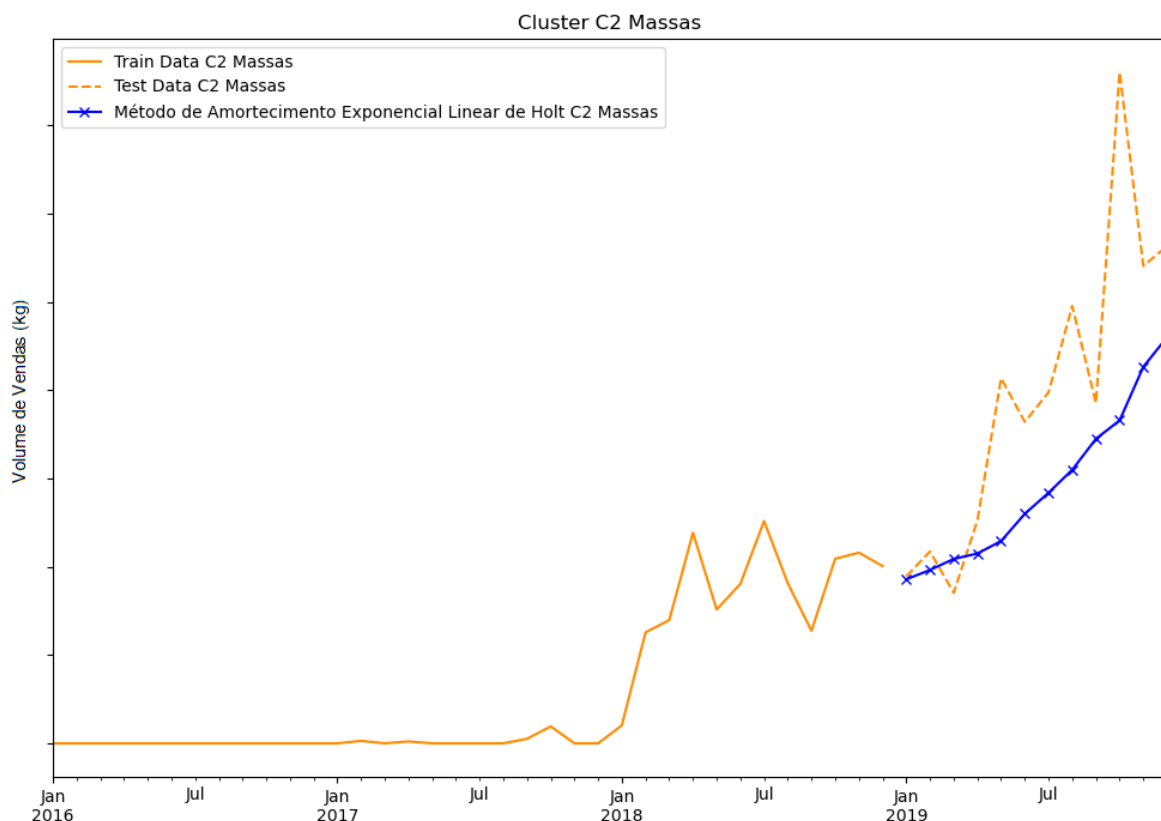


Figura 23 - Previsão de vendas do grupo *Cluster 2* Massas, através do método de amortecimento exponencial linear de Holt.

Na Figura 23 apresentam-se os resultados obtidos para o grupo *Cluster 2* da família das Massas, obtidos através do método de amortecimento exponencial Linear de Holt, com constante de amortecimento α igual a 0,1 e β igual a 0,1.

Uma vez que as previsões realizadas foram de forma rolante, e visto que as vendas observadas apresentam, desde o seu início em 2018, uma tendência crescente, tal foi replicado nos valores obtidos em 2019. No entanto, e dado o baixo valor da constante de amortecimento β , o declive não foi tão acentuado quanto o das observações mais recentes à previsão, pois o modelo atribuiu maior peso às observações mais antigas.

As previsões obtidas apresentam uma correção baixa relativamente ao erro verificado nas observações próximas, o que é explicado pelo baixo valor da constante de amortecimento α , já que quanto mais próximo de zero for o seu valor, menor é a correção do nível da série.

Esta combinação de constantes de amortecimento levou aos resultados preditivos observados. O erro de previsão obtido foi de 79,11%. Este valor elevado pode ser justificado pelo facto de que nos dois primeiros anos não existirem vendas. Este grupo apresenta, ainda, um valor de enviesamento relativamente ao erro elevado, o que se pode verificar no gráfico.

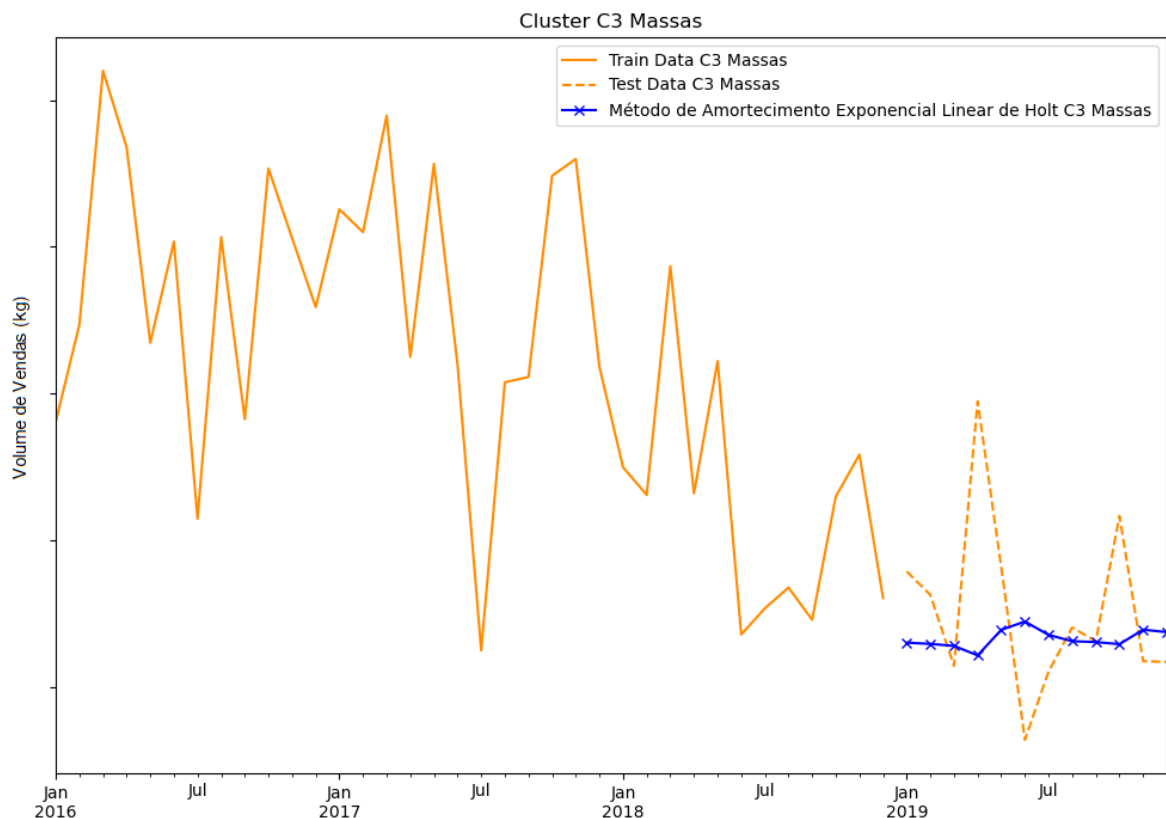


Figura 24 - Previsão de vendas do grupo *Cluster 3 Massas*, através do método de amortecimento exponencial linear de Holt.

Na Figura 24 apresentam-se os resultados obtidos para o grupo *Cluster 3* da família das Massas, obtidos através do método de amortecimento exponencial Linear de Holt, com constante de amortecimento α igual a 0,1 e β igual a 0,3.

Pelo gráfico pode concluir-se que não existe um padrão de sazonalidade, o que suporta a decisão de não seleção do método de amortecimento exponencial de Holt-Winters.

Consegue-se, por outro lado, denotar a existência de uma tendência decrescente do volume de vendas, pelo que a constante de amortecimento β tomou um valor moderado. Tal permitiu que nas previsões de 2019 a linha de declive da tendência fosse semelhante à das observações passadas, como janeiro de 2017, onde se pode inferir que a tendência é quase constante.

As previsões obtidas apresentam uma correção de nível baixa relativamente ao erro verificado nas observações próximas, o que é explicado pelo baixo valor da constante de amortecimento α .

Esta combinação de constantes de amortecimento levou aos resultados preditivos observados. O erro de previsão obtido, de 29,29%, pode ser justificado pelos valores das constantes de α e β , que atribuíram um peso maior às observações passadas, no caso do nível, e a observações um pouco mais recentes, no caso da tendência.

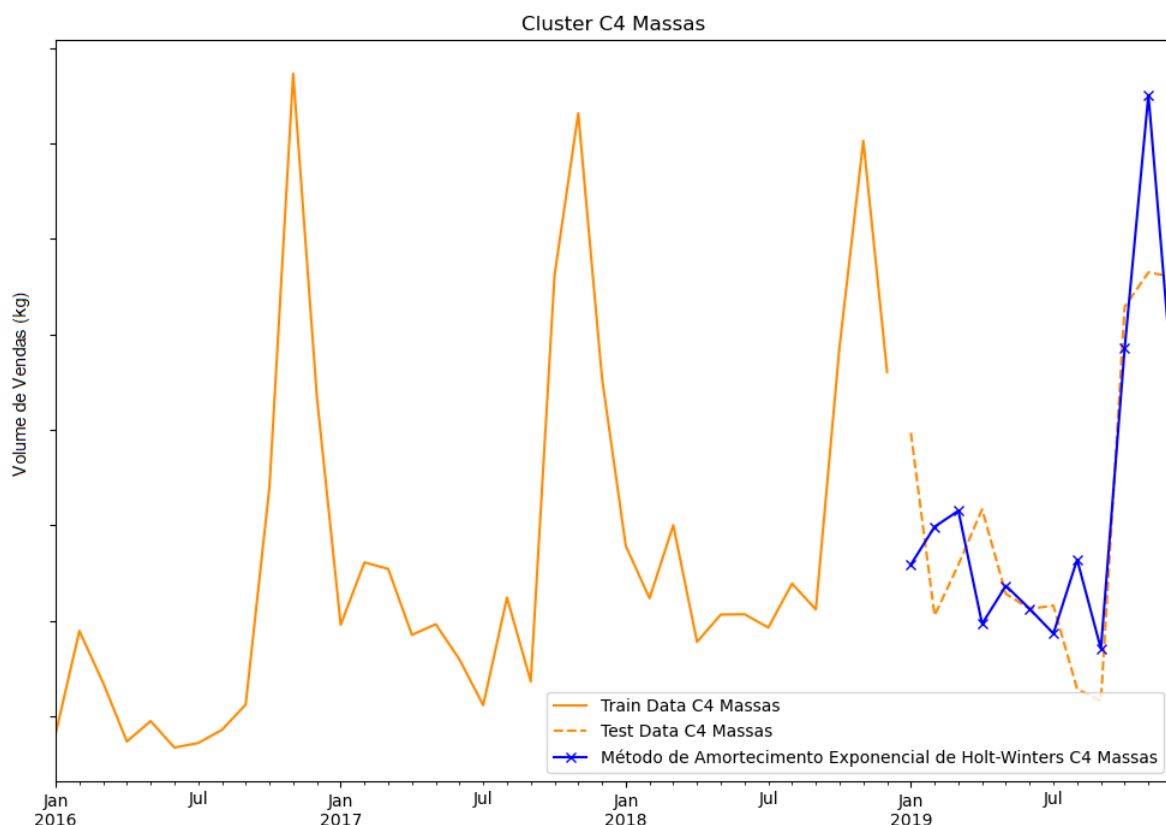


Figura 25 - Previsão de vendas do grupo *Cluster 4* Massas, através do método de amortecimento exponencial linear de Holt-Winters.

Na Figura 25 apresentam-se os resultados obtidos para o grupo *Cluster 4* da família das Massas, obtidos através do método de amortecimento exponencial de Holt-Winters, com constante de amortecimento α igual a 0,1, β igual a 0,1 e γ igual a 0,4.

Observando-se o gráfico, consegue verificar-se a existência, de forma nítida, de um padrão sazonal, com picos de vendas na época de final de ano.

Esta sazonalidade foi replicada nas previsões obtidas, verificando-se que o padrão sazonal de 2019 é semelhante àquele verificado nos anos anteriores. Tal pode ser explicado pelo valor da constante de amortecimento γ , que levou à replicação da sazonalidade verificada ao longo dos anos.

Ao longo do período de *set* de estudo utilizado para modelação da série antes de se simularem os valores das constantes de amortecimento, denota-se que a tendência do volume de vendas é constante. Como tal, o valor obtido para a constante de amortecimento β foi baixo, levando à consideração de uma tendência constante nas previsões do ano de 2019.

As previsões obtidas apresentam uma correção baixa relativamente ao erro verificado nas observações próximas, o que é explicado pelo baixo valor da constante de amortecimento α , já que quanto mais próximo de zero for o seu valor, menor é a correção do nível da série.

Esta combinação de constantes de amortecimento levou aos resultados preditivos observados. O erro de previsão obtido, de 64,26%, pode ser justificado, por um lado, pela variação no padrão de vendas verificado em 2019 e, por outro, pelos valores das constantes de α , β e γ , que atribuíram um peso maior às observações passadas, no caso do nível e tendência da série, e às observações recentes, no caso da sazonalidade.

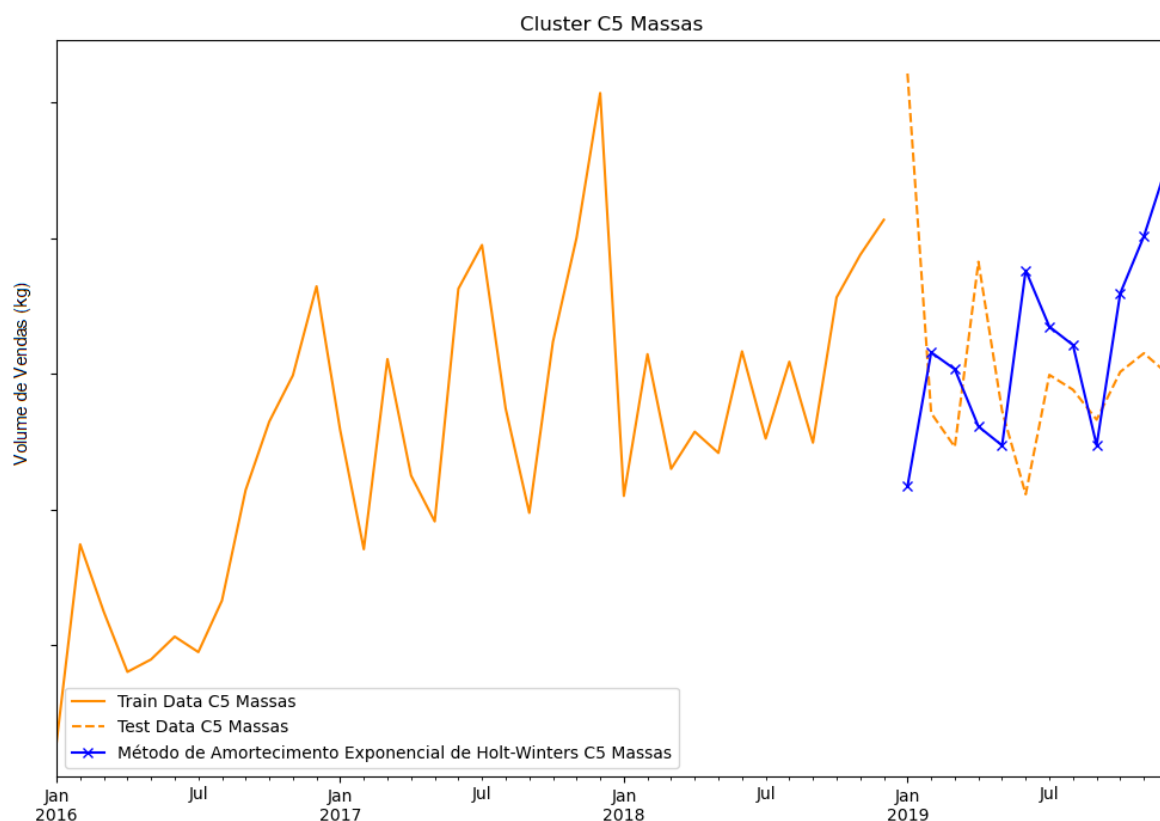


Figura 26 - Previsão de vendas do grupo *Cluster 5 Massas*, através do método de amortecimento exponencial linear de Holt-Winters.

Na Figura 27 apresentam-se os resultados obtidos para o grupo *Cluster 5* da família das Massas, obtidos através do método de amortecimento exponencial de Holt-Winters, com constante de amortecimento α igual a 0,1, β igual a 0,1 e γ igual a 0,4.

Observando-se o gráfico, consegue verificar-se a existência de um padrão sazonal, com picos de vendas também na época de final de ano.

Esta sazonalidade foi replicada nas previsões obtidas, verificando-se que o padrão sazonal de 2019 é muito semelhante àquele verificado no ano anterior. Ou seja, as sazonalidades dos meses do ano previsto são semelhantes aos períodos homólogos dos anos próximos. Tal pode ser explicado pelo valor da constante de amortecimento γ , que levou à atribuição de maior peso ao padrão sazonal das observações mais recentes.

Verifica-se que a tendência do volume de vendas é constante ao longo de todo o período. Como tal, o valor obtido para a constante de amortecimento β foi baixo, levando à consideração de uma tendência constante nas previsões do ano de 2019.

As previsões obtidas apresentam uma correção baixa relativamente ao erro verificado nas observações próximas, o que é explicado pelo baixo valor da constante de amortecimento α , já que quanto mais próximo de zero for o seu valor, menor é a correção do nível da série.

Esta combinação de constantes de amortecimento levou aos resultados preditivos observados. O erro de previsão obtido, de 50,63%, valor justificado pelos valores das constantes de α , β e γ , que atribuíram um peso maior às observações passadas, no caso do nível e tendência da série, e às observações recentes, no caso da sazonalidade.

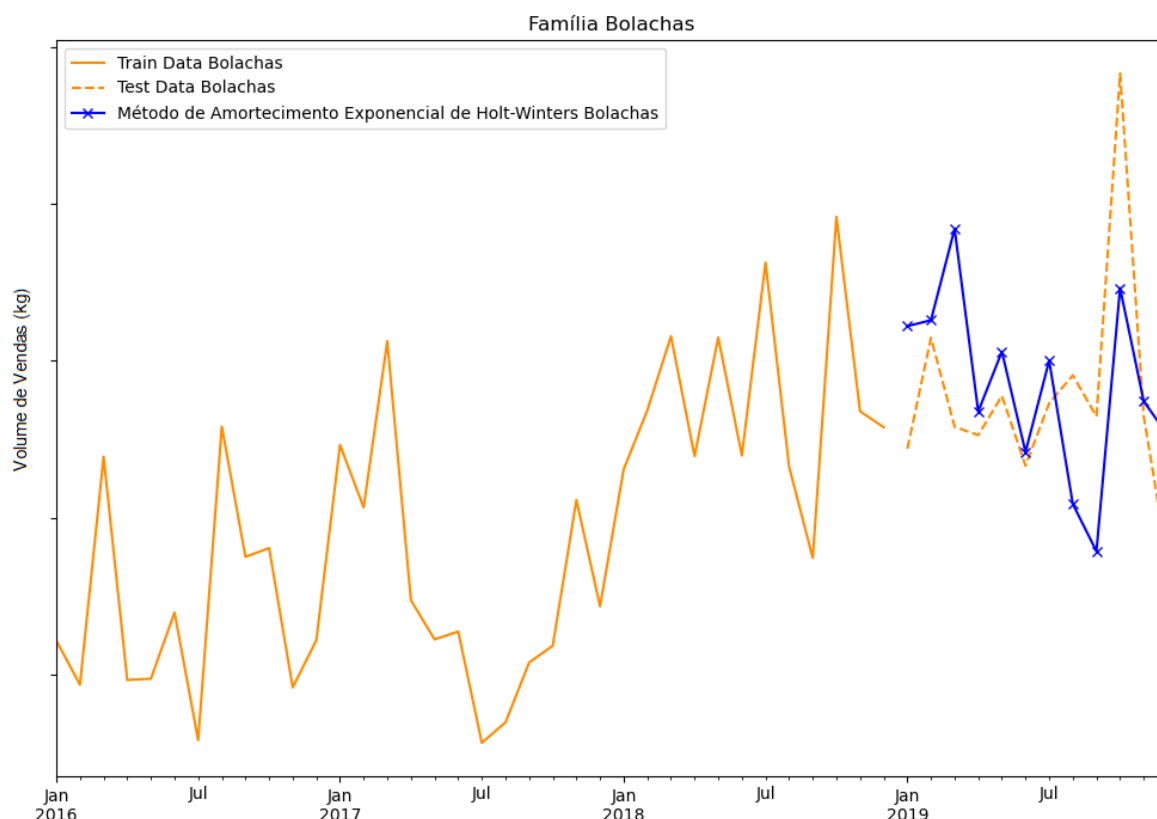


Figura 27 - Previsão de vendas da família das Bolachas, através do método de amortecimento exponencial de Holt-Winters.

Na Figura 27 apresentam-se os resultados obtidos para a família das Bolachas, obtidos através do método de amortecimento exponencial de Holt-Winters, com constante de amortecimento α igual a 0,1, β igual a 0,1 e γ igual a 0,7.

Observando-se o gráfico, consegue verificar-se a existência de um padrão sazonal, com picos de vendas na época de inverno, tal como se constatou na análise de sazonalidade da secção 3.2.2.

Esta sazonalidade foi replicada nas previsões obtidas, verificando-se que o padrão sazonal de 2019 é muito semelhante àquele verificado no ano anterior. Ou seja, as sazonalidades dos meses do ano previsto são semelhantes aos períodos homólogos do ano imediatamente anterior. Tal pode ser explicado pelo valor elevado da constante de amortecimento γ , que levou à atribuição de maior peso ao padrão sazonal das observações mais recentes.

Ao longo do período de set de estudo utilizado para modelação da série antes de se simularem os valores das constantes de amortecimento, denota-se que a tendência do volume de vendas é constante, havendo um abrupto aumento do declive das vendas no início de 2018, que estabilizou ao fim de alguns meses. Como tal, o valor obtido para a constante de amortecimento β foi baixo, levando à consideração de uma tendência constante nas previsões do ano de 2019, visto que se atribuiu maior peso às observações mais antigas, onde o declive da linha de tendência foi nulo.

As previsões obtidas apresentam uma correção baixa relativamente ao erro verificado nas observações próximas, o que é explicado pelo baixo valor da constante de amortecimento α , já que quanto mais próximo de zero for o seu valor, menor é a correção do nível da série.

Esta combinação de constantes de amortecimento levou aos resultados preditivos observados. O erro de previsão obtido, de 40,08%, pode ser justificado pelos valores das

constantes de α , β e γ , que atribuíram um peso maior às observações passadas, no caso do nível e tendência da série, e às observações recentes, no caso da sazonalidade.

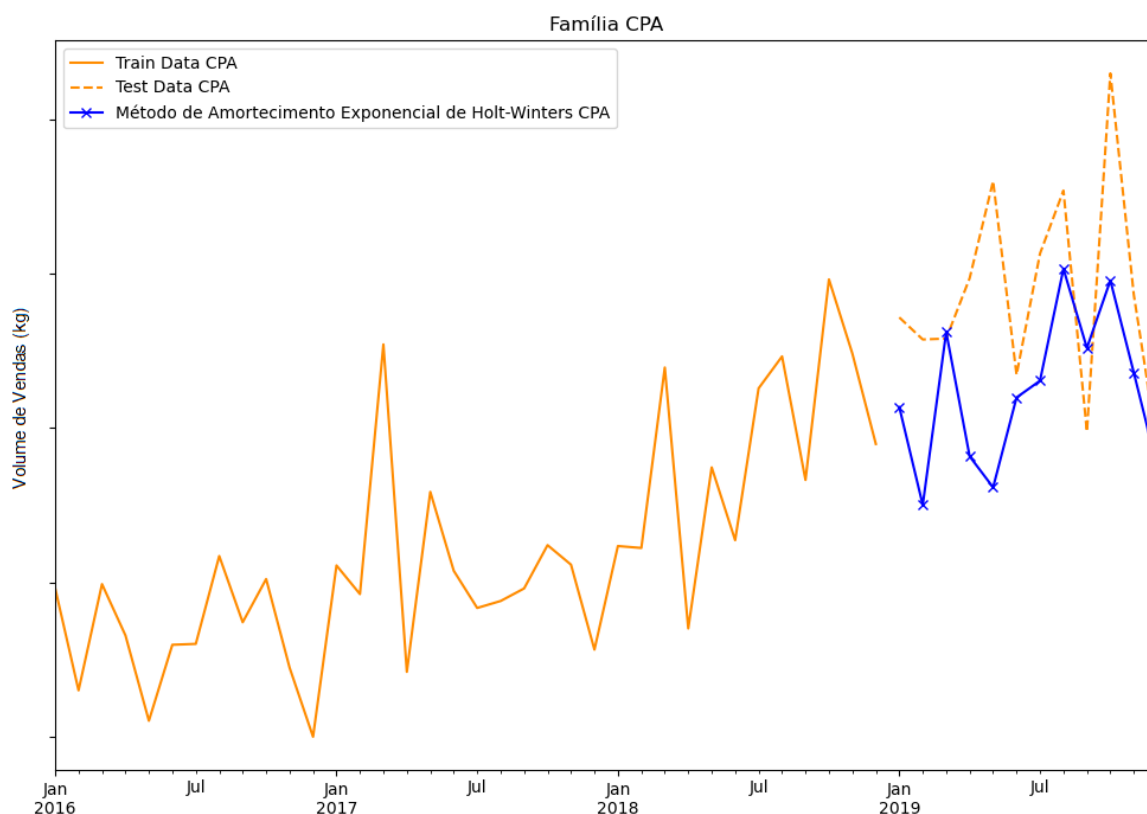


Figura 28 - Previsão de vendas da família dos CPA, através do método de amortecimento exponencial de Holt-Winters.

Na Figura 28 apresentam-se os resultados obtidos para a família dos CPA, obtidos através do método de amortecimento exponencial de Holt-Winters, com constante de amortecimento α igual a 0,1, β igual a 0,1 e γ igual a 0,1.

Consegue-se verificar a existência de um padrão sazonal anual, com picos de vendas nos primeiros meses dos anos, tal como se constatou na análise de sazonalidade realizada previamente neste projeto de dissertação.

Este padrão foi replicado nas previsões obtidas, verificando-se que o padrão sazonal de 2019 é muito semelhante ao daquele verificado no ano das primeiras observações. Ou seja, as sazonalidades dos meses do ano previsto são semelhantes aos períodos homólogos do ano de 2016. Tal pode ser explicado pelo baixo valor da constante de amortecimento γ , que levou à atribuição de maior peso ao padrão sazonal das observações mais antigas.

Ao longo do período de *set* de estudo utilizado para modelação da série antes de se simularem os valores das constantes de amortecimento, denota-se que a tendência do volume de vendas apresenta um crescimento ligeiro ao longo de todo o período. Como tal, o valor obtido para a constante de amortecimento β foi baixo, levando à consideração de uma tendência quase constante nas previsões do ano de 2019, visto que se atribuiu maior peso às observações mais antigas, onde o declive da linha de tendência foi baixo.

As previsões obtidas apresentam uma correção baixa relativamente ao erro verificado nas observações próximas, o que é explicado pelo baixo valor da constante de amortecimento α , já que quanto mais próximo de zero for o seu valor, menor é a correção do nível da série.

Esta combinação de constantes de amortecimento levou aos resultados preditivos observados. O erro de previsão obtido, de 42,00%, pode justificar-se pelos valores das

constantes de α , β e γ , que atribuíram um peso maior às observações passadas, e não tendo sido considerado um maior crescimento no volume de vendas, como se verificou em 2019.

4.2 Implementação de métodos de *Machine Learning*

Machine learning (ML) é um ramo da ciência computacional que implementa algoritmos que analisam um determinado conjunto de dados, conhecendo de forma independente os seus comportamentos (Bourne, 2016).

Aplicando ML na obtenção de previsões, existe a expectativa de se obterem resultados mais precisos devido à variedade de algoritmos disponíveis para a resolução de problemas que contemplam outro tipo de informação disponível sobre os artigos.

Comparativamente com os métodos clássicos de previsão de vendas, tais como os enunciados neste projeto de dissertação, os métodos de ML apresentam uma maior rapidez no processo de previsão, permitem obter resultados mais precisos através da deteção de padrões impercetíveis nas séries temporais e também uma maior capacidade de adaptação a alterações nos dados causadas por fatores externos, como ações promocionais (Tabela 15).

Tabela 15 - Comparação entre previsões por métodos tradicionais e métodos de *machine learning* (adaptado) (Demand Forecasting Methods: Using Machine Learning and Predictive Analytics to See the Future of Sales, 2019).

	Previsão por Métodos Tradicionais	Previsão por ML
Capacidade de análise de um grande número de dados	Baixa	Alta
Necessidade de mão de obra	Alta	Baixa
Mínimo de dados necessários	Pouco	Muito
Complexidade de manutenção	Baixa	Alta
Melhor aplicação	Séries estáveis, produtos com vendas definidas e previsões de médio-longo prazo.	Séries variáveis, novo produtos e previsões de curto-médio prazo.

No entanto os algoritmos utilizados em ML podem ter problemas na modelação dos dados, uma vez que por vezes não realizam a decomposição das séries de forma adequada, levando à obtenção de maus resultados de previsão.

Assim, os métodos de previsão por ML aliados aos métodos clássicos de previsão apresentam um enorme potencial. Estas duas abordagens completam-se na medida em que os métodos clássicos realizam boas decomposições de dados em termos das componentes de erros, sazonalidade e tendência, que são depois utilizadas como *inputs* necessários na modelação pelos algoritmos de ML.

Desta forma, a implementação de ML na previsão de vendas para posterior planeamento de produção, em paralelo com os métodos de amortecimento exponencial, constitui uma proposta de melhoria para a Cerealis.

Não só a rapidez do processo, mas também a sua automatização na análise de dados de históricos de vendas permitia aos colaboradores do departamento de planeamento ter um maior foco na resolução de problemas e otimização de parâmetros, ajudando na constante melhoria do processo de previsões de vendas.

A implementação de ML na rotina de planeamento da Cerealis exigiria repensar a utilização do *SO99+* enquanto método de geração de previsões. necessitando esta mudança de uma avaliação cuidada e ponderada relativamente aos próximos passos a seguir no projeto.

5 Conclusões e perspectivas de trabalho futuro

O projeto desenvolvido no âmbito desta dissertação de mestrado teve como objetivo a otimização dos métodos de previsão de amortecimento exponencial, através da reparametrização das constantes de amortecimento e formulação de níveis de agregação adequados, para posterior implementação na ferramenta de produção *SO99+*, de forma a ser possível a minimização dos erros de previsão, obtendo-se valores mais assertivos, levando ao melhor planeamento de produção dos artigos da Cerealis.

Para tal, foram desenvolvidos programas de simulação dos métodos de amortecimento exponencial, com recurso à linguagem de *Python*, com o objetivo de se testarem várias combinações de valores de constantes, vários níveis de agregação e modelos aditivos e multiplicativos, para os três métodos de amortecimento exponencial simples, linear de Holt e de Holt-Winters.

Primeiramente, foi analisado qual o peso que cada família tem no panorama global da empresa, tendo-se concluído que as Massas é a família que apresenta maior peso, correspondendo a 53,43% dos artigos, seguindo-se a família dos CPA com 34,80%. A família das Bolachas é a que apresenta menor peso, detendo 11,77% dos artigos da empresa. Foi também realizado um estudo detalhado dos históricos de vendas dos artigos, de forma a ser possível a sua caracterização, quer ao nível de ABC/XYZ, quer também quanto à sua sazonalidade e tendência. Através deste estudo foi possível uma melhor interpretação dos métodos de previsão, na medida em que permitiu apurar que 43,68% dos artigos apresentam variabilidade no volume de vendas, implicando que a previsão de vendas destes artigos apresente mais obstáculos.

Relativamente à análise de sazonalidade, verificou-se que os três grupos apresentavam sazonalidade anual, com picos de vendas focadas no mês de dezembro. Quanto à tendência, as Massas apresentam declive nulo, ao passo que as Bolachas e os CPA têm declive positivo na tendência, implicando que, nestas duas últimas famílias, o volume de vendas tem vindo a crescer ao longo do tempo.

De forma a obterem-se resultados preditivos de vendas mais precisos, foram estudados níveis de agregação passíveis de serem implementados, uma vez que foi um dos pontos identificados como possível melhoria para o *SO99+*. Perante tal estudo, foi possível estabelecerem-se grupos para níveis de agregação, tendo-se concluído que, mediante os resultados obtidos dos erros *VWMAPE* e optando-se pelo nível que tinha menor valor de erro, o nível de agregação adequado para os artigos das famílias de Massas e Bolachas seria o de *Clustering* à Família e o da família dos CPA seria de Famílias. A agregação de artigos em níveis superiores torna-se fundamental, pois permite não só minimizar o tempo despendido no processo de planeamento como também diminuir os erros das previsões e obter séries temporais mais estáveis.

Sendo os métodos de previsão utilizados pelo *SO99+* métodos de amortecimento exponencial, durante este projeto foram estes a serem sujeitos a análise. Foram desenvolvidos programas de simulação destes métodos, de forma a serem experimentados vários valores das

constantes de amortecimento α , β e γ , e ser possível escolher o método adequado a cada família e os respectivos valores das constantes.

Verificou-se que, para os grupos *Cluster 1*, *Cluster 2* e *Cluster 3* da família das Massas o método de previsão adequado seria o de Amortecimento Exponencial Linear de Holt e para os grupos *Cluster 4* e *Cluster 5* o método adequado seria o de Amortecimento Exponencial de Holt-Winters.

Apesar de no estudo de tendência se ter concluído que esta família apresenta tendência constante, sendo de esperar que o método que melhor se adaptasse a todos os grupos desta família fosse o de Amortecimento Exponencial Simples, pelas simulações efetuadas e ao desagregar os artigos em níveis inferiores, as séries temporais apresentavam tendência ligeiramente distintas de constante. Comparativamente com o método *naïve*, os erros de previsão apresentam uma diferença de 3 p.p. a 9 p.p. abaixo dos obtidos pelo método *naïve*.

Relativamente à família de Bolachas, o método de previsão selecionado foi o de Amortecimento Exponencial de Holt-Winters, tal como seria esperado. Comparando os resultados obtidos pelas simulações realizadas, e após seleção das melhores combinações, com aqueles adquiridos através do método *naïve*, conclui-se que o método de amortecimento exponencial selecionado apresenta previsões de vendas aceitáveis do ponto de vista de precisão.

Abordando a família dos CPA, o método selecionado foi o de Amortecimento Exponencial de Holt-Winters. Verificou-se inicialmente uma tendência crescente das vendas, bem como uma sazonalidade anual, pelo que neste caso o método de Holt-Winters seria esperado ser o mais adequado. O valor mínimo de erro de previsão foi obtido para valores de constantes otimizados, sendo este de 42,00%. Quando comparado com o método *naïve*, o método de Holt-Winters apresentou de novo um erro de previsão inferior, sendo essa diferença de 4,63 p.p., concluindo-se que as previsões obtidas são boas.

A comparação entre os métodos de amortecimento exponencial selecionados e o método *naïve* mostrou que as previsões obtidas apresentam uma precisão confortável, uma vez que a diferença dos erros entre os dois métodos é de 3 p.p. a 10 p.p..

Sendo o objetivo essencial deste projeto a otimização dos métodos de previsão, através da parametrização com valores ótimos das constantes de amortecimento para posterior implementação das mesmas na ferramenta de planeamento de produção SO99+, a análise final comparativa entre as previsões obtidas através das simulações realizadas neste projeto com as obtidas pelo SO99+ é fundamental. Desta forma, aferiu-se que houve uma diminuição de 8,68 p.p. no erro VWMAPE na família dos CPA, relativamente ao mês de fevereiro de 2020, e de 7,56 p.p. na família das Massas. Quanto à família das Bolachas, a diminuição foi de 2,56 p.p.. A minimização dos valores dos erros VWMAPE permite concluir que os métodos de previsão selecionados conduziram a previsões mais assertivas, relativamente àquelas geradas pela ferramenta de planeamento SO99+.

Pode concluir-se que a reparametrização das constantes de amortecimento dos métodos preditivos, bem como a correta agregação dos artigos conduz à minimização dos erros de previsão e, por fim, a previsões de vendas mais assertivas.

Quanto à implementação de *Machine Learning* no processo de previsão de vendas, foram encontradas bases suficientes para suportar a conclusão de que tal seria uma mais valia a ser explorada no cenário atual da Cerealis. Isto permitiria uma maior otimização das previsões, bem como a automatização deste processo. De notar que esta deverá ser implementada em conjunto com os métodos de amortecimento exponencial, de forma a obterem-se valores mínimos dos erros VWMAPE e valores os mais assertivos possíveis de previsões. Desta forma, seria pertinente alargar os trabalhos realizados neste projeto para o estudo do impacto que teria a implementação de *Machine Learning* nas previsões de vendas.

Todas as propostas de melhoria realizadas neste projeto visam atingir o objetivo de otimizar os métodos de previsão, de forma a obterem-se erros de valor reduzidos ao mínimo possível. Esta otimização das previsões de vendas desencadeia também a otimização do planeamento de produção, levando à melhor gestão de stocks por parte da empresa.

Como tal, de forma a validar os resultados obtidos na realização deste projeto, deverá ser realizada a implementação das conclusões mencionadas na ferramenta *SO99+*, sendo este o principal passo a tomar num futuro próximo. Como proposta futura, deverá ser realizada a aferição dos valores de constantes de amortecimento, bem como os níveis de agregação obtidos, de forma a ser possível obter-se uma validação mais prática dos resultados.

Referências

- Alade, T. (2018). *Tutorial: How to determine the optimal number of clusters for k-means clustering*. Obtido em Junho de 2020, de Cambridge Spark: <https://blog.cambridgespark.com/how-to-determine-the-optimal-number-of-clusters-for-k-means-clustering-14f27070048f>
- Armstrong, J., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, 08, 69-80.
- Ballou, R. H. (1999). *Business Logistics Management*. Prentice-Hall, Inc.
- Bourne, C. (Junho de 2016). *Forecasting with Machine Learning Techniques*. Obtido em Maio de 2020, de CardinalPath: <https://www.cardinalpath.com/blog/forecasting-with-machine-learning-techniques>
- Brockwell, P., & Davis, R. (2016). *Introduction to Time Series and Forecasting* (3rd ed.). Springer International Publishing Switzerland.
- Cerealis. (2018). Obtido em Março de 2020, de HOME | Cerealis: <https://www.cerealis.pt/>
- Demand Forecasting Methods: Using Machine Learning and Predictive Analytics to See the Future of Sales*. (Novembro de 2019). Obtido em Maio de 2020, de Altexsoft: <https://www.altexsoft.com/blog/demand-forecasting-methods-using-machine-learning/>
- developers, s.-l. (s.d.). 2.3. *Clustering*. Obtido em Junho de 2020, de scikit-learn: <https://scikit-learn.org/stable/modules/clustering.html>
- Firth, D. (1993). Bias Reduction of Maximum Likelihood Estimates. *Biometrika*, 27-38.
- Flores, B. E., Olson, D. L., & Dorai, V. K. (1992). Management of Multicriteria Inventory Classification. *Mathl. Comput. Modelling*, 16, 71-82.
- Flowers, A. (1980). A Simulation Study of Smoothing Constant Limits for an Adaptive Forecasting System. *Journal of Operations Management*, 1, 85-94.
- Guedes, A. P. (2017). *Planeamento Integrado & Gestão de Stocks/Materiais*.
- Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20, 5-10.
- How to report Forecast Accuracy to Management*. (s.d.). Obtido em Março de 2020, de arkieva: <https://blog.arkieva.com/how-to-report-forecast-accuracy-to-management/>
- Huneycutt, J. (Maio de 2018). *An Introduction to Clustering Algorithms in Python*. Obtido em Maio de 2020, de Towards Data Science: <https://towardsdatascience.com/an-introduction-to-clustering-algorithms-in-python-123438574097>
- Hyndman, R. J., & Athanasopoulos, G. (2013). *Forecasting: Principles and Practice* (2nd ed.). OTexts. Obtido de <https://otexts.com/fpp2/>

- Hyndman, R. J., Koehler, A. B., Ord, J. K., & Snyder, R. D. (2008). *Forecasting with Exponential Smoothing - The State Space Approach*. Springer.
- Hyndman, R., & Koehler, A. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 41, 679-688.
- Kourentzes, N. (2016). *ABC-XYZ analysis for forecasting*. Obtido em Abril de 2020, de Nikolaos Kourentzes Forecasting research: <https://kourentzes.com/forecasting/2016/10/15/abc-xyz-analysis-for-forecasting/>
- Lizardo, O. (2015). The Effect of Economic and Cultural Globalization on Anti-U.S. Transnational Terrorism 1971– 2000*. *Journal of World-Systems Research*, 149.
- Makridakis, S., Wheelwright, S. C., & Hyndman, R. J. (1998). *Forecasting: Methods and Applications* (3rd ed.). John Wiley & Sons, Inc.
- Portilla, J. (2019). Obtido em Abril de 2020, de Python for Time Series Data Analysis: <https://www.udemy.com/>
- Professionals, C. o. (2013). *CSCMP Supply Chain Management Definitions and Glossary*. Retrieved Março 2020, from https://cscmp.org/CSCMP/Educate/SCM_Definitions_and_Glossary_of_Terms.aspx
- Ramanathan, R. (2006). ABC inventory classification with multiple-criteria using weighted linear optimization. *Computers & Operations Research*, 33, 695-700.
- Rushton, A., Oxley, J., & Croucher, P. (2000). *The Handbook of Logistics and Distribution Management* (2nd ed.). Kogan Page.
- Rutner, S. M., & Jr., C. L. (2000). Logistics Value: Definition, Process and Measurement. *The International Journal of Logistics Management*, 11, 73-82.
- Sayak, P. (2018). *K-Means Clustering in Python with scikit-learn*. Obtido em Junho de 2020, de DataCamp: <https://www.datacamp.com/community/tutorials/k-means-clustering-python>
- Seif, G. (2018). *The 5 Clustering Algorithms Data Scientists Need to Know*. Obtido em Maio de 2020, de Towards Data Science: <https://towardsdatascience.com/the-5-clustering-algorithms-data-scientists-need-to-know-a36d136ef68>
- Singh, S. (2015). *How to Report Forecast Accuracy to Management*. Obtido em Março de 2020, de Arkieva: <https://blog.arkieva.com/how-to-report-forecast-accuracy-to-management/>
- Snapp, S. (2012). *How to best understand forecast bias*. Obtido em Março de 2020, de Brightwork | Demand Planning: https://www.brightworkresearch.com/demandplanning/2012/02/forecastbias/#What_is_Bias
- Swamidass, P. M. (2000). *Encyclopedia of Production and Manufacturing Management* (1st ed.). Springer US.
- Taranenko, L. (2020). *Machine Learning In Demand Forecasting For Retail*. Obtido em Maio de 2020, de Mobidev: <https://mobidev.biz/blog/machine-learning-methods-demand-forecasting-retail>
- ToolsGroup. (2018). *Service Optimizer 99+ User Manual*.
- Wilson, E. (2019). *How Do You Pick The Right Aggregation Level For Forecasting?* Obtido em Abril de 2020, de Institute of Business Forecasting and Planning: <https://demand-planning.com/2019/05/27/aggregation-levels-for-forecasting/>

- Winters, P. R. (1960). Forecasting Sales by Exponentially Weighted Moving Averages. *Management Science*, 6, 324-342.
- Zotteri, G., Kalchschmidt, M., & Caniato, F. (2005). The impact of aggregation level on forecasting performance. *International journal of production economics*, 93-94, 479-491.

ANEXO A: Gráficos do *Elbow Method* para obtenção do valor ótimo de *clusters*

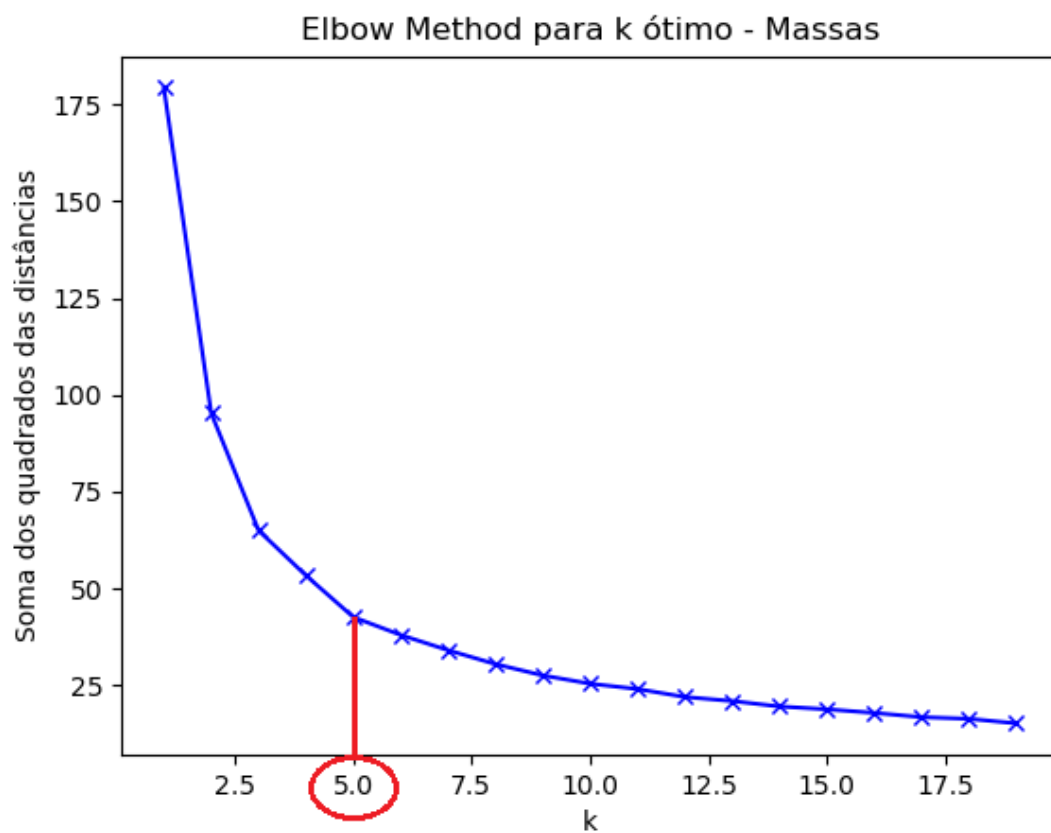


Figura 29 - Gráfico de *Elbow Method* para obtenção de valor ótimo *k* de *clusters*, para a família das Massas.

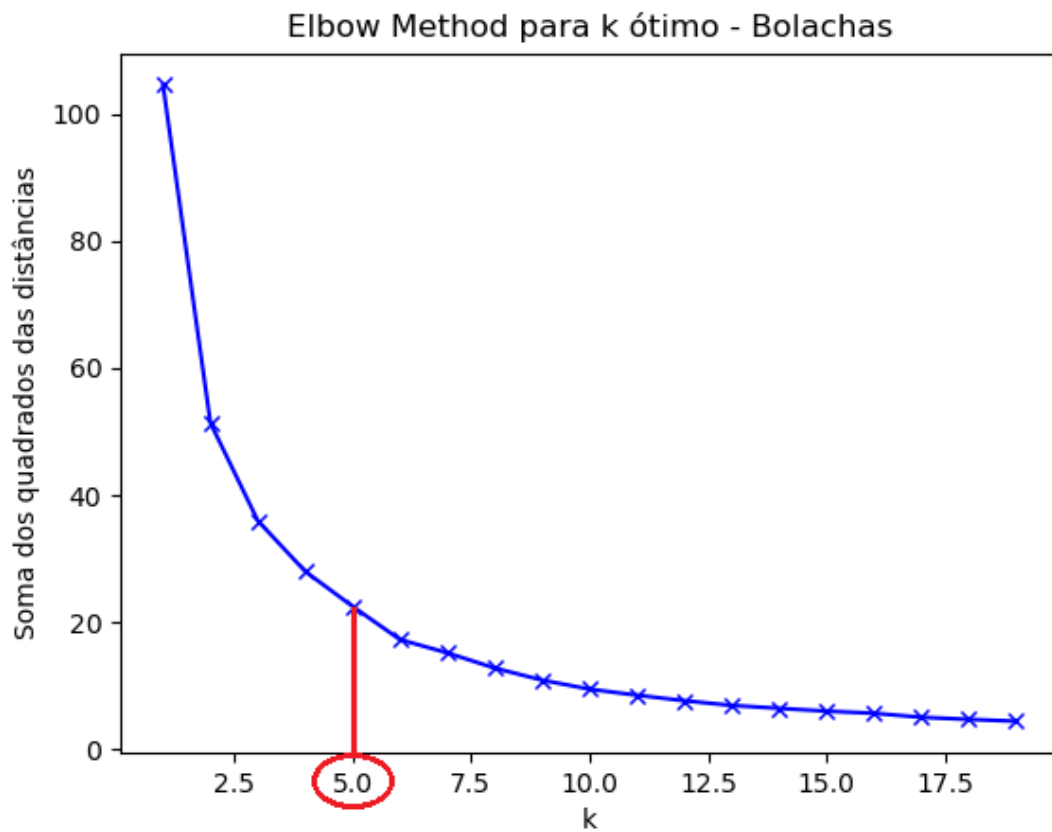


Figura 30 - Gráfico de *Elbow Method* para obtenção de valor ótimo k de *clusters*, para a família das Bolachas.

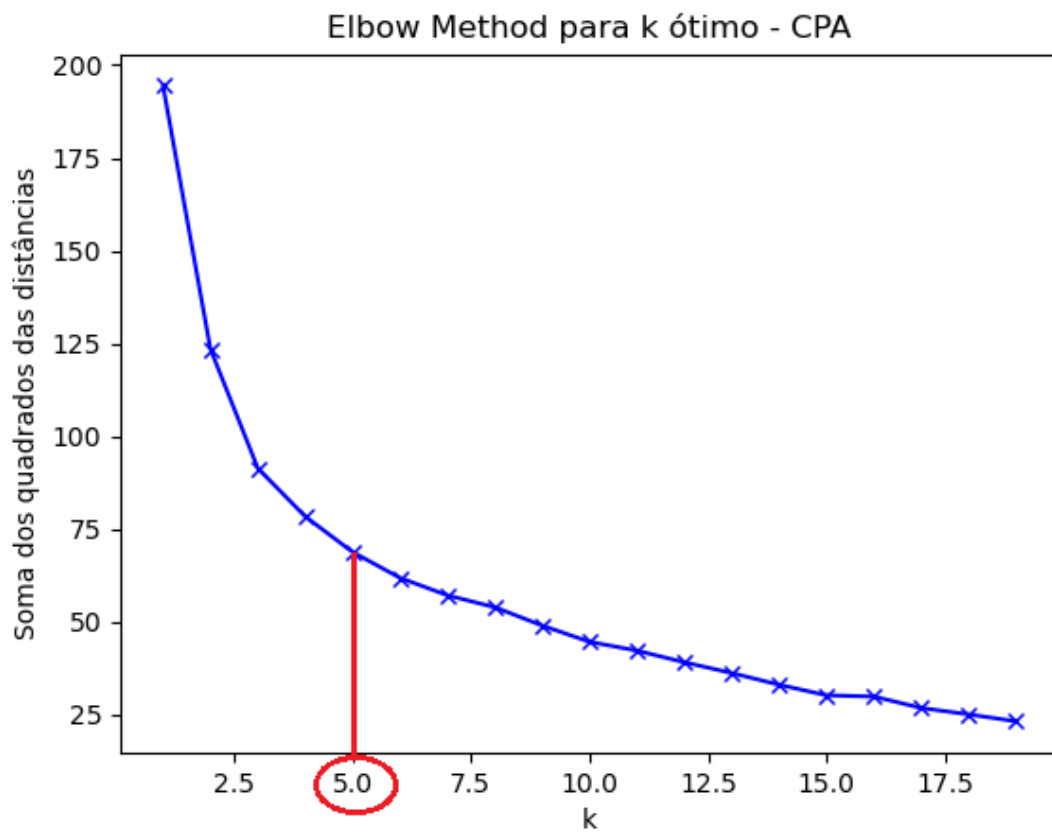


Figura 31 - Gráfico de *Elbow Method* para obtenção de valor ótimo k de *clusters*, para a família dos CPA.

ANEXO B: Constantes de Amortecimento Seleccionadas

Grupo	Nível	Método	Modelo	α	β	γ	MAE (kg)	MAPE (%)	VWMAPE (%)	BIAS (%)
Cluster 1 Massas	Clustering à Família	Simples	Multiplicativo	0,1	-	-	4538,77	45,79	45,11	-1,13
		Holt		0,3	0,6	-	4660,46	47,02	36,43	-2,30
		Holt-Winters		0,1	0,1	0,5	5675,74	52,21	42,12	-12,86
Simples		Aditivo	0,1	-	-	4730,98	180,47	83,08	47,74	
Holt			0,1	0,1	-	4505,07	245,85	79,11	26,60	
Holt-Winters			0,1	0,1	0,1	4831,67	188,24	197,25	114,15	
Cluster 3 Massas		Simples	Aditivo	0,9	-	-	5635,17	34,66	30,54	-0,64
		Holt		0,1	0,3	-	5405,51	33,63	29,29	2,59
		Holt-Winters		0,9	0,1	0,5	6111,64	35,66	30,29	-0,73
Cluster 4 Massas		Simples	Multiplicativo	0,9	-	-	5465,74	74,59	70,28	2,31
		Holt		0,9	0,9	-	5896,36	75,82	68,50	4,93
		Holt-Winters		0,1	0,1	0,4	5034,28	65,53	64,26	-2,63
Cluster 5 Massas		Simples	Multiplicativo	0,1	-	-	6327,76	132,93	57,74	-1,15
		Holt		0,1	0,3	-	7323,81	153,87	58,58	-18,65
		Holt-Winters		0,1	0,1	0,4	6049,72	129,49	50,61	-4,66
Bolachas	Famílias	Simples	Aditivo	0,1	-	-	1055,83	68,67	40,22	4,19
		Holt		0,1	0,1	-	1120,468	76,67	42,68	-7,62
		Holt-Winters		0,1	0,1	0,7	1071,13	66,42	40,08	-1,13
Simples		Multiplicativo	0,1	-	-	1777,10	56,60	42,02	10,19	
Holt			0,1	0,1	-	1887,98	64,65	44,64	-4,15	
Holt-Winters			0,1	0,1	0,1	1776,27	57,34	41,99	9,30	

Figura 32 - Constantes de amortecimento ótimas para os três métodos de previsão em análise.