

UNIVERSIDADE DO VALE DO RIO DOS SINOS — UNISINOS
UNIDADE ACADÊMICA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO E
SISTEMAS
NÍVEL DOUTORADO

JOSÉ EDUARDO DE CARVALHO LIMA

PLANEJAMENTO DA DEMANDA PARA INDÚSTRIA DE MANUFATURA DE BENS A
PARTIR DE FORMALISMOS DE SÉRIES TEMPORAIS

SÃO LEOPOLDO-RS
2021

José Eduardo de Carvalho Lima

PLANEJAMENTO DA DEMANDA PARA INDÚSTRIA DE MANUFATURA DE BENS A
PARTIR DE FORMALISMOS DE SÉRIES TEMPORAIS

Tese apresentada como requisito parcial para a
obtenção do título de Doutor pelo Programa
de Pós-Graduação em Engenharia de
Produção e Sistemas da Universidade do Vale
do Rio dos Sinos — UNISINOS

Orientador:
Prof. Dr. Luiz Alberto Oliveira Rocha

Coorientador:
Prof. Dr. Paulo Renato Alves Firmino

São Leopoldo-RS
2021

L732p

Lima, José Eduardo de Carvalho.

Planejamento da demanda para indústria de manufatura de bens a partir de formalismos de séries temporais / por José Eduardo de Carvalho Lima. – 2021.

126 f. : il. ; 30 cm.

Tese (doutorado) — Universidade do Vale do Rio dos Sinos, Programa de Pós-Graduação em Engenharia de Produção e Sistemas, São Leopoldo, RS, 2021.

Orientador: Dr. Luiz Alberto Oliveira Rocha.

Coorientador: Dr. Paulo Renato Alves Firmino.

1. Planejamento da produção. 2. Previsão de demanda. 3. Séries temporais. 4. Combinação de preditores. 5. Indústria de manufatura. I. Título.

CDU: 658.5

Dedico este trabalho à minha família, pessoas simples e humildes, que desde sempre acreditaram em mim.

AGRADECIMENTOS

A Deus Pai, Filho e Espírito Santo por todas as graças derramadas em minha vida. A Nossa Senhora, pela sua intercessão e proteção de mãe. Obrigado Senhor, pela força e inspiração nos momentos de assumir com responsabilidade, competência e amor a missão a mim confiada.

Aos meus pais, Olavo e Mocinha, pela vida, exemplo de educação, trabalho, orientação e amor. E pelos incentivos nos estudos que me fizeram acreditar que os meus sonhos seriam possíveis de conquistar. Meus irmãos, Alan Roges e Camila Nayane, que com muito amor estão sempre acompanhando meus passos. Meus avós, Ana e Manoel Luiz (*in memorian*) e Francisca e Domingos (*in memorian*), que também são parte do que sou hoje.

A minha esposa, Jocilene (Tamia), por está sempre presente na minha vida, apoiando e incentivando esta jornada acadêmica, oferecendo sempre um porto seguro de compreensão e incentivo. Aos nossos filhos, Pablo Thiago e Hillary Hellen, que com suas alegrias e amor renovam minhas energias.

Agradeço ao meu orientador, professor Dr. Luiz Alberto Oliveira Rocha, por todo conhecimento, paciência, disponibilidade e apoio.

Agradeço ao meu coorientador, professor Dr. Paulo Renato Alves Firmino, inicialmente pela oportunidade de ser membro do grupo de pesquisa Modelagem Estatística, Simulação e Otimização de Risco (MESOR) e por todo conhecimento, paciência, disponibilidade, apoio e exemplo de competência, dedicação e amor à profissão docente.

Ao amigo e colega do grupo de pesquisa MESOR, Jair, que por diversas vezes auxiliou na solução de dúvidas e parceria nos estudos.

Agradeço aos parceiros, colegas e amigos que fiz desde o início da minha vida estudantil na Escola Maria Afonsina Diniz Macedo, passando pela antiga Escola Técnica Federal do Ceará, Universidade Regional do Cariri (URCA), Universidade Federal do Ceará (UFC) e chegando à Universidade do Vale do Rio dos Sinos (UNISINOS). Seu companheirismo e apoio foram diretamente responsáveis pelos meus êxitos.

Agradeço aos meus mestres, que me introduziram nos mais diversos saberes e no desejo de aprender, sempre. A todas as pessoas que apesar de não ter citado os nomes, contribuíram direta ou indiretamente no decorrer deste trabalho.

Obrigado a todos! Entretanto, este não é o fim. “A princesa está em outro castelo”.

RESUMO

Diante de constantes e significativas mudanças nos processos de produção, em que cada vez mais as indústrias de manufatura de bens e serviços são desafiadas, tanto pela complexidade provocada pela variação da demanda, como pela acirrada concorrência imposta pelo mercado, os fabricantes precisam ofertar produtos de alta qualidade, a curto prazo e alta personalização. Neste cenário, o Planejamento e Controle da Produção é uma tarefa importante do sistema produtivo, onde desenvolver um processo confiável de previsão de demanda é o primeiro passo na busca por planos ótimos. Esta pesquisa tem por objetivo propor um sistema de previsão de demanda que auxilie no planejamento da produção, baseando-se em diversos formalismos de séries temporais e do uso de metaheurísticas na otimização dos formalismos. Partindo de uma revisão sistemática de literatura, sumarizada a partir de técnicas de mineração de texto, o trabalho busca compreender como o planejamento da produção e as ferramentas de análise de séries temporais de demanda interagem, possibilitando o seu aprimoramento. Buscou-se, também, identificar os formalismos de séries temporais subjacentes, as técnicas de combinação, as métricas de desempenho utilizadas na previsão da demanda para o planejamento da produção e os tipos de indústria de manufatura aplicados. O estudo possibilitou, entre outras coisas, a elaboração da estrutura conceitual sobre o sistema de previsão a ser proposto. Com o objetivo de avaliar a qualidade do artefato, o sistema de previsão de demanda foi aplicado em uma indústria de manufatura de bens do setor de produtos de higiene, limpeza e saneantes de uso doméstico e profissional. As unidades de análise da presente pesquisa são históricos de demanda, isto é, séries temporais univariadas. A previsão foi realizada considerando formalismos individuais e combinados, bem como a metodologia usada pela empresa. Os resultados das previsões foram avaliados a partir da aplicação das métricas de desempenho e comparando com a metodologia usada pela indústria de manufatura (denominada JUA). Os resultados encontrados demonstraram que o sistema de previsão de demanda formado com o auxílio de combinadores obteve desempenho superior ao modelo de previsão da empresa. Vários dos modelos do sistema de previsão proposto se mostraram adequados para predição de todas as séries temporais da família de produtos analisados. Os resultados da aplicação do artefato sugerem que os formalismos de combinação baseados na média simples (cSA), em redes neurais artificiais (cANN), em vetores de suporte (cSVR) e de mínima variância (cMV), apresentam qualidade superior que a metodologia JUA. Os modelos podem, então, ser utilizados como ferramentas de previsão para indústria de manufatura de produtos de limpeza, a fim de reduzir os problemas de subprodução e superprodução, que ocasionam aumento substancial e contínuo de recursos empatados em estoque, redução da produtividade e aumento da carga de trabalho de natureza de hora extra. Atrelando-se à pesquisa operacional, as previsões foram obtidas por meio de horizonte *one-step-ahead*, o que permite que os gestores otimizem a alocação de recursos, materiais e de custos com foco no planejamento e na programação da produção a curto prazo da empresa objeto de estudo.

Palavras-chave: Planejamento da produção. Previsão de demanda. Séries temporais. Combinação de preditores. Indústria de manufatura.

ABSTRACT

Faced with constant and significant changes in production processes, in which the manufacturing industries of goods and services are increasingly challenged, both by the complexity caused by the variation in demand and by the fierce competition imposed by the market, manufacturers need to offer products of high quality, short term, and high customization. In this scenario, Production Planning and Control is an important task of the production system, where developing a reliable demand forecasting process is the first step in searching for optimal plans. This research proposes a demand forecasting system that helps in production planning based on different formalisms of time series and the use of metaheuristics in the optimization of formalisms. Starting from a systematic literature review, summarized from text mining techniques, the work seeks to understand how production planning and demand time series analysis tools interact, enabling their improvement. It was also sought to identify the underlying time series formalisms, the combination techniques, the performance metrics used in forecasting demand for production planning, and the types of manufacturing industry applied. The study allowed, among other things, the elaboration of the conceptual framework on the forecast system to be proposed. In order to assess the quality of the artifact, the demand forecasting system was applied in a goods manufacturing industry in the sector of hygiene, cleaning, and sanitizing products for domestic and professional use. The unit of analysis of this research is demand histories, that is, univariate time series. The forecast was carried out considering individual and combined formalisms, as well as the methodology used by the company. The results of the forecasts were evaluated based on the application of performance metrics and compared to the methodology used by the manufacturing industry (called JUA). The results showed that the demand forecasting system formed with the help of combiners had a superior performance than the company's forecasting model. Several of the models of the proposed forecast system proved to be adequate for predicting all the time series of the analyzed product family. The results of the application of the artifact suggest that the combination formalisms based on the simple mean (cSA), on artificial neural networks (cANN), on support vectors (cSVR) and on minimum variance (cMV), have superior quality to the JUA methodology. The models can then be used as a forecasting tool for the cleaning product manufacturing industry, in order to reduce the problems of underproduction and overproduction, which cause a substantial and continuous increase in resources tied up in stock, reduced productivity and increased workload of overtime nature. Linking to operational research, forecasts were obtained through a one-step-ahead horizon, which allows managers to optimize the allocation of resources, materials, and costs, focusing on the planning and scheduling of short-term production of the company object of study.

Keywords: Production planning. Demand forecast. Time series. Combination of predictors. Manufacturing Industry.

LISTA DE FIGURAS

Figura 1:	<i>Framework</i> de criação do artefato a partir de modelos individuais, técnicas de combinação e modelos combinados	25
Figura 2:	Arquitetura da abordagem proposta para modelagem e aplicação dos modelos individuais, para o desenvolvimento dos modelos combinados, verificação da qualidade dos modelos a partir das medidas de precisão e da previsão da demanda	31
Figura 3:	Metodologia para a combinação dos modelos individuais de previsão de séries temporais de demanda	33
Figura 4:	Estrutura da Discussão Teórica envolvendo a Revisão Sistemática de Literatura e Revisão Bibliográfica	34
Figura 5:	Constituição da amostra de pesquisa às bases de dados consultadas acerca da literatura internacional envolvendo previsão de demanda, séries temporais e planejamento e controle da produção. Entre parênteses, no cabeçalho de cada ação, encontra-se o número de documentos resultantes	35
Figura 6:	Nuvem de palavras do <i>corpus</i> textual acerca da literatura internacional envolvendo previsão de demanda, séries temporais e planejamento e controle da produção, no período de 2014 a 2019	36
Figura 7:	Dendrograma representativo das repartições em classes e palavras acerca da literatura internacional envolvendo previsão de demanda, séries temporais e planejamento e controle da produção	38
Figura 8:	Análise de Similitude acerca da literatura internacional no período de 2014 a 2019 envolvendo previsão de demanda, séries temporais e planejamento e controle da produção	39
Figura 9:	Série temporal mensal de demanda da família de produtos (DJUF1) sabão em barra, para período de 2015 a 2021	44
Figura 10:	Modelo geral de um neurônio artificial segundo McCulloch e Pitts	51
Figura 11:	Arquitetura de uma rede MLP com duas camadas intermediárias	53
Figura 12:	Arquitetura de Rede Neural adaptada para séries temporais com duas entradas regressas, duas camadas ocultas e uma saída	55
Figura 13:	Representação gráfica da transformação de um problema de regressão em classificação binária. Problema de regressão (a); adição e subtração da constante d_s aos dados de entrada (b); transformação do problema de regressão original em um problema de classificação binária (c); hiperplano que define a função de regressão (d).	58
Figura 14:	Função Perda ϵ -insensível	60
Figura 15:	Arquitetura do modelo de Regressão por Vetores de Suporte (SVR)	62
Figura 16:	Estrutura básica de sistema de inferência <i>fuzzy</i> com <i>fuzzifier</i> e <i>defuzzifier</i>	64
Figura 17:	Estrutura geral de combinação dos formalismos individuais (ANN, ARIMA, ETS, FUZZY, JUA e SVR) por meio das técnicas de combinação (cANN, cMV, cSA, cSM e cSVR) de uma determinada série temporal	71
Figura 18:	Fluxograma do algoritmo da metaheurística <i>Simulated Annealing</i> (SA)	82

Figura 19: Séries temporais de demanda das famílias de produtos DJUF1, DJUF2, DJUF3 e DJUF6	83
Figura 20: Séries temporais de demanda das famílias de produtos DJUF7, DJUF11, DJUF12 e DJUF13	84
Figura 21: Média agregada normalizada de performance dos modelos individuais e combinados para as séries de demanda de produtos DJUF1, DJUF2, DJUF3 e DJUF6	89
Figura 22: Formalismos combinados para as séries temporais de demanda das famílias de produtos DJUF1, DJUF2, DJUF3 e DJUF6 (Fase de treinamento e teste)	90
Figura 23: Média agregada normalizada de performance dos modelos individuais e combinados para as séries de demanda de produtos DJUF7, DJUF11, DJUF12 e DJUF13	94
Figura 24: Formalismos combinados para as séries temporais de demanda das famílias de produtos DJUF7, DJUF11, DJUF12 e DJUF13 (Fase de treinamento e teste)	94

LISTA DE QUADROS

Quadro 1:	Etapas do método de trabalho no processo de criação e avaliação dos modelos (artefato) de previsão	25
Quadro 2:	Teste de associação entre palavras e classes, onde a Classe 1: Contextos e Aplicações; Classe 2: Fundamentação Estatística; Classe 3: Modelagem e Previsão. As Colunas 1-3 representam, nesta ordem, o termo, a estatística χ^2 e o % citações	39
Quadro 3:	A família de modelos ETS, com diferentes tipos de tendência e sazonalidade, de acordo com Hyndman (HYNDMAN et al., 2008)	50
Quadro 4:	Analogia entre o sistema físico e o problema de otimização	81

LISTA DE TABELAS

Tabela 1:	Resumo das características das séries temporais de demanda das famílias de produtos objetos de estudo	30
Tabela 2:	Parâmetros de ajuste dos formalismos de previsão individuais e combinados para cada série temporal considerada (DJUF1, DJUF2, DJUF3, DJUF6, DJUF7, DJUF11, DJUF12, DJUF13)	85
Tabela 3:	Descrição dos modelos individuais e combinados para as séries temporais de demanda das famílias de produtos DJUF1, DJUF2, DJUF3 e DJUF6	87
Tabela 4:	Desempenho dos modelos de previsão (ANN, ARIMA, ETS, FUZZY, SVR, JUA, cSA, cSM, cANN, cMV, cSVR) ao prever as séries temporais DJUF1, DJUF2, DJUF3 e DJUF6 (fase de teste). Os melhores valores encontram-se em negrito	88
Tabela 5:	Descrição dos modelos individuais e combinados para as séries temporais de demanda das famílias de produtos DJUF7, DJUF11, DJUF12 e DJUF13	91
Tabela 6:	Desempenho dos modelos de previsão (ANN, ARIMA, ETS, FUZZY, SVR, JUA, cSA, cSM, cANN, cMV, cSVR) ao prever as séries temporais DJUF7, DJUF11, DJUF12 e DJUF13 (fase de teste). Os melhores valores encontram-se em negrito	93
Tabela 7:	Desempenho normalizado médio agregado dos modelos de previsão (ANN, ARIMA, ETS, FUZZY, JUA, SVR, cSA, cSM, cMV, cANN, cSVR) ao prever as séries temporais DJUF1, DJUF2, DJUF3, DJUF6, DJUF7, DJUF11, DJUF12 e DJUF13 (fase de teste). A classificação de cada modelo está entre parênteses	96
Tabela 8:	Desempenho normalizado médio agregado dos modelos de previsão (ANN, ARIMA, ETS, FUZZY, JUA, SVR, cSA, cSM, cMV, cANN, cSVR) ao prever as séries temporais DJUF1, DJUF2, DJUF3, DJUF6, DJUF7, DJUF11, DJUF12 e DJUF13 (fase de treinamento). A classificação de cada modelo está entre parênteses	97
Tabela 9:	Tempo de treinamento dos modelos (ANN, ARIMA, ETS, FUZZY, SVR, cANN e cSVR) para as séries DJUF1, DJUF2, DJUF3, DJUF6, DJUF7, DJUF11, DJUF12 e DJUF13 (em segundos). Os melhores valores encontram-se em negrito	99

LISTA DE SIGLAS

AF	<i>Activation Function</i> ou Função de Ativação
AM	Abbasov-Mamedova
ANN	<i>Artificial Neural Networks</i> ou Redes Neurais Artificiais
AIC	<i>Akaike Information Criteria</i> ou Critério de Informação <i>Akaike</i>
APM	<i>Aggregate Performance Metrics</i> ou Métrica de Desempenho Agregado
AR	Autorregressivo
ARIMA	Autorregressivo Integrado de Médias Móveis
ARMA	Autorregressivo de Média Móvel
ARV	<i>Average Relative Variance</i> ou Variância Relativa Média
BIC	<i>Bayesian Information Criteria</i> ou Critério de Informação <i>Bayesiano</i>
cANN	Combinador de Redes Neurais Artificiais
CHD	Classificação Hierárquica Descendente
CIS	Critério de Informação de <i>Schwarz</i>
cMV	Combinador de Mínima Variância
cSA	Combinador de Média Simples
cSM	Combinador de Mediana Simples
cSVR	Combinador de Regressão por Vetores de Suporte
DSR	<i>Design Science Research</i>
ERM	<i>Empirical Risk Minimization</i>
ERP	<i>Enterprise Resource Planning</i>
ETS	<i>Error, Tendency, and Seasonality</i> ou Suavização Exponencial
FI	Formalismos Individuais
GA	<i>Genetic Algorithms</i> ou Algoritmos Genéticos
IC	Índice de Concordância
ID	Índice de Desacordo
LA	<i>Learning Algorithm</i> ou Algoritmo de Aprendizagem
MA	Médias Móveis
MAPE	<i>Mean Absolute Percentage Error</i> ou Erro Médio Absoluto Percentual
MLP	<i>Rede Multilayer Perceptron</i>
MSE	<i>Mean Squared Error</i> ou Erro Quadrático Médio
PCP	Planejamento e Controle da Produção
POCID	<i>Prediction on Change of Direction</i> ou Predição de Alteração na Direção
PP	Planejamento da Produção
PPC	<i>Planning and Production Control</i>
PSO	<i>Particle Swarms</i> ou Enxames de Partículas

RBF	<i>Radial Basis Function</i> ou Redes de Base Radial
RSL	Revisão Sistemática da Literatura
RPROP	Retropropagação Resiliente
SAn	<i>Simulated annealing</i> ou Recozimento Simulado
SARIMA	Autorregressivo Integrado de Médias Móveis Sazonal
SIC	<i>Schwarz's Information Criterion</i>
ST	Segmento de Texto
STF	Série Temporal <i>Fuzzy</i>
SVM	<i>Support Vector Machines</i> ou Máquina de Vetores de Suporte
SVR	<i>Support Vector Regression</i> ou Regressão por Vetores de Suporte
Theil'U	Coefficiente U de Theil ou Estatística U de Theil
THF	<i>Hyperbolic Tangent Function</i> ou Função Tangente Hiperbólica
WPOCID	<i>Wrong Prediction on Change of Direction</i> ou Predição Errada de Alteração na Direção

LISTA DE SÍMBOLOS

\mathcal{A}	Axônio
R^2	Coefficiente de determinação do ajuste linear
WR^2	Coefficiente de indeterminação do ajuste linear ($WR^2 = 1 - R^2$)
$\{\gamma_i^{ar}\}_{i \in I_u}$	Coefficientes que ponderam as defasagens temporais referentes à série diferenciada
$\{\gamma_j^{ma}\}_{j \in I_a}$	Coefficientes que ponderam as defasagens temporais referentes aos resíduos do modelo
E	Componente de Erro aleatório do modelo
S	Componente de Sazonalidade do modelo
T	Componente de Tendência do modelo
u_i	Comprimento do i^o intervalo do modelo Fuzzy
$\{\mathcal{P}_h\}_{h=1}^H$	Conjunto de perceptrons da ANN
A_i	Conjunto de variáveis linguísticas
$D_1 e D_2$	Constantes de ajuste do universo de discurso
\mathcal{C}	Constante de regularização ou penalização
C	Constante positiva de conversão de valores quantitativos em valores <i>fuzzy</i>
\mathcal{S}	Corpo celular
$\sigma_{i,j}$	Covariância entre os erros dos modelos
\mathcal{I}	Dendritos
x_m	Entradas do neurônio; Entradas regressas de uma série temporal
u	Elemento genérico do universo de discurso U
ϵ	Erro aleatório (Desvio)
\hat{u}_t^{SA}	Estimativa do modelo combinado
μA_i	Função de associação (pertinência)
$\mathcal{A}(\cdot)$	Função de ativação
$\varphi(\cdot)$	Função de ativação do k -ésimo neurônio
$\mathcal{S}(\cdot)$	Função de ativação intermediária
$e^{-\Delta/kT}$	Função de <i>Boltzmann</i>
$L(\cdot)$	Função de perda ϵ -insensível
$f(\mathbf{x})$	Função de regressão
$K(x_i, x)$	Função kernel
$\mu A_i(u_i)$	Grau de associação
\div	Indica o separador
t	Índice cronológico
γ_0	Intercepto do modelo ARIMA
M^{-1}	Inversa da matriz de covariância

θ_k	Limiar do neurônio k (bias)
V_{max}	Maior variação
$\mathbf{K}(t)$	Matriz de critérios
$O^w(t)$	Matriz de operação
$\mathbf{R}(t)$	Matriz de relações
\bar{u}_t	Média dos valores observados da série temporal
V_{min}	Menor variação
$n.Metric_i$	Métrica i , normalizada
$AR(p)$	Modelo Autoregressivo
$ARMA(p, q)$	Modelo Autoregressivo de Médias Móveis
$ARIMA(p, d, q)$	Modelo Autoregressivo Integrado de Médias Móveis
$MA(q)$	Modelo de Médias Móveis
α_i e α_i^*	Multiplicadores de Lagrange
\mathcal{N}	Neurônio artificial
n	Número de conjuntos <i>fuzzy</i>
NI	Número de entradas autorregressivas (autorregressivas e autorregressivas sazonais PAR +PARS)
K	Número de formalismos individuais
m	Número de métricas
N	Número de nós na camada intermediária
n_o	Número de observações da série temporal regressa
k_{BIC}	Número de parâmetros BIC
NO	Número de saídas da ANN
d_s	Número positivo
\circ	Operador de composição de max-min
Δ	Operador de diferença
\otimes	Operador $\min(\cap)$
"+"	Operação de união
d	Ordem de diferenciação
p	Ordem do modelo autorregressivo (ARIMA); Número de sinais de entrada do neurônio
q	Ordem do modelo de médias móveis
ϕ	Parâmetro de amortecimento do modelo ETS
w	Parâmetro de tempo
γ_i^{ar}	Parâmetros do modelo $AR(p)$
γ_j^{ma}	Parâmetros do modelo $MA(q)$
\mathcal{P}	<i>Perceptron</i>

w_{km}	Peso associado com o m -ésimo sinal de entrada
w_i	Pesos do modelo ANN
u_i^m	Ponto médio do conjunto fuzzy
T_h	Previsão da tendência
$\hat{F}(t)$	Previsão nebulosa das variações
NHL	Quantidade de nodos na camada oculta
R	Relação difusa
y_i	Saídas
$F(t)$	Série temporal <i>fuzzy</i>
U_t	Série temporal univariada
$Y(t)$	Subconjunto dos números reais
N	Tamanho da série observada
(T')	Temperatura atual
b_s	Termo de crescimento do modelo ETS
l	Termo de nível do modelo ETS
U	Universo de discurso
$\mu_{A_i}(V_t)$	Valor da função de associação da variação da previsão
ML	Valor da função de máxima verossimilhança
\hat{u}_t	Valor estimado/previsto
u_t	Valor observado da série temporal
$\{\hat{u}_{ti}\}_{i=1}^k$	Valor previsto por combinadores
$\hat{u}_{t,i}$	Valor previsto pelo o i -ésimo modelo para u_t
$PARS$	Valores autoregressivos defasados sazonais
min_i e max_i	Valores mínimo e máximo observados para $Metric_i$
V_t	Variação
$V(t)$	Variação da previsão defuzzificada
ξ_i, ξ_i^*	Variáveis de folga
$x_i \in x$	Vetores de entradas
b	Viés (ou intercepto) do modelo SVR

SUMÁRIO

1 INTRODUÇÃO	17
1.1 Objeto e Problema de Pesquisa	19
1.2 Objetivos	20
1.2.1 Objetivo Geral	20
1.2.2 Objetivos Específicos	20
1.3 Justificativa	20
1.4 Delimitações	21
1.5 Estrutura do Trabalho	22
2 PROCEDIMENTOS METODOLÓGICOS	23
2.1 Método de pesquisa	23
2.2 Método de Trabalho	24
2.2.1 Revisão Bibliográfica	25
2.2.2 Revisão Sistemática da Literatura	26
2.2.2.1 Bases e descritores	26
2.2.2.2 Critérios de inclusão e exclusão	27
2.2.2.3 Análise léxica e de conteúdo	27
2.2.3 Definição do Caso de Estudo	28
2.2.4 Coleta dos Dados	29
2.2.5 Construção dos modelos e análise dos dados	29
3 DISCUSSÃO TEÓRICA	34
3.1 A interação entre PCP e análise de séries temporais de demanda	34
3.2 Planejamento e Controle da Produção	40
3.3 Previsão de Demanda	42
3.4 Formalismos de Séries Temporais	43
3.4.1 Formalismos Individuais	45
3.4.1.1 Modelos Box & Jenkins	45
3.4.1.2 Modelos de Suavização Exponencial	47
3.4.1.3 Redes Neurais Artificiais	50
3.4.1.3.1 Perceptron de Múltiplas Camadas	53
3.4.1.4 Máquinas de Vetores de Suporte	57
3.4.1.4.1 Regressão por Vetores de Suporte	57
3.4.1.5 Lógica <i>Fuzzy</i>	63
3.4.1.5.1 Série Temporal <i>Fuzzy</i>	64
3.4.1.5.2 Definições Relacionadas a Séries Temporais <i>Fuzzy</i>	66
3.4.1.5.3 Modelo Abbasov-Mamedova	67
3.4.1.6 Metodologia de Previsão JUA	69
3.4.2 Formalismos de Modelagem Combinada	71
3.4.2.1 Combinação via Média Simples	72
3.4.2.2 Combinação via Mediana Simples	72
3.4.2.3 Combinação via Mínima Variância	73

3.4.2.4	Combinação via Redes Neurais Artificiais	73
3.4.2.5	Combinação via Regressão por Vetores de Suporte	74
3.5	Medidas de desempenho de previsão	74
3.5.1	Erro Quadrático Médio	75
3.5.2	Erro Médio Absoluto Percentual	76
3.5.3	O Coeficiente U de Theil	76
3.5.4	Variância Relativa Média	77
3.5.5	Predição de Alteração na Direção	77
3.5.6	Índice de Desacordo	78
3.5.7	Coefficientes do Modelo de Regressão Linear	78
3.5.8	Métrica de Desempenho Agregado	79
3.5.9	Critério de Informação Bayesiano	80
3.5.10	<i>Simulated de Annealing</i>	80
4	APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS	83
4.1	Séries temporais de demanda das famílias de produtos	83
4.2	Parâmetros iniciais das modelagens	84
4.3	Modelagem, Análise e Comparação dos Métodos	85
5	DISCUSSÃO DOS RESULTADOS	95
6	CONSIDERAÇÕES FINAIS	100
	REFERÊNCIAS	103

1 INTRODUÇÃO

O setor de manufatura de bens e serviços vem sofrendo mudanças significativas nas últimas décadas em termos de escala, complexidade e tecnologia, adicionadas aos problemas advindos da pandemia da COVID-19. Para se manterem competitivos no mercado, cada vez mais acirrado e incerto, associado à variação da demanda, os fabricantes precisam produzir produtos de alta qualidade, a curto prazo e alta personalização (ZHANG et al., 2012; LINGITZ et al., 2013). Eles buscam, ainda, manter uma flexibilidade suficiente para atender as demandas dos clientes, que mudam rapidamente, maximizando o lucro e minimizando ou mantendo os custos de produção em um nível razoável (GYULAI; KÁDÁR; MONOSOTORI, 2015; JEON; KIM, 2016; TRATAR; MOJŠKERC; TOMAN, 2016). É neste contexto que as indústrias de manufatura são desafiadas a planejar as atividades envolvidas em seus processos de produção, a fim de melhorar o fluxo e alocação de informações, materiais e pessoas, e atender apropriadamente a demanda pelos seus produtos.

Dentre as diversas atividades de uma organização empresarial, planejar a produção é uma tarefa importante e desafiadora na condução para o sucesso. Em termos simples, em um planejamento da produção (PP) eficaz, deve-se decidir o quanto, como e em quanto tempo produzir e quais recursos necessários serão utilizados nos processos produtivos. Informações como essas podem ser obtidas a partir da previsão da demanda, a ser alinhada com a capacidade de produção (PRAKASH; RAMYA; YUGANDHAR, 2016; MOR et al., 2019). Desse modo, o planejamento da demanda, que é baseado nas previsões das demandas futuras, tem uma conexão muito forte com o Planejamento e Controle da Produção (PCP) (BÓNA; LÉNÁRT, 2014; SCHOLZ-REITER; KÜCK; LAPPE, 2014).

De fato, as previsões de demanda geram dados de entrada para outras áreas de planejamento, que são usadas no planejamento desde processos estratégicos a operacionais (SCHOLZ-REITER; KÜCK; LAPPE, 2014; TADIĆ; ZEČEVIĆ, 2012; MRÓWCZYŃSKA et al., 2012; MARKOVIĆ et al., 2013). Por exemplo, o planejamento operacional da produção elabora o cronograma de produção dos produtos, o planejamento de recursos na manufatura é responsável pela quantidade e composição dos estoques de matérias-primas e máquinas, o planejamento financeiro e orçamentário fornecem a base financeira estável a operação futura, e assim por diante (SCHOLZ-REITER; KÜCK; LAPPE, 2014). Em todas essas áreas, o planejamento baseado na demanda é imprescindível e, se o resultado for impreciso, a operação da empresa será ineficiente e instável. Aliyeva (2017), reforça que o planejamento de produção bem-sucedido depende da modelagem precisa de aspectos relacionados ao problema, incluído a incerteza da demanda, e que os desenvolvimentos recentes nas áreas de previsão, planeja-

mento de produção e modelagem fornecem elementos para soluções efetivas e aproximadas para este difícil problema.

Nos últimos anos, vários métodos, seguindo diferentes teorias e abordagens, têm sido propostos para analisar a complexidade da produção e as incertezas que afetam os processos de produção (EFTHYMIOU et al., 2016). Alguns autores, Claisse, Chu e Jemai (2016), mencionam as dificuldades para planejar a produção em que a demanda possui características dinâmicas e incertas. Para Aliyeva (2017), a previsão da demanda quase sempre é executada em face da incerteza, gerando consequências como a elevação dos níveis de estoque e aumento dos custos de produção. Já Amirkolaii et al. (2017), apontam que isso se deve ao fato de que a demanda continua flutuando em suas classes que resultam em incertezas. Este cenário tem levado engenheiros, gestores de produção e gerentes de negócios a recorrerem a ferramentas capazes de modelar, analisar e projetar matematicamente as demandas. Isto facilita a incorporação de políticas robustas à presença de perturbações, bem como a mitigação dos impactos negativos da turbulência no ambiente de produção (DUFFIE; CHEHADE; ATHAVALE, 2014).

Avanços recentes apresentaram benefícios relacionados à previsão de demanda, incluindo maior precisão, melhor controle de estoque, redução do tempo de reação e da incerteza de modelagem, entre outros (ROSTAMI-TABAR et al., 2013; KOURENTZES; ROSTAMI-TABAR; BARROW, 2017). Porém, pesquisas realizadas durante o presente trabalho revelam desequilíbrio da literatura entre as abordagens sobre o tema, o que eleva as dificuldades de melhores resultados da área de PCP. De forma desagregada, há documentos que apresentam campos conceituais diferentes, como contextos e aplicações em que priorizam estudos e conceitos sobre PCP, PP, Cadeia de Suprimentos e Indústria (AIT-ALLA et al., 2014; HOLLMANN; SCAVARDA; THOMÉ, 2015; LAAN et al., 2016; KIM; DEKKER; HEIJ, 2017; LEE; LIANG, 2018; GOODWIN; GÖNÜL; ÖNKAL, 2019); pesquisas que agrupam conceitos sobre distribuição, parâmetros, variância, erro, mensuração e estimador, voltados à fundamentação estatística (MATSUMOTO; KOMATSU, 2015; ZHU et al., 2017; HASNI et al., 2018; PRAK; TEUNTER, 2019); e estudos que se dedicam à inferência dos fenômenos de demanda, a partir de formalismos (métodos) específicos. Em outros termos, tratam de conceitos relacionados à modelagem e previsão de demanda, envolvendo eficiência e acurácia de modelos lineares e não lineares, combinações e incorporação de métodos (SCHOLZ-REITER; KÜCK; LAPPE, 2014; SCARPEL, 2015; FENG; MEARS; SCHULTE, 2016; CHANG; YU; JIN, 2016; TIRKES; GURAY; CELEBI, 2017; HAMICHE et al., 2018; MOR et al., 2019). Isto pode refletir um distanciamento dos formalismos dedicados ao estudo da demanda em relação às demais áreas da organização. Apesar da utilização de diferentes formalismos de previsão, observa-se

pouca exploração de métodos individuais robustos, menos ainda da combinação de modelos, e do uso de diferentes técnicas de incorporação na previsão da demanda a partir de séries temporais e também da exploração na indústria de manufatura de produtos de limpeza e saneantes.

Diante disso, as empresas reconhecem a necessidade de um sistema de previsão de demanda que busque minimizar as medidas de erro (OHTA; HIRAMOTO; KITAMURA, 2014; MELO MENEZES et al., 2015). A qualidade dos métodos tradicionais de previsão é muito variável no ambiente manufatureiro (BÓNA; LÉNÁRT, 2014). Contudo, uma das formas para aprimoramento é, dentre os métodos de previsão de demanda existentes, selecionar e utilizar aquele adequado às particularidades do produto e, até mesmo, da empresa e do segmento de mercado em que ela está inserida (CECATTO; BELFIORE, 2015). Esta tese dedica-se a este problema de maneira geral.

1.1 Objeto e Problema de Pesquisa

Diversos pesquisadores desenvolveram inúmeras abordagens a fim de contribuir com o sistema de previsão, melhorar a predição da demanda e, por consequência, o planejamento do sistema produtivo e o uso deste. Desde o trabalho pioneiro de Reid (1968) e o seminal de Bates e Granger (1969) sobre a combinação de modelos de previsão, essa abordagem vem ganhando força, na busca por melhores resultados (BATES; GRANGER, 1969; CLEMEN, 1989; TIMMERMANN, 2006; WANG et al., 2018; WANG; LI; LI, 2018; OLIVEIRA; SILVA; MATTOS NETO, 2021).

Autores como Rosienkiewicz, Chlebus e Detyna (2017), propuseram um modelo híbrido dividido em dois estágios que combina análise de regressão e rede neural artificial (ANN). Guo e Shen (2016) combinaram métodos lineares e não lineares, baseados em ANN. Foram também propostas combinações metodológicas e temporais de métodos únicos e múltiplos, considerando pesos iguais nas combinações (PETROPOULOS; KOURENTZES, 2015). Por fim, Prakash, Ramya e Yugandhar (2016), combinaram modelos individuais usando um algoritmo genético modificado para otimizar os pesos dos preditores individuais. Estas pesquisas, apesar de trabalharem modelos e técnicas diferentes, e de terem em comum um aumento da qualidade da previsão, com reflexos diretos sobre o planejamento das atividades de produção, apresentam pontos de melhorias nos modelos desenvolvidos.

A partir disso, a literatura sugere que a qualidade das previsões pode melhorar sensivelmente, combinando apropriadamente métodos diferentes (MATSUMOTO et al., 2016; WANG; PETROPOULOS, 2016). Para tanto, existem diversas técnicas de previsão disponíveis, desde as simples às mais sofisticadas. O grande desafio passa a ser a identificação da melhor estra-

tégia de modelagem, de maneira a promover preditores estatisticamente acurados e eficientes. Assim, o problema que guia esta pesquisa é: **Como avaliar e melhorar preditores de demanda para auxílio no planejamento da produção?** Destaca-se, neste sentido, como objeto de estudo, modelos de previsão de demanda.

1.2 Objetivos

1.2.1 Objetivo Geral

Propor sistema de previsão de demanda que auxilie no planejamento da produção, baseando-se em diversos formalismos de séries temporais e suas combinações.

1.2.2 Objetivos Específicos

- a) Analisar a relação entre os formalismos de previsão de demanda e do planejamento e controle da produção;
- b) Desenvolver modelos de previsão a partir da combinação de preditores de demanda;
- c) Aplicar o(s) modelo(s) desenvolvido(s) em pelo menos um caso de estudo;
- d) Avaliar o desempenho do artefato proposto junto ao caso de estudo; e
- e) Indicar, para o caso de estudo, um modelo ótimo de previsão de demanda para produção e suporte à gestão e programação da produção.

1.3 Justificativa

Conforme constatado na literatura de previsão de demanda para PCP, existem discussões acerca dos métodos e técnicas de previsão disponíveis a serem empregadas quando se trata da complexidade da produção, das incertezas que afetam a demanda e dos processos de produção (MULA et al., 2006; EFTHYMIOU et al., 2016). Isso tem levado pesquisadores a proporem novos métodos e aproximarem campos conceituais. Assim, uma das principais justificativas para esta tese é a importância da previsão de demanda para o planejamento da produção, com especial atenção na sua necessidade de aprimoramento.

Autores como Scholz-Reiter, Kück e Lappe (2014); Mrówczyńska et al. (2012); Marković et al. (2013), destacam que as etapas subsequentes do planejamento da produção baseiam-se em previsões de demanda, sendo capazes de influenciar fortemente no sucesso ou fracasso das empresas, já que direcionam decisões de longo, médio ou curto prazo (REID; SANDERS, 2015; STEVENSON; SUM, 2015).

Após a leitura de vários documentos, como por exemplo, Petropoulos e Kourentzes (2015); Matsumoto et al. (2016); Wang e Petropoulos (2016); Prakash, Ramya e Yugandhar (2016); Guo e Shen (2016); Rosienkiewicz, Chlebus e Detyna (2017); Sergio (2017); Kourentzes, Rostami-Tabar e Barrow (2017); Wang et al. (2018); Wang, Li e Li (2018); Kozłowski et al. (2018); Dombi, Jónás e Tóth (2018) foi possível identificar que a maioria das indústrias de manufatura não utilizam um método mais robusto para realizar previsões.

Um motivo apontado por Armstrong (2001) é que, na maioria das vezes, os tomadores de decisão optam por usar métodos de previsão de demanda sobre os quais possuem familiaridade, mesmo que os resultados sejam precários. Além disso, a seleção de um método por conveniência pode levar a erros graves em situações que envolvem grandes mudanças. Assim, nota-se a importância na seleção dos métodos de previsão de demanda existentes relacionados com as necessidades, particularidades e realidades dos produtos, da empresa e do segmento de mercado, visando minimizar os erros das previsões de demanda quando se trata de planejar a produção. Tais argumentos justificam a relevância do presente estudo.

1.4 Delimitações

A presente pesquisa busca propor um sistema de previsão de demanda para o planejamento da manufatura de bens a partir da combinação de preditores de demanda. Quanto às suas limitações, algumas dizem respeito ao escopo teórico do trabalho e outras estão relacionadas ao caso de estudo apresentado. Especificamente, destaca-se que:

- a) os resultados do trabalho se referem especificamente à indústria de manufatura;
- b) as bases de dados utilizadas na formulação dos modelos são séries temporais univariadas voltadas às demandas por produtos sob estudo, fornecidas pela empresa caso de estudo;
- c) cofatores externos, tais como preço de venda, promoções e outros, não foram abordados nesta pesquisa;
- d) os métodos utilizados no trabalho compreendem formalismos aplicados a séries temporais, como o modelo linear Box Jenkins (ARIMA) (BOX; JENKINS, 1976; BOX, 2013; BOX et al., 2015), a família de modelos de Suavização Exponencial (ETS) (BROWN, 1959; WINTERS, 1960), de aprendizado de máquinas Redes Neurais Artificiais (ANNs) (SEGALL, 1995; HYNDMAN; ATHANASOPOULOS, 2018), Regressão por Vetores de Suporte (SVR) (SMOLA; SCHÖLKOPF, 2004), o modelo de séries temporais *fuzzy* de Abbasov e Mamedova (ABBASOV; MAMEDOVA, 2003) e a metodologia JUA usada pela empresa objeto de estudo. Não são abordadas, de forma detalhada, outras técnicas de previsão de demanda;

- e) os modelos desenvolvidos foram avaliados a partir de um caso de estudo e não por meio de simulação de dados ou de vários casos de estudo;
- f) a aplicação prática acima citada não inclui a mensuração do impacto financeiro, nem será realizada uma análise financeira do investimento necessário para a implantação dos formalismos de séries temporais na empresa, caso de estudo;
- g) no caso de estudo apresentado, foram analisadas as famílias de demandas de produtos (bens), que apresentam maior relevância para a empresa (em conformidade com os critérios estabelecidos pela mesma);
- h) a análise dos resultados da previsão permitiu a avaliação de alternativas relacionadas à utilização ou não da política de previsão.

Assim, a entrega desta pesquisa passa a ser de modelos específicos em conformidade com as características da demanda dos bens sob estudo. Entende-se ser relevante ressaltar o pressuposto considerado na elaboração desta proposta, que é possível desenvolver o artefato, ainda que o fenômeno não seja passível de generalização para as empresas de manufatura. Nesse sentido, o resultado apresentado neste trabalho deve ser visto como uma contribuição incremental para a literatura do planejamento e controle da produção a partir da previsão da demanda.

1.5 Estrutura do Trabalho

O presente trabalho de pesquisa está organizado em cinco capítulos. Os próximos capítulos consideram a seguinte estrutura.

Capítulo 2 – Percurso metodológico: são explicados os procedimentos metodológicos utilizados na pesquisa, a fim de caracterizá-la e de esclarecer as etapas realizadas.

Capítulo 3 – Discussão teórica: traz os arcabouços teóricos que constituem a base para a construção dos modelos de previsão. Nesta parte, abordam-se os conceitos relacionados aos formalismos de séries temporais, planejamento e controle da produção, técnicas de combinação e as medidas de desempenho utilizadas.

Capítulo 4 – Encontra-se a análise de resultados cujo objetivo é direcionar o leitor à conclusão com base nas observações dos dados.

Finalmente, o Capítulo 5, apresenta as considerações finais a fim de apontar os conhecimentos cientificamente comprovados pela pesquisa.

2 PROCEDIMENTOS METODOLÓGICOS

Para garantir o rigor científico, bem como a validade da pesquisa, esta seção descreve o método de pesquisa, abordando suas características, bem como o método de trabalho, que indica os estágios e os procedimentos realizados.

2.1 Método de pesquisa

O problema identificado na presente pesquisa está na complexidade, na seleção do modelo, na incerteza que envolve prever a demanda e na sua importância para programação da produção. Desenvolver um método de previsão confiável é fundamental para facilitar o planejamento e o controle efetivo da produção. Por conseguinte, o trabalho buscou desenvolver um método com esse objetivo. O modelo (artefato) é útil para a previsão da demanda, sendo o primeiro passo e o mais importante para a otimização do processo geral de planejamento da produção.

O método científico *Design Science Research* (DSR), cujo paradigma epistemológico é a ciência do *design*, foi selecionado por ser adequado quando o objetivo final da pesquisa é a proposta de solução para problemas reais (DRESCH; LACERDA; JÚNIOR, 2015). Historicamente, as bases lógicas de investigação são dedutivas, tentando descrever como a realidade deve ser, isto é, explicar o conteúdo das premissas; ou indutivas, procurando explicar como a realidade é, isto é, ampliar o alcance dos conhecimentos (PRODANOV; FREITAS, 2013; GIL, 2017).

A ciência do *design* busca desenvolver soluções úteis ao ser humano, diferentemente das ciências naturais que objetivam fazer descobertas e justificá-las (MARCH; SMITH, 1995; DRESCH; LACERDA; JÚNIOR, 2015). Com isso, o método científico DSR tem como premissa básica a utilidade dos resultados da pesquisa, atribuindo a ela uma perspectiva prescritiva no lugar de descritiva (PEFFERS et al., 2007; GREGOR; HEVNER, 2013). Assim, o DSR pode empregar as bases lógicas de investigação dedutivas e indutivas para produzir o artefato e tentar generalizá-lo como solução para uma classe de problemas.

O método DSR também faz uso do conhecimento descritivo e prescritivo. Sendo que o primeiro passa ser a base construtiva para o segundo, que possibilita a produção de um artefato. Conforme March e Smith (1995), esse artefato pode ser, entre outros, implementações, métodos ou modelos, objeto desta tese. Os autores ainda destacam que o desenvolvimento da ciência do *design* está pautado em duas etapas: na construção e avaliação do artefato.

A próxima seção da tese é referente ao Método de Trabalho e apresenta como a pesquisa

foi operacionalizada, descrevendo as diferentes técnicas de coleta e análise de dados utilizadas.

2.2 Método de Trabalho

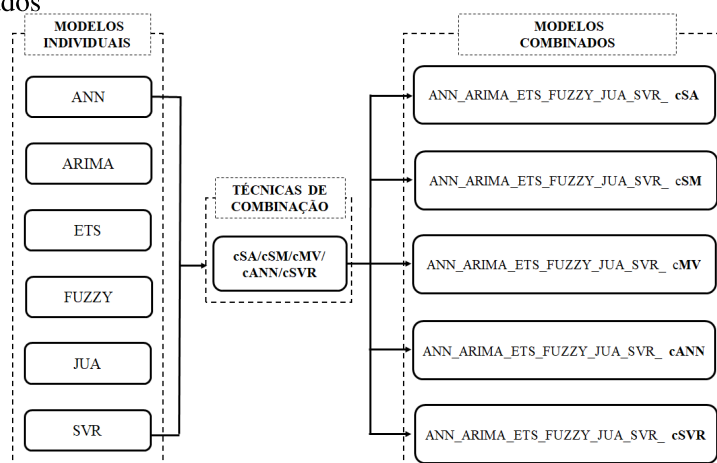
A metodologia de trabalho utilizada segue os procedimentos realizados em cada fase do processo de pesquisa científica do método DSR, ou seja, com foco na criação e avaliação.

A primeira fase, criação, compreende as atividades de definição do problema e a identificação do conhecimento descritivo e prescritivo que existe sobre o problema. A revisão bibliográfica sobre PCP, previsão de demanda, formalismos de séries temporais e medidas de desempenho de previsão, bem como a revisão sistemática da literatura com o auxílio do *software* IRAMUTEQ (*Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*) (MARCHAND; RATINAUD, 2012; CAMARGO; JUSTO, 2013), forneceram subsídios para a elaboração de um *framework* (estrutura) conceitual sobre os modelos de previsão, técnicas de combinação e medidas de performance utilizados no sistema de previsão proposto. Para esta fase, o estudo adota uma abordagem de previsão em dois estágios.

No primeiro estágio, foram realizadas previsões de demanda com aplicação dos modelos ditos individuais: Autorregressivo Integrado de Média Móvel (ARIMA), a família de modelos de Suavização Exponencial (ETS), Redes Neurais Artificiais (ANN), Regressão por Vetores de Suporte (SVR), Séries Temporais Fuzzy (FUZZY) e Modelo JUA (JUA), conforme Figura 1. Estes modelos são apresentados na Seção 3.4.1. No segundo estágio, essas previsões foram agrupadas através de técnicas de combinação de preditores, como a média simples (cSA, do inglês *Simple Average*), mínima variância (cMV, do inglês *Minimal Variance*), mediana simples (cSM, do inglês *Simple Median*), redes neurais (cANN, do inglês *Artificial Neural Networks*) e regressão vetores de suporte (cSVR de *Support Vector Regression*) emergindo novos combinadores, aplicados ao problema do planejamento da produção. Estas técnicas são apresentadas na Seção 3.4.2. Com isso, a demanda foi prevista mediante 11 modelos, sendo, 6 individuais e 5 formados a partir da combinação (agregação) destes.

A segunda fase, avaliação do artefato (modelos), foi realizada por meio de várias medidas de desempenho (acurácia e eficiência), metaheurísticas (otimização dos modelos) e critérios de informação (sobre a parcimônia) usados na seleção dos modelos e a comparação entre métodos estatísticos de previsão individual e combinados. Estas medidas são apresentadas na Seção 3.5. O modelo também foi comparado ao utilizado pela empresa, objeto de caso de estudo. O modelo usado pela empresa é descrito na Seção 3.4.1.6. O Quadro 1 resume as etapas do método de trabalho.

Figura 1: *Framework* de criação do artefato a partir de modelos individuais, técnicas de combinação e modelos combinados



Fonte: Elaborado pelo autor.

Quadro 1: Etapas do método de trabalho no processo de criação e avaliação dos modelos (artefato) de previsão

Etapa	Procedimentos	Objetivo
Criação	Revisão Bibliográfica Revisão Sistemática de Literatura	1. Analisar a relação entre os formalismos de previsão de demanda e o planejamento e controle da produção. 2. Desenvolver modelos de previsão a partir da combinação de preditores de demanda.
Avaliação	Avaliação dos modelos	3. Aplicar o(s) modelo(s) desenvolvido(s) em pelo menos um caso de estudo, mostrando sua eficiência na solução do problema. 4. Avaliar o desempenho do artefato proposto junto a empresa objeto de estudo. 5. Disponibilizar para o caso de estudo um modelo ótimo de previsão de demanda para produção, e suporte à gestão da programação da produção.

Fonte: Elaborado pelo Autor.

2.2.1 Revisão Bibliográfica

A revisão bibliográfica abrange quatro temas: Planejamento e Controle da Produção (PCP), Previsão de Demanda, Séries Temporais e Medidas de Desempenho de Previsão. Nos dois primeiros, exploram-se os conceitos sobre a Previsão da Demanda aplicada ao PCP. Nesta etapa foram considerados artigos científicos, publicações oficiais e livros sobre Previsão de Demanda e PCP. Optou-se por buscar informações a respeito das previsões de demanda para o planejamento da produção, visando compreender o tema independente do termo utilizado pelos autores.

A revisão sobre formalismos de séries temporais subjacentes parte do conceito dos modelos individuais e combinados, trazendo suas técnicas de combinação, medidas de desempenho de previsão e os critérios de informação. Para tanto, pesquisas com as expressões "demand

forecasting", *time series*" e *planning and production control*" foram realizadas em diversas bases de dados e no *Google Scholar*, e os artigos mais relevantes foram lidos. As fontes foram, principalmente, artigos científicos de periódicos de ampla circulação tais como: *The International Journal of Advanced Manufacturing Technology*, *Journal of Remanufacturing*, *Journal of Operations Management*, *Procedia Computer Science*, *International Conference on Industrial Engineering and Operations Management (IEOM)*, *Mathematical Problems in Engineering*, *Journal of Business Research*, *Computers & Industrial Engineering*, *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, *Journal Production & Manufacturing Research*, *European Journal of Operational Research*, *Journal of the Operational Research Society*, *International Journal of Production Research*, *International Journal of Applied Management Science*, entre outros.

2.2.2 Revisão Sistemática da Literatura

Por ser um tema bastante pesquisado, a previsão de demanda como auxílio para o planejamento da produção precisa de definições claras e concisas e os estudos existentes, seja no meio profissional ou acadêmico, atribuem várias abordagens a esse paradigma. A diversidade sobre a quantidade de modelos de previsão associada à incerteza das previsões, apresentam com frequência dúvidas sobre qual utilizar e de que forma modelar os dados da demanda. Tendo em vista o objetivo desta pesquisa, identificou-se a necessidade de sistematização do que já foi publicado. A revisão sistemática de literatura (RSL) teve esse propósito e foi conduzida em etapas como busca, seleção, leitura e análise.

2.2.2.1 Bases e descritores

Para a RSL, definiu-se como ponto de partida a compreensão de "**como o PCP e as ferramentas de análise de séries temporais de demanda interagem**", visando estabelecer *corpus* textual, isto é, o conjunto de artigos, que seriam incluídos no estudo. Em seguida, buscou-se uma amostra representativa da literatura científica internacional que pudesse contribuir com a resposta da questão de pesquisa, a partir de descritores como "*Demand Forecasting, Time Series e Planning and Production Control*". Os termos foram pesquisados nos campos: título, resumo e palavras-chave, nas bases *EBSCOhost*, *Emeraldinsight*, *Jstor*, *Taylor & Francis Online*, *IEEE*, *InderScienceOnline*, Portal de Periódicos da CAPES, *Scopus* e *ScienceDirect*. Realizou-se, também, uma busca no *Google Scholar* à procura de documentos que não estão indexados nas bases mencionadas anteriormente. O período definido para a pesquisa foi de

janeiro de 2014 até dezembro de 2018, incluindo documentos com data de publicação para 2019. Devido a sua representatividade junto à comunidade científica, em termos gerais, e especificamente em se tratando dos temas objetos de estudo, foram considerados apenas artigos e *reviews* escritos na língua inglesa.

2.2.2.2 Critérios de inclusão e exclusão

Para assegurar uma qualidade mínima aos artigos e *reviews* selecionados, foram estabelecidos critérios de inclusão e exclusão dos textos. Foram desconsiderados todos os textos que falavam apenas de PCP ou PP como contexto, mas que não discutiam o paradigma, ou não falavam dos formalismos de séries temporais a ele relacionados. Também foram retiradas da análise pesquisas que tratavam da aplicação do conceito fora do contexto da previsão de demanda para indústria de manufatura. Isto porque o objetivo desta revisão é a compreensão da relação entre os formalismos de previsão de demanda e o PCP. Por outro lado, foram incluídos os textos que têm como tema central formalismos de séries temporais aplicados a demanda para planejamento da produção. Ao final das buscas e filtros, 60 documentos compõem o *corpus* analisado.

2.2.2.3 Análise léxica e de conteúdo

O *corpus* textual construído, versando sobre as relações entre previsão de demanda e PCP, foi analisado em seu conteúdo, porém partindo-se de um estudo léxico a partir do *software* IRAMUTEQ. Da análise de conteúdo, a pesquisa revela o desequilíbrio da literatura, que eleva as dificuldades de melhores resultados da área de PCP. Destaque-se também que poucos trabalhos dedicados ao PCP têm recorrido à combinação de preditores para previsão de demanda, uma das mais bem-sucedidas classes de formalismos para modelagem e predição de séries temporais, seja por sua simplicidade, ou pela qualidade dos seus resultados. Buscou-se, também, identificar os formalismos de séries temporais subjacentes, as técnicas de combinação, as métricas de desempenho utilizados na previsão da demanda para o planejamento da produção e os tipos de indústria de manufatura aplicados.

O estudo possibilitou, entre outras coisas, a elaboração da estrutura conceitual sobre o sistema de previsão a ser proposto. O artefato oferece uma estrutura robusta para a previsão de demanda e avaliação no planejamento da produção, e por isso foi submetido à avaliação por meio das medidas de desempenho (MSE, MAPE, U-Theil, ARV, WPOCID, ID, Reg_Intercept, Reg_Slope, WR^2 e APM), metaheurísticas, critérios de informação e pela me-

metodologia usada pela indústria de manufatura. Finalizada a análise textual, com o uso da ferramenta, os resultados produzidos atenderam aos Objetivos Específicos a e b da presente pesquisa e, adicionadas às etapas seguintes, contribuiu para atender o objetivo geral desta tese.

2.2.3 Definição do Caso de Estudo

Com o objetivo de avaliar a qualidade dos modelos desenvolvidos, sua aplicação a uma indústria de manufatura foi apropriada. Foram utilizados elementos do caso de estudo para a avaliação do artefato (DRESCH; LACERDA; JÚNIOR, 2015). A unidade de análise da presente pesquisa são históricos de demanda, isto é, séries temporais univariadas, da indústria de manufatura de produtos de limpeza e saneantes, escolhida por apresentar características semelhantes ao concluído na RSL, além da dificuldade de ajuste por parte de modelos tradicionais a dados semelhantes (ARMSTRONG, 2001; BÓNA; LÉNÁRT, 2014) e da incerteza da demanda subjacente (CLAISSE; CHU; JEMAI, 2016; ALIYEVA, 2017; AMIRKOLAI et al., 2017).

A empresa objeto de estudo é uma indústria do setor de produtos de higiene, limpeza e saneantes de uso doméstico e profissional, fundada em 1998 e localizada na cidade de Juazeiro do Norte no estado do Ceará. Atualmente a indústria conta com uma área total de aproximadamente $17.600m^2$, é dotada de um moderno parque industrial, com capacidade de produção de 2.100.000kg de produtos de limpeza por mês. Infraestrutura esta que, aliada à sua frota própria, permite atender as demandas dos clientes de vários estados do Nordeste como Bahia, Ceará, Maranhão, Paraíba, Pernambuco e Piauí. A empresa atua na industrialização e comercialização de produtos de limpeza e conta com um *mix* de produtos diversos como Água Sanitária, Amaciante, Desengordurantes, Desinfetante, Lava Louças, Lava Roupas, Limpa Alumínio, Limpa Vidros, Limpadores Perfumados, Sabão de Coco, Sabão em barra e Sabão em Pó.

Este estudo focou na previsão da demanda de famílias de produtos. Entenda-se por família de produtos as classificações atribuídas aos produtos cadastrados no sistema ERP do inglês *Enterprise Resource Planning*, em que se identifica, por exemplo, uma linha de produto, cor, tamanho, finalidade, vasilhame, fragrância, e assim por diante.

A partir de dados coletados nas entrevistas e visitação *in loco*, constatou-se que a indústria possui um departamento específico para estudar e calcular as previsões de demanda, sendo vinculado aos setores de produção e comercial. Este departamento funciona com subdivisões em determinadas unidades, denominadas Planejamento Comercial, que deriva para Planejamento e Programação de Controle da Produção. Este setor visa a otimização do parque fabril e

o controle das matérias-primas e dos estoques através da Gestão Integrada de Planejamento de Produção. Dessa forma, essa espécie de árvore que tem sua raiz no Departamento Comercial, ramifica-se pelo Planejamento Comercial de modo a distribuir de forma assídua e correta as informações necessárias para uma previsão de demanda que esteja de acordo com a realidade da fábrica de produção, assim iniciando o planejamento estratégico.

2.2.4 Coleta dos Dados

Para a pesquisa em questão, foram utilizados registros já disponíveis e acessíveis mediante pesquisa documental na empresa. Autores como Prodanov e Freitas (2013) e Severino (2017), classificam esses dados como secundários, por se tratarem de “dados de segunda-mão”, isto é, não construídos a partir de esforços e instrumentos coordenados pelo próprio pesquisador. As informações referentes à indústria em questão foram colhidas *in loco*. O processo de coleta dos dados da indústria do setor de higiene, limpeza e saneantes, objeto de estudo, foi realizado com dados históricos mensais de janeiro de 2015 a abril de 2021, armazenados no sistema *Winthor*, conforme Tabela 1. A exceção é a série temporal de demanda “DJUF7”, com dados históricos mensais de julho de 2015 a abril de 2021.

A Tabela 1 apresenta um resumo das características das séries temporais de demanda dos produtos. A coluna “Série” apresenta as séries temporais formadas a partir do conceito de família de produtos utilizados pela empresa. As outras colunas dispõem do código dos produtos (Cód. Prod.), descrição e linha de produção ao qual pertence. A seção seguinte apresenta os passos usados na construção e análise dos modelos do sistema de previsão.

2.2.5 Construção dos modelos e análise dos dados

Nesta seção, abordam-se a descrição da metodologia adotada na modelagem, previsão e avaliação dos resultados gerados pelos modelos utilizados e suas respectivas combinações. O tratamento, a modelagem e a análise estatística dos dados foram realizados por meio do *software* gratuito R (R FOUNDATION FOR STATISTICAL COMPUTING, 2019), através da interface *RStudio* e com o auxílio de pacotes disponíveis no programa e do desenvolvimento de algoritmos específicos. O procedimento metodológico seguinte foi dividido em quatro etapas, mostradas na Figura 2. A base de dados em séries temporais da indústria de manufatura foi analisada sobre cada variável de demanda. As métricas de desempenho utilizadas para analisar os resultados de previsão dos modelos são: erro quadrático médio (MSE), erro médio absoluto percentual (MAPE), a estatística U de *Theil* (U-Theil), variância relativa média

Tabela 1: Resumo das características das séries temporais de demanda das famílias de produtos objetos de estudo

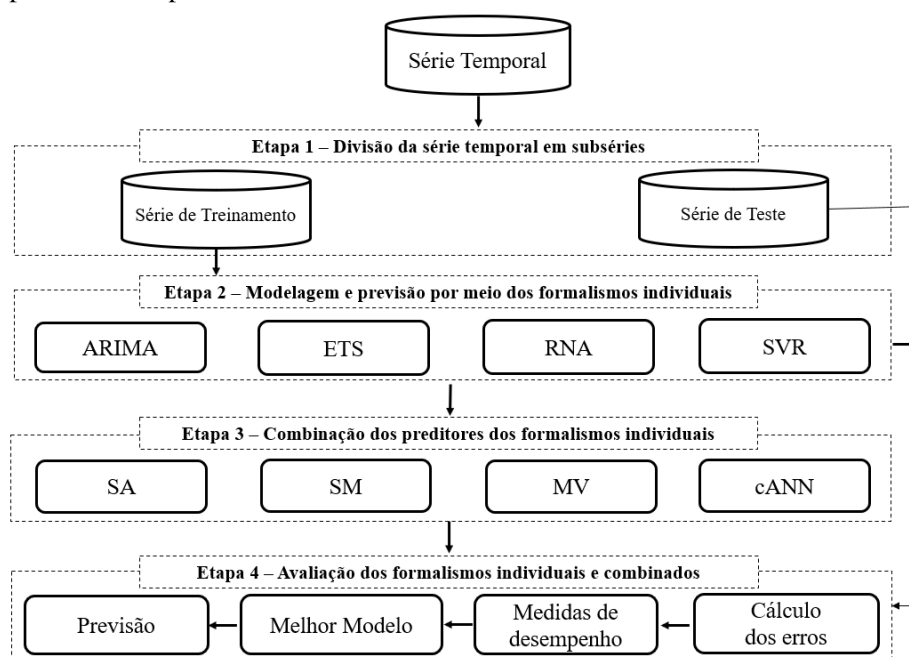
Série	Cód. Prod.	Descrição	Seção	Linha
DJUF1	3	Sabão JUA Glic. Limão CX10X1000G	Sabão em Barra	5
	4	Sabão JUA Glic. Limão CX C/20X500G	Sabão em Barra	
	5	Sabão JUA Glic. Limão CX C/50X200G	Sabão em Barra	
	12	Sabão JUA Glic. Azul CX C/50X200G	Sabão em Barra	
	13	Sabão Baleia Glic. Limão 50X200G	Sabão em Barra	
DJUF2	6	Sabão JUA Glicerinado Neutro CX C/10X1000G	Sabão em Barra	5
	7	Sabão JUA Glic. Neutro CX C/20X500G	Sabão em Barra	
	8	Sabão JUA Glic. Neutro CX C/50X200G	Sabão em Barra	
	10	Sabão JUA Glic. Maça CX C/50X200G	Sabão em Barra	
	14	Sabão Baleia Neutro Glic. 50X200G	Sabão em Barra	
DJUF3	32 1308	Lava Roupas em Pó JUA PC C/24 SACHE 500G Lava Roupas em Pó JUA PC C/12 SACHE 1000G	Sabão em Pó Sabão em Pó	Pó
DJUF6	23	Det. Lava Louças JUA Clean CXC/24X500ML	Detergente	4
	24	Det. Lava Louças JUA Limão CXC/24X500ML	Detergente	
	25	Det. Lava Louças JUA Maça CXC/24X500ML	Detergente	
	26	Det. Lava Louças JUA Neutro CXC/24X500ML	Detergente	
DJUF7	546	Det. Lava Roupas Liq. JUA C/12X1L(S. CONC)	Lava Roupas	2
	763	Det. Louças JUA Limão CX C/12X1000ML	Detergente	
	764	Det. Louças JUA Maça CX C/12X1000ML	Detergente	
	765	Det. Louças JUA Neutro CX C/12X1000ML	Lava Roupas Líq.	
DJUF11	28	Det. Lava Roupas Liq. JUA 24X500ML(S. CONC)	Lava Roupas Líq.	2
DJUF12	16	Des. JUA Floral CX C/12X1000ML	Desinfetante	3
	18	Des. JUA Lavanda CX C/12X1000ML	Desinfetante	
	21	Des. JUA Limão CX C/12X1000ML	Desinfetante	
	69	Des. JUA Herbal CX C/12X1000ML	Desinfetante	
	771	Des. JUA Flor Campo CX C/12X1000ML	Desinfetante	
DJUF13	17	Des. JUA Floral CX C/6X2000ML	Desinfetante	3
	20	Des. JUA Lavanda CX C/6X2000ML	Desinfetante	
	22	Des. JUA Limão CX C/6X2000ML	Desinfetante	
	770	Des. JUA Herbal CX C/6X2000ML	Desinfetante	
	772	Des. JUA Flor Campo CX C/6X2000ML	Desinfetante	

Fonte: Elaborado pelo Autor.

(ARV), erro de predição de alteração na direção da série (WPOCID), índice de desacordo (ID), Reg_Intercept, Reg_Slope, WR^2 e a medida de desempenho agregada (APM), introduzidas na Seção 3.5.

Especificamente, na Etapa 1 (Divisão da série temporal em subséries), a série foi particionada em dois conjuntos, um destinado à fase de modelagem (quando os modelos são treinados), na qual os preditores são construídos, com 80% das observações da série, e outro voltado à fase de teste dos modelos (dados desconhecidos), com 20% das observações restantes, esse processo foi considerado na aplicação dos modelos ARIMA, ETS e FUZZY. Para os modelos ANN e SVR, a série foi dividida em três conjuntos: treinamento, validação e teste. O conjunto de validação foi usado para estimar o erro de predição para seleção de modelo, com 30% das observações do conjunto de treinamento. Considerando o modelo JUA, os dados de previsão referentes ao período especificado foram repassados pela empresa e inseridos na etapa para

Figura 2: Arquitetura da abordagem proposta para modelagem e aplicação dos modelos individuais, para o desenvolvimento dos modelos combinados, verificação da qualidade dos modelos a partir das medidas de precisão e da previsão da demanda



Fonte: Elaborado pelo autor.

combinação.

Todas as séries de treinamento de demanda dos produtos foram normalizadas no intervalo $[0, 4; 0, 6]$. A utilização desse intervalo vem da necessidade das entradas para modelagem via ANN estarem compreendidas entre $[0, 0; 1, 0]$, e, caso haja extrapolação do intervalo $[0, 4; 0, 6]$, os valores ainda tenderão a permanecer no intervalo $[0, 0; 1, 0]$, na etapa de teste. Neste sentido, o conjunto de valores para treinamento foi utilizado para estimação dos modelos, de acordo com critérios de qualidade. Já o conjunto de valores de teste foi utilizado para verificar se os formalismos apresentaram a capacidade de predizer o processo gerador da série temporal. Por fim, o conjunto de validação foi utilizado para verificar a eficiência dos modelos quanto a sua capacidade de generalização durante o treinamento.

Na Etapa 2 (Modelagem e previsão por meio dos formalismos individuais de séries temporais), os modelos ARIMA e ETS foram implementados por meio do pacote de previsão *forecast* do *software* R (HYNDMAN; KHANDAKAR et al., 2007), respectivamente, usando as funções *auto.arima* e *ets*. Os melhores modelos para cada metodologia foram selecionados de acordo com o critério de informação *Bayesiano* (BIC), introduzido na Seção 3.5.9.

Os formalismos individuais de *Machine Learning*, ANN e SVR, foram implementados computacionalmente via pacotes `neuralnet` (FRITSCH; GUENTHER; GUENTHER, 2019) e `e1071` (MEYER et al., 2020), respectivamente, usando as funções `neuralnet` e `svm`. As arquiteturas ANN e SVR foram otimizadas via *simulated annealing* (SAn) (inserido na Seção 3.5.10), de acordo com o pacote `GenSA` (XIANG et al., 2013).

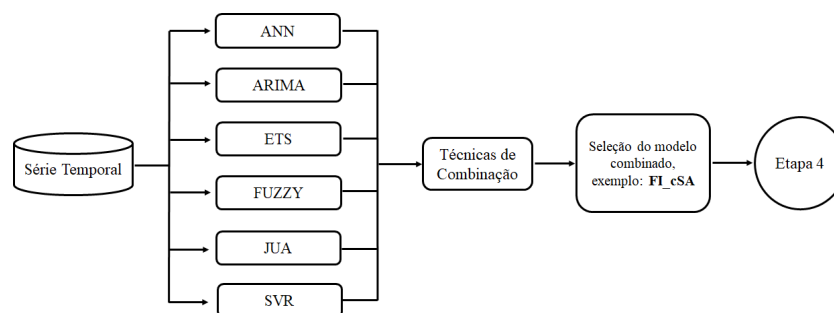
O SAn é uma metaheurística que foi adaptada para resolver problemas de otimização probabilística (FRANZIN; STÜTZLE, 2019). Aqui, o SAn procura minimizar a função BIC. Para variações do algoritmo de aprendizagem (AL) na ANN foram utilizados: BPROP, RPROP+, RPROP-, SAG e SLR e para a função de ativação (FA) usou-se LF e THF. No caso do SVR, o tipo adotado foi *eps-regression*, utilizado para casos de regressão via SVR e as variações para função Kernel foram: Linear (LK), Polinomial (PK), Sigmoides (SK) e Base Radial (RBFK).

O modelo de inteligência computacional de séries temporais fuzzy (STF), que utiliza os princípios fuzzy (introduzido na Seção 3.4.1.5.3), foi implementado por meio da construção de algoritmo em R, aqui denominado de FUZZY. O código foi desenvolvido seguindo as etapas da metodologia AM (ABBASOV; MAMEDOVA, 2003) e os parâmetros do modelo D_1 , D_2 , n , w e C foram selecionados através da metaheurística *simulated de annealing* (SAn) via pacote `GenSA` (XIANG et al., 2013), na determinação dos parâmetros ótimos. O último passo desta etapa foi a realização da previsão dos valores futuros por meio dos modelos produzidos.

Para a Etapa 3 (Combinação dos preditores dos formalismos individuais), após a previsão dos modelos individuais, cinco técnicas de agregação (cSA, cSM, cMV, cANN e cSVR) (inserido na Seção 3.4.2) foram usadas no desenvolvimento dos modelos combinados. As combinações via média simples (cSA), mediana simples (cSM) e mínima variância (cMV) levaram em consideração as previsões dos modelos individuais, calculando-se, respectivamente, a média, a mediana e a covariância destes. A arquitetura do combinador via redes neurais (cANN) segue a mesma estrutura do formalismo individual ANN descrito na Etapa 2. Porém, as entradas serão as previsões dos formalismos individuais.

O cANN é otimizado via *simulated annealing* (SAn) buscando-se minimizar a função BIC. Utiliza-se as mesmas variações para o algoritmo de aprendizado bem como para a função de ativação. A combinação via regressão por vetores de suporte (cSVR) ocorre de forma similar à modelagem individual, no entanto, as entradas serão as previsões dos modelos individuais. Os modelos combinados via cSVR foram otimizados de forma similar aos SVR individuais. A variação para o *custo* ocorreu no intervalo $[10^{-5}, 10^2]$, a do ϵ entre $[10^{-5}, 1]$, γ do em $[10^{-5}, 10^4]$, do *degree* em $[1, 5]$ e para o coef0 tomou-se valores no intervalo $[0, 10]$. Para as funções kernel considerou-se Linear, Polinomial, Sigmoides e Base Radial. Na Figura 3, são mostrados os passos para a obtenção das estimativas das respectivas combinações.

Figura 3: Metodologia para a combinação dos modelos individuais de previsão de séries temporais de demanda



Fonte: Elaborado pelo autor.

A combinação das previsões foi realizada agrupando todos os modelos individuais a partir de cada uma das cinco técnicas de agregação, conforme Figura 3. Para cada combinação, foram utilizados e comparados os resultados dos métodos de incorporação descritos. Para simplificar o entendimento das combinações idealizadas, inicialmente define-se a nomenclatura a ser usada pelo modelo. Os formalismos ditos individuais são representados por FI. Por exemplo, a combinação FI_cSA foi obtida pela agregação dos modelos individuais (ANN, ARIMA, ETS, FUZZY, JUA e SVR) e por meio da técnica cSA. As etapas de 1 a 3 fazem parte do processo de criação do artefato em conformidade com o método DRS.

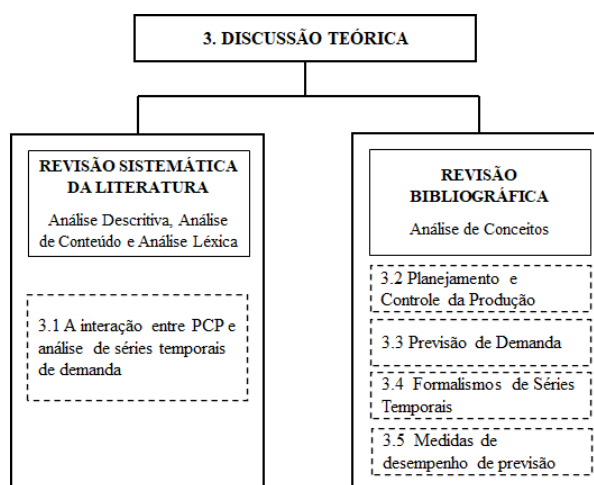
Na Etapa 4 (Avaliação dos formalismos individuais e combinados), foi realizada a avaliação dos modelos de previsão a partir da aplicação das métricas de desempenho MSE, MAPE, U de Theil, ARV, WPOCID, ID, Reg_Intercept, Reg_Slope, WR^2 e APM (inserido na Seção 3.5), para fins de cálculo e comparação dos modelos simples e combinados. O objetivo foi verificar qual dos métodos resultará em um melhoramento significativo na capacidade preditiva.

Por fim, o melhor modelo selecionado a partir da Etapa 4 foi comparado com a metodologia usada pela indústria de manufatura. Caso o desempenho do modelo seja melhor que o da metodologia usada atualmente pela empresa, será sugerida a implementação do modelo proposto na resolução do problema de planejamento e programação da produção. Esta etapa do método de trabalho está associada à segunda parte do processo de pesquisa DSR: a de avaliação do artefato.

3 DISCUSSÃO TEÓRICA

A discussão teórica está organizada em cinco subcapítulos: um apresentando a interação entre o PCP e as ferramentas de análise de séries temporais de demanda, outro caracterizando o PCP, o seguinte explicando sobre previsão de demanda, o quarto destacando os formalismos de séries temporais subjacentes e o último focado nas medidas de desempenho de previsão. A estrutura da fundamentação teórica é ilustrada na Figura 4. O primeiro subcapítulo foi

Figura 4: Estrutura da Discussão Teórica envolvendo a Revisão Sistemática de Literatura e Revisão Bibliográfica



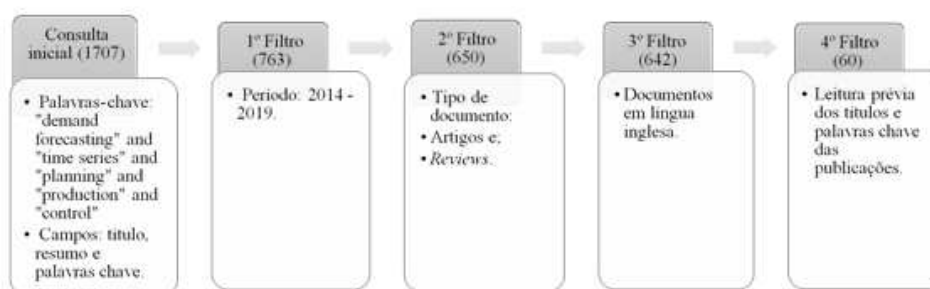
Fonte: Elaborado pelo autor.

elaborado a partir de uma revisão sistemática de literatura com uso do *software* IRAMUTEQ, interface para o *software* estatístico R. Os outros quatro subcapítulos foram desenvolvidos por meio de uma revisão bibliográfica das principais referências na área, envolvendo também o estado da arte no que tange as interações entre PCP e estudos de séries temporais de demanda.

3.1 A interação entre PCP e análise de séries temporais de demanda

Considerando o objetivo de mapear a produção científica internacional sobre os descritores, a construção do *corpus* textual (CAMARGO; JUSTO, 2013; KAMI et al., 2016) foi realizada por meio do cruzamento dos descritores, bem como seus correlatos em inglês. As bases consideradas possuem um repertório significativo de ferramentas que facilitam as avaliações bibliométricas (AGUILLO, 2012). Algumas destas mantêm diversos artigos disponíveis para pesquisadores e interessados, e ainda permitem a exportação de documentos em vários formatos. A Figura 5 resume a metodologia adotada para a seleção de documentos.

Figura 5: Constituição da amostra de pesquisa às bases de dados consultadas acerca da literatura internacional envolvendo previsão de demanda, séries temporais e planejamento e controle da produção. Entre parênteses, no cabeçalho de cada ação, encontra-se o número de documentos resultantes



Fonte: Elaborado pelo autor.

Inicialmente realizou-se uma consulta aos títulos indexados nas bases de dados, baseando-se nos critérios definidos previamente para a busca sistemática dos documentos a serem analisados, e resultou um total de 1.707 textos contendo as expressões de busca nos campos “título”, “resumo” ou “palavras-chave”, considerando registros de todo o período da base de conhecimento até a data de realização da extração – dezembro de 2018, o que mostra a relevância do tema no meio acadêmico.

O primeiro filtro buscou selecionar documentos publicados no período de “2014 até 2019”, sendo encontrados 763. O segundo filtro aplicado limitou a pesquisa às categorias “artigos” e “reviews”, ficando 650. O último filtro incluiu a restrição a documentos em língua inglesa, resultando em 642, sendo considerada uma prévia da amostra estabelecida para a pesquisa. Baseado nesta amostra, foram realizados dois tipos de análise. A primeira se dedicou a uma leitura prévia do título, resumo dos documentos e das palavras-chave a fim de identificar e avaliar as características dos documentos que se relacionam com o tema da pesquisa, sendo selecionados 60. Para tanto, tomou-se de suporte ao *software* Microsoft Excel.

Uma vez construído o *corpus*, deu-se início sua análise léxica, com o suporte do *software* IRAMUTEQ, versão 0.7 alpha 2. Para isso, o conjunto de palavras-chave, resumos, introdução e conclusão dos artigos foram identificados e explorados pela ferramenta “análise de texto”, apresentados por meio de ilustrações e agrupamentos de palavras, possibilitando inferências sobre dados textuais qualitativos de forma sistemática e quantitativa (ABHAYAWANSA, 2011).

Para a análise léxica do *corpus* textual, cada artigo correspondeu a um documento. Por sua vez, cada segmento de texto (ST) desse artigo correspondeu a um fragmento de texto. O *corpus* textual foi então composto por 60 documentos, que deram origem a 1.914 STs, com

dos temas de modelagem e fundamentação pode levar profissionais aplicados e pesquisadores da área de PCP a subutilizarem os formalismos existentes e mesmo auxiliarem menos no suporte ao desenvolvimento e aprimoramento de alternativas.

Do *corpus* textual analisado, a Classe 2 apresenta o menor número de segmentos de textos, indicando uma baixa exploração dos termos formadores do campo conceitual “Modelagem e Previsão” no PCP. Mais uma vez, isto pode refletir um distanciamento dos formalismos dedicados à demanda (considerada a variável mais importante de um sistema de produção, principalmente para as funções desenvolvidas pelo PCP), em relação às demais áreas da organização.

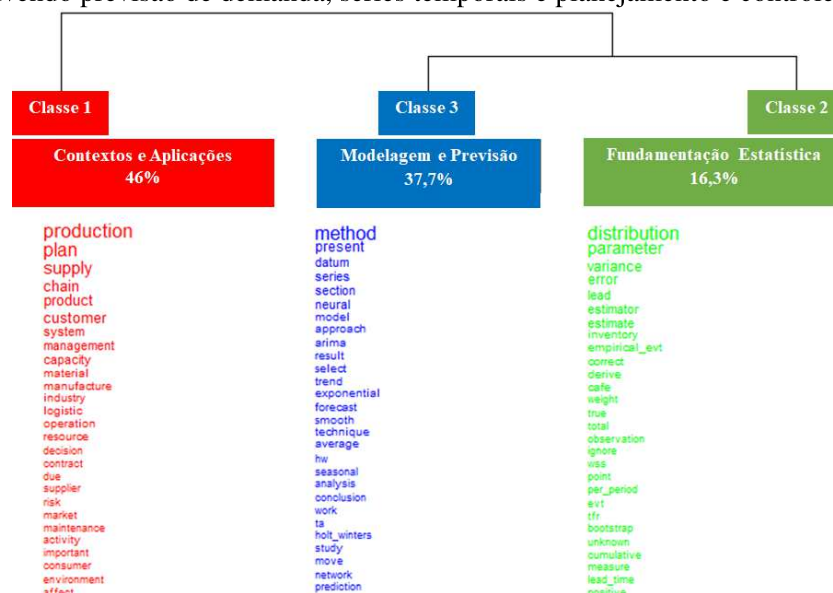
Especificamente, a Classe 1 é composta por artigos que priorizam estudos e conceitos sobre PCP, Planejamento da Produção, Cadeia de Suprimentos e Indústria (AIT-ALLA et al., 2014; SIVARAMAKUMAR; JAIN, 2014; HIBINO et al., 2014; DWEIRI; KHAN; JAIN, 2015; ZULKEPLI; FONG; ABIDIN, 2015; LIU et al., 2015; GANSTERER, 2015; BARBOSA; CHRISTO; COSTA, 2015; HOLLMANN; SCAVARDA; THOMÉ, 2015; LAAN et al., 2016; VRONTAKIS; KAMPIANAKIS; TSINARAKIS, 2016; UZZAMAN et al., 2016; JATTA; KRISHNAN, 2016; VEIGA; VEIGA; TORTATO, 2016; MIRCETIC et al., 2016; KIM; DEKKER; HEIJ, 2017; AMIRKOLAI et al., 2017; ROSIENKIEWICZ; CHLEBUS; DETYNA, 2017; LEE; LIANG, 2018; PEREIRA et al., 2018; UZSOY; FOWLER; MÖNCH, 2018; GOODWIN; GÖNÜL; ÖNKAL, 2019), sendo composta por 23 artigos.

Já a Classe 2 (Figura 7), agrupa conceitos sobre distribuição, parâmetros, variância, erro, mensuração e estimador, voltados à fundamentação estatística. É possível perceber, além dos conceitos encontrados, uma relação significativa entre os termos *bootstrapping*, simulação de distribuição, *lead time*, erro de previsão, precisão ou viés de previsão, qualidade das previsões, distribuições de demanda, parâmetros dos modelos e estimadores de parâmetros (MATSUMOTO; IKEDA, 2015; REGO; MESQUITA, 2015; SCARPEL, 2015; SYNTETOS; BABAI; GARDNER JR, 2015; WANG; PETROPOULOS, 2016; PRAK; TEUNTER; SYNTETOS, 2017; ZHU et al., 2017; HASNI et al., 2018; HA; SEOK; OK, 2018; PRAK; TEUNTER, 2019).

A Classe 3 se dedica à inferência dos fenômenos de demanda, a partir de formalismos como ANNs, ARIMA, ETS e modelos de médias. Em outros termos, ela trata de conceitos relacionados à modelagem e previsão de demanda, envolvendo qualidade de modelos lineares e não lineares, combinações e incorporação de métodos (BÓNA; LÉNÁRT, 2014; SCHOLZREITER; KÜCK; LAPPE, 2014; SCARPEL, 2015; MELO MENEZES et al., 2015; PETROPOULOS; KOURENTZES, 2015; MATSUMOTO et al., 2016; FENG; MEARS; SCHULTE, 2016; TRATAR; MOJŠKERC; TOMAN, 2016; CHANG; YU; JIN, 2016; TIRKES; GU-

RAY; CELEBI, 2017; KOURENTZES; ROSTAMI-TABAR; BARROW, 2017; SONI; SRI-KANTH, 2017; VARGAS; CORTÉS, 2017; SAHIN; ELDEMIR, 2018; HAMICHE et al., 2018; DOMBI; JÓNÁS; TÓTH, 2018; MOR et al., 2019).

Figura 7: Dendrograma representativo das repartições em classes e palavras acerca da literatura internacional envolvendo previsão de demanda, séries temporais e planejamento e controle da produção



Fonte: Elaborado pelo autor.

O Quadro 2 fornece a lista das principais palavras significativamente associadas a essas classes de acordo com a estatística de qui-quadrado (χ^2) de Pearson. Neste caso, o *valor - p* para as associações foi sempre menor que 0,0001. Logo, pode-se considerar que existem associações sob uma baixa probabilidade de equívoco. A coluna (χ^2) do quadro mostra a estatística de teste entre palavra e a classe; quanto maior, maior a probabilidade de dependência.

A partir da representação gráfica das Figuras 8a e 8b, observa-se que ocorreu um leque semântico de palavras mais frequentes no *corpus*: *forecast*, *demand*, *method*, *model* e *time*, agrupadas em zonas centrais e periféricas.

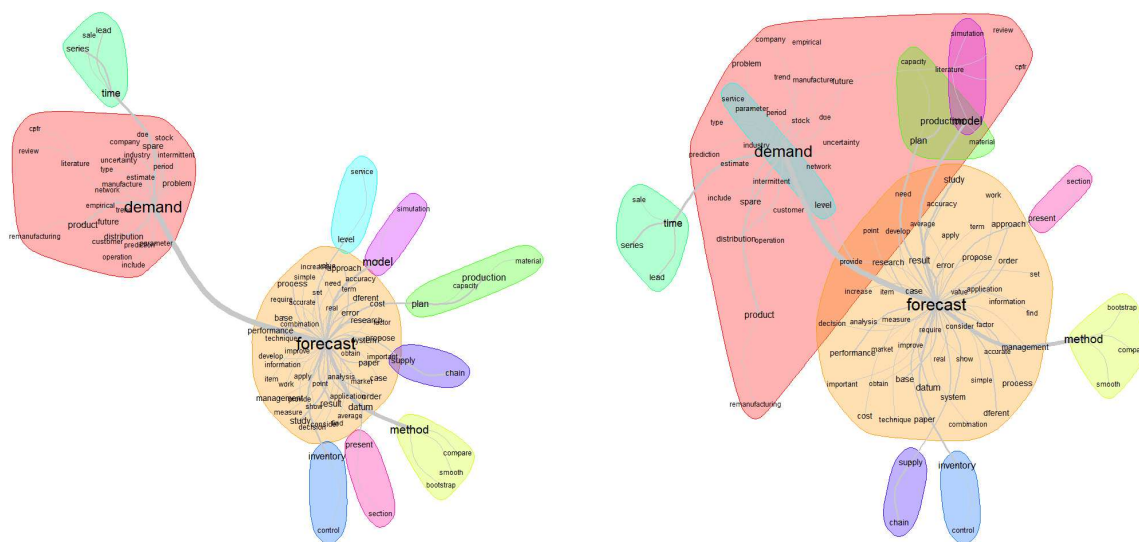
A análise de similaridade realizada Figura 8a fornece uma representação centrada na palavra *forecast* e em grupos de conceitos como *demand*, *plan*, *model* e *method* que estão interligados, porém independentes. Estes podem ser considerados igualmente importantes na literatura para o PCP. Assim, há indícios de que pesquisas anteriores consideram subcampos de estudos específicos aplicados ao PCP como planejamento de produção, controle de estoque, desempenho de inventário, planejamento de recursos, tomada de decisão, cadeias de suprimentos, e assim por diante. Ainda na Figura 8a, percebe-se uma certa deficiência de pesquisas em que

Quadro 2: Teste de associação entre palavras e classes, onde a Classe 1: Contextos e Aplicações; Classe 2: Fundamentação Estatística; Classe 3: Modelagem e Previsão. As Colunas 1-3 representam, nesta ordem, o termo, a estatística χ^2 e o % citações

Classe 1			Classe 2			Classe 3		
Palavra	χ^2	%	Palavra	χ^2	%	Palavra	χ^2	%
<i>Production</i>	161,6	81,3	<i>Distribution</i>	225,8	69,6	<i>Method</i>	206,8	65,0
<i>Planning</i>	142,1	77,6	<i>Parameter</i>	159,7	64,8	<i>Present</i>	89,0	75,2
<i>Supply</i>	125,7	86,6	<i>Variance</i>	126,2	83,8	<i>Data</i>	80,2	64,4
<i>Chain</i>	91,8	85,3	<i>Error</i>	97,0	48,3	<i>Series</i>	75,9	67,2
<i>Product</i>	91,7	74,6	<i>Lead</i>	96,1	49,6	<i>Section</i>	64,4	78,8
<i>Customer</i>	89,1	90,5	<i>Estimator</i>	91,9	95,0	<i>Neural</i>	62,5	100,0
<i>System</i>	65,5	77,1	<i>Estimate</i>	82,8	54,8	<i>Model</i>	52,7	53,3
<i>Management</i>	65,0	80,3	<i>Inventory</i>	65,2	33,5	<i>Approach</i>	52,3	66,7
<i>Capacity</i>	54,3	96,2	<i>Empirical</i>	62,1	100,0	<i>Arima</i>	48,7	94,3
<i>Material</i>	47,8	94,0	<i>Correct</i>	59,9	83,3	<i>Result</i>	47,1	59,5

Fonte: Elaborado pelo autor.

Figura 8: Análise de Similitude acerca da literatura internacional no período de 2014 a 2019 envolvendo previsão de demanda, séries temporais e planejamento e controle da produção



(a) Análise de Similitude entre as palavras do *corpus* (b) Análise de Similitude centrada no subconjunto das envolvendo previsão de demanda, séries temporais e palavras mais usadas planejamento e controle da produção

Fonte: Elaborado pelo autor.

o planejamento da demanda para o PCP ocorre a partir de modelos aplicados (formalismos) a séries temporais. Os métodos aplicados nestes casos são considerados simples e podem apresentar erros de previsão elevados. Uma segunda análise de similaridade lexical foi realizada considerando o subconjunto das palavras mais usadas (ou seja, *demand*, *time*, *method*, *plan* e *model*), e sem o conceito central de *forecast*. Isto permite um olhar mais profundo nos campos

conceituais subjacentes à Figura 8b. Ainda na Figura 8b, observa-se a existência de alguns conceitos interligados e que aparecem em um novo mapeamento. Contudo, nota-se a baixa relação entre os conceitos *plan* (grupo verde) e *demand* (grupo vermelho), que apresentam um baixo volume de pesquisas que utilizam a demanda como forma de planejar a produção.

Fica evidente, nas relações entre o subconjunto das palavras ativas do dendrograma, que o processo de previsão pressupõe uma fundamentação e aplicação prática, o qual se constitui o ponto de partida para a tomada de decisões relacionadas ao Planejamento e Controle da Produção. Assim, percebe-se que as palavras referentes aos elementos associados ao PCP aparecem de forma independentes dos termos pertinentes à previsão. Isso tende a mostrar que os conceitos ainda não estão bem conectados na literatura, diferentemente de outros como *model*, *planning*, *method* e *supply chain* relacionados a previsão, que estão mais claramente conectados.

Relacionados aos modelos, métodos e técnicas de previsão identificados nos documentos analisados, estes fundamentais para o planejamento da demanda e importantes para as etapas subsequentes do planejamento e do controle efetivo da produção. Autores sugerem que a qualidade das previsões pode melhorar sensivelmente combinando apropriadamente métodos diferentes (MATSUMOTO et al., 2016; WANG; PETROPOULOS, 2016), principalmente em comparação com os modelos individuais. Exemplo: Rosienkiewicz, Chlebus e Detyna (2017) usaram a combinação na previsão de demanda irregular, típica de peças de reposição. Guo e Shen (2016), estabeleceram um modelo de predição de demanda para atividades de produção, combinando os mecanismos de previsão dos métodos suavização exponencial e rede neural artificial. Petropoulos e Kourentzes (2015), exploraram empiricamente a eficiência das combinações na previsão da demanda intermitente, produzidas a partir de métodos paramétricos.

Apesar da utilização de diferentes formalismos de previsão, dos 60 documentos analisados, observa-se uma pouca exploração de métodos individuais robustos, e menos ainda da combinação de modelos e do uso de diferentes técnicas de incorporação na previsão da demanda a partir de séries temporais. Ainda, de acordo com a exploração dos autores, nenhuma pesquisa sobre previsão de demanda foi relacionada à implementação na indústria de produtos de higiene, limpeza e saneantes.

3.2 Planejamento e Controle da Produção

O Planejamento e Controle da Produção (PCP) desempenha um papel importante nas indústrias modernas de manufatura de bens (PECHMANN; ZARTE, 2017). Na literatura sobre PCP não há uma definição comum compartilhada por autores como (SLACK; CHAMBERS;

JOHNSTON, 2010; SANDERS; REID, 2013; STEVENSON; SUM, 2015). Porém, apesar de não haver consenso, existe um pensamento compartilhado na direção da eficiência e da eficácia nos sistemas de produção. Tem-se em comum o conceito de que o PCP é um sistema de suporte à produção, que gerencia e coordena o processo produtivo, visando cumprir o planejamento, a programação e controle dos processos de maneira eficaz, a fim de satisfazer os requisitos de tempo, qualidade e quantidade.

O PCP aborda uma função fundamental de produtividade, gerenciamento e utilização de recursos (MOURTZIS; DOUKAS, 2014; WIENDAHL, 2014). Pesquisas revelam que os sistemas de planejamento sofrem com baixa qualidade e baixa faixa nos dados de planejamento, o que resulta em tempos de entrega muitas vezes irreais (GEORGIADIS; POLITOU, 2013). Isso leva ao dilema do PCP, ou seja, alcançar alta eficiência de processo, baixo tempo de produção e boa confiabilidade no planejamento, apesar de um ambiente turbulento, com incertezas como mudanças dinâmicas na fábrica, restrições, ciclos de vida de produto curtos, uma variedade cada vez maior da individualização da demanda (SCHUH; POTENTE; FUCHS, 2012).

Para a Associação de Engenheiros Alemães, o planejamento da produção (PP) tem como propósito determinar os objetivos relevantes para a produção e as tarefas necessárias para alcançá-los (ZIPFEL; BRAUNREUTHER; REINHART, 2019). Em grande parte dos casos, o PP é examinado manualmente por especialistas, mesmo na produção moderna com a disponibilidade de sofisticados sistemas automatizados. Este processo manual torna-se demorado, subótimo (como apenas algumas alternativas são consideradas) e totalmente dependente da habilidade do planejador (FIGUEIRA et al., 2015). Esses desafios acima exigem que a indústria de manufatura seja capaz de atuar em seu PCP de forma flexível, que pode responder a situações imprevistas.

É necessário modelar e simular o PCP com a dinâmica da informação para analisar os riscos causados por múltiplas incertezas (YANG; ARNDT; LANZA, 2016). As necessidades de as empresas acompanharem as mudanças tecnológicas e de mercado para se manterem competitivas, promovem uma série de adequações nos processos de produção e, conseqüentemente, do PCP. Muitas metodologias estão sendo desenvolvidas a fim de auxiliar o PP ideal de uma indústria dada a sua complexidade (VRONTAKIS; KAMPIANAKIS; TSINARAKIS, 2016). Assim, desenvolver um processo confiável de previsão é o primeiro passo para a otimização do processo geral de planejamento e, assim, do sistema de produção (MATSUMOTO; IKEDA, 2015; MATSUMOTO et al., 2016).

No processo de PP para o PCP, a previsão de demanda tem um papel importante no planejamento e no cronograma de produção de uma empresa. As previsões devem ser as mais

precisas possíveis para permitir a otimização da produção, melhorando a programação e o controle de produção, evitando custos extras, satisfazendo as necessidades dos clientes no prazo e para que a oferta possa ser correspondida de perto com a demanda (MELO MENEZES et al., 2015; PRAKASH; RAMYA; YUGANDHAR, 2016; TIRKES; GURAY; CELEBI, 2017).

3.3 Previsão de Demanda

A problemática de prever observações futuras surge naturalmente em muitas áreas relevantes da pesquisa e modelagem estatística, e tem sido muito útil em vários campos de aplicação. Uma dessas áreas é a previsão de demanda, que pode ser definida como a melhor estimativa da empresa da demanda futura, dado um conjunto de suposições (MOON, 2018).

A previsão de demanda desempenha um papel importante nas operações das organizações modernas (FILDES et al., 2008; SYNTETOS et al., 2016). Ela suporta uma variedade de decisões de negócios, desde operacionais, táticas e até estratégicas, como capacidade planejamento (MIYAOKA; HAUSMAN, 2008), planejamento de recursos (BARROW, 2016; JALAL; HOSSEINI; KARLSSON, 2016), planejamento publicitário e promocional (MATSUMOTO; KOMATSU, 2015; DWEIRI; KHAN; JAIN, 2015; SYNTETOS; BABAI; GARDNER JR, 2015; TRAPERO; KOURENTZES; FILDES, 2015; MA; FILDES; HUANG, 2016), planejamento de demanda (TRAPERO; KOURENTZES; FILDES, 2012; AIT-ALLA et al., 2014; BARBOSA; CHRISTO; COSTA, 2015; ALBEY et al., 2015), analisando os efeitos da competição (MERINO; RAMIREZ-NAFARRATE, 2016), planejamento tático da produção (SAGAERT et al., 2018), entre outros. Além disso, é fase importante no planejamento de recursos da empresa. Do ponto de vista do suporte há outras funções de planejamento, especialmente no planejamento de produção (BÓNA; LÉNÁRT, 2014).

Os métodos de previsão de demanda podem ser baseados em modelos matemáticos e estatísticos que utilizam dados históricos (séries temporais) ou em métodos qualitativos, planejados de acordo com a experiência administrativa. Eles também podem ser baseados em uma combinação de métodos quantitativos (BATES; GRANGER, 1969; WINKLER; MARKRIDAKIS, 1983; BOX et al., 2015) ou com métodos qualitativos (FLORES; WHITE, 1988; LAWRENCE et al., 2006; LEWIS, 2012). Obter uma previsão de demanda precisa é o ponto crítico desse processo (JIANG; ZHANG; SONG, 2014).

A qualidade das previsões permite planejar melhor as operações da empresa, como produção, reposição de estoque e planejamento de transporte e planejamento de materiais e recursos humanos (MORETTIN; TOLOI, 2006; SCARPEL, 2015). Nesse contexto, quanto melhor a previsão de demanda, menor será os custos da cadeia produtiva. Por outro lado, o aumento dos

erros de previsão pode comprometer os custos de produção e a competitividade da organização empresarial (ARMSTRONG, 2006; WRIGHT; GOODWIN, 2009; PÉREZ-GALLARDO et al., 2014).

A previsão da demanda como problema de séries temporais tem como objetivo, partindo das observações históricas, identificar padrões e permitir projeções de futuro (QIN; LI; DU, 2017). Na essência, os métodos identificam os padrões da série, separando-o do ruído contido nas observações individuais, e usa para prever os valores futuros (HYNDMAN et al., 2008; BOX et al., 2015). Dessa forma, a previsão de séries temporais envolve incerteza, visto que o comportamento da série, não necessariamente, segue o padrão histórico (MONTGOMERY; JENNINGS; KULAHCI, 2015).

Ao longo da história da previsão de séries temporais, pode-se observar que diferentes abordagens teóricas foram propostas na literatura, podendo ser categorizados em duas classes principais: modelos individuais e combinados, com um aumento significativo dos estudos com interesse em projetar e implementar modelos que utilizem a teoria da combinação (WANG; LI; LI, 2018; HAJIRAHIMI; KHASHEI, 2019). Um dos principais motivadores para combinação é a busca em aumentar a acurácia e eficiência das previsões em relação a cada modelo individual usado separadamente. Essa observação vem do fato de que modelos individuais só podem modelar padrões específicos e dificilmente produzem resultados precisos, quando confrontados com problemas complexos em que nenhum deles pode obter a precisão de previsão desejada (GRANGER; RAMANATHAN, 1984; CLEMEN, 1989; DE MENEZES; BUNN; TAYLOR, 2000; TIMMERMANN, 2006). Assim, ao combinar diferentes modelos, a principal característica do modelo individual é extraída e um modelo abrangente (combinado) é desenvolvido, o qual pode superar as limitações dos modelos que o constitui e, por conseguinte, levar a um melhor desempenho (KIM; JEONG; KO, 2006; WANG et al., 2018).

3.4 Formalismos de Séries Temporais

Autores como Granger e Newbold (1986); Box et al. (2015); Montgomery, Jennings e Kulahci (2015); Hyndman e Athanasopoulos (2018), descrevem uma série temporal como sendo uma sequência de observações ordenadas no tempo, formada a partir de observações medidas de um fenômeno, como: quantidade de chuva anual no Ceará; número mensal de acidentes de trânsito no município de Várzea Alegre; quantidade de atendimentos médicos por semana na região do Cariri; valor diário da cotação de uma ação, quantidade diária de casos de COVID-19 em Juazeiro do Norte, entre outros. Do ponto de vista estatístico, uma série de dados históricos pode ser tratada como uma sequência de instâncias de variáveis aleatórias,

podendo ser mensuradas de forma contínua ou discreta (FIORUCCI et al., 2016).

Assim, a série temporal pode então ser entendida como a observação de um processo estocástico ao longo do tempo discreto (SERGIO, 2017). Uma série temporal univariada, diga-se U_t , é formalizada como uma sequência de observações (Equação 3.1).

$$U_t = \{u_t \in \mathbb{R} \mid t = 1, 2, 3, \dots, n\}, n \in \mathbb{N}, \quad (3.1)$$

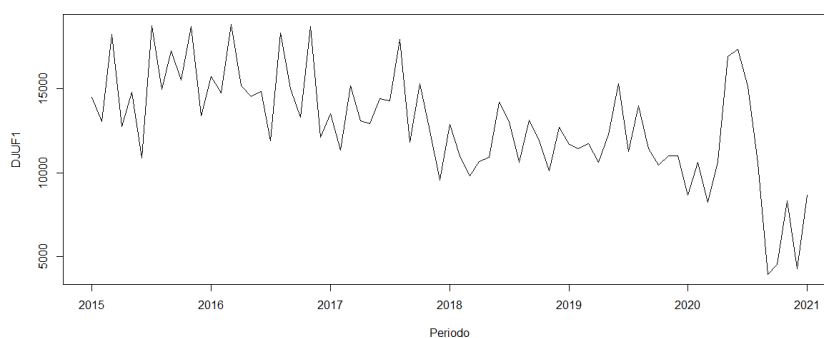
em que u_t é o valor da série temporal no tempo t e n o número de observações da série temporal regressa.

Prever uma série temporal implica em inferir seu valor futuro (BOX et al., 2015), simbolizado genericamente pela Equação 3.2.

$$\hat{u}_{t+1} = x_i(u_t, u_{t-1}, \dots, u_1) \quad (3.2)$$

em que \hat{u}_{t+1} é o valor a ser previsto para a série no instante $t + 1$, a partir dos valores conhecidos atual e passados (u_t, u_{t-1}, \dots, u_1) e x_i é o i^o modelo preditor utilizado, uma função matemática da série histórica. Cada observação u_t pode ser considerada como um ponto em um gráfico bidimensional, onde o eixo das ordenadas determinam as medições dos dados e o eixo das abscissas delimitam em que momento discreto do tempo tais medições foram mensuradas. A Figura 9 ilustra um exemplo de série temporal da família de produtos sabão em barra (ver Tabela 1 na Seção 2.2.4). Essa série apresenta 73 observações de registro mensal de demanda da família de produtos da empresa objeto do caso de estudo, considerando diversos produtos, entre o período de janeiro de 2015 e abril de 2021.

Figura 9: Série temporal mensal de demanda da família de produtos (DJUF1) sabão em barra, para período de 2015 a 2021



Fonte: Elaborado pelo autor.

As previsões baseadas em séries temporais presumem que os fatores que influenciam as observações no passado e no presente tenderão a influenciá-las no futuro. Com isso, é possível

a identificação de padrões relevantes, como tendência, sazonalidade e aleatoriedade (HYNDMAN; ATHANASOPOULOS, 2018). A tendência pode ser definida como a direção de crescimento ou decréscimo da série em um certo período de tempo, enquanto a sazonalidade descreve padrões comportamentais regulares ao longo do tempo. Já a aleatoriedade é a parte da série temporal não explicada pela modelagem aplicada (KITAGAWA, 2010; PEÑA; TIAO; TSAY, 2011; GUJARATI; PORTER, 2011). Esses componentes permitem entender e inferir os fenômenos estocásticos estudados.

A análise de séries temporais pode envolver várias modalidades, como a análise exploratória (ANDRIENKO; ANDRIENKO, 2006), o ajuste de curva (BRADLEY et al., 2007), a classificação (XI et al., 2006), a análise de entropia (STOSIC et al., 2016), a extração de características (CAVALCANTE; MINKU; OLIVEIRA, 2016) e a segmentação (KEOGH et al., 2004). Na literatura existem diversos modelos de previsão de séries temporais disponíveis para analisar os dados de demanda (NAIM; MAHARA; IDRISI, 2018). Vários estudos examinaram, a partir da análise de séries temporais, a demanda na manufatura (SCHOLZ-REITER; KÜCK; LAPPE, 2014; ALIYEVA, 2017) ou remanufatura (MATSUMOTO; KOMATSU, 2015; MATSUMOTO et al., 2016), outros apresentam resultados da qualidade das previsões de demanda (MATSUMOTO; IKEDA, 2015; MATSUMOTO et al., 2016; VARGAS; CORTÉS, 2017; GOODWIN; GÖNÜL; ÖNKAL, 2019). Para Jatta e Krishnan (2016), muitas empresas utilizam séries temporais como base de suas previsões de demanda e para planejar sua produção.

3.4.1 Formalismos Individuais

Nesta seção são apresentados alguns formalismos de previsão ditos individuais. Eles foram escolhidos a partir da revisão sistemática de literatura, abordando suas características, bem como aplicações relevantes na manufatura de bens.

3.4.1.1 Modelos Box & Jenkins

Proposto na década de 70 e publicado por George Box e Gwilyn Jenkins, no livro intitulado *Time series analysis: Forecasting and control*, deram um novo alento à pesquisa e conduziram a uma nova geração de ferramentas de previsão. O modelo Autoregressivo Integrado de Média Móvel (ARIMA), pode ser usado para identificar padrões complexos em dados e gerar previsões (BOX; JENKINS, 1976; BOX et al., 2015). O método prevê um valor em uma série temporal de resposta como uma combinação linear de seus próprios valores e resíduos passados. Assumem que os erros são ruídos brancos por natureza e exigem que os dados sejam

estacionários antes de ajustar uma equação linear (QIN; LI; DU, 2017).

Autores como Newbold e Granger (1974), Pankratz (2009) e Montgomery, Jennings e Kulahci (2015), salientam que o modelo ARIMA tem origem nos Modelos Autorregressivo (AR), Médias Móveis (MA) e na combinação destes para gerar o Modelo Autorregressivo de Médias Móveis (ARMA). O modelo $ARMA(p, q)$ genérico, incorpora as ideias dos modelos $AR(p)$ e $MA(q)$ em uma forma compacta, reduzindo a quantidade de parâmetros. As Equações 3.3, 3.4 e 3.5 descrevem, nesta ordem, os modelos $AR(p)$, isto é, $ARIMA(p, 0, 0)$, os modelos $MA(q)$, isto é, $ARIMA(0, 0, q)$, e os modelos autorregressivo de médias móveis $ARMA(p, q)$, isto é, $ARIMA(p, 0, q)$.

$$U_t = \gamma_1^{ar} U_{t-1} + \dots + \gamma_p^{ar} U_{t-p} + \epsilon_t, \quad (3.3)$$

$$U_t = \epsilon_t - \gamma_1^{ma} \epsilon_{t-1} - \dots - \gamma_q^{ma} \epsilon_{t-q}, \quad (3.4)$$

$$U_t = \gamma_1^{ar} U_{t-1} + \dots + \gamma_p^{ar} U_{t-p} + \epsilon_t - \gamma_1^{ma} \epsilon_{t-1} - \dots - \gamma_q^{ma} \epsilon_{t-q}, \quad (3.5)$$

em que γ_i^{ar} , γ_j^{ma} e ϵ são, respectivamente, os parâmetros $AR(p)$, os parâmetros $MA(q)$ e o erro aleatório. Segundo Cryer e Chan (2008) e Box et al. (2015), a aplicação dos modelos AR, MA e ARMA restringe-se a séries temporais estacionárias. Entretanto, diferenciando uma série temporal d vezes, a tendência da mesma é tornar-se estacionária e ajustável pelo modelo $ARMA(p, q)$, da origem ao modelo $ARIMA(p, d, q)$, ou seja, uma série temporal, em que p denota os números dos termos autorregressivos, d a quantidade de vezes que a série deve ser diferenciada antes de tornar-se estacionária e q o número de termos de média móvel. Assim, a partir da realização de diferenciações, tem-se o modelo $ARIMA(p, d, q)$ Equação, 3.6,

$$\hat{u}_t^d = \gamma_0 + \sum_{i \in I_u} \gamma_i^{ar} u_{t-t_i}^d + \sum_{j \in I_{\hat{u}}} \gamma_j^{ma} e_{t-t_j}^d, \quad (3.6)$$

em que, \hat{u}_t^d é o modelo ajustado a u_t^d ; γ_0 é o intercepto do modelo ARIMA; $\{\gamma_i^{ar}\}_{i \in I_u}$ é o conjunto de coeficientes que ponderam as defasagens temporais referentes à série diferenciada; d é o número de diferenciações para que u_t se torne estacionária; $\{\gamma_j^{ma}\}_{j \in I_{\hat{u}}}$ é o conjunto de coeficientes que ponderam as defasagens temporais referentes aos resíduos do modelo; e_t^d é o resíduo correspondente ao modelo ajustado a u_t^d .

Com base na Equação 3.6, os componentes $AR(p)$ e $MA(q)$ do modelo são constituídos a partir de somatórios. Com relação às diferenciações, tem-se que:

$$u_t^0 = u_t; u_t^1 = u_t - u_{t-1}; \dots; u_t^d = u_t^{d-1} - u_{t-1}^{d-1} \quad (3.7)$$

Para Anderson (1977) e Montgomery, Jennings e Kulahci (2015), o modelo desenvolvido serve não só como método de previsão, mas também para explicar o processo de geração da série. Os autores Cryer e Chan (2008), Pankratz (2009) e Shumway e Stoffer (2011), generalizaram o modelo ARIMA e definiram o ARIMA sazonal, conhecido como SARIMA(p, d, q)(P,D,Q)[s], sendo este aplicado a séries temporais que apresentam padrões de repetição em determinado intervalo de tempo (sazonalidade).

De acordo com Torbat, Khashei e Bijari (2018), o modelo ARIMA provou ser mais aplicável a problemas relacionados a previsão de séries temporais e que as propriedades estatísticas dos modelos são a principal causa de sua popularidade. No entanto, embora possuam a vantagem de produzir estimativas precisas em períodos curtos com relativa facilidade de implementação, sofrem com outras dificuldades que reduzem sua popularidade na previsão de séries temporais com mais variabilidade.

Pesquisas apresentam resultados relevantes com a aplicação do modelo ARIMA, na busca de prever a demanda para o PCP. Autores como Matsumoto e Komatsu (2015) e Junior e Godinho Filho (2016), usaram no PCP na manufatura e remanufatura; Melo Menezes et al. (2015) e Rosienkiewicz, Chlebus e Detyna (2017), na previsão de peças de reposição para indústria; Vargas e Cortés (2017), utilizaram na previsão da demanda na indústria automotiva, considerada estratégica no setor industrial e desenvolvimento econômico do país; Stüker (2014) e Mircetic et al. (2016), no planejamento e programação de atividades logísticas dentro da cadeia de suprimentos e Babai et al. (2013), na previsão e desempenho de inventário.

3.4.1.2 Modelos de Suavização Exponencial

A família dos modelos de Suavização Exponencial (ETS) são técnicas de previsão amplamente utilizadas na análise de séries temporais univariadas, devido sua simplicidade e robustez como mecanismos de previsão automática (HYNDMAN et al., 2002; GARDNER JR, 2006; WU; LIU; YANG, 2016). O modelo foi proposto no final da década de 1950 por Brown (1959), Winters (1960) e Holt (2004) e motivou alguns dos métodos de previsão mais bem-sucedidos (YAFFEE; MCGEE, 2000; HYNDMAN et al., 2008; HYNDMAN; ATHANASOPOULOS, 2018). Para Corberán-Vallet, Bermúdez e Vercher (2011), a forma geral da função de previsão de suavização exponencial, envolvendo um conjunto de coeficientes adaptativos, foi dada, possivelmente, pela primeira vez, por Box e Jenkins (1976).

O método consiste em procedimentos tradicionais para revisar continuamente uma previsão à luz de informações mais recentes sobre os dados estimados. Os modelos são médias ponderadas que atribuem pesos exponencialmente decrescentes à medida em que a observa-

ção se torna mais antiga, ou seja, observações recentes recebem um peso relativamente maior na previsão do que as mais antigas (GARDNER JR, 1985; MAKRIDAKIS; HIBON, 2000; GARDNER JR, 2006). Para Hyndman e Athanasopoulos (2018), esse *framework* produz previsões confiáveis rapidamente e para uma ampla gama de séries temporais, o que é uma grande vantagem, e extremamente importante para aplicações na indústria de manufatura.

Partindo do modelo introduzido inicialmente na década de 50, Holt (1957), estendeu a suavização exponencial linear para permitir a previsão de dados com tendências, e também propôs um método para dados sazonais. Seu método foi estudado por Winters (1960), hoje conhecido como método de Holt-Winters. Posteriormente, foram adaptados para séries temporais com tendência multiplicativa, Pegels (1969), tendência de amortecimento, Gardner Jr (1985); Taylor (2003) e séries temporais de demanda intermitente (SYNTETOS; BOYLAN, 2001; WALLSTRÖM; SEGERSTEDT, 2010). Seu bom desempenho estatístico foi estabelecido quando modelos estocásticos de espaço de estado foram introduzidos (GARDNER JR, 2006; HYNDMAN et al., 2008; MCKENZIE; GARDNER JR, 2010). Posteriormente, os métodos foram adaptados para séries temporais com múltiplos períodos sazonais ou sazonalidade (TAYLOR, 2003; DE LIVERA, 2010; DE LIVERA; HYNDMAN; SNYDER, 2011).

Os modelos de suavização são baseados em suposições de que as séries temporais são construídas a partir de componentes não observados, como nível, tendência e efeitos sazonais. Esses componentes devem ser adaptados ao longo do tempo para seguir o padrão de demanda (WU; LIU; YANG, 2016). Esse acrônimo ETS se refere tanto à *Exponential Smoothing* como aos componentes de Erro (E), Tendência (T) e Sazonalidade (S), que definem um modelo dentro da família ETS. Pode-se incorporar a sazonalidade aditiva ou multiplicativa. As variantes aditiva e multiplicativa dependem se a variação sazonal é considerada independentemente do nível da média local ou proporcional a ela (BERGMEIR; HYNDMAN; BENÍTEZ, 2016).

Os modelos produzidos a partir das combinações podem ser considerados aditivos, apropriado para séries com tendência temporal linear e variação sazonal aditiva (Equação 3.8), ou multiplicativo, adequado para séries com tendência temporal linear e variação sazonal multiplicativa (Equação 3.9), sendo possível outras variações (Equação 3.10).

$$u_t = T + S + \epsilon_t \quad (3.8)$$

$$u_t = T \cdot S \cdot \epsilon_t \quad (3.9)$$

$$u_t = (T + S) \cdot \epsilon_t, \quad (3.10)$$

O componente sistemático T é dado pela combinação entre um termo de nível (l) e um termo de crescimento (b_s), combinando estes, da origem a outros tipos diferentes de tendência

futura como: inexistência, aditiva, aditiva amortecida, multiplicativa e multiplicativa amortecida (Equação 3.11-3.15) (HYNDMAN et al., 2008).

$$T_h = l \quad (3.11)$$

$$T_h = l + b_s h \quad (3.12)$$

$$T_h = l + (\phi + \phi^2 + \dots + \phi^h) b_s \quad (3.13)$$

$$T_h = l b_s^h \quad (3.14)$$

$$T_h = l b_s^{(\phi + \phi^2 + \dots + \phi^h)} \quad (3.15)$$

em que, T_h é considerada a previsão da tendência para o próximo período h e ϕ é o parâmetro de amortecimento, compreendido entre 0 e 1. Assim, após a estimação de T , os componentes S (nenhum, aditivo ou multiplicativo) e E (aditivo ou multiplicativo) são selecionados. Essa classificação foi proposta por McCormick (1969) e na sequência modificada por Gardner Jr (1985), Hyndman et al. (2002, 2008), Taylor (2003), quando a notação ETS foi então introduzida.

A família de modelos ETS podem ser diferenciados pelo tipo de erro, tendência e sazonalidade que cada um utiliza. Geralmente, a tendência pode ser classificada como inexistente, aditiva, multiplicativa, aditiva amortecida ou multiplicativa amortecida. A sazonalidade pode ser inexistente, aditiva ou multiplicativa. O erro pode ser aditivo ou multiplicativo; no entanto, a diferença entre essas duas opções passa a ser relevante somente para intervalos de previsão, não para previsões pontuais. Com isso, um total de quinze modelos com diferentes combinações de tendência e sazonalidade são mostrados no Quadro 3. Assim, como o erro pode ser expresso de forma aditivo ou multiplicativo, um total de trinta modelos podem ser estruturados (HYNDMAN et al., 2002, 2008). Desta forma, um modelo ETS(A,M,N) tem erro aditivo, tendência multiplicativa e nenhuma sazonalidade, por exemplo. As diversas combinações de tendência e sazonalidade são apresentadas no Quadro 3.

Os métodos mais comuns, desconsiderando o erro, são os seguintes: suavização exponencial simples (ausência de tendência e sazonalidade), método linear de Holt (com tendência aditiva e sem sazonalidade), aditivo de Holt-Winters (com tendência e sazonalidade aditiva) e multiplicativo de Holt-Winters (com tendência aditiva e sazonalidade multiplicativa) (McCORMICK, 1969; HYNDMAN et al., 2008; HYNDMAN; ATHANASOPOULOS, 2018).

Diferentes aplicações e resultados importantes são encontrados na literatura com o uso dos métodos ETS, como na previsão de demanda para o planejamento da produção, controle de estoque, demanda ruidosa, demanda intermitente, cadeia de suprimentos (GRIMAUD; DOL-

Quadro 3: A família de modelos ETS, com diferentes tipos de tendência e sazonalidade, de acordo com Hyndman (HYNDMAN et al., 2008)

Componente de Tendência	Componente Sazonal		
	N (Nenhum)	A (Aditivo)	M (Multiplicativo)
N (Nenhum)	NN	NA	NM
A (Aditivo)	AN	AA	AM
Ad (Aditivo amortecido)	AdN	AdA	AdM
M (Multiplicativo)	MN	MA	MM
Md (Multiplicativo amortecido)	MdN	MdA	MdM

Fonte: Elaborado pelo Autor.

GUI; KORYTKOWSKI, 2014; SCARPEL, 2015; BARBOSA; CHRISTO; COSTA, 2015; SYNTETOS; BABAI; GARDNER JR, 2015; TRATAR; MOJŠKERC; TOMAN, 2016; PRAK; TEUNTER; SYNTETOS, 2017; SONI; SRIKANTH, 2017; VARGAS; CORTÉS, 2017; KITTI-UDOMPORN; SUKCHAREONPONG; WITCHAKUL, 2018). Os resultados obtidos podem ajudar a demonstrar a importância de melhorar as metodologias de previsão.

3.4.1.3 Redes Neurais Artificiais

Uma rede neural artificial (ANN), do inglês *Artificial Neural Networks*, é um mecanismo de aprendizado de máquina (*Machine Learning*), que por sua vez, é um subconjunto da inteligência artificial que imita basicamente como o cérebro humano aprende, ou seja, uma espécie de inteligência cognitiva. O cérebro recebe um estímulo, processa-o e gera uma resposta. À medida em que a tarefa se torna complicada, vários neurônios formam uma rede complexa, transmitindo informações entre si. Usando uma rede neural artificial, tenta-se imitar um comportamento semelhante (ZHANG, 2020).

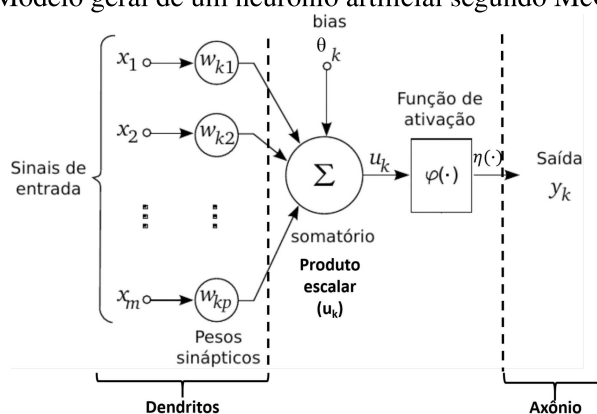
As ANNs são métodos de previsão baseados em modelos matemáticos cuja construção se inspira nos processos biológicos do cérebro humano, propostas por McCulloch e Pitts em 1943, que procuraram simular o funcionamento do cérebro (HAYKIN, 1994; SEGALL, 1995; HYNDMAN; ATHANASOPOULOS, 2018). São flexíveis e com capacidade de processar dados não lineares complexos entre a previsão (variável de resposta) e seus preditores (variável de entrada), acionados por informações, que possuem propriedades interessantes para a previsão (KOURENTZES, 2013). As ANNs possuem características como: (i) capacidade de aprendizado por meio de treinamento; (ii) adaptação por experiência, em que os pesos das sinapses se ajustam conforme a entrada de dados; (iii) habilidade de generalização, já que a rede consegue aferir uma solução mesmo que não tenha sido apresentada certa entrada no treinamento; (iv) a implementação pode ser realizada utilizando diferentes ferramentas computacionais.

O modelo geral de um neurônio artificial, baseado em McCulloch e Pitts, representado na Figura 10, recebe observações (sinais) de entrada x_i combinados por meio do produto escalar com seus pesos w_i , aplica a função φ , para produzir seu estado de ativação. Na sequência, aplica-se o valor resultante da etapa anterior à função η para produzir o valor de saída do neurônio. A soma ponderada dos dados de entrada do neurônio k produz o valor efetivo de entrada (Equação 3.16).

$$u : u_k = \theta_k + \sum_{i=1}^p w_{ki} \cdot x_i \quad (3.16)$$

em que, p é o número de sinais de entrada do neurônio; x_i é o i -ésimo sinal de entrada do

Figura 10: Modelo geral de um neurônio artificial segundo McCulloch e Pitts



Fonte: Adaptado de HAYKIN (2009).

neurônio; w_{ki} é o peso associado com o i -ésimo sinal de entrada, no neurônio k ; θ_k é o limiar do neurônio k (bias); u_k é uma combinação linear dos sinais de entrada e do bias, no k -ésimo neurônio; y_k é o neurônio de saída e $\varphi(\cdot)$ é a função de ativação do k -ésimo neurônio. Neste modelo de ANN, as entradas x_m do neurônio podem ser saídas de outros neurônios, entradas externas, um viés ou qualquer combinação destes elementos. No neurônio biológico, esse último papel é a do axônio, o qual distribui as informações processadas.

Para Giacomini (2003), o nome de confluência é dado à combinação ponderada das entradas de um neurônio. O uso do valor auxiliar θ_k (bias) tem o efeito de ajustar a saída u_k do combinador linear.

Um aspecto importante na implementação de ANNs é o treinamento, usado com a finalidade para que a rede aprenda uma tarefa com sucesso, consistindo no ajuste dos pesos sinápticos de forma iterativa sob algumas regras de aprendizagem, até que ela extraia as características para representar as informações fornecidas. Para o treinamento das ANNs existem vários algoritmos de treinamento que diferem praticamente pelo modo como é realizado o

ajuste dos pesos, dividindo-se em dois tipos, aprendizado supervisionado (existe um supervisor externo que fornece à rede tanto valores de entrada quanto os valores de saída esperados) e aprendizado não supervisionado (fornece à rede somente valores de entrada que são organizados em diferentes classes de acordo com a semelhança de suas propriedades através do ajuste de seus pesos da rede) (HAYKIN, 2010; PRECHELT, 2012). No presente documento de tese recorreu-se ao aprendizado supervisionado.

Existem diferentes métodos para a determinação do momento em que o treinamento de uma ANN deve ser finalizado. Os critérios de parada mais utilizados são o número de ciclos, o erro e a parada antecipada. O número de ciclos define a quantidade de vezes em que o conjunto de treinamento é apresentado à rede. Um número elevado de ciclos pode conduzir a rede à perda do poder de generalização (sobre ajuste, do inglês *overfitting*). Por outro lado, com um pequeno número de ciclos a rede pode não chegar ao seu melhor desempenho (falta de ajuste, do inglês *underfitting*). A parada proveniente do tamanho do erro significa finalizar o treinamento após o erro identificado pela rede e o real ficar abaixo de um valor pré-definido. Na parada antecipada, conhecida em inglês como *early stopping*, o treinamento é interrompido a cada número de ciclos e é realizada uma estimação do erro da rede sobre o conjunto de dados de validação (PRECHELT, 2012).

As funções de ativação são classificadas em: parcialmente diferenciáveis (função degrau, degrau bipolar e rampa simétrica) ou diferenciáveis (função logística, tangente hiperbólica, linear e gaussiana). As funções parcialmente diferenciáveis são mais utilizadas em redes como *Perceptron* e *Adaline*, ao passo que as diferenciáveis são aplicadas em Redes de Base Radial (RBF), Perceptron de Múltiplas Camadas (MLP), entre outras (HAYKIN, 1994, 2007, 2010).

As ANNs possuem diversas características como auto-organização, adaptação e autoaprendizagem, além de não-localidade, não-estacionaridade e não-convexidade, tornando-a uma ferramenta poderosa para resolver problemas complexos (AZADEH; GHADERI; SOHRABKHANI, 2008). Pode ser usada na resolução de problemas de otimização, controle de processo, problemas de classificação, reconhecimento de padrões e modelagem de séries temporais (DANIEL, 2007; KHASHEI; BIJARI, 2012; SZOPLIK, 2015).

Uma rede neural pode ser organizada em camadas. A camada de entrada (ou preditores) forma a camada inferior e a camada de saída (ou previsões), forma a camada superior. Também, pode haver camadas intermediárias contendo “neurônios ocultos” (HYNDMAN; ATHANASOPOULOS, 2018). As principais arquiteturas de redes podem ser divididas em: Rede neural de arquitetura com camada única alimentada adiante; Rede neural de arquitetura de múltiplas camadas alimentadas adiantes; Rede neural de arquitetura recorrente, *feedback* ou retroalimentação e Rede neural em malha. A presente tese utilizou a Rede neural de ar-

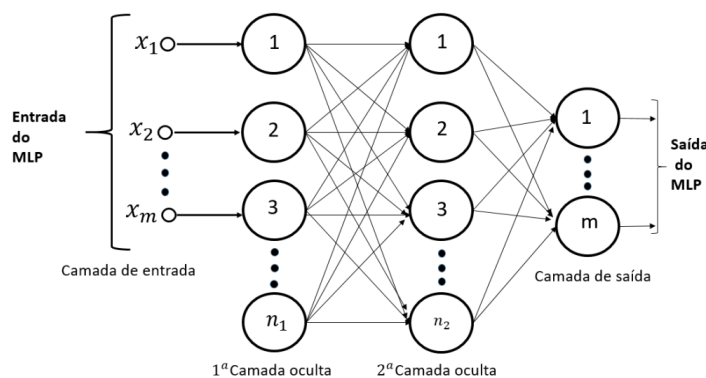
quitetura, tipo “MLP”, redes que usam, na sua arquitetura, múltiplas camadas alimentadas adiante (HAYKIN, 2007; NUNES; DA SILVA, 2018).

3.4.1.3.1 Perceptron de Múltiplas Camadas

Uma rede perceptron de múltiplas camadas do inglês (MLP, *multilayer perceptron*), é uma rede que possui, pelo menos, uma camada intermediária, conhecida como camada de neurônios oculta ou escondida. Ou seja, a rede consiste de um conjunto de unidades sensoriais (nós de fonte) que constituem a camada de entrada, uma ou mais camadas ocultas de nós computacionais e uma camada de saída de nós computacionais. O sinal de entrada se propaga para frente através da rede, camada por camada (HAYKIN, 2007). Para Cybenko (1989), uma rede com uma camada intermediária pode implementar qualquer função contínua. A utilização de duas camadas intermediárias permite a aproximação de qualquer função matemática (CYBENKO, 1988).

A MLP é considerada uma das arquiteturas de rede mais versáteis, podendo ser encontrada em diversos sistemas, como de processamento de sinais, função universal de aproximação (ajuste de curva), reconhecimento de padrões, identificação de processo e controle, otimização de sistema e previsão de séries temporais (NUNES; DA SILVA, 2018). A Figura 11 apresenta a arquitetura de uma rede MLP, com x_m entradas regressas de uma série temporal e duas camadas intermediárias.

Figura 11: Arquitetura de uma rede MLP com duas camadas intermediárias



Fonte: Adaptado de (HAYKIN, 2007).

A informação desse tipo de topologia de rede é passada em sentido único, cuja arquitetura é chamada de *feedforward*. Seu treinamento de forma supervisionada é realizado com um algoritmo conhecido como retropropagação de erro (*error back-propagation*), cujo objetivo é encontrar o mínimo local ou global da função do erro entre entrada e a saída. Este algoritmo é

baseado na regra de aprendizagem por correção de erro. Isto é, o algoritmo utiliza a técnica do gradiente descendente direcionando o ajuste dos pesos (HAYKIN, 2007; PÁDUA BRAGA; LEON FERREIRA; LUDERMIR, 2007).

A aprendizagem por retropropagação de erro consiste de dois passos através das diferentes camadas da rede: um passo para frente (*forward*), a propagação, onde os pesos sinápticos da rede são todos fixos; e um passo para trás (*backward*), a retropropagação, onde os pesos sinápticos são todos ajustados de acordo com uma regra de correção de erro. No passo para frente, um padrão de atividade (vetor de entrada) é aplicado aos nós sensoriais da rede e seu efeito se propaga através da rede, camada por camada. Com isso, um conjunto de saídas é produzido como a resposta real da rede que é subtraída de uma resposta desejada, produzindo um sinal de erro. Por sua vez, este sinal é propagado para trás (retropopagação do erro) contra a direção das conexões sinápticas. Com isso, os pesos sinápticos são ajustados, fazendo com que a resposta real se aproxime da desejada, em um sentido estatístico (HAYKIN, 2010).

A unidade básica de processamento de informações do sistema nervoso central humano é o neurônio (\mathcal{N}), composto por três elementos: um conjunto de dendritos (\mathcal{I}), um corpo celular (\mathcal{S}) e um axônio (\mathcal{A}), cada um com funções específicas, porém complementares (HAYKIN, 2007; DANIEL, 2013). De forma simplificada, esses elementos trabalham ações em série, a fim de fornecer respostas a um conjunto de estímulos (sinais): (i) \mathcal{I} inicialmente recebe os estímulos (sinais) externos a \mathcal{S} (identificados como sendo os terminais ou camada de entrada de \mathcal{N}); (ii) em \mathcal{S} os estímulos são ponderados (combinados) por pesos acoplados em \mathcal{S} , realizando operações agregadas simples; e (iii) o axônio \mathcal{A} processa a informação fornecida por \mathcal{S} , com operações eventualmente mais sofisticadas, gerando as respostas de \mathcal{N} aos estímulos recebidos por \mathcal{I} .

As respostas de um neurônio podem ser o estímulo de outro, chamado de sinapse. É pelas sinapses que os nós se unem funcionalmente, formando redes neurais. Essa topologia de ANN, comumente utilizada para previsão tem sido eficiente, principalmente, na previsão de séries temporais com padrões não lineares, caso em que outros modelos de previsão não são eficazes.

No contexto de utilização das ANNs, para o caso específico de séries temporais u_t , \mathcal{N} (neurônio) é modelado pela rede perceptron (\mathcal{P}). No caso, u_t envolve

$$\mathcal{I}_t = \{u_{t-t_i}\}_{i \in I_u}, \quad (3.17)$$

isto é, os valores anteriores da série e o modelo.

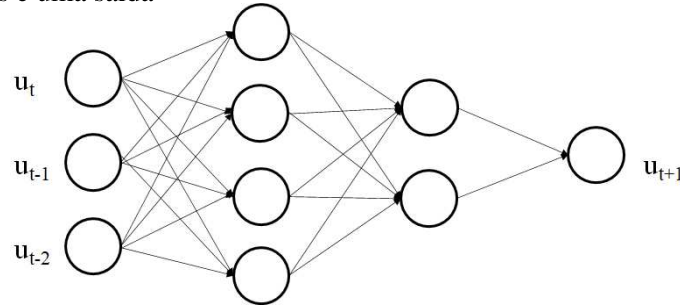
$$\mathcal{S}(\mathcal{I}_t) = \sum_{i \in I_u} \alpha_i u_{t-t_i} + \alpha_0, \quad (3.18)$$

em que α_i é o coeficiente que pesa u_{t-t_i} na camada de entrada de uma ANN (com α_0 sendo a correção de viés ou o intercepto do modelo). Assim, $\mathcal{S}(\cdot)$ envolve uma combinação linear ponderada dos dados de entrada. Além disso, refletindo a resposta à série de dados regressos, tem-se a previsão

$$\hat{u}_t = \mathcal{A}(\mathcal{S}(\mathcal{I}_t)) \quad (3.19)$$

O ajuste $\mathcal{A}(\cdot)$ é chamado de função de ativação e pode assumir diversos modelos de funções como, discreta, contínua, linear ou não linear. Assim, as Equações 3.17 - 3.19 apresentam relações que resumem a aproximação de \mathcal{N} via \mathcal{P} . A Figura 12 ilustra uma rede neural, com duas camadas intermediárias, para duas entradas regressas de uma série temporal.

Figura 12: Arquitetura de Rede Neural adaptada para séries temporais com duas entradas regressas, duas camadas ocultas e uma saída



Fonte: Elaborado pelo autor.

Com o objetivo de melhorar o desempenho dos preditores de ANN de séries temporais, vários pesquisadores adicionaram à arquitetura das redes neurais camadas intermediárias com neurônios ocultos (conjunto de perceptrons), digam-se $\{\mathcal{P}_h\}_{h=1}^H$, entre as camadas de entrada e saída da ANN (HYNDMAN; ATHANASOPOULOS, 2018). Conforme visto, isso é o que define uma rede *feed-forward* multicamadas, onde cada camada de nós recebe entradas das camadas anteriores. Tal arquitetura leva a uma sequência de somas simples,

$$\mathcal{S}_h(\mathcal{I}_t) = \sum_{i \in I_u} \alpha_{hi} u_{t-t_i} + \alpha_{h0}, \quad (3.20)$$

e respectivas funções de ativação intermediárias,

$$\mathcal{A}_h(\mathcal{S}_h(\mathcal{I}_t)) \quad (3.21)$$

Com isso, essas funções intermediárias de ativação devem, por sua vez, ser operadas por uma função final de ativação, dedicada à previsão u_t :

$$\hat{u}_t = \mathcal{A}_o \left(\sum_{h=1}^H \theta_{oh} + \mathcal{A}_h(\mathcal{S}_h(\mathcal{I}_t)) \right) \quad (3.22)$$

As funções de ativação têm o papel de informar às camadas anteriores os erros cometidos pela rede com a maior precisão possível (HAYKIN, 2010). As funções de ativação mais comuns são a Linear, Sigmóide Logística e a Tangente Hiperbólica. A notação aqui adotada para modelos de redes neurais será ANN(NI,No,Nc), com NI indicando o número de entradas autoregressivas (autoregressivas e autoregressivas sazonais $P_{AR} + P_{ARS}$), No é a quantidade de nós na camada oculta e Nc é a quantidade de camadas ocultas. A saída, para a abordagem desse trabalho é sempre igual a 1, uma vez que se trabalha com a previsão de um passo à frente (*one-step-ahead*).

Outras características importantes das ANNs é que elas realizam processamento paralelo da informação fornecida como entrada, envolvendo características anexas como uma taxa de aprendizado e um algoritmo de aprendizado (ZHANG, 2003; HAYKIN, 2007). Alterações nessas variáveis influenciam diretamente o desempenho dos modelos resultantes (CIANCIO et al., 2016; TANG et al., 2018). Contudo, o trabalho de determinar essa arquitetura tem sido uma das questões mais importantes para pesquisadores de ANNs. Normalmente, um procedimento de tentativa e erro é realizado para determinar esses parâmetros (MADIĆ; RADOVANOVIĆ, 2011). Assim, o custo computacional relacionado à construção da ANN ótima passa a ser alto. Já o uso de técnicas de otimização tem sido útil na potencialização desses preditores de séries temporais (HOU; ZHAO; LU, 2018; JIANG; YANG; HENG, 2019).

Zhang, Patuwo e Hu (1998), listam várias aplicações de previsão onde as ANNs foram empregadas com sucesso e Adya e Collopy (1998), constataram que dos 73% dos artigos revisados, as ANNs tiveram bom desempenho, superando os *benchmarks* estabelecidos. Em trabalhos como previsão anual da demanda de gasolina no Irã (KAZEMI et al., 2009), demanda urbana de água (GHIASSI; ZIMBRA; SAIDANE, 2008), demanda na indústria de mineração (ROSIENKIEWICZ; CHLEBUS; DETYNA, 2017), demanda de peças de reposição (MELO MENEZES et al., 2015), as ANNs foram aplicadas como uma técnica de previsão de demanda.

3.4.1.4 Máquinas de Vetores de Suporte

A teoria da Máquina de Vetores de Suporte, do Inglês *Support Vector Machines* (SVMs), foi proposta pela primeira vez por Vapnik na década de 90 (VAPNIK; VAPNIK, 1998; DANESHDOOST et al., 1998; XU; CHAN; ZHANG, 2019). É uma técnica de aprendizado de máquinas que utiliza o princípio da indução, na obtenção de conclusões genéricas, a partir de um conjunto de informações particular (MITCHELL; LEARNING, 1997). Desde então, tem sido bastante utilizado na aprendizagem estatística, tarefas de classificação e análise de regressão para o caso de séries temporais, por produzir bons modelos e com menor demanda computacional (CRISTIANINI; SHAWE-TAYLOR et al., 2000; BLANCHARD et al., 2008).

O SVM foi originalmente desenvolvido com o objetivo de resolver problemas de classificação. Uma versão estendida e adaptada do SVM, conhecida como regressão por vetores de suporte, também foi desenvolvida para problemas complexos de regressão. As máquinas de vetores suporte para regressão, diferem da técnica de classificação no sentido de que, enquanto a segunda busca dividir os dados em diferentes classes e classificar os dados futuros, a primeira busca encontrar um preditor que aproxime bem os dados de amostra. A ideia básica do procedimento é considerar, para cada amostra x_i da regressão (Figura 13a) um número positivo qualquer aqui definido por d_s , que é adicionado e subtraído do respectivo valor correspondente de interesse y_i (Figura 13b) formando dois novos conjuntos.

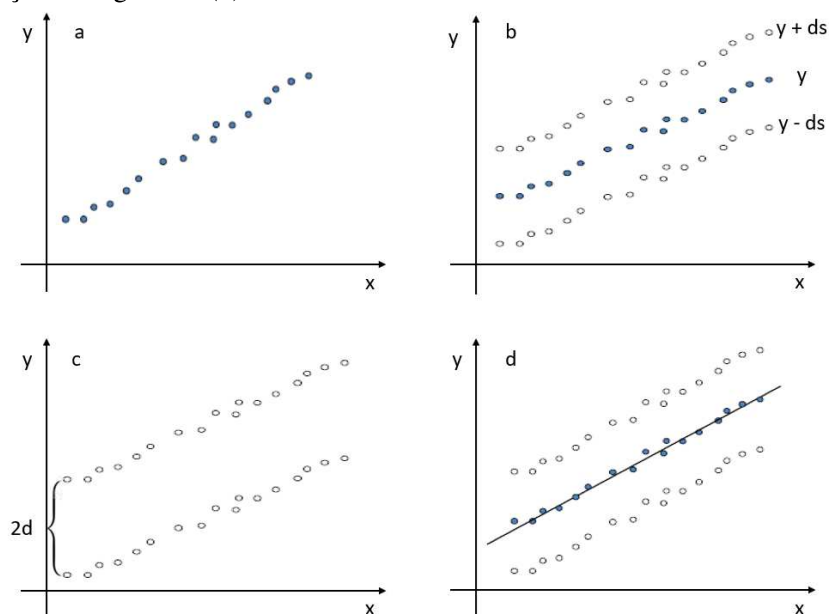
Observando a Figura 13c pode-se ver, facilmente, a formação de duas classes, uma positiva e outra negativa. Assim, o hiperplano de máxima separação obtido para as classes formadas passará exatamente pelos valores de y_i , pois a máxima separação será $y_i + 0$ (Figura 13d). Assim, o problema de regressão é transformado em um problema de classificação binária e o algoritmo SVM pode ser usado e aplicado normalmente.

3.4.1.4.1 Regressão por Vetores de Suporte

A regressão por vetores de suporte, do inglês *Support Vector Regression* (SVR), é uma técnica de aprendizado estatístico supervisionado para modelagem de dados de processos (SMOLA; SCHÖLKOPF, 2004; XU; CHAN; ZHANG, 2019), e também para aplicações de previsão de regressão e de séries temporais, com a obtenção de excelentes desempenhos (MÜLLER et al., 1997; DRUCKER et al., 1997; STITSON et al., 1999; MATTERA; HAYKIN, 1999).

A ideia chave do SVR está em transformar o espaço de entrada (espaço original) (dados de treinamento) em um espaço de recurso de alta dimensão (espaço de característica), por meio

Figura 13: Representação gráfica da transformação de um problema de regressão em classificação binária. Problema de regressão (a); adição e subtração da constante d_s aos dados de entrada (b); transformação do problema de regressão original em um problema de classificação binária (c); hiperplano que define a função de regressão (d).



Fonte: Elaborado pelo autor.

de uma transformação não linear e extrair as informações e a regularidade contida entre os dados (CORTES; VAPNIK, 1995). Ao executar essa transformação, o SVR calcula um espaço de recurso de correlação e as variáveis de saída. Para melhorar o desempenho da regressão, o SVR minimiza a função de erro no treinamento. Como resultado, um relacionamento não linear entre as variáveis de entrada e saída é encontrado no espaço de alta dimensão (VAPNIK, 2013; MA; YE; MA, 2019). O SVR também faz uso da função de perda ϵ -insensível, com a finalidade de desenvolver uma função onde todos os erros estão sob um valor de ϵ pré-definido, mas produzindo a melhor capacidade de generalização possível (LU; WANG, 2010; KHOSRAVI et al., 2018).

O modelo básico especificamente supõe um conjunto de dados (amostra) de treinamento,

$$(x_1, y_1), (x_2, y_2), \dots, (x_{n_t}, y_{n_t}) \in (\mathbb{R}^{p_n} \times \mathbb{R}^{n_t}) \quad (3.23)$$

em que, n_t é o número de observações do conjunto de treinamento; x_i é um vetor de entradas p -dimensional e y_i são as saídas p -dimensionais x_i . O objetivo é encontrar a função de regressão

$f(\mathbf{x})$. Tal função linear, ou seja, a função SVR, é dada pela Equação 3.24:

$$f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x}_i + b \quad (3.24)$$

em que $f(\mathbf{x})$ é o valor predito com no máximo um desvio ϵ de y_i ; \mathbf{w} é um vetor de pesos (indicando a característica das entradas); b é o viés (ou intercepto) do modelo (KANG; LI, 2016).

Considerando a função linear da Equação 3.24, no SVR procura-se uma função com pequeno \mathbf{w} , o que pode ser conseguido pela minimização da norma $\|\mathbf{w}\|$. O algoritmo resolve, então, o problema de otimização apresentado na Equação 3.25 (SMOLA; SCHÖLKOPF, 2004; FACELI et al., 2011).

$$\text{minimizar} \quad \frac{1}{2} \|\mathbf{w}\|^2 \quad (3.25)$$

$$\text{sujeito a} \quad \begin{cases} y_i - \mathbf{w}^T x_i - b \leq \epsilon_i \\ \mathbf{w}^T x_i + b - y_i \leq \epsilon_i \end{cases} \quad (3.26)$$

A minimização da Equação 3.25 significa encontrar o ϵ -tubo mais delgado possível. As restrições das Inequações 3.26 refletem que a função $f(\mathbf{x})$ aproxima os pares (x_i, y_i) com uma precisão constante ϵ , ao mesmo tempo torna o problema de solução convexa possível. Mesmo considerando um erro mínimo nas Inequações 3.26, ainda assim, os dados de treinamento podem extrapolar a margem de precisão. Estes dados podem ser incorporados ao processo de cálculo, ao mesmo tempo em que o problema de minimização da Equação 3.25 pode ser flexibilizado. Para isso, incorpora-se variáveis de folga do tipo (ξ_i) , possibilitando trabalhar com ruídos e *outliers*, aceitando que alguns exemplos fiquem fora da região entre $-\epsilon$ e $+\epsilon$ e tornando o problema solucionável.

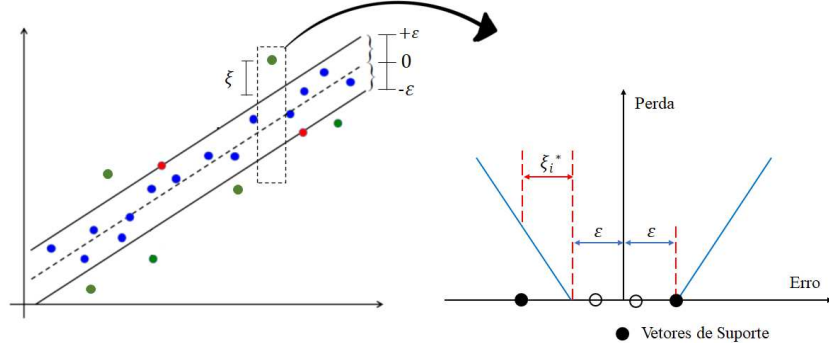
A Equação 3.27 exhibe as variáveis de folga e as restrições são mostradas nas Inequações 3.28. As variáveis de folga ξ_i e ξ_i^* estão associadas com os dados localizados abaixo da margem inferior e acima da margem superior, respectivamente. Assim, a equação do problema de otimização pode ser dada por:

$$\text{minimizar} \quad \frac{1}{2} \|\mathbf{w}\|^2 = C \sum_{i=1}^n [L(\xi_i) + L(\xi_i^*)] \quad (3.27)$$

$$\text{sujeito a } \begin{cases} y_i - f(x_i) \leq \epsilon + \xi_i \\ f(x_i) - y_i \leq \epsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \quad (3.28)$$

em que, $C > 0$ é a constante de regularização que determina a quantidade de erros no conjunto de treinamento e o custo para obtenção da região delgada ϵ -tubo; $\|\cdot\|$ é a norma do vetor; e $L(\cdot)$ é a função de perda ϵ -insensível (Equação 3.29), que introduz variáveis de folga não negativas e considera os pontos situados fora da margem $|f(\mathbf{x}_i) - y_i| \leq \epsilon$ com uma certa penalidade (SMOLA; SCHÖLKOPF, 2004; SAPANKEVYCH; SANKAR, 2009; FACELI et al., 2011). É importante destacar que apenas as amostras localizadas fora do ϵ -tubo, conforme Figura

Figura 14: Função Perda ϵ -insensível



Fonte: Adaptado de (SMOLA; SCHÖLKOPF, 2004; KANG; LI, 2016).

14, contribuem para o cálculo da Equação 3.27, isto é, os pontos que não excedam ϵ são penalizados através de uma função de perda ϵ -insensível apresentada na Equação 3.29. Essas amostras são denominadas vetores de suporte, pois determinam os fundamentos para a função de regressão estimada. Os outros pontos da amostra (dentro do ϵ -tubo) não são incluídos no modelo podendo ser excluídos após a construção do modelo SVR. A função de perda ϵ -insensível é dada pela Equação 3.29:

$$L(\xi) = \begin{cases} 0 & \text{se } |\xi| \leq \epsilon \\ |\xi| - \epsilon & \text{caso contrário} \end{cases} \quad (3.29)$$

Utilizando o método do multiplicador de *Lagrange* no problema de otimização da Equação 3.27, obtém-se a função objetivo apresentada na Equação 3.30.

$$f(\mathbf{x}) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) \mathbf{w} \cdot \mathbf{x} + b \quad (3.30)$$

onde, α_i e α_i^* são os multiplicadores de Lagrange (SMOLA; SCHÖLKOPF, 2004). Em casos em que o modelo não é linear, é necessário expandir essa regressão para um caso de regressão não linear, inserindo uma função Kernel, que mapeia os dados de entrada para um espaço de característica de dimensão superior, onde a regressão linear torna-se possível. Portanto, a função de regressão de SVR é obtida como na Equação 3.31

$$f(\mathbf{x}) = \sum_{i=1}^n (\alpha_i - \alpha_i^*) K(\mathbf{x}_i, \mathbf{x}) + b \quad (3.31)$$

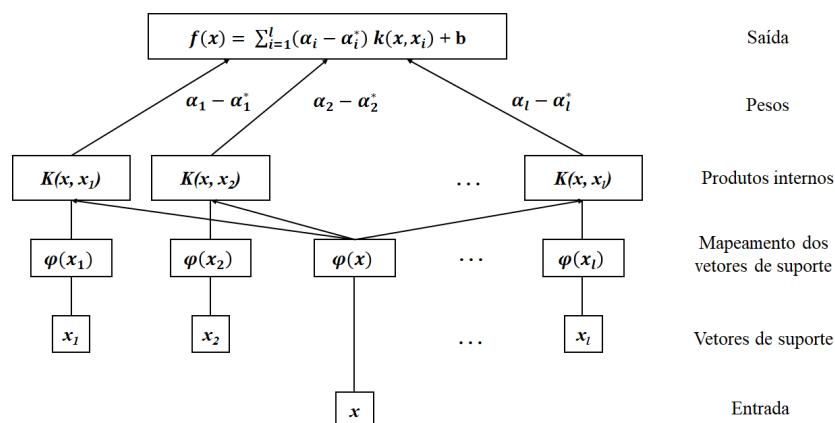
em que o valor de $K(x_i, x)$ é chamada de função kernel aplicada aos dados de entrada, e o valor do kernel é igual ao produto interno de dois vetores, x_i e x , no espaço de recurso $\phi(x_i)$ e $\phi(x)$, respectivamente; isto é $K(x_i, x) = \phi(x_i) \cdot \phi(x)$.

Qualquer função que atenda à condição de Mercer pode ser usada como função do *kernel* (MERCER, 1909; VAPNIK; VAPNIK, 1998; HORNIK; MEYER; KARATZOGLOU, 2006; STEINWART; CHRISTMANN, 2008; VAPNIK, 2013). Existem vários tipos de funções *kernel*, as mais utilizadas na prática em vetores de suporte são a linear, polinomial, sigmoïdal e função de base radial (RBF do inglês, *radial basis function*). A utilização de Funções *kernel* diminui a quantidade de parâmetros livres na computação de problemas envolvendo dados não lineares, fato que confere ao SVM superioridade enquanto técnica de aprendizado de máquinas (CHEN, 2004).

Os parâmetros do SVR exercem muita influência nos resultados da previsão. Assim, o desempenho do SVR e sua capacidade de generalização dependem fortemente da combinação dos parâmetros C , ϵ e do(s) parâmetro(s) da função *kernel*. O parâmetro C , chamado de penalização, é responsável pelo *tradeoff* entre a maximização da margem do ϵ -tubo, quando são aceitos valores com desvios maiores que ϵ , e a minimização do erro no conjunto de treinamento. Se seu valor for muito alto pode haver *overfitting*, caso seja muito baixo *underfitting*. O parâmetro ϵ é responsável pelo raio do ϵ -tubo ao redor da função de regressão. Considere-se também o desvio permitido a previsão em relação aos dados alvos (WANG et al., 2014; KANG; LI, 2016).

Após a identificação dos parâmetros, o modelo pode ser utilizado para prever a saída correspondente a uma entrada desconhecida usando Equação 3.31. A arquitetura do SVR é apresentada na Figura 15. A entrada \mathbf{x} (para a qual uma predição deve ser realizada) e os vetores de suporte $x_i, i = 1, \dots, n$ são mapeados não linearmente pela função ϕ para um espaço característico, onde os produtos internos são calculados. Isto corresponde ao cálculo de funções *kernel* $K(x_i, x)$. Finalmente, os resultados são combinados através dos pesos $(\alpha_i - \alpha_i^*)$ e somados com o termo constante b para fornecer a previsão de \mathbf{x} .

Figura 15: Arquitetura do modelo de Regressão por Vetores de Suporte (SVR)



Fonte: Adaptado de (KANG; LI, 2016).

Uma tarefa importante e ao mesmo tempo complexa na abordagem de SVR é a estimativa ótima dos seus parâmetros e dos parâmetros da função *kernel*, considerando que não há um consenso na literatura entre os pesquisadores da área sobre o melhor método a ser usado. Na literatura é possível encontrar pesquisas que utilizaram Algoritmos Genéticos (GA) (SAPANKEVYCH; SANKAR, 2009) e Enxames de Partículas (PSO) (WANG et al., 2014) na otimização dos parâmetros. Entretanto, para Chen (2004), as abordagens práticas usadas na escolha dos hiperparâmetros do SVR possuem algumas deficiências. No presente trabalho, recorre-se ao Algoritmo *Simulated Annealing* (KIRKPATRICK; GELATT; VECCHI, 1983).

Para Lu e Wang (2010), as SVMs têm um grande potencial e proporcionam desempenho superior em diversas aplicações práticas. Isso se deve, em parte, aos princípios de minimização do risco da estrutura (*structural risk minimization*, SRM) das SVMs, que possui capacidade maior de generalização e são superiores ao princípio da minimização do risco empírico (*empirical risk minimization*, ERM), usado por modelos de aprendizado de máquinas, como redes neurais tradicionais. Assim, as SVMs garantem ótimos globais (CAO; TAY, 2003).

Pesquisas apresentam resultados relevantes com a aplicação do modelo SVR, na busca de prever a demanda. Guanghui (2012) aplicava SVR para prever a demanda da cadeia de suprimentos; Sarhani e El Afia (2014), usaram na previsão da demanda da cadeia de suprimentos, a fim de otimizar o gerenciamento da cadeia de suprimentos e Fattaheian-Dehkordi et al. (2014), na demanda de eletricidade da Companhia de Distribuição de Eletricidade de Teerã.

3.4.1.5 Lógica *Fuzzy*

A lógica *fuzzy*, conhecida como lógica nebulosa, ou difusa, é uma técnica de inteligência computacional cuja ideia é modelar o modo aproximado do raciocínio humano (ROSS, 2005). Com isso, é possível tomar decisões em ambientes de incerteza e imprecisão Zadeh (1965), o que não ocorre com a lógica clássica (lógica discreta ou booleana). A modelagem matemática de conceitos *fuzzy* foi introduzida por Lofti A. Zadeh em seu artigo em 1965, onde argumenta-se que o significado na linguagem natural é uma questão de grau (PIRES et al., 2004). A principal ideia da lógica *fuzzy* é admitir possibilidades de um valor ser "verdadeiro", diferentemente do que se tem na lógica clássica cuja representação binária (0,1) diz se um valor é falso ou verdadeiro, não podendo ser ao mesmo tempo parcialmente verdadeira e parcialmente falsa (RUSSEL; NORVIG, 2013).

As lógicas ditas não clássicas contrariam as suposições binárias que não admitem ambiguidades e contradições. No entanto, o conceito de dualidade estabelece a possibilidade de uma coexistência entre algo e seu oposto, fazendo com que as aplicações das lógicas não clássicas (exemplo, lógica *fuzzy*) parecerem naturais e, por vezes, inevitáveis.

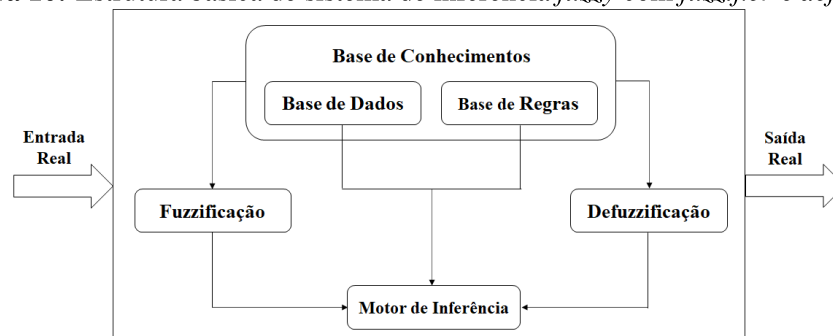
Na lógica *fuzzy*, um valor pode ser classificado de forma linguística como "totalmente verdadeiro", "totalmente falso", "parcialmente verdadeiro", ou "parcialmente falso". Dessa maneira, a representação é dada em um intervalo [0,1] (KLIR; YUAN, 1996). Essas ideias foram ampliadas para a teoria dos conjuntos. Assim, em uma abordagem clássica, a definição de pertinência de um elemento é associá-lo aos valores 1 se pertence a um conjunto específico, caso contrário 0. No contexto *fuzzy*, a pertinência não é uma questão de afirmação ou negação, mas uma questão de grau e é dada no intervalo [0,1], de modo que um mesmo elemento pode pertencer parcialmente a mais de um conjunto *fuzzy* com diferentes graus de pertinência variando nesse intervalo (WANF, 2003).

Uma das principais contribuições da lógica *fuzzy* reside no tratamento de dados registrados a partir de termos linguísticos por meio de funções de pertinência (funções características) aos conjuntos *fuzzy*, permitindo que a linguagem natural seja convertida para um formato numérico, de fácil manipulação. Assim, um conjunto *fuzzy* é um agrupamento impreciso e indefinido onde o processo de transição de não pertinência para pertinência ocorre de forma gradual, não brusca (NGUYEN; WALKER; WALKER, 2018).

A estrutura geral de um sistema *fuzzy* possui alguns componentes conceituais como: uma base de regras, que contém o conjunto de regras nebulosas; uma base de dados, que define as funções de pertinência usadas nas regras nebulosas; e um mecanismo de raciocínio, que realiza um procedimento de inferência (raciocínio nebuloso) para obter a saída ou conclusão,

baseado nas regras e fatos conhecidos (WANF, 2003).

Figura 16: Estrutura básica de sistema de inferência *fuzzy* com *fuzzifier* e *defuzzifier*



Fonte: Adaptado de (WANF, 2003; PIRES et al., 2004).

A Figura 16 mostra a arquitetura básica de um sistema *fuzzy*, formada a partir de um módulo de fuzzificação, cuja principal função é converter os valores de entrada (escalar ou vetorial) do sistema para termos linguísticos representados por conjuntos *fuzzy*; uma base de conhecimento, em que todo o conhecimento sobre o domínio do problema em questão é armazenado. A base de conhecimento, formada por uma base de regras, que contém o conjunto de regras *fuzzy* e por uma base de dados, que define as funções de pertinência usadas nas regras nebulosas; um mecanismo (motor) de inferência responsável pelo desenvolvimento do raciocínio nebuloso baseado no conhecimento representado na base de conhecimento; e por um módulo de defuzzificação, que realiza uma transformação da resposta do sistema nebuloso, a qual está representada por um conjunto nebuloso, em uma resposta não nebulosa (valor real) (ZADEH, 1965). Com isso, diversos pesquisadores propuseram alguns modelos de previsão baseados na teoria *fuzzy* a fim de resolver problemas de séries temporais (ZADEH, 1973).

3.4.1.5.1 Série Temporal *Fuzzy*

Na metodologia tradicional de séries temporais, os valores das observações são representados por números reais. Contudo, na metodologia de séries temporais *fuzzy* (STF), os valores das observações podem também ser representados por valores linguísticos (como muito baixo, baixo, médio, alto e muito alto).

A STF é um método de previsão que utiliza os princípios *fuzzy* desenvolvidos por Zadeh (1965), que mais tarde foi aplicado na resolução de problemas de previsão. Song e Chissom (1991, 1993a, 1994), foram os primeiros a propor um modelo de STF baseado nos conceitos da teoria de sistemas *fuzzy*, para modelar e prever processos cujos valores são descritos por variáveis linguísticas e usaram como aplicação a previsão de matrículas educacionais (matrí-

culas de estudantes da Universidade do Alabama). O modelo usou um operador de composição max-min (\circ) e construiu a relação difusa R (é a união de relações *fuzzy*) para calcular os valores de previsão.

Sullivan e Woodall (1994), revisaram os modelos de STF variante e invariante no tempo de primeira ordem proposto por Song e Chissom (1993a), onde os modelos são comparados entre si e com uma variante do tempo modelo de Markov usando rótulos linguísticos com distribuição de probabilidade. Em 1996, Chen melhorou o modelo de STF de primeira ordem, introduzindo um método aritmético mais eficiente usando a tabela de relações lógicas nebulosas ao invés do operador max-min. O método é computacionalmente mais fácil do que o método de Song e Chissom (CHEN et al., 1996).

Hwang, Chen e Lee melhoraram o modelo de STF de primeira ordem com base na tendência de matrículas dos últimos anos (HWANG; CHEN; LEE, 1998). A ideia do método proposto é que a variação das matrículas deste ano, por exemplo, está relacionada à tendência das matrículas dos últimos anos. Com isso, se a tendência das matrículas dos últimos anos está aumentando, então o número de matrículas deste ano pode aumentar. Hwang estendeu o modelo de Chen com regras adicionais de previsão e propôs um modelo usando o conhecimento heurístico para melhorar a previsão para a inscrição da UA.

Em um estudo posterior, Hwang identificou que a extensão dos intervalos afeta os resultados da previsão e propôs outro modelo para definir o comprimento do intervalo, em que o comprimento era baseado na distribuição e no comprimento médio (HUANG, 2001). Chen (2002), desenvolveu um novo modelo de STF chamado modelo de série temporal difusa de alta ordem para lidar com problemas de previsão. No entanto, esses modelos não podem lidar com o problema de previsão sobre os dados de períodos desconhecidos, limitados à simulação e à previsão aos dados históricos.

Com base nessa limitação, Abbasov e Mamedova tiveram importantes contribuições para essa direção. Os autores propuseram o modelo STF a fim de prever a população do Azerbaijão. No modelo, é utilizado a diferença de dados de primeira ordem. Devido ao seu bom desempenho para diversos tipos de problemas de previsão, o modelo de Abbasov e Mamedova (AM) teve sua aplicação estendida a muitos casos (ABBASOV; MAMEDOVA, 2003). O modelo AM tem vantagens sobre os outros, pois pode prever valores fora do intervalo dos dados originais (ABBASOV; MAMEDOVA, 2003).

Diferentes modelos de STF foram propostos para prever dados não lineares e várias aplicações, como matrículas (SONG; CHISSOM, 1991; CHEN et al., 1996; SONG; CHISSOM, 1994), temperatura (CHEN; HWANG, 2000; WANG; CHEN, 2009), acidentes de carro (JILANI; BURNEY; ARDIL, 2007), demanda turística (MUHAMMAD et al., 2012), concen-

tração de poluição (DOMAŃSKA; WOJTYLAK, 2012), cadeia de suprimentos (TOZAN; VAYVAY, 2008; CHAN; SAMVEDI; CHUNG, 2015), previsão de emprego (MARINOIU, 2016), mercados de ações (CHENG et al., 2008; CHU et al., 2009; CHEN; KAO, 2013), e na previsão de demanda (LI et al., 2012). Da mesma forma, pesquisadores apresentaram métodos de STF para lidar com a previsão de demanda.

3.4.1.5.2 Definições Relacionadas a Séries Temporais *Fuzzy*

Para Song e Chissom (1991, 1993a,b, 1994), a previsão de um valor é realizada a partir das relações de lógica *fuzzy* extraídas de uma STF. As principais definições relacionadas as STF são apresentadas a seguir.

Definição 1. Um conjunto *fuzzy* é uma classe de objetos com graus de associação contínuos. Seja U o universo de discurso ou conjunto universal, $U = \{u_1, u_2, \dots, u_n\}$, onde os elementos de U são considerados seus possíveis valores, portanto um conjunto *fuzzy* de variáveis linguísticas A_i de U é definido pela Equação 3.32.

$$A_i = \frac{\mu A_i(u_1)}{u_1} + \frac{\mu A_i(u_2)}{u_2} + \frac{\mu A_i(u_3)}{u_3} + \dots + \frac{\mu A_i(u_n)}{u_n}, \quad (3.32)$$

em que μA_i é a função de associação (pertinência) do conjunto A_i , tal que $\mu A_i: U \rightarrow [0, 1]$. Se u_i pertence a A_i , então $\mu A_i(u_i)$ indica o grau de associação com que u_i está em A_i e $1 \leq i \leq n$.

Aqui, o símbolo "+" indica a operação de união e o símbolo "÷" indica o separador, em vez do somatório e divisão comumente usados na álgebra, respectivamente (ZADEH, 1965).

Definição 2. Seja $Y(t)$ um subconjunto dos números reais e U universo de discurso no qual os conjuntos *fuzzy* $f_i(t)$, ($i = 1, 2, \dots$) são definidos. E seja $F(t)$ uma coleção de $f_i(t)$. Com isso, $F(t)$ é chamada de série temporal *fuzzy* em $Y(t)$ com ($t = 0, 1, 2, \dots$).

Definição 3. Considere que $F(t)$ é causada apenas por $F(t - 1)$ e é representada por $F(t - 1) \rightarrow F(t)$. A relação existente entre $F(t)$ e $F(t - 1)$ é uma relação *fuzzy* e esse relacionamento pode ser representado pela equação relacional *fuzzy* (Equação 3.33).

$$F(t) = F(t - 1) \circ R(t, t - 1) \quad (3.33)$$

onde "o" originalmente é o operador de composição de máx-min e $R(t, t - 1)$ é a relação difusa entre $F(t)$ e $F(t - 1)$. Assim, $F(t - 1) \rightarrow F(t)$ é chamada de relação *fuzzy* de primeira ordem do modelo $F(t)$.

Definição 4. Se $F(t)$ é causada por mais conjuntos *fuzzy* $F(t - 1), F(t - 2), \dots, F(t - n)$, a relação *fuzzy* é expressa por $F(t - n), \dots, F(t - 2), F(t - 1) \rightarrow F(t)$. Essa relação é

chamada de n -ésima ordem do modelo de série temporal *fuzzy* (CHEN et al., 1996).

Definição 5. Se $R(t, t - 1)$ é uma relação difusa de $F(t)$ independente do tempo t , ou seja, para tempos diferentes t_1 e t_2 , $R(t_1, t_1-1) = R(t_2, t_2-1)$, então $F(t)$ é chamada de série temporal *fuzzy* invariante no tempo. Caso seja dependente de t , ou seja, $R(t_1, t_1-1) \neq R(t_2, t_2-1)$ para algum t , então $F(t)$ é denominada como uma série temporal *fuzzy* com variação temporal ou variante no tempo (CHEN; HWANG, 2000) e a relação pode ser expressa como uma equação *fuzzy* relacional (Equação 3.34).

$$F(t) = F(t - 1) \circ R^w(t, t - 1) \quad (3.34)$$

$w > 1$ é um parâmetro de tempo (número de minutos ou anos) pelo qual a previsão $F(t)$ é afetada.

3.4.1.5.3 Modelo Abbasov-Mamedova

O formalismo de Abbasov e Mamedova (AM), desenvolvido a partir da ideia de STF, foi proposto com o objetivo de prever a população do Azerbaijão (ABBASOV; MAMEDOVA, 2003). O modelo AM tem vantagens sobre outros modelos de STF, pois pode prever valores fora do intervalo dos dados reais. No entanto, o desempenho do modelo depende fortemente de três parâmetros: o número de conjuntos *fuzzy* (n), o número de elementos no conjunto de dados usado como informação anterior para prever os dados (w) e a constante (C) pertencente ao intervalo $(0,1)$, escolhida de forma a garantir a conversão de valores quantitativos definidos em valores fuzzy. Estes parâmetros podem ser dimensionados com base na própria experiência do pesquisador (ABBASOV; MAMEDOVA, 2003; SASU, 2010; MARINOIU, 2016; KHAN; KHAN, 2019), embora seja um problema complexo.

Os parâmetros do modelo não são investigados adequadamente para encontrar os valores ótimos para cada série temporal. Com isso, diversas abordagens foram propostas a fim de identificar os parâmetros apropriados (n , w e C) do modelo AM. Esses métodos apresentaram bons resultados. Porém, autores como Vovan (2019) e Che-Ngoc et al. (2018), propuseram diferentes métodos na busca do valor ideal para cada parâmetro. Che-Ngoc et al. (2018), propôs um algoritmo de dois estágios para o modelo AM, no qual, o primeiro estágio identifica os parâmetros (n , w e C), antes de passá-los para o modelo AM no segundo estágio. A principal lacuna dessa abordagem é que os parâmetros foram examinados separadamente e suas interações não foram investigadas. Os autores Thao (2019), propuseram uma nova abordagem que avalia a qualidade do modelo alterando simultaneamente os três parâmetros com base no

algoritmo de evolução diferencial (DE).

Outro ponto importante a ser destacado é que o modelo AM constrói seu universo de discurso denominado U a partir da variação de dados entre períodos consecutivos de tempo ou de sua variação percentual. Para a construção de U , são atribuídos dois valores D_1 e D_2 , que são números positivos usados para ajustar os limites de U para garantir a suavidade dos limites do intervalo (ABBASOV; MAMEDOVA, 2003). Com base na ideia acima e a fim de melhorar a qualidade do modelo AM, nesta tese foi utilizado o *simulated de annealing* (SAn) na determinação dos parâmetros ótimos. O SAn mimetiza o processo de resfriamento de materiais (DELAHAYE; CHAIMATANAN; MONGEAU, 2019).

A metodologia de previsão proposta por AM é definida com base nos dados históricos da série temporal, u_t correspondendo ao ano $t = 1, 2, \dots, m$. O formalismo Abbasov-Mamedova consiste na realização de seis etapas a seguir.

Etapa 1: Calcule a variação V_t entre dois períodos consecutivos pela Equação 3.35. Na sequência, defina o universo de discurso U pela Equação 3.36.

$$V_t = u_t - u_{t-1} \quad (3.35)$$

$$U = [V_{min} - D_1, V_{max} + D_2] \quad (3.36)$$

em que, V_{min} é a menor variação, V_{max} é a maior variação, D_1 e D_2 são números positivos.

Etapa 2: Particione o universo de discurso U em n intervalos de comprimento igual u_i , $i = 1, 2, \dots, n$, de modo que cada intervalo u_i contenha no mínimo um valor de variação. Em seguida, encontre os pontos médios u_i^m de cada intervalo.

Etapa 3: Defina o conjunto *fuzzy* A_i , $i = 1, 2, \dots, n$, no conjunto universal U pela Equação 3.37:

$$\mu_{A_i}(u) = \frac{1}{1 + [C \times (u - u_i^m)]^2}, \quad (3.37)$$

onde u é um elemento genérico do conjunto universal U , u_i^m é o ponto médio do intervalo correspondente u_i ($i = 1, 2, \dots, n$) identificado na Etapa 2 e C é uma constante e $C \in (0,1)$.

Etapa 4: Converta (fuzzificação) os dados de entrada, variações de ponto no tempo, que são valores reais em valores difusos pela Fórmula 3.37.

Etapa 5: Selecione um número inteiro w , $1 < w < l$, onde l é o número de períodos, antes do período atual incluído na avaliação experimental. Com base no sistema de inferência w e Mamdani escolhido, estabelecemos uma matriz de operação $O^w(t)$ de tamanho $i \times j$ (aqui i é o número de linhas, que está em conformidade com a sequência de anos $t - 2, t - 3, \dots, t - w$, j é o número de colunas, conforme o número de intervalos de variação) e uma matriz

de critérios $K(t)$ de tamanho $1 \times j$ (uma matriz de linhas que corresponde à variação difusa no período $t - 1$) depois disso, a matriz de relações $R(t)$ é calculada conforme as Equações 3.38 e 3.39.

$$R(t) [i, j] = O^w [i, j] \cap K(t) [j] \quad (3.38)$$

$$R(t) = O^w(t) \otimes K(t) = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1j} \\ R_{21} & R_{22} & \cdots & R_{2j} \\ \cdots & \cdots & \cdots & \cdots \\ R_{i1} & R_{i2} & \cdots & R_{ij} \end{bmatrix} \quad (3.39)$$

onde $O^w(t)$ é a matriz de operação, $K(t)$ é a matriz de critérios, \otimes é o operador $\min(\cap)$. Defina $\hat{F}(t)$, a previsão nebulosa das variações para o período t com base na equação 3.40.

$$\begin{aligned} \hat{F}(t) &= [\max(R_{11}, \cdots, R_{i1}), \cdots, \max(R_{1j}, \cdots, R_{ij})] = \\ &= [\mu_{A_1}(V_t), \mu_{A_2}(V_t), \cdots, \mu_{A_m}(V_t)] \end{aligned} \quad (3.40)$$

Etapa 6: Defuzzifique os resultados obtidos da etapa 5 (valores de $\hat{F}(t)$) de acordo com a Equação 3.41.

$$V(t) = \frac{\sum_{i=1}^m \mu_{A_i}(V_t) \times u_i^m}{\sum_{i=1}^m \mu_{A_i}(V_t)} \quad (3.41)$$

em que, $\mu_{A_i}(V_t)$ é o valor da função de associação da variação da previsão no intervalo i , $V(t)$ é a variação da previsão defuzzificada. Para estimar o valor previsto \hat{u}_t para o período t , é utilizada a Equação 3.42:

$$\hat{u}_t = u_{t-1} + V(t) \quad (3.42)$$

onde u_{t-1} é o valor para o período $t - 1$ e $V(t)$ é a variação para o período t .

3.4.1.6 Metodologia de Previsão JUA

Nesta seção é descrita a metodologia de previsão de demanda da indústria de manufatura do caso de estudo, sendo aqui denominada de JUA. A empresa utiliza metodologia de previsão dividida em duas etapas. A primeira etapa com modelos matemáticos e a segunda com modelos cognitivos. Na etapa dos modelos matemáticos são utilizadas as informações fornecidas pelo sistema *Winthor* (Sistema de ERP desenvolvido pela empresa TOTVS) que se baseia em dados e informações do histórico de demanda da empresa (WINTHOR, 2021). A ferramenta utiliza modelos matemáticos para entender padrões do comportamento da demanda no passado e estimar o comportamento da demanda no futuro. No entanto, a ferramenta é

considerada uma caixa preta para o gestor da produção, não sabendo qual(ais) o(s) método(s) utilizado(s), bem como, o nível de assertividade entre os valores previstos e os observados, causando insegurança na previsão realizada. Isso leva o gestor a necessitar de uma segunda etapa de previsão.

Na segunda etapa, são utilizados modelos cognitivos. Esses modelos são usados como fonte de informação para modelagens de estudos onde os dados empíricos não possuem uma relevância considerável ou são parcialmente relevantes. Os modelos cognitivos são as opiniões de especialistas em relação a determinada variável de interesse com base nos seus conhecimentos e experiências (NETO et al., 2020). O método utiliza a análise do conhecimento, experiência e a opinião de especialistas, identificados como os gerentes de vendas e produção e o gestor do PCP, explorando o conhecimento acumulado e a experiência de cada profissional, além de informações como nível de estoques, tempo de esgotamento e tendências de mercado (ARMSTRONG, 2001). Com isso, identificou-se, portanto, uma combinação de métodos matemáticos e cognitivos, corroborando com a afirmação de Hyndman, Makridakis e Wheelwright (1998), para os quais a união de diferentes métodos de previsão contribui para a redução dos erros, aumentando a confiabilidade dos dados gerados.

Ainda assim, a metodologia de previsão causa, em diversos momentos, problemas de sub-produção e ou superprodução como aumento substancial e contínuo de recursos empatados em estoque, que gira em torno de 55.674 caixas de produtos; aumento do uso dos espaços da fábrica com estoque, uma vez que os espaços usados para estocar produtos fabricados é cerca de $1.450m^2$ da área da fábrica; redução da produtividade ocasionada pelo aumento no número de realinhamentos/recalibrações das máquinas usadas para a produção de variados tipos de produtos relacionados a demandas não assertivas; aumento da carga de trabalho de natureza de hora extra, girando em torno de 600hrs extras/mês; e problemas relacionados ao dimensionamento de equipes quanto de equipamentos, máquinas, manutenções e insumos.

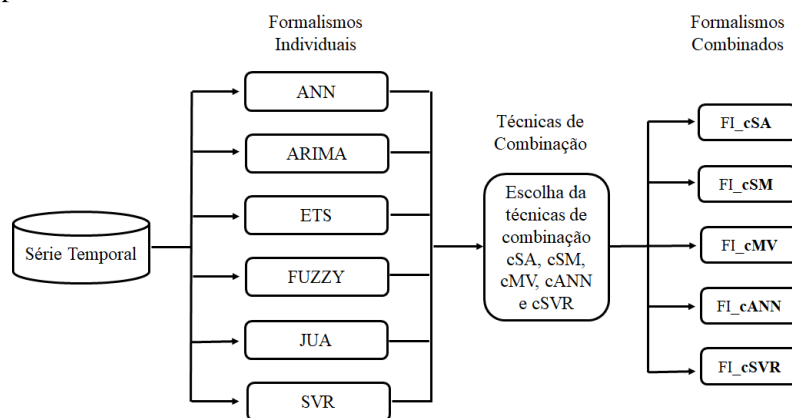
A escolha de metodologias por conveniência, sistemas simples (sistema *Winthor*) e/ou de sistemas sobre os quais possuem familiaridade (experiência dos gestores), podem levar a resultados distantes dos valores reais, ocasionando erros graves em situações que envolvem grandes mudanças (ARMSTRONG, 2001; PETROPOULOS; KOURENTZES, 2015; PRAKASH; RAMYA; YUGANDHAR, 2016; MATSUMOTO et al., 2016; ROSIENKIEWICZ; CHLEBUS; DETYNA, 2017; KOURENTZES; ROSTAMI-TABAR; BARROW, 2017). Por sua vez, o mercado de ferramentas dedicadas a séries temporais carece de alternativas mais atraentes e acessíveis à indústria.

3.4.2 Formalismos de Modelagem Combinada

A combinação de preditores é um método constantemente usado para melhorar a qualidade da previsão (COSTANTINI; PAPPALARDO, 2010; KOZIK; SEP, 2012). Foi apresentada pela primeira vez por Bates e Granger (1969), e é considerado um método atraente para realizar previsões (CLEMEN, 1989). No entanto, melhorar o desempenho da previsão é uma tarefa importante, mas, frequentemente, difícil (ROSIENKIEWICZ; CHLEBUS; DETYNA, 2017). Para situações em que nenhuma abordagem dominante foi determinada, a combinação de preditores tem sido uma das possibilidades de pesquisa que vem produzindo resultados marcantes. Autores como Bates e Granger (1969); Winkler e Makridakis (1983); Firmino, Mattos Neto e Ferreira (2014), contribuem para essa assertiva.

Para modelagem combinada, a previsão para u_t , considera as estimativas produzidas pelos K formalismos individuais, $\{\hat{u}_{ti}\}_{i=1}^K, i = 1, \dots, K$. Baseado nos objetivos desta tese foram utilizados cinco tipos de combinadores: Média Simples (cSA), Mediana Simples (cSM), Mínima Variância (cMV), Regressão por Vetores de Suporte (cSVR) e Rede Neural Artificial (cANN). A Figura 17 apresenta a estrutura geral de combinação dos formalismos individuais através de seus combinadores.

Figura 17: Estrutura geral de combinação dos formalismos individuais (ANN, ARIMA, ETS, FUZZY, JUA e SVR) por meio das técnicas de combinação (cANN, cMV, cSA, cSM e cSVR) de uma determinada série temporal



Fonte: Elaborado pelo autor.

Nessa estrutura, os formalismos individuais (ANN, ARIMA, ETS, FUZZY, JUA e SVR) foram identificados como FI na parte dos formalismos combinados, associados à técnica de combinação escolhida.

3.4.2.1 Combinação via Média Simples

Em combinação de preditores, a média simples tem sido um tópico de pesquisa importante e um considerável ponto da literatura. Resultados indicam que a robustez da combinação via Média Simples, aqui identificada por cSA, decorre de estimativas de peso instáveis de pequenas amostras de treinamento ou características de erro de previsão divergentes entre o treinamento e as amostras de avaliação (BLANC; SETZER, 2016). Assim, a cSA é apontada como um processo de fácil utilização, sendo considerada em determinados casos melhores e mais acurados/assertivos que métodos mais elaborados (MAKRIDAKIS et al., 1982; FLORES B. E. & WHITE, 1989; ARMSTRONG; COLLOPY, 1998; H.; W., 2004; COSTANTINI; GUNTER; M. KUNST, 2017). O método é uma combinação linear que calcula essencialmente a média aritmética simples das previsões dos K formalismos individuais conforme Equação 3.43

$$\hat{u}_t^{SA} = \frac{\sum_{i=1}^K \hat{u}_{t,i}}{K} \quad (3.43)$$

em que, \hat{u}_t^{SA} é a estimativa do modelo combinado para u_t e $\hat{u}_{t,i}$ é o valor previsto pelo i -ésimo modelo para u_t .

Uma característica básica e intuitiva sobre a cSA é que seu desempenho depende da intensidade das variações de erros das previsões, bem como de sua correlação. Assim, espera-se que o método mostre um bom desempenho no caso de variações de erros similares e correlações de erro não elevadas (BUNN, 1985; GUPTA; WILTON, 1987), pois os pesos que são ótimos na amostra de avaliação abordam pesos iguais. Contudo, como apresentado por (WINKLER; CLEMEN, 1992; SMITH; WALLIS, 2009), a cSA pode suplantiar outros métodos, mesmo para diferentes variações de erros ou erros fortemente correlacionados devido a estimativas de peso instáveis (BLANC; SETZER, 2016).

3.4.2.2 Combinação via Mediana Simples

Na combinação por meio da Mediana Simples, identificada por cSM, calcula-se a mediana das previsões (valores estimados) dos K formalismos individuais. Como tal, os pesos são variáveis no tempo, pois cada método de previsão pode ser reordenado ao longo do tempo (MCMILLAN, 2018). Especificamente, a mediana representa a informação que divide os dados ordenados em duas partes iguais, metade abaixo e metade acima. Se o número de observações for par, a mediana está na metade entre os dois valores centrais, como seu ponto médio (MONTGOMERY; RUNGER, 2010; ROSS, 2014).

3.4.2.3 Combinação via Mínima Variância

O método de combinação por Mínima Variância, aqui identificada por cMV, proposto por Bates e Granger (1969), baseia-se na combinação linear de previsões individuais (ARMSTRONG, 2001; WERNER, 2005). O método estima os pesos dos modelos individuais com base nas variações de erro de suas previsões e suas correlações com os demais modelos (MARTINS; WERNER, 2014). Sua expressão é apresentada conforme as Equações 3.44 e 3.45.

$$\hat{u}_t^{MV} = \sum_{i=1}^K \omega_i \hat{u}_{t,i} \quad (3.44)$$

em que

$$\omega_i = \frac{\sum_{l=1}^K a_{li}}{\sum_{l=1}^K \sum_{j=1}^K a_{lj}} \quad (3.45)$$

sendo ω_i o peso atribuído ao i -ésimo modelo e a_{lj} o j -ésimo elemento da j -ésima linha da inversa da matriz de covariâncias M^{-1} com relação aos resíduos dos modelos individuais (DE OLIVEIRA et al., 2013). Os autores Firmino, Mattos Neto e Ferreira (2014) e Martins e Werner (2014), contribuem para uso do método e utilizam as Equações 3.46 e 3.47 para calcular os pesos ω_i , referentes a dois modelos individuais.

$$\omega_1 = \frac{\sigma_2^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}, \quad (3.46)$$

$$\omega_2 = \frac{\sigma_1^2 - \sigma_{12}}{\sigma_1^2 + \sigma_2^2 - 2\sigma_{12}}, \quad (3.47)$$

onde σ_i^2 é a variância relativa ao erro do modelo i e $\sigma_{i,j}$ é a covariância entre os erros dos modelos X_1 e X_2 . Entretanto, é importante destacar que para a utilização dos combinadores cMV os erros de previsão devem seguir distribuição normal, do contrário as estimativas podem ter viés (FIRMINO; MATTOS NETO; FERREIRA, 2014).

3.4.2.4 Combinação via Redes Neurais Artificiais

A combinação de modelos de previsão por técnicas não lineares, a exemplo das Redes Neurais Artificiais, aqui chamada de cANN, tem sido muito explorada por pesquisadores, apresentando resultados expressivos. Autores, como Zhang (2003), em seu modelo híbrido, recorrem a uma combinação de valores estimados dos modelos e seus resíduos. Os resultados obtidos por Zhang foram melhores no caso de modelos mistos, do que os obtidos com

cada modelo individual. Vários outros autores confiaram e utilizaram os modelos de Zhang ou desenvolveram novos modelos baseados no mesmo princípio de hibridização (ABURTO; WEBER, 2007; ZOU et al., 2007; PHAM; YANG et al., 2010).

Alguns desses modelos híbridos, usando redes neurais, foram aplicados diretamente ao problema de prever a demanda de peças sobressalentes em indústrias híbridas de manufatura, combinando regressão linear, ARIMA e ANNs (KHASHEI; BIJARI, 2012; LIU; TIAN; LI, 2012; ROSIENKIEWICZ; CHLEBUS; DETYNA, 2017; VARGAS; CORTÉS, 2017). Guo e Shen (2016), combinaram os preditores do modelo ETS e ANN para prever a demanda de material para o planejamento das atividades de produção.

Basicamente, a única diferença entre o modelo de previsão de redes neurais artificiais (denominado ANN nesta tese) e o combinador de rede neural cANN é a camada de entrada. No caso, o cANN envolve as previsões para u_t dos formalismos ANN, ARIMA, ETS, FUZZY, JUA e SVR na camada de entrada, enquanto a ANN faz uso dos valores defasados da série temporal em estudo. Além disso, a estimativa de parâmetros, bem como a arquitetura da cANN, são realizadas por meio da mesma metodologia aplicada para construir os previsores da ANN individual.

3.4.2.5 Combinação via Regressão por Vetores de Suporte

Para a combinação utilizando a metodologia de *Support Vector Regression*, similarmente a cANN, aqui denominada cSVR, são aplicadas como entradas da máquina de aprendizado as previsões para u_t dos formalismos individuais (ANN, ARIMA, ETS, FUZZY, JUA e SVR) e como alvo u_t . De forma semelhante à modelagem do formalismo individual SVR, os parâmetros serão otimizados através do Algoritmo *Simulated Annealing*.

3.5 Medidas de desempenho de previsão

Os modelos de previsão possuem um certo grau de incerteza associado, sendo necessário a medição deste. Existem diversas métricas que podem ser utilizadas com o objetivo de avaliar o desempenho dos modelos de previsão individuais e combinados. Tais métricas auxiliam a verificar quais modelos produzem a melhor estimativa (FORMIGONI CARVALHO WALTER et al., 2013; SYNTETOS; BABAI; GARDNER JR, 2015; FIRMINO; MATTOS NETO; FERREIRA, 2015). O critério mais importante para a escolha de um método de previsão é o seu desempenho, mensurado pela sua precisão, ou seja, quão próximo ele prevê o evento real.

Um método com pequeno erro de previsão é geralmente aceito como bom e o erro de

previsão é calculado com base na discrepância entre os valores previstos \hat{u}_t e reais ou observados u_t (HYNDMAN et al., 2008). Embora várias instituições dependam, principalmente, de previsões pontuais Klassen e Flores (2001); Danese e Kalchschmidt (2011), idealmente as previsões também devem incluir estimativas da incerteza na demanda futura, expressas como previsões de densidade ou intervalos de previsão. Como o desempenho do planejamento de produção depende muito da precisão da demanda prevista, escolher um método de previsão adequado e uma medida deve ser um pré-requisito para estabelecer a existência de um bom planejamento de produção (HA; SEOK; OK, 2018).

Neste trabalho são usadas as métricas de desempenho a seguir: Erro Quadrático Médio, do inglês *Mean Squared Error* (MSE) (CLEMENTS; HENDRY, 1993), Erro Médio Absoluto Percentual, do inglês *Mean Absolute Percentage Error* (MAPE) (CLEMENTS; HENDRY, 1993), Variância Relativa Média, do inglês *Average Relative Variance* (ARV) (HANN; STEURER, 1996), a estatística U de Theil (Theil) (HANN; STEURER, 1996), a Predição Errada de alteração na Direção, do inglês *Wrong Prediction on Change of Direction* (WPOCID) (YAO; TAN, 2000), Índice de Desacordo (ID), Intercepto do ajuste linear entre \hat{u}_t e u_t (Reg_ Intercept), Coeficiente angular do ajuste linear entre \hat{u}_t e u_t (Reg_Slope) e Coeficiente de indeterminação do ajuste linear entre \hat{u}_t e u_t (WR^2) e Métrica de desempenho agregado (LIMA JUNIOR et al., 2010; FIRMINO; MATTOS NETO; FERREIRA, 2014; FIRMINO et al., 2020; MATTOS NETO et al., 2020). No caso das métricas MSE, MAPE, Theil e ARV, quanto menor o valor dessas métricas, melhor é o desempenho do modelo.

3.5.1 Erro Quadrático Médio

O MSE é uma das medidas de qualidade de previsão mais típicas, verificando de forma ponderada a eficiência e a acurácia de um preditor. É determinado com base na diferença quadrática média entre os valores reais e preditos da série temporal. A métrica estima a variação do erro de previsão e expressa a acurácia e eficiência estatística pela soma do quadrado do erro. Por esse motivo, é um indicador muito mais sensível ao erro. Se o MSE resultante da Equação 3.48 for um valor grande indica a possibilidade de grandes erros de previsão.

$$MSE = \frac{\sum_{t=1}^N (u_t - \hat{u}_t)^2}{N} \quad (3.48)$$

em que, u_t é o valor observado, \hat{u}_t é o valor estimado e N o tamanho da série observada. Para Hyndman, Khandakar et al. (2007), essa medida é adequada para acessar a precisão do método de previsão em uma série de dados, porém não é significativa para determinar a exatidão de

um método de previsão em várias séries de dados.

3.5.2 Erro Médio Absoluto Percentual

O MAPE é uma das medidas mais utilizadas de precisão de previsão, devido às suas vantagens de independência de escala e interpretabilidade. A métrica mede a precisão do preditor em valores relativo (percentuais), conforme Equação 3.49 (KIM; KIM, 2016; HA; SEOK; OK, 2018).

$$MAPE = \frac{100}{N} \sum_{t=1}^N \left| \frac{u_t - \hat{u}_t}{u_t} \right|. \quad (3.49)$$

A medida proporciona uma indicação de quão grandes os erros de previsão estão na comparação com os valores atuais da série. Assim, essa medida é intuitivamente atraente, pois penaliza a sub e a superestimação, em relação ao resultado real, de maneira simétrica. No entanto, o MAPE é prejudicado por valores iguais ou próximos a zero nos dados observados, produzindo valores infinitos ou indefinidos. Com isso, quanto menor o MAPE melhor é o nosso ajuste (MAKRIDAKIS; WHEELWRIGHT; HYNDMAN, 1998; HYNDMAN; KHAN-DAKAR et al., 2007).

3.5.3 O Coeficiente U de Theil

A métrica U de Theil, também conhecida como coeficiente de desigualdade de Theil, é uma medida de acurácia constantemente usada na literatura. O coeficiente de Theil avalia o desempenho da previsão em relação à predição ingênua ou trivial, significando que o valor futuro estimado é igual ao valor atual (MIZRACH, 1992). Para Firmino, Mattos Neto e Ferreira (2015) essa medida é usada para comparar a performance do modelo com a performance de modelo *Random Walk* (passeio aleatório), que utiliza a observação anterior da série, u_{t-1} para prever u_t . A expressão matemática da métrica é dada conforme a Equação 3.50.

$$Theil = \frac{\sum_{t=1}^N (u_t - \hat{u}_t)^2}{\sum_{t=1}^N (u_t - \hat{u}_{t-1})^2} \quad (3.50)$$

A medida U de Theil utiliza-se de uma estratégia simples na análise da qualidade das previsões, tomando a observação mais recente como a melhor previsão do próximo valor da série, ou seja, a previsão um passo à frente do passeio aleatório (MAKRIDAKIS; WHEELWRIGHT; HYNDMAN, 1998). Se o coeficiente for equivalente a 1, o modelo apresenta desempenho igual ao *Random Walk* (desempenho igual a uma predição ingênua), e caso o valor seja menor

que 1, o desempenho do modelo é melhor que o desempenho do modelo *Random Walk* (uma previsão trivial). No entanto, se o valor da métrica for maior que 1, a previsão do modelo é pior que a previsão realizada pelo modelo ingênuo (OLIVEIRA; NOBRE; ZARATE, 2013; FIRMINO; MATTOS NETO; FERREIRA, 2015).

3.5.4 Variância Relativa Média

A medida ARV compara simplesmente as previsões do modelo com a média simples dos valores observados da série temporal. A medida é definida pela Equação 3.51

$$ARV = \frac{\sum_{t=1}^N (u_t - \hat{u}_t)^2}{\sum_{t=1}^N (\hat{u}_t - \bar{u}_t)^2} \quad (3.51)$$

em que, \bar{u}_t é a média dos valores observados da série temporal regressa até o instante t .

Se $ARV = 1$, o desempenho do modelo equivale ao de um modelo que sempre retorna a média da série como previsão. Caso $ARV < 1$, a previsão do modelo é melhor que a média, o desempenho é superior a média da série. E, se $ARV > 1$, o desempenho do modelo é considerado inferior à previsão realizada utilizando a média da série.

3.5.5 Predição de Alteração na Direção

A medida de desempenho POCID considera a precisão de se prever as alterações da direção (previsão da tendência da série), ou seja, se o valor futuro irá subir ou cair em relação ao valor atual, diferente das métricas anteriores, que se busca a minimização. No caso da POCID, a medida assume valores entre 0 e 1, e quanto maior o valor da métrica melhor será o modelo. A métrica é definida pelas Equações 3.52 e 3.53.

$$POCID = \frac{\sum_{t=1}^N D_t}{N} \quad (3.52)$$

$$D_t = \begin{cases} 1, & \text{se } (u_t - u_{t-1})(\hat{u}_t - \hat{u}_{t-1}) \geq 0 \\ 0, & \text{se } (u_t - u_{t-1})(\hat{u}_t - \hat{u}_{t-1}) < 0. \end{cases} \quad (3.53)$$

Da Equação 3.53, se $D_t = 1$ indica que o modelo acertou a predição da tendência da série temporal entre os instantes $t - 1$ e t . No entanto, se $D_t = 0$ indica que o modelo de previsão errou a predição da tendência (FIRMINO; MATTOS NETO; FERREIRA, 2014, 2015). A fim de padronizar a análise em relação às demais métricas de desempenho, o de minimizar essas métricas, adotar-se-á aqui $WPOCID = 1 - POCID$. Com isso, quanto maior o valor do

WPOCID, pior o modelo.

3.5.6 Índice de Desacordo

O Índice de Desacordo (ID) é uma medida complementar ao Índice de Concordância (IC) Willmott (1982), usado na comparação de diversos modelos de um mesmo fenômeno, independente da unidade de medida, uma vez que seu valor é dado no intervalo $[0, 1]$. Essa métrica estabelece a precisão do valor previsto em relação ao observado. Se $IC = 1$ a concordância é considerada perfeita, porém se $IC = 0$ indica total discordância entre observado e previsto. A métrica é apresentada nas Equações 3.54 - 3.56.

$$ID = 1 - IC \quad (3.54)$$

$$= 1 - \left\{ 1 - \frac{\sum_{t=1}^N (\hat{u}_t - u_t)^2}{\sum_{t=1}^N (|\hat{u}_t - \bar{u}_t| + |u_t - \bar{z}_t|)^2} \right\} \quad (3.55)$$

$$= \frac{\sum_{t=1}^N (\hat{u}_t - u_t)^2}{\sum_{t=1}^N (|\hat{u}_t - \bar{u}_t| + |u_t - \bar{z}_t|)^2} \quad (3.56)$$

No caso do ID ocorre o oposto: quanto mais próximo de zero estiver essa medida, melhor é o desempenho do modelo (FIRMINO et al., 2020).

3.5.7 Coeficientes do Modelo de Regressão Linear

Foram consideradas, também, métricas relacionadas a modelos de regressão (ajuste) linear, ajustado através do método de mínimos quadrados, tais como Intercepto do ajuste linear entre valor previsto \hat{u}_t e observado u_t (Reg_Intercept), Coeficiente de inclinação do ajuste linear entre valor previsto \hat{u}_t e observado u_t (Reg_Slope) e Coeficiente de indeterminação do ajuste linear entre valor previsto \hat{u}_t e observado u_t (WR^2). O coeficiente (WR^2) é o valor ajustado do coeficiente de determinação do ajuste linear R^2 , conforme Equação 3.57.

$$WR^2 = 1 - R^2 \quad (3.57)$$

O WR^2 é assim usado com o objetivo de padronizar a análise quando comparada as outras métricas de erros, ou seja, quanto menor o valor de WR^2 , melhor é a performance do modelo.

Considere uma equação da forma $u_t = a + b \cdot \hat{u}_t$, em que a e b representam os coeficientes Reg_Intercept e Reg_Slope, respectivamente. Assim, os coeficientes Reg_Intercept e

Reg_Slope representam os erros aditivos e multiplicativos das previsões \hat{u}_t para u_t geradas pelo modelo, respectivamente. Nesse caso, há um erro constante a independente da previsão e um erro proporcional b relacionado à previsão. Por sua vez, o coeficiente de determinação R^2 , expressa o desempenho do preditor na captura da variabilidade da série temporal observada (FIRMINO et al., 2020; MATTOS NETO et al., 2020). O R^2 é definido como (Equação 3.58)

$$R^2 = \frac{\sum_{t=1}^N (u_t - \hat{u}_t)^2}{\sum_{t=1}^N (u_t - \bar{u})^2} \quad (3.58)$$

em que, \bar{u} é a média da série temporal observada. Portanto, um preditor ideal apresentaria valores como $R^2 = 1$ ou $WR^2 = 0$, Reg_Intercept = 0 e Reg_Slope = 1, obtendo a equação $u_t = 0 + 1 \cdot \hat{u}_t$. O R^2 varia no intervalo de 0 a 1, e quanto maior o valor, melhor será o preditor.

3.5.8 Métrica de Desempenho Agregado

A métrica de desempenho agregado (APM) fornece uma análise geral das outras métricas descritas, normalizadas no intervalo $[0, 1]$. APM gera um resultado único para cada modelo, através do cálculo da média simples das outras métricas, pertencentes ao mesmo intervalo $[0, 1]$. O APM é baseado no raciocínio de que quanto mais próxima de 0 for seu valor, melhor será a performance do modelo, entretanto, quanto mais próxima de 1, pior será o desempenho do modelo (FIRMINO et al., 2020). Neste sentido, para APM (Equação 3.59), adotaram-se $|Reg_Intercept|$ e $||Reg_Slope| - 1|$ em substituição ao Reg_Intercept e Reg_Slope.

$$APM = \frac{\sum_{i=1}^m n.Metric_i}{m} \quad (3.59)$$

onde $n.Metric_i$ é a métrica i , normalizada de acordo com seus valores em relação aos modelos utilizados (ou seja, ANN, ARIMA, ETS, FUZZY e SVR) e m é o número de métricas ($m = 8$ reflete os erros MSE, MAPE, ARV, Theil'U, WPOCID, ID, $|Reg_Intercept|$ e $|Reg_Slope|$). A normalização das métricas de desempenho para o intervalo $[0, 1]$ ocorre por meio da Equação 3.60.

$$n.Metric_i = \frac{Metric_i - \min_i}{\max_i - \min_i} \quad (3.60)$$

em que \min_i e \max_i correspondem nessa ordem, aos valores mínimo e máximo observados para $Metric_i$ ao longo dos formalismos ajustados à série temporal sob estudo. As métricas consideradas no estudo têm como objetivo a minimização das equações que as expressam. De

forma geral, o melhor formalismo será o que minimizar a Equação 3.59.

3.5.9 Critério de Informação Bayesiano

O critério de informação *Bayesiano* (BIC) é também conhecido como critério de informação de *Schwarz* (CIS), do inglês *Schwarz's information criterion* (SIC), desenvolvido por Gideon E. Schwarz (SCHWARZ et al., 1978). É uma medida bem-sucedida para a seleção de modelos que pode ser interpretada como uma aproximação do fator Bayes (KASS; RAFTERY, 1995). Baseado em parte na função de distribuição de probabilidade e intimamente relacionado ao Critério de Informação Akaike (AIC). O BIC é um critério assintoticamente consistente para a seleção de modelos, se a distribuição por trás dos dados for regular (por exemplo, da família exponencial) e as anteriores no espaço de hipóteses e no espaço de parâmetros forem uniformes. A medida BIC é formulada pela Equação 3.61

$$BIC = -2 \log(ML) + k_{BIC} \cdot \log(N), \quad (3.61)$$

em que, k_{BIC} é o número de parâmetros do modelo, ML é o valor da função de máxima verossimilhança subjacente aos resíduos do preditor em face do ajuste da função de densidade de probabilidade correspondente e N é o tamanho da amostra.

O BIC parte do pressuposto de que existe um “modelo verdadeiro” que descreve a relação entre a variável dependente e as diversas variáveis explanatórias entre os diversos modelos sob seleção. Desse modo, o critério procura maximizar a probabilidade de se identificar o verdadeiro modelo dentre os avaliados. O modelo com menor valor para o BIC é considerado o de melhor ajuste (LIU et al., 1992; KONISHI; KITAGAWA, 2008; BOX; TIAO, 2011).

Esta medida tem o compromisso de parcimônia, buscando ponderar tanto a qualidade do modelo, a partir da primeira parcela da soma, quanto sua sofisticação em termos do número de parâmetros (segunda parcela da soma).

3.5.10 *Simulated de Annealing*

O recozimento simulado (do inglês *Simulated Annealing* – SAn) é uma das mais simples e conhecidas metaheurísticas de busca de vizinhança (busca local), usada para tratar de problemas de otimização discretos e, em menor grau, problemas de otimização contínua, baseada na simulação do processo de resfriamento de materiais. Sua origem se baseia na analogia entre o processo físico do resfriamento de um metal em estado de fusão e o problema de otimi-

zação, conforme Quadro 4. O SAn surgiu no contexto da mecânica estatística, desenvolvido

Quadro 4: Analogia entre o sistema físico e o problema de otimização

Simulação termodinâmica	Otimização
Estados do sistema	Soluções viáveis
Energia	Custo
Mudança de estado	Estrutura de vizinhança
Temperatura	Parâmetro de controle
Estado congelado	Solução heurística

Fonte: Elaborado pelo Autor.

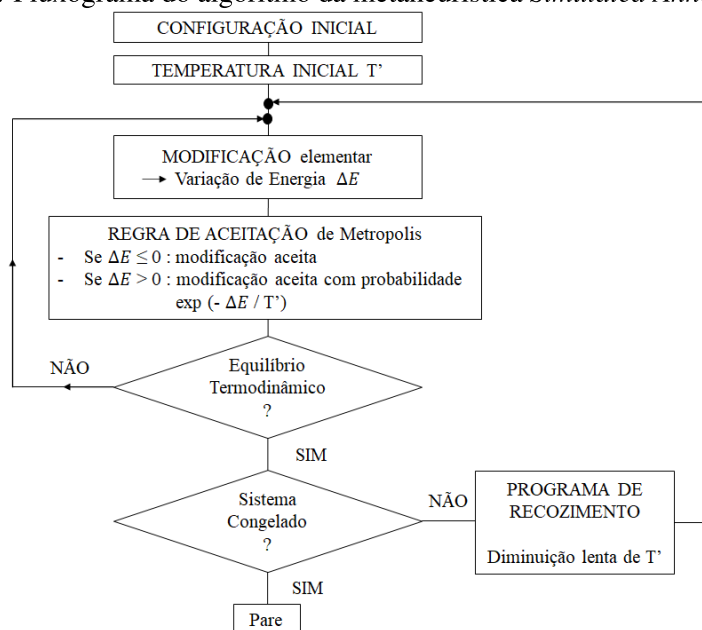
por Kirkpatrick, Gelatt e Vecchi (1983), e independentemente por Černý (1985), utilizando o algoritmo de simulação proposto por Metropolis et al. (1953). Neste contexto, o processo de otimização é executado simulando os níveis de temperatura no resfriamento. Em tese, o algoritmo parte de uma temperatura inicial definida chamada (T_s). Durante o processo, a temperatura (T') diminui gradativamente pelo mecanismo de um programa de resfriamento até atingir a temperatura final (T_e).

Em cada iteração, o algoritmo gera uma nova solução na vizinhança da solução atual. Então, comparando a adequação do vizinho com a solução atual, o SAn decide se aceita ou não a solução. O algoritmo aceita vizinhos superiores como a solução atual; mas no caso de soluções inferiores, ele aceita a solução pela probabilidade que é gerada pela função de Boltzmann ($e^{-\Delta/kT'}$). Nesta função, k é a parte constante, T' é a temperatura atual e Δ é a diferença entre a adequação da solução atual e da nova (METROPOLIS et al., 1953; KIRKPATRICK; GELATT; VECCHI, 1983).

O fluxograma do SAn é apresentado esquematicamente na Figura 18. A principal característica do SAn é que ele fornece um meio para escapar de ótimos locais, permitindo movimentos de escalada (isto é, movimentos que pioram o valor da função objetivo) na esperança de encontrar um ótimo global. Outra vantagem do algoritmo SAn é a facilidade de implementação computacional, possuindo uma baixa dependência de parâmetros (temperatura inicial, número de iterações ou limite de tempo de processamento e fator de redução de temperatura) (DELAHAYE; CHAIMATANAN; MONGEAU, 2019).

Nos casos apresentados nesta tese, a função objetivo que se busca minimizar é o MSE (Equação 3.48), uma métrica utilizada na seleção de modelos. Os sistemas para os casos FUZZY, ANN e SVR (individual e combinado) são, respectivamente, o modelo de ANN e o modelo de SVR. Em ANN, a saída é o vetor $p = (p_{AR}, p_{SAR}, S, N_o, N_c, LA, AF)$ em que p_{AR} são os valores autoregressivos defasados, p_{ARS} são os valores autoregressivos defasados sazonais, S é a sazonalidade, N_o indica o número de nós nas camadas intermediárias, N_c indica o número de camadas intermediárias, LA refere-se ao algoritmo de aprendizagem e AF a função de ativação, que é o conjunto de parâmetros que minimiza a função objetivo. Foram

Figura 18: Fluxograma do algoritmo da metaheurística *Simulated Annealing* (SAn)



Fonte: Adaptado de Dréo et al. (2006).

consideradas as seguintes variações para o algoritmo de aprendizagem (*backpropagation*): BPROP, RPROP+, RPROP-, SAG e SLR. Como variantes da função de ativação foram utilizadas as funções Logística (LF) e Tangente Hiperbólica (THF).

Considerando o SVR, a saída é representada pelo vetor $p = (p_{AR}, p_{ARS}, S, cost, \epsilon, degree, coef0, kernel, SV)$ onde p_{AR} , p_{ARS} , S conforme descritos para ANN, $cost$ é o custo, ϵ é responsável pelo tubo ao redor da função de regressão, γ , $degree$ e $coef0$ são parâmetros das funções Kernel, exceto a linear, $kernel$ é a função Kernel e SV , o número de vetores de suporte, que é o conjunto de parâmetros que minimiza a função objetivo. A variação para o custo se deu no intervalo $[10^{-5}, 10^2]$, a do ϵ entre $[10^{-5}, 1]$, a do γ em $[10^{-5}, 10^4]$, do $degree$ em $[1, 5]$ e para o $coef0$ tomou-se valores no intervalo $[0, 10]$. Esses valores foram considerados a partir de uma busca na literatura pelos intervalos para otimização dos parâmetros de SVR. As funções Kernel diferenciaram entre linear (LK), polinomial (PK), sigmoide (SK) e de base radial (RBFK).

O vetor de saída do FUZZY é composto de 5 componentes $p = (D_1, D_2, n, w e C)$. D_1 e D_2 são usados no ajuste dos limites de U , n representa número de conjuntos fuzzy, w número de elementos do conjunto de dados usados como informação prévia para prever os dados e C uma constante escolhida para garantir a conversão das observações em valores fuzzy, definida no intervalo $(0,1)$.

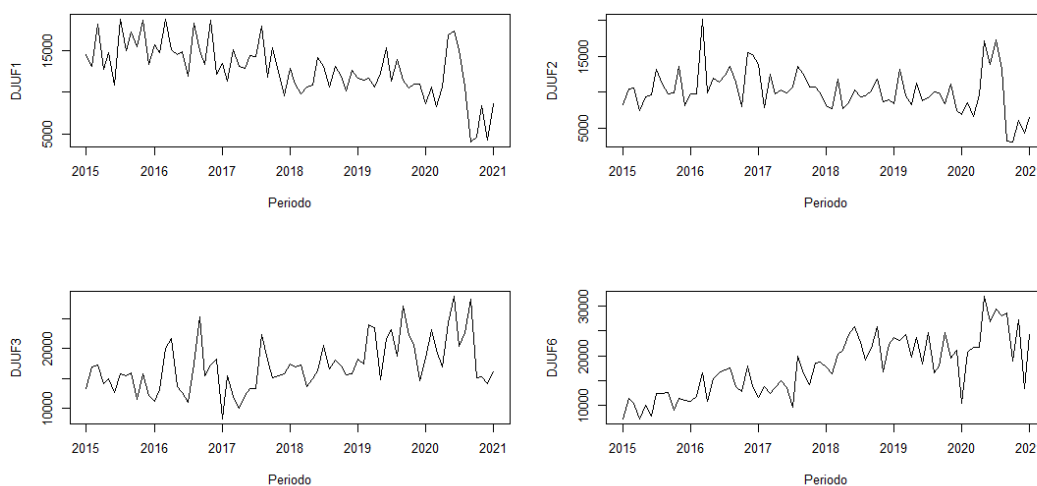
4 APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

Neste capítulo são apresentadas as modelagens das séries temporais de demanda utilizadas na presente pesquisa, conforme os modelos descritos na Seção 3.4 do Capítulo 3. Em um primeiro momento, na Seção 4.1, apresentam-se as séries temporais de demanda da família de produtos. Na sequência, Seção 4.2 são mostrados os parâmetros iniciais das modelagens. Por fim, a Seção 4.3 restringe-se as modelagens, resultados alcançados e comparação dos melhores modelos, dada a experimentação das séries temporais objeto de estudo.

4.1 Séries temporais de demanda das famílias de produtos

As séries temporais de demanda são constituídas por famílias de produtos conforme os critérios estabelecidos pela empresa como tipo de produto, peso, capacidade do recipiente, quantidade de produtos por caixa e/ou fragrância. As séries objeto de estudo aqui identificadas como DJUF1, DJUF2, DJUF3, DJUF6, DJUF7, DJUF11, DJUF12 e DJUF13 possuem 73 observações com periodicidade mensal, exceto a série DJUF7 com 67 observações (ver Tabela 1). Nas Figuras 19 e 20 é possível observar o comportamento das séries temporais de demanda dos produtos graficamente.

Figura 19: Séries temporais de demanda das famílias de produtos DJUF1, DJUF2, DJUF3 e DJUF6



Fonte: Elaborado pelo autor.

Figura 20: Séries temporais de demanda das famílias de produtos DJUF7, DJUF11, DJUF12 e DJUF13



Fonte: Elaborado pelo autor.

4.2 Parâmetros iniciais das modelagens

Para realização da pesquisa foram utilizadas 8 séries temporais de demanda, com o intuito de analisar o comportamento dos modelos propostos sob as perspectivas das famílias de produtos da empresa objeto de estudo.

As séries temporais foram separadas em conjuntos de treinamento e de teste, estratégia que reflete a prática de treinar o modelo usando dados do passado, gerando as previsões de demanda e comparando com os valores reais, por meio das medidas de acurácia. Os dados de treinamento são usados para estimar os parâmetros do método de previsão e os dados de teste são usados para avaliar a precisão. Assim, os dados de teste não são usados para treinar os preditores e fornecem uma indicação confiável de quão bem o modelo gera as previsões (HYNDMAN; ATHANASOPOULOS, 2018).

Com isso, foram propostos e gerados 80 modelos no total, sendo 10 para cada série temporal de demanda, 5 individuais e 5 combinados. Além disso, tem-se o uso da metodologia de previsão da empresa objeto de estudo, gerando mais 8 modelos preditivos. Usada com o objetivo de avaliar a qualidade do modelo em comparação aos propostos.

Na sequência, serão apresentadas a estrutura e a performance de cada modelo em cada série. Para experimentação, processamento e geração dos modelos foi utilizado um notebook com sistema operacional *Windows 10 Home* (64 bits), processador Intel i5, com 2.3GHz e 8GB de memória RAM de configuração.

Para geração dos modelos, foram considerados alguns parâmetros usados no ajuste dos formalismos na etapa de modelagem das séries, mostrados na Tabela 2.

Tabela 2: Parâmetros de ajuste dos formalismos de previsão individuais e combinados para cada série temporal considerada (DJUF1, DJUF2, DJUF3, DJUF6, DJUF7, DJUF11, DJUF12, DJUF13)

Parâmetro	Valor	Formalismo
Proporção de treinamento	0.8	Todos
Proporção AR e ARS	5e-02	ANN e SVR
ANN.threshold	1e-03	ANN
ANN.stepmax	6e+06	ANN
Número de modelos	5e+03	ARIMA e ETS
FUZZY.D ₁	1	FUZZY
FUZZY.D ₂	1	FUZZY
FUZZY.n	5	FUZZY
FUZZY.w	2	FUZZY
FUZZY.C	0.5	FUZZY
GSA.max.call	5e+03	ANN, FUZZY, SVR, cANN e cSVR
GSA.max.time	600	ANN, FUZZY, SVR, cANN e cSVR
GSA.maxit	1e+03	ANN, FUZZY, SVR, cANN e cSVR
GSA.temperature	1e+03	ANN, FUZZY, SVR, cANN e cSVR
GSA.nb.stop.improvement	20	ANN, FUZZY, SVR, cANN e cSVR

Fonte: Elaborado pelo Autor.

Os formalismos combinados cMV, cSA e cSM não precisam de ajuste de parâmetros na etapa de modelagem, sendo assim, não foram considerados para a construção da Tabela 2. Também, na apresentação que se segue dos modelos e suas respectivas descrições, esses não aparecem por não apresentarem parâmetros passíveis de interpretação, exceto cMV.

4.3 Modelagem, Análise e Comparação dos Métodos

Na Tabela 3, encontram-se descritas as arquiteturas dos modelos individuais e combinados propostos no sistema de previsão de demanda para as séries temporais das famílias de produtos DJUF1, DJUF2, DJUF3 e DJUF6. O modelo ARIMA apresentou para as quatro séries a mesma estrutura, ordem de integração igual a um, indicando a necessidade de apenas uma diferenciação para estabilizar os valores observados da série e um termo de média móvel. Esse modelo gera previsões equivalentes ao modelo de suavização exponencial simples, onde $q = 1$ é o parâmetro de suavização. O modelo ETS obteve um arcabouço para a série DJUF1 com erro multiplicativo, tendência aditiva e nenhuma sazonalidade. O modelo equivale ao método linear de Holt com erros multiplicativos. Assim, os modelos estatísticos ARIMA e ETS foram matematicamente semelhantes para série DJUF1. Já para as séries DJUF2, DJUF3 e DJUF6, o modelo ETS indicou um erro multiplicativo e nenhuma sazonalidade ou tendência (suavização exponencial simples com erros multiplicativos). Logo, o modelo se assemelha ao modelo de *random walk*.

Por sua vez, os modelos de *Machine Learning*, ANN e SVR apresentaram de uma a três entradas autorregressiva para as séries de demanda, o que indica que foram utilizados valores anteriores como u_{t-1} a u_{t-3} para realizar as previsões das séries no instante t (u_t). A rede neural ainda usou a função de ativação Tangente Hiperbólica nas séries DJUF1 e DJUF2 e logística nas séries DJUF3 e DJUF6. Já o SVR foi estruturado com *kernel* polinomial (DJUF1), linear (DJUF2), base radial (DJUF3) e sigmoideal (DJUF6) e com uma quantidade de vetores de suporte que variam de 11 a 37. O modelo de inteligência computacional FUZZY, dividiu as séries temporais em 12 a 17 conjuntos difusos, com uma constante de associação $C = 1$ e utilizou de 5 a 27 diferenças nebulosas para predição.

Os modelos combinados consideram como entradas as previsões dos seis formalismos individuais, sendo 5 propostos mais o preditor da empresa. As combinações cANN apresentam arquitetura que levam de 4 a 10 nós em cada uma das camadas intermediárias, a função de ativação logística é utilizada na modelagem das quatro séries temporais de demanda das famílias de produtos. O combinador cSVR obteve modelos com *kernel* linear (DJUF2 e DJUF3) e polinomial (DJUF1 e DJUF6) e com um número de vetores de suporte que variam de 11 a 21. Já o combinador cMV atribuiu pesos próximos para os modelos ANN, ARIMA e ETS e para FUZZY e JUA na predição da série DJUF2. Indicando que esses modelos apresentam variâncias de erros de suas previsões e correlações próximas. Em séries como DJUF3 e DJUF6 o combinador cMV tem forte influência dos preditores ANN, ARIMA e ETS. Ainda é possível verificar por meio do cMV nas séries a diversidade dos pesos associados aos formalismos individuais.

Na Tabela 4, são apresentadas as métricas de desempenho dos formalismos individuais e combinadores do sistema de previsão de demanda ora proposto, para efeitos de comparação das séries de demanda dos produtos DJUF1, DJUF2, DJUF3 e DJUF6. A metodologia usada pela empresa combinando dados quantitativos com informações subjetivas para prever a demanda tem gerado resultados de boa qualidade. A metodologia apresenta erros de previsão baixos e quando comparados com os erros produzidos por outros preditores identifica-se que os erros são os menores no geral (ver série DJUF2). O modelo exibe um WPOCID e Reg_WR^2 baixo para as quatro séries, indicando que o JUA acerta bem a previsão de tendência das séries, como também é um dos que mais se ajusta linearmente aos dados das séries. No entanto, o modelo não apresenta boa performance mensurada pelo MSE, que é um indicador muito mais sensível ao erro. Quando se compara o MSE do JUA com os combinadores percebe-se uma diferença significativa, o que pode prejudicar o modelo no agregado. Dentre os preditores individuais propostos, o FUZZY se destaca nas séries DJUF1 e DJUF2, SVR na DJUF3 e o ANN na DJUF6. Os modelos apresentam ID pequeno demonstrando uma baixa

Tabela 3: Descrição dos modelos individuais e combinados para as séries temporais de demanda das famílias de produtos DJUF1, DJUF2, DJUF3 e DJUF6

Série	Formalismo	Modelo	Descrição
DJUF1	ARIMA	ARIMA(0, 1, 1)	-
	ETS	ETS(M, A, N)	$\alpha = 1e - 04$ e $\beta = 1e - 04$
	FUZZY	-	$D_1 = 8e - 02$, $D_2 = 0.0528$, $C = 1$, $n = 17$ e $w = 26$
	ANN	ANN(5, 3, 1)	$P_{AR} = 3$, $P_{ARS} = 2$, $S = 6$, $AF = \text{TANH}$ e $LA = \text{BACKPROP}$
	SVR	-	$P_{AR} = 2$, $P_{ARS} = 1$, $cost = 26.0025$, $\epsilon = 0.0943$, $\gamma = 2653.7542$, $Kernel = PK$, $SV = 37$
	cANN	cANN(6, 4, 2)	$AF = LF$ e $LA = SAG$
cMV	-	$\omega_{ANN} = 0.7967$, $\omega_{ARIMA} = -0.2095$, $\omega_{ETS} = 0.3801$, $\omega_{FUZZY} = 0.2303$, $\omega_{JUA} = 0.1909$ e $\omega_{SVR} = -0.3886$	
cSVR	-	$cost = 29.1408$, $\epsilon = 0.7281$, $\gamma = 8201.9293$, $Kernel = PK$ e $SV = 11$	
DJUF2	ARIMA	ARIMA(0, 1, 1)	-
	ETS	ETS(M, N, N)	$\alpha = 0.1284$
	FUZZY	-	$D_1 = 8e - 02$, $D_2 = 0.0788$, $C = 1$, $n = 12$ e $w = 15$
	ANN	ANN(3, 2, 2)	$P_{AR} = 2$, $P_{ARS} = 1$, $S = 8$, $AF = \text{TANH}$ e $LA = \text{RPROP+}$
	SVR	-	$P_{AR} = 1$, $P_{ARS} = 2$, $cost = 28.9455$, $\epsilon = 0.8064$, $\gamma = 4403.1882$, $Kernel = LK$ e $SV = 15$
	cANN	cANN(6, 8, 1)	$AF = LF$ e $LA = \text{RPROP+}$
cMV	-	$\omega_{ANN} = 0.3775$, $\omega_{ARIMA} = 0.3623$, $\omega_{ETS} = -0.3699$, $\omega_{FUZZY} = 0.2911$, $\omega_{JUA} = 0.3043$ e $\omega_{SVR} = 0.0348$	
cSVR	-	$cost = 32.2695$, $\epsilon = 0.584$, $\gamma = 6618.6286$, $Kernel = LK$ e $SV = 21$	
DJUF3	ARIMA	ARIMA(0, 1, 1)	-
	ETS	ETS(M, N, N)	$\alpha = 0.1424$
	FUZZY	-	$D_1 = 8e - 02$, $D_2 = 0.0987$, $C = 1$, $n = 11$ e $w = 27$
	ANN	ANN(4, 6, 2)	$P_{AR} = 2$, $P_{ARS} = 2$, $S = 8$, $AF = LF$ e $LA = SAG$
	SVR	-	$P_{AR} = 3$, $P_{ARS} = 2$, $cost = 28.5561$, $\epsilon = 0.9325$, $\gamma = 5265.6064$, $Kernel = \text{RBFK}$, $SV = 11$
	cANN	cANN(6, 10, 1)	$AF = LF$ e $LA = \text{RPROP-}$
cMV	-	$\omega_{ANN} = -2.7588$, $\omega_{ARIMA} = 2.2843$, $\omega_{ETS} = 0.5081$, $\omega_{FUZZY} = 0.021$, $\omega_{JUA} = 0.2581$ e $\omega_{SVR} = 0.6874$	
cSVR	-	$cost = 53.5328$, $\epsilon = 0.7175$, $\gamma = 5361.837$, $Kernel = LK$ e $SV = 9$	
DJUF6	ARIMA	ARIMA(0, 1, 1)	-
	ETS	ETS(M, N, N)	$\alpha = 0.3212$
	FUZZY	-	$D_1 = 8e - 02$, $D_2 = 0.1032$, $C = 1$, $n = 16$ e $w = 5$
	ANN	ANN(2, 6, 3)	$P_{AR} = 2$, $P_{ARS} = 0$, $S = 8$, $AF = LF$ e $LA = \text{RPROP+}$
	SVR	-	$P_{AR} = 3$, $P_{ARS} = 1$, $cost = 0.1124$, $\epsilon = 0.7587$, $\gamma = 2287.5689$, $Kernel = SK$ e $SV = 16$
	cANN	cANN(6, 8, 1)	$AF = LF$, $LA = \text{RPROP-}$
cMV	-	$\omega_{ANN} = 0.2453$, $\omega_{ARIMA} = 1.7712$, $\omega_{ETS} = -0.7738$, $\omega_{FUZZY} = -0.7699$, $\omega_{JUA} = 0.5472$ e $\omega_{SVR} = -0.0199$	
cSVR	-	$cost = 70.3051$, $\epsilon = 0.5643$, $\gamma = 9093.3094$, $Kernel = PK$ e $SV = 11$	

Fonte: Elaborado pelo autor.

discordância entre os valores previstos e observados para as séries. Geram previsões melhores que o modelo de passeio aleatório (Theil < 1). E, assim como JUA, apresentam WPOCID tendendo a zero.

Considerando os combinadores observa-se boa performance destes, como se pode verificar cSVR nas séries DJUF1 e DJUF6, cMV nas séries DJUF2 e DJUF3 e cSA nas séries DJUF1 e DJUF3. Os modelos apresentam os menores valores para MSE, que é um indicador

mais sensível ao erro. Quando verifica-se a proporção dos erros de previsão em comparação com os valores atuais da série mensurado pelo MAPE, nota-se um bom ajuste dos combinadores. Ainda assim, quando se analisa a precisão entre o previsto e observado os formalismos apresentam um bom nível de concordância mensurado pelo ID. Essas descobertas demonstram a qualidade preditiva desses modelos quando comparados com os outros, apontando que a previsão realizada pelas combinações é superior à dos preditores individuais, incluindo a metodologia JUA da indústria de manufatura.

Tabela 4: Desempenho dos modelos de previsão (ANN, ARIMA, ETS, FUZZY, SVR, JUA, cSA, cSM, cANN, cMV, cSVR) ao prever as séries temporais DJUF1, DJUF2, DJUF3 e DJUF6 (fase de teste). Os melhores valores encontram-se em negrito

Série	Métrica	FUZZY	ANN	ARIMA	ETS	SVR	JUA	cSA	cSM	cANN	cMV	cSVR
DJUF1	MSE	12994910	20590891	17681669	16214436	14487579	19218222	9252133	13108878	11209007	10190982	11012551
	MAPE	0.400	0.472	0.508	0.454	0.405	0.357	0.349	0.388	0.377	0.358	0.309
	ARV	1.040	4.804	20.511	7.483	3.116	0.526	3.624	5.303	6.892	4.251	1.839
	ID	0.279	0.648	0.855	0.649	0.469	0.217	0.342	0.495	0.486	0.395	0.307
	Theil	0.947	1.530	1.313	1.205	1.077	1.428	0.688	0.974	0.833	0.758	0.812
	WPOCID	0.571	0.429	0.500	0.500	0.500	0.071	0.429	0.357	0.214	0.214	0.143
	Reg_ Intercept	3662	7036	2760	-36666	2935	2393	-2941	339	-5151	-2921	-3340
	Reg_ Slope	0.632	0.258	0.644	4.672	0.686	0.576	1.217	0.927	1.431	1.235	1.463
	Reg_ WR2	0.614	0.975	0.944	0.810	0.790	0.201	0.487	0.745	0.578	0.554	0.559
	DJUF2	MSE	14462983	25075380	20665264	20679826	23650283	11913113	11419780	18387247	8993069	7096219
MAPE		0.437	0.740	0.625	0.625	0.719	0.419	0.487	0.594	0.428	0.394	0.443
ARV		1.020	4.638	22.319	22.740	9.755	0.394	9.350	27.278	3.533	1.435	5.692
ID		0.272	0.785	0.900	0.911	0.899	0.140	0.454	0.860	0.295	0.191	0.394
Theil		0.938	1.635	1.355	1.355	1.541	0.767	0.749	1.205	0.589	0.465	0.705
WPOCID		0.714	0.429	0.500	0.500	0.571	0.071	0.214	0.214	0.071	0.143	0.143
Reg_ Intercept		3295	10874	5959	6152	84076	1867	-12504	-10433	-6238	-5544	-5776
Reg_ Slope		0.638	-0.190	0.306	0.285	-7.369	0.650	2.166	1.999	1.604	1.423	1.581
Reg_ WR2		0.599	0.998	0.993	0.995	0.661	0.189	0.330	0.860	0.345	0.207	0.461
DJUF3		MSE	33224760	28050772	28209131	28633719	32200932	29077384	20431770	22834385	23845825	44066052
	MAPE	0.235	0.246	0.281	0.282	0.244	0.233	0.224	0.337	0.203	0.354	0.280
	ARV	1.094	2.676	12.171	12.921	1.852	0.982	3.446	4.156	1.603	4.469	6.716
	ID	0.340	0.476	0.790	0.826	0.435	0.286	0.406	0.476	0.332	0.749	0.651
	Theil	0.968	0.806	0.854	0.866	0.932	0.903	0.606	0.667	0.681	1.365	0.825
	WPOCID	0.714	0.643	0.571	0.571	1.000	0.357	0.429	0.429	0.429	0.214	0.643
	Reg_ Intercept	11150	6484	2902	3186	19134	5610	-5070	-7222	-10906	-8449	4190
	Reg_ Slope	0.445	0.691	0.805	0.792	0.612	1.253	1.372	1.710	1.163	1.163	0.731
	Reg_ WR2	0.808	0.940	0.942	0.959	1.000	0.506	0.697	0.778	0.691	0.801	0.875
	DJUF6	MSE	56685346	38796578	35869976	35861876	41927151	25442206	28653819	32433840	27123679	23079167
MAPE		0.289	0.247	0.242	0.242	0.221	0.216	0.214	0.231	0.202	0.191	0.179
ARV		1.802	2.137	6.378	6.791	2.377	0.728	3.408	4.668	2.038	1.745	2.151
ID		0.707	0.552	0.789	0.800	0.567	0.256	0.577	0.705	0.420	0.368	0.413
Theil		0.951	0.701	0.676	0.677	0.807	0.450	0.523	0.599	0.496	0.444	0.401
WPOCID		0.643	0.429	0.571	0.571	0.786	0.357	0.643	0.500	0.500	0.286	0.643
Reg_ Intercept		20012	12095	18626	18996	63858	10769	9175	14341	-866	7775	5110
Reg_ Slope		0.154	0.562	0.210	0.194	-1.975	0.543	0.651	0.411	1.164	0.682	0.759
Reg_ WR2		0.970	0.912	0.983	0.987	0.874	0.582	0.869	0.947	0.706	0.730	0.686

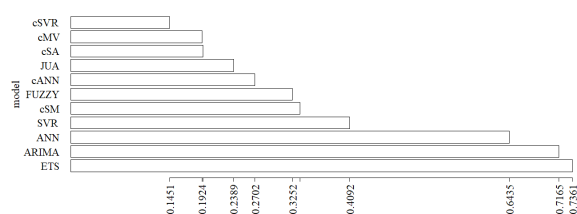
Fonte: Elaborado pelo autor.

A Figura 21 apresenta a média agregada normalizada de desempenho dos formalismos individuais e combinados propostos no sistema de previsão de demanda para as séries temporais dos produtos DJUF1, DJUF2, DJUF3 e DJUF6. Para agregar a análise dessas medidas de desempenho é usada a Média Agregada Normalizada que agrega todas as medidas de erro, normalizadas no intervalo [0,1], e calcula a média simples destas, gerando um resultado único para cada modelo, também com valores entre 0 e 1.

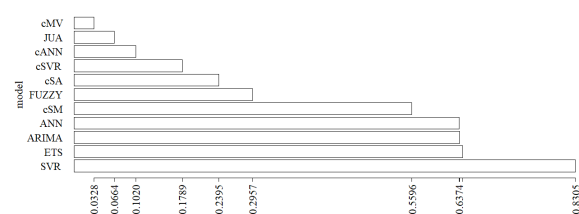
Dada a análise da performance agregada, é possível confirmar o bom desempenho dos formalismos combinados nas séries DJUF1, DJUF2, DJUF3 e DJUF6. Os modelos propostos, a partir do sistema de previsão de demanda objeto desta tese, se qualificam como bons estimadores para a demanda da indústria, descrevendo bem o processo gerador da série. Em

resumo, o uso dos combinadores cSVR, cMV e cSA na série DJUF1 proporcionam a empresa uma redução nos erros de predição de até 64,64% (Figura 22(a)) quando comparado a metodologia usada atualmente pela empresa. Para a série DJUF2 o combinador apresenta uma média agregada de 0,0328 versus 0,0664 do JUA. Uma redução do erro de predição de mais de 102% (Figura 22(b)). Considerando a série DJUF3, o combinador cSA produz um erro 0,1695 o que equivale a uma redução de aproximadamente 13% (Figura 22(c)). O modelo cSVR do sistema de previsão proposto é mais qualificado para processo de previsão da demanda da família DJUF6. O modelo promove uma redução de aproximadamente 29% do erro médio agregado (Figura 22(d)). Com isso, entende-se a relevância dos resultados encontrados a partir do uso do sistema de previsão ora proposto e que a adoção do sistema em substituição a metodologia adotada pela indústria é sugerida, com o propósito de melhorar a previsão da demanda das famílias de produtos, bem como o planejamento da produção.

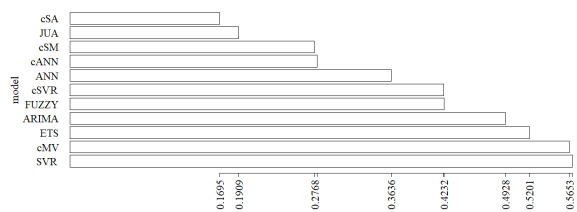
Figura 21: Média agregada normalizada de performance dos modelos individuais e combinados para as séries de demanda de produtos DJUF1, DJUF2, DJUF3 e DJUF6



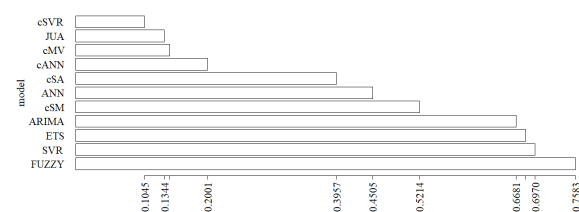
(a) Média agregada normalizada de performance para série DJUF1



(b) Média agregada normalizada de performance para série DJUF2



(c) Média agregada normalizada de performance para série DJUF3



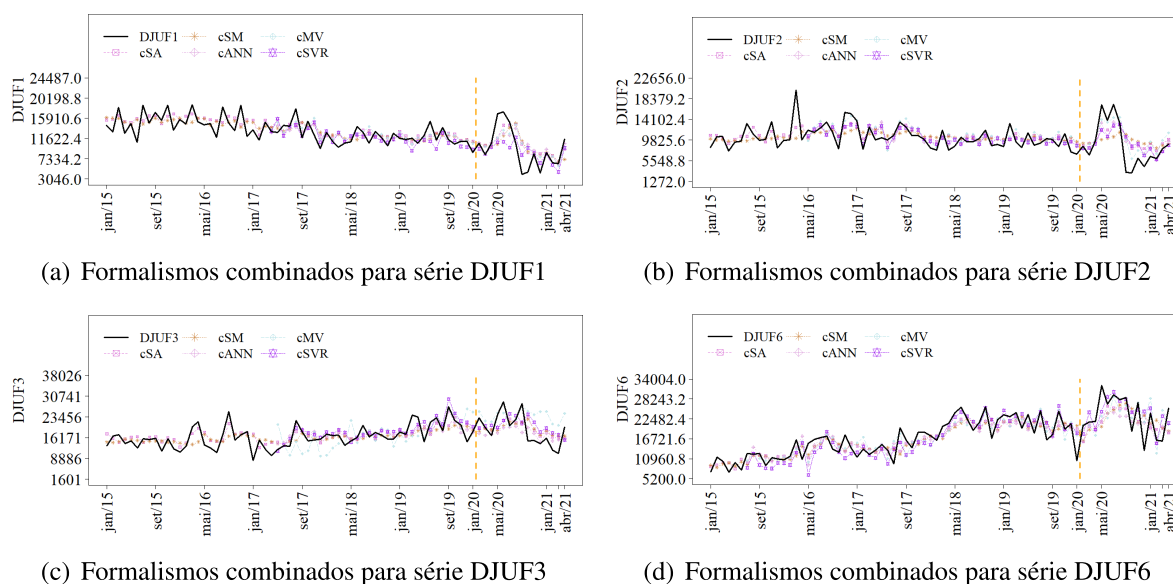
(d) Média agregada normalizada de performance para série DJUF6

Fonte: Elaborado pelo autor.

Na Figura 22 são mostradas graficamente as séries de demanda e as previsões para as famílias de produtos DJUF1, DJUF2, DJUF3 e DJUF6 realizadas por meio dos combinadores. É possível verificar que os formalismos acompanham a tendência da série, assim como é possível visualizar a aderência dos modelos aos dados da série.

Na Tabela 5 são mostrados os parâmetros dos modelos individuais e combinados para as séries temporais de demanda da família de produtos DJUF7, DJUF11, DJUF12 e DJUF13. O

Figura 22: Formalismos combinados para as séries temporais de demanda das famílias de produtos DJUF1, DJUF2, DJUF3 e DJUF6 (Fase de treinamento e teste)



Fonte: Elaborado pelo autor.

modelo de Box & Jenkins apresentou uma estrutura igual para as séries DJUF7, DJUF12 e DJUF13. A estrutura consiste de ordem de integração igual a um, necessária para estabilizar a série, e um termo de média móvel. Esse modelo é equivalente ao ETS(A,N,N) com erros aditivos, sem tendência ou sazonalidade. Para a série DJUF11 o modelo ARIMA não apresentou termos autorregressivos ou indicou necessidade de diferenciações, tornando-se um modelo de médias móveis de 1 termo.

O modelo ETS diagnosticou erro e tendência multiplicativos para a série DJUF7. Para as outras séries o modelo é uma suavização exponencial simples com erros multiplicativos. A modelagem de lógica difusa exibiu o valor da constante de associação $C = 0.5537$ para a série DJUF7, para as demais o $C = 1$. O modelo obteve de 12 a 16 conjuntos difusos e para a série DJUF11 usou 18 diferenças nebulosas para previsão. Os formalismos de inteligência artificial ANN e SVR apresentaram arquitetura com até três entradas autorregressivas e oito nós nas camadas intermediárias para realizar a previsão de um passo a frente. O SVR ainda obteve kernel linear (DJUF7 e DJUF11), de base radial radial (DJUF12) e sigmoidal (DJUF13).

Considerando os combinadores, o cMV aponta para uma diversidade de pesos atribuídos aos estimadores individuais. Em poucos casos, o modelo exhibe ponderação próxima para os preditores individuais. Essa diversidade de ponderação indica um distanciamento nas variações de erro de suas previsões e correlações. A arquitetura do cANN utiliza de 4 a 9 nós nas camadas intermediárias, uma função de ativação logística comum às quatro séries e uma

variedade para o algoritmo de aprendizagem usado no treinamento da rede. Por fim, a técnica de aprendizado de máquinas como combinador cSVR, aplicada às séries, apresenta na sua estrutura de 10 a 20 vetores de suporte usados para determinar os fundamentos para a função de regressão estimada.

Tabela 5: Descrição dos modelos individuais e combinados para as séries temporais de demanda das famílias de produtos DJUF7, DJUF11, DJUF12 e DJUF13

Série	Formalismo	Modelo	Descrição
DJUF7	ARIMA	ARIMA(0, 1, 1)	-
	ETS	ETS(M, M, N)	$\alpha = 1e - 04$ e $\beta = 1e - 04$
	FUZZY	-	$D_1 = 0.0799, D_2 = 0.0714, C = 0.5537, n = 15$ e $w = 2$
	ANN	ANN(4, 7, 3)	$P_{AR} = 3, P_{ARS} = 1, S = 7, AF = LF$ e $LA = SLR$
	SVR	-	$P_{AR} = 2, P_{ARS} = 1, cost = 32.2801, \epsilon = 0.3281, \gamma = 8988.3696, Kernel = LK$ e $SV = 12$
	cANN	cANN(6, 4, 1)	$AF = LF, LA = SLR$
cMV	-	$\omega_{ANN} = 0.8704, \omega_{ARIMA} = -1.574, \omega_{ETS} = 1.0357, \omega_{FUZZY} = 0.3466, \omega_{SVR} = -0.0123$ e $\omega_{JUA} = 0.3337$	
cSVR	-	$cost = 0.9005, \epsilon = 0.2805, \gamma = 9575.7606, Kernel = LK$ e $SV = 10$	
DJUF11	ARIMA	ARIMA(0, 0, 1)	-
	ETS	ETS(M, N, N)	$\alpha = 0.1058$
	FUZZY	-	$D_1 = 8e - 02, D_2 = 0.0950, C = 1, n = 16$ e $w = 18$
	ANN	ANN(1, 8, 3)	$P_{AR} = 1, P_{ARS} = 0, S = 3, AF = TANH$ e $LA = RPROP-$
	SVR	-	$P_{AR} = 2, P_{ARS} = 0, cost = 88.2847, \epsilon = 0.5444, \gamma = 1566.8523, Kernel = LK$ e $SV = 24$
	cANN	cANN(6, 4, 1)	$AF = LF, LA = RPROP+$
cMV	-	$\omega_{ANN} = 0.8602, \omega_{ARIMA} = -0.0985, \omega_{ETS} = 2.1501, \omega_{FUZZY} = -0.1514, \omega_{SVR} = -1.9525$ e $\omega_{JUA} = 0.1921$	
cSVR	-	$cost = 0.0408, \epsilon = 0.2682, \gamma = 4605.4605, Kernel = RBFK$ e $SV = 20$	
DJUF12	ARIMA	ARIMA(0, 1, 1)	-
	ETS	ETS(M, N, N)	$\alpha = 0.1389$
	FUZZY	-	$D_1 = 0.0712, D_2 = 0.12, C = 1, n = 13$ e $w = 4$
	ANN	ANN(3, 6, 2)	$P_{AR} = 3, P_{ARS} = 0, S = 9, AF = TANH$ e $LA = SLR$
	SVR	-	$P_{AR} = 3, P_{ARS} = 2, cost = 33.9724, \epsilon = 0.9998, \gamma = 6409.9987, Kernel = RBFK$ e $SV = 42$
	cANN	cANN(6, 6, 5)	$AF = LF$ e $LA = RPROP-$
cMV	-	$\omega_{ANN} = 0.6263, \omega_{ARIMA} = -0.4384, \omega_{ETS} = -0.6556, \omega_{FUZZY} = 0.1326, \omega_{SVR} = 1.1585$ e $\omega_{JUA} = 0.1763$	
cSVR	-	$cost = 77.4589, \epsilon = 0.6578, \gamma = 8335.6102, Kernel = LK$ e $SV = 10$	
DJUF13	ARIMA	ARIMA(0, 1, 1)	-
	ETS	ETS(M, N, N)	$\alpha = 0.1953$
	FUZZY	-	$D_1 = 8e - 02, D_2 = 0.1082, C = 1, n = 12$ e $w = 15$
	ANN	ANN(4, 4, 3)	$P_{AR} = 3, P_{ARS} = 1, S = 8, AF = LF$ e $LA = SLR$
	SVR	-	$P_{AR} = 1, P_{ARS} = 0, cost = 3.011, \epsilon = 0.3529, \gamma = 4087.6465, Kernel = SK, SV = 33$
	cANN	cANN(6, 9, 6)	$AF = LF$ e $LA = SLR$
cMV	-	$\omega_{ANN} = 0.3288, \omega_{ARIMA} = -1.8167, \omega_{ETS} = 1.9079, \omega_{FUZZY} = 5e - 04, \omega_{SVR} = 0.13$ e $\omega_{JUA} = 0.4496$	
cSVR	-	$cost = 11.2751, \epsilon = 0.9003, \gamma = 143.5928, Kernel = PK$ e $SV = 11$	

Fonte: Elaborado pelo autor.

A Tabela 6 mostra o desempenho de todos os formalismos individuais e combinados integrantes do sistema de previsão de demanda proposto para as séries de demanda das famílias

de produtos DJUF7, DJUF11, DJUF12 e DJUF13. Considerando os preditores individuais, a metodologia de previsão JUA adotada pela indústria de manufatura apresenta uma boa performance. Verifica-se que o modelo estima valores para métrica ARV < 1 para todas as séries, indicando que seu desempenho é superior a média da série, único entre todos os formalismos a apresentar essa condição. O modelo ainda apresenta o menor valor do ID para séries DJUF7, DJUF12 e DJUF13, sinalizando uma baixa discordância entre os valores previstos e observados para as séries. No entanto, o modelo exibe um Theil > 1 para 3 séries, apontando que a previsão realizada pelo estimador é pior que a previsão de um modelo de *random walk*.

Dos formalismos individuais propostos no sistema de previsão de demanda o ETS (DJUF7), o SVR (DJUF11) e ANN (DJUF12 e DJUF13), foram os que apresentaram melhor desempenho. Os modelos acertam a predição da tendência da série temporal entre os instantes $t - 1$ e t (WPOCID < 1). No entanto, os modelos mostram desempenho considerado inferior à previsão realizada utilizando a média da série (ARV > 1) e também a metodologia de previsão usada pela empresa.

Quando verifica-se os preditores combinados, observa-se baixo erro mensurado por diversas das métricas utilizadas como MSE, MAPE e Theil para cSVR na série DJUF11, MSE, MAPE, Theil e WPOCID para cSA na série DJUF12 e MSE, Theil, WPOCID, Reg_ Intercept e Reg_Slop para cMV na série DJUF13, sendo estes valores os menores no geral para as séries. Esse achado mostra a estabilidade e os bons resultados preditivos desses modelos em comparação com os demais, embora apresentem ARV superior a um. A descoberta indica que a qualidade da predição realizada pelos combinadores é superior a dos preditores individuais, incluindo a metodologia JUA. A Figura 23 mostra a performance preditiva dos formalismos individuais e combinados propostos no sistema de previsão, considerando a média agregada normalizada das métricas de desempenho para as séries de demanda DJUF7, DJUF11, DJUF12 e DJUF13. É importante lembrar que a métrica de desempenho agregado fornece uma análise geral das outras métricas descritas, normalizadas no intervalo $[0; 1]$.

A partir disso, é possível constatar os bons resultados preditivos dos combinadores para as séries de demanda DJUF7, DJUF11, DJUF12 e DJUF13 considerados no sistema de previsão de demanda proposto, no qual, os combinadores se qualificam como os melhores modelos no *rank* geral de cada série, descrevendo bem o processo gerador da série e muitas vezes precedido de outro combinador. Tal resultado pode estar atrelado à dificuldade dos modelos individuais se distanciarem de um *random walk*.

Em síntese, os combinadores cANN, cMV e cSVR apresentam, nesta ordem, uma redução nos erros de previsão para série DJUF7 (Figura 24(a)) de 92,28%, 74,10% e 22,66% em comparação com modelo JUA. Nas séries DJUF11 e DJUF12 (Figuras 24(b) e 24(c)) essa redução

Tabela 6: Desempenho dos modelos de previsão (ANN, ARIMA, ETS, FUZZY, SVR, JUA, cSA, cSM, cANN, cMV, cSVR) ao prever as séries temporais DJUF7, DJUF11, DJUF12 e DJUF13 (fase de teste). Os melhores valores encontram-se em negrito

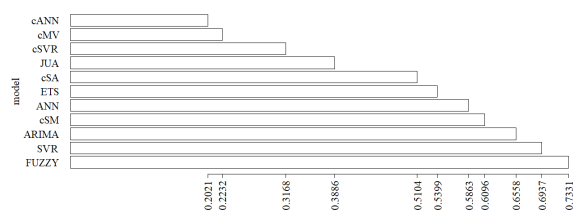
Série	Métrica	FUZZY	ANN	ARIMA	ETS	SVR	JUA	cSA	cSM	cANN	cMV	cSVR
DJUF7	MSE	586846	440657	382428	309987	451827	479266	266137	292126	280890	271672	273359
	MAPE	0.177	0.148	0.148	0.136	0.166	0.164	0.124	0.131	0.113	0.112	0.113
	ARV	1.748	3.118	4.160	5.048	2.366	0.897	5.297	5.226	1.786	1.867	2.569
	ID	0.746	0.822	0.958	0.830	0.726	0.423	0.858	0.862	0.483	0.459	0.564
	Theil	0.998	0.990	0.799	0.704	1.024	1.014	0.546	0.624	0.593	0.617	0.609
	WPOCID	0.538	0.692	0.538	0.385	0.538	0.385	0.538	0.615	0.385	0.308	0.385
	Reg_Intercept	3415	6042	4283	4430	3591	2653	3056	3592	2326	2643	2887
	Reg_Slope	0.035	-0.773	-0.222	-0.258	-0.017	0.248	0.139	-0.018	0.363	0.260	0.188
	Reg_WR2	0.998	0.825	0.974	0.978	1.000	0.838	0.994	1.000	0.928	0.964	0.984
	DJUF11	MSE	116338	57126	1979296	58730	54606	263839	100442	54860	53979	132296
MAPE		0.305	0.212	1.387	0.226	0.194	0.392	0.226	0.192	0.191	0.356	0.189
ARV		2.651	5.632	0.974	5.730	5.581	0.929	1.303	6.032	3.187	1.560	4.780
ID		0.796	0.785	0.781	0.912	0.733	0.508	0.480	0.738	0.571	0.651	0.680
Theil		0.951	0.440	15.233	0.411	0.417	2.164	0.824	0.434	0.422	0.894	0.407
WPOCID		0.786	0.786	0.500	0.857	0.500	0.500	0.571	0.786	0.500	0.500	1.000
Reg_Intercept		1146	1988	1090	1411	1757	857	512	754	554	25	1016
Reg_Slope		-0.129	-0.966	0.209	-0.365	-0.739	0.163	0.639	0.265	0.459	0.758	0.188
Reg_WR2		0.984	0.973	0.928	0.998	0.992	0.843	0.878	0.996	0.936	0.884	1.000
DJUF12		MSE	1428289	1126263	802039	801550	1051008	1306913	570726	753137	1231420	1199473
	MAPE	0.217	0.163	0.147	0.147	0.165	0.195	0.117	0.134	0.178	0.177	0.170
	ARV	1.627	2.052	3.800	3.778	1.755	0.785	3.351	3.321	1.602	1.263	1.816
	ID	0.711	0.706	0.910	0.904	0.629	0.391	0.613	0.769	0.628	0.534	0.671
	Theil	1.022	0.936	0.590	0.591	0.873	1.089	0.456	0.568	1.064	1.089	1.063
	WPOCID	0.857	0.429	0.714	0.714	1.000	0.357	0.286	0.571	0.429	0.286	0.500
	Reg_Intercept	4647	5770	10025	10212	5048	3267	-821	4048	-491759	-275	5691
	Reg_Slope	0.073	-0.157	-1.002	-1.042	0.363	1.199	0.205	115.257	1.266	-0.144	0.995
	Reg_WR2	0.995	0.994	0.950	0.948	1.000	0.585	0.787	0.997	0.598	0.697	0.995
	DJUF13	MSE	2767058	1653754	1756079	1749441	3112169	1356544	1481062	1548334	1173615	1089084
MAPE		0.273	0.204	0.222	0.221	0.241	0.163	0.199	0.201	0.174	0.174	0.343
ARV		2.659	4.106	13.905	14.136	1.722	0.873	5.749	6.761	1.772	2.393	1.224
ID		0.720	0.630	0.908	0.911	0.568	0.292	0.649	0.706	0.371	0.395	0.778
Theil		0.923	0.551	0.574	0.571	1.038	0.398	0.494	0.516	0.381	0.357	3.036
WPOCID		0.714	0.286	0.786	0.786	1.000	0.357	0.500	0.500	0.286	0.286	0.571
Reg_Intercept		5252	2845	7560	7912	11193	2578	1948	2723	585	241	6059
Reg_Slope		0.040	0.519	-0.386	-0.452	-1.328	0.535	0.679	0.527	0.977	1.018	-0.145
Reg_WR2		0.999	0.957	0.985	0.982	0.914	0.611	0.920	0.956	0.639	0.655	0.937

Fonte: Elaborado pelo autor.

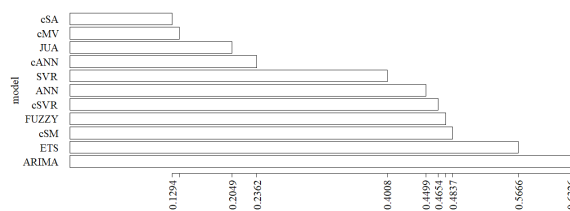
proporcionada pelo cSA no erro médio agregado é de 58,34% e 55,11%, respectivamente. Quando verifica-se o cANN na série DJUF13 (Figura 24(d)), o formalismo proporciona uma redução no erro de mais de 118%. Com isso, entende-se que esses resultados são significativos não só para a literatura, mas também para a indústria de manufatura objeto de estudo. Esses resultados possibilitam o uso desses preditores pela indústria de manufatura, por apresentarem uma acuracidade melhor, quando comparado ao seu modelo. Logo, o sistema de previsão de demanda desenvolvido pode ser adotado pela indústria objeto de estudo com a finalidade de melhorar a predição da demanda e também o planejamento da produção.

Na Figura 24 estão as representações gráficas das séries de demanda e das previsões realizadas pelos combinadores para as famílias de produtos DJUF7, DJUF11, DJUF12 e DJUF13. É possível verificar que os formalismos acompanham a tendência da série, ao mesmo tempo em que se visualiza a aderência dos modelos aos dados da série.

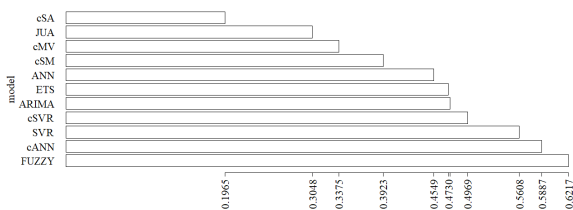
Figura 23: Média agregada normalizada de performance dos modelos individuais e combinados para as séries de demanda de produtos DJUF7, DJUF11, DJUF12 e DJUF13



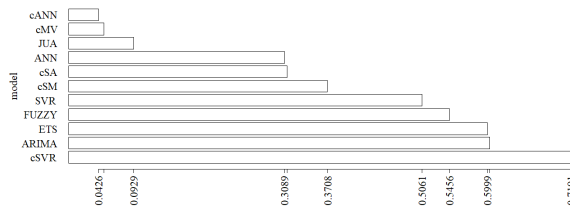
(a) Média agregada normalizada de performance para série DJUF7



(b) Média agregada normalizada de performance para série DJUF11



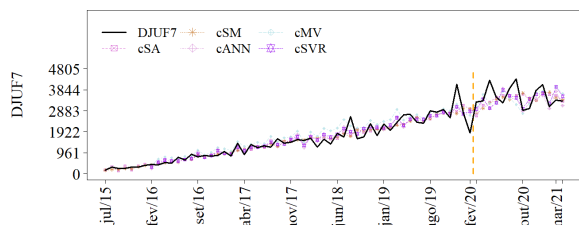
(c) Média agregada normalizada de performance para série DJUF12



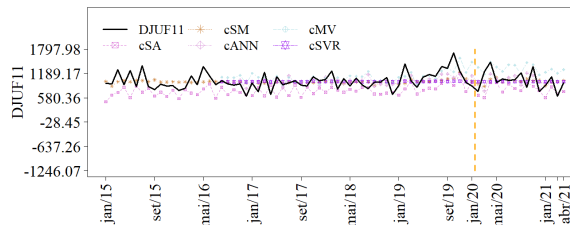
(d) Média agregada normalizada de performance para série DJUF13

Fonte: Elaborado pelo autor.

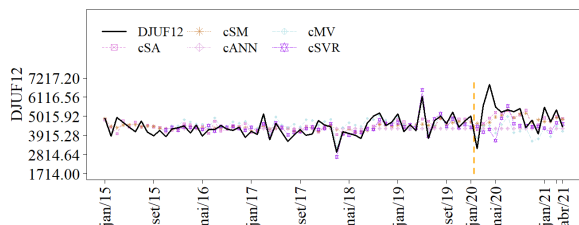
Figura 24: Formalismos combinados para as séries temporais de demanda das famílias de produtos DJUF7, DJUF11, DJUF12 e DJUF13 (Fase de treinamento e teste)



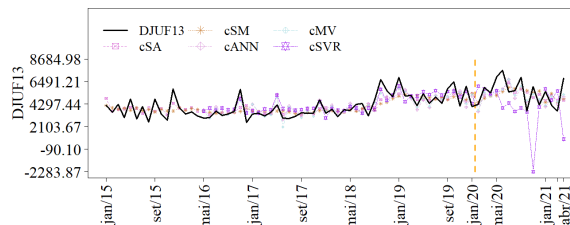
(a) Formalismos combinados para série DJUF7



(b) Formalismos combinados para série DJUF11



(c) Formalismos combinados para série DJUF12



(d) Formalismos combinados para série DJUF13

Fonte: Elaborado pelo autor.

5 DISCUSSÃO DOS RESULTADOS

Esta seção tem como objetivo descrever uma discussão acerca dos resultados obtidos com os formalismos individuais e combinados ora propostos, por meio de uma análise de seus desempenhos e das vantagens e dificuldades na utilização do sistema de previsão de demanda.

Neste estudo foram usadas oito séries temporais de demanda para avaliar o desempenho dos modelos e do sistema de previsão propostos. Foram gerados, ao todo, 80 modelos, sendo 40 individuais com as metodologias ANN, ARIMA, ETS, FUZZY e SVR e 40 combinados, cANN, cMV, cSA, cSM e cSVR. Além disso, tem-se o uso da metodologia de previsão da empresa objeto de estudo, gerando mais 8 modelos preditivos. O método aqui empregado buscou analisar o potencial de cada preditor na previsão da demanda da família de produtos da indústria de manufatura e comparar com o modelo usado por ela. O objetivo é combinar as previsões geradas pelos diferentes modelos de previsão como uma tentativa de capturar as contribuições das características de previsão mais importantes de cada modelo individual a cada etapa de previsão.

As Tabelas 7 e 8 a seguir mostram os desempenhos dos modelos gerados sob a ótica da média agregada das métricas de desempenho normalizadas, para os conjuntos de teste e de treinamento. Sua análise permite inferir que os formalismos combinados do sistema de previsão de demanda se sobressaíram diante dos individuais, vencendo em todas as oito séries de demanda estudadas. Isso quer dizer que a escolha ideal para realizar a predição da demanda analisando série a série seria por meio de um combinador. Esse achado corrobora com a ideia de autores como Bates e Granger (1969); Newbold e Granger (1974); Clemen (1989); Timmermann (2006); Matsumoto et al. (2016); Wang e Petropoulos (2016); Wang, Li e Li (2018); Wang et al. (2018), que sugerem que a combinação apropriada de preditores tende a melhorar sensivelmente a qualidade das previsões.

Entretanto, nota-se que a metodologia JUA no agregado geral (considerando todos os modelos e séries) apresenta desempenho mediano na fase de treinamento, ocupando a nona posição, quando verifica-se os resultados na fase de teste o modelo se apresenta como o melhor preditor, passando a ocupar a primeira posição (ver Tabelas 7 e 8). Com isso, o método JUA ganha no agregado geral, mas nunca em algum caso específico. Isso permite refletir sobre qual resultado ofereceria mais ganhos para indústria, entre adotar um modelo único para predição ou modelos que se ajustam às séries temporais individualmente. Entende-se que considerar a aplicação de um modelo único como é o caso do JUA, na previsão de diversas séries temporais não seja ideal, partindo do pressuposto que as séries possuem características, padrões e propriedades estatísticas diferentes e que o tipo de dado determina qual método de previsão usar.

A identificação mais refinada desses padrões não aleatórios das séries temporais de demanda de uma variável de interesse e a observação desse comportamento passado permitem fazer previsões com maior acuracidade sobre o futuro, orientando melhor o processo de tomada de decisão.

Outro ponto de discussão a ser considerado é sobre a conexão entre a previsão da demanda e planejamento da produção. O setor de manufatura sofre mudanças constantes de escala, complexidade e tecnologia que buscam atender às demandas dos clientes que mudam rapidamente. Ao mesmo tempo, as indústrias são desafiadas a planejar as atividades envolvidas em seus processos de produção na condução para sucesso. Com isso, as previsões de demanda geram dados de entrada para outras áreas de planejamento, incluindo o de produção. Assim, a escolha e uso de formalismos que se ajustem bem aos dados de demanda em detrimento ao uso de formalismo único é preferível e fundamental. Uma forma de aprimoramento do processo de previsão de demanda é escolher métodos adequados às particularidades do produto (CECATTO; BELFIORE, 2015). O planejamento de produção bem-sucedido depende da modelagem precisa de aspectos relacionados ao problema, incluído a incerteza da demanda contida nas séries das famílias de produtos.

Verificando os formalismos individuais, os resultados da análise demonstram que a modelagem fuzzy de STF não apresenta bons resultados. No entanto, autores como Che-Ngoc et al. (2018), Vovan (2019) e Thao (2019), ressaltam que o desempenho do modelo depende fortemente dos parâmetros (n , w e C). Possivelmente, o baixo desempenho do formalismo está associado a escolha desses parâmetros. Neste caso, foi usado o conhecimento heurístico do *Simulated Annealing* com esse objetivo. Por certo, investigar adequadamente os parâmetros na busca por valores ideais para cada conjunto de dados é indicado.

Tabela 7: Desempenho normalizado médio agregado dos modelos de previsão (ANN, ARIMA, ETS, FUZZY, JUA, SVR, cSA, cSM, cMV, cANN, cSVR) ao prever as séries temporais DJUF1, DJUF2, DJUF3, DJUF6, DJUF7, DJUF11, DJUF12 e DJUF13 (fase de teste). A classificação de cada modelo está entre parênteses

Série	Pior	Melhor	FUZZY	ANN	ARIMA	ETS	SVR	JUA	cSA	cSM	cANN	cMV	cSVR
DJUF1	ETS	cSVR	0.325 (6)	0.644 (9)	0.716 (10)	0.736 (11)	0.409 (8)	0.239 (4)	0.194 (3)	0.336 (7)	0.270 (5)	0.192 (2)	0.145 (1)
DJUF2	SVR	cMV	0.296 (6)	0.637 (8)	0.639 (9)	0.643 (10)	0.831 (11)	0.066 (2)	0.239 (5)	0.559 (7)	0.102 (3)	0.033 (1)	0.179 (4)
DJUF3	cMV	cSA	0.410 (5)	0.634 (10)	0.487 (7)	0.511 (9)	0.257 (4)	0.141 (2)	0.141 (1)	0.201 (3)	0.506 (8)	0.686 (11)	0.458 (6)
DJUF6	FUZZY	cSVR	0.758 (11)	0.451 (6)	0.668 (8)	0.682 (9)	0.697 (10)	0.134 (2)	0.396 (5)	0.521 (7)	0.200 (4)	0.142 (3)	0.105 (1)
DJUF7	FUZZY	cANN	0.733 (11)	0.586 (7)	0.656 (9)	0.539 (6)	0.694 (10)	0.389 (4)	0.510 (5)	0.609 (8)	0.202 (1)	0.223 (2)	0.317 (3)
DJUF11	ARIMA	cSA	0.475 (8)	0.449 (6)	0.633 (11)	0.567 (10)	0.401 (5)	0.205 (3)	0.129 (1)	0.484 (9)	0.236 (4)	0.138 (2)	0.465 (7)
DJUF12	FUZZY	cSA	0.621 (11)	0.455 (5)	0.475 (7)	0.473 (6)	0.561 (9)	0.305 (2)	0.196 (1)	0.392 (4)	0.589 (10)	0.338 (3)	0.497 (8)
DJUF13	cSVR	cANN	0.546 (8)	0.309 (4)	0.603 (10)	0.599 (9)	0.506 (7)	0.093 (3)	0.313 (5)	0.371 (6)	0.043 (1)	0.050 (2)	0.719 (11)
GERAL	ARIMA	JUA	0.522 (8)	0.488 (7)	0.610 (11)	0.595 (10)	0.583 (9)	0.203 (1)	0.268 (4)	0.444 (6)	0.240 (3)	0.210 (2)	0.356 (5)

Fonte: Elaborado pelo autor.

O modelo SVR individual não apresenta bom desempenho no geral, sendo necessário um ajuste mais refinado dos hiperparâmetros cost , ϵ e os parâmetros da função kernel para melhorar o desempenho. Reforçando a ideia de autores como Liu, Cai e Tan (2007) e Tsirikoglou

et al. (2017) em que, apesar de suas vantagens, os modelos de SVR exigem uma seleção precisa dos parâmetros de configuração para alcançar um bom desempenho de generalização. Vale salientar que, em alguns casos, o modelo SVR se aproxima dos valores alvos das séries no conjunto de treinamento, entretanto, no conjunto de teste, tem-se um resultado ruim. O modelo não consegue descrever o processo gerador da série.

O ETS, acompanhado do ARIMA, falham em determinar um padrão e sofrem com ajuste insuficiente. Os modelos apresentam aparentemente um caso típico de *underfitting*. Esse comportamento também é observado no modelo FUZZY. Ocorre quando o modelo não se adapta bem sequer aos dados com os quais foi treinado. Com isso, resulta em resultados problemáticos ou errôneos em novos dados, ou dados nos quais não foi treinado, e muitas vezes tem um desempenho ruim, mesmo em dados de treinamento.

Tabela 8: Desempenho normalizado médio agregado dos modelos de previsão (ANN, ARIMA, ETS, FUZZY, JUA, SVR, cSA, cSM, cMV, cANN, cSVR) ao prever as séries temporais DJUF1, DJUF2, DJUF3, DJUF6, DJUF7, DJUF11, DJUF12 e DJUF13 (fase de treinamento). A classificação de cada modelo está entre parênteses

Séries	Pior	Melhor	FUZZY	ANN	ARIMA	ETS	SVR	JUA	cSA	cSM	cANN	cMV	cSVR
DJUF1	JUA	ANN	0.641 (10)	0.072 (1)	0.329 (9)	0.083 (2)	0.265 (8)	0.751 (11)	0.176 (4)	0.224 (6)	0.204 (5)	0.228 (7)	0.163 (3)
DJUF2	ARIMA	cANN	0.576 (9)	0.307 (5)	0.599 (11)	0.592 (10)	0.473 (7)	0.566 (8)	0.251 (4)	0.451 (6)	0.025 (1)	0.159 (3)	0.027 (2)
DJUF3	JUA	cSA	0.479 (8)	0.299 (4)	0.459 (6)	0.467 (7)	0.480 (9)	0.618 (11)	0.087 (1)	0.312 (5)	0.157 (2)	0.489 (10)	0.222 (3)
DJUF6	FUZZY	cMV	0.959 (11)	0.531 (9)	0.431 (7)	0.439 (8)	0.583 (10)	0.326 (6)	0.191 (2)	0.322 (5)	0.205 (3)	0.064 (1)	0.315 (4)
DJUF7	FUZZY	cSVR	0.932 (11)	0.257 (7)	0.415 (10)	0.072 (2)	0.388 (8)	0.407 (9)	0.119 (4)	0.146 (5)	0.114 (3)	0.245 (6)	0.069 (1)
DJUF11	ARIMA	JUA	0.257 (5)	0.306 (7)	0.648 (11)	0.435 (10)	0.298 (6)	0.113 (1)	0.249 (4)	0.323 (8)	0.128 (2)	0.129 (3)	0.389 (9)
DJUF12	cANN	cSVR	0.572 (10)	0.288 (5)	0.439 (8)	0.436 (7)	0.184 (3)	0.524 (9)	0.257 (4)	0.321 (6)	0.669 (11)	0.130 (2)	1e-04 (1)
DJUF13	FUZZY	cMV	0.773 (11)	0.382 (6)	0.477 (8)	0.476 (7)	0.731 (10)	0.512 (9)	0.255 (4)	0.370 (5)	0.104 (3)	0.045 (1)	0.103 (2)
GERAL	FUZZY	cSVR	0.653 (11)	0.317 (5)	0.478 (10)	0.379 (7)	0.413 (8)	0.477 (9)	0.207 (4)	0.319 (6)	0.189 (2)	0.189 (3)	0.142 (1)

Fonte: Elaborado pelo autor.

O modelo ANN, como sistema de inteligência artificial para extração de padrões estatísticos, não obteve boa performance como preditor individual. Uma explicação razoável para isso pode estar associada ao tamanho da série temporal. Um ponto fraco desses sistemas ou modelos de aprendizado de máquinas é que eles precisam de muitos dados para conseguir entender os padrões que se apresentam na complexa realidade dos dados. No caso dessa pesquisa, temos um conjunto de dados para treinamento com 54 e 58 observações.

O formalismo JUA usado pela indústria de manufatura apresenta resultados que precisam ser melhor avaliados e por consequência melhorados. A realização de previsões de demanda por essa metodologia é capaz de influenciar fortemente no sucesso ou fracasso das previsões da empresa, já que direciona decisões estratégicas de longo, médio ou curto prazo. Um motivo apontado é que os tomadores de decisão da empresa optam por usar esse método de previsão de demanda por possuir familiaridade, mesmo gerando em diversos momentos problemas de subprodução ou superprodução.

Esses problemas causam, entre outros, acréscimo dos gastos da empresa com aumento do nível e do uso de espaços da fábrica com estoque (55.674 caixas de produtos e cerca de

1.450m² da área da fábrica), elevação na carga de trabalho ocasionando um aumento na quantidade de horas extras (aproximadamente 600hrs extras/mês) e redução da produtividade devido o aumento no número de realinhamentos/recalibrações das máquinas usadas para a produção de variados tipos de produtos relacionados a demandas não assertivas, demonstrando a importância na seleção dos métodos de previsão de demanda relacionados com as necessidades, particularidades e realidades dos produtos, da empresa e do segmento de mercado, visando minimizar os erros das previsões de demanda quando se trata de planejar a produção.

Diante da existência de diversas técnicas de previsão disponíveis, desde as simples às mais sofisticadas, o grande desafio passou a ser a identificação da melhor estratégia de modelagem, de maneira a promover preditores estatisticamente acurados e eficientes. O que entende-se com os achados é que o desafio foi superado, considerando a identificação de preditores como cSA, cANN, cSVR e cMV que se mostram mais eficientes e acurados que a metodologia usada atualmente pela empresa. O sistema de previsão de demanda proposto e desenvolvido nesta tese garante e assegura uma melhoria significativa na qualidade das previsões e, por consequência, para outras áreas de planejamento como, por exemplo, o planejamento da produção.

A informação fornecida pelo modelo, a previsão de demanda, é a informação que o especialista necessita para adequar o planejamento e programação da produção, auxiliando na redução dos problemas ocasionados pelo modelo JUA. Isso significa que os modelos têm potencial para melhorar o nível de estoques, reduzir a quantidade de horas extras e aumentar a produtividade a partir da diminuição do número de realinhamentos/recalibrações das máquinas da empresa. Outra contribuição a considerar com o melhoramento da qualidade das previsões é a redução do tempo dedicado às reuniões (objeto da segunda etapa da metodologia de previsão da empresa) e/ou diminuição da frequência com que a mesma ocorre.

Outro desafio, foi o gerado pelo advento da pandemia da Covid-19 que impulsionou um crescimento pela comercialização de itens como água sanitária, desinfetante e sabão em barra. A indústria de produtos de limpeza logo viu seus números subirem. De janeiro a julho de 2020, o setor cresceu 5,9%. No entanto, com o passar do tempo e os desafios, os números começaram a cair e o segmento fechou 2020 com estabilidade (HIGIENE, 2021). Ainda assim, vários dos modelos preditivos, que fazem parte do sistema de previsão demanda, superaram essas dificuldades, alcançando e gerando os resultados demonstrados.

É importante considerar, também, a parcimônia dos formalismos. O princípio da parcimônia estabelece que quando há concorrência entre modelos equivalentes, deve-se optar pelo mais simples. No âmbito do aprendizado de máquina, este princípio considera a aplicabilidade do modelo às diferentes bases de dados, ao custo computacional e ao tempo gasto no

processo de modelagem (PAULA, 2016).

Outro ponto de discussão com vistas ao melhoramento da performance dos modelos propostos, é o tempo de treinamento dos formalismos. A Tabela 9 exhibe o tempo gasto para o treinamento de cada formalismo, assim como o tempo médio por série temporal de demanda da família de produtos, por formalismo e geral. Em termos gerais, o tempo médio de modelagem das alternativas adotadas foi de 1.259,66 segundos. Dentre os formalismos de modelagem individual, o que consumiu menor tempo, em média, foi o ETS; próximo de 0,05 segundos. Já dentre os combinadores paramétricos, cANN foi o mais rápido, envolvendo em média 1.852,60 segundos. Por outro lado, em relação às séries temporais, DJUF1 foi a que demandou mais tempo de modelagem; em média 3.447,94 segundos e a série DJUF2 foi a que levou menor tempo para modelagem; uma média de 521,53 segundos. Em média, o modelo que demandou mais tempo na etapa de treinamento foi o cSVR, seguido do cANN com respectivos tempos de 3.675,23 e 1.852,60 segundos.

Tabela 9: Tempo de treinamento dos modelos (ANN, ARIMA, ETS, FUZZY, SVR, cANN e cSVR) para as séries DJUF1, DJUF2, DJUF3, DJUF6, DJUF7, DJUF11, DJUF12 e DJUF13 (em segundos). Os melhores valores encontram-se em negrito

Série	ANN	ARIMA	ETS	FUZZY	SVR	cANN	cSVR	Média
DJUF1	216.81	0.07	0.08	2111.36	6398.89	562.76	14845.59	3447.94
DJUF2	786.25	0.07	0.07	1005.97	645.78	611.30	601.28	521.53
DJUF3	22.06	0.14	0.05	1547.20	70.81	3874.52	15.47	790.04
DJUF6	610.08	0.06	0.03	741.16	643.91	1001.77	11504.43	2071.63
DJUF7	1097.91	0.03	0.05	602.33	638.64	2256.28	611.50	743.82
DJUF11	1536.17	0.14	0.05	2186.62	628.32	3660.36	606.61	1231.18
DJUF12	623.30	0.10	0.05	601.50	633.60	613.57	614.38	440.93
DJUF13	1108.18	0.13	0.05	1220.42	639.64	2240.24	602.61	830.18
Média	750.10	0.09	0.05	1252.07	1287.45	1852.60	3675.23	1259.66

Fonte: Elaborado pelo autor.

Finalizada a aplicação e análise, os resultados produzidos atenderam aos objetivos específicos da presente pesquisa e contribuíram para o atendimento do objetivo geral desta tese.

6 CONSIDERAÇÕES FINAIS

Na presente pesquisa foi abordado o tema de previsão de demanda para indústria de manufatura de bens. O planejamento da demanda, que é baseado nas previsões das demandas futuras, é fundamental (KILGER; WAGNER, 2008). A identificação da melhor estratégia de modelagem, de maneira a promover preditores estatisticamente acurados e eficientes passou a ser o maior desafio. Nesta direção, propôs-se, como objetivo geral, e que foi alcançado, um sistema de previsão de demanda que auxilie no planejamento da produção, baseando-se em diversos formalismos de séries temporais e suas combinações.

Como base teórica deste estudo, foi apresentada uma revisão sistemática da literatura. As sínteses da revisão foram realizadas a partir de uma análise léxica, com o auxílio do *software* IRAMUTEQ, sobre a Previsão de Demanda aplicada a Séries Temporais e sua interação com o Planejamento da Produção. A análise permitiu identificar diferentes modelos e métodos de previsão de demanda, seguindo diferentes teorias e abordagens. Apesar da utilização de diferentes formalismos de previsão, observou-se a pouca exploração de métodos individuais robustos e da combinação de modelos e do uso de diferentes técnicas de incorporação na previsão da demanda a partir de séries temporais.

Neste sentido, foi possível formular o sistema de previsão de demanda (artefato) com uso de 6 formalismos individuais (ANN, ARIMA, ETS, FUZZY, JUA e SVR) e 5 combinadores (cANN, cSA, cSM, cMV e cSVR), utilizando aspectos, conceitos e características dos diferentes modelos conceituais. Entretanto, é fundamental validar o artefato adotado e sempre mantê-lo sob controle, de maneira a poder efetuar correções dentro do menor prazo possível. Portanto, é necessário testar diversos modelos até encontrar o mais adequado ao caso específico que se está analisando.

O artefato, então, foi aplicado a um Caso de Estudo, por meio das séries temporais de demanda das famílias de produtos de uma indústria de manufatura do setor de produtos de higiene, limpeza e saneantes de uso doméstico e profissional. A empresa possui uma ferramenta de previsão que é considerada uma caixa preta para o gestor da produção, causando insegurança na previsão realizada, o que leva os gestores a utilizarem uma metodologia sobre os quais possuem familiaridade, levando por vezes a resultados distantes dos valores reais, além de problemas de subprodução e superprodução.

Os resultados da aplicação do artefato sugerem que os formalismos cSA, cANN, cSVR e cMV apresentam maior acuracidade que o modelo JUA usado pela empresa. Os modelos podem, então, ser utilizados como ferramenta de previsão para indústria de manufatura de produtos de limpeza. O artefato proposto está apto a realizar previsões com acuracidade superior

para todas as famílias de produtos produzidos e comercializados. Assim, tende a atender a necessidade da empresa por um sistema de previsão de demanda eficiente.

Os formalismos ANN, ARIMA, ETS, FUZZY e SVR, também propostos, apesar de apresentarem acuracidade menor que o modelo JUA, também foram importantes para a construção dos formalismos combinados. O critério de seleção do modelo matemático foi o melhor ajuste do formalismo aos dados da série histórica, sendo mensurado por várias métricas de desempenho.

As previsões foram obtidas por meio de horizonte *one-step-ahead* (um passo a frente), permitindo aos gestores otimizar a alocação de recursos, materiais e de custos com foco no planejamento e na programação da produção. Os modelos de previsão de demanda propostos foram construídos a partir de dados comumente disponíveis e, desse modo, podem ser facilmente implementados e utilizados na previsão da demanda da indústria de manufatura de bens. Com o aumento do conjunto de dados das séries, o princípio de parcimônia, ajuste refinado nos parâmetros dos modelos e de informações que possivelmente possam influenciar a demanda da empresa, acredita-se que a utilização destes modelos pode ser estendida às outras indústrias de manufatura.

Outra contribuição é que o estudo demonstra a importância do uso de modelos matemáticos de previsão para indústrias de manufatura de bens. A aplicação dos modelos de previsão de demanda deve ser vista pelos gestores como uma ferramenta para auxiliar as decisões, apoiando processos de planejamento, programação, gestão e avaliação no ambiente manufatureiro, sendo um instrumento que gera dados de entrada para outras áreas de planejamento, que são usadas no planejamento desde processos estratégicos a operacionais. Entender o comportamento da demanda da indústria objeto de estudo em sua totalidade, permitirá que a gestão planeje com antecedência, garanta que os recursos necessários estejam disponíveis, gerencie melhor os orçamentos e venha a reduzir situações de subprodução ou superprodução.

Como contribuição acadêmica, destaca-se a aceitação do artigo "*Demand forecasting, production planning, and control: a systematic literature review*" para publicação na Springer Science + Business Media dos autores José Eduardo de Carvalho Lima, Paulo Renato Alves Firmino e Luiz Alberto Oliveira Rocha. Destaca-se também a construção de mais dois artigos em fase final, dos mesmos autores, com títulos "*Optimising Abbasov-Mamedov fuzzy time series models*" e "Planejamento da Demanda para Indústria de Manufatura de Bens a partir de Formalismos de Séries Temporais".

Entende-se que o trabalho atendeu aos objetivos propostos, na medida em que propôs um sistema de previsão de demanda a partir da combinação de preditores individuais, e que utilizou os dados históricos de demanda e teve desempenho superior ao modelo usado pela

indústria objeto de estudo. A acuracidade do sistema de previsão proposto pode ser verificada, bem como a viabilidade do mesmo pode ser analisada.

Considerando as limitações da presente pesquisa, o sistema de previsão de demanda (artefato) não pode ser considerado como validado. Como foi aplicado, há somente um caso de estudo, não se pode esperar que os mesmos resultados sejam encontrados em casos diferentes. No entanto, os resultados alcançados podem ser úteis à indústria manufatureira de onde os dados foram coletados. Além de que a continuidade dos estudos pode refinar o sistema de previsão, tornando-o mais robusto.

Como sugestões para futuros trabalhos, numa primeira instância, são propostos aprimoramentos para o artefato ora proposto, como por exemplo a parcimônia dos modelos ANN e SVR individuais e combinados. Estudos relacionados à previsão de demanda de um modo geral apresentam oportunidades de maiores pesquisas, sendo oportuna a incorporação de conceitos adicionais, de outras abordagens individuais como modelos caóticos e cópulas, bem como combinar por meio de sistemas nebulosos. Explorar o desempenho do artefato nas demandas dos produtos individuais. Aplicar testes estatísticos a fim de determinar se as previsões são significativamente diferentes.

A aplicação do sistema de previsão proposto à prática, tomando decisões com base nas informações do modelo, também é sugerida. A partir das implicações de como são realizadas as previsões pela empresa, também emergem sugestões para futuros trabalhos baseados no artefato ora proposto. Considerando que a tomada de decisão com base nas informações fornecidas pelo artefato é aconselhada a ser feita por um especialista, sugere-se a integração da previsão com uma previsão de demanda por julgamento, definindo como, quando e onde o especialista deveria atuar. E por fim, estudar o desempenho de modelos considerando a previsão *multi-step-ahead*.

Considerando implicações relacionadas ao tamanho das séries, sugere-se a realização de novos estudos com um conjunto maior de observações, em outras indústrias de manufatura de produtos com características diferentes, envolvendo a previsão de demanda de produtos individuais e da família. O modelo utilizado nesta pesquisa é recomendado a ser aplicado em empresas que possuam características semelhantes as do caso de estudo, visando trazer maior qualidade nas previsões de demanda.

REFERÊNCIAS

- ABBASOV, A.; MAMEDOVA, M. Application of fuzzy time series to population forecasting. **Vienna University of Technology**, v. 12, p. 545–552, 2003.
- ABHAYAWANSA, S. A methodology for investigating intellectual capital information in analyst reports. **Journal of Intellectual capital**, v. 12, n. 3, p. 446–476, 2011.
- ABURTO, L.; WEBER, R. Improved supply chain management based on hybrid demand forecasts. **Applied Soft Computing**, v. 7, n. 1, p. 136–144, 2007.
- ADYA, M.; COLLOPY, F. How effective are neural networks at forecasting and prediction? A review and evaluation. **Journal of forecasting**, v. 17, n. 5-6, p. 481–495, 1998.
- AGUILLO, I. F. Is Google Scholar useful for bibliometrics? A webometric analysis. **Scientometrics**, v. 91, n. 2, p. 343–351, 2012.
- AIT-ALLA, A.; TEUCKE, M.; LÜTJEN, M.; BEHESHTI-KASHI, S.; KARIMI, H. R. Robust production planning in fashion apparel industry under demand uncertainty via conditional value at risk. **Mathematical Problems in Engineering**, v. 2014, 2014.
- ALBEY, E.; NOROUZI, A.; KEMPF, K. G.; UZSOY, R. Demand modeling with forecast evolution: an application to production planning. **IEEE Transactions on Semiconductor Manufacturing**, v. 28, n. 3, p. 374–384, 2015.
- ALIYEVA, K. Demand forecasting for manufacturing under Z-Information. **Procedia computer science**, v. 120, p. 509–514, 2017.
- AMIRKOLAI, K. N.; BABOLI, A.; SHAHZAD, M.; TONADRE, R. Demand Forecasting for Irregular Demands in Business Aircraft Spare Parts Supply Chains by using Artificial Intelligence (AI). **IFAC-PapersOnLine**, v. 50, n. 1, p. 15221–15226, 2017.
- ANDERSON, O. The Interpretation of Box-Jenkins Time Series Models. **Journal of the Royal Statistical Society: Series D (The Statistician)**, v. 26, n. 2, p. 127–145, 1977.
- ANDRIENKO, N.; ANDRIENKO, G. **Exploratory analysis of spatial and temporal data: a systematic approach**. Springer Science & Business Media, 2006.
- ARMSTRONG, J. S. **Principles of forecasting: a handbook for researchers and practitioners**. Springer Science & Business Media, 2001. v. 30.
- ARMSTRONG, J. S. Findings from evidence-based forecasting: methods for reducing forecast error. **International Journal of Forecasting**, v. 22, n. 3, p. 583–598, 2006.
- ARMSTRONG, J. S.; COLLOPY, F. Integration of statistical methods and judgment for time series forecasting: principles from empirical research. , 1998.

- AZADEH, A.; GHADERI, S.; SOHRABKHANI, S. A simulated-based neural network algorithm for forecasting electrical energy consumption in Iran. **Energy Policy**, v. 36, n. 7, p. 2637–2644, 2008.
- BABAI, M. Z.; ALI, M. M.; BOYLAN, J. E.; SYNTETOS, A. A. Forecasting and inventory performance in a two-stage supply chain with ARIMA (0, 1, 1) demand: theory and empirical analysis. **International Journal of Production Economics**, v. 143, n. 2, p. 463–471, 2013.
- BARBOSA, N. d. P.; CHRISTO, E. d. S.; COSTA, K. Demand forecasting for production planning in a food company. **ARN Journal of Engineering and Applied Sciences**, v. 10, n. 16, p. 7137–7141, 2015.
- BARROW, D. K. Forecasting intraday call arrivals using the seasonal moving average method. **Journal of Business Research**, v. 69, n. 12, p. 6088–6096, 2016.
- BATES, J. M.; GRANGER, C. W. The combination of forecasts. **Journal of the Operational Research Society**, v. 20, n. 4, p. 451–468, 1969.
- BERGMEIR, C.; HYNDMAN, R. J.; BENÍTEZ, J. M. Bagging exponential smoothing methods using STL decomposition and Box–Cox transformation. **International journal of forecasting**, v. 32, n. 2, p. 303–312, 2016.
- BLANC, S. M.; SETZER, T. When to choose the simple average in forecast combination. **Journal of Business Research**, v. 69, n. 10, p. 3951–3962, 2016.
- BLANCHARD, G.; BOUSQUET, O.; MASSART, P. et al. Statistical performance of support vector machines. **The Annals of Statistics**, v. 36, n. 2, p. 489–531, 2008.
- BÓNA, K.; LÉNÁRT, B. Supporting demand planning process with Walsh-Fourier based techniques. **Periodica Polytechnica Transportation Engineering**, v. 42, n. 2, p. 97–102, 2014.
- BOX, G. Box and Jenkins: time series analysis, forecasting and control. In: **A Very British Affair**. Springer, 2013. p. 161–215.
- BOX, G. E.; JENKINS, G. M. Time series analysis: forecasting and control san francisco. **Calif: Holden-Day**, 1976.
- BOX, G. E.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. **Time series analysis: forecasting and control**. John Wiley & Sons, 2015.
- BOX, G. E.; TIAO, G. C. **Bayesian inference in statistical analysis**. John Wiley & Sons, 2011. v. 40.
- BRADLEY, B. A.; JACOB, R. W.; HERMANCÉ, J. F.; MUSTARD, J. F. A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. **Remote Sensing of Environment**, v. 106, n. 2, p. 137–145, 2007.

- BROWN, R. G. **Statistical forecasting for inventory control**. McGraw/Hill, 1959.
- BUNN, D. W. Statistical efficiency in the linear combination of forecasts. **International Journal of Forecasting**, v. 1, n. 2, p. 151–163, 1985.
- CAMARGO, B. V.; JUSTO, A. M. IRAMUTEQ: um software gratuito para análise de dados textuais. **Temas em psicologia**, v. 21, n. 2, p. 513–518, 2013.
- CAO, L.-J.; TAY, F. E. H. Support vector machine with adaptive parameters in financial time series forecasting. **IEEE Transactions on neural networks**, v. 14, n. 6, p. 1506–1518, 2003.
- CAVALCANTE, R. C.; MINKU, L. L.; OLIVEIRA, A. L. Fedd: feature extraction for explicit concept drift detection in time series. In: INTERNATIONAL JOINT CONFERENCE ON NEURAL NETWORKS (IJCNN), 2016., 2016. **Anais...** 2016. p. 740–747.
- CECATTO, C.; BELFIORE, P. Demand forecasting methods in the Brazilian food industries. **Gestão & Produção**, v. 22, n. 2, p. 404–418, 2015.
- ČERNÝ, V. Thermodynamical approach to the traveling salesman problem: an efficient simulation algorithm. **Journal of optimization theory and applications**, v. 45, n. 1, p. 41–51, 1985.
- CHAN, F. T.; SAMVEDI, A.; CHUNG, S. H. Fuzzy time series forecasting for supply chain disruptions. **Industrial Management & Data Systems**, 2015.
- CHANG, C.-J.; YU, L.; JIN, P. A mega-trend-diffusion grey forecasting model for short-term manufacturing demand. **Journal of the Operational Research Society**, v. 67, n. 12, p. 1439–1445, 2016.
- CHE-NGOC, H.; VO-VAN, T.; HUYNH-LE, Q.-C.; HO, V.; NGUYEN-TRANG, T.; CHU-THI, M.-T. An Improved Fuzzy Time Series Forecasting Model. In: INTERNATIONAL ECONOMETRIC CONFERENCE OF VIETNAM, 2018. **Anais...** 2018. p. 474–490.
- CHEN, N. **Support vector machine in chemistry**. World Scientific, 2004.
- CHEN, S.-M. Forecasting enrollments based on high-order fuzzy time series. **Cybernetics and Systems**, v. 33, n. 1, p. 1–16, 2002.
- CHEN, S.-M. et al. Forecasting enrollments based on fuzzy time series. **Fuzzy sets and systems**, v. 81, n. 3, p. 311–319, 1996.
- CHEN, S.-M.; HWANG, J.-R. Temperature prediction using fuzzy time series. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, v. 30, n. 2, p. 263–275, 2000.

- CHEN, S.-M.; KAO, P.-Y. TAIEX forecasting based on fuzzy time series, particle swarm optimization techniques and support vector machines. **Information Sciences**, v. 247, p. 62–71, 2013.
- CHENG, C.-H.; CHEN, T.-L.; TEOH, H. J.; CHIANG, C.-H. Fuzzy time-series based on adaptive expectation model for TAIEX forecasting. **Expert systems with applications**, v. 34, n. 2, p. 1126–1132, 2008.
- CHU, H.-H.; CHEN, T.-L.; CHENG, C.-H.; HUANG, C.-C. Fuzzy dual-factor time-series for stock index forecasting. **Expert systems with applications**, v. 36, n. 1, p. 165–171, 2009.
- CIANCIO, C.; AMBROGIO, G.; GAGLIARDI, F.; MUSMANNO, R. Heuristic techniques to optimize neural network architecture in manufacturing applications. **Neural Computing and Applications**, v. 27, n. 7, p. 2001–2015, 2016.
- CLAISSE, M.; CHU, C.; JEMAI, Z. Production planning under uncertainties and forecasts updates. In: INTERNATIONAL CONFERENCE ON CONTROL, DECISION AND INFORMATION TECHNOLOGIES (CODIT), 2016., 2016. **Anais...** 2016. p. 490–495.
- CLEMEN, R. T. Combining forecasts: a review and annotated bibliography. **International journal of forecasting**, v. 5, n. 4, p. 559–583, 1989.
- CLEMENTS, M. P.; HENDRY, D. F. On the limitations of comparing mean square forecast errors. **Journal of Forecasting**, v. 12, n. 8, p. 617–637, 1993.
- CORBERÁN-VALLET, A.; BERMÚDEZ, J. D.; VERCHER, E. Forecasting correlated time series with exponential smoothing models. **International Journal of Forecasting**, v. 27, n. 2, p. 252–265, 2011.
- CORTES, C.; VAPNIK, V. Support-vector networks. **Machine learning**, v. 20, n. 3, p. 273–297, 1995.
- COSTANTINI, M.; GUNTER, U.; M. KUNST, R. Forecast Combinations in a DSGE-VAR Lab. **Journal of Forecasting**, v. 36, n. 3, p. 305–324, 2017.
- COSTANTINI, M.; PAPPALARDO, C. A hierarchical procedure for the combination of forecasts. **International journal of forecasting**, v. 26, n. 4, p. 725–743, 2010.
- CRISTIANINI, N.; SHAWE-TAYLOR, J. et al. **An introduction to support vector machines and other kernel-based learning methods**. Cambridge university press, 2000.
- CRYER, J. D.; CHAN, K.-S. **Time series analysis: with applications in r**. Springer Science & Business Media, 2008.
- CYBENKO, G. Continuous valued neural networks with two hidden layer are sufficient. Medford: tufts university. , 1988.
- CYBENKO, G. Approximation by superpositions of a sigmoidal function. **Mathematics of control, signals and systems**, v. 2, n. 4, p. 303–314, 1989.

DANESE, P.; KALCHSCHMIDT, M. The role of the forecasting process in improving forecast accuracy and operational performance. **International Journal of Production Economics**, v. 131, n. 1, p. 204–214, 2011.

DANESHDOOST, M.; LOTFALIAN, M.; BUMROONGGIT, G.; NGOY, J. Neural network with fuzzy set-based classification for short-term load forecasting. **IEEE transactions on power systems**, v. 13, n. 4, p. 1386–1391, 1998.

DANIEL, G. **Principles of Artificial Neural Networks (Advanced Series in Circuits and Systems)**. World Scientific Pub Co Inc, 2007.

DANIEL, G. **Principles of artificial neural networks**. World Scientific, 2013. v. 7.

DE LIVERA, A. M. **Modeling time series with complex seasonal patterns using exponential smoothing**. 2010. Tese Monash University, 2010.

DE LIVERA, A. M.; HYNDMAN, R. J.; SNYDER, R. D. Forecasting time series with complex seasonal patterns using exponential smoothing. **Journal of the American Statistical Association**, v. 106, n. 496, p. 1513–1527, 2011.

DE MENEZES, L. M.; BUNN, D. W.; TAYLOR, J. W. Review of guidelines for the use of combined forecasts. **European Journal of Operational Research**, v. 120, n. 1, p. 190–204, 2000.

DE OLIVEIRA, R. T.; OLIVEIRA, T. F.; FIRMINO, P. R. A.; FERREIRA, T. A. Combining time series forecasting models via gumbel-hougaard copulas. In: BRICS CONGRESS ON COMPUTATIONAL INTELLIGENCE & 11TH BRAZILIAN CONGRESS ON COMPUTATIONAL INTELLIGENCE (BRICS-CCI & CBIC), 2013., 2013. **Anais...** 2013. p. 568–573.

DELAHAYE, D.; CHAIMATANAN, S.; MONGEAU, M. Simulated annealing: from basics to applications. In: **Handbook of Metaheuristics**. Springer, 2019. p. 1–35.

DOMAŃSKA, D.; WOJTYLAK, M. Application of fuzzy time series models for forecasting pollution concentrations. **Expert Systems with Applications**, v. 39, n. 9, p. 7673–7679, 2012.

DOMBI, J.; JÓNÁS, T.; TÓTH, Z. E. Modeling and long-term forecasting demand in spare parts logistics businesses. **International Journal of Production Economics**, v. 201, p. 1–17, 2018.

DRÉO, J.; PÉROWSKI, A.; SIARRY, P.; TAILLARD, E. **Metaheuristics for hard optimization: methods and case studies**. Springer Science & Business Media, 2006.

DRESCH, A.; LACERDA, D. P.; JÚNIOR, J. A. V. A. **Design science research: método de pesquisa para avanço da ciência e tecnologia**. Bookman Editora, 2015.

DRUCKER, H.; BURGESS, C. J.; KAUFMAN, L.; SMOLA, A. J.; VAPNIK, V. Support vector regression machines. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS, 1997. **Anais...** 1997. p. 155–161.

DUFFIE, N.; CHEHADE, A.; ATHAVALE, A. Control theoretical modeling of transient behavior of production planning and control: a review. **Procedia Cirp**, v. 17, p. 20–25, 2014.

DWEIRI, F.; KHAN, S. A.; JAIN, V. Production planning forecasting method selection in a supply chain: a case study. **International Journal of Applied Management Science**, v. 7, n. 1, p. 38–58, 2015.

EFTHYMIU, K.; MOURTZIS, D.; PAGOROPOULOS, A.; PAPAKOSTAS, N.; CHRYSOLOURIS, G. Manufacturing systems complexity analysis methods review. **International Journal of Computer Integrated Manufacturing**, v. 29, n. 9, p. 1025–1044, 2016.

FACELI, K.; LORENA, A. C.; GAMA, J.; CARVALHO, A. C. P. d. L. et al. *Inteligência Artificial: uma abordagem de aprendizado de máquina.* , 2011.

FATTAHEIAN-DEHKORDI, S.; FERREIDUNIAN, A.; GHOLAMI-DEHKORDI, H.; LESANI, H. Hour-ahead demand forecasting in smart grid using support vector regression (SVR). **International transactions on electrical energy systems**, v. 24, n. 12, p. 1650–1663, 2014.

FENG, L.; MEARS, L.; SCHULTE, J. Energy Demand Forecasting in an Automotive Manufacturing Plant. , 2016.

FIGUEIRA, G.; AMORIM, P.; GUIMARÃES, L.; AMORIM-LOPES, M.; NEVES-MOREIRA, F.; ALMADA-LOBO, B. A decision support system for the operational production planning and scheduling of an integrated pulp and paper mill. **Computers & Chemical Engineering**, v. 77, p. 85–104, 2015.

FILDES, R.; NIKOLOPOULOS, K.; CRONE, S. F.; SYNTETOS, A. A. Forecasting and operational research: a review. **Journal of the Operational Research Society**, v. 59, n. 9, p. 1150–1172, 2008.

FIORUCCI, J. A. et al. Time series forecasting: advances on theta method. , 2016.

FIRMINO, P. R. A.; MATTOS NETO, P. S. de; FERREIRA, T. A. Correcting and combining time series forecasters. **Neural networks**, v. 50, p. 1–11, 2014.

FIRMINO, P. R. A.; MATTOS NETO, P. S. de; FERREIRA, T. A. Error modeling approach to improve time series forecasters. **Neurocomputing**, v. 153, p. 242–254, 2015.

FIRMINO, P. R. A.; SALES, J. P. de; JÚNIOR, J. G.; SILVA, T. A. da. A non-central beta model to forecast and evaluate pandemics time series. **Chaos, Solitons & Fractals**, v. 140, p. 110211, 2020.

- FLORES B. E. & WHITE, E. M. Combining forecasts: why, when and how. **Journal of Business Forecasting Methods & Systems**, v. 8, p. 2–5, 1989.
- FLORES, B. E.; WHITE, E. M. A Framework for the Combination of Forecasts. **Journal of the Academy of Marketing Science**, v. 16, n. 3-4, p. 95–103, 1988.
- FORMIGONI CARVALHO WALTER, O. M.; HENNING, E.; MORO, G.; SAMOHYL, R. W. Aplicação de um modelo SARIMA na previsão de vendas de motocicletas. **Exacta**, v. 11, n. 1, p. 77–88, 2013.
- FRANZIN, A.; STÜTZLE, T. Revisiting simulated annealing: a component-based analysis. **Computers & operations research**, v. 104, p. 191–206, 2019.
- FRITSCH, S.; GUENTHER, F.; GUENTHER, M. F. Package ‘neuralnet’. **Training of Neural Networks. Recuperado de <https://cran.r-project.org/web/packages/neuralnet/neuralnet.pdf>**, 2019.
- GANSTERER, M. Aggregate planning and forecasting in make-to-order production systems. **International Journal of Production Economics**, v. 170, p. 521–528, 2015.
- GARDNER JR, E. S. Exponential smoothing: the state of the art. **Journal of forecasting**, v. 4, n. 1, p. 1–28, 1985.
- GARDNER JR, E. S. Exponential smoothing: the state of the art—part ii. **International journal of forecasting**, v. 22, n. 4, p. 637–666, 2006.
- GEORGIADIS, P.; POLITOU, A. Dynamic Drum-Buffer-Rope approach for production planning and control in capacitated flow-shop manufacturing systems. **Computers & Industrial Engineering**, v. 65, n. 4, p. 689–703, 2013.
- GHIASSI, M.; ZIMBRA, D. K.; SAIDANE, H. Urban water demand forecasting with a dynamic artificial neural network model. **Journal of Water Resources Planning and Management**, v. 134, n. 2, p. 138–146, 2008.
- GIACOMINI, E. **Neural networks in quantitative finance**. 2003. Dissertação Humboldt-Universität zu Berlin, Wirtschaftswissenschaftliche Fakultät, 2003.
- GIL, A. C. Como elaborar projetos de pesquisa. **São Paulo**, v. 6, n. 61, p. 16–17, 2017.
- GOODWIN, P.; GÖNÜL, M. S.; ÖNKAL, D. When providing optimistic and pessimistic scenarios can be detrimental to judgmental demand forecasts and production decisions. **European Journal of Operational Research**, v. 273, n. 3, p. 992–1004, 2019.
- GRANGER, C. W. J.; NEWBOLD, P. **Forecasting Economic Time Series (2nd)**. **San Diego: Academic**, 1986.
- GRANGER, C. W.; RAMANATHAN, R. Improved methods of combining forecasts. **Journal of forecasting**, v. 3, n. 2, p. 197–204, 1984.

- GREGOR, S.; HEVNER, A. R. Positioning and presenting design science research for maximum impact. **MIS quarterly**, p. 337–355, 2013.
- GRIMAUD, F.; DOLGUI, A.; KORYTKOWSKI, P. Exponential smoothing for multi-product lot-sizing with heijunka and varying demand. **Management and Production Engineering Review**, v. 5, n. 2, p. 20–26, 2014.
- GUANGHUI, W. Demand forecasting of supply chain based on support vector regression method. **Procedia Engineering**, v. 29, p. 280–284, 2012.
- GUJARATI, D. N.; PORTER, D. C. **Econometria Básica-5**. Amgh Editora, 2011.
- GUO, G.; SHEN, A. Combined forecasting technology in the material demand forecast based on neural networks. **Journal of the Balkan Tribological Association**, v. 22, n. 4, 2016.
- GUPTA, S.; WILTON, P. C. Combination of forecasts: an extension. **Management Science**, v. 33, n. 3, p. 356–372, 1987.
- GYULAI, D.; KÁDÁR, B.; MONOSOTORI, L. Robust production planning and capacity control for flexible assembly lines. **IFAC-PapersOnLine**, v. 48, n. 3, p. 2312–2317, 2015.
- H., S. J.; W., W. M. Combination forecasts of output growth in a seven-country data set. **Journal of Forecasting**, v. 23, n. 6, p. 405–430, 2004.
- HA, C.; SEOK, H.; OK, C. Evaluation of forecasting methods in aggregate production planning: a cumulative absolute forecast error (cafe). **Computers & Industrial Engineering**, v. 118, p. 329–339, 2018.
- HAJIRAHIMI, Z.; KHASHEI, M. Weighted sequential hybrid approaches for time series forecasting. **Physica A: Statistical Mechanics and its Applications**, p. 121717, 2019.
- HAMICHE, K.; ABOUAÏSSA, H.; GONCALVES, G.; HSU, T. A robust and easy approach for demand forecasting in supply chains. **IFAC-PapersOnLine**, v. 51, n. 11, p. 1732–1737, 2018.
- HANN, T. H.; STEURER, E. Much ado about nothing? Exchange rate forecasting: neural networks vs. linear models using monthly and weekly data. **Neurocomputing**, v. 10, n. 4, p. 323–339, 1996.
- HASNI, M.; AGUIR, M.; BABAI, M.; JEMAI, Z. Spare parts demand forecasting: a review on bootstrapping methods. **International Journal of Production Research**, p. 1–14, 2018.
- HAYKIN, S. **Neural networks: a comprehensive foundation**. Prentice Hall PTR, 1994.
- HAYKIN, S. **Redes neurais: princípios e prática**. Bookman Editora, 2007.
- HAYKIN, S. **Neural Networks and Learning Machines, 3/E**. Pearson Education India, 2010.

HIBINO, H.; TANAKA, K.; UMEZAWA, Y.; FUKUDA, Y. Method for production forecast in demand-synchronized production. **Journal of Advanced Mechanical Design, Systems, and Manufacturing**, v. 8, n. 6, p. JAMDSM0076–JAMDSM0076, 2014.

HIGIENE, L. e. S. d. U. D. e. d. U. P. ABIPLA. Associação Brasileira das Indústrias de Produtos de. Anuário da ABIPLA, 2020. , p. Acesso em 12 junho 2021, 2021.

HOLLMANN, R. L.; SCAVARDA, L. F.; THOMÉ, A. M. T. Collaborative planning, forecasting and replenishment: a literature review. **International Journal of Productivity and Performance Management**, v. 64, n. 7, p. 971–993, 2015.

HOLT, C. C. Forecasting seasonals and trends by exponentially weighted moving averages. **International journal of forecasting**, v. 20, n. 1, p. 5–10, 2004.

HOLT, E. B. G. **A documentary history of art**. 1957.

HORNIK, K.; MEYER, D.; KARATZOGLOU, A. Support vector machines in R. **Journal of statistical software**, v. 15, n. 9, p. 1–28, 2006.

HOU, Y.; ZHAO, L.; LU, H. Fuzzy neural network optimization and network traffic forecasting based on improved differential evolution. **Future Generation Computer Systems**, v. 81, p. 425–432, 2018.

HUARNG, K. Effective lengths of intervals to improve forecasting in fuzzy time series. **Fuzzy sets and systems**, v. 123, n. 3, p. 387–394, 2001.

HWANG, J.-R.; CHEN, S.-M.; LEE, C.-H. Handling forecasting problems using fuzzy time series. **Fuzzy sets and systems**, v. 100, n. 1-3, p. 217–228, 1998.

HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. OTexts, 2018.

HYNDMAN, R. J.; KHANDAKAR, Y. et al. **Automatic time series for forecasting: the forecast package for r**. Monash University, Department of Econometrics and Business Statistics . . . , 2007. n. 6/07.

HYNDMAN, R. J.; KOEHLER, A. B.; SNYDER, R. D.; GROSE, S. A state space framework for automatic forecasting using exponential smoothing methods. **International Journal of forecasting**, v. 18, n. 3, p. 439–454, 2002.

HYNDMAN, R. J.; MAKRIDAKIS, S.; WHEELWRIGHT, S. **Forecasting—Methods and Applications**. , 1998.

HYNDMAN, R.; KOEHLER, A. B.; ORD, J. K.; SNYDER, R. D. **Forecasting with exponential smoothing: the state space approach**. Springer Science & Business Media, 2008.

- JALAL, M. E.; HOSSEINI, M.; KARLSSON, S. Forecasting incoming call volumes in call centers with recurrent neural networks. **Journal of Business Research**, v. 69, n. 11, p. 4811–4814, 2016.
- JATTA, J. S.; KRISHNAN, K. K. An empirical assessment of a univariate time series for demand planning in a demand-driven supply chain. **International Journal of Business Forecasting and Marketing Intelligence**, v. 2, n. 3, p. 269–290, 2016.
- JEON, S. M.; KIM, G. A survey of simulation modeling techniques in production planning and control (PPC). **Production Planning & Control**, v. 27, n. 5, p. 360–377, 2016.
- JIANG, C.; ZHANG, J.; SONG, F. Selecting single model in combination forecasting based on cointegration test and encompassing test. **The Scientific World Journal**, v. 2014, 2014.
- JIANG, P.; YANG, H.; HENG, J. A hybrid forecasting system based on fuzzy time series and multi-objective optimization for wind speed forecasting. **Applied energy**, v. 235, p. 786–801, 2019.
- JILANI, T. A.; BURNEY, S. A.; ARDIL, C. Multivariate high order fuzzy time series forecasting for car road accidents. **International Journal of Computational Intelligence**, v. 4, n. 1, p. 15–20, 2007.
- JUNIOR, M. L.; GODINHO FILHO, M. Production planning and control for remanufacturing: exploring characteristics and difficulties with case studies. **Production Planning and Control**, v. 27, n. 3, p. 212–225, 2016.
- KAMI, M. T. M.; LAROCCA, L. M.; CHAVES, M. M. N.; LOWEN, I. M. V.; SOUZA, V. M. P. d.; GOTO, D. Y. N. Trabalho no consultório na rua: uso do software iramuteq no apoio à pesquisa qualitativa. **Escola Anna Nery**, v. 20, n. 3, 2016.
- KANG, F.; LI, J. Artificial bee colony algorithm optimized support vector regression for system reliability analysis of slopes. **Journal of Computing in Civil Engineering**, v. 30, n. 3, p. 04015040, 2016.
- KASS, R. E.; RAFTERY, A. E. Bayes factors. **Journal of the american statistical association**, v. 90, n. 430, p. 773–795, 1995.
- KAZEMI, A.; GANJAVI, H. S.; MENHAJ, M.; MEHREGAN, M.; TAGHIZADEH, M.; ASL, A. F. A multi-level artificial neural network for gasoline demand forecasting of Iran. In: SECOND INTERNATIONAL CONFERENCE ON COMPUTER AND ELECTRICAL ENGINEERING, 2009., 2009. **Anais...** 2009. v. 1, p. 61–64.
- KEOGH, E.; CHU, S.; HART, D.; PAZZANI, M. Segmenting time series: a survey and novel approach. In: **Data mining in time series databases**. World Scientific, 2004. p. 1–21.
- KHAN, M. Z.; KHAN, M. F. Application of ANFIS, ANN and fuzzy time series models to CO2 emission from the energy sector and global temperature increase. **International Journal of Climate Change Strategies and Management**, 2019.

- KHASHEI, M.; BIJARI, M. A new class of hybrid models for time series forecasting. **Expert Systems with Applications**, v. 39, n. 4, p. 4344–4357, 2012.
- KHOSRAVI, A.; PABON, J.; KOURY, R.; MACHADO, L. Using machine learning algorithms to predict the pressure drop during evaporation of R407C. **Applied Thermal Engineering**, v. 133, p. 361–370, 2018.
- KILGER, C.; WAGNER, M. Demand planning. In: **Supply chain management and advanced planning**. Springer, 2008. p. 133–160.
- KIM, S.; KIM, H. A new metric of absolute percentage error for intermittent demand forecasts. **International Journal of Forecasting**, v. 32, n. 3, p. 669–679, 2016.
- KIM, T. Y.; DEKKER, R.; HEIJ, C. Spare part demand forecasting for consumer goods using installed base information. **Computers & Industrial Engineering**, v. 103, p. 201–215, 2017.
- KIM, Y.-O.; JEONG, D.; KO, I. H. Combining rainfall-runoff model outputs for improving ensemble streamflow prediction. **Journal of Hydrologic Engineering**, v. 11, n. 6, p. 578–588, 2006.
- KIRKPATRICK, S.; GELATT, C. D.; VECCHI, M. P. Optimization by simulated annealing. **science**, v. 220, n. 4598, p. 671–680, 1983.
- KITAGAWA, G. **Introduction to time series modeling**. Chapman and Hall/CRC, 2010.
- KITTI-UDOMPORN, S.; SUKCHAREONPONG, P.; WITCHAKUL, S. Forecasting and purchasing planning for shelf life-limited spare parts. In: INTERNATIONAL CONFERENCE ON BUSINESS AND INDUSTRIAL RESEARCH (ICBIR), 2018., 2018. **Anais...** 2018. p. 264–269.
- KLASSEN, R. D.; FLORES, B. E. Forecasting practices of Canadian firms: survey results and comparisons. **International journal of production economics**, v. 70, n. 2, p. 163–174, 2001.
- KLIR, G. J.; YUAN, B. Fuzzy sets and fuzzy logic: theory and applications. **Possibility Theory versus Probab. Theory**, v. 32, n. 2, p. 207–208, 1996.
- KONISHI, S.; KITAGAWA, G. **Information criteria and statistical modeling**. Springer Science & Business Media, 2008.
- KOURENTZES, N. Intermittent demand forecasts with neural networks. **International Journal of Production Economics**, v. 143, n. 1, p. 198–206, 2013.
- KOURENTZES, N.; ROSTAMI-TABAR, B.; BARROW, D. K. Demand forecasting by temporal aggregation: using optimal or multiple aggregation levels? **Journal of Business Research**, v. 78, p. 1–9, 2017.
- KOZIK, P.; SEP, J. Aircraft engine overhaul demand forecasting using ANN. **Management and Production Engineering Review**, v. 3, n. 2, p. 21–26, 2012.

- KOZŁOWSKI, E.; KOWALSKA, B.; KOWALSKI, D.; MAZURKIEWICZ, D. Water demand forecasting by trend and harmonic analysis. **Archives of Civil and Mechanical Engineering**, v. 18, n. 1, p. 140–148, 2018.
- LAAN, E. van der; DALEN, J. van; ROHRMOSER, M.; SIMPSON, R. Demand forecasting and order planning for humanitarian logistics: an empirical assessment. **Journal of Operations Management**, v. 45, p. 114–122, 2016.
- LAWRENCE, M.; GOODWIN, P.; O’CONNOR, M.; ÖNKAL, D. Judgmental forecasting: a review of progress over the last 25 years. **International Journal of forecasting**, v. 22, n. 3, p. 493–518, 2006.
- LEE, C.-Y.; LIANG, C.-L. Manufacturer’s printing forecast, reprinting decision, and contract design in the educational publishing industry. **Computers & Industrial Engineering**, v. 125, p. 678–687, 2018.
- LEWIS, C. **Demand forecasting and inventory control**. Routledge, 2012.
- LI, B.; LI, J.; LI, W.; SHIRODKAR, S. A. Demand forecasting for production planning decision-making based on the new optimised fuzzy short time-series clustering. **Production Planning & Control**, v. 23, n. 9, p. 663–673, 2012.
- LIMA JUNIOR, A. R.; SILVA, D. A.; MATTOS NETO, P. S.; FERREIRA, T. A. An experimental study of fitness function and time series forecasting using artificial neural networks. In: GENETIC AND EVOLUTIONARY COMPUTATION, 12., 2010. **Proceedings...** 2010. p. 2015–2018.
- LINGITZ, L.; MORAWETZ, C.; GIGLOO, D. T.; MINNER, S.; SIHN, W. Modelling of flexibility costs in a decision support system for mid-term capacity planning. **Procedia CIRP**, v. 7, p. 539–544, 2013.
- LIU, H.; TIAN, H.-q.; LI, Y.-f. Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. **Applied Energy**, v. 98, p. 415–424, 2012.
- LIU, J.-x.; CAI, H.-p.; TAN, Y.-j. Heuristic algorithm for tuning hyper parameters in support vector regression. **Journal of System Simulation**, v. 7, 2007.
- LIU, L.-M.; HUDAK, G. B.; BOX, G. E.; MULLER, M. E.; TIAO, G. C. **Forecasting and time series analysis using the SCA statistical system**. Scientific Computing Associates DeKalb, IL, 1992. v. 1, n. 2.
- LIU, Y.; JU, W.; WANG, K.; GAO, J. Demand forecasting for footwear products using wavelet transform and Artificial Bee Colony algorithm optimized Polynomial Fitting. In: NATURAL COMPUTATION (ICNC), 2015 11TH INTERNATIONAL CONFERENCE ON, 2015. **Anais...** 2015. p. 1146–1150.

- LU, C.-J.; WANG, Y.-W. Combining independent component analysis and growing hierarchical self-organizing maps with support vector regression in product demand forecasting. **International Journal of Production Economics**, v. 128, n. 2, p. 603–613, 2010.
- MA, S.; FILDES, R.; HUANG, T. Demand forecasting with high dimensional data: the case of sku retail sales forecasting with intra-and inter-category promotional information. **European Journal of Operational Research**, v. 249, n. 1, p. 245–257, 2016.
- MA, Z.; YE, C.; MA, W. Support vector regression for predicting building energy consumption in southern China. **Energy Procedia**, v. 158, p. 3433–3438, 2019.
- MADIĆ, M. J.; RADOVANOVIĆ, M. R. Optimal selection of ANN training and architectural parameters using Taguchi method: a case study. **FME Transactions**, v. 39, n. 2, p. 79–86, 2011.
- MAKRIDAKIS, S.; ANDERSEN, A.; CARBONE, R.; FILDES, R.; HIBON, M.; LEWANDOWSKI, R.; NEWTON, J.; PARZEN, E.; WINKLER, R. The accuracy of extrapolation (time series) methods: results of a forecasting competition. **Journal of forecasting**, v. 1, n. 2, p. 111–153, 1982.
- MAKRIDAKIS, S.; HIBON, M. The M3-Competition: results, conclusions and implications. **International journal of forecasting**, v. 16, n. 4, p. 451–476, 2000.
- MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; HYNDMAN, R. J. Forecasting methods and applications, John Wiley & Sons. **Inc, New York**, 1998.
- MARCH, S. T.; SMITH, G. F. Design and natural science research on information technology. **Decision support systems**, v. 15, n. 4, p. 251–266, 1995.
- MARCHAND, P.; RATINAUD, P. L'analyse de similitude appliquée aux corpus textuels: les primaires socialistes pour l'élection présidentielle française (septembre-octobre 2011). **Actes des 11eme Journées internationales d'Analyse statistique des Données Textuelles. JADT**, v. 2012, p. 687–699, 2012.
- MARINOIU, C. Forecast of the evolution of employment in Romanian Agriculture Using FuzzyTime Series. **Economic Insights-Trends and Challenges**, v. 5, n. 3, p. 2016, 2016.
- MARKOVIĆ, G.; GAŠIĆ, M.; KOLAREVIĆ, M.; SAVKOVIĆ, M.; MARINKOVIĆ, Z. Application of the MODIPROM method to the final solution of logistics centre location. **Transport**, v. 28, n. 4, p. 341–351, 2013.
- MARTINS, V. L. M.; WERNER, L. Comparação de previsões individuais e suas combinações: um estudo com séries industriais. **Produção, Porto Alegre**, v. 24, n. 3, p. 618–627, 2014.
- MATSUMOTO, M.; IKEDA, A. Examination of demand forecasting by time series analysis for auto parts remanufacturing. **Journal of Remanufacturing**, v. 5, n. 1, p. 1, 2015.

- MATSUMOTO, M.; KOMATSU, S. Demand forecasting for production planning in remanufacturing. **The International Journal of Advanced Manufacturing Technology**, v. 79, n. 1-4, p. 161–175, 2015.
- MATSUMOTO, M.; UMEDA, Y.; TSUCHIYA, S.; TANG, L. Development of demand forecasting model for automotive electric component remanufacturing. In: **ELECTRONICS GOES GREEN 2016+(EGG)**, 2016., 2016. **Anais...** 2016. p. 1–7.
- MATTERA, D.; HAYKIN, S. Support vector machines for dynamic reconstruction of a chaotic system. In: **ADVANCES IN KERNEL METHODS**, 1999. **Anais...** 1999. p. 211–241.
- MATTOS NETO, P. S. de; CAVALCANTI, G. D.; FIRMINO, P. R.; SILVA, E. G.; NOVA FILHO, S. R. V. A temporal-window framework for modelling and forecasting time series. **Knowledge-Based Systems**, v. 193, p. 105476, 2020.
- MCCORMICK, G. P. Communications to the editor—exponential forecasting: some new variations. **Management Science**, v. 15, n. 5, p. 311–320, 1969.
- MCKENZIE, E.; GARDNER JR, E. S. Damped trend exponential smoothing: a modelling viewpoint. **International Journal of Forecasting**, v. 26, n. 4, p. 661–665, 2010.
- MCMILLAN, D. G. Which Variables Predict and Forecast Stock Market Returns? In: **Predicting Stock Returns**. Springer, 2018. p. 77–101.
- MELO MENEZES, B. A. de; SIQUEIRA BRAGA, D. de; HELLINGRATH, B.; LIMA NETO, F. B. de. An evaluation of forecasting methods for anticipating spare parts demand. In: **LATIN AMERICA CONGRESS ON COMPUTATIONAL INTELLIGENCE (LA-CCI)**, 2015., 2015. **Anais...** 2015. p. 1–6.
- MERCER, J. **Functions of positive and negative type and their connection with the theory of integral equations**, **Philosophical Transactions of the Royal Society of London, Ser. A**, 1909.
- MERINO, M.; RAMIREZ-NAFARRATE, A. Estimation of retail sales under competitive location in Mexico. **Journal of Business Research**, v. 69, n. 2, p. 445–451, 2016.
- METROPOLIS, N.; ROSENBLUTH, A. W.; ROSENBLUTH, M. N.; TELLER, A. H.; TELLER, E. Equation of state calculations by fast computing machines. **The journal of chemical physics**, v. 21, n. 6, p. 1087–1092, 1953.
- MEYER, D.; DIMITRIADOU, E.; HORNIK, K.; WEINGESSEL, A.; LEISCH, F. **Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: e1071)**. tu wien. r package version 1.7-2.: e1071. 2020.
- MIRCETIC, D.; NIKOLICIC, S.; MASLARIC, M.; RALEVIC, N.; DEBELIC, B. Development of S-ARIMA Model for Forecasting Demand in a Beverage Supply Chain. **Open Engineering**, v. 6, n. 1, 2016.

- MITCHELL, T. M.; LEARNING, M. McGraw-hill science. **Engineering/Math**, v. 1, p. 27, 1997.
- MIYAOKA, J.; HAUSMAN, W. H. How improved forecasts can degrade decentralized supply chains. **Manufacturing & Service Operations Management**, v. 10, n. 3, p. 547–562, 2008.
- MIZRACH, B. The distribution of the Theil U-statistic in bivariate normal populations. **Economics Letters**, v. 38, n. 2, p. 163–167, 1992.
- MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. **Introduction to time series analysis and forecasting**. John Wiley & Sons, 2015.
- MONTGOMERY, D. C.; RUNGER, G. C. **Applied statistics and probability for engineers**. John Wiley & Sons, 2010.
- MOON, M. A. **Demand and supply integration: the key to world-class demand forecasting**. Walter de Gruyter GmbH & Co KG, 2018.
- MOR, R. S.; JAISWAL, S. K.; SINGH, S.; BHARDWAJ, A. Demand Forecasting of the Short-Lifecycle Dairy Products. In: **Understanding the Role of Business Analytics**. Springer, 2019. p. 87–117.
- MORETTIN, P. A.; TOLOI, C. **Análise de séries temporais**. Blucher, 2006.
- MOURTZIS, D.; DOUKAS, M. Design and planning of manufacturing networks for mass customisation and personalisation: challenges and outlook. **Procedia Cirp**, v. 19, p. 1–13, 2014.
- MRÓWCZYŃSKA, B.; ŁACHACZ, K.; HANISZEWSKI, T.; ŚLADKOWSKI, A. A comparison of forecasting the results of road transportation needs. **Transport**, v. 27, n. 1, p. 73–78, 2012.
- MUHAMMAD, H. L.; MARIA, E. N.; HOSSAIN, J. S.; NUR, H.; NUR, A. et al. Fuzzy time series: an application to tourism demand forecasting. **American Journal of Applied Sciences**, v. 9, n. 1, p. 132–140, 2012.
- MULA, J.; POLER, R.; GARCÍA-SABATER, J. P.; LARIO, F. C. Models for production planning under uncertainty: a review. **International journal of production economics**, v. 103, n. 1, p. 271–285, 2006.
- MÜLLER, K.-R.; SMOLA, A. J.; RÄTSCH, G.; SCHÖLKOPF, B.; KOHLMORGEN, J.; VAPNIK, V. Predicting time series with support vector machines. In: INTERNATIONAL CONFERENCE ON ARTIFICIAL NEURAL NETWORKS, 1997. **Anais...** 1997. p. 999–1004.
- NAIM, I.; MAHARA, T.; IDRISI, A. R. Effective Short-Term Forecasting for Daily Time Series with Complex Seasonal Patterns. **Procedia computer science**, v. 132, p. 1832–1841, 2018.

- NETO, A. B. S.; FERREIRA, T. A. E.; BATISTA, M. C. M.; FIRMINO, P. R. A. Studying the Performance of Cognitive Models in Time Series Forecasting. **Revista de Informática Teórica e Aplicada**, v. 27, n. 1, p. 83–91, 2020.
- NEWBOLD, P.; GRANGER, C. W. Experience with forecasting univariate time series and the combination of forecasts. **Journal of the Royal Statistical Society. Series A (General)**, p. 131–165, 1974.
- NGUYEN, H. T.; WALKER, C. L.; WALKER, E. A. **A first course in fuzzy logic**. CRC press, 2018.
- NUNES, I.; DA SILVA, H. S. **Artificial neural networks: a practical course**. Springer, 2018.
- OHTA, S.; HIRAMOTO, R.; KITAMURA, A. Strategic Decision Making of the Product-Mix Using a New Demand Forecasting Model in the Manufacturing Industry. **Journal of Japan Industrial Management Association**, v. 64, n. 4E, p. 614–619, 2014.
- OLIVEIRA, F. A. de; NOBRE, C. N.; ZARATE, L. E. Applying Artificial Neural Networks to prediction of stock price and improvement of the directional prediction index—Case study of PETR4, Petrobras, Brazil. **Expert Systems with Applications**, v. 40, n. 18, p. 7596–7606, 2013.
- OLIVEIRA, J. F. de; SILVA, E. G.; MATTOS NETO, P. S. de. A hybrid system based on dynamic selection for time series forecasting. **IEEE Transactions on Neural Networks and Learning Systems**, 2021.
- PÁDUA BRAGA, A. de; LEON FERREIRA, A. C. P. de; LUDERMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. LTC Editora Rio de Janeiro, Brazil:, 2007.
- PANKRATZ, A. **Forecasting with univariate Box-Jenkins models: concepts and cases**. John Wiley & Sons, 2009. v. 224.
- PAULA, R. O. d. Máquinas de suporte vetorial como instrumento de priorização de investimentos aplicado ao Programa de Investimento em Logística–Aeroportos. , 2016.
- PECHMANN, A.; ZARTE, M. Procedure for generating a basis for PPC systems to schedule the production considering energy demand and available renewable energy. **Procedia CIRP**, v. 64, p. 393–398, 2017.
- PEFFERS, K.; TUUNANEN, T.; ROTHENBERGER, M. A.; CHATTERJEE, S. A design science research methodology for information systems research. **Journal of management information systems**, v. 24, n. 3, p. 45–77, 2007.
- PEGELS, C. C. Exponential forecasting: some new variations. **Management Science**, p. 311–315, 1969.
- PEÑA, D.; TIAO, G. C.; TSAY, R. S. **A course in time series analysis**. John Wiley & Sons, 2011. v. 322.

PEREIRA, M. M.; MACHADO, R. L.; PIRES, S. R. I.; DANTAS, M. J. P.; ZALUSKI, P. R.; FRAZZON, E. M. Forecasting scrap tires returns in closed-loop supply chains in Brazil. **Journal of Cleaner Production**, v. 188, p. 741–750, 2018.

PÉREZ-GALLARDO, J. R.; HERNÁNDEZ-VERA, B.; MORAS SÁNCHEZ, C. G.; AGUILAR LASSERRE, A. A.; POSADA-GÓMEZ, R.; JUÁREZ-MARTÍNEZ, U.; ALOR-HERNÁNDEZ, G. Methodology for supply Chain integration: a case study in the artisan industry of footwear. **Mathematical Problems in Engineering**, v. 2014, 2014.

PETROPOULOS, F.; KOURENTZES, N. Forecast combinations for intermittent demand. **Journal of the Operational Research Society**, v. 66, n. 6, p. 914–924, 2015.

PHAM, H. T.; YANG, B.-S. et al. A hybrid of nonlinear autoregressive model with exogenous input and autoregressive moving average model for long-term machine state forecasting. **Expert Systems with Applications**, v. 37, n. 4, p. 3310–3317, 2010.

PIRES, M. G. et al. Aprendizado genético de funções de pertinência na modelagem nebulosa. , 2004.

PRAK, D.; TEUNTER, R. A general method for addressing forecasting uncertainty in inventory models. **International Journal of Forecasting**, v. 35, n. 1, p. 224–238, 2019.

PRAK, D.; TEUNTER, R.; SYNTETOS, A. On the calculation of safety stocks when demand is forecasted. **European Journal of Operational Research**, v. 256, n. 2, p. 454–461, 2017.

PRAKASH, P. B.; RAMYA, V.; YUGANDHAR, M. A Statistical GA Based Demand Forecasting Model for Automotive Batteries Manufacturing Company. **i-Manager's Journal on Mechanical Engineering**, v. 6, n. 3, p. 18, 2016.

PRECHELT, L. Neural Networks: tricks of the trade. chapter “early stopping—but when. , 2012.

PRODANOV, C. C.; FREITAS, E. C. de. **Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico-2ª edição**. Editora Feevale, 2013.

QIN, M.; LI, Z.; DU, Z. Red tide time series forecasting by combining ARIMA and deep belief network. **Knowledge-Based Systems**, v. 125, p. 39–52, 2017.

R Core Team. **R: a language and environment for statistical computing**. Vienna, Austria, 2019.

REGO, J. R. do; MESQUITA, M. A. de. Demand forecasting and inventory control: a simulation study on automotive spare parts. **International Journal of Production Economics**, v. 161, p. 1–16, 2015.

REID, D. J. Combining three estimates of gross domestic product. **Economica**, v. 35, n. 140, p. 431–444, 1968.

- REID, R. D.; SANDERS, N. R. **Operations Management, Binder Ready Version: an integrated approach.** John Wiley & Sons, 2015.
- ROSIENKIEWICZ, M.; CHLEBUS, E.; DETYNA, J. A hybrid spares demand forecasting method dedicated to mining industry. **Applied Mathematical Modelling**, v. 49, p. 87–107, 2017.
- ROSS, S. M. **Introduction to probability and statistics for engineers and scientists.** Academic Press, 2014.
- ROSS, T. J. **Fuzzy logic with engineering applications.** John Wiley & Sons, 2005.
- ROSTAMI-TABAR, B.; BABAI, M. Z.; SYNTETOS, A.; DUCQ, Y. Demand forecasting by temporal aggregation. **Naval Research Logistics (NRL)**, v. 60, n. 6, p. 479–498, 2013.
- RUSSEL, S.; NORVIG, P. *Inteligência Artificial.* 3a Edição. **Editora Campus**, 2013.
- SAGAERT, Y. R.; AGHEZZAF, E.-H.; KOURENTZES, N.; DESMET, B. Temporal big data for tire industry tactical sales forecasting. **Interfaces**, v. 48, n. 2, p. 121–129, 2018.
- SAHIN, M.; ELDEMIR, F. Application of QR Policy for Non-smooth Demand in the Aviation Industry. In: **Industrial Engineering in the Industry 4.0 Era.** Springer, 2018. p. 159–171.
- SANDERS, N. R.; REID, R. D. **Operations management: an integrated approach.** John Wiley & Sons, 2013.
- SAPANKEVYCH, N. I.; SANKAR, R. Time series prediction using support vector machines: a survey. **IEEE Computational Intelligence Magazine**, v. 4, n. 2, p. 24–38, 2009.
- SARHANI, M.; EL AFIA, A. Intelligent system based support vector regression for supply chain demand forecasting. In: **SECOND WORLD CONFERENCE ON COMPLEX SYSTEMS (WCCS), 2014., 2014.** *Anais...* 2014. p. 79–83.
- SASU, A. An application of fuzzy time series to the Romanian population. **Bulletin of the Transilvania University of Brasov. Mathematics, Informatics, Physics. Series III**, v. 3, p. 125, 2010.
- SCARPEL, R. A. An integrated mixture of local experts model for demand forecasting. **International Journal of Production Economics**, v. 164, p. 35–42, 2015.
- SCHOLZ-REITER, B.; KÜCK, M.; LAPPE, D. Prediction of customer demands for production planning—Automated selection and configuration of suitable prediction methods. **CIRP Annals**, v. 63, n. 1, p. 417–420, 2014.
- SCHUH, G.; POTENTE, T.; FUCHS, S. Shifting bottlenecks in production control. In: **Enabling Manufacturing Competitiveness and Economic Sustainability.** Springer, 2012. p. 505–511.

- SCHWARZ, G. et al. Estimating the dimension of a model. **The annals of statistics**, v. 6, n. 2, p. 461–464, 1978.
- SEGALL, R. S. Some mathematical and computer modelling of neural networks. **Applied mathematical modelling**, v. 19, n. 7, p. 386–399, 1995.
- SERGIO, A. T. Seleção dinâmica de combinadores de previsão de séries temporais. , 2017.
- SEVERINO, A. J. **Metodologia do trabalho científico**. Cortez editora, 2017.
- SHUMWAY, R. H.; STOFFER, D. S. Time series regression and exploratory data analysis. In: **Time series analysis and its applications**. Springer, 2011. p. 47–82.
- SIVARAMAKUMAR, G.; JAIN, S. Production planning with uncertain demands and real options. In: IEEE INTERNATIONAL CONFERENCE ON AUTOMATION SCIENCE AND ENGINEERING (CASE), 2014., 2014. **Anais...** 2014. p. 671–676.
- SLACK, N.; CHAMBERS, S.; JOHNSTON, R. **Operations management**. Pearson education, 2010.
- SMITH, J.; WALLIS, K. F. A simple explanation of the forecast combination puzzle. **Oxford Bulletin of Economics and Statistics**, v. 71, n. 3, p. 331–355, 2009.
- SMOLA, A. J.; SCHÖLKOPF, B. A tutorial on support vector regression. **Statistics and computing**, v. 14, n. 3, p. 199–222, 2004.
- SONG, Q.; CHISSOM, B. S. Forecasting Enrollments with Fuzzy Time Series. , 1991.
- SONG, Q.; CHISSOM, B. S. Forecasting enrollments with fuzzy time series—part I. **Fuzzy sets and systems**, v. 54, n. 1, p. 1–9, 1993.
- SONG, Q.; CHISSOM, B. S. Fuzzy time series and its models. **Fuzzy sets and systems**, v. 54, n. 3, p. 269–277, 1993.
- SONG, Q.; CHISSOM, B. S. Forecasting enrollments with fuzzy time series-Part II. **Fuzzy sets and systems**, v. 62, n. 1, p. 1–8, 1994.
- SONI, R.; SRIKANTH, D. Inventory forecasting model using genetic programming and Holt-Winter's exponential smoothing method. In: IEEE INTERNATIONAL CONFERENCE ON RECENT TRENDS IN ELECTRONICS, INFORMATION & COMMUNICATION TECHNOLOGY (RTEICT), 2017., 2017. **Anais...** 2017. p. 2086–2091.
- STEINWART, I.; CHRISTMANN, A. **Support vector machines**. Springer Science & Business Media, 2008.
- STEVENSON, W.; SUM, C. C. **Operations management**. McGraw-Hill Higher Education, 2015.

STITSON, M.; GAMMERMAN, A.; VAPNIK, V.; VOVK, V.; WATKINS, C.; WESTON, J. Support vector regression with ANOVA decomposition kernels. **Advances in kernel methods—Support vector learning**, p. 285–292, 1999.

STOSIC, D.; STOSIC, D.; LUDERMIR, T.; OLIVEIRA, W. de; STOSIC, T. Foreign exchange rate entropy evolution during financial crises. **Physica A: Statistical Mechanics and its Applications**, v. 449, p. 233–239, 2016.

STÜKER, T. A. Modelo para atualização da previsão de demanda em cadeia de suprimentos de moda rápida na indústria calçadista. , 2014.

SULLIVAN, J.; WOODALL, W. H. A comparison of fuzzy forecasting and Markov modeling. **Fuzzy Sets and Systems**, v. 64, n. 3, p. 279–293, 1994.

SYNTETOS, A. A.; BABAI, M. Z.; GARDNER JR, E. S. Forecasting intermittent inventory demands: simple parametric methods vs. bootstrapping. **Journal of Business Research**, v. 68, n. 8, p. 1746–1752, 2015.

SYNTETOS, A. A.; BABAI, Z.; BOYLAN, J. E.; KOLASSA, S.; NIKOLOPOULOS, K. Supply chain forecasting: theory, practice, their gap and the future. **European Journal of Operational Research**, v. 252, n. 1, p. 1–26, 2016.

SYNTETOS, A. A.; BOYLAN, J. E. On the bias of intermittent demand estimates. **International journal of production economics**, v. 71, n. 1-3, p. 457–466, 2001.

SZOPLIK, J. Forecasting of natural gas consumption with artificial neural networks. **Energy**, v. 85, p. 208–220, 2015.

TADIĆ, S.; ZEČEVIĆ, S. DEVELOPMENT OF INTERMODAL TRANSPORT AND LOGISTICS IN SERBIA. **International Journal for Traffic & Transport Engineering**, v. 2, n. 4, 2012.

TANG, R.; FONG, S.; DEB, S.; VASILAKOS, A. V.; MILLHAM, R. C. Dynamic group optimisation algorithm for training feed-forward neural networks. **Neurocomputing**, v. 314, p. 1–19, 2018.

TAYLOR, J. W. Exponential smoothing with a damped multiplicative trend. **International journal of Forecasting**, v. 19, n. 4, p. 715–725, 2003.

THAO, N.-T. An improved fuzzy time series forecasting model using the differential evolution algorithm. **Journal of Intelligent & Fuzzy Systems**, v. 36, n. 2, p. 1727–1741, 2019.

TIMMERMANN, A. **Forecast Combinations, Handbook of Economic Forecasting, vol. 1, chap. 4, 135–196.** Elsevier, 2006.

TIRKES, G.; GURAY, C.; CELEBI, N. DEMAND FORECASTING: a comparison between the holt-winters, trend analysis and decomposition models/predvidanje potraznje: usporedba između holt-winters modela, analize trenda i modela dekompozicije. **Tehniski Vjesnik-Technical Gazette**, v. 24, n. S2, p. 503–510, 2017.

TORBAT, S.; KHASHEI, M.; BIJARI, M. A hybrid probabilistic fuzzy ARIMA model for consumption forecasting in commodity markets. **Economic Analysis and Policy**, v. 58, p. 22–31, 2018.

TOZAN, H.; VAYVAY, O. Fuzzy forecasting applications on supply chains. **WSEAS Transactions on Systems**, v. 7, n. 5, p. 600–609, 2008.

TRAPERO, J. R.; KOURENTZES, N.; FILDES, R. Impact of information exchange on supplier forecasting performance. **Omega**, v. 40, n. 6, p. 738–747, 2012.

TRAPERO, J. R.; KOURENTZES, N.; FILDES, R. On the identification of sales forecasting models in the presence of promotions. **Journal of the operational Research Society**, v. 66, n. 2, p. 299–307, 2015.

TRATAR, L. F.; MOJŠKERC, B.; TOMAN, A. Demand forecasting with four-parameter exponential smoothing. **International Journal of Production Economics**, v. 181, p. 162–173, 2016.

TSIRIKOGLU, P.; ABRAHAM, S.; CONTINO, F.; LACOR, C.; GHORBANIASL, G. A hyperparameters selection technique for support vector regression models. **Applied Soft Computing**, v. 61, p. 139–148, 2017.

UZSOY, R.; FOWLER, J. W.; MÖNCH, L. A survey of semiconductor supply chain models Part II: demand planning, inventory management, and capacity planning. **International Journal of Production Research**, v. 56, n. 13, p. 4546–4564, 2018.

UZZAMAN, I.; RAHMAN, M.; ALAM, M.; ALAM, S. et al. Simulation of Cement Manufacturing Process and Demand Forecasting of Cement Industry. **Global Journal of Researches in Engineering: G Industrial Engineering**, v. 16, n. 2, 2016.

VAPNIK, V. **The nature of statistical learning theory**. Springer science & business media, 2013.

VAPNIK, V.; VAPNIK, V. **Statistical learning theory** Wiley. **New York**, v. 1, 1998.

VARGAS, C.; CORTÉS, M. E. Automobile spare-parts forecasting: a comparative study of time series methods. **International Journal of Automotive & Mechanical Engineering**, v. 14, n. 1, 2017.

VEIGA, C. P. da; VEIGA, C. R. P. da; TORTATO, U. Demand forecasting strategies: understanding the most important concepts. **Revista ESPACIOS| Vol. 37 (Nº 05) Año 2016**, 2016.

VOVAN, T. An improved fuzzy time series forecasting model using variations of data. **Fuzzy Optimization and Decision Making**, v. 18, n. 2, p. 151–173, 2019.

VRONTAKIS, K. E.; KAMPIANAKIS, A. N.; TSINARAKIS, G. J. A Petri net based methodology for modelling, analysis, demand forecast and optimal planning of batch production systems. In: IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN, AND CYBERNETICS (SMC), 2016., 2016. **Anais...** 2016. p. 001136–001141.

WALLSTRÖM, P.; SEGERSTEDT, A. Evaluation of forecasting error measurements and techniques for intermittent demand. **International Journal of Production Economics**, v. 128, n. 2, p. 625–636, 2010.

WANF, L. A course in Fuzzy Systems and control [M]. , 2003.

WANG, L.; WANG, Z.; QU, H.; LIU, S. Optimal forecast combination based on neural networks for time series forecasting. **Applied Soft Computing**, v. 66, p. 1–17, 2018.

WANG, N.-Y.; CHEN, S.-M. Temperature prediction and TAIFEX forecasting based on automatic clustering techniques and two-factors high-order fuzzy time series. **Expert Systems with Applications**, v. 36, n. 2, p. 2143–2154, 2009.

WANG, Q.; LI, S.; LI, R. Forecasting energy demand in China and India: using single-linear, hybrid-linear, and non-linear time series forecast techniques. **Energy**, v. 161, p. 821–831, 2018.

WANG, X.; PETROPOULOS, F. To select or to combine? The inventory performance of model and expert forecasts. **International Journal of Production Research**, v. 54, n. 17, p. 5271–5282, 2016.

WANG, X.; WEN, J.; ZHANG, Y.; WANG, Y. Real estate price forecasting based on SVM optimized by PSO. **Optik**, v. 125, n. 3, p. 1439–1443, 2014.

WERNER, L. Um modelo composto para realizar previsão de demanda através da integração da combinação de previsões e do ajuste baseado na opinião. , 2005.

WIENDAHL, H.-P. **Betriebsorganisation für Ingenieure**. Carl Hanser Verlag GmbH Co KG, 2014.

WILLMOTT, C. J. Some comments on the evaluation of model performance. **Bulletin of the American Meteorological Society**, v. 63, n. 11, p. 1309–1313, 1982.

WINKLER, R. L.; CLEMEN, R. T. Sensitivity of weights in combining forecasts. **Operations research**, v. 40, n. 3, p. 609–614, 1992.

WINKLER, R. L.; MAKRIDAKIS, S. The combination of forecasts. **Journal of the Royal Statistical Society. Series A (General)**, p. 150–157, 1983.

WINTERS, P. R. Forecasting sales by exponentially weighted moving averages. **Management science**, v. 6, n. 3, p. 324–342, 1960.

WINTHOR, T. L. Base de Conhecimento PC Sistemas. , p. Acesso em 30 junho 2021, 2021.

WRIGHT, G.; GOODWIN, P. Decision making and planning under low levels of predictability: enhancing the scenario method. **International Journal of Forecasting**, v. 25, n. 4, p. 813–825, 2009.

WU, L.; LIU, S.; YANG, Y. Grey double exponential smoothing model and its application on pig price forecasting in China. **Applied Soft Computing**, v. 39, p. 117–123, 2016.

XI, X.; KEOGH, E.; SHELTON, C.; WEI, L.; RATANAMAHATANA, C. A. Fast time series classification using numerosity reduction. In: MACHINE LEARNING, 23., 2006. **Proceedings...** 2006. p. 1033–1040.

XIANG, Y.; GUBIAN, S.; SUOMELA, B.; HOENG, J. Generalized Simulated Annealing for Global Optimization: the gensa package. **The R Journal**, v. 5, n. 1, p. 13–28, 2013.

XU, S.; CHAN, H. K.; ZHANG, T. Forecasting the demand of the aviation industry using hybrid time series SARIMA-SVR approach. **Transportation Research Part E: Logistics and Transportation Review**, v. 122, p. 169–180, 2019.

YAFFEE, R. A.; MCGEE, M. **An introduction to time series analysis and forecasting: with applications of sas® and spss®**. Elsevier, 2000.

YANG, S.; ARNDT, T.; LANZA, G. A flexible simulation support for production planning and control in small and medium enterprises. **Procedia CIRP**, v. 56, p. 389–394, 2016.

YAO, J.; TAN, C. L. A case study on using neural networks to perform technical forecasting of forex. **Neurocomputing**, v. 34, n. 1-4, p. 79–98, 2000.

ZADEH, L. A. Fuzzy sets. **Information and control**, v. 8, n. 3, p. 338–353, 1965.

ZADEH, L. A. Outline of a new approach to the analysis of complex systems and decision processes. **IEEE Transactions on systems, Man, and Cybernetics**, n. 1, p. 28–44, 1973.

ZHANG, B.-x.; HU, S.-q.; SONG, J.; CHENG, S. Analysis about medium-long-term demands forecasting and capacity decision based on the grey GM (1, 1) improved models. In: INTERNATIONAL CONFERENCE ON MANAGEMENT SCIENCE & ENGINEERING 19TH ANNUAL CONFERENCE PROCEEDINGS, 2012., 2012. **Anais...** 2012. p. 276–281.

ZHANG, G. P. Time series forecasting using a hybrid ARIMA and neural network model. **Neurocomputing**, v. 50, p. 159–175, 2003.

ZHANG, G.; PATUWO, B. E.; HU, M. Y. Forecasting with artificial neural networks:: the state of the art. **International journal of forecasting**, v. 14, n. 1, p. 35–62, 1998.

ZHANG, X.-D. **A Matrix Algebra Approach to Artificial Intelligence**. Springer, 2020.

ZHU, S.; DEKKER, R.; VAN JAARVELD, W.; RENJIE, R. W.; KONING, A. J. An improved method for forecasting spare parts demand using extreme value theory. **European Journal of Operational Research**, v. 261, n. 1, p. 169–181, 2017.

ZIPFEL, A.; BRAUNREUTHER, S.; REINHART, G. Approach for a Production Planning and Control System in Value-Adding Networks. **Procedia CIRP**, v. 81, p. 1195–1200, 2019.

ZOU, H.; XIA, G.; YANG, F.; WANG, H. An investigation and comparison of artificial neural network and time series models for Chinese food grain price forecasting. **Neurocomputing**, v. 70, n. 16-18, p. 2913–2923, 2007.

ZULKEPLI, J.; FONG, C. H.; ABIDIN, N. Z. Demand forecasting for automotive sector in Malaysia by system dynamics approach. In: AIP CONFERENCE PROCEEDINGS, 2015. **Anais...** 2015. v. 1691, n. 1, p. 030031.