

Pontifícia Universidade Católica de Goiás
Programa de Mestrado em Engenharia de Produção e Sistema

**AVALIAÇÃO E MENSURAÇÃO DO DESEMPENHO DO NÍVEL
DE SERVIÇO LOGÍSTICO**

LUCAS DE SOUZA CAMPOS

2020

AVALIAÇÃO E MENSURAÇÃO DO DESEMPENHO DO NÍVEL DE SERVIÇO LOGÍSTICO

LUCAS DE SOUZA CAMPOS

Coletânea de artigos apresentado ao programa de Pós-Graduação em Engenharia de Produção e Sistemas da Pontifícia Universidade Católica de Goiás, como parte dos requisitos para obtenção do título de Mestre em Engenharia de Produção e Sistemas.

Orientador: Prof. José Elmo de Menezes, *Dr.*

Goiânia

Março de 2020

AVALIAÇÃO E MENSURAÇÃO DO DESEMPENHO DO NÍVEL DE SERVIÇO LOGÍSTICO

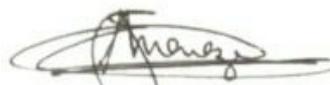
LUCAS DE SOUZA CAMPOS

Esta coletânea de artigos julgada adequada para obtenção do título de Mestre em Engenharia de Produção e Sistemas, e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia de Produção e Sistemas da Pontifícia Universidade Católica de Goiás em março de 2020.



Prof. Marcos Lajovic Carneiro, Dr.
Coordenador do Programa de Pós-Graduação em
Engenharia de Produção e Sistemas

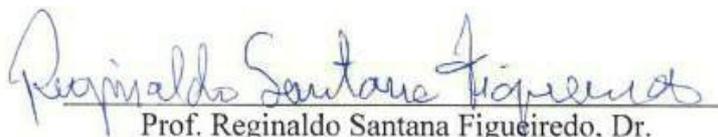
Banca Examinadora:



Prof. José Elmo de Menezes, Dr.
Orientador



Prof. Clóves Gonçalves Rodrigues, Dr.



Prof. Reginaldo Santana Figueiredo, Dr.

Goiânia

Março de 2020

C198a Campos, Lucas de Souza

Avaliação e mensuração do desempenho do nível de serviço logístico / Lucas de Souza Campos. -- 2020.

Goiás/MEPROS, 2020

83 f.; il.

Texto em português, com resumo em inglês.

Dissertação (mestrado) – Pontifícia Universidade Católica de Goiás, Escola de Engenharia, Goiânia, 2020
Inclui referências: f. 69-78

Dissertação (mestrado) - PUC Goiás/ MEPROS/
Programa de Pós-Graduação em Engenharia de Produção e Sistemas, 2020.

1. Mínimos quadrados. 2. Programação genética (Computação).
3. Análise de séries temporais. I. Menezes, José Elmo de. II Pontifícia Universidade Católica de Goiás – Programa de Pós-Graduação em Engenharia de Produção e Sistemas – 2020. III. Título.

CDU: Ed. 2007 – 658.78:004(043)

Dedico este trabalho a minha infável esposa Rafaela Ferreira da Silva Garrido Paiva, pelo companheirismo, incentivo, alegria e entendimento nos momentos de minha ausência. A meu pai Ailton Campos Pereira, que apesar de todas as dificuldades me fortaleceu e que para mim foi muito importante. A minha mãe Elza Maria de Souza Campos, heroína, que me deu apoio, incentivou nas horas difíceis de desânimo e cansaço. A meus irmãos João Elias de Souza Neto e Lorena de Souza Campos, que fizeram entender nos momentos de ausência.

AGRADECIMENTOS

Agradeço ao soberano e eterno Deus, pelo dom da vida e pela maravilhosa graça. Por ter colocado em minha vida pessoas que contribuíram para minha formação como pessoa e como profissional. Agradeço a Ele também por ter dado a oportunidade de realizar este tão sonhado curso.

Aos meus pais Ailton Campos Pereira e Elza Maria de Souza Campos que sempre estiveram ao meu lado me apoiando ao longo de toda a minha trajetória.

A minha esposa Rafaela Ferreira da Silva Garrido Paiva pela compreensão, companheirismo e paciência demonstrada durante o período do projeto.

A meu orientador, professor Dr. José Elmo de Menezes, pelo aprendizado compartilhado, confiança, paciência e empenho dedicado à elaboração deste trabalho.

Aos professores Clarimar José Coelho, Clóves Gonçalves Rodrigues, Maria José Pereira Dantas, Ricardo Luiz Machado e Solange da Silva pelos conhecimentos técnico e científico passado. Ao fiel amigo Leonardo Rodrigues de Oliveira Merelles pelos momentos de estudos juntos, pelo carinho e atenção demonstrado. Sou extremamente grato.

Aos colegas corporativos, André Luiz Nogueira, Hemerson Luciano Pereira, Paulo Henrique Silva dos Santos e Rafael Gomes Vilar Ferreira pela compreensão e ajuda na realização deste trabalho.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro.

“Ainda que a figueira não floresça, nem haja fruto na vide; o produto da oliveira minta, e os campos não produzam mantimento; as ovelhas sejam arrebatadas do aprisco, e nos currais não haja gado, todavia, eu me alegro no Senhor, exulto no Deus da minha salvação. O Senhor Deus é a minha fortaleza, e faz os meus pés como os da corça, e me faz andar altaneiramente.”

Habacuque 3:17-1

Resumo da coletânea de artigos apresentada ao MEPROS/PUC Goiás como parte dos requisitos necessários para a obtenção do grau de Mestre em Engenharia de Produção e Sistemas (M.Sc.)

AVALIAÇÃO E MENSURAÇÃO DO DESEMPENHO DO NÍVEL DE SERVIÇO LOGÍSTICO

Março / 2020

RESUMO:

Esta pesquisa propõe analisar aspectos cruciais no atendimento logístico de uma empresa estudada, a fim de haver uma melhora percebida no nível de serviço, com o propósito de se manter competitiva no mercado em que atua. O caso em estudo é de uma empresa localizada na região Centro-Oeste do Brasil, que produz, comercializa e distribui produtos alimentícios. A pesquisa parte da elaboração de um método matemático para identificação das variáveis que influenciam a *performance* do atendimento de pedidos. Compõe-se também da construção de um modelo para mensurar o atendimento destes pedidos. Por fim, propõe um modelo de previsão de demanda, para execução do planejamento e preparação para atendimento de vendas. Foram duas abordagens metodológicas arrematadas nesta pesquisa, sendo a modelagem utilizando o software R e o estudo de caso, onde foi analisado relatórios de pedidos e documentações de demanda. Logo, esta pesquisa se tornou exequível identificar variáveis que impactam a *performance* do atendimento dos pedidos, tanto de forma negativa como de forma positiva, foi possível também medir o desempenho através do modelo aplicado, logo, possibilitou traçar uma previsão de demanda com o modelo de séries temporais, a fim de planejar a capacidade de distribuição.

Palavras-chave: Mínimos quadrados parciais (PLS). Programação genética. Séries temporais. Nível de serviço.

Abstract of the collectanea of articles presented to MEPROS / PUC Goiás as part of the requirements for obtaining a Master's degree in Production and Systems Engineering (M.Sc.)

EVALUATION AND MEASUREMENT OF THE PERFORMANCE OF THE LOGISTICS SERVICE LEVEL

February / 2020

ABSTRACT:

This research proposes analyze aspects that are crucial in the logistical service of a studied company, in order to have a perceived improvement in the level of service, in order to remain competitive in the market in which it operates. The case under study is of a company located in the Midwest region of Brazil, which produces, sells and distributes food products. The research starts from the elaboration of a mathematical method to identify the variables that influence the performance of order fulfillment. It also consists of building a model to measure the fulfillment of these requests. Finally, it proposes a demand forecasting model, for executing the planning and preparing for sales service. There were two methodological approaches taken in this research, being the modeling using the R software and the case study, where order reports and demand documentation were analyzed. Therefore, this research became feasible to identify variables that impact the performance of order fulfillment, both negatively and positively, it was also possible to measure performance through the applied model, therefore, it was possible to trace a demand forecast with the time series in order to plan distribution capacity.

Key-words: Partial least squares (PLS). Genetic programming. Time series. Service level.

LISTA DE TABELAS

TABELA 1 - Sequência de artigos, autores e submissões	17
TABELA 2 - <i>Coding of each predictor variable</i>	24
TABELA 3 - <i>Grouping of variables impacting OTIF</i>	32
TABELA 4 - Conjuntos de dados do repositório	43
TABELA 5 - Parâmetros de entrada da programação genética.....	44
TABELA 6 - Resumo dos resultados da programação genética no conjunto de dados de validação cruzada (RMSE).....	45
TABELA 7 - Resumo da precisão da programação genética para testar conjuntos de dados.....	45
TABELA 8 - Resumo dos resultados da regressão linear múltipla.....	46
TABELA 9 - Acurácia do modelo.....	46
TABELA 10 - Acurácia do modelo.....	64

LISTA DE FIGURAS

FIGURE 1 - <i>PLS correlation between OTIF and factors that reduce efficiency. 1a) Behavior of the variables in relation to OTIF over time. 1b) Behavior of OTIF in relation to the week. 1c) dispersion between the X and Y blocks. 1d) Variables that impact the OTIF</i>	29
FIGURE 2 - <i>Cluster analysis by the Ward method.</i>	31
FIGURA 3 - Operador de crossover. (a) árvore de pais. (b) árvores de descendentes..	40
FIGURA 4 - Operador de mutação. (a) árvores pais. (b) árvores de descendentes.....	41
FIGURA 5 - Gráfico da modelagem de índice.....	47
FIGURA 6 - Fluxo do processo de atendimento de pedido	59
FIGURA 7 - Gráfico temporal.....	60
FIGURA 8 - Box-Plot da série de histórico de vendas. Figura 3A Box-Plot com Outliers. Figura 3B Box-plot sem Outliers.	61
FIGURA 9 - Estimativa de vendas	62
FIGURA 10 - Comparativo do modelo e série histórica	65

LISTA DE ABREVEATURAS E SIGLAS

ARIMA	<i>Autoregressive Integrated Moving Average</i>
ARMA	<i>Autoregressive Moving Average</i>
CD	Centro de Distribuição
CV	Validação Cruzada
DMAIC	<i>Define, Measure, Analyze, Improve, Control</i>
EA	Algoritmos Evolutivos
EDDIE	<i>Evolutionary Dynamic Data Investment Evaluator</i>
GA	Algoritmo Genético
GP	Programação Genética
JIT	<i>Just In Time</i>
KPI	<i>Key Performance Indicator</i>
MAE	<i>Mean Absolute Error</i>
MAPE	<i>Mean Absolute Percentage Error</i>
NSE	<i>Nash-Sutcliffe</i>
OTIF	<i>On Time In Full</i>
PBIAS	<i>Percent Bias</i>
PLS	<i>Partial Least Square</i>
RMSE	<i>Root Mean Square Error</i>

SUMÁRIO

1 INTRODUÇÃO	14
1.1 Contexto.....	14
1.2 Problematização.....	15
1.3 Justificativa	16
1.4 Objetivos.....	16
1.5 Estrutura do trabalho.....	17
2 METODOLOGIA.....	18
2.1 Abordagem e método de pesquisa	18
2.2 Objeto de estudo	18
2.3 Método de pesquisa	19
3 VARIABLES THAT INFLUENCE LOGISTIC PERFORMANCE: APPLICATION OF PARTIAL LEAST SQUARE.....	22
3.1 Abstract.....	22
3.2 INTRODUCTION	22
3.3 Materials and methods	24
3.3.1 Correlation PLS.....	26
3.4 Results and discussion	28
3.5 Conclusion	34
4 MODELAGEM DE ÍNDICE COM PROGRAMAÇÃO GENÉTICA.....	36
4.1 Resumo	36
4.2 INTRODUÇÃO	36
4.3 Programação genética	38
4.4 Resultado e implementação de programação genética	43
4.4.1 Modelagem do índice	46
4.5 Conclusões	47
5 PREVISÃO DE DEMANDA NA CADEIA DE SUPRIMENTOS: APLICAÇÃO DO MODELO DE SÉRIES TEMPORAIS.....	48
5.1 Resumo	48
5.2 INTRODUÇÃO	48
5.3 Materiais e métodos	51
5.3.1 Calibração e validação do modelo	56
5.3.2 Caso de teste.....	57
5.4 Resultados e discussão.....	61

5.5 Conclusão.....	65
6 CONSIDERAÇÕES FINAIS.....	67
REFERENCIAS:	69
APÊNDICE:	79

1 INTRODUÇÃO

Neste capítulo são abordados o contexto, a problemática, a justificativa e o objetivo da dissertação a qual foi organizada em forma de coletâneas.

1.1 Contexto

Fazer mais com menos, maximizar lucro e reduzir o custo, otimização de processos, maior eficiência, prospecção, automação, todas essas ações e demandas tem apenas um só objetivo, rentabilidade. Todas as empresas, companhias ou órgãos cujo objetivo é de gerar lucro, foca em diversas atividades do seu processo visando enxugar e reduzir custos. Só que além disso tudo, é necessário investir em relacionamento com o cliente, entender a sua real necessidade, dar foco para o que é de interesse do cliente e assim oferecer o serviço/produto que realmente o satisfaz. É extremamente importante que as empresas estejam focadas e atentas as expectativas dos seus clientes, pois a partir daí a confiabilidade, credibilidade e fidedignidade irão garantir prospecção de clientes e aumento significativo para o seu negócio (CAMPOS & MENEZES, 2019).

A logística pode aumentar o valor percebido pelos clientes e executivos da área na comercialização de produtos ou serviços, em que grande parte desse valor é gerado a partir da capacidade de reduzir custos e fornecer soluções de entrega de acordo com as necessidades dos clientes (GRAWE, 2009).

O gerenciamento da cadeia de fornecimento oferece um grande potencial para as organizações reduzirem custos e melhorarem o desempenho do atendimento ao cliente. No mercado competitivo de hoje, as empresas são pressionadas a atingir altos níveis de serviço ao cliente com menos recursos. Pressões adicionais de maior variedade de produtos, ciclos de vida de produto mais curtos e prazos de entrega mais curtos tornaram

cada vez mais difícil atingir altos níveis de serviço com recursos limitados (JEFFERY, BUTLER, & MALONE, 2008).

A distribuição do produto está atrelada ao nível de serviço e pode impactar na rentabilidade das empresas pela descontinuidade das compras (CHAPMAN, BERNON, & HAGGETT, 2011). As empresas são obrigadas a manter altos níveis de serviço ao cliente e, ao mesmo tempo, são forçadas a reduzir custos e manter as margens de lucro (ALTIPARMAK, GEN, LIN, & PAKSOY, 2006).

Os clientes avaliam as ofertas de qualquer empresa em termos de qualidade, preço e serviço, e dependendo da sua conveniência aproveita ou rejeita tal produto/serviço. O termo serviço ao cliente abrange uma pluralidade de elementos, desde a disponibilidade do produto até o pós-venda, sendo todas as atividades logísticas e processo de suprimentos compondo o termo, logo, o sistema logístico oferece o nível de serviço a ser entregue ao cliente. Até mesmo a estrutura operacional é essencial para atender aos níveis de atendimento ao cliente entregando o produto certo no prazo (ONSTEIN, TAVASSZY, & VAN DAMME, 2019).

1.2 Problematização

Assim, o problema desta pesquisa está concentrado em mensurar, evidenciar e adequar atributos e dimensões que venham contribuir com a *performance* do nível de serviço ao cliente, cumprindo com o critério tomador de pedido. A análise avalia os atributos e dimensões identificados no termo serviço ao cliente. Para tanto, o resultado deste trabalho deve responder à questão de pesquisa:

Como avaliar e mensurar o desempenho do nível de serviço prestado ao cliente final?

A resposta dessa questão de pesquisa é a base para proposta de novos modelos de mensuração e análise de *performance* do serviço prestado, segundo o critério proposto pelo cliente no ato do envio do seu pedido.

1.3 Justificativa

A grande frequência nos atrasos das entregas da empresa estudada e as inúmeras rupturas nos pedidos feitos, onde se pede o que não tem disponível em estoque, desencadeou a descontinuidade de compra por parte do cliente, migrando para o concorrente. Logo, gerou contastes reivindicações pela área de vendas, despertando no setor de logística a identificação das causas e aplicação de melhorias.

O nível de serviço da distribuição logística se tornou a justificativa de estudo para a empresa, com o designo de identificar as não conformidades no processo, mensurar e tratar, a fim de garantir uma melhor *performance* no atendimento ao cliente.

Decidir o nível de serviço oferecido ao cliente é fundamental na concretização da rentabilidade projetada pela organização.

Existem pesquisadores que avaliam o nível de serviço através de métodos como mapeamento do fluxo de valor, outros por modelos de equação estrutural. Outros usam técnicas da Gestão da Qualidade (SUAREZ-BARRAZA, MIGUEL-DAVILA, & VASQUEZ-GARCÍA, 2016; GOH & ELDRIDGE, 2019; CHAPMAN, BERNON, & HAGGETT, 2011). Esta pesquisa usou modelos matemáticos para evidenciar variáveis que impactam o nível de serviço e modelos para auxiliar na medição da *performance* e prever a demanda futura.

1.4 Objetivos

Este trabalho tem como objetivo principal exercer modelos para auxiliar e mensurar na identificação de variáveis que impactam o nível de atendimento ao cliente.

Como objetivo específico, busca: (1) modelar a demanda de vendas; (2) modelar o índice de desempenho da distribuição logística; e (3) identificar as variáveis que impactam o atendimento dos pedidos.

1.5 Estrutura do trabalho

Este trabalho está estruturado em seis capítulos. O capítulo dois é a metodologia aplicada. Os capítulos três, quatro e cinco, são artigos inseridos no compêndio da dissertação. O capítulo seis são as considerações finais.

Na Tabela 1 logo abaixo, evidencia cada artigo com seus devidos autores e revistas submetidas.

TABELA 1 - Sequência de artigos, autores e submissões

Sequência	Título do artigo	Autores	Revista submetida
Capítulo 1	Variáveis que influenciam a <i>performance</i> logística: Aplicação do <i>Partial Least Square</i>	Campos, Lucas Souza; Menezes, José Elmo; Merelles, Leonardo Rodrigues de Oliveira;	<i>International Journal of Physical Distribution & Logistics Management</i>
Capítulo 2	Modelagem de índice com programação genética	Campos, Lucas Souza; Menezes, José Elmo; Merelles, Leonardo Rodrigues de Oliveira;	Gestão & Produção
Capítulo 3	Previsão de demanda na cadeia de suprimentos: Aplicação do modelo de séries temporais	Campos, Lucas Souza; Menezes, José Elmo; Merelles, Leonardo Rodrigues de Oliveira;	Gestão & Produção

2 METODOLOGIA

Neste capítulo será abordada a metodologia que correlaciona os artigos nos demais capítulos. Inicialmente é descrito a abordagem de pesquisa, seguido por objeto de estudo e método de pesquisa.

2.1 Abordagem e método de pesquisa

Com o propósito de medir o impacto do nível de serviço logístico, mensurar o índice de desempenho da distribuição e prever a demanda de vendas, foi utilizado a abordagem de pesquisa quantitativa.

O método de pesquisa aplicado foi o estudo de caso e a modelagem estocástica, pois permite aplicar e testar modelos de previsões em dados históricos.

2.2 Objeto de estudo

A empresa estudada se encontra na região centro-oeste do Brasil, ela produz e comercializa produtos alimentícios congelados, realiza distribuição em todo o estado de Goiás, atendendo lanchonetes, restaurantes, supermercadistas e atacadistas. A empresa detém *market share* de 20% do seu negócio.

Após a produção do produto ele é destinado ao Centro de Distribuição (CD), onde está o estoque posicionado e disponível para venda. A área comercial tem visibilidade deste estoque através do *palmtop*, onde o vendedor digita o pedido do cliente atendido.

Em seguida, estes pedidos entram em uma carteira para serem roteirizados, os quais passarão por alguns critérios de atendimento, tais como disponibilidade do produto em estoque, se o preço está de acordo com a tipologia do cliente, e se o cliente está com credibilidade na praça. Essas e mais variáveis foram utilizadas para identificação e entendimento da *performance* logística de cada pedido.

O histórico dessas vendas foi utilizado para modelagem da previsão da demanda. Os dados foram acessados em maio de 2017 até maio de 2019. Estes dados foram coletados semanalmente, ou seja, foi acumulado os volumes de vendas (em toneladas) durante a semana.

O setor logístico é mensurado pelo atendimento do critério tomador de pedido, métrica denominada como *On Time In Full* (OTIF). Como se calcular o OTIF é descrito no capítulo três. O OTIF captura informações do atendimento do pedido, calculando se o pedido foi atendido no tempo e quantidade correta. O OTIF pode variar de zero a um, onde zero é o menor valor e um o maior valor.

Nesta empresa o OTIF é calculado diariamente, os dados foram coletados de janeiro até dezembro de 2018. Esses dados foram utilizados nos capítulos três e quatro.

2.3 Método de pesquisa

A partir do capítulo três cada artigo evidencia um método diferente. No terceiro capítulo evidencia-se um *Key Performance Indicator* (KPI) que mensura a qualidade do atendimento do pedido, *On-Time/In-Full* (OTIF), métrica adotada por diversas companhias (GODSELL & VAN HOEK, 2009; MACHADO, BARROSO, & MACHADO, 2014; SEHGAL, SAHAY, & GOYAL, 2006).

Neste capítulo, evidencia um estudo de caso em que o OTIF está em patamares insatisfatórios. Com objetivo de elevar a performance do KPI foi aplicado um método estatístico multivariado *Partial Least Square* (PLS), para determinar quais as variáveis que devem ser atacadas para que aumente a performance do OTIF.

Foi possível identificar o comportamento das variáveis e as semanas que mais afetam o indicador comprometendo o resultado de forma direta no OTIF. Este

comportamento das variáveis fica claro quanto ao período que fecha o relatório de vendas e o abastecimento do estoque no centro de distribuição.

No segundo capítulo apresenta o método de Programação Genética (GP), no qual tem sido aplicado em problemas computacionais e de modelagem (LIU, TONG, XIE, & YEE NG, 2015; REZANIA & JAVADI, 2008). Quando existe um conjunto de dados que são os históricos de pedidos e a saída que é o KPI, a programação genética retorna uma expressão ou modelo para o conjunto de dados. Neste capítulo foi proposto selecionar e otimizar expressões para modelar índices usando algoritmos evolucionários.

Com dados contínuos e/ou discretos a Programação Genética apresentou excelente desempenho ao modelar índice, desempenho este que ocorre devido às transformações matemáticas que advêm nos dados durante a modelagem. Assim, a Programação Genética pode ser aplicada para modelagem de expressões que retornam um índice (KPI).

Já no terceiro capítulo trazemos um método de previsão baseado em abordagens de séries temporais, essas abordagens de previsão preveem demanda futura com base em dados históricos de séries temporais, sequência de pontos de dados medidos em intervalos sucessivos em tempo.

A previsão é de fundamental importância para auxiliar em informações e ter maior assertividade em situações complexas de decisões, desde capacidade de frota na distribuição até a mão-de-obra operacional, podendo assim atender o cliente no dia certo e na quantidade certa conforme critério tomador de pedido. A previsão de demanda é fundamental e crucial para programar todo o processo de produção e distribuição ao cliente final.

As oscilações de volume de vendas impactam de forma direta os custos logísticos, e conseqüentemente o nível de serviço prestado ao cliente. Volume baixo de distribuição deixa a frota de veículos e mão-de-obra ocioso, já o elevado número de distribuição pode comprometer a operação e atrasar as entregas deixando o cliente insatisfeito, correndo o risco de ele migrar para o concorrente.

O objetivo do capítulo é de preparar a operação logística para a demanda da distribuição, e minimizar a incerteza do número de distribuição, estando preparado para as oscilações no departamento de vendas.

O modelo evidenciado é o auto-regressivo integrado de médias móveis (*autoregressive integrated moving average*) ou ARIMA, modelo este que permite trazer uma previsão de demanda.

Na prática o ARIMA é raramente usado em cadeias de suprimentos, porém ele pode ser eficientemente utilizado no contexto *de supply chain* (SVETUNKOV & BOYLAN, 2019).

A previsão da cadeia de suprimentos como um descritor tem como objetivo capturar todos os aspectos da previsão para apoiar as operações, desde o varejista ou prestador de serviços através do sistema de distribuição aos fabricantes e fornecedores de terceiro nível (BOONE, GANESHAN, JAIN, & SANDERS, 2019).

3 VARIABLES THAT INFLUENCE LOGISTIC PERFORMANCE: APPLICATION OF PARTIAL LEAST SQUARE

3.1 Abstract

Understanding how punctuality and adherence to orders can reduce supply chain logistics performance is important to improve organizational productivity and customer satisfaction. Thus, metrics are often applied in logistical processes to measure the performance of products and/or services. To improve logistics performance and to increase customer satisfaction a method has been proposed to determine and explain the variables that can reduce the performance of the logistics sector. This case study is about a food company that produces and distributes refrigerated products in the state of Goiás in Brazil. The variables that reduce the performance of the logistics distribution are delay in delivery, product inventory, and credit.

Keywords: Partial Least Square; Logistics Distribution; Service level; On-Time/In-Full (OTIF);

3.2 INTRODUCTION

Supply chain integration is widely regarded by professionals and researchers as a vital contributor to supply chain performance (PRAJOGO & OLHAGER, 2012). In the modern world, competition is no longer between organizations, but between supply chains (TRKMAN, MCCORMACK, DE OLIVEIRA, & LADEIRA, 2010). The supply chain has become a potential and extremely valuable method to ensure competitive advantage and to improve organizational performance.

The distribution of final products to customers can play an important role in supply chain management (WANG, FUNG, & CHAI, 2004). The distribution of a product is linked to the quality of the services and it can affect the companies' profitability because the costumers may discontinue the purchases (CHAPMAN, BERNON, & HAGGETT, 2011). Therefore, many management models and metrics have been introduced in order to improve customer perception (GUPTA & BOYD, 2008).

On-Time/In-Full (OTIF) has been adopted by several companies as a logistic process metric (GODSELL & VAN HOEK, 2009; SEHGAL, SAHAY, & GOYAL, 2006). These system solutions measure and monitor key performance indicators (KPI), which are crucial to optimizing supply chain performance. Measuring performance is important for companies to improve supply chain efficiency. Monitoring and improving the performance of a supply chain has become an increasingly complex task. A complex system of performance management includes many management processes, such as identifying measures, setting goals, planning, communication, monitoring, reporting, and feedback (CAI, LIU, XIAO, & LIU, 2009).

The OTIF calculation is the linear combination of metrics On-Time and In-Full (Equation 1) and it goes from zero to one; zero is the worst performance and one is the best performance.

$$OTIF = \left(\frac{1}{n} \sum_{i=1}^n OnTime_i \right) \left(\frac{1}{n} \sum_{i=1}^n InFull_i \right) \quad (1)$$

OTIF is the performance measurement of the logistics service. On-Time is the delivery of the product and/or service previously agreed with the customer (punctuality of the distribution). In-Full is the delivery of the product and/or service in its entirety (distribution integrity), and n is the total of orders served in one day. The metrics On-

Time and In-Full are measured in percentage; for each service the values of zero (failure) and one (success) are resumed.

A new concept in the supply chain has been applied, in which increasing quality management appears to be applicable to the supply chain; therefore, this new concept successfully integrates the area (SUAREZ-BARRAZA, MIGUEL-DAVILA, & VASQUEZ-GARCÍA, 2016). Quality tools have been applied to enhance the distribution performance, and methods such as Lean Six Sigma and its DMAIC (Define, Measure, Analyze, Improve, and Control) mental model places OTIF at the highest levels of customer satisfaction.

The supply chain practices point out that customer service and satisfaction improve results, such as increasing revenue, reducing inventory costs and product costs, as well as improving delivery and timeframe (SAHAY & MOHAN, 2003).

The Measuring performance in the distribution process allows us to evaluate the effectiveness and efficiency of the organization. In the supply chain, the goal is to minimize cost and deliver value to the customer. OTIF is sensitive to these goals, distributing goods in the right quantity, in the right place, and at the right time, which delivers value to the customer. The maximization of OTIF can leverage strategic advantage over other organizations. To better understand the reasons that reduce efficiency of OTIF, the application of PLS was proposed to qualify and measure the influences of punctuality and integrity of the logistic distribution in the OTIF metric.

3.3 Materials and methods

To verify the influence of activities that reduce the performance of the logistics sector, a case study was carried out in a food company founded in the 1950s. The company supplies refrigerated food in countries such as Brazil and the organization of

this study is located in Aparecida de Goiânia in the state of Goiás. The case study was conducted from January to May in 2018. To verify the behavior of OTIF during those months the data were divided into four groups. The analysis in groups verifies the behavior of the variables that make up the On-Time and In-Full. In addition, it allows the evaluation of performance strategy such as pricing policy. Hence, days 1 to 7 were labeled as the first week, and for the second week, the dates were between days 8 and 14. From day 15 to 21 the third week and in the fourth week the dates were between days 22 and 28. To understand the groups, Partial Least Squares Correlation (PLSC) was applied.

The causes studied in this research have a direct impact on OTIF performance. The predictors and causes were: (1) Minimum quantity corresponds to a lower volume demanded by the company. Low volume orders do not guarantee the cost of the product; therefore, a minimum quantity for the orders have been established. (2) Credit represents the denial of the customer's credit. The order is analyzed, and the credit area approves or disapproves depending on the customer's situation, they analyze if the customer has any financial problems in the market. (3) Price is when the value of the product does not conform to the standard table. (4) Return occurs when the order is not what the customer and the seller agreed upon and the discrepancy causes the return of the order. (5) Lack of stock occurs when there is no stock to meet the request, causing the rupture. (6) In Full Loss is when only a portion of the order is met, that is, when only part of the requested order is approved in the stock. (7) Loss On Time occurs when requests are not met at the time agreed with the customer, which generates delays. (8) Shelf Life happens when the life of the product has not been met. In general each customer has a percentage of shelf life that is accepted, and even when the product is in stock and the shelf life does not go according to the criteria demanded by the customer the request is not met. (9) Others include logistic problems such as lack of vehicles to fulfill an order, lack of labor,

inconsistency in the billing system, or defective machinery. (10) Financial indicates when the order is denied due to a question of customer viability, i.e., when the customer's purchase value is high compared to other orders. (11) Fill Rate occurs when the order is partially attended with volume. (12) The variables from 12 to 36 represent the 25 products that most cause effects in the OTIF measurement. The coding of each predictor variable is in Table 2.

TABELA 2 - Coding of each predictor variable

Variables	Represents	Variables	Represents	Variables	Represents
Variable 1	Minimum quantity	Variable 13	Product 2	Variable 25	Product 14
Variable 2	Credit	Variable 14	Product 3	Variable 26	Product 15
Variable 3	Price	Variable 15	Product 4	Variable 27	Product 16
Variable 4	Return	Variable 16	Product 5	Variable 28	Product 17
Variable 5	Lack of stock	Variable 17	Product 6	Variable 29	Product 18
Variable 6	Loss in Full	Variable 18	Product 7	Variable 30	Product 19
Variable 7	Loss on Time	Variable 19	Product 8	Variable 31	Product 20
Variable 8	<i>Shelf Life</i>	Variable 20	Product 9	Variable 32	Product 21
Variable 9	Others	Variable 21	Product 10	Variable 33	Product 22
Variable 10	Financial	Variable 22	Product 11	Variable 34	Product 23
Variable 11	Loss Fill Rate	Variable 23	Product 12	Variable 35	Product 24
Variable 12	Product 1	Variable 24	Product 13	Variable 36	Product 25

Source: Prepared by the authors (2020)

3.3.1 Correlation PLS

PLS was proposed in economics and popularized in chemistry (MATEOS-APARICIO, 2016; WOLD, SJÖSTRÖM, & ERIKSSON, 2001). The advantages of PLS are: first, it can be applied when there is collinearity in the data, second, the number of variables is greater than the number of observations, and finally, when the number of

predictors is greater than one. The PLS reduces dimensions and relationship matrices X and Y (PÉREZ-ENCISO & TENENHAUS, 2003).

The partial least squares correlation (PLSC) analyzes the relationship between the X and Y matrices (KRISHNAN, WILLIAMS, MCINTOSH, & ABDI, 2011). These input matrices store the measurements collected in the logistics sector. The relationship between the X and Y column is measured by the scalar product between these two matrices. When the matrices are centered on the estimable, the scalar product provides the covariance matrix. When the two matrices are normalized (z-scores), the scalar product expresses the correlation matrix. Since correlation and covariance matrices do not depend on the order of variables, the functions of X and Y are symmetric, and the analyses are concentrated on shared information.

Formally, the relation between the X_i and Y_i matrices is stored in an matrix with cross products, denoted as R_i , which is calculated as:

$$R=[y_1, y_i, y_n]^t [x_1, x_i, x_n] \quad (2)$$

X is a matrix with the causes; Y is a matrix with the OTIF; X_i and Y_i corresponds to the i groups of weeks, where $i=1, \dots, 4$. In general, R_i is a matrix of correlations, because X and Y were centered and normalized to z-scores. The normalization for z-scores was used because the cause variables are measured in different units (for example, price in monetary unit, shelf-life in percentage).

The PLSC finds pairs of latent vectors that maximize covariance between the X and Y matrices. The latent vectors can permit the analysis of the similarity and dissimilarity between the causes and the OTIF. To classify the variables that influence the OTIF in groups, Cluster analysis was performed. The method of grouping was the one

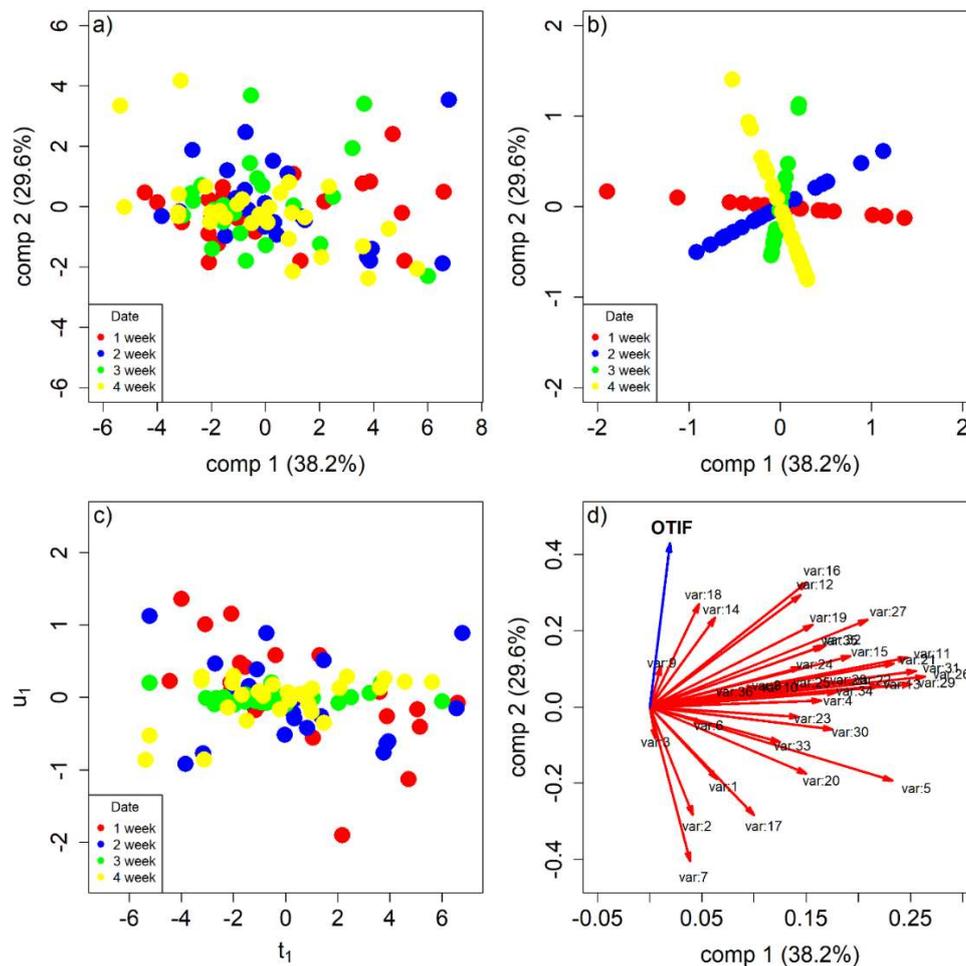
of Ward. The main function of Ward's method is to minimize the variance of Euclidean distance. In this way, the Ward approach maximizes the variability between the groups and minimizes the variability within the groups (DE OLIVEIRA et al., 2018; ESZERGÁR-KISS & CAESAR, 2017; JOE H. WARD, 1963).

3.4 Results and discussion

To verify the causes that reduce the level of service in the logistic distribution of this case study, the PLS correlation method was proposed to identify the behavior of the variables that directly interfere in the OTIF. To verify the OTIF behavior within months, the data were separated into four groups, each group representing a week of the month. The analysis in groups allows the behavior of variables, which influence the performance of the logistic distribution, to be verified.

The graphical representation of the PLS correlation is in Figure 1. The variability of the variables influencing the On-Time/In-Full per week is in Figure 1a. It is possible to see that the groups do not have well defined behavior. However, the OTIF behavior per week presents linear clustering (Figure 1b). The first week suffers with the lack of products from the stock due to the large volume of orders in the fourth and last week of the month. The second week has enough time to supply and to put in position the products in stock, being able to attend the orders in their entirety and in the time agreed by the customer. In the third week, the behavior is different because the customer holds back to the request to wait until the fourth week, where better prices and deadlines are expected. The fourth week has smaller OTIF because the logistics operation is affected by the business strategy. The non-correlation of the OTIF behavior between the weeks can be observed when the weeks form a ninety degree angle ($\cos 90^\circ = 0$). Therefore, the first week is not correlated with the third week, as the second week is also not correlated with the fourth week.

FIGURE 1 - PLS correlation between OTIF and factors that reduce efficiency. 1a) behavior of the variables in relation to OTIF over time. 1b) behavior of OTIF in relation to the week. 1c) dispersion between the X and Y blocks. 1d) variables that impact the OTIF



Source: Prepared by the authors (2020)

The dispersion between the predictor and response blocks (OTIF) is in Figure 1c, where the Y axis is OTIF and the X is the set of variables represented in time (latent variable). The first and second week do not present well defined behavior, and there is randomness in the variables. In the third and fourth week, this behavior is different, it presents linearity, and it is more evident in the third week. This happens because the sales department intensifies the capture of orders near to the final period of the report, normally in the end of the month, which marks the execution of the length of their goal. This period refers to the end of the month, this intensification occurs due to the flexibility in prices

and payment. When the goal is accomplished, the sales department receive a remuneration according to its performance.

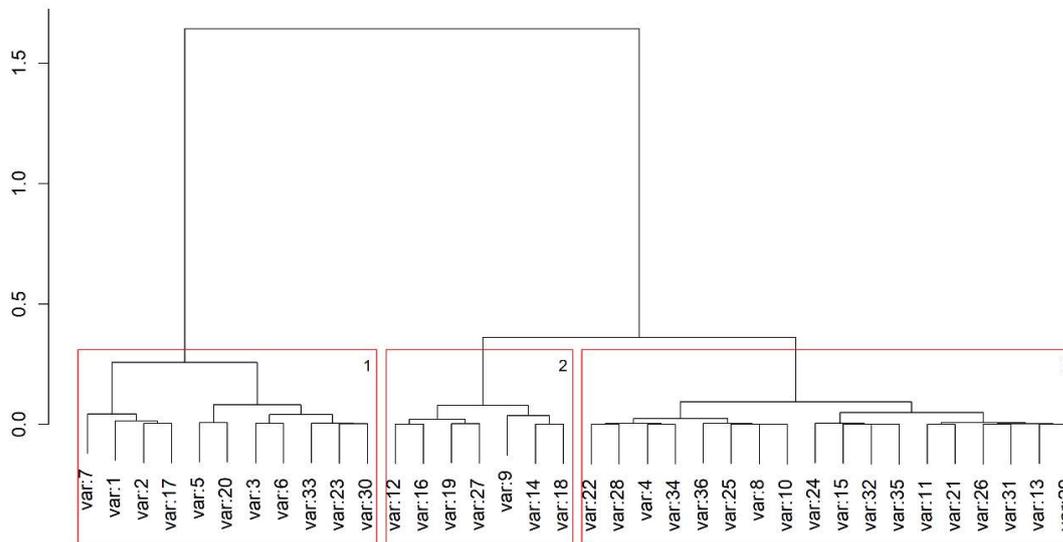
Figure 1d shows a Cartesian plane, where the variables are pointed towards their significance, the higher the arrow and the inclination in the 90° axis, the greater the significance, with a positive impact on the indicator. If the arrow is directed to the 270° axis the variable has a negative impact. Variable 7 has a negative impact, reducing OTIF. Variable 7 means the loss of On Time, that is, the failure to meet deadlines in the service. Product 7, represented by variable 18, had a positive impact, increasing OTIF. This variable has a stock placement that suits its demand, so it performs well.

The intensification that occurs at the end of the period (month) has a direct impact on logistics as mentioned by GODSELL & VAN HOEK (2009), it usually results in additional costs in the supply chain (hiring of vehicles, freight, and overtime), loss of sales revenue (discounts, rebates, combos), and reduced time for cash cycle (extended deadlines). The power of this period also affects in the beginning of the next month, because some products will be out of stock due to the high demand at the end of the previous period. The increasing number of orders concentrated in the last week ends up leading to product rupture and no sale, which causes poor performance in OTIF. At the beginning of the next period, the customers will also have a high inventory level, because they have absorbed products in the last week and they have enjoyed all the conditions proposed by the sales department.

To segregate a set of variables, Cluster analysis was proposed; it also can be identified as conglomerate analysis. Similar and homogeneous groups, following the measurement characteristic, classified the X causes. Strategically, the variables were

organized into three groups (Figure 2). Group-1 negative influences and group-2 positive influences OTIF. In Group-3, there are the variables that are in the neutral zone.

FIGURE 2 - Cluster analysis by the Ward method.



Source: Prepared by the authors (2020)

In the first cluster there are the variables that influence the OTIF in a negative way. The low performance of these variables negatively impacted OTIF (7; 1; 2; 17; 5; 20; 3; 6; 33; 23; and 30). The variable represents the causes of greater evidence of nonconformity, loss of On Time, caused by the delay of the request. The customer criteria (deadline) is not being met, and this aspect reduces customer satisfaction. Requests with the volume below the minimum quantity (represented as variable 1) contributed to the minimization of OTIF. This happens because the commercial area places orders that do not meet the minimum volume required to make the sale viable. There are also orders that are not met due to default or credit, (variable 2). All these variables could potentially raise OTIF performance.

The second grouping contains the variables that influence the OTIF positively, which are the variables: 12; 16; 19; 27; 9; 14; and 18. These variables represent the products positioned in time, in the demand of sale, and in the correct shelf life. The supply

of product in the correct time potentiates the sale, because the customer's request will always be attended.

In the third grouping there are neutral variables, which do not have a significant negative or positive impact. These variables are 22; 28; 4; 34; 36; 25; 8; 10; 24; 15; 32; 35; 11; 21; 26; 31; 13; and 29. The grouping of the variables and their representativeness are in Table 3. These causes have their share of impact, but they do not compromise the OTIF. These variables, in fact, can be used to increase the OTIF.

TABELA 3 - Grouping of variables impacting OTIF.

1 - Negatively Influence	Represents	2 - Positively Influence	Represents	3 - Neutral Zone	Represents
Variable 1	Minimum quantity	Variable 9	Others	Variable 4	Return
Variable 2	Credit	Variable 12	Product 1	Variable 8	Shelf Life
Variable 3	Price	Variable 14	Product 3	Variable 10	Financial
Variable 5	Lack of stock	Variable 16	Product 5	Variable 11	Loss Fill Rate
Variable 6	Loss in Full	Variable 18	Product 7	Variable 13	Product 2
Variable 7	Loss on Time	Variable 19	Product 8	Variable 15	Product 4
Variable 17	Product 6	Variable 27	Product 16	Variable 21	Product 10
Variable 20	Product 9			Variable 22	Product 11
Variable 23	Product 12			Variable 24	Product 13
Variable 30	Product 19			Variable 25	Product 14
Variable 33	Product 22			Variable 26	Product 15
				Variable 28	Product 17
				Variable 29	Product 18
				Variable 31	Product 20
				Variable 32	Product 21
				Variable 34	Product 23
				Variable 35	Product 24
				Variable 36	Product 25

Source: Prepared by the authors (2018)

The OTIF is an integrating indicator of the areas and can measure all the processes of the company in a viable way, from the acquisition of raw material suppliers to

production, sales, and even the distribution of final products. Consequently, it is possible to align customer service at all parts in the chain, resulting in satisfied customer and higher sales volume and ensuring the continuation of other orders. The retail sector has high OTIF values (GODSELL & VAN HOEK, 2009); however, there are problems such as poor inventory, delinquent customers, and delayed delivery. The OTIF has a lower performance in channel model wholesale, because this requires more volume, and to meet this volume it takes more time, since the production will be in the Make to Order format (the production starts right after the sales order is received and approved by the company). Moreover, some problems may occur in the performance of the sales functions and in the distribution time, which makes OTIF a driving force to process improvement initiatives throughout the organization.

SEHGAL et al. (2006) observed that it is possible to increase the KPI with Fault Tree analysis. This analysis tracks possible causes of non-compliance in OTIF. JIWA & TARIGAN (2018) evidenced the impact of organizations' commitment to the performance of the planning process and controlling product innovation using PLS. CHIANG et al. (2012) used the PLS as a methodology to test for possible effects in mediating the company's strategic flexibility. BRAUNSCHEIDEL & SURESH (2009) aimed at ensuring the continuity of operations without interruption, even in the midst of risks or uncertainties in the market. The research evidences practices in favor of increasing the performance of a company's supply chain. PLS was used as a method to assert means to guarantee better assertiveness in the supply chain process. NG, DING, & YIP (2013) proposed a business model of contracts based on productivity and results, in which PLS was used to identify relations between assets and contract performance through quantitative research. ZHANG, MU, LI, & NING (2009) proposed a

predictability study of energy transport demand in China for the years 2010, 2015, and 2020 based on the PLS regression method.

3.5 Conclusion

This study proposes a method to maximize the results of the logistics sector. Among many different methods to increase OTIF performance, PLS correlation can be applied to determine which variables should be addressed to increase the performance of the logistics sector. Using OTIF metric, a company in the refrigerator segment was used as a field in this case study. The PLS correlation allowed us to determine the variables that positively and negatively influence the results of the logistic sector. The behavior of the logistics sector is different each week. In practice, there are different strategies for each week of the month. Thus, the behavior of the variables during the weeks that comprise the month and its correlation with the OTIF became evident.

For the logistic sector, the commercial strategy contributes to the randomness of the process that influences the variables that damage the OTIF. This strategy can influence the customer service and supply of the distribution center. The sale in the correct period following the sales phase and the sales being executed according to the timeline and without oscillations effectively contribute to the customer being served in the time predefined in the order. That is because the loss of On-Time was one of the most aggravating variables in the service level, contributing negatively to the OTIF. The article by GODSELL, BIRTWISTLE, & VAN HOEK (2010) highlighted seven lessons to define the main aspects of the business alignment model and it showed how to set up their supply chain. In the seventh lesson, the OTIF measure based on a standard lead-time was exemplified from the moment the order was accepted into the planning system.

To ensure that OTIF achieves significant performance, customer orders may only be accepted if the product is in stock or if there is an arrival scheduled of this product, preventing the order from suffering back stock. This action guarantees a low risk for lack of products in stock. Moreover, it would make the company presented in this document a successful case. This adjustment ensures better KPI performance, and certainly ensures greater customer satisfaction.

4 MODELAGEM DE ÍNDICE COM PROGRAMAÇÃO GENÉTICA

4.1 Resumo

Um índice ou *Key Performance Indicator* (KPI) pode ser extraído por uma ferramenta multivariada e métodos têm sido propostos para sua modelagem. Neste artigo foi proposto selecionar e otimizar expressões para modelar índices usando algoritmos evolucionários. Quando existe um conjunto de dados de entrada e saída, a programação genética retorna uma expressão ou modelo para o conjunto de dados. Nos resultados com os operadores básicos e algumas funções trigonométricas a precisão de modelar foi R^2 e ENS maior que 0.9. Concluiu-se que a programação genética pode ser aplicada para modelar indicadores com variáveis numéricas e/ou categóricas.

Palavras-chave:

Modelagem matemática; Algoritmo genético; Algoritmo evolutivo

4.2 INTRODUÇÃO

Muitos pesquisadores têm se dedicado a criar índices que auxiliem na tomada de decisões, como a evasão fiscal (GOKALP, LEE, & PENG, 2017; W DIDIMO, L GIAMMINONNI, LIOTTA, & F MONTECCHIANI, 2018) classificação de câncer e outros. Depois que um índice é criado, uma expressão ou função é necessária para replicá-lo (ELIS REGINA DE OLIVEIRA et al., 2018). Geralmente, entre várias técnicas aplicadas são similares a uma regressão, onde a seleção de um conjunto de variáveis é executada e então seus pesos são determinados.

Assim, similarmente a programação genética (GP) tem sido aplicada em problemas computacionais e de modelagem (REZANIA & JAVADI, 2008). O artigo de

(TSANG, YUNG, & LI, 2004) evidencia o *Evolutionary Dynamic Data Investment Evaluator* (EDDIE) ferramenta de apoio à decisão baseada em programação genética (GP) para previsão financeira. Este artigo descreve uma arquitetura que incorpora o EDDIE para aprender e monitorar o mercado de ações. O EDDIE serve para aprimorar a produtividade da especialista, com objetivo de buscar no espaço de árvores, com foco de melhorar as chances em favor do usuário.

Inicialmente a GP foi introduzida na década de 1990 por KOZA (J R KOZA, 1995) e pertence à classe de algoritmos evolutivos (EA), que são técnicas baseadas em mecanismos evolutivos encontrados na natureza (como seleção, cruzamento e mutação) para encontrar soluções exatas ou aproximadas para problemas de otimização e busca (MOLINA, LATORRE, & HERRERA, 2018).

Na GP, o objetivo é pesquisar e / ou otimizar uma função (programa) que se ajusta a um conjunto de dados. GP é similarmente como Algoritmo Genético (GA), onde há um grupo de indivíduos competindo entre si (LIU, TONG, XIE, & YEE NG, 2015; REZANIA & JAVADI, 2008). A programação genética é uma extensão do algoritmo genético no qual as estruturas na população não são cadeias de caracteres de tamanho fixo que codificam soluções candidatas a um problema, mas programas que, quando executados, são as soluções candidatas para o problema (KOZA & POLI, 1983). Os programas são expressos em programação genética como árvores de sintaxe e não como linhas de código. Enquanto o GA procura otimizar uma sequência ou valores para melhorar um determinado modelo, o GP procura combinar um conjunto de modelos possíveis para formar o melhor modelo possível. A principal diferença está na estrutura do indivíduo: GA tem indivíduos em estrutura de cordas, enquanto que em GP os indivíduos são árvores.

Na literatura, vários métodos têm sido propostos para modelagem (E.R. DE OLIVEIRA et al., 2018). No entanto, neste artigo, foi proposto selecionar e otimizar expressões para modelar índices.

4.3 Programação genética

A representação de uma árvore no GP é composta por nós terminais e não-terminais (ver Figura 03 e 04). Um nó não terminal (funcional) é um nó interno com nós filhos. E um nó terminal é um nó folha sem nós filhos. Geralmente, os nós terminais são elementos primitivos (incluindo constantes ou variáveis) e nós não terminais são operadores e/ou funções para combinar os elementos primitivos.

Dado um sistema físico que tem uma variável dependente Y e algumas variáveis independentes $X = [X_1, \dots, X_m]$, onde x_j é a j th variável do conjunto terminal, então o problema de minimizar os parâmetros funcionais para uma solução pode ser matematicamente formulado (MOLINA et al., 2018; SCRUCCA, 2015) como:

$$\Theta^* \equiv \underset{\theta \in \Theta}{\text{arg min}} f(\mathbf{X}, \theta) = \{\theta^* \in \Theta : f(\mathbf{X}, \theta^*) \leq f(\mathbf{X}, \theta), \quad \forall \theta \in \Theta\}$$

Em que Θ é um conjunto de parâmetros de solução candidatos no espaço de pesquisa viável, isto é, o domínio dos parâmetros $\theta = (\theta_1, \dots, \theta_p)$ onde cada θ_i é um parâmetro de conjunto não terminal Θ^* é o conjunto de soluções $f(\mathbf{X}, \theta)$ é uma função em um espaço m -dimensional, onde m é o número de variáveis de entrada. A Programação Genética gera uma expressão $f(\mathbf{X}, \theta)$ codificada em uma estrutura de árvore de tamanho variável e procura por uma expressão global que melhor se ajuste ao conjunto de dados $f(\mathbf{X}, \theta^*)$.

O processo de evolução começa com uma população de n árvores geradas aleatoriamente por Full, Grow e *ramped-half-and-half* e outros métodos

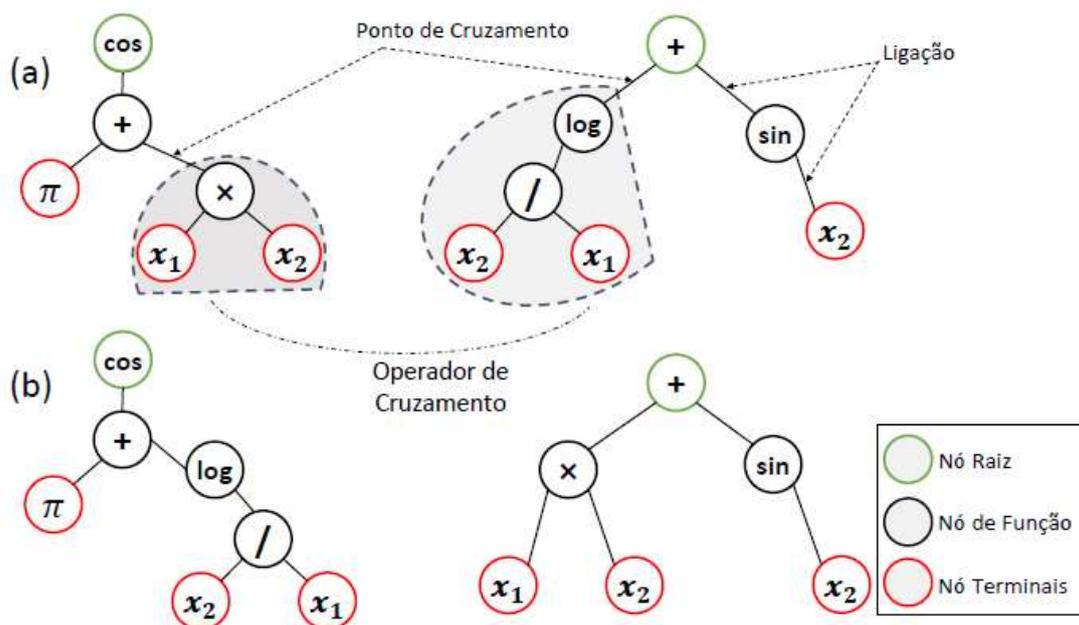
(CHELLAPILLA, 1997); (J R KOZA, 1995); (LUKE & PANAIT, 2001). Cada árvore é treinada e testada usando treinamento de conjunto de dados (80%) e teste (20%). A aptidão de todos os indivíduos é medida para minimizar o erro médio quadrático (função objetivo), dado por:

$$RMSE = \left(\frac{1}{N} \sum_{a=1}^N \delta_a^2 \right)^{1/2} \quad (3)$$

onde N é o número de instâncias e δ é erro absoluto.

Algumas árvores são selecionadas (torneios) e os valores mais altos de aptidão sobrevivem e produzem descendentes. Assim, operadores de crossover, mutação e cópia são aplicados aos indivíduos sobreviventes (ver Figura 03 e 04). Operador de crossover cria a chance de se juntar a dois indivíduos diferentes, o que é feito mudando os galhos de duas árvores (indivíduos). O operador de mutação evita os mínimos locais no processo evolutivo por aleatoriedade. Esse processo ocorre alterando aleatoriamente alguns nós terminais ou não-terminais da árvore. Junto com o cruzamento e a mutação, o operador de cópia envia algumas árvores para se juntar aos indivíduos sobreviventes. O processo com os operadores é repetido por λ vezes, então $\mu = n$ indivíduos são selecionados, completando uma nova geração. As iterações (gerações) devem continuar até que um critério de terminação seja satisfeito.

FIGURA 3 - Operador de crossover. (a) árvore de pais. (b) árvores de descendentes.



Fonte: Elaborado pelos autores (2020)

Antes de iniciar o GP, os dados foram pré-processados:

1. A variável de resposta foi escalonada $[0, 1]$ para retornar ao índice (equação 4).

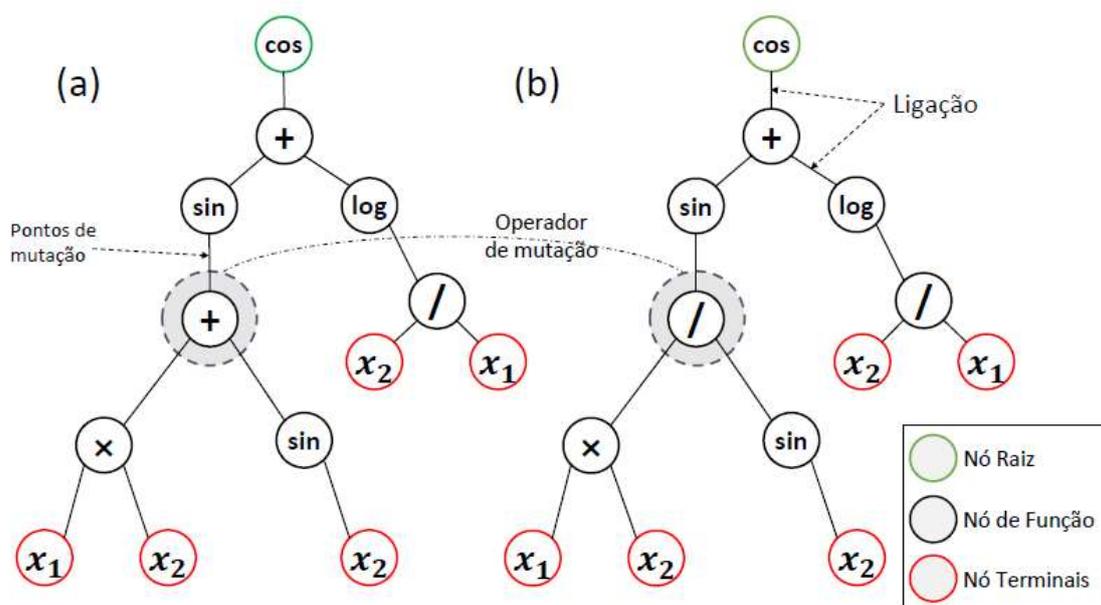
$$\hat{y}_i = \frac{y_i - \min_y}{\max_y - \min_y} \quad (4)$$

O y é o valor original, \hat{y} é o valor escalado, \min_y é o limite inferior do valor do recurso e \max_y é o limite superior do valor do recurso;

2. As variáveis numéricas do preditor foram escalonadas para a variância unitária.

Assim, as variáveis têm o mesmo peso (equação 5).

FIGURA 4 - Operador de mutação. (a) árvores pais. (b) árvores de descendentes.



Fonte: Elaborado pelos autores (2020)

$$\hat{x}_{ij} = \frac{x_{ij} - \bar{x}_j}{S_j} \quad (5)$$

x_{ij} é o valor original, \hat{x}_{ij} é um valor escalado, \bar{x}_j é um vetor com as médias das variáveis e S_j é um vetor com o desvio padrão das variáveis;

3. As variáveis predictoras categóricas (fatores) foram escalonadas para zero e uma equação 6.

$$x'_{i,l} = \begin{cases} 1, & \text{se o caso } X_{ij} \text{ for verdadeiro} \\ 0, & \text{se o caso } X_{ij} \text{ for falso} \end{cases}, \forall l \quad (6)$$

x_{il} é o valor original, $x'_{i,l}$ é o valor escalado para todos os níveis l ($l = 1, \dots, \text{níveis}$).

Após o pré-processamento, para garantir que o GP não retorne valores maiores que o índice (equação 4), a função *Sigmoid* foi aplicada na saída do modelo (equação 7).

$$\hat{y} = \frac{1}{1 + e^{-1/2f(x,\theta)}} \quad (7)$$

Durante o processo de evolução, é necessário avaliar o desempenho dos modelos (indivíduos) e manter a capacidade de generalização. Assim, para evitar *overfitting* foi aplicado a validação cruzada (CV). O CV divide parte do conjunto de treinamento em subconjuntos de validação. Em CV usamos 4 dobras para evitar pequenos grupos na validação. Assim, o CV é aplicado em quatro iterações. Durante cada iteração, 3 dobras representam o conjunto de treinamento e 1-dobra o conjunto de validação. Conjunto de dados foram divididos entre os grupos de forma aleatória e uniforme.

Os programas (árvores) foram avaliados separadamente para cada conjunto de dados (Tabela 4). Foi utilizada a regressão linear múltipla para comparar os resultados com a Programação Genética. Todos os procedimentos para modelagem do índice foram implementados em Python com o pacote DEAP (<https://github.com/DEAP>).

A precisão do modelo foi medida pelo coeficiente de determinação (R^2) e pelo coeficiente de eficiência de Nash-Sutcliffe (NSE). R^2 indica quanto o modelo pode explicar a amostra e assume valores entre zero e um, se o próximo é que o modelo é *good-of-fit* (MYERS, MYERS, WALPOLE, & KEYING, 2015). O NSE avalia a magnitude da variância do resíduo com a variância da amostra (NASH & SUTCLIFFE, 1970). NSE assume valores entre $-\infty$ até um, se perto de um modelo é *good-of-fit*, quando menor que zero a média é melhor que o modelo proposto (MORIASI et al., 2007).

TABELA 4 - Conjuntos de dados do repositório

N ^a	Nomes	Não de instâncias	Características nominais	N ^a de níveis	N ^a de fatores	Total fatores
1	Credit card	400	4	5	7	11
2	SML20102	4137	0	0	24	24

1 James et al. (2013), 2 UCL *repositor* (James, Witten, Hastie, & Tibshirani, 2013)

4.4 Resultado e implementação de programação genética

Considerando o problema de minimização (Equação 4) foram determinados parâmetros para modelar um problema de regressão simbólica. Os parâmetros de entrada do GP estão na Tabela 5. O foco do GP é desenvolver um programa ou função que minimize o RMSE. A função possui algumas variáveis de entrada selecionadas pelo GP do terminal X, que retorna um índice. Operadores matemáticos básicos e algumas funções trigonométricas foram incluídos no conjunto funcional.

TABELA 2 - Parâmetros de entrada da programação genética

Parâmetros	Valores
Objetivo	Evoluir uma função
Função objetivo	Minimize RMSE *
Conjunto de terminais	Conjunto de variáveis preditoras
Conjunto funcional	+, -, *, %, cos, sin
Constante transitória (efêmera)	Uniforme (-1 a 1)
Tamanho da população ($\mu = n$)	600
Filhos produzidos em cada geração (λ)	900
Taxa de cruzamento	50%
Taxa de mutação	20%
Critério de seleção	Torneio, tamanho 3
Crítérios de terminação	últimas 100 gerações sem melhorar
Max. de gerações	600
Max. profundidade da árvore após o cruzamento	50
Max. profundidade da árvore após mutação	50
Método de inicialização	Crescer

*Erro quadrático médio da raiz, % é a divisão protegida por zero.

Durante a inicialização, as árvores foram criadas aleatoriamente pelo método Grow (profundidade 1-10, 20%; 11-20, 20%; 21-30, 20%; 31-40, 20% e 41-50, 20%). O operador de crossover utilizado foi o enviesado da folha, em que os pontos de cruzamento são selecionados 90% para nós não terminais e 10% para nós terminais. O operador de mutação usado foi: (1) substituir um nó por um nó; (2) substituir um nó por um ramo; e (3) substituir um nó efêmero. A probabilidade das três mutações é a mesma (um terço).

TABELA 3 - Resumo dos resultados da programação genética no conjunto de dados de validação cruzada (RMSE).

N ^a	Nomes	Conjunto de treinamento	Conjunto de validação	Recursos selecionados	Profundidade da árvore	Comprimento do conjunto de parâmetros
1	Credit card	0.00113	0.00176	12	23	184
2	SML2010	0.07104	0.0754	6	11	20

Fonte: Elaborado pelos autores (2020)

Os critérios de terminação são 600 gerações ou que o valor de aptidão não melhora durante as últimas 100 gerações. A melhor árvore (modelo) é obtida quando os critérios de terminação são satisfeitos. Para o cartão de crédito do conjunto de dados, o RMSE do conjunto de treinamento, validação e teste foi de 0,00113, 0,00176 e 0,00115, respectivamente. A Tabela 6 mostra os resultados resumidos para os modelos e a Tabela 7 mostra os resultados resumidos para precisão.

TABELA 4 - Resumo da precisão da programação genética para testar conjuntos de dados

N ^a	Nomes	RMSE	R ² *	NSE
1	Credit card	0.00115	0.99999	0.99999
2	SML2010	0.07574	0.99996	0.99997

*Coeficiente de determinação

O resultado da regressão linear múltipla para o primeiro dataset Credit card, no treino, o R² foi de 0,91 e no teste foi de 0,87. Para o segundo dataset SML20102, no treino, o R² foi de 0,98 e no teste 0,99. No caso do segundo *dataset* que não possuíam as variáveis nominais o resultado foi melhor. Entretanto, a Programação Genética apresentou melhor desempenho que a regressão linear múltipla em ambos os casos. Porém, foi observado que a programação genética apresentou excelente performance, mesmo com o *dataset* que possui variáveis nominais.

TABELA 5 - Resumo dos resultados da regressão linear múltipla

Data Set	R ²	RMSE	NSE
Treino			
Credit card	0,9198	10,2599	0,9197
SML 20102	0,9878	0,3154	0,9877
Teste			
Credit card	0,8425	12,8122	0,8253
SML 20102	0,9929	0,2151	0,9928

Fonte: Elaborado pelos autores (2020)

4.4.1 Modelagem do índice

Para otimizar expressão de algoritmo evolucionários foi utilizado os dados do armazém e a eficiência do processo.

A programação genética retorna um modelo para o conjunto de dados, sendo a precisão deste modelo explicitada na Tabela 9.

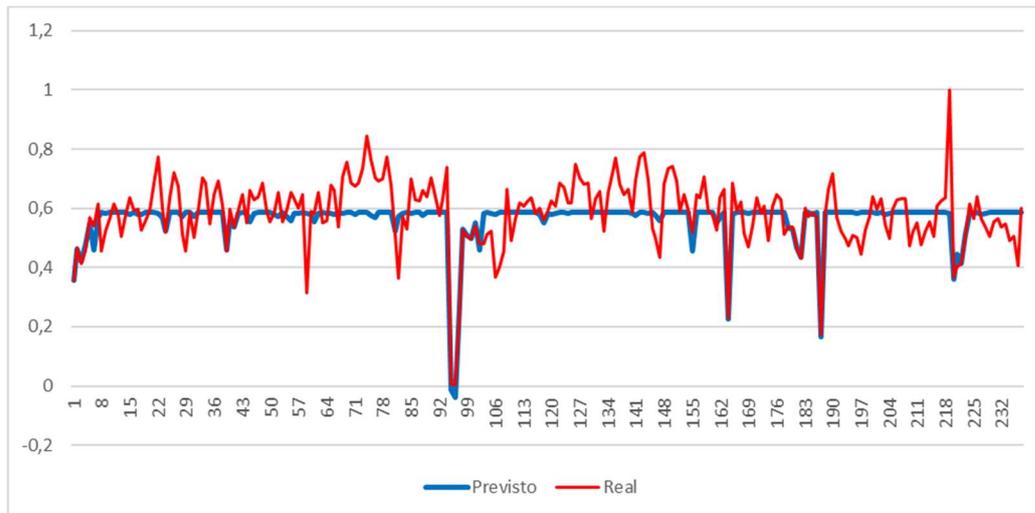
Para analisar a eficácia do modelo é necessário que a modelagem possua um bom ajuste. Os valores de PBIAS, MAE, RMSE e MAPE foi utilizado para calibração e validação.

O gráfico da modelagem do índice se encontra na Figura 5. O índice oscila em torno da média que é 0,582. O modelo de programação genética conseguiu representar os três principais picos. Por este motivo, os valores de MAE e RMSE ficaram baixos.

TABELA 6 - Acurácia do modelo

PBIAS (%)	MAE	RMSE	MAPE (%)
10,8936	0,0634	0,0854	16,1228

Fonte: Elaborado pelos autores (2020)

FIGURA 5 - Gráfico da modelagem de índice

Fonte: Elaborado pelos autores (2020)

4.5 Conclusões

Durante a análise de dados multivariados, alguns pesquisadores criam um índice. Assim, é necessário modelar o índice para que ele possa ser replicado e aplicado a outros conjuntos de dados. Com dados contínuos e / ou discretos a Programação Genética apresentou excelente desempenho ao modelar índice. Assim, GP pode ser aplicado para modelagem de expressão para retornar um índice.

O desempenho do método GP ocorre devido às transformações matemáticas que ocorrem nos dados durante a modelagem. As transformações utilizadas foram as funções trigonométricas seno e cosseno. Sugerimos para trabalhos futuros adicionar outras funções matemáticas e redes neurais para aprender a constante transitória.

5 PREVISÃO DE DEMANDA NA CADEIA DE SUPRIMENTOS: APLICAÇÃO DO MODELO DE SÉRIES TEMPORAIS

5.1 Resumo

Identificar e compreender variáveis que influenciam na qualidade do atendimento logístico, é fundamental para melhora do nível de serviço e crucial para a companhia, sendo este aspecto fator essencial e estratégico para a rentabilidade da empresa e garantia de continuidade de vendas e pedidos. Neste estudo, foi utilizado algumas ferramentas como Séries Temporais e o modelo ARIMA para modelagem de previsibilidade de demanda, com objetivo de prever o volume a ser expedido e garantir o número suficiente de produtos em estoque, de forma a atender todos os pedidos sem rupturas e ter a quantidade de veículos suficiente para distribuição, sem que haja atrasos. O *software* R versão 3.4.1 (R DEVELOPMENT CORE TEAM, 2017) e da sua IDE (*Integrated Development Environment*) R STUDIO, ambos gratuitos e de código aberto, foram utilizados neste trabalho. Os resultados da aplicação dessas ferramentas evidenciam variáveis que impactam na qualidade do nível de serviço logístico, e auxiliam os gestores em tomada de decisão.

Palavras-chave: Séries temporais; Previsão de demanda; ARIMA;

5.2 INTRODUÇÃO

Fazer mais com menos, maximizar lucro e reduzir o custo, otimização de processos, buscar maior eficiência, prospecção, automação, todas essas ações e demandas tem apenas um só objetivo, a rentabilidade. Todas as empresas, companhias ou órgãos

cujo objetivo é gerar lucro, focar em diversas atividades do seu processo visando enxugar e reduzir (CAMPOS & MENEZES, 2019). Esses fatores se tornam cada dia mais desafiador para as organizações desenvolverem novos métodos, ferramentas e continuarem vivas no mercado. A estes requisitos o *Just in Time* (JIT) apresenta o modelo a cumprir tais condições. O modelo JIT tem sido bastante relevante principalmente no aspecto financeiro. Estudos forneceram evidências de rentabilidade (BALAKRISHNAN, LINSMEIER, & VENKATACHALAM, 1996; BOYD, 2001; KINNEY & WEMPE, 2002; MISTRY, 2005). As pesquisas de (CALLEN, MOREL, & FADER, 2003) indicam que a fabricação pelo método de JIT é mais lucrativa do que a fabricação convencional.

Ainda (CALLEN et al., 2003) existem explicações potenciais para a rentabilidade no modelo JIT em comparação ao modelo de produção convencional, sendo uma explicação postula seria o mínimo de estoque em processo e o mínimo de estoque em produto acabado, a fabricação pelo método JIT está sujeita a maiores riscos operacionais do que a manufatura convencional, por justamente produzir de acordo com a demanda, sendo este fator primordial para a rentabilidade e sucesso do modelo.

O JIT tornou-se uma das mais conhecidas e debatidas, metodologias de melhoria de processos, que ainda fomenta diversas pesquisas (HASLE, BOJESSEN, JENSEN, & BRAMMING, 2012; NAGA VAMSI KRISHNA JASTI, 2014). Para o sucesso do método JIT é necessário ter o conhecimento prévio da demanda, para então cumprir com a filosofia *Lean*, caso contrário o cliente aguardará longos prazos para entrega do produto final, o que provocará impactos no nível de serviço e insatisfações com o fornecedor. Se os clientes não encontrarem os produtos à procura, eles poderão optar por outro competidor ou comprar itens substitutos (MCGOLDRICK & ANDRE, 1997).

O planejamento de produção, estoque, logística, marketing e finanças nas empresas da cadeia de fornecimento dependem de previsões (FILDES, GOODWIN, & ONKAL, 2016). A previsibilidade da demanda garante o gerenciamento eficiente da capacidade a longo prazo sendo vital para qualquer empresa de manufatura. Isso tem implicações no desempenho competitivo em termos de custo, confiabilidade, flexibilidade e velocidade de entrega. A estratégia de manufatura e o planejamento de vendas e operações fornecem perspectivas sobre o gerenciamento da capacidade, levantando e tratando diferentes questões, como desempenho competitivo em termos de custo, velocidade de entrega e confiabilidade (OLHAGER, RUDBERG, & WIKNER, 2001).

A previsão de demanda é uma das principais questões das cadeias de suprimentos. Tendo conhecimento da previsão, pode-se otimizar estoques, reduzir custos, aumentar as vendas, o lucro e a fidelidade do cliente. Para isso, os dados históricos podem ser analisados para melhorar a previsão de demanda usando vários métodos, como técnicas de aprendizado de máquina e análise de séries temporais (KILIMCI et al., 2019).

Faz-se necessário e de vital importância a previsão para o alcance do objetivo da companhia. Essas abordagens de previsão preveem demanda futura com base em dados históricos de séries temporais, sequência de pontos de dados medidos em intervalos sucessivos em tempo. Métodos de séries temporais incluem o método Naïve, método médio, suavização exponencial, tendência linear de Holt, método de tendência exponencial, métodos de tendência amortecida, Método sazonal de Holt-Winters, médias móveis, ARMA, (Média móvel autorregressiva) e modelos ARIMA (Média móvel autorregressiva integrada) (ROB J HYNDMAN & ATHANASOPOULOS, 2013). ARMA e ARIMA (também chamado de método Box-Jenkins nomeado após os estatísticos George Box e Gwilym Jenkins) são os métodos mais comuns que são

aplicados para encontrar o melhor ajuste de um modelo para os valores históricos de uma série temporal (GUJARATI, 2003).

No trabalho (RODRIGUES et al., 2018) utiliza séries temporais com o modelo ARIMA para previsão de geração de resíduos sólidos para o aterro sanitário de Aparecida de Goiânia no estado de Goiás. A estimativa serve como base para adequação de infraestrutura, desenvolvimento sustentável, otimização de recursos utilizados no tratamento dos resíduos sólidos urbanos. A pesquisa de (SVETUNKOV & BOYLAN, 2019a) evidencia o uso do modelo ARIMA, no qual pode ser usado eficientemente no contexto da cadeia de suprimentos. O ARIMA de espaço de estados simplifica algumas das etapas da previsão e pode ser usado até mesmo em dados com um histórico curto. No trabalho de (XU, CHAN, & ZHANG, 2019) usa-se séries temporais para prever indicadores estatísticos no setor de aviação que podem ser usados para posterior gerenciamento de capacidade e planejamento. A pesquisa de (DELLINO, LAUDADIO, MARI, MASTRONARDI, & MELONI, 2018) evidencia o modelo ARIMA para previsão de vendas com planejamento de pedidos, que inclui um sistema de seleção de modelo individual. Com esses modelos a logística pode se preparar de forma prévia quanto a sua real necessidade para atender aos pedidos no tempo certo e na sua totalidade. O objetivo deste trabalho é fomentar informações que contribuam para a gestão da cadeia de suprimentos. Será utilizado o modelo ARIMA para prever o volume de vendas a ser distribuído nos próximos meses.

5.3 Materiais e métodos

A previsão da cadeia de suprimentos como um descritor tem como objetivo capturar todos os aspectos da previsão para apoiar as operações, desde o varejista ou prestador de serviços através do sistema de distribuição aos fabricantes e fornecedores de terceiro nível (BOONE, BOYLAN, FILDES, GANESHAN, & SANDERS, 2019).

No livro (RICHARD G. BROWN, 1959) conectando previsão com gerenciamento de estoques, figurou proeminentemente na literatura antiga. As inovações de Brown levaram ao desenvolvimento e implementação em software nos anos 60. No entanto, o tema tem sido largamente negligenciado desde as primeiras contribuições, com relativamente poucos artigos sendo identificado como influente (FILDES, 2006). Em vez disso, o foco tem sido ao nível da unidade de manutenção de estoque individual, *Stock Keeping Unit* (SKU), com menos atenção ao nível agregado (BOONE, BOYLAN, et al., 2019).

Alguns métodos estatísticos como o Box-Jenkins e, mais recentemente, a formulação do espaço estatal de (HYNDMAN, KOEHLER, ORD, SNYDER, & HAND, 2008) provaram influente nos métodos de previsão implementados na oferta empresas da cadeia, através de software comercial para Métodos ARIMA (*autoregressive integrated moving average*) e através de software de código aberto no caso do espaço de estado (BOONE, BOYLAN, et al., 2019). No artigo de (BOONE, GANESHAN, JAIN, & SANDERS, 2019) é relatado que nos últimos dez anos, avanços em tecnologia e coleta de dados em sistemas resultaram na geração de grandes volumes de dados em uma ampla variedade de tópicos e em uma grande velocidade. Em seu trabalho (BOONE, GANESHAN, et al., 2019) analisa o impacto que esta explosão de dados está tendo sobre previsão do produto e como ele está melhorando, o método utilizado são séries temporais, e baseado nisso explorou-se como esses dados podem ser usados para obter informações sobre o comportamento do consumidor e o impacto de tais dados na previsão organizacional.

As tecnologias de *big data* permitiram que o escalão da cadeia de fornecimento deve estar "conectada" e "comunicar" uns com os outros, criando a cadeia de fornecimento "inteligente" (SANDERS, 2016). Ainda Sanders, os aplicativos de análise

que podem fornecer uma vantagem competitiva e aparecem ao longo de todo o espectro de decisões e auxílios da cadeia de suprimentos, desde o marketing direcionado com base na localização até a otimização dos estoques da cadeia de suprimentos para permitir a avaliação de risco do fornecedor. Embora muitas empresas a tenham usado para extrair novos insights e criar novas formas de valor, outras empresas ainda precisam alavancar grandes volumes de dados para transformar suas operações na cadeia de suprimentos. Se faz necessário para a logística o uso de aplicativos analíticos para roteamento de programação de demanda, podendo assim prever a real necessidade de mão de obra extra e até mesmo contratar frota extra.

A previsão é de fundamental importância para auxiliar em informações e ter maior assertividade em situações complexas de decisões. Para (BRESSAN, 2002) as previsões econômicas buscam reduzir os cenários de incertezas, principalmente no setor agropecuário que está constantemente sujeito a distúrbios devido a fatores de natureza sazonal, cíclica e irregular e que influenciam o preço final dos produtos. As projeções de preços representam instrumentos de grande relevância para a tomada de decisões, pois podem reduzir o risco na tomada de decisão dos agentes econômicos. Como base para planejar e executar as atividades da cadeia de suprimentos, as previsões servirão para se estruturar a operação e atender o cliente conforme suas expectativas, quantidade correta e prazos solicitados.

Um dos maiores interesses no estudo de séries temporais é conseguir fazer previsões das observações no tempo. Previsão de séries temporais é uma alternativa para redução de incertezas no processo de tomada de decisão. As estimativas são feitas através de processos estocásticos, no qual existe padrões sistemático ou persistente no comportamento da variável, ou seja, processos controlados por leis probabilísticas. Modelos de séries univariadas de tempo representam uma classe específica, onde se

objetiva modelar e prever variáveis utilizando apenas informações contidas em seus próprios valores passados e atuais e um termo de erro. Considerando-se atóricos, o que implica que a sua construção e utilização não é baseada em qualquer teoria subjacente ao comportamento de outra variável, ou seja, não existe qualquer independência e causalidade com outras séries temporais (BROOKS, 2002).

A notação usada para uma série temporal é x_1, \dots, x_{t-1}, x_t que indica uma série de tamanho t . Uma grande quantidade de fenômenos de natureza física, biológica, econômica, pode ser enquadrada nesta categoria. A maneira tradicional de analisar uma série temporal é através da sua decomposição em componentes de tendência, ciclo e sazonalidade (MORETTIN & TOLOI, 1987).

O método consiste em ajustar modelos auto-regressivos integrados de médias móveis ARIMA (p,d,q) a um conjunto de dados, onde p , d e q são parâmetros de números inteiros não negativos, p é a ordem do modelo auto-regressivo, d é o grau de diferenciação (o número de vezes em que os dados tiveram valores passados subtraídos) e q é a ordem do modelo de média móvel. A etapa auto-regressiva (AR) do modelo ARIMA indica que a variável evoluinte de interesse é regressada em seus próprios valores defasados, isto é, anteriores. A parte de média móvel (MA) indica que o erro de regressão é combinação linear dos termos de erro, cujos valores ocorreram contemporaneamente e em vários momentos no passado. A etapa integrada (I) indica que os valores de dados foram substituídos com a diferença entre seus valores e os valores anteriores e este processo diferenciador pode ter sido realizado mais de uma vez. O propósito da cada uma destas características é fazer o modelo se ajustar aos dados da melhor forma possível (WILLIAM H. GREENE, 1997).

Para a construção do modelo seguimos um algoritmo no qual a escolha da estrutura do modelo é baseada nos próprios dados. A partir do momento que conseguimos identificar e estimar um modelo ARIMA adequado às observações, então podemos utilizar a modelagem ARIMA para prever os valores das observações h passos a frente. O desígnio principal é prever um valor X_{t+h} , $h \geq 1$, tendo todas as observações até o instante t , isto é, ..., X_{t-2} , X_{t-1} , X_t , sendo t as origem das previsões. Denotamos por $\hat{X}_t(h)$ a previsão de h passos a partir da origem t . Modelo ARIMA(p,d,q) estacionário, inversível e com os parâmetros conhecidos, então a equação de previsão, considerada como uma função de h , com origem t fixa é dada por:

$$\hat{X}_t(h) = \sum_{i=1}^{p+q} \phi_i \hat{X}_t(h-1), \quad h > q \quad (8)$$

Temos que para $h > q - p - d$, a função $\hat{X}_t(h)$ consistirá de uma mistura de polinômios, exponenciais e senóides amortecida, com sua forma exata determinada pelas raízes do operador de retardo $\phi(B) = 0$. A solução geral terá a forma:

$$\hat{X}_t(h) = c_1^{(t)} f_1(h) + c_2^{(t)} f_2(h) + \dots + c_{p+q}^{(t)} f_{p+q}(h), \quad h > q - p - d \quad (9)$$

Onde, $f_i(h)$, $h = 1, \dots, p + q$, são funções de h e c_1^t, \dots, c_{p+q}^t são coeficientes adaptados que dependem da origem da previsão e são determinados por $\hat{X}_t(1), \hat{X}_t(2), \dots, \hat{X}_t(p + d)$. A atualização destas previsões considerando X_{t+h+1} feitas a partir de duas ascendências:

- (i) $t + 1: \hat{X}_{t+1}(h) = \psi_h a_{t+1} + \psi_{h+1} a_t + \psi_{h+2} a_{t-1} + \dots$
- (ii) $t: \hat{X}_t(h + 1) = \psi_{h+1} a_t + \psi_{h+2} a_{t-1} + \dots$

Subtraindo (ii) por (i) temos:

$$\hat{X}_{t+1}(h) = \hat{X}_t(h+1) + \psi_h a_{t+1} \quad (10)$$

Contudo, a previsão de X_{t+h+1} , feita no instante t , pode ser atualizada quando um novo dado X_{t+1} for observado. Assim sendo, faremos a previsão de X_{t+h+1} , na origem $t+1$, adicionando à previsão $\hat{X}(h+1)$ um múltiplo do erro de previsão $a_{t+1} = X_{t+1} - \hat{X}_t(1)$ (MORETTIN & TOLOI, 2006).

5.3.1 Calibração e validação do modelo

Será modelado a previsão do volume de vendas em quilograma (kg) das próximas 8 semanas que serão expedidos pela área de logística. O método a ser utilizado será o ARIMA, os dados históricos irão se iniciar no período de maio de 2017 a junho de 2019, agrupado por volume semana, totalizando 114 dados. Existem determinados critérios que podem ser aplicados para avaliar a qualidade de ajuste do modelo (MORIASI et al., 2007; KRAUSE et al., 2005). Neste trabalho, a desempenho do modelo de previsão proposto foi avaliado aplicando as seguintes métricas: média absoluta percentual dos erros (MAPE); percentual de viés (PBIAS); erro absoluto médio (MAE); raiz do erro quadrático médio (RMSE)

O MAPE mensura a porcentagem de erro em média, podendo assumir valores maiores que zero. Apesar do MAPE ser aplicado com frequência, em circunstâncias onde exista discrepância nos erros, pode ocasionar um resultado acima do valor real (GOODWIN; LAWTON, 1999). Também com unidade em porcentagem o PBIAS mede quanto dados estimados estão acima ou abaixo da amostra em média. PBIAS com valor positivo indica que o modelo possui viés de subestimar os dados e quando o valor é negativo, indica que o modelo possui viés de superestimar.

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (11)$$

$$PBIAS = 100 \left[\frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{\sum_{i=1}^n y_i} \right] \quad (12)$$

O MAE mensura a magnitude média dos erros sem considerar sua direção, onde todos os erros têm o mesmo peso. De forma análoga, o RMSE também mede a magnitude média dos erros, porém os erros são elevados ao quadrado. Isso faz com que modelos que possuam um elemento com erro alto seja penalizado. Ambos, possuem valores que podem variar de zero a ∞ , e casos onde o erro é zero, o ajuste é ideal.

Entretanto, o MAE é a medida mais simples aplicada, além de ser adicionada a outras métricas (WILLMOTT; MATSUURA, 2005). De acordo com Chai e Draxler (2014), a presença de *outliers* é melhor descrita pelo RMSE, porém o MAE é mais apropriado para dados com distribuição normal.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (13)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|^2} \quad (14)$$

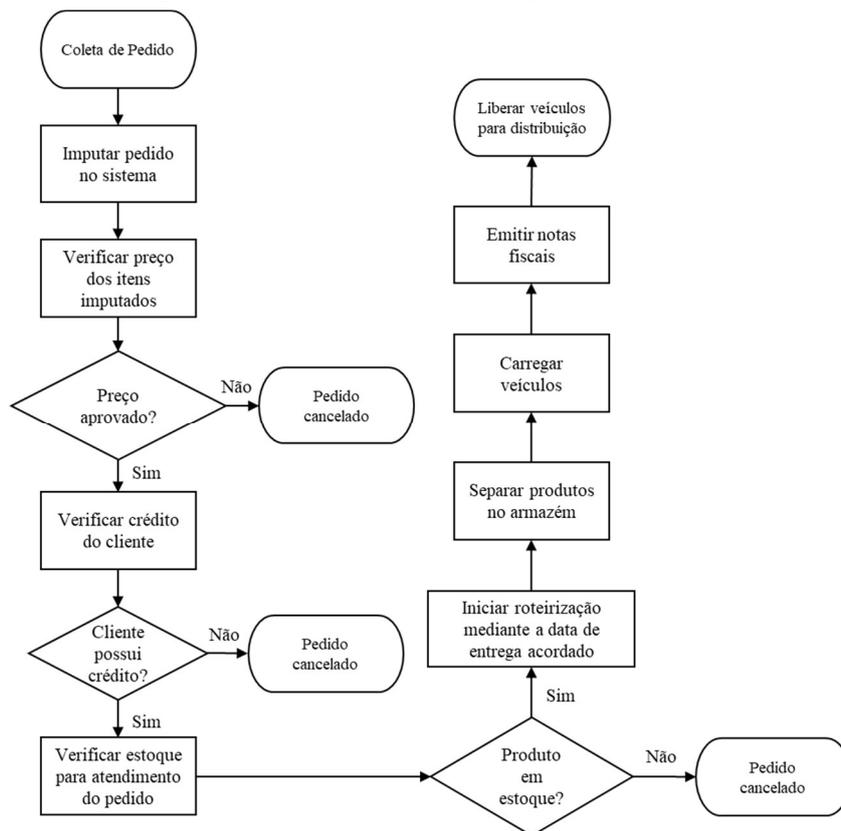
5.3.2 Caso de teste

O ARIMA é raramente usado na cadeia de suprimentos, porém ele pode ser eficientemente utilizado no contexto de *supply chain* (SVETUNKOV & BOYLAN, 2019b). Usamos este modelo para aplicação da previsão de volume de vendas em uma indústria de produtos alimentícios. A principal justificativa para adoção do sistema é

prever o volume de vendas, para garantir melhor nível de serviço com o menor custo possível na expedição e distribuição dos produtos aos clientes, de tal forma que a logística se prepare para a capacidade do volume a ser movimentado, contratando o número certo de veículos, evitando ociosidade de frota ou até mesmo custeando por estes veículos parados, ou, deixando de atender clientes no dia e hora alinhado por falta de veículos, negligenciado o critério tomador de pedido. Outro fator preponderante seria a mão-de-obra para atender a capacidade, evitando pessoas ociosas ou ausência de gente para garantir a expedição e uma operação eficiente.

No contexto específico desta pesquisa a indústria estudada se refere a uma multinacional, porém, os dados coletados são somente da unidade que está localizada no estado de Goiás, região Centro-Oeste do Brasil, considerando o período de maio de 2017 a junho de 2019. Os dados foram coletados através dos controles existentes do supervisor de logística, mediante ao fluxo demonstrado do processo (Figura 6) “*coleta do pedido*”, “*aprovação dos pedidos*”, “*roteirização de pedido*”, “*separação*”, “*carregamento*” e “*distribuição*”.

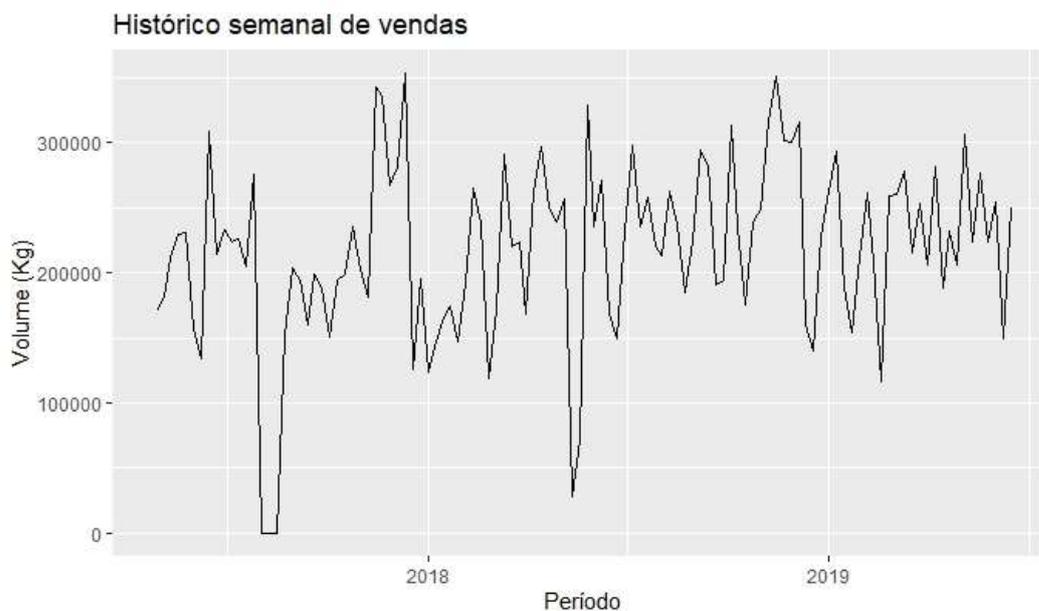
FIGURA 6 - Fluxo do processo de atendimento de pedido



Fonte: Elaborado pelos autores (2020)

Os dados coletados por semana estão no Apêndice A. As datas demonstradas apontam para o domingo, período t , em que representa o volume realizado daquela semana, ponto X_t . O volume especificado é representado pela unidade em quilograma (Kg). Logo, o gráfico temporal está na Figura 2.

Para extração das análises estatísticas dos dados, fez-se uso do software R versão 3.4.1 (R DEVELOPMENT CORE TEAM, 2017) e da sua IDE (*Integrated Development Environment*) R STUDIO, ambos gratuitos e de código aberto.

FIGURA 7 - Gráfico temporal

Fonte: Elaborado pelos autores (2020)

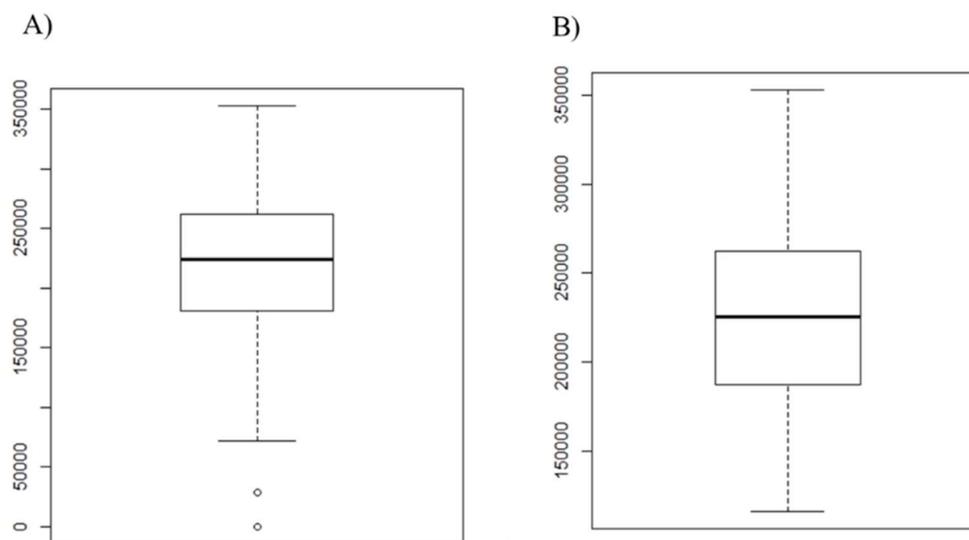
Na série temporal de vendas é observado valores atípicos, *outliers*, devido a duas circunstâncias:

- (i) nas semanas “06/08/2017”, “13/08/2017” e “20/08/2017” os dados foram extraviados;
- (ii) nas semanas “20/05/2018” e “27/05/2018” greve dos caminhoneiros, impactando o volume de vendas e expedição;

Os *outliers* estão na Figura 8. Os *outliers* foram substituídos, a métrica utilizada foi a média dos últimos três meses de acordo com a sequência da semana, ou seja, a semana que se iniciou no dia “06/08/2019” foi substituída pela média das últimas três primeiras semanas dos últimos meses “01/05/2019”, “04/06/2019” e “02/07/2019”, da mesma forma ocorreu para as outras variáveis. Contudo o Figura 8B mostra a não existência de *outliers*.

Essa métrica de substituição foi adotada devido as diferenças de volume de vendas de acordo com o período de fechamento do relatório, quanto mais próximo ao fechamento maior é o volume de vendas.

FIGURA 8 - Box-Plot da série de histórico de vendas. Figura 3A Box-Plot com Outliers. Figura 3B Box-plot sem Outliers.



Fonte: Elaborado pelos autores (2020)

Procedimento para calibração, validação e teste do modelo. Métricas utilizadas para verificar o ajuste do modelo.

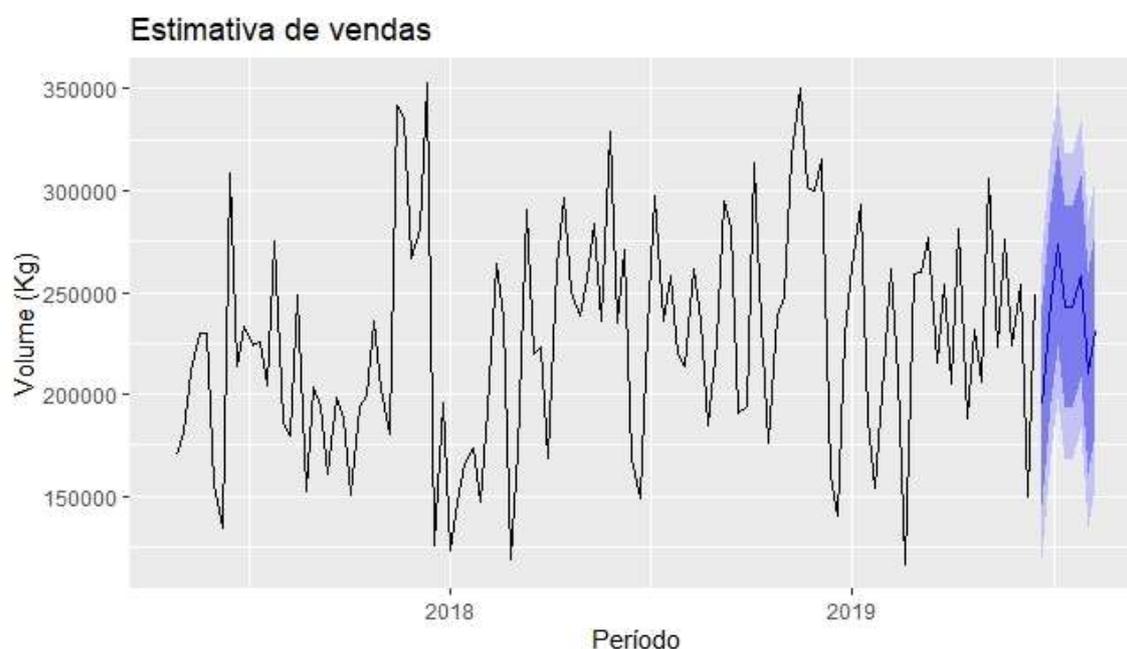
5.4 Resultados e discussão

A dura realidade é que certas práticas comuns nas organizações sacrificam o gerenciamento da cadeia de suprimentos para cumprimento de metas de curto prazo, isso é mais frequente no gerenciamento de vendas (GODSELL & VAN HOEK, 2009). Algumas práticas do gerenciamento de vendas vão contra o modelo JIT, afim de cumprir metas e alcançar resultados que possam garantir uma maior rentabilidade para área de vendas e assim motivá-los a vender cada dia mais. Porém, quando se vende em períodos fora do que foi planejado e em volumes divergentes ao traçado como meta, impacta toda a cadeia de suprimentos. A previsibilidade do que será vendido é de extrema importância

para a área de suprimentos e distribuição, adequação a este volume e preparar equipamentos para atender todos os clientes de acordo com o acordado no momento de coleta do pedido. Diminuir a ociosidade de mão-de-obra e de frota de veículos é fundamental para os objetivos de qualquer companhia, pois esta ociosidade impacta diretamente em custo.

A partir da relação com o histórico de vendas no caso de teste do período anterior e por meio do método ARIMA, é possível estimar o padrão de comportamento da venda futura, no qual exerce influência em todo o processo da cadeia de suprimentos. A previsão do comportamento de vendas nas próximas semanas é representada na Figura 9. Dando maior estimativa de volume a ser operado para a organização. Contudo, este número dará maior ciência de recursos necessários para atender a demanda.

FIGURA 9 - Estimativa de vendas



Fonte: Elaborado pelos autores (2020)

Os resultados obtidos através do modelo ARIMA apresenta um bom desempenho preditivo e parâmetros consistentes em suas significâncias estatísticas. A previsibilidade

do volume de vendas auxiliará na tomada de decisão, na contratação de veículos e impulsionar o transportador a realizar novos investimentos afim de atender a demanda futura. A previsibilidade do volume dará uma melhor assertividade à gestão dos níveis de estoque do produto final, posicionando itens de forma correta e que terão giro. Estoques elevados geram custos adicionais, enquanto estoques desbalanceados e muito baixo podem comprometer o atendimento dos clientes, gerando rupturas e corte em seus pedidos.

O estudo de (LIN, JANG, & WONG, 2005) tem como objetivo criar políticas de pedidos que minimizem o estoque sem criar atrasos que sacrifiquem a satisfação do cliente. No entanto, a operação logística real de uma cadeia de suprimentos é inerentemente um sistema dinâmico discreto, e é mais conveniente usar um modelo de séries temporais. Em um sistema de cadeia de suprimentos, é muito importante prever as mudanças no mercado para manter um nível de estoque que seja suficiente apenas para satisfazer a demanda do cliente. Já o estudo de (LO & WU, 2003) apresenta uma avaliação de risco na incerteza da previsão da demanda, neste mesmo estudo várias previsões de carga foram realizadas e a previsão de carga é baseada em modelos de séries temporais (ARIMA). A pesquisa de (DHAHRI & CHABCHOUB, 2007) visa reduzir, ou mesmo eliminar, o efeito chicote, distorção na demanda do mercado quando esta demanda se propaga de empresa para empresa. Finalmente, no final da cadeia, o fornecedor de matérias-primas recebe comandos completamente incertos. A solução que os autores propuseram ao chicote, baseia-se, no uso das funções de preferência baseadas em uma análise estatística de séries cronológicas (método de Box e Jenkins) para construir os diferentes modelos como demanda, estoque e quantidade da ordem. O artigo de (LEE, LIM, & KIM, 2017) desenvolve um método prático para prever o volume potencial de contêineres que pode ser induzido por um projeto de desenvolvimento de portos em uma

rede de transporte de contêineres combinando a escolha de portas e um modelo de média móvel integrado autorregressivo (ARIMA). O modelo ARIMA são ajustados aos dados da série temporal para interpretar melhor os dados ou para prever pontos futuros da série, este modelo vem sendo usado para prever a demanda e se torna usual em diversas companhias como mencionado acima.

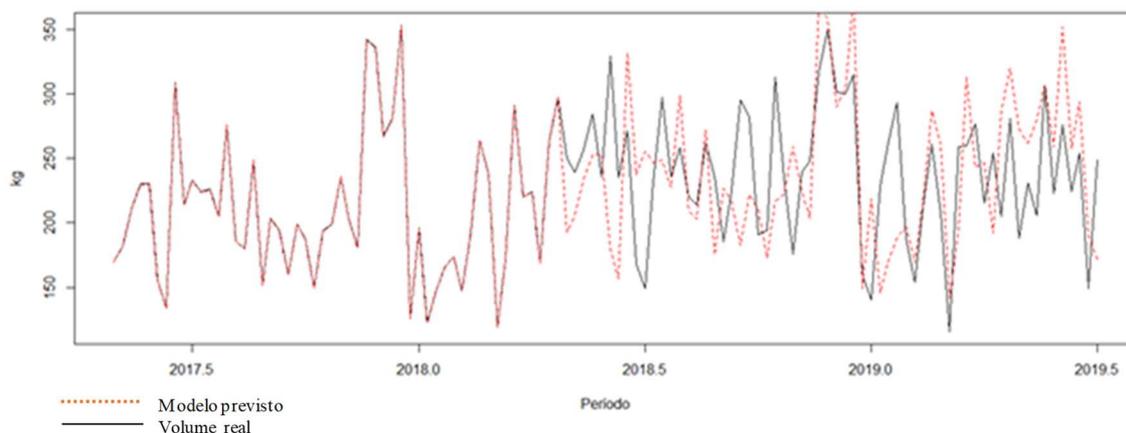
Para ajustar os parâmetros do modelo foi realizado testes, no qual o algoritmo percorreu um espaço de busca com os modelos possíveis e selecionou o que apresentou menor critério de AIC (*Akaike Information Criterion*) estimador de erro de previsão. O parâmetro com menor AIC foi o ARIMA (2,1,3), logo, a previsão do volume de vendas nas próximas 8 semanas e com o devido intervalo de confiança está no Apêndice C.

Para analisar a previsão do volume de vendas é necessário que a modelagem possua um bom ajuste. Os valores do MAPE, PBIAS, MAE e RMSE para calibração e validação da previsão de vendas estão na Tabela 10. A diferença entre o MAPE ficou baixo (< 2%), o PBIAS apresentou número baixo também com 2,09%.

O modelo foi aplicado para série histórica para análise comparativa de seu comportamento e está na Figura 10. Quando analisado o modelo com o período já realizado em maio de 2017 a início de 2018, apresentou melhores resultados, porém quando avaliado o período de maio de 2018 até o início de 2019 a modelagem se distancia do realizado, devido a greve dos caminhoneiros, a modelagem foi se aproximando do realizado na proximidade no meio do ano de 2019.

TABELA 7 - Acurácia do modelo

Modelo	MAPE (%)	PBIAS (%)	MAE	RMSE
Calibração	10,903	11,044	24,996	41,427
Validação	12,704	13,138	36.867,500	45.128,61

FIGURA 10 - Comparativo do modelo e série histórica

Fonte: Elaborado pelos autores (2020)

5.5 Conclusão

A importância da logística dos serviços ao cliente é hoje algo mais do que evidente e claro. As tomadas de decisões logísticas poderiam ser aperfeiçoadas com uma informação fidedigna como se processam as mudanças em vendas. É extremamente complexo a implantação do método JIT nas operações de distribuição, os picos de venda, as oscilações da demanda e o gerenciamento de vendas são totalmente contraditórias ao método japonês. Atender a demanda dos clientes e solicitar os insumos (inputs) necessários, com base nas previsões, requer um grau de evolução importante. Nas organizações geralmente existem problemas enormes para fazê-lo. Uma previsão superestimada da demanda eleva os estoques, e com isso os custos também se acentuam. Trabalhar com demandas abaixo dos patamares impacta o nível de serviço, deixando de atender diversos clientes. A previsão se torna ferramenta primordial para o objetivo do negócio, o modelo ARIMA apresenta um bom modelo de predição.

Podemos concluir que o ARIMA na forma de espaço de estado é uma opção prática, eficiente para a previsão da cadeia de suprimentos e, de fato, para qualquer contexto, onde a série histórica é limitada com poucos ciclos sazonais completos. No

entanto, esta não é a única área de aplicação possível, e acreditamos que desenvolver e explorar os algoritmos eficientes para a aplicação ARIMA em outros contextos de negócios é uma direção interessante para futuras pesquisas. Isso significa que, como um trabalho futuro, o ARIMA do espaço de estado deve ser testado em outros conjuntos de dados e comparado com outros métodos de previsão populares.

6 CONSIDERAÇÕES FINAIS

Este trabalho utilizou modelos para auxiliar e mensurar na identificação de variáveis que impactam o nível de atendimento ao cliente. Primeiramente, foi identificado quais variáveis que impactam o atendimento dos pedidos, depois foi modelado o índice de desempenho da distribuição logística. Por último foi modelado a demanda de vendas.

A principal variável que influencia de forma negativa o OTIF é o *On Time*, ou seja, atraso na entrega. Isso acontece devido aos elevados volumes de vendas que são maiores que a capacidade de distribuição, essa oscilação interfere de forma direta. A variável que contribui de forma positiva é o produto 11, pois este sempre esteve disponível em estoque e não ocorreu atrasos em sua distribuição. Foi possível identificar estas variáveis com aplicação da modelagem de equações estruturais com o método PLS (*Partial Least Square*). Ficou evidente também como a estratégia comercial impacta o nível de serviço, no qual é evidente o comportamento das semanas antecedentes ao fechamento do relatório de vendas.

Assim, com a Programação Genética foi possível modelar o índice OTIF, podendo avaliar o desempenho logístico. A Programação Genética exposta neste documento apresentou excelente desempenho ao modelar.

Ainda neste documento fica evidenciado como o setor logístico é frágil às oscilações de vendas da equipe comercial, e como prever uma demanda é de extrema importância para tomadas de decisões que poderão impactar nos custos operacionais e nos custos de distribuição. Ferramentas de séries temporais com modelo de previsão ARIMA é uma opção prática, eficiente para a previsão da demanda. Este método apresentou um bom desempenho preditivo e parâmetros consistentes em suas significâncias estatísticas.

Pode-se concluir, que o serviço logístico é fator preponderante para o sucesso do negócio da organização. Este parâmetro pode definir a continuidade de compra do cliente e até mesmo a migração para o concorrente.

As organizações estão se adequando e mensurando se todos os pedidos a serem entregues aos seus clientes estão sendo atendidos em sua totalidade, se existe atraso na entrega e se as condições e critérios que geraram o pedido foram atendidas. Essa análise multivariada de dados é possível gerar índices.

Os atributos perceptíveis que provocam a frustração, decepção e a inconformidade do cliente certamente impacta o nível de serviço, tão logo, que a mensuração do atendimento do pedido, desde a coleta do mesmo até a entrega do produto se tornam primordiais para análise e correção do que não está conforme.

Diante do que foi exposto cabe a organização dar foco no que é de interesse ao cliente, satisfazendo e atendendo o critério tomador de pedido, entregando no dia correto, nas condições estabelecidas, na hora correta e na totalidade com que foi pedido, tendo a certeza da continuidade de compras pelo nível de serviço prestado.

Como trabalho futuro é possível realizar estudo de impacto financeiro na concentração do volume de vendas nos últimos dias do mês, com o intuito de atingimento de metas de vendas. Este vício impacta diretamente na distribuição do produto, gerando demanda maior que a capacidade de distribuição, e para atendimento dos pedidos acaba tendo que ter um custo adicional com contratação a mais de veículo e mão-de-obra.

REFERENCIAS:

GOH, Shao Hung; ELDRIDGE, Stephen. Sales and Operations Planning: The effect of coordination mechanisms on supply chain performance. **International Journal of Production Economics**, v. 214, n. April, p. 80–94, 2019.

ALTIPARMAK, Fulya; GEN, Mitsuo; LIN, Lin; et al. A genetic algorithm approach for multi-objective optimization of supply chain networks. **Computers and Industrial Engineering**, v. 51, n. 1, p. 196–215, 2006.

BALAKRISHNAN, R; LINSMEIER, T.J; VENKATACHALAM, M. Financial benefits from JIT adoption: effects of customer concentration and cost structure. **The Accounting Review**, v. 71, n. 2, p. 183–205, 1996.

BOONE, Tonya; BOYLAN, John E.; FILDES, Robert; et al. Perspectives on supply chain forecasting. **International Journal of Forecasting**, v. 35, n. 1, p. 121–127, 2019.

BOONE, Tonya; GANESHAN, Ram; JAIN, Aditya; et al. Forecasting sales in the supply chain: Consumer analytics in the big data era. **International Journal of Forecasting**, v. 35, n. 1, p. 170–180, 2019.

BOYD, David T. Corporate Adoption of JIT: The Effect of Time and Implementation on Selected Performance Measures. **Southern Business Review**, v. 26, n. 2, p. 20–26, 2001.

BRAUNSCHEIDEL, Michael J.; SURESH, Nallan C. The organizational antecedents of a firm's supply chain agility for risk mitigation and response. **Journal of Operations Management**, v. 27, n. 2, p. 119–140, 2009.

BRESSAN, Aureliano Angel. Modelos de previsão de preços aplicados aos contratos futuros de boi gordo na BM&F. **Nova Economia**, v. 12, n. 1, p. 117–140, 2002.

BROOKS, Chris. **Introductory Econometrics for Finance**. 3rd Editio. New York: [s.n.], 2002.

CAI, Jian; LIU, Xiangdong; XIAO, Zhihui; et al. Improving supply chain performance management: A systematic approach to analyzing iterative KPI accomplishment. **Decision Support Systems**, v. 46, p. 512–521, 2009.

CALLEN, Jeffrey L.; MOREL, Mindy; FADER, Chris. The profitability-risk tradeoff of just-in-time manufacturing technologies. **Managerial and Decision Economics**, v. 24, n. 5, p. 393–402, 2003. Disponível em: <<http://doi.wiley.com/10.1002/mde.1104>>.

CAMPOS, Lucas de Souza; MENEZES, José Elmo De. Uso gerencial das ferramentas da qualidade para tomada de decisão: Um estudo de caso sobre o elevado índice de devolução na distribuição logística e a criação da área de Customer Service. **Sodebrás**, v. 14, n. ISSN-1809-3957, p. 203–208, 2019.

CHAPMAN, Paul; BERNON, Michael; HAGGETT, Paul. Applying selected quality management techniques to diagnose delivery time variability. **International Journal of Quality and Reliability Management**, v. 28, n. 9, p. 1019–1040, 2011.

Chellapilla, K. Evolving computer programs without subtree crossover. **IEEE Transactions on Evolutionary Computation**, v. 1, n. 3, p. 209–216, 1997.

CHIANG, Chung Yean; KOCABASOGLU-HILLMER, Canan; SURESH, Nallan. An empirical investigation of the impact of strategic sourcing and flexibility on firm's supply chain agility. **International Journal of Operations and Production Management**, v. 32, n. 1, p. 49–78, 2012.

DE OLIVEIRA, E.R.; SILVA, J.R.; BAUMANN, L.R.F.; et al. Development of a technological index for the assessment of the beef production systems of the vermelho river basin in Goiás, Brazil. **Pesquisa Operacional**, v. 38, n. 1, 2018.

DELLINO, Gabriella; LAUDADIO, Teresa; MARI, Renato; et al. A reliable decision support system for fresh food supply chain management. **International Journal of Production Research**, v. 56, n. 4, p. 1458–1485, 2018.

DHAHRI, Issam; CHABCHOUB, Habib. Nonlinear goal programming models quantifying the bullwhip effect in supply chain based on ARIMA parameters. **European Journal of Operational Research**, v. 177, n. 3, p. 1800–1810, 2007.

ESZERGÁR-KISS, Domokos; CAESAR, Bálint. Definition of user groups applying Ward's method. **Transportation Research Procedia**, v. 22, n. 2016, p. 25–34, 2017.

FILDES, Robert. The forecasting journals and their contribution to forecasting research: Citation analysis and expert opinion. **International Journal of Forecasting**, v. 22, n. 3, p. 415–432, 2006.

FILDES, Robert; GOODWIN, Paul; ONKAL, Dilek. Use and misuse of information in supply chain forecasting of promotion events. **International Journal of Forecasting**, v. 4, n. June, p. 144–156, 2016.

GODSELL, Janet; BIRTWISTLE, Andy; VAN HOEK, Remko. Building the supply chain to enable business alignment: lessons from British American Tobacco (BAT). **Supply Chain Management: An International Journal**, v. 15, n. 1, p. 10–15, 2010.

GODSELL, Janet; VAN HOEK, Remko. Fudging the supply chain to hit the number: Five common practices that sacrifice the supply chain and what financial analysts should ask about them. **Supply Chain Management**, v. 14, n. 3, p. 171–176, 2009.

Gokalp, O. N., Lee, S. H., & Peng, M. W. Competition and corporate tax evasion: An institution-based view. **Journal of World Business**, v. 52, n. 2, p. 258–269, 2017.

GRAWE, Scott J. Logistics innovation: A literature-based conceptual framework. **The International Journal of Logistics Management**, v. 20, n. 3, p. 360–377, 2009.

GUJARATI, Damodar N. **Basic econometrics**. New York, USA: [s.n.], 2003.

GUPTA, Mahesh C.; BOYD, Lynn H. Theory of constraints: a theory for operations management. **International Journal of Operations & Production Management**, v. 28, n. 10, p. 991–1012, 2008.

HASLE, Peter; BOJESEN, Anders; JENSEN, Per Langaa; et al. Lean and the working environment: A review of the literature. **International Journal of Operations and Production Management**, v. 32, n. 7, p. 829–849, 2012.

HYNDMAN, Rob; KOEHLER, Anne B.; ORD, J. Keith; et al. **Forecasting with Exponential Smoothing: The State Space Approach**. [s.l.: s.n.], 2008.

J R KOZA. Genetic programming II, automatic discovery of reusable subprograms. Cambridge, MA: **Harvard Business School**, 441, 439–441, 1995.

JAMES, G., WITTEN, D., HASTIE, T., & TIBSHIRANI, R. **An Introduction to Statistical Thermodynamics** (1st ed.; Springer, ed.), 2013.

JEFFERY, Mariah M.; BUTLER, Renee J.; MALONE, Linda C. Determining a cost-effective customer service level. **Supply Chain Management**, v. 13, n. 3, p. 225–232, 2008.

JIWA, Zeplin; TARIGAN, Husada. the Impact of Organization Commitment To Process and Product Innovation in Improving. v. 19, n. 2, p. 335–346, 2018.

JOE H. WARD, Jr. Hierarchical Grouping to Optimize an Objective Function. **Journal of the American Statistical Association**, v. 58, n. 301, p. 236–244, 1963.

KILIMCI, Zeynep Hilal; AKYUZ, A. Okay; UYSAL, Mitat; et al. An improved demand forecasting model using deep learning approach and proposed decision integration strategy for supply chain. **Complexity**, v. 2019, n. March, 2019.

KINNEY, Michael R.; WEMPE, William F. Further evidence on the extent and origins of JIT's profitability effects. **Accounting Review**, v. 77, n. 1, p. 203–225, 2002.

Koza, J. R., & Poli, R. **Genetic programming: on the programming of computers by means of natural selection. In Computer**, 1983.

KRISHNAN, Anjali; WILLIAMS, Lynne J.; MCINTOSH, Anthony Randal; et al. Partial Least Squares (PLS) methods for neuroimaging: A tutorial and review. **NeuroImage**, v. 56, n. 2, p. 455–475, 2011.

LEE, Sang Yoon; LIM, Hyunwoo; KIM, Hwa Joong. Forecasting container port volume: Implications for dredging. **Maritime Economics and Logistics**, v. 19, n. 2, p. 296–314, 2017.

LIN, Pin Ho; JANG, Shi Shang; WONG, David Shan Hill. Predictive control of a decentralized supply chain unit. **Industrial and Engineering Chemistry Research**, v. 44, n. 24, p. 9120–9128, 2005.

LIU, Kun-Hong; TONG, Muchenxuan; XIE, Shu-Tong; et al. Genetic Programming Based Ensemble System for Microarray Data Classification. **Computational and Mathematical Methods in Medicine**, v. 2015, n. 1, p. 1–11, 2015.

LO, K. L.; WU, Yuan-Kang. Risk assessment due to local demand forecast uncertainty in the competitive supply industry. **IEE Proceedings-Generation, Transmission and Distribution**, v. 150, n. 3, p. 573–582, 2003

LUKE, S., & PANAIT, L. **A survey and comparison of tree generation algorithms**. GECCO-2001: Proceedings of the Genetic and Evolutionary Computation Conference, 81–88, 2001.

MACHADO, V. H.; BARROSO, A. P.; MACHADO, V. Cruz. An Analytic Network Process model to support decision making in a pharmaceutical supply chain. **IEEE International Conference on Industrial Engineering and Engineering Management**, p. 1612–1616, 2014.

MATEOS-APARICIO, Gregoria. Communications in Statistics - Theory and Methods Partial Least Squares (PLS) Methods : Origins , Evolution , and Application to Social Sciences Partial Least Squares (PLS) Methods : Origins, Evolution, and Application to Social Sciences. v. 0926, n. January, p. 1–18, 2016.

MCGOLDRICK, Peter J.; ANDRE, Elisabeth. Consumer misbehaviour: Promiscuity or loyalty in grocery shopping. **Journal of Retailing and Consumer Services**, v. 4, n. 2, p. 73–81, 1997.

MISTRY, Jamshed J. Origins of profitability through JIT processes in the supply chain. **Industrial Management and Data Systems**, v. 105, n. 6, p. 752–768, 2005.

MOLINA, D., LaTorre, A., & Herrera, F. An Insight into Bio-inspired and Evolutionary Algorithms for Global Optimization: Review, Analysis, and Lessons Learnt over a Decade of Competitions. **Cognitive Computation**, v. 10 n.4, p. 517–544, 2018.

MORETTIN, Pedro Alberto; TOLOI, Clélia Maria de Castro. **Análise de Séries Temporais**. 2rd. ed. [s.l.: s.n.], 2006.

MORETTIN, Pedro Alberto; TOLOI, Clélia Maria de Castro. **Previsão de Séries Temporais**. 2rd. ed. [s.l.: s.n.], 1987.

MORIASI, D. N., ARNOLD, J. G., LIEW, M. W. Van, BINGNER, R. L., HARMEL, R. D., & Veith, T. L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. **American Society of Agricultural and Biological Engineers**, v. 50 n.3, p. 885–900, 2007.

MYERS, R. H., MYERS, S. L., WALPOLE, R. E., & KEYING, E. Y. Probability & Statistics for Engineers & Scientists, MyStatLab. (2 edition; Pearson Higher Ed, ed.), 2015.

NAGA VAMSI KRISHNA JASTI, Rambabu Kodali. A literature review of empirical research methodology in lean manufacturing. **International Journal of Operations & Production Management**, v. 34, n. 8, p. 1080–1122, 2014.

NASH, E., & SUTCLIFFE, V. Part I- a discussion of principles. The problem of determining river flows from rainfall, evaporation, and other factors, occupies a central place in the technology of applied hydrology . It is not only the essential problem of flood forecasting but a. 10, p. 282–290, 1970.

NG, Irene C.L.; DING, David Xin; YIP, Nick. Outcome-based contracts as new business model: The role of partnership and value-driven relational assets. **Industrial Marketing Management**, v. 42, n. 5, p. 730–743, 2013.

OLHAGER, Jan; RUDBERG, Martin; WIKNER, Joakim. Long-term capacity management: Linking the perspectives from manufacturing strategy and sales and

operations planning. **International Journal of Production Economics**, v. 69, n. 2, p. 215–225, 2001.

OLIVEIRA, Elis Regina de, SILVA, J. R., BAUMANN, L. R. F., MIZIARA, F., FERREIRA Júnior, L. G., MERELLES, L. R. de O., & COUTO, V. R. M. Development of a Technological Index for the Assessment of the Beef Production Systems of the Vermelho River Basin in Goiás, Brazil. **Pesquisa Operacional**, 2018.

ONSTEIN, Alexander T.C.; TAVASSZY, Lóránt A.; VAN DAMME, Dick A. Factors determining distribution structure decisions in logistics: a literature review and research agenda. **Transport Reviews**, v. 39, n. 2, p. 243–260, 2019.

PÉREZ-ENCISO, Miguel; TENENHAUS, Michel. Prediction of clinical outcome with microarray data: a partial least squares discriminant analysis (PLS-DA) approach. **Human Genetics**, v. 112, n. 5, p. 581–592, 2003.

PRAJOGO, Daniel; OLHAGER, Jan. Supply chain integration and performance: The effects of long-term relationships, information technology and sharing, and logistics integration. **International Journal of Production Economics**, v. 135, n. 1, p. 514–522, 2012.

REZANIA, Mohammad; JAVADI, Akbar A. A new genetic programming model for predicting settlement of shallow foundations. **Canadian Geotechnical Journal**, v. 44, n. 12, p. 1462–1473, 2008.

RICHARD G BROWN. **Statistical forecasting for inventory control**. New York: [s.n.], 1959.

ROB J HYNDMAN; ATHANASOPOULOS, George. **Forecasting: principles and practice**. Melbourne, Australia: OTexts, 2013.

RODRIGUES, Leonardo; MERELLES, De Oliveira; OLIVEIRA, Cláudio De; et al. Previsão de geração de resíduos sólidos para o aterro de Aparecida de Goiânia (GO) por séries temporais. **Engenharia Sanitária Ambiental**, p. 1–10, 2018.

SAHAY, B.S.; MOHAN, Ramneesh. Supply chain management practices in Indian industry. **International Journal of Physical Distribution & Logistics Management**, v. 33, n. 7, p. 582–606, 2003

SANDERS, Nada R. How to Use Big Data to Drive Your Supply Chain. **California Management Review**, v. 58, n. 3, p. 26–48, 2016.

Scrucca, L. GA: A Package for Genetic Algorithms in R . **Journal of Statistical Software**, v. 53 n.4, 2015.

SEHGAL, Sanjay; SAHAY, B. S.; GOYAL, S. K. Reengineering the supply chain in a paint company. **International Journal of Productivity and Performance Management**, v. 55, n. 8, p. 655–670, 2006.

SUAREZ-BARRAZA, Manuel F.; MIGUEL-DAVILA, José-Á; VASQUEZ-GARCÍA, C. Fabiola. Supply chain value stream mapping: a new tool of operation management. **International Journal of Quality & Reliability Management**, v. 33, n. 4, p. 518–534, 2016.

SVETUNKOV, Ivan; BOYLAN, John E. State-space ARIMA for supply-chain forecasting. **International Journal of Production Research**, n. June, 2019.

TRKMAN, Peter; MCCORMACK, Kevin; DE OLIVEIRA, Marcos Paulo Valadares; et al. The impact of business analytics on supply chain performance. **Decision Support Systems**, v. 49, n. 3, p. 318–327, 2010.

Tsang, E., Yung, P., & Li, J. EDDIE-automation, a decision support tool for financial forecasting. **Decision Support Systems**, v. 37, n. 4, 559–565, 2004.

W Didimo, L Giamminonni, Liotta, G., & F Montecchiani. A visual analytics system to support tax evasion discovery. **Decision Support Systems**, 110, 71–83, 2018.

WANG, Wei; FUNG, Richard Y K; CHAI, Yueting. Approach of just-in-time distribution requirements planning for supply chain management. **International Journal of Production Economics**, v. 91, n. 2, p. 101–107, 2004.

WILLIAM H. GREENE. **Econometric Analysis**. 3rd. ed. [s.l.: s.n.], 1997.

WOLD, Svante; SJÖSTRÖM, Michael; ERIKSSON, Lennart. PLS-Regression - A basic tool of chemometrics. **Chemometrics and Intelligent Laboratory Systems**, v. 58, p. 109–130, 2001.

XU, Shuojian; CHAN, Hing Kai; ZHANG, Tiantian. Forecasting the demand of the aviation industry using hybrid time series SARIMA-SVR approach. **Transportation Research Part E: Logistics and Transportation Review**, v. 122, n. October 2018, p. 169–180, 2019.

ZHANG, Ming; MU, Hailin; LI, Gang; et al. Forecasting the transport energy demand based on PLSR method in China. **Energy**, v. 34, n. 9, p. 1396–1400, 2009.

APÊNDICE:

Apêndice A – Volume semanal de vendas

t	X_t										
01/05/2017	170.278	10/09/2017	193.801	21/01/2018	165.886	03/06/2018	328.848	14/10/2018	313.498	24/02/2019	208.213
07/05/2017	181.662	17/09/2017	160.473	28/01/2018	173.660	10/06/2018	235.237	21/10/2018	235.290	03/03/2019	116.317
14/05/2017	211.495	24/09/2017	198.549	04/02/2018	146.757	17/06/2018	270.720	28/10/2018	175.932	10/03/2019	259.108
21/05/2017	229.906	01/10/2017	188.311	11/02/2018	192.216	24/06/2018	168.151	04/11/2018	238.901	17/03/2019	259.541
28/05/2017	230.302	08/10/2017	150.284	18/02/2018	264.230	01/07/2018	149.135	11/11/2018	248.328	24/03/2019	277.168
04/06/2017	155.962	15/10/2017	194.209	25/02/2018	237.757	08/07/2018	231.992	18/11/2018	317.152	31/03/2019	215.436
11/06/2017	134.129	22/10/2017	199.410	04/03/2018	118.874	15/07/2018	297.499	25/11/2018	350.156	07/04/2019	253.540
18/06/2017	308.706	29/10/2017	235.615	11/03/2018	171.617	22/07/2018	235.874	02/12/2018	301.334	14/04/2019	205.431
25/06/2017	214.082	05/11/2017	202.847	18/03/2018	290.711	29/07/2018	258.026	09/12/2018	300.363	21/04/2019	281.252
02/07/2017	233.057	12/11/2017	180.730	25/03/2018	219.760	05/08/2018	219.611	16/12/2018	315.437	28/04/2019	188.429
09/07/2017	224.110	19/11/2017	341.960	01/04/2018	223.514	12/08/2018	213.501	23/12/2018	160.157	05/05/2019	231.364
16/07/2017	226.047	26/11/2017	335.755	08/04/2018	168.516	19/08/2018	261.953	30/12/2018	140.489	12/05/2019	206.098
23/07/2017	204.622	03/12/2017	266.596	15/04/2018	262.813	26/08/2018	236.137	06/01/2019	228.073	19/05/2019	306.214
30/07/2017	275.583	10/12/2017	281.181	22/04/2018	296.832	02/09/2018	185.106	13/01/2019	262.111	26/05/2019	223.052
06/08/2017	NA	17/12/2017	352.964	29/04/2018	250.191	09/09/2018	225.352	20/01/2019	292.926	02/06/2019	276.323
13/08/2017	NA	24/12/2017	126.212	06/05/2018	238.623	16/09/2018	294.792	27/01/2019	187.749	09/06/2019	224.188
20/08/2017	NA	31/12/2017	195.796	13/05/2018	256.707	23/09/2018	281.950	03/02/2019	154.187	16/06/2019	253.853
27/08/2017	151.974	07/01/2018	123.392	20/05/2018	28.693	30/09/2018	190.906	10/02/2019	212.266	23/06/2019	149.288
03/09/2017	203.777	14/01/2018	146.905	27/05/2018	71.714	07/10/2018	194.456	17/02/2019	261.351	30/06/2019	248.929

NA: Dados faltantes; t : Período no tempo; X_t : Ponto no período.

Apêndice B – Medidas descritivas

Média	Mediana	Desvio padrão	Coefficiente de variação	Coefficiente de assimetria	Curtose
226342,20	225699,50	54582,65	24,12%	0,1565	-0,5226

Apêndice C – Volume previsto de vendas para as próximas 8 semanas

Período	Previsão de vendas	Intervalo de confiança			
		Limite inferior 80%	Limite superior 80%	Limite inferior 95%	Limite superior 95%
07/07/2019	254.871	182.292	327.449	143.871	365.870
14/07/2019	319.871	247.292	392.449	208.871	430.870
21/07/2019	258.871	186.292	331.449	147.871	369.870
28/07/2019	280.871	208.292	353.449	169.871	391.870
04/08/2019	242.871	170.292	315.449	131.871	353.870
11/08/2019	236.871	164.292	309.449	125.871	347.870
18/08/2019	284.871	212.292	357.449	173.871	395.870
25/08/2019	258.871	186.292	331.449	147.871	369.870

Apêndice D – Teste do modelo

Modelo	ϕ_1	ϕ_2	ϕ_3	θ_1	θ_2	θ_3	Intercept	AIC
Arima(0,0,1)				0,2434			226,3062	1234.33
Arima(1,0,0)	0,2506						226,2177	1234.16
Arima(1,0,1)	0.1624			0.0941			226.2563	1236.10
Arima(0,0,2)				0.2591	0.0408		226.2641	1236.08
Arima(1,0,2)	0.6457			-0.3906	-0.1067		226.1418	1237.83
Arima(2,0,0)	0.2555	-0.0193					226.2517	1236.11
Arima(2,0,1)	0.1275	0.0099		0.1287			226.2617	1238.10
Arima(2,0,2)	0.5958	0.0314		-0.3425	-0.1258		226.1379	1239.82
Arima(0,0,3)				0.2697	0.0326	-0.0341	226.3058	1237.93
Arima(1,0,3)	0.6116			-0.3652	-0.0979	0.0254	226.1057	1239.78
Arima(2,0,3)	0.0473	0.8111		0.2775	-0.7455	-0.3349	226.1948	1238.59
Arima(3,0,0)	0.2568	-0.0326	0.0495				226.1878	1237.84
Arima(3,0,1)	0.6412	-0.1384	0.0857	-0.3862			226.0814	1239.31
Arima(3,0,2)	0.5030	-1.0620	0.2806	-0.2330	0.9972		226.5359	1228.52
Arima(3,0,3)	-0.6405	0.1052	0.4829	1.0236	0.1564	-0.5397	226.1914	1231.60
Arima(0,1,1)				-0.9469			0	1231.75
Arima(1,1,0)	-0.3641						0	1258.79
Arima(1,1,1)	0.2363			-0.9681			0	1227.75
Arima(0,1,2)				-0.7320	-0.2240		0	1227.72
Arima(1,1,2)	0.1038			-0.8279	-0.1337		0	1229.63
Arima(2,1,0)	-0.4742	-0.3142					0	1249.23
Arima(2,1,1)	0.2416	-0.0333		-0.9659			0	1229.64
Arima(2,1,2)	0.1269	-0.0079		-0.8507	-0.1114		0	1231.62
Arima(0,1,3)				-0.7218	-0.2076	-0.0279	0	1229.60
***Arima(1,1,3)								
Arima(2,1,3)	-1.2564	-0.7030		0.6068	-0.4979	-0.9575	0	1222.59
Arima(3,1,0)	-0.5784	-0.4666	-0.3272				0	1238.56
Arima(3,1,1)	0.2456	-0.0410	0.0391	-0.9685			0	1231.48
***Arima(3,1,2)								
Arima(3,1,3)	-1.5195	-0.4471	0.179	0.9034	-0.8132	-0.9684	0	1225.30
Arima(0,2,1)				-1.0000			0	1269.14
Arima(1,2,0)	-0.5819						0	1330.53

Arima(1,2,1)	-0.3583			-1.0000			0	1256.01
Arima(0,2,2)				-1.9419	-1.9419		0	1233.29
Arima(1,2,2)	0.2592			-1.9898	0.9914		0	1228.68
Arima(2,2,0)	-0.8234	-0.4224					0	1310.80
Arima(2,2,1)	-0.4681	-0.3081		-1.0000			0	1247.08
Arima(2,2,2)	0.2614	-0.0212		-1.9827	0.9841		0	1230.66
Arima(0,2,3)				-1.7259	0.4865	0.2410	0	1228.87
Arima(1,2,3)	0.1414			-1.8619	0.7445	0.1190	0	1230.65
Arima(2,2,3)	0.6193	-0.0792		-2.3582	1.7208	-0.3616	0	1232.65
Arima(3,2,0)	-1.0436	-0.8389	-0.5094				0	1279.40
Arima(3,2,1)	-0.5723	-0.4597	-0.3212	-1.0000			0	1237.08
Arima(3,2,2)	0.2679	-0.0290	0.0550	-1.9976	0.9994		0	1232.33
Arima(3,2,3)	-0.3888	0.1654	-0.0955	-1.3319	-0.3151	0.6492	0	1233.48
Arima(0,3,1)				-1.0000			0	1369.14
Arima(1,3,0)	-0.6989						0	1417.52
Arima(1,3,1)	-0.5780			-1.0000			0	1327.31
Arima(0,3,2)				-1.9941	1.0000		0	1270.25
Arima(1,3,2)	-0.3493			-1.9955	1.0000		0	1258.18
Arima(2,3,0)	-1.0352	-0.4845					0	1390.08
Arima(2,3,1)	-0.8182	-0.4172		-1.0000			0	1308.47
Arima(2,3,2)	-0.4576	-0.2983		-1.9967	1.0000		0	1250.18
Arima(0,3,3)				-2.7071	2.4167	-0.7091	0	1240.24
Arima(1,3,3)	0.3065			-2.9775	2.9626	-0.9850	0	1236.60
Arima(2,3,3)	0.2602	-0.0011		-2.9318	2.8681	-0.9361	0	1238.79
Arima(3,3,0)	-1.3248	-1.0849	-0.5785				0	1347.39
Arima(3,3,1)	-1.0392	-0.8328	-0.5050	-1.000			0	1278.19
Arima(3,3,2)	-0.5613	-0.4477	-0.3113	-1.9977	0.9999		0	1241.15
Arima(3,3,3)	-1.0074	-0.6358	-0.4066	-1.4878	-0.0180	0.5096	0	1243.43

*** Com este vetor de parâmetros a série não é estacionária.