



CENTRO UNIVERSITÁRIO DO ESTADO PARÁ - CESUPA
ESCOLA DE NEGÓCIOS TECNOLOGIA E INOVAÇÃO DO CESUPA - ARGO
CURSO DE ENGENHARIA DE PRODUÇÃO

RODRIGO SIMÕES TEIXEIRA
SIRIUS RAFFAEL JANSEN COSTA SIQUEIRA

**PREVISÃO DE DEMANDA EM UMA REDE DE POSTOS DE COMBUSTÍVEIS,
COM O AUXÍLIO DE SÉRIES TEMPORAIS, MÉTODOS CAUSAIS E REDES
NEURAS ARTIFICIAIS**

BELÉM
2018



RODRIGO SIMÕES TEIXEIRA
SIRIUS RAFFAEL JANSEN COSTA SIQUEIRA

**PREVISÃO DE DEMANDA EM UMA REDE DE POSTOS DE COMBUSTÍVEIS,
COM O AUXÍLIO DE SÉRIES TEMPORAIS, MÉTODOS CAUSAIS E REDES
NEURAS ARTIFICIAIS**

Trabalho de conclusão de curso apresentado à Área de Ciências Exatas e Tecnológicas do Centro Universitário do Estado do Pará como requisito para obtenção do título de Bacharel em Engenharia de Produção.

Orientador: MSc. Felipe Fonseca Tavares de Freitas

BELÉM
2018

Rodrigo Simões Teixeira

Sirius Raffael Jansen Costa Siqueira

**PREVISÃO DE DEMANDAS EM UMA REDE DE POSTOS DE COMBUSTÍVEIS,
COM AUXÍLIO DE SÉRIES TEMPORAIS, MÉTODOS CAUSAIS E REDES
NEURAIIS ARTIFICIAIS**

Trabalho de Curso apresentado na modalidade monografia, apresentado como requisito parcial para obtenção do grau em Bacharelado em Engenharia de Produção do Centro Universitário do Estado do Pará – CESUPA.

Data da Defesa:/...../.....

Banca Examinadora:

Prof. Orientador Msc. Felipe Fonseca Tavares De Freitas – CESUPA

Prof.^a Msc Polyana Santos Fonseca Nascimento – Membro Interno - CESUPA

Prof. Msc Carlos Gilberto Vieira da Silva Junior – Membro Interno - CESUPA

Belém

2018

AGRADECIMENTOS

Agradeço primeiramente a minha família, por ter me ensinado desde sempre a batalhar por tudo aquilo que eu desejo e quero para mim, por me apoiar e incentivar nas minhas decisões e pelo suporte que me deram durante todos esses anos. Em especial, quero agradecer ao meu pai, Antônio Carlos dos Santos Teixeira, por toda a ajuda e por me dar a oportunidade de estudar em uma boa instituição. Agradecer também a minha avó, Maria Leonor dos Santos Teixeira, por ter desempenhado um papel muito importante na minha vida, pelo seu carinho, atenção, dedicação, apoio e amor durante toda essa caminhada.

Agradecer aos meus amigos de longa data e aos que fiz durante o curso, pelo apoio e confiança que recebi de todos vocês. Agradeço também pelos momentos únicos e divertidos e que vocês puderam me proporcionar durante o curso. Em especial pelo meu amigo e dupla de TC, Sirius Raffael Jansen Costa Siqueira, pela amizade e dedicação que apresentou desde o início do curso até o final do mesmo.

Agradecer também a todos os professores que fizeram parte e contribuíram para o meu aprendizado e conhecimento durante todo o curso, em especial ao professor, coordenador e orientador Felipe Fonseca Tavares de Freitas, por sua ajuda, atenção e tempo para a conclusão deste trabalho.

AGRADECIMENTOS

Ao orientador, Mestre Felipe Fonseca Tavares de Freitas, por toda sua ajuda, apoio, conselhos e disposição para elaboração deste trabalho.

À professora Mestre, Polyana Santos Fonseca, por todo o carisma, gentileza e boa vontade em auxiliar em qualquer momento requerido.

Ao professor, Carlos Benedito Pereira da Paixão, por toda paciência em colaborar com o andamento deste trabalho.

Aos demais professores e professoras que ao longo do curso de alguma forma agregaram conhecimento em mim, para que pudesse seguir confiante de meus potenciais.

Ao meu pai, Raimundo Nonato Oliveira Siqueira, e à minha mãe, Symone Celeste Jansen Costa Siqueira, por me darem a melhor educação possível, além de todas as oportunidades para minha formação. Sem contar tudo que fizeram por mim ao longo da minha vida. Serei eternamente grato por tudo e estarei sempre disposto a recompensar todo o esforço e confiança investido em mim.

Ao meu irmão, Daniel Henrique Jansen Costa Siqueira, e à minha irmã, Suyane Priscila Jansen Costa Siqueira, por todo apoio e confiança depositados em mim.

Aos meus colegas de classe, por todo auxílio e troca de experiências ao longo do curso. Em especial aos meus amigos, Marcos Valente, Michel Moreira, Gabriel Oliveira e Raphaela Gallo, por toda companhia, ajuda, carinho e acima de tudo o privilégio de os ter conhecido e poder contar com vocês para tudo.

Ao meu amigo e parceiro desta monografia, Rodrigo Simões Teixeira, pela amizade verdadeira e por toda colaboração e esforço investido neste trabalho.

Sirius Raffael Jansen Costa Siqueira

RESUMO

O presente trabalho visa apresentar o estudo acerca de três métodos quantitativos existentes sobre previsão de demanda, que são as séries temporais, métodos causais e redes neurais artificiais. E a partir de seus resultados, realizar análises comparativas, tanto de dados diários, quanto semanais, com o intuito de se conhecer qual destes conseguem atingir melhor desempenho na previsão de demanda de uma rede de postos de combustíveis. Tendo como base em uma medição de erro, onde aquele que apresentar o menor será considerado o de melhor método. Para somente assim, atingir o objetivo deste estudo, que se baseia em avaliar qual dos métodos matemáticos consegue ser mais eficientes numa previsão de demanda. Para tal, serão aplicados nas ferramentas Excel®, Crystal Ball® e Matlab®. Os resultados alcançados mostraram que as redes neurais artificiais conseguiram obter melhor eficiência nas previsões, seguida pela média móvel dupla, das séries temporais, e por fim os métodos causais. As análises semanais obtiveram os melhores resultados, pelo fato de agruparem os dados e eliminarem boa parte dos *outliers*. Podendo assim concluir que as redes neurais artificiais conseguem ter melhor desempenho para realizar uma previsão de demandas.

Palavras-chave: Previsão de demanda. Séries Temporais. Métodos Causais. Redes Neurais Artificiais.

ABSTRACT

The object of this study is demonstrate with three quantity ways about ideas of demand, what is: time series, casual method and artificial neural system. And with this results, organize comparations about routine data and find the better method to be apply in a gas station. Based on an error measurement, where the one presenting the lowest is considered the best method. In order to achieve this, the objective of this study is to evaluate which of the mathematical methods can be more efficient in forecasting demand. In analyses will be use Excel®, Crystal Ball® and Matlab®. After all work, it is possible see artificial neural system with the best results, in second place we have time series and for last casual method. When we consider the week analyses, it's verified the best results, because of data control and outliers finish. In this form, it is correct consider the artificial neural with the best method when the subject is ideas of demand.

Keywords: Demand Forecast. Times Series. Causal Methods. Artificial Neural Networks.

LISTA DE ILUSTRAÇÕES

Figura 1 - Representação de uma rede neural	33
Figura 2 - Representação de um aprendizado não supervisionado	36
Figura 3 - Rede de Camada Única	38
Figura 4 - Rede de Múltiplas Camadas	39
Figura 5 - Rede recorrente sem neurônios ocultos	40
Figura 6 - Rede recorrente com neurônios ocultos	41
Figura 7 - Fluxograma dos procedimentos metodológicos realizados.....	49
Figura 8 - Fórmula da média no Excel®.....	50
Figura 9 - Onde encontrar o suplemento Solver® no Excel®	51
Figura 10 - Aplicação parcial da média móvel simples em dados diários	52
Figura 11 - Aplicação parcial da média móvel dupla em dados diários	52
Figura 12 - Aplicação parcial da média móvel ponderada em dados diários.....	53
Figura 13 - Aplicação parcial da suavização exponencial simples em dados diários.....	53
Figura 14 - Aplicação parcial da suavização exponencial com tendência em dados diários ...	54
Figura 15 - Aplicação parcial da suavização exponencial com tendência e sazonalidade em dados diários	54
Figura 16 - Onde encontrar a opção Predictor® no Excel®	55
Figura 17 - Seleção da série de dados inseridos no Predictor®	55
Figura 18 - Descrição da característica dos dados inseridos no Predictor®	56
Figura 19 - Seleção dos métodos de previsão que serão aplicados aos dados inseridos no Predictor®	57
Figura 20 - Seleção do tipo de erro e de previsão dos dados inseridos no Predictor®.....	58
Figura 21 - Previsão de MMS realizada pelo Crystal Ball® para o etanol do posto D	59
Figura 22 - Seleção parcial dos dados inseridos no Predictor®	60
Figura 23 - Seleção da série de dados inseridos no Predictor®	60
Figura 24 - Descrição da característica dos dados inseridos no Predictor®	61
Figura 25 - Seleção dos métodos de previsão que serão aplicados aos dados do combustível etanol do posto D e definição das variáveis dependentes e independentes inseridos no Predictor®	62
Figura 26 - Seleção do método de previsão no Predictor®.....	63

Figura 27 - Seleção do tipo de erro e da técnica de previsão dos dados inseridos no Predictor®	64
Figura 28 - Previsão de regressão padrão realizada pelo Crystal Ball® para o etanol do posto D.....	65
Figura 29 - Parte do agrupamento dos dados, no Excel®, do combustível Etanol do Posto D	66
Figura 30 - Seleção das primeiras 7 linhas e 814 colunas para agrupamento, no Excel®, do combustível Etanol do Posto D	67
Figura 31 - Dados de saída baseando-se no 7º valor da coluna seguinte, no Excel®, do combustível Etanol do Posto D	67
Figura 32 - Saída igual a 0, por não ter referência na coluna seguinte. Dados do Etanol do Posto D, no Excel®.....	68
Figura 33 - Encontrando os menores e maiores valores, dentre os dados diários do Etanol do Posto D, no Excel®.....	69
Figura 34 - Exemplo de dados diários do Etanol do Posto D, antes e depois da normalização, no Excel®.....	69
Figura 35 - Criação de cada matriz necessária, no software Matlab®, como os dados do Etanol do Posto D	70
Figura 36 - Ordens das matrizes criadas no software Matlab®, com os dados do Etanol do Posto D.....	71
Figura 37 - Estrutura dos dados diários do Etanol do Posto D, no Excel®	71
Figura 38 - Script para utilização em todas aplicações, no Matlab®	72
Figura 39 - Parte do agrupamento de dados semanais do Etanol do Posto D, no Excel®	72
Figura 40 - Ordem das matrizes dos dados semanais do Etanol do Posto D, no Matlab®.....	73
Figura 41 - Estrutura dos dados semanais do Etanol, no Excel®.....	73

LISTA DE QUADROS

Quadro 1 - Exemplos de tipos de padrões de demanda por diferentes autores	18
Quadro 2 - Exemplos de tipos de padrões de demanda por diferentes autores (Continuação)	19
Quadro 3 - Erros obtidos por meio das fórmulas aplicadas no Excel® e do Crystal Ball® para os dados diários e semanais referentes ao etanol do posto D.....	76
Quadro 4 - Erros obtidos por meio do Crystal Ball® para os dados diários referentes ao etanol do posto D	77
Quadro 5 - Erros obtidos por meio do Crystal Ball® para os métodos de regressão linear simples e múltipla sobre os dados semanais referentes ao etanol do posto D.....	78
Quadro 6 - Parâmetros utilizados para obtenção dos melhores resultados, dos dados do etanol do posto D, no Matlab®.....	78
Quadro 7 - Resultados obtidos a partir dos melhores parâmetros encontrados, nos dados do etanol do posto D. No Excel®	79
Quadro 8 - Parâmetros que resultaram nas melhores taxas de acerto, na análise semanal, do etanol do posto D. No Excel®	79
Quadro 9 - Resultados obtidos nos dados semanais do etanol do posto D, no Excel®	80
Quadro 10 - Comparação entre os melhores resultados obtidos em cada método, diária e semanalmente, para os dados do etanol do posto D, no Excel®	81
Quadro 11 - Análise comparativa do ponto de vista dos autores entre os métodos utilizados	83
Quadro 12 - Análise comparativa do ponto de vista dos autores entre as ferramentas utilizadas para os métodos	84

SUMÁRIO

1 INTRODUÇÃO	14
2 REFERENCIAL TEÓRICO	18
2.1 PREVISÃO DE DEMANDA	18
2.2 PADRÃO DE DEMANDA	18
2.3 MÉTODOS QUANTITATIVOS.....	19
2.3.1 Séries temporais	19
2.3.1.1 Média móvel simples	20
2.3.1.2 Média móvel dupla	21
2.3.1.3 Média móvel ponderada.....	21
2.3.1.4 Suavização exponencial simples	22
2.3.1.5 Suavização exponencial com tendência (<i>Holt</i>).....	23
2.3.1.6 Suavização exponencial com tendência e sazonalidade (<i>Holt-Winters</i>).....	23
2.3.2 Métodos Causais	24
2.3.2.1 Análise de Regressão Linear Simples.....	25
2.3.2.2 Análise de Regressão Linear Múltipla	27
2.3.2.3 Estimação de modelos.....	29
2.4 REDES NEURAIS ARTIFICIAIS	30
2.4.1 Modelo MCP de neurônio artificial	31
2.4.2 Rede neural do tipo <i>feedforward</i>	32
2.4.3 Processos de aprendizagem	34
2.4.3.1 Aprendizado Supervisionado	35
2.4.3.2 Aprendizado não-supervisionado	36
2.4.4 Arquitetura de Rede	37
2.4.4.1 Redes de camada única	37
2.4.4.2 Redes de múltiplas camadas.....	38
2.4.4.3 Redes recorrentes.....	39
2.4.5 Algoritmo de Retropropagação (<i>backpropagation</i>)	41
2.5 ERROS DE PREVISÃO	43
3 METODOLOGIA	45
3.1 CARACTERIZAÇÃO DA PESQUISA.....	45

3.2 LÓCUS DA PESQUISA	45
3.3 PROCEDIMENTOS METODOLÓGICOS	46
3.3.1 Dados diários	46
3.3.2 Dados semanais	47
4 APLICAÇÃO DOS MÉTODOS	50
4.1 APLICAÇÃO DE SÉRIES TEMPORAIS NO EXCEL®	50
4.2 APLICAÇÃO DE SÉRIES TEMPORAIS NO CRYSTAL BALL®	54
4.3 APLICAÇÃO DE MÉTODOS CAUSAIS NO CRYSTAL BALL®.....	59
4.4 APLICAÇÃO DE REDES NEURAS ARTIFICIAIS NO MATLAB®.....	65
5 RESULTADOS	74
5.1 COLETA DE DADOS	74
5.2 SÉRIES TEMPORAIS	75
5.3 MÉTODOS CAUSAIS.....	77
5.4 REDES NEURAS ARTIFICIAIS	78
6 ANÁLISES COMPARATIVAS	81
7 CONSIDERAÇÕES FINAIS	85
REFERÊNCIAS	87
APÊNDICE A - QUADRO REFERENTE À APLICAÇÃO DE SÉRIES TEMPORAIS, EM DADOS DIÁRIOS, COM E SEM UTILIZAÇÃO DO CRYSTAL BALL® NO POSTO A	89
APÊNDICE B - QUADRO REFERENTE À APLICAÇÃO DE SÉRIES TEMPORAIS, EM DADOS DIÁRIOS, COM E SEM UTILIZAÇÃO DO CRYSTAL BALL® NO POSTO B	90
APÊNDICE C - QUADRO REFERENTE À APLICAÇÃO DE SÉRIES TEMPORAIS, EM DADOS DIÁRIOS, COM E SEM UTILIZAÇÃO DO CRYSTAL BALL® NO POSTO C	91
APÊNDICE D - QUADRO REFERENTE À APLICAÇÃO DE SÉRIES TEMPORAIS, EM DADOS DIÁRIOS, COM E SEM UTILIZAÇÃO DO CRYSTAL BALL® NO POSTO D	92
APÊNDICE E - QUADRO REFERENTE À APLICAÇÃO DE SÉRIES TEMPORAIS, EM DADOS DIÁRIOS, COM E SEM UTILIZAÇÃO DO CRYSTAL BALL® NO POSTO E	93
APÊNDICE F - QUADRO REFERENTE À APLICAÇÃO DE SÉRIES TEMPORAIS, EM DADOS SEMANAIS, COM E SEM UTILIZAÇÃO DO CRYSTAL BALL® NO POSTO A	94

APÊNDICE G - QUADRO REFERENTE À APLICAÇÃO DE SÉRIES TEMPORAIS, EM DADOS SEMANAIS, COM E SEM UTILIZAÇÃO DO CRYSTAL BALL® NO POSTO B.....	95
APÊNDICE H - QUADRO REFERENTE À APLICAÇÃO DE SÉRIES TEMPORAIS, EM DADOS SEMANAIS, COM E SEM UTILIZAÇÃO DO CRYSTAL BALL® NO POSTO C.....	96
APÊNDICE I - QUADRO REFERENTE À APLICAÇÃO DE SÉRIES TEMPORAIS, EM DADOS SEMANAIS, COM E SEM UTILIZAÇÃO DO CRYSTAL BALL® NO POSTO D.....	97
APÊNDICE J - QUADRO REFERENTE À APLICAÇÃO DE SÉRIES TEMPORAIS, EM DADOS SEMANAIS, COM E SEM UTILIZAÇÃO DO CRYSTAL BALL® NO POSTO E.....	98
APÊNDICE K - QUADRO COM ERROS MAPE OBTIDOS REFERENTES À APLICAÇÃO DE MÉTODOS CAUSAIS, EM DADOS DIÁRIOS, COM UTILIZAÇÃO DO CRYSTAL BALL® EM TODOS OS COMBUSTÍVEIS DE CADA POSTO.....	99
APÊNDICE L - QUADRO COM ERROS MAPE OBTIDOS REFERENTES À APLICAÇÃO DE MÉTODOS CAUSAIS, EM DADOS SEMANAIS, COM A UTILIZAÇÃO DO CRYSTAL BALL® EM TODOS OS COMBUSTÍVEIS DE CADA POSTO	100
APÊNDICE M - QUADRO COM O DESVIO PADRÃO CALCULADO NAS APLICAÇÕES EM REDES NEURAS ARTIFICIAIS, EM CADA COMBUSTÍVEL, DE CADA POSTO.....	101
APÊNDICE N - QUADRO REFERENTE AOS MELHORES RESULTADOS ALCANÇADOS, DIÁRIOS E SEMANAIS, DE CADA UM DOS MÉTODOS UTILIZADOS (SÉRIES TEMPORAIS, MÉTODOS CAUSAIS E REDES NEURAS ARTIFICIAIS) NO POSTO A	102
APÊNDICE O - QUADRO REFERENTE AOS MELHORES RESULTADOS ALCANÇADOS, DIÁRIOS E SEMANAIS, DE CADA UM DOS MÉTODOS UTILIZADOS (SÉRIES TEMPORAIS, MÉTODOS CAUSAIS E REDES NEURAS ARTIFICIAIS) NO POSTO B.....	104
APÊNDICE P - QUADRO REFERENTE AOS MELHORES RESULTADOS ALCANÇADOS, DIÁRIOS E SEMANAIS, DE CADA UM DOS MÉTODOS UTILIZADOS (SÉRIES TEMPORAIS, MÉTODOS CAUSAIS E REDES NEURAS ARTIFICIAIS) NO POSTO C	106
APÊNDICE Q - QUADRO REFERENTE AOS MELHORES RESULTADOS ALCANÇADOS, DIÁRIOS E SEMANAIS, DE CADA UM DOS MÉTODOS UTILIZADOS (SÉRIES TEMPORAIS, MÉTODOS CAUSAIS E REDES NEURAS ARTIFICIAIS) NO POSTO D	108
APÊNDICE R - QUADRO REFERENTE AOS MELHORES RESULTADOS ALCANÇADOS, DIÁRIOS E SEMANAIS, DE CADA UM DOS MÉTODOS	

UTILIZADOS (SÉRIES TEMPORAIS, MÉTODOS CAUSAIS E REDES NEURAIAS ARTIFICIAIS) NO POSTO E.....	110
--	------------

1 INTRODUÇÃO

A previsão de demandas em uma empresa se faz necessária para que a mesma possa conhecer melhor o mercado e estar apta para possíveis mudanças, sem que assim possa sofrer complicações prejudiciais ao seu funcionamento. Tubino (2007) diz que previsão de demanda é a possibilidade dos administradores preverem o futuro e assim planejarem adequadamente suas ações. No mercado atual, aquele que compreende o cenário no qual atua, consegue se tornar mais flexível e preparado para agir em diferentes situações.

A aplicação de uma previsão de demandas, ainda deve levar em consideração fatores internos da empresa, já que a mesma é responsável por moldar a maneira de trabalho, alterando toda uma filosofia de exercício.

Uma boa previsão de demanda auxilia tanto nas tomadas de decisões quanto na posição da empresa frente ao mercado, que se torna mais competitiva uma vez que conhece melhor as necessidades de seus clientes. Contudo, para que uma previsão de demandas seja positiva e satisfatória, é requerida uma boa análise das diversas ferramentas e métodos para tal.

Neste trabalho foram adotados modelos quantitativos para fazerem essas previsões, caracterizando-se pela análise que fazem sobre um conjunto de dados e informações de um histórico. Logo, se entende que foram adotados modelos matemáticos capazes de gerar resultados a partir de comportamentos passados, que auxiliariam nas tomadas de decisões citadas anteriormente.

Dentre elas, podemos citar as séries temporais, que segundo Ragsdale (2017), se baseiam em valores anteriores para estimar valores futuros, envolvendo Média Móvel Simples, Dupla e Ponderada, Suavização Exponencial Simples, com tendência e com tendência e sazonalidade.

No que diz respeito aos métodos causais, sua importância se deve justamente por ampliar o conhecimento da empresa em relação ao andamento de suas ações e atividades ao longo dos anos, e com isso tendo base para conhecer quais suas principais variáveis independentes que são responsáveis por influenciar o andamento do sistema e como as mesmas se relacionam com as variáveis dependentes. Como afirma Ragsdale (2017, p.351), “O objetivo na análise de regressão é identificar uma função que descreva o mais aproximado possível, a relação entre essas variáveis”. Pode-se destacar a utilização da Análise de Regressão Linear Simples e Múltipla.

Em paralelo a isso, as redes neurais se tornam ainda mais interessantes como meio de usufruir das informações levantadas e buscar padrões de comportamento entre as variáveis do sistema, obtendo assim conhecimento e automatizando ideias usando como referência as variáveis definidas anteriormente. A RNA (REDE NEURAL ARTIFICIAL) é um modelo que se assemelha com estruturas neurais biológicas e que tem capacidade computacional adquirida por meio do aprendizado e generalização, que melhora seu desempenho à medida que interage com o ambiente. (BRAGA, CARVALHO; LUDEMIR, 2000).

Esses métodos quando comparados, permitem a empresa destacar dentre elas qual aquela que venha a se tornar mais eficiente no que diz respeito à previsão de demandas. Com o método ideal encontrado, a rede de postos de combustíveis pode ser capaz de implementar uma previsão eficiente, já conhecendo as características e funcionalidades da mesma. Conseguindo assim avaliar melhor seus custos e ter maior controle de suas ações, assim como a capacidade de analisar melhor os investimentos necessários no dia-a-dia.

1.1 SITUAÇÃO PROBLEMA

Sendo assim, este trabalho tem como principal problema a ser solucionado, encontrar o melhor método de previsão de demanda, dentre os estudados (séries temporais, métodos causais e redes neurais), para a previsão de demanda em uma rede de postos de combustíveis, situada em Belém, a partir da análise de dois tipos de comportamento (diário e semanal).

1.2 OBJETIVOS DO ESTUDO

1.2.1 Objetivo Geral

Dito isso, o objetivo geral deste trabalho é o de avaliar por meio da aplicação de séries temporais, métodos causais e redes neurais artificiais, como os mesmos podem contribuir para a previsão de demanda de uma rede de postos de combustíveis.

1.2.2 Objetivos Específicos

Para tal, alguns objetivos específicos foram determinados para alcançar esse objetivo geral, tais como:

- a) Analisar e definir os métodos mais relevantes envolvendo séries temporais e métodos causais.
- b) Analisar e definir a estrutura de redes neurais que obteve os melhores resultados.
- c) Analisar comparativamente, de maneira geral, todos os resultados obtidos.

1.3 JUSTIFICATIVA DO TRABALHO

O tema deste respectivo trabalho de conclusão do curso, voltado para previsão de demanda, foi escolhido pela sua grande importância no que diz respeito à tomada de decisões dentro de uma empresa, o que auxilia na redução de estoque e conseqüentemente na redução dos custos envolvidos. Logo, é necessário conhecer métodos que proporcionem atingir tal resultado, e mais do que isso, identificar como os mesmos podem ser implementados numa empresa.

Sendo assim, o problema deste trabalho, que circunda em conhecer qual o melhor método quantitativo de previsão vai contribuir para uma rede de postos de combustíveis, se torna de extrema importância, uma vez que visa demonstrar a aplicação de cada método e observar como uma empresa poderia ter um maior controle sobre suas vendas, caso tivesse uma boa previsão de demandas.

Com os resultados alcançados, espera-se obter benefícios que vão além do conhecimento acerca do desempenho de cada um. O ideal almejado é o de estabelecer um método que num futuro possa servir de base para a previsão de demandas em uma rede de postos de combustíveis.

Os maiores beneficiários com a conclusão deste respectivo trabalho podem ser classificados em: os autores, que irão expandir seus conhecimentos acerca do tema e de todos os métodos e ferramentas aqui discutidos; os proprietários da rede de postos de combustíveis, que em algum momento, poderão ter tais métodos de previsão implementados, já com uma boa base acerca de seus resultados; e todos aqueles que tiverem acesso à esse presente trabalho para que possam utilizar o mesmo como base para estudos e/ou modelo de trabalho de conclusão de curso.

Contudo, este trabalho se limita apenas a demonstrar as aplicações diárias e semanais de cada ferramenta de previsão de demanda citadas anteriormente, assim como as análises comparativas entre as mesmas. Em nenhuma hipótese serão criados ou inventados novos

métodos de previsão de demanda pelos autores, assim como, os dados coletados ou atingidos futuramente.

Por fim, a pesquisa aqui realizada para possível posterior aplicação e acompanhamento de análises, servirá de demonstração prática acerca de quais ferramentas ou metodologias que realizam previsão de demandas, obtém melhores resultados e eficácia numa empresa que pretende fazer uso de tal previsão.

1.4 ESTRUTURA DO TRABALHO

Neste contexto, este trabalho visa conhecer melhor a demanda de uma rede de postos de combustíveis, situado na cidade de Belém do Pará, onde neste primeiro momento tem-se a introdução já comentada, assim como seu respectivo tema, problema, objetivos gerais, objetivos específicos e justificativa. A segunda parte deste trabalho, é composta pelo referencial teórico que servirá de base para o estudo. No terceiro tópico, será abordada a metodologia que será utilizada, assim como o processo de obtenção de todos os dados necessários para elaboração do estudo. E o lócus da empresa e suas características. Na quarta etapa, serão modelados todos os dados e aplicados todos os métodos previamente comentados. Já na quinta parte, os resultados obtidos serão analisados. Por fim, a sexta parte consiste nas considerações finais e respectivas referências.

2 REFERENCIAL TEÓRICO

2.1 PREVISÃO DE DEMANDA

Segundo Ritzman e Krajewski (2008), previsão é uma avaliação de eventos futuros para fins de planejamento. Sendo a mesma necessária para auxiliar a determinar que recursos são necessários, na programação dos recursos existentes e na aquisição de recursos adicionais.

Para Lustosa et al (2008), demanda é a disposição dos clientes ao consumo de bens e serviços de uma organização. Sendo influenciada por diversos fatores que se estendem desde condições macroeconômica até operacionais, como disponibilidade do produto e preço no ponto de venda.

Conforme Moreira (2006), a previsão de demanda é um processo que busca informações acerca do valor das vendas futuras de um item ou de um conjunto de itens.

“As previsões têm uma função muito importante nos processos de planejamento dos sistemas de produção, pois permitem que os administradores destes sistemas antevejam o futuro e planejem adequadamente suas ações” (TUBINO, 2007, p. 15).

Constata-se que previsão de demanda é prever as necessidades de um determinado bem ou serviço a fim de se definir um planejamento que consiga suprir as mesmas.

2.2 PADRÃO DE DEMANDA

Segundo Ritzman e Krajewski (2008) as séries temporais são fruto de várias observações constantes sobre a demanda, que formam um padrão conhecido por essa nomenclatura. Os padrões básicos podem ser vistos no Quadro 1 e no Quadro 2:

Quadro 1 - Exemplos de tipos de padrões de demanda por diferentes autores

Nº de padrões de demanda do autor	Ritzman e Krajewski (2008)	Davis, Aquilano e Chase (2008)	Lustosa et al (2008)
1	<u>Horizontal</u> : flutuação de dados em torno de uma média constante;	Demanda média para o período;	Estacionária;
2	<u>Tendência</u> : aumento ou diminuição, de maneira sistêmica, na média das séries ao longo do tempo;	Tendência;	Com tendência;

Fonte: Autores (2018)

Quadro 2 - Exemplos de tipos de padrões de demanda por diferentes autores (Continuação)

Nº de padrões de demanda do autor	Ritzman e Krajewski (2008)	Davis, Aquilano e Chase (2008)	Lustosa et al (2008)
3	<u>Sazonal</u> : repetição de aumento ou diminuição da demanda, o que vai depender da hora do dia, da semana, mês ou estação;	Influência sazonal;	Estacionária sazonal;
4	<u>Cíclico</u> : aumento ou diminuição gradual de demandas menos previsíveis em períodos de maior tempo (anos ou décadas);	Elementos cíclicos;	Com tendência e sazonalidade.
5	<u>Aleatório</u> : variação da demanda que não pode ser prevista.	Variação aleatória.	

Fonte: Autores (2018)

2.3 MÉTODOS QUANTITATIVOS

Para Lustosa et al (2008) e Moreira (2006), caracterizam-se pelo emprego de modelos matemáticos para obtenção de previsões.

Ainda segundo Moreira (2006), os métodos matemáticos (quantitativos) subdividem-se em métodos causais, onde a demanda de um ou vários itens é relacionada a uma ou mais variáveis internas ou externas a empresa; e séries temporais, tendo como expectativa da observação do padrão dos valores passados que gerem informações adequadas para a previsão de valores futuros da demanda.

2.3.1 Séries temporais

As séries temporais são conhecidas por utilizarem como base para suas previsões demandas passadas de um determinado produto ou serviço, ou seja, suas previsões são feitas a partir do passado, da demanda histórica que aquele determinado produto ou serviço possui e então espera-se que o mesmo padrão se mantenha para o futuro.

Ritzman e Krajewski (2008) dizem que as séries temporais falam a respeito somente da variável dependente. Baseando-se na hipótese de que o padrão anterior da variável dependente continuará futuramente.

"Pressupõem-se que o padrão de demanda observado no passado deve repetir-se no futuro e, com base nessa premissa, são feitas as novas previsões" (LUSTOSA et al, 2008, p. 59). Moreira (2006), Tubino (2007), Ragsdale (2017) e Davis, Aquilano e Chase (2001) também dizem que os valores futuros das séries podem ser estimados com base nos valores passados.

2.3.1.1 Média móvel simples

Conforme Lustosa et al (2008), é o método mais simples que vem à mente, baseado em prever que a demanda do período atual seria igual ao período anterior.

Segundo Ritzman e Krajewski (2008), o mesmo é empregado para estimar a média de uma série temporal de demanda, eliminando os efeitos da flutuação aleatória.

De acordo com Davis, Aquilano e Chase (2001), se a demanda de um produto não cresce ou decresce rapidamente e não apresenta característica sazonal, uma média móvel simples pode ser útil na determinação de uma tendência dentro da flutuação dos dados. Sendo a formula da mesma a equação 1:

$$F_t = \frac{A_{t-1} + A_{t-2} \dots + A_{t-n}}{n} \quad (1)$$

Onde:

F_t = Vendas previstas no período t;

A_{t-1} = Vendas realizadas no período t-1;

n = Número de períodos considerados na média.

"Ele é de grande utilidade quando a demanda não possui influências sazonais ou de tendências acentuadas" (RITSMAN; KRAJEWSKI, 2008, p. 271).

Pode-se concluir que a simplicidade deste método limita sua utilização em vários casos, porém, é o método mais simples utilizado, onde o mesmo apenas tenta imitar a média dos últimos valores de uma demanda para torná-la em previsão.

2.3.1.2 Média móvel dupla

Para Corrar e Theófilo (2004), conforme citado por Sousa e Guimarães (2010), o método é aplicável a séries de dados que apresentam tendência, logo, algum tipo de crescimento linear ou exponencial ao longo do tempo. Primeiramente é necessário calcular a média móvel simples como apresentado na equação 1 anteriormente. Sendo utilizada posteriormente na fórmula que calcula a média das médias móveis, como mostra a equação 2. Depois, deve-se calcular a diferença entre o dobro da média móvel e a média móvel das médias moveis, segundo a equação 3. Como o método é sensível à tendência existente na série de dados, necessita-se adicionar um fator de tendência, como pode ser visto na equação 4. Ao fim, a previsão é então calculada por meio da equação 5.

$$F' = \frac{F_{t-1} + F_{t-2} \dots F_{t-n}}{n} \quad (2)$$

$$a_t = F_t + (F_t - F'_t) = 2 * F_t - F'_t \quad (3)$$

$$b_t = \frac{2}{n-1} * (F_t - F'_t) \quad (4)$$

$$P_{t+p} = a_t + b_{tp} \quad (5)$$

Onde:

P_{t+1} = Vendas previstas para o próximo período;

n = Número de períodos utilizados;

F_t = Média móvel do período t ;

F' = Média móvel das médias móveis.

2.3.1.3 Média móvel ponderada

Ritzman e Krajewski (2008) e Moreira (2006) dizem que a vantagem desse método é que ele permite enfatizar a demanda recente em vez da antiga, propiciando uma resposta mais rápida do que a previsão da média móvel simples.

Davis, Aquilano e Chase (2001) ressaltam ainda que a média móvel simples atribui pesos iguais para cada componente da série de dados, já a média móvel ponderada permite que cada um dos elementos seja ponderado por um fator, sendo a soma destes, iguais a um. Sendo sua fórmula a equação 6:

$$F_t = \frac{w_{t-1}A_{t-1} + w_{t-2}A_{t-2} \dots + w_{t-n}A_{t-n}}{n} \quad (6)$$

Onde:

F_t = Vendas previstas no período t;

A_{t-1} = Vendas realizadas no período t-1;

w_{t-1} = Peso atribuído ao período t-1;

n = Número de períodos considerados na média.

A média móvel ponderada acaba sendo uma adaptação da média móvel simples. Logo, por conta de seus pesos, sua aplicabilidade se torna maior por aceitar que certos valores possuem uma importância maior que outros, tornando as previsões mais assertivas.

2.3.1.4 Suavização exponencial simples

"O método de suavização exponencial é um método sofisticado de média móvel ponderada que calcula a média de uma série temporal atribuindo às demandas recentes maior peso do que as demandas iniciais" (RITSMAN; KRAJEWSKI, 2008, p. 272).

Como é colocado por Lustosa et al (2008), neste método pressupõe-se que a demanda oscila em torno de um patamar ou demanda base constante. Começando por um valor inicial, a "base" é corrigida no passar dos períodos, conforme os dados são adicionados na série. Isso pode ser constatado por meio das equações 7 e 8:

$$B_t = \alpha * D_t + (1 - \alpha) * B_{t-1} \quad (7)$$

$$F_t(t + k) = B_t \quad (8)$$

Onde:

D_t = demanda do período t;

B_t = base ao final do instante t;

T_t = tendência ao final do instante t;

α = constante de suavização para Base;

$F_t(u)$ = previsão ao final do período t para o período u ($u > t$);

k = número de períodos adiante.

2.3.1.5 Suavização exponencial com tendência (*Holt*)

Como falam Ritzman e Krajewski (2008), tendência é um aumento ou diminuição na média da série ao longo do tempo. Logo os métodos de suavização exponencial precisam ser modificados para que as previsões não fiquem acima nem abaixo da demanda real.

Lustosa et al (2008) coloca uma segunda variável que reflete o crescimento de demandas entre períodos. Sendo essa variável atualizada exponencialmente e aplicada no cálculo de previsão conforme as equações 9, 10 e 11 a seguir:

$$B_t = \alpha * D_t + (1 - \alpha) * (B_{t-1} + T_{t-1}) \quad (9)$$

$$T_t = \beta(B_t - B_{t-1}) + (1 - \beta) * T_{t-1} \quad (10)$$

$$F_t(t + k) = B_t + kT_t \quad (11)$$

Onde:

D_t = demanda do período t;

B_t = base ao final do instante t;

T_t = tendência ao final do instante t;

α = constante de suavização para Base;

β = constante de suavização para Tendência;

$F_t(u)$ = previsão ao final do período t para o período u ($u > t$);

k = número de períodos adiante.

A equação 9 calcula a média ponderada entre a nova base e a demanda real, que incorpora uma parcela de crescimento ou redução da demanda expressa na variável tendência. Na equação 10, tem-se a suavização da tendência, calculada na variação da base dos dois últimos períodos e a estimativa anterior. Por último a equação 11 fornecendo a previsão de demanda para períodos adiante "k", segundo progressão linear.

2.3.1.6 Suavização exponencial com tendência e sazonalidade (*Holt-Winters*)

Segundo Ritzman e Krajewski (2008), padrões sazonais são repetições constantes para cima ou para baixo com intervalos inferiores a um ano.

Para Lustosa et al (2008), o modelo consiste em projetar a demanda da base, extraindo o efeito de sazonalidade, representada pelo índice de sazonalidade (I_t). A equação 12 mostra

o cálculo final da base no período t. A equação 13 permite o cálculo da tendência. Na equação 14 tem-se o uso do índice de sazonalidade, onde terá uma média ponderada entre o real observado e o índice anterior. A constante L representa a extensão do ciclo sazonal. Finalizando, para obter uma previsão de demanda para um período futuro, multiplica-se a projeção da demanda da base pelo índice de sazonalidade correspondente, segundo a equação 15.

$$B_t = \alpha * \left(\frac{D_t}{I_{t-L}} \right) + (1 - \alpha) * (B_{t-1} + T_{t-1}) \quad (12)$$

$$T_t = \beta(B_t - B_{t-1}) + (1 - \beta) * T_{t-1} \quad (13)$$

$$I_t = \gamma * \left(\frac{D_t}{B_t} \right) + (1 - \gamma) * I_{t-1} \quad (14)$$

$$F_t(t + k) = (B_t + kT_t) * I_{t-L+k} \quad (15)$$

Onde:

D_t = demanda do período t;

B_t = base ao final do instante t;

T_t = tendência ao final do instante t;

I_t = índice de sazonalidade do instante t;

α = constante de suavização para Base;

β = constante de suavização para Tendência;

γ =constante de suavização para Sazonalidade;

$F_t(u)$ = previsão ao final do período t para o período u ($u > t$);

k = número de períodos adiante.

2.3.2 Métodos Causais

Segundo Davis, Aquilano, Chase (2001), as series causais tem como foco todo e qualquer fator associado ao produto final que possa afetar a sua venda, tais como; a qualidade percebida, *marketing*, localização, entre outros. Sendo assim responsável por observar dentre um conjunto de dados, quais são as variáveis que influenciam diretamente ou indiretamente nos resultados de uma empresa.

2.3.2.1 Análise de Regressão Linear Simples

De acordo com Davis, Aquilano, Chase (2001), a análise de regressão linear tem como função definir o relacionamento entre duas ou mais variáveis que estejam correlacionadas. Geralmente, tal relacionamento entre uma variável independente e variável dependente é considerado como sendo de uma linha reta.

Para Corrar, Paulo e José Filho (2007), o pesquisador é quem classifica as variáveis, sendo normalmente, as variáveis dependentes não-passíveis de controle e as variáveis independentes podendo serem controladas.

Segundo Ritzman e Krajewski (2004), a ideia é encontrar valores de a e b que possam minimizar a somatória dos desvios quadrados dos dados reais da linha reta demonstrada em gráfico, como demonstra a equação 16 adiante:

$$Y = a + bX \quad (16)$$

Onde:

Y = Variável dependente;

a = Interseção no eixo Y ;

b = Inclinação;

X = Variável independente (representa a unidade de tempo).

Para Ragsdale (2017), um exemplo simples seria a relação entre vendas para uma empresa e o custo da mesma com propaganda, sendo as vendas a variável dependente e a propaganda como variável independente. Tem-se na equação 17 de uma linha reta a seguir, uma equação apropriada que demonstra tal relação entre venda e propaganda:

$$Y_i = \beta_0 + \beta_1 x_1 + \varepsilon_i \quad (17)$$

Onde:

Y_i = Variável real de venda;

x_1 = Investimento em propaganda;

β_0 = Representa um valor constante;

β_1 = Inclinação da linha;

ε = Representa o resíduo ou erro da regressão.

De acordo com Davis, Aquilano, Chase (2001, p.228):

Este método de previsão de demanda é útil para previsões de longo prazo na maioria das ocorrências e para planejamento agregado. Por exemplo, a regressão linear seria muito útil para produtos individuais, dentro de uma família de produtos. Mesmo imaginando que a demanda para produtos individuais, dentro de uma família, possa variar amplamente ao longo de um período de tempo, a demanda para a família total de produtos é surpreendentemente suave.

Conforme Davis, Aquilano, Chase (2001), a maior limitação nessa técnica de previsão é a de considerar tanto os dados do passado como as previsões futuras se situando próximos a uma linha reta. Podendo ser maior eficaz se considerado um período menor de tempo.

Segundo Davis, Aquilano, Chase (2001), a regressão linear serve tanto para previsão de séries temporais, quando a variável Y muda como resultado da variável X, como para previsão de relacionamentos causais, quando a variável dependente sofre alteração devido a mudanças ocorridas na variável independente.

Para Ritzman e Krajewski (2004), existem certas medidas que precisam ser analisadas para que a regressão seja verificada. Tais medidas se dividem em: coeficiente de correlação, o coeficiente de determinação de amostra e o erro-padrão da estimativa.

Ainda segundo Ritzman e Krajewski (2004), o coeficiente de correlação é responsável por medir a quantidade da variação da variável dependente Y em relação a sua média, explicada pela linha de regressão. Como afirmam Martins e Laugeni (2005), o valor de r varia entre +1 e -1, sendo valores positivos indicados por retas ascendentes e valores negativos indicados por retas descendentes. Logo, quando mais próximo de +1 se encontra o valor do coeficiente de correlação, melhor será a linha de regressão.

Conforme Davis, Aquilano, Chase (2001), o coeficiente de determinação vai medir a quantidade de variação da variável dependente como resultado das alterações da variável independente, e é calculada como mostra a equação 18 a seguir:

$$r^2 = \frac{\sum(y_i - \bar{Y})^2 - \sum(y_i - \hat{y}_i)^2}{\sum(y_i - \bar{Y})^2} \quad (18)$$

Onde:

r^2 = Coeficiente de determinação;

y_i = Valor real de Y observado por um dado valor de X;

\bar{Y} = Média aritmética de todos os valores de y ;

\hat{y}_i = Valor de Y correspondente a um dado valor de X , calculado pela equação da regressão linear.

Como se pode observar, o termo $\sum (y_i - \bar{Y})^2$ se encontra tanto no numerador como no denominador na equação acima. E, de acordo com Davis, Aquilano, Chase (2001), esse termo em questão se refere a variação total da variável dependente em torno da média aritmética. Enquanto que o segundo termo $\sum (y_i - \hat{y}_i)^2$ se refere ao erro ou a variação na variável dependente sem que tenha relação com a equação de regressão.

Segundo Ritzman e Krajewski (2004), o coeficiente de determinação com valor próximo de 1 pode ser considerado como desejável, pois indica que as variações da variável dependente junto com as previsões geradas pela equação estão correlacionadas.

Ainda de acordo com Ritzman e Krajewski (2004) e Davis, Aquilano, Chase (2001), uma outra medida de confiabilidade da previsão de regressão é a chamada erro padrão da estimativa, que tem como função medir a proximidade entre os dados sobre a variável dependente e a linha de regressão. Ou seja, entender como os erros estão dispersos próximos da linha de regressão. E tem como equação 19 adiante:

$$S_{YX} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - 2}} \quad (19)$$

Onde:

S_{YX} = Erro padrão de estimativa;

Como visto, a análise de regressão linear simples é bastante recomendada quando se buscar conhecer a relação existente entre uma única variável dependente com uma variável independente, e como tal relação influencia no comportamento e resultado da outra.

2.3.2.2 Análise de Regressão Linear Múltipla

Uma outra alternativa de previsão de demanda é a chamada análise de regressão múltipla, onde se leva em consideração mais de uma variável independente. Para Corrar, Paulo e José Filho (2007), uma regressão linear múltipla permite analisar a relação que existe

entre uma única variável dependente e mais de uma variável independente. E a partir dela, fazer projeções. Logo, só se faz previsão se houver regressão.

Ainda de acordo com Corrar, Paulo e José Filho (2007), a combinação linear de variáveis independentes usadas para prever uma variável dependente é conhecida como modelo de regressão, que pode ser visto na equação 20 adiante:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon \quad (20)$$

Onde:

Y = Variável dependente;

$x_1, x_2 \dots x_n$ = Variáveis independentes;

β_0 = Coeficiente linear (representa o valor de Y quando X é igual a zero);

$\beta_1, \beta_2 \dots \beta_n$ = Parâmetros da regressão (Coeficientes angulares);

ε = Representa o resíduo ou erro da regressão;

Segundo um exemplo demonstrado por Davis, Aquilano, Chase (2001), num negócio voltado para mobílias para residências, diversos fatores contribuem para a expectativa de vendas, podendo serem expressados perfeitamente pela equação 21 de regressão múltipla, como visto a seguir:

$$S = B + B_m (M) + B_h (H) + B_i (I) + B_t (T) \quad (21)$$

Onde:

S = Venda bruta ao ano;

B = Base de venda (onde outros fatores têm influência);

M = Matrimônios durante o ano;

H = Moradias novas durante o ano;

I = Renda pessoal disponível ao ano;

T = Tendência do Tempo;

B_m, B_h, B_i, B_t = Representam a influência na expectativa de vendas em função das variáveis, respectivamente.

Ainda segundo Davis, Aquilano, Chase (2001), esse tipo de previsão de demandas é bastante recomendado quando se tem diversos fatores possíveis de influência direta em uma

determinada variável. Como visto acima no exemplo, um mercado voltado para venda de móveis pode sofrer variação positiva ou negativa dependendo de quais variáveis externas estiverem mais atuantes, que no caso se estendem para diversos fatores. Onde todos podem impactar diretamente nos resultados de vendas daquele ano ou período.

Conforme Corrar, Paulo e José Filho (2007), os modelos de regressão apresentam como pressupostos básicos:

1. Variável Y ser aleatória;
2. Média dos resíduos ser igual a zero;
3. A variância de ε é constante e igual a σ^2 ;
4. Resíduos são independentes entre si;
5. Resíduos têm distribuição normal.

Davis, Aquilano, Chase (2001), ainda afirmam que a maior complexidade na utilização de tal técnica de previsão de demanda é a da fase de coleta de dados e informações e questão que envolve muitos cálculos matemáticos.

2.3.2.3 Estimação de modelos

Segundo Corrar, Paulo e José Filho (2007, p.137):

A explicação estatística de um fenômeno da natureza fundamenta-se na análise dos erros ou resíduos, que se caracterizam pela diferença entre as observações reais e os valores estimados para cada observação da amostra, por qualquer das técnicas estatísticas utilizadas. A estimativa mais adequada será aquela que apresentar o menor resíduo estatístico, ou seja, a menor diferença entre os valores reais observados e os valores estimados pelo modelo.

Ainda segundo Corrar, Paulo e José Filho (2007), o Método dos Mínimos Quadrados (MMQ) se destaca dentre os métodos de estimação de modelos utilizados em regressão linear. Sendo seu objetivo obter a menor soma de quadrados dos resíduos (SQR) possível.

De acordo com Ragsdale (2017), várias observações devem ser feitas num SQR, já que cada erro de estimativa é elevado ao quadrado, fazendo com que seu valor sempre seja positivo e logo, seu menor valor seja 0 (zero). E para ter um valor igual a 0 (zero), é

necessário que todos os erros de estimativa individuais sejam iguais a 0 (zero), caso que faria a linha de regressão estimada ser ajustada perfeitamente.

2.4 REDES NEURAIS ARTIFICIAIS

As redes neurais artificiais representam uma tecnologia que possui raízes em diversas disciplinas, como neurociência, matemática, engenharia, física, estatística e ciência da computação. Podendo ser aplicadas em modelagem, análise de séries temporais, reconhecimento de padrões e processamento de sinais e controles. Sendo tudo isso fruto de sua importante habilidade de “aprender” a partir de conjuntos de entrada, com monitoramento ou não de um professor/supervisor. (HAYKIN, 2001).

De acordo com Ragsdale (2017), as redes neurais são consideradas uma técnica que tem por finalidade reconhecer padrões, quando existentes, entre conjuntos de variáveis de entrada e saída. E sua característica de aprendizado tem relação com a forma com que seus dispositivos foram originados, cujo ideal era de funcionamento baseado na maneira como o cérebro humano aprende.

Ainda segundo Ragsdale (2017, p.423):

O processamento de estímulos no cérebro ocorre por meio de um conjunto de neurônios massivamente interconectados que respondem aos estímulos com sinais de saída de intensidades diferentes (excitados ou inibidos por neurotransmissores químicos), que são então enviados através de sinapses para outros neurônios.

Logo, pode-se entender que as redes neurais artificiais são programas que funcionam “imitando” a maneira como o cérebro humano atua no corpo, por mais complexo que se possa imaginar. Por isso, tais redes são conhecidas por reconhecerem padrões e “aprenderem”, similar ao que o cérebro é capaz.

Conforme afirmam Davis, Aquilano, Chase (2001), as redes neurais são uma técnica de previsão de demandas ainda recentes e se diferenciam das demais por focarem no aprendizado humano, que consiste em “ensinar” para a máquina padrões de comportamento nas variáveis de entrada e saída de um modelo de previsão. Em um serviço, por exemplo, as entradas seriam as vendas anteriores, horas de movimento, dias da semana mais agitados e

meses com maiores demandas. Enquanto que as saídas seriam a quantidade de clientes esperados em um determinado dia, ou mês.

Logo, percebe-se que a ideia por trás das redes neurais é a de observar e identificar a função que melhor mapeie os conjuntos de variáveis de entrada e seus correspondentes de saída. E por mais que se assemelhe muito com outras técnicas de previsão, ela consegue ter maior destaque dentre as demais existentes, como afirma Ragsdale (2017), que diz que a Análise de Regressão, por exemplo, exige que a relação dentre os conjuntos de variáveis tenha uma forma funcional já definida pelo analista. Enquanto que as redes neurais fazem a maior parte do trabalho, já que definem o tipo de relação, através dos dados, sem que o analista se preocupe com isso.

Ainda segundo Davis, Aquilano, Chase (2001), tais redes neurais são capazes de realizar cálculos de forma muito mais rápida que os métodos anteriores. Além de poderem ser divididas em duas classes, àquelas que são “treinadas” por dados históricos e às que observam e identificam padrões em conjuntos de dados.

2.4.1 Modelo MCP de neurônio artificial

Segundo Braga, Carvalho e Ludemir (2000), o modelo de McCulloch e Pitts é uma representação de um modelo neural simplificado (MCP). No qual apresenta valores de entrada $X = x_1, x_2, x_3, \dots, x_n$, com tamanho n . Cada entrada x possui pesos correspondentes w . A soma ponderada de todas as entradas com os seus pesos resulta uma saída linear u . Enquanto que a saída do neurônio y recebe o nome de saída de ativação, e é obtida através da aplicação de uma função de ativação $f(.)$ na saída u . Tanto a saída linear quanto a saída de ativação podem ser melhor visualizadas nas equações 22 e 23 adiante:

$$u = \sum_i x_i w_i \quad (22)$$

Onde:

u = Saída linear;

x_i = Valores de entrada;

w_i = Pesos correspondentes.

$$y = f(\hat{u}) \quad (23)$$

Onde:

y = Saída de ativação;

$f(\hat{u})$ = Função de ativação.

Ainda de acordo com Braga, Carvalho e Ludemir (2000), existem várias formas de função de ativação, sendo geralmente não-lineares.

2.4.2 Rede neural do tipo *feedforward*

Segundo Braga, Carvalho e Ludemir (2000), os elementos antes demonstrados no modelo MCP são considerados do tipo simples. Porém, as redes neurais têm capacidade de resolver problemas mais complexos, sendo capaz de solucionar situações de regressão, classificação e predição.

De acordo com Ragsdale (2017), a função que mapeia um registro de entrada em relação com as de saída, recebe o nome de *feedforward*, e pode ser visualizada pela equação 24 a seguir:

$$y_i = f(x_{i1}, x_{i2}, \dots, x_{ip}) \quad (24)$$

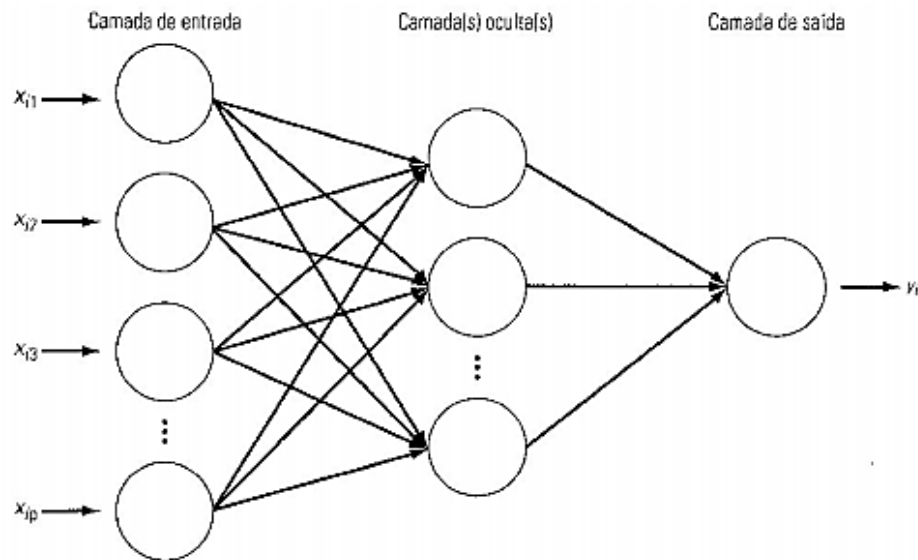
Onde:

$f()$ = Função de mapeamento;

x_{ip} = Registro de entrada;

y_i = Valor de saída.

Figura 1 - Representação de uma rede neural



Fonte: Ragsdale (2017, p. 423)

Pode-se observar na Figura 1, que cada valor de entrada é enviado para cada nó de entrada (camada oculta). Sendo que cada nó calcula sua entrada de rede N_i através da soma ponderada de cada valor de entrada, como mostra a equação 25 adiante:

$$N_i = b_i + \sum_k a_{ik} x_{ik} \quad (25)$$

Onde:

N_i = Entrada de rede;

b_i = Constante representando um valor de viés para o nó oculto i ;

i = Nó oculto.

Segundo Ragsdale (2017), cada nó oculto recebe uma resposta R_i de forma sigmoidal. Tal resposta então é enviada para para cada um dos demais nós existentes na camada oculta, onde novamente será calculado as entradas e respostas, respectivamente. Pode ser representado da seguinte maneira, como mostra a equação 26 a seguir:

$$R_i = (1 + EXP(- N_i))^{-1} \quad (26)$$

Onde:

R_i = Resposta do nó oculto.

Conforme Ragsdale (2017), para finalizar, o nó de saída é responsável por calcular sua saída y_i da mesma forma que a entrada fez anteriormente, se baseando na mesma equação utilizada na soma ponderada para o valor de entrada em cada nó, com a diferença de que nesse momento se utiliza a resposta de todos os nós, como visto pela equação 27:

$$y_i = b_i + \sum_k a_{ik} R_{ik} \quad (27)$$

Onde:

y_i = Saída de rede;

b_i = Constante representando um valor de viés para o nó oculto i ;

i = Nó oculto;

Ragsdale (2017) e Rezende (2005), ainda afirmam, que todas as camadas, seja de entrada, oculta e saída, podem possuir uma quantidade aleatória de nós presentes em cada uma. Assim como a camada oculta em particular pode ter mais de uma camada, o que deixa a rede neural não-linear. O analista é quem define quantas camadas ocultas e quantos nós em cada uma delas sua rede terá. Como não existe padrão certo de maior acerto, é necessário confiar na tentativa e erro até se achar a quantidade certa a se utilizar, uma vez que os números de camadas causam impactos diretamente no nível de complexidade da rede. Assim como o número de nós podem impactar nos ajustes dos dados e na modelagem de relações. O mesmo serve para o número de saídas, que podem ser definidos a partir da classificação de quantos grupos serão levados em conta.

Pode-se concluir, que as redes do tipo *feedforward* se tornam de muita importância uma vez que não apresentam conexões entre neurônios de forma cíclica, mas sim de forma contínua do início até o fim da rede.

2.4.3 Processos de aprendizagem

De acordo com Ragsdale (2017), a aprendizagem de uma rede neural artificial se dá através dos ajustes dos pesos para cada conjunto de entradas, de forma com que o valor de saída estimado da rede se aproxime da saída real. Para tal, é necessário um treinamento na rede neural que mostre conjuntos de dados com valores conhecidos e emparelhados, ajustando os pesos para então ser capaz de reduzir os erros na previsão da rede. Porém, a capacidade de

generalização da rede neural pode ser comprometida quando há excesso de treinamento (*overfitting*) ou falta (*underfitting*).

Pode-se afirmar então, que o principal objetivo do treinamento é o de encontrar o ajuste ideal na fronteira entre *overfitting* e *underfitting*, uma vez que os pesos são a “chave” para o funcionamento de toda a rede neural artificial, já que eles permitem a RNA identificar padrões de comportamento.

Segundo Braga, Carvalho e Ludemir (2000), o que determina a eficiência de um modelo de rede neural e o momento ideal para término do seu treinamento são os parâmetros pré-definidos do treinamento, que podem ter relação com o valor de erro quadrático médio das respostas de um determinado conjunto de dados inseridos na RNA. A generalização de uma rede neural artificial está relacionada com a capacidade da mesma em responder de forma coerente dados ainda não conhecidos por ela durante a fase de treinamento. Os fatores responsáveis por afetarem a generalização, têm sido estudados para que novas técnicas de aprendizado consigam ter controle sobre a generalização.

De acordo com Haykin (2001, p.75):

Aprendizagem é um processo pelo qual os parâmetros livres de uma rede neural são adaptados através de um processo de estimulação pelo ambiente no qual a rede está inserida. O tipo de aprendizagem é determinado pela maneira pelo qual a modificação dos parâmetros ocorre.

Ainda segundo Haykin (2001), os algoritmos de aprendizagem são conjuntos de regras definidas para solucionar problemas de aprendizagem. Logo, existem diversos tipos de algoritmos, cada um com vantagens próprias. E como afirma Rezende (2005), tais algoritmos se diferem na forma de ajustar seus pesos sinápticos de um neurônio. Enquanto que os paradigmas de aprendizagem são um modelo do ambiente no qual uma rede neural se relaciona. Podendo ser de três classificações: Aprendizado Supervisionado, Aprendizado não supervisionado e Aprendizado por reforço.

2.4.3.1 Aprendizado Supervisionado

De acordo com Haykin (2001) e Braga, Carvalho e Ludemir (2000), um aprendizado supervisionado é aquele que possui o acompanhamento de um professor, ou supervisor, que já

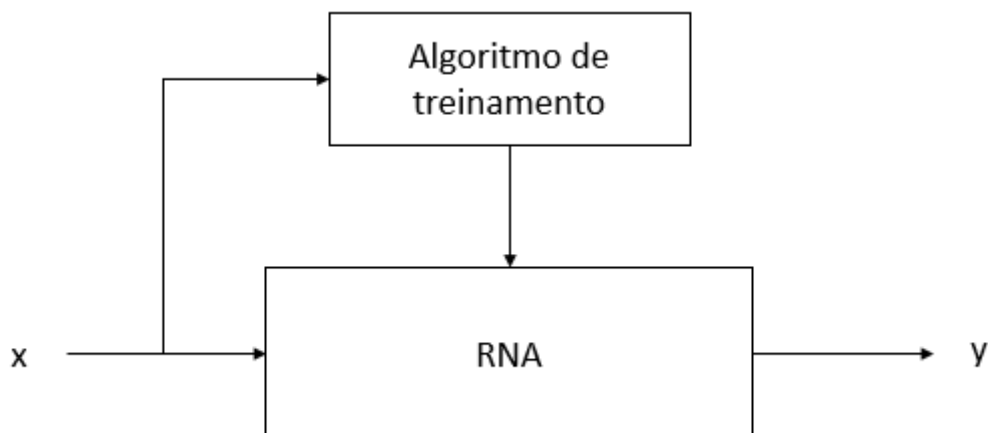
tem conhecimento prévio acerca do ambiente relacionado àquela rede neural, porém, sem que tal rede conheça esse mesmo ambiente. Esse conhecimento se refere aos valores de entrada e saída. Quando expostos em um vetor de treinamento, o supervisor fornece à rede uma resposta desejada. Quando a resposta obtida pela rede for diferente da desejada, tem-se a diferença pelo sinal de erro. Este ajuste vai sendo realizado, de forma iterativa, com o intuito de fazer a rede neural se comportar como um “emulador” do professor. Logo, o conhecimento acerca do ambiente vai sendo repassado do professor à rede, da maneira mais completa possível, até que o supervisor possa ser dispensado.

2.4.3.2 Aprendizado não-supervisionado

O aprendizado não-supervisionado, como o nome sugere, não possui supervisor e nem valores desejados de saída, apenas vetores de entrada. Em contrapartida, são fornecidas certas condições para desenvolver uma medida sem relação com aquilo que a rede deve aprender, sendo os parâmetros otimizados em torno desta medida. No momento em que a rede se ajusta com os dados de entrada, ela passa a poder formar internamente, representações codificadoras das características da entrada, de forma a criar novos grupos. (HAYKIN, 2001).

Como exemplo, tem-se a categorização de usuários de cartões de crédito com base nos seus perfis de compras. Onde o perfil seria os vetores de entrada. Assim, a rede iria retornar o número de categorias de usuários e quais pertencem a cada categoria. (BRAGA; CARVALHO; LUDEMIR, 2000). Pode ser melhor visualizado na Figura 2 adiante:

Figura 2 - Representação de um aprendizado não supervisionado



Fonte: Rezende (2005, p. 151)

2.4.4 Arquitetura de Rede

De acordo com Rezende (2005), a arquitetura de uma rede limita em quais problemas a mesma pode ser utilizada, uma vez que quando maior for a complexidade da situação, mais camadas devem ser necessárias para que a rede possa solucioná-la.

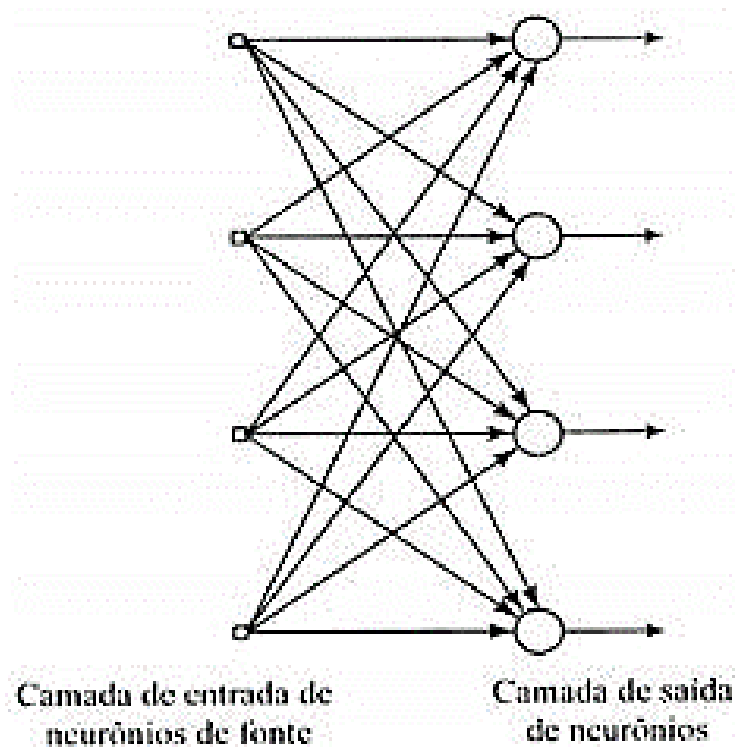
Para Corrar, Paulo e José Filho (2007), redes sem camadas ocultas conseguem apenas fazer separações lineares. Enquanto que redes com múltiplas camadas são apropriadas para resolverem problemas de mais complexos.

Os neurônios em uma rede estão estruturados conforme o algoritmo de aprendizagem utilizado no treinamento de uma rede neural artificial. E podem ser classificados em três tipos de arquiteturas distintas, Redes de camada única, Redes de múltiplas camadas e Redes recorrentes. (HAYKIN, 2001).

2.4.4.1 Redes de camada única

Segundo Haykin (2001), numa rede neural com camadas, os neurônios se estruturam na forma de camadas. Nesse tipo de rede, há uma quantidade de nós na camada de entrada se projetando para uma camada de saída. Na figura 3 se observa 4 nós na camada de entrada e outros 4 nós na camada de saída. O fato de ser chamado de “camada única”, representa a quantidade de saída de neurônios. Vale ressaltar que só há computação nas camadas de saída.

Figura 3 - Rede de Camada Única

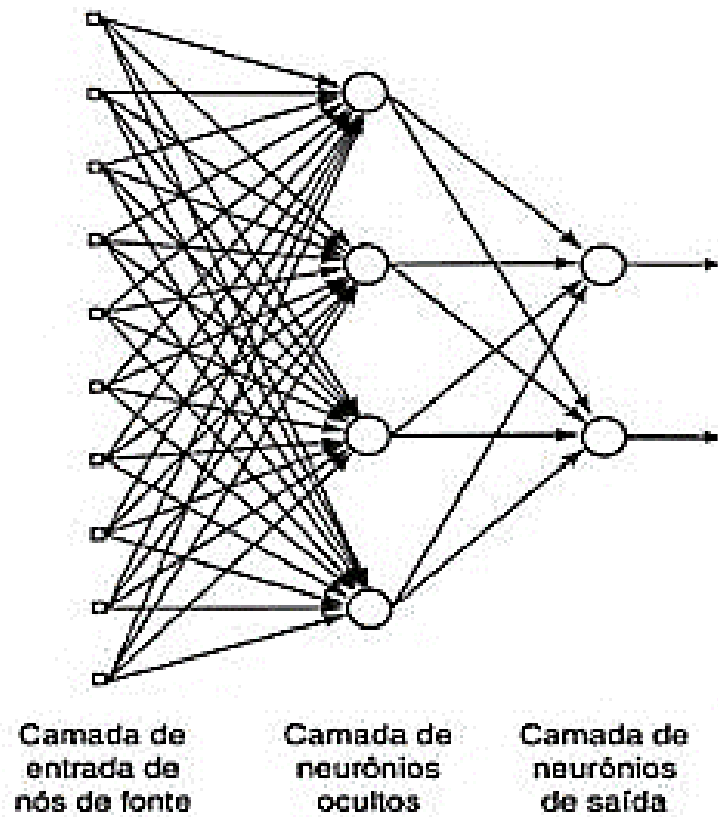


Fonte: Haykin, (2001, p. 47)

2.4.4.2 Redes de múltiplas camadas

Nas redes alimentadas com múltiplas camadas, tem-se a presença de uma ou mais camadas ocultas, cujos neurônios são chamados de nós ocultos. Sua função é a de intervir entre as camadas de entrada e de saída de forma útil. Quando mais camadas ocultas numa rede, maior será a capacidade da mesma de alcançar valores estatísticos de ordem elevada, essa habilidade é valiosa quando o tamanho da camada de entrada é grande. Na figura 4 a seguir pode-se observar uma rede de múltipla camada, onde se tem 10 nós de entrada, 4 neurônios ocultos e 2 nós de saída, sendo que todos eles estão conectados com os nós da camada adjacente. (HAYKIN, 2001).

Figura 4 - Rede de Múltiplas Camadas



Fonte: Haykin, (2001, p. 48)

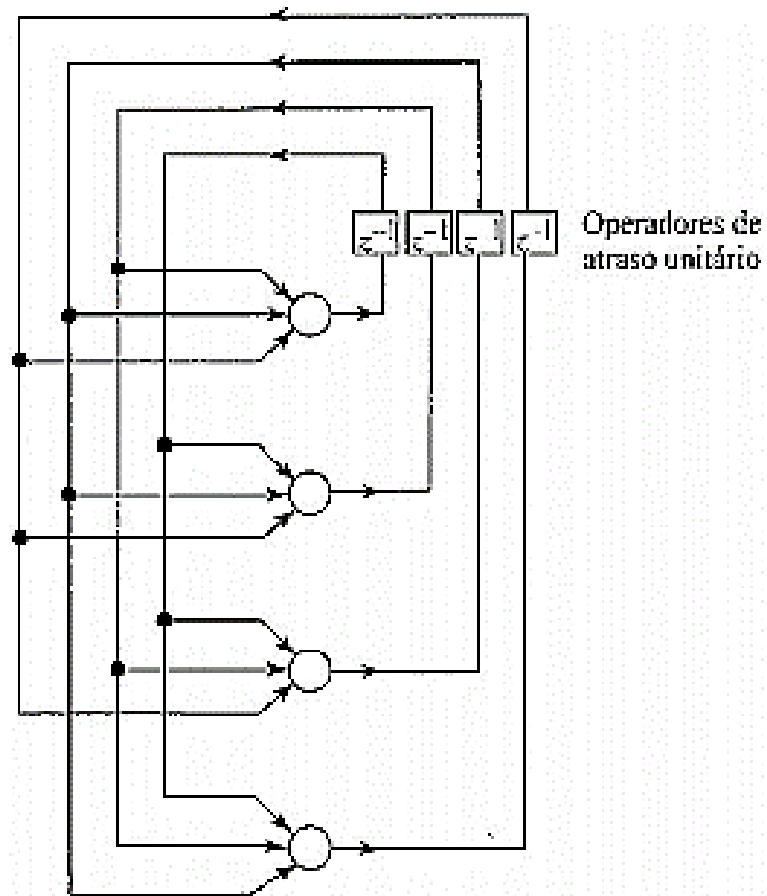
Como visto, usualmente as camadas em uma rede neural são divididas em três grupos:

- Camadas de entrada: Onde padrões são apresentados para a rede;
- Camadas ocultas ou intermediárias: Onde ocorre a maior parte do processamento, através das conexões ponderadas. São responsáveis por extrair características;
- Camadas de saída: Onde se tem a resposta final da rede.

2.4.4.3 Redes recorrentes

De acordo com Haykin (2001), as redes recorrentes se diferenciam das demais por terem, no mínimo, um laço de realimentação. Nessa rede, os neurônios podem estar isolados, mas sendo capazes de alimentar a rede através dos seus sinais de saída indo em direção as entradas de outros neurônios. Na figura 5 adiante, está ilustrado como funciona uma rede recorrente, mas sem a presença de neurônios ocultos.

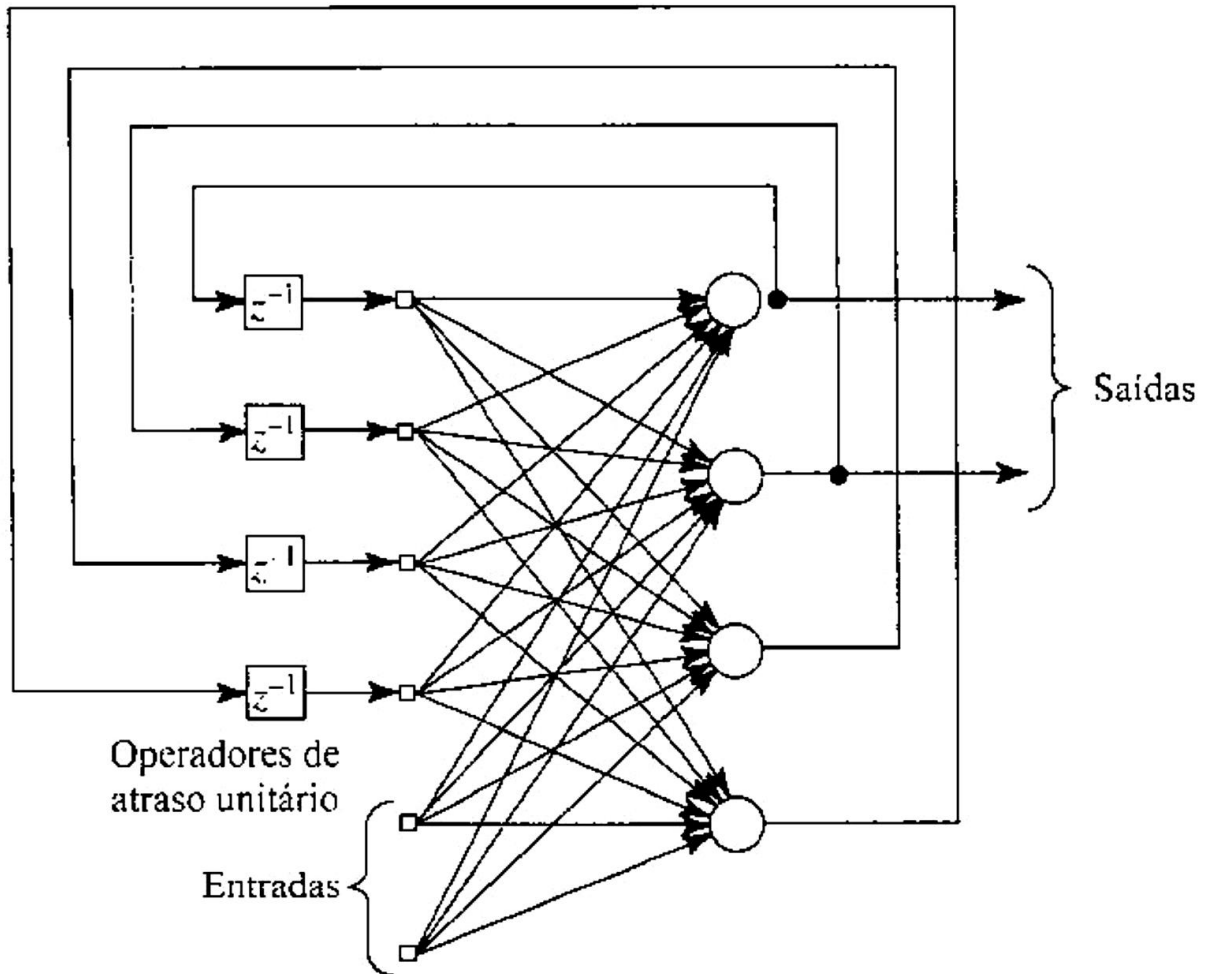
Figura 5 - Rede recorrente sem neurônios ocultos



Fonte: Haykin, (2001, p. 48)

Na figura 6 a seguir, tem-se uma representação de rede recorrente com a presença de neurônios ocultos, onde as mesmas, junto com os neurônios de saída são responsáveis pela realimentação. Esses ciclos impactam no aprendizado e no desempenho da rede. Além, de envolverem o uso de ramos particulares que resulta num comportamento dinâmico não-linear. (HAYKIN, 2001).

Figura 6 - Rede recorrente com neurônios ocultos



Fonte: Haykin, (2001, p. 49)

2.4.5 Algoritmo de Retropropagação (*backpropagation*)

Segundo Braga, Carvalho e Ludemir (2000), os modelos de redes com múltiplas camadas eram considerados de muita dificuldade de serem treinados, até que surgiu o algoritmo de retropropagação de erro (*backpropagation*), nos anos 80, desenvolvido por Rumelhart, Hinton e Williams, que viria a sanar tal dificuldade. Sendo este baseado na regra de aprendizagem por correção de erro para realizar ajustes nos pesos das camadas ocultas. A forma com que calcula as derivadas parciais dos erros de saída relacionados a cada peso correspondente é o que caracteriza esse algoritmo. De acordo com Haykin (2001), o algoritmo de retropropagação representa um marco nas redes neurais.

O treinamento feito por um algoritmo de retropropagação pode ser separado em 2 etapas. Na primeira delas, um sinal de entrada se propaga para frente (*forward*), indo da entrada até a saída da rede, camada por camada. Resultando num conjunto de saídas como

resposta real da rede. Como o valor de saída desejado para a entrada corrente já é conhecido, o erro para a camada de saída pode ser calculado, ajustando assim seus respectivos pesos. Já na segunda etapa, no processo para trás (*backward*), como não existem valores de saídas desejados para camadas intermediárias, os ajustes de seus pesos sinápticos são feitos através da propagação no sentido para trás do erro da camada de saída. A resposta real da rede é subtraída de uma resposta desejada para se produzir um sinal de erro que se propaga pela rede neural contra a direção das camadas. Daí vem o nome de retropropagação de erro (*backpropagation*). (HAYKIN, 2001; BRAGA; CARVALHO; LUDEMIR, 2000).

Conforme Rezende (2005), se considerando uma rede neural com duas camadas, a equação que demonstra o ajuste de pesos para um neurônio da camada de saída pode ser vista na equação 28:

$$w_{ji}(n + 1) = w_{ji}(n) + \eta e_j(n) f_j'(u_j(n)) x_i(n) \quad (28)$$

Onde:

$f_j'(u_j(n))$ = Derivada da função de ativação do neurônio j em relação à sua saída linear $u_j(n)$;

$x_i(n)$ = Saída do neurônio i da camada escondida;

η = Taxa de aprendizado.

Já para os pesos da camada intermediária, a equação 29 representa o ajuste que é realizado. (Rezende, 2005).

$$w_{ji}(n + 1) = w_{ji}(n) + \eta f_j'(u_j(n)) \sum_l e_l(n) f_l'(u_l(n)) w_{lj}(l) x_i(n) \quad (29)$$

Onde:

$\sum_l e_l(n) f_l'(u_l(n)) w_{lj}(l)$ = Estimativa do erro do neurônio j da camada escondida.

Segundo Rezende (2005, p.153), o fato da superfície do erro ser irregular por conta das não-linearidades das funções de ativação, causa uma certa complexidade nos treinamentos como o *backpropagation*. Já que tais não-linearidades podem provocar o nascimento de locais mínimos e regiões planas de gradiente nulo.

De acordo com Corrar, Paulo e José Filho (2007), o ponto crucial desse algoritmo de retropropagação é o de utilizar o valor do erro para ajuste dos pesos. Onde os neurônios buscam esse ajuste na intenção de reduzir os erros, e não os eliminar. Após passar por uma quantidade suficiente de exemplos, tanto os pesos como os erros não se alteram mais, e é nesse ponto que o treinamento para, já que a rede neural reconheceu as entradas.

Porém, tal algoritmo de *backpropagation* pode ter seu aprendizado afetado, quando maior for a quantidade de exemplos realizados no mesmo. Desde sua criação, o algoritmo vem sendo alterado na tentativa de acelerar o seu tempo de treinamento, sendo criadas adaptações como o *quick-propagation*, entre outras. Mas tais algoritmos acabam por não serem tão satisfatórios como o *backpropagation*. (CORRAR; PAULO; JOSÉ FILHO, 2007). Logo, tal algoritmo ainda continua sendo o mais utilizado e o mais indicado nas redes neurais artificiais.

Pode-se concluir então que as redes neurais artificiais acabam por serem consideradas de maior precisão quanto o assunto é previsão de demandas, já que conseguem se basear na experiência e automatizar ideias. Algo não feito pelos métodos anteriores citados. Se uma rede neural for bem programada, ela passa a entender todas as situações possíveis e resolvê-las.

2.5 ERROS DE PREVISÃO

Erros de previsão, segundo Ragsdale (2017), medem diferenças entre valores reais e valores previstos, ou ajustados, pela técnica de previsão. A seguir estão listadas as equações 30, 31, 32, 33 e 34 conhecidas, respectivamente, como erro de previsão (E_t), erro médio (EM) (Lustosa *et al*, 2008), desvio absoluto médio (DAM), erro quadrado médio (EQM) e erro percentual absoluto ($EPAM$) (Ragsdale, 2017).

$$E_t = D_t - F_t \quad (30)$$

$$EM = \frac{\sum_{i=1}^n (D_t - F_t)}{n} \quad (31)$$

$$DAM = \frac{\sum_{i=1}^n |D_t - F_t|}{n} \quad (32)$$

$$EQM = \frac{\sum_{i=1}^n (D_t - F_t)^2}{n} \quad (33)$$

$$EPAM = \frac{\sum_{i=1}^n \left| \frac{D_t - F_t}{D_t} \right|}{n} \quad (34)$$

Onde:

D_t = Valor real da demanda no período correspondente;

F_t = Valor previsto da demanda no período correspondente;

n = Número de períodos analisados.

A importância da utilização deste cálculo de erro na previsão como última etapa na aplicação, se dar por servir de meio para observar se uma previsão está atingindo resultados satisfatórios ou não. Quanto maior o valor do erro estimado, menos correto será a previsão realizada por tal método. E quando menor for tal erro, melhor e mais precisa será a previsão estimada. Sem este cálculo não seria possível determinar se um método é eficiente ou não.

3 METODOLOGIA

3.1 CARACTERIZAÇÃO DA PESQUISA

De acordo com Gil (2010), para avaliar a qualidade dos resultados de uma pesquisa é necessário saber a proveniência dos seus dados, assim como os procedimentos adotados em sua análise e interpretação. Quanto ao tipo de abordagem a pesquisa se classifica como quantitativa, pelo fato do levantamento de dados passados, com o intuito de prevê-los com a aplicação de métodos matemáticos e também por conta do levantamento feito por meio de entrevista, observação e pesquisas das variáveis que mais impactam na previsão de demanda, que serão transformadas em números no segundo momento.

Se caracteriza por ser uma pesquisa pura, por não se preocupar com a aplicação dos resultados gerados.

Levando em consideração a procedência dos dados, os mesmos se configuram como secundários por conta de serem obtidos por meio de planilhas de dados gerados pelo próprio *software* utilizado nos postos e também por planilhas obtidas no site da ANP (Agência Nacional de Petróleo, Gás Natural e Biocombustíveis).

Tendo como foco agora o objetivo, a pesquisa é descritiva, pois a mesma será analisada com o intuito de se verificar como as informações levantadas impactarão na previsão de demanda em uma maneira geral. Segundo Gil (2010), pesquisas descritivas possuem a finalidade de descrever as características de determinada população, podendo serem elaboradas também com o intuito de identificar possíveis relações entre as variáveis.

Quanto a abrangência de tempo a pesquisa se caracteriza por fazer um estudo transversal, avaliando como diferentes fatores influenciam na demanda dos postos no decorrer de diferentes anos.

3.2 LÓCUS DA PESQUISA

A pesquisa foi realizada em uma rede de postos de combustíveis localizada na região de Belém, Pará. Com um total de cinco postos de combustíveis, onde quatro trabalham com quatro tipos de produto (Óleo diesel B S-10 original aditivado, Etanol hidratado comb marcado, Gasolina aditivada DT clean e Gasolina original C) e um com somente três (Óleo diesel B S-10 original aditivado, Etanol hidratado comb marcado e Gasolina original C).

A mesma pertence a um grupo empresarial com mais de 60 anos inserido no cenário comercial da cidade de Belém. Em relação aos postos, eles tiveram início a cerca de 40 anos atrás, com a construção do primeiro posto de gasolina e posteriormente com a aquisição de mais quatro. Além, da venda de combustíveis o grupo também trabalha com transporte de combustíveis, lojas de serviços automotivos, materiais de construção, conveniência e um centro comercial de médio porte.

3.3 PROCEDIMENTOS METODOLÓGICOS

3.3.1 Dados diários

O primeiro passo a ser feito foi a coleta de dados, que no caso deste trabalho foi feita através do *software* EMSys® e convertidos em planilha Excel® (versão 2010). Estes dados foram coletados entre 01 de Janeiro de 2016 e 31 de Março de 2018, dos relatórios de venda de uma rede formada por cinco postos de combustíveis referentes a quatro tipos de mercadorias diferentes (Óleo diesel B S-10 original aditivado, Etanol hidratado comb. marcado, Gasolina aditivada DT clean e Gasolina original C). No total foram coletadas 15.599 amostras de dados referentes à demanda de todos os produtos vendidos em cada um dos postos.

Após coletados, os dados necessitaram de tratamento. Para tal, os dados foram transplantados do Excel® para o documento online Planilhas Google® com o intuito de tornar o tratamento dos dados mais rápido por meio do compartilhamento do documento ao mesmo tempo entre os autores. O tratamento foi constituído pela tabulação dos dados de maneira a facilitar seu uso posteriormente e adequação dos mesmos.

Durante o ajuste observou-se dois problemas. O primeiro, dizia respeito à falta de dados em períodos de dias aleatórios, que poderiam ser provenientes da falta de um produto no posto ou a falta do mesmo pelo fornecedor. Não descartando a demanda para os mesmos em tais respectivos dias. Já o segundo, foi observado quando em determinados dias as vendas de todos os combustíveis não foram computadas, mas percebeu-se que os valores de vendas do dia seguinte eram consideravelmente altos em relação ao normal. Logo, concluiu-se que de alguma forma, os dados haviam sido somados no dia posterior.

Para contornar tais situações, foi planejado o seguinte: No primeiro caso, seria utilizado algum método estatístico que visasse definir valores com base em algum fundamento específico para eliminar tais *gaps* nas tabelas. E conforme discutido com o

orientador e com um professor especializado na área, foram sugeridos dois métodos: o intervalo interquartil e a interpolação polinomial de 6º grau. Contudo, a partir de uma breve aplicação e devido a algumas dúvidas e incertezas quanto aos resultados, o método envolvendo quartis foi descartado. Enquanto que a interpolação foi aplicada nos dados de um posto, porém apresentando resultados abaixo do esperado, já que as lacunas ficavam com resultados negativos. Restou-se então, aplicar os dados de forma bruta, sem tratamento, inserindo o número 0 (zero) nos *gaps*, já que em tais dias de fato não houve vendas, por conta da falta de tal produto e não por falta da demanda em si.

No segundo problema, a situação foi resolvida de maneira bem simples. Uma média foi tirada do dia com valores somados e então seu resultado distribuído para os dias com lacunas. Vale ressaltar, que em alguns casos a soma englobou mais de dois dias, ou seja, a lacuna inicial era de quatro dias ou mais, e seus valores todos somados no 5º dia. Para isso, o resultado do quinto dia foi dividido por cinco e seu novo valor distribuído assim para os cinco dias em questão.

Também por meio do software EMSys®, foi possível obter os dados referentes aos valores dos combustíveis para venda de todos os postos, onde os mesmos foram transcritos para o Excel®. É importante ressaltar que a rede de postos trabalhava com valores padronizados de combustíveis, ou seja, os valores dos combustíveis eram os mesmos para todos os postos. Ex: o valor da Gasolina aditivada DT clean no posto A no dia 25/01/2017 era o mesmo para os demais postos.

Todos os dados são encontrados nos apêndices S, T, U, V, W e AC pertencentes a este trabalho e podem ser visualizados e baixados no repositório digital Zenodo, conforme link a seguir: <https://bit.ly/2A8ZtEL>.

3.3.2 Dados semanais

Por meio do site da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP), foi possível obter duas planilhas em formato Excel®: uma continha informações referentes aos anos de 2012 até 2017 e a outra a respeito de 2018, de onde foram extraídos os dados de 03/01/2016 até 31/03/2018 referentes à média de preços dos combustíveis (óleo diesel S10, etanol hidratado e gasolina comum) no município de Belém entre vários postos. A gasolina aditivada não possuía quaisquer informações a respeito do preço de venda ao longo deste período.

Após a adequação dos dados diários, foram feitos a partir destes os dados semanais, onde foram somados os dados de venda de sete em sete dias, resultando em uma quantidade menor de valores, com um total de 2223 amostras, onde cada produto possui 117. Os valores necessitaram ser somados duas vezes. Na primeira a soma de sete em sete dias se iniciou desde o dia 01/01/2016 e finalizou no dia 29/03/2018, sendo descartados os dias 30 e 31 de março de 2018, pois não formavam a somatória de sete dias que era requisito. Esse primeiro conjunto de dados precisaram ser totalmente descartados, pois enquanto se aplicavam os métodos causais, observou-se que os mesmos não estavam alinhados com os dados colhidos no site da Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP), onde a mesma iniciou sua análise a partir do dia 03/01/2016. Logo, a segunda soma precisou obedecer a esse critério, onde a soma de sete em sete dias se iniciou desde o dia 03/01/2016 e finalizou no dia 31/03/2018. Sendo assim, os dias 01 e 02 de janeiro de 2016 precisaram ser descartados, pelo fato de não formarem a somatória de sete dias que era requisito.

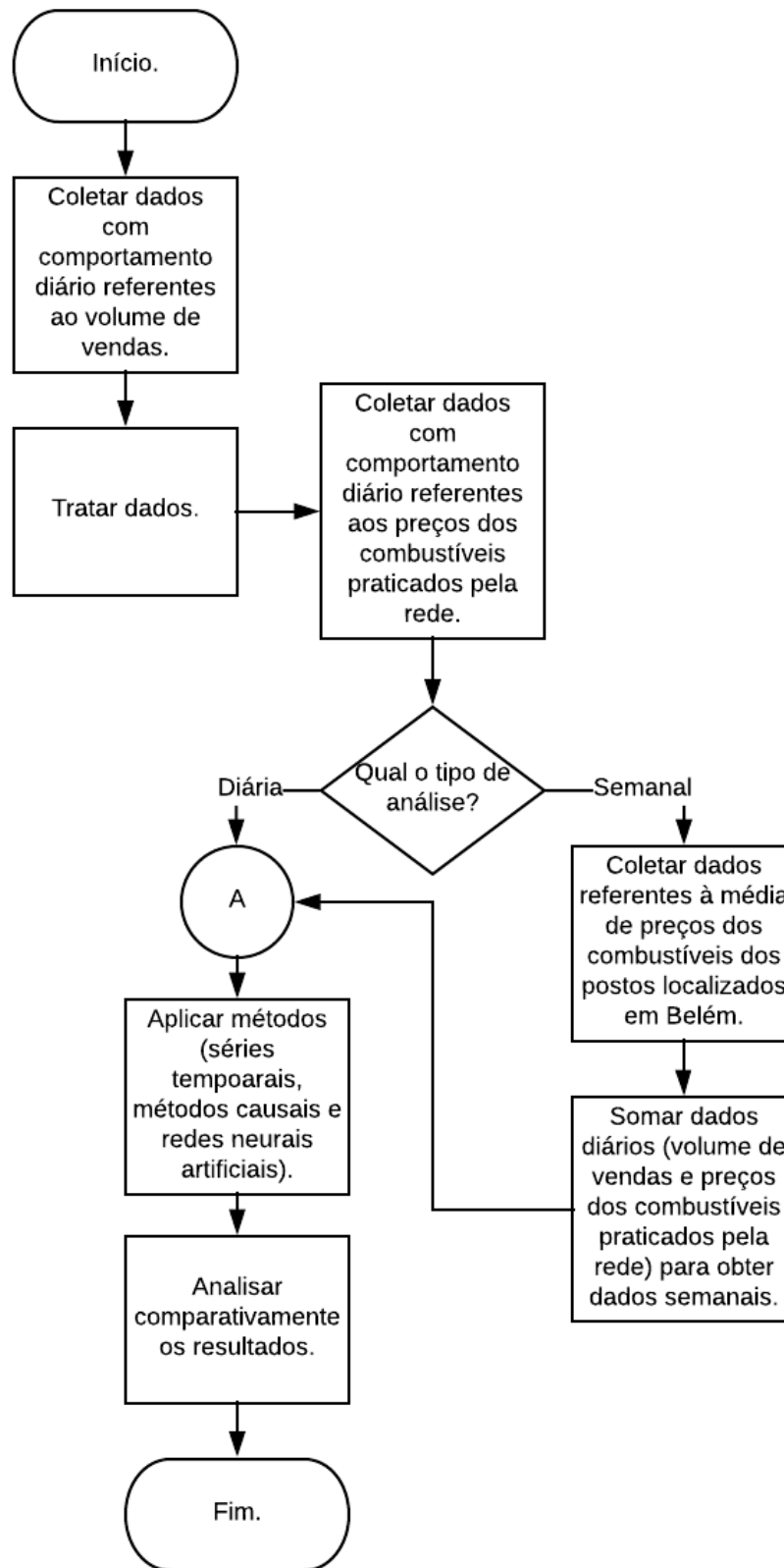
Analogamente ao descrito no parágrafo anterior, para os valores dos combustíveis da rede de postos foi realizada uma média de sete em sete dias, que se iniciou no dia 03/01/2016 e finalizou no dia 31/03/2018. Os dias 01 e 02 de janeiro de 2016 precisaram ser descartados, por conta das mesmas não formarem a somatória de sete dias que era requisito.

Todos os dados são encontrados nos apêndices X, Y, Z, AA, AB e AD pertencentes a este trabalho e podem ser visualizados e baixados no repositório digital Zenodo, conforme link a seguir: <https://bit.ly/2A8ZtEL>.

Concluídas estas etapas, foi realizada a aplicação, nos dados diários e semanais, de diferentes métodos para a previsão de demanda que são eles: Séries Temporais, que envolvem a Média Móvel Simples, Dupla e Ponderada, Suavização Exponencial Simples, com Tendência e com Tendência e Sazonalidade; Métodos Causais envolvendo Análise de Regressão Linear Simples e Múltipla, e por fim, Redes Neurais Artificiais com aprendizado *backpropagation*. Tais aplicações serão detalhadas mais abaixo neste trabalho.

Ao final foram realizadas análises envolvendo as séries de dados diários e semanais, os métodos aplicados e as ferramentas utilizadas para realização dos mesmos, por meio dos resultados obtidos e pelo que foi vivenciado durante a realização do trabalho, com o intuito de se definir qual o melhor para basear uma previsão demanda. Todos os procedimentos podem ser melhor visualizados na Figura 7.

Figura 7 - Fluxograma dos procedimentos metodológicos realizados



Fonte: Autores (2018)

4 APLICAÇÃO DOS MÉTODOS

Para fins de facilitar o entendimento geral do trabalho, foi selecionado um único combustível de um único posto, sendo ele o etanol hidratado comb. marcado do posto D, na tentativa de demonstrar em cada método os procedimentos utilizados durante a aplicação. Logo, o que foi aplicado neste produto, foi replicado para todos os demais combustíveis dos outros postos.

4.1 APLICAÇÃO DE SÉRIES TEMPORAIS NO EXCEL®

Após as mudanças citadas anteriormente em relação às planilhas, houve então a aplicação dos dados em cada um dos métodos de previsões. A primeira a ser trabalhada foi a de séries temporais. Suas aplicações em cada método dos já mencionados tiveram como bases o conteúdo teórico já explicado mais acima neste trabalho. Já nas tabelas, foi inserido o cálculo do erro percentual absoluto (MAPE), que faz uma média entre todos os erros obtidos das previsões. A Figura 8 adiante demonstra a fórmula utilizada na planilha Excel®.

Figura 8 - Fórmula da média no Excel®

MÉDIA

Núm1 = número

Núm2 = número

=

Retorna a média (aritmética) dos argumentos que podem ser números ou nomes, matrizes ou referências que contêm números.

Núm1: núm1;núm2;... de 1 255 argumentos numéricos cuja média se deseja obter.

Resultado da fórmula =

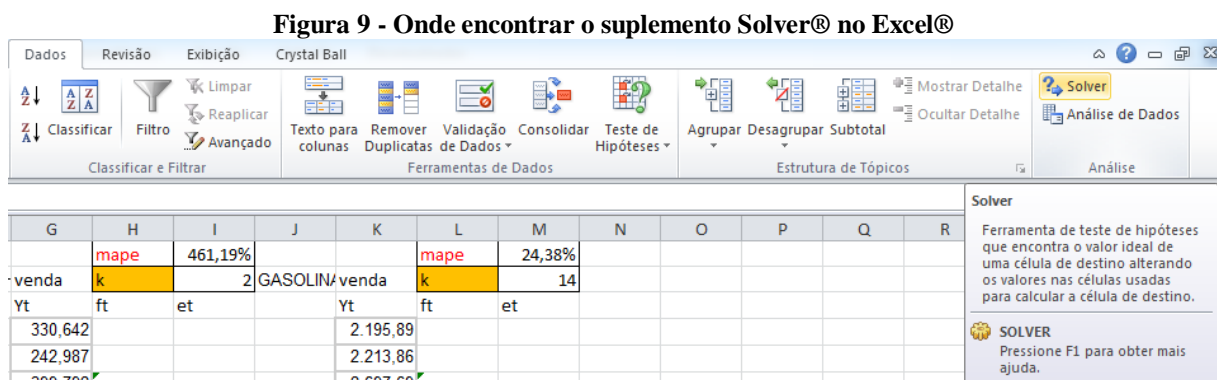
[Ajuda sobre esta função](#)

Fonte: Autores (2018)

Nas Séries Temporais, foram aplicados os dados de duas formas: análises diárias e semanais. Não foram aplicadas de maneira mensal e anual, justamente porque o intuito era o de prever a demanda da maneira mais clara à realidade da empresa, uma vez que a compra

dos pedidos é realizada em intervalos de poucos dias. Nessas análises, tem-se o objetivo de prever a demanda com base no dia(s)/semana(s) anterior(es) ou no mesmo dia(s)/semana(s), porém da(s) semana(s)/mês(es) anterior(es). Mas isso depende do método utilizado para previsão.

Para todos os métodos foi utilizado o suplemento Solver® da ferramenta Excel®, com o intuito de se obter o menor erro possível em cada valor de MAPE. Em todas as aplicações, com exceção da média móvel simples e dupla, os valores de alpha, beta e gama foram deixados em um valor padrão de início equivalente a 0,5, como orientado aos autores, para que após a execução do suplemento, o resultado fosse o mais apropriado possível. A Figura 9 adiante demonstra onde se encontra esta ferramenta nas planilhas:



Nas médias móveis simples (MMS) e dupla (MMD), o Solver® buscou dentre um número determinado de dias ou semanas que seriam as restrições (no caso deste trabalho para o diário foi de 2 a 15 dias e semanal de 4 a 8 semanas), qual seria o melhor período para a previsão, reduzindo o erro. Porém, durante a aplicação, percebeu-se que os resultados fornecidos pelo Solver® não eram o ideal, por conta de um ótimo local, chegando a um resultado que não era o melhor possível. Por conta disso, foi feita a busca pelo melhor resultado de maneira manual, uma vez que o número de possibilidades era pequeno. Vale ressaltar que nos demais métodos, este mesmo problema pode ter ocorrido, porém não foi realizada uma busca de maneira manual, já que as possibilidades eram inúmeras. A Figura 10 e a Figura 11 mostram parte da aplicação de tais métodos:

Figura 10 - Aplicação parcial da média móvel simples em dados diários

	mape	461,19%
venda	k	2
Yt	ft	et
330,642		
242,987		
399,792	286,8145	0,282591
537,596	321,3895	0,402173
341,004	468,694	0,374453
146,882	439,3	1,990836
301,954	243,943	0,192119
359,452	224,418	0,375666

Fonte: Autores (2018)

Figura 11 - Aplicação parcial da média móvel dupla em dados diários

		mape	279,00%
venda		k	2
Yt	mms	mmd	et
330,642			
242,987	286,8145		
399,792	321,3895	304,102	0,239349
537,596	468,694	395,0418	0,26517
341,004	439,3	453,997	0,331354
146,882	243,943	341,6215	1,325823
301,954	224,418	234,1805	0,22445
359,452	330,703	277,5605	0,227823

Fonte: Autores (2018)

Na média móvel ponderada, a previsão foi feita tendo como base o dia anterior na análise diária ou a semana anterior na análise semanal desejada (Exemplo: Prevendo quarta-feira com base na terça-feira; terceira semana baseada na segunda semana). Uma exclusividade deste método é uma restrição existente para o funcionamento do suplemento do Solver® é que a somatória das duas variáveis (alfa e beta) precisam ser iguais, obrigatoriamente, a 1. A Figura 12 a seguir, mostra como ficou parte da aplicação utilizando o combustível etanol hidratado comb. marcado, do posto D:

Figura 12 - Aplicação parcial da média móvel ponderada em dados diários

	mape	187,52%
venda		1
Yt	ft	et
330,642	alfa	0
242,987	beta	1
399,792	242,987	0,392216
537,596	399,792	0,256334
341,004	537,596	0,576509
146,882	341,004	1,321619
301,954	146,882	0,513562
359,452	301,954	0,15996

Fonte: Autores (2018)

Na suavização exponencial simples e suavização exponencial com tendência, a previsão foi feita tendo como base o dia anterior na análise diária ou a semana anterior na análise semanal desejada (Exemplo: Prevendo quarta-feira com base na terça-feira; terceira semana baseada na segunda semana). Por sua vez, no método suavização exponencial com tendência e sazonalidade, foram aplicados a partir dos ciclos, ou seja, prevendo segunda-feira com base no mesmo dia da semana anterior na análise diária; e prevendo a segunda semana do mês com base na segunda semana do mês anterior na análise semanal. E assim se seguiu como modelo para cada combustível de cada posto. As Figuras 13, 14 e 15 mostram parte da aplicação destes três métodos citados:

Figura 13 - Aplicação parcial da suavização exponencial simples em dados diários

		mape	187,22%
venda		alfa	1
Yt	bt	ft	et
	655,0868		
330,642	330,642	655,0868	0,981257
242,987	242,987	330,642	0,360739
399,792	399,792	242,987	0,392216
537,596	537,596	399,792	0,256334
341,004	341,004	537,596	0,576509
146,882	146,882	341,004	1,321619
301,954	301,954	146,882	0,513562
359,452	359,452	301,954	0,15996

Fonte: Autores (2018)

Figura 14 - Aplicação parcial da suavização exponencial com tendência em dados diários

			mape	177,26%
venda	alfa	0,586729		
	beta	0,436144		
Yt	Bt	Tt	Ft	erro
	316,5949	0,17441		
330,642	324,9088	3,724401	316,7694	0,041957
242,987	278,3821	-18,1923	328,6332	0,352473
399,792	342,0984	17,53166	260,1898	0,349187
537,596	464,0478	63,07288	359,6301	0,33104
341,004	417,9207	15,44588	527,1207	0,54579
146,882	265,2778	-57,8651	433,3665	1,95044
301,954	262,8828	-33,6721	207,4127	0,313098
359,452	305,627	-0,34355	229,2107	0,362333

Fonte: Autores (2018)

Figura 15 - Aplicação parcial da suavização exponencial com tendência e sazonalidade em dados diários

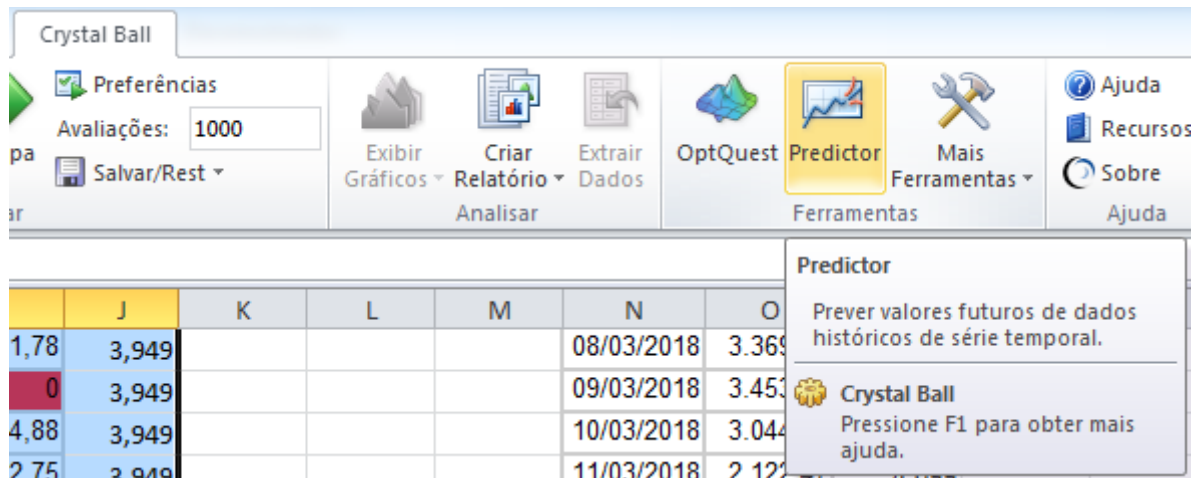
			mape	307,71%		
vendas	alfa	0,499197737	beta	0,50427255	gama	0,507135
Yt	Bt	Tt	It	Ft	erro	
330,642	328,6938571		1,005927			
242,987	328,6938571		0,73925			
399,792	328,6938571		1,216305			
537,596	328,6938571		1,635552			
341,004	328,6938571		1,037452			
146,882	328,6938571		0,446866			
301,954	301,0353519	0	0,918648			
359,452	329,1395638	14,17218262	1,00661	302,8196	0,157552	

Fonte: Autores (2018)

4.2 APLICAÇÃO DE SÉRIES TEMPORAIS NO CRYSTAL BALL®

Nas aplicações já mencionadas, os dados não sofreram qualquer tipo de retirada de *outliers* (valores afastados dos demais da série), logo era natural que seus resultados apresentassem inconsistências ou valores muito exorbitantes de erro. Logo, para uma melhor análise, foram-se também aplicados os dados no Oracle Crystal Ball®, por meio da opção Predictor®, com a intenção de observar como os dados se comportam com a ausência de tais *outliers*. A Figura 16 mostra onde se encontra essa funcionalidade:

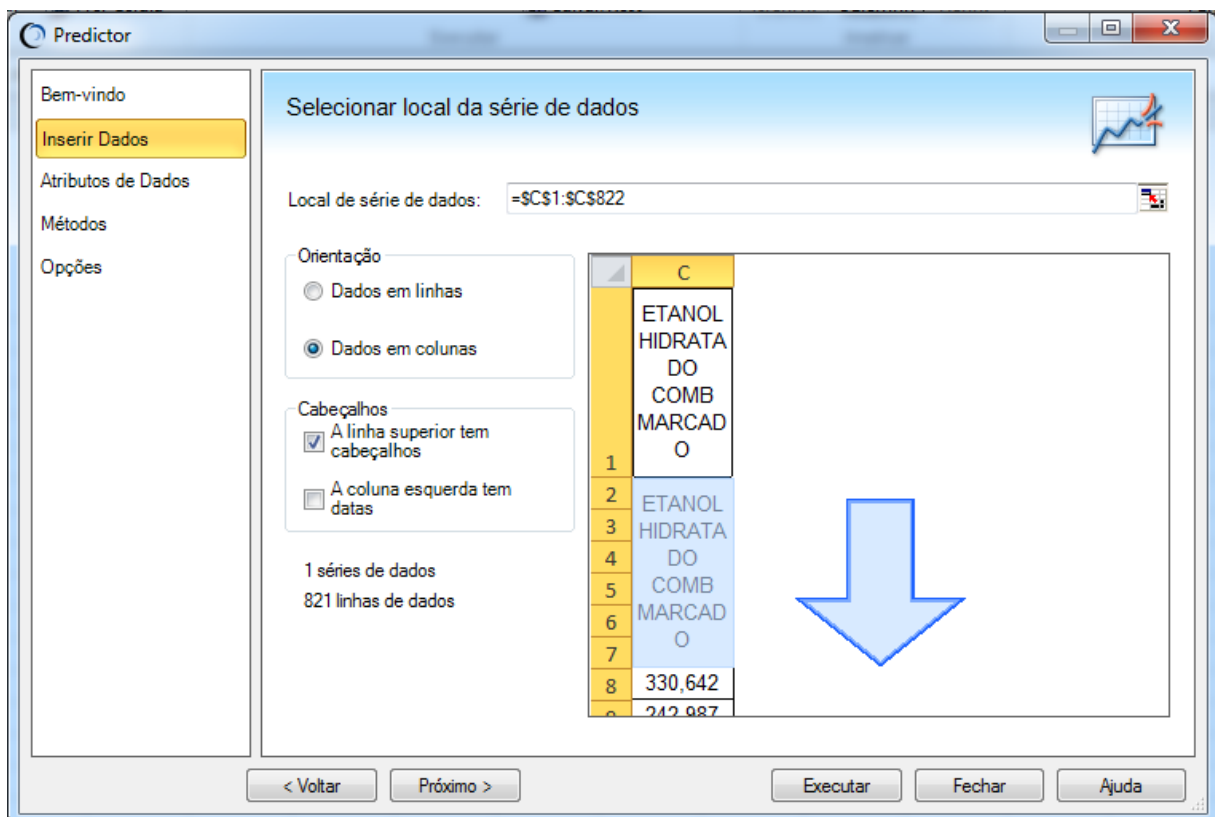
Figura 16 - Onde encontrar a opção Predictor® no Excel®



Fonte: Autores (2018)

Na Figura 17, foi preenchido segundo as necessidades dos dados, que foram previamente selecionados, sendo elas “Dados em colunas” e “Linha superior tem cabeçalhos”.

Figura 17 - Seleção da série de dados inseridos no Predictor®

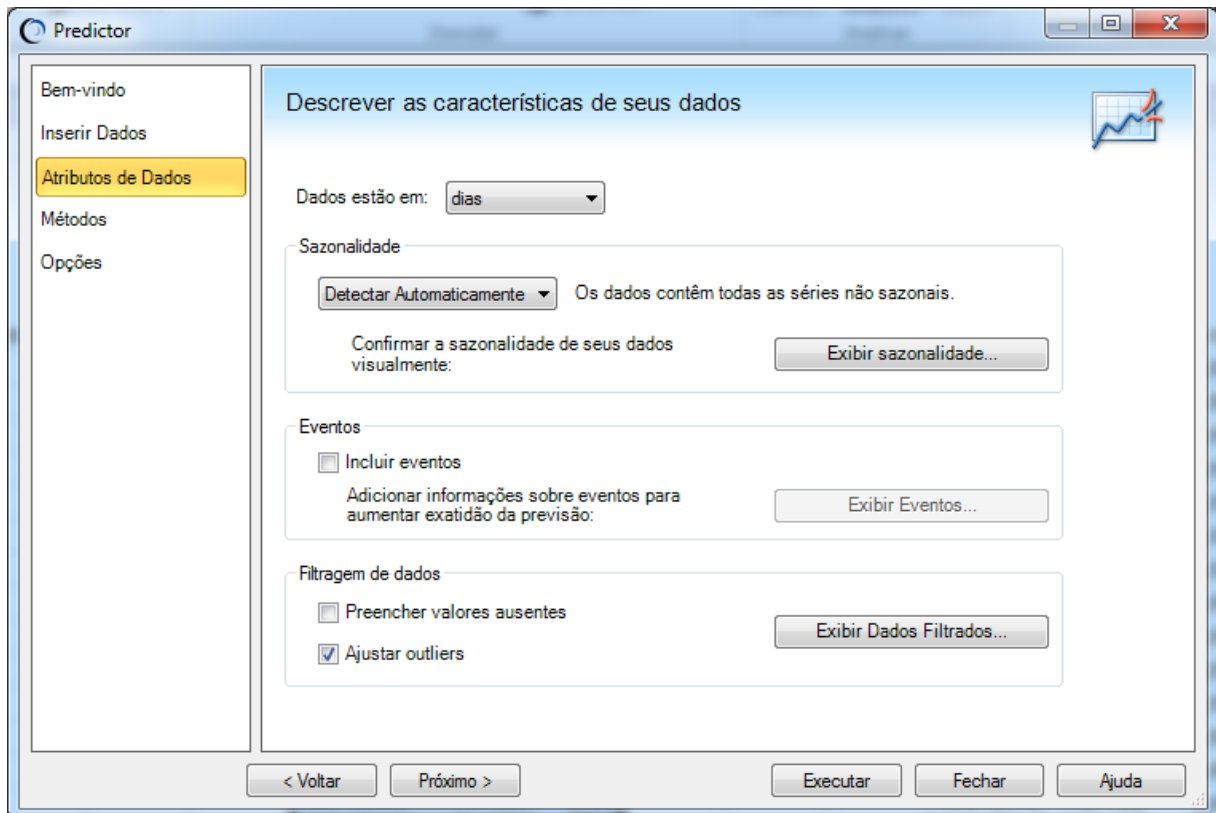


Fonte: Autores (2018)

Na Figura 18 adiante, foi classificado a forma como os dados estão (Dias) e foi marcada a opção de ajustar os *outliers*. Com o intuito de eliminar tais valores e tornar a

previsão mais assertiva. Também se escolheu detectar sazonalidade automaticamente, para que caso não haja sazonalidade, o próprio sistema “ignore” a utilização de métodos sazonais.

Figura 18 - Descrição da característica dos dados inseridos no Predictor®

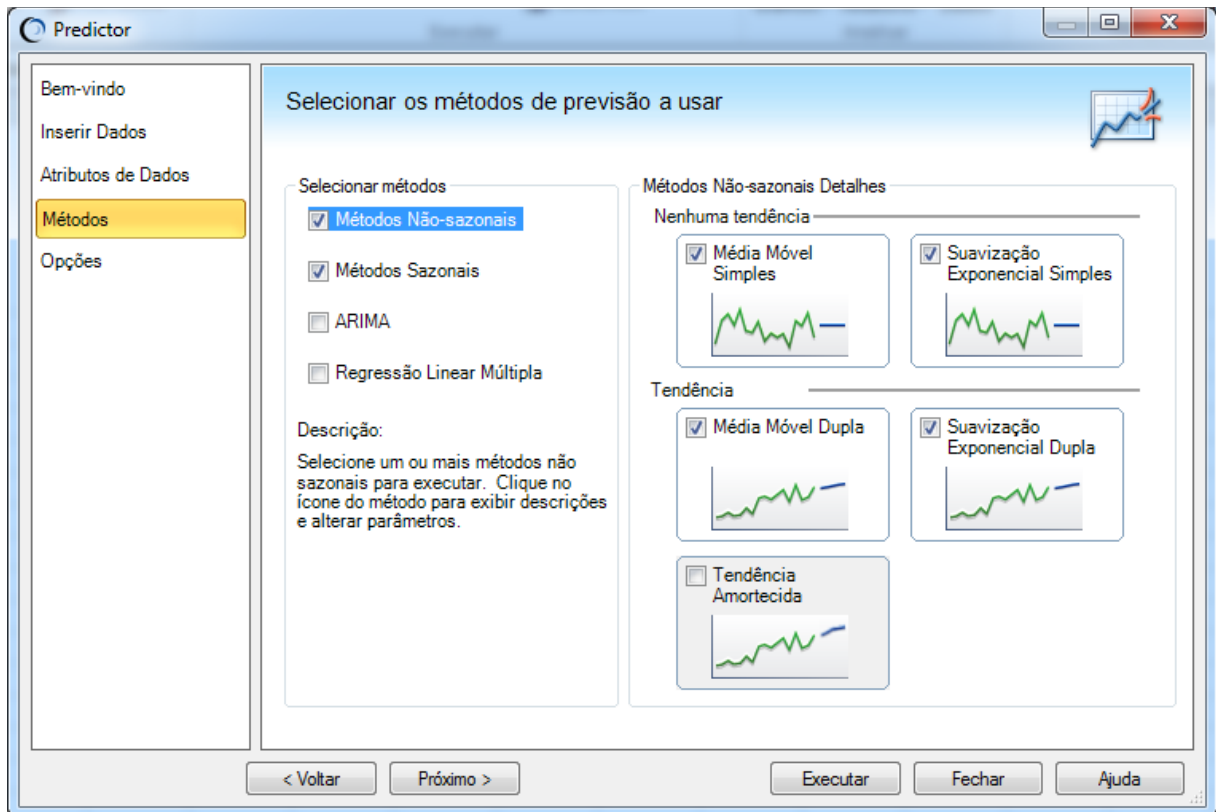


Fonte: Autores (2018)

Na Figura 19 a seguir, foram feitas as seleções dos métodos matemáticos que seriam interessantes, como os métodos não-sazonais e métodos sazonais. Dentre os não-sazonais, foram selecionados as médias móveis simples e dupla e as suavizações exponenciais simples e dupla. Para os sazonais foi escolhido apenas o método multiplicativo de Holt-Winter. Vale ressaltar, que nesta ferramenta, não existe a opção de se trabalhar com a média móvel ponderada.

O intuito de se escolher esses métodos foi o de, após serem aplicados, fazer uma comparação entre os resultados alcançados por meio das formulas aplicadas no Excel® e os obtidos por meio dos mesmos métodos através do Crystal Ball®.

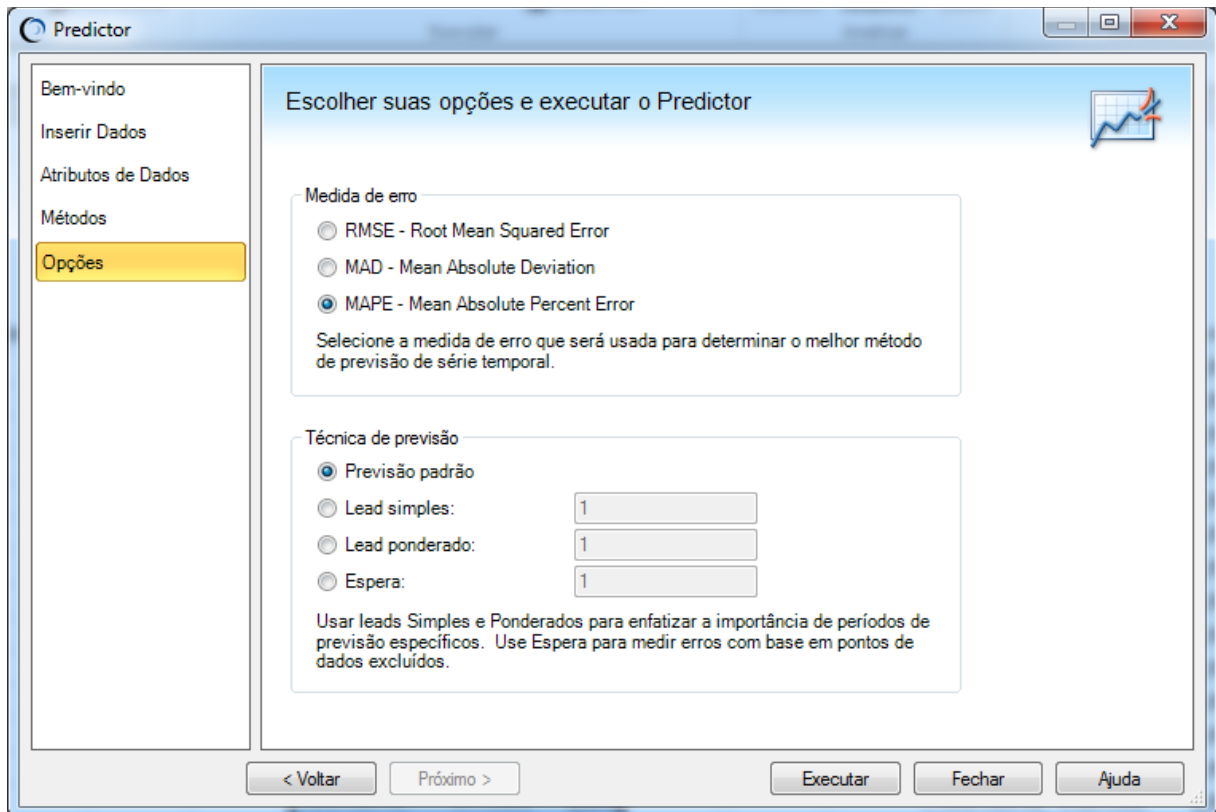
Figura 19 - Seleção dos métodos de previsão que serão aplicados aos dados inseridos no Predictor®



Fonte: Autores (2018)

E por último foi definida a medida de erro, no caso o MAPE, para que a análise final seja feita da mesma forma que foi escolhida nas séries temporais com aplicação de formulas no Excel®. Também se selecionou o tipo de previsão padrão, pois não houve a necessidade de se ressaltar lead-times específicos ou medir erros de dados excluídos. Como mostra a Figura 20 adiante:

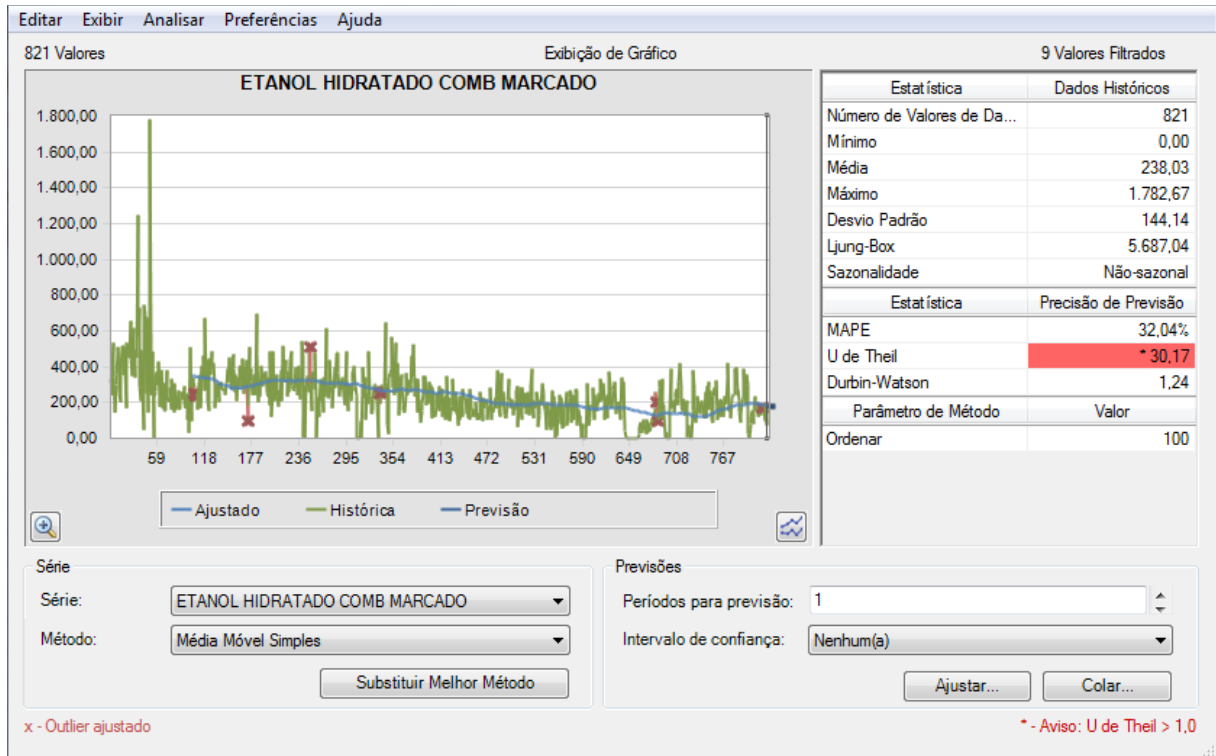
Figura 20 - Seleção do tipo de erro e de previsão dos dados inseridos no Predictor®



Fonte: Autores (2018)

Na Figura 21, tem-se o exemplo da previsão feita de Média Móvel Simples, onde o mesmo foi realizado para os demais métodos e combustíveis. Como se pode observar, o próprio sistema faz um ajuste de onde é melhor iniciar sua previsão, informa o número de *outliers* encontrado, assim como o erro MAPE.

Figura 21 - Previsão de MMS realizada pelo Crystal Ball® para o etanol do posto D



Fonte: Autores (2018)

Todos os passos mostrados acima, também foram utilizados nas análises semanais, sem qualquer alteração, com o intuito de poder fazer uma comparação entre os dois tipos de análises.

4.3 APLICAÇÃO DE MÉTODOS CAUSAIS NO CRYSTAL BALL®

Inicialmente, foi necessário encontrar as variáveis que serviriam de base para a previsão. No caso dos dados diários, foi obtido o histórico dos preços de venda, como comentado mais acima. Feito isso, foram aplicados os dados referentes ao volume de vendas e seus preços na função Predictor® da ferramenta Crystal Ball® do Microsoft Excel®. Como demonstra a Figura 22 a seguir:

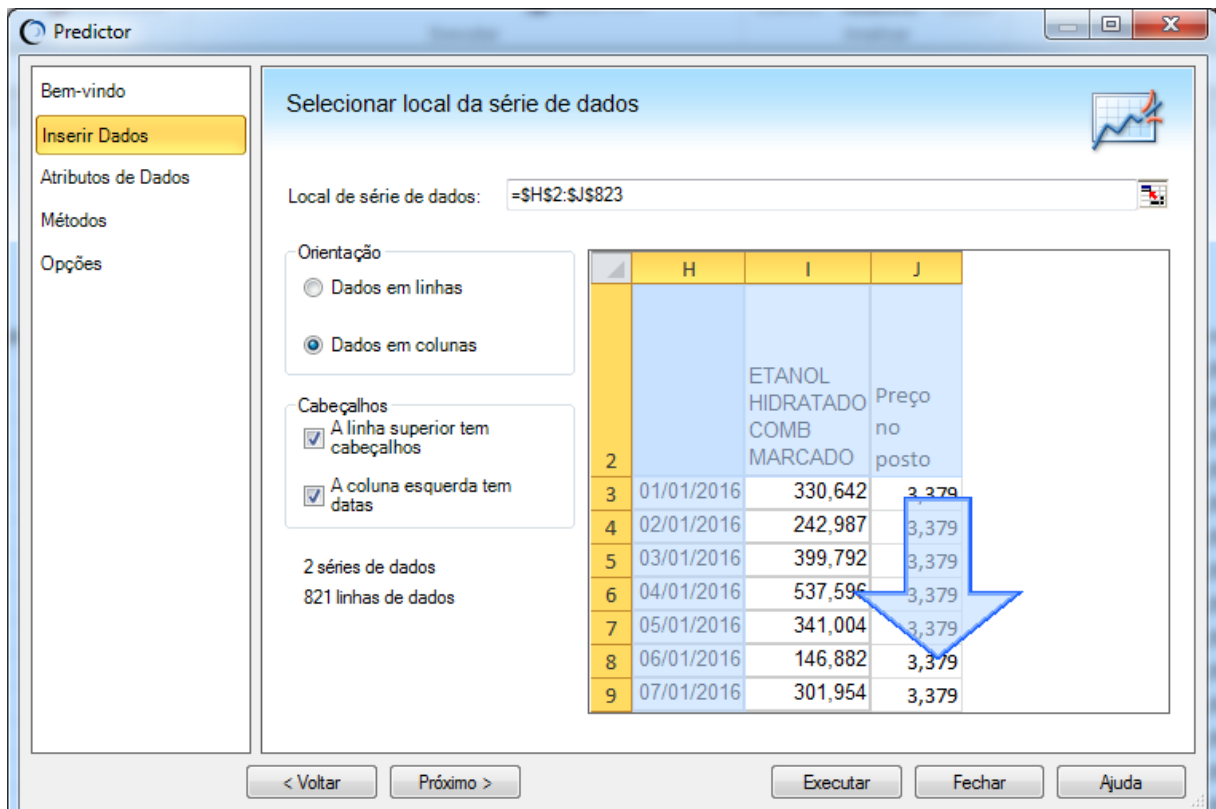
Figura 22 - Seleção parcial dos dados inseridos no Predictor®

	ETANOL HIDRATADO COMB MARCADO	Preço no posto
01/01/2016	330,642	3,379
02/01/2016	242,987	3,379
03/01/2016	399,792	3,379
04/01/2016	537,596	3,379
05/01/2016	341,004	3,379
06/01/2016	146,882	3,379
07/01/2016	301,954	3,379
08/01/2016	359,452	3,379

Fonte: Autores (2018)

Após selecionar os dados e a função Predictor®, foram preenchidas as necessidades dos dados, que seriam “Dados em colunas”, e marcado as opções “Linha superior tem cabeçalhos” e “Coluna esquerda tem datas”. Como mostra a Figura 23:

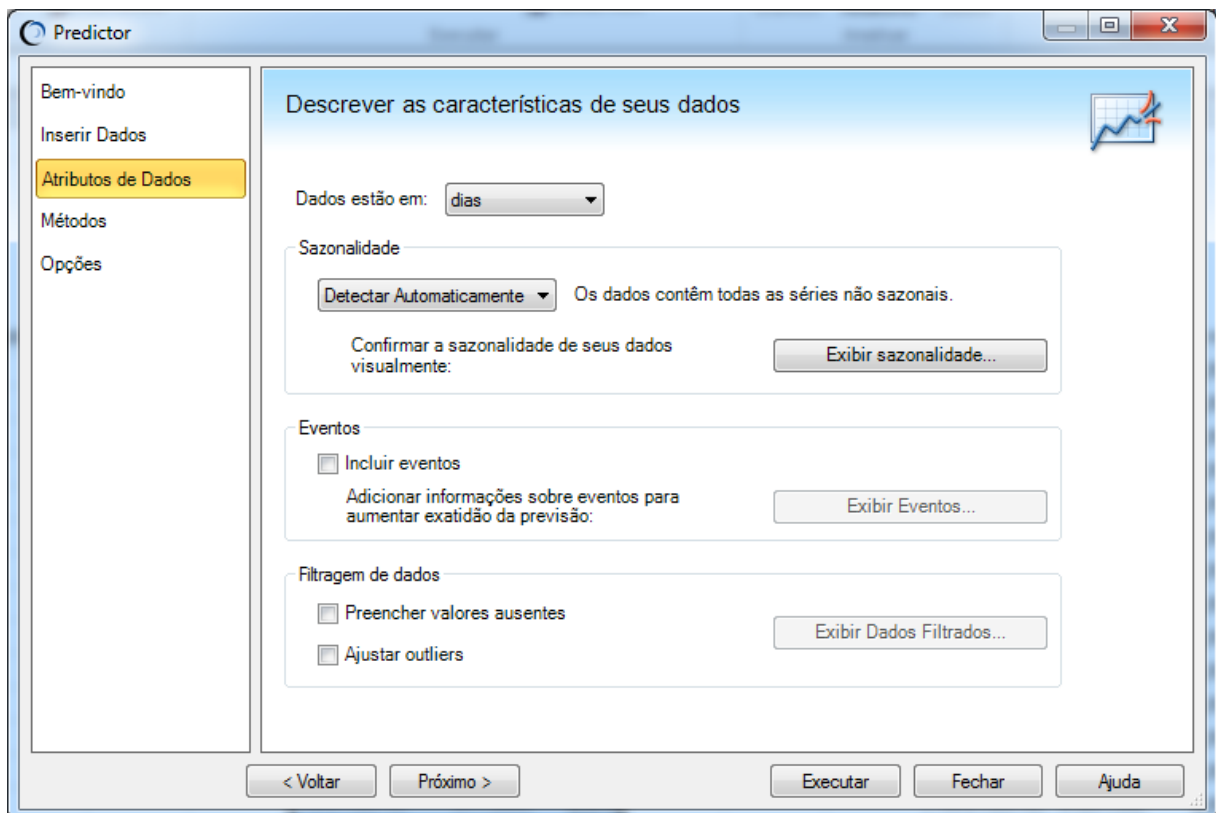
Figura 23 - Seleção da série de dados inseridos no Predictor®



Fonte: Autores (2018)

Na próxima etapa, é onde selecionada os atributos de dados. Foi selecionado que os dados estão em dias, na análise diária, e semanas, na análise semanal. E escolhido detectar sazonalidade automaticamente, para que caso não haja sazonalidade, o próprio sistema “ignore” a utilização de métodos sazonais. Neste caso, não foi marcado a opção “Ajustar outliers”, pois o recurso de filtragem de dados não se aplica a séries de variáveis dependentes em uma análise de regressão. Como mostra a Figura 24 adiante:

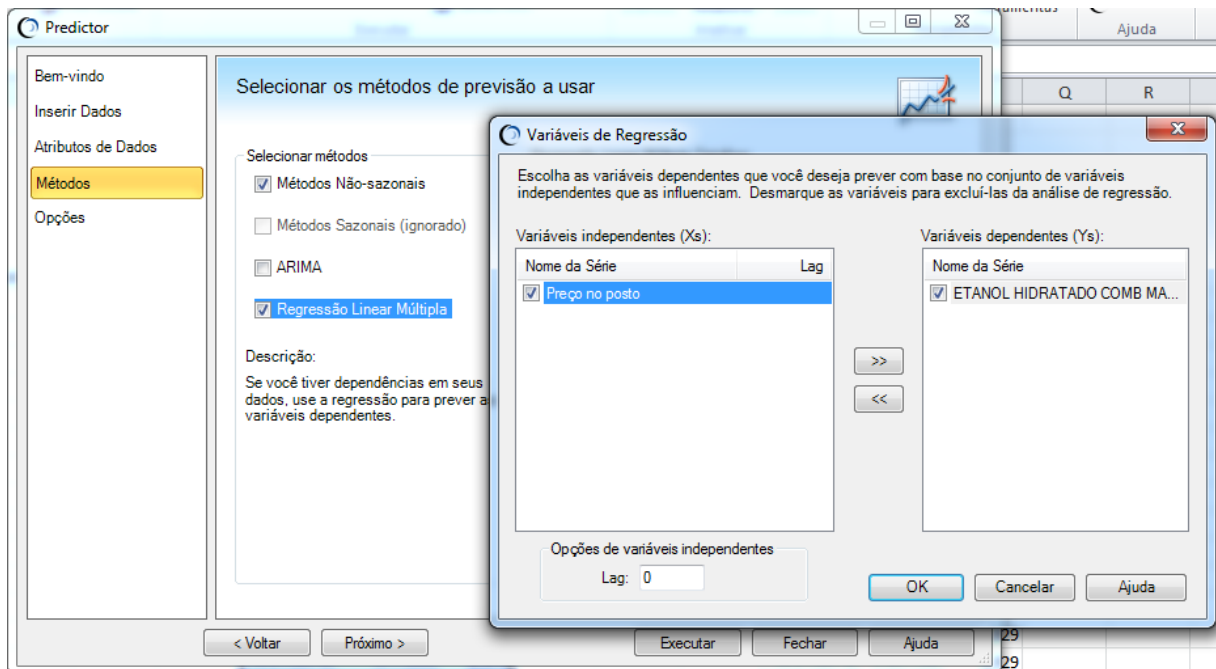
Figura 24 - Descrição da característica dos dados inseridos no Predictor®



Fonte: Autores (2018)

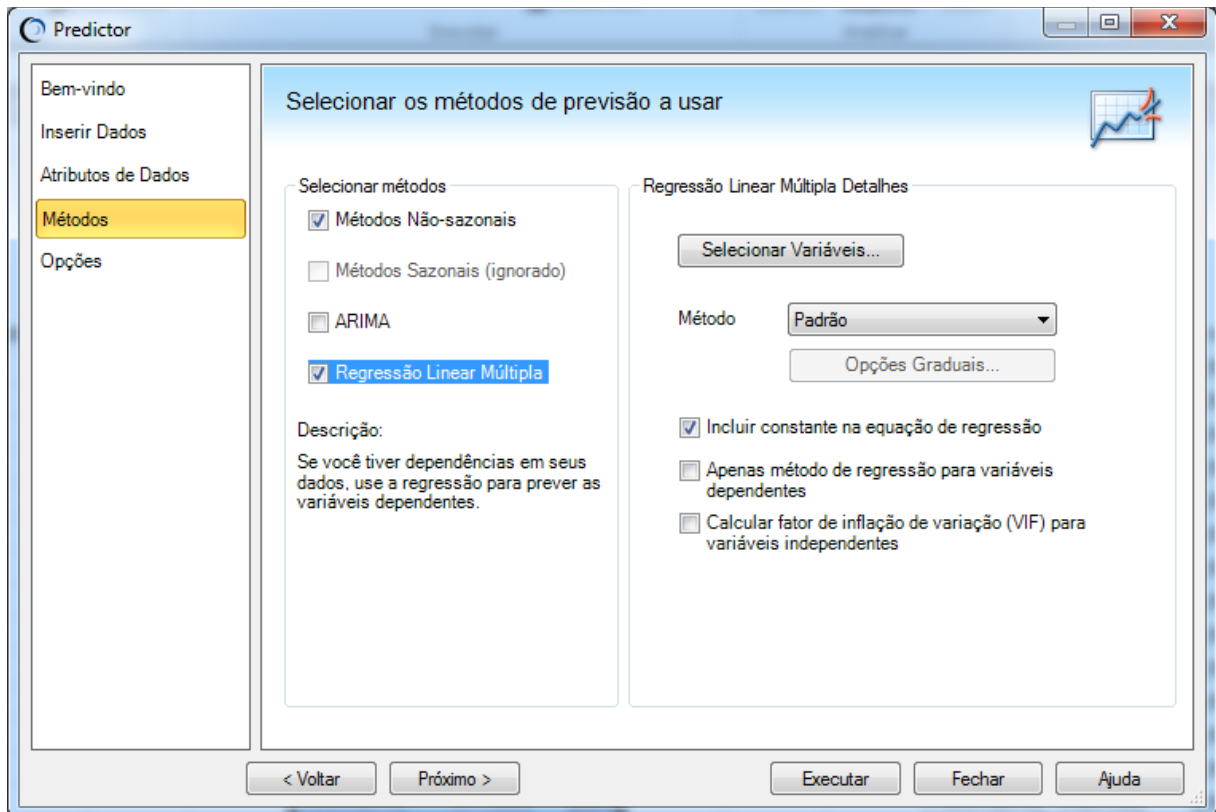
Feito isso, é necessário selecionar o método desejado, no caso o de regressão linear múltipla. Porém, caso seja selecionado apenas este método, um erro é mostrado na tela informando que não foi escolhido nenhum tipo de previsão das séries de tempo. Logo, foi selecionado o de métodos não-sazonais. Após isso, as variáveis dependentes e independentes foram definidas. Como mostra a Figura 25 a seguir:

Figura 25 - Seleção dos métodos de previsão que serão aplicados aos dados do combustível etanol do posto D e definição das variáveis dependentes e independentes inseridos no Predictor®



Fonte: Autores (2018)

A Figura 26 mostra a marcação de “Incluir constante na equação de regressão”, cuja constante representa o intercepto da reta com o eixo dos Y. Ressaltando que quando marcado, apresentava resultados ligeiramente melhores.

Figura 26 - Seleção do método de previsão no Predictor®

Fonte: Autores (2018)

Na Figura 27 adiante, demonstra-se onde se escolhe o tipo de medida de erro, que neste caso foi o MAPE, para que a análise final seja feita da mesma forma que foi escolhida nas séries temporais. Também se selecionou o tipo de previsão padrão, pois não houve a necessidade de se ressaltar lead-times específicos ou medir erros de dados excluídos.

Figura 27 - Seleção do tipo de erro e da técnica de previsão dos dados inseridos no Predictor®

The image shows a software window titled "Predictor". On the left is a navigation pane with the following items: "Bem-vindo", "Inserir Dados", "Atributos de Dados", "Métodos", and "Opções" (which is highlighted in yellow). The main area is titled "Escolher suas opções e executar o Predictor" and contains two sections:

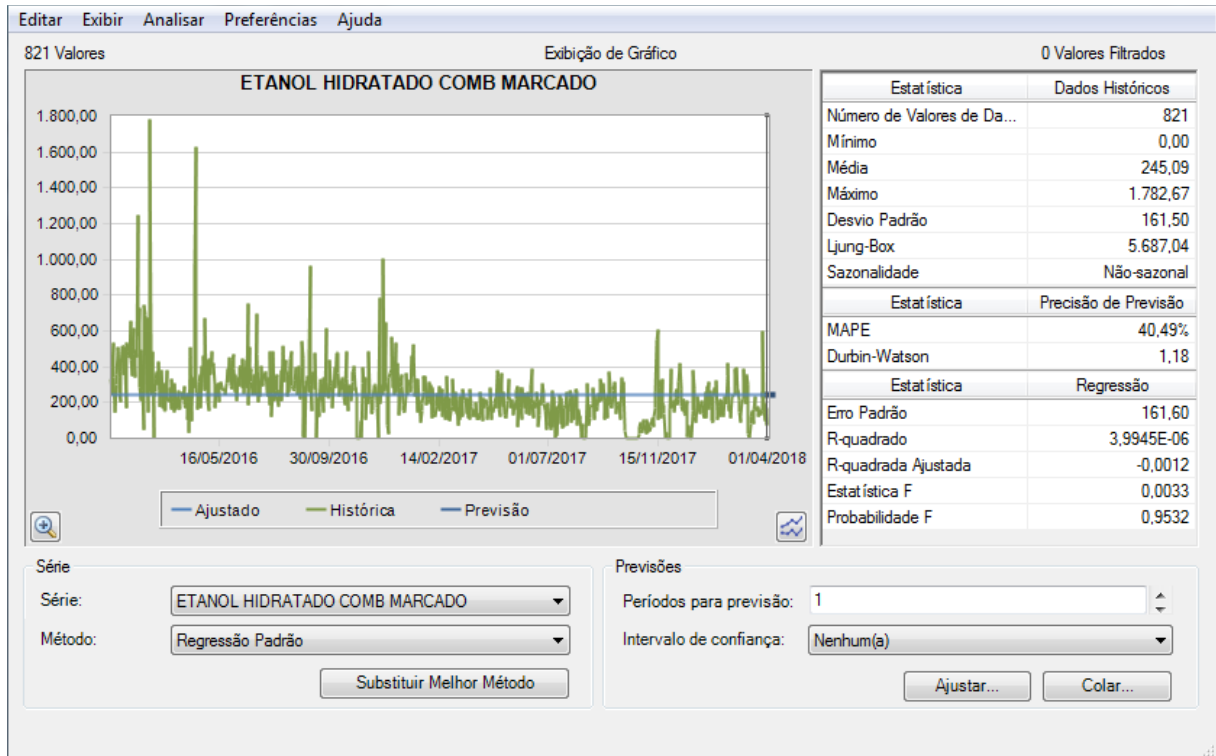
- Medida de erro**: Three radio buttons are present: "RMSE - Root Mean Squared Error", "MAD - Mean Absolute Deviation", and "MAPE - Mean Absolute Percent Error". The "MAPE" option is selected. Below these is the text: "Selecione a medida de erro que será usada para determinar o melhor método de previsão de série temporal."
- Técnica de previsão**: Four radio buttons are present: "Previsão padrão", "Lead simples:", "Lead ponderado:", and "Espera:". The "Previsão padrão" option is selected. The other three options have associated input fields, each containing the number "1". Below these is the text: "Usar leads Simples e Ponderados para enfatizar a importância de períodos de previsão específicos. Use Espera para medir erros com base em pontos de dados excluídos."

At the bottom of the window are five buttons: "< Voltar", "Próximo >", "Executar", "Fechar", and "Ajuda".

Fonte: Autores (2018)

Realizado cada passo anterior, chega o momento da execução do Crystal Ball®. Como mostra a Figura 28. Tem-se selecionado o método de regressão padrão.

Figura 28 - Previsão de regressão padrão realizada pelo Crystal Ball® para o etanol do posto D



Fonte: Autores (2018)

No que diz respeito às análises semanais, a aplicação foi feita de acordo como mostrado no diário, porém tendo como variáveis independentes o preço praticado pela rede de posto de combustíveis e a média de preços de demais postos localizados no município de Belém. Vale ressaltar, que tal método só foi possível de ser realizado para os combustíveis óleo diesel B S-10 original aditivado, etanol hidratado comb. marcado e gasolina original C, pois a gasolina aditivada DT clean não possuía dados média de preços de demais postos localizados em Belém para a análise. E neste caso, neste combustível foi aplicado tendo como variável apenas o preço praticado pela rede. O procedimento de aplicação segue a mesma linha de raciocínio da diária, demonstrada mais acima, onde foi aplicada regressão linear simples aos dados de comportamento diário e semanal. Enquanto a regressão linear múltipla foi aplicada aos dados de comportamento semanal.

4.4 APLICAÇÃO DE REDES NEURAIIS ARTIFICIAIS NO MATLAB®

No que diz respeito à aplicação dos dados nas redes neurais artificiais, o *software* toolbox do Matlab® (versão 2015) foi escolhido pela familiaridade com o programa e pela eficiência do mesmo. Porém, para serem implementados, os dados passaram por alguns

ajustes até serem inseridos na rede. A primeira etapa foi a de reorganizar os dados de maneira a agrupar os mesmos da seguinte forma: de sete em sete dias, sendo o primeiro valor da próxima semana, o segundo valor da semana anterior e assim sucessivamente. Na segunda etapa, a ideia é a de agrupar os dados de maneira que obedecem às condições de agrupamento necessárias para implementação nas redes neurais. Para uma melhor compreensão, a Figura 29 mostra parte do que foi realizado no combustível Etanol hidratado comb. Mercado do Posto D:

Figura 29 - Parte do agrupamento dos dados, no Excel®, do combustível Etanol do Posto D

330,642	1	330,642	242,987	399,792	537,596	341,004	146,882	301,954	359,452	489,331
242,987	2	242,987	399,792	537,596	341,004	146,882	301,954	359,452	489,331	510,441
399,792	3	399,792	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751
537,596	4	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872
341,004	5	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242
146,882	6	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006
301,954	7	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006	518,667
359,452	2	242,987	399,792	537,596	341,004	146,882	301,954	359,452	489,331	510,441
489,331	3	399,792	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751
510,441	4	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872
342,751	5	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242
251,872	6	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006
203,242	7	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006	518,667
377,006	8	359,452	489,331	510,441	342,751	251,872	203,242	377,006	518,667	507,519
518,667	3	399,792	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751
507,519	4	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872
360,181	5	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242
507,92	6	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006
377,25	7	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006	518,667
166,714	8	359,452	489,331	510,441	342,751	251,872	203,242	377,006	518,667	507,519
534,009	9	489,331	510,441	342,751	251,872	203,242	377,006	518,667	507,519	360,181

Fonte: Autores (2018)

Feito isso, a próxima parte consistiu em copiar todas as primeiras sete linhas e as colunas respectivas para finalizar o processo de agrupamento. A Figura 30 adiante demonstra como foi feita:

Figura 30 - Seleção das primeiras 7 linhas e 814 colunas para agrupamento, no Excel®, do combustível Etanol do Posto D

330,642	1	330,642	242,987	399,792	537,596	341,004	146,882	301,954	359,452	489,331
242,987	2	242,987	399,792	537,596	341,004	146,882	301,954	359,452	489,331	510,441
399,792	3	399,792	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751
537,596	4	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872
341,004	5	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242
146,882	6	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006
301,954	7	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006	518,667
359,452	2	242,987	399,792	537,596	341,004	146,882	301,954	359,452	489,331	510,441
489,331	3	399,792	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751
510,441	4	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872
342,751	5	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242
251,872	6	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006
203,242	7	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006	518,667
377,006	8	359,452	489,331	510,441	342,751	251,872	203,242	377,006	518,667	507,519
518,667	3	399,792	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751
507,519	4	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872
360,181	5	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242
507,92	6	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006
377,25	7	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006	518,667
166,714	8	359,452	489,331	510,441	342,751	251,872	203,242	377,006	518,667	507,519
534,009	9	489,331	510,441	342,751	251,872	203,242	377,006	518,667	507,519	360,181

Fonte: Autores (2018)

Com os dados agrupados, têm-se então a camada de entrada de 7 dias. Para encontrar a camada de neurônios de saída, foi levado em consideração o sétimo valor da coluna seguinte, como demonstra a Figura 31 a seguir:

Figura 31 - Dados de saída baseando-se no 7º valor da coluna seguinte, no Excel®, do combustível Etanol do Posto D

	330,642	242,987	399,792	537,596	341,004	146,882	301,954	359,452	489,331
	242,987	399,792	537,596	341,004	146,882	301,954	359,452	489,331	510,441
	399,792	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751
	537,596	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872
	341,004	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242
	146,882	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006
	301,954	359,452	489,331	510,441	342,751	251,872	203,242	377,006	518,667
saída	=C8	489,331	510,441	342,751	251,872	203,242	377,006	518,667	507,519

Fonte: Autores (2018)

Contudo, ao finalizar tal procedimento percebeu-se que o último resultado de saída apresentava valor igual a 0 (zero), pelo fato de não se ter um dado de referência na coluna seguinte, como pode ser visto na Figura 32. E como foi orientado durante as aulas que neste trabalho as matrizes tivessem mesma quantidade de colunas, esta última foi excluída.

Figura 32 - Saída igual a 0, por não ter referência na coluna seguinte. Dados do Etanol do Posto D, no Excel®

155,63	175,36	118,43	144,39	162,42	132,21	149,68	156,44	
175,36	118,43	144,39	162,42	132,21	149,68	156,44	602,74	
118,43	144,39	162,42	132,21	149,68	156,44	602,74	145,57	
144,39	162,42	132,21	149,68	156,44	602,74	145,57	98,81	
162,42	132,21	149,68	156,44	602,74	145,57	98,81	192,74	
132,21	149,68	156,44	602,74	145,57	98,81	192,74	70,77	
149,68	156,44	602,74	145,57	98,81	192,74	70,77	171,76	
156,44	602,74	145,57	98,81	192,74	70,77	171,76	0	

Fonte: Autores (2018)

Feito todo o agrupamento dos dados, foi necessária uma última etapa de tratamento, já que os dados estavam muito dispersos, o que levaria ao *software* Matlab® e as redes a fazerem análises inconclusivas, prejudicando assim o tratamento dos dados e suas previsões. Vale ressaltar, que a rede neural ainda consegue ter um bom aprendizado mesmo com tais valores dispersos, mas que para obter um nível de acerto ainda maior, pode-se realizar o procedimento de normalização a seguir, de acordo com a equação 35 adiante:

$$y = \frac{(x - x_{\min})(d2 - d1)}{x_{\max} - x_{\min}} + d1 \quad (35)$$

Onde:

y = Valor normalizado;

x = Valor a ser normalizado;

x_{\max} = Valor máximo do conjunto de dados;

x_{\min} = Valor mínimo do conjunto de dados;

$d1$ = Novo valor mínimo do conjunto de dados;

$d2$ = Novo valor máximo do conjunto de dados;

Esta normalização tem por intenção deixar todos os valores dentre o intervalo 0 e 1, para facilitar o aprendizado da rede neural e a consequente previsão oriunda do programa.

Primeiramente, deve-se achar o menor e o maior valor dentre as camadas de entrada, que servirão de base tanto para a normalização quanto para a matriz R, que será abordada mais adiante. A Figura 33 mostra esses valores sendo encontrados:

Figura 33 - Encontrando os menores e maiores valores, dentre os dados diários do Etanol do Posto D, no Excel®

	330,642	242,987	399,792	537,596	341,004	146,882	301,954
	242,987	399,792	537,596	341,004	146,882	301,954	359,452
	399,792	537,596	341,004	146,882	301,954	359,452	489,331
	537,596	341,004	146,882	301,954	359,452	489,331	510,441
	341,004	146,882	301,954	359,452	489,331	510,441	342,751
	146,882	301,954	359,452	489,331	510,441	342,751	251,872
	301,954	359,452	489,331	510,441	342,751	251,872	203,242
saída	=C8	=D8	=E8	=F8	=G8	=H8	=I8
r	0	2000				Mínimo	=MÍNIMO(B2:AEI9)
	0	2000				Máximo	=MÁXIMO(B2:AEI9)
	0	2000					
	0	2000					
	0	2000					
	0	2000					
	0	2000					
	0	2000					

Fonte: Autores (2018)

Após encontrado os menores e maiores valores, chega o momento de normalização. A Figura 34 a seguir mostra essa transformação:

Figura 34 - Exemplo de dados diários do Etanol do Posto D, antes e depois da normalização, no Excel®

330,642		0,185476
242,987		0,136305
399,792		0,224266
537,596	→	0,301568
341,004		0,191288
146,882		0,082394
301,954		0,169383
359,452		0,201637

Fonte: Autores (2018)

Concluído esta etapa, foi definido os novos mínimos e máximos das variáveis. Por terem sido normalizados, tais valores passaram a ser de 0 e 1, respectivamente. Feito a normalização, foi necessário apenas substituir todas as vírgulas presentes nos dados por pontos, uma vez que o *software* Matlab® não compreende valores com virgulas. Esta etapa foi realizada no Excel®, selecionando todos os dados normalizados e utilizando a opção “Localizar e Selecionar”, seguido de “Substituir...” da aba “Página Inicial”. Por fim, pôde-se inserir os valores no programa.

Já no Matlab®, foram criadas as matrizes necessárias. A primeira delas foi a matriz x , que recebeu os valores de entrada e que possuía uma ordem de 7×407 , onde 407 são os volumes de vendas que foram utilizados para ensinar a RNA durante o treinamento e 7, o valor correspondente em dias, representando o número de neurônios da camada de entrada da rede.

Feito isso, foi criada a matriz R de entrada de dados, que tem por intuito definir os números mínimos e máximos aceitos pelo problema em cada neurônio. Sendo constituída na ordem de 7 x 2.

Já a matriz Dx foi constituída da ordem de 1 x 407, sendo cada coluna a saída desejada para cada coluna da matriz x. E suas linhas representando o número de neurônios da camada de saída da rede neural.

E por fim, foi inserido a matriz teste, de ordem 4 x 407, e a matriz dteste, de ordem 1 x 407. Sendo ambas formadas por conjuntos de valores distintos daqueles utilizados no treinamento, para assim validar a rede. A Figura 35 mostra, parcialmente, como ficaram cada matriz aqui acima mencionadas:

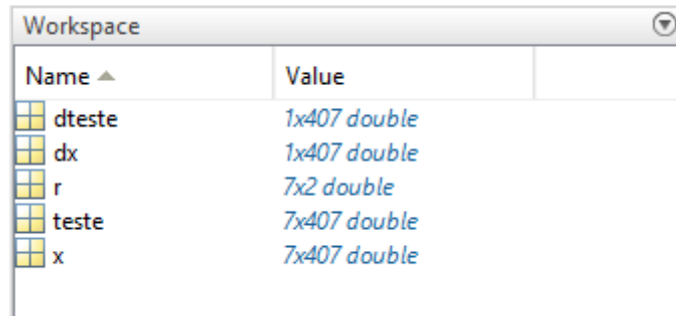
Figura 35 - Criação de cada matriz necessária, no software Matlab®, como os dados do Etanol do Posto D

```
>> x = [0.185475719005761    0.136305092922414
0.136305092922414    0.224265848418384    0.301567872909737
0.224265848418384    0.301567872909737    0.191288348376312
0.301567872909737    0.191288348376312    0.0823943859491661
0.191288348376312    0.0823943859491661    0.169383004145467
0.0823943859491661    0.169383004145467    0.201636870536891
0.169383004145467    0.201636870536891    0.274493316205467
];
>> dx = [0.201636870536891    0.274493316205467
];
>> teste = [0.142999545625382    0.0967099911929858
0.0967099911929858    0.0984130545754402    0.164163866559711
0.0984130545754402    0.164163866559711    0.0876668143851638
0.164163866559711    0.0876668143851638    0.156013171254354
0.0876668143851638    0.156013171254354    0.124840828644673
0.156013171254354    0.124840828644673    0.154282621012302
0.124840828644673    0.154282621012302    0.069894035353
];
>> dteste = [0.154282621012302    0.069894035353
];
>> r = [0    1
0    1
0    1
0    1
0    1
0    1
0    1
];
```

Fonte: Autores (2018)

Enquanto que na Figura 36 a seguir, pode ser visualizada as ordens de cada matriz inserida na rede:

Figura 36 - Ordens das matrizes criadas no software Matlab®, com os dados do Etanol do Posto D



Name ▲	Value
dteste	1x407 double
dx	1x407 double
r	7x2 double
teste	7x407 double
x	7x407 double

Fonte: Autores (2018)

A Figura 37 adiante mostra como ficou a estrutura dos dados para a previsão e para a validação da rede, nas análises diárias:

Figura 37 - Estrutura dos dados diários do Etanol do Posto D, no Excel®

x(7x407)	1ºDia/2016	2ºDia/2016	3ºDia/2016	4ºDia/2016	...	406ºDia/2018	407ºDia/2018
	2ºDia/2016	3ºDia/2016	4ºDia/2016	5ºDia/2016		407ºDia/2018	408ºDia/2018
	3ºDia/2016	4ºDia/2016	5ºDia/2016	6ºDia/2016		408ºDia/2018	409ºDia/2018
	4ºDia/2016	5ºDia/2016	6ºDia/2016	7ºDia/2016		409ºDia/2018	410ºDia/2018
	5ºDia/2016	6ºDia/2016	7ºDia/2016	8ºDia/2016		410ºDia/2018	411ºDia/2018
	6ºDia/2016	7ºDia/2016	8ºDia/2016	9ºDia/2016		411ºDia/2018	412ºDia/2018
	7ºDia/2016	8ºDia/2016	9ºDia/2016	10ºDia/2016		412ºDia/2018	413ºDia/2018
dx(1x407)	8ºDia/2016	9ºDia/2016	10ºDia/2016	11ºDia/2016	...	413ºDia/2018	414ºDia/2018
teste(7x407)	408ºDia/2018	409ºDia/2018	410ºDia/2018	411ºDia/2018	...	813ºDia/2018	814ºDia/2018
	409ºDia/2018	410ºDia/2018	411ºDia/2018	412ºDia/2018		814ºDia/2018	815ºDia/2018
	410ºDia/2018	411ºDia/2018	412ºDia/2018	413ºDia/2018		815ºDia/2018	816ºDia/2018
	411ºDia/2018	412ºDia/2018	413ºDia/2018	414ºDia/2018		816ºDia/2018	817ºDia/2018
	412ºDia/2018	413ºDia/2018	414ºDia/2018	415ºDia/2018		817ºDia/2018	818ºDia/2018
	413ºDia/2018	414ºDia/2018	415ºDia/2018	414ºDia/2018		818ºDia/2018	819ºDia/2018
	414ºDia/2018	415ºDia/2018	414ºDia/2018	415ºDia/2018		819ºDia/2018	820ºDia/2018
dteste(1x407)	415ºDia/2018	414ºDia/2018	415ºDia/2018	416ºDia/2018	...	820ºDia/2018	821ºDia/2018

Fonte: Autores (2018)

Realizado tudo isso, foi utilizado então o *script*, que serve para treinar a rede com diferentes tipos de parâmetros ou pesos. A Figura 38 mostra como ficou definido os parâmetros para utilização em cada aplicação. Na primeira linha, tem-se quantos neurônios terá na camada intermediária (pode ou não ser diferente para cada exemplo). Seguido das funções de saída dos neurônios, tanto da camada de entrada, quanto da de saída. Por fim, tem-se o tipo de treinamento, que em todos os casos serão de *backpropagation*. É definido

também a taxa de aprendizado, o MEMQ e o número máximo de épocas que a rede irá treinar. Esses são os ajustes padrões que precisam ser realizados em cada aplicação, no intuito de se obter os melhores resultados.

Figura 38 - Script para utilização em todas aplicações, no Matlab®

```
net=newff(r, ['?', 1], {'?', '?'}, 'tansig');
net.trainParam.lr = '?';
net.trainParam.goal = '?';
net.trainParam.epochs = '?';
net=init(net);
net=train(net, x, dx);
Y=sim(net, teste)
erro=abs(dteste-Y);
n_acertos=sum(erro<0.2);
perc_acertos=n_acertos/length(erro)
```

Fonte: Matlab® (2018)

Vale ressaltar, que o material de *script* foi apresentado para os autores durante as aulas em sala com uma professora especializada na área.

No que diz respeito as análises semanais, o mesmo já comentado acima foi realizado, com diferenças apenas nos números de entrada, que neste caso foram de apenas 4, em relação ao número de semanas para prever a 5ª Semana. Conforme mostra a Figura 39:

Figura 39 - Parte do agrupamento de dados semanais do Etanol do Posto D, no Excel®

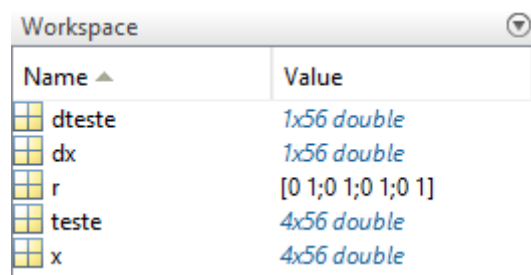
2576,011	1	2576,011	2711,498	2898,766	3492,567	4123,232	2727,451
2711,498	2	2711,498	2898,766	3492,567	4123,232	2727,451	4207,925
2898,766	3	2898,766	3492,567	4123,232	2727,451	4207,925	1930,066
3492,567	4	3492,567	4123,232	2727,451	4207,925	1930,066	2200,309
4123,232	2	2711,498	2898,766	3492,567	4123,232	2727,451	4207,925
2727,451	3	2898,766	3492,567	4123,232	2727,451	4207,925	1930,066
4207,925	4	3492,567	4123,232	2727,451	4207,925	1930,066	2200,309
1930,066	5	4123,232	2727,451	4207,925	1930,066	2200,309	1595,254
2200,309	3	2898,766	3492,567	4123,232	2727,451	4207,925	1930,066
1595,254	4	3492,567	4123,232	2727,451	4207,925	1930,066	2200,309
1645,811	5	4123,232	2727,451	4207,925	1930,066	2200,309	1595,254
1451,604	6	2727,451	4207,925	1930,066	2200,309	1595,254	1645,811

Fonte: Autores (2018)

Como o total de dados semanais equivale a 117 dados, quando agrupados, estes dados completam 114 colunas, mas como mencionado na aplicação diária, a última coluna recebe o

valor 0 (zero) na camada de saída por não ter referencial na coluna seguinte. Logo, ao excluir tal coluna, o resultado restante é de 113 colunas. Entretanto, para tentar igualar o feito realizado na análise diária, a 113ª coluna foi removida, com o intuito de se ter o mesmo número de colunas tanto para treinamento da rede, quanto para testes e validação. As demais etapas seguiram a mesma sequência feitas nas análises diárias. A única diferença foi na ordem das matrizes, que neste caso ficaram como mostra a Figura 40:

Figura 40 - Ordem das matrizes dos dados semanais do Etanol do Posto D, no Matlab®



Name	Value
dteste	1x56 double
dx	1x56 double
r	[0 1;0 1;0 1]
teste	4x56 double
x	4x56 double

Fonte: Autores (2018)

A Figura 41 a seguir mostra como ficou a estrutura dos dados semanais do combustível etanol hidratado comb. marcado do Posto D para a previsão e para a validação da rede:

Figura 41 - Estrutura dos dados semanais do Etanol, no Excel®

x(4x56)	1ªSemana/2016	2ªSemana/2016	3ªSemana/2016	...	56ªSemana/2018	57ªSemana/2018
	2ªSemana/2016	3ªSemana/2016	4ªSemana/2016		57ªSemana/2018	58ªSemana/2018
	3ªSemana/2016	4ªSemana/2016	5ªSemana/2016		58ªSemana/2018	59ªSemana/2018
	4ªSemana/2016	5ªSemana/2016	6ªSemana/2016		59ªSemana/2018	60ªSemana/2018
dx(1x56)	5ªSemana/2016	6ªSemana/2016	7ªSemana/2016	...	60ªSemana/2018	61ªSemana/2018
teste(4x56)	58ªSemana/2018	59ªSemana/2018	60ªSemana/2018	...	111ªSemana/2018	112ªSemana/2018
	59ªSemana/2018	60ªSemana/2018	61ªSemana/2018		112ªSemana/2018	113ªSemana/2018
	60ªSemana/2018	61ªSemana/2018	62ªSemana/2018		113ªSemana/2018	114ªSemana/2018
	61ªSemana/2018	62ªSemana/2018	63ªSemana/2018		114ªSemana/2018	115ªSemana/2018
dteste(1x56)	62ªSemana/2018	63ªSemana/2018	64ªSemana/2018	...	115ªSemana/2018	116ªSemana/2018

Fonte: Autores (2018)

5 RESULTADOS

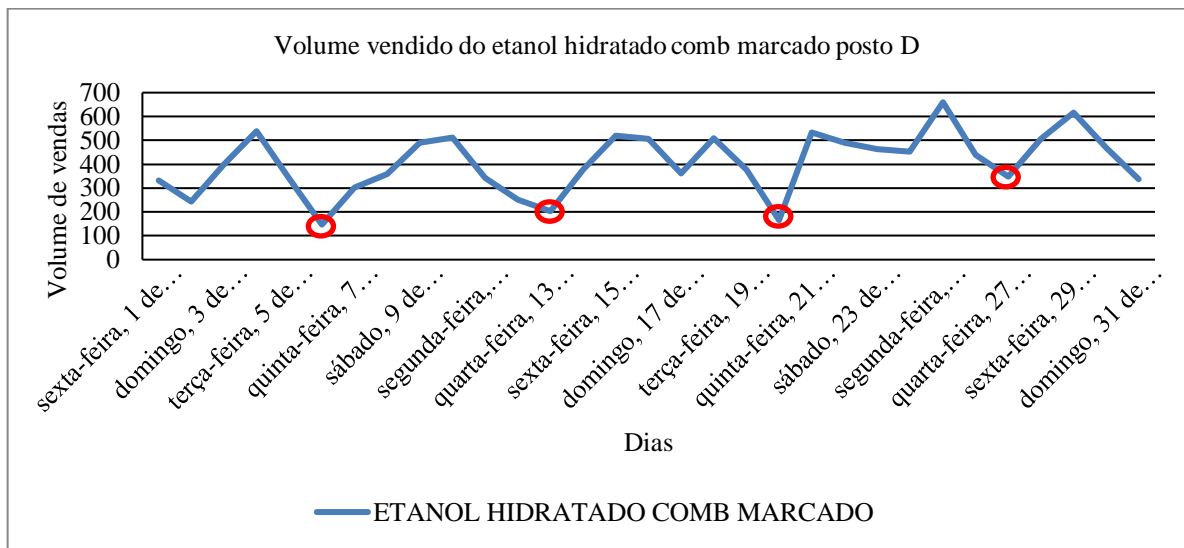
Abaixo serão listados e analisados todos os resultados alcançados a partir das simulações realizadas, seguindo a mesma metodologia feita durante a aplicação dos métodos matemáticos, onde foram listados os resultados alcançados para o combustível Etanol hidratado comb. Marcado do posto D durante o desenvolvimento do trabalho desde a coleta dos dados, passando pelas aplicações e finalizando nos valores de erro alcançados em cada um dos métodos. Logo, têm-se todos os resultados referentes aos demais combustíveis de cada posto nos apêndices A, B, C, D, E, F, G, H, I, J, K, L, M.

5.1 COLETA DE DADOS

Durante a coleta de dados diários, foi possível observar que quando se colocou os dados referentes ao etanol hidratado comb. Marcado do posto D em um gráfico, foi possível observar, como no Gráfico 1, uma leve sazonalidade durante a semana onde em uma parcela das quartas-feiras costumam possuir um menor volume de vendas que os demais dias. Essa sazonalidade acaba por variar de produto para produto e no decorrer do tempo, porém pode ser observada principalmente nos produtos que possuem um maior volume de vendas. Essa foi uma análise que contribuiu para a adoção de uma análise diária dos dados.

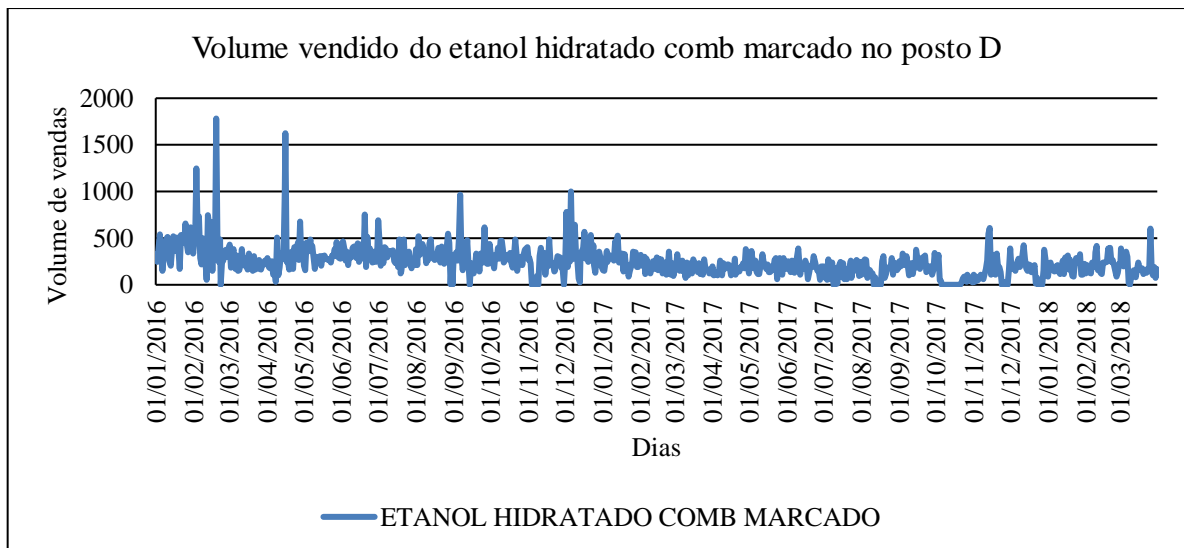
Em relação aos dados semanais, foi menor o número de casos onde é possível observar a ocorrência de sazonalidade. Fato que pode ser confirmado durante a aplicação dos modelos matemáticos. O Gráfico 2 apresenta a séria histórica de dados referentes ao etanol hidratado comb. marcado do posto D.

Gráfico 1 - Dados do etanol hidratado comb. Marcado do posto D de 01/01/2016 até 31/01/2016, contendo sazonalidades em destaque



Fonte: Autores (2018)

Gráfico 2 - Dados do etanol hidratado comb. Marcado do posto D



Fonte: Autores (2018)

5.2 SÉRIES TEMPORAIS

Depois de feita toda a etapa de aplicação, referente aos dados diários, foram obtidos resultados muito elevados, com erros de previsões muito discrepantes, em relação às fórmulas aplicadas no Excel®, podendo ser explicado por conta dos *outliers*. Problema este que foi resolvido com a utilização da ferramenta Crystal Ball®, sendo esta responsável por realizar a identificação e extração dos *outliers* e fazer uma aplicação mais eficiente das fórmulas, que

acabou por culminar em uma grande redução do erro de previsão. O Quadro 3 demonstra a porcentagem de erro obtido para cada método aplicado, tanto para análise diário, quanto semanal, ao combustível etanol hidratado comb. Marcado do posto D, tanto das formulas utilizadas no Excel®, quanto da utilização do Crystal Ball®.

Quadro 3 - Erros obtidos por meio das fórmulas aplicadas no Excel® e do Crystal Ball® para os dados diários e semanais referentes ao etanol do posto D

Posto D	Diário		Semanal	
	Etanol			
	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®
Média Móvel Simples	461,19%	32,04%	24,25%	20,53%
Média Móvel Dupla	279%	33,60%	22,16%	21,95%
Média Móvel Ponderada	187,52%	X	23,49%	X
Suavização Exponencial Simples	187,22%	31,07%	23,28%	20,30%
Suavização Exponencial com Tendência	177,26%	31,14%	50,63%	20,32%
Suavização Exponencial com Tendência e Sazonalidade	307,71%	Ignorado	23,39%	Ignorado

Fonte: Autores (2018)

Em relação a análise semanal é possível observar que a mesma alcançou resultados melhores, em todos os métodos, que os demonstrados na análise diária. Isso ocorreu, pois quando se faz a somatória dos dados diários para gerar um semanal, essa soma acaba por tornar os dados mais homogêneos, facilitando assim a previsão dos mesmos.

Comparando os resultados obtidos por meio das fórmulas aplicadas no Excel® e do Crystal Ball® é possível perceber que todos os resultados provenientes deste último foram melhores em comparação aos anteriores. Por conta do Crystal Ball® fazer a eliminação de *outliers* que restaram após a somatória e por também fazer uma aplicação mais eficiente das fórmulas, como já havia sido para os dados diários.

5.3 MÉTODOS CAUSAIS

Foi possível observar que mesmo com a escassez da variedade de dados diários para serem utilizados como variáveis independentes foi possível obter resultados parcialmente satisfatórios. O Quadro 4 mostra a porcentagem de erro obtido por meio da regressão para os dados diários referentes ao etanol hidratado comb. marcado do posto D.

Quadro 4 - Erros obtidos por meio do Crystal Ball® para os dados diários referentes ao etanol do posto D

	Etanol
Posto D	40,49%

Fonte: Autores (2018)

Em relação à aplicação de métodos causais por meio do Crystal Ball® aos dados semanais é possível constatar que este obteve melhores resultados em relação aos dados diários, o que pode ser explicado novamente pela homogeneização dos dados, proveniente da somatória dos dados diários para gerar os semanais. Comparando-se os modelos matemáticos utilizados, nota-se que a regressão linear múltipla alcançou um resultado ligeiramente melhor que a simples, podendo ser explicado pelo fato da múltipla basear sua previsão em uma

quantidade maior de variáveis independentes que exercem influência no volume de etanol vendido.

O Quadro 5 mostra os erros obtidos após a aplicação dos métodos causais aos dados semanais.

Quadro 5 - Erros obtidos por meio do Crystal Ball® para os métodos de regressão linear simples e múltipla sobre os dados semanais referentes ao etanol do posto D

	Etanol	
	R. L. Simples	R. L. Múltipla
Posto D	31,52%	31,01%

Fonte: Autores (2018)

5.4 REDES NEURAIS ARTIFICIAIS

Nas aplicações diárias, do combustível etanol hidratado comb. marcado do Posto D, os parâmetros de treinamento foram ajustados e executados diversas vezes até se atingir o melhor resultado possível em taxas de acerto. Os melhores desempenhos foram obtidos com os seguintes parâmetros, como mostra o Quadro 6, onde pode-se observar que o resultado atingido superou os 90% de acerto.

Quadro 6 - Parâmetros utilizados para obtenção dos melhores resultados, dos dados do etanol do posto D, no Matlab®

Nº de Neurônios	Taxa de Aprendizado	MEMQ	Nº máximo de épocas	Treinamento Utilizado	Função de Ativação	Margem de erro	Resultados
30	0,001	0,001	3500	<i>Backpropagation com momentum</i>	Sigmoidal	0,2	0,95749

Fonte: Autores (2018)

Como visto acima, o resultado obteve 96% de acerto, com apenas 4% de erro na previsão. Um resultado muito satisfatório, uma vez que o problema em si é bastante complexo. Porém, este resultado é uma média de 10 resultados distintos obtidos a partir dos melhores parâmetros encontrados, como pode ser visualizado no Quadro 7 adiante. Assim como o desvio padrão obtido desses mesmos 10 resultados, que foi de 0,0566.

Quadro 7 - Resultados obtidos a partir dos melhores parâmetros encontrados, nos dados do etanol do posto D. No Excel®

POSTO D	RESULTADOS
Etanol	0,9558
Etanol	0,9877
Etanol	0,9951
Etanol	0,9607
Etanol	0,9877
Etanol	0,8452
Etanol	0,8624
Etanol	0,9926
Etanol	0,9926
Etanol	0,9951
MÉDIA	0,95749
DESVIO PADRÃO	0,0566

Fonte: Autores (2018)

No que diz respeito à análise semanal, os mesmos testes em relação aos parâmetros foram realizados, com o intuito de se encontrar os resultados mais satisfatórios. O Quadro 8 adiante mostra os parâmetros definidos para o exemplo semanal do combustível etanol hidratado comb. marcado do posto D.

Quadro 8 - Parâmetros que resultaram nas melhores taxas de acerto, na análise semanal, do etanol do posto D. No Excel®

Nº de Neurônios	Taxa de Aprendizado	MEMQ	Nº máximo de épocas	Treinamento Utilizado	Função de Ativação	Margem de erro	Resultados
30	0,1	0,001	3000	<i>Backpropagation com momentum</i>	Sigmoidal	0,2	0,78393

Fonte: Autores (2018)

Já em relação aos seus resultados, os mesmos podem ser vistos no Quadro 9 a seguir, onde a média foi de 78%, com erro de 22% na previsão. E desvio padrão de 0,0660.

Quadro 9 - Resultados obtidos nos dados semanais do etanol do posto D, no Excel®

POSTO D	RESULTADOS
Etanol	0,7679
Etanol	0,875
Etanol	0,7321
Etanol	0,7679
Etanol	0,7321
Etanol	0,7679
Etanol	0,75
Etanol	0,7857
Etanol	0,9286
Etanol	0,7321
MÉDIA	0,78393
DESVIO PADRÃO	0,0660

Fonte: Autores (2018)

6 ANÁLISES COMPARATIVAS

Após a obtenção de todos os resultados alcançados por cada um dos métodos, tanto diária, quanto semanalmente, faz-se necessário então, comparar os melhores desempenhos de cada, para assim determinar qual método quantitativo de previsão conseguiu ser mais eficiente no combustível etanol hidratado comb. marcado, do posto D. Conforme demonstra o Quadro 10 adiante, na análise diária, quem obteve o melhor resultado dentre todos foram as redes neurais artificiais, com 4,25% de erro apenas. Tendo um desempenho bem acima quando comparado com os demais. Enquanto que o método suavização exponencial com tendência, de séries temporais, obteve o pior resultado, com 177,26% de erro na previsão. O que pode ser explicado pelo fato de terem sido executados ainda com a presença dos *outliers*, no Excel®.

Quadro 10 - Comparação entre os melhores resultados obtidos em cada método, diária e semanalmente, para os dados do etanol do posto D, no Excel®

Posto D		Séries temporais		Métodos causais	Redes Neurais Artificiais
		Fórmulas Excel®	Crystal Ball®		
Etanol	Diário	177,26%	31,07%	40,49%	4,25%
		Suavização Exponencial com tendência	Suavização Exponencial Simples		
	Semanal	22,16%	20,30%	31,01%	21,61%
		Média Móvel Dupla	Suavização Exponencial Simples	Regressão Linear Múltipla	

Fonte: Autores (2018)

Já nas análises semanais, os métodos tiveram resultados bem próximos e satisfatórios, e quem teve melhor desempenho na previsão foi o método de suavização exponencial

simples, das séries temporais, com erro de 20,30%, com o uso do Crystal Ball®. O que pode ser explicado pelo motivo da retirada dos *outliers*. Enquanto aquele que demonstrou menor acerto foi a análise de regressão linear múltipla, dos métodos causais, com 31,01% de erro.

A partir destes resultados, é possível concluir então que para o combustível etanol hidratado comb. marcado do posto D, o resultado obtido pelas redes neurais artificiais, na análise diária, seria utilizado como base para a previsão de demanda para este combustível em especial.

Tendo agora uma visão geral de todos os resultados obtidos para todos os postos, de acordo com os apêndices N, O, P, Q e R, pode-se concluir que o método mais eficiente nas previsões foram as redes neurais artificiais, que conseguiram obter os melhores resultados em grande maioria dos exemplos, sendo 20 de um total de 38. Seguido pelas séries temporais, aplicadas por meio do Crystal Ball®, que alcançaram um total de 15 melhores desempenhos em prever a demanda. Os 3 exemplos restantes, tiveram o melhor resultado por meio das séries temporais, aplicadas pelo Excel®. Enquanto; os métodos causais, em nenhum caso, conseguiram obter o menor erro de previsão.

No que diz respeito às análises diárias e semanais, percebeu-se que as semanais conseguiram se destacar, e dentre as 19 previsões realizadas para cada combustível de cada posto, em 14 situações o melhor resultado foi consequência dessa análise. O que pode ser explicado pelo fato do agrupamento dos dados, que os torna mais homogêneos, facilitando assim a previsão.

Porém, mesmo que consigam obter melhores resultados, as análises semanais não são capazes de identificar detalhes relacionados aos seus dados, dificultando na tomada de decisões, uma vez que não se tem conhecimento do real comportamento dos dados, quando comparado às análises diárias. Um exemplo disso seria nos dias em que há menor número de vendas, onde poderiam ser adotadas políticas que atraíssem um número maior de clientes. Aumentando assim o volume de vendas. Por fazer o agrupamento desses dados de forma semanal, não teria como saber com exatidão quais dias obteve menor número de vendas.

Enquanto as análises diárias obtiveram 5 melhores resultados, sendo 4 deles referentes ao etanol hidratado comb. marcado e 1 referente a gasolina aditivada DT clean, sendo relacionado ao baixo volume de vendas deste combustível, em comparação com os demais. A partir de uma análise feita nos dados presentes nos apêndices S, T, V e W, deduz-se então, que os melhores resultados em tais combustíveis, se deu por conta da baixa variabilidade ao longo da série histórica, o que permitiu a análise ser bem precisa em suas previsões.

Já em relação aos melhores métodos de séries temporais, pode-se dar destaque à média móvel dupla, que obteve os melhores desempenhos ao longo dos exemplos. Enquanto média móvel ponderada obteve o pior desempenho quando comparado com os demais métodos.

Por fim, nos métodos causais, a regressão linear múltipla conseguiu atingir um desempenho muito melhor, quando comparada com regressão linear simples, pela mesma ser influenciada por mais de uma variável independente.

Para finalizar as análises, foram preenchidos alguns pontos positivos e negativos em relação a cada método, como pode ser visto no Quadro 11:

Quadro 11 - Análise comparativa do ponto de vista dos autores entre os métodos utilizados

Séries temporais	Métodos causais	Redes neurais artificiais
Antes da utilização	Antes da utilização	Antes da utilização
Necessário realizar o levantamento dos dados que se deseja prever e analisar se os mesmos necessitam de ajustes.	Necessário realizar o levantamento dos dados que se deseja prever e de dados que servirão como variáveis independentes. Também é necessário analisar se os mesmos precisam de ajustes.	Necessário realizar o levantamento dos dados que se deseja prever e analisar se os mesmos necessitam de ajustes.
Durante a utilização	Durante a utilização	Durante a utilização
Possui métodos simples de serem aplicados, utilizando apenas um conjunto de dados e as formulas necessárias; Podem ser aplicados por meio de vários tipos de ferramentas, desde as mais simples até mais complexas; É necessário se aplicar vários modelos matemáticos para poder se alcançar o melhor valor de previsão, o que geralmente acaba demandando um maior tempo que os demais métodos;	Podem ser aplicados por meio de vários tipos de ferramentas, desde as mais simples até as mais complexas; Possui uma dificuldade mediana de aplicação, pois são necessários no mínimo dois conjuntos de dados que possuam correlação entre si, caso contrário não se chegará a um resultado satisfatório;	Possui a maior dificuldade de aplicação entre os três métodos, por conta das várias etapas de aplicação que vão desde o agrupamento dos dados até os ajustes de parâmetros na rede neural. (Pode ser que seja diferente para outros <i>softwares</i> pelo fato de apenas ter sido aplicado no <i>software</i> MATLAB®); Apenas podem ser aplicadas em um número limitado de ferramentas que possuem as características necessárias para aplicar este método.
Após a utilização	Após a utilização	Após a utilização
Os resultados finais alcançados acabam sendo muito variáveis (A maioria dos melhores resultados foi com a utilização do Crystal Ball®, já dos piores foram com as formulas aplicadas no Excel®).	Foi o único método que não alcançou nenhum melhor resultado dentre todos os outros obtidos para prever os volumes vendidos de todos os combustíveis.	Alcançou os melhores resultados na maioria das previsões feitas.

Fonte: Autores (2018)

Para complementar a análise, tem-se pontos positivos e negativos de cada ferramenta utilizada nas implementações de cada método, afim de compartilhar as opiniões sobre as mesmas. Conforme pode ser visualizado no Quadro 12 adiante:

Quadro 12 - Análise comparativa do ponto de vista dos autores entre as ferramentas utilizadas para os métodos

Excel®	Crystal Ball®	Matlab®
Antes da utilização	Antes da utilização	Antes da utilização
É facilmente encontrado instalado na grande maioria dos computadores, porém sua licença possui um valor muito elevado.	Precisa ser instalado no computador e o valor de sua licença é relativamente menor em comparação aos outros.	Precisa ser instalado no computador, seu programa é relativamente pesado e possui uma licença de valor elevada.
Durante a utilização	Durante a utilização	Durante a utilização
É simples de se utilizar quando já se possui um conhecimento básico, porém é necessário se ter atenção, pois o mesmo, para a aplicação de formulas fica muito suscetível a erros e demanda bastante tempo para montá-las e revisá-las.	É extremamente simples e rápido para gerar resultados, além de oferecer um grande número de auxílios durante as aplicações.	Possui comandos com os quais não estamos habituados, além de serem mais complexos.
Após a utilização	Após a utilização	Após a utilização
A ferramenta apresentou resultados em sua maioria muito elevados.	Os resultados gerados por essa ferramenta foram em sua maioria interessantes para a previsão de demandas.	A maioria dos melhores resultados encontrados foram obtidos por meio desta ferramenta.

Fonte: Autores (2018)

A partir dos quadros acima é possível se ter um bom entendimento, a partir do ponto de vista dos autores, sobre a utilização de cada um dos métodos aplicados e das ferramentas utilizadas juntamente com os mesmos.

7 CONSIDERAÇÕES FINAIS

Conforme visualizado ao longo do trabalho, e principalmente após os resultados finais obtidos, uma boa previsão de demandas pode ser responsável por guiar uma empresa para o caminho mais eficiente possível, uma vez que possibilita a mesma ter maior controle de suas ações, permitindo assim uma organização maior no que diz respeito aos seus processos como um todo.

A diminuição de estoque, custos demasiados, aumento da satisfação dos clientes, baixos índices de perda de produto, são apenas algumas das inúmeras vantagens que uma previsão adequada é responsável por alcançar. Situações nas quais um produto não é vendido por dias, semanas, poderia facilmente ocorrer em menor escala, uma vez que a empresa consegue compreender de maneira mais inteligente o mercado que o circunda. Sendo assim, um fator de extrema relevância dentro de uma organização que busca se destacar ainda mais no ramo em que atua.

A partir dos resultados alcançados é possível então fazer diversas observações acerca dos métodos utilizados neste trabalho. De maneira geral, o método de redes neurais artificiais conseguiu alcançar os melhores resultados na maioria das previsões realizadas. Seguido pelas séries temporais, que alcançaram resultados bem satisfatórios nas aplicações sem a presença de *outliers*. Enquanto os métodos causais, em nenhum exemplo, conseguiram atingir um desempenho superior aos demais. É possível então inferir, que o método baseado em aprendizagem acaba por se tornar a melhor alternativa para fazer uma previsão de demanda, por apresentar um maior percentual de acerto que os demais métodos.

Sobre os tipos de análise diária e semanal, a última se destaca por obter os melhores resultados, pois como foi possível observar durante o desenvolvimento deste trabalho, a mesma é responsável por realizar uma somatória que eventualmente acarreta numa homogeneização dos dados e, logo, na diminuição da quantidade de *outliers*. Este fato acaba facilitando as previsões realizadas por meio dessa análise, o que explica seus melhores resultados.

Em relação aos métodos causais é importante salientar que o fato de seus resultados não terem alcançado o que se esperava, foi proveniente da falta de disponibilidade de outros dados que poderiam ser utilizados como variáveis independentes, o que poderia resultar em melhores resultados de previsão. Pode-se inferir também que o volume de vendas não possui

uma grande correlação com o valor de venda praticado na rede de postos. Esta situação poderia ser diferente, caso fossem utilizadas variáveis como: localização e atendimento.

Ressalta-se ainda que seria interessante a aplicação dos dados obtidos no presente trabalho em uma política de ressurgimento, com o intuito de se verificar como a mesma impactaria nos custos, logo, ficando como sugestão para novos trabalhos analisarem quais seriam os resultados alcançados a partir dessa aplicação.

Este trabalho fica a disposição para servir de base para novas pesquisas relacionadas a previsão de demandas, que tenham como objetivo avaliar e analisar diferentes modelos matemáticos que venham a obter resultados cada vez melhores. Com o intuito de se expandir o conhecimento relacionado a esta tão importante área da engenharia de produção.

REFERÊNCIAS

- ANP. **Série histórica do levantamento de preços e de margens de comercialização de combustíveis.** Disponível em <http://www.anp.gov.br/precos-e-desfesa/234-precos/levantamento-de-precos/868-serie-historica-do-levantamento-de-precos-e-de-margens-de-comercializacao-de-combustiveis>. Acesso em: 19 de novembro de 2018.
- BRAGA, Antônio; CARVALHO, André; LUDERMIR, Teresa. **Redes neurais artificiais: teoria e aplicações.** Rio de Janeiro: Livros Técnicos e Científicos, 2000.
- CHASE, Richard B.; JACOBS, Robert F.; AQUILANO, Nicholas J. **Fundamentos da Administração da Produção.** 3. ed. Porto Alegre: Bookman Editora, 2001.
- CORRAR, Luiz; PAULO, Edilson; FILHO, José. **Análise Multivariada para os Cursos de Administração, Ciências Contábeis e Economia.** São Paulo: Atlas, 2007.
- CORRAR, Luiz; THEÓFILO, Carlos R. **Pesquisa operacional para decisão em contabilidade e administração.** São Paulo: Editora Atlas, 2004.
- DAVIS, Mark; AQUILANO, Nicolas; CHASE, Richard. **Fundamentos da administração da produção.** 3. ed. Porto Alegre: Bookman Editora, 2001.
- DIAS, Marco Aurélio P. **Administração de materiais: uma abordagem logística.** 5. ed. São Paulo: Atlas, 2010.
- GIL, Antônio. **Como elaborar projetos de pesquisa.** 5. ed. São Paulo: Atlas, 2010.
- HAYKIN, Simon. **Redes Neurais Princípios e prática.** 2. ed. Porto Alegre: Bookman Editora, 2001.
- LUSTOSA, Leonardo *et al.* **Planejamento e controle da produção.** 2. ed. Rio de Janeiro: Elsevier, 2008.
- MARTINS, Petrônio G.; ALT, Paulo. **Administração de materiais e recursos patrimoniais.** 1. ed. São Paulo: Saraiva, 2004.
- MARTINS, Petrônio G.; LAUGENI, Fernando P. **Administração da produção.** São Paulo: Saraiva, 2005.
- MOREIRA, Daniel. **Administração da produção e operações.** 1. ed. São Paulo: Thonson Learning, 2006.
- RAGSDALE, Cliff. **Modelagem de planilha e análise de decisão: Uma introdução prática a *business analytics*.** 7. ed. São Paulo: Cengage Learnig, 2017.
- REZENDE, Solange. **Sistemas inteligentes: fundamentos e aplicações.** Barueri, SP: Manole, 2005.

RITZMAN, Barbara; KRAJEWSKI, Judie. **Administração da Produção e Operações**. 3. ed. São Paulo: Pearson Prentice Hall, 2008.

SOUSA, Tulio Franco de; GUIMARAES, Thayana Araujo. **Escolha e utilização de um modelo de previsão de demanda em serviços: estudo de caso em uma empresa prestadora de serviços de recapagem de pneus**. XXX Encontro Nacional de Engenharia de Produção, São Carlos, p. 4, 12 a 15 de outubro de 2010.

TUBINO, Dalvio. **Planejamento e controle da produção: teoria e prática**. 1. ed. São Paulo: Atlas, 2007.

**APÊNDICE A - Quadro referente à aplicação de séries temporais, em dados diários,
com e sem utilização do Crystal Ball® no posto A**

	Óleo Diesel		Etanol		Gasolina Aditivada		Gasolina Comum	
	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®
Média Móvel Simples	1591,46%	34,71%	112,88%	30,16%	3095,31%	21,10%	349,84%	20,95%
Média Móvel Dupla	1475,99%	35,66%	104,02%	29,53%	2848,12%	20,67%	357,21%	21,03%
Média Móvel Ponderada	1762,59%	X	109,71%	X	989,11%	X	349,84%	X
Suavização Exponencial Simples	1620,23%	36,44%	123,28%	29,80%	3317,97%	22,39%	3773,31%	22,61%
Suavização Exponencial com Tendência	749,38%	36,23%	138,34%	29,77%	2359,67%	21,87%	128,15%	22,33%
Suavização Exponencial com Tendência e Sazonalidade	590,77%	106,93%	473,25%	Ignorado	2246,20%	Ignorado	2669,81%	87%

**APÊNDICE B - Quadro referente à aplicação de séries temporais, em dados diários,
com e sem utilização do Crystal Ball® no posto B**

	Óleo Diesel		Etanol		Gasolina Aditivada		Gasolina Comum	
	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®
Média Móvel Simples	94,92%	29,33%	67,31%	28,26%	31,05%	21,54%	36,32%	20,54%
Média Móvel Dupla	95,57%	30,36%	48,36%	28,52%	24,38%	22,03%	31,68%	20,54%
Média Móvel Ponderada	126,20%	X	58,63%	X	32,25%	X	41,63%	X
Suavização Exponencial Simples	93,78%	29,80,%	59,03%	27,72%	31,21%	22,22%	36,86%	21,42%
Suavização Exponencial com Tendência	86,04%	29,75%	58,75%	27,71%	31,27%	22,18%	36,47%	21,21%
Suavização Exponencial com Tendência e Sazonalidade	148,75%	27,95%	63,41%	Ignorado	39,10%	19,78%	34,41%	14,47%

APÊNDICE C – Quadro referente à aplicação de séries temporais, em dados diários, com e sem utilização do Crystal Ball® no posto C

	Óleo Diesel		Etanol		Gasolina Aditivada		Gasolina Comum	
	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®
Média Móvel Simples	2402,99%	47,09%	174,15%	31,22%	2233,59%	28,36%	30,19%	21,87%
Média Móvel Dupla	1740,32%	47,96%	103,97%	32,91%	1475,90%	28,77%	27,75%	22,27%
Média Móvel Ponderada	2100,17%	X	66,56%	X	1415,37%	X	36,63%	X
Suavização Exponencial Simples	2096,86%	48,79%	66,61%	31,77%	1412,04%	29,12%	33,39%	22,70%
Suavização Exponencial com Tendência	2096,69%	48,44%	67,44%	31,86%	535,19%	28,89%	33,32%	22,64%
Suavização Exponencial com Tendência e Sazonalidade	1270,49%	36,77%	200,81%	Ignorado	3224,96%	189,93%	41,57%	18,43%

APÊNDICE D - Quadro referente à aplicação de séries temporais, em dados diários, com e sem utilização do Crystal Ball® no posto D

	Óleo Diesel		Etanol		Gasolina Aditivada	
	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®
Média Móvel Simples	891,52%	48,41%	461,19%	32,04%	24,38%	17,53%
Média Móvel Dupla	682,52%	48,48%	279%	33,60%	21,50%	17,78%
Média Móvel Ponderada	829,67%	X	187,52%	X	28,43%	X
Suavização Exponencial Simples	927,14%	49,53%	187,22%	31,07%	24,37%	17,92%
Suavização Exponencial com Tendência	194,61%	49,11%	177,26%	31,14%	24,38%	17,90%
Suavização Exponencial com Tendência e Sazonalidade	822,73%	41,54%	307,71%	Ignorado	23,98%	14,93%

APÊNDICE E - Quadro referente à aplicação de séries temporais, em dados diários, com e sem utilização do Crystal Ball® no posto E

	Óleo Diesel		Etanol		Gasolina Aditivada		Gasolina Comum	
	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®
Média Móvel Simples	1781,94%	42,49%	90,66%	36,86%	258,47%	27,52%	261,50%	26,06%
Média Móvel Dupla	1356,11%	42,67%	66,98%	40,12%	260,22%	27,06%	243,60%	25,99%
Média Móvel Ponderada	1769,55%	X	84,55%	X	270,94%	X	324,80%	X
Suavização Exponencial Simples	1775,29%	45,41%	86,32%	37,70%	263,48%	28,38%	274,36%	27,56%
Suavização Exponencial com Tendência	1785,26%	45,28%	85,07%	37,67%	204,93%	28,42%	194,40%	27,16%
Suavização Exponencial com Tendência e Sazonalidade	1493,27%	35,70%	418,78%	Ignorado	115,63%	22,88%	122,12%	17,27%

APÊNDICE F - Quadro referente à aplicação de séries temporais, em dados semanais, com e sem utilização do Crystal Ball® no posto A

	Óleo Diesel		Etanol		Gasolina Aditivada		Gasolina Comum	
	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®
Média Móvel Simples	16,21%	14,55%	27,25%	18,77%	12,75%	10,58%	17,67%	9,33%
Média Móvel Dupla	15,39%	15,93%	23,32%	18,97%	10,91%	12,86%	14,94%	13%
Média Móvel Ponderada	16,82%	X	25,75%	X	11,98%	X	12,43%	X
Suavização Exponencial Simples	15,97%	14,34%	24,94%	17,58%	11,48%	10,61%	12,26%	9,26%
Suavização Exponencial com Tendência	19,09%	14,34%	38,29%	17,61%	12,36%	10,61%	12,39%	9,26%
Suavização Exponencial com Tendência e Sazonalidade	17%	Ignorado	23,63%	Ignorado	11,23%	Ignorado	12,18%	Ignorado

APÊNDICE G - Quadro referente à aplicação de séries temporais, em dados semanais, com e sem utilização do Crystal Ball® no posto B

	Óleo Diesel		Etanol		Gasolina Aditivada		Gasolina Comum	
	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®
Média Móvel Simples	15,76%	12,15%	18,09%	15,67%	15,55%	12,48%	11%	9,63%
Média Móvel Dupla	14,28%	11,96%	16,56%	17,04%	13,49%	11,62%	9,51%	11,18%
Média Móvel Ponderada	16,77%	X	19,85%	X	15,56%	X	9,65%	X
Suavização Exponencial Simples	15,44%	13,34%	18,20%	16,08%	14,90%	12,71%	9,85%	9,47%
Suavização Exponencial com Tendência	18,35%	13,35%	46,50%	16,19%	51,48%	12,72%	9,90%	9,47%
Suavização Exponencial com Tendência e Sazonalidade	15,50%	Ignorado	17,91%	Ignorado	14%	Ignorado	9,59%	Ignorado

APÊNDICE H - Quadro referente à aplicação de séries temporais, em dados semanais, com e sem utilização do Crystal Ball® no posto C

	Óleo Diesel		Etanol		Gasolina Aditivada		Gasolina Comum	
	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®
Média Móvel Simples	16,98%	15,47%	25,78%	20,14%	17,61%	14,16%	11,05%	10,09%
Média Móvel Dupla	15,79%	15,03%	24,70%	22,10%	16,09%	14,85%	9,58%	10,90%
Média Móvel Ponderada	20,75%	X	24,65%	X	19,73%	X	11,28%	X
Suavização Exponencial Simples	17,36%	15,63%	25,42%	19,65%	17,18%	14,45%	10,49%	10,30%
Suavização Exponencial com Tendência	23,15%	15,31%	46,79%	19,71%	54,62%	14,45%	12,20%	10,31%
Suavização Exponencial com Tendência e Sazonalidade	16,63%	Ignorado	32,89%	Ignorado	17,21%	Ignorado	10,27%	Ignorado

**APÊNDICE I - Quadro referente à aplicação de séries temporais, em dados semanais,
com e sem utilização do Crystal Ball® no posto D**

	Óleo Diesel		Etanol		Gasolina Aditivada	
	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®
Média Móvel Simples	26,43%	19,34%	24,25%	20,53%	8,81%	8,42%
Média Móvel Dupla	23,93%	20,39%	22,16%	21,95%	7,84%	8,63%
Média Móvel Ponderada	28,06%	X	23,49%	X	8,91%	X
Suavização Exponencial Simples	25,82%	19,66%	23,28%	20,30%	8,50%	8,04%
Suavização Exponencial com Tendência	33,40%	19,55%	50,63%	20,32%	47,46%	8,04%
Suavização Exponencial com Tendência e Sazonalidade	27,52%	Ignorado	23,39%	Ignorado	8,12%	Ignorado

APÊNDICE J - Quadro referente à aplicação de séries temporais, em dados semanais, com e sem utilização do Crystal Ball® no posto E

	Óleo Diesel		Etanol		Gasolina Aditivada		Gasolina Comum	
	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®	Formula Excel®	Crystal Ball®
Média Móvel Simples	42,07%	13,71%	66,57%	23,32%	14,25%	12,15%	11,08%	9,77%
Média Móvel Dupla	38,50%	13,16%	61,03%	26,23%	13,39%	11,03%	9,71%	10,16%
Média Móvel Ponderada	41,53%	X	68,61%	X	15,03%	X	10,74%	X
Suavização Exponencial Simples	41,32%	14,73%	63,41%	25,84%	14,45%	12,41%	10,38%	9,87%
Suavização Exponencial com Tendência	45,73%	14,75%	73,20%	25,92%	16,47%	12,42%	10,98%	9,88%
Suavização Exponencial com Tendência e Sazonalidade	49,33%	Ignorado	68,03%	Ignorado	14,23%	Ignorado	10,15%	Ignorado

APÊNDICE K - Quadro com erros MAPE obtidos referentes à aplicação de métodos causais, em dados diários, com utilização do Crystal Ball® em todos os combustíveis de cada posto

	Óleo Diesel	Etanol	Gasolina Aditivada	Gasolina Comum
Posto A	37,46%	51,28%	37,75%	35,02%
Posto B	30,19%	32,67%	38,23%	23,77%
Posto C	48,41%	43,56%	36,46%	24,33%
Posto D	50,78%	40,49%	20,13%	X
Posto E	44,33%	47,14%	36,12%	27,28%

APÊNDICE L - Quadro com erros MAPE obtidos referentes à aplicação de métodos causais, em dados semanais, com a utilização do Crystal Ball® em todos os combustíveis de cada posto

	Óleo Diesel		Etanol		Gasolina Aditivada		Gasolina Comum	
	R. L. Simples	R. L. Múltipla	R. L. Simples	R. L. Múltipla	R. L. Simples	R. L. Múltipla	R. L. Simples	R. L. Múltipla
Posto A	17,15%	16,58%	46,80%	44,81%	28,63%	X	24,77%	23,60%
Posto B	13,56%	13,55%	21,87%	21,65%	31,70%	X	13,60%	13,63%
Posto C	19,37%	18,87%	37,65%	34,88%	25,90%	X	15,13%	15,18%
Posto D	22,92%	22,96%	31,52%	31,01%	12,07%	X	X	X
Posto E	20,06%	20,06%	38,85%	38,91%	25,02%	X	10,83%	10,75%

APÊNDICE M - Quadro com o desvio padrão calculado nas aplicações em redes neurais artificiais, em cada combustível, de cada posto

	Óleo Diesel		Etanol		Gasolina Adivada		Gasolina Comum	
	Diário	Semanal	Diário	Semanal	Diário	Semanal	Diário	Semanal
Posto A	0,0824	0,3042	0,0067	0,1179	0,0917	0,5048	0,2040	0,0860
Posto B	0,0151	0,0804	0,0251	0,0603	0,1421	0,0702	0,0092	0,0606
Posto C	0,0188	0,0367	0,0686	0,0168	0,0586	0,1648	0,0379	0,0320
Posto D	0,0241	0,0546	0,0566	0,0660	0,0889	0,0441	X	X
Posto E	0,0147	0,0294	0,0109	0,0975	0,0496	0,0686	0,0098	0,0473

APÊNDICE N - Quadro referente aos melhores resultados alcançados, diários e semanais, de cada um dos métodos utilizados (séries temporais, métodos causais e redes neurais artificiais) no posto A

		Séries temporais		Métodos causais	Redes Neurais Artificiais
		Formulas Excel®	Crystal Ball®		
Óleo diesel	Diário	590,77%	34,71%	37,46%	23,54%
		Suavização Exponencial com tendência e sazonalidade	Média Móvel Simples		
	Semanal	15,39%	14,34%	16,58%	49,27%
		Média Móvel Dupla	Suavização Exponencial Simples e Suavização Exponencial com tendência	Regressão Linear Múltipla	
Etanol	Diário	104,02%	29,53%	51,28%	0,37%
		Média Móvel Dupla	Média Móvel Dupla		
	Semanal	23,32%	17,58%	44,81%	5,36%
		Média Móvel Dupla	Suavização Exponencial Simples	Regressão Linear Múltipla	
Gasolina aditivada	Diário	989,11%	20,67%	37,75%	21,20%

		Média Móvel Ponderada	Média Móvel Dupla		
	Semanal	10,91%	10,58%	28,63%	14,11%
		Média Móvel Dupla	Média Móvel Simples		
Gasolina comum	Diário	128,15%	20,95%	35,02%	31,70%
		Suavização Exponencial com tendência	Média Móvel Simples		
	Semanal	12,18%	9,26%	23,60%	16,25%
		Suavização Exponencial com tendência e sazonalidade	Suavização Exponencial Simples e Suavização Exponencial com tendência	Regressão Linear Múltipla	

APÊNDICE O - Quadro referente aos melhores resultados alcançados, diários e semanais, de cada um dos métodos utilizados (séries temporais, métodos causais e redes neurais artificiais) no posto B

		Séries temporais		Métodos causais	Redes Neurais Artificiais
		Formulas Excel®	Crystal Ball®		
Óleo diesel	Diário	86,04%	27,95%	30,19%	23,64%
		Suavização Exponencial com tendência	Suavização Exponencial com tendência e sazonalidade		
	Semanal	14,28%	11,96%	13,55%	27,68%
		Média Móvel Dupla	Média Móvel Dupla	Regressão Linear Múltipla	
Etanol	Diário	48,36%	27,71%	32,67%	1,62%
		Média Móvel Dupla	Suavização Exponencial com tendência		
	Semanal	16,56%	15,67%	21,65%	4,47%
		Média Móvel Dupla	Média Móvel Simples	Regressão Linear Múltipla	
Gasolina aditivada	Diário	24,38%	19,78%	38,23%	15,95%

		Média Móvel Dupla	Suavização Exponencial com tendência e sazonalidade		
	Semanal	13,49%	11,62%	31,70%	14,11%
		Média Móvel Dupla	Média Móvel Dupla		
Gasolina comum	Diário	31,68%	14,47%	23,77%	15,70%
		Média Móvel Dupla	Suavização Exponencial com tendência e sazonalidade		
	Semanal	9,51%	9,47%	13,60%	27,14%
		Média Móvel Dupla	Suavização Exponencial Simples e Suavização Exponencial com tendência	Regressão Linear Simples	

APÊNDICE P - Quadro referente aos melhores resultados alcançados, diários e semanais, de cada um dos métodos utilizados (séries temporais, métodos causais e redes neurais artificiais) no posto C

		Séries temporais		Métodos causais	Redes Neurais Artificiais
		Formulas Excel®	Crystal Ball®		
Óleo diesel	Diário	1270,49%	36,37%	48,41%	22,68%
		Suavização Exponencial com tendência e sazonalidade	Suavização Exponencial com tendência e sazonalidade		
	Semanal	15,79%	15,03%	18,87%	15,54%
		Média Móvel Dupla	Média Móvel Dupla	Regressão Linear Múltipla	
Etanol	Diário	66,56%	31,22%	43,56%	8,30%
		Média Móvel Ponderada	Média Móvel Simples		
	Semanal	24,65%	19,65%	34,88%	7,73%
		Média Móvel Ponderada	Suavização Exponencial Simples	Regressão Linear Múltipla	
Gasolina aditivada	Diário	535,19%	28,36%	36,46%	14,18%

		Suavização Exponencial com tendência	Média Móvel Simples		
	Semanal	16,09%	14,16%	25,90%	32,11%
		Média Móvel Dupla	Média Móvel Simples		
Gasolina comum	Diário	27,75%	18,43%	24,33%	12,65%
		Média Móvel Dupla	Suavização Exponencial com tendência e sazonalidade		
	Semanal	9,58%	10,09%	15,13%	17,14%
		Média Móvel Dupla	Média Móvel Simples	Regressão Linear Simples	

APÊNDICE Q - Quadro referente aos melhores resultados alcançados, diários e semanais, de cada um dos métodos utilizados (séries temporais, métodos causais e redes neurais artificiais) no posto D

		Séries temporais		Métodos causais	Redes Neurais Artificiais
		Formulas Excel®	Crystal Ball®		
Óleo diesel	Diário	682,52%	41,54%	50,78%	20,52%
		Média Móvel Dupla	Suavização Exponencial com tendência e sazonalidade		
	Semanal	23,93%	19,34%	22,92%	32,14%
		Média Móvel Dupla	Média Móvel Simples		
Etanol	Diário	177,26%	31,07%	40,49%	4,25%
		Suavização Exponencial com tendência	Suavização Exponencial Simples		
	Semanal	22,16%	20,30%	31,01%	21,61%
		Média Móvel Dupla	Suavização Exponencial Simples		
Gasolina aditivada	Diário	21,50%	14,93%	20,13%	20,88%

		Média Móvel Dupla	Suavização Exponencial com tendência e sazonalidade		
	Semanal	7,84%	8,04%	12,07%	26,96%
		Média Móvel Dupla	Suavização Exponencial Simples e Suavização Exponencial com tendência		

APÊNDICE R - Quadro referente aos melhores resultados alcançados, diários e semanais, de cada um dos métodos utilizados (séries temporais, métodos causais e redes neurais artificiais) no posto E

		Séries temporais		Métodos causais	Redes Neurais Artificiais
		Formulas Excel®	Crystal Ball®		
Óleo diesel	Diário	1356,11%	35,70%	44,33%	16,81%
		Média Móvel Dupla	Suavização Exponencial com tendência e sazonalidade		
	Semanal	38,50%	13,16%	20,06%	8,21%
		Média Móvel Dupla	Média Móvel Dupla	Regressão Linear Simples e Regressão Linear Múltipla	
Etanol	Diário	66,98%	36,86%	47,4%	0,54%
		Média Móvel Dupla	Média Móvel Simples		
	Semanal	61,03%	23,32%	38,85%	13,22%
		Média Móvel Dupla	Média Móvel Simples	Regressão Linear Simples	
Gasolina aditivada	Diário	115,63%	22,88%	36,12%	9,56%

		Suavização Exponencial com tendência e sazonalidade	Suavização Exponencial com tendência e sazonalidade		
	Semanal	13,39%	11,03%	25,02%	18,04%
		Média Móvel Dupla	Média Móvel Dupla		
Gasolina comum	Diário	122,12%	17,27%	27,28%	11,99%
		Suavização Exponencial com tendência e sazonalidade	Suavização Exponencial com tendência e sazonalidade		
	Semanal	9,71%	9,77%	10,75%	27,69%
		Média Móvel Dupla	Média Móvel Simples	Regressão Linear Múltipla	