

UNIVERSIDADE DE SÃO PAULO  
ESCOLA DE ENGENHARIA DE SÃO CARLOS  
DEPARTAMENTO DE ENGENHARIA ELÉTRICA

PEDRO HENRIQUE MARQUES TREVISANI

ESTIMAÇÃO DE PREÇO E DEMANDA DE ENERGIA ELÉTRICA BASEADO EM  
DADOS NO DOMÍNIO DO TEMPO E DA FREQUÊNCIA

SÃO CARLOS

2017



PEDRO HENRIQUE MARQUES TREVISANI

ESTIMAÇÃO DE PREÇO E DEMANDA DE ENERGIA ELÉTRICA BASEADO EM  
DADOS NO DOMÍNIO DO TEMPO E DA FREQUÊNCIA

Trabalho de Conclusão de Curso apresentado  
à Escola de Engenharia de São Carlos da  
Universidade de São Carlos.

Curso de Engenharia Elétrica com Ênfase em  
Sistemas de Energia e Automação.

ORIENTADOR: Prof. Dr. Ricardo A. S. Fernandes

SÃO CARLOS

2017

AUTORIZO A REPRODUÇÃO TOTAL OU PARCIAL DESTE TRABALHO,  
POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS  
DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

T814e

Trevisani, Pedro Henrique Marques  
Estimação de Preço e Demanda de Energia Elétrica  
baseado em Dados no Dominínio do Tempo e da Frequênci  
/ Pedro Henrique Marques Trevisani; orientador Ricardo  
Fernandes. São Carlos, 2017.

Monografia (Graduação em Engenharia Elétrica com  
ênfase em Sistemas de Energia e Automação) -- Escola de  
Engenharia de São Carlos da Universidade de São Paulo,  
2017.

1. Estimadores. 2. Redes Neurais Artificiais. 3.  
Preço. 4. Demanda. 5. Frequênci. I. Título.

# FOLHA DE APROVAÇÃO

Nome: Pedro Henrique Marques Trevisani

Título: "Estimação de preço e demanda de energia elétrica baseado em dados no domínio do tempo e da frequência"

Trabalho de Conclusão de Curso defendido e aprovado

em 30/06/2017,

com NOTA 9,5 (nove, cinco), pela Comissão Julgadora:

*Prof. Dr. Ricardo Augusto Souza Fernandes - Orientador -  
DEE/UFSCar*

*Mestre Fabbio Anderson Silva Borges - Doutorando - SEL/EESC/USP*

*Prof. Dr. Giuseppe Antonio Cirino - DEE/UFSCar*

Coordenador da CoC-Engenharia Elétrica - EESC/USP:  
Prof. Associado José Carlos de Melo Vieira Júnior



## **Dedicatória e Agradecimentos**

Este trabalho representa a conclusão dos anos em que vivi em São Carlos, e eu não poderia deixar de nomear e agradecer as pessoas que foram imprescindíveis para a conquista dos meus objetivos.

Aos meus pais e irmã, Rose, Henrique e Ana Laura, agradeço por serem a base do respeito, caráter e carinho que me trouxeram até aqui. À minha namorada, Patrícia, agradeço pelo companheirismo e motivação.

Agradeço aos meus queridos amigos de sala, por dividirem comigo as preocupações e alegrias proporcionadas pela graduação, sendo muitas vezes o apoio para que eu não desistisse dos meus objetivos. Agradeço especialmente, ao Pedro, amigo que morou comigo durante estes 5 anos, e foi exemplo de humildade e inteligência.

Agradeço também ao meu orientador Prof. Dr. Ricardo Fernandes, pelos ensinamentos e paciência oferecidos durante a execução deste trabalho.

Por fim, dedico este trabalho aos meus avós, Anézio, Mariana e Laura, os quais ergueram minha família sobre os pilares da fé, bondade e resiliência.



## Resumo

**TREVISANI, P. H. M., Estimação de Preço e Demanda de Energia Elétrica baseado em Dados no Domínio do Tempo e da Frequência.** Trabalho de Conclusão de Curso – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2017.

O trabalho intitulado como “Estimação de Preço e Demanda de Energia Elétrica baseado em Dados no Domínio do Tempo e da Frequência” tem como propósito a implementação de um sistema baseado em Redes Neurais Artificiais (RNA) para previsão de preço e demanda de energia elétrica no horizonte de uma hora a frente. Para tal, foram utilizados dados amplamente empregados em pesquisas semelhantes, de modo a facilitar a comparação que determinaria a eficácia do método utilizado. A partir destes dados, foram realizadas análises no domínio da frequência, responsáveis pela obtenção de novos atributos, os quais, juntamente com os originais, passaram por uma seleção de atributos capaz de determinar quais dados apresentavam a maior relevância para as estimações desejadas. Os resultados obtidos permitiram uma discussão sobre a eficácia da metodologia utilizada e da sua empregabilidade em pesquisas futuras.

**Palavras-chave:** Estimação. Redes Neurais Artificiais. Preço. Demanda. Frequência



## Abstract

TREVISANI, P. H. M., **Forecasting Price and Demand of Electrical Energy based on Time and Frequency Domain Data.** Trabalho de Conclusão de Curso – Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2017.

The work named “Forecasting price and demand of electrical energy based on time and frequency domain data” has by purpose the implementation of an Intelligent System based on Artificial Neural Network (ANN) for forecasting price and demand in one hour ahead horizon. Thus, a widely-employed data was used, to make easier the comparison that determinates the method accuracy. From these data, analyzes were carried out in the frequency domain, that generates new attributes, which, together with the original, underwent a attributes selection capable of determinate which data show the greatest relevance for the desired estimates. The acquired results allowed an argument about the effectiveness of the methodology used and its application in future searches.

**Keywords:** Forecasting. Artificial Neural Network. Price. Demand. Frequency.



## **Lista de Ilustrações**

Figura 1: Gráfico de Demanda de energia elétrica para os Dias Úteis .....	6
Figura 2: Gráfico de Demanda de energia elétrica para Feriados e Finais de Semana .....	6
Figura 3: Gráfico de Demanda de energia elétrica para Todos os Dias .....	6
Figura 4: Gráfico de Preço de energia elétrica para os Dias Úteis .....	7
Figura 5: Gráfico de Preço de energia elétrica para os Feriados e Finais de Semana .....	7
Figura 6: Gráfico de Preço de energia elétrica para Todos os Dias .....	7
Figura 7: Esquema de um neurônio artificial .....	19



## **Lista de Tabelas**

Tabela 1: Dados organizados após a etapa de pré-processamento .....	8
Tabela 2: Nomenclatura para os atributos no domínio da frequência .....	11
Tabela 3: Atributos selecionados, por ambos os algoritmos, para as estimativas de demanda considerando o grupo de bases #1 .....	14
Tabela 4: Atributos selecionados, por ambos os algoritmos, para as estimativas de preço considerando o grupo de bases #1 .....	15
Tabela 5: Atributos selecionados, por ambos os algoritmos, para as estimativas de demanda considerando o grupo de bases #2 .....	15
Tabela 6: Atributos selecionados, por ambos os algoritmos, para as estimativas de preço considerando o grupo de bases #2 .....	16
Tabela 7: Atributos selecionados, por ambos os algoritmos, para as estimativas de demanda considerando o grupo de bases #3 .....	17
Tabela 8: Atributos selecionados, por ambos os algoritmos, para as estimativas de preço considerando o grupo de bases #3 .....	18
Tabela 9: Resultados obtidos pelas RNAs considerando os atributos selecionados do grupo de base de dados #1 para as estimativas de demanda.....	24
Tabela 10: Resultados obtidos pelas RNAs considerando os atributos selecionados do grupo de base de dados #1 para as estimativas de preço .....	24
Tabela 11: Resultados obtidos pelas RNAs considerando os atributos selecionados do grupo de base de dados #2 para as estimativas de demanda.....	25
Tabela 12: Resultados obtidos pelas RNAs considerando os atributos selecionados do grupo de base de dados #2 para as estimativas de preço .....	25
Tabela 13: Resultados obtidos pelas RNAs considerando os atributos selecionados do grupo de base de dados #3 para as estimativas de demanda.....	26

Tabela 14: Resultados obtidos pelas RNAs considerando os atributos selecionados do grupo de base de dados #3 para as estimativas de preço ..... 27

Tabela 15: Média de MAPEs das estimativas de demanda para Todos os Dias ..... 27

## Sumário

<b>CAPÍTULO 1</b>	<b>Introdução ao Trabalho de Conclusão de Curso .....</b>	<b>1</b>
1.1.	Introdução .....	1
1.2.	Motivação .....	2
1.3.	Objetivos e Metodologia.....	2
1.4.	Organização do Trabalho .....	3
<b>CAPÍTULO 2</b>	<b>Organização Inicial dos Dados.....</b>	<b>5</b>
<b>CAPÍTULO 3</b>	<b>Extração de Características no Domínio da Frequência .....</b>	<b>9</b>
3.1.	Introdução .....	9
3.2.	Características Obtidas no Domínio da Frequência.....	10
<b>CAPÍTULO 4</b>	<b>Seleção de Atributos.....</b>	<b>13</b>
4.1.	Fundamentos da Seleção de Atributos .....	13
4.2.	Resultados da Seleção de Atributos.....	14
<b>CAPÍTULO 5</b>	<b>Redes Neurais Artificiais .....</b>	<b>19</b>
5.1.	Introdução .....	19
5.2.	Características da RNA Utilizada .....	20
<b>CAPÍTULO 6</b>	<b>Resultados e Discussões .....</b>	<b>23</b>
6.1.	Resultados Obtidos com os Dados no Domínio do Tempo .....	23
6.2.	Resultados Obtidos com os dados apenas no domínio da frequência.....	24
6.3.	Resultados Obtidos com os Dados em Ambos os Domínios.....	26
<b>CAPÍTULO 7</b>	<b>Conclusões.....</b>	<b>29</b>
	<b>Referências Bibliográficas.....</b>	<b>31</b>



# CAPÍTULO 1

## Introdução ao Trabalho de Conclusão de Curso

### 1.1. Introdução

A energia elétrica é amplamente comercializada e tratada como commodity em diversos mercados ao redor do mundo. Porém, algumas de suas características naturais tornam o comércio e o fornecimento de energia diferentes daqueles das demais mercadorias.

Por não ser possível uma armazenagem, há a necessidade de um constante balanço entre geração e demanda por energia elétrica [2]. Além disto, o preço da energia ainda apresenta alta dependência de outros fatores ligados, principalmente, a matriz energética do mercado em questão. Dentre os quais destacam-se meteorologia, hidrografia, demanda por energia, preço de combustíveis que dificultam a estabilidade e previsibilidade dos mercados de energia.

Assim, sempre existiu a necessidade de bons estimadores que fossem capazes de proporcionar vantagens econômicas, e até ambientais, através da previsão do preço e demanda da energia.

Partindo desta necessidade, pesquisadores de todo o mundo têm desenvolvido pesquisas em cima da previsão de preço e demanda [1-7]. Estas pesquisas apresentam soluções baseadas na utilização de métodos como regressão dinâmica [1], Redes Neurais Artificiais [2-3] e sistemas Fuzzy [5].

A pesquisa realizada em [2] utilizou os dados do mercado de Ontario, no Canadá e propôs a estimação do preço da energia no dia seguinte. O método proposto utiliza uma versão modificada do algoritmo Relief, para ponderar os atributos que seriam apresentados a uma Rede Neural Artificial em cascata. O método obteve resultados ótimos de previsão, apresentando melhora significativa quando comparados aos resultados obtidos em [5], que também utilizou redes neurais para prever os preços de Ontario.

Uma outra abordagem dos dados do mercado de Ontario é proposta em [3], onde a transformada wavelet é utilizada para fornecer dados melhores para os métodos de estimação utilizados. A análise no domínio da frequência também foi utilizada em [7], onde os dados do mercado de Nova York foram divididos em suas componentes de baixa, média e alta frequência. Tal separação foi feita através da Transformada Discreta de Fourier e possibilitou a análise de um modelo de previsão linear para preço, demanda e geração de energia eólica.

O recente aumento da geração distribuída criou a necessidade de modelos de previsões de demanda específicos para alguns casos, como em [16], onde um estudo foi feito sobre a possibilidade do aumento de micro redes, um conceito onde pequenos sistemas, separados da rede de distribuição, se mantêm sozinhos, através da micro geração distribuída. O estudo ainda mostra como a demanda nestes pequenos sistemas é diferente das usuais, e propõe métodos de evolução diferencial para agregar aos métodos já conhecidos de previsão de demanda.

A previsão de preço e demanda também se mostra importante para aquilo que é considerado o futuro do consumo em energia elétrica. O estudo apresentado em [4] associa um estimador de preço simples ao sistema de automação de uma casa, possibilitando que o consumo de energia seja dividido ao longo do dia através de uma avaliação dos preços previstos para aquele dia.

## **1.2. Motivação**

O contexto exposto na introdução já justifica o interesse pela pesquisa de estimadores e dos métodos utilizados, porém, deve-se ainda acrescentar o panorama do mercado brasileiro.

O ano de 2016 apresentou crescimento de 87% no número de ativos que migraram do ambiente de contratação regulado para o mercado livre de energia elétrica [20]. Tal crescimento está diretamente ligado a tentativa de indústrias e empresas de reduzir gastos em meio à crise econômica.

Uma vez migrado, o consumidor passa a ter responsabilidades, ligadas ao seu planejamento de consumo de energia que antes não existiam ou eram de responsabilidade da distribuidora. Gerando assim a necessidade cada vez maior de uma antecipação de seu consumo e também do preço que será pago pela energia adquirida em contratos de longo prazo.

## **1.3. Objetivos e Metodologia**

Este trabalho de conclusão de curso tem como objetivo final a apresentação de resultados satisfatórios na estimação de preço e demanda de energia para tanto, foram estabelecidas metas menores, onde foi possível desenvolver conhecimentos adquiridos ao longo do curso de engenharia elétrica.

- Obter e organizar os dados históricos de preço e demanda de energia elétrica afim de criar uma base que fosse capaz de alimentar um sistema de previsão;

- Realizar uma análise destes dados no domínio da frequência afim de aumentar a quantidade de atributos a serem oferecidos para o sistema inteligente;
- Selecionar os atributos mais relevantes para a estimação do preço e demanda de energia;
- Implementar um Rede Neural Artificial capaz de estimar valores em um horizonte de curtíssimo prazo e discutir os resultados obtidos afim de determinar o impacto causado pela mudança na entrada de dados.

Para atingir tais objetivos este trabalho teve como ponto de partida os dados que foram trabalhados e organizados em [17], os quais foram cedidos gentilmente pela autora. No referido trabalho, foram utilizados os dados do mercado de Ontario no Canadá, assim como em [1-3], [5], [9], [11], [16], [17].

Após a obtenção destas informações foram realizadas análises destes dados no domínio da frequência, onde foram gerados novos atributos, que, juntamente com os iniciais, foram selecionados através dos algoritmos de seleção, Relief e CFS, capazes de elencar quais os atributos mais relevantes na estimação dos valores de preço e demanda.

Na sequência foram desenvolvidas redes neurais idênticas para serem alimentadas com as diferentes bases de dados criadas a partir da organização dos atributos selecionados. Assim, foi possível discutir e analisar quais entradas permitiriam uma eficácia maior das redes.

#### **1.4. Organização do Trabalho**

O trabalho de conclusão de curso foi dividido em 7 capítulos principais. O capítulo 1 apresenta uma introdução sobre o assunto tratado durante o trabalho, citando as pesquisas que foram utilizadas com embasamento teórico e motivação para a realização deste trabalho. Ainda no capítulo 1 são descritos os objetivos bem como a metodologia utilizada para alcançá-los.

No capítulo 2 é apresentado ao leitor as características dos dados do mercado de energia de Ontário no Canadá, que iniciaram este trabalho. O objetivo deste capítulo é o de contextualizar o leitor sobre as informações presentes nos dados, e a maneira como estes foram organizados.

O capítulo 3 descreve brevemente os conceitos de análise no domínio da frequência utilizados, neste trabalho, para obter informações adicionais sobre os dados iniciais. Os resultados obtidos pela análise feita também estão explicitados no capítulo.

No capítulo 4 são descritos os fundamentos da seleção de atributos, bem como os resultados obtidos pelos algoritmos de seleção que foram escolhidos pelo autor.

O capítulo 5 traz uma breve introdução sobre Redes Neurais Artificiais, explicando rapidamente o funcionamento destes sistemas inteligentes. Há também neste capítulo a descrição das arquiteturas e topologias escolhidas para as RNAs utilizadas para realizar as estimativas de preço e demanda.

O capítulo 6 discute os resultados finais obtidos pelos estimadores desenvolvidos ao longo do trabalho, apontando os destaques e as falhas destes estimadores.

No capítulo 7 são descritas as implicações e conclusões retiradas da realização do trabalho.

# CAPÍTULO 2

## Organização Inicial dos Dados

Como dito anteriormente, este trabalho utiliza os dados do mercado de Ontario que foram trabalhados em [17], portanto, esta seção mostra a organização final obtida pela autora no referido trabalho.

Todos os dados utilizados são valores horários de cada um dos atributos e correspondem ao período de pouco mais de 3 anos, que se inicia em julho de 2009 e termina em julho de 2012. Além disto, foi feita a separação entre as horas que correspondem aos dias úteis do período e as horas que correspondem aos dias de final de semana e feriado, gerando assim, 3 bases distintas de dados:

- Dias Úteis: Com o total de 18720 horas;
- Feriado e FDS: Com o total de 8112 horas;
- Todos os Dias: Com o total de 26832 horas.

Cada uma dessas bases é composta por 5 atributos, os quais são discretizados no tempo:

Demand: Valor horário da demanda do sistema de Ontario;

Dólar: cotação do dólar canadense em relação ao dólar americano da hora em questão;

Energia eólica: quantidade de energia eólica gerada nos parques presentes no sistema no intervalo horário;

Balanço: balanço energético que representa as importações e exportações de energia elétrica;

Preço: preço horário da energia elétrica no mercado de Ontário.

Como o objetivo do trabalho era desenvolver estimadores de preço e demanda de energia, cada valor horário destes atributos foi considerado como a saída de dados desejada para a previsão realizada. Desta forma, as figuras na sequência mostram os gráficos dos valores horários destes dois atributos, ao longo do período analisado. Os gráficos estão divididos conforme as bases de dados citadas anteriormente: feriados e finais de semana, dias úteis e todos os dias, proporcionando uma análise visual do comportamento dos dados à serem estimados, nas diferentes bases.

Figura 1: Gráfico de Demanda de energia elétrica para os Dias Úteis

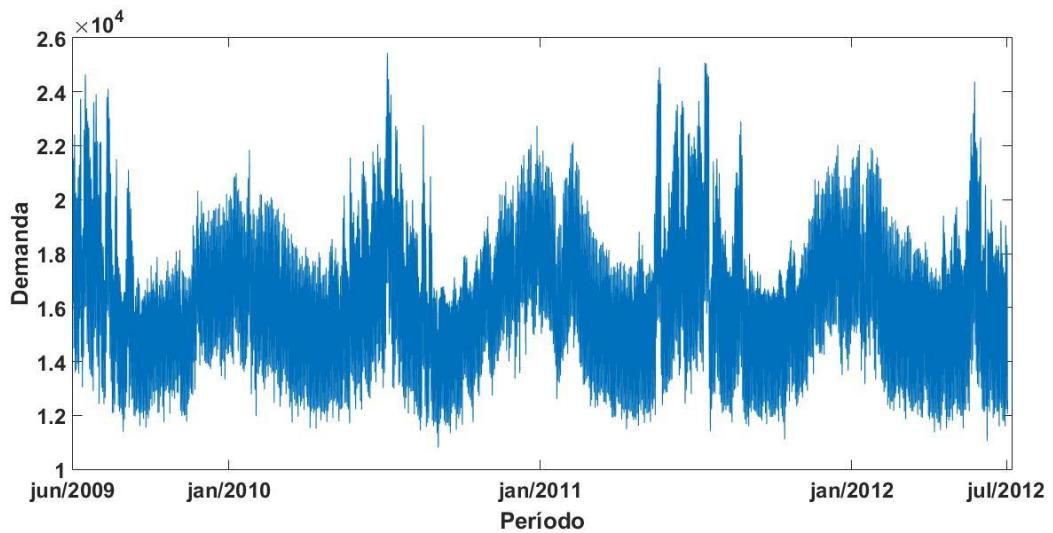


Figura 2: Gráfico de Demanda de energia elétrica para Feriados e Finais de Semana

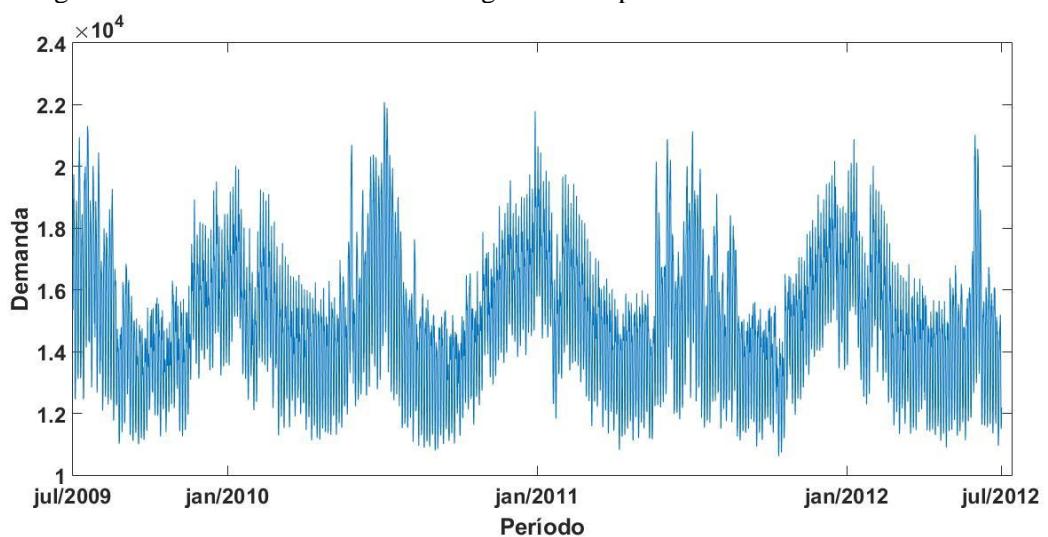


Figura 3: Gráfico de Demanda de energia elétrica para Todos os Dias

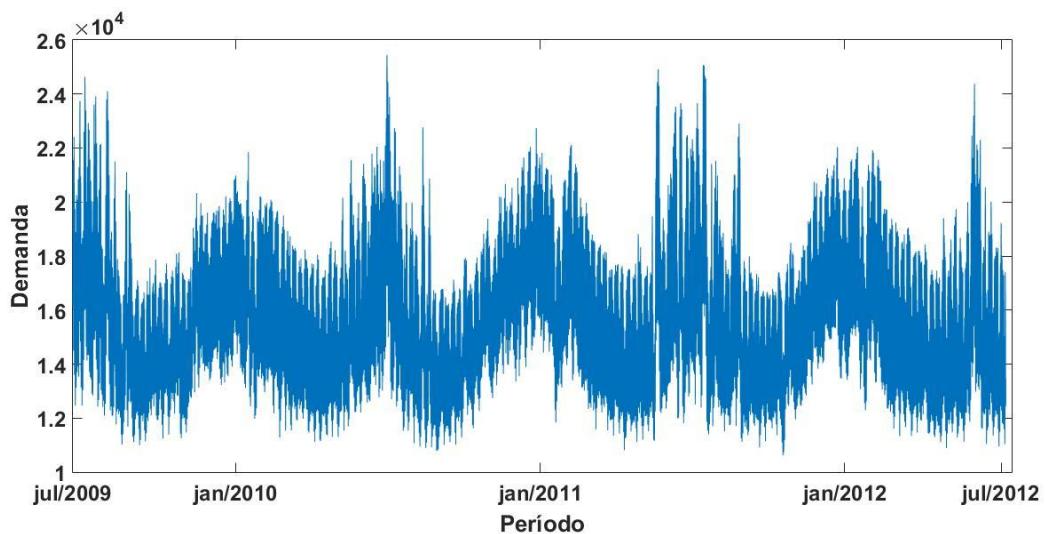


Figura 4: Gráfico de Preço de energia elétrica para os Dias Úteis

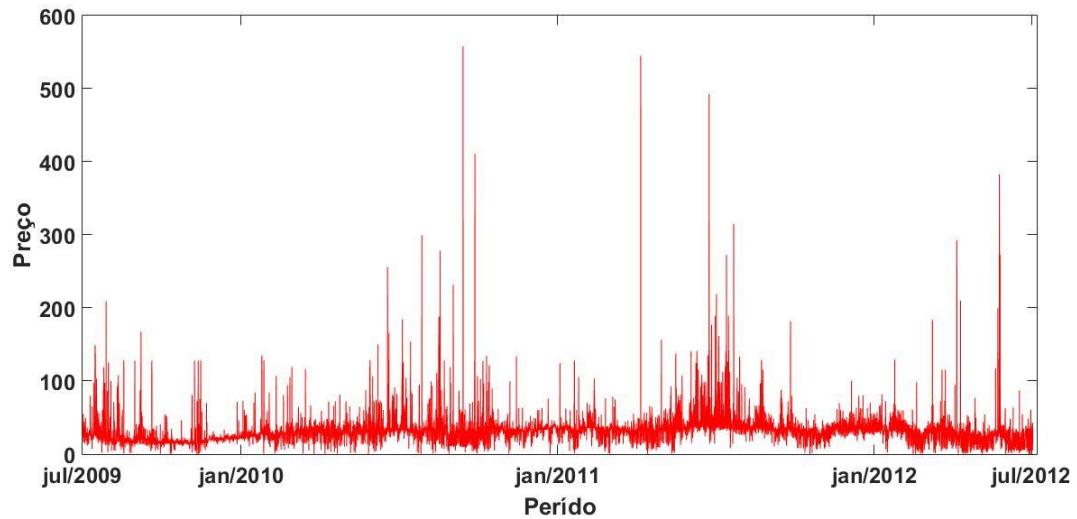


Figura 5: Gráfico de Preço de energia elétrica para os Feriados e Finais de Semana

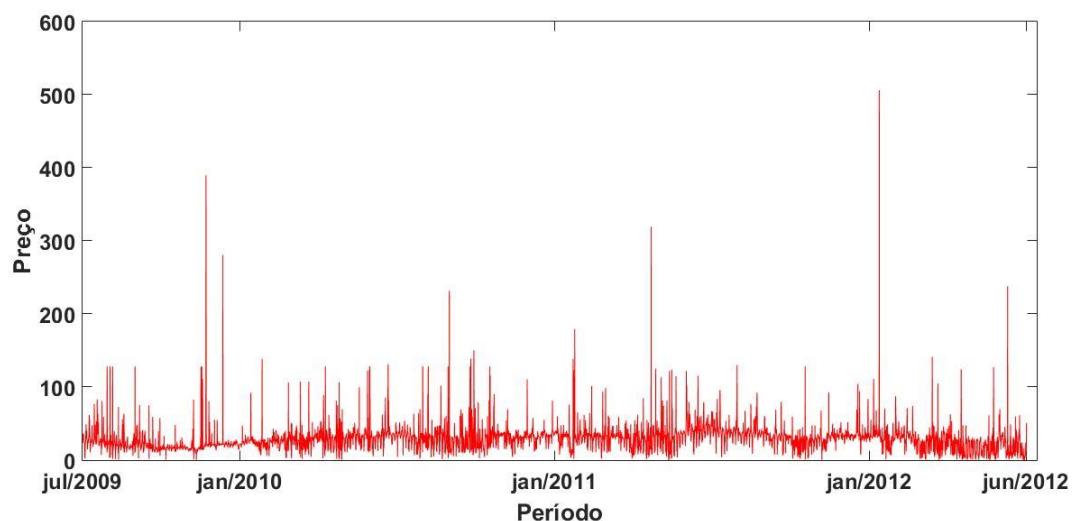
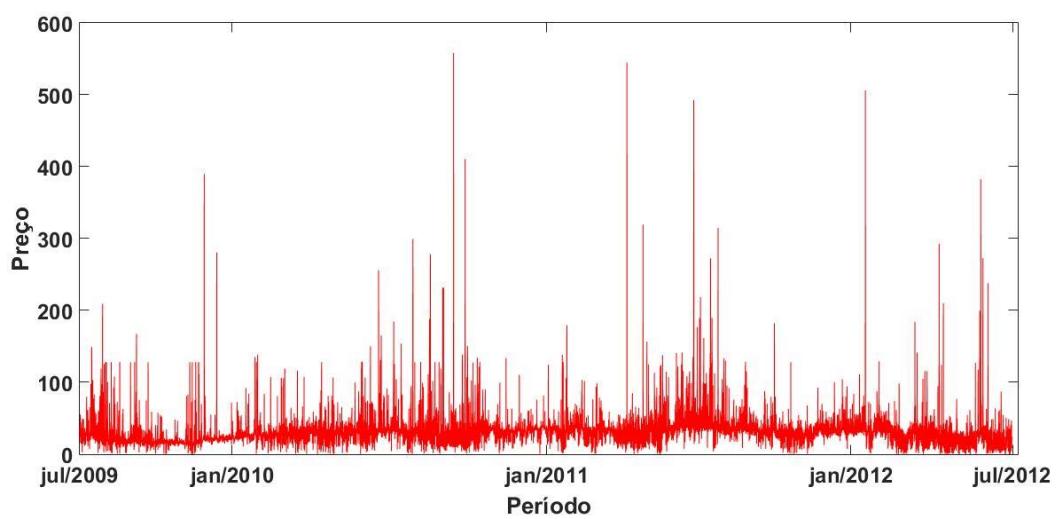


Figura 6: Gráfico de Preço de energia elétrica para Todos os Dias



Visualmente é possível notar que os valores de demanda apresentam um padrão de variação ao longo do ano, o qual se mantém para as 3 bases de dados. Já os valores de preço apresentam uma volatilidade maior, apresentando picos que aparentemente não seguem nenhum padrão.

A entrada de dados foi criada a partir dos valores das horas anteriores à hora a ser estimada  $h$ . Para cada valor de  $i$ , variando de 1 até 23, foi considerado um novo atributo, que continha o valor de um dos atributos originais, na hora  $h - i$ . Desta maneira, cada um dos atributos de preço, demanda, energia eólica e balanço energético geraram 23 atributos que foram considerados como entrada. Para os valores do dólar, foram gerados apenas 5 novos atributos, onde  $i = [24, 48, 72, 96, 120]$ , pois a cotação era alterada apenas de um dia para o outro.

Além disto, os dados encontrados no trabalho [17] apresentam também como atributos o cálculo de Médias Móveis (MM) ponderadas que consideram 12 janelas móveis com intervalos de 24 horas. É importante destacar, assim como foi destacado em [17], que o objetivo do cálculo destas médias móveis é o de filtrar eventuais flutuações demaisadas e mostrar mais explicitamente uma tendência, caso essa exista.

Tabela 1: Dados organizados após a etapa de pré-processamento

Entradas	$h - i$	Preço Demanda Energia Eólica Balanço Dólar
	Média móvel de 24 horas	Preço Demanda Energia Eólica Balanço
Saídas	desejadas para $h$	Preço Demanda

Fonte: [17]

# CAPÍTULO 3

## Extração de Características no Domínio da Frequência

### 3.1. Introdução

Um sinal discreto é qualquer sequência de valores que contenham informações sobre estado ou comportamento de um sistema físico, portanto, quando esta sequência representa o intervalo de tempo entre uma amostra e outra do sinal, diz-se que o sinal é discreto no tempo.

Dentro da análise de sinais, o domínio da frequência compreende uma alternativa de representação dos dados, onde os valores obtidos representam a amplitude que o sinal possui em determinadas frequências.

Para realizar a conversão entre os domínios do tempo e frequência é necessário utilizar-se de uma espécie particular de operadores matemáticos chamados de transformadas integrais. As transformadas de Fourier são, indubitavelmente, as transformadas mais comumente utilizadas para fins de análise no domínio da frequência, sendo que, as características do sinal a ser transformado determinam qual transformada de Fourier será utilizada.

Quando se deseja analisar sinais de duração finita, é possível utilizar uma das variações da transformada de Fourier, chamada de Transformada Discreta de Fourier (TDF). A TDF é, na verdade, uma sequência ao invés de uma função de uma variável contínua, e corresponde a amostras, igualmente espaçadas em frequência, da transformada de Fourier do sinal. Além de sua importância teórica a TDF tem grande importância dentro das diferentes implementações de algoritmos utilizados para processamento digital de sinais. Isto ocorre devido a existência de algoritmos eficientes capazes de calcular a TDF dos sinais [18].

$$X_k = \sum_{n=0}^{N-1} x_n * e^{-\frac{i2\pi kn}{N}} \quad (1)$$

A equação (1) representa a definição da TDF, que transforma uma sequência de  $n$  números  $[x_0, x_1, \dots, x_{n-1}]$ , contidos no domínio do tempo, em outra sequência também de  $n$  números  $[X_0, X_1, \dots, X_{n-1}]$ .

### 3.2. Características Obtidas no Domínio da Frequência

A organização dos dados explicada no capítulo 2 mostra que, para cada hora pertencente ao intervalo disponível, foram considerados as informações das 23 horas anteriores. Portanto, cada um destes conjuntos foi encarado como um sinal discreto no tempo, com intervalo de amostragem de 1 hora, gerando assim, os sinais que foram analisados no domínio da frequência.

É importante lembrar que para os dados de cotação do dólar foram considerados apenas 5 dados de valores anteriores, portanto, a análise foi realizada apenas para os dados de demanda, energia eólica, balanço e preço da energia elétrica.

Para realizar as análises foi utilizado o algoritmo da Transformada Rápida de Fourier disponível no software Mathworks MATLAB®. Este algoritmo representa uma maneira computacionalmente eficiente de calcular a TDF, sendo largamente utilizado para processamento digital de sinais.

Desta maneira, cada conjunto de 23 dados foi utilizado como entrada para o referido algoritmo, gerando um espectro de frequência onde eram expostas as amplitudes do sinal para cada valor de frequência. Porém estes valores separadamente não poderiam ser considerados como atributos relevantes para as previsões que seriam realizadas nas próximas etapas deste trabalho.

Para que estes valores fossem então considerados como atributos, o espectro de frequência foi dividido entre os pontos que representavam as amplitudes das baixas frequências do sinal e os pontos que representavam as amplitudes das altas frequências do sinal. Após separados os pontos, foram calculadas as energias associadas para cada um dos dois conjuntos de pontos gerado. A equação (2) sintetiza o cálculo realizado, onde  $N$  é o número de pontos do conjunto, e  $X[i]$  são as amplitudes de cada ponto.

$$E = \sum_{i=1}^N |X[i]|^2 \quad (2)$$

Portanto, foram gerados 8 novos atributos para cada uma das 3 bases de dados disponíveis, tais atributos foram organizados juntamente aos apresentados no capítulo 2, completando assim, a totalidade dos dados que foram analisados na seleção de atributos descrita no próximo capítulo. A tabela 2 foi criada para determinar a nomenclatura que será utilizada na continuação deste trabalho para facilitar o entendimento.

Tabela 2: Nomenclatura para os atributos no domínio da frequência

Atributo	Siglas
Energia associada às Altas Frequências	Demanda
	Balanço
	Energia Eólica
	Preço
Energia associada às Baixas Frequências	Demanda
	Balanço
	Energia Eólica
	Preço

É importante destacar que além das 3 bases de dados contendo a totalidade dos atributos, originais e obtidos pela análise no domínio da frequência, foram criados outros dois grupos de 3 bases cada, que continham, respectivamente, dados estritamente no domínio do tempo (dados originais) e dados estritamente no domínio da frequência.

- Grupo de Bases #1: Dados apenas no domínio do tempo.
- Grupo de Bases #2: Dados apenas no domínio da frequência;
- Grupo de Bases #3: Dados em ambos os domínios;

O objetivo da criação destes grupos de bases é permitir que a relevância dos dados obtidos pela análise em frequência seja verificada no caso destes dados serem utilizados separadamente dos dados no domínio do tempo. Além disso, a manutenção da base original e os testes realizados com esta, facilitam a comparação dos resultados obtidos no final deste trabalho.



# CAPÍTULO 4

## Seleção de Atributos

### 4.1. Fundamentos da Seleção de Atributos

Segundo [21], “atributos são as colunas de uma base de dados que representam as diferentes características das diversas instâncias contidas em uma base de dados”. Partindo deste conceito, a seleção de atributos pode ser definida como o processo responsável por limitar o conjunto de dados apenas aos dados relevantes. Desta maneira, o subconjunto de atributos selecionado deverá ser necessariamente igual ou menor que o conjunto original, porém o subconjunto deve ser capaz de generalizar as informações previamente compreendida.

A utilização dos algoritmos de seleção geralmente ocorre quando um sistema inteligente se mostra incapaz de generalizar as informações contidas em uma base de dados, isto pode ocorrer devido a uma grande quantidade de atributos fornecidos, ou ainda quando se tornam inviáveis devido ao custo computacional associado ao exagero de informações.

Os algoritmos podem ser separados conforme o seu funcionamento, sendo que, usualmente é utilizada uma observação individual de cada atributo, ou uma análise dos possíveis subconjuntos ótimos. Os seletores podem ser ainda divididos em filtros e wrappers, sendo que a diferença entre as categorias se dá apenas pela independência dos filtros com relação aos algoritmos de aprendizagem utilizados posteriormente à seleção.

Foram utilizados neste trabalho os mesmos filtros apresentados em [17]: Relief e CFS (Correlation-based Feature Selection). A diferença básica entre os métodos está no funcionamento, onde o método Relief utiliza uma observação individual de todos os atributos da base de dados, ranqueando-os de acordo com a sua relevância, atribuindo valores para estas relevâncias. As vantagens apresentadas por este método estão na possibilidade de eliminar facilmente atributos irrelevantes para a base, determinando um valor mínimo de relevância associada ao atributo para que este seja mantido na base.

O algoritmo CFS apresenta um método de busca por subconjuntos considerados ótimos para a estimação proposta. A vantagem deste está em sua rápida execução e também na possibilidade de eliminação de atributos redundantes para a base. Porém, por não apresentar a relevância isolada de cada atributo, corre-se o risco deste algoritmo selecionar a maior parte da base original para o subconjunto considerado otimizado.

## 4.2. Resultados da Seleção de Atributos

A seleção dos atributos foi realizada aplicando os algoritmos de seleção descritos na seção anterior (Relief e CFS). Estes algoritmos estão presentes no software WEKA (Waikato Environment for Knowledge Analysis), o qual é licenciado publicamente, e, portanto, é largamente utilizado para pesquisas relacionadas à análise de dados e estimadores.

É importante mencionar que o algoritmo CFS apresentava em sua saída um subconjunto conciso, e considerado ótimo para os estimadores, tais subconjuntos tiveram todos os seus atributos considerados como entrada de dados para os estimadores. Diferente disto, o algoritmo Relief apresentou na saída uma lista ranqueada dos atributos, e, portanto, foi necessário estabelecer um valor mínimo de relevância para determinar os atributos à serem utilizados. Tal valor foi obtido empiricamente e fixado em 0,01 para as seleções realizadas nos grupos de bases #2 e #3, o grupo #1 apresentou valores drasticamente menores de relevância e por isso teve o limiar determinado em 0,002.

Nas tabelas 3 e 4 são apresentados os atributos selecionados a partir do grupo #1 de bases de dados. Assim, a tabela 3 exibe as escolhas de ambos os algoritmos de seleção para as previsões de demanda, enquanto que os dados relacionados às previsões de preço encontram-se na tabela 4.

Tabela 3: Atributos selecionados, por ambos os algoritmos, para as estimativas de demanda considerando o grupo de bases #1

Relief			CFS		
Feriados e FDS	Dias Uteis	Todos os Dias	Feriados e FDS	Dias Uteis	Todos os Dias
Demand h-1	Demand h-1	Demand h-1	Demand h-1	Demand h-1	Demand h-1
Demand h-23	Demand h-23	Demand h-23	Demand h-23	Demand h-12	Demand h-12
Demand h-2	Demand h-2	Demand h-2	Eólica h-3	Demand h-23	Demand h-23
Demand h-22	Demand h-22	Demand h-22	Balanço h-13	Eólica h-4	Eólica h-4
Demand h-3	Demand h-16	-	-	-	-
Demand h-15	Demand h-17	-	-	-	-
Demand h-9	Demand h-7	-	-	-	-
Demand h-16	Demand h-8	-	-	-	-
Demand h-14	Demand h-3	-	-	-	-
Demand h-8	-	-	-	-	-

É importante lembrar que este grupo de bases representa os mesmos dados utilizados em [17], além disto, os algoritmos de seleção utilizados também são os mesmos, portanto, como

era esperado, os atributos selecionados foram os mesmo que a autora obteve no referido trabalho.

Tabela 4: Atributos selecionados, por ambos os algoritmos, para as estimativas de preço considerando o grupo de bases #1

Relief			CFS		
Feriados e FDS	Dias Uteis	Todos os Dias	Feriados e FDS	Dias Uteis	Todos os Dias
Preço h-1	Preço h-1	Preço h-1	Demand h-23	Demand h-23	Demand h-23
1 <sup>a</sup> MM	1 <sup>a</sup> MM Preço	1 <sup>a</sup> MM Preço 1	Balanço h-1	Balanço h-1	Balanço h-1
2 <sup>a</sup> MM Preço	Preço h-2	Preço h-2	Preço h-1	Preço h-1	Preço h-1
Preço h-2	2 <sup>a</sup> MM Preço	2 <sup>a</sup> MM Preço	11 <sup>a</sup> MM Preço	12 <sup>a</sup> MM Preço	12 <sup>a</sup> MM Preço
3 <sup>a</sup> MM Preço	Dólar h-24	-	-	-	-
Dólar h-120	Dólar h-120	-	-	-	-
-	3 <sup>a</sup> MM Preço	-	-	-	-

A tabela 3 deixa claro que os dados das horas, anteriores à hora a ser estimada, são de extrema relevância para a realização das estimativas de demanda, sendo que a demanda das horas imediatamente anterior e de 23 horas antes, foram selecionadas para todas as bases em ambos os seletores. Além destes, apenas um valor horário de balanço e três de energia eólica foram selecionados, demonstrando a irrelevância dos valores da cotação do dólar e das janelas de médias móveis para se prever a demanda de energia elétrica.

Os atributos escolhidos para as previsões de preço mostraram uma variedade maior de atributos base, sendo escolhidos valores baseados na cotação do dólar, e balanço energético. Também na tabela 4 é possível perceber que os algoritmos se diferenciaram muito nas escolhas, onde apenas o valor de preço da hora imediatamente anterior à hora a ser estimada foi escolhido por ambos os seletores. Ainda nestes dados é importante notar a presença das janelas de médias móveis dos valores anteriores de preço, revelando a importância da série histórica de preços para a estimativa de preço de energia elétrica.

Tabela 5: Atributos selecionados, por ambos os algoritmos, para as estimativas de demanda considerando o grupo de bases #2

Relief			CFS		
Feriados e FDS	Dias Uteis	Todos os Dias	Feriados e FDS	Dias Uteis	Todos os Dias
EBF - D	EBF - D	EBF - D	EBF - D	EBF - D	EBF - D
EBF - B	EBF - P	EBF - P	EAF - D	EAF - D	EAF - D
EBF - P	EBF - B	EBF - B	-	-	-

As tabelas 5 e 6, seguindo o padrão das duas anteriores, exibem os resultados obtidos para a seleção dos atributos estritamente relacionados à frequência, presentes no grupo de bases de dados #2.

Tabela 6: Atributos selecionados, por ambos os algoritmos, para as estimativas de preço considerando o grupo de bases #2

Relief			CFS		
Feriados e FDS	Dias Uteis	Todos os Dias	Feriados e FDS	Dias Uteis	Todos os Dias
EBF - P	EBF - P	EBF - P	EBF - D	EBF - P	EBF - P
EAF - P	EAF - P	EAF - P	EAF - D	-	-
EBF - D	EBF - D	EBF - D	EAF - B	-	-
EBF - B	-	EBF - B	EBF - P	-	-
EAF - B	-	-	-	-	-

Os resultados das tabelas 5 e 6 permitem uma análise sobre a relevância dos atributos associados estritamente ao domínio da frequência, sem a influência dos demais, e ainda por este motivo, serão as estimativas realizadas com estes atributos selecionados como entrada, que permitirão uma análise sobre a eficácia dos dados no domínio da frequência quando estes são utilizados separadamente dos demais.

Os resultados deixam claro a irrelevância das energias associadas aos dados de energia eólica, as quais não foram selecionadas por nenhum algoritmo. Já as energias associadas aos dados de demanda foram selecionadas tanto para estimativas de preço quanto de demanda, em ambos os algoritmos. Outro ponto importante foram as escolhas do algoritmo Relief para as estimativas de demanda, onde foram selecionadas apenas energias associadas às baixas frequências.

As tabelas 7 e 8 mostram os principais resultados obtidos nesta seção, onde os seletores de atributos foram expostos ao grupo de base de dados #3, na qual estão contidos todos os atributos considerados ao longo do trabalho, para os domínios do tempo e da frequência.

A tabela 7 mostra que os atributos selecionados para as estimativas de demanda, pelo algoritmo Relief, são os mesmos apresentados na tabela 3, com a adição da energia associada às altas frequências dos dados de demanda (EAF - D). O algoritmo CFS selecionou, também para as estimativas de demanda, a energia das altas frequências para as bases de dados de feriados e todos os dias. A comparação entre os resultados obtidos pelas redes treinadas pelos atributos selecionados por Relief e CFS, permitirá determinar se a utilização de dados de

entrada com origem em outros atributos base será, de fato, benéfico para as estimativas em demanda, pois apenas o algoritmo CFS selecionou atributos destas origens.

Tabela 7: Atributos selecionados, por ambos os algoritmos, para as estimativas de demanda considerando o grupo de bases #3

Relief			CFS		
Feriados e FDS	Dias Uteis	Todos os Dias	Feriados e FDS	Dias Uteis	Todos os Dias
Demandas h-1	Demandas h-1	Demandas h-1	Demandas h-1	Demandas h-1	Demandas h-1
Demandas h-23	Demandas h-23	Demandas h-23	Demandas h-23	Demandas h-23	Demandas h-23
Demandas h-2	Demandas h-2	Demandas h-2	EAF - D	Demandas h-12	EAF - D
Demandas h-22	Demandas h-22	EAF - D	Eólica h-3	EAF - B	Eólica h-4
Demandas h-15	Demandas h-16	Demandas h-22	Balanço h-13	-	-
Demandas h-3	Demandas h-17	Demandas h-16	-	-	-
Demandas h-9	Demandas h-7	-	-	-	-
Demandas h-16	Demandas h-8	-	-	-	-
Demandas h-14	Demandas h-3	-	-	-	-
Demandas h-8	Demandas h-15	-	-	-	-
EAF - D	EAF - D	-	-	-	-

A tabela 8 apresenta os atributos selecionados para as estimativas de preço, onde, para o algoritmo Relief, as escolhas foram as mesmas apresentadas na tabela 4, com a adição da energia associada as baixas frequências dos sinais de preço. Estas escolha, associadas à análise feita sobre a tabela 7, permitem concluir que os atributos gerados, a partir dos dados em frequência, utilizando atributos originais diferentes de preço e demanda, são pouco ou irrelevantes para as estimativas desejadas. É importante destacar também, que o algoritmo CFS não selecionou nenhum atributo associado à frequência, o que fará com que as estimativas realizadas para estes atributos de entrada sejam muito próximas das que forem realizadas para os atributos escolhidos para o grupo #1 de base de dados, impossibilitando uma análise comparativo dos resultados.

Tabela 8: Atributos selecionados, por ambos os algoritmos, para as estimativas de preço considerando o grupo de bases #3

Relief			CFS		
Feriados e FDS	Dias Uteis	Todos os Dias	Feriados e FDS	Dias Uteis	Todos os Dias
Preço h-1	Preço h-1	Preço h-1	Demanda h-23	Demanda h-23	Demanda h-23
1 <sup>a</sup> MM Preço	1 <sup>a</sup> MM Preço	1 <sup>a</sup> MM Preço	Balanço h-1	Balanço h-1	Balanço h-1
2 <sup>a</sup> MM Preço	EBF - P	Preço h-2	Preço h-1	Preço h-1	Preço h-1
EBF - P	Preço h-2	EBF - P	11 <sup>a</sup> MM Preço	12 <sup>a</sup> MM Preço	12 <sup>a</sup> MM Preço
Preço h-2	2 <sup>a</sup> MM Preço	2 <sup>a</sup> MM Preço	-	-	-
3 <sup>a</sup> MM Preço	EAF - P	-	-	-	-
Dólar h-120	Dólar h-24	-	-	-	-
-	Dólar h-120	-	-	-	-
-	3 <sup>a</sup> MM Preço	-	-	-	-

# CAPÍTULO 5

## Redes Neurais Artificiais

### 5.1. Introdução

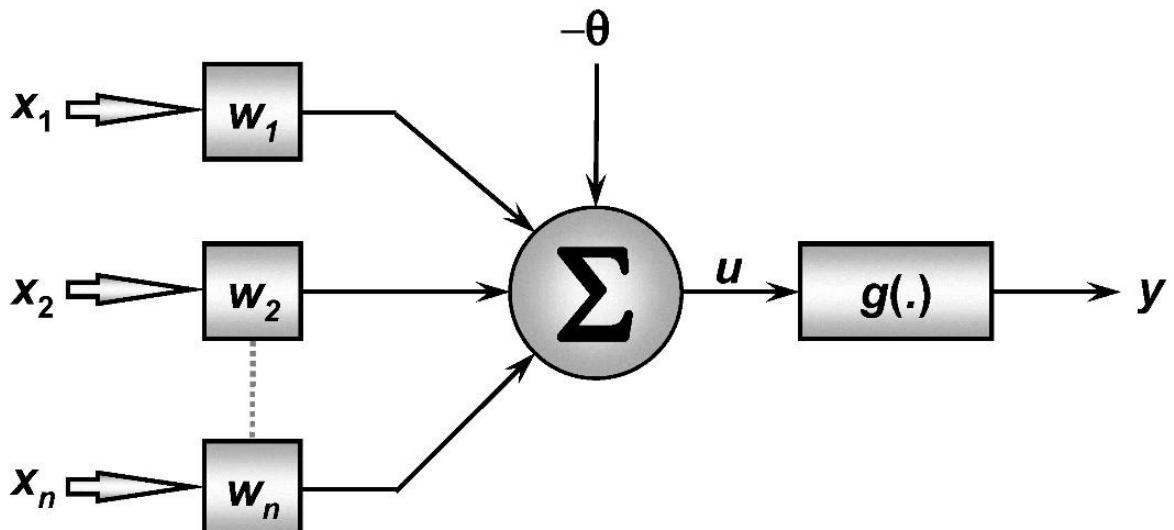
As Redes Neurais Artificiais são modelos computacionais que se baseiam no funcionamento das redes neurais presente no sistema nervoso dos seres vivos [19]. Portanto, tais modelos apresentam capacidades e características semelhantes às de um cérebro humano, dentre as quais podemos citar:

- A presença de neurônios como unidades de processamento;
- A capacidade de aprendizado baseado em informações previamente fornecidas;
- Generalização da informação, sendo capaz de estimar soluções antes desconhecidas

Os modelos de neurônios artificiais, utilizados em RNA, apresentam funcionamento semelhante aos neurônios biológicos encontrados na natureza. Porém, devido às incertezas que ainda existem quanto ao funcionamento dos neurônios biológicos, os modelos artificiais utilizados são limitados àquilo que já é conhecido.

Uma maneira sucinta de entender o neurônio artificial está explicada na figura abaixo.

Figura 7: Esquema de um neurônio artificial



Fonte: [19]

O neurônio recebe as informações de entradas contidas em  $x_n$ , pondera estes valores de acordo com os seus pesos sinápticos presentes em  $w_n$  e os informa para o combinador linear  $\sum$ . O limiar de ativação  $\theta$  determina se o resultado produzido no combinador linear tem um potencial de ativação,  $u$ , positivo ou negativo. Por fim, o potencial de ativação  $u$  serve de entrada para a função de ativação  $g$ , responsável por determinar, e manter em níveis aceitáveis, o valor do sinal de saída  $y$ .

O modelo de neurônio explicado na imagem e texto acima, foi proposto por McCulloch e Pitts [19], e o resultado produzido por este modelo pode ser sintetizado pelas equações:

$$u = \sum_{i=1}^n w_i \cdot x_i - \theta \quad (3)$$

$$y = g(u) \quad (4)$$

Cabe ainda ressaltar que os neurônios artificiais possuem dois modos de operação: treinamento e uso. Durante o treinamento são apresentados aos neurônios conjuntos ordenados de dados de entrada e saídas desejadas. A partir destes, os pesos sinápticos e os limiares de ativação dos neurônios são ajustados para que, durante a fase de uso, o neurônio seja capaz de generalizar as informações, produzindo saídas próximas as desejadas quando alimentados com entradas anteriormente desconhecidas.

Além do modelo do modelo de neurônio, outros aspectos que determinam as características de uma RNA são sua arquitetura, topologia e o processo de treinamento utilizado. Estas características podem variar de acordo com as particularidades apresentadas pelo sistema físico ao qual a rede será inserida, sendo que, será a escolha correta destes aspectos que determinará boa parte do sucesso da rede na resolução do problema apresentado.

## 5.2. Características da RNA Utilizada

Como dito anteriormente, o objetivo do trabalho é o de desenvolver um estimador capaz de prever a demanda e o preço da energia de uma hora a frente aos dados de entrada, se tratando então de uma previsão de série temporal.

Portanto, assim como em [17] a rede neural utilizada foi a *Perceptron* multicamadas, rede que pertence à arquitetura *feedforward* e apresenta uma alta aplicabilidade na resolução de séries temporais.

O algoritmo de treinamento utilizado foi o *Levenberg-Marquardt*, assim como em outras pesquisas semelhantes, [1]-[3], devido à eficiência apresentada durante os processos de treinamento das redes.

Quanto ao número de camadas presentes na arquitetura da rede, cabe dizer que logo nos primeiros testes realizados ficou claro para o aluno que apenas uma camada oculta era o suficiente para o estimador. Sendo assim, foram utilizadas redes com apenas uma camada oculta.



# CAPÍTULO 6

## Resultados e Discussões

A partir da arquitetura descrita no Capítulo 5, foram criadas redes neurais idênticas, que foram treinadas de acordo com cada base de dados de entrada obtidas na seleção de atributos. Cada uma das redes foi analisada para 10 números diferentes de neurônios na camada oculta, variando de 5 até 50 neurônios, incrementando 5 destes a cada análise.

Deve-se ainda mencionar que os dados foram divididos em dois conjuntos, sendo que o conjunto destinado à aprendizagem continha 60% dos dados, e o conjunto de validação os 40% restante.

As seções seguintes irão apresentar o erro percentual absoluto médio (MAPE) de cada uma das redes neurais em seus períodos de validação. Cabe ainda dizer que os dados foram agrupados em tabelas onde os algoritmos de seleção, Relief e CFS, são comparados para cada uma das bases de dados.

### 6.1. Resultados Obtidos com os Dados no Domínio do Tempo

Como dito anteriormente, ao analisar os dados do domínio do tempo os algoritmos de seleção utilizados, determinaram que os atributos relevantes para a estimação de preço e demanda eram os mesmo que aqueles obtidos em [17]. Portanto, esta seção apresentou resultados semelhantes aos encontrados na pesquisa citada.

É importante notar nesta seção que ambos os algoritmos de seleção se mostraram precisos nas estimativas de demanda da energia elétrica. Os erros mostrados na tabela 9 deixam claro que, apesar de próximos, os valores obtidos pelo método de seleção Relief foram mais precisos em todas as divisões de dias.

Diferente da demanda, os valores de erros obtidos para as estimativas de preço se mostraram muito maiores e mais instáveis quando analisados as diferenças entre o maior e menor MAPE apresentado. Os resultados mostram ainda que a separação dos dados em dias úteis torna a estimativa mais eficiente, onde a média de erros para o algoritmo Relief para estes dias foi de 29,18%, já para os dias de finais de semana a média foi de 63,13% para o mesmo algoritmo.

Tabela 9: Resultados obtidos pelas RNAs considerando os atributos selecionados do grupo de base de dados #1 para as estimativas de demanda

Neurônios da Camada Oculta	Relief			CFS		
	Feriados e FDS MAPE(%)	Dias Uteis MAPE(%)	Todos os Dias MAPE(%)	Feriados e FDS MAPE(%)	Dias Uteis MAPE(%)	Todos os Dias MAPE(%)
5	1,21912	0,97213	1,26414	1,54236	1,55565	1,81699
10	1,03533	0,87930	1,06693	1,44681	1,36484	1,75440
15	1,05970	0,83295	1,03205	1,44004	1,36767	1,78575
20	1,06692	0,86534	1,10311	1,49434	1,45778	1,75417
25	1,15641	0,92449	1,12690	1,53305	1,50877	1,83583
30	1,13701	0,90395	1,12816	1,56256	1,49136	1,82754
35	1,11877	0,89395	1,10054	1,48698	1,53770	1,82254
40	1,19840	0,98403	1,16168	1,59598	1,57112	1,86818
45	1,15743	0,92047	1,11309	1,47999	1,50847	1,83264
50	1,14648	0,89131	1,06603	1,67899	1,45723	1,78044
Média	1,1296	0,9068	1,1163	1,5261	1,4821	1,8078

Tabela 10: Resultados obtidos pelas RNAs considerando os atributos selecionados do grupo de base de dados #1 para as estimativas de preço

Neurônios da Camada Oculta	Relief			CFS		
	Feriados e FDS MAPE(%)	Dias Uteis MAPE(%)	Todos os Dias MAPE(%)	Feriados e FDS MAPE(%)	Dias Uteis MAPE(%)	Todos os Dias MAPE(%)
5	43,17941	28,79142	38,64197	31,48718	25,47374	27,16508
10	32,92239	25,63072	32,56173	31,87440	27,73746	27,57371
15	63,86321	29,69251	36,59875	31,41086	26,68251	26,20979
20	42,21748	30,07094	35,68224	34,74937	27,17465	37,27511
25	45,13317	28,44223	33,42781	37,56383	35,33737	28,74637
30	33,04778	25,61342	35,75520	59,54800	29,70712	29,06112
35	41,21787	29,50887	40,94649	43,21417	29,51054	28,34908
40	89,16683	34,69696	41,87525	37,20012	34,16325	31,72256
45	74,40898	30,98599	41,79667	46,74925	28,92833	27,08829
50	166,19295	28,38906	39,45920	37,98999	31,77708	37,71481
Média	63,1350	29,1822	37,6745	39,1787	29,6492	30,0906

## 6.2. Resultados Obtidos com os dados apenas no domínio da frequência

Ao analisar os atributos selecionados no capítulo 4 em ambos os domínios, foi notória a maior relevância dos atributos temporais. Os dados desta seção têm o objetivo de deixar claro esta análise, mostrando que quando as redes são treinadas apenas com os atributos no domínio da frequência, os erros apresentados são maiores para as estimativas de preço e demanda.

Tabela 11: Resultados obtidos pelas RNAs considerando os atributos selecionados do grupo de base de dados #2 para as estimativas de demanda

Neurônios da Camada Oculta	Relief			CFS		
	Feriados e FDS MAPE(%)	Dias Uteis MAPE(%)	Todos os Dias MAPE(%)	Feriados e FDS MAPE(%)	Dias Uteis MAPE(%)	Todos os Dias MAPE(%)
5	10,72498	12,37948	11,59643	10,73678	12,21324	11,41401
10	11,89606	14,40233	11,70008	10,84056	12,19292	11,37118
15	11,65259	12,51976	13,34152	10,82689	12,90208	11,41188
20	13,10390	13,01970	12,68289	11,03193	12,38867	11,44081
25	13,16147	13,49251	13,99626	11,12478	12,26919	11,45951
30	15,34432	15,37498	13,17618	10,86375	12,57456	12,14211
35	17,86036	15,27632	13,32182	12,00871	12,33248	11,41447
40	14,12265	15,79585	15,65460	11,24134	12,33256	11,36566
45	13,31387	17,43833	16,56546	10,86164	12,61299	11,37094
50	13,98019	15,99840	16,68383	10,84870	12,29454	11,47374
Média	13,5160	14,5698	13,8719	11,0385	12,4113	11,4864

Os erros apresentados na tabela 11 são de 10 à 15 vezes maiores que os dados apresentados para as demais estimativas de demanda. Já os erros obtidos para a estimativa de preço, expostos na tabela 12, apresentaram MAPE médio de 62,27%, enquanto as demais estimativas de preço obtiveram média de erro de 36,04%.

Tabela 12: Resultados obtidos pelas RNAs considerando os atributos selecionados do grupo de base de dados #2 para as estimativas de preço

Neurônios da Camada Oculta	Relief			CFS		
	Feriados e FDS MAPE(%)	Dias Uteis MAPE(%)	Todos os Dias MAPE(%)	Feriados e FDS MAPE(%)	Dias Uteis MAPE(%)	Todos os Dias MAPE(%)
5	58,06918	44,92070	55,28382	59,99715	43,41327	49,16837
10	57,98405	49,67108	57,35684	63,25642	49,91761	50,37537
15	57,04776	47,99008	49,20141	70,33623	51,15419	50,25103
20	68,89304	51,07410	47,45198	66,40879	48,20126	49,67033
25	58,22580	45,00769	50,15864	63,64078	116,9491	53,32676
30	60,18634	43,99680	56,59159	93,99459	87,10001	50,78802
35	63,76518	12,00141	52,13329	76,46823	233,2421	50,58855
40	59,59765	52,25735	57,84570	70,91179	92,02006	53,44611
45	68,45809	46,47543	57,79918	76,97391	56,75038	49,63140
50	62,86755	46,69843	63,45644	74,21845	79,99289	51,79500
Média	61,5095	47,5657	54,7279	71,6206	85,8741	50,9041

### 6.3. Resultados Obtidos com os Dados em Ambos os Domínios

Como dito no capítulo 4, a seleção dos dados revelou que os atributos obtidos na análise em frequência dos dados possuem certa relevância para o treinamento das redes. Desta maneira, esta seção apresenta os erros obtidos para as estimativas feitas com dados em ambos os domínios.

É importante mencionar que, como todos os dados a serem apresentados nesta seção se mostraram muito próximos dos dados obtidos na seção 6.1, o aluno realizou seguidos testes como as mesmas entradas, para que os possíveis resultados extremos fossem eliminados.

Como exibido na tabela 13, os MAPEs obtidos, para estimativas de demanda, a partir dos atributos selecionados pelo algoritmo Relief, se mostraram, novamente, mais precisos do que aqueles obtidos a partir do algoritmo CFS. Quando comparados aos dados obtidos na tabela da seção 6.1, a coluna referente a todos os dias apresenta erros ligeiramente menores, que podem ser encarados como uma possível melhora causada pelos dados obtidos no domínio da frequência.

Tabela 13: Resultados obtidos pelas RNAs considerando os atributos selecionados do grupo de base de dados #3 para as estimativas de demanda

Neurônios da Camada Oculta	Relief			CFS		
	Feriados e FDS MAPE(%)	Dias Uteis MAPE(%)	Todos os Dias MAPE(%)	Feriados e FDS MAPE(%)	Dias Uteis MAPE(%)	Todos os Dias MAPE(%)
5	1,14777	1,01450	1,17510	1,98290	1,52441	1,75868
10	1,09007	0,92116	1,08544	2,02368	1,46396	1,73589
15	1,14306	0,88298	1,05626	2,01251	1,43769	1,73984
20	1,08601	0,87744	1,00947	2,04540	1,43782	1,72848
25	1,12755	0,89862	0,88476	2,04184	1,43391	1,73149
30	1,12317	0,89113	0,99578	2,06662	1,46315	1,73791
35	1,16284	0,97845	0,77582	2,07417	1,43995	1,74447
40	1,13231	0,978435	0,98052	2,21980	1,42435	1,75444
45	1,18592	0,87997	0,79408	2,06159	1,44467	1,74038
50	1,12502	0,92072	0,98848	2,15307	1,47461	1,75127
Média	1,1324	0,9243	0,9746	2,0682	1,4545	1,7423

Novamente os MAPEs apresentados para estimativas de preço se mostraram bem maiores que aqueles observados para estimativas de demanda. A diferença entre os algoritmos de seleção também seguiu os padrões apresentados na seção 6.1, onde o algoritmo CFS apresentou o menor MAPE, 24,92%, e média de 28,37% para os dias úteis, enquanto o

algoritmo Relief apresentou média de 35,85% para os mesmos dias. Porém, é importante lembrar que o algoritmo CFS não selecionou dados no domínio da frequência para a estimativa de preço, o que torna os resultados obtidos nesta seção, para este algoritmo, iguais, em sua origem dos dados de entrada, aos obtidos na seção 6.1.

Tabela 14: Resultados obtidos pelas RNAs considerando os atributos selecionados do grupo de base de dados #3 para as estimativas de preço

Neurônios da Camada Oculta	Relief			CFS		
	Feriados e FDS MAPE(%)	Dias Uteis MAPE(%)	Todos os Dias MAPE(%)	Feriados e FDS MAPE(%)	Dias Uteis MAPE(%)	Todos os Dias MAPE(%)
5	38,54751	30,64825	28,13423	35,97496	28,29985	29,38587
10	40,27990	33,45915	28,43051	33,67583	24,92128	27,49878
15	44,74877	33,44274	27,38789	34,59171	28,37133	27,46555
20	72,85994	41,72809	29,41148	40,10237	28,04239	30,05669
25	39,92144	34,42736	29,40532	37,76698	29,36648	28,57969
30	38,66177	34,03669	35,07950	48,41101	29,03378	27,37042
35	42,31526	33,56323	28,41938	37,23810	27,97853	27,50911
40	40,86236	42,21219	31,46702	42,51070	27,97420	29,79106
45	50,70725	36,71655	28,48941	38,21958	29,29579	28,45528
50	41,24519	38,26638	28,03249	33,54996	28,46149	31,28844
Média	45,0149	35,8501	29,4257	38,2041	28,1745	28,7401

Como comparação final, é importante apontar a melhora, mesmo que pequena, apresentada para as estimativas de demanda de Todos os Dias, dos dois seletores de atributos, quando utilizados os dados no domínio da frequência. As médias dos MAPEs para as estimativas destes dias estão expostas na tabela 15.

Tabela 15: Média de MAPEs das estimativas de demanda para Todos os Dias

	Relief	CFS
Dados no domínio do tempo	1,116%	1,807%
Dados em ambos os domínios	0,974%	1,745%

O destaque desta análise foi a diminuição de 12,72%, entre as médias de erro nas estimativas de demanda do seletor Relief para Todos os dias. Desta maneira, pode-se concluir que, ao adicionar os dados do domínio da frequência na base de atributos fornecidos para a rede neural, houve uma melhora significativa nas estimativas de demanda para todos os dias.



# CAPÍTULO 7

## Conclusões

Este trabalho apresentou uma análise sobre os estimadores de preço e demanda de energia elétrica, bem como sobre os dados de entrada fornecidos para estes estimadores. Um dos objetivos deste trabalho era o de demonstrar a importância da análise no domínio da frequência e a relevância dos dados obtidos nestas análises.

O cenário de pesquisas apresentadas como embasamento teórico deste trabalho deixou claro a necessidade eminente por estimadores que sejam aplicáveis em sistemas com características diferentes das apresentadas pelos sistemas de energia comumente conhecidos. Além disto, as recentes aplicações de redes inteligentes irão aumentar cada vez a quantidade de dados fornecidos sobre os sistemas de energia, que quando alinhados com a imprevisibilidade dos sistemas irá gerar uma demanda por estimadores cada vez mais precisos e concisos na utilização dos dados.

Desta maneira, a etapa de seleção de atributos apresentada neste trabalho foi de extrema importância para a eliminação de dados irrelevantes para estas estimações, deixando claro que sem esta prévia análise, os custos computacionais tornariam as estimações inviáveis para as futuras aplicações mencionadas acima. Este capítulo revelou ainda que a observação e criação de dados no domínio da frequência são uma alternativa para aumentar a qualidade dos dados de entrada, possibilitando um sucesso maior dos estimadores.

A redução na taxa de erros, cometidos nas estimações de demanda de energia elétrica, confirma a relevância dos dados em frequência, levando a concluir que a utilização de dados em ambos os domínios pode ser a chave para se obter uma base de dados concisa e eficiente para as previsões desejadas.



# Referências Bibliográficas

- [1] ZAREIPOUR, Hamidreza et al. Application of Public-Domain Market Information to Forecast Ontario's Wholesale Electricity Prices. **IEEE Transactions on Power Systems**, v. 21, n. 4, nov. 2006.
- [2] AMJADY, Nima; DARAEEPOUR, Ali; KEYNIA, Farshid. Day-ahead electricity price forecasting by modified Relief algorithm and hybrid neural network. **IET Generation, Transmission & Distribution**, v. 4, Iss. 3, pp. 432– 444, 2010.
- [3] AGGARWAL, Sanjeev Kumar; SAINI, Lalit Mohan; KUMAR, Ashwani. Electricity Price Forecasting in Ontario Electricity Market Using Wavelet Transform in Artificial Neural Network Based Model. **International Journal of Control, Automation and Systems**, v. 6, n. 5, pp. 639 – 650, oct. 2008.
- [4] MOHSENIAN-RAD, Amir-Hamed; GARCIA, Alberto Leon. Optimal Residential Load Control With Price Prediction in Real-Time Electricity Pricing Environments. **IEEE Transactions on Smart Grid**, v. 1, n. 2, set. 2010.
- [5] RODRIGUEZ, Claudia P.; ANDERS, George J.. Energy Price Forecasting in the Ontario Competitive Power System Market. **IEEE Transactions on Power Systems**, v. 19, n. 1, feb. 2004.
- [6] HUANG, Dongliang et al. Data Mining for Electricity Price Classification and the Application to Demand-Side Management. **IEEE Transactions on Smart Grid**, v. 3, n. 2, jun. 2012.
- [7] ABDEL-KARIM, Noha; ILIC, Marija. Modeling Uncertain Load and Wind Power in the Electric Energy Systems. **Power and Energy Society General Meeting IEEE**, San Diego, 2012.
- [8] HERNÁNDEZ, Luis et al. A Multi-Agent System Architecture for Smart Grid Management and Forecasting of Energy Demand in Virtual Power Plants. **Ultimate Technologies and Advances for Future Smart Grid – UTASG. IEEE Communications Magazine**, jan. 2013.
- [9] SANDHUA, Harmanjot Singh; FANGA, Liping; GUANB, Ling. Forecasting day-ahead price spikes for the Ontario electricity market. **Electric Power Systems Research**, 2016.
- [10] HOBBY, John D.; SHOSHITAISHVILI, Alexander; TUCCI, Gabriel H.. Analysis and Methodology to Segregate Residential Electricity Consumption in Different Taxonomies. **IEEE Transactions on Smart Grid**, v. 3, n. 1, mar. 2012.

- [11] ZAREIPOUR, Hamidreza et al. Classification of Future Electricity Market Prices. **IEEE Transactions on Power Systems**, v. 26, n. 1, feb. 2011.
- [12] NYENG, Preben; ØSTERGAARD, Jacob. Information and Communications Systems for Control-by-Price of Distributed Energy Resources and Flexible Demand. **IEEE Transactions on Smart Grid**, v. 2, n. 2, jun. 2011.
- [13] CHAN, S.C. et al. Load/ Price Forecasting and Managing Demand Response for Smart Grids. **IEEE Signal Processing Magazine**, sep. 2012.
- [14] MATHIEU, Johanna L. et al. Quantifying Changes in Building Electricity Use, With Application to Demand Response. **IEEE Transactions on Smart Grid**, v. 2, n. 3, sep. 2011.
- [15] GATSIS, Nikolaos; GIANNAKIS, Georgios B.. Residential Load Control: Distributed Scheduling and Convergence With Lost AMI Messages. **IEEE Transactions on Smart Grid**, v. 3, n. 2, jun. 2012.
- [16] AMJADY, Nima; KEYNIA, Farshid; ZAREIPOUR, Hamidreza. Short-Term Load Forecast of Microgrids by a New Bilevel Prediction Strategy. **IEEE Transactions on Smart Grid**, v. 1, n. 3, dec. 2010.
- [17] DOS SANTOS, Maristela Silva. **Desenvolvimento de Estimador de Preço e Demanda de Energia Elétrica Baseado em Sistema Inteligente aplicado para Smart Grids**. 57 f. Monografia – Departamento de Engenharia Elétrica e de Computação, Universidade de São Paulo, 2016.
- [18] OPPENHEIM, Alan V.; SCHAFER, Ronald W. **Discrete-Time Signal Processing**. 2. Ed. New Jersey: Prentice Hall, 1999.
- [19] DA SILVA, Ivan Nunes; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. **Redes Neurais Artificiais para Engenharia e Ciências Aplicadas – Curso Prático**. São Paulo: Artliber, 2010.
- [20] Câmera de Comercialização de Energia Elétrica. **CCEE Atinge a Marca de 10 mil ativos modelados**. Disponível em: <[https://www.ccee.org.br/portal/faces/pages\\_publico/noticias-opiniao/noticias/noticialeitura?contentid=CCEE\\_386153&\\_afrLoop=202096453666019%40%3Fcontentid%3DCCEE\\_386153%26\\_afrLoop%3D202096453666019%26adf.ctrl-state%3Dwx9fobm3\\_4](https://www.ccee.org.br/portal/faces/pages_publico/noticias-opiniao/noticias/noticialeitura?contentid=CCEE_386153&_afrLoop=202096453666019%40%3Fcontentid%3DCCEE_386153%26_afrLoop%3D202096453666019%26adf.ctrl-state%3Dwx9fobm3_4)>. Acesso em: 02 Mar. 2017.
- [21] Fernandes, R. A. S., Identificação de fontes de correntes harmônicas por redes neurais artificiais, Dissertação de Mestrado, Escola de Engenharia de São Carlos, Universidade de São Paulo, 2009.