

**UNIVERSIDADE FEDERAL DE JUIZ DE FORA  
FACULDADE DE ENGENHARIA  
ENGENHARIA ELÉTRICA – HABILITAÇÃO EM ROBÓTICA E AUTOMAÇÃO  
INDUSTRIAL**

**Fabiane Barbosa do Nascimento**

**REDES NEURAIIS ARTIFICIAIS APLICADAS À PREDIÇÃO DO PREÇO DE  
LIQUIDAÇÃO DAS DIFERENÇAS NO MERCADO DE ENERGIA**

**Juiz de Fora  
2017**

**Fabiane Barbosa do Nascimento**

**REDES NEURAS ARTIFICIAIS APLICADAS À PREDIÇÃO DO PREÇO DE  
LIQUIDAÇÃO DAS DIFERENÇAS NO MERCADO DE ENERGIA**

Trabalho de Conclusão de Curso apresentado ao programa de graduação em Engenharia Elétrica, da Universidade Federal de Juiz de Fora, como requisito parcial a obtenção do grau de Engenheira Eletricista.

Orientador: Leonardo Rocha Olivi  
Coorientador: Luís Henrique Lopes Lima

**Juiz de Fora  
2017**

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Nascimento, Fabiane Barbosa do.

REDES NEURAIAS ARTIFICIAIS APLICADAS À PREDIÇÃO DO PREÇO DE LIQUIDAÇÃO DAS DIFERENÇAS NO MERCADO DE ENERGIA / Fabiane Barbosa do Nascimento. -- 2017.

54 p.

Orientador: Leonardo Rocha Olivi

Coorientador: Luis Henrique Lopes Lima

Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Juiz de Fora, Faculdade de Engenharia, 2017.

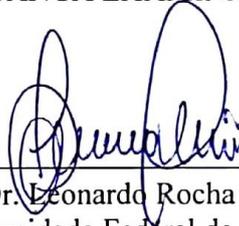
1. Redes Neurais Artificiais. 2. Preço de Liquidação das Diferenças. 3. Mercado Livre de Energia. 4. Predição do Preço de Energia. I. Olivi, Leonardo Rocha, orient. II. Lima, Luis Henrique Lopes, coorient. III. Título.

**Fabiane Barbosa do Nascimento**

**REDES NEURAIAS ARTIFICIAIS APLICADAS À PREDIÇÃO DO PREÇO DE  
LIQUIDAÇÃO DAS DIFERENÇAS NO MERCADO DE ENERGIA**

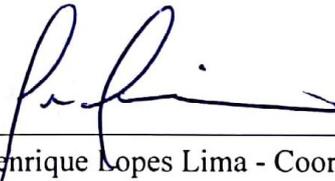
Trabalho de Conclusão de Curso apresentado ao programa de graduação em Engenharia Elétrica, da Universidade Federal de Juiz de Fora, como requisito parcial a obtenção do grau de Engenheira Eletricista.

**BANCA EXAMINADORA**



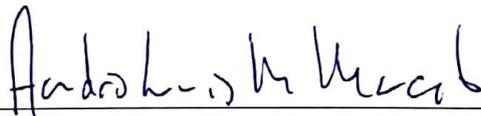
---

Professor Dr. Leonardo Rocha Olivi - Orientador  
Universidade Federal de Juiz de Fora



---

Professor Dr. Luís Henrique Lopes Lima - Coorientador  
Universidade Federal de Juiz de Fora



---

Professor Dr. André Luís Marques Marcato  
Universidade Federal de Juiz de Fora



## ATA DE APRESENTAÇÃO DE TRABALHO FINAL DE CURSO

DATA E HORA DA DEFESA: 27/11/2017 – 15H00

CANDIDATA: **FABIANE BARBOSA DO NASCIMENTO**

ORIENTADOR: **PROF. LEONARDO ROCHA OLIVI**

COORDENADOR: **PROF. LUÍS HENRIQUE LOPES LIMA**

TÍTULO DO TRABALHO: **REDES NEURAIAS ARTIFICIAIS APLICADAS À PREDIÇÃO DO PREÇO DE LIQUIDAÇÃO DAS DIFERENÇAS NO MERCADO DE ENERGIA**

BANCA EXAMINADORA/INSTITUIÇÃO:

PREDIDENTE: **PROF. LEONARDO ROCHA OLIVI - UFJF**

AVALIADOR 1: **PROF. LUÍS HENRIQUE LOPES LIMA - UFJF**

AVALIADOR 2: **PROF. ANDRÉ LUÍS MARQUES MARCATO - UFJF**

LOCAL: **ANFITEATRO 2 – EDIFÍCIO ITAMAR FRANCO – UNIVERSIDADE FEDERAL DE JUIZ DE FORA**

Nesta data, em sessão pública, após exposição oral de 25 minutos, a candidata foi arguida pelos membros da banca. Em decorrência desta arguição, a banca considerou a candidata:

**APROVADA**

**REPROVADA**

Na forma regulamentar foi lavrada a presente Ata que é abaixo assinada pelos membros da banca na ordem determinada e pela candidata:

PREDIDENTE: Leonardo Rocha Olivi  
AVALIADOR 1: Luís Henrique Lopes Lima  
AVALIADOR 2: André Luís Marques Marcato  
CANDIDATA: Fabiane B. do Nascimento

Juiz de Fora, 27 de novembro de 2017.

Vistos: (coordenador do curso ou presidente da comissão de TFCs do curso).

Engenharia Elétrica  
Robótica e Automação  
COORDENAÇÃO DE CURSO  
Faculdade de Engenharia - UFJF

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus, que me ouviu em momentos difíceis, me confortou e me deu forças para chegar onde estou.

Agradeço aos meus pais, José Luiz e Eva, que são e sempre foram meus grandes amigos e incentivadores. Eles nunca mediram esforços e empenharam suas vidas para que eu pudesse chegar até aqui. Agradeço também a minha irmã Jociane, que esteve sempre à disposição para me ajudar, demonstrando confiança, respeito e compreensão nas minhas escolhas.

Ao professor, coordenador e orientador Leonardo Rocha Olivi, primeiramente por ser o introdutor da área de redes neurais no meu percurso acadêmico e também pela excelência na orientação e dedicação concedidos para a realização deste trabalho. Além dos ensinamentos, contribuições e segurança transmitida, fatores fundamentais não só para a realização desta monografia, mas também para minha formação profissional.

Quero também agradecer ao professor Luís Henrique Lopes Lima. Suas sugestões contribuíram muito para o desenvolvimento deste trabalho.

Aos amigos e professores do curso de graduação que direta ou indiretamente contribuíram nessa caminhada, em especial à Fabiana, Alessandra, Bruna e Ana Carolina, que me apoiaram, consolaram e ensinaram.

"Obstáculo é aquilo que você enxerga,  
quando tira os olhos do seu objetivo."  
(Henry Ford)

## RESUMO

O Preço de Liquidação das Diferenças (PLD) é o preço de curto prazo divulgado semanalmente pela Câmara de Comercialização de Energia Elétrica (CCEE), sendo ele uma variável relevante para o mercado livre de energia elétrica. Acompanhá-lo e entender seu funcionamento permite que os participantes do mercado consigam definir suas estratégias de maneira assertiva, resultando em contratações energéticas bem-sucedidas, com os montantes adequados, no momento certo e com menores preços. O PLD sofre influência de diversos fatores, ligados principalmente à incerteza da demanda e à hidrologia, fazendo com que o mesmo sofra uma excessiva variabilidade, adquirindo características de sistemas dinâmicos. O presente trabalho propõe a utilização de Redes Neurais Artificiais visando auxiliar o processo de decisão da compra de energia no Mercado Livre de Energia. O modelo proposto se trata de uma rede neural artificial recorrente, treinada por meio do algoritmo *resilient-backpropagation*, aplicada ao mercado brasileiro de energia. A rede mostrou desempenho satisfatório, sendo capaz de realizar uma predição de 24 semanas à frente para o submercado Sudeste/Centro-Oeste, com nível de acerto relevante. Por conseguinte, o modelo se torna uma ferramenta atrativa no que concerne à atenuação de riscos da comercialização de energia.

Palavras-chave: Redes Neurais Artificiais, Preço de Liquidação das Diferenças, Mercado Livre de Energia, Predição de Preço de Energia.

## **ABSTRACT**

The Settlement Price of Differences (PLD) is a weekly short-term price published by the Brazilian Electric Energy Trading Chamber (CCEE), being a relevant variable for the Free Electricity Market. Accompanying and understanding its operation allows market participants to define optimal strategies, resulting in a successful energy contract, with the right amounts, at the right time, and at lower prices. Several factors influence the PLD, as the uncertainty of demand and hydrology, causing excessive variability and configuring characteristics of dynamical systems for the PLD. The present work proposes an Artificial Neural Network for the decision process of the energy purchase in the Free Energy Market. The proposed model is a recurrent artificial neural network, trained through the resilient-backpropagation algorithm, applied to the Brazilian energy market. The network showed satisfactory performance, being able to predict 24 weeks ahead for the Southeast/Center-West submarket, with a relevant success level. Therefore, the model becomes an attractive tool for the risk mitigation of the commercialization of energy in the free market.

**Keywords:** Artificial Neural Networks, Settlement Price of Differences, Free Energy Market, Energy Price Prediction.

## LISTA DE ILUSTRAÇÕES

Figura 1.1 – Duração dos contratos .....	16
Figura 2.1 – Regiões do Sistema Interligado Nacional .....	20
Figura 2.2 – Média mensal do PLD – submercado SE/CO .....	21
Figura 2.3 – Média mensal do PLD – submercado S .....	22
Figura 2.4 – Média mensal do PLD – submercado NE .....	22
Figura 2.5 – Média mensal do PLD – submercado N .....	23
Figura 2.6 – Cronograma de decisão para um sistema hidrotérmico .....	23
Figura 2.7 – Neurônio biológico .....	26
Figura 2.8 – Neurônio Artificial .....	27
Figura 2.9 – Função de ativação degrau .....	28
Figura 2.10 – Função de ativação sinal .....	29
Figura 2.11 – Função de ativação rampa simétrica .....	29
Figura 2.12 – Função de ativação logística .....	30
Figura 2.13 – Função de ativação tangente hiperbólica .....	30
Figura 2.14 – Exemplo de rede <i>feedforward</i> de camada simples .....	32
Figura 2.15 – Exemplo de rede <i>feedforward</i> de camada múltiplas .....	32
Figura 2.16 – Exemplo de rede recorrente .....	33
Figura 3.1 – Normalização dos dados a serem aprendidos .....	35
Figura 3.2 – Treinamento da Rede Neural Artificial recorrente.....	40
Figura 3.3 – Descrição do processo de treinamento e predição por meio da rede neural recorrente .....	42
Figura 4.1 – Desempenho da rede (RMSE x $n_p$ ) para $N_1=25$ neurônios e 1 passo no futuro...45	
Figura 4.2 – Desempenho da rede (RMSE x $n_p$ ) para $N_1=40$ neurônios e 1 passo no futuro...45	
Figura 4.3 – Desempenho da rede (RMSE x $n_p$ ) para $N_1=70$ neurônios e 1 passo no futuro...46	
Figura 4.4 – Fase de treinamento para $N_1=40$ neurônios intermediários e $n_q=3$ saídas realimentadas .....	47
Figura 4.5 – Predição para $N_1=40$ neurônios intermediários e $n_q=3$ saídas realimentadas .....	47
Figura 4.6 – Fase de treinamento para $N_1=70$ neurônios intermediários e $n_q=6$ saídas realimentadas .....	48
Figura 4.7 – Predição para $N_1=70$ neurônios intermediários e $n_q=6$ saídas realimentadas .....	48
Figura 4.8 – Indicadores para a rede com 25 neurônios na camada intermediária x número de saídas realimentadas .....	49

Figura 4.9 – Indicadores para a rede com 40 neurônios na camada intermediária x número de saídas realimentadas .....	50
Figura 4.10 – Indicadores para a rede com 25 neurônios na camada intermediária x número de saídas realimentadas .....	50
Figura 4.11 – Predição para $N_1=40$ neurônios intermediários e $n_q= 3$ saídas realimentadas ...	53
Figura 4.12 – Predição para $N_1=70$ neurônios intermediários e $n_q= 6$ saídas realimentadas ...	54

## LISTA DE TABELAS

Tabela 4.1 – Topologias das RNAs aplicadas à solução do problema .....	44
Tabela 4.2 – Indicadores para horizontes de 4, 12 e 24 semanas com 40 neurônios e 1 passo à frente para cada rodada da rede .....	51
Tabela 4.3 – Indicadores para horizontes de 4, 12 e 24 semanas à frente com $N_1 = 40$ e $n_q = 3$ .....	53
Tabela 4.4 – Indicadores para horizontes de 4, 12 e 24 semanas à frente com $N_1 = 70$ e $n_q = 6$ .....	54

## LISTA DE ABREVIATURAS E SIGLAS

ACR	Ambiente de Contratação Regulada
ACL	Ambiente de Contratação Livre
PLD	Preço de Liquidação das Diferenças
CCEE	Câmara de Comercialização de Energia Elétrica
RNA	Redes Neurais Artificiais
RE-SEB	Reestruturação do Setor Elétrico Brasileiro
ANEEL	Agência Nacional de Energia Elétrica
ONS	Operador Nacional do Sistema
EPE	Empresa de Pesquisa Energética
CMSE	Comitê de Monitoramento do Setor Elétrico
SIN	Sistema Interligado Nacional
S	Sul
SE/CO	Sudeste/Centro-Oeste
NE	Nordeste
N	Norte
ARMAX	<i>ARMA with explanatory variables</i>
ARIMA	<i>Autoregressive integrated moving average</i>
AR	<i>Autoregressive</i>
RMSE	<i>Root Mean Squared Error</i>
MAE	<i>Mean Absolute Error</i>
IC	Intervalo de Confiança

## SUMÁRIO

<b>1. INTRODUÇÃO .....</b>	<b>15</b>
<b>2. REVISÃO BIBLIOGRÁFICA.....</b>	<b>18</b>
2.1    O SETOR ELÉTRICO BRASILEIRO .....	18
2.2    O MERCADO DE CURTO PRAZO E A VOLATILIDADE DO PLD.....	19
2.3    PREDIÇÃO DO PLD.....	24
2.4    REDES NEURAIS ARTIFICIAIS.....	25
<b>2.4.1    Neurônio Biológico .....</b>	<b>26</b>
<b>2.4.2    Modelagem do Neurônio Artificial .....</b>	<b>27</b>
<b>2.4.3    Arquiteturas de Redes Neurais Artificiais .....</b>	<b>31</b>
<b>3. METODOLOGIA .....</b>	<b>34</b>
<b>4. RESULTADOS.....</b>	<b>43</b>
4.1    DEFINIÇÃO DA TOPOLOGIA.....	43
4.2    TREINAMENTO E VALIDAÇÃO PARA UM PASSO NO FUTURO .....	44
4.3    TREINAMENTO E VALIDAÇÃO PARA VÁRIOS PASSOS NO FUTURO ...	51
<b>5. CONCLUSÕES .....</b>	<b>56</b>
5.1    CONCLUSÕES.....	56
5.2    PROPOSTAS DE DESENVOLVIMENTOS FUTUROS .....	57
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>58</b>

## 1. INTRODUÇÃO

A energia elétrica é um fator determinante para que um país cresça e desenvolva sua indústria e comércio para criar oportunidades para seus cidadãos. Para algumas empresas, a energia ocupa um papel tão essencial quanto ao da matéria-prima, tecnologia ou mão de obra. Por se tratar de um insumo fundamental para a economia de uma nação e para seu parque industrial e comercial, a comercialização da energia elétrica será objeto de estudo do presente trabalho.

A comercialização de energia elétrica no Brasil está segmentada em dois ambientes de contratação a saber: O Ambiente de Contratação Regulada (ACR), onde o consumidor é denominado cativo pois adquire a energia das concessionárias de distribuição às quais estão conectados via ponto de entrega do ramal de entrada; e o Ambiente de Contratação Livre (ACL), formado pelos consumidores denominados livres e especiais, pois estes podem comprar a energia diretamente dos entes geradores, como as hidrelétricas, ou via entidades comercializadoras de energia, empresas que compram e negociam energia [1].

O Mercado livre de energia é um ambiente competitivo de negociação, através do qual se realizam operações de compra e venda de energia elétrica. Os contratos são bilaterais, livremente negociados entre o vendedor e o consumidor, que não é obrigado a comprar da concessionária local. Este ambiente de negociação gera muitos benefícios para as empresas consumidoras, como a possibilidade de conseguir custos menores que aqueles que pagariam no mercado regulado. Outra vantagem é que ele propicia ao comprador uma maior gestão sobre a energia por ele demandada [1].

A flexibilidade dos contratos no Mercado Livre permite ao consumidor uma compra de acordo com o seu perfil de consumo, que melhor atenda suas expectativas na relação custo-benefício. Por se tratarem de contratos livremente negociados, as contratações podem ser de curto, médio ou longo prazo. O gráfico da Figura 1.1 ilustra a duração dos contratos de compra por consumidores livres e especiais no Ambiente de Contratação Livre.

Dessa maneira, os fatores preço, índices de correção pré-estabelecidos e a flexibilidade dos contratos possibilitam que a empresa aproveite todas as oportunidades para planejar, economizar e racionalizar sua demanda energética.

Os Artigos 15 e 16 da Lei do Setor Elétrico, Lei nº 9.074/95, definem que todos os consumidores, que possuem uma demanda igual ou superior a 3.000 kW, atendidos em tensão maior ou igual a 69kV, se conectados à rede elétrica antes de 08/07/1995, ou, atendidos em qualquer tensão, se conectados à rede elétrica após 08/07/1995, podem optar por contratar seu

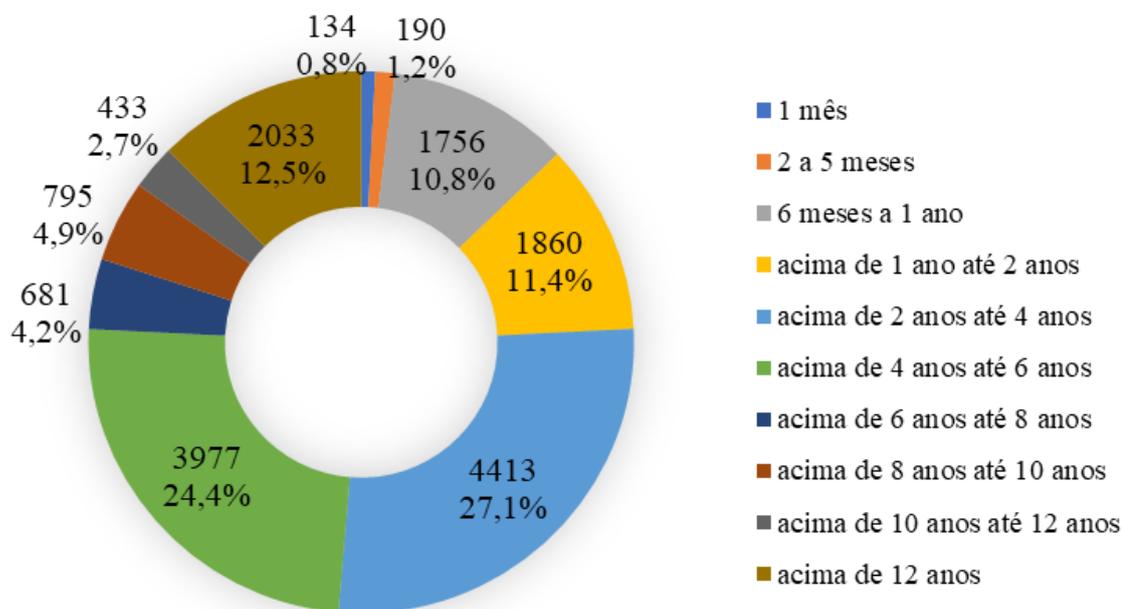


Figura 1.1 – Duração dos contratos  
 Fonte: INFOMERCADO MENSAL CCEE, agosto de 2017 [2]

fornecimento, no todo ou em parte, com qualquer concessionário, permissionário ou autorizado de energia elétrica do mesmo sistema interligado [3]. Além desses, existe o consumidor livre especial, que tem uma demanda média entre 500 kW e 3.000 kW e pode comprar energia apenas de fontes alternativas, como hidrelétricas, usinas eólicas ou termelétricas de pequeno porte.

Para os casos de clientes que consumirem uma quantidade diferente da que foi previamente contratada, a eles é possível realizar contratos no mercado de curto prazo até o nono dia útil do mês subsequente ao consumo de forma a quitar essas diferenças, que são valoradas por meio do Preço de Liquidação das Diferenças (PLD) [4]. O PLD é publicado semanalmente pela CCEE (Câmara de Comercialização de Energia Elétrica) [5] e a maneira como ele é calculado está intrinsecamente conectada com as características operativas do setor elétrico brasileiro.

O sistema elétrico brasileiro é predominantemente hidrelétrico [6]. Um aspecto importante da operação diz respeito ao fato que é impossível prever com exatidão as aflúncias, isto é, a quantidade de água que chega aos reservatórios. Conseqüentemente, dado que as condições de suprimento dependem de quando, quanto e onde chove, agrega-se assim um caráter probabilístico ao processo de geração.

A limitação da geração de energia nas usinas hidrelétricas, causada pela capacidade dos reservatórios, adicionado à incerteza das aflúncias, introduz uma relação entre uma decisão

operativa tomada no presente e as consequências futuras desta decisão. Por exemplo, se os reservatórios forem utilizados no presente e ocorrer um período com baixas afluências será necessário um elevado nível de utilização termelétrica. Em contrapartida, se os reservatórios são pouco utilizados no presente e ainda, um elevado nível de afluências vir a ocorrer será preciso abrir as comportas do vertedouro para liberar as águas do reservatório.

Por sua natureza estocástica, existe uma excessiva variabilidade do PLD. Sendo assim, esta deve ser uma variável a ser considerada no processo de migração para o Mercado Livre de Energia. Se o consumidor possuir um conhecimento prévio da tendência de variação do PLD, é possível definir estratégias de contratação de energia que podem refletir em um generoso aumento no lucro final da empresa, quando bem empregadas.

Como o problema probabilístico relatado leva a um modelo de predição de resultados futuros, se faz necessário estudar técnicas que atendam a esta especificação. As Redes Neurais Artificiais (RNA) têm grande aplicabilidade à predição de séries temporais. São redes desenvolvidas com base em modelos de sistemas de neurônios biológicos. O modelo de RNA empregado ao trabalho proposto obteve resultados baseados na entrada de dados históricos do PLD, que possibilitaram o cálculo de valores futuros (problema de predição de dados) de maneira dinâmica e acurada.

A organização do presente trabalho se dá em cinco capítulos. O primeiro capítulo trata-se desta introdução, apresentando o tema proposto. O segundo capítulo trata de uma revisão bibliográfica dos temas abordados no trabalho. O terceiro capítulo descreverá a modelagem matemática da técnica utilizada na resolução do problema. O quarto capítulo mostra os resultados e suas implicações no tema proposto. O quinto capítulo conclui os resultados e aponta trabalhos futuros.

## 2. REVISÃO BIBLIOGRÁFICA

Este capítulo faz uma revisão técnica acerca dos assuntos que serão tratados neste trabalho. Segue uma descrição do setor elétrico brasileiro, o foco de aplicação do trabalho, e também uma descrição das técnicas utilizadas para a análise.

### 2.1 O SETOR ELÉTRICO BRASILEIRO

No final do século XX, ocorreram importantes transformações no setor elétrico brasileiro. A primeira modificação institucional na estrutura do setor elétrico ocorreu no governo de Itamar Franco, quando foi promulgada a Lei nº 8.631/93, que extinguiu a equalização tarifária vigente e criou contratos de suprimento de energia entre geradores e distribuidores [7] e [8]. Esta Lei foi responsável por abrir caminho para a reestruturação da indústria de energia elétrica. Em 07 de julho de 1995, a Lei nº 9.074 estabeleceu a criação do Produtor Independente de Energia e do Consumidor Livre de Energia, estabelecendo uma nova formatação da comercialização de energia elétrica no país [3].

De acordo com o autor de [7], consultores nacionais e internacionais, liderados pela Coopers&Lybrand [9] foram contratados em 1996 para a condução do programa de reforma do setor elétrico. Os resultados deste trabalho foram consolidados no Projeto Reestruturação do Setor Elétrico Brasileiro (RE-SEB), que previa a separação das empresas de energia em áreas específicas: geração, transmissão, distribuição e comercialização de energia e incentivar a competição nos segmentos de geração e comercialização, mantendo sob regulação do Estado os setores de distribuição e transmissão de energia elétrica [10].

O Projeto RE-SEB apresentou a necessidade da criação de uma estrutura regulatória composta pelo órgão regulador e fiscalizador, a Agência Nacional de Energia Elétrica (ANEEL), o Operador Nacional do Sistema (ONS) que cria e aplica regras e normas para o gerenciamento harmônico da produção e transmissão de eletricidade e o Mercado Atacadista de Energia Elétrica, um ambiente para a realização das transações de compra e venda de energia elétrica. Contudo, a grave crise de abastecimento sofrida em 2001 pelo setor elétrico resultou na implementação de um plano de racionamento de energia elétrica bem como em propostas de alterações no modelo proposto, criadas pelo Comitê de Revitalização do Modelo do Setor Elétrico.

Segundo o autor de [10] o novo modelo do setor elétrico consolidou-se entre os anos de 2003 e 2004, sendo definida a criação de uma entidade responsável pelo planejamento do setor

elétrico a longo prazo, a Empresa de Pesquisa Energética (EPE), o Comitê de Monitoramento do Setor Elétrico (CMSE), responsável por avaliar permanentemente a segurança do suprimento de energia elétrica e uma instituição para a comercialização de energia elétrica no sistema interligado, dando continuidade às atividades do MAE, a Câmara de Comercialização de Energia Elétrica (CCEE).

O novo modelo do setor elétrico brasileiro tem como bases institucionais e legais, as Leis nº 10.847 e 10.848, ambas de 2004. A última dispõe sobre a comercialização de energia elétrica, estabelecendo regras para a comercialização de energia entre concessionários, permissionários e autorizados de serviços e instalações de energia elétrica, bem como destes com seus consumidores. Em relação à comercialização, foram criados dois ambientes para a celebração de contratos de compra e venda de energia, com lógicas e estruturas distintas: O Ambiente de Contratação Livre (ACL) e o Ambiente de Contratação Regulada (ACR) [11].

A CCEE define que no Ambiente de Contratação Regulada os participantes são geradoras, distribuidoras e comercializadoras, visando atender as demandas dos consumidores denominados cativos, que engloba os consumidores residenciais e indústrias com nível de consumo menores. Neste ambiente a contratação é realizada por meio de leilões de energia que são promovidos pela CCEE, sob delegação da ANEEL. Os contratos também são regulados pela ANEEL e o preço da energia é estabelecido em leilões de compra e venda de energia elétrica.

Já no Ambiente de Contratação Livre, além das geradoras e comercializadoras, participam os consumidores livres e especiais, caracterizados por empresas com um consumo mais elevado. Os consumidores livres tradicionais são grandes consumidores, com contratos de energia acima de 3000 kW já os consumidores livres especiais têm contratos entre 500 kW e 3000 kW sendo empresas de pequeno e médio portes, de acordo com a Lei nº 9.074/95. Para este grupo, a energia contratada é advinda de fontes limpas, como solar, eólica e pequenas centrais hidrelétricas. No ACL, os contratos são livremente negociados entre compradores e vendedores por meio de contratos bilaterais, definindo-se preços, volumes e prazos.

## 2.2 O MERCADO DE CURTO PRAZO E A VOLATILIDADE DO PLD

O Mercado de Curto Prazo é um segmento da CCEE onde são contabilizadas as diferenças entre os montantes de energia elétrica contratados pelos agentes e os montantes de geração e de consumo efetivamente verificados e atribuídos aos respectivos agentes. Todos os contratos de compra e venda de energia celebrados no mercado, tanto no Ambiente de

Contratação Regulado como no Ambiente de Contratação Livre devem ser registrados na CCEE, que realiza a medição dos montantes efetivamente produzidos/consumidos por cada agente. As diferenças apuradas, positivas ou negativas, são contabilizadas para posterior liquidação financeira no Mercado de Curto Prazo e valoradas ao Preço de Liquidação das Diferenças (PLD) [5].

O Sistema Interligado Nacional (SIN) é o sistema de produção e transmissão de energia elétrica do Brasil. Ele é constituído por quatro submercados: Sul (S), Sudeste/Centro-Oeste (SE/CO), Nordeste (NE) e a maior parte da região Norte (N), conforme mostrado na Figura 2.1. A capacidade instalada de geração do SIN é composta, principalmente, por usinas hidrelétricas. A interconexão dos sistemas elétricos permite a transferência de energia entre os submercados. A integração dos recursos de geração e transmissão permite o atendimento ao mercado com segurança e economicidade [12].



Figura 2.1 – Regiões do Sistema Interligado Nacional  
Fonte: CCEE [5]

O Preço de Liquidação das Diferenças (PLD) é calculado individualmente para cada um dos quatro submercados e em três patamares de carga: leve, médio e pesado. Esses patamares de carga representam horários de baixo, médio e alto consumo, respectivamente. A CCEE realiza mensalmente o cálculo da média mensal do Preço de Liquidação das Diferenças por submercado. O cálculo considera os preços semanais por patamar de carga, ponderado pelo número de horas em cada patamar e em cada semana do mês.

Conforme os autores de [13] a energia elétrica no Brasil possui características singulares, as quais pode-se citar:

- Não é possível armazená-la ou estocá-la no atacado para futuro uso;
- Incerteza da demanda;
- Dependente da hidrologia;
- Expansão da geração exige longo prazo.

Características estas que contribuem para variações do preço no mercado de curto prazo de energia. A seguir, as Figuras 2.2, 2.3, 2.4 e 2.5 representam a média mensal do PLD para os submercados SE/CO, S, NE e N respectivamente, no período de maio/13 a maio/17, onde é possível verificar a elevada variação citada. Nota-se, por inspeção visual, que, para todas as regiões do país, as características de variação do PLD são qualitativamente equivalentes.

### Evolução média mensal do PLD - SE/CO

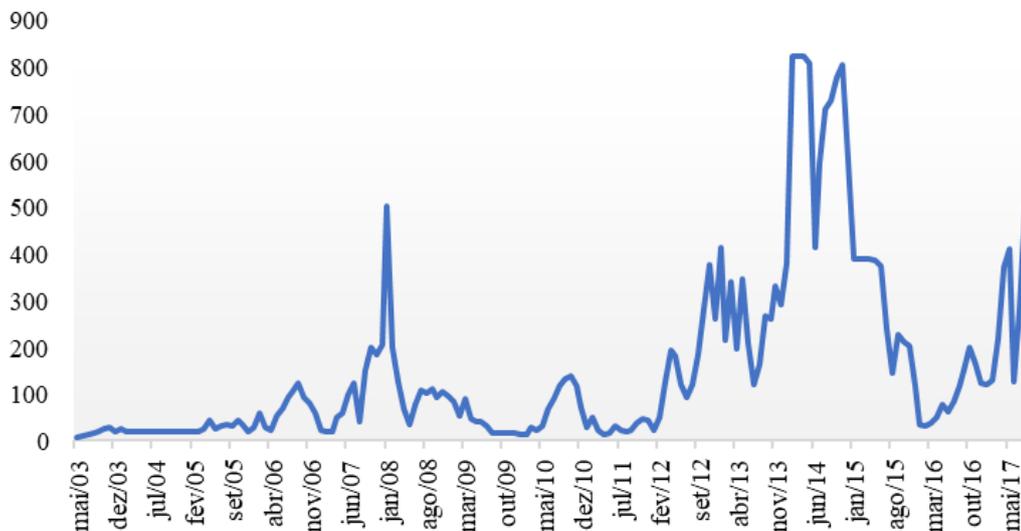


Figura 2.2 – Média mensal do PLD – submercado SE/CO

Devido a predominância hidráulica do parque gerador brasileiro, o PLD tem como objetivo encontrar a solução ótima entre a utilização da energia provida das hidrelétricas, ou o armazenamento de água para utilização futura. Sendo assim, toda decisão tomada na etapa de planejamento em sistemas com base hidráulica agrega consequências futuras, pois a afluência futura é incerta e existe uma limitação da geração causada pela capacidade de armazenamento dos reservatórios.

### Evolução da média mensal do PLD - S

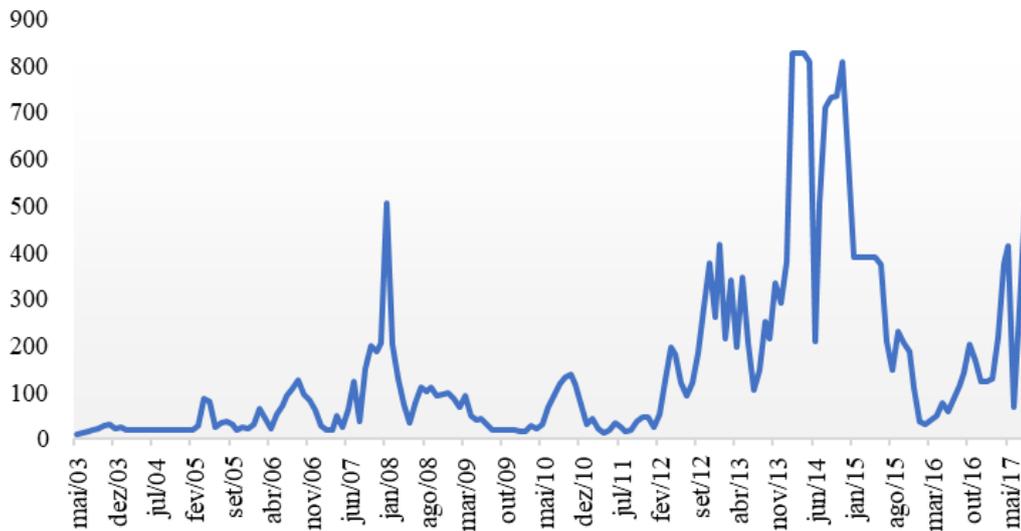


Figura 2.3 – Média mensal do PLD – submercado S

### Evolução da média mensal do PLD - NE

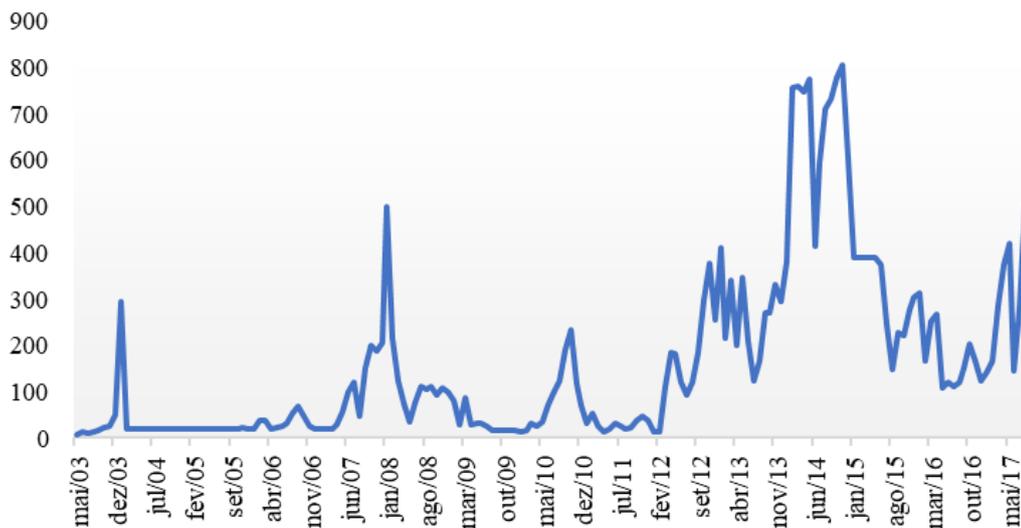


Figura 2.4 – Média mensal do PLD – submercado NE

A decisão de utilizar as usinas hidrelétricas prioritariamente implica em menores custos, uma vez que se trata de uma fonte barata. Estas podem ser caracterizadas como usinas a fio d'água ou usinas com reservatório, onde a primeira não possui a capacidade de armazenar água para uma geração futura, ou seja, o que não for gerado será vertido. Nas usinas com reservatório, ao optar pela utilização da água armazenada, o sistema fica suscetível às características hidrológicas que, caso seja seguido por um período de baixas afluências, será necessário recorrer as usinas termelétricas, encarecendo o preço da energia.

## Evolução da média mensal do PLD - N

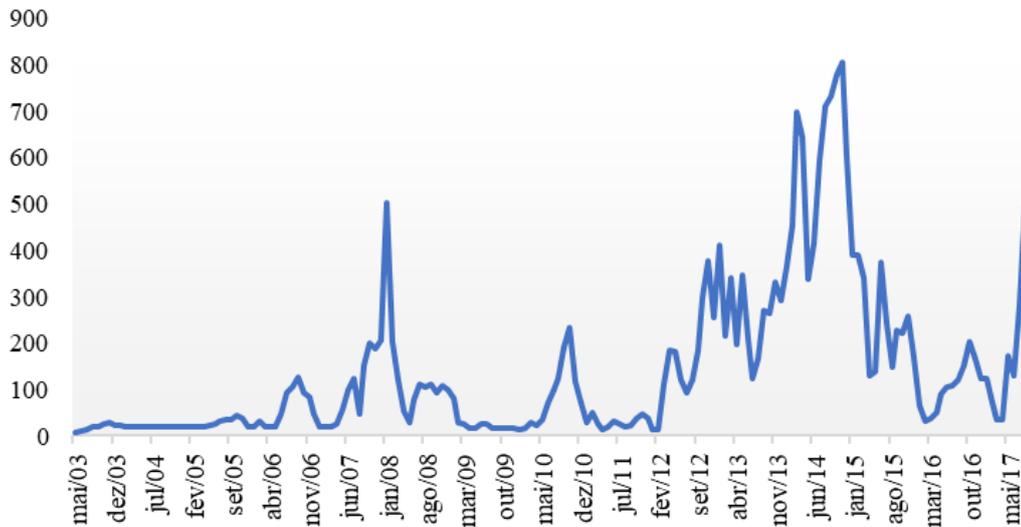


Figura 2.5 – Média mensal do PLD – submercado N

Por outro lado, se os reservatórios são pouco utilizados no presente e as previsões climáticas são favoráveis, ou seja, um elevado nível de afluições vir a ocorrer tem-se vertimento, isto é, abertura das comportas do vertedouro para liberação de água, indicando desperdício de energia. Em contrapartida, se a previsão de chuvas é desfavorável, têm-se operações econômicas. A Figura 2.6 mostra este processo de decisão e planejamento.

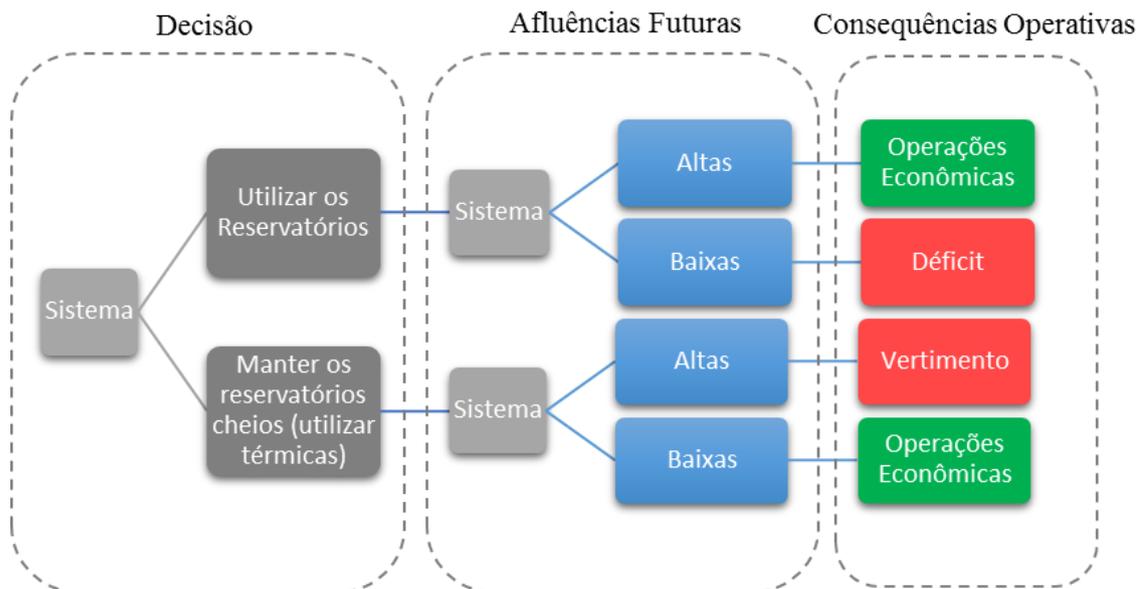


Figura 2.6 – Cronograma de decisão para um sistema hidrotérmico

O Preço de Liquidação das Diferenças é uma variável relevante para o Mercado Livre de Energia a Curto Prazo. Acompanhá-lo e entender seu funcionamento se torna essencial para uma contratação bem-sucedida: dos montantes adequados, no momento certo e com menores preços. Sendo assim, faz-se necessária a utilização de ferramentas adequadas que possam auxiliar no processo de decisão da compra de energia no Mercado Livre.

### 2.3 PREDIÇÃO DO PLD

A predição do Preço de Liquidação das Diferenças é um ponto importante para os consumidores que contratam seu fornecimento, no todo ou em parte, no Mercado Livre de Energia. Conhecer os valores futuros do PLD auxiliam na decisão das estratégias mais adequadas, estabelecendo assim, contratos que visam maximizar os lucros e minimizar os riscos. Diversas pesquisas têm sido desenvolvidas visando a predição de preços de energia no mercado de curto prazo. Segundo [14], os modelos de predição dos preços de eletricidade podem ser classificados principalmente em três grupos: teoria dos jogos, modelos fundamentais e modelos de séries temporais.

A teoria dos jogos, primeiro grupo, consiste em analisar o comportamento estratégico dos agentes e seu impacto nos preços de eletricidade, considerando-se como chave os modelos de equilíbrio do mercado de eletricidade. Os autores de [15] apresentaram em seu trabalho um método que permite obter os preços de liquidação do mercado chileno, utilizando a teoria dos jogos. Já os autores de [16] desenvolveram um modelo computacional baseado na Teoria dos Jogos para auxiliar os agentes do setor elétrico brasileiro nos leilões de energia elétrica de contratos bilaterais que ocorrem no Brasil.

O segundo grupo é fundamentado em modelos que simulam sistemas de potência capazes de satisfazer as demandas a um custo mínimo. Nesta abordagem, os preços da energia são modelados com base nos custos de geração, considerando congestionamento de transmissão, perdas, e outras solicitações de serviços auxiliares em mercado de energia [14].

Por fim, tem-se o terceiro grupo que é baseado em modelos de séries temporais. São modelos focados principalmente na análise de dados, dos quais vale citar: modelos autorregressivos (AR, ARMAX, ARIMA), Redes Neurais Artificiais, Sistemas *Fuzzy*, dentre outros. As Redes Neurais Artificiais são uma alternativa com grande capacidade de aprendizado, com possibilidades de aprender as dinâmicas e nuances dos problemas, é generalista e, depois de bem calibrada, possui complexidade computacional bastante reduzida, sendo então uma excelente escolha para o problema.

Reston Filho [14] desenvolveu em seu trabalho uma coletânea de técnicas para o problema de predição de dados PLD no mercado brasileiro de energia elétrica, dentre as quais constam modelos autorregressivos, modelos de médias móveis e Redes Neurais Artificiais. O autor desenvolveu uma técnica híbrida, mesclando um modelo autorregressivo integrado de médias móveis com redes neurais artificiais, com capacidade de predição em até doze passos (semanas) à frente, aproximadamente três meses. Os resultados são bons para os doze passos à frente, mostrando resultados de validação coerentes com os preços futuros reais.

## 2.4 REDES NEURAIAS ARTIFICIAIS

Uma Rede Neural Artificial é um modelo computacional inspirado no sistema nervoso de seres vivos. Redes Neurais são atrativas devido à reconhecida capacidade de aprender através de exemplos e de generalizar o conhecimento aprendido, fornecendo respostas adequadas para dados que não estão no conjunto de exemplos [17].

Os primeiros estudos envolvendo Redes Neurais Artificiais (RNAs) surgiram no ano de 1943, em um artigo publicado por McCulloch e Pitts [18]. Os autores foram os primeiros a descrever um modelo artificial para um neurônio biológico. Em 1949, o psicólogo Donald Hebb demonstrou o que ocorre durante o processo de aprendizado no cérebro, a nível celular [19]. Esta teoria deu origem a “Regra de Hebb”, que é a base do aprendizado das Redes Neurais Artificiais, sendo utilizada em vários algoritmos de treinamento [17].

Desde então, diversos pesquisadores voltaram-se para esta linha de pesquisa. Na década de 50, uma topologia de rede foi proposta por Frank Rosenblatt, denominada Perceptron [20]. Bernard Widrow foi responsável pelo desenvolvimento de um modelo novo de processamento de redes neurais, evidenciado por sua poderosa lei de aprendizado, a qual ficou conhecida como Adaline (acrônimo do inglês, *ADaptive Linear NEuron*) [21]. Minsky e Papert [22], publicaram em 1969 um livro no qual apresentaram limitações do uso das redes neurais citadas, como o problema “ou-exclusivo”, que revelou a falta de capacidade dos neurônios artificiais na solução de problemas que não fossem linearmente separáveis.

Com isso, gerou-se um retraimento na evolução das pesquisas sobre as RNAs e somente a partir de 1986, com a proposta de treinamento de redes alimentadas adiante (*feedforward*) com múltiplas camadas, por Rumelhart e McClelland [23], foi despertado o interesse em estudos sobre RNAs novamente. Deste ponto, até os dias atuais as Redes Neurais Artificiais vêm sendo aplicadas em diversas áreas do conhecimento humano.

### 2.4.1 Neurônio Biológico

Conforme descrito pelos autores de [24], as RNAs foram desenvolvidas a partir de modelos conhecidos de sistemas biológicos e do cérebro humano. A célula elementar do sistema nervoso cerebral é o neurônio, que conduz estímulos elétricos advindos de reações físico-químicas sob determinadas condições de operação [25]. Sendo assim, este serve de modelo computacional. O neurônio biológico pode ser dividido em três partes principais: os dendritos, o corpo celular e o axônio, como está representado na Figura 2.7.

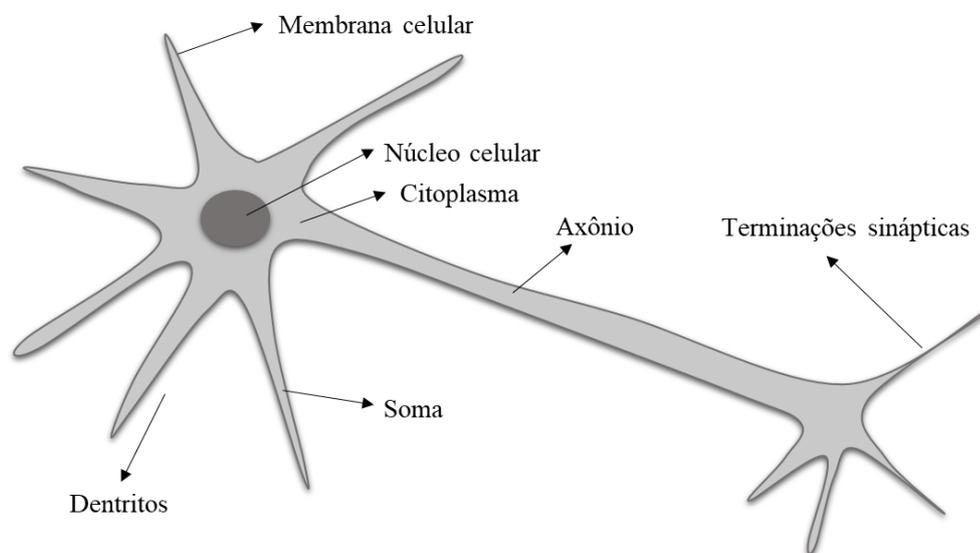


Figura 2.7 – Neurônio biológico

Os dendritos formam uma região onde os sinais de entrada são recebidos, sendo, portanto, os receptores dos sinais de entradas. O corpo celular é responsável por coletar e combinar informações vindas de outros neurônios, sendo o fator agregador de informações advindas das entradas. Os axônios são as vias através das quais o sinal é transportado até a outra extremidade do neurônio [26], sendo os transmissores das saídas do neurônio. Em relação à organização anatômica do cérebro, as sinapses são fundamentais, pois viabilizam a transferência de impulsos elétricos do axônio de um neurônio para os dendritos de outros, sendo a via de comunicação entre os neurônios.

Apesar de se mostrarem bastante simples, os neurônios, atuando em conjunto são responsáveis por todos os processamentos executados e gerenciados pelo cérebro humano, sendo uma poderosa forma de aprendizado e processamento de informações. Estima-se que haja, aproximadamente, 10 bilhões de neurônios no córtex humano e 60 trilhões de sinapses ou conexões [27].

## 2.4.2 Modelagem do Neurônio Artificial

Assim como o neurônio biológico, o neurônio artificial funciona recebendo e transmitindo informações. O modelo mais simples e que integra as características principais de uma rede neural biológica, foi proposto por McCulloch e Pitts [18]. O modelo do neurônio artificial é mostrado na Figura 2.8. Os diversos sinais de entrada  $\{x_1, x_2, \dots, x_n\}$  são equivalentes aos impulsos elétricos captados pelos dendritos no neurônio biológico. Os pesos sinápticos  $\{\omega_{11}, \omega_{21}, \dots, \omega_{n1}\}$  são propriedades de grande importância, pois representam a relevância que determinada entrada possui.

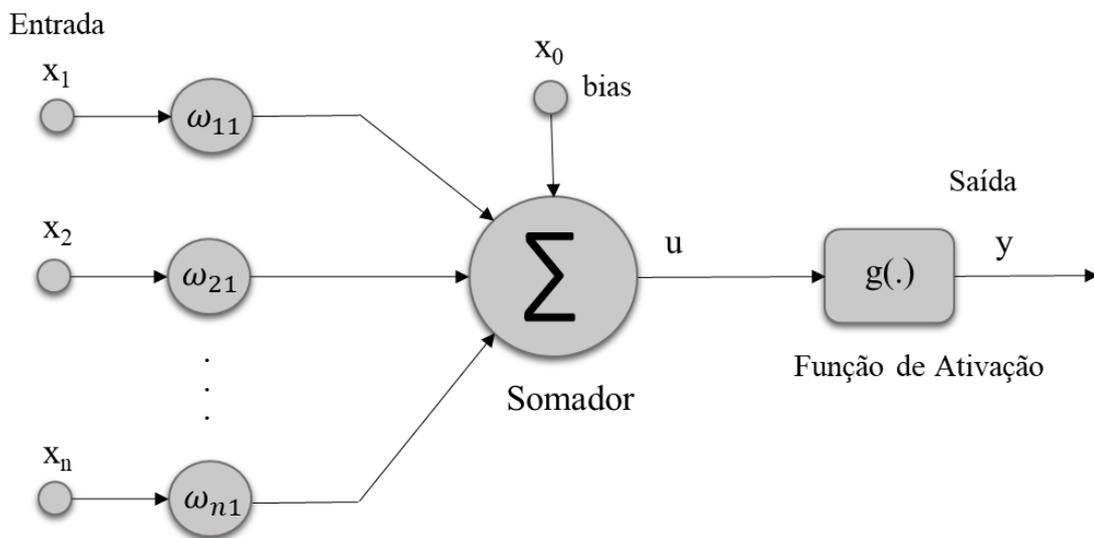


Figura 2.8 – Neurônio Artificial

Os pesos sinápticos têm seus valores alterados em função da relevância de cada uma das entradas  $\{x_i\}$  do neurônio, e este processo de alteração é denominado como processo de aprendizagem. Quanto maior o grau de importância de uma entrada, mais estimulado será o peso equivalente. Desta forma, mais significante este peso será para o resultado da saída  $\{y\}$  do neurônio. O sinal de ativação do neurônio  $\{u\}$  é decorrente do somatório do produto dos sinais de entrada pelos pesos sinápticos.

Por meio da função de ativação  $\{g(\cdot)\}$  é que o neurônio pode decidir se o resultado da soma ponderada das entradas será ativado ou não. A escolha adequada da função de ativação e de suas propriedades contribuem para melhorar o processo de convergência da rede durante seu treinamento. Como exemplos de função de ativação, pode-se citar funções degrau, sinal, rampa simétrica, logística e tangente hiperbólica. A função degrau está ilustrada na Figura 2.9 e é definida pela Equação 2.1.

$$g(u) = \begin{cases} 1, & \text{se } u \geq 0 \\ 0, & \text{se } u < 0 \end{cases} \quad (2.1)$$

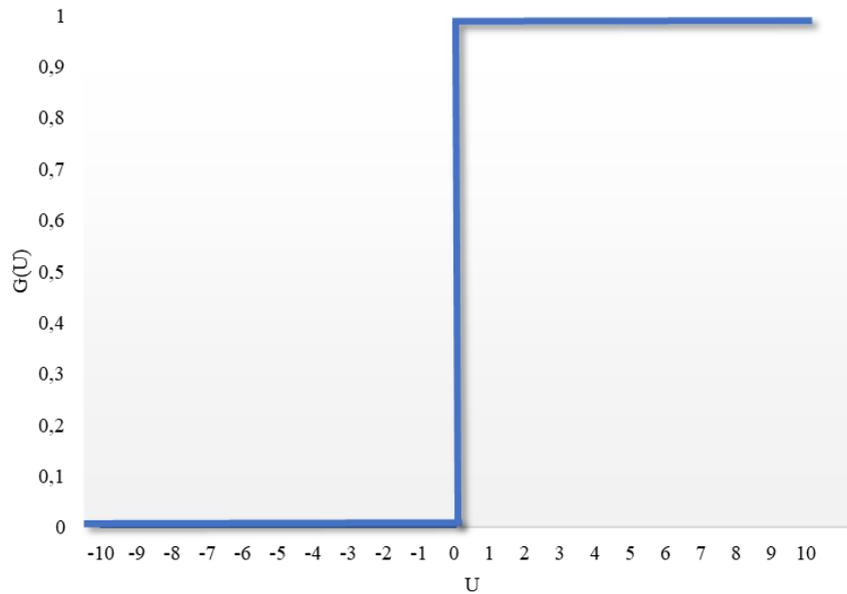


Figura 2.9 – Função de ativação degrau

O resultado produzido pela aplicação desta função assumirá valores unitários positivos quando o potencial de ativação do neurônio for maior ou igual a zero, caso contrário, o resultado assumirá valores nulos.

A função sinal está representada na Figura 2.10, Equação 2.2. O resultado produzido pela aplicação da função sinal assumirá valores positivos quando o potencial de ativação do neurônio for maior que zero; valor nulo quando o potencial for nulo; e valores unitários negativos quando o potencial for menor que zero, ou seja:

$$g(u) = \begin{cases} 1, & \text{se } u > 0 \\ 0, & \text{se } u = 0 \\ -1, & \text{se } u < 0 \end{cases} \quad (2.2)$$

Na função rampa simétrica, os valores retornados são iguais aos próprios valores dos potenciais de ativação quando estes estão definidos no intervalo  $[-a, a]$ , restringindo-se aos valores limites em caso contrário, conforme a Equação 2.3. A Figura 2.11 representa uma rampa simétrica para  $a = 2$ . Matematicamente, tem-se:

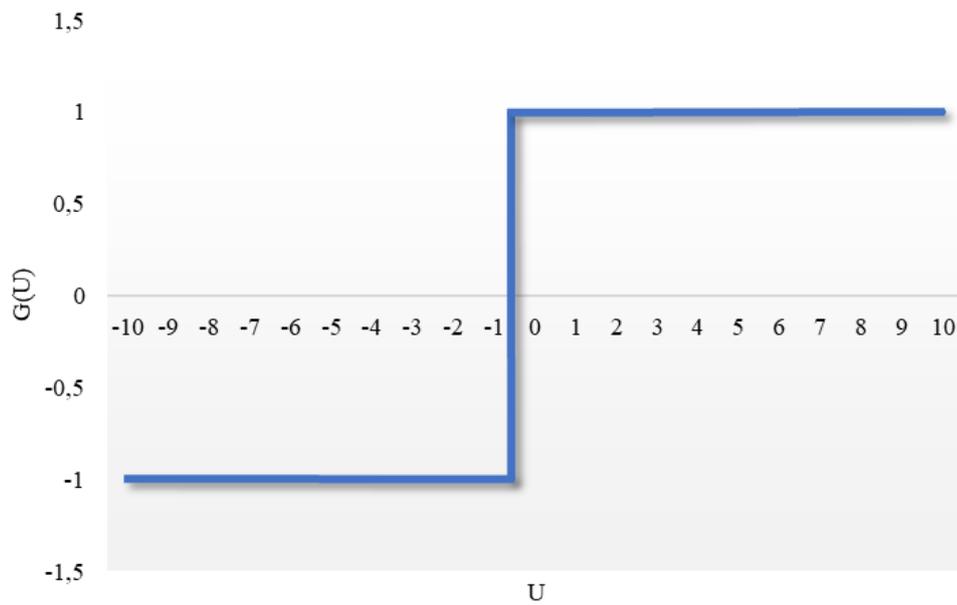


Figura 2.10 – Função de ativação sinal

$$g(u) = \begin{cases} a, & \text{se } u > a \\ u, & \text{se } -a \leq u \leq a \\ -a, & \text{se } u < -a \end{cases} \quad (2.3)$$

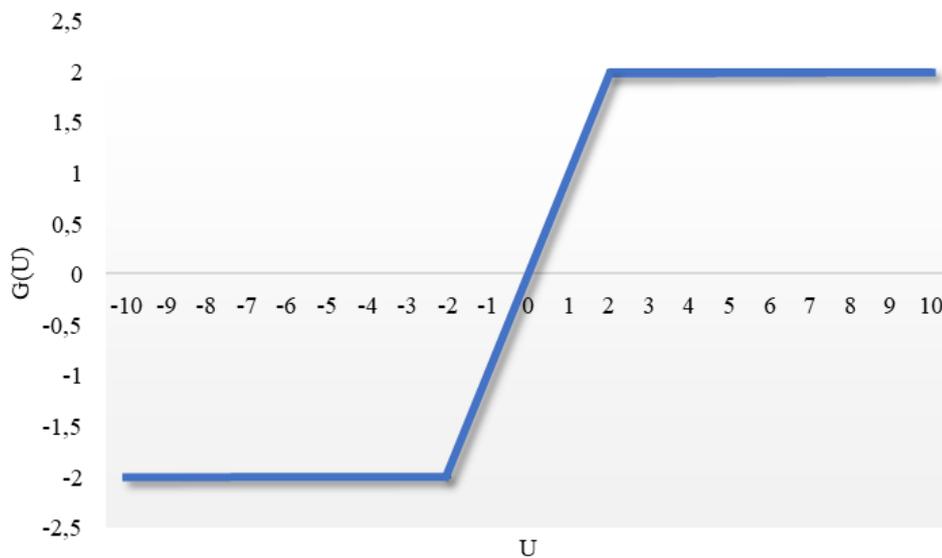


Figura 2.11 – Função de ativação rampa simétrica

A função logística, ilustrada na Figura 2.12, tem em sua saída valores reais entre zero e 1, tendo-se sua expressão matemática dada pela Equação 2.4. Tem-se que  $\beta$  é uma constante real que está associada ao nível de inclinação da função em relação ao seu ponto de inflexão.

$$g(u) = \frac{1}{1 + e^{-\beta u}} \quad (2.4)$$

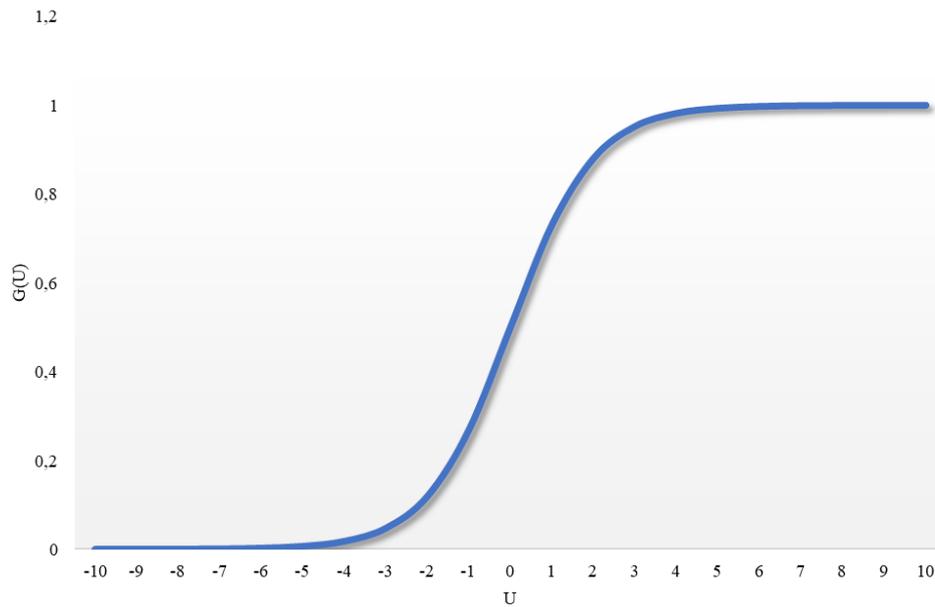


Figura 2.12 – Função de ativação logística

Diferentemente da função logística, a função tangente hiperbólica sempre assumirá valores reais entre -1 e 1. A representação gráfica desta função é ilustrada na Figura 2.13.

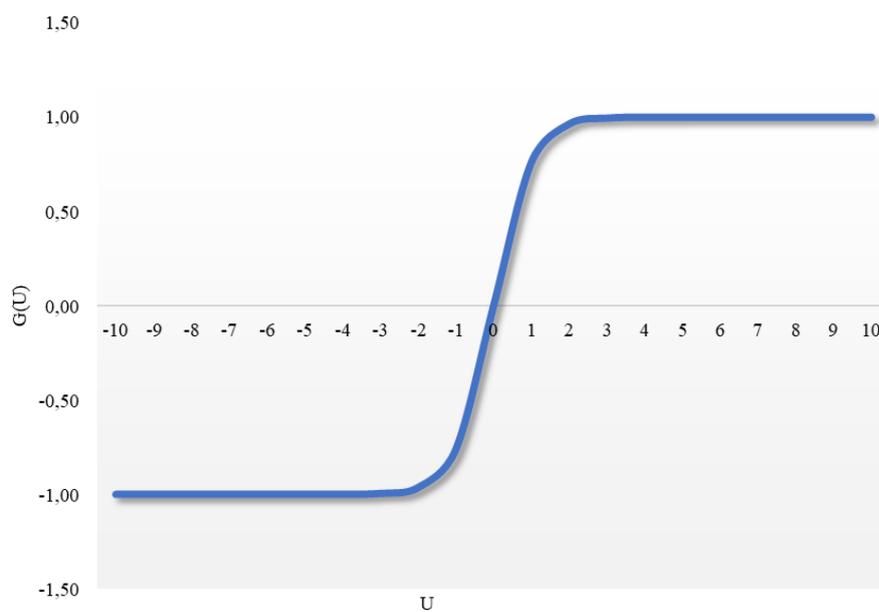


Figura 2.13 – Função de ativação tangente hiperbólica

A expressão matemática da tangente hiperbólica está representada na Equação 2.5, em que  $\beta$  está associado ao nível de inclinação da função em relação ao seu ponto de inflexão.

$$g(u) = \frac{1 - e^{-\beta u}}{1 + e^{-\beta u}} \quad (2.5)$$

Complementarmente, o bias  $\{-x_0\}$  permite aumentar os graus de liberdade fazendo com que a rede neural tenha uma melhor adaptação ao conhecimento fornecido à ela. Por fim, como definido por McCulloch e Pitts [18] o resultado produzido por um neurônio pode ser resumido por meio das Equações 2.6 e 2.7:

$$u = \omega_1 x_1 + \omega_2 x_2 + \dots + \omega_n x_n - \omega_0 x_0 = \sum_{i=1}^n \omega_i x_i - \omega_0 x_0 \quad (2.6)$$

$$y = g(u) \quad (2.7)$$

### 2.4.3 Arquiteturas de Redes Neurais Artificiais

A definição da arquitetura de uma Rede Neural Artificial é de grande importância, pois o seu arranjo está intimamente ligado ao tipo do problema a ser resolvido e ao algoritmo de aprendizado que será utilizado na fase de treinamento. As principais arquiteturas, considerando a disposição dos neurônios, as formas que os mesmos são interligados e ainda a constituição de suas camadas, são: redes *feedforward* de camada simples, de camadas múltiplas e redes recorrentes.

A arquitetura *feedforward* de camada simples possui apenas uma camada de entrada e uma única camada de neurônios, que se trata da camada de saída. Esta rede tem aplicabilidade em problemas que envolvem classificação de padrões e filtragem linear. O *Perceptron* e o *Adaline* são tipos de redes com essa arquitetura. Um exemplo desta arquitetura está representado na Figura 2.14.

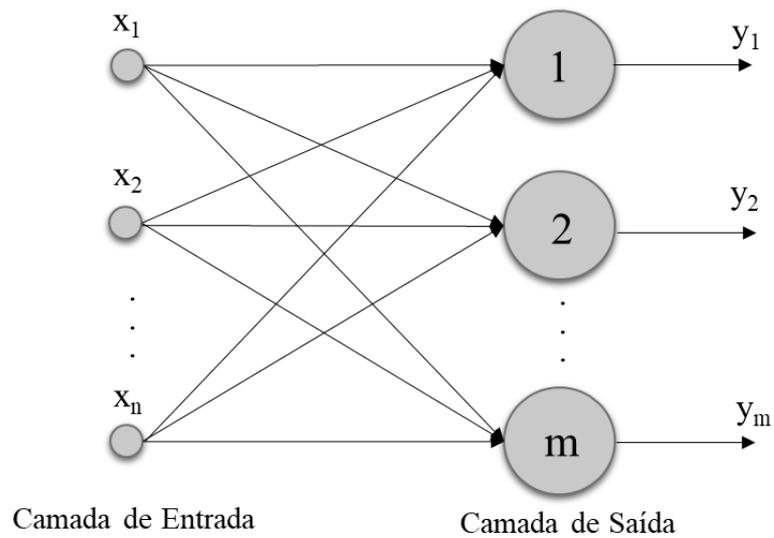


Figura 2.14 – Exemplo de rede *feedforward* de camada simples

Já as redes *feedforward* de camadas múltiplas, como mostra a Figura 2.15, são constituídas de uma ou mais camadas escondidas de neurônios. Devido ao conjunto extra de conexões sinápticas e da riqueza de interações neurais, as camadas ocultas são capazes de extrair características complexas dos dados sobre os quais atuam [27]. Estas redes são tipicamente empregadas em problemas relacionados à aproximação de funções, classificação de padrões, otimização, robótica, dentre muitos outros.

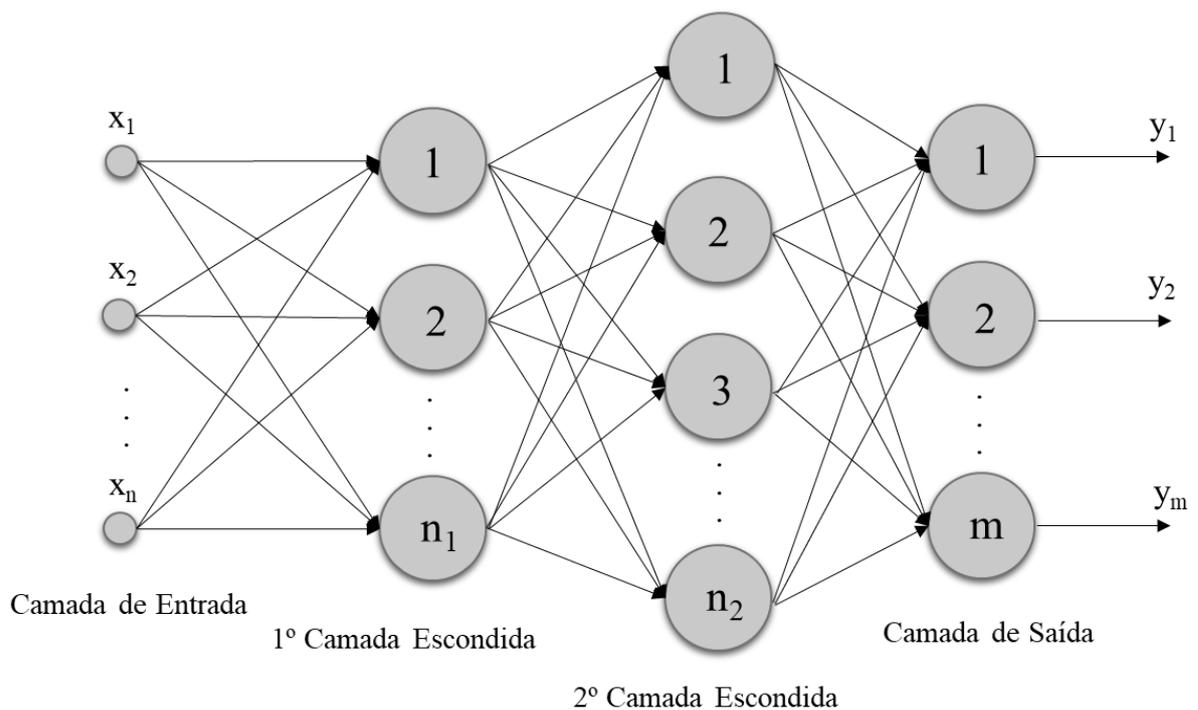


Figura 2.15 – Exemplo de rede *feedforward* de camada múltiplas

As redes recorrentes possuem as saídas dos neurônios realimentadas como sinais de entrada para outros neurônios. A presença da realimentação torna este tipo de arquitetura adequado para processamento dinâmico de informações, isto é, em sistemas variantes no tempo, como a previsão de séries temporais, que é o caso do presente trabalho. Um exemplo deste tipo de arquitetura está representado na Figura 2.16.

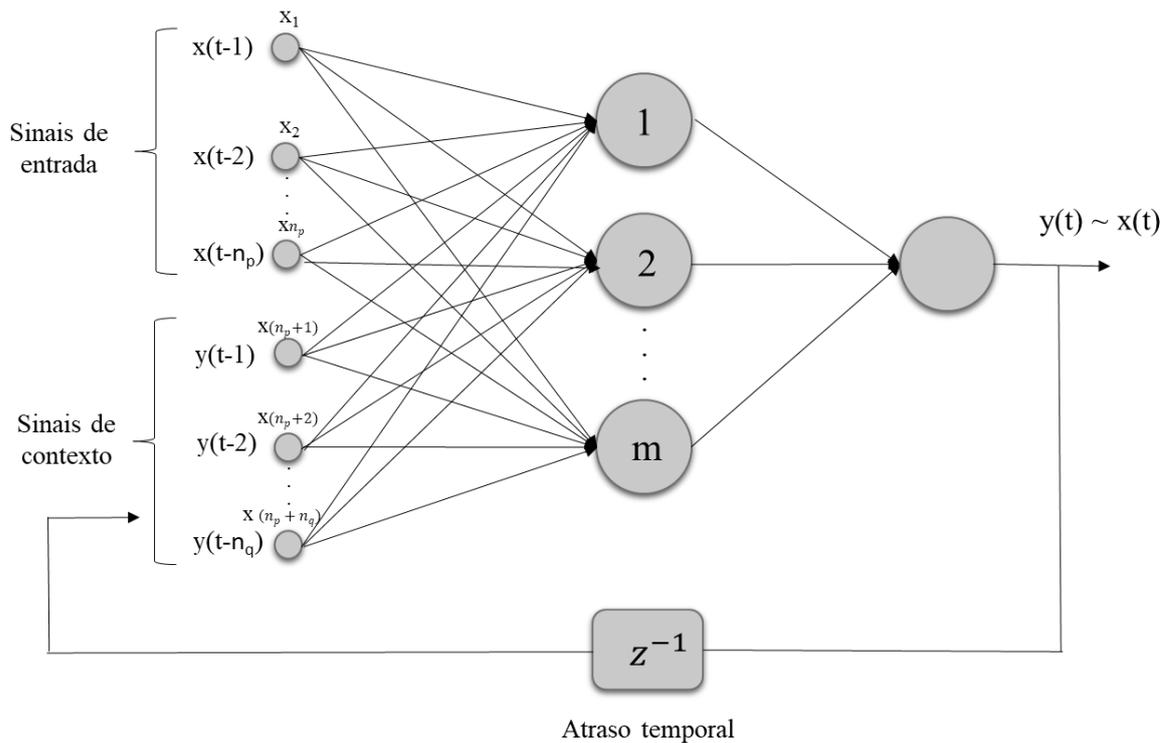


Figura 2.16 – Exemplo de rede recorrente

Vale ressaltar ainda que a topologia de uma Rede Neural, considerando determinada arquitetura, pode ser definida como sendo as diferentes formas de composições estruturais que esta poderá assumir [25], variando o número de neurônios, o número de camadas e a função de ativação, sendo que todos esses parâmetros influenciarão no resultado alcançado pela rede.

Em contrapartida, o treinamento de uma determinada arquitetura consiste na aplicação de um conjunto de passos com o objetivo de fazer o ajuste dos pesos sinápticos dos neurônios, tendo em vista sintonizar a rede para que as respostas produzidas por ela estejam próximas dos valores desejados. Este treinamento pode se tornar mais complexo com o crescimento da rede, e diferentes métodos de treinamento deverão ser empregados para acelerar a convergência e desprender-se de possíveis mínimos locais que venham a existir.

### 3. METODOLOGIA

Neste capítulo serão apresentados os detalhes da RNA implementada para a predição do Preço de Liquidação das Diferenças. Por se tratarem de dados dinâmicos, a Rede Neural Artificial tipo Elman, uma rede que realimenta as próprias saídas passadas como entradas, é melhor adaptada para lidar com o problema. Nesta topologia, os resultados produzidos pelos neurônios de saída da rede são realimentados às suas entradas por meio de atrasos temporais. Como apresentado anteriormente, tal rede pertence à classe de arquiteturas recorrentes.

A base de dados utilizada no presente trabalho foi obtida na Câmara de Comercialização de Energia Elétrica. Os dados abrangem o período de 21/03/2015 à 08/09/2017, totalizando 132 semanas. Os dados que estão contidos na base se tratam do PLD médio referente ao submercado Sudeste/Centro-Oeste. Escolheu-se a região Sudeste/Centro-Oeste por conta da localização da Universidade Federal de Juiz de Fora, no entanto, a mesma arquitetura de rede poderia ter sido utilizada para quaisquer outras regiões.

A primeira etapa se trata da normalização dos dados, que implica em escalonar as amostras para a faixa de variação dinâmica da função de ativação. Esta etapa é importante, pois mantém os valores das entradas proporcionais aos limites da função de ativação utilizada na rede, evitando que a saída do neurônio seja saturada nos valores limites da função, quando os potenciais de ativação possuem valores elevados.

No caso deste trabalho, utilizou-se a função de ativação tangente hiperbólica, com uma inclinação obtida de maneira empírica no valor de 0,26. Por se tratar de uma função ímpar, a tangente hiperbólica contribui para melhorar o processo de convergência da rede durante o treinamento [25]. De forma a evitar as regiões de saturação da função tangente hiperbólica, optou-se por normalizar os dados no intervalo de 0 a 0,5. A Equação 3.1 demonstra como é feita a normalização dos dados, por meio de uma relação linear:

$$s_t = s_{max} - \frac{(s_{max} - s_{min})(f_{max} - f_t)}{f_{max} - f_{min}} \quad (3.1)$$

Na Equação 3.1,  $t$  é o tempo;  $f$  é a função de treinamento, no caso, os valores de PLD disponibilizados pela CCEE de 21/03/2015 à 08/09/2017, os quais se deseja aprender com a rede, sendo então uma série temporal;  $s$  é a função de treinamento normalizada;  $s_{max}$  é o valor máximo permitido para a normalização e  $s_{min}$  é o respectivo valor mínimo permitido, e então  $s$  está contido no intervalo  $[s_{min}, s_{max}]$ ; de mesma maneira,  $f_{max}$  é o valor máximo permitido para a série temporal e  $f_{min}$  o respectivo valor mínimo, conforme Figura 3.1.

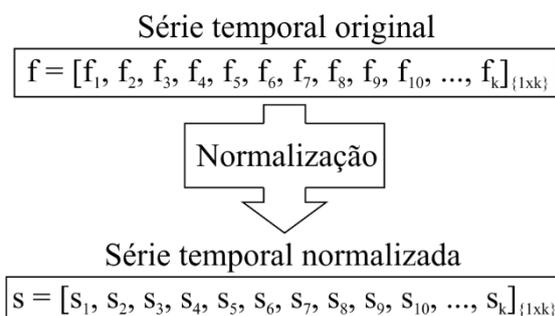


Figura 3.1 – Normalização dos dados a serem aprendidos

Para o treinamento, dividiu-se o conjunto de dados normalizados  $s$ , sendo os 108 primeiros para treinamento (81,82%) e os 24 últimos para validação (18,18%), uma vez que se deseja prever até 24 semanas no futuro com a rede neural, aproximadamente seis meses. O conjunto de treinamento é responsável pelo aprendizado da rede e ajuste dos pesos sinápticos. O segundo conjunto é utilizado para avaliar o desempenho de generalização da rede.

O conjunto de treinamento, composto pelas entradas e saídas desejadas é obtido a partir do conjunto de dados normalizados. Considerando a topologia representada na Figura 2.16, página 33, no primeiro instante a rede recorrente utilizará os sinais de entrada atrasadas no tempo em conjunto com as saídas realimentadas iguais a zero, uma vez que as mesmas inexistem no momento. No instante posterior, o primeiro valor produzido pela saída da rede estará disponível e pode realimentar uma entrada da rede, enquanto as outras ainda são zero. Sucessivamente, os próximos instantes das saídas estarão disponíveis e serão introduzidas nas entradas. Este processo é repetido para todos os instantes posteriores de forma a construir o conjunto de treinamento.

O treinamento de uma rede neural, normalmente segue o seguinte encadeamento: calculam-se as saídas da rede por meio dos pesos sinápticos da mesma em uma fase denominada *forward*, as quais são comparadas com os resultados desejados por meio de uma equação de erro quadrático. Caso a amplitude do erro seja alta, significa que as saídas da rede são muito diferentes dos resultados desejados, e então se faz necessário ajustar os pesos sinápticos em uma fase denominada *backward*, pois os pesos são ajustados de trás para frente. O algoritmo mais comum de treinamento é o *Backpropagation* [17], [25] e [27].

O método de treinamento utilizado no presente trabalho é o *Resilient Backpropagation* [17], [25] e [27]. Este método é uma variação computacionalmente mais eficiente do clássico *Backpropagation*. Nele, iniciam-se os primeiros vetores de pesos sinápticos com valores aleatórios e assim é possível calcular as primeiras saídas da na fase *forward*. A saída gerada é

comparada com a saída desejada e o erro obtido é comparado com um erro quadrático médio pré-estabelecido, o qual define se a rede convergiu ou não.

Se, por ventura, o erro quadrático médio for menor que um limiar pré-estabelecido, tem-se que a resposta da rede é satisfatória e generaliza bem os dados que se deseja aprender, sendo o algoritmo encerrado. No entanto, caso a resposta seja negativa é preciso fazer o ajuste dos pesos sinápticos na fase *backward* e o processo se repete até que a condição de erro mínimo (limiar) seja alcançada. Neste trabalho será utilizada uma camada escondida para a rede neural pois, usualmente, uma camada escondida é suficiente para a aproximação universal de funções, sendo este o caso do trabalho.

Na fase *forward*, obtém-se a saída da rede por meio do equacionamento a seguir:

$$u_1 = \omega_1 x + \omega_{01} x_0 \quad (3.2)$$

$$y_1 = \tanh(\beta u_1) \quad (3.3)$$

Na Equação 3.2,  $\omega_1$  são os pesos sinápticos atuais;  $x$  é o conjunto de entradas, sendo elas os valores da função  $s$  com  $m$  valores atrasados no tempo, a média e o desvio padrão desses valores atuais de entrada  $s$ , e os valores atrasados das saídas da rede;  $\omega_{01}$  é o bias da rede;  $x_0$  é um vetor de valores iguais a 1; Na equação 3.3,  $\beta$  é a inclinação da função tangente hiperbólica, a saída  $y_1$  da camada de entrada. A saída  $y_1$  é utilizada como entrada da próxima camada e assim, obter a saída da mesma, como mostram as Equações 3.4 e 3.5. É preciso realizar estes cálculos até que se atinja a última camada da rede.

$$u_2 = \omega_2 y_1 + \omega_{02} x_0 \quad (3.4)$$

$$y_2 = \tanh(\beta u_2) \quad (3.5)$$

Os símbolos das Equações 3.4 e 3.5 são análogos aos das Equações 3.2 e 3.3 e, para o caso de uma rede neural artificial com uma única camada escondida, como é o caso do presente trabalho, a Equação 3.5 é a saída final da rede. A rede em questão possui apenas uma camada escondida em decorrência do Teorema da Aproximação Universal, derivado do Teorema de Andrey Kolmogorov (1957), também verificado por George Cybenko (1989) e Kurt Hornik (1991), que estabelece a capacidade de aproximação universal de funções para redes neurais artificiais com apenas uma camada escondida e número finito de neurônios [25] e [27].

O erro entre a saída da rede ( $y_2$ ) e o sinal desejado ( $d$ , aquele que se deseja aprender, e no caso do presente trabalho é a série temporal do PLD), é calculado a partir da diferença entre ambos. Como função objetivo desta comparação utiliza-se o Erro Quadrático Médio ( $EQM$ ), dado pela Equação 3.6, pois, por ser uma função quadrática possui um único ponto ótimo. A otimização dos pesos sinápticos e bias da rede ( $\omega_1, \omega_{01}, \omega_2, \omega_{02}$ ) que, em última análise, são os detentores do conhecimento dos dados, é feito por meio do *Resilient Backpropagation*, que é uma versão elaborada do método Gradiente Descendente [17], [25] e [27].

$$EQM = \sum \frac{1}{2} (d - y_2)^2 \quad (3.6)$$

O gradiente é um vetor que aponta na direção de máximo crescimento da função. Logo, é possível obter a região de mínimo crescimento através da direção oposta do gradiente e assim garantir que os pesos obtidos ajustem o erro em seu menor valor.

A fase *backward* é realizada de trás para frente, ou seja, o ajuste dos pesos se inicia na última camada da rede. Para ajustar os pesos da última camada, é necessário calcular o gradiente de  $EQM$  com relação à  $\omega_2$  e  $\omega_{02}$  que são os pesos relativos à última camada. As derivadas calculadas são mostradas a seguir:

$$\frac{\partial EQM}{\partial \omega_2} = \frac{\partial EQM}{\partial y_2} \frac{\partial y_2}{\partial u_2} \frac{\partial u_2}{\partial \omega_2} \quad (3.7)$$

$$\frac{\partial EQM}{\partial y_2} = -(d - y_2) \quad (3.8)$$

$$\frac{\partial y_2}{\partial u_2} = g'(u_2) \quad (3.9)$$

$$\frac{\partial u_2}{\partial \omega_2} = y_1 \quad (3.10)$$

Associa-se as Equações 3.7 e 3.8, de forma a evitar esforço computacional, uma vez que este cálculo será utilizado posteriormente.

$$\delta_2 = -(d - y_2) .* g'(u_2) \quad (3.11)$$

Em que  $.*$  denota a multiplicação matricial elemento a elemento. Por fim, a Equação 3.12 representa a derivada do erro  $EQM$  em relação a  $\omega_2$ . É importante atentar-se às dimensões das matrizes, de forma que as operações sejam possíveis.

$$\frac{\partial EQM}{\partial \omega_2} = \delta_2 y_1^T \quad (3.12)$$

Os mesmos cálculos são realizados em relação a  $\omega_{02}$ , obtendo-se o resultado representado pela Equação 3.13:

$$\frac{\partial EQM}{\partial \omega_{02}} = \delta_2 x_0^T \quad (3.13)$$

De posse destes dados, é possível fazer o ajuste dos pesos:

$$\omega_2^{\{t+1\}} = \omega_2^{\{t\}} - \eta_{\omega_2} \frac{\partial EQM_t}{\partial \omega_2} \quad (3.14)$$

$$\omega_{02}^{\{t+1\}} = \omega_{02}^{\{t\}} - \eta_{\omega_{02}} \frac{\partial EQM_t}{\partial \omega_{02}} \quad (3.15)$$

Em que  $\eta_{\omega_2}$  e  $\eta_{\omega_{02}}$  são as matrizes das taxas de aprendizado do algoritmo, de mesma dimensão dos pesos que ajustam. Estas constantes têm relação com a ponderação das mudanças realizadas nos pesos sinápticos, influenciando na trajetória percorrida pelo gradiente descendente no espaço de busca. A escolha correta desta taxa tem um efeito importante no treinamento da rede, uma vez que um valor baixo causa demora no aprendizado e um valor elevado pode levar à oscilação.

O algoritmo *Resilient Backpropagation* ajusta as taxas de aprendizado  $\eta_{\omega}$  a cada interação, conforme a inclinação dos gradientes. Quando a derivada parcial do peso  $\omega_2$  muda de sinal, a taxa é decrementada, pois esta mudança no sinal do gradiente significa que a última atualização do peso foi muito grande fazendo com que a busca pulasse uma região de mínimo

local ou então fosse transladada para outra região da função objetivo. A Equação 3.16 mostra como se dá a atualização dos pesos sinápticos.

$$\eta_{\omega}(i, j) = \begin{cases} \eta^+ \eta_{\omega}(i, j), & \text{se } \frac{\partial EQM_{k-1}}{\partial \omega} \frac{\partial EQM_k}{\partial \omega} > 0 \\ \eta^- \eta_{\omega}(i, j), & \text{se } \frac{\partial EQM_{k-1}}{\partial \omega} \frac{\partial EQM_k}{\partial \omega} < 0 \\ \eta_{\omega}(i, j) & , \text{se } \frac{\partial EQM_{k-1}}{\partial \omega} \frac{\partial EQM_k}{\partial \omega} = 0 \end{cases} \quad (3.16)$$

A mesma análise é válida para todos os pesos  $\omega$ . É esta abordagem que permite ao *Resilient Backpropagation* lidar melhor que outros métodos com possíveis mínimos locais durante o treinamento, fazendo com que a convergência seja bastante acelerada quando comparada com outros métodos baseados em *Backpropagation*.

Uma vez que os pesos sinápticos  $\omega_2$  e  $\omega_{02}$  forem atualizados, são realizados os cálculos para a atualização dos pesos da camada anterior:  $\omega_1$  e  $\omega_{01}$ . Todo o processo aplicado na camada anterior é repetido, iniciando com o cálculo da derivada parcial de *EQM* com relação a  $\omega_1$  e  $\omega_{01}$ , por meio da regra da cadeia aplicada anteriormente, seguido da atualização dos pesos. Este processo pode ser acompanhado nas Equações a seguir:

$$\frac{\partial EQM}{\partial \omega_1} = \frac{\partial EQM}{\partial y_2} \frac{\partial y_2}{\partial u_2} \frac{\partial u_2}{\partial y_1} \frac{\partial y_1}{\partial u_1} \frac{\partial u_1}{\partial \omega_1} \quad (3.17)$$

$$\frac{\partial u_2}{\partial y_1} = \omega_2 \quad (3.18)$$

$$\frac{\partial y_1}{\partial u_1} = g'(u_1) \quad (3.19)$$

$$\frac{\partial u_1}{\partial \omega_1} = x \quad (3.20)$$

Associa-se as Equações 3.11, 3.18 e 3.19 de forma a evitar esforço computacional, uma vez que este cálculo será utilizado posteriormente.

$$\delta_1 = \omega_2^T (\delta_2 .* g'(u_1)) \quad (3.21)$$

Por fim, tem-se na Equação 3.22 a derivada do erro  $EQM$  com relação a  $\omega_1$ .

$$\frac{\partial EQM}{\partial \omega_1} = \delta_1 x^T \quad (3.22)$$

Os mesmos cálculos são realizados em relação a  $\omega_{01}$ , obtendo-se o resultado representado pela Equação 3.23.

$$\frac{\partial EQM}{\partial \omega_{01}} = \delta_1 x_0^T \quad (3.23)$$

Por fim, é possível fazer o ajuste dos pesos:

$$\omega_1^{\{t+1\}} = \omega_1^{\{t\}} - \eta_{\omega_1} \frac{\partial EQM_t}{\partial \omega_1} \quad (3.24)$$

$$\omega_{01}^{\{t+1\}} = \omega_{01}^{\{t\}} - \eta_{\omega_{01}} \frac{\partial EQM_t}{\partial \omega_{01}} \quad (3.25)$$

Têm-se que  $\eta_{\omega_1}$  e  $\eta_{\omega_{01}}$  seguem os mesmos critérios apresentados para os pesos anteriores. O treinamento se dá conforme a Figura 3.2.

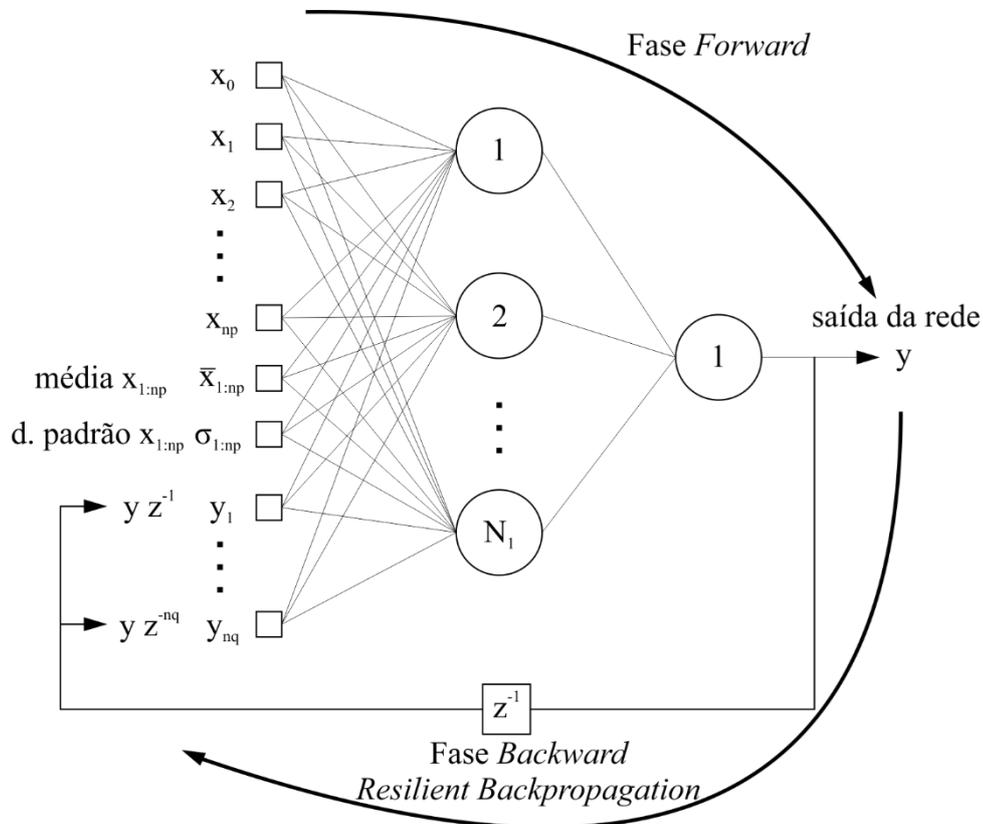


Figura 3.2 – Treinamento da Rede Neural Artificial recorrente

O critério de parada é estabelecido em função do erro quadrático médio, conforme calculado na Equação 3.6. O algoritmo converge quando este erro for suficientemente pequeno, respeitando uma precisão requerida para o processo de convergência. Em conjunto com o erro, pode-se utilizar a variável época como critério de parada. Esta variável contabiliza o número de épocas necessárias para a efetivação do treinamento da rede. Em situações onde a precisão se torna inatingível, o número de épocas limita o processo de treinamento.

Após o treinamento da rede, é possível realizar a predição dos valores futuros. A predição é efetuada utilizando a mesma configuração de entradas que foram assumidas no aprendizado. Dessa forma, considerando quaisquer instantes futuros, a predição de seus valores sempre levará em conta as últimas entradas atrasadas no tempo, bem como as últimas saídas produzidas pela rede.

Por trabalhar com os dados normalizados, os resultados obtidos também serão normalizados. Para encontrar o valor real previsto pela rede, é preciso executar a função inversa da normalização, representada na Equação 3.26.

$$f_{predito} = f_{max} - \frac{(f_{max} - f_{min})(s_{max} - s_{predito})}{s_{max} - s_{min}} \quad (3.26)$$

Em que  $f_{predito}$  é o valor desnormalizado predito para a série temporal a partir de  $s_{predito}$ , o qual é o valor normalizado obtido a partir da rede neural.

Como descrito anteriormente,  $s_{max}$  é o intervalo máximo da normalização (no presente trabalho é igual a 0,5),  $s_{min}$  é o intervalo mínimo da normalização (igual a 0),  $ft$  (função de treinamento) é o valor da entrada,  $f_{max}$  é o valor máximo das variáveis de entrada e  $f_{min}$  o valor mínimo das variáveis de entrada, considerando toda a série temporal. A Figura 3.3 mostra o processo descrito ao longo do capítulo, de treinamento e predição dos dados desejados.

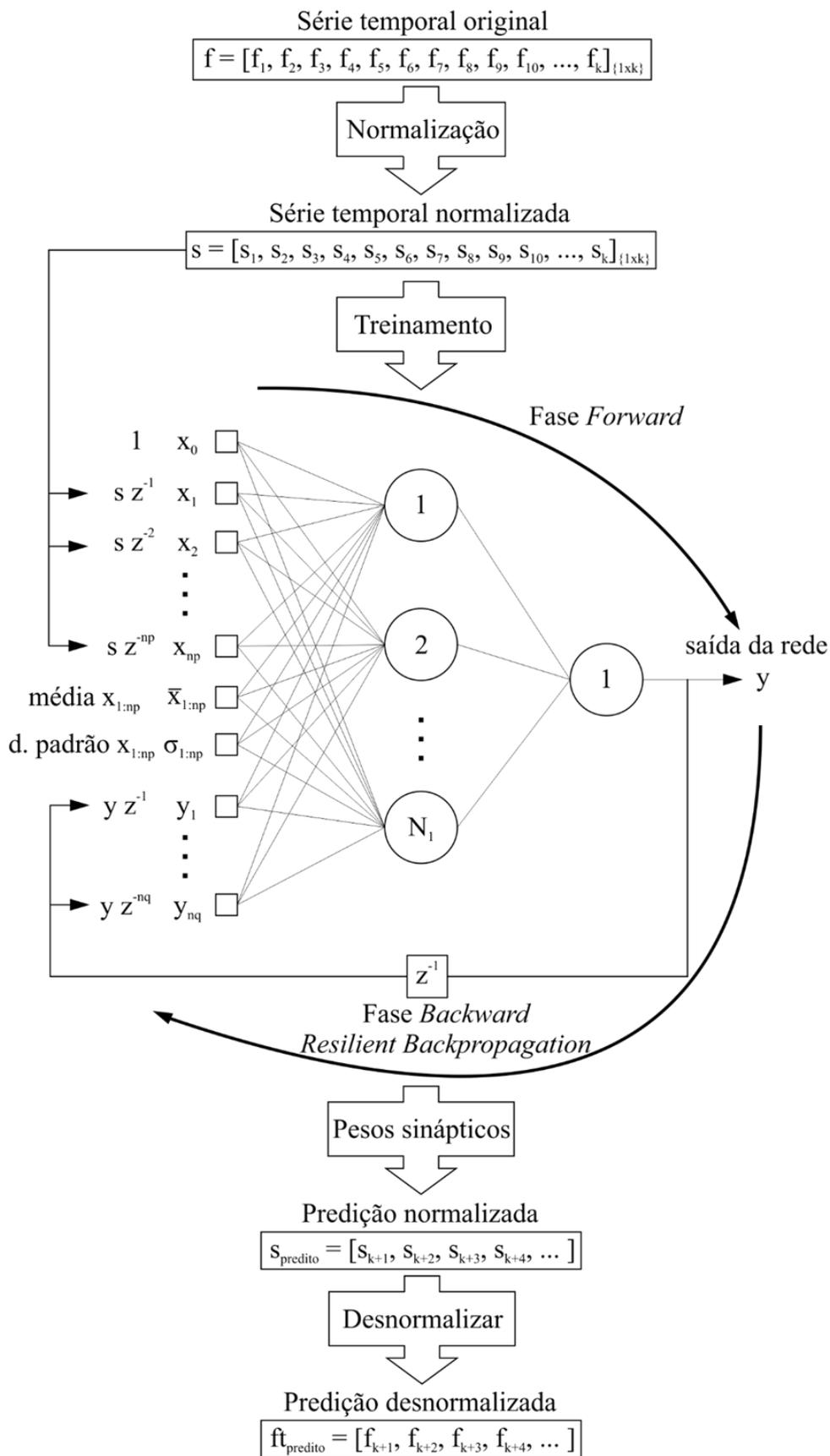


Figura 3.3 – Descrição do processo de treinamento e predição por meio da rede neural recorrente

## 4. RESULTADOS

Este capítulo apresenta os resultados obtidos através da aplicação da metodologia proposta no capítulo anterior nos dados reais de PLD obtidos da Câmara de Comercialização de Energia Elétrica (CCEE) e demonstra a eficácia do método utilizado. Serão aplicadas duas metodologias para explorar as capacidades da rede recorrente: predição de apenas um passo (semana) no futuro e; predição de vários passos no futuro, no caso, até vinte e quatro semanas, equivalente a aproximadamente seis meses.

### 4.1 DEFINIÇÃO DA TOPOLOGIA

Para avaliar a performance de uma RNA, utiliza-se um conjunto de dados que não está presente no processo de treinamento. Após treinar a rede, estes dados são comparados com os dados previstos pela rede, permitindo obter os erros de predição através da Equação 4.1, a raiz quadrada do erro quadrático médio (*Root Mean Squared Error*, RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n |P_i - Q_i|^2} \quad (4.1)$$

Em que  $n$  é a quantidade de amostras no conjunto,  $P_i$  e  $Q_i$  são o  $i$ -ésimo valor desejado e obtido pela rede, respectivamente. O *RMSE* é um parâmetro que permite a avaliação da configuração da rede neural, possibilitando a escolha da melhor configuração.

De acordo com a topologia representada na Figura 2.16, página 33, definiu-se  $n_p = 22$  de maneira empírica após diversos testes, o número de entradas da rede advindos da série temporal. Deste quantitativo, 20 são obtidos através da base de dados e 2 se tratam da média e do desvio padrão das entradas atuais. O valor  $n_q$ , que representa a quantidade de saídas passadas que serão utilizadas na predição foi variado em experimentos distintos, de forma a obter uma análise quantitativa para os melhores resultados. O número de neurônios na camada intermediária também sofreu variações. A Tabela 4.1 representa as topologias das RNAs utilizadas, objetivando a predição do PLD para 24 semanas.

Tabela 4.1 – Topologias das RNAs aplicadas à solução do problema

	Configuração 1			Configuração 2			Configuração 3		
Número de atrasos na entrada ( $n_p$ )	22	22	22	22	22	22	22	22	22
Número de atrasos na saída ( $n_q$ )	1	3	6	1	3	6	1	3	6
Neurônios na camada intermediária ( $N_1$ )	25	25	25	40	40	40	70	70	70
Neurônios na camada de saída ( $N_2$ )	1	1	1	1	1	1	1	1	1

Conforme citado previamente, para gerar os resultados serão consideradas duas abordagens. A primeira, a abordagem clássica, é a utilização da rede neural recorrente para a predição do futuro a um passo, ou seja, a rede preverá apenas o valor do PLD uma semana à frente. A segunda abordagem é a adaptação da rede recorrente para a predição de quantos passos se queiram no futuro, e, no caso, serão preditos até 24 passos (semanas) no futuro. A comparação de ambas as estratégias será objeto de discussão.

#### 4.2 TREINAMENTO E VALIDAÇÃO PARA UM PASSO NO FUTURO

A utilização clássica de redes recorrentes se dá para a predição de um passo no futuro. As redes neurais artificiais possuem muitos parâmetros que podem influenciar o treinamento e a convergência da rede, podendo causar, inclusive, a divergência da mesma. Assim, definir as topologias que melhor generalizam os dados é de fundamental importância para o sucesso da rede frente à aplicação.

Para os subconjuntos de treinamento e validação, calculou-se o RMSE a fim de escolher a melhor configuração de rede neural. Os gráficos das Figuras 4.1, 4.2 e 4.3 representam o comportamento do RMSE para cada rede com  $N_1$  sendo 25, 40 e 70 neurônios na camada intermediária e  $n_p$  sendo 1, 3 e 6, respectivamente.

Diversas outras configurações para a topologia da rede foram testadas. No entanto, notou-se que os três números de neurônios para as camadas intermediárias supracitados concentraram em si, em média, os comportamentos dominantes da rede com relação aos dados testados, e por isso foram escolhidos como objeto principal de estudo.

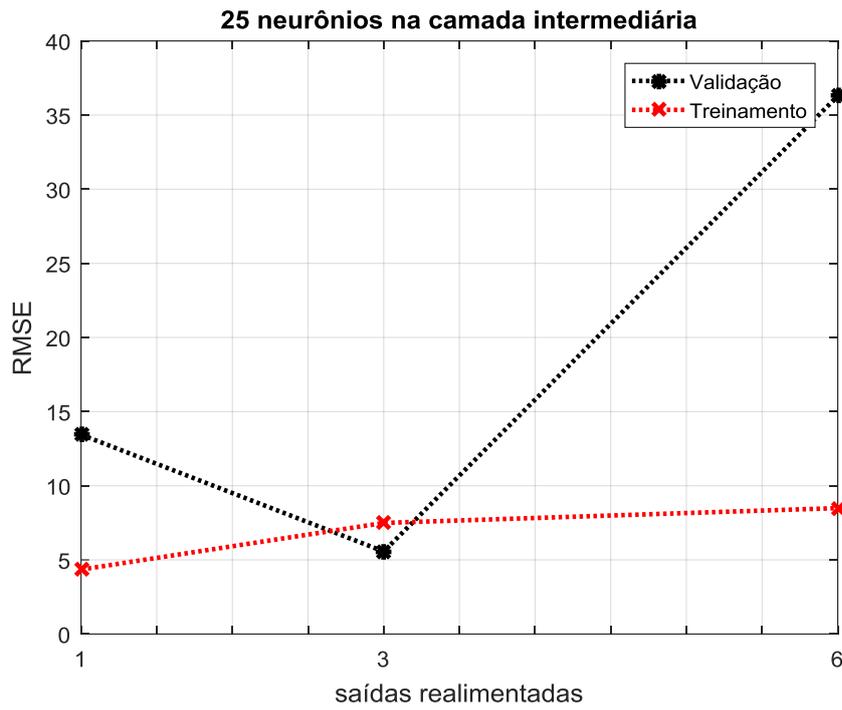


Figura 4.1 – Desempenho da rede (RMSE  $\times$   $n_p$ ) para  $N_1=25$  neurônios e 1 passo no futuro

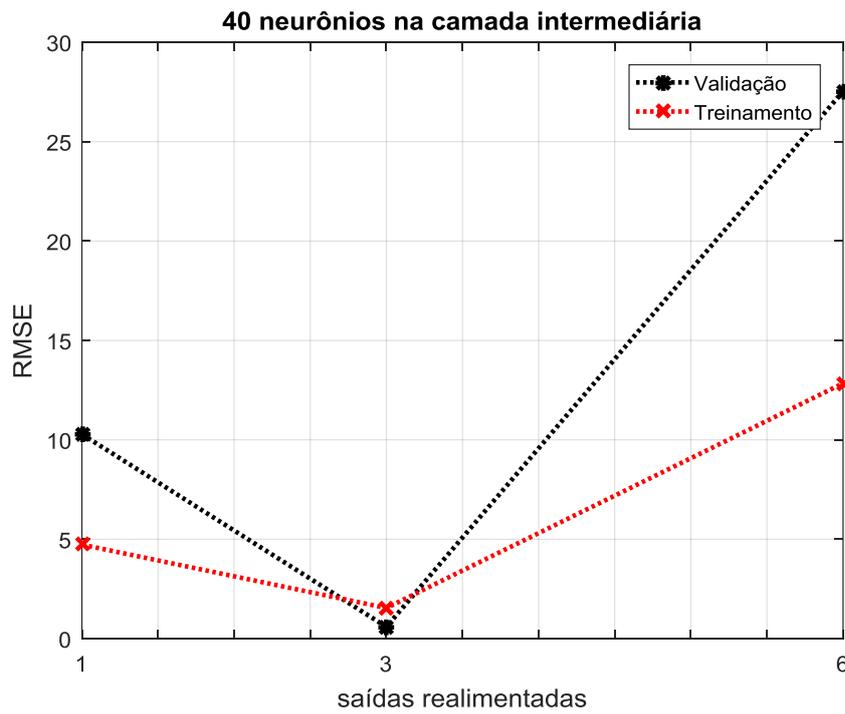


Figura 4.2 – Desempenho da rede (RMSE  $\times$   $n_p$ ) para  $N_1=40$  neurônios e 1 passo no futuro

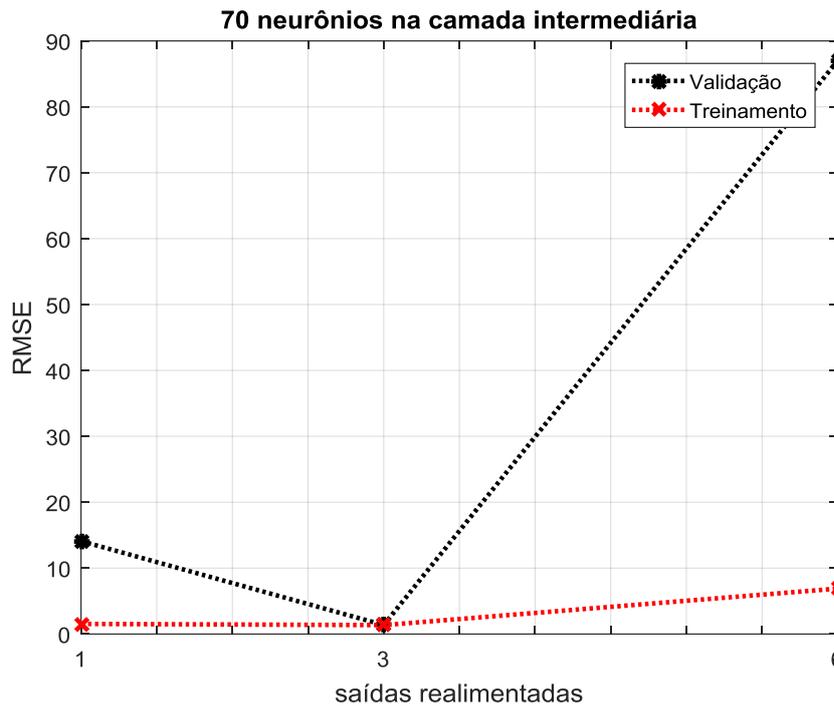


Figura 4.3 – Desempenho da rede (RMSE x  $n_p$ ) para  $N_1=70$  neurônios e 1 passo no futuro

Analisando os gráficos é possível perceber que ao utilizar 3 saídas realimentadas ( $n_q$ ), a rede apresenta menor RMSE, sendo este considerado o melhor desempenho em termos do erro produzido pela rede. Em outras palavras, o RMSE mostra o quão bem a rede neural generalizou a série temporal, pois a diferença entre os dados da série temporal e os dados produzidos pela rede produzem um RSME pequeno.

Tendo as configurações que produzem o menor erro (RSME), então, as mesmas são utilizadas para a predição dos dados. As Figuras 4.4 e 4.5 ilustram respectivamente, as fases de treinamento e predição para a topologia com melhor desempenho, com 40 neurônios na camada intermediária ( $N_1$ ) e 3 realimentações das saídas ( $n_q$ ). A rede com 70 neurônios na camada intermediária ( $N_1$ ) e 6 saídas realimentadas ( $n_q$ ) foi a que apresentou piores resultados, conforme pode ser visto nas Figuras 4.6 e 4.7. Assim, é possível uma comparação entre o melhor e o pior resultado para a predição de um passo no futuro.

Para interpretar os dados, é necessário, primeiramente, compreender que redes neurais subdimensionadas ou superdimensionadas em termos de números de neurônios nas camadas intermediárias tendem a perder a qualidade de aprendizado e também de predição. Este foi o motivo dos vários testes realizados na Tabela 4.1, a identificação de um número de dimensionamento satisfatório. A configuração 1, apresentada nesta tabela, foi o resultado que apresentou eficácia intermediária. Nota-se que mesmo o pior resultado de predição, com  $n_q=6$ ,

ainda assim foi capaz de acompanhar a tendência dos dados, mostrando que apesar do erro ter sido o menos satisfatório, a rede incorporou o conhecimento da variação dos dados.

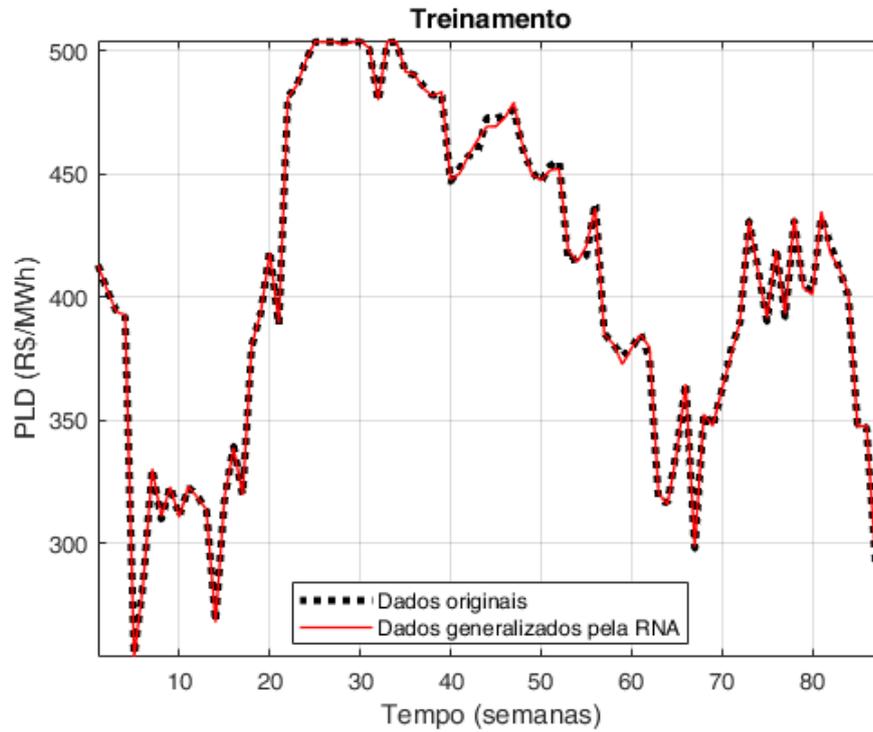


Figura 4.4 – Fase de treinamento para  $N_1=40$  neurônios intermediários e  $n_q=3$  saídas realimentadas

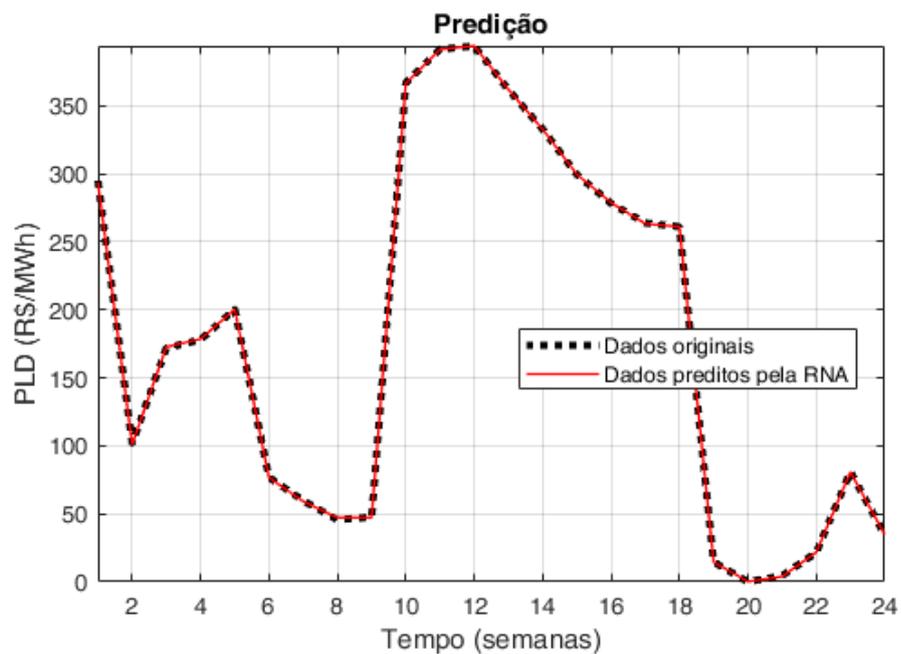


Figura 4.5 – Predição para  $N_1=40$  neurônios intermediários e  $n_q=3$  saídas realimentadas

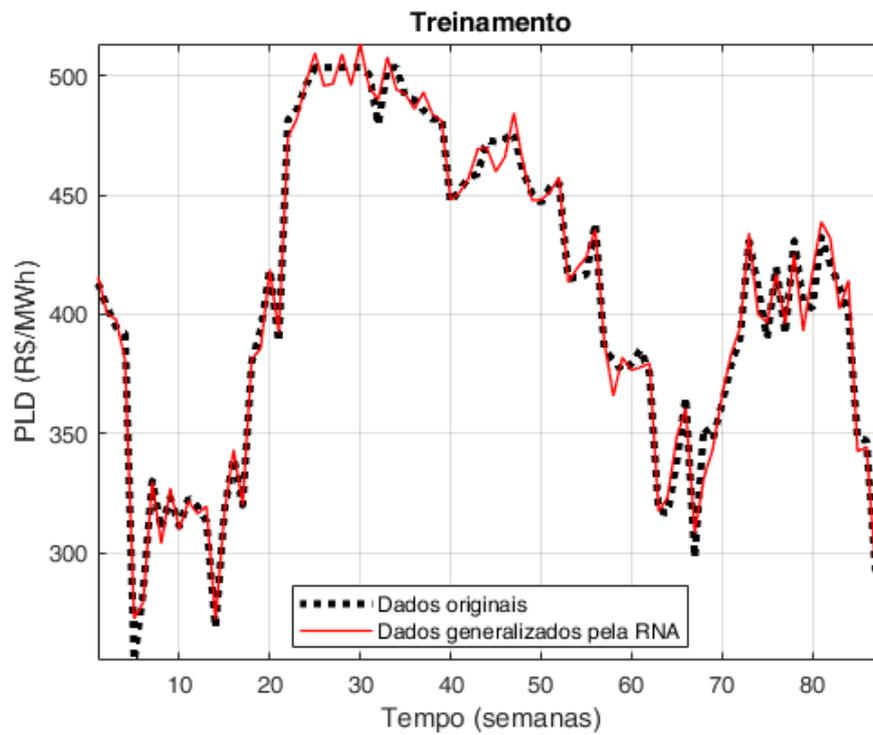


Figura 4.6 – Fase de treinamento para  $N_1=70$  neurônios intermediários e  $n_q=6$  saídas realimentadas

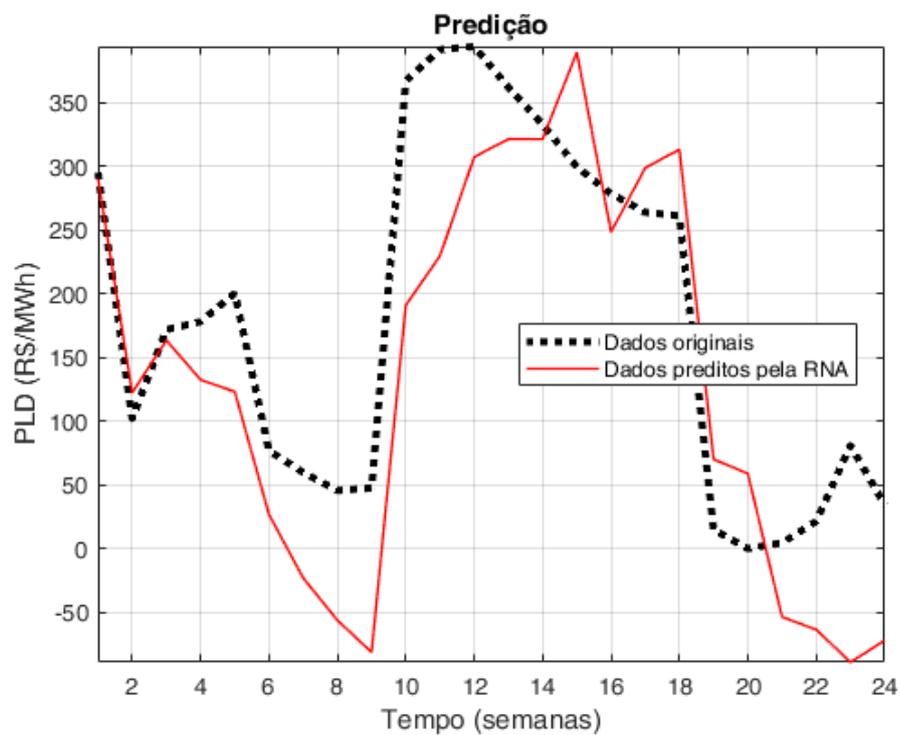


Figura 4.7 – Predição para  $N_1=70$  neurônios intermediários e  $n_q=6$  saídas realimentadas

Assim, é possível observar, de maneira qualitativa, que os resultados entre ambas as topologias são consideráveis. Para um passo no futuro a topologia com menor quantidade de neurônios foi capaz de generalizar tanto o treinamento quanto a predição de maneira bastante satisfatória, evidenciando o comportamento altamente não-linear das redes neurais, em que não necessariamente a rede com maior número de neurônios terá melhor desempenho.

Além do RMSE, outros critérios comumente utilizados na avaliação da predição da RNA são: MAE (*Mean Absolute Error*), o desvio padrão do MAE e o IC (Intervalo de Confiança). As Figuras 4.8, 4.9 e 4.10 apresenta os resultados obtidos de RMSE, MAE, Desvio Padrão do MAE, IC a 90%, 95% e 99% para a predição do PLD 24 semanas à frente, com um passo de predição para cada rodada da rede.

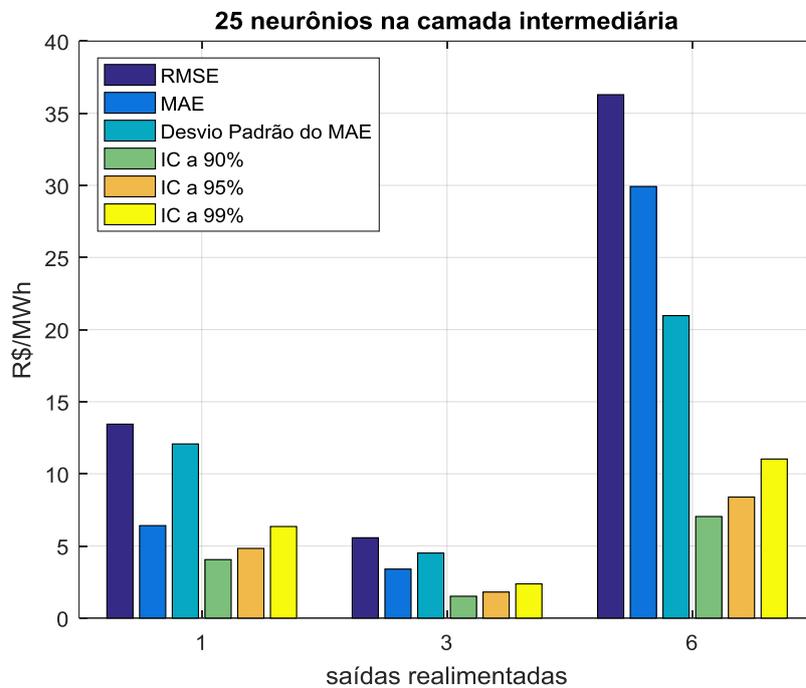


Figura 4.8 – Indicadores para a rede com 25 neurônios na camada intermediária x número de saídas realimentadas

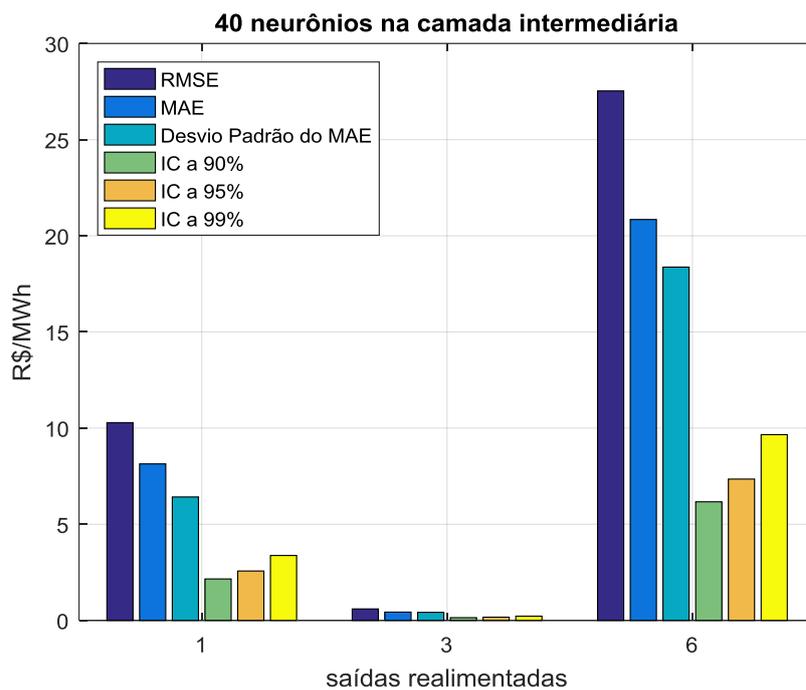


Figura 4.9 – Indicadores para a rede com 40 neurônios na camada intermediária x número de saídas realimentadas

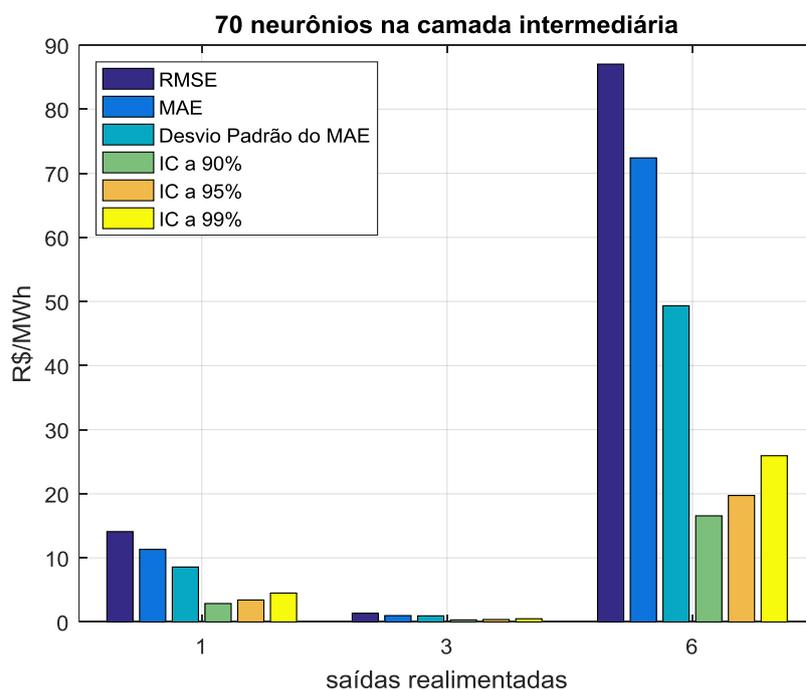


Figura 4.10 – Indicadores para a rede com 25 neurônios na camada intermediária x número de saídas realimentadas

Estes gráficos ilustram que para as diferentes quantidades de neurônios na camada intermediária, ao utilizar 3 neurônios de saídas realimentadas ( $n_q$ ), os indicadores de erros indicam os melhores resultados, evidenciando a não-linearidade das tendências dos dados. Ao elevar este número, os indicadores aumentam demasiadamente em função do superdimensionamento

da rede, fazendo-a entrar em uma fase de supertreinamento (*overtraining*), onde os dados são generalizados de forma redundante, gerando anomalias de aprendizado.

A rede apresentada no presente trabalho também foi testada em horizontes temporais menores. A Tabela 4.2 mostra os valores de indicadores de erros para horizontes de 4, 12 e 24 semanas à frente, considerando a rede com 40 neurônios na camada intermediária e 3 saídas realimentadas.

Tabela 4.2 – Indicadores para horizontes de 4, 12 e 24 semanas com 40 neurônios e 1 passo à frente para cada rodada da rede

Indicadores (R\$/MWh)	Horizonte Temporal		
	4 semanas	12 semanas	24 semanas
RMSE	0,9783	0,7719	0,5962
MAE	0,9305	0,5895	0,4311
Desvio padrão MAE	0,3418	0,2945	0,1683
IC a 90%	0,2869	0,2471	0,1413
IC a 95%	0,3418	0,2945	0,1683
IC a 99%	0,4492	0,3870	0,2212

Nota-se que os indicadores diminuem com o número de amostras preditas no futuro. Isto se deve à maneira de calcular as estatísticas, uma vez que se os indicadores utilizam a divisão pelo número de amostras preditas. Naturalmente, quando o número de amostras cresce, o indicador tende a diminuir pela quantidade numérica de elementos em denominadores de indicadores.

#### 4.3 TREINAMENTO E VALIDAÇÃO PARA VÁRIOS PASSOS NO FUTURO

A estratégia da seção anterior mostra a predição semanal da rede, em que apenas o próximo passo (semana) é predita sem o conhecimento dos dados reais. Além dos dados de predição, pôde-se analisar a eficácia do método *Resilient-Backpropagation* na fase de treinamento da rede, fazendo uma boa aproximação dos dados, como pode-se evidenciar pelo RMSE respectivo de cada topologia.

Nota-se que, em particular para a topologia de 40 neurônios na camada escondida, a validação foi efetiva, conseguindo uma predição com erro irrisório. No entanto, lembra-se que

os dados tratam da predição para apenas um passo afrente, evidenciando ainda a não-linearidade de resultados com a variação da quantidade de neurônios da camada escondida.

Nesta seção, mostram-se os resultados para a predição de vários passos (semanas) no futuro a cada rodada da rede. Isto significa que o conhecimento do futuro não está disponível para a rede durante a predição, a qual será realimentada com suas próprias saídas preditas para os N passos no futuro. Desta maneira, a rede contaminará suas predições com os próprios erros gerados pelo treinamento.

De maneira a expor a diferença entre ambas as metodologias, tem-se a seguinte situação: a predição do dado futuro do tempo  $t + 1$  resulta na saída  $y_{t+1}$  cujo resultado verdadeiro é  $d_{t+1}$ . Quão menor é o erro entre  $y_{t+1}$  e  $d_{t+1}$ , melhor a generalização da rede. Para a topologia da seção anterior, que prediz apenas um passo no futuro, a predição do resultado do próximo tempo,  $t + 2$ , utilizaria como dado realimentado o dado verdadeiro  $d_{t+1}$  para compor a próxima predição, uma vez que este está disponível para a rede. Já a topologia desta seção realimentará a própria saída passada da rede,  $y_{t+1}$ , para compor a próxima considerando que os dados verdadeiros não estão disponíveis. E assim sucessivamente, fazendo com que os erros entre os dados preditos e os dados verdadeiros se acumulem ao longo das iterações futuras.

No escopo deste trabalho, escolheu-se predizer até vinte e quatro passos (semanas) no futuro, aproximadamente entre cinco e seis meses no futuro, utilizando as mesmas topologias analisadas para verificação da efetividade das mesmas a longo prazo, com o objetivo de verificar se as características adquiridas na seção anterior se mantêm na nova modalidade de utilização dos dados.

A primeira topologia utilizou 40 neurônios na camada intermediária (escondida) da rede neural, com três realimentações das saídas (parte recorrente). A Figura 4.11 apresenta os resultados qualitativos e a Tabela 4.3 apresenta os resultados quantitativos para a predição de até 24 passos no futuro.

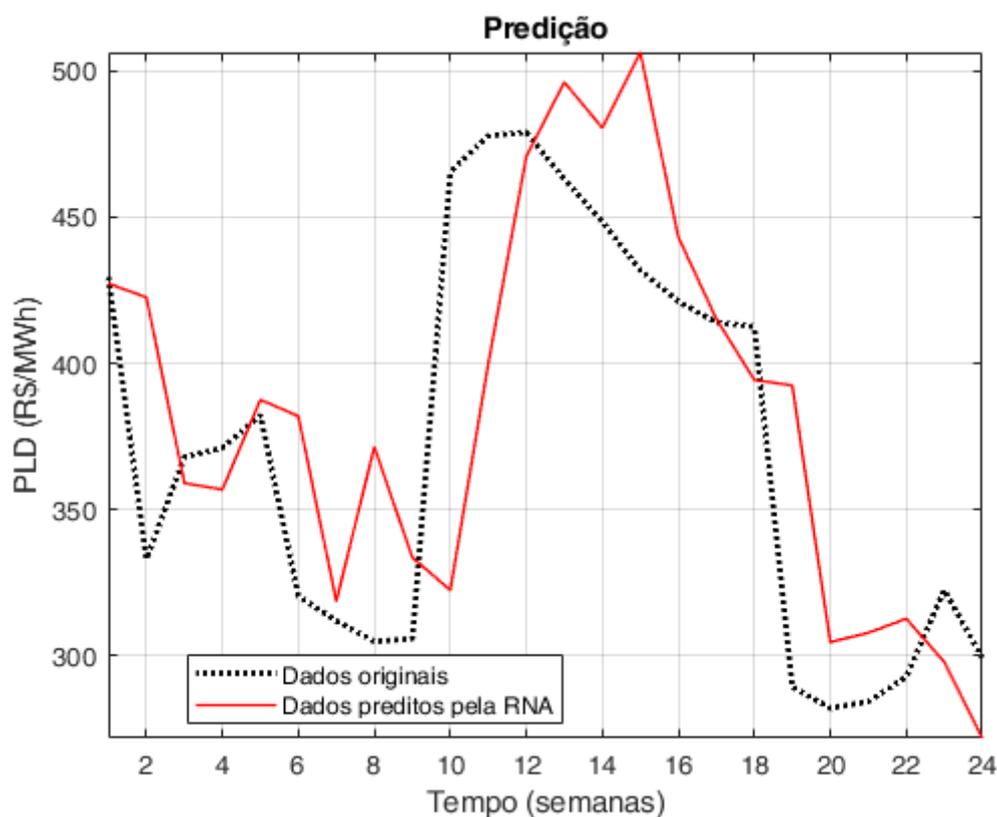


Figura 4.11 – Predição para  $N_1=40$  neurônios intermediários e  $n_q=3$  saídas realimentadas

Tabela 4.3 – Indicadores para horizontes de 4, 12 e 24 semanas à frente com  $N_1 = 40$  e  $n_q = 3$

Indicadores (R\$/MWh)	Horizonte Temporal		
	4 semanas	12 semanas	24 semanas
RMSE	5,3645	6,5331	6,1737
MAE	28,7781	42,6819	38,1144
Desvio padrão MAE	40,8575	44,8616	36,7657
IC a 90%	67,2106	73,7973	60,4796
IC a 95%	80,0808	87,9287	72,0608
IC a 99%	105,2490	115,5635	94,7084

A segunda topologia utilizou 70 neurônios na camada intermediária (escondida) da rede neural, com seis realimentações das saídas (parte recorrente). A Figura 4.12 apresenta os resultados qualitativos e a

Tabela 4.4 apresenta os resultados quantitativos para a predição de até 24 passos no futuro.

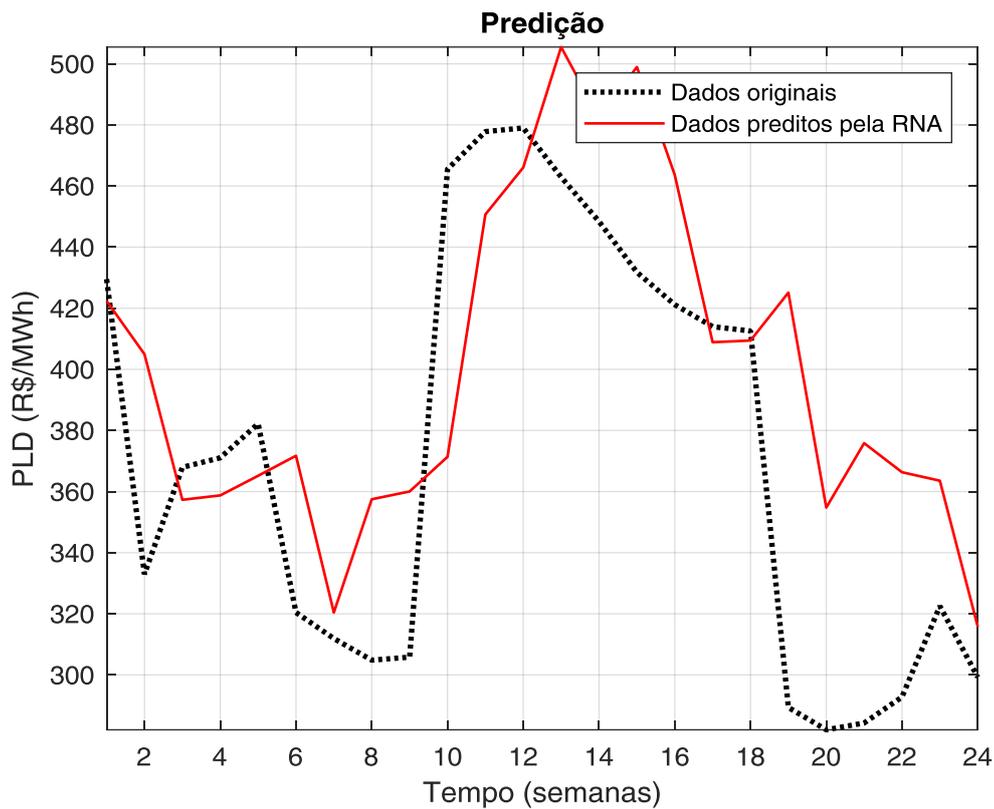


Figura 4.12 – Predição para  $N_1=70$  neurônios intermediários e  $n_q=6$  saídas realimentadas

Tabela 4.4 – Indicadores para horizontes de 4, 12 e 24 semanas à frente com  $N_1 = 70$  e  $n_q = 6$

Indicadores (R\$/MWh)	Horizonte Temporal		
	4 semanas	12 semanas	24 semanas
RMSE	5,0509	5,9160	6,6007
MAE	25,5112	34,9992	43,5693
Desvio padrão MAE	31,0680	29,0322	34,3954
IC a 90%	51,1068	47,7580	56,5805
IC a 95%	60,8932	56,9031	67,4150
IC a 99%	80,0311	74,7870	88,6026

Têm-se que os resultados obtidos para a previsão de apenas uma semana no futuro (seção passada) não se repetem, em virtude da grande variabilidade nos dados, do acúmulo de erros na realimentação de saídas da rede a longo prazo e, especialmente, da grande quantidade de mínimos locais presentes na função objetivo do treinamento, que, eventualmente sobrepõem a capacidade de inovação nos pesos do método *Resilient-Backpropagation*, levando à estagnação do treinamento.

Ainda assim, é possível verificar nas Figuras 4.11 e 4.12, de maneira qualitativa, que a rede recorrente é capaz de generalizar, em média, a tendência dos dados, sendo um indicativo da tendência dos dados a longo prazo. Apesar dos dados quantitativos da rede com 70 neurônios terem se destacado com relação à de 40 neurônios, a vantagem numérica não foi significativa, mostrando que ambas tiveram comportamento semelhante. Entretanto, qualitativamente, a rede de 40 neurônios teve ligeira vantagem na predição da tendência dos dados, conforme observa-se por inspeção visual.

Ressalta-se que ambas têm um desempenho inicial semelhante, acompanhando a tendência dos dados, e este resultado vai se deteriorando com o passar das semanas, aumentando a disparidade entre os dados reais e os dados preditos. Embora os resultados alcançados neste tipo de metodologia não estejam no mesmo patamar alcançado pelos resultados com predição a um passo à frente, considera-se o fato da rede manter a tendência dos dados verdadeiros como um sinal de que é possível utilizá-la, em conjunto com outras técnicas preditivas, para compor a estratégia de contratação da energia para as empresas participantes do mercado de curto prazo.

## 5. CONCLUSÕES

### 5.1 CONCLUSÕES

Esta monografia teve como objetivo a aplicação de técnicas de redes neurais artificiais, na expectativa da predição do Preço de Liquidação das Diferenças (PLD) usando como fonte de conhecimento sobre os dados o próprio histórico (série temporal) de preços do PLD para o submercado Sudeste/Centro-Oeste.

Primeiramente esta monografia apresentou diversos conceitos técnicos acerca do assunto de redes neurais artificiais, bem como a modelagem e algoritmos de treinamento de neurônios artificiais, conceitos que serviram de suporte para a aplicação de redes neurais no decorrer do trabalho. Em seguida apresentou-se a modelagem da rede específica utilizada, a rede realimentada tipo Elman, a qual adquire características autorregressivas mais apropriadas no aprendizado e predição de dados provenientes de sistemas dinâmicos.

Com relação ao desempenho da rede, constatou-se um desempenho elevado para os testes efetuados no modelo de predição de um passo à frente por rodada da rede, permitindo a predição de até 24 semanas, sendo uma predição de aproximadamente seis meses no futuro, com resultados satisfatórios. Os resultados obtidos foram analisados por meio de indicadores RMSE, MAE, Desvio Padrão do MAE e IC a 90%, 95% e 99%. A topologia que concebeu os melhores indicadores foi a rede com 40 neurônios na camada intermediária e 25 entradas, sendo estas 22 realimentações atrasadas da série temporal, a média e desvio padrão dessas realimentações e, finalmente, 3 realimentações atrasadas das saídas da própria rede.

Já a metodologia de utilização da rede com predição para diversos passos no futuro para cada rodada da rede não foi capaz de manter o mesmo desempenho da metodologia anterior, no entanto, foi capaz de manter a tendência dos dados, sendo, portanto, uma boa metodologia para compor, juntamente com outras técnicas preditivas, uma coletânea de indicativos para auxiliar os estrategistas a decidirem pela melhor tomada de decisão na contratação de energia no mercado livre de curto prazo.

O histórico de preços utilizados compreende o período de 21/03/2015 à 08/09/2017, totalizando 132 semanas, perfazendo um histórico de dados de dois anos para a predição dos próximos seis meses. A escolha desse período temporal foi empírica, considerando que os dados de treinamento não poderiam ser grandes demais para não incorporar características antigas ou ultrapassadas do sistema, nem pequenos demais que não propiciassem o aprendizado correto das tendências dos dados prejudicando a predição do PLD.

Por último, vale ressaltar que existem momentos nos dados do PLD em que os valores atingem o teto máximo permitido para o dado no respectivo ano, nos períodos próximos a julho de 2014 nas Figuras 2.2, 2.3, 2.4 e 2.5, páginas 21 a 23. Nestes momentos específicos, os modelos de identificação de sistemas mais usuais, tais como as redes neurais, modelos AR, ARIMA, dentre outros, perdem a referência uma vez que dados constantes não permitem aos modelos o conhecimento da próxima variação, e por isso esses momentos foram excluídos do trabalho.

## **5.2 PROPOSTAS DE DESENVOLVIMENTOS FUTUROS**

Seguindo a linha de pesquisa desenvolvida nesta monografia e tendo em vista os resultados obtidos, os tópicos a seguir apresentam perspectivas no desenvolvimento de trabalhos futuros:

- A fim de criar uma metodologia com maior capacidade de generalização para períodos mais longos que 24 semanas no futuro, inserir outras variáveis de entrada na rede, tais como: Energia de entrada em reservatórios, Energia armazenada em reservatórios e Geração hidrelétrica total, disponibilizados em outras fontes, como no Operador Nacional do Sistema Elétrico (ONS);
- Novas maneiras de treinamento que sejam mais robustas à presença de mínimos locais;
- Utilizar outras técnicas de redes neurais em conjunto com outros métodos de inteligências artificiais, como sistemas *fuzzy*.

## REFERÊNCIAS BIBLIOGRÁFICAS

[1] AMBIENTE LIVRE E AMBIENTE REGULADO. Câmara de Comercialização de Energia Elétrica (CCEE). Disponível em:

Acesso em: 20/08/17.

[2] INFOMERCADO MENSAL Nº 122. Câmara de Comercialização de Energia Elétrica (CCEE). Disponível em: <[https://www.ccee.org.br/portal/faces/pages\\_publico/o-que-fazemos/infomercado?\\_afLoop=39783755035248#!%40%40%3F\\_afLoop%3D39783755035248%26\\_adf.ctrl-state%3Dz1p0hfiqh\\_106](https://www.ccee.org.br/portal/faces/pages_publico/o-que-fazemos/infomercado?_afLoop=39783755035248#!%40%40%3F_afLoop%3D39783755035248%26_adf.ctrl-state%3Dz1p0hfiqh_106)>.

<[https://www.ccee.org.br/portal/faces/pages\\_publico/o-que-fazemos/infomercado?\\_afLoop=39783755035248#!%40%40%3F\\_afLoop%3D39783755035248%26\\_adf.ctrl-state%3Dz1p0hfiqh\\_106](https://www.ccee.org.br/portal/faces/pages_publico/o-que-fazemos/infomercado?_afLoop=39783755035248#!%40%40%3F_afLoop%3D39783755035248%26_adf.ctrl-state%3Dz1p0hfiqh_106)>.

Acesso em: 20/08/17.

[3] BRASIL, Lei nº 9.074, de 7 de julho de 1995.

[4] CARTILHA MERCADO LIVRE DE ENERGIA. Associação Brasileira dos Comercializadores de Energia (ABRACEEL). Disponível em:

<[http://www.abraceel.com.br/archives/files/Abraceel\\_Cartilha\\_MercadoLivre\\_V9.pdf](http://www.abraceel.com.br/archives/files/Abraceel_Cartilha_MercadoLivre_V9.pdf)>.

Acesso em: 20/08/17.

[5] COMERCIALIZAÇÃO. Câmara de Comercialização de Energia Elétrica (CCEE).

Disponível em: <[https://www.ccee.org.br/portal/faces/pages\\_publico/onde-atuamos/comercializacao?\\_afLoop=789045258616065#%40%3F\\_afLoop%3D789045258616065%26\\_adf.ctrl-state%3D154c19nmpw\\_34](https://www.ccee.org.br/portal/faces/pages_publico/onde-atuamos/comercializacao?_afLoop=789045258616065#%40%3F_afLoop%3D789045258616065%26_adf.ctrl-state%3D154c19nmpw_34)>.

<[https://www.ccee.org.br/portal/faces/pages\\_publico/onde-atuamos/comercializacao?\\_afLoop=789045258616065#%40%3F\\_afLoop%3D789045258616065%26\\_adf.ctrl-state%3D154c19nmpw\\_34](https://www.ccee.org.br/portal/faces/pages_publico/onde-atuamos/comercializacao?_afLoop=789045258616065#%40%3F_afLoop%3D789045258616065%26_adf.ctrl-state%3D154c19nmpw_34)>.

Acesso em: 20/08/17.

[6] RELATÓRIO SÍNTESE, ANO BASE 2016. Balanço Energético Nacional (BEN).

Disponível em:

<[https://ben.epe.gov.br/downloads/S%C3%ADntese%20do%20Relat%C3%B3rio%20Final\\_2017\\_Web.pdf](https://ben.epe.gov.br/downloads/S%C3%ADntese%20do%20Relat%C3%B3rio%20Final_2017_Web.pdf)>.

Acesso em: 20/08/17.

- [7] GARCIA, Leandro Henrique Cascaldi. *A reestruturação do setor elétrico brasileiro e a crise do racionamento (1981-2002)*. Monografia. Universidade de São Paulo, 2011.
- [8] Brasil, Lei nº 8.631, de 4 de março de 1993.
- [9] PWC. Disponível em: <http://www.pwc.com/>. Acesso em: 31/08/17.
- [10] SOUZA, Fábio Cavaliere. *Dinâmica da gestão de riscos no ambiente de contratação regulada no setor elétrico brasileiro*. Tese de Doutorado. Universidade Federal do Rio de Janeiro, 2008.
- [11] Brasil, Lei nº 10.848, de 15 de março de 2004.
- [12] O SISTEMA INTERLIGADO NACIONAL. Operador Nacional do Sistema (ONS). Disponível em: <<http://www.ons.org.br/pt/paginas/sobre-o-sin/o-que-e-o-sin>>. Acesso em: 27/08/17.
- [13] SOUSA, Antônio José Sobrinho de. *Análise e Previsão da volatilidade do preço de liquidação das diferenças no mercado brasileiro utilizando o modelo GARCH*. Dissertação de Mestrado. Universidade Federal da Bahia, 2013.
- [14] RESTON FILHO, José Carlos. *Previsão Multi-passos à frente do preço de energia elétrica de curto prazo no mercado brasileiro*. Tese de Doutorado. Universidade Federal do Pará, 2014.
- [15] MOLINA, Juan Pablo et al. *Nash-Cournot equilibria in hydrothermal electricity markets*. IEEE Transactions on Power Systems, v. 26, n. 3, p. 1089-1101, 2011.
- [16] AZEVEDO, Erick Menezes de et al. *Modelo computacional de teoria dos jogos aplicado aos leilões brasileiros de energia elétrica*. Tese de Doutorado. Universidade Estadual de Campinas, 2004.
- [17] BRAGA, A.DE P.; LUDEMIR, T.B.; CARVALHO, A. C. P. L. F. *Redes Neurais Artificiais: Teorias e Aplicações*. São Paulo: LTC, 2000.262 p.

- [18] MCCULLOCH, Warren S.; PITTS, Walter. *A logical calculus of the ideas immanent in nervous activity*. *The bulletin of mathematical biophysics*, v. 5, n. 4, p. 115-133, 1943.
- [19] HEBB, Donald Olding. *The Organization Of Behavior A Neuropsychological Theory*. John Wiley And Sons Inc; New York; Chapman And Hall Limited; London, 2013.
- [20] ROSENBLATT, Frank. *The perceptron: A probabilistic model for information storage and organization in the brain*. *Psychological review*, 1958.
- [21] WIDROW, Bernard. *Generalization and information storage in network of Adaline 'neurons'*. *Self-organizing systems-1962*, 1962.
- [22] MINSKY, Marvin e PAPERT, Seymour. *Perceptrons*, MIT Press, Cambridge, MA, 1969.
- [23] RUMELHART, D. E, MCCLELLAND, J. L., *Parallel Distributed Processing*. MT Press, Vol.1, 1986.
- [24] TEIXEIRA, Ana Carolina Caetano. *Controle de Prótese Robótica por Comandos de Voz utilizando Redes Neurais Artificiais*. Monografia. Universidade Federal de Juiz de Fora, 2017.
- [25] DA SILVA, Ivan Nunes; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. *"Redes Neurais Artificiais: para engenharia e ciências aplicadas"*. Editora Artliber, 2010.
- [26] OLIVEIRA, ÂNGELO ROCHA. *Redes Neurais Artificiais Aplicadas na Detecção, Classificação e Localização de Defeitos em Linhas de Transmissão*. Dissertação de Mestrado. Universidade Federal de Juiz de Fora, 2005.
- [27] HAYKIN, Simon. *Redes neurais: princípios e prática*. Bookman Editora, 2001.