

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

Gabriela Bevilaqua Costa Oliveira

Predição das Variáveis que Compõem o Preço de Liquidação das Diferenças utilizando
Redes Neurais Artificiais

Juiz de Fora

2018

Gabriela Bevilaqua Costa Oliveira

**Predição das Variáveis que Compõem o Preço de Liquidação das Diferenças utilizando
Redes Neurais Artificiais**

Trabalho de conclusão de curso apresentado à
Faculdade de Engenharia da Universidade
Federal de Juiz de Fora, como requisito para
obtenção do grau de Engenheira Eletricista.

Orientador: Leonardo Rocha Olivi

Juiz de Fora

2018

Gabriela Bevilaqua Costa Oliveira

**Predição das Variáveis que Compõem o Preço de Liquidação das Diferenças utilizando
Redes Neurais Artificiais**

Trabalho de conclusão de curso apresentado à
Faculdade de Engenharia da Universidade
Federal de Juiz de Fora, como requisito para
obtenção do grau de Engenheira Eletricista.

Aprovada em (dia) de (mês) de (ano)

BANCA EXAMINADORA

Professor Leonardo Rocha Olivi
Universidade Federal de Juiz de Fora
Orientador

Professor Exuperry Barros Costa
Universidade Federal de Juiz de Fora

Professor Luís Henrique Lopes Lima
Universidade Federal de Juiz de Fora

AGRADECIMENTOS

É difícil colocar em palavras a minha gratidão a todos os que contribuíram para que este ciclo que se conclui fosse possível.

Agradeço primeiramente a Deus, que sempre me ajudou nos momentos difíceis, me amando, me dando suporte e me lembrando que o meu melhor futuro não é aquele que eu planejo sozinha, mas sim o que planejo com Ele.

Também agradeço profundamente à minha mãe Flávia, que nunca mediu esforços para que eu pudesse ir além. Sempre me poupando, me entendendo e colocando as minhas necessidades como prioridade às dela para que eu pudesse chegar até aqui. Sem ela, que em todos os momentos sorriu e chorou comigo, estou certa de que nada disso seria possível. Juntamente com ela, agradeço também a todos os meus familiares, que sempre torceram por mim e me incentivaram.

Ao meu professor e orientador Leonardo Olivi que foi mais que excelente em tudo o que fez nestes cinco anos de graduação. Sempre paciente, bem humorado, dedicado e didático, fazendo qualquer coisa parecer fácil. O precursor das redes neurais em minha caminhada acadêmica.

Agradeço ao meu noivo Guilherme e aos meus amigos, que sempre entenderam meus momentos de ausência para estudar, e me apoiaram e acreditaram em mim em todos os momentos.

“O passado é uma cortina de vidro. Felizes o que observam o passado para poder caminhar no futuro.”

(Augusto Cury)

RESUMO

O Preço de Liquidação das Diferenças (PLD) tem grande relevância para o mercado livre de energia elétrica brasileiro, cuja matriz energética é predominantemente de geração hidrelétrica. Assim, este preço visa encontrar o ponto de equilíbrio entre o benefício presente do uso da água dos reservatórios e o benefício futuro de seu armazenamento. Entender o processo de comercialização de energia elétrica e as variáveis envolvidas na valoração deste insumo permite contratações energéticas bem sucedidas, a menores preços e com os montantes adequados. O presente trabalho propõe, então, a predição de três das variáveis que compõem o PLD, são elas: Energia Armazenada, Energia Natural Afluyente por Bacia e Energia Natural Afluyente por Subsistema. O método de predição das variáveis se baseia na implementação de uma Rede Neural Artificial do tipo TDNN para cada variável, cujo algoritmo de treinamento é um híbrido de *Levenberg-Marquardt* com *Momentum* e *Resilient-Backpropagation*. As redes mostraram bom desempenho para a predição das variáveis a uma semana no futuro e o acompanhamento da tendência dos dados a quatro semanas o futuro.

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

ABSTRACT

The Settlement Price of Differences (PLD) has great relevance for the free market of Brazilian electric energy, whose energy matrix is predominantly hydroelectric generation. Thus, this price aims to find the balance between the present benefit of the reservoirs's water use and its future storage benefits. Understanding the commercialization process of electric energy and the variables involved in the valuation of this input allows for successful energy contracting, at lower prices and with the appropriate amounts. The present work proposes, then, the prediction of three of the variables that compose the PLD, they are: Stored Energy, Influent Natural Energy by Basin and Subsluent Natural Energy by Subsystem. The variables prediction method is based on TDNN Artificial Neural Network implementation for each variable, whose training algorithm is a Levenberg-Marquardt hybrid with Momentum and Resilient-Backpropagation. The networks showed good performance for the variables prediction for one week in the future and the monitoring of the data trending for four weeks in the future.

Keywords: Artificial Neural Network, Settlement Price of Differences, Prediction.

LISTA DE ILUSTRAÇÕES

Figura 1 - Cronograma de decisão para um sistema hidrotérmico	16
Figura 2 - Neurônio Biológico	19
Figura 3 - Neurônio Artificial	20
Figura 4 - Exemplo de rede feedforward de camada simples	22
Figura 5 - Exemplo de rede Perceptron de Múltiplas camadas	22
Figura 6 - Fase de treinamento para $N_1 = 50$ neurônios intermediários	30
Figura 7 - Predição para $N_1 = 50$ neurônios intermediários	30
Figura 8 - Fase de treinamento para $N_1 = 100$ neurônios intermediários	31
Figura 9 - Predição para $N_1 = 100$ neurônios intermediários	31
Figura 10 - Fase de treinamento para $N_1 = 120$ neurônios intermediários	32
Figura 11 - Predição para $N_1 = 120$ neurônios intermediários	32
Figura 12 - Fase de treinamento para $N_1 = 50$ neurônios intermediários	34
Figura 13 - Predição para $N_1 = 50$ neurônios intermediários	34
Figura 14 - Fase de treinamento para $N_1 = 100$ neurônios intermediários	35
Figura 15 - Predição para $N_1 = 100$ neurônios intermediários	35
Figura 16 - Fase de treinamento para $N_1 = 101$ neurônios intermediários	36
Figura 17 - Predição para $N_1 = 101$ neurônios intermediários	36
Figura 18 - Fase de treinamento para $N_1 = 50$ neurônios intermediários	38
Figura 19 - Predição para $N_1 = 50$ neurônios intermediários	38
Figura 20 - Fase de treinamento para $N_1 = 100$ neurônios intermediários	39
Figura 21 - Predição para $N_1 = 100$ neurônios intermediários	39
Figura 22 - Fase de treinamento para $N_1 = 120$ neurônios intermediários	40
Figura 23 - Predição para $N_1 = 120$ neurônios intermediários	40

LISTA DE TABELAS

Tabela 1 - Indicadores para configurações de 50, 100 e 120 neurônios intermediários para a variável Demanda Máxima	33
Tabela 2 - Indicadores para configurações de 50, 100 e 101 neurônios intermediários para a variável Energia Natural Afluente por Bacia	33
Tabela 3 - Indicadores para configurações de 50, 100 e 101 neurônios intermediários para a variável Energia Natural Afluente por Subsistema	37

LISTA DE ABREVIATURAS E SIGLAS

PLD	Preço de Liquidação das Diferenças
MCP	Mercado de Curto Prazo
RNA	Rede Neural Artificial
ACR	Ambiente de Contratação Regulada
ACL	Ambiente de Contratação Livre
CCEE	Câmara de Comercialização de Energia Elétrica
ANEEL	Agência Nacional de Energia Elétrica
SIN	Sistema Interligado Nacional
ONS	Operador Nacional do Sistema Elétrico
AR	<i>Autoregressive</i>
ARMAX	<i>ARMA with explanatory variables</i>
ARIMA	<i>Autoregressive integrated moving average</i>
PMC	<i>Perceptron de Múltiplas Camadas</i>
TDNN	<i>Time Delay Neural Network</i>
EQM	Erro Quadrático Médio

SUMÁRIO

1 INTRODUÇÃO	11
2 REVISÃO BIBLIOGRÁFICA	13
2.1 O Setor Elétrico Brasileiro	13
2.2 O Preço de Liquidação das Diferenças	14
2.3 Predição das Variáveis que Compõem o PLD	16
2.4 Redes Neurais Artificiais	17
2.4.1 Introdução	17
2.4.2. O Neurônio Biológico	18
2.4.3 O Neurônio Artificial	20
2.4.4 A Rede Perceptron	21
2.4.5 A Rede Perceptron Multicamadas (PMC)	22
3 METODOLOGIA	24
3.1 Treinamento	24
3.2 Validação	28
4 RESULTADOS	29
4.1 Demanda Máxima	29
4.2 Energia Natural Afluente por Bacia	33
4.3 Energia Natural Afluente por Subsistema	37
5 CONCLUSÕES E TRABALHOS FUTUROS	41
5.1 Conclusões	41
5.2 Trabalhos Futuros	42

1 INTRODUÇÃO

A energia elétrica é essencial para a vida moderna e, a sua comercialização impacta diretamente no crescimento do país, uma vez que os gastos com a mesma influenciam fortemente na estratégia operacional de empresas de todos os portes e segmentos. Assim, entender o processo de comercialização da energia elétrica e as variáveis envolvidas na valoração deste insumo permite contratações energéticas bem sucedidas, a menores preços e com os montantes adequados [1].

Para grandes consumidores há, em geral, duas formas de se adquirir energia elétrica: por meio do mercado cativo ou do mercado livre. O mercado cativo corresponde à forma tradicional de comercialização de energia elétrica, onde o consumidor de pequeno porte, considerado cativo, é atendido pela concessionária distribuidora local a preços pré-fixados e normalmente controlados pelas agências reguladoras [2].

Em contrapartida, o conceito de mercado livre de energia é restrito aos consumidores de grande porte, denominados livres ou especiais, os quais não são obrigados a comprar energia da concessionária local. Esta forma de mercado está relacionada a uma possível competição, onde a energia pode ser ofertada a preços com valores reduzidos por meio de contratos bilaterais entre vendedor e consumidor. Assim, o mercado livre de energia torna-se mais atrativo, pois permite às empresas consumidoras e aos compradores, maiores economias e gestões energéticas [1] e [2].

Estes contratos bilaterais podem ser livremente negociados, com contratações de curto, médio ou longo prazo. Porém, pelo fato de a energia elétrica não ser um insumo que possa ser embalado e entregue na quantidade exata à que foi contratada, é possível que hajam pequenas diferenças entre energia medida e contratada. Essas diferenças são valoradas a um preço chamado de Preço de Liquidação das Diferenças (PLD) e podem ser contratadas no Mercado de Curto Prazo (MCP) até nono dia útil do mês seguinte ao consumo de forma a quitar essas diferenças [2].

Segundo [3], o PLD é um cálculo complexo que visa precificar os máximos e mínimos dos preços de geração de energia no Brasil. Pelo país ter na matriz energética predominantemente a geração hidrelétrica e o regime de água ser imprevisível, gera um caráter probabilístico ao processo de geração, com diversas incertezas contaminando as diversas variáveis utilizadas no cálculo do PLD.

Assim, o cálculo do PLD visa encontrar o ponto de equilíbrio entre o benefício presente do uso da água e o benefício futuro de seu armazenamento, medido em termos da economia esperada dos combustíveis das usinas termelétricas. Ou seja, deve-se usar a água dos reservatórios no presente para gerar energia a um custo mais baixo, já que a água é um bem gratuito, ou deve-se utilizar mais as termelétricas a um custo operacional elevado para preservar os reservatórios, minimizando os riscos de racionamentos futuros? Questões como estas podem ser respondidas caso o preço do PLD possa ser predito com confiabilidade.

Sendo assim, é evidente que o estudo e a análise das variáveis que compõem o PLD, como nível dos reservatórios, condições hidrológicas, preços do combustível, demanda e intercâmbio de energia, dentre outras variáveis, auxiliam nas tomadas de decisões tanto na geração, onde analisa-se usar hidrelétricas ou térmicas, quanto no consumo, onde as empresas, cientes dos momentos de altas e baixas dos preços, podem contratar energia a um melhor custo-benefício.

Portanto, devido ao caráter probabilístico destas variáveis, torna-se necessária a utilização de ferramentas adequadas que possam auxiliar no processo de tomada de decisão. Por esta razão, o presente trabalho faz uso das Redes Neurais Artificiais (RNA), que é uma técnica de inteligência artificial com grande aplicabilidade em predição de séries temporais não-determinísticas. São redes baseadas no comportamento do cérebro humano, ou seja, nas redes neurais naturais. Neste trabalho, cada rede neural recebe como entrada dados históricos de uma variável que compõe o PLD e, tem como saída, a previsão desta no futuro.

A organização do trabalho se dá em cinco capítulos, sendo o primeiro esta introdução. No segundo capítulo é apresentada a revisão bibliográfica descrevendo os temas apontados no trabalho através da literatura já existente. No terceiro capítulo é apresentado o desenvolvimento da rede neural artificial implementada para o problema em questão. No quarto capítulo são mostrados os resultados dos experimentos realizados. E finalmente, o quinto capítulo contém a conclusão do trabalho e aponta possíveis trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

O objetivo deste capítulo é apresentar uma revisão sobre os temas presentes neste trabalho. Serão abordados os assuntos: Setor Elétrico Brasileiro, Comercialização da Energia Elétrica e Redes Neurais Artificiais. Ao final do capítulo, será exposto, ainda, uma explicação do problema proposto e como pretende-se resolvê-lo.

2.1 O Setor Elétrico Brasileiro

Segundo [4], o setor elétrico brasileiro pode ser dividido em três períodos: o primeiro, correspondente à primeira metade do século XX, foi quando se deram início às atividades envolvendo energia elétrica no país; o segundo, que foi até o final do século XX, foi marcado pela forte intervenção do Estado nas atividades do setor elétrico; e o terceiro, que se caracterizou pela reestruturação do setor elétrico para um modelo com maior participação da iniciativa privada, segue até os dias atuais.

No último período, de reestruturação do setor elétrico, o governo deixou, em grande parte, o papel de executor dos serviços para ter a função de regulador do mercado. Assim, procurou tomar medidas que pudessem atrair investidores para a ampliação do parque gerador de energia elétrica do país, e dessa forma estimular o desenvolvimento econômico da nação [4].

Em 1995 foi sancionada a lei nº 9.074, que estabeleceu a criação do Produtor Independente de Energia e do Consumidor Livre de Energia, e assim, foi capaz de modificar a comercialização de energia elétrica no Brasil. Uma das consequências foi a separação das empresas de energia em áreas específicas: geração, transmissão, distribuição e comercialização de energia. Nos setores de geração e comercialização incentivou-se o surgimento de concorrência; já a transmissão e a distribuição continuaram sendo tratadas como serviços públicos regulados [1].

Porém, devido a problemas nas estratégias políticas antigas, de acordo com [5], bastou um período hidrológico desfavorável em 2001 para que o sistema brasileiro, fortemente dependente de hidroeletricidade, gerasse menos do que era necessário, forçando o governo federal a implementar programas de racionamento de energia elétrica, paralelamente a outras ações que buscavam a melhoria do sistema.

Foi criado, então, em 2004, o Novo Modelo do Setor Elétrico através das leis nº 10.847 e nº 10.848, a fim de solucionar estes problemas e melhorar o sistema. Em relação à comercialização, uma das ações foi a criação de dois ambientes para a celebração de contratos de compra e venda de energia, com lógicas e estruturas distintas: o Ambiente de Contratação Regulada (ACR), composto de agentes de geração, distribuição e comercialização; e o Ambiente de Contratação Livre (ACL), formado por geradores, distribuidores, comercializadores e consumidores livres e especiais [2] e [4].

O Ambiente de Contratação Regulada visa abastecer os consumidores cativos, que adquirem a energia das concessionárias de distribuição às quais estão conectados, via ponto de entrega do ramal de entrada. Neste ambiente, a contratação é realizada por meio de leilões de energia que são promovidos pela Câmara de Comercialização de Energia Elétrica (CCEE), sob delegação da Agência Nacional de Energia Elétrica (ANEEL). Esses leilões são realizados pelo critério de menor tarifa, ou seja, vencem o leilão os empreendedores que oferecerem o menor preço. Com isso, a tarifa repassada ao consumidor é a menor possível, garantindo a modicidade tarifária [2].

Já no Ambiente de Contratação Livre, podem comprar energia os consumidores livres e especiais, segundo a Lei nº 9.074/95 [6]. Os consumidores livres tradicionais são grandes consumidores, com contratos de energia acima de 3000 kW. Já os consumidores livres especiais, considerados empresas de pequeno e médio portes, têm contratos entre 500 kW e 3000 kW e a energia contratada deve ser advinda de fontes limpas, como solar, eólica e pequenas centrais hidrelétricas. Os contratos são livremente negociados, com cláusulas contratuais firmadas entre as partes, onde define-se o prazo, o preço, o volume de energia, entre outros.

Para a CCEE realizar a contabilização dos agentes, ela utiliza a quantidade de energia contratada total de cada um desses agentes e a quantidade de energia medida total de cada um deles. Este ambiente onde a diferença entre energia gerada e consumida por cada agente é contabilizada é chamado de Mercado de Curto Prazo (MCP). Essa diferença é valorada a um preço chamado Preço de Liquidação das Diferenças (PLD) [3].

2.2 O Preço de Liquidação das Diferenças

O funcionamento do MCP é possível devido ao Sistema Interligado Nacional (SIN), que é responsável por todo controle de produção e transmissão de energia elétrica no Brasil. Em toda a extensão do país estão inseridas usinas hidráulicas e térmicas de diferentes tipos. Desta forma, o SIN é caracterizado como um sistema hidrotérmico de grande porte, com a predominância de usinas hidrelétricas, interligado por linhas de transmissão. Essa característica da malha brasileira ser interligada que permite aos agentes comercializar energia livremente entre si, independentemente dos submercados onde estão as usinas e as cargas [2] e [9].

Atualmente, o Brasil possui 4 submercados: Norte, Nordeste, Sudeste/Centro-Oeste, e Sul; que não possuem as mesmas divisões geopolíticas conhecidas, mas são divididos baseados em critérios geoeletricos, considerando as características da malha de transmissão; o que influencia diretamente no cálculo do PLD [2]. Assim, o PLD é calculado semanalmente para cada um dos quatro submercados e em três patamares de carga: leve, médio e pesado. Estes patamares correspondem aos horários de baixo, médio e alto consumo, respectivamente.

Segundo [8], o cálculo e divulgação do PLD é feito pela CCEE, regulamentado pela ANEEL. O valor divulgado semanalmente, é válido para a próxima semana de operação do sistema elétrico. Assim, no final de cada mês é realizada a média ponderada mensal para cada submercado, conhecido como PLD Mensal, preço no qual as contratações de energia no MCP são baseadas.

Devido ao predomínio da fonte hídrica no SIN, o PLD tem como objetivo encontrar o ponto ótimo entre a utilização da energia provinda das hidrelétricas, ou o armazenamento de água para utilização futura. A maior utilização da energia hidrelétrica traria menores custos imediatos ao sistema, pois esta fonte é mais barata, porém seu uso demasiado aumentaria os preços futuros e aumentaria o risco de déficit de energia. A Figura 1 mostra este processo de tomada de decisão e planejamento.

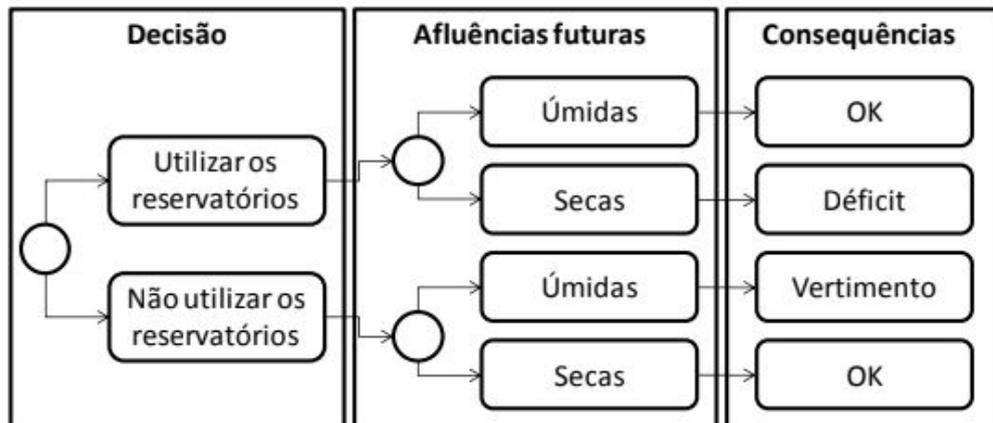


Figura 1 - Cronograma de decisão para um sistema hidrotérmico [11].

Assim, o PLD é calculado baseado principalmente no estado atual dos reservatórios, previsão de chuvas, demanda de energia, preço de combustíveis das usinas térmicas, entrada de novos projetos e disponibilidade de transmissão e geração. Estas são algumas das variáveis que compõem o PLD [3].

O planejamento da operação é feito pelo Operador Nacional do Sistema Elétrico (ONS) [10] com base nas previsões destas variáveis. Então, quando os reservatórios estão com alto armazenamento e as previsões climáticas são favoráveis, ou seja, maior incidência de chuvas nas principais bacias hidrográficas para a geração, o PLD é menor. Caso os reservatórios estejam com volumes mais baixos e a previsão de chuvas é desfavorável, seu preço aumenta.

Como o ONS é responsável pela operação física do sistema, ele acionará, de acordo com a demanda, o melhor conjunto de usinas disponíveis para atender essa demanda, visando o menor custo de operação.

Sendo assim, tendo em vista que o Preço de Liquidação das Diferenças (PLD) é a variável mais importante para o Mercado Livre de Energia, entender as variáveis que o compõem e acompanhar o funcionamento das mesmas, é essencial para uma predição mais precisa e, conseqüente contratação bem sucedida, no momento certo e com menores preços [8].

2.3 Predição das variáveis que compõem o PLD

Tendo em vista a importância do Preço de Liquidação das Diferenças (PLD) no Mercado de Curto Prazo (MCP), a sua predição tem sido alvo de diversas pesquisas. Como modelos de predição para esta aplicação, [11] cita: teoria dos jogos, modelos fundamentais e modelos de séries temporais, onde se encontram os modelos autorregressivos (AR, ARMAX, ARIMA), Redes Neurais Artificiais, Sistemas Fuzzy, dentre outros.

Fabiane Barbosa [1] utiliza em seu trabalho uma Rede Neural Artificial tipo Elman como modelo de série temporal. Nesta, a rede realimenta as próprias saídas passadas como entradas, a fim de prever o PLD até 24 semanas no futuro. Utilizando o valor desta variável como entrada da rede, a autora faz a predição do futuro a um passo e também até 24 passos (semanas) no futuro. Os resultados apresentam bom desempenho inicial para as duas abordagens e, apenas na segunda, vão se deteriorando com o passar das semanas, mas sendo uma relevante ferramenta para verificar a tendência dos dados verdadeiros.

Já Reston Filho [11] faz a predição de até 24 semanas no futuro de algumas das variáveis que compõem o PLD utilizando os filtros autorregressivos integrados de médias móveis (ARIMA) como modelo de série temporal. A predição das variáveis é então combinada a uma RNA em cascata, onde tem-se uma predição aceitável do Preço de Liquidação das Diferenças.

Sabe-se que o PLD é uma variável muito volátil e de difícil previsão. Grande parte deste caráter probabilístico se dá pelo fato de a matriz energética brasileira ter predominantemente a geração hidrelétrica como fonte de energia e o regime de água ser imprevisível. Assim, sem um conhecimento do comportamento das variáveis que compõem o PLD, a predição se torna difícil, como concluído em [1]. No entanto, a escolha de métodos pouco robustos para a predição com o conhecimento do comportamento das variáveis também torna a predição difícil, como é visto em [11]. Uma alternativa, então, seria a predição destas variáveis explanatórias com um método robusto, como as Redes Neurais Artificiais.

2.4 Redes Neurais Artificiais

2.4.1 Introdução

As Redes Neurais Artificiais (RNAs) são modelos computacionais que apresentam um modelo matemático inspirado no sistema nervoso dos seres vivos. Possuem a capacidade de

adquirir conhecimento através da experiência. Com isso, são capazes de generalizar o conhecimento aprendido, fornecendo respostas adequadas para dados desconhecidos, os quais não estão no conjunto de exemplos [12].

O primeiro modelo matemático inspirado no neurônio biológico foi desenvolvido por McCulloch e Pitts [13] em 1943. Logo após, em 1949, Hebb propôs o primeiro método de treinamento, o qual chamou de regra de aprendizado de Hebb, baseado em hipóteses e observações de caráter fisiológicos [14].

Estes trabalhos fizeram com que, no decorrer dos anos 50, diversos pesquisadores se interessassem nesta linha de pesquisa. Dentre eles, se destacou o trabalho de Frank Rosenblatt [15], que no final da década de 50 desenvolveu o primeiro neurocomputador, idealizando o modelo básico do Perceptron, o qual despertou interesse devido à sua capacidade de reconhecer padrões simples. Logo depois, em 1960, Widrow e Hoff [16] desenvolveram um tipo de rede chamada Adaline (Adaptative Linear Element) e, posteriormente, propôs-se a Madaline, a Adaline múltipla, rede cujo algoritmo de aprendizado é fundamentado no Método dos Mínimos Quadrados (Least Mean Square, LMS).

Após estes trabalhos pioneiros, muitas pesquisas se voltaram para este assunto. Porém, em 1969, Minsky e Papert [17] demonstram a limitação das redes neurais artificiais constituídas de apenas uma camada, como o Perceptron e a Adaline, em resolverem problemas de classificação como o Ou Exclusivo (XOR). Com esta publicação, demonstraram que os neurônios artificiais não eram capazes de solucionar problemas que não fossem linearmente separáveis.

Com isso, por alguns anos, poucas pesquisas foram desenvolvidas sobre as RNAs. Somente em 1989, Rumelhart, Hinton e Williams [18] desenvolveram um algoritmo capaz de solucionar o problema do XOR. Este algoritmo, chamado de *Backpropagation*, permitia ajustar os pesos em uma rede com mais de uma camada e fez com que fosse despertado o interesse em estudos sobre RNAs novamente. Assim, destes dias até hoje, as Redes Neurais Artificiais têm sido aplicadas em diversas áreas do conhecimento, como aproximador universal de funções, reconhecimento e classificação de padrões, sistemas de predição, dentre outras aplicações [12].

2.4.2 O Neurônio Biológico

Conforme citado no tópico anterior, 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, cuja função é basicamente conduzir estímulos elétricos advindos de reações físico-químicas sob determinadas condições de operação [12]. Em sua estrutura, o neurônio biológico pode ser dividido em três partes principais: os dendritos, o corpo celular (também chamado de soma) e o axônio, conforme mostrado na Figura 2.

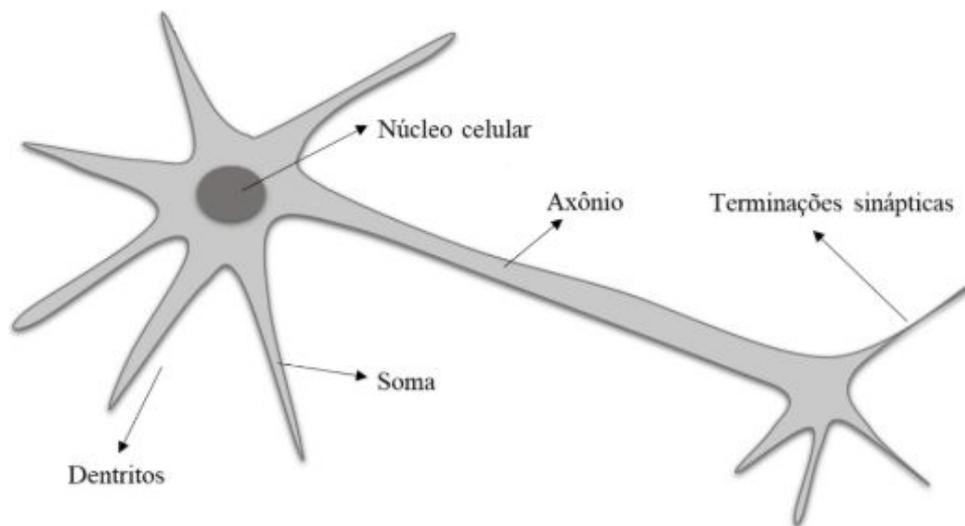


Figura 2 - Neurônio Biológico.

A principal função dos dendritos é captar os estímulos vindos de outros neurônios, ou seja, são os receptores dos sinais de entradas. Já o corpo celular tem a função de processar todas as informações que vêm dos dendritos, a fim de produzir um potencial de ativação que indica se o neurônio pode disparar um impulso elétrico ao longo de seu axônio. E o axônio, por sua vez, conduz os impulsos elétricos para outros neurônios, sendo então, os transmissores das saídas do neurônio [12].

As ligações entre os neurônios são feitas através de sinapses e possuem um grau de importância. O núcleo agrega todos os sinais recebidos pelos dendritos e, caso ele seja ativado por um sinal de ativação, faz a tomada de decisão: se deve enviar ou não um sinal para os outros neurônios ligados ao axônio.

Ainda que as atividades relacionadas aos neurônios pareçam simples, quando estes atuam em conjunto, são os principais responsáveis por todos os processamentos executados e gerenciados pelo cérebro humano. Estima-se que cada rede neural biológica seja constituída

por cerca de 100 bilhões de neurônios, onde cada um deles é interligado em média a outros 6.000 neurônios, totalizando assim, 600 trilhões de sinapses [12].

2.4.3 O Neurônio Artificial

Os neurônios artificiais, propostos por McCulloch e Pitts [13], são não-lineares e funcionam recebendo e transmitindo informações, assim como os neurônios biológicos. Fornecem saídas contínuas e produzem uma resposta de acordo com as entradas coletadas dos sinais externos, a qual é dependente da sua função de ativação. O modelo de um neurônio artificial pode ser visto na Figura 3.

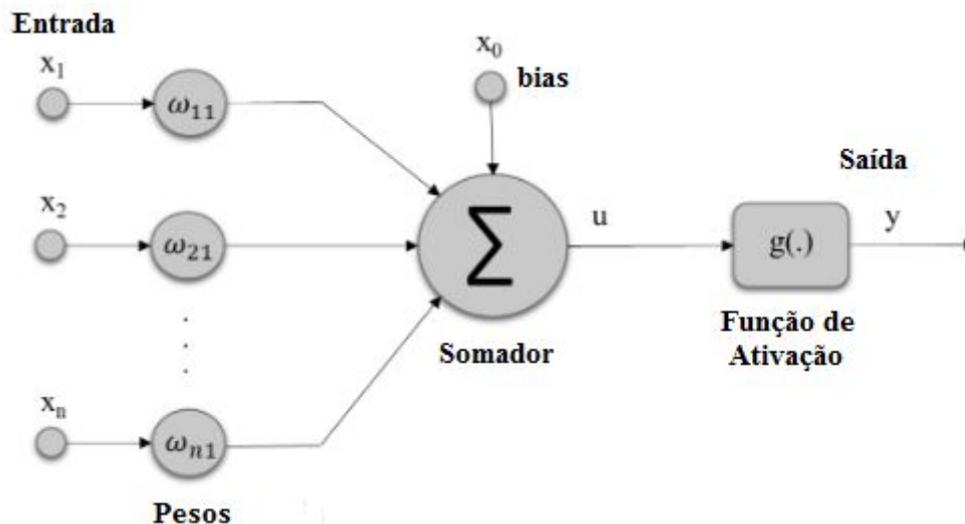


Figura 3 - Neurônio Artificial.

Os diversos sinais de entrada $\{x_1, x_2, \dots, x_n\}$ são análogos aos impulsos elétricos externos captados pelos dendritos no neurônio biológico. Estes sinais representam os valores assumidos pelas variáveis de determinada aplicação e normalmente são normalizados, a fim de aumentar a eficiência computacional dos algoritmos de aprendizagem.

As ponderações exercidas pelas junções sinápticas do modelo biológico são representadas pelos pesos sinápticos $\{\omega_1, \omega_2, \dots, \omega_n\}$. Estes, por sua vez, representam a relevância que determinada entrada possui. O processo de alteração dos pesos sinápticos é chamado de processo de aprendizagem. Neste processo, quanto maior o grau de importância

de uma entrada $\{x_i\}$, maior será o valor do peso sináptico $\{\omega_i\}$ referente a ela e, conseqüentemente, maior influência esta entrada terá na saída $\{y\}$.

Os sinais ponderados são agregados, ou seja, somados, e produzem um valor de potencial de ativação $\{u\}$. Este potencial de ativação é produzido pela diferença entre o valor do combinador linear e do limiar de ativação. O limiar de ativação, também chamado de bias (θ), especifica qual será o patamar apropriado para que o valor gerado pelo combinador linear (Σ) possa gerar um valor de disparo para a saída do neurônio.

Por meio da função de ativação $\{g(\cdot)\}$ o neurônio pode decidir se o resultado da soma ponderada das entradas será ativado ou não. Segundo [1], a escolha adequada desta função é capaz de melhorar o processo de convergência da rede durante seu treinamento. São exemplos de função de ativação as funções degrau, sinal, rampa simétrica, logística e tangente hiperbólica. O resultado produzido pela aplicação da função escolhida corresponde ao sinal de saída $\{y\}$.

Por fim, o resultado produzido pelo neurônio proposto por [13] pode ser resumido por:

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

$$y = g(u) \quad (2)$$

2.4.4 A Rede Perceptron

O *Perceptron*, idealizado por [15], é a forma mais simples de configuração de uma Rede Neural Artificial. Similar ao ilustrado na Figura 3, a simplicidade do *Perceptron* está associada à sua condição de ser constituído de apenas uma camada neural e um neurônio artificial nesta única camada.

A principal arquitetura associada a esta rede é a rede *feedforward* de camada simples. Nesta, tem-se apenas uma camada de entrada e uma única camada de neurônios, que é a própria camada de saída. Assim, o fluxo de informações segue sempre em uma única direção, ou seja, da camada de entrada em direção à camada de saída. Esta rede, com grande aplicabilidade em problemas que envolvem classificação de padrões e filtragem linear, é mostrada na Figura 4.

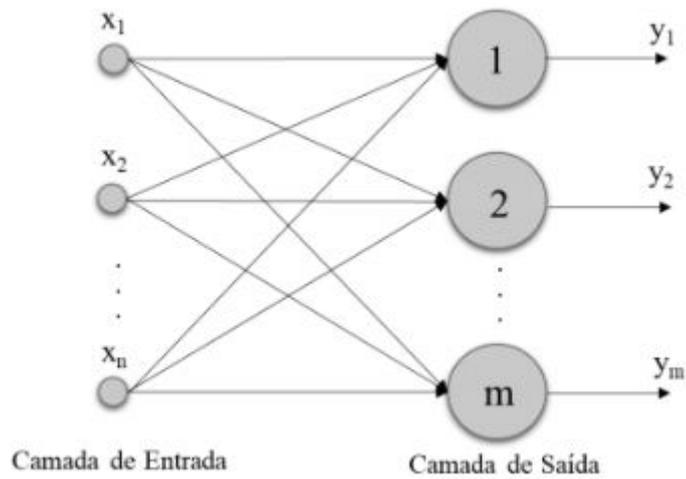


Figura 4 - Exemplo de rede *feedforward* de camada simples.

2.4.5 A Rede Perceptron Multicamadas (PMC)

As redes Perceptron de Múltiplas Camadas (PMC) são caracterizadas pela presença de uma ou mais camadas intermediárias de neurônios entre as camadas de entrada e de saída da rede, como mostrado na Figura 5.

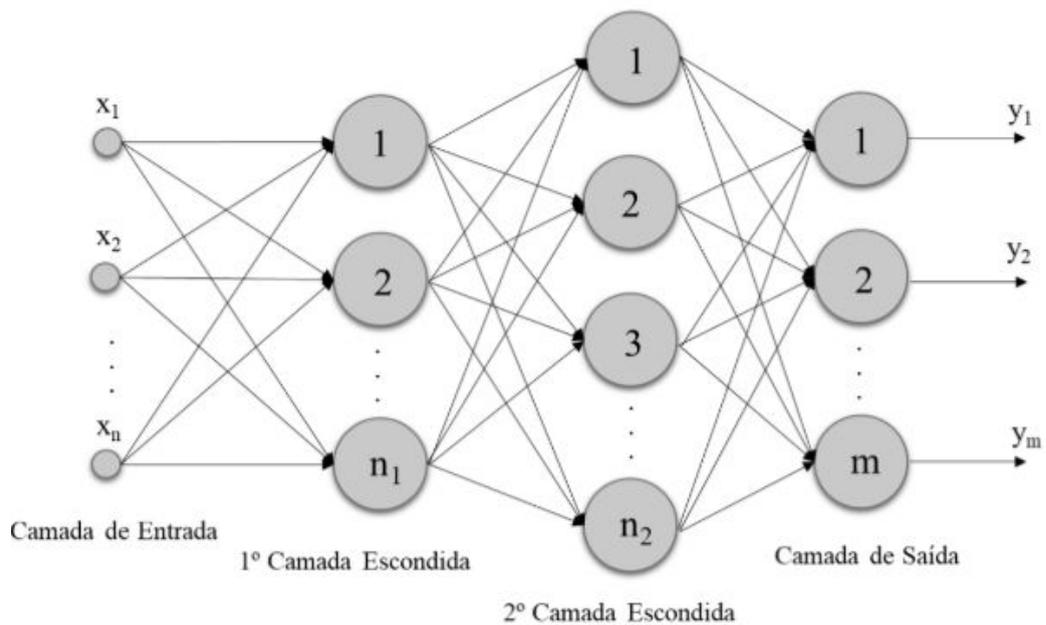


Figura 5 - Exemplo de rede *Perceptron* de Múltiplas camadas.

Estas redes são dotadas de uma vasta quantidade de aplicações, dentre elas, pode-se citar os problemas relacionados à aproximação de funções, classificação de padrões,

otimização, robótica, dentre outros. O fato de a rede ter mais de um neurônio na saída, diferentemente das redes Perceptron e Adaline, permite uma maior quantidade de aplicações possíveis.

A principal arquitetura associada às redes PMC é a rede *feedforward* de camadas múltiplas, cujo treinamento é realizado de forma supervisionada, por um método conhecido como *backpropagation*. Este método é composto por duas fases: *forward* e *backward*. A primeira fase a ser aplicada é a “propagação adiante” (*forward*), onde um padrão é apresentado à camada de entrada da rede e é propagado entre as camadas, uma a uma, até que a resposta seja produzida pela camada de saída.

Na segunda fase, a fase de “propagação reversa” (*backward*), a saída obtida é comparada à saída desejada para esse padrão e então é calculado o erro, caso este exista. Existindo, o erro é propagado a partir da camada de saída até a camada de entrada, e os pesos sinápticos vão sendo modificados conforme o erro é retropropagado. Assim, diminui-se a cada iteração a soma dos erros entre a resposta obtida e a resposta gerada.

Existem algumas derivações do algoritmo *backpropagation* descrito acima, como o *Backpropagation com Momentum*, *Resilient-Backpropagation* e *Levenberg-Marquardt*. Estes algoritmos de treinamento, conhecidos como “deep-learning”, apresentam dificuldades em treinar redes profundas, conhecidas como “deep-networks”. Além do tempo de treinamento tender a ser muito lento, pode acontecer o que a literatura chama de “problema dos gradientes explodindo ou desvanecendo”, pois quando as RNAs são muito profundas, a chance da regra da cadeia produzir uma multiplicação instável aumenta.

Assim, como consequência, a atualização dos parâmetros das camadas mais distantes da saída tendem a sofrer atualizações: instáveis; ou insignificantes, o que é mais comum; ou até mesmo atualizações muito grandes, o que é mais raro de se acontecer. Assim, embora a rede neural seja mais poderosa quanto mais profunda for, mais difícil se torna de ser treinada.

3 METODOLOGIA

Neste capítulo serão apresentados os detalhes das RNAs implementadas para a predição de 3 variáveis que compõem o PLD. São elas: Demanda Máxima, Energia Natural Afluyente por Bacia e Energia Natural Afluyente por Subsistema. Por se tratarem de dados dinâmicos variantes no tempo, as RNAs utilizadas foram do tipo TDNN (*Time Delay Neural Network*), as quais introduzem atraso no tempo.

As bases de dados utilizadas neste trabalho foram obtidas através do Operador Nacional do Sistema Elétrico (ONS). Para cada variável utilizada, os dados abrangem o período de 23/07/2016 à 09/11/2018, totalizando 120 semanas operativas. Para a variável Demanda Máxima, medida em MWh/h, foi selecionado o subsistema Sudeste/Centro-Oeste. Para a Energia Natural Afluyente por Bacia, foi utilizada a bacia Tietê e para a variável Energia Natural Afluyente por Subsistema foi escolhido o subsistema Sudeste/Centro-Oeste. Essas duas últimas são medidas em ENA Bruta (MWmed).

3.1 Treinamento

Nesta seção serão apresentados os detalhes da RNA implementada para a variável Demanda Máxima. A metodologia é a mesma para as demais variáveis, alterando apenas alguns parâmetros da rede, os quais serão detalhados nos resultados.

Em primeiro momento, os dados devem ser normalizados entre 0 e 1, pois, de acordo com o teorema da aproximação universal de Kolmogorov e de Cybenko [19], toda RNA converge com apenas uma camada escondida desde que os dados estejam contidos num hipercubo unitário.

Dos 120 dados normalizados, os 116 primeiros foram usados para treinamento e os 4 últimos para validação. Apenas um dado para validação seria suficiente, pois se deseja prever a próxima semana no futuro. Mas a fim de observar a tendência dos dados no decorrer de 4 semanas, foram utilizados 4 dados para validação. Assim, tendo sido normalizados os dados, os principais parâmetros da rede são escolha do projetista. Três dos principais parâmetros são o número de entradas da rede, a quantidade de camadas escondidas e o número de neurônios na camada escondida.

Para a aplicação específica deste trabalho, o número de entradas da rede foi estabelecido de forma empírica, sendo testadas, por exemplo, 30, 40, 45 e 50 entradas. Devido à normalização, uma camada escondida foi suficiente. Várias configurações de neurônios na camada escondida foram testados, como 30, 40, 50, 70, 100 e 120 neurônios, por exemplo. Já a camada de saída é composta por um neurônio, uma vez que se deseja prever uma variável por vez.

Após o ajuste destes parâmetros, inicia-se a fase de *backpropagation*, onde a rede de fato começa a ser treinada. Os pesos sinápticos da RNA são iniciados aleatoriamente para então serem ajustados. Após o ajuste, calcula-se a saída da rede (fase *forward*) e esta é comparada com o valor final desejado. Se o erro entre elas for menor do que o estipulado, significa que a rede convergiu, caso contrário, os pesos devem ser ajustados (fase *backward*) e o processo reiniciado.

O algoritmo mais comum de treinamento para redes TDNN é o *Backpropagation*, em que o ajuste dos pesos da rede é feito em relação à direção oposta do gradiente da função erro quadrático, ou seja, os valores de erro caminham para a direção de máximo decrescimento da função. No entanto, na prática, este algoritmo tende a convergir muito lentamente, o que foi o caso deste trabalho, exigindo um elevado esforço computacional. Para contornar este inconveniente, dentre as técnicas utilizadas, destaca-se o algoritmo de *Levenberg-Marquardt*.

Ao invés de fazer uso da descida do gradiente na função erro quadrático, o algoritmo de *Levenberg-Marquardt* é uma aproximação do método de Newton, que consegue conduzir o treinamento de redes PMC na ordem de 10 a 100 vezes mais rápido que o *Backpropagation*. No entanto, dependendo da aplicação, a rede pode não convergir, como também foi a realidade desta aplicação. Uma ferramenta alternativa é a inserção do termo de *momentum*, que se baseia na teoria de momento da física, com o objetivo de aumentar a rapidez com que o erro evolui para seu valor mínimo.

Após alguns testes de convergência, o método de treinamento escolhido para este trabalho foi um híbrido de *Levenberg-Marquardt* com *Momentum* e *Resilient Backpropagation*. Este último, ao invés de considerar as variações das magnitudes do gradiente da função erro, leva em conta somente a variação de seu sinal, o que faz com que a taxa de aprendizado do método se torne dinâmica.

Assim, com base nas Equações 1 e 2 do neurônio proposto por [13], uma descrição do equacionamento utilizado para o cálculo e ajuste dos pesos será detalhado a seguir.

Sendo x as entradas da rede, x_0 o limiar de ativação e w_1 e w_{01} os pesos da primeira camada, é calculado o potencial de ativação (u_1) e a saída da primeira camada (y_1) por meio do equacionamento a seguir:

$$u_1 = w_1x + w_{01}x_0 \quad (3)$$

$$y_1 = \tanh(\beta u_1) \quad (4)$$

A função de ativação escolhida foi a tangente hiperbólica, pois, por ser uma função ímpar, contribui para melhorar o processo de convergência da rede durante o treinamento. Na Equação 4, o termo β corresponde à inclinação da função e foi setado, por experimentação, com o valor de 0,4. Contendo este termo, a saída y_1 é utilizada como entrada da próxima camada da rede que, neste caso, é a camada de saída:

$$u_2 = w_2y_1 + w_{02}x_0 \quad (5)$$

$$y_2 = \tanh(\beta u_2) \quad (6)$$

A saída y_2 da rede deve ser comparada com a saída desejada d , saída esta que a rede deseja aprender, que é a série temporal da variável trabalhada. Para o cálculo do erro referente a esta comparação, utiliza-se o Erro Quadrático Médio (EQM), descrito pela Equação 7.

$$EQM = \frac{1}{2}(d - y_2)^2 \quad (7)$$

O ajuste dos pesos é feito, então, por meio do método híbrido criado, que mescla *Levenberg-Marquardt* com *Momentum* e *Resilient Backpropagation*. O processo é realizado de trás para frente (fase *backward*), ou seja, o ajuste nos pesos começa da última camada da rede. Para isto, calcula-se o gradiente do EQM com relação a w_2 e w_{02} através do emprego da regra da cadeia para derivadas parciais de função composta, como indicado nas Equações 8 e 9.

$$\frac{\partial EQM}{\partial w_2} = \frac{\partial EQM}{\partial y_2} \frac{\partial y_2}{\partial u_2} \frac{\partial u_2}{\partial w_2} \quad (8)$$

$$\frac{\partial EQM}{\partial w_{02}} = \frac{\partial EQM}{\partial y_2} \frac{\partial y_2}{\partial u_2} \frac{\partial u_2}{\partial w_{02}} \quad (9)$$

Assim, com algumas articulações matemáticas, é possível fazer o ajuste dos pesos através das Equações 10 e 11, onde w_2^t , w_2^{t-1} , w_{02}^t , w_{02}^{t-1} e α são parâmetros relativos ao momentum e n_{w_2} e $n_{w_{02}}$ são as matrizes das taxas de aprendizado do algoritmo. Assim, o novo peso recebe o valor do peso atual, mas considerando a diferença observada entre as respostas produzidas.

$$w_2^{t+1} = w_2^t + \alpha \left(w_2^t - w_2^{t-1} \right) - n_{w_2} \frac{dEQM_t}{dw_2} \quad (10)$$

$$w_{02}^{t+1} = w_{02}^t + \alpha \left(w_{02}^t - w_{02}^{t-1} \right) - n_{w_{02}} \frac{dEQM_t}{dw_{02}} \quad (11)$$

De *Levenberg-Marquardt* vem a taxa de *momentum* (α), cujo valor está compreendido entre 0 e 1. Quando a solução atual estiver longe da solução final, a variação na direção oposta ao gradiente da função erro quadrático entre duas iterações sucessivas será grande, então pode-se adicionar um passo maior de incremento para o peso em direção ao mínimo da função erro. Porém, quando a solução atual estiver próxima da solução final, a contribuição do termo de *momentum* é pequena.

Já do método *Resilient Backpropagation* vêm as taxas de aprendizado n_w , que são matrizes da mesma dimensão dos pesos. Estas taxas são ajustadas a cada iteração, conforme a inclinação do gradiente e, ao invés de considerar as variações das magnitudes do gradiente da função erro, leva em conta apenas a variação de seu sinal.

Desta forma, quando os sinais dos gradientes forem os mesmos, considerando duas iterações sucessivas, significa que se pode incrementar a taxa, pois o ponto mínimo da função erro ainda está distante. Porém, quando os sinais do gradiente forem diferentes, o ponto mínimo da função foi ultrapassado, então a taxa de aprendizado é reduzida para que o algoritmo possa convergir de forma suave. A Equação 12 mostra como se dá esta dinâmica.

$$\eta_w(i,j) = \begin{cases} \eta^+ \eta_w(i,j), & \text{se } \frac{\partial EQM_{k-1}}{\partial w} \frac{\partial EQM_k}{\partial w} > 0 \\ \eta^- \eta_w(i,j), & \text{se } \frac{\partial EQM_{k-1}}{\partial w} \frac{\partial EQM_k}{\partial w} < 0 \\ \eta_w(i,j), & \text{se } \frac{\partial EQM_{k-1}}{\partial w} \frac{\partial EQM_k}{\partial w} = 0 \end{cases} \quad (12)$$

Assim, com os pesos atualizados, são realizados os mesmos cálculos para a atualização dos pesos da camada anterior: w_1 , w_{01} .

A rede será considerada treinada quando a diferença entre as respostas produzidas por ela e as respostas desejadas estiver dentro de valores aceitáveis, escolhidos pelo projetista. Em conjunto com este erro, pode-se utilizar também o número de épocas como critério de parada. Assim, define-se um ponto de parada no processo de treinamento para casos onde a precisão se torna inatingível.

3.2 Validação

Após a primeira etapa, referente ao de treinamento da rede, é necessária uma segunda etapa onde os pesos obtidos na primeira serão verificados. Este processo consiste, basicamente, em aplicar a fase *forward* às entradas que não foram utilizadas no treinamento e submetê-las aos pesos obtidos, como mostrado nas Equações 3 a 12.

Nesta etapa é verificada a eficiência da rede quanto à sua capacidade de generalizar os dados que não foram utilizados na fase de treinamento, fazendo assim a predição destes valores no futuro.

4 RESULTADOS

Este capítulo tem o objetivo de relatar os resultados obtidos no treinamento e na validação de dados reais da Demanda Máxima, da Energia Natural Afluyente por Bacia e da Energia Natural Afluyente por Subsistema, além de detalhar o comportamento das RNAs para as respectivas variáveis.

4.1 Demanda Máxima

Nesta seção estão os resultados obtidos durante a fase de treinamento e de validação (predição) da variável Demanda Máxima. Na fase de treinamento foram feitos diversos testes quanto ao número de neurônios na camada de entrada (N) e na camada intermediária (N_1). Na camada de entrada, as configurações com 30, 40 e 45 neurônios, por exemplo, não obtiveram sucesso, enquanto com 50 neurônios os resultados foram melhores. Assim, nesta e nas próximas seções será abordada apenas a configuração com 50 neurônios na camada de entrada e três configurações da camada de saída, a fim de sintetizar os resultados. Como critério de parada foi utilizado o erro quadrático médio menor que 0,0005 (0,05%) e o número de épocas máximo de 5000 épocas.

As Figuras 6 a 11 mostram o comportamento da rede para $N = 50$ e $N_1 = 50, 100$ e 120 neurônios, respectivamente. E a Tabela 1 mostra os resultados quantitativos.

Contudo, antes de interpretar os resultados deve-se ter em mente 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.

Assim, é possível observar que com 50 neurônios na camada intermediária a rede aprende o comportamento dos dados sem alta precisão e não é capaz de prever o próximo passo no futuro, nem o comportamento durante 4 semanas, o que é comprovado também pelos EQMs e pela variância dos dados. Já com 100 neurônios, o treinamento é altamente assertivo e a predição, além de acompanhar a tendência dos dados, acerta a próxima semana no futuro, apresentando valores de EQMs e variância baixos. No entanto, como visto na Figura 11, uma rede sobredimensionada não significa melhores resultados, ainda que a fase de treinamento tenha um erro pequeno.

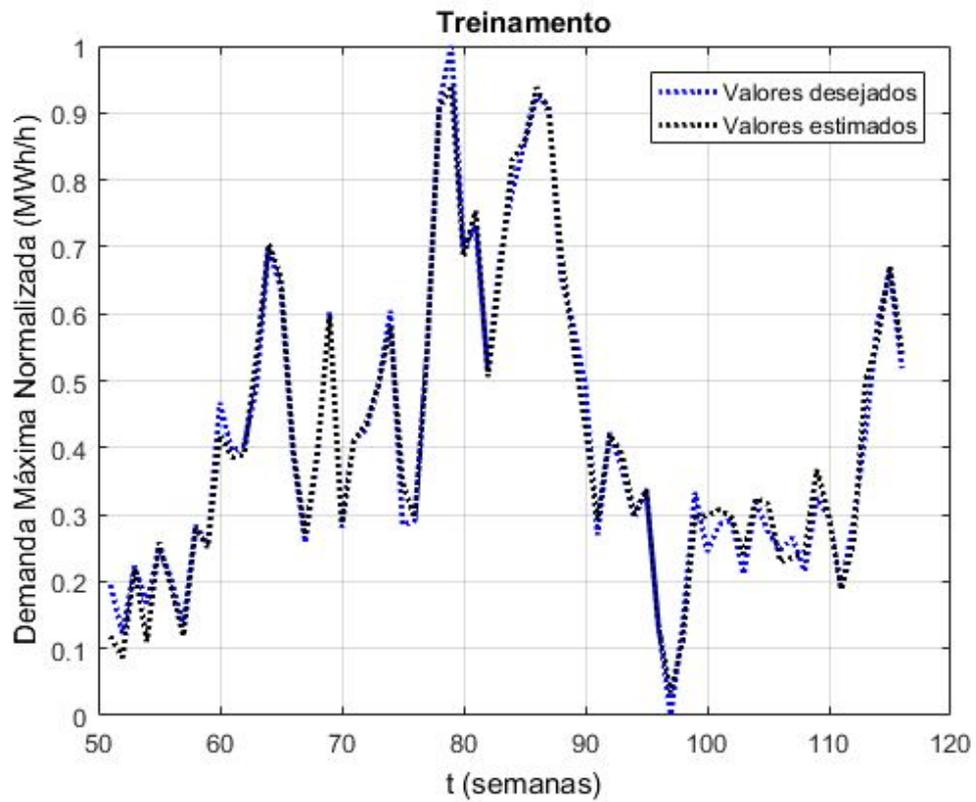


Figura 6 - Fase de treinamento para $N_1 = 50$ neurônios intermediários.

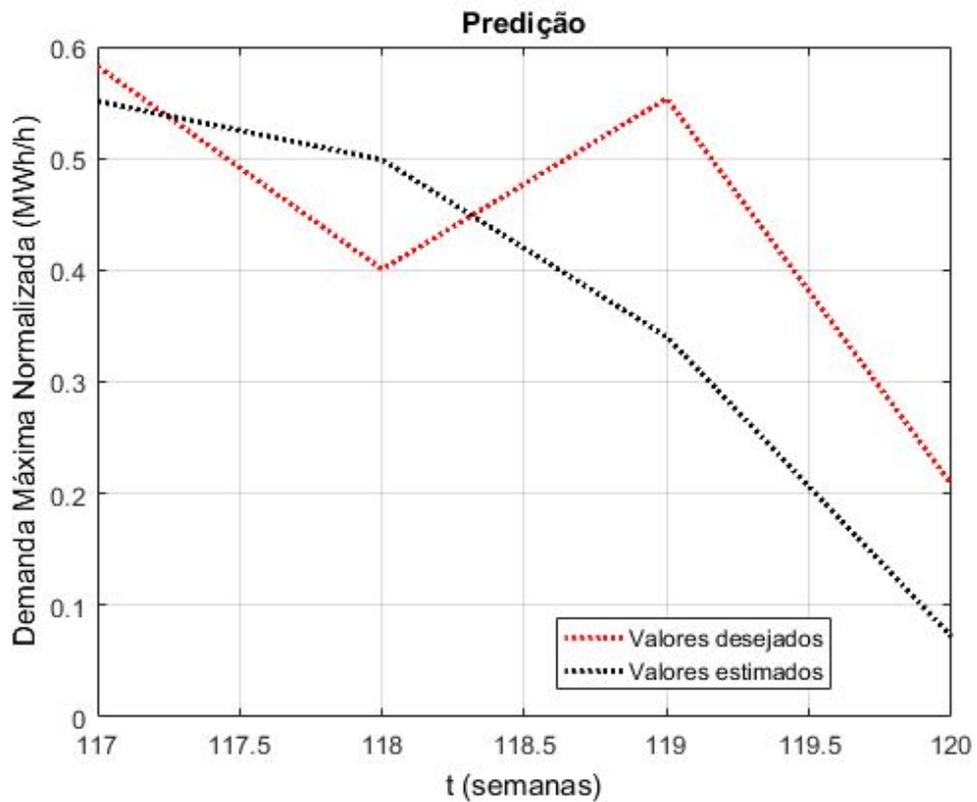


Figura 7 - Predição para $N_1 = 50$ neurônios intermediários.

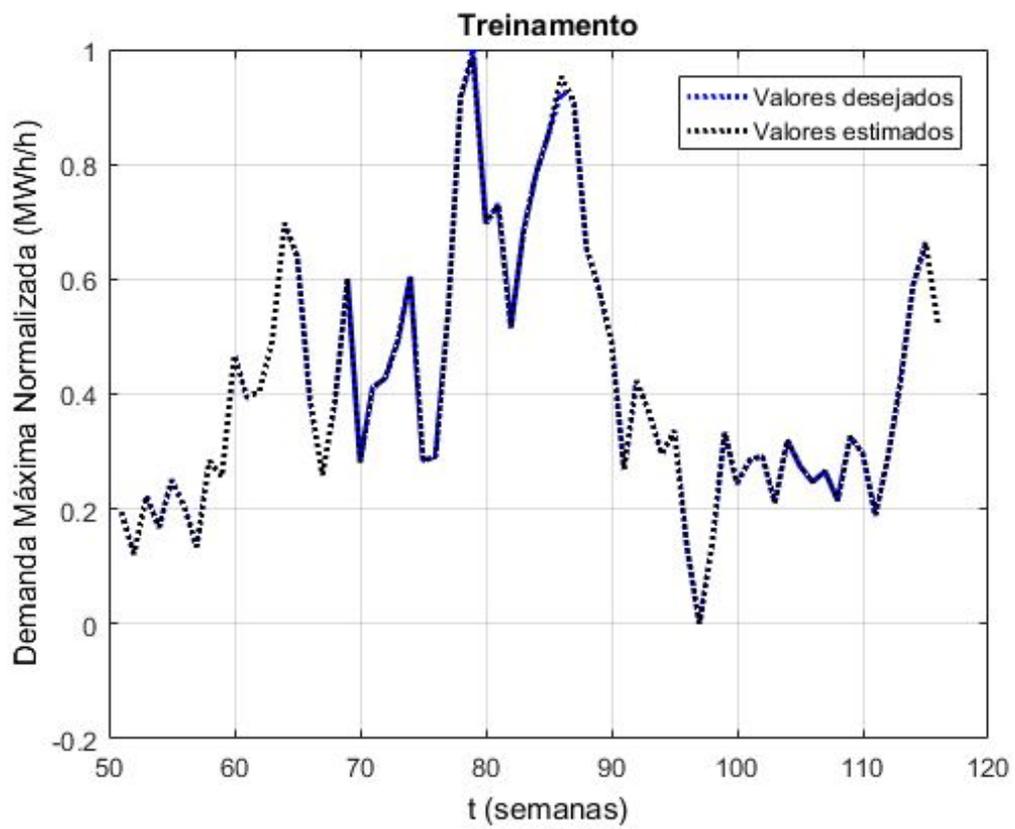


Figura 8 - Fase de treinamento para $N_1 = 100$ neurônios intermediários.

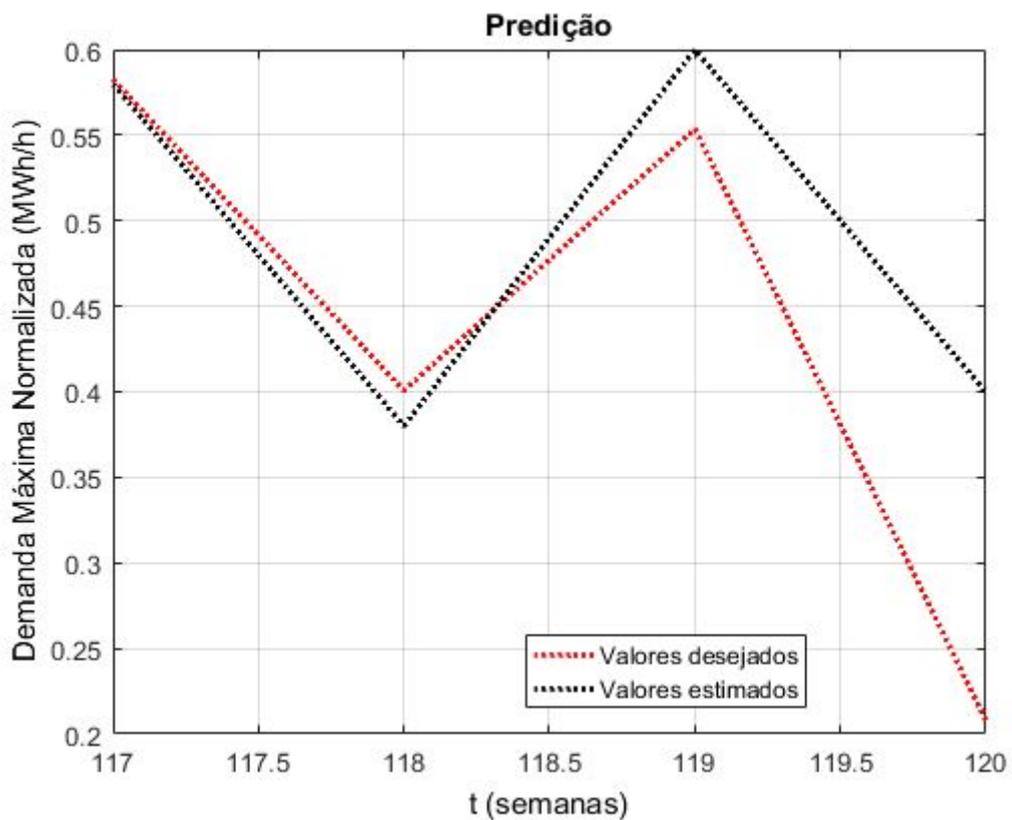


Figura 9 - Predição para $N_1 = 100$ neurônios intermediários.

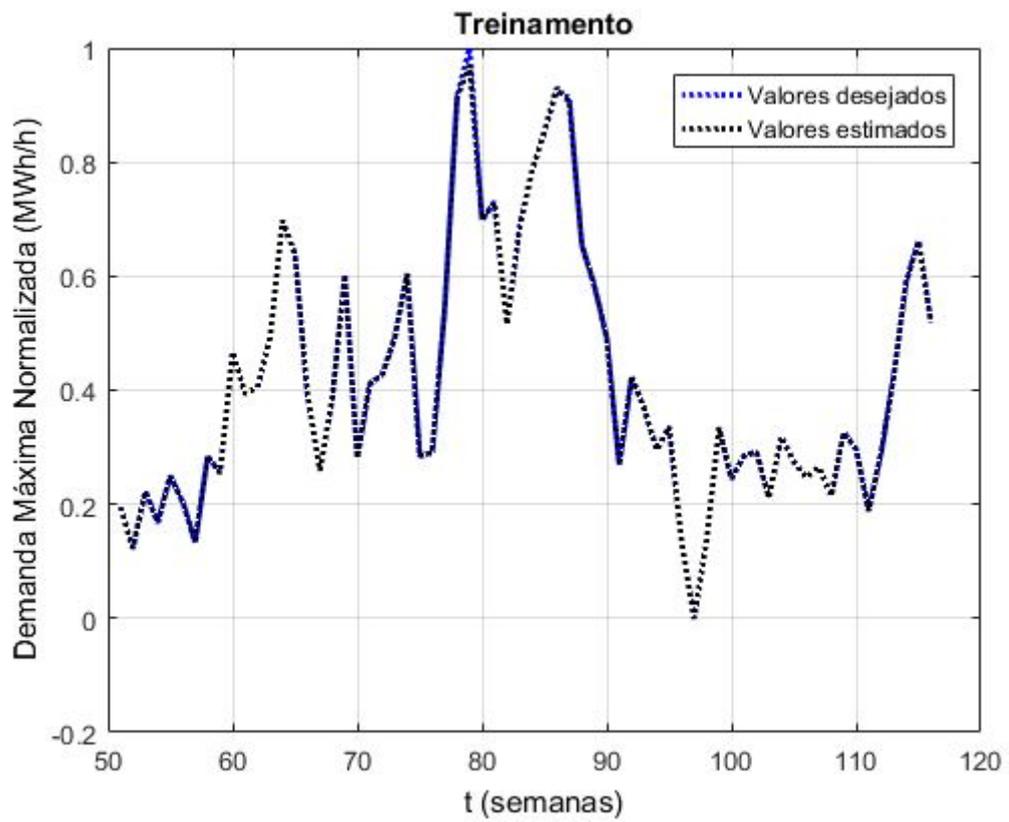


Figura 10 - Fase de treinamento para $N_1 = 120$ neurônios intermediários.

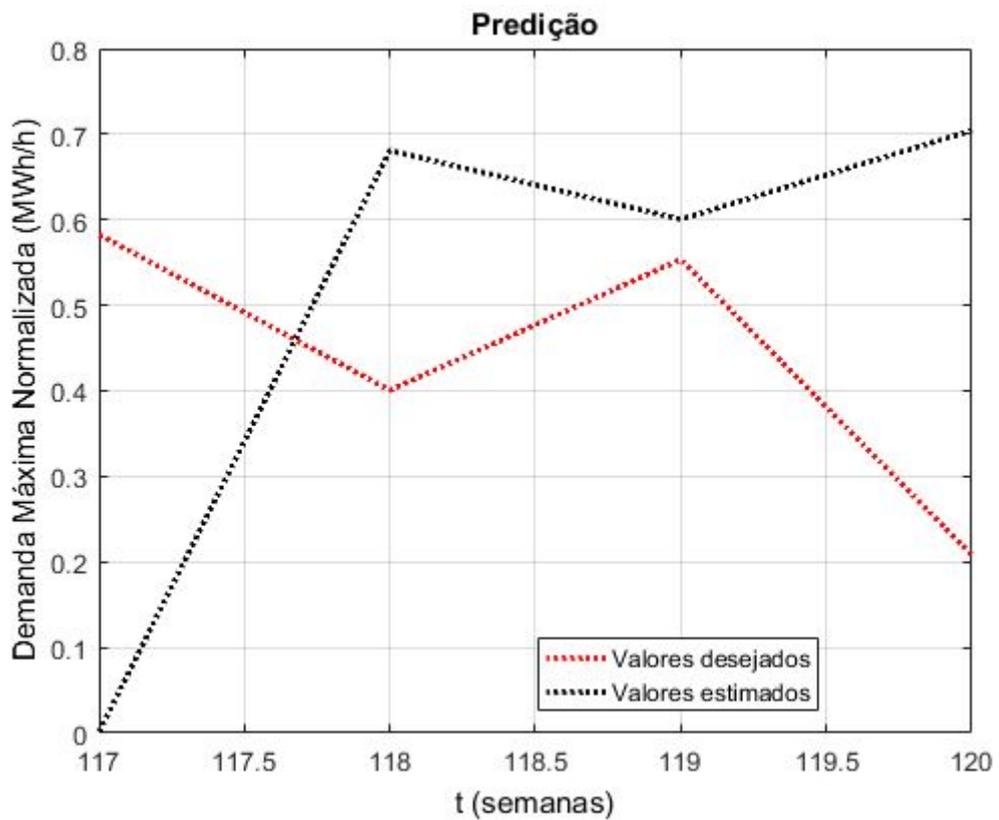


Figura 11 - Predição para $N_1 = 120$ neurônios intermediários.

Indicadores	Número de neurônios na camada intermediária		
	50 neurônios	100 neurônios	120 neurônios
Número de épocas	5000	2326	1214
EQM	0.028190	0.000500	0.000500
EQM da predição	0.037333	0.019725	0.332685
Variância da predição	0.034789	0.010100	0.083391

Tabela 1 - Indicadores para configurações de 50, 100 e 120 neurônios intermediários para a variável Demanda Máxima.

4.2 Energia Natural Afluyente por Bacia

Os resultados da variável Energia Natural Afluyente por Bacia são semelhantes aos da Demanda Máxima quanto ao retratar a perda da qualidade de aprendizado e de predição de redes subdimensionadas ou superdimensionadas em termos de números de neurônios nas camadas intermediárias.

As Figuras 12 a 17 mostram o comportamento da rede para $N = 50$ e $N_1 = 50, 100$ e 101 neurônios, respectivamente, e a Tabela 2 mostra estes resultados de forma quantitativa. Com 50 neurônios o treinamento é bom e a tendência dos dados na predição também, porém a rede não é capaz de predizer um passo no futuro, o que justifica o EQM da predição razoável. Com 100 neurônios na camada intermediária o treinamento é ainda melhor, e a predição, além de conseguir acertar a próxima semana, acerta também a segunda e a terceira semanas, indo além do esperado por esta rede, que era acompanhar a tendência dos dados futuros. No entanto, bastou apenas 1 neurônio a mais na camada intermediária para a rede não fazer uma boa predição, como na Figura 17.

Indicadores	Número de neurônios na camada intermediária		
	50 neurônios	100 neurônios	101 neurônios
Número de épocas	3975	1294	2276
EQM	0.000500	0.000500	0.000500
EQM da predição	0.007360	0.003695	0.029322
Variância da predição	0.001114	0.009350	0.001332

Tabela 2 - Indicadores para configurações de 50, 100 e 101 neurônios intermediários para a variável Energia Natural Afluyente por Bacia.

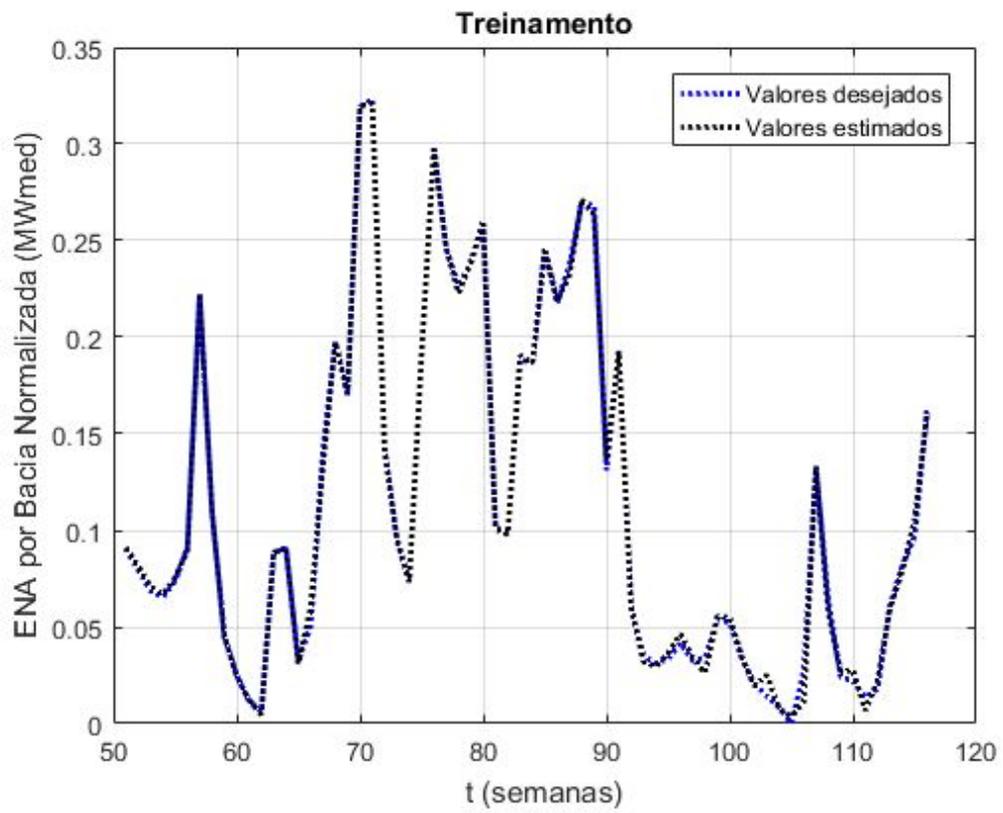


Figura 12 - Fase de treinamento para $N_1 = 50$ neurônios intermediários.

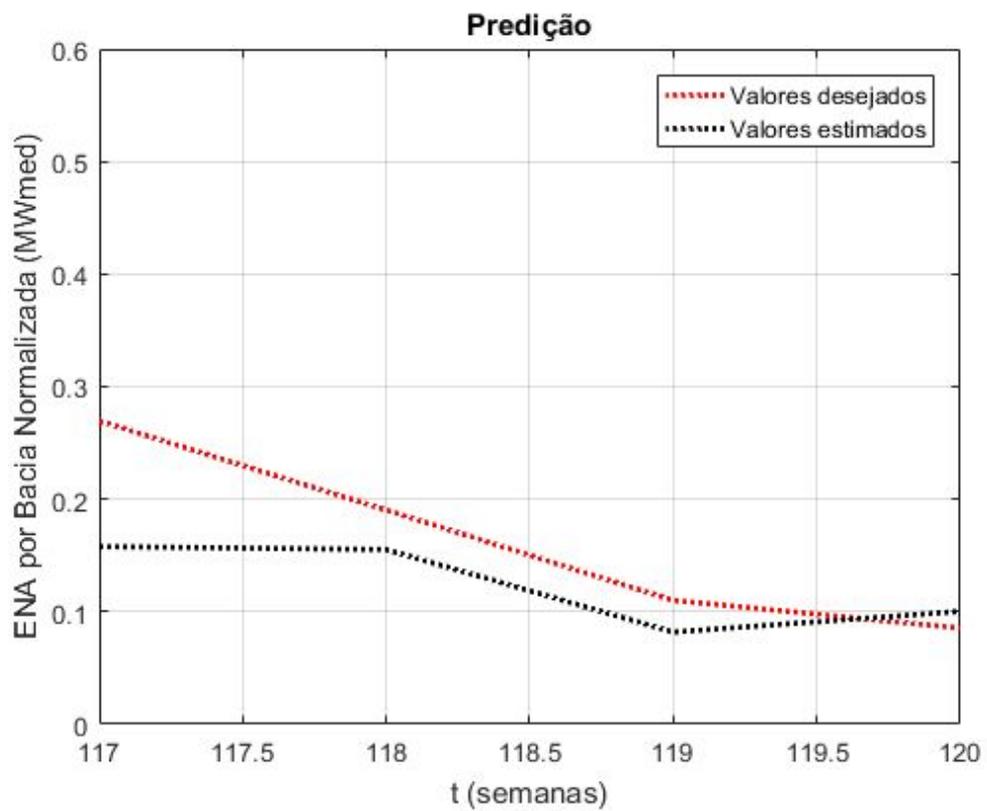


Figura 13 - Predição para $N_1 = 50$ neurônios intermediários.

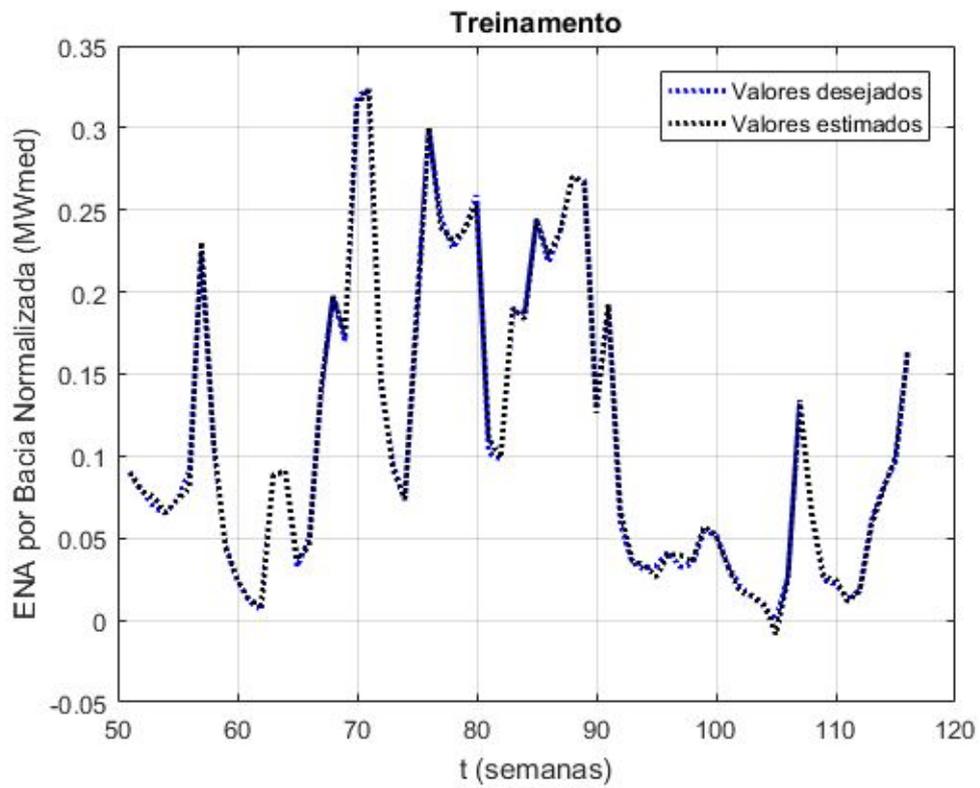


Figura 14 - Fase de treinamento para $N_1 = 100$ neurônios intermediários.

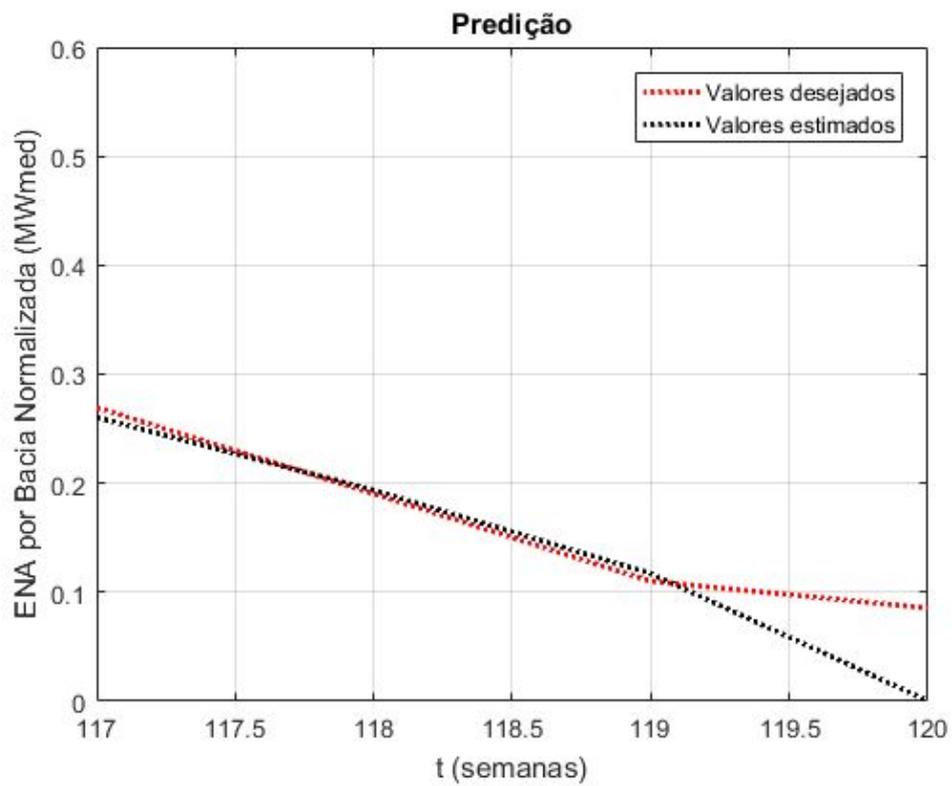


Figura 15 - Predição para $N_1 = 100$ neurônios intermediários.

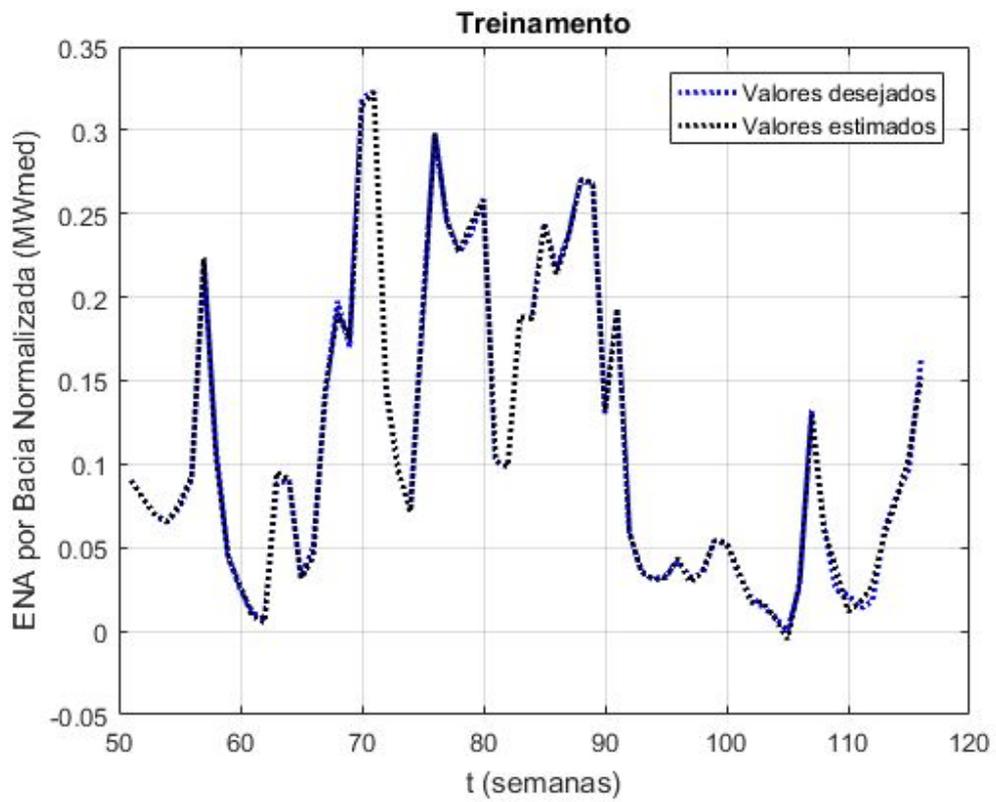


Figura 16 - Fase de treinamento para $N_1 = 101$ neurônios intermediários.

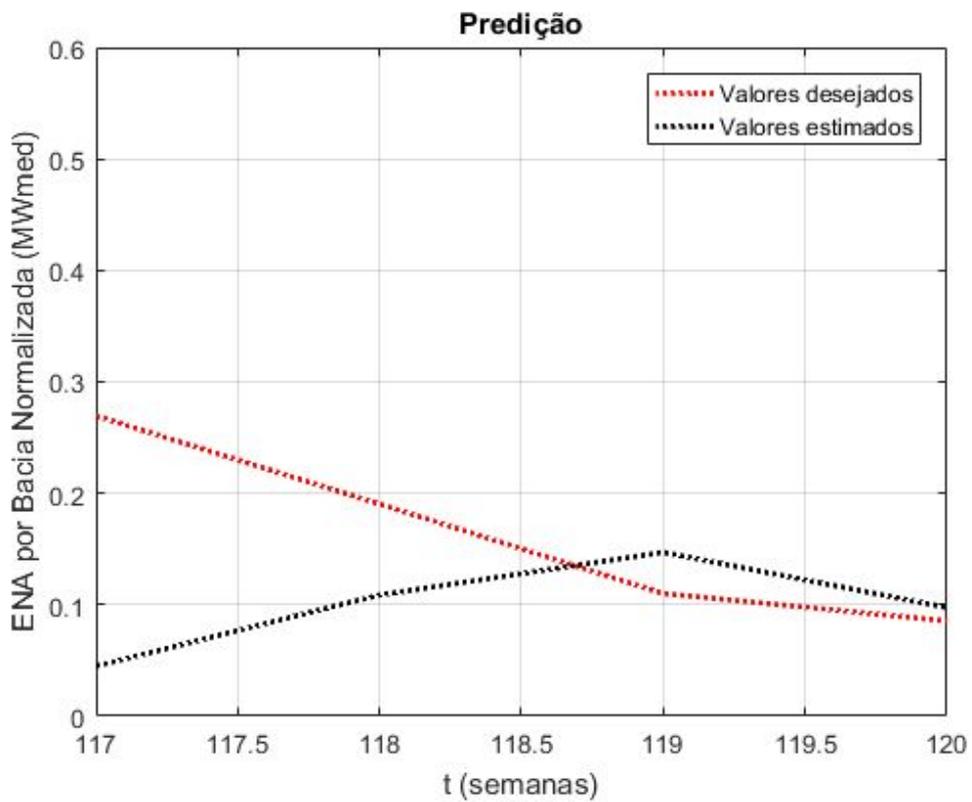


Figura 17 - Predição para $N_1 = 101$ neurônios intermediários.

4.3 Energia Natural Afluente por Subsistema

Já para a variável Energia Natural Afluente por Subsistema, referente ao subsistema Sudeste/Centro-Oeste, a rede não obteve bons resultados. Para as três configurações de neurônios na camada escondida, apesar de acompanhar o comportamento dos dados de treinamento, a rede não atingiu o erro máximo de 0,0005, sendo interrompida pelo número de épocas máximo pré-estabelecido, como pode ser observado na Tabela 3.

A tabela também mostra os altos valores de EQM da rede, quando comparado aos das outras duas variáveis abordadas neste trabalho. Quanto aos valores referentes à variância, o menor se dá com 100 neurônios na camada intermediária, pois, como é possível ver na Figura 21, a rede acompanha a tendência dos dados e, como a variância mede a dispersão de cada valor comparado com a média dos valores, seu resultado é baixo. No entanto, o objetivo de prever 1 semana no futuro não foi atingido.

Outros limites de época também foram testados para acompanhar o desenvolvimento do EQM, mas o comportamento dos resultados se manteve, o que fica evidente que a questão se baseia em encontrar os parâmetros corretos dentre os diversos que as RNAs possuem.

Indicadores	Número de neurônios na camada intermediária		
	50 neurônios	100 neurônios	120 neurônios
Número de épocas	5000	5000	5000
EQM	0.017215	0.028495	0.025300
EQM da predição	0.050038	0.013519	0.034032
Variância da predição	0.019728	0.002867	0.027003

Tabela 3 - Indicadores para configurações de 50, 100 e 120 neurônios intermediários para a variável Energia Natural Afluente por Subsistema.

As Figuras 18 a 23 mostram o comportamento da rede para $N = 50$ e $N_1 = 50, 100$ e 120 neurônios, respectivamente.

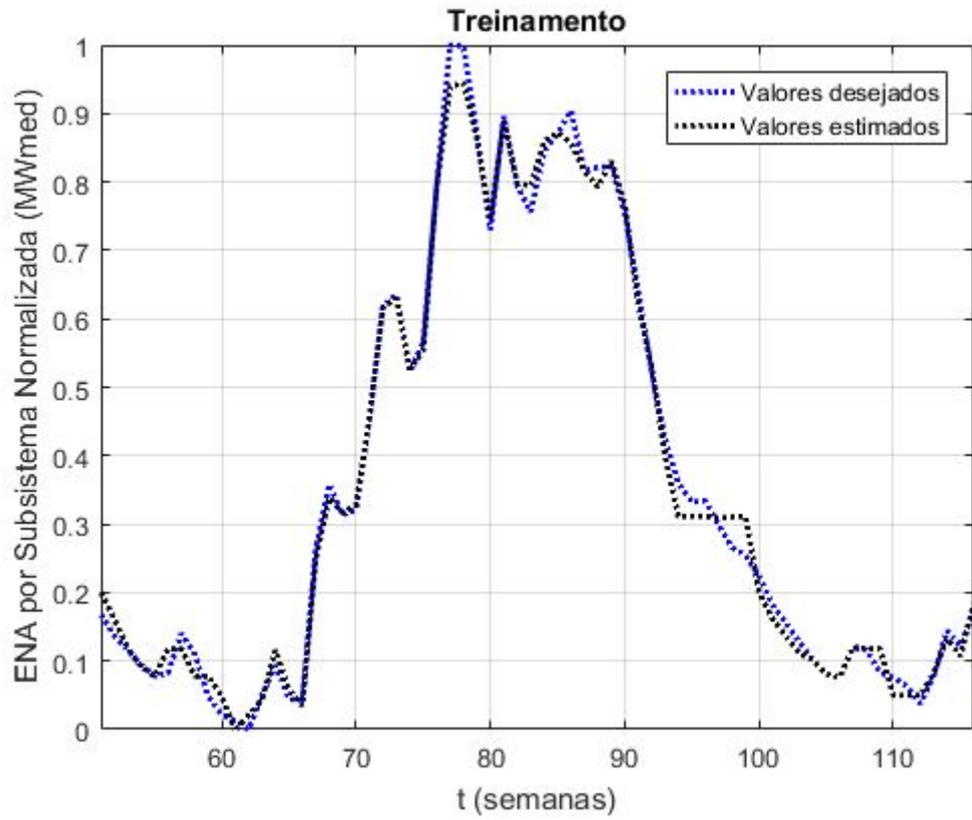


Figura 18 - Fase de treinamento para $N_1 = 50$ neurônios intermediários.

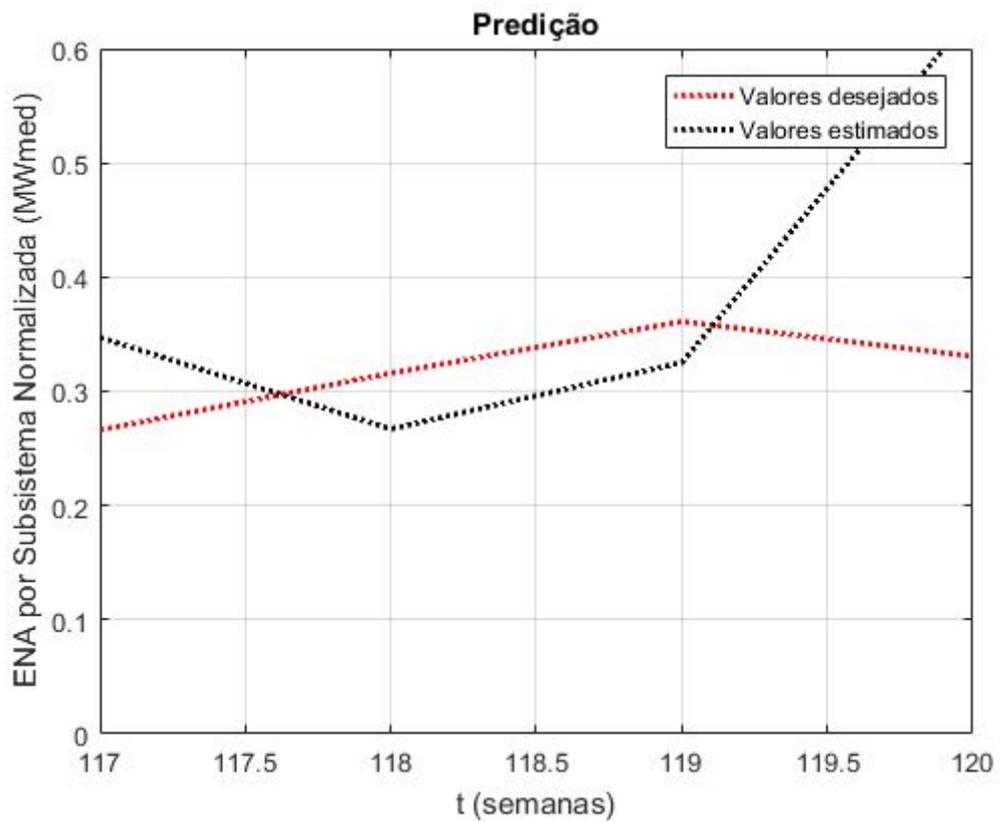


Figura 19 - Predição para $N_1 = 50$ neurônios intermediários.

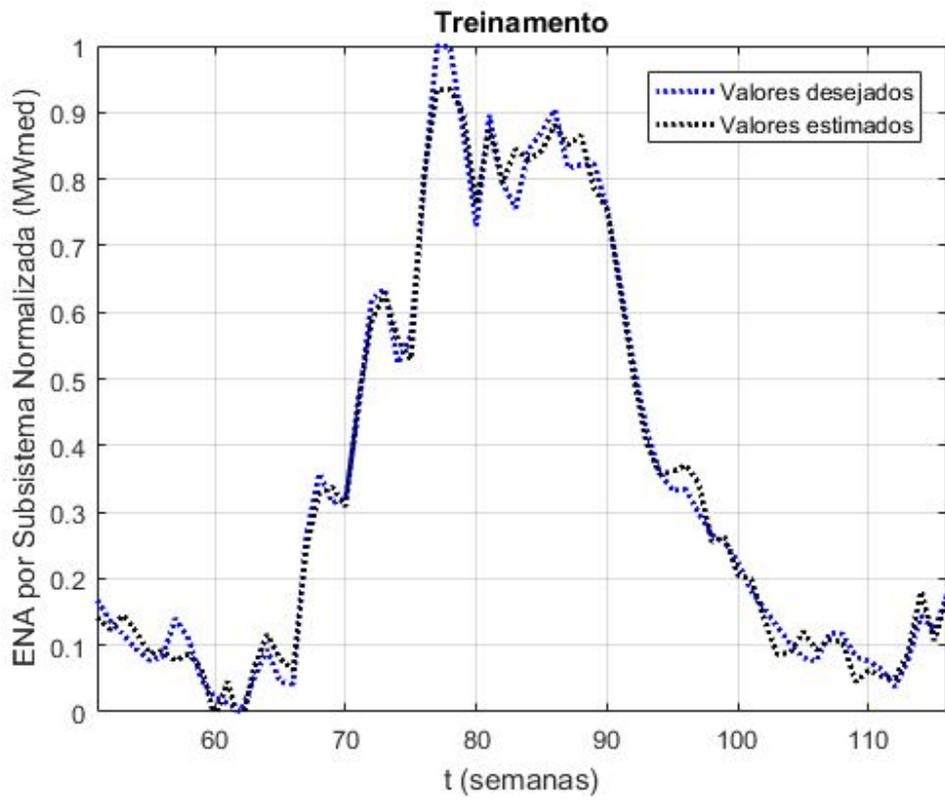


Figura 20 - Fase de treinamento para $N_1 = 100$ neurônios intermediários.

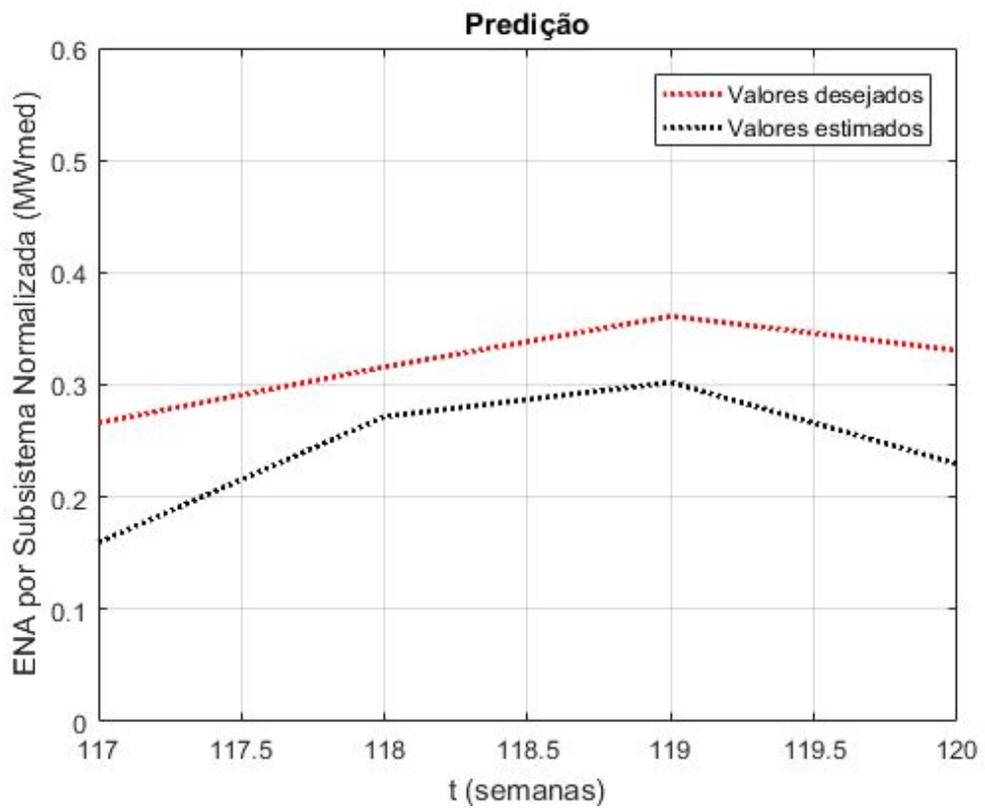


Figura 21 - Predição para $N_1 = 100$ neurônios intermediários.

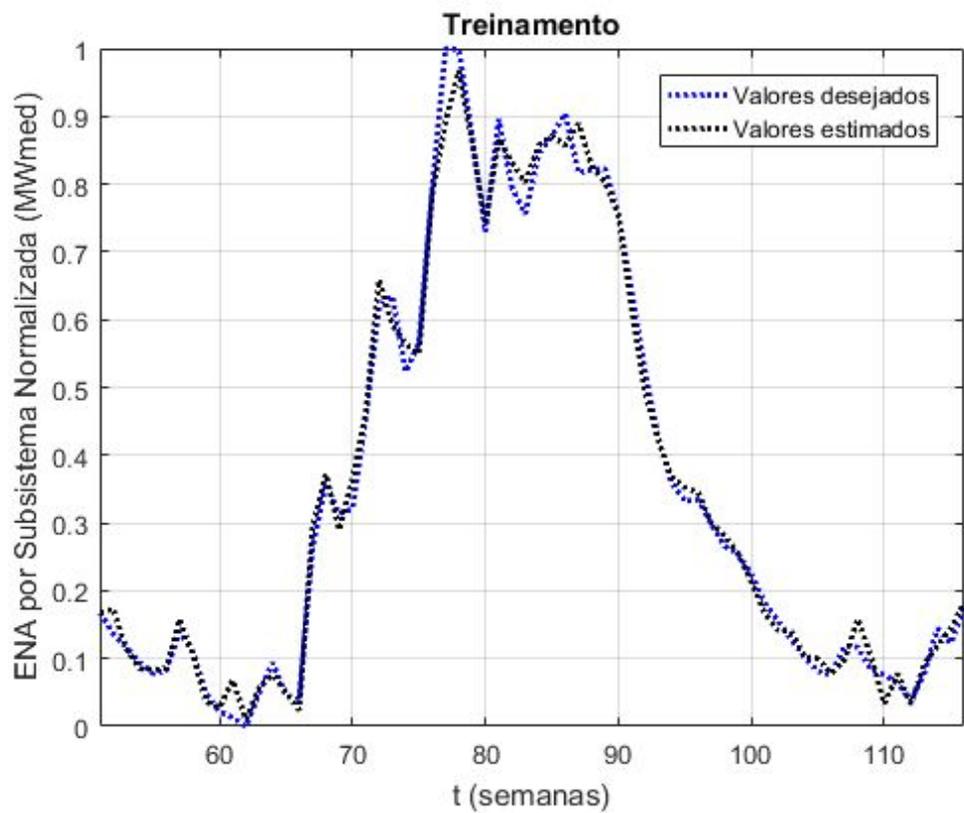


Figura 22 - Fase de treinamento para $N_1 = 120$ neurônios intermediários.

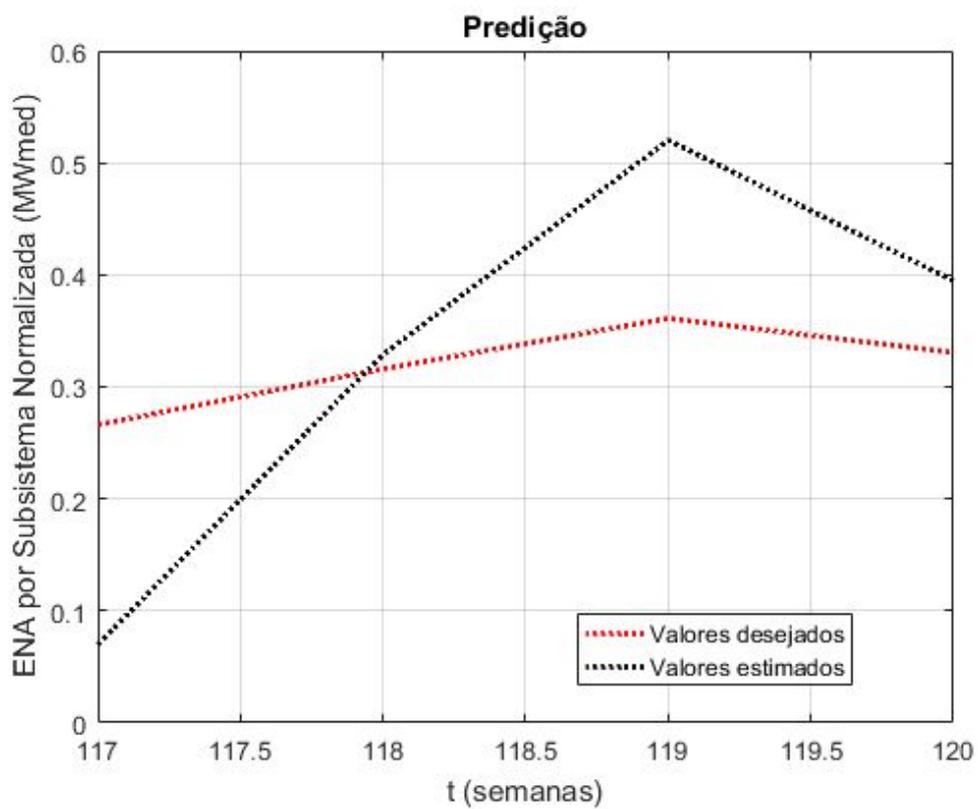


Figura 23 - Predição para $N_1 = 120$ neurônios intermediários.

5 CONCLUSÕES E TRABALHOS FUTUROS

5.1 Conclusões

Inicialmente esta monografia apresentou diversos conceitos sobre o Setor Elétrico Brasileiro e contextualizou o leitor sobre a comercialização de energia no mercado livre de curto prazo, considerando também a importância não só da predição do PLD, mas a predição das variáveis que o compõem. Três variáveis escolhidas para predição foram a Demanda Máxima horária do subsistema Sudeste/Centro-Oeste, a Energia Natural Aflente por Bacia da bacia Tietê e a Energia Natural Aflente por Subsistema do subsistema Sudeste/Centro-Oeste.

Além disso, foram apresentados conceitos técnicos sobre Redes Neurais Artificiais, como as topologias mais utilizadas nas mais variadas aplicações e os algoritmos de treinamento de neurônios artificiais. Em seguida apresentou-se a modelagem matemática da rede específica utilizada, a rede com atrasos temporais, do tipo TDNN. Para as duas primeiras variáveis citadas, as fases de treinamento para todas as configurações de neurônios na camada intermediária foram excelentes, atingindo o mínimo EQM estabelecido de 0.0005 sem que o número máximo de épocas estipulado fosse alcançado. Já o treinamento da terceira variável foi interrompido pelo número de épocas máximo, pois não convergia para o EQM setado.

Quanto à validação dos resultados, a predição em si, as configurações que mais obtiveram sucesso, para todas as três variáveis, foram com 50 neurônios na camada de entrada e 100 neurônios na camada intermediária. As Tabelas 1, 2 e 3 mostraram de forma quantitativa que redes subdimensionadas ou superdimensionadas em termos de números de neurônios nas camadas intermediárias tendem a perder a qualidade de aprendizado e também de predição.

No entanto, a terceira variável não foi capaz de manter o desempenho para nenhuma das três configurações abordadas, apresentando valores mais altos de EQM e variância. Apesar da configuração com 100 neurônios acompanhar a tendência dos dados durante 4 semanas, não foi capaz de prever a primeira semana no futuro, não cumprindo o objetivo da rede, que era prever a próxima semana e seguir a tendência durante 4 passos no futuro.

Assim, conclui-se que resultados ótimos são função da escolha adequada dos diversos parâmetros da RNA. Como para as duas primeiras variáveis a rede obteve sucesso, fica

evidente que para a terceira o problema foi a dificuldade na escolha dos parâmetros apropriados. Para contornar esta situação, pode-se investir em uma busca inteligente dos parâmetros em detrimento de uma busca empírica.

5.2 Trabalhos Futuros

Assim, para trabalhos futuros, podem ser realizadas melhorias na busca por parâmetros ótimos, optando por técnicas inteligentes ao invés de uma busca empírica. Contornado isto, seguindo a linha de pesquisa desenvolvida, os tópicos a seguir apresentam propostas de desenvolvimentos futuros:

- Investir em outra topologia de RNA, como redes recorrentes tipo Elman, para a predição das variáveis mais maior capacidade de generalização para períodos mais longos que 4 semanas no futuro;
- Mesclar técnicas de redes neurais com outras técnicas de inteligência artificial, como algoritmos genéticos e sistemas fuzzy;
- Investir em novas maneiras de treinamento que sejam mais robustas à presença de mínimos locais;
- Utilizar as saídas das redes desenvolvidas como entradas de outra RNA juntamente com dados históricos do PLD para a predição do mesmo de forma eficaz.

REFERÊNCIAS

- [1] NASCIMENTO, Fabiane Barbosa. “*Redes Neurais Artificiais Aplicadas à Predição do Preço de Liquidação das Diferenças no Mercado de Energia*”. Trabalho de Conclusão de Curso, Faculdade de Engenharia, Universidade Federal de Juiz de Fora, UFJF 2017.
- [2] Primeiros Passos sobre o Setor Elétrico e a CCEE. Capacita CCEE, Portal de Aprendizado. Disponível em <<http://ccee.micropower.com.br/Performa/Web/Portal/Main/Home.aspx>>. Acesso em 08/10/2018.
- [3] PREÇOS. Câmara de Comercialização de Energia Elétrica (CCEE). Disponível em <https://www.ccee.org.br/portal/faces/oquefazemos_menu_lateral/precos?_adf.ctrl-state=1c66z9bn7u_1&_afLoop=10013224718894#!%40%40%3F_afLoop%3D10013224718894%26_adf.ctrl-state%3D1c66z9bn7u_5>. Acesso em 08/10/2018.
- [4] SANTOS, Felipe Marques. “*Sistema Elétrico Brasileiro: Histórico, Estrutura e Análise de Investimentos no Setor*”. Trabalho de Conclusão de Curso, Universidade Federal de Santa Catarina - Campus Araranguá, UFSC 2015.
- [5] DE ALMEIDA, José Álvaro Jardim. “*P&D no Setor Elétrico Brasileiro: Um Estudo de Caso na Companhia Hidroelétrica do São Francisco*”. Dissertação (Mestre em Economia), Universidade Federal de Pernambuco, UFPE 2008.
- [6] Brasil, Lei nº 9.074, de 7 de julho de 1995.
- [7] 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: 15/10/18.
- [8] *O que é o PLD*. MERCADO LIVRE DE ENERGIA - GRUGREEN. Disponível em <<http://grugreen.eng.br/o-que-e-o-pld-preco-de-liquidacao-das-diferencas/>>. Acesso em 22/10/2018.
- [9] 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>. Acesso em: 15/10/18.
- [10] HISTÓRICO DA OPERAÇÃO. Operador Nacional do Sistema Elétrico (ONS). Disponível em <<http://ons.org.br/paginas/resultados-da-operacao/historico-da-operacao>>. Acesso em 08/10/2018.
- [11] 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.

- [12] DA SILVA, Ivan Nunes; SPATTI, Danilo Hernane; FLAUZINO, Rogério Andrade. “*Redes Neurais Artificiais: para engenharia e ciências aplicadas*”. Editora Artliber, 2010.
- [13] 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.
- [14] HEBB, Donald Olding. *The Organization Of Behavior A Neuropsychological Theory*. John Wiley And Sons Inc; New York; Chapman And Hall Limited; London, 2013.
- [15] ROSENBLATT, Frank. *The perceptron: A probabilistic model for information storage and organization in the brain*. *Psychological review*, 1958.
- [16] WIDROW, Bernard. *Generalization and information storage in network of Adaline 'neurons'*. *Self-organizing systems-1962*, 1962.
- [17] MINSKY, Marvin e PAPERT, Seymour. *Perceptrons*, MIT Press, Cambridge, MA, 1969.
- [18] RUMELHART, D. E, MCCLELLAND, J. L., *Parallel Distributed Processing*. MT Press, Vol.1, 1986.
- [19] CYBENKO, G., *Approximation by Superpositions of a Sigmoidal Function*. *Math. Control Signals Systems*, 1989.
- [20] UOL EDUCAÇÃO. Disponível em <<https://educacao.uol.com.br/disciplinas/matematica/desvio-padrao-entenda-este-conceito-de-estatistica.html>>. Acesso em 20/11/2018.