

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
CAMPUS CORNÉLIO PROCÓPIO
DIRETORIA DE GRADUAÇÃO E EDUCAÇÃO PROFISSIONAL
CURSO SUPERIOR DE ENGENHARIA ELÉTRICA

LUCAS MOSCARDINI AMARAL

**ESTIMAÇÃO DO PREÇO DE LIQUIDAÇÃO DAS DIFERENÇAS NO
MERCADO DE ENERGIA ELÉTRICA DE CURTO PRAZO
UTILIZANDO REDES NEURAS ARTIFICIAIS**

TRABALHO DE CONCLUSÃO DE CURSO

CORNÉLIO PROCÓPIO

2019

LUCAS MOSCARDINI AMARAL

**ESTIMAÇÃO DO PREÇO DE LIQUIDAÇÃO DAS DIFERENÇAS NO
MERCADO DE ENERGIA ELÉTRICA DE CURTO PRAZO
UTILIZANDO REDES NEURAS ARTIFICIAIS**

Trabalho de conclusão de curso de graduação, apresentado à disciplina Trabalho de Conclusão de Curso, do curso de Engenharia Elétrica da Universidade Tecnológica Federal do Paraná – UTFPR, como requisito para a obtenção do título de Engenheiro Eletricista.

Orientador: Prof. Dr. Edson Aparecido Rozas Theodoro

Coorientadora: Profa. Dra. Gabriela Helena Bauab Shiguemoto

CORNÉLIO PROCÓPIO

2019



Universidade Tecnológica Federal do Paraná
Campus Cornélio Procópio
Departamento Acadêmico de Elétrica
Curso de Engenharia Elétrica



FOLHA DE APROVAÇÃO

Lucas Moscardini Amaral

**Estimação do preço de liquidação das diferenças no mercado de energia elétrica de curto prazo
utilizando redes neurais artificiais**

Trabalho de conclusão de curso apresentado às 17:30hs do dia 26/06/2019 como requisito parcial para a obtenção do título de Engenheiro Eletricista no programa de Graduação em Engenharia Elétrica da Universidade Tecnológica Federal do Paraná. O candidato foi arguido pela Banca Avaliadora composta pelos professores abaixo assinados. Após deliberação, a Banca Avaliadora considerou o trabalho aprovado.

Prof(a). Dr(a). Edson Aparecido Rozas Theodoro - Presidente (Orientador)

Prof(a). Dr(a). Gabriela Helena Bauab Shiguemoto - (Coorientador)

Prof(a). Dr(a). Danilo Hernane Spatti - (Membro)

Prof(a). Dr(a). Murilo da Silva - (Membro)

RESUMO

AMARAL, Lucas M. **Estimação do preço de liquidação das diferenças no mercado de energia elétrica de curto prazo utilizando redes neurais artificiais**. 2019. 59 f. Trabalho de conclusão de curso (Graduação) – Engenharia Elétrica. Universidade Tecnológica Federal do Paraná. Cornélio Procópio, 2019.

Após a reestruturação dos modelos e regras de comercialização de energia elétrica no Brasil, surgiu uma nova forma de negociar a compra e venda de energia elétrica no mercado de curto prazo com base no preço de liquidação das diferenças (PLD). Devido as características da matriz energética brasileira, que é na sua maioria composta por geração hidráulica, o PLD tende a ser um valor volátil devido as variações nas condições hidrológicas, condições de demanda, custo de déficit e preço dos combustíveis usados nas termelétricas. Neste trabalho o PLD será estimado através de uma rede neural artificial utilizando séries históricas de diversas variáveis que compõem ou influenciam a formação do preço de energia elétrica no mercado de curto prazo. Esse trabalho vai utilizar no processo de estimação as redes neurais artificiais *Time Delay Neural Network* (TDNN) e *Perceptron* multicamadas (PMC), também será utilizado o coeficiente de Pearson como proposta para a seleção das variáveis de entrada na rede PMC.

Palavras-chave: Preço de liquidação das diferenças (PLD), mercado de curto prazo, estimação, rede neural artificial, *Time Delay Neural Network* (TDNN), *Perceptron* multicamadas (PMC), coeficiente de Pearson.

ABSTRACT

AMARAL, Lucas M. **Estimation of the differences settlement price in short-term electricity market using artificial neural networks**. 2019. 59 p. Graduation Work - Electrical Engineering. Federal Technological University of Paraná. Cornélio Procópio, 2019.

After the restructuring of the commercialization models and rules of electric energy in Brazil, a new way of negotiating electricity purchase and sale in the short term market based on the differences settlement price (PLD).

Due to the characteristics of the Brazilian energy matrix, which is mostly composed of hydraulic generation, the PLD tends to be a volatile value due to variations in hydrological conditions, demand requirements, cost of deficit and price of fuels used in thermoelectric plants. In this work the PLD will be estimated through an artificial neural network using historical series of several variables that make up or influence the formation of electric energy price in short term market. This work will use two artificial neural networks in the estimation process, the Time Delay Neural Network (TDNN) and the Multilayer Perceptron (MLP), also the Pearson coefficient will be used as a proposal for the selection of input variables in the MLP network.

Key words: *Preço de liquidação das diferenças* (PLD), short-term market, estimation, Artificial Neural Network (ANN), Temporal Delay Neural Network (TDNN), multilayer Perceptron (MLP), Pearson's coefficient.

LISTA DE FIGURAS

FIGURA 1 – VARIAÇÃO DO PLD DO SUBMERCADO SE/CO (2014-2015)	11
FIGURA 2 – MINIMIZAÇÃO DO CUSTO TOTAL DE OPERAÇÃO	15
FIGURA 3 – ENERGIA COMERCIALIZADA NO MERC. DE CURTO PRAZO	16
FIGURA 4 – REDE <i>FEEDFORWARD PERCEPTRON</i> MULTICAMANDAS	19
FIGURA 5 – TOPOLOGIA TDNN COM CONF. FOCADA	21
FIGURA 6 – FLUXOGRAMA TREINAMENTO TDNN	22
FIGURA 7 – FLUXOGRAMA TREINAMENTO PMC DE UMA CAMADA	23
FIGURA 8 – FLUXOGRAMA TREINAMENTO PMC DE DUAS CAMADAS	24
FIGURA 9 – ESCALA DE CORRELAÇÃO POR COEF. DE PEARSON	25
FIGURA 10 – EXEMPLO DE ASSOCIAÇÃO LINEAR	26
FIGURA 11 – PLD MÉDIO MENSAL SE-CO	29
FIGURA 12 – PLD MÉDIO MENSAL SE-CO DE 2016 (SUBCONJ. DE TESTES)...	29
QUADRO 1 – VARIÁVEIS DE CADA SUBMERCADO	31
FIGURA 13 – COMP. VALOR REAL DO PLD E APROX. DA REDE TDNN	38
FIGURA 14 – COMPARAÇÃO PMC (CÓD. 41)	40
FIGURA 15 – COMPARAÇÃO PMC (CÓD. 42)	41
FIGURA 16 – COMPARAÇÃO PMC (CÓD. 47)	44
FIGURA 17 – COMPARAÇÃO PMC (CÓD. 52)	46
FIGURA 18 – COMPARAÇÃO PMC (CÓD. 57)	48
FIGURA 19 – COMPARAÇÃO PMC (CÓD. 62)	50
FIGURA 20 – COMPARAÇÃO PMC (CÓD. 68)	52
FIGURA 21 – MELHOR RESULTADO ENCONTRADO (CÓD. 52)	55

LISTA DE TABELAS

TABELA 1 – PLD MÉDIO MENSAL DO SUBMERCADO SE-CO.....	28
TABELA 2 – PLD MÉDIO MENSAL DO SUB. SE-CO 2016.....	28
TABELA 3 – TREINAMENTOS REALIZADOS COM A REDE TDNN.....	32
TABELA 4 – TREIN. REDE PMC COM UMA CAMADA INTERMEDIÁRIA.....	33
TABELA 5 – TREIN. REDE PMC COM DUAS CAMADAS INTERMEDIÁRIAS.....	33
TABELA 6 – COEF. DE PEARSON ENTRE VAR. DE ENT. E PLD DO SE-CO.....	34
TABELA 7 – TREIN. PMC DE DUAS C. COM SELEÇÃO PRÉVIA DE VARIÁVEIS....	35
TABELA 8 – TREIN. PMC DE UMA C. COM SELEÇÃO PRÉVIA DE VARIÁVEIS....	36
TABELA 9 – RES. OBTIDOS NO TREIN. DA REDE TDNN (CÓD. 01 A 40).....	37
TABELA 10 – RESULTADOS TREINAMENTO CÓD. 09.....	39
TABELA 11 – RESULTADO OBTIDO NO TREINAMENTO DA REDE PMC DE UMA CAMADA (CÓD. 41).....	39
TABELA 12 – RESULTADOS TREINAMENTO CÓD. 41.....	40
TABELA 13 – RESULTADO OBTIDO NO TREINAMENTO DA REDE PMC DE DUAS CAMADAS (CÓD. 42 A 46).....	41
TABELA 14 – RESULTADOS TREINAMENTO CÓD. 42.....	42
TABELA 15 – VAR. COM MÓDULO DO COEF. DE PEARSON MAIOR QUE 0,4....	42
TABELA 16 – RESULTADO OBTIDO NO TREINAMENTO DA REDE PMC DE DUAS CAMADAS (CÓD. 47 A 51).....	43
TABELA 17 – RESULTADOS TREINAMENTO CÓD. 47.....	44
TABELA 18 – VAR. COM MÓDULO DO COEF. DE PEARSON MAIOR QUE 0,5....	45
TABELA 19 – RESULTADO OBTIDO NO TREINAMENTO DA REDE PMC DE DUAS CAMADAS (CÓD. 52 A 56).....	45
TABELA 20 – RESULTADOS TREINAMENTO CÓD. 52.....	46
TABELA 21 – VAR. COM MÓD. DO COEF. DE PEARSON MAIOR QUE 0,6.....	47
TABELA 22 – RESULTADO OBTIDO NO TREINAMENTO DA REDE PMC DE DUAS CAMADAS (CÓD. 57 A 61).....	47
TABELA 23 – RESULTADOS TREINAMENTO CÓD. 57.....	48
TABELA 24 – VAR. COM MÓD. DO COEF. DE PEARSON MAIOR QUE 0,7.....	49
TABELA 25 – RES. OBTIDO NO TREINAMENTO DA REDE PMC DE DUAS CAMADAS (CÓD. 62 A 66).....	49

TABELA 26 – RESULTADOS TREINAMENTO CÓD. 62.....	50
TABELA 27 – RESULTADO OBTIDO NO TREINAMENTO DA REDE PMC DE UMA CAMADA (CÓD. 67 A 70).....	51
TABELA 28– RESULTADOS TREINAMENTO CÓD. 68.....	52
TABELA 29 – MELHOR RESULTADO ENCONTRADO PARA CADA TOPOLOGIA..	53
TABELA 30 – MENORES E MAIORES ERROS MENSAIS E RES. MÉDIOS.....	54

SUMÁRIO

1 INTRODUÇÃO.....	10
1.1 OBJETIVOS.....	12
2 REVISÃO DA LITERATURA.....	13
2.1 PRECIFICAÇÃO DA ENERGIA ELÉTRICA	13
2.2 MERCADO DE CURTO PRAZO	15
2.3 PREÇO DE LIQUIDAÇÃO DAS DIFERENÇAS.....	16
2.4 REDES NEURAS ARTIFICIAIS.....	18
2.4.1 Arquitetura.....	19
2.4.2 Treinamento e Aprendizagem	20
2.4.3 TDNN com Configuração Focada	21
2.4.4 <i>Perceptron</i> Multicamadas.....	23
3 METODOLOGIA.....	27
3.1 VARIÁVEIS UTILIZADAS	27
3.2 CONFIGURAÇÕES DA REDE TDNN UTILIZADA.....	32
3.3 CONFIGURAÇÕES DA REDE PMC UTILIZADA.....	33
3.4 SELEÇÃO DE VARIÁVEIS POR COEFICIENTE DE PEARSON	34
4 RESULTADOS OBTIDOS	37
4.1 TDNN COM CONFIGURAÇÃO FOCADA	37
4.2 PMC COM UMA CAMADA OCULTA.....	39
4.2 PMC COM DUAS CAMADAS OCULTAS.....	40
4.3 PMC DUAS CAMADAS COM SELEÇÃO DE ENTRADAS.....	42
4.3.1 Entradas com coeficiente de Pearson abaixo de -0,4 ou acima de 0,4.....	42
4.3.2 Entradas com coeficiente de Pearson abaixo de -0,5 ou acima de 0,5.....	44
4.3.3 Entradas com coeficiente de Pearson abaixo de -0,6 ou acima de 0,6.....	46
4.3.4 Entradas com coeficiente de Pearson abaixo de -0,7 ou acima de 0,7.....	48

4.4 PMC UMA CAMADA COM SELEÇÃO DE ENTRADAS	51
5 CONCLUSÕES E CONSIDERAÇÕES FINAIS	53
5.1 Trabalhos Futuros	56
REFERÊNCIAS	57

1 INTRODUÇÃO

Devido à crise de abastecimento no setor elétrico em 2001, que resultou em um plano de racionamento de energia, foi necessário um incentivo ainda maior para reestruturação do sistema elétrico brasileiro, já em andamento desde os anos noventa, visando o planejamento do setor a longo prazo, uma avaliação permanente da segurança do suprimento de energia elétrica e mudanças nas atividades do Mercado Atacadista de Energia (MAE), até então, agente público responsável por realizar as transações de compra e venda de energia elétrica. Com tal intuito foram criadas a Empresa de Pesquisa Energética (EPE), o Comitê de Monitoramento do Setor Elétrico (CMSE) e a Câmara de Comercialização de Energia Elétrica (CCEE) (SOUZA, 2014).

A forma de comercializar energia elétrica teve uma significativa alteração com a implementação da CCEE. Foram instituídos dois ambientes de contratação, o Ambiente de Contratação Regulado (ACR) e o Ambiente de Contratação Livre (ACL), os contratos realizados no ACR possuem regulação específica de alguns aspectos como preço da energia, submercado de registro do contrato e vigência de suprimento, que não são passíveis de alterações bilaterais por parte dos agentes compradores e vendedores de energia elétrica (CCEE, 2010).

No ACL, esses agentes têm a liberdade de negociar a compra de energia, estabelecendo volumes, preços e prazos de suprimento em contratos bilaterais. Com a criação desses contratos, os agentes (compradores e vendedores de energia elétrica) passaram a ter proteções de eventuais incertezas do sistema nas transações comerciais (CRUZ, 2017).

Porém a alteração mais significativa foi na forma de realizar os leilões de energia, que a partir da criação da CCEE passaram a utilizar o critério de menor tarifa, que visa a redução do custo de aquisição da energia elétrica a ser repassada aos consumidores cativos. Esse critério apresenta tarifas viáveis para o consumidor, asseguram o retorno satisfatório dos investimentos realizados pelos agentes do setor elétrico e viabilizam a expansão da economia, qualidade de vida e desenvolvimento para sociedade (CCEE, 2010).

Entre as funções absorvidas da MAE, a CCEE passou a ser responsável pela apuração do Preço de Liquidação das Diferenças (PLD). Segundo a CCEE é

“utilizado para valorar as transações realizadas no mercado de curto prazo” (CCEE, 2010, p. 12).

O Operador Nacional do Sistema (ONS), implantado a partir do projeto de Reestruturação do Setor Elétrico brasileiro (projeto RE-SEB) em 1996, utiliza uma cadeia de modelos de otimização para determinar o despacho de mínimo custo de operação (Custo Marginal de Operação, CMO) (ANEEL 2015).

O CMO é determinado a partir de um modelo de precificação que obtém o despacho ótimo para o período em estudo, definindo a geração para cada submercado com base em condições hidrológicas, na demanda de energia, nos preços de combustíveis, no custo de déficit, na entrada de novos projetos e na disponibilidade de equipamentos de geração e transmissão. A variação dessas condições interferem diretamente no CMO e conseqüentemente no PLD que o toma como base para ser calculado (CCEE, 2010).

A figura 1 apresenta um gráfico com a variação do PLD em R\$/MWh durante o ano de 2014 e 2015 no submercado Sudeste/Centro-Oeste, utilizando valores fornecidos pela CCEE.

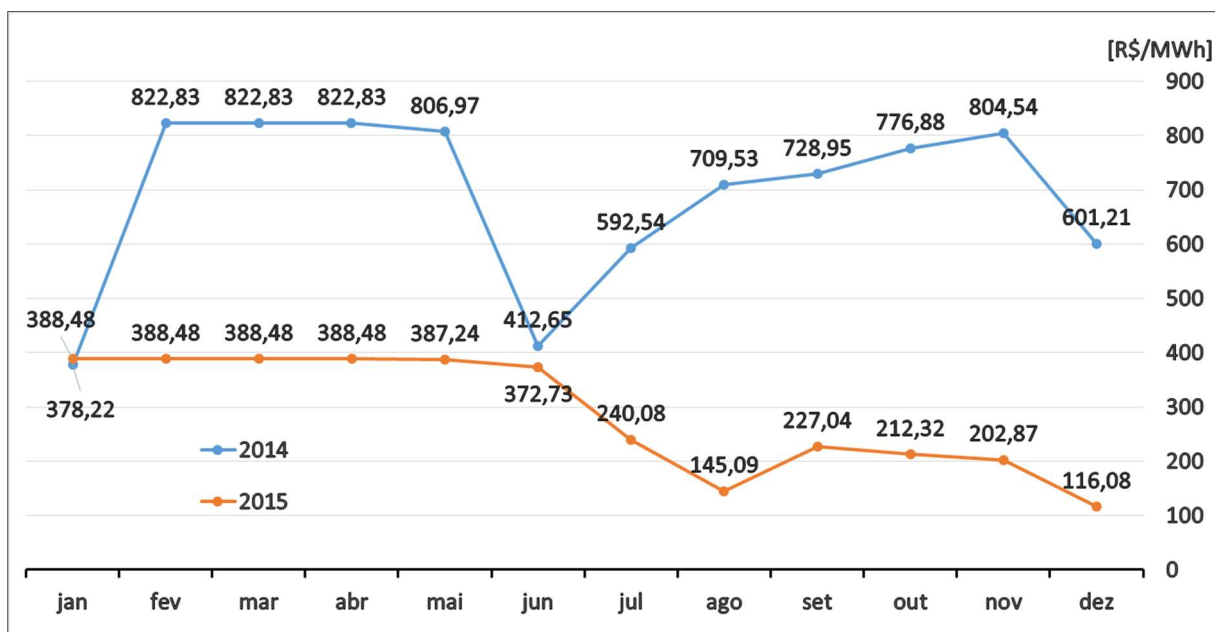


Figura 1 – Variação do PLD no submercado SE/CO (2014-2015)

Fonte: Autoria própria

Nota-se na figura 1 que não só existe variação ao decorrer de um ano, devido as condições climáticas diferentes das estações, como também existe uma

grande variação para o mesmo mês de 2014 e 2015, que chegaram a ter mais de cinquenta por cento de redução de um ano para outro. Fatores imprevisíveis que influenciam na composição do PLD, como o comportamento meteorológico no submercado SE/CO em 2014 (onde a presença de uma massa de ar de alta pressão, seca e quente não permitiu a entrada das frentes frias típicas do período chuvoso), podem variar o PLD de uma forma acentuada e representa uma grande diferença financeira na comercialização de energia elétrica (CCEE, 2016).

Como o setor elétrico brasileiro é composto predominantemente por sistemas hidráulicos, é natural que o PLD seja um ativo volátil. Essa volatilidade é ligada a dinâmica das afluentes que visa garantir a oferta de carga mesmo em condições hidrológicas adversas (LEITE; CASTRO, 2010).

A indústria da energia elétrica no mundo vem se tornando cada vez mais competitiva. Isso explica a importância de prever o comportamento do mercado, consequentemente do preço da energia, para que o produtor ou consumidor de energia elétrica possa negociá-la em momentos economicamente vantajosos. Essa previsão de mercado pode ser realizada por estimativas de eventos futuros, e teria o objetivo de reduzir riscos em tomadas de decisões. As redes neurais artificiais foram definidas para esse trabalho por terem sido muito utilizadas para problemas de previsão ligados a energia elétrica, baseando-se em séries temporais (ALANIS, 2018).

Visto a necessidade dos agentes comercializadores de energia elétrica do mercado brasileiro de curto prazo competir de forma vantajosa, esse trabalho irá estimar o valor do PLD do submercado sudeste/centro-oeste utilizando redes neurais artificiais. Do exposto anteriormente, justifica-se a importância da presente proposta de trabalho.

1.1 OBJETIVOS

O trabalho tem objetivo geral de estimar o PLD no mercado de energia de curto prazo utilizando redes neurais artificiais, e como objetivos específicos, verificar através do uso de uma rede TDNN se o problema de estimação do PLD é dependente de seu próprio histórico de valores e se existe correlação linear entre variáveis consideradas no planejamento energético e o PLD.

2 REVISÃO DA LITERATURA

Neste capítulo será abordado o embasamento teórico do trabalho, abordando a precificação da energia elétrica, bem como o PLD e o mercado de curto prazo; as redes neurais utilizadas, *Time Delay Neural Networking* (com configuração foada) e *Perceptron* Multicamadas; e também os conceitos sobre o Coeficiente de Correlação Linear (ou Coeficiente de Pearson) utilizado como uma proposta para selecionar variáveis do Sistema Interligado Nacional (SIN) por grau de correlação, entre as mesmas e o PLD do submercado Sudeste/Centro-Oeste, afim de reduzir o número de variáveis utilizadas no processo de estimação.

2.1 PRECIFICAÇÃO DA ENERGIA ELÉTRICA

Como citado anteriormente, no Brasil, o que determina o modelo de precificação são as características da matriz energética, que tem a sua maior parcela na geração hídrica. Portanto, é essencial que se controle os níveis dos reservatórios, para que entre cheias e secas seja mantido o atendimento da demanda em todo o país. A ONS é o órgão responsável por tomar decisões para cumprir esse atendimento. Para isso é necessário um planejamento operacional que visa minimizar os custos de todos os recursos utilizados (SOUZA, 2014).

Entre as decisões que visam o custo mínimo de operação estão: os despachos hídrico e térmico, determinando quais usinas, em sua maioria hídricas e térmicas, devem gerar ou não energia elétrica para o sistema; os intercâmbios entre mercados, controlando a comercialização nacional e internacional (como o caso de Itaipu); e determinando os cortes de carga necessários em casos de geração insuficiente para o atendimento do sistema (SOUZA, 2014).

Para análise dessas decisões e controle desses custos variáveis a ONS conta com modelos computacionais que auxiliam a formação da precificação da energia. Entre eles existem o NEWAVE, DECOMP e DESSEM. O DESSEM é fruto de um projeto realizado pelo Centro de Pesquisas de Energia Elétrica (CEPEL) que desenvolve modelos computacionais relacionados à programação diária da operação

de sistemas de energia elétrica, esse modelo programa a operação do sistema com um horizonte de até dois dias, e não influencia a formação do PLD, objeto de estudo do trabalho que tem foco no mercado de curto prazo (CCEE, 2010).

Segundo a CCEE (2010) “O processo completo de cálculo do PLD - Preço de Liquidação das Diferenças consiste na utilização dos modelos computacionais NEWAVE e DECOMP, os quais produzem como resultado o Custo Marginal de Operação de cada submercado, respectivamente em base mensal e semanal”.

Portanto, para o cálculo do PLD a ONS utiliza dois modelos computacionais para auxiliar a otimização (minimizar custos) da operação, o NEWAVE é utilizado para análise de médio e longo prazo, subdividido em base mensal (horizonte de cinco anos), e será utilizado como informação de entrada para o DECOMP, que é utilizado para curto prazo (horizonte de um ano) determinando o despacho ótimo por usina, o que minimiza custos semanais de operação. Ambos os modelos buscam o equilíbrio entre custos imediatos e futuros na determinação dos custos marginais de operação (VITORINO, 2011).

Como visto anteriormente, o custo total de operação depende de custos imediatos e futuros de operação. O custo imediato é relacionado ao custo de operação imediata para despacho hídrico, geração térmica e déficit de atendimento, a máxima utilização da energia hidrelétrica disponível é a ação mais econômica do ponto de vista imediato, pois minimiza os custos de combustíveis na geração térmica, isso resulta em maiores riscos de déficits futuros. Por outro lado, a máxima confiabilidade de fornecimento é obtida conservando o nível dos reservatórios o mais elevado possível, o que significa utilizar mais geração térmica e, portanto, aumento dos custos de operação. (CCEE, 2010).

Os custos futuros, ao contrário dos custos imediatos, representam uma variação decrescente em relação ao aumento dos níveis de armazenamento de água. Para o atendimento futuro com níveis próximos de zero, o custo é alto, se resumindo em geração térmica e custos de déficit. A medida que os níveis de armazenamento de água para o futuro são maiores se diminui a expectativa de geração térmica (VITORINO, 2011).

Os custos imediatos e futuros podem ser então representados como curvas, crescente e decrescente, respectivamente (figura 2).

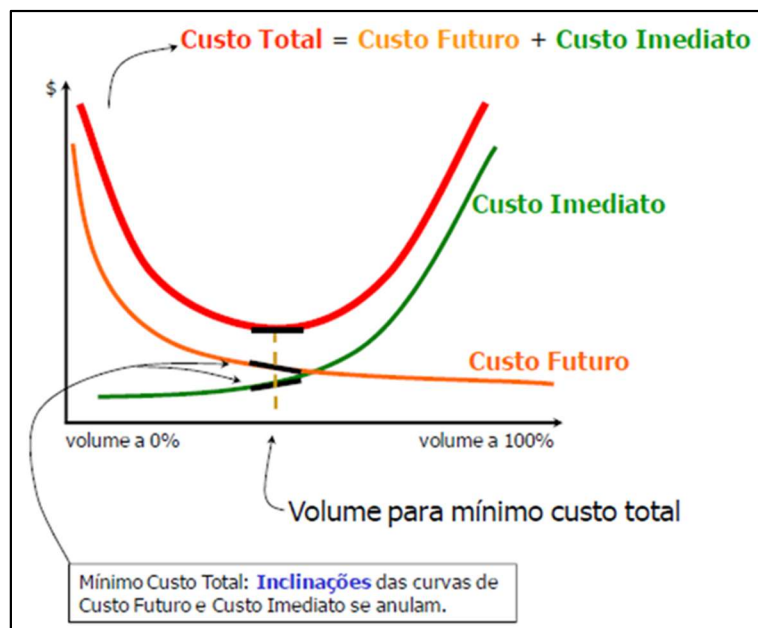


Figura 2 – Minimização do custo total de operação

Fonte: VITORINO, 2011 (modificado)

Segundo Vitorino (2011), “o uso ótimo da água é aquele que minimiza a soma do custo de geração térmica no presente, e o valor esperado do custo de geração desta fonte até o fim do período de estudo”. Esse ponto ótimo de operação, entre as variações de custo imediato e futuro, é refletido no ponto mínimo da curva que representa o custo total, como apresentado na figura 2. Os modelos computacionais NEWAVE e DECOMP vão utilizar uma base de dados levantados pelo monitoramento da ONS para estimar os custos futuros, estimar a curva de custos totais e determinar o ponto ótimo de operação.

2.2 MERCADO DE CURTO PRAZO

Tanto no Ambiente de Contratação Regulada, do qual participam agentes de geração, comercialização e distribuição de energia elétrica, quanto no Ambiente de Contratação Livre, do qual participam agentes de geração, comercialização, importadores e exportadores de energia e consumidores livres, todos os contratos de compra e venda de energia devem ser registrados na CCEE, que tem a função de

medir a real quantidade de energia produzida e consumida para cada agente do setor elétrico (CCEE, 2010).

A figura 3 ilustra de forma simples a energia comercializada no mercado de curto prazo, que atua nas diferenças entre energia verificada (produzida) e energia contratada (consumida).

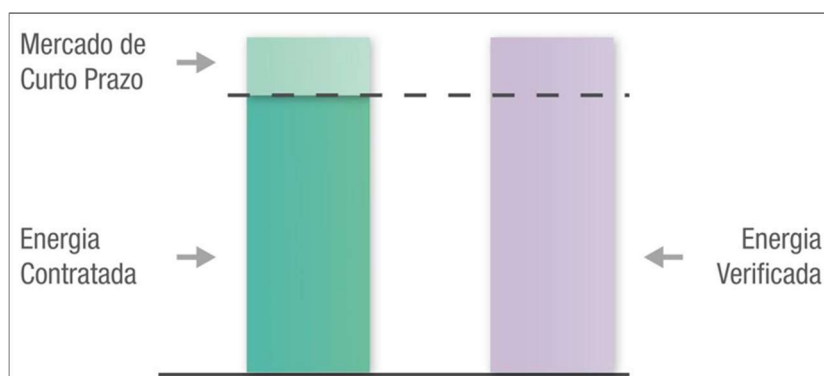


Figura 3 – Energia comercializada no mercado de curto prazo

Fonte: CCEE, 2010

As diferenças entre energia contratada e consumida (positiva ou negativa) são contabilizadas para serem liquidadas financeiramente no Mercado de Curto Prazo (MCP) e valoradas ao Preço de Liquidação das Diferenças (CCEE, 2010).

2.3 PREÇO DE LIQUIDAÇÃO DAS DIFERENÇAS

O preço de liquidação das diferenças (PLD), utilizado no mercado de curto prazo de energia elétrica, é calculado a partir dos dados considerados pelo ONS que otimizam o sistema interligado nacional (SIN). Como na matriz energética do Brasil a geração é majoritariamente hídrica, modelos matemáticos são utilizados para calcular o PLD, que tem por objetivo o equilíbrio entre o benefício presente no uso da água e o benefício futuro de seu armazenamento. Esse preço é apurado com base em informações previstas, anteriores à operação real do sistema, considerando os valores de disponibilidades declaradas de geração e o consumo previsto de cada submercado (CCEE, 2010).

Baseando-se nas condições hidrológicas, na demanda de energia, nos preços dos combustíveis usados na geração térmica, no custo de déficit (custo estimado para um possível não atendimento à carga do sistema elétrico), na entrada de novos projetos e na disponibilidade de transmissão e geração, o modelo calcula o despacho ótimo para o período em estudo, definindo assim a geração hidráulica e térmica para cada submercado e, conseqüentemente, Custo Marginal de Operação (CMO) (CCEE, 2010).

O operador do mercado (CCEE) utiliza os mesmos modelos de otimização do ONS e o CMO para determinar semanalmente o PLD, que é limitado por um valor mínimo e por um valor máximo, para cada patamar de carga e para cada submercado, com base no CMO (ANEEL, 2015).

O valor máximo é definido como sendo o menor preço estrutural da termelétrica mais cara do programa mensal de operação (PMO) entre janeiro do ano corrente e o preço limite do ano anterior corrigido pelo IGP-DI (Índice Geral de Preços – Disponibilidade Interna, utilizado em reajustes ligados a inflação de tarifas públicas) acumulado entre novembro de um ano e novembro do ano seguinte. O valor mínimo é atualizado anualmente na primeira semana operativa de janeiro, pelo custo variável da usina Binacional de Itaipu, e é valorado pela média geométrica da taxa de câmbio do dólar do ano anterior. Esses limites são regulamentados pela Resolução ANEEL nº377/3 e não levam em consideração as restrições de transmissão interna dos submercados, porém, consideram os limites de intercâmbio entre eles (CCEE, 2010).

2.4 REDES NEURAIS ARTIFICIAIS

As redes neurais artificiais (RNAs) são modelos computacionais que imitam o sistema nervoso dos seres vivos, dentro da área de inteligência computacional, onde são capazes de adquirir e processar conhecimentos e informações por um conjunto de unidades de processamento comparadas aos neurônios e por um grande número de interconexões comparadas as sinapses do cérebro (SILVA et al., 2010).

Essas redes tem como características principais a adaptação por experiência (onde seus pesos sinápticos são ajustados a partir da apresentação sucessiva de padrões relacionados ao comportamento do processo) possibilitando a aprendizagem por experimentação, a capacidade de aprendizado através de um processo de treinamento (fazendo com que a rede consiga extrair a relação entre as variáveis do processo) e a habilidade de generalizar o conhecimento adquirido, possibilitando a estimação de resultados até então desconhecidos (SILVA et al., 2010).

A utilização da rede neural artificial foi definida a partir da sua capacidade de organização interna visando possibilitar o agrupamento de padrões que apresentam particularidades em comum, da capacidade de incremento da robustez da arquitetura no eventual não funcionamento de um neurônio e da tolerância a falhas, quando parte da estrutura é sensivelmente corrompida, devido ao elevado número de interconexões entre os neurônios artificiais (SILVA et al., 2010).

Foi definido para o problema de estimação do PLD o uso de duas topologias de redes neurais, a rede TDNN (*time delay neural network*) com configuração focada (também chamada de *focused time lagged feedforward network*) onde a entrada é a saída atrasada no tempo e a rede *Perceptron* de Múltiplas Camadas (PMC), consideradas por Silva, Spatti e Flauzino (2010, p. 25) como “arquiteturas versáteis quanto à aplicabilidade”.

2.4.1 Arquitetura

A arquitetura de uma rede neural artificial depende do número de camadas e do número e disposição de neurônios e suas interligações. Pode-se dividir as arquiteturas em quatro arranjos topológicos, *feedforward* de camada simples, *feedforward* de múltiplas camadas, rede realimentada e rede reticulada. Nas redes *feedforward* o fluxo de informações é sempre unidirecional, as redes realimentadas possuem saídas dos neurônios realimentadas como sinais de entrada para outros neurônios e as redes reticuladas tem camadas ramificadas e disposição bidimensional dos neurônios (SILVA et al., 2010).

De forma geral, as redes podem ser dividida em três partes denominadas como camadas (ilustrado na figura 4 por uma rede *perceptron* multicamadas): A camada de entrada, que é responsável pelo recebimento das informações, sinais, características ou medições do meio externo (dados, amostras ou padrões); as camadas escondidas (também chamadas de camadas intermediárias, ocultas ou invisíveis) são responsáveis por extrair as características associadas ao processo ou sistema analisado; e a camada de saída, responsável pela produção e apresentação dos resultados finais da rede (SILVA et al., 2010).

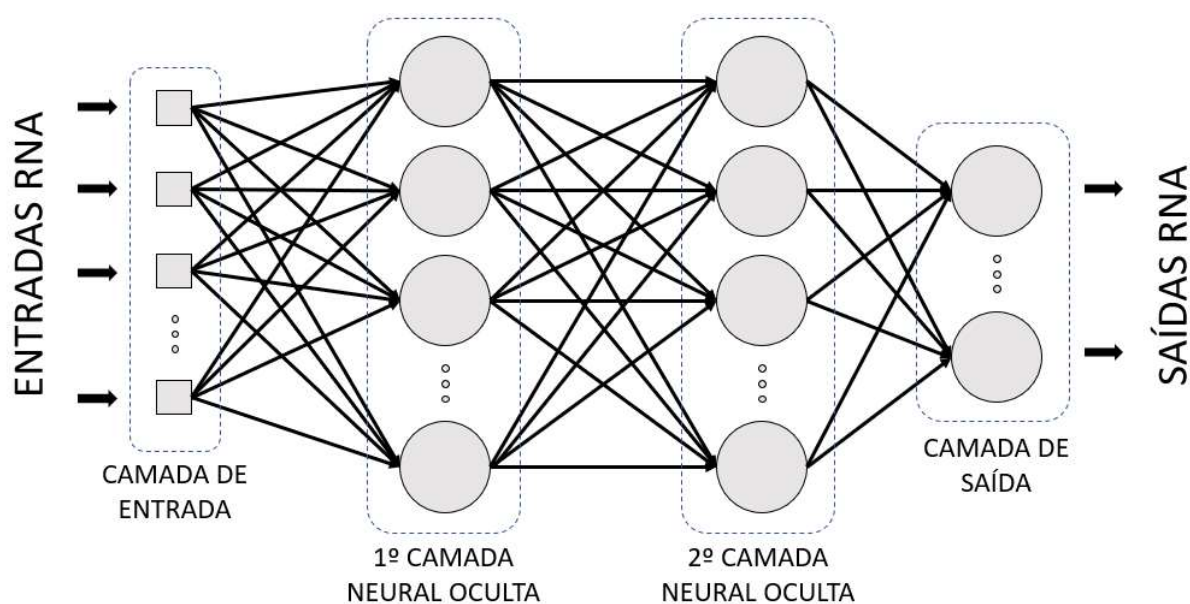


Figura 4 – Rede *feedforward perceptron* multicamadas

Fonte: Autoria própria

2.4.2 Treinamento e Aprendizagem

Como dito anteriormente, a relevância do uso das redes neurais está na sua capacidade de aprender a partir da apresentação de padrões que demonstram o comportamento do sistema, ou seja, a rede aprende o relacionamento entre a camada de entrada e camada de saída, se tornando capaz de generalizar soluções para o sistema com a apresentação de quaisquer informações desejadas (SILVA et al., 2010).

Essa apresentação sucessiva de padrões no treinamento é responsável por ajustar os pesos sinápticos (peso das relações entre neurônios) e as limiares dos neurônios. Assim para cada melhor topologia encontrada no processo de validação da rede, cada neurônio terá um valor atribuído de peso sináptico (bias) que se inicia aleatoriamente com um valor pequeno e se ajusta de acordo com a regra Delta Generalizada (*Backpropagation*), apresentada na equação 1 (THEODORO, 2019).

$$W_{ji}^{(L) \text{ novo}} = W_{ji}^{(L) \text{ antigo}} + \eta \delta_j^{(L)} x_{i(k)}, i = 0, 1, 2 \dots n \quad (1)$$

Onde “ W ” é o peso sináptico “ j ” da camada “ L ” respectivo à sua conexão com o neurônio “ i ” da camada “ $L - j$ ”, “ η ” é a taxa de aprendizagem ($0 < \eta < 1$), “ $\delta_j^{(L)}$ ” é calculado para cada neurônio “ j ” da camada “ L ”, e x é o elemento i do vetor de entradas da amostra k . Essa regra conduz a rede à incrementar os pesos sinápticos sempre que há concordância entre a resposta da rede e a saída desejada e a decrementar os pesos quando ocorra o contrário (SILVA et al., 2010).

Durante esse processo, cada apresentação completa de amostras é denominada como uma época, e as amostras são divididas em dois subconjuntos, o subconjunto de treinamento, onde acontecerá a aprendizagem, e o subconjunto de testes que será comparado aos valores desejados e obtidos pela generalização (SILVA et al., 2010).

Duas topologias serão utilizadas no trabalho, na TDNN será utilizado um treinamento supervisionado, onde as entradas são as saídas atrasadas no tempo, assim a rede se auto organiza em relação as particularidades do conjunto de amostras para produzir uma estimativa de valores futuros da variável. Já na rede PMC,

o treinamento supervisionado é realizado para cada um conjunto de amostras de entrada constituído de variáveis correlacionadas à variável de saída que se deseja estimar e um conjunto de amostras de saída, esse processo é comparado por Silva, Spatti e Flauzino (2010, p. 52) à um “professor ensinando para a rede qual a resposta correta para cada amostra apresentada em suas entradas”.

2.4.3 TDNN com Configuração Focada

Nessa rede a previsão ou predição de valores posteriores a partir de um instante é realizada em função do conhecimento de seus valores, temporalmente, anteriores (equação 2).

$$x(t) = f(x(t - 1), x(t - 2), \dots, x(t - n_p)) \quad (2)$$

Diferente de uma TDNN (*time delay neural network*) convencional, onde todos os neurônios recebem atraso temporal, a TDNN com configuração focada, receberá o atraso temporal apenas na camada de entrada (SILVA et al., 2010).

Na figura 5 vemos a ilustração da rede, onde n_p é a quantidade de amostras passadas necessárias para a estimação do valor $x(t)$.

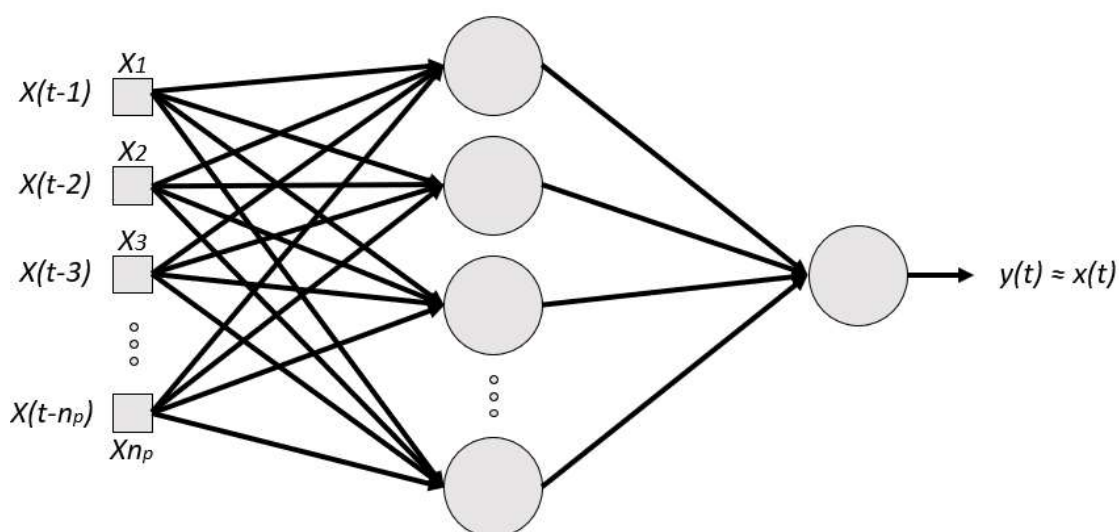


Figura 5 – Topologia TDNN com configuração focada

Fonte: SILVA et al., 2010 (modificado)

A figura 6 apresenta o método utilizado neste trabalho, em Matlab® com auxílio de seu *toolbox*, no treinamento da rede TDNN, o incremento de um neurônio ($n+1$) na camada intermediária ocorre a cada ciclo.

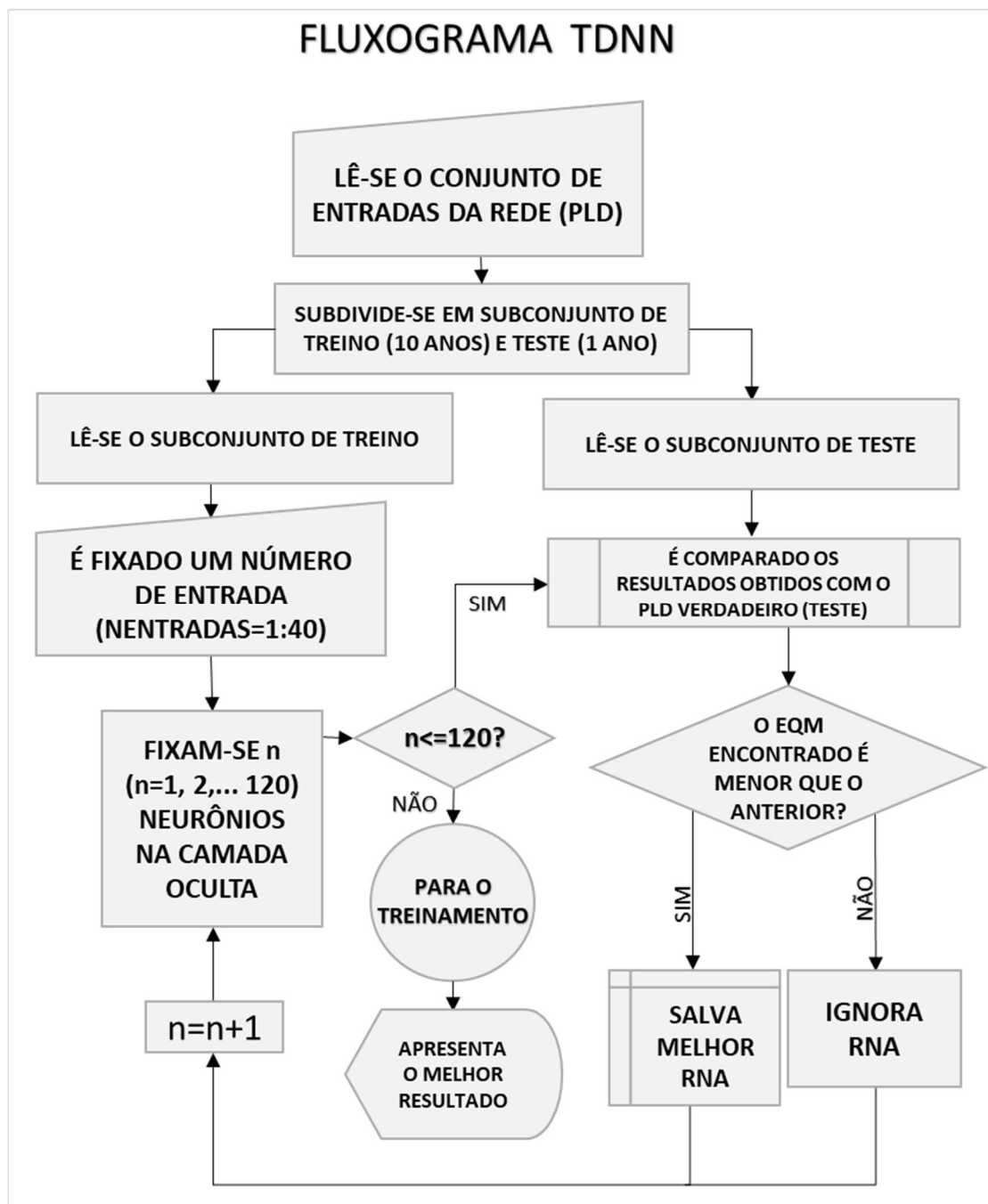


Figura 6 – Fluxograma treinamento TDNN

Fonte: Autoria própria

2.4.4 Perceptron Multicamadas

Como visto na figura 4, sinais obtidos da aplicação são as entradas do processo, que serão propagadas de neurônio a neurônio na direção da camada neural de saída. Observa-se também que as entradas em uma camada neural oculta são as saídas da camada neural anterior, ou seja, as sinapses são realizadas de forma unidirecional. A quantidade de entradas, de números de camadas ocultas e de neurônios na camada de saída são especificações que serão características advindas do processo e serão definidas de forma a otimiza-lo (SILVA, et al., 2010).

A figura 7 apresenta o método do sistema de treinamento da rede PMC de uma camada utilizada neste trabalho.

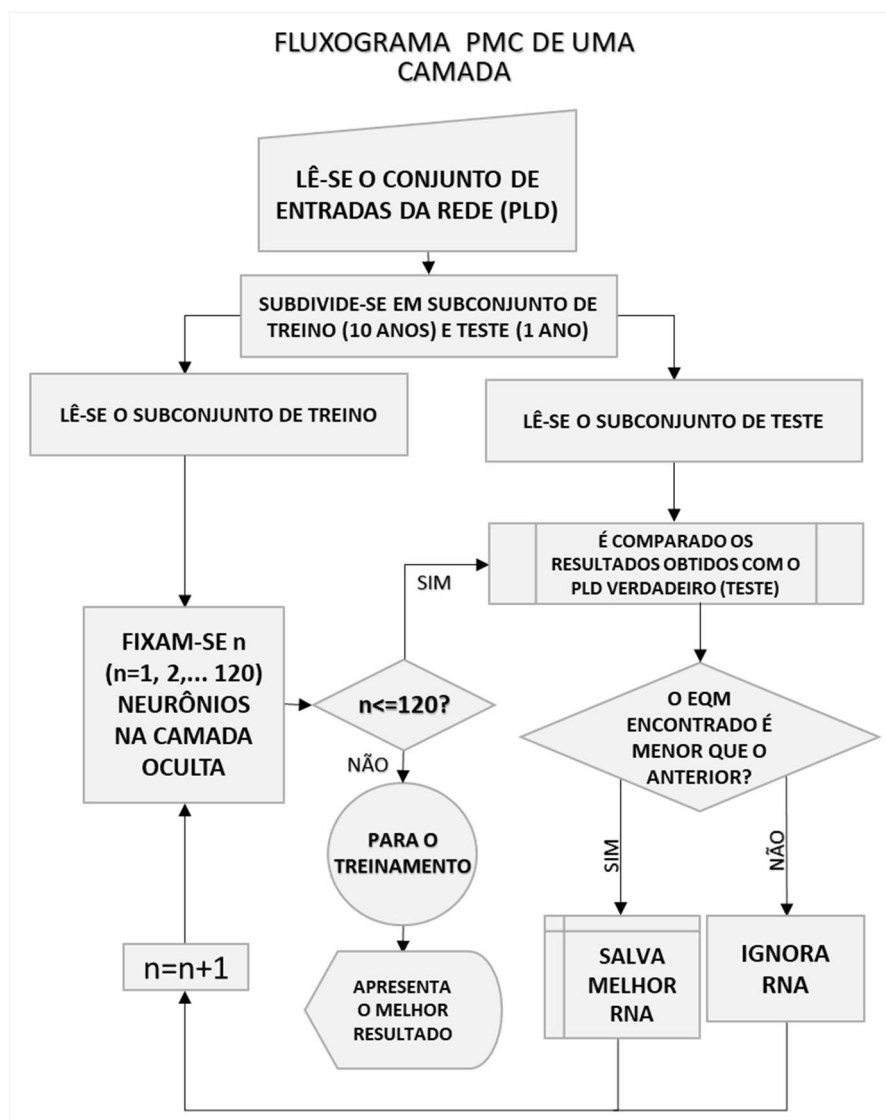


Figura 7 – Fluxograma treinamento PMC de uma camada

Fonte: Autoria própria

Assim como na figura 6, no fluxograma apresentado na figura 7 possuem incremento de um neurônio ($n+1$) na camada intermediária. A figura 8 apresenta o método do sistema de treinamento para rede PMC de duas camadas intermediárias utilizada neste trabalho, que terá o incremento de neurônios na primeira ($n+1$) e segunda ($m+1$) camadas intermediárias.

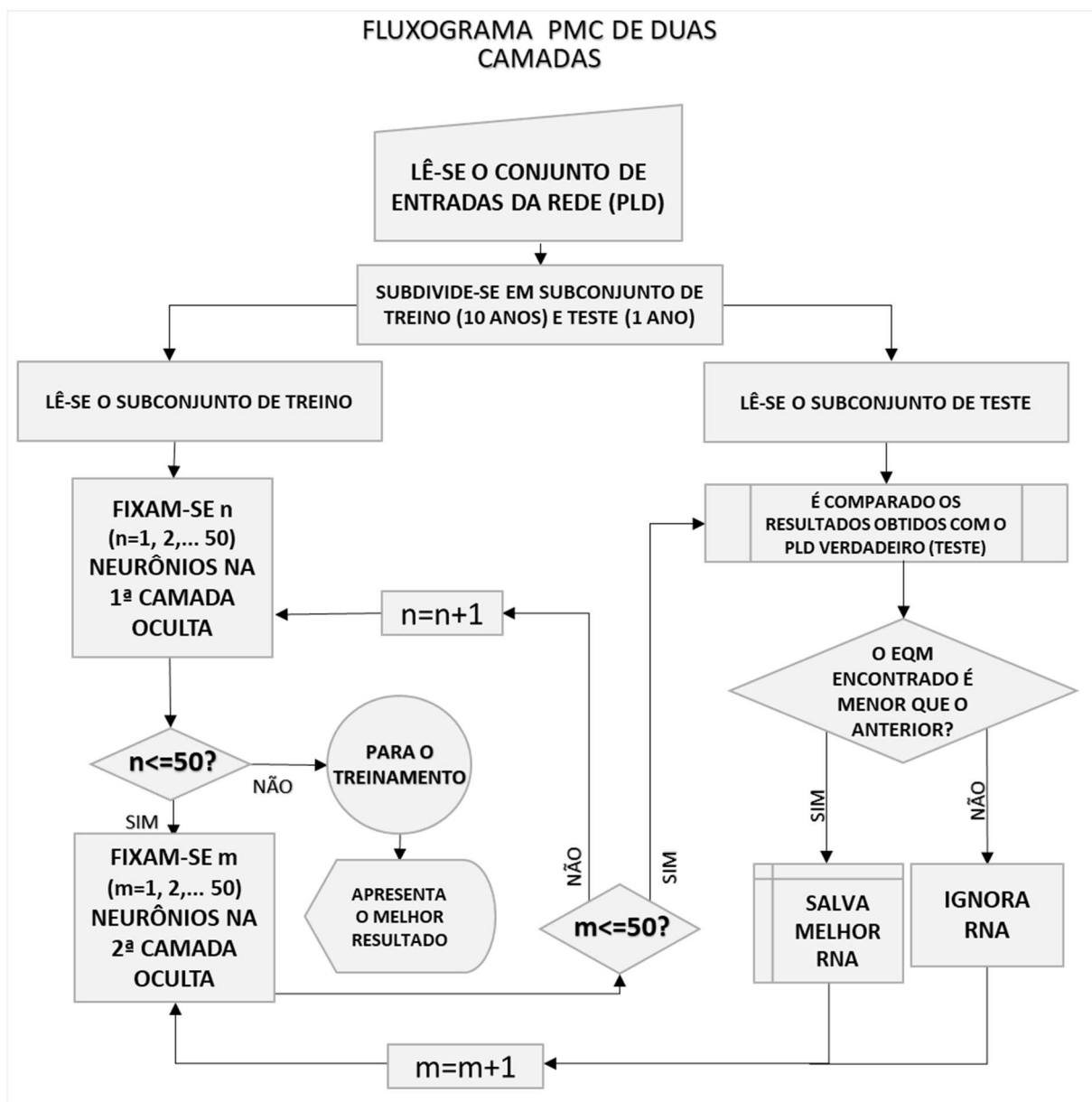


Figura 8 – Fluxograma treinamento PMC de duas camadas

Fonte: Autoria própria

2.5 COEFICIENTE DE PEARSON

O Coeficiente de Pearson será utilizado nesse trabalho na tentativa de selecionar as variáveis de entrada com maior influência na estimativa do PLD, e será calculado via *software* Matlab®. Os coeficientes de correlação linear (coeficientes de Pearson) variam entre -1 e 1, indicando uma forte associação inversa quando próximo de -1, uma fraca associação quando próximo de 0 e uma forte associação direta quando próximo de 1, como descrito na escala apresentada na figura 9.

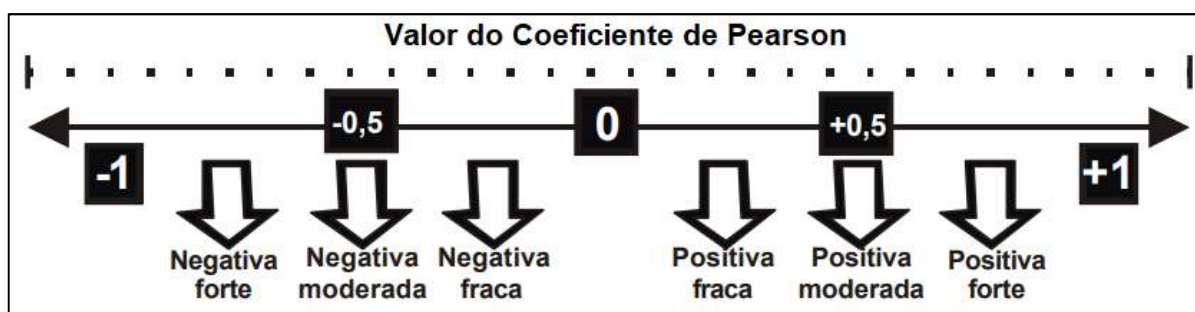


Figura 9 – Escala de correlação por coeficiente de Pearson

Fonte: LOPES, 2003

Consideremos os gráficos de dispersão da figura 10 (gráfico onde pontos são representados ao longo de um plano coordenado, providos de uma escala) e um conjunto de variáveis x e y (no nosso caso x é a entrada analisada e y a saída da rede). Ao analisar a dispersão da figura 10 (a), afirma-se uma associação linear inversa entre as variáveis, visto que o produto entre as coordenadas será sempre negativo, visto que o aumento da variável y implica na redução da variável x . De modo análogo, na figura 10(b), vemos que os dados possuem uma associação linear direta, porque a medida que a variável y aumenta, x tende a aumentar. Esse produto sempre resultará em um valor positivo, portanto a correlação direta tende a +1. Entretanto, na figura 10(c) não existe um padrão para o comportamento de uma variável em relação a outra, o que implica na não correlação linear (LARSON, 2009).

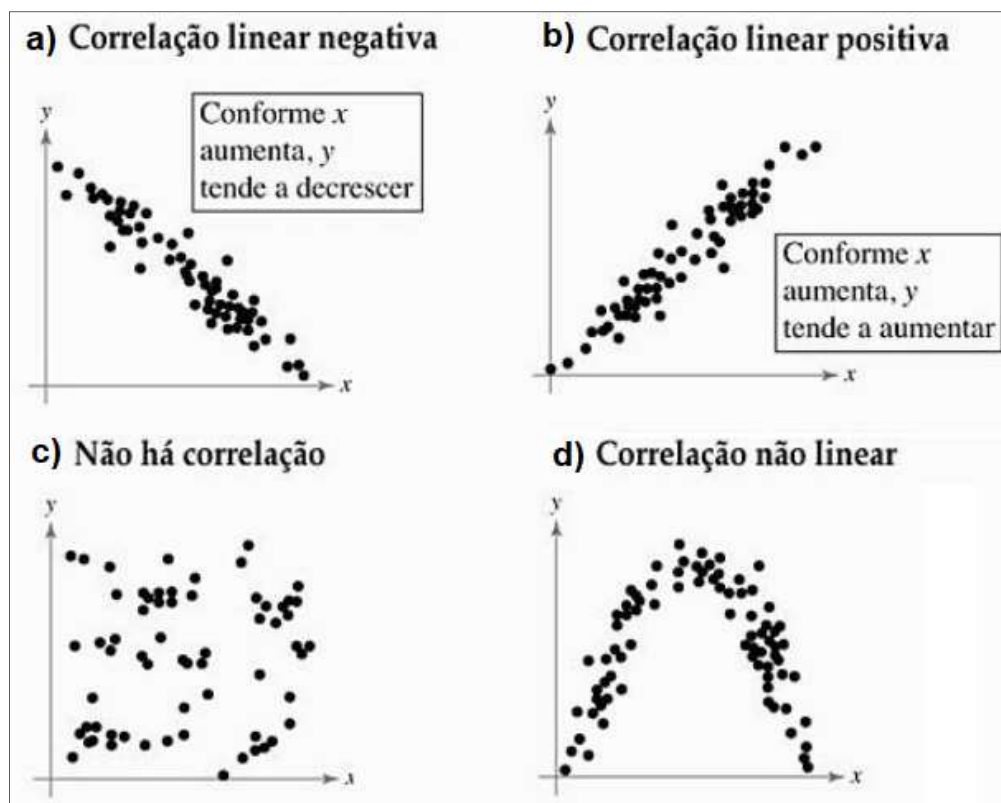


Figura 10 – Exemplos de associação linear

Fonte: LARSON, 2010, Pg. 395

É importante evidenciar que existem relações não lineares entre as variáveis do SIN, assim, essa análise linear não poderia identificar a correlação entre as mesmas, como por exemplo, uma relação parabólica ilustrada na figura 10(d). Nela vemos claramente uma relação parabólica entre as variáveis no gráfico de dispersão de pontos, porém uma análise de correlação linear resultaria em um coeficiente fraco, o que não representaria de forma verídica a correlação entre as variáveis em questão.

3 METODOLOGIA

Neste capítulo será apresentado como foram utilizados os métodos descritos anteriormente para o processo de estimação do PLD do submercado Sudeste - Centro Oeste. As redes foram testadas ao longo do trabalho com a finalidade de buscar um erro quadrático médio (medida de erro utilizada na comparação entre o valor real e estimado) menor. Assim, utilizou-se inicialmente topologias mais simples que posteriormente foram alteradas com aumento do número de informações de entradas, número de camadas, e por fim, seleção de variáveis.

3.1 VARIÁVEIS UTILIZADAS

As amostras de entrada e saída utilizadas são médias mensais oriundas da base de dados técnica do ONS e refletem a carga atendida por usinas supervisionadas e programadas pelo ONS durante onze anos, de janeiro de 2006 a dezembro de 2016. Esse conjunto total de informações foram divididas em dois subconjuntos, um subconjunto de dez anos para treinamento e um ano (último ano) para o subconjunto de testes, tanto na rede TDNN onde as variáveis utilizadas são as médias mensais do PLD do submercado Sudeste – Centro Oeste, quanto na PMC onde as saídas são as mesmas da rede anterior, porém a entrada é formada por vinte e oito séries amostrais retiradas do SIN, sendo diversas variáveis distintas para os quatro submercados e uma para a usina binacional de Itaipu.

A tabela 1 apresenta o subconjunto de treinamento extraído do PLD do submercado Sudeste – Centro Oeste em R\$/MWh (reais por megawatt hora). Na rede TDNN, no treinamento supervisionado, esses dados serão apresentados na entrada como um padrão a ser identificado para estimação futura, e na PMC, no treinamento supervisionado, serão apresentados como soluções para um subconjunto de variáveis (descritas adiante) inserido na entrada.

A tabela 2 apresenta o subconjunto de testes do PLD médio mensal do submercado SE-CO, utilizado nas duas redes para comparação entre o resultado obtido através da estimação e o valor real no ano de 2016.

Tabela 1 – PLD médio mensal do submercado SE-CO de Jan. de 2006 à Dez. de 2015

PLD MÉDIO MENSAL DO SUBMERCADO SE-CO [R\$/MWh] - SUBCONJUNTO DE TREINAMENTO									
jan/06	28,64	jan/08	502,45	jan/10	12,91	jan/12	23,14	jan/14	378,22
fev/06	58,02	fev/08	200,42	fev/10	13,82	fev/12	50,67	fev/14	822,83
mar/06	28,56	mar/08	124,7	mar/10	27,24	mar/12	124,97	mar/14	822,83
abr/06	20,87	abr/08	68,8	abr/10	21,47	abr/12	192,7	abr/14	822,83
mai/06	51,91	mai/08	34,18	mai/10	32,34	mai/12	180,94	mai/14	806,97
jun/06	67,89	jun/08	76,2	jun/10	67,7	jun/12	118,49	jun/14	412,65
jul/06	90,9	jul/08	108,42	jul/10	89,61	jul/12	91,24	jul/14	592,54
ago/06	104,98	ago/08	102,79	ago/10	116,66	ago/12	119,08	ago/14	709,53
set/06	123,88	set/08	109,93	set/10	132,1	set/12	182,94	set/14	728,95
out/06	92,42	out/08	92,43	out/10	137,78	out/12	280,39	out/14	776,88
nov/06	80,82	nov/08	106,14	nov/10	116,68	nov/12	375,54	nov/14	804,54
dez/06	58,75	dez/08	96,97	dez/10	71,62	dez/12	259,57	dez/14	601,21
jan/07	22,62	jan/09	83,64	jan/11	28,19	jan/13	413,95	jan/15	388,48
fev/07	17,59	fev/09	52,08	fev/11	49,59	fev/13	214,54	fev/15	388,48
mar/07	17,59	mar/09	90,87	mar/11	23,41	mar/13	339,75	mar/15	388,48
abr/07	49,36	abr/09	46,46	abr/11	12,2	abr/13	196,13	abr/15	388,48
mai/07	59,96	mai/09	39	mai/11	17,35	mai/13	344,84	mai/15	387,24
jun/07	97,15	jun/09	40,84	jun/11	31,8	jun/13	207,62	jun/15	372,73
jul/07	122,59	jul/09	30,43	jul/11	23,08	jul/13	121,29	jul/15	240,08
ago/07	39,27	ago/09	16,31	ago/11	19,61	ago/13	163,38	ago/15	145,09
set/07	149,53	set/09	16,31	set/11	21,18	set/13	266,16	set/15	227,04
out/07	198,13	out/09	16,31	out/11	37,14	out/13	260,99	out/15	212,32
nov/07	185,11	nov/09	16,31	nov/11	45,55	nov/13	331,07	nov/15	202,87
dez/07	204,93	dez/09	16,31	dez/11	44,47	dez/13	290,72	dez/15	116,08

Fonte: CCEE, 2016

Tabela 2 – PLD médio mensal do submercado SE-CO de Janeiro de 2016 à Dezembro de 2016

PLD MÉDIO MENSAL DO SUBMERCADO SE-CO [R\$/MWh] - SUBCONJUNTO DE TESTES											
jan/16	fev/16	mar/16	abr/16	mai/16	jun/16	jul/16	ago/16	set/16	out/16	nov/16	dez/16
35,66	30,42	37,73	49,42	75,93	61,32	83,43	115,58	149	200,2	166,1	122,2

Fonte: CCEE, 2016

A figura 11 apresenta graficamente o conjunto dos 11 anos de dados médios mensais do PLD do submercado SE-CO utilizados apresentados nas tabelas 1 e 2, em verde (primeiros 10 anos) o subconjunto de treinamento e em azul (últimos 12 meses) o subconjunto de testes.

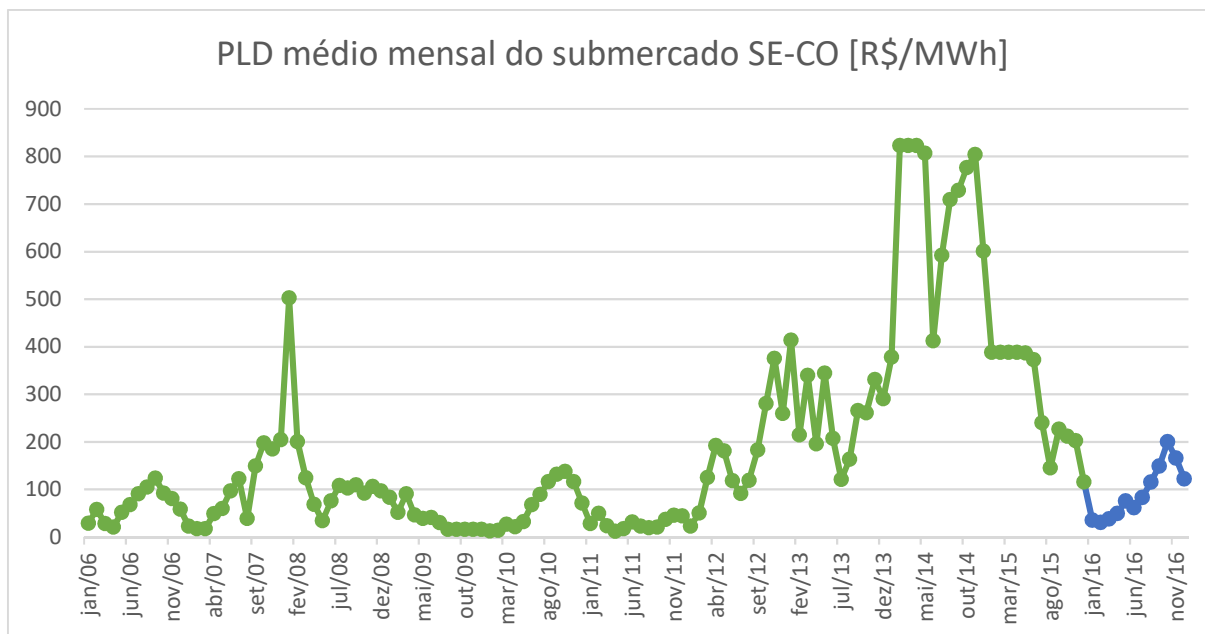


Figura 11 – PLD médio mensal SE-CO

Fonte: Autoria própria

A figura 12 apresenta apenas os últimos 12 meses (demonstrados também em azul na figura 11), período em que a rede foi testada, comparando os valores reais aos valores estimados.

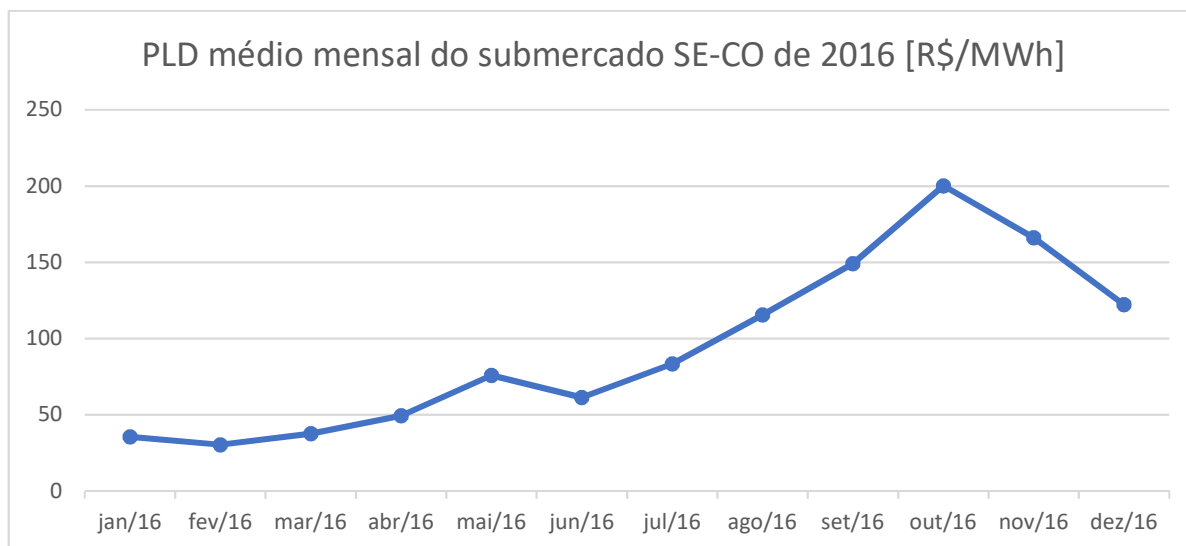


Figura 12 – PLD médio mensal SE-CO de 2016 (subconjunto de testes)

Fonte: Autoria própria

Na rede PMC, oito variáveis distintas foram inseridas como entradas no processo de estimação do PLD do submercado SE-CO, são elas:

- Carga de Demanda (MWh/h);
- Carga de Energia (GWh);
- Geração Hidráulica (GWh);
- Geração Térmica (GWh);
- Geração Termonuclear (GWh);
- Geração Eólica (GWh);
- Energia Natural Afluente (MWmed);
- Níveis Hidrológicos (GWh).

A carga de demanda tem MWh/h como unidade, e representa a demanda de cada submercado e a carga de energia tem GWh como unidade e é a energia distribuída ao decorrer do período estipulado. Essas variáveis históricas trazem uma prospecção do cenário energético no período estudado.

Quanto a geração, utilizou-se as fontes mais representativas nos quatro submercado do SIN, hidráulica, térmica, termo-nuclear, eólica e hidráulica da usina binacional de Itaipu. Esses dados tem GWh como unidade, e representam a geração das geradoras supervisionadas e programadas pela ONS.

A energia natural afluente (ENA) que tem como unidade MWmed (energia média no intervalo considerado) e os níveis hidrológicos (ou energia armazenada) que tem GWh como unidade, são variáveis climáticas que irão informar o comportamento físico de cada submercado ao decorrer do período estudado.

Visto que os quatro submercados e a usina de Itaipu são interligados pelo SIN, utilizou-se o conjunto de dados de cada submercado (Sudeste – Centro Oeste, Sul, Norte e Nordeste) e de Itaipu, relacionados ao PLD do submercado SE-CO (saídas demonstradas para rede), resultando em 28 vetores de entrada com 132 meses em cada. Assim eliminou-se a chance de se perder dados de outros submercados que colaborassem para a estimação do PLD do submercado SE-CO. O quadro 1 apresenta quais das oito variáveis os submercados (ou Itaipu) possuem, demarcado por um X no cruzamento entre linhas e colunas.

Quadro 1 – Variáveis de cada submercado.

	N	NE	SE-CO	S	Itaipu
CARGA DE DEMANDA	X	X	X	X	
CARGA DE ENERGIA	X	X	X	X	
ENERGIA NATURAL AFLUENTE	X	X	X	X	
GERAÇÃO EÓLICA		X		X	
GERAÇÃO HIDRÁULICA	X	X	X	X	X
GERAÇÃO TÉRMICA	X	X	X	X	
GERAÇÃO TERMONUCLEAR			X		
NÍVEIS HIDROLÓGICOS	X	X	X	X	

Portanto, os vinte e oito vetores de entrada utilizados para estimativa da rede PMC são os seguintes:

- a1 - Carga de Demanda SE-CO;
- a2 - Carga de Demanda S;
- a3 - Carga de Demanda NE;
- a4 - Carga de Demanda N;
- b1 - Carga de Energia SE-CO;
- b2 - Carga de Energia S;
- b3 - Carga de Energia NE;
- b4 - Carga de Energia N;
- c1 - ENA SE-CO;
- c2 - ENA S;
- c3 - ENA NE;
- c4 - ENA N;
- d1 - Geração Eólica S;
- d2 - Geração Eólica NE;
- e1 - Geração Hidráulica SE-CO;
- e2 - Geração Hidráulica S;
- e3 - Geração Hidráulica NE;
- e4 - Geração Hidráulica N;
- e5 - Geração Hidráulica Itaipu;
- f1 - Geração Térmica SE-CO;
- f2 - Geração Térmica S;
- f3 - Geração Térmica NE;
- f4 - Geração Térmica N;
- g1 - Geração Termonuclear SE-CO;
- h1 – Níveis Hidrológicos SE-CO;
- h2 – Níveis Hidrológicos S;
- h3 - Níveis Hidrológicos NE;
- h4 - Níveis Hidrológicos N.

3.2 CONFIGURAÇÕES DA REDE TDNN UTILIZADA

Como descrito no capítulo 2, a rede TDNN de configuração focada (saídas atrasadas no tempo utilizadas nas entradas) utilizou apenas os dados das tabelas 1 e 2, respectivamente, para treinamento e teste. A camada inicial variou de 1 a 40 entradas (cada variação para um teste) e a camada oculta de 1 a 120 neurônios em todos os testes realizados. A tabela 3 apresenta as configurações das redes TDNN treinadas e testadas.

Tabela 3 – Treinamentos realizados com a rede TDNN

CÓD.	NENTRADAS	NEURO	TEMPO EM HORAS	CÓD.	NENTRADAS	NEURO	TEMPO EM HORAS
01	1	1:120	0,02	20	20	1:120	0,22
02	2	1:120	0,04	21	21	1:120	0,21
03	3	1:120	0,04	22	22	1:120	0,24
04	4	1:120	0,04	23	23	1:120	0,21
05	5	1:120	0,05	24	24	1:120	0,23
06	6	1:120	0,05	25	25	1:120	0,22
07	7	1:120	0,06	26	26	1:120	0,25
08	8	1:120	0,07	27	27	1:120	0,25
09	9	1:120	0,09	28	28	1:120	0,27
10	10	1:120	0,10	29	29	1:120	0,27
11	11	1:120	0,10	30	30	1:120	0,28
12	12	1:120	0,12	31	31	1:120	0,29
13	13	1:120	0,13	32	32	1:120	0,31
14	14	1:120	0,14	33	33	1:120	0,32
15	15	1:120	0,14	34	34	1:120	0,36
16	16	1:120	0,15	35	35	1:120	0,37
17	17	1:120	0,16	36	36	1:120	0,41
18	18	1:120	0,16	37	37	1:120	0,49
19	19	1:120	0,21	38	38	1:120	0,47
20	20	1:120	0,22	39	39	1:120	0,46
21	21	1:120	0,21	40	40	1:120	0,50

Na tabela 3 têm-se as informações de CÓDIGO (coluna que identifica o treinamento), NENTRADAS (número de entradas), NEURO (número de neurônios na camada oculta) e o TEMPO EM HORAS para treinamento e teste de cada configuração (totalizando oito horas e meia para os resultados apresentados da rede TDNN).

3.3 CONFIGURAÇÕES DA REDE PMC UTILIZADA

A rede PMC foi utilizada em três momentos distintos, no primeiro foram utilizados os vinte e oito vetores de entrada e apenas uma camada intermediária, como apresenta a tabela 4.

Tabela 4 – Treinamento rede PMC com uma camada intermediária

CÓD.	NENTRADAS	NEURO	T. EM HORAS
41	28	1:120	0,18

Para essa rede as entradas foram fixadas para os 28 vetores de variáveis distintas (NENTRADAS) e variou-se os neurônios da camada intermediária de 1 a 120.

Posteriormente, para o mesmo número de entradas, realizou-se o teste para uma rede de duas camadas intermediárias, onde a entrada se manteve fixa, e variou-se os neurônios das camadas intermediárias de 1 a 50.

Tabela 5 – Treinamento rede PMC com duas camadas intermediárias

CÓD.	NENTRADAS	NEURO 1C	NEURO 2C	T. EM HORAS
42	28	1:50	1:10	0,34
43	28	1:50	11:20	0,41
44	28	1:50	21:30	0,56
45	28	1:50	31:40	0,76
46	28	1:50	41:50	1,06

Na tabela 5 a primeira camada intermediária foi nomeada como NEURO 1C e a segunda camada intermediária como NEURO 2C. A variação na segunda camada ocorreu de forma fracionada (10 em 10 neurônios) afim de exibir o processo de treinamento e avaliação dos resultados. A partir do treinamento cód. 46, as variáveis de entrada foram selecionadas gradativamente de acordo com o coeficiente de Pearson, correlacionando os 28 vetores de entrada ao PLD médio mensal do submercado SE-CO (apresentados no item 3.4).

3.4 SELEÇÃO DE VARIÁVEIS POR COEFICIENTE DE PEARSON

O Coeficiente de Pearson será utilizado nesse trabalho na tentativa de selecionar as variáveis de entrada com maior influência na estimativa do PLD. Essa redução de variáveis de entrada (além de diminuir as estimativas necessárias do processo) tem por finalidade a tentativa de reduzir o erro quadrático médio entre a estimação e os valores reais de teste. A tabela 6 apresenta o coeficiente de Pearson entre os vinte e oito vetores de entrada e o PLD médio do submercado SE-CO.

Tabela 6 – Coeficiente de Pearson entre variáveis de entrada e PLD do SE-CO

CÓD.	VARIÁVEL - SUBMERCADO	UNIDADE	COEF. PEARSON
a1	Carga de Demanda SE-CO	MWh/h	0,4851
a2	Carga de Demanda S	MWh/h	0,4752
a3	Carga de Demanda NE	MWh/h	0,4989
a4	Carga de Demanda N	MWh/h	0,5021
b1	Carga de Energia SE-CO	GWh	0,4436
b2	Carga de Energia S	GWh	0,5013
b3	Carga de Energia NE	GWh	0,5011
b4	Carga de Energia N	GWh	0,5125
c1	ENA SE-CO	MWmed	-0,2956
c2	ENA S	MWmed	0,414
c3	ENA NE	MWmed	-0,2886
c4	ENA N	MWmed	0,0354
d1	Geração Eólica S	GWh	0,1719
d2	Geração Eólica NE	GWh	0,1769
e1	Geração Hidráulica SE-CO	GWh	-0,4473
e2	Geração Hidráulica S	GWh	0,2108
e3	Geração Hidráulica NE	GWh	-0,4592
e4	Geração Hidráulica N	GWh	0,162
e5	Geração Hidráulica Itaipu	GWh	-0,3205
f1	Geração Térmica SE-CO	GWh	0,728
f2	Geração Térmica S	GWh	0,6866
f3	Geração Térmica NE	GWh	-0,4592
f4	Geração Térmica N	GWh	0,162
g1	Geração Termo-Nuclear SE-CO	GWh	0,0945
h1	Energia Armazenada SE-CO	GWh	-0,6831
h2	Energia Armazenada S	GWh	-0,1833
h3	Energia Armazenada NE	GWh	-0,5263
h4	Energia Armazenada N	GWh	-0,000236

Reduzindo de forma gradativa o número de entradas na rede é possível monitorar seu comportamento quanto a redução ou aumento do erro quadrático médio, assim será verificado se (ou até que ponto) a redução é benéfica ou não no processo de estimação do PLD. Foram realizados então, treinamentos para variáveis de entrada com coeficiente de correlação maior que o módulo de 0,4, 0,5, 0,6, e 0,7, com número de entradas, respectivamente, iguais a 16, 8, 3 e 1 (tabela 7).

Tabela 7 – Treinamento PMC de duas camadas com seleção prévia de variáveis

VARIÁVEIS DE ENTRADA COM COEFICIENTE DE PEARSON > ±0,4 CORRELACIONADAS AO PLD DO SUBMERCADO SE-CO				
CÓD.	NENTRADAS	NEURO 1C	NEURO 2C	T. EM HORAS
47	16	1-50	1-10	0,18
48	16	1-50	11-20	0,27
49	16	1-50	21-30	0,42
50	16	1-50	31-40	0,63
51	16	1-50	41-50	0,91
VARIÁVEIS DE ENTRADA COM COEFICIENTE DE PEARSON > ±0,5 CORRELACIONADAS AO PLD DO SUBMERCADO SE-CO				
CÓD.	NENTRADAS	NEURO 1C	NEURO 2C	T. EM HORAS
52	8	1-50	1-10	0,13
53	8	1-50	11-20	0,21
54	8	1-50	21-30	0,37
55	8	1-50	31-40	0,59
56	8	1-50	41-50	0,85
VARIÁVEIS DE ENTRADA COM COEFICIENTE DE PEARSON > ±0,6 CORRELACIONADAS AO PLD DO SUBMERCADO SE-CO				
CÓD.	NENTRADAS	NEURO 1C	NEURO 2C	T. EM HORAS
57	3	1-50	1-10	0,08
58	3	1-50	11-20	0,19
59	3	1-50	21-30	0,38
60	3	1-50	31-40	0,70
61	3	1-50	41-50	1,27
VARIÁVEIS DE ENTRADA COM COEFICIENTE DE PEARSON > ±0,7 CORRELACIONADAS AO PLD DO SUBMERCADO SE-CO				
CÓD.	NENTRADAS	NEURO 1C	NEURO 2C	T. EM HORAS
62	1	1-50	1-10	0,08
63	1	1-50	11-20	0,17
64	1	1-50	21-30	0,37
65	1	1-50	31-40	0,69
66	1	1-50	41-50	1,13

Na tabela 7, apresentada anteriormente, tem-se a configuração das redes PMC de duas camadas intermediárias utilizadas nos treinamentos com seleção gradativa de variáveis de entradas, na primeira coluna têm-se o código de cada treinamento, na segunda coluna o número de variáveis de entrada (selecionadas por coeficiente de Pearson), na terceira coluna o número de neurônios na primeira camada intermediária (variando de 1 a 50 neurônios), na quarta coluna o número de neurônios da segunda camada intermediária (variando de 1 a 50 neurônios) e na última coluna o tempo em horas necessário para o treinamento.

Para efeito de comparação, e afirmação do menor erro quadrático médio, realizou-se também os treinamentos para PMC de uma camada intermediária com a mesma seleção gradativa de variáveis de entrada utilizada nos treinamentos anteriores (tabela 8).

Tabela 8 – Treinamento PMC de uma camada com seleção prévia de variáveis

CÓD.	NENTRADAS	NEURO	T. EM HORAS	COEF. DE PEARSON
67	16	1:120	0,07	> ±0,4
68	8	1:120	0,03	> ±0,5
69	3	1:120	0,02	> ±0,6
70	1	1:120	0,01	> ±0,7

A tabela 8 apresenta a configuração utilizada para a rede PMC de uma camada com seleção gradativa de variáveis, onde a primeira coluna é o código de cada treinamento, a segunda é o número de entradas (selecionadas) a terceira o tempo em horas para realizar cada treinamento e a última o coeficiente de Pearson mínimo para cada treinamento.

4 RESULTADOS OBTIDOS

Este capítulo apresentará os resultados obtidos nos treinamentos, evidenciando o erro quadrático médio encontrado, bem como suas configurações e a comparação gráfica entre o valor real da série do PLD médio mensal do submercado SE-CO e a aproximação obtida pelas melhores redes de cada topologia testada.

4.1 TDNN COM CONFIGURAÇÃO FOCADA

O melhor erro quadrático médio encontrado (EQMBest) para cada uma das 40 configurações associadas ao melhor número de neurônios encontrados (NEUROBEST) para camada intermediária estão apresentados na tabela 9.

Tabela 9 – Resultados obtidos no treinamento da rede TDNN (CÓD. 01 a 40)

TREINAMENTO TDNN									
CÓD.	NENTRADAS	NEURO	NEUROBEST	EQMBEST	CÓD.	NENTRADAS	NEURO	NEUROBEST	EQMBEST
01	1	1:120	48	2,64E+03	21	21	1:120	1	2,92E+03
02	2	1:120	30	2,27E+03	22	22	1:120	2	4,90E+03
03	3	1:120	17	2,53E+03	23	23	1:120	1	6,21E+03
04	4	1:120	15	2,71E+03	24	24	1:120	3	3,39E+03
05	5	1:120	2	3,20E+03	25	25	1:120	2	9,95E+03
06	6	1:120	4	3,04E+03	26	26	1:120	13	1,25E+04
07	7	1:120	2	2,28E+03	27	27	1:120	5	1,11E+04
08	8	1:120	2	3,05E+03	28	28	1:120	1	6,37E+03
09	9	1:120	25	1,53E+03	29	29	1:120	2	1,27E+04
10	10	1:120	17	2,98E+03	30	30	1:120	9	2,62E+04
11	11	1:120	3	1,83E+03	31	31	1:120	2	1,03E+04
12	12	1:120	1	2,56E+03	32	32	1:120	3	6,20E+03
13	13	1:120	1	3,67E+03	33	33	1:120	2	8,39E+03
14	14	1:120	13	2,83E+03	34	34	1:120	2	6,21E+03
15	15	1:120	1	2,11E+03	35	35	1:120	8	8,16E+03
16	16	1:120	1	7,47E+03	36	36	1:120	4	3,06E+03
17	17	1:120	12	4,31E+03	37	37	1:120	5	1,17E+04
18	18	1:120	1	3,08E+03	38	38	1:120	3	6,15E+03
19	19	1:120	1	5,17E+03	39	39	1:120	4	1,25E+04
20	20	1:120	10	6,22E+03	40	40	1:120	4	4,14E+03

A tabela 9 possui 5 colunas que apresentam o código de cada treinamento, o número da entradas, a variação de neurônios na camada intermediária, o melhor número de neurônios na camada intermediária (considerando o melhor EQM para o número fixo de entradas) e o melhor erro quadrático médio encontrado para cada configuração (coluna EQMBest). Como pode-se notar, o melhor EQMBest encontrado foi $1,53E+3$ no treinamento CÒD. 09, com 9 entradas e 25 neurônios na camada intermediária (NEUROBEST).

A figura 13 apresenta graficamente a comparação entre o PLD médio mensal do submercado SE-CO (ou seja, o conjunto de testes, apresentado na figura 8) e a aproximação da rede para o mesmo no ano de 2016 (estimação da rede) do treinamento CÒD. 09, que tem NENTRADAS=9, NEUROBEST=25 e EQMBest= $1,53E+03$.

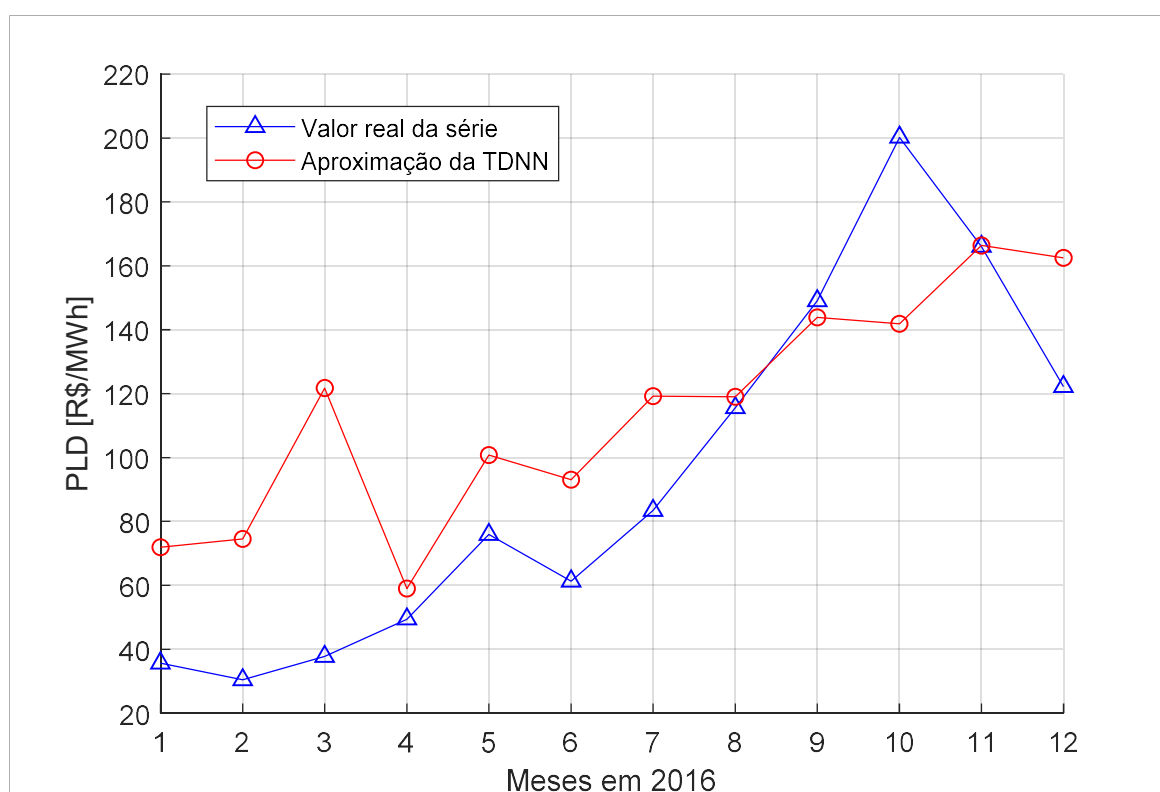


Figura 13 – Comparação entre valor real do PLD e aproximação da rede TDNN (CÒD. 09)

Fonte: Autoria própria

A tabela 10 apresenta os resultados obtidos no treinamento CÓD. 09.

Tabela 10 – Resultados treinamento CÓD. 09

CÓD. 09	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
PLD Real [R\$]	35,66	30,42	37,73	49,42	75,93	61,32	83,43	115,58	149,02	200,21	166,05	122,19
PLD Estimado [R\$]	71,90	74,50	121,76	58,97	100,76	93,09	119,23	118,98	143,86	141,87	166,36	162,47
Diferença [R\$]	36,24	44,08	84,03	9,55	24,83	31,77	35,80	3,40	-5,16	-58,34	0,31	40,28
Diferença [%]	101,64	144,89	222,71	19,32	32,71	51,81	42,91	2,94	-3,46	-29,14	0,19	32,96

4.2 PMC COM UMA CAMADA OCULTA

Finalizado o treinamento da rede TDNN, que utilizava como entrada a saída atrasada no tempo, treinou-se uma rede neural artificial *Perceptron* multicamadas com uma única camada intermediária, utilizando a mesma saída da rede anterior, porém na entrada, 28 séries de dados, contemplando as 8 variáveis utilizadas para todo o SIN (apresentadas no quadro 1). A tabela 11 apresenta o resultado obtido.

Tabela 11 – Resultado obtido no treinamento da rede PMC de uma camada (CÓD. 41)

PMC DE UMA CAMADA				
CÓD.	NENTRADAS	NEURO	NEUROBEST	EQMBEST
41	28	1:120	3	4,82E+03

A figura 14 apresenta a comparação gráfica entre o valor real da série do PLD do submercado SE-CO e a aproximação da melhor rede PMC de uma camada encontrada com 28 vetores de entrada, tendo 3 neurônios na camada intermediária e erro quadrático médio de 4,82E+03.

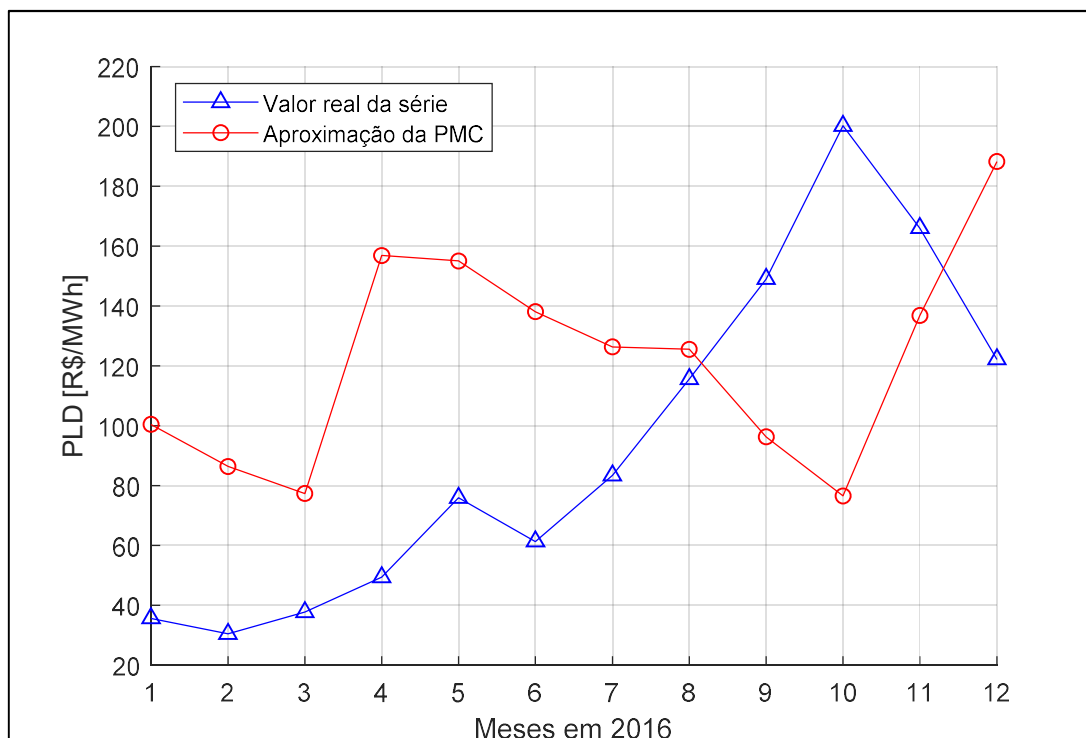


Figura 14 – Comparação entre valor real do PLD e aproximação da rede PMC (CÓD. 41)

Fonte: Autoria própria

A tabela 12 apresenta os resultados obtidos no treinamento CÓD. 41.

Tabela 12 – Resultados treinamento CÓD. 41

CÓD. 41	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
PLD Real [R\$]	35,66	30,42	37,73	49,42	75,93	61,32	83,43	115,58	149,02	200,21	166,05	122,19
PLD Estimado [R\$]	100,45	86,37	77,36	156,89	155,08	138,09	126,32	125,56	96,28	76,57	136,81	188,31
Diferença [R\$]	64,79	55,95	39,63	107,47	79,15	76,77	42,89	9,98	-52,74	-123,64	-29,24	66,12
Diferença [%]	181,68	183,93	105,02	217,46	104,24	125,20	51,41	8,64	-35,39	-61,76	-17,61	54,11

4.2 PMC COM DUAS CAMADAS OCULTAS

Após realizar o treinamento da rede PMC de uma camada, realizou-se o treinamento para a PMC de duas camadas. A tabela 13 apresenta os resultados obtidos para NENTRADAS=28, evidenciando o melhor EQM encontrado (CÓD. 42).

Tabela 13 – Resultado obtido no treinamento da rede PMC de duas camadas (CÓD. 42 a 46)

PMC DE DUAS CAMADAS (NENTRADAS=28)					
CÓD.	NEURO 1C	NEURO 2C	NEUROBEST 1C	NEUROBEST 2C	EQMbest
42	1:50	1:10	5	7	1,07E+03
43	1:50	11:20	18	16	1,75E+03
44	1:50	21:30	2	23	1,78E+03
45	1:50	31:40	9	35	1,25E+03
46	1:50	41:50	3	49	1,66E+03

A figura 15 apresenta a comparação gráfica entre o valor real da série do PLD do submercado SE-CO e a aproximação da melhor rede PMC de duas camadas encontrada com 28 vetores de entrada.

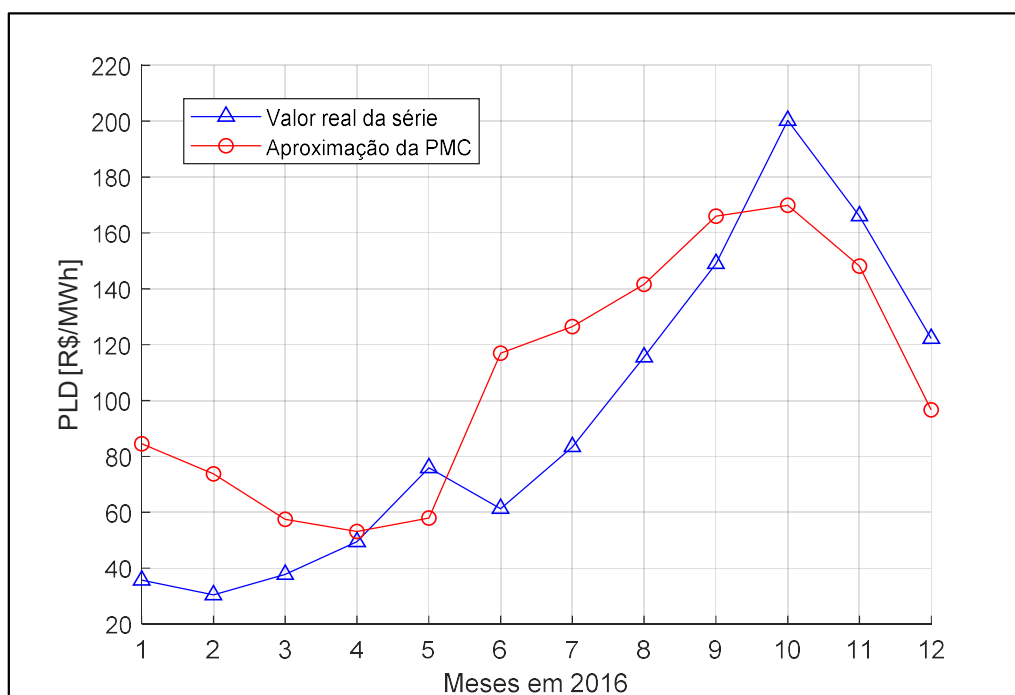


Figura 15 – Comp. entre valor real do PLD e aproximação da rede PMC (CÓD. 42)

Fonte: Autoria própria

O treinamento CÓD. 42, com melhor erro quadrático médio encontrado na primeira etapa de treinamentos da rede PMC de duas camadas, possui configuração com 5 neurônios na primeira camada intermediária, 7 neurônios da segunda camada intermediária e erro quadrático médio de 1,07E+03. A tabela 14 apresenta os resultados obtidos no treinamento CÓD. 42.

Tabela 14 – Resultados treinamento CÓD. 42

CÓD. 42	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
PLD Real [R\$]	35,66	30,42	37,73	49,42	75,93	61,32	83,43	115,58	149,02	200,21	166,05	122,19
PLD Estimado [R\$]	84,48	73,75	57,45	53,11	57,95	116,93	126,50	141,61	166,03	169,93	148,10	96,72
Diferença [R\$]	48,82	43,33	19,72	3,69	-17,98	55,61	43,07	26,03	17,01	-30,28	-17,95	-25,47
Diferença [%]	136,91	142,44	52,25	7,46	-23,68	90,69	51,63	22,52	11,42	-15,12	-10,81	-20,85

Após realizado o treinamento com as 28 séries da dados na entrada da rede, fez-se a seleção de variáveis pelo coeficiente de Pearson, afim de, além de diminuir as estimações iniciais necessárias para o processo, diminuir o erro quadrático médio da previsão do PLD.

4.3 PMC DUAS CAMADAS COM SELEÇÃO DE ENTRADAS

4.3.1 Entradas com coeficiente de Pearson abaixo de -0,4 ou acima de 0,4

Iniciou-se então a seleção de variáveis correlacionadas ao PLD. A primeira seleção com variáveis de coeficiente Pearson com módulo superior a 0,4 resultou em 16 vetores de entrada para a rede PMC de duas camadas (Tabela 15).

Tabela 15 – Variáveis com módulo do coef. de Pearson maior que 0,4 (continua)

CÓD.	VARIÁVEL - SUBMERCADO	COEF.
a1	Carga de Demanda - SE-CO	0,4851
a2	Carga de Demanda - S	0,4752
a3	Carga de Demanda - NE	0,4989
a4	Carga de Demanda - N	0,5021
b1	Carga de Energia - SE-CO	0,4436
b2	Carga de Energia - S	0,5013
b3	Carga de Energia - NE	0,5011
b4	Carga de Energia - N	0,5125
c2	Energia Natural Afluente - S	0,414
e1	Geração Hidráulica - SE-CO	-0,4473
e3	Geração Hidráulica - NE	-0,4592
f1	Geração Térmica - SE-CO	0,728
f2	Geração Térmica - S	0,6866
f3	Geração Térmica - NE	-0,4592

Tabela 15 – Variáveis com módulo do coef. de Pearson maior que 0,4 (continuação)

CÓD.	VARIÁVEL - SUBMERCADO	COEF.
h1	Níveis Hidrológicos - SE-CO	-0,6831
h3	Níveis Hidrológicos - NE	-0,5263

Na primeira coluna têm-se o código utilizado para nomear cada entrada, na segunda a variável e seu submercado e na terceira o coeficiente de Pearson que representa a correlação linear entre a variável e o PLD, calculado via MATLAB. A tabela 16 apresenta os resultados obtidos nos treinamentos da PMC de duas camadas com as 16 entradas selecionadas anteriormente (NENTRADAS=16).

Tabela 16 – Resultado obtido no treinamento da rede PMC de duas camadas (CÓD. 47 a 51)

PMC 2 CAMADAS P/ VARIÁVEIS COM COEF. DE PEARSON > 0,4 CORRELACIONADAS AO PLD DO SE-CO (NENTRADAS=16)					
CÓD.	NEURO 1C	NEURO 2C	NEUROBEST 1C	NEUROBEST 2C	EQMBEST
47	1:50	1:10	8	4	5,98E+02
48	1:50	11:20	21	11	1,09E+03
49	1:50	21:30	12	23	1,31E+03
50	1:50	31:40	7	34	1,52E+03
51	1:50	41:50	18	47	1,63E+03

A configuração da rede que obteve o menor erro quadrático médio de 5,98E+02 (CÓD. 47), possui 8 neurônios na primeira camada intermediária e 4 neurônios na segunda camada intermediária. A comparação gráfica entre o valor real da série do PLD e a aproximação da rede PMC de duas camadas com melhor EQMBest está apresentada na figura 16.

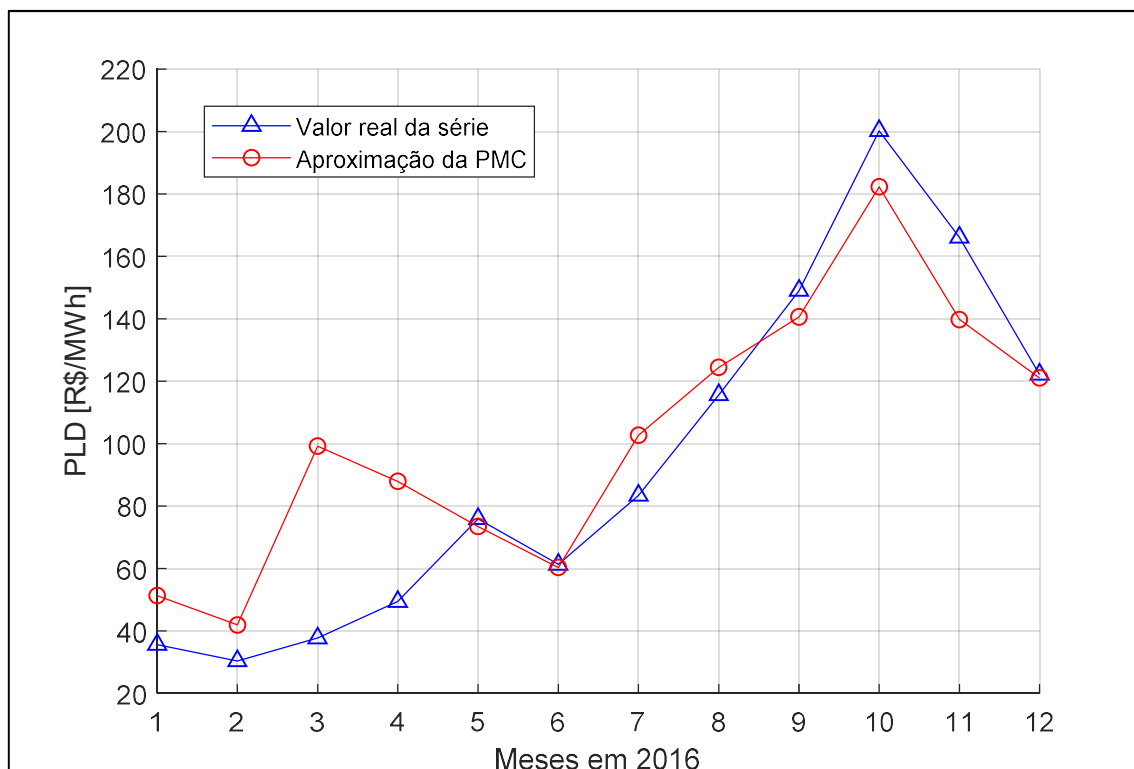


Figura 16 – Comp. entre valor real do PLD e aproximação da rede PMC (CÓD. 47)

Fonte: Autoria própria

A tabela 17 apresenta os resultados obtidos no treinamento CÓD. 47.

Tabela 17 – Resultados treinamento CÓD. 47

CÓD. 47	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
PLD Real [R\$]	35,66	30,42	37,73	49,42	75,93	61,32	83,43	115,58	149,02	200,21	166,05	122,19
PLD Estimado [R\$]	51,32	41,93	99,19	87,93	73,49	60,34	102,71	124,43	140,61	182,27	139,76	121,07
Diferença [R\$]	15,66	11,51	61,46	38,51	-2,44	-0,98	19,28	8,85	-8,41	-17,94	-26,29	-1,12
Diferença [%]	43,91	37,84	162,90	77,92	-3,22	-1,60	23,11	7,66	-5,65	-8,96	-15,83	-0,92

4.3.2 Entradas com coeficiente de Pearson abaixo de -0,5 ou acima de 0,5

Prosseguindo com a seleção gradativa de variáveis, fez-se a seleção de variáveis com módulo do coeficiente de Pearson superior a 0,5, resultando em 8 vetores de entrada apresentados na tabela 18.

Tabela 18 – Variáveis com módulo do coef. de Pearson maior que 0,5

CÓD.	VARIÁVEL - SUBMERCADO	COEF.
a4	Carga de Demanda - N	0,5021
b2	Carga de Energia - S	0,5013
b3	Carga de Energia - NE	0,5011
b4	Carga de Energia - N	0,5125
f1	Geração Térmica - SE-CO	0,728
f2	Geração Térmica - S	0,6866
h1	Níveis Hidrológicos - SE-CO	-0,6831
h3	Níveis Hidrológicos - NE	-0,5263

A tabela 19 apresenta os resultados obtidos nos treinamentos das redes PMC de duas camadas com os 8 vetores de entrada com módulo do coeficiente de Pearson maior que 0,5, selecionados anteriormente.

Tabela 19 – Resultado obtido no treinamento da rede PMC de duas camadas (CÓD. 52 a 56)

PMC 2 CAMADAS P/ VARIÁVEIS COM COEF. DE PEARSON > 0,5 CORRELACIONADAS AO PLD DO SE-CO (NENTRADAS=8)					
CÓD.	NEURO 1C	NEURO 2C	NEUROBEST 1C	NEUROBEST 2C	EQMBEST
52	1:50	1:10	28	10	4,33E+02
53	1:50	11:20	5	16	9,65E+02
54	1:50	21:30	41	27	1,89E+03
55	1:50	31:40	4	40	1,92E+03
56	1:50	41:50	2	46	1,17E+03

A configuração da rede que obteve o menor erro quadrático médio de 4,33E+02 (CÓD. 52), possui 28 neurônios na primeira camada intermediária e 10 neurônios na segunda camada intermediária. A comparação gráfica entre o valor real da série do PLD e a aproximação da rede PMC de duas camadas com melhor EQMBest está apresentada na figura 17.

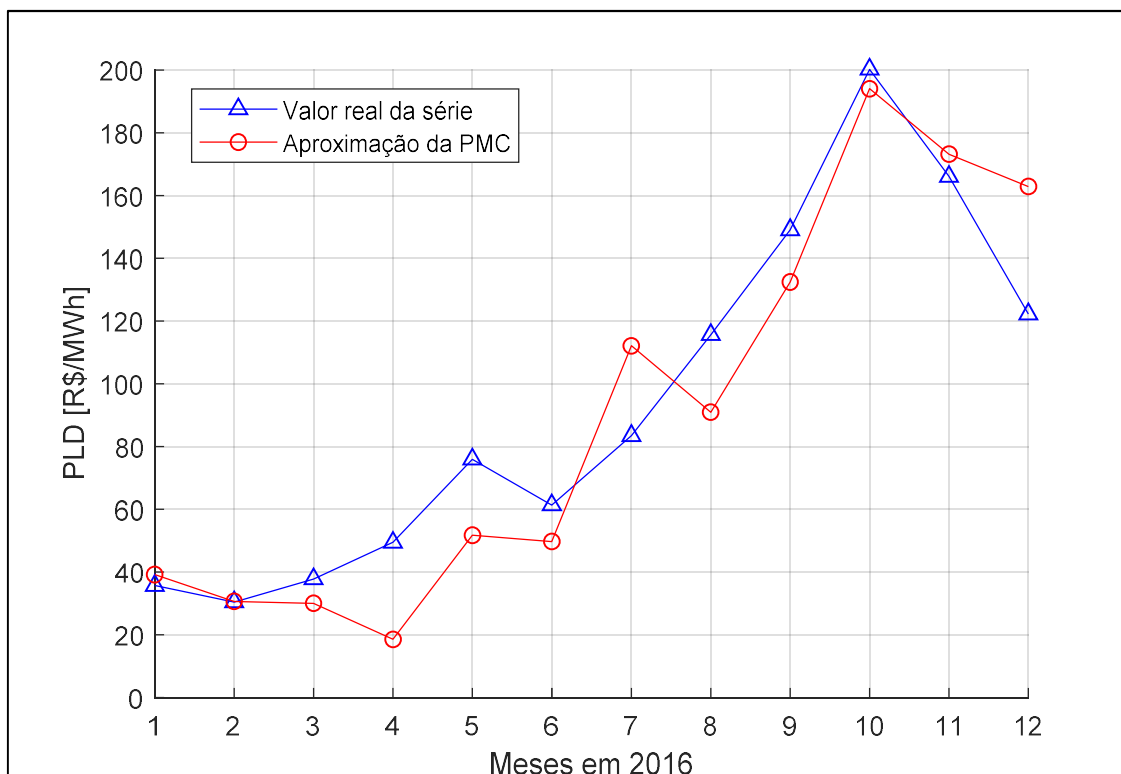


Figura 17 – Comp. entre valor real do PLD e aproximação da rede PMC (CÓD. 52)

Fonte: Autoria própria

A tabela 20 apresenta os resultados obtidos no treinamento CÓD. 52.

Tabela 20 – Resultados treinamento CÓD. 52

CÓD. 52	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
PLD Real [R\$]	35,66	30,42	37,73	49,42	75,93	61,32	83,43	115,58	149,02	200,21	166,05	122,19
PLD Estimado [R\$]	39,17	30,63	30,02	18,52	51,69	46,72	112,09	90,93	132,41	194,01	173,18	162,87
Diferença [R\$]	3,51	0,21	-7,71	-30,90	-24,24	-14,60	28,66	-24,65	-16,61	-6,20	7,13	40,68
Diferença [%]	9,84	0,68	-20,45	-62,52	-31,92	-23,82	34,35	-21,33	-11,14	-3,10	4,29	33,29

4.3.3 Entradas com coeficiente de Pearson abaixo de -0,6 ou acima de 0,6

Prosseguindo com a seleção gradativa de variáveis, fez-se a seleção de variáveis com módulo do coeficiente de Pearson superior a 0,6, resultando em 3 vetores de entrada apresentados na tabela 21.

Tabela 21 – Variáveis com módulo do coef. de Pearson maior que 0,6

CÓD.	VARIÁVEL - SUBMERCADO	COEF.
f1	Geração Térmica SE-CO	0,728
f2	Geração Térmica S	0,6866
h1	Energia Armazenada SE-CO	-0,6831

A tabela 22 apresenta os resultados obtidos nos treinamentos das redes PMC de duas camadas com os 3 vetores de entrada com módulo do coeficiente de Pearson maior que 0,6, selecionados anteriormente.

Tabela 22 – Resultado obtido no treinamento da rede PMC de duas camadas (CÓD. 57 a 61)

PMC CAMADAS P/ VARIÁVEIS COM COEF. DE PEARSON > 0,6 CORRELACIONADAS AO PLD DO SE-CO (NENTRADAS=3)					
CÓD.	NEURO 1C	NEURO 2C	NEUROBEST 1C	NEUROBEST 2C	EQMBEST
57	1:50	1:10	49	7	1,15E+03
58	1:50	11:20	14	20	1,18E+03
59	1:50	21:30	10	23	2,54E+03
60	1:50	31:40	6	40	1,83E+03
61	1:50	41:50	8	43	3,48E+03

A configuração da rede que obteve o menor erro quadrático médio de 1,55E+03 (CÓD. 57), possui 49 neurônios na primeira camada intermediária e 7 neurônios na segunda camada intermediária. A comparação gráfica entre o valor real da série do PLD e a aproximação da rede PMC de duas camadas com melhor EQMBest está apresentada na figura 18.

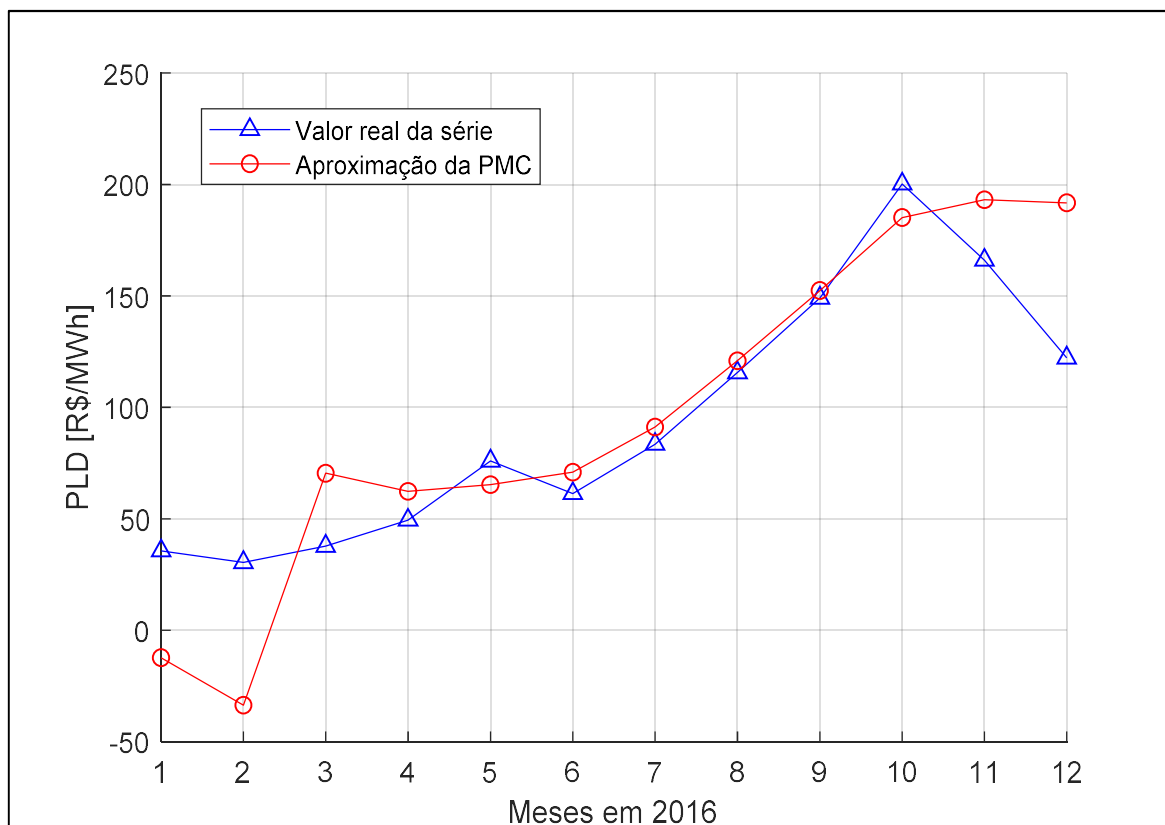


Figura 18 – Comp. entre valor real do PLD e aproximação da rede PMC (CÓD. 57)

Fonte: Autoria própria

A tabela 23 apresenta os resultados obtidos no treinamento CÓD. 57.

Tabela 23 – Resultados treinamento CÓD. 57

CÓD. 57	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
PLD Real [R\$]	35,66	30,42	37,73	49,42	75,93	61,32	83,43	115,58	149,02	200,21	166,05	122,19
PLD Estimado [R\$]	-12,30	-33,62	70,45	62,32	65,31	70,92	91,17	120,90	152,41	185,17	193,14	191,78
Diferença [R\$]	-47,96	-64,04	32,72	12,90	-10,62	9,60	7,74	5,32	3,39	-15,04	27,09	69,59
Diferença [%]	-134,49	-210,51	86,73	26,11	-13,98	15,65	9,28	4,60	2,27	-7,51	16,32	56,95

4.3.4 Entradas com coeficiente de Pearson abaixo de -0,7 ou acima de 0,7

Na última variação para PMC de duas camadas, fez-se a seleção de variáveis com módulo do coeficiente de Pearson superior a 0,7, resultando em apenas um vetor de entrada, apresentado na tabela 24.

Tabela 24 – Variável com módulo do coef. de Pearson maior que 0,7

CÓD.	VARIÁVEL - SUBMERCADO	COEF.
f1	Geração Térmica SE-CO	0,728

A tabela 25 apresenta os resultados obtidos nos treinamentos das redes PMC de duas camadas com o único vetor de entrada com módulo do coeficiente de Pearson maior que 0,7.

Tabela 25 – Resultado obtido no treinamento da rede PMC de duas camadas (CÓD. 62 a 66)

PMC 2 CAMADAS P/ VARIÁVEIS COM COEF. DE PEARSON > 0,7 CORRELACIONADAS AO PLD DO SE-CO (NENTRADAS=1)					
CÓD.	NEURO 1C	NEURO 2C	NEUROBEST 1C	NEUROBEST 2C	EQMBEST
62	1:50	1:10	40	2	2,32E+03
63	1:50	11:20	2	20	2,97E+03
64	1:50	21:30	2	24	2,80E+03
65	1:50	31:40	2	36	6,49E+03
66	1:50	41:50	5	45	4,46E+03

A configuração da rede que obteve o menor erro quadrático médio de 2,33E+03 (CÓD. 62), possui 40 neurônios na primeira camada intermediária e 2 neurônios na segunda camada intermediária. A comparação gráfica entre o valor real da série do PLD e a aproximação da rede PMC de duas camadas com melhor EQMBest está apresentada na figura 19.

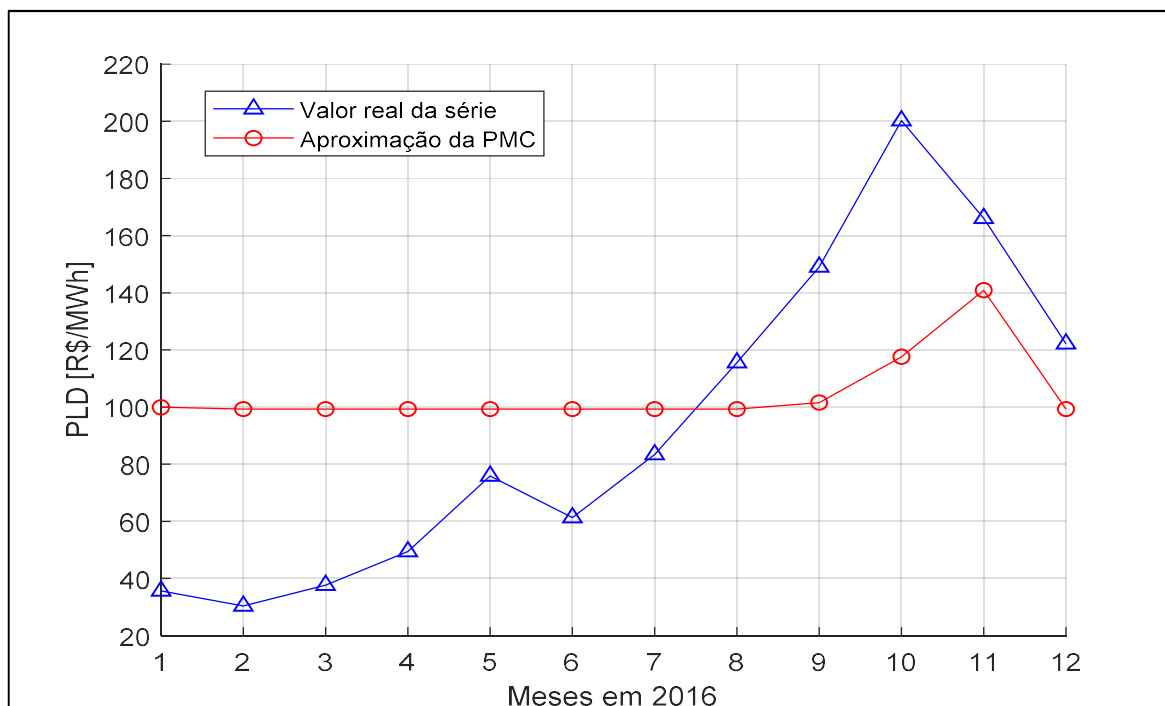


Figura 19 – Comp. entre valor real do PLD e aproximação da rede PMC (CÓD. 62)

Fonte: Autoria própria

A tabela 26 apresenta os resultados obtidos no treinamento CÓD. 62.

Tabela 26 – Resultados treinamento CÓD. 62

CÓD. 62	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
PLD Real [R\$]	35,66	30,42	37,73	49,42	75,93	61,32	83,43	115,58	149,02	200,21	166,05	122,19
PLD Estimado [R\$]	99,94	99,29	99,29	99,29	99,29	99,29	99,29	99,29	101,56	117,57	140,91	99,29
Diferença [R\$]	64,28	68,87	61,56	49,87	23,36	37,97	15,86	-16,29	-47,46	-82,64	-25,14	-22,90
Direção [%]	180,27	226,41	163,17	100,92	30,77	61,93	19,01	-14,09	-31,85	-41,28	-15,14	-18,74

4.4 PMC UMA CAMADA COM SELEÇÃO DE ENTRADAS

Para efeito de comparação realizou-se a seleção de entradas por coeficiente de Pearson para a rede PMC de uma camada. O resultado obtido nos treinamentos estão apresentados na tabela 27

Tabela 27 – Resultado obtido no treinamento da rede PMC de uma camada (CÓD. 67 a 70)

PMC DE 1 CAMADA P/ MÓDULO DO COEF. DE PEARSON MAIOR QUE 0,4, 0,5, 0,6 E 0,7				
CÓD.	NENTRADAS	NEURO	NEUROBEST	EQMBEST
67	16	1:120	2	2,78E+03
68	8	1:120	1	1,36E+03
69	3	1:120	2	3,40E+03
70	1	1:120	2	8,59E+03

A configuração da rede PMC com uma camada oculta com menor erro quadrático médio encontrado (1,36E+03), tem oito entradas e um neurônio na camada intermediária.

O melhor resultado obtido (CÓD. 68) para a última etapa de treinamentos está apresentado na comparação entre o valor real da série do PLD e a aproximação da rede PMC com uma camada oculta e entradas com coeficiente de Pearson menores que -0,7 ou maiores que 0,7 em relação ao PLD do submercado SE-CO (Figura 20).

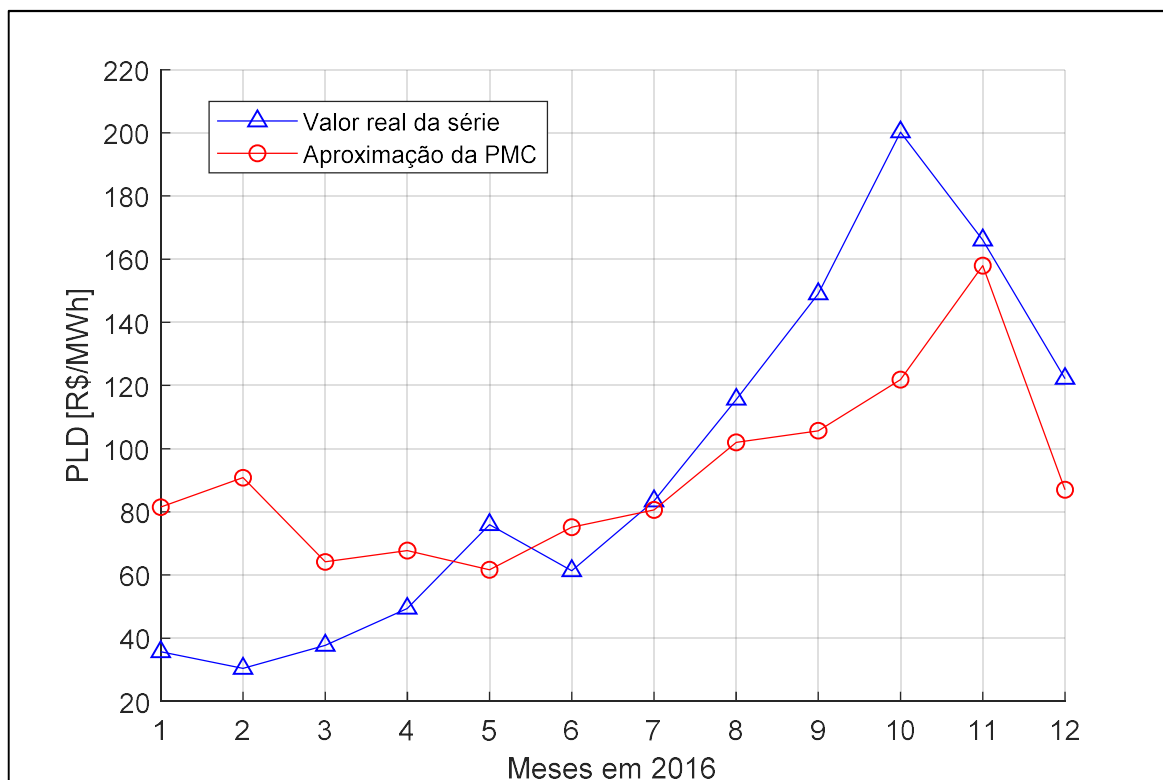


Figura 20 – Comp. entre valor real do PLD e aproximação da rede PMC (CÓD. 68)

Fonte: Autoria própria

A tabela 28 apresenta os resultados obtidos no treinamento CÓD. 68, encerrando a etapa de apresentação de resultados.

Tabela 28 – Resultados treinamento CÓD. 68

CÓD. 68	JAN	FEV	MAR	ABR	MAI	JUN	JUL	AGO	SET	OUT	NOV	DEZ
PLD Real [R\$]	35,66	30,42	37,73	49,42	75,93	61,32	83,43	115,58	149,02	200,21	166,05	122,19
PLD Estimado [R\$]	81,48	90,81	64,14	67,72	61,68	75,16	80,60	101,99	105,68	121,87	157,96	86,99
Diferença [R\$]	45,82	60,39	26,41	18,30	-14,25	13,84	-2,83	-13,59	-43,34	-78,34	-8,09	-35,20
Diferença [%]	128,48	198,52	70,01	37,03	-18,77	22,57	-3,40	-11,75	-29,08	-39,13	-4,87	-28,81

5 CONCLUSÕES E CONSIDERAÇÕES FINAIS

Iniciou-se os treinamentos pela topologia de rede TDNN com configuração focada, com um grande erro quadrático médio encontrado ($1,53E+3$) e erros de previsão que mesmo inferiores a 5% nos meses de agosto, setembro e novembro, foram altos na maioria dos restante dos meses chegando a 222% no mês de março entre o PLD real e estimado como apresentado na tabela 10. Concluiu-se que o problema de estimação do PLD não é dependente de seu próprio histórico de valores.

Na tentativa de melhorar o desempenho de estimação, utilizou-se em seguida uma topologia PMC de uma camada, que diferente da anterior, agora contava com 28 séries de dados como entrada ao invés de apenas a saída (PLD) atrasada no tempo como a topologia anterior. O resultado porém não se mostrou positivo, com erro quadrático médio encontrado de $4,38E+3$, essa topologia teve um erro superior ao encontrado na rede TDNN de configuração focada.

Iniciou-se então os treinamentos na rede PMC de duas camadas intermediárias, que superou o erro das duas topologias citadas anteriormente mesmo antes da seleção de variáveis de entrada através do coeficiente de Pearson. A tabela 29 apresenta o melhor resultado encontrado para cada topologia.

Tabela 29 – Melhor resultado encontrado para cada topologia

MELHORES TOPOLOGIAS DE RNA				
CÓD.	NENTRADAS	NEUROBEST (1C 2C)		EQMBEST
09	9	25		$1,53E+03$
41	28	3		$4,82E+03$
42	28	5	7	$1,07E+03$
47	16	8	4	$5,98E+02$
52	8	28	10	$4,33E+02$
57	3	49	7	$1,15E+03$
62	1	40	2	$2,32E+03$
68	8	1		$1,36E+03$

Como apresentado na Tabela 29, a rede PMC de duas camadas intermediárias reduziu o erro quadrático médio inicialmente para $1,07E+3$, apresentando valores de PLD com erro inferior a 10% em metade dos meses do ano, porém ainda apresentou erro pontual de 162% no mês de março (tabela 21). Iniciou-se então a seleção de variáveis através do coeficiente de Pearson (treinamentos CÓD.

47 em diante) confirmando a existência de correlação linear (moderada e forte) entre as variáveis de entrada da rede PMC levadas em consideração no trabalho e o PLD.

O melhor resultado encontrado (EQMBEST = 433) foi para variáveis de entrada com coeficiente de correlação linear (coeficiente de Pearson) inferiores a -0,5 e superiores a 0,5, resultando em oito entradas (treinamento CÓD. 52). A partir daí a seleção de variáveis fez com que o erro de estimação ao decorrer dos meses começasse a aumentar, a Tabela 30 apresenta o maior e menor erro encontrado para cada topologia e também os valores médios obtidos para o ano de teste.

Tabela 30 – Menores e maiores erros mensais e resultados médios

CÓD.	MENOR ERRO ENCONTRADO		MAIOR ERRO ENCONTRADO		DIFERENÇA MÉDIA		PLD MED. REAL	PLD MÉD. ESTIMADO	ERRO PLD MÉDIO
	%	Mês	%	Mês	R\$	%	R\$	R\$	%
09	0,19	NOV	222,71	MAR	20,57	51,62	93,91	114,48	21,90
41	8,64	AGO	217,46	ABR	28,09	76,41	93,91	122,01	29,92
42	7,46	ABR	90,69	JUN	13,80	37,07	93,91	107,71	14,69
47	-1,60	JUN	162,90	MAR	8,17	26,43	93,91	102,09	8,71
52	0,68	FEV	-62,52	ABR	-3,73	-7,65	93,91	90,91	-3,19
57	2,27	SET	-210,51	FEV	2,56	-12,38	93,91	96,47	2,73
62	-14,09	AGO	226,41	FEV	10,61	55,11	93,91	104,53	11,31
68	-3,40	JUL	198,52	FEV	-2,57	26,73	93,91	91,34	-2,74

Como apresentado na Tabela 30, o treinamento CÓD. 52 não só apresentou o menor erro quadrático médio, como também apresentou os menores erros no decorrer dos doze meses do ano de teste, apresentando erros inferiores a 1% e maior erro de 62% que chegou a essa proporção apenas no mês de Abril (figura 21). Mesmo que o valor médio tenha ficado mais próximo do real para os treinamentos 57 e 68, estes tiveram erros de até 200% ao decorrer do ano, o que tornaria a estimação do PLD, que é um valor mensal, impossível de ser utilizada para fins práticos.

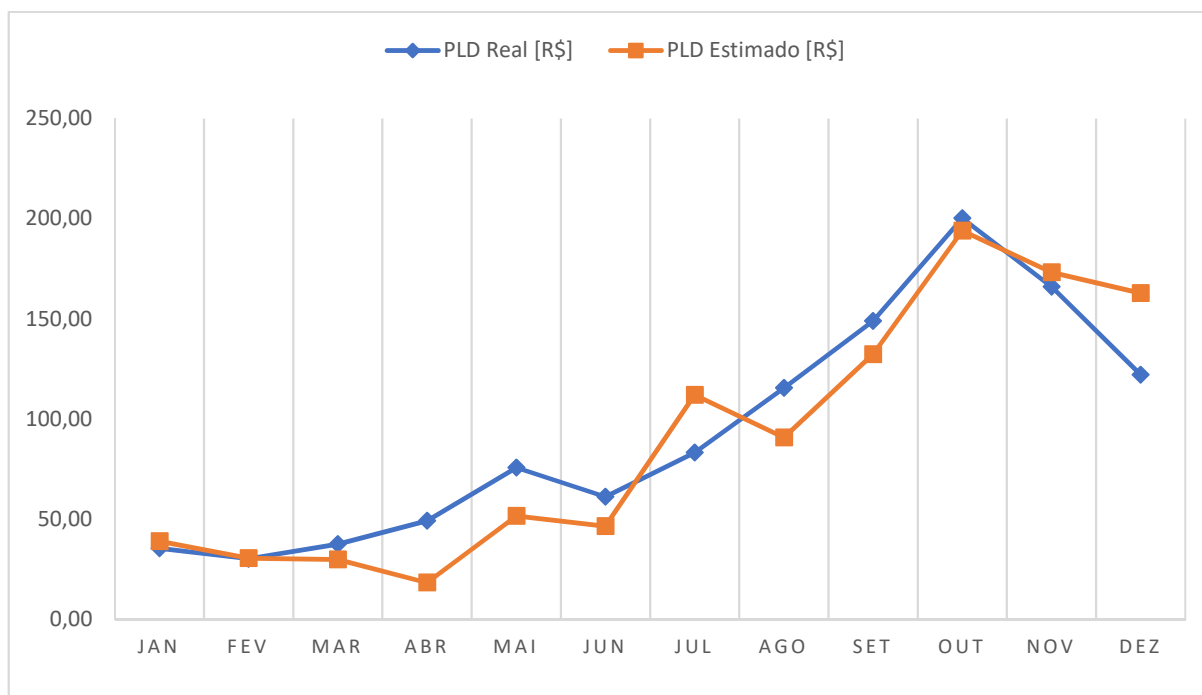


Figura 21 – Melhor estimaco encontrada (CD. 52)

Fonte: Autoria prpria

Na figura 21 confirmou-se tambm que a estimaco do treinamento 52 consegue prever (com melhor desempenho que os outros treinamentos) o comportamento do PLD ao decorrer do ano, que s destoa da realidade entre os meses de maro e abril e entre julho e agosto, onde o valor estimado decresce enquanto o valor real cresce.

Concluiu-se por fim que a seleo de variveis auxiliou de modo eficiente a estimaco do PLD mesmo sabendo que as variveis possuem correlaces no lineares, dado que o erro foi reduzido em mais de 50% do melhor treinamento realizado com todas as 28 entradas (CD. 42) para o treinamento com oito variveis de entrada (CD. 52). Verificou-se tambm na PMC de uma camada, que mesmo no tendo atingido resultados to bons quanto a rede de duas camadas intermedirias, com as mesmas oito entradas superou a qualidade da estimativa da rede TDNN (CD.09) e d prpria PMC de uma camada sem seleo de variveis (CD. 41), confirmando a existncia de correlaco linear entre as variveis de entrada levadas em considerao no trabalho e o PLD.

5.1 Trabalhos Futuros

Trabalhos futuros podem ser realizados tanto para aprimorar as ferramentas utilizadas quanto na aplicação de novas ferramentas para estimar o PLD, substituindo a RNA por outra ferramenta de inteligência artificial. Além disso, o foco da estimação pode ser alterado para os outros três submercados, ou até mesmo para o SIN num todo.

Como forma de aprimorar as ferramentas utilizadas, têm-se as mudanças nas topologias de RNA utilizadas, como testar diferentes topologias com número de camadas e neurônios diferentes, e também a forma de utilizar e tratar as variáveis, podendo ter acréscimo de novas entradas e também uma análise de correlação diferente da utilizada para variáveis não correlacionadas linearmente.

REFERÊNCIAS

ANEEL, **Limites do PLD para 2016**. 2015. Disponível em:<
http://www2.aneel.gov.br/aplicacoes/noticias/Output_Noticias.cfm?Identidade=9010&id_area=>. Acesso em 15 de Outubro de 2016.

ALANIS, Alma y.. **Electricity Prices Forecasting using Artificial Neural Networks**. IEEE LATIN AMERICA TRANSACTIONS, VOL. 16, NO. 1, Janeiro de 2018. Universidade de Guadalajara, México, 2018.

CCEE - CÂMARA DE COMERCIALIZAÇÃO DE ENERGIA ELÉTRICA. **Visão geral das operações na CCEE**. Manual de operações. 2010.

CCEE, **Preços médios – PLD**. 2016. Disponível em: <https://www.ccee.org.br/portal/faces/pages_publico/o-que-fazemos/como_ccee_atua/precos/precos_medios?_afLoop=1238293706122064#%40%3F_afLoop%3D1238293706122064%26_adf.ctrl-state%3Dpvg0prh6r_105>. Acesso em 13 de Outubro de 2016.

CRUZ, Matheus Palma. **Estratégias de ofertas em mercados competitivos de energia elétrica com predominância de geração hidrelétrica**. 2017. 131 f. Tese (Doutorado em Engenharia Elétrica) - Universidade Federal de Santa Catarina, Florianópolis, 2017.

LARSON, RON. e FARBER, BETSY. **Estatística Aplicada**. Editora Pearson, 4ed. São Paulo, 2009.

LEITE, A. L. S.; CASTRO, N. J. **Preço spot de eletricidade: teoria e evidências do caso brasileiro**. IV Encontro de Economia Catarinense, Criciúma, 2010.

LOPES, L. F. D. **Apostila Estatística**. Departamento de Engenharia. Universidade Federal de Santa Maria, 2003.

SILVA, I.N., SPATTI D., e FLAUZINO, R. **Redes neurais artificiais para engenharia e ciências aplicadas: curso prático**. Artliber Editora Ltda, São Paulo, 2010.

SOUZA, F. C.; LEGEY, L. F. L.. **Brazilian Electricity Market Structure and Risk Management Tools**. IEEE PES General Meeting, Pittsburgh, 2014.

THEODORO, Edson A. R.. **Introdução às RNAs e a rede Perceptron**, 15 de maio de 2019. Notas de Aula, curso de Engenharia Elétrica. Universidade Federal Tecnológica do Paraná. Cornélio Procópio, 2019.

VITORINO, Roney N.. **Percepção de preços e aferição dos riscos de exposição financeira no ambiente de contratação livre**. Dissertação de mestrado, Universidade Politécnica de São Paulo. São Paulo, 2011.