UNIVERSIDADE FEDERAL DA PARAÍBA CENTRO DE ENERGIAS ALTERNATIVAS E RENOVÁVEIS DEPARTAMENTO DE ENGENHARIA ELÉTRICA

Matheus Henrique Gomes de Freitas

Previsão de Demanda em Subestações Utilizando Inteligência Artificial Considerando a Potência e a Energia

João Pessoa 2019 Matheus Henrique Gomes de Freitas

Previsão de Demanda em Subestações Utilizando Inteligência Artificial Considerando a Potência e a Energia

Trabalho de Conclusão de Curso apresentado à Universidade Federal da Paraíba como exigência para a obtenção do título de Bacharel em Engenharia Elétrica.

Universidade Federal da Paraíba Centro de Energias Alternativas e Renováveis Curso de Graduação em Engenharia Elétrica

Orientador: Prof. Dr. Juan Moises Maurício Villanueva

João Pessoa

2019

© Matheus Henrique Gomes de Freitas

Catalogação na publicação Seção de Catalogação e Classificação

F866p Freitas, Matheus Henrique Gomes de. Previsão de Demanda em Subestações Utilizando Inteligência Artificial Considerando a Potência e a Energia / Matheus Henrique Gomes de Freitas. - João Pessoa, 2019. 81 f. : il. Orientação: Juan Moises Mauricio Villanueva. TCC (Especialização) - UFPB/CEAR. 1. Previsão de Demanda. 2. Distribuição de Energia Elétrica. 3. Previsão de Séries Temporais. 4. Redes Neurais Artificiais. 5. Inteligência Artificial. I. Villanueva, Juan Moises Mauricio. II. Título. UFPB/BC Matheus Henrique Gomes de Freitas

Previsão de Demanda em Subestações Utilizando Inteligência Artificial Considerando a Potência e a Energia

Trabalho de Conclusão de Curso apresentado à Universidade Federal da Paraíba como exigência para a obtenção do título de Bacharel em Engenharia Elétrica.

Trabalho aprovado. João Pessoa, 20 de setembro de 2019:

Prof. Dr. Juan Moises Maurício Villanueva Orientador

Prof. Dr. Helon David de Macêdo Braz Examinador Interno UFPB

Prof. Dr. Euler Cássio Tavares de Macedo Examinador Interno UFPB

MSc. Marcelo Renato de Cerqueira Paes Junior Examinador Externo

> João Pessoa 2019

Agradecimentos

Agradeço primeiramente a Deus por tudo que Ele tem feito, e por ter me permitido chegar até aqui.

Aos meus pais, pelo apoio indescritível e constante durante toda a minha existência. Sem vocês, jamais teria alcançado meus objetivos.

A todos os demais membros da família que se fizeram presente na minha trajetória.

Aos amigos que fiz na graduação Bruno, Felipe do Ó, Felipe Vidal, Gabriel, Isaac, João Victor, Jonyelison, Henrique, Lucas, Manuella, Matheus Silveira, Mikaelly, Sávio, Thiago e Vinícius, pela companhia nos momentos de estudos e nos momentos de lazer, vocês fizeram meus dias de graduação mais felizes.

Ao professor Juan Maurício Villanueva, por ter sido um orientador extremamente solícito e compreensivo.

A todos os professores por todo o ensinamento transmitido, que foi essencial para minha formação profissional.

Por fim, a todos os que contribuíram direta ou indiretamente para a realização deste trabalho.

Se trabalharmos sobre o mármore, um dia ele acabará. Se trabalharmos sobre o metal, um dia o tempo o consumirá. Se erguermos templos, um dia se tornarão pó. Mas se trabalharmos sobre almas jovens e imortais, se nós as imbuirmos com os princípios do justo temor ao criador e amor à humanidade, daqui a cem anos pouco importará o quanto tenhamos acumulado no banco, que tipo de casa, palacete ou carro possuímos. Mas o mundo poderá ser diferente, talvez porque fomos importantes na vida dos jovens. (Frank Sherman Land)

Resumo

A previsão de demanda é uma ferramenta fundamental no planejamento das ações de uma concessionária de energia elétrica, pois existe a necessidade de alocar os seus recursos com antecedência para expandir o seu sistema de distribuição, realizar manutenção, substituição e compra de equipamentos, além de adotar estratégias no mercado de energia. Com esta premissa, foi desenvolvida uma ferramenta computacional no software MATLAB® capaz de realizar a previsão de demanda utilizando redes neurais artificias para auxiliar especialistas da área de planejamento estratégico de sistemas de distribuição de energia elétrica na tomada de decisões. Na metodologia proposta, optou-se por utilizar a energia acumulada e variações da energia como variável externa, com a finalidade de melhorar o desempenho da previsão. A ferramenta de previsão proposta foi utilizada na previsão de demanda da subestação de João Pessoa, onde foram utilizados dados reais de medições de potência ativa entre os anos de 2008 e 2013, a rede neural foi treinada com dados de 2008 a 2012 para prever o ano de 2013. A previsão foi comparada com os dados reais, onde foram calculadas algumas métricas como erro médio quadrático, erro absoluto e erro percentual médio absoluto para aferir a acurácia da previsão.

Palavras-chave: Previsão de Demanda, Distribuição de Energia Elétrica, Previsão de Séries Semporais, Redes Neurais Artificiais, Inteligência Artificial.

Abstract

Demand forecasting is an essential tool in planning the actions of a eletric utility, as there is a need to allocate their resources in advance to expand its distribution system, perform maintenance, replace and purchase equipment, in addition to adopt strategies in the energy market. With this premise, a computational tool was developed in the MATLAB® software capable of forecasting demand using artificial neural networks to assist specialists in the strategic planning area of power distribution systems in decision making. In the proposed methodology, it was decided to use the energy as an external variable in order to improve the forecasting performance. The proposed prediction tool was used to forecast the demand of the João Pessoa substation, where real active power measurement data from 2008 to 2013 were used, the neural network was trained with data from 2008 to 2012 to predict the year of 2013. The forecast was compared with actual data, where some metrics such as mean square error, absolute error and absolute mean percentage error were calculated to measure the accuracy of the forecast.

Keywords: Demand Forecasting, Electric Power Distribution, Time Series Prediction, Artificial Neural Networks, Artificial Intelligence.

Lista de ilustrações

Figura	1	_	Representação de um sistema elétrico de potência	20
Figura	2	_	Composição da matriz energética brasileira.	21
Figura	3	_	Sistemas Isolados - Ciclo 2018	22
Figura	4	_	Sistemas Interligado Nacional para o horizonte de 2017	23
Figura	5	_	Diagrama unifilar simplificado de uma subestação de distribuição	25
Figura	6	_	Sistema de aquisição de dados.	26
Figura	7	_	Exemplos de séries temporais: a) leitura a cada duas horas de uma	
			concentração "x" em processo químico; b) variação diária do preço das	
			ações da IBM; c) leitura de temperatura em processo químico a cada	
			minuto; d) leitura horária da viscos idade em processo químico	26
Figura	8	_	Perfil de carga de um alimentador da subestação de João Pessoa	28
Figura	9	_	Neurônio biológico.	32
Figura	10	_	Neurônio artificial.	33
Figura	11	_	Rede alimentada adiante com camada única	36
Figura	12	_	Rede alimentada adiante com múltiplas camadas	37
Figura	13	_	Rede alimentada adiante com múltiplas camadas	38
Figura	14	_	Sentido de propagação de sinais de entrada e de sinais de erro em uma	
			rede neural	40
Figura	15	_	Exemplo de previsão de séries temporais	42
Figura	16	_	Esquema de uma RNA com memória de linha de atraso	43
Figura	17	_	Diagrama simplificado da subestação JPS	44
Figura	18	_	Série temporal antes da correção de <i>outliers</i>	47
Figura	19	_	Série temporal após a correção de <i>outliers</i>	47
Figura	20	_	Dados de potência máxima semanal da SE JPS, alimentador 21L1, de	
			2008 até 2013	51
Figura	21	_	Entradas e saídas da RNA e janelamento	52
Figura	22	_	Comportamento da potência máxima semanal, energia acumulada na	
			semana e variação de energia entre semanas para um período de um ano.	54
Figura	23	_	Valores médios de MAPE das redes do primeiro modelo	58
Figura	24	_	Previsão da rede de menor MAPE e objetivo	59
Figura	25	_	Valores médios de RMSE das redes do primeiro modelo	60
Figura	26	_	Valores médios de MAPE das redes do segundo modelo	61
Figura	27	_	Comparação da previsão da rede de menor MAPE e objetivo de saída.	61
Figura	28	_	Valores médios da RMSE das redes do segundo modelo	62
Figura	29	_	Comparação da previsão da rede de menor RMSE e objetivo de saída	63

Figura 30 – Comparação entre o erro absoluto da rede de menor MAPE com o erro	
da rede de menor RMSE. $\ldots \ldots $	53
Figura 31 – Valores médios de MAPE das redes do terceiro modelo 6	35
Figura 32 – Comparação da previsão da rede de menor MAPE e objetivo de saída. 6	35
Figura 33 – Valores médios da RMSE das redes do terceiro modelo 6	36
Figura 34 – Comparação da previsão da rede de menor RMSE e objetivo de saída $$ 6	37
Figura 35 – Comparação entre o erro da rede de menor MAPE com o erro da rede	
de menor RMSE. \ldots 6	37
Figura 36 – Valores médios de MAPE das redes do quarto modelo 6	38
Figura 37 – Comparação da previsão da rede de menor MAPE e objetivo de saída. 6	39
Figura 38 – Valores médios da RMSE das redes do segundo modelo	70
Figura 39 – Comparação da previsão da rede de menor RMSE e objetivo de saída 7	71
Figura 40 – Comparação entre o erro da rede de menor MAPE com o erro da rede	
de menor RMSE. \ldots 7	71
Figura 41 – Comparação entre o resultado da melhor rede com e sem realimentação. 7	74

Lista de tabelas

Tabela 1 –	Organização do banco de dados utilizado na previsão de demanda	48
Tabela 2 –	Organização do banco com dados diários de potência máxima e energia	
	acumulada	49
Tabela 3 –	Organização do banco com dados semanais de potência máxima, energia	
	acumulada e variação da energia	50
Tabela 4 –	Tabela de dados de entrada para treino da RNA do terceiro modelo. $\ .$	53
Tabela 5 –	Tabela com os alvos de saída do treino da RNA do terceiro modelo. $\ .$	53
Tabela 6 –	Especificações técnicas da máquina utilizada	57
Tabela 7 $$ –	Valores das métricas de desempenho para a rede de melhor MAPE e	
	RMSE do primeiro modelo.	60
Tabela 8 –	Valores das métricas de desempenho para a rede de menor MAPE e	
	para a rede de menor RMSE do segundo modelo	64
Tabela 9 –	Valores das métricas de desempenho para a rede de menor MAPE e	
	para a rede de menor RMSE do terceiro modelo	68
Tabela 10 –	Valores das métricas de desempenho para a rede de menor MAPE e	
	para a rede de menor RMSE do quarto modelo	72
Tabela 11 –	Comparação das melhores redes de cada um dos modelos propostos $\ .$	73
Tabela 12 –	Exemplo de como foi feita a realimentação da rede	73

Lista de abreviaturas e siglas

ANEEL	Agência Nacional de Energia Elétrica
ARIMA	Autoregressive Integrated Moving Average
EA	Erro absoluto
EPE	Empresa de Pesquisa Energética
ET	Estação Transformadora
ER	Erro relativo
HP	Horizonte de Previsão
IA	Inteligência Artificial
ID	Identificador
INNS	International Neural Networks Society
JP	Janela de Previsão
JPS	João Pessoa
LTLF	Long-term Load Forecasting
MAPE	Erro percentual médio absoluto
MCP	Memória de curto prazo
MTLF	Medium-term Load Forecasting
n1	Número de neurônios na primeira camada
n2	Número de neurônios na segunda camada
ONS	Operador Nacional do Sistema Elétrico
PB	Paraíba
PRODIST	Procedimentos de Distribuição
RMSE	Raiz do erro médio quadrático
RNA	Rede Neural Artificial

RR Roraima

SCADA Supervisory Control and Data Acquisition

- SEP Sistema Elétrico de Potência
- SIN Sistema Interligado Nacional
- SVM Support Vector Machine
- STLF Short-term Load Forecasting

Lista de símbolos

∂	Derivada parcial
β	Letra grega Beta
Δ	Letra grega Delta
θ	Letra grega Theta
ρ	Letra grega Rho
Σ	Letra grega Sigma
Φ	Letra grega Phi
R	Marca registrada

Sumário

1	INTRODUÇÃO	16
1.1	Pertinência e motivação do trabalho	16
1.2	Objetivos	18
1.3	Organização do trabalho	18
2	FUNDAMENTAÇÃO TEÓRICA	20
2.1	O Sistema Elétrico de Potência	20
2.2	Séries temporais	26
2.3	Técnica de previsão utilizada neste trabalho	29
2.4	Redes Neurais Artificiais	30
2.5	Breve histórico	31
2.6	O Neurônio Artificial	32
2.6.1	Tipos de função de ativação tradicionais	34
2.7	Arquitetura das redes neurais	35
2.7.1	Redes alimentadas adiante com camada única	36
2.7.2	Redes alimentadas adiante com múltiplas camadas	36
2.7.3	Redes recorrentes	38
2.8	Treinamento da RNA	38
2.8.1	Algoritmo de retropropagação	39
2.8.2	Algoritmo de retropropagação resiliente	40
2.8.3	Algoritmo de Levenberg-Marquardt	41
2.9	Previsão de séries temporais utilizando RNA	42
3	METODOLOGIA UTILIZADA PARA PREVISÃO DE DEMANDA .	44
3.1	Objeto de estudo	44
3.2	Visão geral	45
3.3	Implementação do banco de dados	45
3.4	Treinamento da RNA	50
3.5	Correlação das variáveis	53
3.6	Validação e teste da RNA	55
3.7	Parâmetros de treinamento	56
4	RESULTADOS	58
4.1	Modelo considerando apenas a potência máxima semanal	58
4.2	Modelo considerando a potência máxima semanal e a energia acu-	
	mulada na semana	60

4.3	Modelo considerando a potência máxima semanal e variação de	
	energia entre semanas	64
4.4	Modelo considerando a potência máxima semanal, a energia acu-	
	mulada na semana e variação de energia entre semanas	68
4.5	Comparação entre os modelos propostos	72
4.6	Teste de robustez da melhor rede	73
5	CONCLUSÃO	75
6	REFERÊNCIAS BIBLIOGRÁFICAS	77

1 Introdução

1.1 Pertinência e motivação do trabalho

Prever o futuro sempre foi um desejo dos seres humanos. Desde épocas antigas, quando se tentava prever um período de chuva ou a época de uma colheita, o homem buscava predizer o futuro para tomar decisões no presente. Uma das formas mais elementares de se fazer isso é observar, estudar e analisar informações passadas. Prever comportamentos futuros é uma ferramenta de grande importância em diversas áreas do conhecimento, tais como: economia, meteorologia, comércio e demanda de energia elétrica.

A demanda por energia elétrica vivencia um notável crescimento ao redor do mundo. Particularmente no Brasil, no ano de 2018, por exemplo, aproximadamente 1 milhão de novas unidades consumidoras foram incorporadas à base de consumidores residenciais (EPE, 2019). Além disso, estudos indicam que, entre 2016 e 2026, o consumo de energia elétrica deve crescer em média 3,7% ao ano (EPE, 2017). Sem contar que a energia elétrica foi a modalidade energética mais consumida no Brasil em 2017 (EPE, 2018a).

Frente à necessidade de se atender a estas demandas, empresas do setor elétrico são cada vez mais pressionadas a seguir os mais rigorosos padrões de qualidade, sob pena de multas da parte de órgãos regulatórios. Estes órgãos exigem das concessionárias de energia elétrica relatórios periódicos de previsão de demanda das cargas localizadas na sua área de concessão em horizontes de previsão pré-definidos (ANEEL, 2018).

Além dos estudos exigidos pelos órgãos regulatórios, as empresas de distribuição de energia procuram fazer análises internas para que seja possível alocar os recursos disponíveis com antecedência. Entre esses recursos estão a compra de novos equipamentos, a ampliação e construção de linhas de transmissão, as manutenções programadas, a compra e venda de energia, construção de novas plantas de geração, etc. Isto leva à necessidade de se adotar boas práticas de gestão a fim de melhorar a eficiência e reduzir custos.

Nesse contexto, a previsão de demanda surge como uma importante ferramenta no planejamento estratégico de sistemas de potência. A previsão de demanda de energia elétrica pode ser classificada, com base no horizonte de previsão, em três categorias: previsão de longo prazo (LTLF — *Long-term load forecasting*), previsão de médio prazo (MTLF — *Medium-term load forecasting*) e previsão de curto prazo (STLF — *Short-term load forecasting*) (KHATOON, 2014). A previsão de demanda de longo prazo refere-se a previsões com horizonte de anos ou décadas, e objetiva o planejamento para incrementar a capacidade das linhas de transmissão, construção de novas plantas de geração, contrução de subestações, entre outros. A previsão de demanda de médio prazo envolve previsões de um período de semanas até um ano e permite às concessionárias e empresas de energia elétrica alocar seus recursos como: manutenções nas redes elétricas, a compra ou substituição de equipamentos, expansão do seu sistema, estratégias no mercado de energia (PAN; LEE, 2012). A previsão de demanda de curto prazo refere-se a um horizonte de previsão de alguns minutos até alguns dias, neste tipo de previsão pretende-se avaliar a segurança do sistema, a sua confiabilidade, planejamento e execução de manutenções, também é muito importante na regulação entre a demanda e oferta por energia elétrica a fim de evitar flutuações, ajustando a geração de energia à constantes mudanças nas cargas (CARMONA *et al.*, 2002).

Antes do surgimento de métodos baseados em inteligência artificial (IA), apenas métodos estatísticos e probabilísticos eram utilizados para a previsão de séries temporais, com destaque para o modelo Box e Jenkins (BOX; JENKINS; REINSEL, 2008). O desenvolvimento de técnicas baseadas em inteligência artificial, como as Redes Neurais Artificiais (RNA) e a Lógica Fuzzy, ganharam notoriedade pela capacidade de trabalhar com grandes quantidades de dados não lineares e por não dependerem de modelos matemáticos complexos (GHANBARI *et al.*, 2010). Essas técnicas são inspiradas em aspectos biológicos de aprendizagem e raciocínio, e são aplicadas em diferentes campos da ciência e da engenharia (MAKRIDAKIS; WHEELWRIGHT; HYNDMAN, 1998). Além desses, foram desenvolvidos métodos híbridos, que combinavam dois ou mais métodos de previsão, sendo eles estatísticos ou baseados em técnicas de inteligência artificial.

Apesar de todas as classes de métodos citados conseguirem, ao seu modo, alcançar resultados satisfatórios, estas técnicas continuam sendo aperfeiçoadas e novos métodos têm sido desenvolvidos com a finalidade de melhorar a exatidão e a confiabilidade da previsão, haja vista que alguns estudos mostram que um aumento de 1% no erro de previsão causa um aumento de milhões de reais por ano nos custos operacionais de uma empresa de distribuição de energia elétrica (GHANBARI *et al.*, 2010; HOBBS *et al.*, 1999).

Como explicado anteriormente, diversas abordagens de modelos de previsão de demanda são apresentadas na literatura, entre elas as mais promissoras são aquelas baseadas em Redes Neurais Artificiais (RNA), com a capacidade de relacionar múltiplas variáveis e permitindo a criação de modelos não lineares. (SINGH; HUSSAIN; BAZAZ, 2017). Com o intuito de aperfeiçoar a previsão de demanda de energia elétrica baseada em RNAs, além de dados históricos da demanda de energia elétrica, pode-se trabalhar com algumas de suas variáveis de influência. São vários os fatores que influenciam na demanda de energia, e identificar todas essas variáveis é um desafio. Desta maneira, neste trabalho será analisada o impacto do uso da energia acumulada e variações de energia nos modelos de previsão baseadas em RNA.

1.2 Objetivos

Este trabalho tem como objetivo geral desenvolver uma ferramenta computacional para previsão de demanda, destinada a sevir como embasamento para especialistas da área de planejamento estratégico em sistemas de distribuição, com enfoque em redes neurais artificiais. Além disso, houve a análise da influência da utilização da energia na previsão, com uma posterior comparação entre os resultados considerando ou não esta variável, para concluir se a inclusão dessa variável, de fato, aumenta a acurácia da previsão.

Como objetivos específicos deste trabalho, podemos citar:

- Aplicação de técnicas de inteligência artificial para a construção de modelos de previsão;
- Uso de banco de dados de medições de potências associadas a subestações de energia elétrica;
- Verificação da influência da energia e das variações da energia nos modelos de previsão;
- Comparação do modelo proposto com modelos anteriores de previsão de demanda, sem consideração da energia e variação da energia.

1.3 Organização do trabalho

Este trabalho foi estruturado em cinco capítulos. No primeiro capítulo, será feita uma introdução, abordando a importância e motivação do tema. Além disso, foram citados os objetivos gerais e específicos.

No segundo capítulo, é realizada uma fundamentação teórica e considerações a respeito do sistema elétrico de potência, apontando características dos sistemas de geração, transmissão e distribuição de energia elétrica. Também serão tratados conceitos relacionados à previsão de séries temporais, com enfoque na utilização de redes neurais para realizar esta previsão, apresentando as principais características de uma rede neural, bem como seu príncipio de funcionamento.

O terceiro capítulo aborda a metodologia de previsão utilizada, partindo da descrição do objeto de estudo. Dando continuidade, será exposta a forma de implementação do banco de dados, treinamento da rede neural e seus parâmetros, com a fundamentação dos testes e da validação da rede neural proposta. Por fim, há um tópico abordando a correlação entre os dados de demanda, da energia acumulada e variações da energia, apontando a motivação para a escolha dessa variável para ajudar na previsão. No quarto capítulo, os resultados serão apresentados. Para facilitar o entendimento, os resultados dos modelos serão apresentados separadamente, com apresentação de gráficos da previsão das melhores redes e dos valores das métricas de desempenho. Após, isso será feita uma comparação entre os modelos.

No quinto capítulo constará da conclusão do trabalho, na qual é verificado se objetivos traçados foram alcançados, se a escolha da variável da energia de fato aumentou a acuracia da previsão de demanda e são sugerido trabalhos futuros que poderiam ter como base este trabalho.

Por fim, serão apontadas as referências bibliográficas.

2 Fundamentação teórica

Para realizar a previsão de demanda de energia elétrica é fundamental conhecer o fenômeno com que se está trabalhando. Neste capítulo será feita uma fundamentação téorica sobre o sistema elétrico de potência e seus subsistemas, sobre séries temporais e seus principais componentes, bem como sobre técnicas de previsão de séries temporais, com enfoque nas redes neurais artificiais, que foi a técnica de previsão de demanda utilizada neste trabalho.

2.1 O Sistema Elétrico de Potência

A principal função do Sistema Elétrico de Potência (SEP) é fornecer energia elétrica aos usuários, com qualidade apropriada, na medida em que essa for solicitada, convertendo energia de naturezas diversas (hidráulica, mecânica, térmica, etc.) em energia elétrica e assim, distribuindo aos consumidores (KAGAN; OLIVEIRA; ROBBA, 2005). Um sistema elétrico de potência pode ser subdivididos em três grandes subsistemas, são eles: geração, transmissão e distribuição. Na Figura 1, é possível visualizar a representação de um sistema elétrico de potência, assim como os consumidores a ele conectados.





Fonte: adaptado de (ANEEL, 2008).

O subsistema de geração é responsável por converter alguma forma de energia em energia elétrica. A geração de energia elétrica se faz através diversas fontes que compõem uma matriz energética, essas fontes podem ser hidráulica, eólica, nuclear, solar, dentre outras. Na Figura 2 pode-se observar a distribuição da matriz energética brasileira, que é composta predominantemente por fontes de energia renovável, que representam 80,4% (EPE, 2018a) da oferta interna de eletricidade, considerando tanto a produção nacional quanto a importação de energia, que é essencialmente de origem renovável. Das fontes que compõem a matriz energética brasileira destaca-se a geração hidrelétrica que corresponde a 65,2% (EPE, 2018a) da oferta interna de eletricidade, isso se deve aos grandes recursos hídricos distribuídos ao longo de boa parte do território.





Fonte: (EPE, 2018a).

A energia gerada pelo subsistema de geração necessita ser distribuída em todo o país e, para isso, existe o subsistema de transmissão, que também é responsável por elevar o nível de tensão da geração para níveis de tensão compatíveis com a transmissão. Este subsistema, por sua vez, é dividido em dois grandes blocos: os Sistemas Isolados e o Sistema Interligado Nacional (SIN).

Os sistemas isolados recebem esse nome por não estarem conectados ao SIN, devido a condições geográficas dessas localidades que impedem esta conexão, e, consequentemente, pelo impedimento de permutar energia elétrica com outras regiões. Segundo a EPE (2018b) existem 270 sistemas isolados no Brasil, distribuídos em 8 estados, maioria dos deles situam-se na região Norte, onde destaca-se a cidade de Boa Vista-RR que é a única capital do país ainda não conectada ao SIN. Na Figura 3, é possível visualizar a disposição dos sistemas isolados no país, representados pelos pontos em amarelo.



Figura 3 – Sistemas Isolados - Ciclo 2018.

Fonte: EPE, 2018b.

O SIN abrange quase a totalidade do território brasileiro. Estando presente em todas as regiões do Brasil, o SIN concentra aproximadamente 900 linhas de transmissão que se estendem por cerca de 90 mil quilômetros, além disso abriga 96,6% de toda a capacidade de produção de energia elétrica do país (ANEEL, 2008). A grande dimensão da rede de transmissão no Brasil é justificada pela configuração do sistema de geração, formado, em sua maior parte, por usinas hidrelétricas instaladas em locais distantes das grandes centrais consumidoras. Na Figura 4, ilustra-se o SIN para o horizonte de 2017.



Figura 4 – Sistemas Interligado Nacional para o horizonte de 2017.

Fonte: ONS, 2017.

Segundo a EPE (2018c), as principais funções do SIN são: (i) a transmissão da energia gerada pelas usinas para os grandes centros de carga; (ii) a integração entre os diversos elementos do sistema elétrico para garantir estabilidade e confiabilidade da rede; (iii) a interligação entre as bacias hidrográficas e regiões com características hidrológicas heterogêneas de modo a otimizar a geração hidrelétrica; e (iv) a integração energética com os países vizinhos. O ONS é o órgão responsável pelo controle da operação e pela coordenação do SIN, sob regulamentação e fiscalização da ANEEL.

Quando a energia elétrica proveniente da transmissão chega próximo aos grandes centros de consumo o subsistema de distribuição faz a distribuição dessa energia para os consumidores finais. É através dos denominados pontos de fronteira em que os subsistemas de transmissão e distribuição são conectados. Nos pontos de fronteira existem subestações que abaixam a tensão para os níveis de subtransmissão. O sistema de distribuição é de suma importância, pois além de ser responsável pela entrega de energia ao usuário final, demanda elevada quantidade de investimentos, assim como qualidade e garantia de serviço destinado aos consumidores. Dessa forma, o planejamento estratégico torna-se indispensável, buscando se adequar ao crescimento de carga e um melhor aproveitamento dos recursos financeiros.

O sistema de distribuição de energia elétrica no Brasil é regulado por um conjunto de resoluções da ANEEL, que através das normas elaboradas nos Procedimentos de Distribuição (PRODIST) normatizam e padronizam as atividades técnicas relacionadas ao funcionamento e ao desempenho dos sistemas de distribuição de energia elétrica. A finalidade do PRODIST é disciplinar o planejamento expansão e operação do sistema, a medição e a qualidade da energia elétrica, de forma a garantir que os sistemas de distribuição operem com segurança, eficiência, qualidade e confiabilidade.

Conforme Kagan *et al.*, (2005), o sistema de distribuição pode ser dividido nos seguintes blocos:

- Sistema de subtransmissão: este tem a função de captar a energia das subestações de transmissão e transferi-la para as subestações de distribuição e aos consumidores em tensão de subtransmissão, usualmente, 138 kV ou 69 kV ou, mais raramente, em 34,5 kV. Os consumidores em tensão de subtransmissão são representados, usualmente, por grandes instalações industriais, estações de tratamento e bombeamento de água.
- Sistema de distribuição primária: As redes de distribuição primária, ou de média tensão, emergem das subestações de distribuição e operam, no caso da rede aérea, radialmente, com possibilidade de transferência de blocos de carga entre circuitos para o atendimento da operação em condições de contingência, devido à manutenção corretiva ou preventiva. Estas redes operam usualmente com tensão de 13,8 kV e atendem aos consumidores primários e aos transformadores de distribuição, estações transformadoras (ETs) que suprem a rede secundária, ou de baixa tensão. Dentre os consumidores primários destacam-se indústrias de porte médio, conjuntos comerciais, instalações de iluminação pública, etc. Podem ser aéreas ou subterrâneas, as primeiras de uso mais difundido, pelo seu menor custo, e as segundas, encontrando grande aplicação em áreas de maior densidade de carga, por exemplo zona central de uma metrópole, ou onde há restrições paisagísticas.
- Redes de distribuição secundária: As redes de distribuição primária, ou de média tensão, emergem das estações transformadoras, onde deriva-se a rede de baixa tensão, 220/127 V ou 380/220 V, que pode operar em malha ou radial e que supre os consumidores baixa tensão, consumidores residenciais, pequenos comércios e indústrias. Alcança, por circuito, comprimentos da ordem de centenas de metros. Destaca-se o predomínio, nesta rede, de consumidores residenciais.

Pode-se dizer que as subestações de distribuição, que rebaixam o nível de tensão para 13,8 kV e fornecem energia elétrica a rede de distribuição primária "alimentam" parte da rede de distribuição, dessa maneira, os circuitos de saída dessas subestações são denominados de alimentadores. Na Figura 5, pode-se visualizar um diagrama unifilar simplificado de uma subestação de distribuição que abaixa o nível de tensão de 69 kV para 13,8 kV e possui 4 alimentadores na saída.

Figura 5 – Diagrama unifilar simplificado de uma subestação de distribuição.



Fonte: MEDEIROS, 2016.

As subestações de distribuição de energia são monitoradas por um Sistema de Aquisição de Dados (SCADA - *Supervisory Control and Data Acquisition*) que coleta dados de natureza elétrica (tensão, corrente, fator de potência, etc.), de diferentes pontos de medição localizados em equipamentos (disjuntores e religadores) e barramentos (MEDEI-ROS, 2016). Os dados obtidos nestas medições são enviados a um servidor que, por sua vez, armazena-os em um banco de dados. Essas medições, que são realizadas ao longo do tempo com periodicidade de 15 minutos, são denominadas séries temporais. Na Figura 6, é ilustrada a conexão entre as medições nas subestações à um sistema SCADA.



Figura 6 – Sistema de aquisição de dados.

Fonte: TEPCO, 2015.

2.2 Séries temporais

Série temporal pode ser definida como uma sequência ordenada de dados coletados em intervalos de tempo constantes durante determinado período. Para ser classificada como uma série temporal os dados deverão apresentar dependência entre as observações vizinhas (EHLERS, 2009). De modo com que dados de uma variável aleatória em um instante t, contenha informações que possibilitem determinar o valor desta variável no instante t+1. Uma vasta quantidade de fenômenos de natureza física, biológica, econômica, entre outras, estão inclusos nesta categoria e possuem suas características estudadas através da análise de séries temporais (FAVA, 2000). Na Figura 7, é possível visualizar exemplos reais de séries temporais.

Figura 7 – Exemplos de séries temporais: a) leitura a cada duas horas de uma concentração "x" em processo químico; b) variação diária do preço das ações da IBM; c) leitura de temperatura em processo químico a cada minuto; d) leitura horária da viscosidade em processo químico.



Fonte: MEDEIROS, 2016.

A maneira tradicional de analisar uma série temporal é através da sua decomposição em quatro componentes não observáveis: tendência, sazonalidade, ciclo e erro. Segundo essa abordagem, as séries podem ser representadas como a combinação desses quatro componentes básicos (MORETTIN, 1981).

- A tendência de uma série apresenta o comportamento dos dados no tempo, ou seja, se a sequência de dados é crescente, decrescente ou estável, além da velocidade de tais variações.
- A sazonalidade aponta a recorrência de um padrão em uma série dentro de um período. Esse padrão corresponde a um movimento oscilatório que exprime a influência de fatores cujo comportamento é periódico, que pode ser diário, semanal, anual, etc.
- Ciclo é um deslocamento aleatório de longa duração que expõe a influência de fatores causais. Ele aponta fases de contração e expansão de uma série.
- Erro é um movimento instável que indica a influência de fatores aleatórios.

Um dos métodos de análise de séries temporais que é extensivamente utilizado é a previsão. As previsões, de forma geral, oferecem estimativas de grande relevância em processos de tomada de decisão (FRANCO JUNIOR, 2013). No campo da engenharia elétrica, um caso típico de previsão de séries temporais é a previsão de carga e/ou demanda. Os termos carga e demanda serão considerados equivalentes em todo este trabalho, tendo em vista a definição encontrada em ANEEL (2005) que estabelece que carga é a caracterização da demanda do sistema. Na Figura 8, é ilustrado o perfil de carga de um alimentador da subestação de João Pessoa, na Paraíba, no ano de 2008.



Figura 8 – Perfil de carga de um alimentador da subestação de João Pessoa.

Fonte: elaborado pelo autor, 2019.

Na Figura 8 é apresentada uma quinzena de medições de demanda do alimentador 21L1 da subestação de João Pessoa. É possível notar um padrão que se repete diariamente, com demanda mais baixa na madrugada (entre zero hora e seis da manhã) e mais elevada durante o dia, tendo uma leve queda por volta do meio dia e atingindo o valor máximo do dia no ínicio da noite. Além disso, o consumo é similar entre os dias úteis da semana, mas cai aos finais de semana. Temos, portanto, uma sazonalidade de extensão diária e outra semanal. Sem contar, que se fosse traçada uma curva com as medições de demanda anuais também seria possível notar uma sazonalidade anual. É importante que o modelo de previsão de demanda de energia elétrica seja capaz de tratar a complexa estrutura multi-sazonal de série temporal fruto da demanda de energia elétrica.

Segundo Guirelli (2006), desde a década de 80 diversos modelos tem sido propostos para a realização da previsão de carga. Segundo Khatoon *et al.* (2014), as técnicas de previsão podem ser divididas em duas amplas categorias: as técnicas de previsão tradicionais e as técnicas de previsão baseadas em inteligência artificial.

As técnicas tradicionais de previsão de carga fazem uso de modelos estatísticos, que envolvem equações matemáticas para predizer valores futuros. Essas técnicas possuem a vantagem de serem técnicas matemáticas já bem desenvolvidas e estudadas, elas são rápidas e representam bem séries lineares, porém, dependem de um modelo matemático complexo, além de ter dificuldade de representar séries não-lineares (NIU *et al.*, 2010). Dentre as técnicas de previsão tradicionais destacam-se:

- Regressão Linear;
- ARIMA (Autoregressive Integrated Moving Average);
- Suavização exponencial;
- Análise espectral;

Já as técnicas de previsão de carga baseadas em inteligência artificial tratamse de algoritmos cujo princípio de funcionamento é inspirado em aspectos biológicos de aprendizagem e raciocínio. Esses algoritmos devem ser capazes de armazenar conhecimento, adquirir novo conhecimento através da experiência e aplicar o conhecimento armazenado para resolver problemas. Apesar dos métodos baseados em inteligência artificial não dependerem de um modelamento matemático complexo ou conhecimento profundo do sistema, os mesmos apresentam suas dificuldades. A codificação do conhecimento humano não é tão trivial, ainda existe dificuldade na determinação das variáveis de entrada e nos parâmetros e os processos de treinamento são complexos e exigem esforço computacional e tempo (MEDEIROS, 2016). Dentre as técnicas de previsão baseadas em inteligência artificial destacam-se:

- Sistemas especialistas;
- Redes Neurais Artificias;
- Lógica Fuzzy;
- Algoritmos genéticos.

Adicionalmente, pode-se considerar outra categoria de técnicas de previsão de carga, a das técnicas híbridas, que utilizam duas ou mais técnicas simultaneamente, com a finalidade de aperfeiçoar os resultados das previsões. As técnicas híbridas mais utilizadas são a Neuro-Fuzzy e a Neural-genético.

2.3 Técnica de previsão utilizada neste trabalho

Nos últimos anos, os métodos mais utilizados para previsão de carga no setor elétrico baseiam-se principalmente em técnicas de inteligência artificial (CHEMETOVA; SANTOS; NEVES, 2016). Isso porque essas técnicas apresentam uma ótima maneira para lidar com incertezas, não linearidades e relações de alta complexidade. Existem diversas publicações científicas que comprovam a qualidade e a robustez de previsões de demanda baseadas nessas técnicas (SKOLTHANARAT *et al.*, 2014).

Dentre as técnicas de previsão baseadas em inteligência artificial, a técnica baseada em redes neurais emergiu como o método de previsão que detém a maior parte da atenção e é o mais utilizado na literatura julgando pela quantidade de publicações (BALIYAN *et al.*, 2015; HIPPERT *et al.*, 2001). Sendo assim, este trabalho concentrou-se na utilização das Redes Neurais Artificiais (RNAs) para realizar a previsão de carga, devido aos motivos elencados a seguir:

- Muitos anos de estudos publicados na área;
- Capacidade de lidar com modelos complexos e não lineares;
- Por ser uma técnica renomada e, portanto, mais facilmente aceita pelo setor elétrico;
- Implementação relativamente simples;
- Disponibilidade de um conjunto amplo de dados para a aprendizagem das RNAs.

2.4 Redes Neurais Artificiais

O cérebro humano é um sistema de processamento de informações altamente complexo, não-linear e paralelo. (HAYKIN, 2001). Ele processa informações de maneira totalmente diferente de um computador digital convencional, a sua capacidade de aprencer e de reconhecer, associar e generalizar padrões serviram como motivação para a criação de um modelo computacional baseado em uma rede de neurônios biológicos, esse modelo computacional é denominado rede neural artificial, que pode ser definida como um sistema paralelo, composto por unidades de processamento simples, denominadas neurônios, dispostas em camadas e altamente interligadas, inspiradas no cérebro humano.

Uma metodologia que é geralmente adotada para fazer previsão utilizando RNA em que os dados são baseados em séries temporais é a separação de uma parte da série para o processo de treinamento e uma parte para o processo de validação. Da parte de treinamento é realizada a extração dos componentes de tendência e ciclo da série. A previsão é feita com base nos ciclos da série e o resultado final é obtido adicionando-se o componente da tendência, que deve ser extrapolado.

O resultado da previsão é então comparado com a parte da série temporal separada para validação, onde avalia-se se a previsão se aproxima dos dados reais. Isso demonstra que o modelo de redes neurais é uma alternativa para a realização de previsões, porque possui a capacidade de aprendizagem através da generalização, com base em registros do histórico da série.

2.5 Breve histórico

O surgimento das redes neurais artificiais se deu na decáda de 40, mais precisamente em 1943, ano em que Warren McCulloch e Walter Pitts publicaram em artigo cujo resultado foi tão significativo que é considerado o nascimento dos estudos sobre inteligência artificial e redes neurais (HAYKIN, 2001).

No modelo proposto por McCulloh e Pitts os neurônios recebem um conjunto de entradas ponderadas por pesos, a soma dessas entradas ponderadas é submetida a uma função de ativação que determina se a soma é maior que um valor númerico, se sim o neurônio é ativado, caso contrário o neurônio é desativado (MCCULLOH; PITTS, 1943). De forma simplificada, o que o neurônio faz é avaliar se a soma recebida é maior ou menor que um valor numérico. Em seu trabalho eles também sugeriram que redes de neurônios definidas adequadamente seriam capazer de aprender (NORVIG; RUSSELL, 2004). Apesar de exetremamente simples, o modelo proposto por McCulloh e Pitts influenciou o trabalho de diversos outros pesquisadores.

Em 1949, Donald Hebb propôs o princípio de que a intensidade das ligações entre os neurônios se alteram em resposta a experiência. Ele sugeriu que quando dois neurônios são ativados simultaneamente a conexão entre eles é fortalecida, enquanto outras conexões tendem a se enfraquecer (HEBB, 1949). Essa hipótese estimulou a evolução da teoria da aprendizagem de redes neurais artificiais.

Em 1957, Rosenblatt apresentou o *perceptron* uma das primeiras RNAs com apenas um neurônio e aprendizagem supervisionada (ROSENBLATT, 1957). O ajuste de pesos das conexões da rede proposta eram realizados da seguinte forma: se a saída for igual a saída desejada, os pesos das conexões não sofrem alteração, caso contrário, os pesos são ajustados repassando a eles valores atualizados em função da dimenção do erro na saída.

Ao mesmo tempo em que Rosenblatt trabalhava no *perceptron*, Widrow desenvolveu um novo modelo de processamento de redes neurais chamado de Adaline (*Adaptive Linear Element*), a qual se destacava pela sua poderosa lei de aprendizado. O princípio de treinamento para as redes Adalines ficou conhecido como a Regra Delta, o qual consiste em obter o ponto de mínimo, através de um processo de iteração local, que foi mais tarde generalizada para redes com modelos neurais mais sofisticados. Mais tarde, Widrow criou a Madaline (*Multiple Adaptive Linear Element*), que era uma generalização multidimensional do Adaline.

Apesar do sucesso do perceptron, a pesquisa sobre redes neurais passou a conviver com um grande problema, a expectativa exagerada criada pelos próprios pesquisadores da área, não acompanhada de resultados à altura, acelerou a queda de financiamentos para pesquisa. Entretanto foi em 1969, após a publicação do livro de Minsky e Papert, que as pesquisas na área de redes neurais sofreram uma retração significativa. Como resultado, os autores demonstraram que o *perceptron*, apesar de ser capaz de executar as operações booleanas AND e OR, não é capaz de executar outras operações elementares,como XOR (MINSKY; PAPERT, 1969).

Foi apenas em 1986 que as pesquisas sobre redes neurais cresceram consideravelmente, isso aconteceu devido a publicação de Rumelhart e McClelland onde eles apresentam um modelo computacional que permite o treinamento supervisionado dos neurônios artificiais (RUMELHART *et al.*, 1986). O algoritmo de treinamento ficou conhecido como retropropagação (*backpropagation*), um algoritmo de otimização global e sem restrições. As redes de retropropagação aprendem de maneira mais lenta, pois necessitam, possivelmente, de milhares de iterações para aprender, já que elas se tratam de um método multicamadas, mas geram um resultado muito preciso.

Em 1987 ocorreu a Primeira Conferência de Redes Neurais. Em 1989 foi criada a Sociedade Internacional de Redes Neurais (International Neural Networks Society - INNS) juntamente com o INNS Journal, já em 1990 foram criados o IEEE Transactions on Neural Networks e o Neural Computation. Desde então, muitas instituições formaram institutos de pesquisa e programas de educação em neurocomputação.

2.6 O Neurônio Artificial

O neurônio é o elemento básico do cérebro humano, ele é uma célula especializada no processamento e transmissão de informações. Nele são identificadas três partes principais: o corpo celular, os dendritos e o axônio. Na Figura 9, é possível visualizar um neurônio biológico.





Fonte: BEZERRA, 2016.

A função dos dendritos é receber informações, ou impulsos nervosos, provenientes de outros neurônios e encaminha-los até o corpo celular, já a função do corpo celular é coletar, combinar e processar as informações recebidas pelos dendritos e gerar novas informações que serão transmitidas a outros neurônios, por fim, a função do axônio é transmitir a informação gerada pelo corpo celular para os dendritos dos neurônios adjacentes.

Nosso cérebro é formado por bilhões de neurônios, quando conectados entre si eles formam uma enorme rede neural. O ponto de ligação entre a terminação axônica de um neurônio e o dendrito de outro neurônio é chamado de sinapse. É através das sinapses que os neurônios se unem funcionalmente, formando as redes neurais. As sinapses funcionam como válvulas, sendo capazes de controlar a transmissão de impulsos, isto é, o fluxo da informação entre os neurônios que compõem a rede neural. A informação quando passa na sinapse sofre o efeito de pesos variáveis, ou seja, a informação que está sendo transmitida pode ser amplificada ou atenuada, e é esta variação que dá ao neurônio capacidade de adaptação e de memorização.

Cada região do cérebro é especializada no processamento de uma função específica (sinais sonoros, visuais, auditivos, etc), esse processamento acontece por meio de redes particulares interligadas entre si, que realizam processamentos de forma paralela.

De forma semelhante ao cérebro humano, a unidade básica de uma rede neural artificial é o neurônio artificial, um modelo de neurônio artificial pode ser visto na Figura 10.





Fonte: CASTRO; ZUBEN, 2015.

Segundo Haykin (2001) um modelo de neurônio artificial qualquer apresenta três elementos básicos:

• Um conjunto de sinapses (conexões), no qual, cada uma delas está associada a um peso sináptico. Nesta representação o primeiro subescrito k do peso sináptico w_{kj} corresponde ao neurônio em que ele está ligado e o segundo sobrescrito corresponde

à sinapse ligada a ele. Então, um sinal x_j na entrada da sinapse j conectada ao neurônio k é multiplicado pelo peso sináptico w_{kj} .;

- Um somador para somar os sinais de entrada, ponderados pelos pesos sinápticos das sinapses pelos quais eles passa. O somador basicamente funciona como um combinador linear dos sinais de entrada;
- Uma função de ativação, para restringir a amplitude de saída e para introduzir não-linearidade ao modelo.

Adicionalmente, o modelo de neurônio artificial mostrado na Figura 10, inclui uma aplicação externa denominada *bias* (polarização), que é retratado como b_k . O *bias* tem a finalidade de aumentar ou diminuir a influência do valor da entrada para a ativação do neurônio k, seu valor pode ser positivo ou negativo.

Matematicamente a saída do neurônio k pode ser descrita pela seguinte equação:

$$y_k = f(u_k) = f(\sum_{j=1}^m w_{kj} x_j + b_k)$$
(2.1)

onde, $x_1, x_2, ..., x_m$ são sinais de entrada; $w_{k1}, w_{k2}, ..., w_{km}$ são os pesos sinápticos do neurônio k; u_k representa a saída por combinação linear dos sinais de entrada com a adição do bias b_k ; f(.) é a função de ativação; e y_k é o sinal de saída do neurônio.

2.6.1 Tipos de função de ativação tradicionais

A primeira função de ativação utilizada em redes neurais foi a função degrau, proposta por (MCCULLOH; PITTS, 1943). A função degrau f(x) utiliza um limiar de ativação θ para produzir uma saída que assume valores binários, essa função é definida como:

$$f(x) = \begin{cases} 1, \text{ se } x \ge \theta \\ 0, \text{ se } x < \theta \end{cases}$$
(2.2)

Por consequência da descontinuidade intrínseca à função degrau, não é possível realizar o treinamento através de métodos de minimização de erro baseadas em gradiente.

Posteriormente, a função degrau foi generalizada para a função logística, que possui um formato sigmoidal, seu gráfico possui forma parecida com a letra 's' que se comporta como uma versão suavizada da função degrau. A função logística $\varphi(x)$ é definida como:

$$\varphi(x) = \frac{1}{1 + e^{-ax}} \tag{2.3}$$

em que a é o parâmetro de inclinação da função logística e deve ser maior do que zero, alterando o valor deste parâmetro podem ser obtidas funções logísticas com inclinação diferentes. Quando a tende ao infinito, a função logística se comporta como uma função degrau. Enquanto a função degrau assume apenas os valores de 0 ou 1, a função logística assume valores contínuos contigos no intervalo de 0 a 1.

A função logística foi a primeira função sigmoidal a ser utilizada neste contexto, de forma que seu sucesso levou à utilização de diversas outras funções de ativação sigmoidais (DUCH; JANKOWSKI, 1999). As funções sigmoidais são as formas mais comuns de funções de ativação aplicadas em redes neurais artificiais.

Uma importante função sigmoidal que é comumente recomendada como substituta da função logística é a tangente hiperbólica (LECUN *et al.*, 2012). A função tangente hiperbólica é definida em função das funções cosseno e seno hiperbólicos, como pode ser visto a seguir:

$$tanh(x) = \frac{senh(x)}{cosh(x)}$$
(2.4)

A função tangente hiperbólica pode ser encarada como uma versão escalonada da função logística, já que ela pode atuar no intervalo de -1 a 1, este simples fato pode levar a uma melhor convergência da rede neural, especialmente quando as entradas são normalizadas e o intervalo coincide. Podemos destacar também, outras funções sigmoidais simétricas em relação à origem, como a função gaussiana de erro e o arco tangente.

Outra classe relevante de funções de ativação utilizadas em redes neurais artificiais são as funções radiais. Redes neurais com funções de ativação desta categoria tem uma única camada intermediária, os neurônios de saída são sempre lineares. A principal diferença das redes que tem funções de ativação radiais é a maneira com que entradas são processadas pelos neurônios da camada intermediária. A ativação interna de cada neurônio da camada intermediária não se dá mais pelo produto escalar entre o vetor de entradas e o vetor de pesos, como no caso do perceptron, e sim a partir de uma norma ponderada da diferença entre ambos os vetores. Um exemplo de função radial é a gaussiana, cuja definição pode ser vista a seguir:

$$r(x) = e^{(ax)^2} (2.5)$$

2.7 Arquitetura das redes neurais

O termo "arquitetura" de uma rede neural corresponde a maneira com que os neurônios e suas conexões estão conectados, a arquitetura da rede está estreitamente atrelada ao algoritmo de treinamento utilizado para aprendizagem da rede e é um fator
determinante na sua capacidade de processamento (ZANINI, 2000). Nas seções a seguir, serão abordadas três classes de diferentes de arquiteturas de redes neurais.

2.7.1 Redes alimentadas adiante com camada única

Redes alimentadas adiante ou redes *feedforward*, constituem um caso particular de redes neurais orgazinadas em várias camadas, ela é composta por uma camada de entrada de dados, que se projeta em uma camada única de saída, porém não vice-versa. As camadas são os modos com que os neurônios são estruturados, a camada de saída também é conhecida, neste caso, como camada de processamento, que calcula e fornece as saídas. A designação "rede de camada única" desconsidera a camada de entrada, já que ela não realiza nenhuma computação, referindo-se, então, apenas a camada de processamento.

Na Figura 11 ilustra-se uma rede alimentada adiante com camada única, com quatro neurônios nas camadas de entrada e saída.

Figura 11 – Rede alimentada adiante com camada única.



Fonte: adaptado de HAYKIN, 2008.

2.7.2 Redes alimentadas adiante com múltiplas camadas

Esta classe de rede neural é mais conhecida como *perceptrons* de múltiplas camadas (MLP), sua principal característica é a presença de uma ou mais camadas ocultas, compostas

por neurônios ocultos. A camada oculta tem a finalidade de intervir entre a camada de entrada e a camada de saída de uma maneira útil. Adicionando-se uma ou mais camadas ocultas, a rede torna-se capaz de extrair estatísticas de ordem elevada acerca das informações de entrada. Em um sentido livre, a rede adquire um panorama global apesar da conectividade local, devido ao conjunto extra de conexões e da dimensão extra de interações neurais (CHURCHLAND; SEJNOWSKI, 1992).

Nessa arquitetura, os neurônios da camada de entrada da rede provém os sinais de entrada aos neurônios da primeira camada oculta. Os sinais de saída da primeira camada oculta são utilizados como entrada para a camada oculta seguinte, e assim por diante pelo restante da rede. O conjunto de sinais de saída dos neurônios que compõem a última camada da rede constituem a resposta global da rede.

Na Figura 12 ilustra-se uma rede alimentada adiante com múltiplas camadas, composta por quatro neurônios na camada de entrada, quatro neurônios na camada oculta e dois neurônios na camada de saída.

Figura 12 – Rede alimentada adiante com múltiplas camadas.



Fonte: adaptado de HAYKIN, 2008.

A rede neural apresentada na Figura 12 é dita completamente conectada, pois cada neurônio de uma camada está conectado a todos os neurônios da camada posterior. De forma análoga, caso isto não seja verdade, a rede é dita parcialmente conectada.

2.7.3 Redes recorrentes

Este tipo de arquitetura de redes neurais distingue-se das redes neurais com alimentação adiante por possuir ao menos um laço de realimentação. Esta característica impacta significamente na capacidade de aprendizagem da rede e, consequentemente, no seu desempenho. Além disso, esses laços de realimentação, que são compostos por blocos de atraso unitário, admitem um comportamento não-linear a rede neural. Essa arquitetura de rede neural pode ser vista na Figura 13.

Figura 13 – Rede alimentada adiante com múltiplas camadas.



Fonte: adaptado de HAYKIN, 2008.

2.8 Treinamento da RNA

O conhecimento de uma rede neural está armazenado nos seus pesos sinápticos. Assim, ao se modificar os pesos, atualiza-se o conhecimento da rede. Treinamento é o nome dado ao processo iterativo de atualização dos valores dos pesos sinápticos por um algoritmo de aprendizado. A principal finalidade do treinamento é fazer com que um conjunto de entradas produza um conjunto de saídas desejadas, ou próximo disso.

O treinamento de RNAs pode ser classificado em duas categorias:

• Aprendizado supervisionado: A aprendizagem supervisionada necessita de um vetor

de entrada, um vetor de saída e um vetor alvo. Nesse processo, o vetor de entrada é aplicado na RNA e é então obtido um vetor de saída, que é comparado com o vetor alvo. O erro fruto desta comparação é realimentado através da rede e os pesos sinápticos são atualizados conforme a regra estabelecida no algoritmo de aprendizagem, com o intuito de minimizar o erro. Este processo ocorre repetidamente até que o erro obtido esteja dentro de uma faixa tolerável.

 Aprendizagem não-supervisionada: Neste caso não há um vetor alvo, ou seja, não são feitas comparações para determinar o erro na saída. O algoritmo de treinamento atualiza os pesos da rede de forma a produzir na saída resultados consistentes, este processo ocorre até que um limite de iterações seja atingido.

A seguir serão apresentados alguns dos principais algoritmos de aprendizado supervisionado.

2.8.1 Algoritmo de retropropagação

O algoritmo de retropropagação do erro (*backpropagation*) é o algoritmo mais conhecido para a realização do treinamento de *perceptrons* de múltiplas camadas. Este método de aprendizado pode ser utilizado em qualquer rede que utilize uma função de ativação diferenciável e não-linear com um treinamento supervisionado. Sua otimização tem como base a regra gradiente descendente, que ajusta os pesos a partir do erro da rede.

A aprendizagem por retropropagação consistem em duas etapas: uma etapa para frente (sentido da entrada para a saída), que é a propagação, e uma etapa para trás (sentido da saída para a entrada), que é a retropropagação.

Na primeira etapa, os sinais de entrada são apresentados a rede, que é propagado adiante, camada após camada, até produzir um sinal de saída. O sinal de saída da rede é comparado com o sinal de saída desejado, visto que este é um treinamento supervisionado, esta comparação irá gerar um valor de erro. Na tentativa de minimizar a função erro, calculam-se os valores dos gradientes para cada peso da rede, do cálculo sabemos que o vetor gradiente fornece a direção de crescimento da função, como queremos minimizar a função, basta tomar o sentido contrário ao do gradiente. Na segunda etapa, quando o vetor gradiente estiver calculado, os pesos são ajustados de modo a minimizar o erro, a atualização dos pesos sinápticos de cada camada é feito no sentido oposto a propagação dos sinais de entrada. Na Figura 14, tem se a representação do sentido de propagação dos sinais de entrada e dos sinais de erro em uma rede neural.

Figura 14 – Sentido de propagação de sinais de entrada e de sinais de erro em uma rede neural.



Fonte: HAYKIN, 2011.

A fórmula geral de atualização dos pesos sinápticos a cada iteração é a seguinte:

$$w_{ji}(t+1) = w_{ji}(t) - \eta \frac{\partial \epsilon(t)}{\partial w_{ji}(t)}$$
(2.6)

Pela equação conclui-se que o valor do peso sináptico, representado por w_{ji} , na iteração atual, será o valor do peso sináptico da iteração anterior, corrigido por um valor proporcional ao gradiente do erro, representado por $\partial \epsilon(t)/\partial w_{ji}(t)$. O sinal negativo indica que caminha-se na direção oposta ao gradiente, conforme mencionado. O parâmetro η representa a taxa de aprendizado da rede neural, este parâmetro influência na trajetória do gradiente. Quando menor esta taxa, menores são as atualizações dos pesos sinápticos de uma iteração para outra, com isso é possível que sejam necessárias mais iterações do algoritmo para encontrar uma solução aceitável. O aumento deste parâmetro pode acelerar o processo de convergência da rede, porém pode também torná-la instável, para contornar isso, pode ser utilizada uma taxa de momentum, que tem o objetivo de estabilizar o sinal de erro do gradiente em regiões de instabilidade (CERRI, 2014).

Este processo de aprendizagem repete-se de maneira iterativa, a cada iteração é calculado o gradiente do erro e os valores pesos sinápticos são atualizados, até o erro ficar abaixo de um patamar estabelecido ou até que se atinja o número máximo de iterações.

2.8.2 Algoritmo de retropropagação resiliente

O algoritmo de treinamento de retropropagação resiliente possui uma estrutura muito semelhante ao algoritmo de retropropagação, sendo considerado uma otimização deste último. No algoritmo de retropropagação, a atualização dos pesos sinápticos depende do valor da derivada parcial $\partial \epsilon / \partial w_{ij}$, dessa forma o comportamento da derivada parcial pode influenciar nos resultados. Com o intuito de acabar esse problema, Riedmiller e Braun (1993), propuseram o algoritmo de retropropagação resiliente. Este algoritmo considera apenas o sinal da derivada parcial ou gradiente de erro e não o seu valor absoluto, então a cada iteração cada peso sináptico é incrementado ou decrementado caso o sinal da derivada parcial seja positivo ou negativo, respectivamente. O fator de incremento é denominado η^+ , já o fator de decremento é denominado η^- . A condição de atualização dos pesos sinápticos para o algoritmo de retropropagaçõ resiliente é apresentado na equação a seguir:

$$\Delta w_{ij}(t) = \begin{cases} \eta^+ \Delta w_{ij}(t-1), \text{ se } \frac{\partial \epsilon}{\partial w_{ij}}(t-1) * \frac{\partial \epsilon}{\partial w_{ij}}(t) > 0\\ \eta^- \Delta w_{ij}(t-1), \text{ se } \frac{\partial \epsilon}{\partial w_{ij}}(t-1) * \frac{\partial \epsilon}{\partial w_{ij}}(t) < 0\\ \Delta w_{ij}(t-1), \text{ caso contrário} \end{cases}$$
(2.7)

onde, w_{ij} é o valor peso sináptico $\partial \epsilon / \partial w_{ij}$ é o gradiente do erro, η^+ é o fator de incremento, η^- é o fator de decremento e Δw_{ij} é a variação do peso sináptico. De forma que se o erro não mudar de sinal entre duas iterações $(\partial \epsilon / \partial w_{ij}(t-1) * \partial \epsilon / \partial w_{ij}(t) > 0)$ a variação dos pesos sinápticos é aumentada para acelerar a convergência, caso contrário, entende-se que a atualização dos pesos foi muito grande e a variação dos pesos sinápticos é diminuida.

2.8.3 Algoritmo de Levenberg-Marquardt

Também conhecido como método de amortecimento dos mínimos quadrados, o algoritmo de Levenberg-Marquardt é um método de otimização e aceleração de convergência, cuja minimização da função erro é realizada pelo método de Newton, sendo mais eficiente que a técnica do gradiente descendente utilizada no algoritmo de retropropagação. Este algoritmo foi proposto originalmente por Kenneth Levenberg e aperfeiçoado por Donald Marquardt.

A minimização da função erro é obtida por meio da introdução do parâmetro μ , que representa a taxa de aprendizagem e é inversamente proporcional ao tamanho do passo de atualização dos pesos sinápticos, no método de Gauss-Newton, conforme a equação:

$$\Delta x = [J^T(w)J(w) + \mu I]^{-1}J^T(w)e(w)$$
(2.8)

onde J(w) é a matriz Jacobiana, I é uma matriz identidade, e(w) é o erro e $J^{T}(w)J(w)$ é a matriz Hessiana. A fórmula de atualização dos pesos sinápticos a cada iteração para este algoritmo é a seguinte:

$$w_{ji}(t+1) = w_{ji}(t) - [J^T(w)J(w) + \mu I]^{-1}J^T(w)e(w_{ji}(t))$$
(2.9)

O parâmetro μ no algoritmo de Levenberg-Marquardt é dinamicamente alterado a medida que a função que se deseja minimizar diminui, este parâmetro também diminui, se

a função aumenta, este parâmetro também aumenta. Quando μ é pequeno o algoritmo se aproxima do método de Newton, já quando μ é grande o algoritmo se aproxima método gradiente descendente. É importante que a cada iteração o parâmetro μ seja atualizado para garantir a rápida convergência da solução.

2.9 Previsão de séries temporais utilizando RNA

Dada uma série temporal com uma sequência de observações $y_{(k)}$, com k = 1,2,3,etc.A lógica da previsão pode ser sintetizada da seguinte maneira: para valores conhecidos $y_{(k-n+1)}, y_{(k-n+2)}, ..., y_{(k-1)}, y_{(k)}$, pretende-se estimar o valor desta variável l períodos a frente da posição atual k, isto é, deseja-se determinar o valor de y_{k+l} , onde $n \in l$ são, respectivamente, a janela de previsão (valores passados) e o horizonte de previsão (valores futuros). Deste modo, pretende-se realizar o seguinte mapeamento:

$$f: [y_{(k-n+1)}, y_{(k-n+2)}, \dots, y_{(k-1)}, y_{(k)}] \to [y_{(k+1)}, y_{(k+2)}, \dots, y_{(k+l-1)}, y_{(k+l)}]$$
(2.10)

Na Figura 15 é possível visualizar um exemplo de previsão com uma janela composta de quatro valores passados da variável utilizados para fazer a previsão com horizonte de um valor futuro. Por meio da janela $y_{(k-3)}$, $y_{(k-2)}$, $y_{(k-1)} \in y_{(k)}$, pode-se determinar a medida para um horizonte $y_{(k+1)}$.

Figura 15 – Exemplo de previsão de séries temporais.



Fonte: MEDEIROS, 2019.

Para utilizarmos uma rede neural como instrumento de previsão de séries temporais, é necessário levar em consideração a variável do tempo, de tal forma que seja possível a modelagem das variações estatísticas de uma série temporal de dados. Segundo Haykin (1994), este problema pode ser solucionado através da inclusão de um mecanismo denominado memória de curto prazo (MCP). A forma de MCP mais comum é chamada memória de linha de atraso. Uma rede *perceptron* multicamada, com uma memória de linha de atraso derivada, aplicada à entrada, é a rede mais simples e uma das mais aplicadas na previsão de séries temporais.

Na Figura 16 é ilustrado um esquema de uma rede MLP com memória de linha de atraso. Os blocos Z^{-1} consituem unidades operadoras de atraso e a entrada (k + 1) representa o valor alvo da saída no seu horizonte de previsão.

Figura 16 – Esquema de uma RNA com memória de linha de atraso.



Fonte: adaptado de MEDEIROS, 2016.

3 Metodologia utilizada para previsão de demanda

O método proposto para previsão de demanda, a configuração dos parâmetros da rede neural e seu treinamento, até os testes preliminares para obtenção dos resultados são apresentados neste capítulo.

3.1 Objeto de estudo

A subestação escolhida como objeto de estudo foi a subestação João Pessoa (JPS), localizada na cidade de João Pessoa-PB. Na Figura 17 é ilustrado um diagrama simplificado da subestação JPS.



Figura 17 – Diagrama simplificado da subestação JPS.

Fonte: elaborado pelo autor, 2019.

A subestação JPS trata-se de uma subestação abaixadora, de 69 kV para 13,8 kV. As medições da energia injetada na subestação estão localizadas no lado de alta (69 kV) de ambos os transformadores da subestação, mais especificamente nos pontos 12B1 e 12B2. Já as medições de energia fornecida pela subestação estão localizadas no lado de baixa (13,8 kV) dos transformadores da subestação, mais especificamente nos pontos 11B1 e 11B2, ou nos alimentadores desta subestação, neste caso, nos pontos 21L1, 21L2, 21L3, 21L4, 21L5, 21L6, 21L7, 21L8, 21L9 e 21C1. Em um cenário ideal, a soma das medições nos lados de baixa do transformadores deve ser igual a soma das medições de todos os alimentadores da subestação.

As medições utilizadas neste trabalho foram coletadas no alimentador identificado por 21L1 da subestação JPS, este alimentador foi escolhido por ter uma característica sazonal bem definida, como já foi apresentado anteriormente na Figura 8.

3.2 Visão geral

Para construção do modelo de previsão de demanda, foram utilizados os softwares Microsoft® Excel e MATLAB®, além do histórico de medição de potência ativa.

O processo de previsão de demanda ocorre, basicamente, em quatro etapas: préprocessamento, implementação da rede neural, treinamento e validação. A etapa de pré-processamento consiste na seleção e tratamento dos dados que serão apresentados a rede neural. Na etapa de implementação serão definidas as características da rede neural, tais como, sua arquitetura, sua função de ativação, número de camadas e neurônios, dentre outros. Na etapa de treinamento é definido o algoritmo de aprendizagem da rede, são estabelecidos critérios de parada como erro máximo desejado, número máximo de iterações, também é nessa etapa que os dados de entrada são apresentados a rede neural. Na etapa de validação a saída da rede neural é confrontada com dados nunca vistos antes pela rede, com a finalidade de mensurar o desempenho da previsão, algumas métricas são calculadas para mensurar a exatidão da previsão, tais como erro percentual médio absoluto e a raiz do erro médio quadrático.

Como entradas das redes neurais, foram utilizados uma janela de valores de potência, energia acumulada e variação de energia. Há estudos que indicam maior exatidão no resultado das previsões quando outras variáveis altamente correlacionadas com a demanda são considerados. Dessa forma, este trabalho teve a intenção de incluir as variáveis de energia acumulada e variação de energia, buscando um melhor modelo de previsão.

Para efeito de comparação, serão feitas análises de alguns modelos com arranjos diferentes das variáveis citadas anteriormente, a fim de verificar se estas tem impacto positivo ou negativo nos resultados esperados das redes treinadas.

3.3 Implementação do banco de dados

A base de dados utilizada para subsidiar a previsão de demanda realizada neste trabalho consiste de uma série de medições de potência ativa, realizadas entre os anos de 2008 e 2013, com um intervalo de quinze minutos entre medições. Um sistema de aquisição de dados (SCADA) recebeu as medições nos equipamentos da subestação JPS e os enviou para um banco de dados.

Uma ánalise prévia deste banco de dados identificou a existência de inconsistências nas medições, tais como: intervalos sem medição, medições zeradas e degraus nas medições. Essas medições atípicas são conhecidas na literatura como *outliers* e podem ser ocasionados por diversos fatores, tais como: problemas na telemetria, que podem resultar em intervalos sem medição, troca de medidor sem o ajuste correto da sua constante interna, que pode resultar em medições muito acima ou muito abaixo do normal, medições zeradas, que pode ser problema do equipamento de medição, que neste caso deverá ser trocado. Porém, essas medições incomuns também podem não se tratar de erros, o desligamento de um alimentador para algum tipo de manutenção na subestação, pode gerar medições zeradas, como também pode gerar uma medição bem acima do normal em outro alimentador que irá fornecer energia para alimentar as cargas ligadas ao alimentador desligado.

As medições contidas no banco de dados, passaram por uma etapa de pré-tratamento, pois, independente da natureza, um *outlier* em uma série temporal deve ser corrigido a fim de que este não comprometa o algoritmo de previsão, já que a qualidade e confiabilidade dos dados são essenciais neste processo (GUO et al., 2012).

O escopo deste trabalho não é a identificação e a correção de *outliers*, no entanto, esta é uma etapa importante no pré-tratamento de dados, como fica evidenciado nas Figuras 18 e 19, onde temos, respectivamente, uma curva antes e depois do processo de identificação e correção de *outliers*.

CAPÍTULO 3. METODOLOGIA UTILIZADA PARA PREVISÃO DE DEMANDA 47



Figura 18 – Série temporal antes da correção de outliers.





Figura 19 – Série temporal após a correção de outliers.

Fonte: elaborado pelo autor, 2019.

Na Tabela 1 é apresentado uma parte do banco de dados, da maneira como que ele foi cedido, já considerando a correção de *outliers*, para que se tenha uma ideia de como os

DIA	MES	ANO	HORA	MINUTO	POTÊNCIA ATIVA (MW)
1	1	2008	0	0	1,961
1	1	2008	0	15	1,938
1	1	2008	0	30	1,933
1	1	2008	0	45	1,892
1	1	2008	1	0	1,86
1	1	2008	1	15	1,869
1	1	2008	1	30	1,83
1	1	2008	1	45	1,831
1	1	2008	2	0	1,886
1	1	2008	2	15	1,833
1	1	2008	2	30	1,816
1	1	2008	2	45	1,773
1	1	2008	3	0	1,78
1	1	2008	3	15	1,811

dados estão organizados.

Tabela 1 – Organização do banco de dados utilizado na previsão de demanda.

Fonte: elaborado pelo autor, 2019.

Partindo da Tabela 1, foi criado um algoritmo no MATLAB® para identificar a potência máxima diária, para representar o dia da semana através de um número identificador (ID), onde o número 1 corresponde ao domingo, o número 2 corresponde a segunda-feira e assim sucessivamente, até o sábado que corresponde ao número 7, e para calcular energia acumulada em um dia, que foi calculada pelo somatório das potências medidas em um dia multiplicadas por 0,25 (já que as medições tem um intervalo de um quarto de hora), o que equivale a fazer uma integração númerica pelo método mais simples, o método dos retângulos, além disso, o errro fruto desse processo de integração é tolerável. Na Tabela 2 são apresentados parcialmente os dados resultantes do algoritmo criado.

ID	Potência máxima (MW)	Energia acumulada (MWh)
3	3,246	55,694
4	3,933	65,951
5	4,328	69,028
6	4,022	68,413
7	2,598	52,908
1	2,516	47,897
2	4,893	74,096
3	4,877	76,219
4	4,866	76,174
5	4,847	75,115
6	4,469	72,106
7	2,559	53,273
1	2,282	46,621
2	3,969	65,251

Tabela 2 – Organização do banco com dados diários de potência máxima e energia acumulada.

Fonte: elaborado pelo autor, 2019.

A partir da Tabela 2 com informações diárias, complementou-se o algoritmo do MATLAB® para agora identificar a potência máxima semanal, para obter energia acumulada em uma semana, que foi calculada pelo somatório das energias acumuladas nos dias que compõem a semana, para obter a variação semanal de energia, que foi calculada pela subtração da energia acumulada na semana atual pela energia acumulada na semana anterior e para representar a semana através de um ID, onde o número 1 corresponde a primeira semana do ano, o numéro 2 a segunda semana do ano e assim sucessivamente, até o número 52 que representa a última semana do ano. Foi considerado o ínicio da primeira semana no primeiro domingo de cada ano, finalizando no sábado subsequente, assim sucessivamente, até o final da semana cinquenta e dois que coincidiu com o primeiro sábado do ano posterior, e a partir daí a contagem das semanas foi reiniciado. Na Tabela 3 são apresentados uma parcela dos dados resultantes do algoritmo criado.

	Potência	Energia	Variação da				
ID	máxima (MW)	acumulada (MWh)	energia (MWh)				
1	4,893	474,882	0				
2	4,559	445,261	-29,621				
3	4,953	487,958	42,697				
4	5,008	474,494	-13,464				
5	5,126	436,204	-38,289				
6	$5,\!309$	493,130	56,925				
7	5,409	511,636	18,505				
8	5,511	534,824	$23,\!187$				
9	5,564	535,466	0,642				
10	5,513	528,293	-7,173				
11	5,383	$450,\!657$	-77,725				
12	5,353	489,078	38,511				
13	5,209	489,430	0,352				
14	5,341	492,391	2,96				

Tabela 3 – Organização do banco com dados semanais de potência máxima, energia acumulada e variação da energia.

Fonte: elaborado pelo autor, 2019.

Desta forma, o banco de dados finalizado, formado pelo identificador da semana, pelas potências máximas semanais, pela energia acumulada semanal e pela variação de energia entre semanas, alcançou um total de 312 linhas, já que são dados de seis anos (correspondete aos anos de 2008 a 2013), com cada ano tendo cinquenta e duas semanas.

3.4 Treinamento da RNA

Os dados referentes aos cinco primeiros anos, foram utilizados para treinamento da RNA e os dados referentes ao sexto ano foram utilizados para validação da RNA. A divisão dos dados de treinamento e de validação para a potência máxima semanal podem ser vistos na Figura 20.



Figura 20 – Dados de potência máxima semanal da SE JPS, alimentador 21L1, de 2008 até 2013.

Fonte: elaborado pelo autor, 2019.

Para o treinamento da rede neural artificial, foram adotadas janela e horizonte de previsão variáveis, assim como o número de neurônios em cada uma das duas camadas ocultas, com o intuito de identificar com qual combinação desses parâmetros a rede neural apresentaria uma previsão mais exata. Esses parâmetros foram variados da seguinte forma:

- Horizonte de previsão: Variou-se de 4 a 20 semanas, com passo de 2 semanas;
- Janela de previsão: Variou-se de 2 a 12 semanas, com passo de 2 semanas;
- Número de neurônios na primeira e segunda camadas: Variou-se de 1 a 20 neurônios, com passo unitário;

Essa variação de parâmetros gerou um total de 21600 redes para cada modelo. Serão testados quatro modelos diferentes. O primeiro modelo terá apenas a potência máxima semanal com variável de entrada da RNA, o segundo modelo terá potência máxima semanal e energia acumulada na semana como variáveis de entrada da RNA, o terceiro modelo terá potência máxima semanal e variação da energia entre semanas como variáveis de entrada da RNA, por fim, o quarto modelo terá potência máxima semanal, energia acumulada na semana e variação da energia entre semanas como variáveis de entrada da RNA. Na Figura 21, é apresentado um fluxograma sintetizando o quarto modelo de previsão proposto, considerando uma janela de previsão de quatro amostras e um horizonte de previsão de duas amostras.



Figura 21 – Entradas e saídas da RNA e janelamento.

Fonte: elaborado pelo autor, 2019.

Para representar os demais modelos de RNA propostos, basta ajustar as variáveis de entrada da RNA conforme o modelo desejado. Para realizar o treino foi criada uma tabela com dados de entrada e outra tabela com dados que seriam os alvos da previsão da rede neural. Na Tabela 4 pode-se ver como ficaram organizados os dados de entrada para o treino da RNA do terceiro modelo de previsão proposto. Já na Tabela 5 pode-se ver como ficaram organizados os alvos de saída para o treino da RNA do terceiro modelo de previsão proposto

P(i)	P(i-1)	P(i-2)	P(i-3)	$\Delta E(i)$	$\Delta E(i-1)$	$\Delta E(i-2)$	$\Delta E(i-3)$
5,008	4,953	4,559	4,893	-13,464	42,696	-29,621	0
5,126	5,008	4,953	4,559	-38,289	-13,464	42,696	-29,621
5,309	5,126	5,008	4,953	56,925	-38,289	-13,464	42,696
5,409	5,309	5,126	5,008	18,505	$56,\!925$	-38,289	-13,464
5,511	5,409	5,309	5,126	23,187	18,505	56,925	-38,289
5,564	5,511	5,409	5,309	0,642	23,187	18,505	56,925
5,513	5,564	5,511	$5,\!409$	-7,173	$0,\!642$	$23,\!187$	18,505
5,383	5,513	5,564	5,511	-77,725	-7,173	0,642	23,187
5,353	$5,\!383$	5,513	5,564	38,511	-77,725	-7,173	$0,\!642$
5,209	$5,\!353$	5,383	5,513	0,352	38,511	-77,725	-7,173
5,341	5,209	$5,\!353$	$5,\!383$	2,96	$0,\!352$	38,511	-77,725
$5,\!407$	$5,\!341$	5,209	$5,\!353$	6,629	2,96	$0,\!352$	38,511
5,162	5,407	5,341	5,209	-41,727	6,629	2,96	$0,\!352$
5,019	5,162	5,407	5,341	18,340	-41,727	6,629	2,96

Tabela 4 – Tabela de dados de entrada para treino da RNA do terceiro modelo.

Fonte: elab	oorado	pelo	autor,	2019.
-------------	--------	-----------------------	--------	-------

Tabela 5 – Tabela com os alvos de saída do treino da RNA do terceiro modelo.

P(i+1)	P(i+2)
5,126	5,309
5,309	5,409
5,409	5,511
5,511	5,564
5,564	5,513
5,513	5,383
5,383	5,353
5,353	5,209
5,209	5,341
5,341	5,407
5,407	5,162
5,162	5,019
5,019	5,053
5,053	4,794

Fonte: elaborado pelo autor, 2019.

3.5 Correlação das variáveis

Com o intuito de mensurar o nível de relação entre as variáveis escolhidas para subsidiar a previsão de demanda pela RNA, foram realizadas duas abordagens: análise gráfica e cálculo da correlação. Inicialmente executou-se uma análise gráfica das variáveis em questão, na Figura 22 têm-se a representação do comportamento da máxima potência semanal, da energia acumulada na semana e da variação de energia entre semanas, para uma intervalo de um ano, neste caso, o ano de 2008.

Figura 22 – Comportamento da potência máxima semanal, energia acumulada na semana e variação de energia entre semanas para um período de um ano.



Fonte: elaborado pelo autor, 2019.

Pelas curvas apresentadas na Figura 22, nota-se um comportamento anual semelhante entre a potência máxima semanal e a energia acumulada semanal. Exemplo dessa semelhança são o decaimento nas últimas semanas do ano e o decréscimo por volta da semana vinte e cinco, ambos atingindo seu menor valor no ano, seguido por um período de crescimento. Já o comportamento entre a máxima potência semanal e a variação de energia semanal, aparentemente, não apresentam semelhança, indicando que possivelmente essas variáveis são pouco relacionadas.

A análise gráfica, fornece apenas uma indicação se há relação entre séries temporais, para identificar essa relação com maior exatidão é calculada a correlação entre duas séries temporais.

A correlação mensura o nível de relação entre variáveis, então desenvolveu-se um código em MATLAB® que fez uso do comando *corr*, para retornar o coeficiente de correlação de Pearson entre duas séries temporais. Este coeficiente é o coeficiente de

correlação mais comumente utilizado e é dado pela seguinte equação:

$$\rho(X,Y) = \frac{\sum_{i=1}^{N} (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{N} (X_i - \overline{X})^2 \sum_{i=1}^{N} (Y_i - \overline{Y})^2}}$$
(3.1)

onde, $\overline{X} \in \overline{Y}$ são as médias das séries e N é o tamanho das séries.

Os valores do coeficiente de correlação podem variar de -1 a +1. Um valor de -1 indica correlação negativa perfeita, enquanto um valor de +1 indica correlação positiva perfeita. Um valor 0 indica que não há correlação entre as séries. Para Cohen (1988), valores de correlação até 0,29 indicam correlação fraca, entre 0,3 e 0,69 indicam correlação moderada e acima de 0,7 indicam correlação forte.

Quando executado o código desenvolvido para calcular a correlação, obteve-se o valor de 0,948 para a correlação entre a potência máxima semanal e a energia acumulada na semana, indicando uma correlação forte, e o valor de 0,119 para a correlação entre a potência máxima semanal e a variação de energia semanal, indicando uma correlação fraca.

Devido a baixa correlação entre potência máxima semanal e a variação de energia semanal, decidiu-se examinar se a correlação seria maior caso fosse calculada a correlação entre duas variações, sendo uma a variação da potência máxima semanal e a outra a variação de energia semanal, a correlação obtida neste caso foi de 0,22, uma correlação maior que a encontrada entre potência máxima semanal e a variação de energia semanal, porém ainda classificada como correlação fraca.

É possível presumir que a informação da variação de energia semanal, que tem fraca correlação com a potência máxima semanal, tenha pouca influência na melhora na exatidão da previsão de demanda, essa hipótese será melhor investigada no próximo capítulo.

3.6 Validação e teste da RNA

Para avaliação do desempenho das previsões obtidas pelos modelos de previsão propostos, foram consideradas as seguintes métricas: erro absoluto (EA), erro relativo (ER), raiz do erro médio quadrático (RMSE), erro percentual médio absoluto (MAPE) e o coeficiente U de Theil.

O erro absoluto é o módulo da diferença entre o valor real e o valor previsto. Se designarmos por x_0 o valor previsto e x o valor real, o erro absoluto será:

$$EA = |x - x_0| \tag{3.2}$$

O erro relativo é o módulo da diferença entre o valor real e o valor previsto, dividido pelo valor real, este erro é expresso na forma percentual. Se designarmos por x_0 o valor previsto e x o valor real, o erro absoluto será:

$$ER = \frac{|x - x_0|}{x} * 100 \tag{3.3}$$

A raiz do erro médio quadrático possui dimensão semelhante a dos valores reais e previstos, esta é uma métrica que penaliza erros maiores, já que as diferenças são elevadas ao quadrado (SHCHERBAKOV *et al.*, 2013). Se designarmos por x_0 o valor previsto, x o valor real e N o tamanho do vetor, a raiz do erro médio quadrático será:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (x - x_0)^2}{N}}$$
(3.4)

O erro percentual médio absoluto é uma das métricas mais utilizadas para aferir a acurácia de métodos de previsão e corresponde ao valor médio do erro percentual das previsões (KIM; KIM, 2016). Se designarmos por x_0 o valor previsto, x o valor real e N o tamanho do vetor, o erro percentual médio absoluto será:

$$MAPE = \frac{1}{N} \sum_{j=1}^{N} \left| \frac{x - x_0}{x} \right| * 100$$
(3.5)

O coeficiente U de Theil, mensura quanto os resultados obtidos são melhores do que uma previsão ingênua ou trivial, isto é, quando a melhor estimativa do próximo valor é o próprio valor atual (MAKRIDAKIS *et al.*, 1998). A qualidade da previsão é interpretada da seguinte forma:

- U > 1: o erro do modelo é maior do que o erro de uma previsão ingênua;
- U < 1: o erro do modelo é menor do que o erro de uma previsão ingênua.

Quanto mais próximo de zero o coeficiente de U de Theil, melhor o resultado da previsão. O coeficiente de U de Theil é dado pela seguinte equação:

$$U = \sqrt{\frac{\sum_{i=1}^{N} (x_{(i)} - x_{0(i)})^2}{\sum_{i=1}^{N} (x_{(i)} - x_{0(i-1)})^2}}$$
(3.6)

3.7 Parâmetros de treinamento

O ambiente computacional para realização de todas as simulações foi o MATLAB®, onde utilizou-se o *toolbox* (biblioteca) NEURAL NETWORK TOOLBOX do MATLAB®

2018 para a criação da RNA. Os parâmetros foram definidos baseados em experiências de previsão anteriores, como em Medeiros (2019), sendo assim, os parâmetros ficaram da seguinte forma:

- Número de camadas ocultas: duas camadas.
- Número de neurônios: para o treinamento da rede, variou-se de 1 a 20 o número de neurônios de cada camada, afim de encontrar a combinação de números de neurônios que apresentasse o melhor desempenho.
- Algoritmo de aprendizado: Algoritmo de Levenberg-Marquardt (LM).
- Função de ativação: Utilizou-se a função tangente hiperbólica (*tansig*), do tipo sigmoidal.
- Número máximo de iterações: 2500 iterações.

Para treinamento, processamento dos dados e análise, foi utilizado um computador cujas características estão presentes na Tabela 6.

Fabricante	Dell							
Modelo	Optiplex 7050							
Processador	Intel \mathbb{R} Core TM i7-6700T CPU @3.60 GHz							
Memória RAM	8 GB							

Tabela 6 – Especificações técnicas da máquina utilizada

4 Resultados

Neste capítulo, são apresentados os resultados da previsão dos modelos de previsão propostos. Inicialmente, será tratado cada modelo de forma individual e posteriormente será realizada uma comparação entre modelos. A acurácia da previsão das redes foi avaliada segundo as métricas de desempenho apresentadas no capítulo anterior, optou-se por selecionar as redes de melhor MAPE e RMSE, deixando as outras métricas para análise posterior das redes selecionadas. Para a avaliação das redes foram utilizados dados de teste equivalentes a um período de cinquenta e duas semanas.

4.1 Modelo considerando apenas a potência máxima semanal

Uma vez realizado o treinamento de todas as redes, foi criado um algoritmo em MATLAB® para percorrer os arquivos de todas as redes deste modelo e computar seus respectivos valores de MAPE. Conforme Figura 23, foram esboçados os valores médios do MAPE de todas as redes, em função do número de neurônios na primeira e segunda camadas.



Figura 23 – Valores médios de MAPE das redes do primeiro modelo.

Fonte: elaborado pelo autor, 2019.

Na Figura 23 a barra de cor vertical localizada na direita, associa uma cor do mapa de cores a um valor, neste caso, um valor de MAPE. Por esta figura nota-se que as regiões em que há muitos neurônios em uma camada e pouco neurônios na outra camada apresentam, em média, valores de MAPE mais alto que os demais, sendo a região que apresentou menores valores de MAPE médio foi a região com poucos neurônios em ambas camadas.

O menor valor de MAPE obtido entre todas as redes foi de 2,65%. Esta rede tinha vinte semanas como horizonte de previsão, duas semanas como janela de previsão, dezoito neurônios na primeira camada e nove neurônios na segunda camada. A comparação da previsão da rede de melhor MAPE com os valores reais, que correspondem ao objetivo de saída, pode ser visto na Figura 24.



Figura 24 – Previsão da rede de menor MAPE e objetivo.

Fonte: elaborado pelo autor, 2019.

A avaliação da rede com menor RMSE foi feita de forma análoga a avaliação da rede com menor MAPE. Na Figura 25, são apresentados os valores médios da RMSE de todas as redes, em função do número de neurônios na primeira e segunda camadas.

O comportamento dos valores médios da RMSE se assemelhou bastante ao comportamento dos valores médios de MAPE, no qual, as regiões em que há muitos neurônios em uma camada e pouco neurônios na outra camada apresentam, em média, valores de RMSE mais alto que os demais, sendo a região que apresentou menores valores de RMSE médio foi a região com poucos neurônios em ambas camadas.



Figura 25 – Valores médios de RMSE das redes do primeiro modelo.

Fonte: elaborado pelo autor, 2019.

Dentre todas as RNAs a rede com a menor valor de RMSE foi de 0,182. Esta rede tinha vinte semanas como horizonte de previsão, duas semanas como janela de previsão, dezoito neurônios na primeira camada e nove neurônios na segunda camada, ou seja, a rede que apresentou menor MAPE também foi a rede que apresentou menor RMSE. Na Tabela 7 são apresentados os valores de todas as métricas de desempenho para esta rede.

Tabela 7 – Valores das métricas de desempenho para a rede de melhor MAPE e RMSE do primeiro modelo.

MAPE	RMSE	EA máximo	ER máximo	U de Theil
$2,\!65\%$	0,182	0,363	$5{,}22\%$	0,731

Fonte: elaborado pelo, 2019.

4.2 Modelo considerando a potência máxima semanal e a energia acumulada na semana

Para o modelo de RNA com potência máxima semanal e energia acumulada na semana como variáveis de entrada, os valores médios de MAPE de todas as redes deste modelo em função do número de neurônios na primeira e segunda camadas podem ser vistos na superfície da Figura 26.



Figura 26 – Valores médios de MAPE das redes do segundo modelo.

Fonte: elaborado pelo autor, 2019.

Dentre todas as redes neurais deste modelo, o menor MAPE obtido foi de 2,849%. Esta rede tinha dezoito semanas como horizonte de previsão, seis semanas como janela de previsão, um neurônio na primeira camada e sete neurônios na segunda camada. A comparação da previsão desta rede com o objetivo da saída pode ser visto na Figura 27.

Figura 27 – Comparação da previsão da rede de menor MAPE e objetivo de saída.



Fonte: elaborado pelo autor, 2019.

De forma semelhante ao MAPE, na Figura 28, é apresentada uma superfície com valores médios da RMSE de todas as redes em função do número de neurônios na primeira e segunda camadas.



Figura 28 – Valores médios da RMSE das redes do segundo modelo.

Fonte: elaborado pelo autor, 2019.

Pelas Figuras 27 e 28 nota-se, novamente, que as regiões em que há muitos neurônios em uma camada e pouco neurônios na outra camada apresentam, em média, valores de MAPE e RMSE mais alto que os demais, sendo a região que apresentou menores valores de MAPE e RMSE médio foi a região com poucos neurônios em ambas camadas, porém, de forma geral, as redes do segundo modelo apresentaram valores médios menores que as redes do primeiro modelo, tanto para MAPE quanto para RMSE.

A rede com menor RMSE foi de 0,189. Esta rede tinha vinte semanas como horizonte de previsão, seis semanas como janela de previsão, três neurônios na primeira camada e seis neurônios na segunda camada. A comparação da previsão desta rede com o objetivo da saída pode ser visto na Figura 29.



Figura 29 – Comparação da previsão da rede de menor RMSE e objetivo de saída.

Fonte: elaborado pelo autor, 2019.

Na Figura 30 é apresentada uma comparação entre os erro absoluto das previsões da rede de menor MAPE com a rede de menor RMSE.

Figura 30 – Comparação entre o erro absoluto da rede de menor MAPE com o erro da rede de menor RMSE.



Fonte: elaborado pelo autor, 2019.

Na Figura 30 é possível observar que o erro absoluto das duas redes varia bastante, pórem a média do erro absoluto da rede de menor MAPE é menor que a média deste

mesmo erro para a rede de menor RMSE.

Na Tabela 8 são apresentados os valores de todas as métricas de desempenho para a rede de menor MAPE e para a rede de menor RMSE.

Tabela 8 – Valores das métricas de desempenho para a rede de menor MAPE e para a rede de menor RMSE do segundo modelo.

	MAPE	RMSE	EA máx	ER máx	U de Theil
Rede de menor MAPE	2,849%	0,194	0,345	$5{,}303\%$	0,792
Rede de menor RMSE	2,920%	0,189	0,360	5,510%	0.785

Fonte: elaborado pelo autor, 2019.

Com os dados da Tabela 8, conclui-se que a rede de menor MAPE obteve erro absoluto máximo e erro relativo máximo menores que a rede de menor RMSE, porém a rede de menor RMSE apresentou um U de Theil menor que a rede de menor MAPE.

4.3 Modelo considerando a potência máxima semanal e variação de energia entre semanas

Para o modelo de RNA com potência máxima semanal e variação da energia entre semanas como variáveis de entrada, os valores médios de MAPE de todas as redes deste modelo em função do número de neurônios na primeira e segunda camadas podem ser vistos na superfície da Figura 31.



Figura 31 – Valores médios de MAPE das redes do terceiro modelo.

Fonte: elaborado pelo autor, 2019.

Dentre todas as redes neurais deste modelo, o menor MAPE obtido foi de 3,05%. Esta rede tinha vinte semanas como horizonte de previsão, duas semanas como janela de previsão, onze neurônios na primeira camada e dois neurônios na segunda camada. A comparação da previsão desta rede com o objetivo da saída pode ser visto na Figura 32.

Figura 32 – Comparação da previsão da rede de menor MAPE e objetivo de saída.



Fonte: elaborado pelo autor, 2019.

De forma semelhante ao MAPE, na Figura 33, é apresentada uma superfície com valores médios da RMSE de todas as redes em função do número de neurônios na primeira e segunda camadas.



Figura 33 – Valores médios da RMSE das redes do terceiro modelo.

Fonte: elaborado pelo autor, 2019.

Conforme pode ser visto pelas Figuras 31 e 33 as redes do segundo modelo apresentaram valores médios de MAPE e RMSE maiores que as redes do primeiro e segundo modelos.

A rede com menor RMSE foi de 0,198. Esta rede tinha vinte semanas como horizonte de previsão, seis semanas como janela de previsão, três neurônios na primeira camada e onze neurônios na segunda camada. A comparação da previsão desta rede com o objetivo da saída pode ser visto na Figura 34.



Figura 34 – Comparação da previsão da rede de menor RMSE e objetivo de saída.

Fonte: elaborado pelo autor, 2019.

Na Figura 35 é apresentada uma comparação entre os erro absoluto das previsões da rede de menor MAPE com a rede de menor RMSE.

Figura 35 – Comparação entre o erro da rede de menor MAPE com o erro da rede de menor RMSE.



Fonte: elaborado pelo autor, 2019.



MAPE é levemente menor que a média deste mesmo erro para a rede de menor RMSE (0,188 contra 0,189, respectivamente).

Na Tabela 9 são apresentados os valores de todas as métricas de desempenho para a rede de menor MAPE e para a rede de menor RMSE.

Tabela 9 – Valores das métricas de desempenho para a rede de menor MAPE e para a rede de menor RMSE do terceiro modelo.

	MAPE	RMSE	EA máx	ER máx	U de Theil
Rede de menor MAPE	$3,\!050\%$	0,206	0,342	5,960%	0,802
Rede de menor RMSE	3,110%	0,198	0,323	5,321%	0,830

	20	[
da	18																			
nai	10																			6.5
cal	16	\vdash			-															
Ida																				6
gur	14																			Ŭ
Sei	12	-																		
na																				5.5
ios	10																			
ŝ	8	-																		5
nen					-															Ŭ
de	6																			
ē	4																			4.5
me					-															
ž	2																			4
	0																			
	ΰ	C				ł	5				1	0		1	5		2	0		
Número de neurônios na primeira camada																				

Fonte: elaborado pelo autor, 2019.

Com os dados da Tabela 9, conclui-se que a rede de menor RMSE obteve erro absoluto máximo e erro relativo máximo menores que a rede de menor MAPE, esta, obteve um U de Theil menor que a rede de menor RMSE.

4.4 Modelo considerando a potência máxima semanal, a energia acumulada na semana e variação de energia entre semanas

Para o quarto modelo de RNA, os valores médios de MAPE de todas as redes deste modelo em função do número de neurônios na primeira e segunda camadas podem ser vistos na Figura 36.

Figura 36 – Valores médios de MAPE das redes do quarto modelo.

Fonte: elaborado pelo autor, 2019.

Dentre todas as redes neurais deste modelo, o menor MAPE obtido foi de 2,851%. Esta rede tinha vinte semanas como horizonte de previsão, duas semanas como janela de previsão, um neurônio na primeira camada e um neurônio na segunda camada. A comparação da previsão desta rede com o objetivo da saída pode ser visto na Figura 37.

Figura 37 – Comparação da previsão da rede de menor MAPE e objetivo de saída.



Fonte: elaborado pelo autor, 2019.

De forma semelhante ao MAPE, na Figura 38, são apresentados os valores médios da RMSE de todas as redes em função do número de neurônios na primeira e segunda camadas.



Figura 38 – Valores médios da RMSE das redes do segundo modelo.

Fonte: elaborado pelo autor, 2019.

Pelas Figuras 36 e 38, nota-se, em relação aos demais modelos, maior predominância de blocos azuis, indicando médias baixas e de valores semelhantes.

A rede com menor RMSE foi de 0,196. Esta rede tinha dezesseis semanas como horizonte de previsão, quatro semanas como janela de previsão, três neurônios na primeira camada e quinze neurônios na segunda camada. A comparação da previsão desta rede com o objetivo da saída pode ser visto na Figura 39.



Figura 39 – Comparação da previsão da rede de menor RMSE e objetivo de saída.

Fonte: elaborado pelo autor, 2019.

Na Figura 40 é apresentada uma comparação entre os erro absoluto das previsões da rede de menor MAPE com a rede de menor RMSE.

Figura 40 – Comparação entre o erro da rede de menor MAPE com o erro da rede de menor RMSE.



Fonte: elaborado pelo autor, 2019.


pórem, desta vez, a rede de menor RMSE tem a média do erro absoluto levemente menor que a rede de menor MAPE.

Na Tabela 10 são apresentados os valores de todas as métricas de desempenho para a rede de menor MAPE e para a rede de menor RMSE.

Tabela 10 – Valores das métricas de desempenho para a rede de menor MAPE e para a rede de menor RMSE do quarto modelo.

	MAPE	RMSE	EA máx	ER máx	U de Theil
Rede de menor MAPE	2,851%	0,200	0,401	6,77%	0,814
Rede de menor RMSE	2,890%	0,196	0,340	5,76%	0,788

de menor MAPE	$2,\!851\%$	0,200	0,401	6,77%	0,814
de menor RMSE	$2,\!890\%$	0,196	0,340	$5{,}76\%$	0,788

Com os dados da Tabela 10, conclui-se que a rede de menor RMSE obteve erro absoluto máximo, erro relativo máximo e U de Theil menores que a rede de menor MAPE.

4.5 Comparação entre os modelos propostos

Quanto ao número de neurônios nas camadas da RNA, em geral, os valores médios de MAPE e RMSE foram menores em redes com poucos neurônios nas camadas, os valores médios de MAPE e RMSE foram altos se tinha muitos neurônios em uma camada e poucos na outra, conforme pode ser comprovado pelas Figuras 23, 25, 26, 28, 31, 33, 36 e 38.

Das redes dos modelos propostos que apresentaram menores MAPE e RMSE, o horizonte de previsão (HP) variou entre duas e seis semanas, o que demonstra que a previsão se saiu melhor com horizontes de previsão curtos, já a janela de previsão (JP) das melhores redes ficou entre a faixa de dezesseis a vinte semanas, o que indica que a previsão se sai melhor quando há mais dados anteriores disponíveis. Na Tabela 11 são apresentados os valores das métricas de desempenho e alguns dos parâmetros das melhores redes de cada modelo.

	Rede de menor MAPE				Rede de menor RMSE			
Modelo	1	2	3	4	1	2	3	4
MAPE	$2,\!650\%$	2,849%	3,050%	2,851%	$2,\!650\%$	2,920%	3,111%	2,890%
RMSE	0,182	0,194	0,206	0,200	0,182	0,188	0,198	0,196
EA máx	0,363	0,345	0,342	0,401	0,363	0,360	0,323	0,340
ER máx	5,220%	5,303%	5,960%	6,770%	5,220%	5,510%	5,321%	5,760%
U de Theil	0,731	0,792	0,802	0,814	0,731	0,785	0,830	0,788
HP	20	18	20	20	20	20	20	16
JP	2	6	6	2	2	6	2	4
n1	18	1	3	11	18	3	1	3
n2	9	7	6	2	9	11	1	15

Tabela 11 – Comparação das melhores redes de cada um dos modelos propostos.

Fonte: elaborado pelo autor, 2019.

Com os dados da Tabela 11, é visto que a rede do modelo com apenas potência como variável de entrada, apresenta os menores valores de MAPE, RMSE, ER máximo e U de Theil, o que a elege como a melhor rede entre todos os modelos propostos. Vale ressaltar também, que as redes do segundo modelo apresentam métricas de desempenho melhores que as redes do terceiro modelo, um fator que pode explicar isso é a correlação maior entre a energia acumulada na semana com a potência máxima semanal, do que a correlação entre a variação de energia entre semanas e a potência máxima semanal.

4.6 Teste de robustez da melhor rede

Para avaliar a robustez da melhor rede neural encontrada, decidiu-se utilizar como dados de entrada da rede, valores que já haviam sido previstos pela mesma. Na Tabela 12 pode ser vista como foi feito esse processo.

P(i)	P(i-1)	P(i-2)	P(i-3)	P(i+1)
$5,\!070$	4,794	$5,\!054$	5,019	5,202
5,202	$5,\!070$	4,794	$5,\!054$	4,654
$4,\!654$	$5,\!202$	$5,\!070$	4,794	4,621
$4,\!621$	$4,\!654$	$5,\!202$	$5,\!070$	4,450
$4,\!450$	4,621	$4,\!654$	5,202	4,317
4,317	4,450	4,621	4,654	5,202

Tabela 12 – Exemplo de como foi feita a realimentação da rede.

Fonte: elaborado pelo autor, 2019.

Pelos dados apresentados na Tabela 12, é possível notar que o valor previsto em uma iteração é utilizado como valor de entrada na iteração seguinte, até que chega um ponto em que todos os dados de entrada são frutos de previsões anteriores. O resultado desse processo de realimentação da RNA pode ser visto na Figura 41.



Figura 41 – Comparação entre o resultado da melhor rede com e sem realimentação.

Fonte: elaborado pelo autor, 2019.

Pelas curvas da Figura 41, vemos que o resultado da previsão da rede com realimentação e da rede sem alimentação são muito próximos, nas primeiras semanas os valores diferem muito pouco e nas semanas finais a diferença é maior, mas ainda assim é uma diferença pequena, com MAPE menor que 1% em relação a rede sem realimentação, isso demonstra a consistência dos resultados de previsão da melhor rede.

5 Conclusão

Neste trabalho, foi proposto um modelo de previsão de demanda, visando criar uma ferramenta que auxilie especialistas do setor de planejamento estratégico de sistemas de distribuição na tomada de decisões. O modelo proposto tem como diferencial a utilização de dados da energia acumulada e variações da energia para tentar melhorar o desempenho da previsão.

Para a realizar a previsão de demanda optou-se pelo uso de técnicas de inteligência artificial, mais especificamente, redes neurais artificiais. A rede utilizada foi do tipo *multi-layer perceptron*, com função de ativação tangente hiperbólica, algoritmo de aprendizado de Levenberg-Marquardt e duas camadas de neurônios ocultos.

Para validar o modelo proposto, foi feito um estudo de caso para a previsão de demanda no alimentador 21L1 da subestação João Pessoa, localizada em João Pessoa-PB. Nesse estudo, foi utilizado um conjunto de medições de potência ativa cedidas pela concessionária de energia, para criar um banco de dados com informações de potência máxima semanal, energia acumulada na semana e variações da energia. Com o banco de dados prontos foram definidos os parâmetros da rede neural, a rede foi treinada e foram realizados testes para a validação do modelo proposto.

Para avaliação da acurácia da previsão do modelo proposto foram calculadas algumas métricas tais como MAPE, RMSE, erro absoluto, erro relativo e U de Theil. Em relação a janela de previsão e sua influência na previsão, pode-se concluir que previsões com uma menor janela de previsão, entre duas e seis semanas, apresentam, em média, resultados de previsão melhores, quanto ao horizonte de previsão, pode-se concluir que com mais dados anteriores a se torna, em média, mais precisa.

Em relação a comparação entre modelos propostos neste trabalho, a adição da variação de energia afetou negativamente o desempenho da previsão, o mesmo foi visto com a adição da energia acumulada na semana, porém o modelo que utilizou a energia acumulada na semana apresentou melhor previsão quando comparado ao modelo que usou a variação de energia, isto pode ser explicado pela maior correlação entre demanda e energia acumulada do que entre demanda e da variação de energia. Por fim, o modelo que apresentou melhor desempenho utilizou apenas dados anteriores de demanda para prever a demanda futura, este modelo apresentou MAPE mínimo de 2,65% e erro relativo máximo de 5,22%.

Como trabalhos futuros sugere-se fazer a previsão de um alimentador que não apresente tanta sazonalidade quanto o alimentador 21L1, utilizado neste trabalho, ou também um alimentador que apresente característica de carga industrial ou residencial, já que ao tipo da carga do alimentador 21L1 parece mais comercial, visto sua maior demanda em hórarios comerciais e também a sua queda no consumo aos finais de semana quando comparado aos dias úteis. Também sugere-se utilizar variáveis externas com maior correlação com a demanda, para tentar melhorar a acurácia da previsão, além de utilizar técnicas híbridas, como a Neuro-Fuzzy, também com o intuito de aperfeiçoar a previsão.

6 Referências bibliográficas

ANEEL. Atlas de energia elétrica do Brasil. Agência Nacional de Energia Elétrica, 2008.

ANEEL. Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST Módulo 2 – Planejamento da Expansão do Sistema de Distribuição, 2016. Disponível em: https://www.aneel.gov.br/prodist. Acesso em: 20 de agosto de 2019.

ANEEL. Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST Módulo 6 – Informações Requeridas e Obrigações, 2018. Disponível em: https://www.aneel.gov.br/prodist. Acesso em: 20 de agosto de 2019.

BALIYAN, A. B.; GAURAV, K.; MISHRA, S. K. A Review of Short Term Load Forecasting using Artificial Neural Network Models. International Conference on Intelligent Computing, Communication & Convergence, p. 121–125, 2015.

BEZERRA, S. G. T. A. Reservoir Computing com Hierarquia para Previsão de Vazões Médias Diárias. Dissertação (Mestrado em Engenharia da Computação) - Universidade de Pernambuco, Recife, 2016.

BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. Time series analysis: forecasting and control. 4 edition ed. Hoboken, N.J: Wiley, 2008.

CARMONA, D.; JARAMILLO, M. A.; GONZALEZ, E.; ALVAREZ, J. A. Electric energy demand forecasting with neural networks. IEEE 28th Annual Conference of the Industrial Electronics Society. IECON 02, Sevilla, vol 3, p. 1860-1865, 2002.

CASTRO, L. N.; ZUBEN, F. J. V. **Redes Neurais Artificiais**. Universidade Estadual de Campinas - Departamento de Engenharia de Computação e Automação Industrial, 2016.

CERRI, R. Redes neurais e algorítmos genéticos para problemas de classificação hierárquica multirrótulo.. Universidade de São Paulo, São Carlos - SP, 2014.

CHEMTOVA, S.; SANTOS, P. J.; NEVES, M. V. Load peak forecasting in different load patterns situations. 10th International Conference on Compatibility, Power Electronics and Power Engineering (CPE-POWERENG), Bydgoszcz, p. 148-151, 2016.

CHURCHLAND, P. S. ; SEJNOWSKI, T. J. Computational neuroscience: The computational brain. The MIT Press, Massachusetts US, 1992.

COHEN, J. Statistical power analysis for the behavioral sciences. Hillsdale, New Jersey, Erlbaum, 1988.

DUCH, W.; JANKOWSKI, N. Survey of neural transfer functions. Neural Computing Surveys, v. 2, p. 163–212, 1999.

EHLERS, R. S. Análise de séries temporais. 2009. Disponível em: http://conteudo.icmc.usp.br/pessoas/ehlers/stemp.pdf>. Acesso em: 17 de agosto de 2019.

EPE. Resenha mensal do mercado de energia elétrica - Janeiro de 2019. Empresa de Pesquisa Energética, 2019.

EPE. Projeção de demanda de energia elétrica para os próximos 10 anos (2017-2026). Empresa de Pesquisa Energética, 2017.

EPE. Balanço Energético Nacional 2018 - Ano Base 2017. Empresa de Pesquisa Energética, 2018a.

EPE. Planejamento do Atendimento aos Sistemas Isolados Horizonte 2023 - Ciclo 2018. Empresa de Pesquisa Energética, 2018b.

EPE. **Plano Decenal de Expansão de Energia 2027.** Empresa de Pesquisa Energética, 2018c.

FAVA, V. L. Metodologia de Box-Jenkins para Modelos Univariados. In: Vasconcellos, M.A.S. and Alves, D., Eds., Manual de econometria: nível intermediário, Atlas, São Paulo, 205-232, 2000. FRANCO JUNIOR, E. F. Modelo para previsão de demanda ativa usando rna. Universidade Estadual de Campinas, 2013.

GHANBARI, A.; HADAVANDI, E.; ABBASIAN-NAGHNEH, S. Comparison of Artificial Intelligence Based Techniques for Short Term Load Forecasting. Third International Conference on Business Intelligence and Financial Engineering, Hong Kong, p. 6-10, 2010.

GUIRELLI, C. R. **Previsão da carga de curto prazo de áreas elétricas através de técnicas de inteligência artificial.** Biblioteca Digital de Teses e Dissertações da USP, 2006.

GUO, Z.; LI, W.; LAU, A.; INGA-ROJAS, T.; WANG, K. Detecting X-Outliers in Load Curve Data in Power Systems. IEEE Transactions on Power Systems, v. 27, n. 2, p. 875–884, 2012.

HAYKIN, S. S. Neural networks: a comprehensive foundation. MacMillan Publishing Company, 1994.

HAYKIN, S. S. **Redes neurais - principios e prática**. 2 edição. Porto Alegre: Bookman, 2001.

HEBB, D. O. The organization of behavior. New York: Wiley, 1949.

HIPPERT, H. S.; PEDREIRA, C. E.; SOUZA, R. C. Neural networks for short-term load forecasting: a review and evaluation. IEEE Transactions on Power Systems, vol. 16, no. 1, p. 44-55, Feb 2001.

HOBBS, B. F.; JITPRAPAIKULSARN, S.; KONDA, S.; CHANKONG, V.; LOPARO, K. A.; MARATUKULUM, D. J. Analysis of the value for unit commitment of improved load forecasts. IEEE Transactions Power Systems, v. 14, n. 4, p. 1342-1348, 1999.

KAGAN, N.; OLIVEIRA, C. C. B. de; ROBBA, E. J. Introdução aos sistemas de distribuiçao de energia. Edição: 1 ed. São Paulo: Edgard Blucher, 2005.

KHATOON, I. S.; IGRAGEEEM; SINGH, A. K.; PRITI. Analysis and comparison of various methods available for load forecasting: An overview. Innovative Applications of Computational Intelligence on Power, Energy and Controls with their impact on Humanity (CIPECH), Ghaziabad, p. 243-247, 2014.

KIM, S.; KIM, H. new metric of absolute percentage error for intermittent demand forecasts. International Journal of Forecasting, v. 32, n.3, p. 669-679, 2016.

LECUN, Y.; BOTTOU, L.; ORR, G. B.; MULLER, K. Efficient BackProp. In: Neural Networks: Tricks of the Trade, Springer-Verlag, Londres, 9-50, 2012.

MACRINI, J. L. R. Modelo de previsão de carga utilizando redes neurais: otimização camada a camada. PUC-Rio, 2000.

MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; HYNDMAN, R. J. Forecasting: Methods and Applications. 3. ed. John Wiley Sons, 1998.

MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. Innovative Applications of Computational Bulletin of Math. Biophys, v. 5, p. 115–133, 1943.

MEDEIROS, I. E. A. de. Aplicação de redes neurais artificiais para a previsão de demanda em subestações de distribuição considerando a influência da temperatura. Trabalho de conclusão de curso - Universidade Federal da Paraíba, João Pessoa, 2019.

MEDEIROS, R. Á. O. de. **Previsão de demanda no médio prazo utilizando redes neurais artificiais em sistemas de distribuição de energia elétrica**. Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal da Paraíba, João Pessoa, 2016.

MINSKY, M.; PAPERT, S. Perceptrons. Oxford, England: M.I.T. Press, 1969.

MORETTIN, P. A.; TOLOI, C. M. C. Modelos para Previsão de Séries Temporais. IMPA. Rio de Janeiro, 1981.

 $\rm NIU, \, D.; \, SHI, \, H.; \, LI, \, J. \; ; \; XU, \, C.$ Research on power load forecasting based on

combined model of Markov and BP neural networks. 8th World Congress on Intelligent Control and Automation, Jinan, p. 4372-4375, 2010.

NORVIG, P.; RUSSELL, S. Inteligência artificial. Edição: 2a ed. [s.l.] Elsevier Acadêmico, 2004.

ONS. Disponível em: <http://www.ons.org.br/paginas/sobre-o-sin/mapas>. Acesso em: 17 de agosto de 2019.

PAN, X.; LEE, B.A comparison of support vector machines and artificial neural networks for mid-term load forecasting. mar. 2012.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychological review, v. 65, n. 6, p. 386, 1958.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. Nature, v. 323, p. 533–536, 1986.

SHCHERBAKOV, M.V.; BREBELS, A.; SHCHERBAKOVA, N. L.; TYUKOV, A. P.; JANOVSKY, T. A.; KAMAEV, V. A. **A Survey of Forecast Error Measures.** World Applied Sciences Journal, v.24, n.24 (Information Technologies in Modern Industry, Education Society), p.171-176, 2013.

SINGH, S.; HUSSAIN, S.; BAZAZ, M. A. Short term load forecasting using artificial neural network. Fourth International Conference on Image Information Processing (ICIIP), Shimla, p. 1-5, 2017.

SKOLTHANARAT, S.; LEWLOMPHAISARL, U.; TUNGPIMOLRUT, K. Short-term load forecasting algorithm and optimization in smart grid operations and planning. IEEE Conference on Technologies for Sustainability (SusTech), Portland, OR, p. 165-171, 2014.

TEPCO. Disponível em: <http://www.tepco.co.jp/index-j.html>. Acesso em: 17 de agosto de 2019.