

UNIVERSIDADE FEDERAL DE SANTA CATARINA
JHUAN DE SOUZA OLIVEIRA

**DESENVOLVIMENTO E TREINAMENTO DE REDES NEURAS ARTIFICIAIS PARA
PROCESSAMENTO DE DADOS DE RADIAÇÃO SOLAR**

ARARANGUÁ

2017

JHUAN DE SOUZA OLIVEIRA

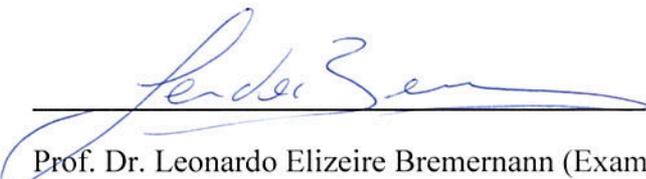
**TRABALHO DE CONCLUSÃO DE CURSO DE GRADUAÇÃO EM
ENGENHARIA DE ENERGIA**

Trabalho de conclusão de curso na modalidade de artigo científico para a aprovação na disciplina de Trabalho de Conclusão de Curso do curso de Engenharia de Energia da Universidade Federal de Santa Catarina – Campus Araranguá

Aprovado em: 06/07/2017



Prof. Dr. Giuliano Arns Rampinelli (Orientador)



Prof. Dr. Leonardo Elizeire Bremernann (Examinador)



Prof. Dr. Luciano Lopes Pfitscher (Examinador)

DESENVOLVIMENTO E TREINAMENTO DE REDES NEURAS ARTIFICIAIS PARA PROCESSAMENTO DE DADOS DE RADIAÇÃO SOLAR

Jhuan de Souza Oliveira*

RESUMO

A energia solar fotovoltaica está sendo inserida gradualmente na matriz energética mundial como uma fonte de energia limpa e sustentável. Atualmente no Brasil, para que um projeto de parque ou usina fotovoltaica seja habilitado em um leilão de energia é obrigatória a apresentação do histórico das medições de radiação solar, por um período contínuo de doze meses. Os dados de radiação solar são fundamentais para as estimativas de geração de energia elétrica dos empreendimentos. A precisão das estimativas está diretamente correlacionada com a qualidade e confiabilidade dos dados de radiação solar. Este trabalho apresenta o desenvolvimento, treinamento e validação de uma rede neural artificial (RNA) para identificação e preenchimento de lacunas em dados de radiação solar. Os dados de radiação solar foram obtidos através da base de dados meteorológicos do INMET para a cidade de Araranguá, Santa Catarina. O desempenho da rede neural artificial será comparado com métodos empíricos de replicação usualmente utilizados para o preenchimento de lacuna nos dados de radiação solar. O critério de avaliação do desempenho é o erro médio quadrático dos dois métodos testados em relação a populações de teste criadas para a validação. Foi possível verificar que a rede neural artificial apresentou um erro quadrático médio 24,6% menor quando comparado ao método de referência. Portanto, do ponto de vista energético a RNA apresenta maior precisão no tratamento dos dados de radiação solar, garantindo maior confiabilidade na estimativa de geração de energia elétrica.

Palavras-chave: Radiação solar. Rede neural artificial. Métodos preditivos.

DEVELOPMENT AND TRAINING OF ARTIFICIAL NEURAL NETWORKS TO PROCESSING SOLAR RADIATION DATA

Jhuan de Souza Oliveira

ABSTRACT

Solar energy is being gradually inserted into the world energy base in the last years worldwide providing a clean and renewable source of energy. Nowadays in Brazil, for projects of solar power plants to be qualified in an auction of energy it is mandatory to present the history of measurements of solar radiation, in a continuous period of twelve months. The solar radiation data are essential to estimate of the electric power generation. This work project presents the development, training and validation of an artificial neural network (RNA) to identify and fill gaps in solar radiation data. The solar radiation data were obtained through the INMET meteorological database for the city of Araranguá, Santa Catarina. The performance of the artificial neural network will be compared with empirical methods of replication usually used to fill the gap in the solar radiation data. The criterion of performance evaluation is the mean square error (MSE) of the two methods tested in relation to test populations created for the validation. It was possible to verify that the artificial neural network presented a mean square error 24.6% smaller when compared to the reference method. Therefore, from the energy point of view the RNA tool presents greater precision in the treatment of solar radiation data, guaranteeing greater reliability in the generation of electric energy.

Keywords: Solar irradiation. Artificial neural network. Predictive Methods.

1 INTRODUÇÃO

O sol é a fonte de energia primordial para a criação e manutenção da vida na Terra. Diversos processos são afetados direta ou indiretamente pela radiação solar, tais como: processos de fotossíntese, agropecuária, agricultura e processos de geração de energia. O recurso solar é gratuito, abundante e inesgotável, por isso tem-se tornado cada dia mais importante na matriz energética nacional como fonte de energia limpa e distribuída (COLLE; PEREIRA, 1998).

A Radiação solar apresenta características de não-linearidade e natureza estocástica, em decorrência disso, os dados de radiação solar apresentam alta variabilidade (GUARNIERI, 2006). Se faz necessário uma análise computacional que seja capaz de rastrear e interpretar o perfil da variação dos dados de radiação, afim de, auxiliar a tomada de decisão das estações solarimétricas. O tratamento dos dados de radiação solar encontra uma aplicação na área de energia visando o planejamento e a otimização do recurso para os sistemas fotovoltaicos, solar térmicos e híbridos. No Brasil, a energia solar é uma das mais promissoras opções energéticas, uma vez que a maior parte do seu território está localizado na região inter-tropical e recebe elevada irradiação solar durante todo o ano (MARTINS, 2007).

O último leilão de energia, com participação de empreendimentos fotovoltaicos, foi o 2º leilão de energia de reserva de 2015 realizado pelo Ministério de Minas e Energia. Neste leilão, foram 649 projetos cadastrados, sendo que 493 projetos foram habilitados para concorrência. A potência contratada foi de 1.115,9 MW_p. O início do suprimento da energia elétrica dos projetos contratados deve ocorrer em 1º de novembro de 2018, com prazo de suprimento garantido em pelo menos 20 anos. Neste leilão ainda não era obrigatória a apresentação do histórico de medição contínuo da irradiação global horizontal no local do empreendimento, porém, o art. 6º-B6, inciso II, da Portaria MME nº 21/2008, já estabelece que, para empreendimentos fotovoltaicos, apenas a partir de 2016 terão que obrigatoriamente apresentar os dados de radiação solar (EPE, 2016).

Redes Neurais Artificiais são ferramentas que computam dados de maneira semelhante àquela realizada pelas redes de neurônios biológicos. As redes são compostas por elementos de processamento interconectados (neurônios). Na RNA empregada, chamada Perceptron de múltiplas camadas (MLP), os neurônios são dispostos em camadas e os sinais passam desde a camada de entrada até a camada de saída através de conexões, chamadas sinapses (HAYKIN, 1994).

Este trabalho apresenta o desenvolvimento, treinamento e validação de redes neurais artificiais (RNAs) para tratamento e processamento de dados de radiação solar. Os dados de treinamento da RNA são dados horários e contínuos provenientes de uma estação meteorológica localizada na cidade de Araranguá, Santa Catarina. O período de integralização dos dados é de um ano. O intuito é provar que a RNA é capaz de inferir valores nas lacunas com maior precisão, se comparada ao método de replicação de dados utilizados atualmente.

2 REFERENCIAL TEÓRICO

A seguir serão descritos os conceitos básicos para o entendimento deste trabalho.

2.1 Radiação solar

O Brasil, por ser um país localizado na sua maior parte na região inter-tropical, possui grande potencial para aproveitamento de energia solar durante todo ano (COLLE 1998; TIBA 2000). A radiação solar pode ser aproveitada de diversas formas como por exemplo: captura de biomassa, aquecimento de ar e água para fins domésticos e industriais, ciclos termodinâmicos variados e, como principal justificativa no ponto de vista energético, a produção de energia elétrica (PEREIRA; VRISMAN; GALVANI, 2002).

A utilização da energia solar na matriz energética nacional traz benefícios ao país viabilizando o desenvolvimento de regiões remotas (onde o custo da eletrificação pela rede convencional é demasiadamente alto com relação ao retorno financeiro do investimento), regulando a oferta de energia em períodos de estiagem, diminuindo a dependência do mercado de petróleo e reduzindo as emissões de gases poluentes. Existe um grande leque de possibilidades a médio e longo prazo para aproveitamento dessa abundante forma de energia renovável, que vai desde pequenos sistemas fotovoltaicos autônomos até as grandes centrais que empregam energia solar.

As primeiras centrais fotovoltaicas com potências da ordem de 1 MWp foram instaladas recentemente no Brasil. A chamada nº 13/2011 da Agência Nacional de Energia Elétrica (ANEEL), denominada Projeto Estratégico: Arranjos Técnicos e Comerciais para Inserção da Geração Solar Fotovoltaica na Matriz Energética Brasileira aprovou diversos projetos de pesquisa,

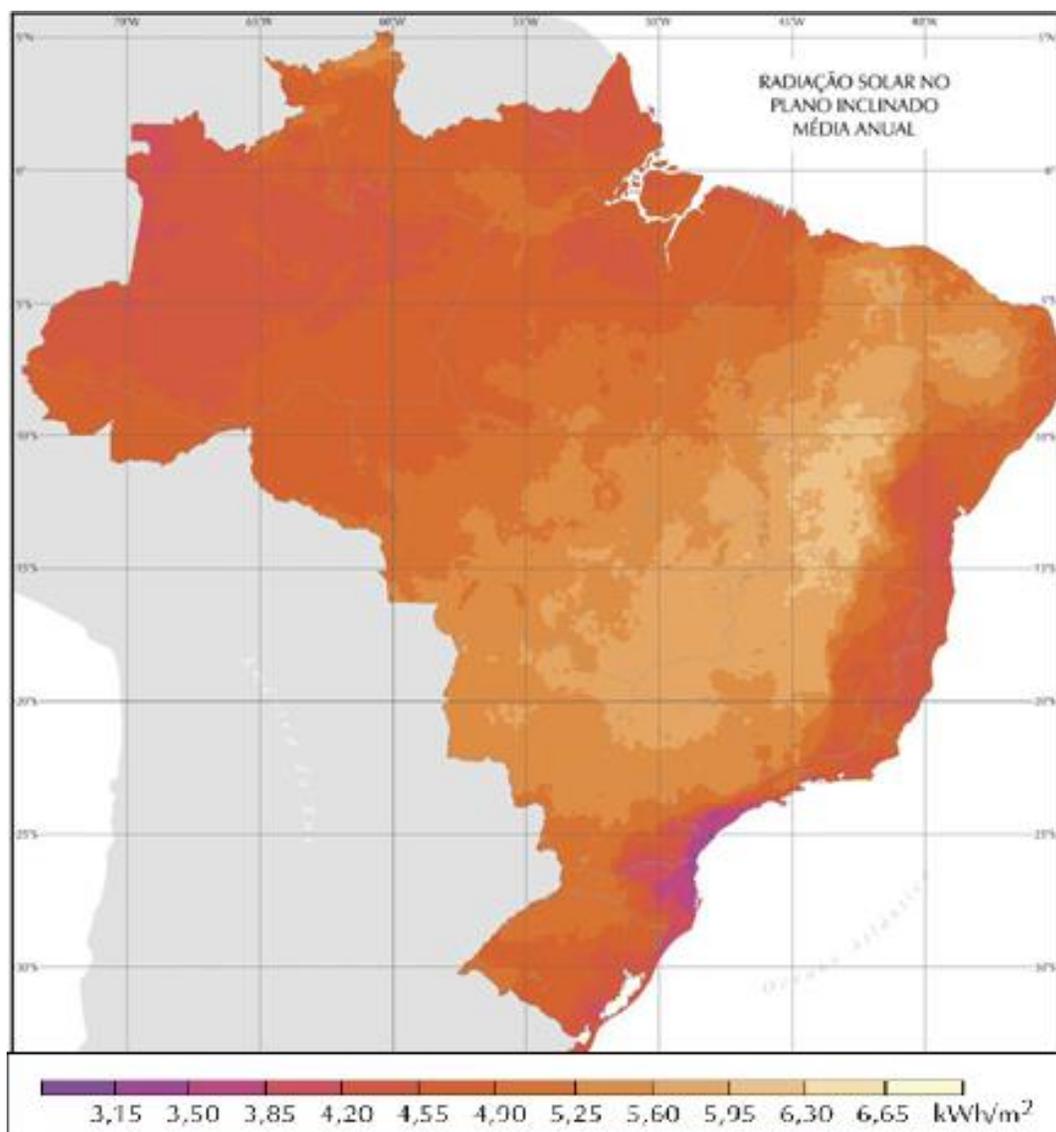
desenvolvimento e inovação de centrais fotovoltaicas entre 0,5 MW e 3 MW a serem desenvolvidos nos próximos anos, totalizando 24,4 MW (BÜHLER, RAMPINELLI, 2015).

Também se destacam, neste cenário promissor para o aproveitamento da energia solar fotovoltaica, iniciativas como a resolução normativa 482/2012 da ANEEL, na qual estabelece as condições gerais para o acesso de mini e microgeração distribuída de energia. Essa normativa foi atualizada pela resolução normativa 687/2015 da ANEEL, onde foi inserido novos modelos de geração distribuída como autoconsumo remoto, geração compartilhada e empreendimentos com múltiplas unidades consumidoras. Além disso, a 687/2015 alterou os limites de microgeração e minigeração distribuída aumentando os limites de geração, de 1MW para 5MW de potência instalada (EPE, 2016).

A radiação solar que incide na superfície terrestre é composta por três componentes: Direta, Difusa e o Albedo. A radiação direta é aquela que provém diretamente da direção do sol e produz sombras nítidas. A radiação difusa é aquela que sofre espalhamento pela atmosfera terrestre. E o albedo é a componente da radiação que é refletida pelo ambiente do entorno (solo, vegetação, obstáculos, terrenos rochosos, etc.). Mesmo em um dia totalmente sem nuvens, pelo menos 20% da radiação que atinge uma superfície é difusa e durante os dias nublados 100% da radiação será difusa (CEPEL, 2014).

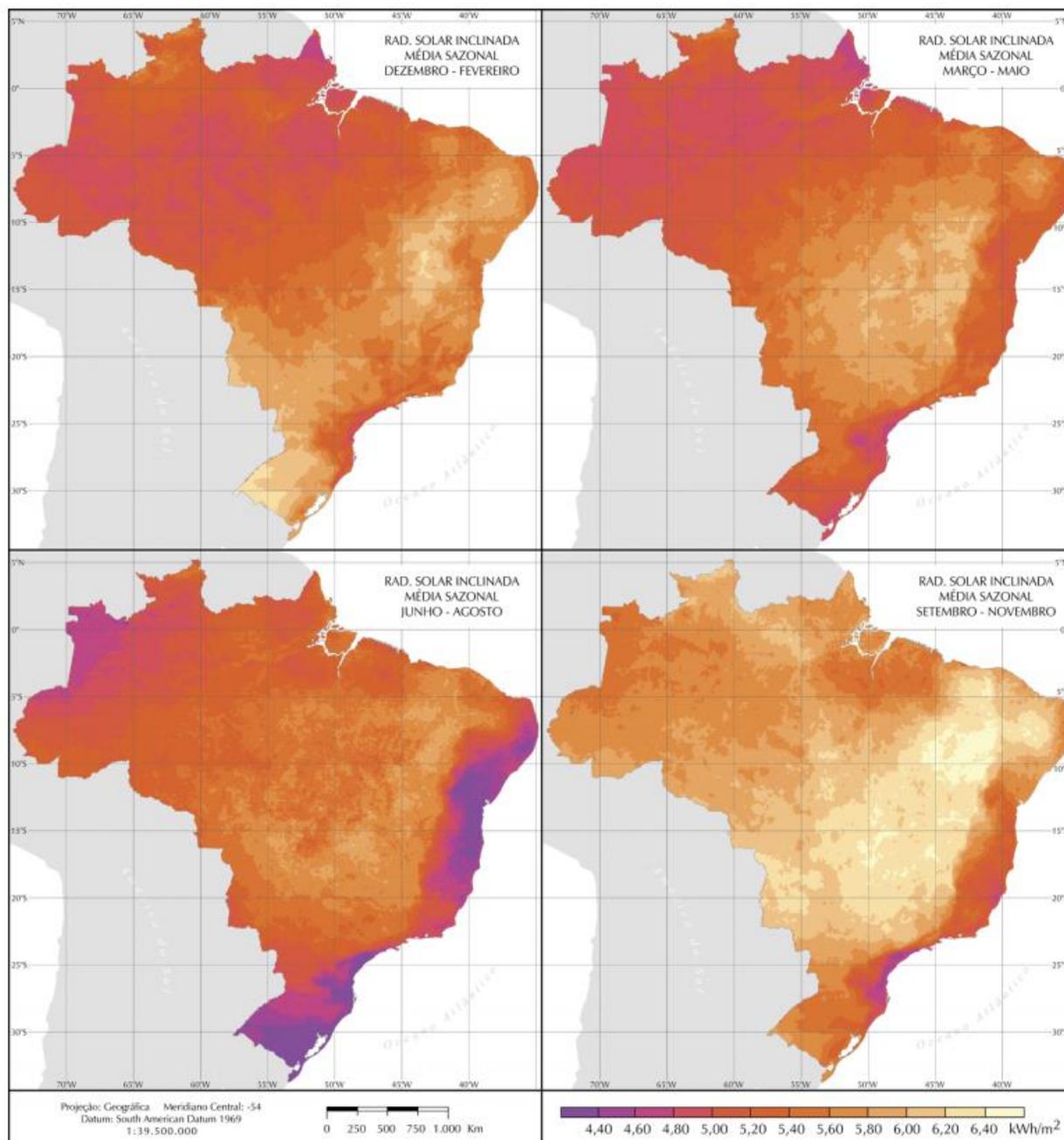
A Figura 1 apresenta o mapa solarimétrico brasileiro onde é possível observar o valor de irradiação solar média diária estimado para cada região do país. É possível observar que o sul do Brasil possui uma Irradiação (quantidade de energia solar incidente por m^2) menor do que comparada a outras regiões como Norte e Nordeste. Isso ocorre pelo distanciamento dos trópicos, onde fica mais concentrada a irradiação solar. Porém, isso não chega a ser um problema para instalação de sistemas fotovoltaicos no Sul do país. Países como Alemanha, por exemplo, possuem índices menores de irradiação solar incidentes e mesmo assim possuem uma fração significativa da sua geração de energia elétrica suprida pela geração solar (INPE, 2006).

Figura 1 - Mapa Brasileiro de Irradiação Solar Média diária anual. Fonte (INPE, 2006)



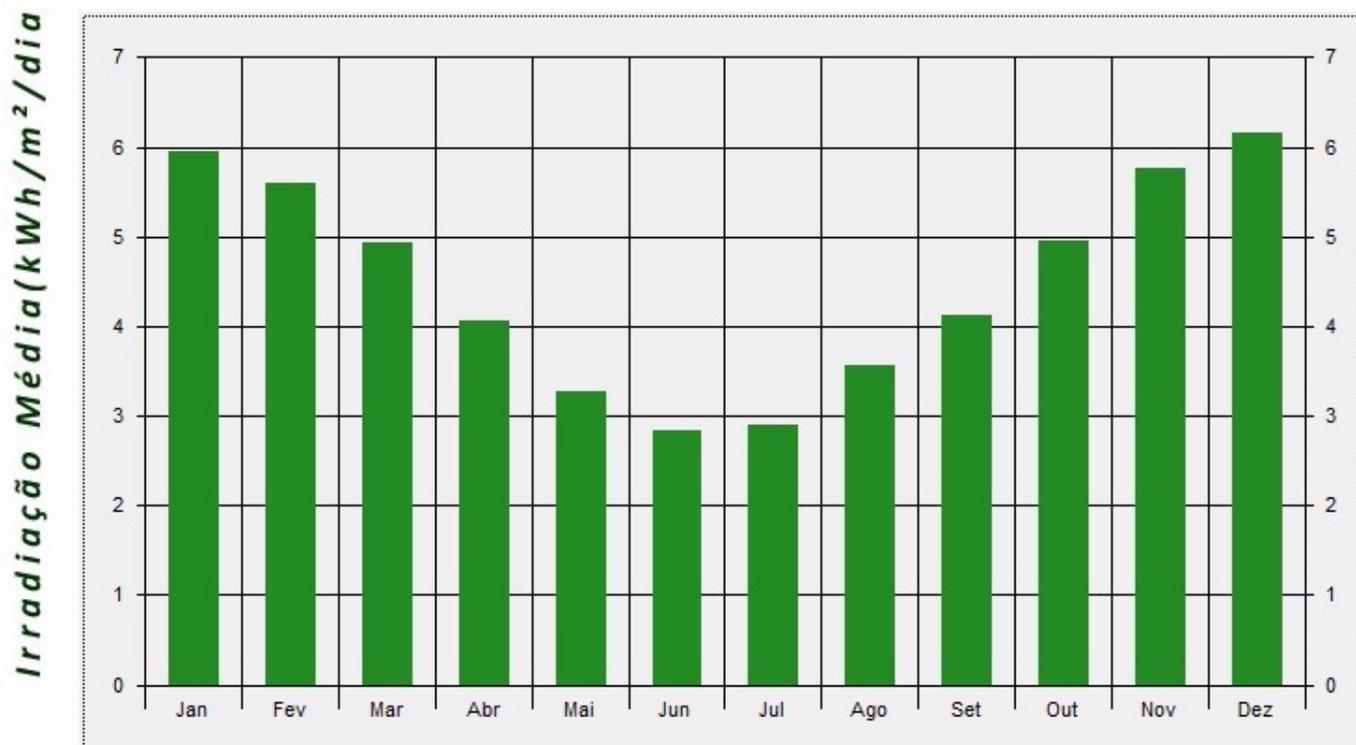
A radiação solar também possui características de sazonalidade associadas com o movimento de translação terrestre, variando sua intensidade com a aproximação solar nos períodos de inverno e verão. Essa variabilidade pode ser observada na Figura 2 (INPE, 2006). A variabilidade diária da radiação solar implica em variabilidade de geração de energia elétrica de um sistema ou usina fotovoltaica. Neste contexto, modelos confiáveis de previsibilidade de curto prazo da radiação solar podem auxiliar no planejamento e operação de uma usina fotovoltaica.

Figura 2- Médias diárias de irradiação solar sazonal. (Fonte: INPE, 2006)



Analisando a região onde este trabalho é focado, pode-se observar que a região sul do Brasil apresenta uma forte variação sazonal na irradiação solar média. Na Figura 3 é possível observar que a irradiação solar média diária por metro quadrado de área é maior durante os meses de dezembro a março, no verão e menor durante os meses de junho a agosto, no inverno.

Figura 3-Irradiação média diária anual do plano horizontal para cada mês na região sul de Santa Catarina. (RADIASOL2,2016)



2.2 Estações da rede SONDA e rede INMET

O Sistema de Organização Nacional de Dados Ambientais (SONDA) para o setor de energia é um projeto desenvolvido e coordenado pelo Centro de Previsão do Tempo e Estudos Climáticos do Instituto Nacional de Pesquisas Espaciais (CPTEC/INPE) para implementação de uma rede de coleta de dados destinada levantar e melhorar a base de dados sobre os recursos de energia solar e eólica no Brasil (GUARNIERI 2006). A base de dados gerada pelo projeto SONDA tem como objetivo principal dar suporte técnico-científico ao setor energético através do atendimento à demanda de informações confiáveis sobre os recursos renováveis de energia, e à capacitação de recursos humanos. O projeto SONDA possibilita a aquisição e aprimoramento do conhecimento científico sobre o potencial disponível dos recursos solar e eólico, bem como sobre a variabilidade dos recursos em função de causas naturais e antropogênicos. A rede SONDA possui um total de 25 estações, sendo que algumas ainda se encontram em fase de implantação. As

estações SONDA estão divididas em quatro categorias: estações de referência, estações solares avançadas, estações solares básicas e estações eólicas.

A rede SONDA como um todo tem sua base de dados dividida em 2 grupos apresentados na Tabela 1. Na Figura 4 é possível observar a localização das estações da rede SONDA e o tipo de estação: referência, solares avançadas, solares básicas e torres eólicas.

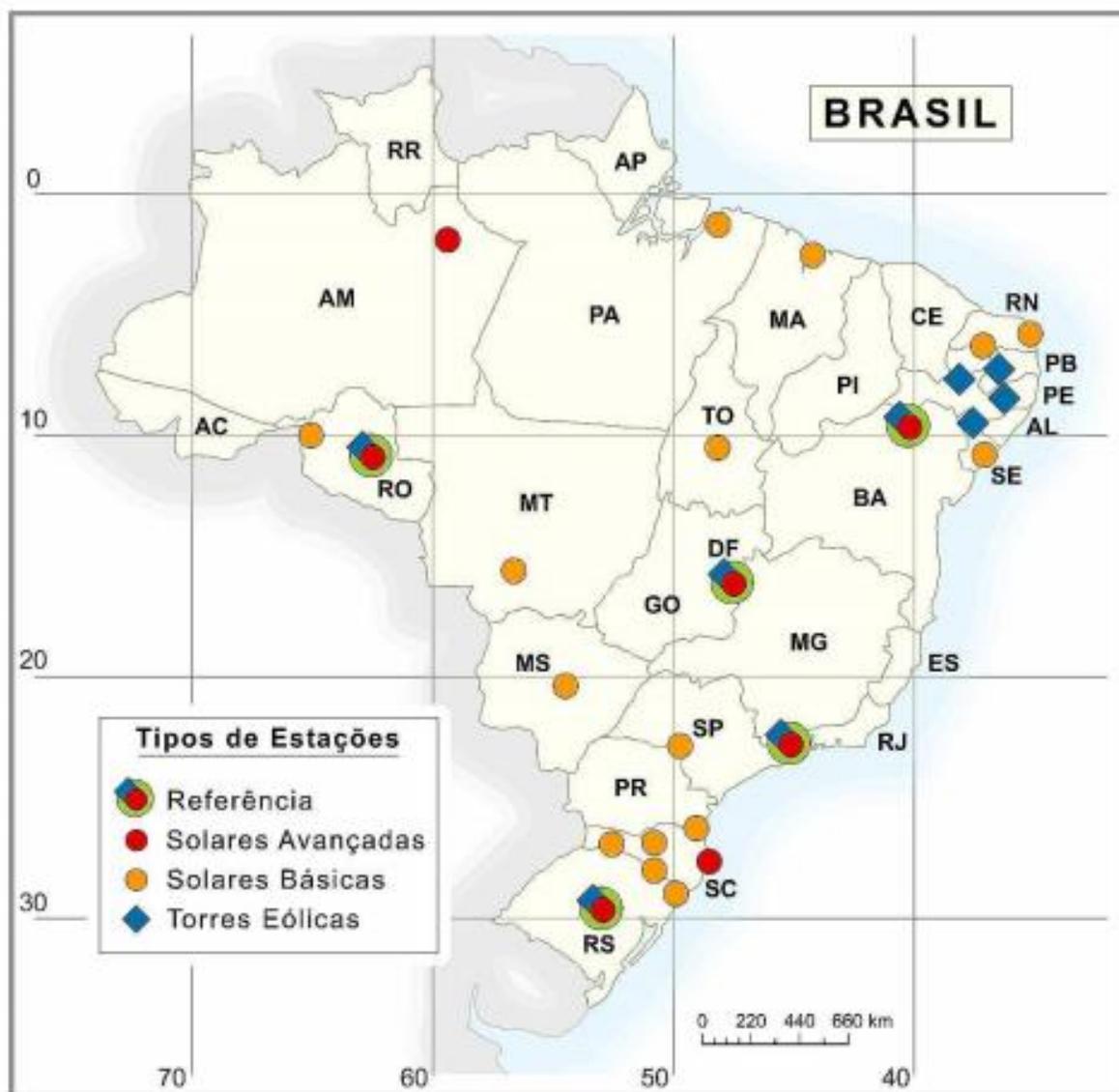
Tabela 1- Dados de coleta SONDA.

	Dados meteorológicos	Dados anemométricos
Periodicidade	Médias de 1 em 1 hora	Médias de 10 em 10 minutos
Tipo de dado	Radiométricos: . Radiação global horizontal; . Radiação direta Normal; . Radiação difusa; . Radiação de onda longa descendente; . Radiação fotossinteticamente ativa (PAR); . Iluminância.	. Velocidade do vento para 25 e 50 m de altura; . Direção do vento para 25 e 50 m de altura; . Temperatura para 25 e 50 m de altura.

Fonte: MARTINS, 2007.

A base de dados da rede SONDA, embora de alta qualidade e confiabilidade, ainda é bastante reduzida temporal e espacialmente. A série de dados tem início no segundo semestre de 2004, com exceção da estação de Florianópolis que já operava sob a coordenação do Laboratório de Energia Solar da Universidade Federal de Santa Catarina antes de sua incorporação à rede SONDA.

Figura 4 - Localização das estações de superfície da rede SONDA (Fonte: MARTINS, 2007).



O sistema de informação meteorológica automática de superfície do INMET compreende o: sub-sistema de coleta de dados, através de sensores que medem as variáveis meteorológicas; sub-sistema de controle e armazenamento local em *data-logger*; sub-sistema de energia; sub-sistema de comunicações; sub-sistema de banco de dados; e sub-sistema de disseminação de dados aos usuários, de forma aberta e gratuita pela internet (INPE, 2011).

Uma estação meteorológica automática (EMA) coleta, de minuto a minuto, as informações meteorológicas (temperatura, umidade, pressão atmosférica, precipitação, direção e velocidade dos ventos, radiação solar) representativas da área em que está localizada. A cada hora, estes dados

são integralizados e disponibilizados para serem transmitidos, via satélite ou telefonia celular, para a sede do INMET, em Brasília. O conjunto dos dados recebidos é validado, através de um controle de qualidade e armazenado em um banco de dados. Além disto, os dados são disponibilizados gratuitamente, em tempo real, através da internet (http://www.inmet.gov.br/sonabra/maps/pg_automaticas.php) para a elaboração de previsão do tempo e dos produtos meteorológicos diversos de interesse de usuários setoriais e do público em geral e para uma vasta gama de aplicações em pesquisa em meteorologia, hidrologia e oceanografia (INPE,2011).

A estação utilizada como base para a busca dos dados é a estação Araranguá-A867. Os dados coletados podem ser observados na Tabela 2. A Figura 5 apresenta a localização das estações meteorológicas automáticas do INMET. A disponibilidade dos dados para consulta popular é restrita aos últimos 365 dias. Porém, há possibilidade da obtenção de dados históricos para fins acadêmicos.

Tabela 2- Dados e unidades coletados pela estação A867.

Dados	Unidades
Temperatura	°C
Umidade	%
Ponto de orvalho	°C
Pressão atmosférica	Pa
Velocidade de vento	m/s
Direção do vento	°
Radiação	kJ/m ²
Precipitação	mm

Figura 5- Mapa das estações meteorológicas automáticas do INMET (Fonte: INMET, 2011).



2.3 Métodos de referência

As lacunas de dados são intervalos de falha de medição de um equipamento causadas por fatores externos, por exemplo: falhas elétricas, falhas humanas, intervenção do meio ambiente (pássaros, roedores, plantas) e até mesmo erro de comunicação entre os sensores e a unidade de memória central (*datalogger*). A identificação dessas lacunas nem sempre é feita de forma rápida, o que acarreta em incertezas na confiabilidade dos dados e pode até comprometer todo o conjunto de medições (ABREU,2000).

A utilização de um método para o tratamento dessas lacunas se faz necessário. Atualmente existem diversos métodos estatísticos e matemáticos para o tratamento e refinamento das lacunas de radiação solar. O método mais utilizado, em larga escala, para o preenchimento dessas lacunas é o método de replicação (ABREU 2000).

O método de replicação de dados consiste em replicar o valor de radiação anterior a 24h do momento da falha, em caso de lacunas pontuais. Já para o caso de lacunas maiores, como dias ou semanas, a replicação é aplicada integralmente pelo período de falha. Esse método de preenchimento apresenta melhor desempenho em dados que apresentam natureza cíclica em sua dispersão ao longo do tempo.

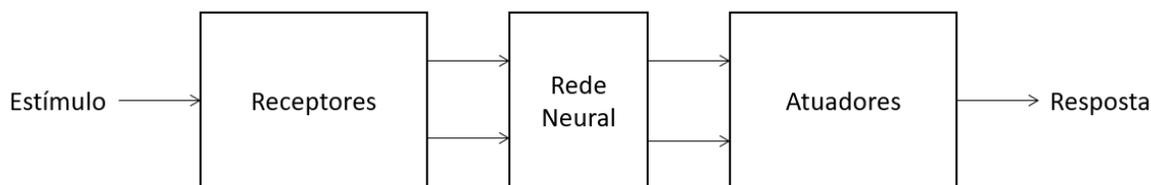
O método de replicação tem um desempenho satisfatório quando levado em consideração lacunas em dias ensolarados consecutivos. Esse desempenho é causado pela natureza dos dados de radiação solar, que apresentam um comportamento cíclico ao longo do dia e baixa variação em sua intensidade em intervalos curtos de comparação. Porém, o método não apresenta bom desempenho em casos de grande variação no padrão de radiação (dias chuvosos ou de rápida mudança climática) (ABREU, 2000). Além disso, ele apresenta um custo computacional muito baixo o que o torna muito atrativo para o tratamento das falhas de medição em estações meteorológicas.

2.4 Redes Neurais

As Rede Neurais Artificiais (RNAs) são modelos matemáticos inspirados no princípio de funcionamento do neurônio biológico. Uma rede neural é um processador maciçamente paralelamente distribuído constituído de unidades de processamento simples, que tem a capacidade natural para armazenar conhecimento experimental e torna-lo disponível para o uso. O cérebro é um computador (sistema de processamento de informação) altamente complexo, não-linear e paralelo. Ele tem a capacidade de organizar seus constituintes estruturais, conhecidos como neurônios, de forma a realizar certos processamentos (por exemplo, reconhecimento de padrões, percepção e controle motor) muito mais rapidamente que o mais rápido computador digital hoje existente (HAYKIN,2001; BRUMATTI, 2005). Uma rede neural é projetada para mimetizar essas características de processamento do cérebro biológico.

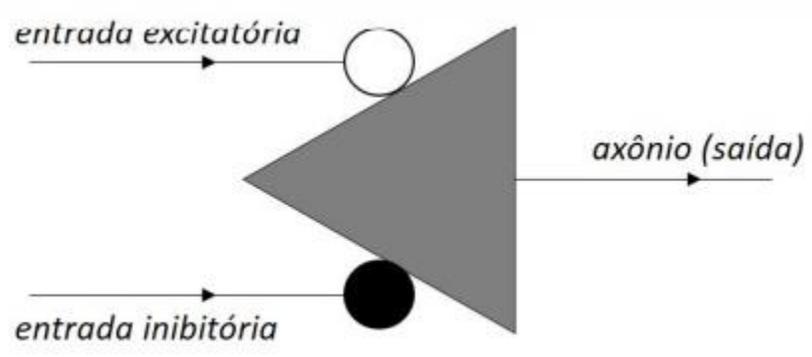
Em geral, uma rede neural é formada por um conjunto de neurônios (unidades de processamento) conectados por ligações. A força dessas ligações entre neurônios é conhecida como peso sináptico e são utilizados para o armazenamento do conhecimento adquirido (HAYKIN,2001). A solução de um problema com RNA's passa inicialmente por uma fase de aprendizagem, onde os pesos das conexões são ajustados até proporcionarem à rede a capacidade de representar o problema. O procedimento utilizado para realizar o processo de aprendizado é chamado de algoritmo de aprendizado, cuja função é modificar os pesos sinápticos da rede de uma forma ordenada para alcançar um objetivo de projeto desejado. Uma vez treinada, os pesos são fixados e a rede pode ser empregada como um modelo, estimando saídas a partir de um conjunto de dados de entrada (GUARNIERI, 2006).

Figura 6- Representação em diagrama de blocos do sistema nervoso (Fonte: HAYKIN,2001).



Os estudos das redes neurais começaram com o trabalho pioneiro de McCulloch e Pitts (1943). Em seu artigo clássico eles descrevem um cálculo lógico das redes neurais que unificava os estudos de neurofisiologia e da lógica matemática. Os autores também propuseram um neurônio artificial booleano ilustrado na Figura 7. Cerca de 15 anos após da publicação de McCulloch e Pitts, uma nova abordagem para o problema de reconhecimento de padrões foi introduzida por Rosenblatt (1958) em seu trabalho sobre *perceptron*, método inovador de aprendizagem supervisionada (ROSENBLATT, 1958). Em 1986 Rumellart, Hinton e Willians desenvolveram o algoritmo de retropropagação (*back-propagation*).

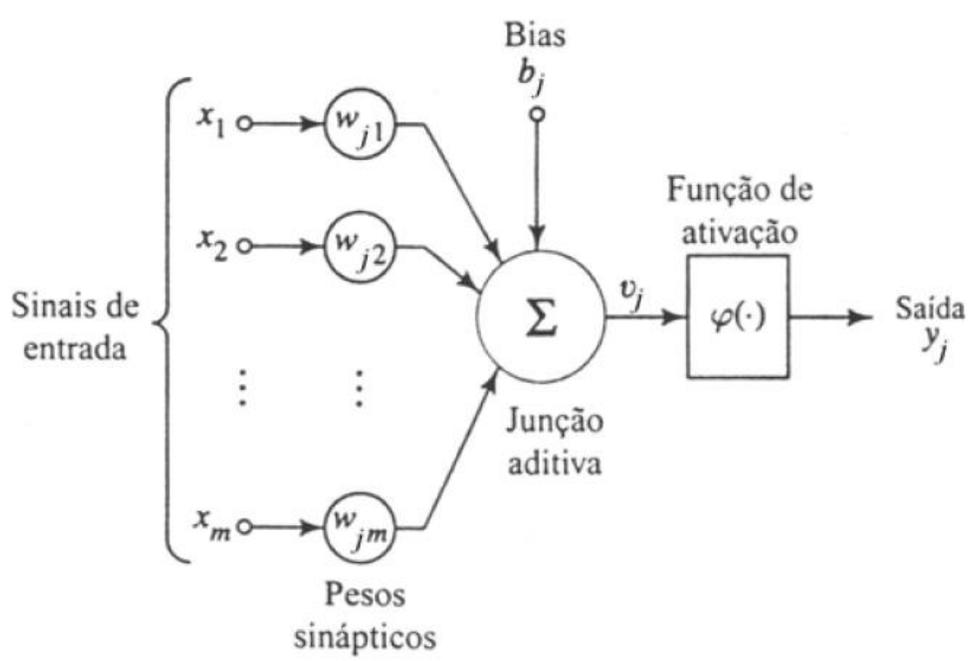
Figura 7 - Neurônio Booleano de McCulloch (Fonte: KOVÁCS, 2002).



2.4.1 Modelos de um neurônio

Neurônios são elementos processadores que desempenham um papel fundamental para a operação de uma rede neural. Os conjuntos de neurônios podem ser dispostos em uma ou mais camadas e estão conectados por ligações com diferentes pesos sinápticos (VELLASCO, 2007). A Figura 8 apresenta um modelo de neurônio, que forma a base para o projeto de redes neurais artificiais.

Figura 8 - Modelo de neurônio (Fonte: HAYKIN, 2001).



Os valores de entrada de um neurônio (x_i) são ponderados por valores associados com cada sinapse (w_{ji}), chamados pesos sinápticos. Todos os valores ponderados são somados, juntamente com um valor chamado bias (b_j). O bias tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se ele é positivo ou negativa, respectivamente. Esta soma é o nível de atividade do neurônio (v_j). A saída do neurônio é finalmente computada pela função de ativação ($\varphi(v_j)$), geralmente uma função linear ou tangente hiperbólica. O uso de uma função não-linear, como a função tangente-hiperbólica, permite às RNAs aprender comportamentos não-lineares e padrões complexos (HAYKIN, 2001)

Em termos matemáticos pode-se descrever o neurônio pelas equações 1 e 2.

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (1)$$

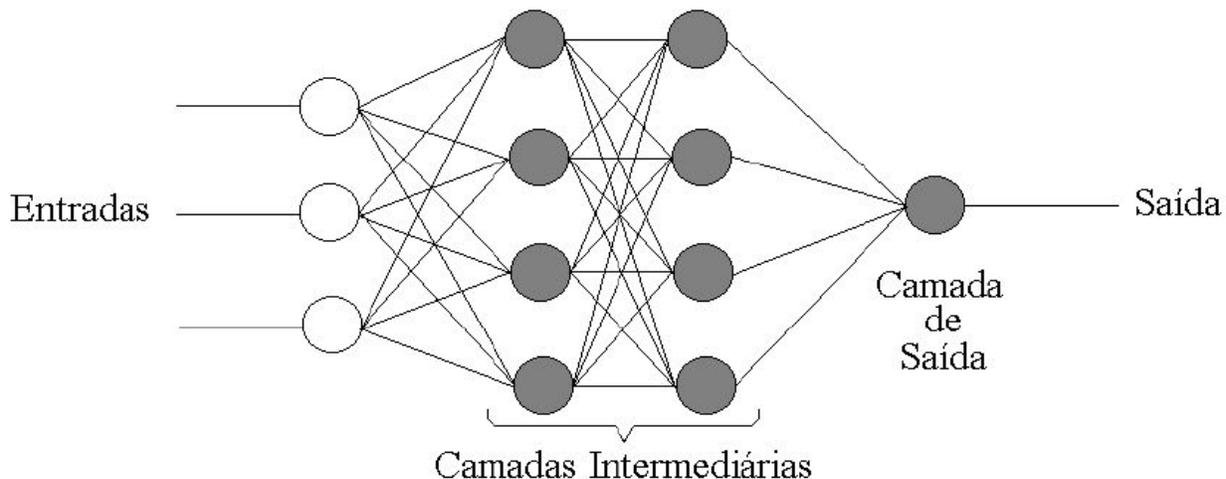
$$y_k = \varphi(u_k + b_k) \quad (2)$$

Onde x_1, x_2, \dots, x_m são os sinais de entrada; w_1, w_2, \dots, w_m são os pesos sinápticos do neurônio k ; u_k é a saída do combinador linear devidos aos sinais de entrada; b_k é o bias; $\varphi(\cdot)$ é a função de ativação; e y_k é o sinal de saída do neurônio (HAYKIN, 2001).

2.4.2 Arquitetura de redes neurais

Em uma rede neural os neurônios podem estar dispostos em uma ou mais camadas. Quando duas ou mais camadas são utilizadas, um neurônio pode receber em seus terminais de entrada valores de saída de neurônios da camada anterior e/ou enviar seu valor de saída para terminais de entrada de neurônios da camada seguinte (FACELI et al, 2011). A Figura 9 ilustra uma RNA com múltiplas camadas.

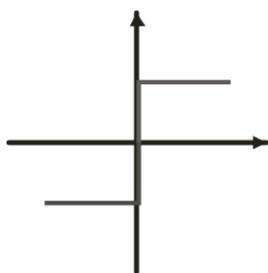
Figura 9 - Exemplo de rede neural de múltiplas camadas (GUARNIERI, 2006).



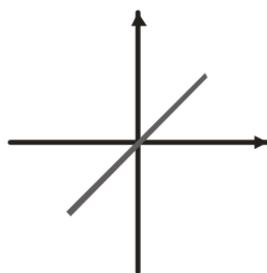
Após os somatórios dos sinais de entrada com seus respectivos pesos sinápticos a saída do neurônio é definida por meio da aplicação de uma função de ativação. As funções de ativação mais comuns são: funções de limiar, função linear e função sigmoial. A Figura 10 ilustra as funções de ativação. É a função de ativação que tem o papel de restringir a amplitude do sinal de saída do neurônio (HAYKIN, 2001).

Figura 10 - Exemplos de funções de ativação (Fonte: FACELI, 2011).

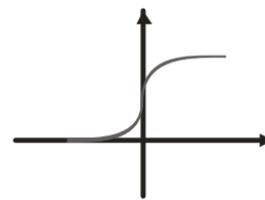
Degrau (ou limiar):



Linear:



Sigmoial:



2.4.3 Aprendizado e treinamento

Na fase de aprendizado e treinamento de uma rede neural, uma das principais características topológicas da rede é a escolha do algoritmo de retropropagação. A rede neural

pode ser classificada em relação a direção de propagação do sinal como *feedforward* (Figura 11) ou *feedback* (Figura 12). Em uma rede neural *feedforward* o sinal é propagado apenas unidirecionalmente da entrada para a saída. Já em uma rede *feedback* o sinal é propagado da entrada até a saída e repropagado para a entrada dos neurônios novamente (FACELI,2011).

Figura 11 - Rede neural *Feedforward* (Fonte: FACELI, 2011).

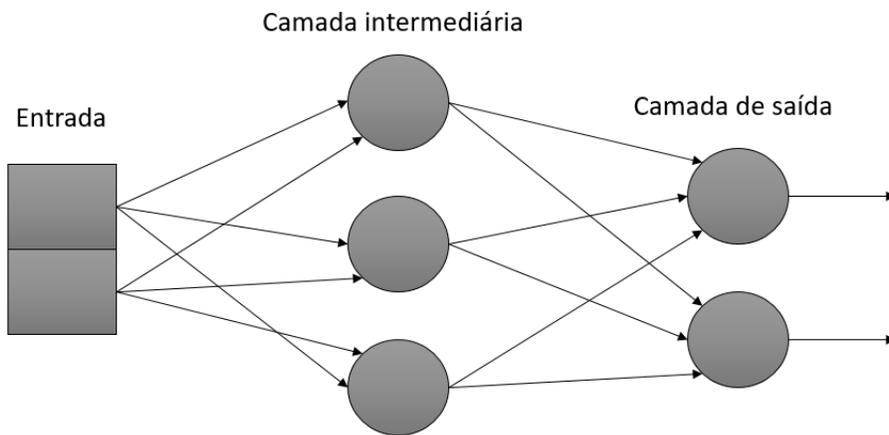
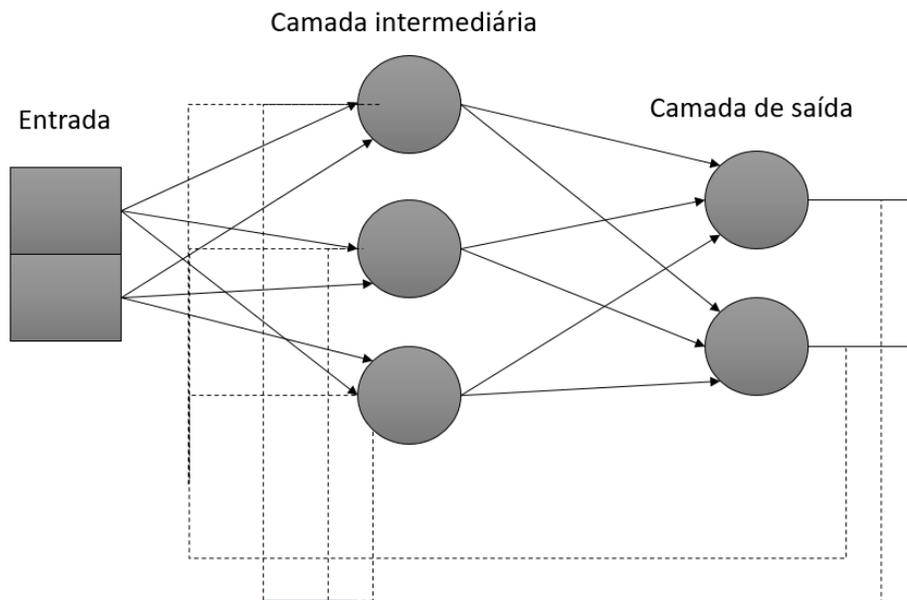


Figura 12 - Rede Neural *feedback* (Fonte: FACELI, 2011).



O aprendizado da rede neural é a atualização dos pesos sinápticos em função de um algoritmo de aprendizagem. O grau de participação do supervisor no processo de aprendizado classifica o método de aprendizado em: aprendizado supervisionado, aprendizado não supervisionado e aprendizado por reforço (NASCIMENTO,2000).

No aprendizado supervisionado, a rede neural é treinada partindo de um conjunto de dados de entrada e seus respectivos padrões de saída. Os pesos sinápticos são então reajustados conforme o valor produzido pela rede neural e o valor fornecido pelo supervisor. Este aprendizado tem como objetivo a minimização do valor de erro quadraticos da saída (FACELI, 2011).

No aprendizado por reforço o supervisor fornece uma avaliação grosseira da saída da rede, por exemplo: certo ou errado, sucesso ou fracasso, erro grande ou erro pequeno. Nessa categoria estão os algoritmos de aprendizado por reforço ou de punição e recompensa (NASCIMENTO, 2000).

Por fim, o aprendizado não supervisionado é utilizado normalmente quando não há uma série histórica de dados com entradas e saídas correspondentes. Nesse aprendizado o supervisor não possui as responsabilidades de oferecer os dados de saída para comparação, bem como, de avaliar o desempenho da rede neural. Os algoritmos utilizados frequentemente no aprendizado não supervisionado são o Hebbiano e o competitivo. O algoritmo Hebbiano é baseado na regra de Hebb, que diz que, se dois neurônios estão simultaneamente ativos, a conexão entre eles deve ser reforçada. O algoritmo competitivo promove uma competição entre os neurônios para determinar qual ou quais devem ter seus pesos reajustados. (FACELI, 2011)

3 MATERIAIS E MÉTODOS

Neste capítulo será explanado a metodologia utilizada para o tratamento dos dados de radiação, treinamento da rede neural artificial e o algoritmo de comparação dos resultados com o método de referência.

3.1 Tratamento dos dados de Radiação

Os dados de radiação foram coletados da plataforma do Instituto Nacional de Meteorologia (INMET) para a região de Araranguá no período de 2015 a 2016. Os dados foram fornecidos em

base em kJ/m^2 , com intervalos de medição de uma hora. A tabela de dados foi então ordenada de maneira linear agrupando os dados mensais de janeiro a dezembro de forma a obter uma matriz de 8760 dados de radiação, representando cada uma das 24 horas dos 365 dias do ano.

Para o treinamento da rede neural foi necessário um processo de tratamento nos dados de radiação já ordenados. O tratamento consistiu na retirada dos valores de radiação negativos, caracterizados pela falta de radiação solar no momento da medição. Esses valores de radiação representam os dados horários de períodos noturnos, ou seja, dados que não teriam valor significativo para o treinamento da rede neural, vide que o objetivo da rede é o preenchimento de lacunas de radiação validas.

A tabela com os dados de radiação validos foi então dividida em duas tabelas para o treinamento da rede neural. A tabela de *input* que consiste nos valores de entrada da rede e a tabela *output* que consiste com os valores de radiação para cada hora do ano. A tabela *input* foi organizada de modo a ter três variáveis para cada posição da tabela representando a hora, o dia e o mês. A Figura 12 mostra um esquemático de como foi organizada a tabelas *input*.

Figura 13 - Tabela input.



3.2 Topologia da RNA

O número de camadas, número de neurônios em cada camada, grau de conectividade e a presença ou não de conexões de retropropagação definem a topologia de uma rede neural artificial (FACELI, 2011). A definição da arquitetura da rede neural é uma etapa muito importante, pois, afeta diretamente a capacidade de processamento da rede. Atualmente, não é possível determinar teoricamente o número de camadas ou quantidade de neurônios que são necessários para resolução

de cada problema com redes neurais. É necessária cautela para que a rede não sofra com números de neurônios/camadas excessivos, causando *overfitting*, ou fique carente de unidades de processamento causando o *underfitting* da rede (HAYKIN, 2001). Além disso, é preciso analisar qual função de ativação e função de treinamento se ajustam melhor ao comportamento dos dados.

Portanto, para determinar o número de neurônios e camadas necessários foram utilizadas diversas configurações de redes neurais, contendo as mesmas tabelas de treinamento. Buscou-se encontrar a melhor configuração através da análise do erro quadrático. As redes foram testadas utilizando a ferramenta de redes neurais artificiais (*toolbox “nntool”*) do MATLAB (*Matrix Laboratory*), *software* de alto desempenho para modelagem e cálculo numérico. As topologias testadas podem ser observadas na Tabela 3.

Tabela 3 - Topologias da RNA.

Parâmetro	Configuração
Número de neurônios	1 a 500
Número de camadas ocultas	1 a 2
Função de ativação	‘Logsig’ e ‘Tansig’ para as camadas intermediárias, ‘Purelin’ para a camada de saída.
Função de treinamento	‘Trainbfg’, ‘Trainrp’, ‘Trainlm’, ‘Trainscg’, ‘Traincgb’, ‘Traincgf’, ‘Traincgp’, ‘Traincgp’, ‘Trainoss’, ‘Traingdx’ e ‘Traingd’

3.3 Teste e comparação

Os testes de desempenho da RNA foram realizados com base na simulação de lacunas de radiação solar. As lacunas foram aleatoriamente simuladas utilizando outra base de dados para o mesmo local de treinamento da rede neural artificial. A base de dados utilizada no processo de teste e comparação foi a estação meteorológica automática de superfície do Núcleo Tecnológico de Engenharia Elétrica (NTEEL) da Universidade Federal de Santa Catarina, Centro Araranguá.

Os dados da estação do NTEEL sofreram o mesmo processo de tratamento que os dados de treinamento da rede, buscando identificar e retirar os valores não válidos de radiação. Após o

tratamento dos dados foram criadas 500 lacunas horarias entre os meses de outubro de 2016 e abril de 2017. Essas lacunas foram então formatadas para o padrão de entrada da rede neural artificial (Figura 12) e testadas.

O método de referência utilizado foi o método da replicação de valores. Para cada lacuna que foi criada na base de dados foram inseridos valores de radiação validos do dia anterior a lacuna, no mesmo horário. Por exemplo, se uma lacuna foi inserida no dia 20 de novembro as 15:00h o método de referência replicava o valor de radiação do dia 19 de novembro as 15:00h para uma tabela de comparação.

A RNA e o método de replicação foram então comparados com o valor de referência da base de dados NTEEL. A análise de performance foi feita a partir dos valores de erro quadrático médio, frequência de erro e amplitude do erro.

4 RESULTADOS E DISCUSSÕES

Primeiramente, o teste de topologia para definição do número de camadas, número de neurônios e funções de ativação foi realizado metodologicamente com as configurações descritas na Tabela 3. Os testes de configurações podem ser visualizados no Anexo A. O critério de escolha da rede neural foi a regressão linear dos dados de treinamento, validação e teste.

As Figuras 14 e 15 e a Tabela 4 apresentam a melhor configuração obtida.

Tabela 4 - Melhor configuração da obtida.

Parâmetro	Configuração
Número de neurônios	15 e 45
Número de camadas ocultas	2
Função de ativação	Tansig, Tansig e Purelin
Função de treinamento	Trainlm

Figura 14- Topologia da Rede Neural.

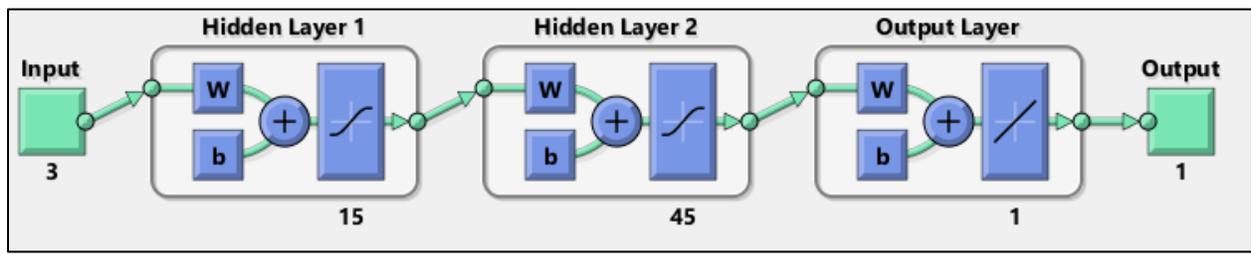
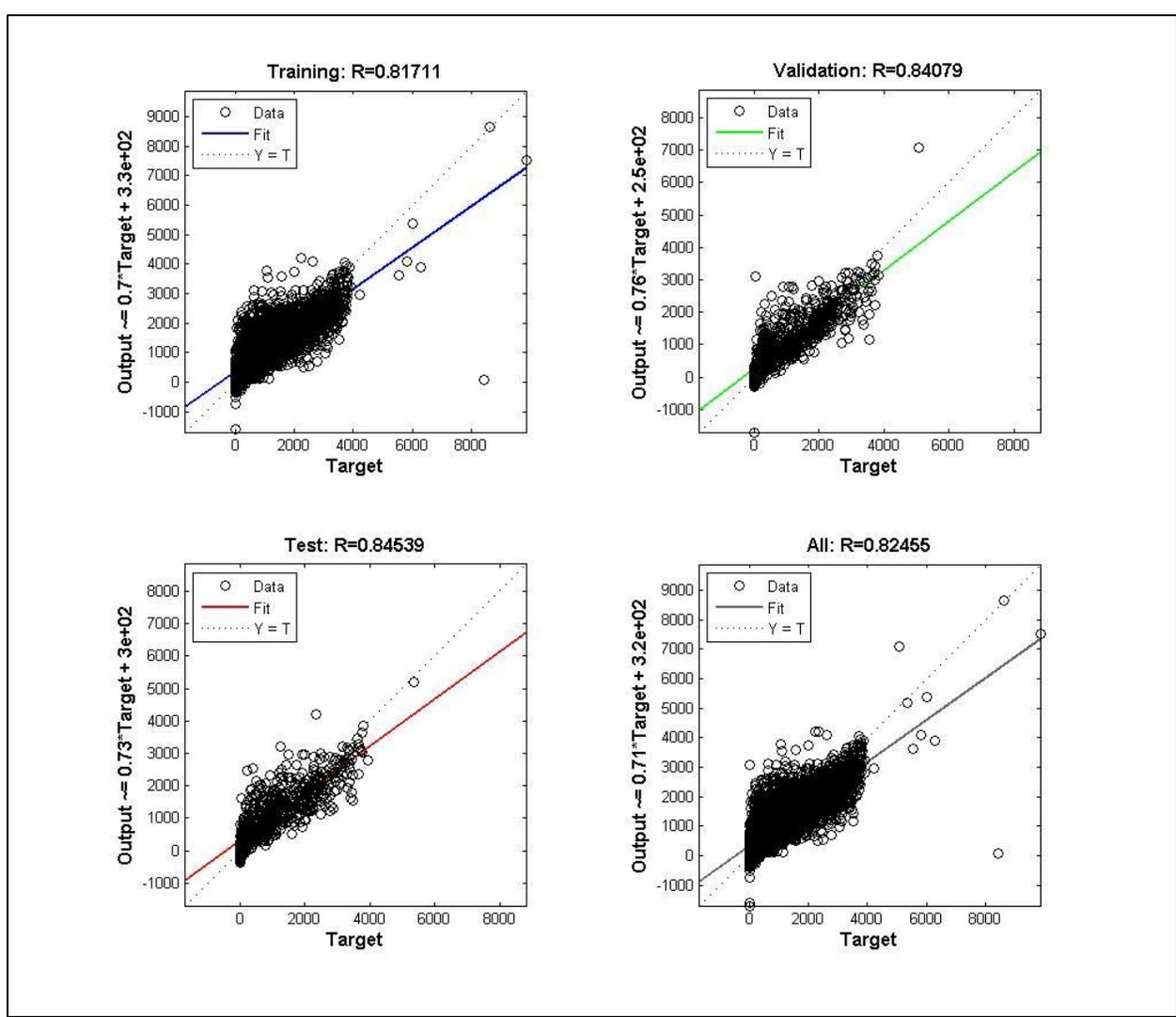


Figura 15 - Regressão Linear.



Os testes foram realizados a partir da metodologia descrita no capítulo anterior, onde foram criadas 10 populações de 500 lacunas cada para avaliar o desempenho da rede neural artificial utilizando o método de replicação de dados como base comparativa. Os dados foram comparados utilizando outra base de dados para evitar viciar a rede com dados que já foram utilizados para o treinamento da mesma.

O parâmetro utilizado para comparação de desempenho foi o erro quadrático médio. O erro quadrático médio foi escolhido como critério de comparação de desempenho por representar melhor a dispersão dos dados, avaliando qual a distância entre o valor obtido pelos métodos e o valor real, ao quadrado. A Equação 3 apresenta a fórmula utilizada para o cálculo do erro quadrático médio (EQM).

$$EQM = \frac{\sum_{i=1}^n (\hat{\theta} - \theta)^2}{N} \quad (3)$$

Onde N é o número de dados da população de teste; $\hat{\theta}$ é o valor encontrado pelo método de preenchimento de lacunas (replicação ou rede neural artificial); θ é o valor real obtido pela base de dados NTEEL.

Além da média, o desvio padrão médio também foi utilizado para quantificar a dispersão do conjunto de dados de erro. A Equação 4 apresenta a formulação matemática utilizada para calcular o desvio padrão (DPM).

$$DPM = \sqrt{\frac{\sum_{i=1}^n (x - \bar{x})^2}{N}} \quad (4)$$

Onde x é o valor de erro obtido; \bar{x} é a média do erro de todas as lacunas de teste e N é número de lacunas testadas.

A Tabela 5 sintetiza os valores resultantes da comparação entre o método de replicação de dados e a rede neural artificial para as 10 populações de teste. Todos os valores de erro estão na unidade kJ/m² que é a unidade de radiação solar utilizada pela base de dados de treinamento e validação.

Tabela 5 - Erro médio quadrático e desvio padrão médio das 10 populações de teste, em kJ/m².

População	Método	Erro médio quadrático (kJ/m²)	Desvio padrão médio (kJ/m²)
1	Replicação	224,31	219,71
	RNA	187,16	145,89
2	Replicação	270,26	252,55
	RNA	195,50	148,41
3	Replicação	263,79	251,04
	RNA	200,98	149,70
4	Replicação	262,91	258,24
	RNA	202,53	161,14
5	Replicação	254,92	244,08
	RNA	194,09	154,35
6	Replicação	255,89	248,03
	RNA	194,61	143,23
7	Replicação	278,64	264,08
	RNA	198,72	154,62
8	Replicação	234,79	229,13
	RNA	175,52	138,76
9	Replicação	273,11	253,51
	RNA	201,80	151,98
10	Replicação	272,86	244,66
	RNA	203,59	155,50

Para nível de comparação, a amplitude da radiação da base de dados varia entre 0 a 4000 kJ/m² ou 0 a 1111 Wh/m². A Tabela 6 apresenta o desempenho geral dos métodos de preenchimento e as Figuras 16 a 25 apresentam os diagramas *blox plot* da dispersão dos dados.

Tabela 6 - Comparação de desempenho entre a RNA e o método de replicação.

Método	Erro médio Quadrático (kJ/m²)	Desvio Padrão Médio (kJ/m²)
Replicação	259,15	246,92
RNA	195,45	150,45

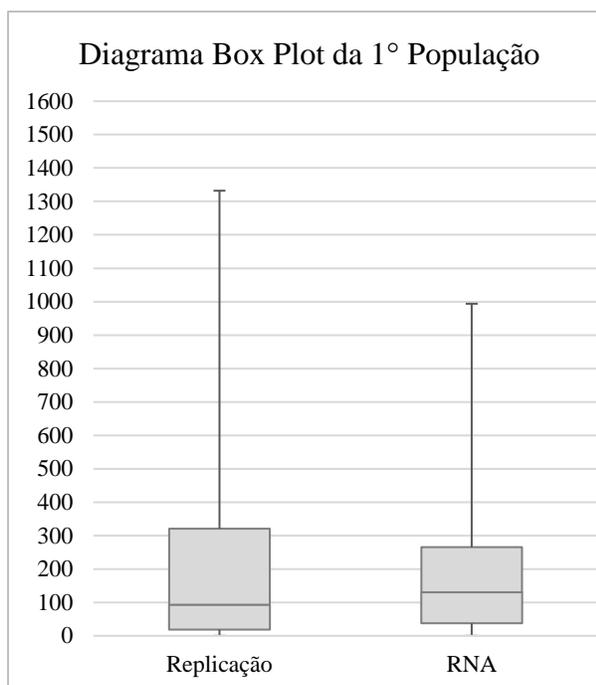
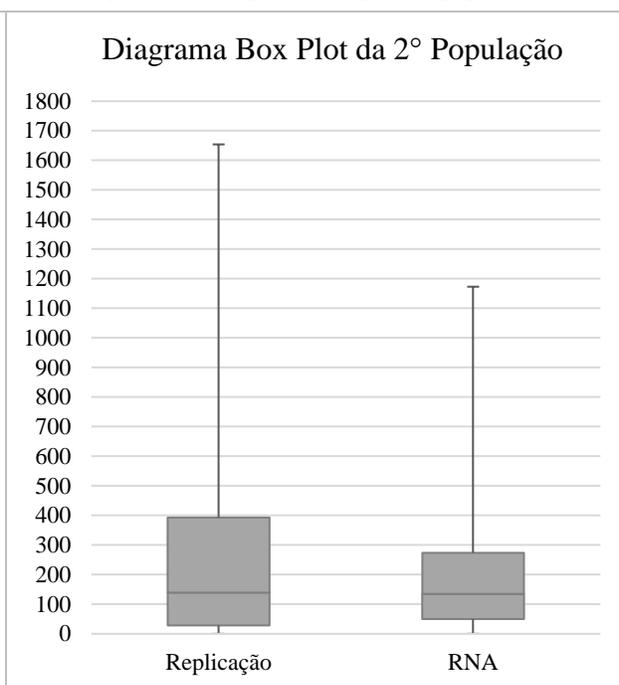
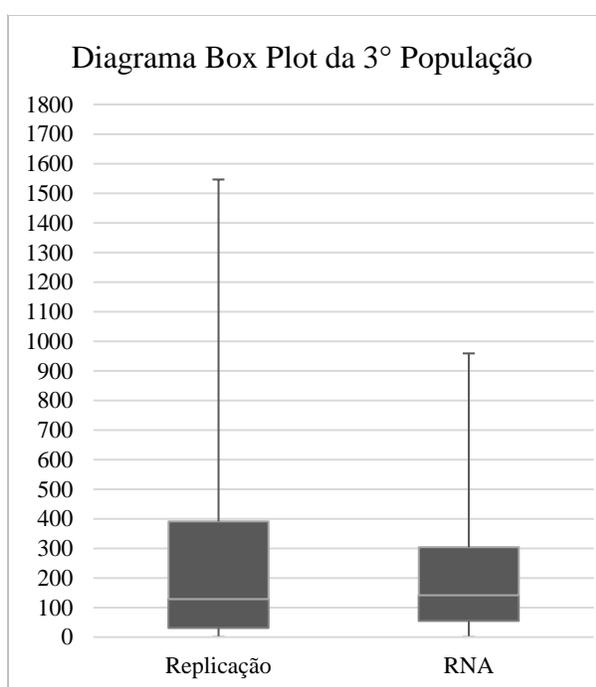
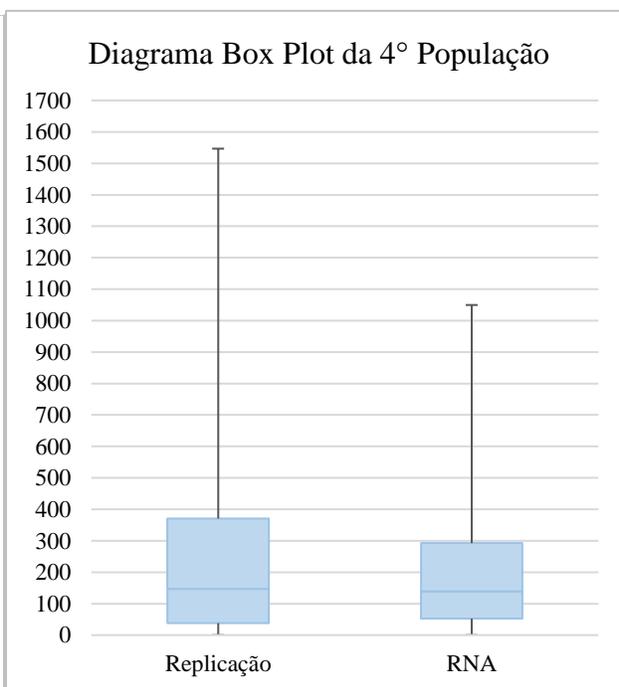
Figura 16 - Diagrama *box plot* 1º população.Figura 17 - Diagrama *box plot* 2º população.Figura 18 - Diagrama *box plot* 3ª população.Figura 19 - Diagrama *box plot* 4ª população.

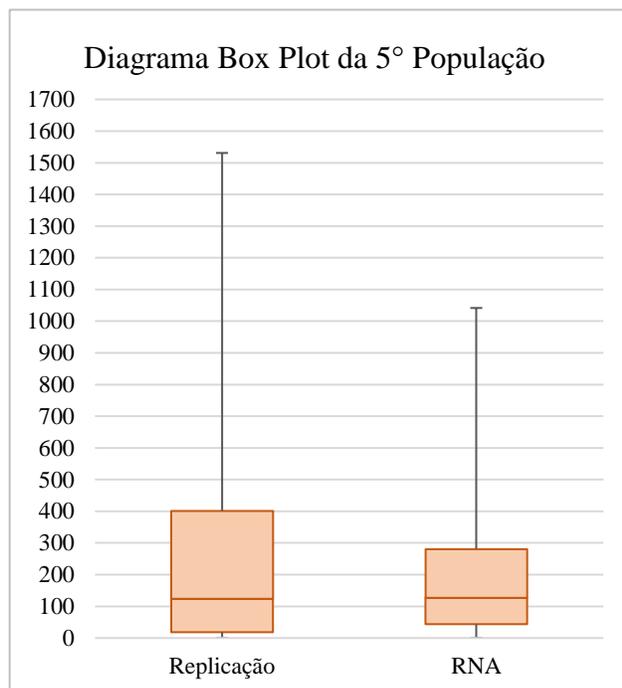
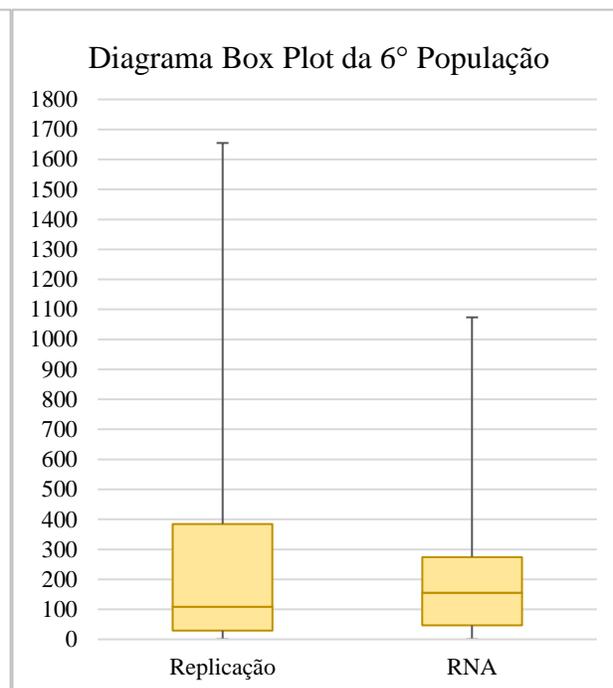
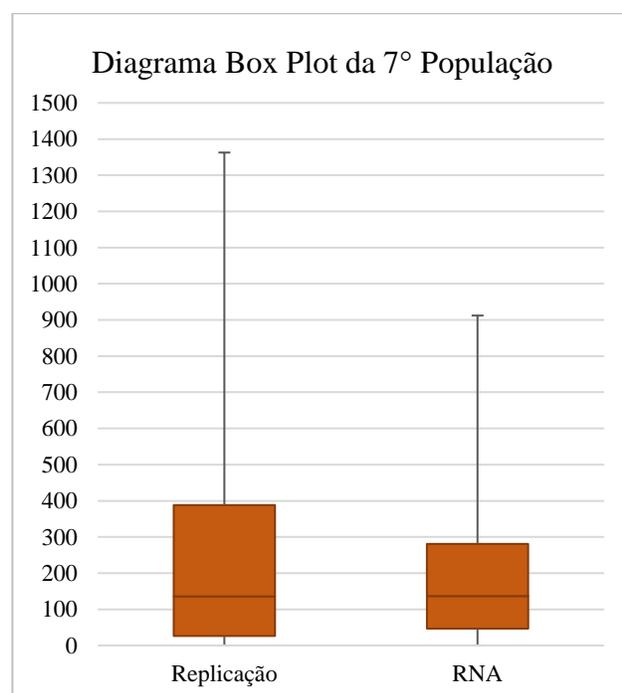
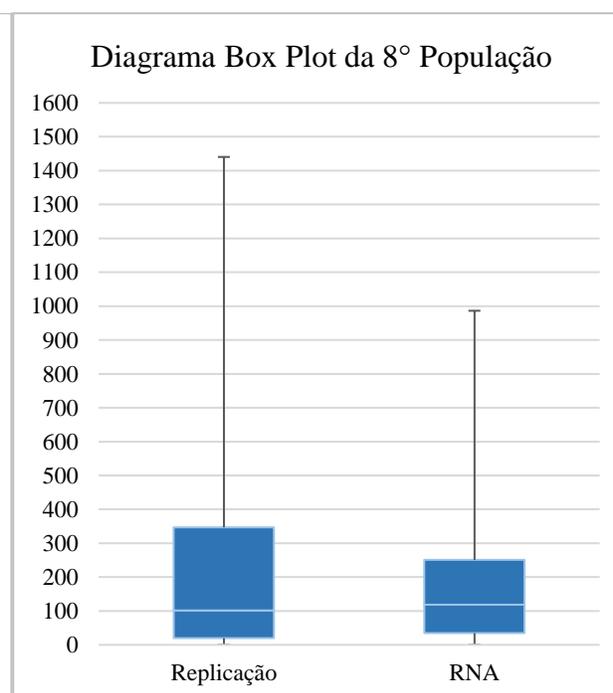
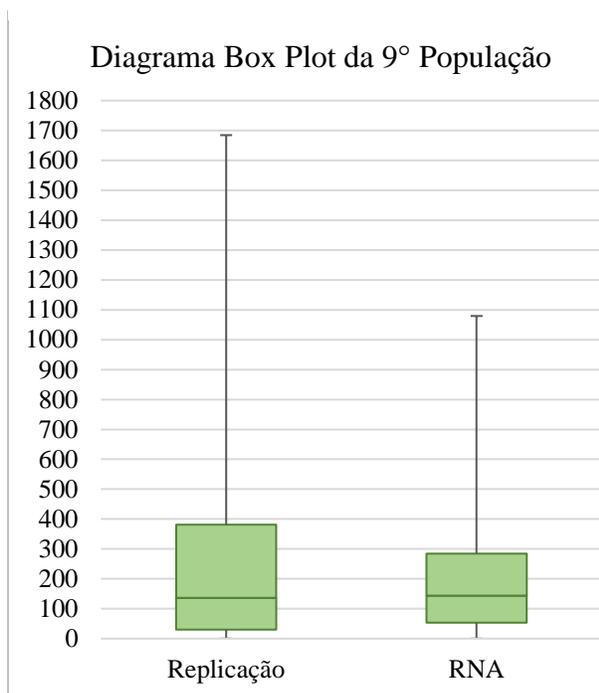
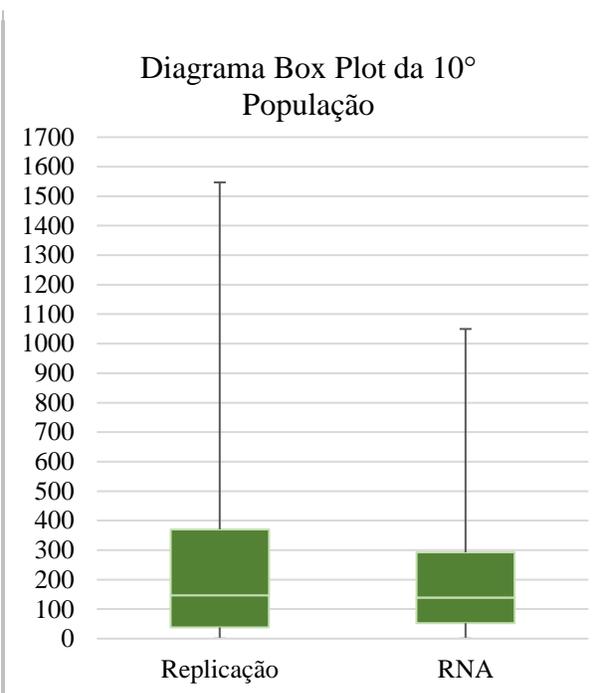
Figura 20- Diagrama *box plot* 5º população.Figura 21- Diagrama *box plot* 6º população.Figura 22 - Diagrama *box plot* 7º população.Figura 23 - Diagrama *box plot* 8º população.

Figura 24- Diagrama *box plot* 9º população.Figura 25- Diagrama *box plot* 10º população.

O diagrama *box plot* é um gráfico de estatística descritiva utilizado para demonstrar visualmente a dispersão de um conjunto de dados. Analisando os diagramas é possível observar que o método de replicação apresenta valor máximo de erro maior em todas as populações testadas. As áreas destacadas nos diagramas representam as regiões interquartil. Nelas estão contidos 50% de todos os valores de erro observados, concentrados na tendência central dos valores, ou seja, eliminando os 25% menores valores e 25% maiores valores.

Em todas as populações a região interquartil da RNA teve menor área quando comparada a região interquartil do método de replicação. Isso caracteriza uma distribuição mais concentrada e precisa por parte da RNA. Porém, analisando o erro mínimo de cada população o método de replicação conseguiu atingir valores menores, devido principalmente a capacidade do método de replicação de inferir valores mais exatos em períodos de baixa variabilidade de radiação solar.

A rede neural artificial mostrou menor média de erros em comparação com o método de replicação, apresentando um desempenho superior da ordem de 24,6 %. Além disso, a RNA apresentou menor dispersão dos dados o que garante maior precisão na inferência de valores.

Os testes mostraram que o método de replicação apresenta vantagem sobre a rede neural artificial, quando comparado em períodos onde a radiação solar não sofre tanta variação. Normalmente períodos de dias ensolarados. Isso acontece devido o padrão de radiação solar apresentar comportamento similar durante uma sequência de dias ensolarados, ou seja, espera-se que dois dias ensolarados consecutivos tenham um padrão de radiação solar similar. Este fator garante um erro médio quadrático menor para o método de replicação, quando comparado em lacunas ensolaradas consecutivas.

Entretanto, a rede neural artificial não deixa de inferir valores com erros aceitáveis durante períodos ensolarados. Isso garante que a variação do erro médio da rede neural artificial seja menor que a do método de replicação. Além disso, a rede neural apresenta melhor performance durante períodos onde há variabilidade na radiação solar, como por exemplo dias parcialmente nublados.

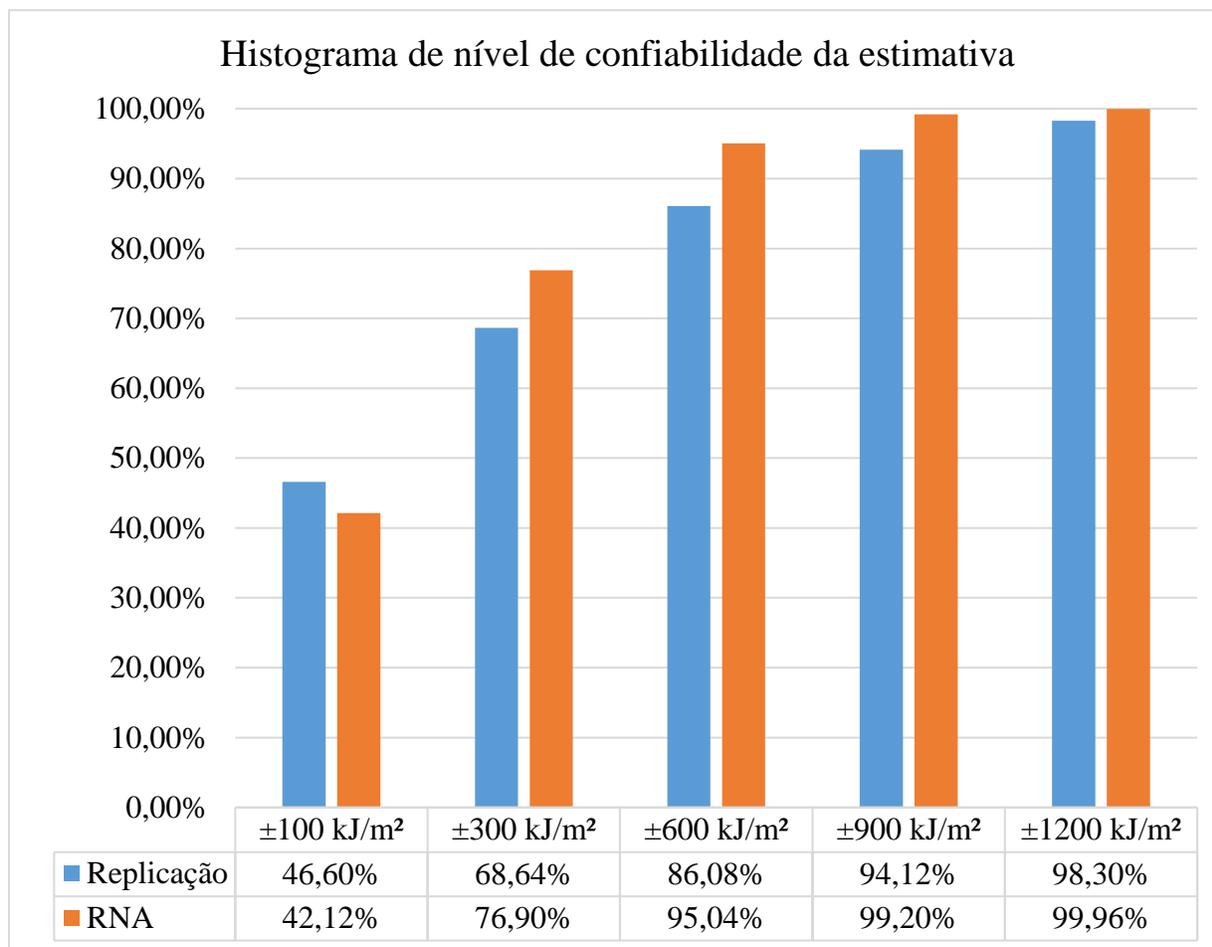
Portanto, apesar de não inserir o valor mais exato nos períodos ensolarados a rede neural infere valores mais precisos em todas as lacunas de dados testadas. Este comportamento é esperado, devido a capacidade da rede de se adaptar ao comportamento caótico da radiação solar.

A Figura 26 apresenta o histograma do nível de confiança da estimativa da radiação solar (probabilidade de acerto na estimativa do parâmetro) nos dois métodos utilizados. Os dados de radiação solar estão compreendidos entre 0 e 4000 kJ/m². A probabilidade do valor da radiação solar estar no intervalo de estimativa do parâmetro, pode ser mensurado pelo nível de confiança da estimativa.

A rede neural artificial apresenta estimativas para as lacunas de radiação solar com um erro padrão de até ± 600 kJ/m² com 95 % de probabilidade. Este intervalo apresenta variações de até 16,7% no parâmetro estimado. Em contraponto, o método de replicação apresenta estimativas para as lacunas de radiação solar com um erro padrão de até ± 600 kJ/m² com 86 % de probabilidade. Portanto, a rede neural artificial apresenta maior precisão na inferência de dados de radiação solar.

O nível de confiança para erros padrão da ordem de ± 100 kJ/m², sendo que este intervalo apresenta variações de até 2,7% no parâmetro estimado, também é apresentada na Figura 26. O método de replicação apresenta desempenho superior ao da RNA inferindo valores dentro do intervalo com probabilidade da ordem de 46,5 % enquanto que a probabilidade do intervalo estimado pela RNA conter o parâmetro é da ordem de 42 %.

Figura 18 - Histograma de nível de confiança da estimativa.



Do ponto de vista energético, considerando que os dois métodos fossem utilizados para tratar dados de radiação solar, cuja função seria a utilização em leilões de energia. O método de preenchimento de lacunas utilizando redes neurais artificiais apresentará uma estimativa de dados de radiação mais precisa. Por consequência, a curva de geração de energia esperada estará mais realista utilizando a RNA para o tratamento dos dados. Outro ponto que deve ser mencionado é o custo computacional. Uma das principais vantagens da rede neural artificial é que depois de treinada a utilização da mesma apresenta custos computacionais baixíssimos. A partir da tabela de pesos sinápticos adquirida basta uma simples rotina de somatórios e funções de ativações para que a rede neural seja empregada de forma rápida e sem a necessidade de mais treinamentos.

CONCLUSÃO

Este trabalho apresentou o desenvolvimento, treinamento e validação de uma rede neural artificial para processamento de dados de radiação solar. A partir dos testes de 10 populações de 500 lacunas nos dados de radiação solar foi possível comparar o desempenho da rede neural artificial com o método de replicação de dados. A rede neural artificial foi previamente treinada baseada em uma topologia adquirida em testes e utilizando dados de radiação solar horária para a cidade de Araranguá, Santa Catarina.

Os critérios de avaliação de desempenho foram os erros médios quadráticos, a dispersão dos erros e as frequências de erro. Os testes de validação comprovaram que o desempenho da rede neural artificial foi superior cerca de 24,6 % ao método de replicação. Analisando a dispersão dos erros foi possível constatar que a rede neural artificial garante maior precisão no preenchimento das lacunas de dados. Essa precisão na inferência de dados de radiação solar é de suma importância do ponto de vista energético, uma vez que garante uma estimativa de geração de energia elétrica mais confiável, viabilizando empreendimentos e garantindo espaço na matriz energética para a geração solar de energia.

REFERÊNCIA

ABREU, SAMUEL L. DE. **Qualificação e recuperação de dados de radiação solar medidos em Florianópolis - SC**. 2000. 37 f. TCC (Graduação) - Curso de Engenharia Elétrica, Universidade Federal de Santa Catarina, Florianópolis, 2000.

BRUMATTI, M. **Redes neurais artificiais**. Vitória, Espírito Santo, ca, 2005.

CENTRO DE PESQUISA DE ENERGIA ELÉTRICA. **Manual de Engenharia para Sistemas Fotovoltaicos**. Rio de Janeiro: CEPEL - CRESESB, 2014. 530 p.

COLLE, S.; PEREIRA, E. B. **Atlas de irradiação solar do Brasil** – 1ª. versão para irradiação global derivada de satélite e validada na superfície. Brasília: Instituto Nacional de Meteorologia, 1998.

EMPRESA DE PESQUISA ENERGÉTICA. **Balço Energético Nacional 2016**: Ano base 2015. Rio de Janeiro: EPE, 2016. 292 p.

EMPRESA DE PESQUISA ENERGÉTICA. **Expansão da Geração**: 2º Leilão de Energia de Reserva de 2015. Rio de Janeiro: EPE, 2016. 28 p.

FACELI, Katti, et al. **Inteligência Artificial**: Uma abordagem de aprendizado de máquina. *Rio de Janeiro: LTC 2* (2011): 192.

INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS. **Atlas Brasileiro de Energia Solar**. São José dos Campos: INPE, 2006. 64 p.

INSTITUTO NACIONAL DE METEOROLOGIA. NOTA TECNICA NÚMERO 001/2011/SEGER/LAIME/CSC/INMET: **Rede de Estações Meteorológicas Automáticas do INMET**, 2011. 11 p.

MARTINS, FERNANDO RAMOS et al. **Projeto sonda – rede nacional de estações para coleta de dados meteorológicos aplicados ao setor de energia**. Fortaleza: Cbens, 2007. p. 1 - 9.

BUHLER, A. J. ; RAMPINELLI, GIULIANO A. ; GASPARIN, FABIANO P. ; KRENZINGER, A. . **Influência do Fator de Correção Espectral na Caracterização de Módulos Fotovoltaicos de Diferentes Tecnologias**. Revista Brasileira de Energias Renováveis, v. 4, p. 38-51, 2015.

HAYKIN, S. **Neural networks**: a comprehensive foundation. New York: Macmillan College Publishing Company, 1994.

HAYKIN, S. **Redes neurais**: princípios e prática, 2. ed. Porto Alegre: Bookman Companhia Editora, 2001. 900p.

KOVACS, Z. L. **Redes neurais artificiais**. Editora Livraria da Física, 2002.

NASCIMENTO, CAIRO L.; Yoneyama, Takashi. **Inteligência artificial em controle e automação**.1ª. São Paulo:Edgard Blücher, 2000

R.A. GUARNIERI, **Emprego de Redes Neurais Artificiais e Regressão Linear Múltipla no Refinamento das Previsões de Radiação Solar do Modelo Eta**. Dissertação de Mestrado, INPE, São José dos Campos (2006), 120 p.

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

PEREIRA, ANDRÉ B.; VRISMAN, AUGUSTO L.; GALVANI, EMERSON. **Estimativa da radiação solar global diária em função do potencial de energia solar na superfície do solo**. 2002.

VELLASCO, M. M. B. R. **Redes neurais artificiais**. PUC, Rio de Janeiro, Brasil, 2007.

ANEXO A

N° Camadas	N° Neurônios	N° de Épocas	R ²
1	10	71	0,73152
1	20	39	0,75174
1	30	204	0,76869
1	40	20	0,74386
1	50	46	0,76141
1	60	31	0,76739
1	70	23	0,75967
1	80	53	0,76335
1	90	13	0,75569
1	100	26	0,76797
1	150	16	0,77123
1	200	14	0,78183
1	250	13	0,76325
1	300	16	0,77072
1	400	22	0,66589
1	500	13	0,70658
2	5 e 5	56	0,73953
2	5 e 10	29	0,73205
2	5 e 15	27	0,74477
2	5 e 20	94	0,75952
2	10 e 15	136	0,80056
2	10 e 20	35	0,78614
2	10 e 25	69	0,7886
2	10 e 30	63	0,77673
2	10 e 40	59	0,77524
2	10 e 50	146	0,80419
2	10 e 60	27	0,77774
2	10 e 70	91	0,80432
2	10 e 80	62	0,77852
2	15 e 25	37	0,77672
2	15 e 30	19	0,75695
2	15 e 35	57	0,79124
2	15 e 40	71	0,80193
2	15 e 45	53	0,82455
2	15 e 50	48	0,80321
2	15 e 55	28	0,7731
2	15 e 60	16	0,7633
2	15 e 65	17	0,76271
2	15 e 70	63	0,75254

AGRADECIMENTOS

Gostaria de agradecer primeiramente a Deus, por ter me dado tantas oportunidades maravilhosas e por estar presente em mim e na minha vida desde sempre e para todo o sempre.

Quero agradecer a minha mãe Andreia, meu exemplo maior, a pessoa que mais admiro no mundo e a quem devo absolutamente tudo em minha vida. Obrigado mãe por ter me criado e educado com todo amor e carinho. Esse trabalho e essa graduação eu dedico a você. Te amo mãe.

Agradecer também aos meus amigos que estiveram presente comigo durante a faculdade me dando apoio sempre, não seria possível ter feito essa caminhada sem a ajuda de vocês. Desde as festas, até as horas de estudos juntos. Agradeço do fundo do meu coração por cada amigo que tenho. Em especial a minha amiga Simone por ter salvo minha graduação e por estar comigo em diversas boas memórias. Aos meus colegas de trabalho da ENEjr por terem enriquecido minha experiência na graduação com conhecimento e sonhos. E aos meus amigos Conrado, Gustavo, Norberto, Raul, Renan e Vitor. Também conhecidos como os mineradores do século 21.

Agradeço especialmente também, a minha namorada Carmem que me deu total suporte emocional e esteve comigo em todas as noites mal dormidas dedicadas e este trabalho, obrigado por tudo amor.

Quero agradecer Em especial ao Prof. Dr. Giuliano Arns Rampinelli por todo amparo e conhecimento disponibilizado em prol da realização deste trabalho de conclusão, além de toda dedicação e esforço para repassar da maneira mais correta e coerente seus ensinamentos. À Universidade Federal de Santa Catarina por toda vivência e conhecimento disponibilizados dentro e fora da sala de aula, propiciando grandes momentos. A todos os professores que contribuíram para o meu crescimento pessoal e profissional. A Banca de avaliação formada pelos professores Leonardo, Luciano e Giuliano (orientador) pela disponibilidade e ensinamentos durante a graduação.

Agradece a todos que disponibilizaram informações para que este trabalho pudesse acontecer. O INMET pela disponibilização dos dados meteorológicos. Ao Professor Anderson e ao Políbio pela ajuda com o desenvolvimento da rede neural artificial. E ao NTEEL pela disponibilização dos dados para validação da rede