



UTILIZANDO RESERVOIR COMPUTING PARA A PREVISÃO DA VELOCIDADE DO VENTO

Trabalho de Conclusão de Curso
Engenharia da Computação

Bruna Cavalcanti Galle de Aguiar

Orientador: Prof. Dr. Mêuser Jorge Silva Valença



Bruna Cavalcanti Galle de Aguiar

***UTILIZANDO RESERVOIR COMPUTING
PARA A PREVISÃO DA VELOCIDADE DO
VENTO***

Monografia apresentada como requisito parcial para obtenção do diploma de Bacharel em Engenharia de Computação pela Escola Politécnica de Pernambuco - Universidade de Pernambuco

Orientador:

Prof. Dr. Mêuser Jorge Silva Valença

UNIVERSIDADE DE PERNAMBUCO
ESCOLA POLITÉCNICA DE PERNAMBUCO
GRADUAÇÃO EM ENGENHARIA DE COMPUTAÇÃO

Recife - PE, Brasil

20 de novembro de 2012

MONOGRAFIA DE FINAL DE CURSO

Avaliação Final (para o presidente da banca)*

No dia 5 de 12 de 2012, às 9:00 horas, reuniu-se para deliberar a defesa da monografia de conclusão de curso do discente BRUNA CAVALCANTI GALLE DE AGUIAR, orientado pelo professor Mêuser Jorge Silva Valença, sob título Utilizando Reservoir Computing para a previsão da velocidade do vento, a banca composta pelos professores:

Sérgio Murilo Maciel Fernandes

Mêuser Jorge Silva Valença

Após a apresentação da monografia e discussão entre os membros da Banca, a mesma foi considerada:

Aprovada Aprovada com Restrições* Reprovada

e foi-lhe atribuída nota: 10 (Dez)

*(Obrigatório o preenchimento do campo abaixo com comentários para o autor)

O discente terá 08 dias para entrega da versão final da monografia a contar da data deste documento.

SÉRGIO MURILO MACIEL FERNANDES

MÊUSER JORGE SILVA VALENÇA

Dedico este trabalho aos meus pais, Cirano e Rosângela Galle, à minha irmã, Rebeca Galle, à minha avó, Elzani Cavalcanti, ao meu namorado, Breno Soares e à minha melhor amiga, Lara Dantas.

Agradecimentos

Agradeço, em primeiro lugar, aos meus pais, Cirano Galle de Aguiar e Rosângela Cavalcanti Galle de Aguiar por sempre me propiciarem as melhores condições e oportunidades na construção do caminho que me levou a minha formação acadêmica, assim como todo o apoio em cada etapa da minha vida. À minha irmã, Rebeca Cavalcanti Galle de Aguiar e minha avó, Elzani Albuquerque Cavalcanti, pela presença e força constantes.

Agradeço, também, ao meu namorado, Breno Soares dos Santos, por me acompanhar durante todos esses anos, pela paciência nos momentos mais difíceis, pela ajuda sempre que necessária e pelo amor e carinho que facilitaram essa longa jornada.

Agradeço à minha melhor amiga, Lara Dantas, aquela que desde o primeiro dia de aula nunca me deixou sozinha para enfrentar qualquer desafio que a faculdade ou a vida me lançaram e a responsável por nunca ter me deixado desistir quando essa parecia ser a única opção.

Agradeço ao meu melhor amigo, Murilo Raphael de Souza Lira, por esses mais de 10 anos de amizade e companheirismo e por ser um dos grandes colaboradores no longo caminho para chegar onde hoje estou.

Agradeço aos três amigos conquistados durante o curso, Breno Costa Inojosa, Breno Menezes e Carlos Castor, e por todos os dias em que rimos de uma situação ou nos amparávamos para enfrentar outra e todos os almoços no meio.

Agradeço às minhas amigas da escola por trazerem momentos alegres a cada encontro e à todas que residem em outros Estados brasileiros, mas que sempre se fizeram presente, acreditando em mim e encorajando-me a seguir adiante.

Agradeço a todos que foram meus professores, em especial ao meu orientador, o professor Dr. Mêuser Jorge Silva Valença e por todo o auxílio fundamental para o desenvolvimento e conclusão deste projeto.

Finalmente, agradeço a Deus por toda força que precisei nos momentos mais difíceis e pela oportunidade de manter e incluir na minha vida todas as pessoas citadas acima.

Resumo

Um dos principais desafios da atualidade é a crescente demanda de energia mundial e, a fim de suprir essa necessidade, as fontes de energia mais utilizadas são o petróleo, gás natural e carvão mineral. O grande problema com essas fontes deve-se ao fato de, além de serem extremamente poluidoras, elas são de origem não renovável. Por causa disso, energias renováveis estão se tornando cada vez mais essenciais para a humanidade e entre elas, está o vento e sua escolha é uma das mais promissoras. Os parques eólicos tem seu potencial diretamente ligado à velocidade do vento, sendo necessária boas estimativas dessa variável para a construção de estratégias e planejamentos eficientes. Entretanto, essa tarefa apresenta grandes dificuldades devido às complexas características do vento, como a alta variabilidade de sua velocidade e direção. Este trabalho tem como objetivo utilizar a técnica de *Reservoir Computing* para a previsão da velocidade do vento e, pelo fato da *Multi-Layer Perceptron* ser a mais utilizada para tal fim, realizar a comparação entre esses dois tipos de Redes Neurais Artificiais e analisar qual possui o melhor desempenho.

Palavras-chave: Energia Eólica, *Reservoir Computing*, MLP, Previsão da Velocidade do Vento.

Abstract

One of the main challenges today is the growing global energy demand and in order to meet this need, the most widely used energy sources are oil, natural gas and coal. The main problem with these sources is due to the fact that, besides being extremely polluting, they are non-renewable sources. Therefore, renewable sources are becoming essential for humanity and among many of them, there is the wind and its choice is the most promising. Wind farms have their potential directly related to the wind speed, which requires good estimates of this variable in order to build effective strategies and plans. However, this task presents great difficulties due to the complex characteristics of the wind, as the high variability of its velocity direction. This paper aims to use the technique of Reservoir Computing for the prediction of wind speed and compare its performance with the one produced by the Multi-Layer Perceptron, another type of artificial neural network and the most widely used for this purpose. At the end, it will be possible to conclude which one is better at predicting wind speed.

Keywords: Wind Energy, Reservoir Computing, MLP, Wind Speed Forecasting

Sumário

Lista de Figuras	p. ix
Lista de Tabelas	p. x
Lista de Abreviaturas e Siglas	p. xi
1 Introdução	p. 12
1.1 Motivação	p. 12
1.2 Objetivos	p. 13
1.2.1 Objetivo Geral	p. 13
1.2.2 Objetivos Específicos	p. 13
1.3 Estrutura da Monografia	p. 14
2 Fundamentação Teórica	p. 15
2.1 Energia Eólica	p. 15
2.1.1 Introdução	p. 15
2.1.2 Turbina Eólica	p. 16
2.1.3 Potencial Eólico Brasileiro	p. 20
2.2 Redes Neurais Artificiais	p. 21
2.3 <i>Multi-Layer Perceptron</i>	p. 23
2.3.1 Algoritmo <i>Backpropagation</i>	p. 24
2.3.2 Critério de Parada	p. 25
2.4 <i>Reservoir Computing</i>	p. 26

2.4.1	Introdução	p. 26
2.4.2	Criação e Uso da Técnica de RC	p. 27
2.4.2.1	Criação das entradas e das conexões do <i>reservoir</i>	p. 27
2.4.2.2	Simulando o <i>reservoir</i> e treinando e testando as saídas	p. 28
2.4.3	Aplicações de <i>Reservoir Computing</i>	p. 30
3	Metodologia	p. 31
3.1	Base de Dados	p. 31
3.2	Método de Seleção de Variáveis	p. 31
3.3	Pré-processamento dos Dados	p. 34
3.4	Predizendo Velocidade do Vento com MLP	p. 35
3.5	Predizendo a Velocidade do Vento com RC	p. 36
3.6	Testes Estatísticos	p. 40
3.6.1	Teste de Shapiro-Wilk	p. 40
3.6.2	Teste F	p. 41
3.6.3	Teste T-Student para amostras independentes	p. 42
3.6.4	Teste da Soma dos Postos de Wilcoxon	p. 43
4	Resultados	p. 44
5	Considerações Finais	p. 47
5.1	Conclusões	p. 47
5.2	Trabalhos Futuros	p. 47
	Referências	p. 49
	Apêndice A – Resultados dos Testes Estatísticos no Software R	p. 50
	Apêndice B – <i>Script</i> dos Testes Estatísticos	p. 52

Lista de Figuras

1	Turbina Eólica de Eixo Vertical	p. 19
2	Turbina Eólica de Eixo Horizontal	p. 20
3	Neurônio Biológico	p. 21
4	Neurônio Artificial de McCulloch e Pitts	p. 22
5	<i>Multi-Layer Perceptron</i>	p. 24
6	Estrutura de uma rede alimentada adiante (esquerda) e de uma rede recorrente (direita)	p. 26
7	Teste de Shapiro-Wilk para as amostras de EPMAs com as topologias RC e MLP respectivamente.	p. 50
8	Teste F para as amostras de EPMAs com as topologias RC e MLP res- pectivamente.	p. 50
9	Teste T-Student	p. 51
10	Teste da Soma dos Postos de Wilcoxon.	p. 51

Lista de Tabelas

1	Correlação linear entre a velocidade do vento e a mesma variável em tempos anteriores	p. 33
2	Correlação linear entre a velocidade e a direção do vento em tempos anteriores	p. 33
3	Execuções para determinação do menor EPMA	p. 44
4	Quantidade de neurônios no <i>reservoir</i> associado aos seus respectivos EP-MAs	p. 45
5	Média do Erro Percentual Médio Absoluto para as duas arquiteturas de RNAs	p. 45

Lista de Abreviaturas e Siglas

MLP	<i>Multi-Layer Perceptron</i>
RNA	<i>Rede Neural Artificial</i>
EMQ	<i>Erro Médio Quadrático</i>
EPMA	<i>Erro Percentual Médio Absoluto</i>
RC	<i>Reservoir Computing</i>
ANEEL	<i>Agência Nacional de Energia Elétrica</i>
MTC	<i>Ministério de Ciência e Tecnologia</i>
CBEE	<i>Centro Brasileiro de Energia Eólica</i>
PROINFA	<i>Programa de Incentivo às Fontes Alternativas de Energia</i>
RNR	<i>Rede Neural Recursiva</i>
LSM	<i>Liquid State Machine</i>
ESN	<i>Echo State Networks</i>
CI	<i>Computação Inteligente</i>

1 *Introdução*

Este capítulo inicia com a motivação para a realização desta monografia. Em seguida, são definidos os principais objetivos a serem alcançados. Finalmente, seu encerramento detalha os conteúdos a serem abordados nos capítulos seguintes.

1.1 *Motivação*

Um dos principais desafios da atualidade é a crescente demanda de energia mundial e, a fim de suprir essa necessidade, as fontes de energia mais utilizadas são o petróleo, gás natural e carvão mineral. O grande problema com essas fontes deve-se ao fato de, além de serem extremamente poluidoras, elas são de origem não renovável, ou seja, irão exaurir-se da natureza dentro de alguns anos (FRANCISCO, 2008). Segundo a Agência Internacional de Energia (AIE), caso não se reduza a média de consumo registrada nas últimas décadas, as reservas mundiais de petróleo e gás natural deverão se esgotar em 100 anos, e de carvão, em 200 (FRANCISCO, 2008). Dessa forma, tornou-se essencial a utilização de fontes renováveis e que ajudem no combate à degradação sofrida pelo meio ambiente.

As energias renováveis, pelos motivos citados acima, estão se tornando cada vez mais essenciais para a humanidade e sua maior vantagem é por serem seguras, limpas, abundantes e conseqüentemente não impactarem de forma negativa nas questões ambientais. Entre as diversas fontes disponíveis no mundo está o vento e sua escolha é uma das mais promissoras. Isso se explica devido a sua disponibilidade constante em qualquer lugar e por sua produção ser considerada hoje de custo competitivo (ALBADÓ, 2002).

Devido à aleatoriedade da geração eólica, não é possível garantir um montante fixo de energia ao sistema elétrico. Adicionado a isso está o investimento cada vez mais alto de diversos governos neste tipo de energia para suprir o alto consumo de eletricidade nos últimos anos. Desta forma, e a fim de ajudar os países cuja matriz energética passou a

incluir o vento como fonte alternativa, a previsão de energia eólica mostra-se fundamental para definição de estratégias e planejamentos adequados, eficientes e não dispendiosos. Essa previsão depende principalmente da velocidade do vento. Existem diversos modelos utilizados para realizá-la, vários deles incluindo inteligência artificial. Este trabalho tem como objetivo utilizar uma arquitetura de Rede Neural Artificial (RNA) chamada *Reservoir Computing* (RC) e analisar seu desempenho nessa tarefa.

Embora já existam modelos utilizando tipos mais comuns de RNAs, como o *Multi-Layer Perceptron* (MLP), o *Reservoir Computing* foi escolhido por possuir uma arquitetura em que os neurônios artificiais estão interligados e organizados de forma mais semelhante ao cérebro humano (metáfora que é a origem das RNAs, como o próprio nome indica).

Essa característica da arquitetura do RC permite que essa técnica represente sistemas com comportamento dinâmico, o que redes neurais estáticas como o MLP possuem dificuldade ao fazê-lo (LUKOSEVICIUS; JAEGER, 2009). Consequentemente, aprender as características temporais dinâmicas que representam a série histórica da velocidade do vento, torna-se uma tarefa mais adequada para o RC.

Devido a isso, espera-se que seus resultados de desempenho das previsões sejam melhores dos que aqueles obtidos pelas outras RNAs. Logo, a previsão seria mais precisa e aumentaria a eficiência no planejamento e uso da energia eólica, incentivando seu uso em diversos lugares e preservando o meio ambiente.

1.2 **Objetivos**

1.2.1 **Objetivo Geral**

O objetivo geral deste projeto é avaliar o desempenho da rede neural cuja arquitetura é o *Reservoir Computing* na previsão de velocidade do vento e comparar com os resultados existentes providos de outros modelos de previsão.

1.2.2 **Objetivos Específicos**

Os objetivos específicos deste trabalho são:

1. Implementar a técnica de *Reservoir Computing*
2. O *Reservoir Computing* por ser uma técnica recente possui diversos parâmetros

que ainda estão em fase de pesquisa. Devido à isso, um dos objetivos específicos é explorar alguns desses parâmetros como o raio espectral, o tamanho do reservatório (reservoir) e a conectividade entre os neurônios.

1.3 Estrutura da Monografia

O capítulo 2 apresenta a fundamentação teórica essencial para a compressão do trabalho proposto, a saber: energia eólica e redes neurais artificiais, com enfoque nas duas arquiteturas escolhidas para essa monografia, a *Multi-Layer Perceptron* e o *Reservoir Computing*. No capítulo 3, é descrita a metodologia definida com o objetivo de alcançar o objetivo geral do projeto, tais como o pré-processamento da base de dados e o uso das duas arquiteturas de RNAs escolhidas, incluindo os parâmetros utilizados. Em seguida, o capítulo 4 mostra os resultados obtidos após a simulação das técnicas selecionadas. Por fim, o capítulo 5 apresenta um resumo do trabalho final, descrevendo as dificuldades encontradas, as conclusões obtidas, assim como melhorias e a proposta de trabalhos futuros.

2 Fundamentação Teórica

Este capítulo provê uma breve explicação de todo o conteúdo teórico utilizado como base na tentativa de resolver o problema descrito no capítulo 1. A seção 2.1 faz considerações a respeito da energia eólica. A seção 2.2 discorre sobre RNAs, uma técnica da área de Computação Inteligente (CI) que vem provando ser uma excelente alternativa para problemas de previsão e classificação. A seção 2.3 aborda uma arquitetura específica de uma RNA, o MLP, com seus conceitos e utilização. Por fim, a seção 2.4 discorre sobre a nova técnica RC, seus conceitos e também a sua utilização.

2.1 Energia Eólica

2.1.1 Introdução

A energia eólica tem origem na energia solar. Isso se explica a partir do fato que o vento é causado pelo aquecimento desigual da superfície da Terra pelo Sol. As diferentes superfícies no planeta, terra e água, são responsáveis por esse aquecimento desigual. Durante o dia, o ar sobre a terra se aquece muito mais rápido em comparação ao ar acima da água. À medida que o ar sobre a terra sobe e expande-se devido ao aquecimento, o ar refrigerado pela água irá tomar o espaço vazio deixado por ele. Esse movimento é o que chamamos de vento e pode ser forte o suficiente para a produção de outra forma de energia. A energia eólica é, então, aquela obtida pelo movimento do ar (vento) e é uma fonte abundante, renovável, limpa e disponível em todos os lugares (ALBADÓ, 2002).

A energia total disponível dos ventos ao redor do planeta é estimada baseada na hipótese de que aproximadamente 2% da energia solar absorvida pela Terra é convertida em energia eólica. Apesar de ser um percentual pequeno, esse valor representa centenas de vezes a potência anual instalada nas centrais elétricas do mundo (LEITE DE SÁ, 2011).

A velocidade instantânea do vento V é calculada como o valor médio acrescido de um desvio a partir da média (flutuação), como pode ser visto na equação 2.1:

$$V = \bar{V} + \Delta V \quad (2.1)$$

Onde \bar{V} é a velocidade média do vento e ΔV , a flutuação. Na prática, algumas aplicações levam em consideração apenas a velocidade média do vento.

A velocidade do vento decresce à medida que se aproxima da superfície da Terra devido à fricção entre o ar e o solo. Logo, quanto maior a aspereza do solo, maior o decréscimo da velocidade. Por isso, as medições em estações meteorológicas são tomadas em duas alturas padrão: 2 metros para propósitos agrícolas e 10 metros, o padrão internacional para medições meteorológicas. No Brasil, realizam-se também medições em alturas de 25 e 50 metros (RODRIGUES, 2007).

2.1.2 Turbina Eólica

Turbinas eólicas, também denominadas aerogeradores, são máquinas responsáveis por converter a energia cinética dos ventos em energia elétrica através da rotação das suas pás. Elas são compostas pelo rotor e pela torre que o sustenta, pela transmissão/multiplicação e pelo conversor.

Essas turbinas podem ser classificadas de acordo com a posição do eixo de rotor em turbinas eólicas de eixo horizontal, ou seja, as que possuem pás que giram em um plano perpendicular à direção principal do vento e as turbinas eólicas de eixo vertical cujas pás giram em um plano paralelo à direção do vento. A energia cinética do ar é extraída apenas quando o ar passa através da área interceptada pelas pás rotativas. A energia cinética bruta por unidade de tempo (potência) do vento passando por uma área A perpendicular ao seu vetor velocidade instantânea V , é dada pela seguinte equação 2.2:

$$P = C_p \cdot \frac{1}{2} \cdot \rho \cdot A \cdot V^3 \quad (2.2)$$

Onde:

- ρ = densidade do ar e varia com a latitude e as condições atmosféricas;
- C_p = é o coeficiente de desempenho que se relaciona com a energia cinética de saída e depende do modelo e na relação entre a velocidade do rotor e a velocidade do vento;
- V = velocidade do vento em m/s^2 .

Como pode ser observado na equação 2.2, a energia potencial de uma turbina eólica depende do cubo da velocidade do vento. Por exemplo, se a velocidade do vento em um determinado local dobrar, a energia potencial de saída de uma turbina é multiplicada por 8 (2^3). Essa forte conexão entre a energia potencial com a velocidade do vento evidencia a importância da obtenção de dados de vento precisos para a estimativa da energia disponível.

Segundo Rodrigues, diversos fatores contribuem na complexidade do campo de escoamento das turbinas eólicas, com características acentuadamente não estacionárias. Entre os fatores mais importantes podem-se assinalar os seguintes:

- Efeitos da camada limite atmosférica modificando a magnitude e direção da velocidade de corrente livre;
- Giro num plano inclinado em relação ao vento principal, produzindo variações de ângulo de ataque;
- Perturbações do fluxo de suporte e eixo do rotor;
- Deformação estrutural das pás originando velocidades induzidas de translação e rotação dos elementos da pá.

Outro parâmetro importante que deve ser analisado é a direção do vento, uma vez que mudanças de direção frequentes indicam situações de rajada de vento. Além disso, saber a direção do vento auxilia na determinação da localização das turbinas em um parque eólico (RODRIGUES, 2007).

Quando se trata do aproveitamento da energia eólica, é fundamental distinguir os vários tipos de variações temporais da velocidade dos ventos, como: variações anuais, sazonais, diárias e de curta duração (RODRIGUES, 2007).

- **Variações Anuais:** Para se obter um bom conhecimento do regime dos ventos não é suficiente basear-se na análise de dados de vento de apenas um ano; o ideal é dispor de dados referentes a vários anos. À medida que uma maior quantidade de dados anuais é coletada, as características levantadas do regime local dos ventos tornam-se mais confiáveis.
- **Variações Sazonais:** O aquecimento não uniforme da superfície terrestre resulta em significativas variações no regime dos ventos, resultando na existência de diferentes

estações do ano. Considerando que, em função da relação cúbica entre a potência eólica e a velocidade do vento (na altura do eixo da turbina), como pode ser observado na equação Y, em algumas faixas de potência, uma pequena variação na velocidade implica numa grande variação na potência. Sendo assim, a utilização de médias anuais (ao invés de médias sazonais) pode levar a resultados que se afastam da realidade.

- **Variações Diárias** - As variações diárias na velocidade do vento (brisas marítimas e terrestres, por exemplo) também são causadas pelo aquecimento não uniforme da superfície da Terra. Essas variações são importantes quando, após a escolha de uma região, procura-se o local mais adequado para a instalação do sistema eólico dentro dessa área. Ao comparar a evolução da velocidade média ao longo do dia percebe-se que há uma significativa variação de um mês para os outros. Com esse tipo de informação pode-se projetar melhor o sistema eólico. Por exemplo, nos locais em que os ventos no período do dia são mais fortes do que os ventos no período da noite e a carga de pico ocorre durante o dia, a carga base pode ser fornecida pelo sistema existente e a carga adicional pelo sistema eólico. Entretanto, se a carga de pico ocorre durante a noite, provavelmente a demanda será maior que a geração disponível e um sistema de estocagem pode se fazer necessário.
- **Variações de Curta Duração** - As variações de curta duração estão associadas tanto às pequenas flutuações quanto às rajadas de vento. Num primeiro momento, essas variações não são consideradas na análise do potencial eólico de uma região, desde que não assumam grandes proporções. As flutuações e a turbulência do vento podem afetar a integridade estrutural do sistema eólico, devido à fadiga que ocorre especialmente nas pás da turbina. Por outro lado, as rajadas, caracterizadas por aumentos bruscos de curta duração da velocidade do vento, geralmente acompanhadas por mudanças de direção, merecem maior atenção, pois a previsão de rajadas em tempo real não se revela eficaz no controle da potência das turbinas eólicas.

A geração de eletricidade através do vento começou em torno do início do século XX e alguns dos primeiros desenvolvimentos foram feitos pelos dinamarqueses. Em 1930, em torno de uma dúzia de firmas americanas estavam produzindo e comercializando primitivos geradores eólicos. Estas máquinas poderiam fornecer até 1000 watts de corrente contínua enquanto o vento estivesse soprando (RODRIGUES, 2007).

Muitos países europeus começaram a construir desenhos avançados de geradores eólicos, durante os anos 1950 e 1960. A primeira turbina eólica comercial ligada à rede

elétrica pública foi instalada em 1976, na Dinamarca. Uma das mais memoráveis máquinas de vento foi construída perto de Rutland, Vermon, USA e foi projetada para fornecer 1250kW para a malha elétrica de Vermont (RODRIGUES, 2007).

A turbina de vento de eixo vertical foi inventada pelo engenheiro francês Darrieus, incluindo uma convencional de duas lâminas, conforme a Figura 1. Diferente das turbinas mais utilizadas e que são reorientadas de acordo com o vento, esta é unidirecional, ou seja, aceita o vento de qualquer direção. Outra vantagem é a sua manutenção mais prática, uma vez que o rotor e suas partes elétricas estão localizados na parte inferior da turbina.



Figura 1: Turbina Eólica de Eixo Vertical

[Fonte: Retirada do site: <http://homem.net>]

As modernas turbinas eólicas de eixo horizontal, ilustradas na Figura 2, foram criadas nas primeiras décadas do século XX e sua criação está ligada ao desenvolvimento da aerodinâmica dos aviões e, especialmente, das hélices para sua propulsão.



Figura 2: Turbina Eólica de Eixo Horizontal

[Fonte: Retirada do site: <http://www.dforcesolar.com/pt/turbinas-eolicas/>]

A primeira grande turbina eólica de eixo horizontal foi construída na Rússia em 1931, junto ao Mar Negro. Tinha um rotor de três pás, com um diâmetro de 30 m e uma potência nominal de 100 kW e funcionou cerca de 2 anos (RODRIGUES, 2007).

Atualmente, existem mais de 30 mil turbinas eólicas em operação no mundo e a Dinamarca é o país que mais investiu em energia eólica, responsável por 60% de toda demanda mundial de turbinas. Nos EUA, apenas 1% de toda energia produzida no país é de origem eólica, no entanto, espera-se que nos próximos 15 anos o vento possa ser responsável por 10% de toda energia norte-americana (RODRIGUES, 2007).

2.1.3 Potencial Eólico Brasileiro

No Brasil, as primeiras avaliações de potencial eólico foram realizadas no Ceará e em Fernando de Noronha (PE) e os resultados positivos permitiram a instalação das primeiras turbinas eólicas do Brasil. Em 1998, com o apoio da ANEEL (Agência Nacional de Energia Elétrica) e do Ministério de Ciência e Tecnologia (MCT), o Centro Brasileiro de Energia Eólica (CBEE), da Universidade Federal de Pernambuco (UFPE), publicou a primeira versão do Atlas Eólico da Região Nordeste. A continuidade deste trabalho gerou

o Panorama do Potencial Eólico no Brasil (ANEEL, 2012).

Os resultados positivos do alto poder eólico brasileiro impulsionou a criação do Programa de Incentivo às Fontes Alternativas de Energia (PROINFA) através da Lei 10.438 de 2002. Essa lei estabelece a instalação de parques eólicos com potências de aproximadamente 454 MW no Sul, 483 MW no Nordeste e 163 MW no Sudeste do país (ELETROBRÁS, 2012).

Atualmente, o potencial eólico do Brasil soma 300 GW e está concentrado principalmente no Nordeste e no Sul, com destaque para os estados da Bahia, do Rio Grande do Norte, Ceará e Rio Grande do Sul. Até o fim de 2016, a meta é inserir no sistema elétrico nacional 8,4 GW de potência eólica e isso significará 5,4% de participação na matriz energética brasileira contra os atuais 1,5%. E até 2020, estima-se um investimento de R\$40 bilhões no setor eólico brasileiro (ANEEL, 2012).

2.2 Redes Neurais Artificiais

As redes neurais artificiais são modelos matemáticos inspirados no funcionamento do cérebro humano e adquirem conhecimento através da experiência. Estes modelos tem como base as redes neurais biológicas associadas ao processamento paralelo do cérebro humano (VALENÇA SILVA, 2011).

O sistema nervoso é composto por um conjunto extremamente complexo e numeroso de células, os neurônios. Eles são formados pelos dendritos, que são os terminais de entrada e por onde recebem sinais de outro neurônio; pelo corpo central, onde essa informação é processada, e é propagada até os axônios, que são os longos terminais de saída.

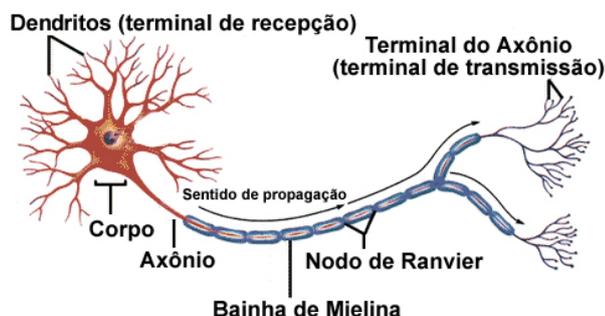


Figura 3: Neurônio Biológico

[Fonte: Retirada do site: <http://meus-projetos.blogspot.com.br>]

Os neurônios se comunicam através das sinapses. Sinapse é a região onde dois neurô-

nios entram em contato e através da qual os impulsos nervosos são transmitidos entre eles. Segundo Valença, um dos aspectos importantes para o entendimento de como os neurônios biológicos funcionam é aquele relacionado à Lei do Tudo ou Nada. Esta lei está ligada com a intensidade do estímulo, chamado limiar excitatório, a partir do qual o neurônio dispara ou não o impulso nervoso. Portanto, se o estímulo for muito pequeno, isto é, se sua intensidade for inferior ao limiar excitatório, não ocorrerá impulso nervoso. Caso passado esse limiar, o potencial de ação do neurônio continuará o mesmo, não importando a intensidade do estímulo.

A primeira representação matemática para esse fato biológico, chamado de neurônio matemático, foi feita por McCulloch e Pitts (1943) (BRAGA; PONCE DE LEON; BERNARDA LUDERMIR, 2000).

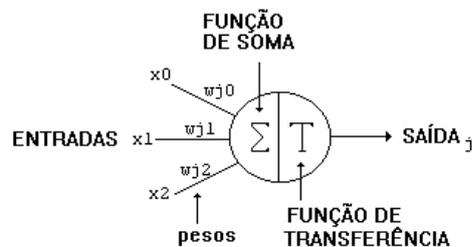


Figura 4: Neurônio Artificial de McCulloch e Pitts

[Fonte: Retirada do site: <http://meus-projetos.blogspot.com.br>]

Como pode ser observado na Figura 4, o modelo proposto é simples e constituído por um conjunto de entradas x_1, x_2, \dots, x_n (correspondente aos dendritos), uma unidade de processamento (representada pelo corpo celular) e uma saída y (correspondente ao axônios). O modelo tem como objetivo representar o neurônio biológico utilizando uma regra de propagação e uma função de ativação.

Cada entrada x_i é multiplicada por um peso w_i a fim de representar o efeito de uma sinapse. Os pesos representam a relevância da entrada ao qual está associado para o neurônio. Ou seja, quanto mais importante aquela entrada, maior será seu peso. A soma do produto $x_i.w_i$, dada pela equação 2.3, é usada como parâmetro para a função de ativação $f(\text{net})$. Esta função, por sua vez, retorna o valor de saída y do neurônio.

Os pesos do neurônio devem possuir valores que, ao multiplicá-los pela entrada, o valor de saída obtido se aproxime ao máximo da saída desejada. Logo, a inteligência do modelo de um neurônio artificial está nos pesos que o constitui (HAYKIN, 2007).

$$net_i = \sum_{i=0}^n x_i \cdot w_i \quad (2.3)$$

Os primeiros modelos propostos foram o Perceptron e o Adaline. Enquanto o Perceptron lida apenas com saídas discretas, Adaline permite respostas no universo contínuo. Essa diferença é ocasionada por causa da função de ativação escolhida.

Dessa forma, a partir da união de neurônios artificiais, em uma ou mais camadas, é possível a construção das Redes Neurais Artificiais (RNAs). Existem vários tipos de RNAs e uma das mais famosas é a *Multi-Layer Perceptron* ou MLP.

2.3 *Multi-Layer Perceptron*

As redes *Multi-Layer Perceptron* (MLP) podem ser consideradas uma generalização do Perceptron. Isso se explica por este tipo de rede apresentar, além das típicas camadas de entrada e saída, pelo menos uma camada intermediária ou camada escondida.

A camada de entrada é onde os neurônios representam as variáveis consideradas entradas do problema. A camada intermediária é a responsável pela solução de problemas não lineares, uma das principais características deste tipo de rede e que a difere do Perceptron e Adaline. Por causa disso, problemas do mundo real, que são considerados funções linearmente não separáveis, podem ser resolvidos através de uma rede MLP. Os neurônios desta camada possuem geralmente uma função de ativação sigmoidal que pode ser a logística ou a tangente hiperbólica.

A camada de saída representa a resposta da rede e é onde se encontra a variável desejada. Em geral, essa variável é aquela que se deseja prever ou classificar. Os neurônios desta camada, além da função de ativação sigmoidal, também podem apresentar uma função do tipo linear.

A Figura 5 ilustra uma rede MLP com três camadas.

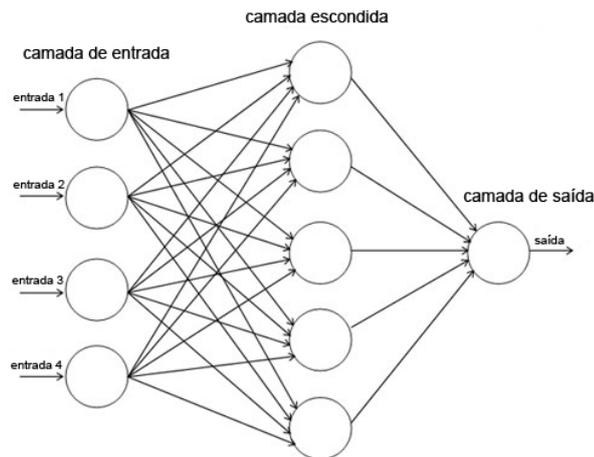


Figura 5: Multi-Layer Perceptron

[Fonte: elaboração própria]

Toda rede neural, a fim de alcançar bons resultados, deve ajustar seus pesos até que um conjunto de pesos ótimo seja estabelecido. Isso é possível na fase de treinamento da rede e é feito com a execução de um algoritmo de otimização. O mais tradicionalmente utilizado é uma generalização da regra delta conhecido como algoritmo *Backpropagation*.

2.3.1 Algoritmo *Backpropagation*

Existem três formas de aprendizado: supervisionado, não supervisionado e por reforço. A MLP aprende de forma supervisionada. Isto é, existe o papel do “professor” e ele conhece as respostas corretas para cada momento da rede. Dessa forma, é possível calcular o erro na saída da rede e, em seguida, ajustar os pesos para que se possa alcançar o valor mais próximo do desejado. Devido à adição da camada intermediária, a complexidade no treinamento aumentou, uma vez que o erro desta camada não é conhecido. O algoritmo *Backpropagation* resolve este problema ao realizar uma propagação recursiva dos erros (HAYKIN, 2007).

A aprendizagem com o uso desse algoritmo acontece em duas etapas: a etapa forward e a etapa backward. Durante a primeira fase, os sinais são propagados no sentido progressivo, ou seja, da camada de entrada para a camada de saída. Ao final desta etapa, o erro é calculado pela diferença entre o valor calculado e o desejado, e os pesos permanecem fixos. Já na segunda fase, o sinal se propaga da camada de saída para a camada de entrada e os pesos são ajustados de acordo com a regra delta generalizada.

A equação 2.4 é responsável pelo ajuste dos pesos ou regra delta generalizada:

$$W_{i,j}^m(t+1) = W_{i,j}^m(t) + \alpha \delta_i^m f^{m-1}(net_j^{m-1}) + \beta \Delta W_{i,j}^m(t-1) \quad (2.4)$$

Onde $w_{i,j}(t+1)$ é o novo valor atribuído ao i -ésimo peso do neurônio j , $w_{i,j}(t)$ é o valor atual do i -ésimo peso do neurônio j no instante t , α é a taxa de aprendizagem, δ_i^m mede a sensibilidade do neurônio ao qual o peso em questão liga o neurônio j e β é o momento.

Caso o neurônio de índice j seja um da camada de saída, a sensibilidade δ_i é calculada de acordo com a seguinte equação:

$$\delta_i^m = (d_i - y_i) f'(net_i) \quad (2.5)$$

Onde d_i é a saída desejada, y_i é a saída encontrada após o treinamento da rede e $f'(net_i)$ é a derivada da função de ativação da camada de saída.

Caso contrário, δ_j é obtido através da equação abaixo:

$$\delta_j^{m-1} = f'^{m-1}(net_j^{m-1}) \sum_{i=l}^N w_{ij}^m \delta_i^m \quad (2.6)$$

Onde δ_i^m é a sensibilidade propagada pelo i -ésimo neurônio da camada imediatamente a frente, w_{ij}^m é o peso j do neurônio i e N é o número de neurônios na camada.

2.3.2 Critério de Parada

Durante o treinamento, um conjunto de padrões é mostrado repetidas vezes a rede neural. Dessa forma, torna-se possível determinar o número de épocas do treinamento. Esse número é variável e depende muito do problema submetido a RNA. Determinar a quantidade de épocas e o momento de parada ideal não é simples, pois pode causar dois problemas.

O primeiro deles está associado ao treinamento demasiado da RNA, que irá decorar padrões e acarretará na perda da sua capacidade de generalização. Esse acontecimento é chamado de *overfitting*. O segundo problema que pode ocorrer é se a RN é treinada por menos tempo que o necessário, a RNA não será capaz de aprender ou de generalizar o suficiente. Um dos critérios de parada mais usados é a validação cruzada.

Para utilizar esse critério se divide o conjunto de padrões em três partes: treinamento,

validação e teste. O conjunto de treinamento é utilizado para o ajuste dos pesos de acordo com o algoritmo de treinamento. Após cada época, a rede previamente treinada é testada utilizando o conjunto de validação cruzada, quando se verifica a diferença entre o valor calculado e o desejado. Levando-se em conta que o conjunto de validação não é utilizado no ajuste dos pesos, ele será sempre inédito para a rede neural, ou seja, enquanto o erro de validação cruzada continuar diminuindo, a rede continua sendo capaz de generalizar.

A partir do momento em que o erro de validação começa a aumentar enquanto o de treinamento continua a diminuir, é sinal de que a rede pode estar decorando os padrões de treinamento. É nesse momento que o treinamento deve ser parado. Posteriormente, o conjunto de testes é introduzido na rede para avaliar o seu desempenho.

2.4 *Reservoir Computing*

2.4.1 Introdução

Além das arquiteturas com alimentação adiante (feedforward), como o MLP, citada na seção anterior, começaram a surgir as Redes Neurais Recorrentes (RNR). Neste novo tipo de rede, existe a adição de conexões recorrentes às já existentes arquiteturas feedforward. Essas conexões transformam o sistema em um sistema dinâmico complexo e ainda mais adequado para a resolução de problemas temporais (ARAÚJO FERREIRA, 2011). No caso deste projeto, torna-se uma opção atraente visto que o problema a ser resolvido, a previsão da velocidade do vento, é de natureza temporal. A Figura 6 apresenta a estrutura de uma rede alimentada adiante e de uma rede recorrente:

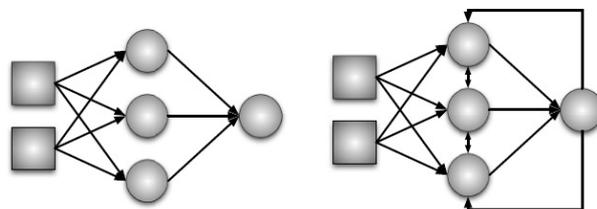


Figura 6: Estrutura de uma rede alimentada adiante (esquerda) e de uma rede recorrente (direita)

[Fonte: (LUKOSEVICIUS; JAEGER, 2009)]

As RNRs são modelos computacionais capazes de criar internamente a memória necessária para armazenar o histórico dos padrões de entrada através das suas conexões recorrentes (ARAÚJO FERREIRA, 2011).

Em 2001, uma nova proposta para o design e treinamento de RNRs foi sugerida, de forma independente, por Wolfgang Maass com o nome *Liquid State Machine* (LSM) e por Herbert Jaeger com o nome *Echo State Networks* (ESN). Verstraeten propôs a unificação dessas duas abordagens em um único termo chamado *Reservoir Computing*. Desde então, RC passou a ser adotado na literatura como um nome genérico para sistemas de aprendizado que consistem de uma rede recorrente dinâmica com simples nós computacionais combinada com uma simples função de saída (VERSTRAETEN, 2009).

Um sistema de *Reservoir Computing* é composto por duas partes principais: o *reservoir* e uma camada de saída linear. O *reservoir* é um sistema dinâmico não linear com uma topologia recorrente de nós de processamento. As conexões entre os nós são geradas aleatoriamente e são globalmente reescaladas a fim de se alcançar um estado dinâmico adequado. Uma propriedade importante do RC é que o *reservoir* possui pesos fixos, isto é, não se faz necessário o seu treinamento. Apenas a camada de saída é treinada e é por este motivo que ela possui uma função de saída. Essa função pode ser, por exemplo, um classificador linear ou um algoritmo de regressão (VERSTRAETEN, 2009).

O fato de apenas a camada de saída precisar ser treinada possibilita, mantendo-se as mesmas entradas, o uso do mesmo *reservoir* para a solução de diferentes tarefas simultaneamente.

Uma característica interessante do RC é aquela baseada nas ESNs, a chamada propriedade do eco. Essa propriedade define que os efeitos de um estado anterior $x(n)$ e um valor de entrada em um futuro estado $x(n+1)$ devem desaparecer gradualmente com o passar do tempo k (isto é, $k \rightarrow$ infinito) e não devem persistir, nem serem amplificados.

Devido às conexões recorrentes, informações sobre entradas passadas são armazenadas na rede. Por causa disso, a rede contém um rico conjunto de transformações não lineares e misturas de sinais de entrada de tempos passados e do presente (os chamados ecos).

2.4.2 Criação e Uso da Técnica de RC

Na explicação a seguir, admite-se que o sistema RC é constituído de N nós do *reservoir*, M entradas e P saídas.

2.4.2.1 Criação das entradas e das conexões do *reservoir*

- 1) Cria-se uma matriz de peso W_{in} $M \times N$ da camada de entrada para o *reservoir*. Os pesos são gerados a partir de uma distribuição aleatória ou de um conjunto discreto.

Se todos os sinais de entrada devem alimentar o *reservoir*, então toda a matriz é preenchida por valores diferentes de zero. Caso contrário, haverá elementos nulos.

- 2) Cria-se uma matriz de peso W_{res} NxN com as interconexões do *reservoir*. Os valores dos pesos são novamente gerados a partir de uma distribuição (exemplo, a gaussiana) ou um conjunto de valores discretos (exemplo, -1,1).
- 3) Reescala-se a matriz de pesos W_{res} globalmente a fim de o *reservoir* possuir uma estabilidade numérica. O modo mais comum de fazer isso é ajustando o raio espectral de W_{res} . O raio espectral de uma matriz é o maior valor absoluto entre os autovalores. Um valor perto de 1 é geralmente considerado um bom ponto inicial para a otimização de ESNs.

2.4.2.2 Simulando o *reservoir* e treinando e testando as saídas

- 1) Criar uma base de dados BD e dividi-la em três subconjuntos: treinamento, validação cruzada e testes.
- 2) O estado da rede no tempo k é denotado como $x[k]$ e uma entrada no mesmo tempo como $u[k]$. Para cada amostra, inicializamos $x[0] = 0$. Antes de iniciar-se o treinamento, os primeiros 100 ciclos são chamados de “*warm up*”, onde a rede neural esquece seus estados iniciais e perde a influência do valor zero (0) atribuído aos estados iniciais. Ao fim do “*warm up*”, acontece o treinamento e a rede é simulada recursivamente através da equação 2.7:

$$x[k + 1] = f(W_{res}x[k] + W_{in}u[k]) \quad (2.7)$$

- 3) Depois que todas as amostras forem simuladas, as matrizes de estado do conjunto de treinamento são concatenadas em uma matriz maior A.
- 4) Calculam-se os pesos da camada de saída. Neste projeto, para realizar este cálculo, foi utilizada a matriz inversa de Moore -Penrose ou pseudoinversa A^+ da matriz A, que é definida como: $x = A^+b$. Toda matriz A do tipo m x n pode ser escrita na forma de:

$$A = U \Sigma V^T \quad (2.8)$$

Onde $U_{m \times m}$ e $V_{n \times n}$ são ortogonais. Pode-se dizer que $U \Sigma V^T$ é a factorização em valores singulares de A.

A pseudo-inversa de A é a matriz nxm: $A^+ = V \Sigma^{-1} U^T$. Existem quatro condições que Penrose deve atender:

- a $AXA = A$
- b $XAX = X$
- c $(AX)^T = AX$
- d $(XA)^T = XA$

Se A é do tipo mxn, então existe uma única matriz X, nxm, que satisfaz estas condições. Se é (quadrada) não singular, então $A^+ = A^{-1}$.

Esses cálculos serão feitos pela rotina do Java chamada JAMA de forma automática.

- 5 Após um exemplo de entrada ser treinado, é feita a validação cruzada a fim de checar se o treinamento já pode ser finalizado. O Erro Médio Quadrático (EMQ) é calculado e armazenado a cada validação cruzada feita de acordo com a fórmula 2.9:

$$EMQ = \frac{1}{N} \sum_{p=1}^N (Z_p - Z_o)^2 \quad (2.9)$$

Onde Z_p é o valor previsto, Z_o é o valor ocorrido e N é o número de valores do conjunto de verificação.

- 6) Ao fim do treinamento, simula-se a rede com o conjunto de teste e calcula-se o Erro Percentual Médio Absoluto (EPMA). O EPMA é definido assim:

$$EPMA = \frac{\sum_{i=1}^n \left| \frac{d_i - y_i}{d_i} \right| 100}{n} \% \quad (2.10)$$

Onde:

- n = número de previsões realizadas;
- d_i = saída desejada para a i-ésima predição;
- y_i = saída obtida para i-ésima predição.

Esses erros são armazenados para posterior realização de testes estatísticos.

Através dos EPMA's calculados para o desempenho da rede com a técnica RC e a MLP, será possível a comparação de qual arquitetura é a melhor escolha para a previsão da velocidade do vento.

2.4.3 Aplicações de *Reservoir Computing*

Várias aplicações de sucesso, sejam com dados sintéticos ou dados reais do mundo da engenharia, do RC já foram relatadas na literatura. Algumas aplicações sintéticas foram classificação de padrões dinâmicos, geração de sinais autônomos ou computação de funções não lineares em taxas instantâneas de trens pulsantes. Na robótica, sistemas de RC foram utilizados para o controle de um braço robótico simulado, para modelar um controlador de robôs, para realizar a previsão da trilha e do movimento de um objeto, detecção de eventos ou em algumas aplicações na competição Robocup, especialmente controle de motor. Sistemas RC tem sido utilizados no contexto de aprendizado por reforço (VERSTRAETEN, 2009).

Há também aplicações de sucesso no campo de processamento de sinais digitais, como o reconhecimento de fala ou modelagem de ruído . Por fim, há relato do uso de RC para geração e previsão de séries temporais caóticas (VERSTRAETEN, 2009).

Em áreas como previsão de séries temporais caóticas e reconhecimento de dígitos isolados, técnicas de RC são consideradas o estado da arte e mais um motivo para o seu uso e avaliação do seu desempenho neste projeto de conclusão de curso (VERSTRAETEN, 2009).

3 Metodologia

Este capítulo detalha como o projeto foi desenvolvido de acordo com a aplicação das teorias apresentadas no capítulo anterior.

3.1 Base de Dados

A base de dados utilizada nos experimentos foi cedida pelo Operador Nacional de Sistema Elétrico (ONS). O ONS é o órgão responsável pela coordenação e controle da operação das instalações de geração e transmissão de energia elétrica no Sistema Interligado Nacional (SIN), sob a fiscalização e regulação da ANEEL (ONS, 2012).

Os dados de velocidade média dos ventos são diários e o período em que foram medidos e coletados vai de 01 de dezembro de 2011 até 31 de julho de 2012. Eles foram observados de 30 em 30 minutos e dentre os dados disponíveis, serão utilizados os valores da velocidade média do vento e da direção para cada um desses instantes.

Esta base de dados apresenta falhas em alguns períodos dos dias, de tal forma que estes valores faltosos foram preenchidos com os valores médios em dias similares. Por exemplo, se o valor de uma segunda-feira estava faltando, os valores das duas segundas-feiras passadas foram somados e divididos por dois e tornaram-se o novo valor a ser colocado na base de dados.

3.2 Método de Seleção de Variáveis

A velocidade do vento pode estar relacionada à outras variáveis anemométricas. Isto é, alterações na velocidade podem ser correlacionadas com a própria variável bem como com outras variáveis tais como: direção do vento, umidade, temperatura, etc. Neste trabalho se utilizará apenas a própria velocidade do vento, em instantes anteriores, e a direção do vento. A fim de verificar o grau de dependência entre estas variáveis, faz-se uso de uma

métrica conhecida como correlação linear.

O termo correlação significa a relação mútua entre dois termos (co + relação) e é utilizado em estatística para indicar a força e direção do relacionamento linear entre duas variáveis aleatórias. A verificação da existência e do grau de relação entre as variáveis é o objeto de estudo da correlação.

A correlação entre duas variáveis x e y pode ser medido de acordo com a equação 3.1:

$$r = \left(\frac{\sum \frac{(x-\bar{x})(y-\bar{y})}{S_x S_y}}{n-1} \right) \quad (3.1)$$

Onde:

- r = correlação linear;
- x = valor da variável x ;
- y = valor da variável y ;
- \bar{x} = valor médio de x ;
- \bar{y} = valor médio de y ;
- S_x = soma dos valores de x ;
- S_y = soma dos valores de y ;
- n = quantidade de valores.

A correlação linear nunca pode ser menor do que -1 ou maior do que 1. Uma correlação próxima à zero indica que as duas variáveis não estão relacionadas. Uma correlação positiva indica que as duas variáveis movem-se juntas e a relação fica mais forte quando se aproxima de 1. Uma correlação negativa significa que as variáveis movem-se em direções opostas e a ligação fica mais forte quando se aproxima de menos 1.

O quadrado da correlação é conhecido como coeficiente de determinação e tem seu valor entre 0 e 1, sempre positivo.

A tabela 1 mostra as correlações lineares existentes entre a velocidade do vento no tempo t e a mesma variável nos tempos: $t - 30$ minutos, $t - 60$ minutos, $t - 90$ minutos, $t - 120$ minutos, $t - 150$ minutos, $t - 180$ minutos e $t - 210$ minutos para todos os dados colhidos pelo ONS.

Tabela 1: Correlação linear entre a velocidade do vento e a mesma variável em tempos anteriores

[Fonte: elaboração própria]

	t - 210 min	t - 180 min	t - 150 min	t - 120 min	t - 90 min	t - 60 min	t - 30 min	t
t - 210 min	1							
t - 180 min	0,944	1						
t - 150 min	0,883	0,944	1					
t - 120 min	0,828	0,883	0,944	1				
t - 90 min	0,777	0,828	0,833	0,944	1			
t - 60 min	0,728	0,777	0,828	0,833	0,944	1		
t - 30 min	0,684	0,728	0,777	0,828	0,833	0,944	1	
t	0,645	0,684	0,728	0,777	0,828	0,833	0,944	1

Através da análise da tabela 1, pode observar que existe uma grande dependência linear entre a velocidade do vento no tempo t e os seus 7 valores anteriores.

A tabela 2 apresenta as correlações lineares existentes entre a velocidade do vento no tempo t e a direção do vento em três tempos anteriores para a mesma base de dados e durante o mesmo período de tempo.

Tabela 2: Correlação linear entre a velocidade e a direção do vento em tempos anteriores

[Fonte: elaboração própria]

	t - 90 min	t - 60 min	t - 30 min	t
t - 90 min	1			
t - 60 min	0,681	1		
t - 30 min	0,493	0,681	1	
t	-0,202	-0,215	-0,236	1

Nesse caso, nota-se que a dependência linear entre essas variáveis é bastante baixa. No entanto, devido a possível correlação não linear entre elas, foram feitos treinamentos à princípio não usando, e depois utilizando três valores passados da direção a fim de observar se há influência no resultado final. Observou-se que a previsão da velocidade do vento foi um pouco melhor com os três valores passados de direção.

De acordo com os cálculos realizados e as conclusões retiradas através da análise das tabelas 1 e 2, decidiu-se que a base de dados a ser utilizada no projeto conteria as seguintes entradas:

- Velocidade do vento no tempo $t - 210$ minutos;
- Velocidade do vento no tempo $t - 180$ minutos;
- Velocidade do vento no tempo $t - 150$ minutos;
- Velocidade do vento no tempo $t - 120$ minutos;
- Velocidade do vento no tempo $t - 90$ minutos;
- Velocidade do vento no tempo $t - 60$ minutos;
- Velocidade do vento no tempo $t - 30$ minutos;
- Velocidade do vento no tempo t ;
- Direção do vento no tempo $t - 90$ minutos;
- Direção do vento no tempo $t - 60$ minutos;
- Direção do vento no tempo $t - 30$ minutos;

As saídas da base serão 2 valores da velocidade do vento ($t + 30$ minutos e $t + 60$ minutos), o que significa a previsão de 1 hora à frente.

3.3 Pré-processamento dos Dados

O primeiro passo durante a fase de pré-processamento é a normalização dos valores. Essa etapa tem como objetivo evitar que valores altos influenciam de forma demasiada nos cálculos da RNA enquanto os valores baixos passam despercebidos. Ou seja, é feita para garantir que as variáveis em intervalos diferentes recebam a mesma atenção durante o treinamento. Além disso, as variáveis devem ter seus valores proporcionais aos limites da função de ativação usada na camada de saída. Se a função de ativação escolhida for a sigmoide logística, seus valores estão limitados entre $[0, 1]$, logo os dados são geralmente normalizados entre $[0,10$ e $0,90]$ ou $[0,15$ e $0,85]$ (VALENÇA SILVA, 2011).

A normalização é feita através da fórmula descrita na equação 3.2:

$$y = \frac{(b - a)(x_i - x_{min})}{(x_{max} - x_{min})} + a \quad (3.2)$$

Onde:

- y = valor normalizado;
- x_i = valor original;
- x_{min} = valor mínimo de x ;
- x_{max} = valor máximo de x ;
- a e b = limites escolhidos. Neste trabalho, $a = 0,15$ e $b = 0,85$.

3.4 Predizendo Velocidade do Vento com MLP

Embora seja bastante utilizada em diversas pesquisas, o MLP necessita que vários de seus parâmetros sejam configuráveis e a escolha de cada um influencia diretamente no resultado final da previsão.

Abaixo, seguem os principais parâmetros do MLP e do algoritmo *Backpropagation*:

- Número de neurônios na camada de entrada;
- Número de neurônios na camada escondida (apenas uma camada escondida);
- Número de neurônios na camada de saída;
- Função de ativação;
- Critério de Parada;
- Taxa de Aprendizado;
- Momento.

A quantidade de entradas foi escolhida de acordo com os resultados da correlação linear feitos na seção 3.2. e foi determinada que seriam 10 entradas (7 valores passados da velocidade e 3 da direção). Os experimentos serão feitos para a base de dados cedida pelo ONS.

A saída são 2 valores de velocidade de vento, ou seja, a previsão para 1 hora a frente.

O algoritmo utilizado é o *Backpropagation* (seção 2.3.2) e a função de ativação escolhida para os neurônios é a sigmoidal logística, descrita na equação 3.3:

$$y = \frac{1}{1 + e^{-net_i}} \quad (3.3)$$

Onde:

- y = valor de saída;
- net_i = a média ponderada dos pesos com as entradas do i -ésimo neurônio

Essa função retorna valores no intervalo $[0, 1]$.

O critério de parada adotado foi o de validação cruzada, com 50% do conjunto de valores para treinamento, 25% para validação cruzada e os restantes 25% para teste.

Diversos testes foram realizados para definição da taxa de aprendizado e do momento sendo que os melhores desempenhos corresponderam aos valores de 0,8 para a taxa de aprendizado e de 0,2 para o momento.

O MLP utilizada foi implementado na linguagem de programação JAVA no ambiente de desenvolvimento Eclipse (SANTOS MEDEIROS, 2010).

3.5 Predizendo a Velocidade do Vento com RC

Assim como o MLP, a técnica de *Reservoir Computing* possui diversos parâmetros que necessitam de configuração. Como é uma área de pesquisa recente, a escolha para essas configurações não pode ser considerada ideal e muitas vezes é realizada de forma aleatória, avaliando-se a cada valor escolhido se ele foi melhor ou pior para o desempenho da rede até encontrar-se um valor considerado ótimo, o que não necessariamente significa o melhor, algo que seria possível alcançar caso houvesse uma forma mais adequada de defini-lo.

Abaixo seguem os parâmetros cujas configurações foram necessárias serem feitas ao longo deste projeto:

- Quantidade de neurônios na camada de entrada;
- Quantidade de neurônios na camada de saída;
- Quantidade de neurônios do reservóir;
- Função de ativação do reservóir;

- Função de ativação da camada de saída;
- Inicialização dos pesos;
- Taxa de conexão do *reservoir*;
- Quantidade de ciclos do “*warm up*”;
- Critério de parada.

A quantidade de entrada permanece a mesma utilizada no MLP visto que o critério de escolha é o mesmo, ou seja, a correlação linear. Além disso, é preciso que esse parâmetro mantenha-se com o mesmo valor da rede neural anterior a fim de realizar-se os testes estatísticos. O mesmo aplica-se para a quantidade de saídas. Ou seja, o RC receberá 10 entradas e terá 2 saídas para prever 1 hora à frente.

A quantidade de neurônios no *reservoir* é um dos parâmetros para o qual não existe um critério fixo que o determine. Foi escolhido aleatoriamente e após a verificação do EPMA, ao fim de cada treinamento, o melhor número de neurônios. Observou-se que para a solução do problema em questão, o número ideal encontrado foi de 100 neurônios.

Como citado na seção 2.4.2.1., os pesos da camada de entrada para o *reservoir* e os pesos do *reservoir* são gerados aleatoriamente a partir de uma distribuição aleatória.

Os estados do *reservoir* são inicializados como zero (0). Por causa disso, como citado na seção 2.4.2.1., decidiu-se adicionar a rede uma fase denominada “*warm up*”. Durante o “*warm up*”, não há necessidade de encontrar os pesos da camada de saída, nem de calcular um valor de saída. Isso se explica pois esta fase “*warm up*” é apenas para atualizar os estados do *reservoir* e retirar a dependência do estado inicial. O número de ciclos do “*warm up*” escolhido foi de 5.

A taxa de conexão dos neurônios do *reservoir* foi de 15%. Ou seja, apenas 15% das ligações possuem pesos diferentes de zero (0) associadas a elas.

O critério de parada escolhido também foi o de validação cruzada, com 50% da base de dados para o conjunto de treinamento, 25% para a validação cruzada e os outros 25% para teste.

A função de ativação escolhida no *reservoir* foi a sigmoide logística e sua fórmula foi descrita na seção anterior. Na camada de saída, a função selecionada foi a linear que é descrita como na equação 3.4:

$$y = net_i \tag{3.4}$$

Onde:

- y = valor de saída;
- net_i = a média ponderada dos pesos com a entrada do i -ésimo neurônio

Durante este projeto foi implementada uma Rede Neural com a técnica de RC na linguagem de programação JAVA e no ambiente de desenvolvimento Eclipse. O pseudo-código abaixo é a forma sintetizada de como funciona essa RNA:

Algoritmo 1: Pseudocódigo do RC

```
1 Definir a quantidade de neurônios na camada de entrada ;
2 Definir a quantidade de neurônios no reservoir ;
3 Definir a quantidade de neurônios na camada de saída ;
4 Gerar aleatoriamente os pesos da matriz Win entre -1 e 1;
5 Gerar aleatoriamente os pesos da matriz Wres entre -1 e 1;
6 Normalizar os pesos de Wres para que o raio espectral da matriz fique menor ou
  igual a 1;
7 while até o final da quantidade de ciclos de warm up do
8   | atualiza os estados dos neurônios do RC;
9 end
10 while até que o critério de parada seja atingido do
11   | for cada valor do conjunto de entrada do
12     | atualiza os estados dos neurônios do RC;
13   end
14   Calcula a pseudo-inversa de Moore-Penrose para encontrar a matriz de pesos
     que conecta o RC à camada de saída;
15   for cada valor do conjunto de validação cruzada do
16     | atualiza os estados dos neurônios do RC;
17   end
18   Calcula os valores de saída do RC;
19   Calcula o EMQ;
20   Verifica se o critério de parada já foi atingido;
21 end
22 for cada valor do conjunto de testes do
23   | atualiza os estados dos neurônios do RC;
24 end
25 Calcula os valores de saída do RC;
26 Calcula o EPMA // para o caso de previsão;
27 Calcula a taxa de acerto //para classificação;
```

3.6 Testes Estatísticos

Após os 30 treinamentos (JURISTO; MORENO M., 2001) com cada tipo de rede neural, foram realizados testes estatísticos a fim de avaliar qual técnica apresenta o melhor desempenho na previsão da velocidade do vento ou se seus resultados podem ser considerados estatisticamente iguais.

Entre os diversos testes existentes na literatura, se encontram o teste T-student e o teste de Wilcoxon. Para aplicação do teste T-student deve-se começar utilizando o teste de Shapiro-Wilk com o objetivo de checar se os dados estão normalmente distribuídos e, caso sejam consideradas amostras normais, é realizado em seguida o teste F para verificar se a variância pode ser considerada como originada de uma mesma população. Ao passar nestes testes é aplicado o teste T-student para dados não emparelhados. Caso as amostras não sejam consideradas normais ou não passem no teste F, é aplicado o teste da Soma dos Postos de Wilcoxon.

3.6.1 Teste de Shapiro-Wilk

A fim de realizar o teste de Shapiro-Wilk, são seguidos os seguintes passos:

1. Formulação da hipótese:

$$\begin{cases} H_0 : \text{A amostra provém de uma população Normal} \\ H_1 : \text{A amostra não provém de uma população Normal} \end{cases}$$

2. É estabelecido o nível de significância do teste (α) em 0,05;

3. Calcula-se a estatística do teste:

- a. Ordenar as n observações da amostra: $x_{(1)}, x_{(2)}, x_{(3)}, \dots, x_{(n)}$;

- b. Calcular:

$$\sum_{i=1}^n (x_{(i)} - \bar{x})^2; \quad (3.5)$$

- c. Calcular b:

$$\begin{cases} \sum_{i=1}^{\frac{n}{2}} a_{n-i+1} (x_{(n-i+1)} - x_{(1)}) & \text{se } n \text{ é par} \\ \sum_{i=1}^{\frac{(n+1)}{2}} a_{n-i+1} (x_{(n-i+1)} - x_{(i)}) & \text{se } n \text{ é ímpar} \end{cases} \quad (3.6)$$

Onde $a_{(n-i+1)}$ são constantes geradas pelas médias, variâncias e covariâncias das estatísticas de ordem de uma amostra de tamanho n de uma distribuição Normal.

d. Calcular W :

$$W = \frac{b^2}{\sum_{i=1}^n (x_{(i)} - \bar{x})^2} \quad (3.7)$$

Onde, x_i são os valores da amostra ordenados ($x_{(1)}$ é o menor). Menores valores de W são evidências de que os dados são normais.

4. Toma-se a decisão: Rejeitar H_0 ao nível de significância α se $W_{calculado} < W_{\alpha}$.

3.6.2 Teste F

O teste F é uma técnica estatística cujo objetivo é testar a igualdade entre duas variâncias. Permite testar se a variabilidade dentro dos grupos é maior que a existente entre os grupos.

$$F = \frac{S_1^2}{S_2^2} \quad (3.8)$$

No numerador colocamos a maior e no denominador a menor das variâncias encontradas nas amostras. O valor de $F_{calculado}$ será comparado com o valor de $F_{critico}$ obtido em uma tabela de distribuição de F, considerando o nível de significância adotado, os graus de liberdade do numerador (n_1-1) e os graus de liberdade do denominador (n_2-1).

No teste F, os seguintes passos são seguidos:

1. Formulação das hipóteses:

$$\begin{cases} H_0 : S_1^2 = S_2^2 \\ H_1 : S_1^2 \neq S_2^2 \end{cases}$$

2. Calcula-se o valor de F e seu valor crítico foi procurado na tabela;

3. Compara-se os valores de F:

- a. Se o valor de F é menor que o encontrado na tabela, não se pode rejeitar a hipótese H_0 (significa que a probabilidade de F é maior que o nível de significância adotado), concluindo que as amostras têm variâncias iguais;
- b. Quando o valor de F for maior ou igual ao valor da sua tabela rejeitamos H_0 , concluindo que as variâncias comparadas são diferentes e não pertencem a mesma população.

3.6.3 Teste T-Student para amostras independentes

O teste T Student é empregado na comparação das médias de duas amostras independentes, com variáveis medidas nas escalas intervalar ou de razão, com distribuição normal.

Para aplicar esse teste, deve-se ter garantido anteriormente que a amostra é normalmente distribuída e que elas pertencem a uma mesma população, ou seja, suas variâncias são estatisticamente iguais.

$$T = \frac{\bar{x}_1 - \bar{x}_2}{EP_{(\bar{x}_1 - \bar{x}_2)}} \quad (3.9)$$

Onde o numerador é a diferença entre as médias das duas amostras e o denominador é o erro padrão que pode ser calculado através da seguinte fórmula:

$$EP_{(\bar{x}_1 - \bar{x}_2)} = \sqrt{\frac{S_p^2}{n_1} + \frac{S_p^2}{n_2}} \quad (3.10)$$

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \quad (3.11)$$

O valor de T_{obtido} , calculado através da fórmula 3.9 deve ser comparado ao valor de $T_{critico}$ obtido na tabela de distribuição de T de Student, considerando $(n_1 + n_2 - 2)$ graus de liberdade e o tipo de teste de hipótese, neste caso um teste de hipótese unilateral.

Para realizar o teste T é feito:

1. Formula-se as hipóteses: A hipótese nula inclui a igualdade das médias das amostras, enquanto a hipótese alternativa, que elas sejam diferentes, ou ainda que $m_1 > m_2$ ou $m_1 < m_2$. No caso deste experimento, será utilizado como hipótese alternativa $m_1 < m_2$.

$$\begin{cases} H_0 : m_1 \geq m_2 \\ H_1 : m_1 < m_2 \end{cases}$$

2. Calcula o valor de T e procura-se o valor do $T_{critico}$ na tabela.
 - a. Nos testes unilaterais do tipo $m_1 < m_2$, se o valor de T é maior que o encontrado na tabela, não rejeitamos a hipótese H_0 (significa que a probabilidade de T é

maior que o nível de significância adotado). Quando T for menor ou igual ao valor da tabela, rejeitamos H_0 .

3.6.4 Teste da Soma dos Postos de Wilcoxon

Esse é um teste não paramétrico que usa os postos, ou seja, as posições que os valores dos dados amostrais de duas populações independentes ocupam quando colocados em ordem crescente. Ele é equivalente ao teste U de Mann-Whitney e é usado para testar a hipótese nula de que duas amostras independentes provêm de populações com medianas iguais. A hipótese alternativa é a afirmativa de que as duas populações tem medianas diferentes.

Quando um teste não-paramétrico é utilizado, supõe-se que a distribuição de seus dados experimentais não seja normal, ou que não existem elementos suficientes para poder afirmar que seja. É apenas necessário que os dados sejam ordenáveis. Os requisitos para que se possa utilizar esse teste são:

- Existir duas amostras independentes de dados selecionados aleatoriamente.
- Se ambos os tamanhos amostrais forem maiores que 10, então a distribuição amostral (R) será aproximadamente normal, com média μ e desvio padrão α .
- Não há qualquer exigência de que as duas populações tenham uma distribuição normal ou qualquer outra distribuição.

De posse dos dados, aplica-se seus valores nas seguintes fórmulas:

$$z = \frac{R - \mu_R}{\sigma_R} \quad (3.12)$$

$$\mu_R = \frac{n_1(n_1 + n_2 + 1)}{2} \quad (3.13)$$

$$\sigma_R = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} \quad (3.14)$$

Onde n_1 é o tamanho da amostra 1, n_2 é o tamanho da amostra 2, R é a soma dos postos da amostra de tamanho n_1 , μ_R é a média dos valores amostrais R, que é esperada quando duas populações tem medianas iguais e σ_R é o desvio padrão dos valores amostrais de R, que também é esperado quando as duas populações tem medianas iguais.

4 *Resultados*

Este capítulo mostra os resultados alcançados através da aplicação da metodologia detalhada no Capítulo 3.

As configurações utilizadas na rede MLP e na RC são as definidas na seção 3.4 e 3.5 respectivamente. Na técnica MLP, a quantidade de neurônios da camada escondida foi definida realizando-se o treinamento da rede com vários valores até encontrar-se os três que possuíam o menor EMPA. Em seguida, alterou-se os valores de taxa de aprendizado e momento. A tabela abaixo mostra as execuções da rede com várias combinações de valores:

Tabela 3: Execuções para determinação do menor EPMA

[Fonte: elaboração própria]

Execução	Taxa de aprendizado	Momento	Qtde. Neurônios da Camada Escondida	EPMA
1	0,2	0,8	15	40,45%
2	0,8	0,2	15	18,67%
3	0,2	0,8	6	23,51%
4	0,8	0,2	6	17,40%
5	0,2	0,8	30	19,54%
6	0,8	0,2	30	18,24%

Ao utilizar a técnica RC, algumas execuções foram feitas para testar a quantidade de neurônios no reservóir e qual valor apresentava o menor EPMA. A tabela a seguir mostra esses testes:

Tabela 4: Quantidade de neurônios no *reservoir* associado aos seus respectivos EPMA

Fonte: [elaboração própria]

Qtd. Neurônios no <i>reservoir</i>	EPMA
40	7,1609%
50	7,1608%
65	7,1585%
80	7,1936%
100	7,1505%

Através da análise da tabela 4, nota-se que os valores dos EPMA não se alteram muito com diferentes quantidades de neurônios no *reservoir*. Dessa forma, a fim de manter o máximo de recursões possíveis e dessa forma, deixar a rede mais dinâmica, escolheu-se deixar 100 neurônios como citado no capítulo anterior.

Foram realizadas 30 simulações com cada uma das topologias e a média dos EPMA pode ser encontrada abaixo:

Tabela 5: Média do Erro Percentual Médio Absoluto para as duas arquiteturas de RNAs

[Fonte: elaboração própria]

Topologia	Média do Erro Percentual Médio Absoluto
RC	7,17 %
MLP	18,55%

Em seguida, para cada conjunto de 30 simulações, foi realizado o teste de Shapiro-Wilk para checar se as amostras podem ser consideradas normais. A formulação das hipóteses para a rede RC pode ser encontrada a seguir:

$$\begin{cases} H_0 : \text{A amostra dos EPMA com a arquitetura RC provém de uma população Normal} \\ H_1 : \text{A amostra dos EPMA com a arquitetura RC não provém de uma população Normal} \end{cases}$$

As hipóteses pertinentes à rede MLP foram definidas da seguinte forma:

$$\begin{cases} H_0 : \text{A amostra dos EPMA com a arquitetura MLP provém de uma população Normal} \\ H_1 : \text{A amostra dos EPMA com a arquitetura MLP não provém de uma população Normal} \end{cases}$$

Utilizando o software R, que como padrão define o nível de significância como 0,05, observou-se que ambas as amostras não provém de uma população normal, ou seja, não

são normalmente distribuídas, uma vez que o p-value calculado durante o teste foi menor que 0,05, seus valores sendo 0,01 para a topologia RC e 0,03 para o MLP.

A partir deste resultado, tornou-se desnecessário a realização do teste T de Student, cuja pré-condição para seu uso é de que as amostras fossem normais. Por causa disso, o próximo passo foi à aplicação do teste da Soma dos Postos de Wilcoxon.

Esse teste é não-paramétrico, ou seja, supõe-se que a distribuição dos dados não seja normal e as hipóteses foram definidas a seguir:

$$\begin{cases} H0 : \text{O desempenho das redes neurais com as arquiteturas RC e MLP são consideradas estatisticamente iguais} \\ H1 : \text{O desempenho da rede neural com a arquitetura RC é melhor do que com a arquitetura MLP} \end{cases}$$

O resultado deste teste mostrou que o p-value possui um valor muito menor do que o nível de significância, logo, refuta-se a hipótese nula e conclui-se que o uso da técnica proposta de RC possui um melhor desempenho.

Um *script* com todos os testes citados neste capítulo pode ser encontrado no Apêndice, assim como imagens mostrando os p-values calculados e sua comparação com o nível de significância.

5 *Considerações Finais*

5.1 Conclusões

Este trabalho de conclusão teve como principal objetivo prever a velocidade do vento para 1 hora a frente utilizando a técnica de *Reservoir Computing* e comparar com os resultados existentes providos por outros modelos de previsão. Como a rede neural MLP é a mais utilizada neste tipo de aplicação, essa foi a técnica escolhida e com a qual os resultados do RC foram comparados.

A fim de alcançar esse objetivo, foi implementada uma rede neural com a topologia RC e uma base de dados disponibilizada pelo Operador Nacional do Sistema Elétrico foi utilizada. Realizou-se diversas simulações com as duas arquiteturas e seus resultados foram comparados.

Através dos testes estatísticos, foi comprovado que o desempenho com o RC é melhor do que aquele encontrado pelo MLP e abre-se um campo de pesquisa para o estudo de redes neurais dinâmicas como o RC para previsão de séries temporais.

Uma grande dificuldade encontrada no decorrer do projeto foi na implementação e utilização da técnica de *Reservoir Computing*. Por ser uma abordagem nova, não há muitos trabalhos disponíveis que auxiliem no melhor entendimento de alguns aspectos da técnica. Além disso, para diversos parâmetros não existe uma forma científica para defini-los, sendo muitos deles escolhidos de forma aleatória e baseado em simulações que apresentaram o melhor resultado para o problema sendo tratado. Conseqüentemente, foram necessários diversos treinamentos apenas com o intuito de definir a melhor configuração da rede.

5.2 Trabalhos Futuros

Como trabalhos futuros, pretende-se a realização de um estudo mais aprofundado dos parâmetros do RC a fim de encontrarem-se as melhores maneiras de defini-los e que

possam influenciar de forma positiva no desempenho da RNA.

Para a base utilizada, há a possibilidade de preencher os valores ausentes com o método de k-vizinhos ou curvas típicas, o que pode aproximar tais valores dos que deveriam ter sido coletados pelo ONS e ter como consequência uma melhora na previsão da velocidade.

Por fim, é necessário obter-se bases de dados de outros lugares do país e do mundo e observar se a técnica de RC continua sendo a melhor escolha na previsão da velocidade do vento.

Referências

- ALBADÓ, R. *Energia Eólica*. 1. ed. [S.l.]: ArtLiber, 2002. 156 p.
- ANEEL. *Atlas da Energia Eólica do Brasil*. 2012. Acessado em 01 de setembro de 2012. Disponível em: <<http://www.aneel.gov.br/>>.
- ARAÚJO FERREIRA, A. *Um método para design e treinamento de Reservoir Computing aplicado à previsão de séries temporais*. Tese (Doutorado) — Universidade de Pernambuco, 2011.
- BRAGA, A. de P.; PONCE DE LEON, F. d. C. A.; BERNARDA LUDERMIR, T. *Redes Neurais Artificiais: Teorias e Aplicações*. 2. ed. [S.l.]: LTC, 2000.
- ELETROBRÁS. *Proinfa*. 2012. Acessado em 29 de agosto de 2012. Disponível em: <<http://www.eletrabras.com/elb/data/Pages/LUMISABB61D26PTBRIE.htm>>.
- FRANCISCO, W. de Cerqueira e. *Fontes de Energia*. 2008. Acessado em 29 de agosto de 2012. Disponível em: <<http://www.mundoeducacao.com.br/geografia/fontes-energia.htm>>.
- HAYKIN, S. *Redes Neurais: Princípios e Práticas*. 2. ed. [S.l.]: Bookman, 2007.
- JURISTO, N.; MORENO M., A. *Basics of Software Engineering Experimentation*. [S.l.]: Kluwer Academic Publisher, 2001.
- LEITE DE SÁ, A. *Energia Eólica: Princípios e Aplicações*. 2011. Acessado em 09 de agosto de 2012. Disponível em: <<http://www.cresesb.cepel.br/>>.
- LUKOSEVICIUS, M.; JAEGER, H. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, v. 3, p. 127–149, 2009.
- ONS. *Operador Nacional do Sistema Elétrico*. 2012. Acessado em 11 de outubro de 2012. Disponível em: <<http://www.ons.org.br/home/>>.
- RODRIGUES, G. *Utilização de Redes Neurais para Previsão de Ventos no Horizonte de 24 Horas*. Dissertação (Mestrado) — Universidade Federal do Rio de Janeiro, 2007.
- SANTOS MEDEIROS, d. O. C. S. dos. *Um Sistema Híbrido Baseado em Rede Neural e Colônia de Formigas*. Dissertação (Mestrado) — Universidade de Pernambuco, 2010.
- VALENÇA SILVA, J. M. *Fundamentos das Redes Neurais*. [S.l.]: Livro Rápido, 2011.
- VERSTRAETEN, D. *Reservoir Computing: computation with dynamical systems*. Tese (Doutorado) — Ghent University, 2009.

APÊNDICE A – Resultados dos Testes Estatísticos no Software R

Segue abaixo os resultados dos testes estatísticos no Software R realizados no decorrer deste projeto.

```

> shapiro.test(ventoRC)

      Shapiro-Wilk normality test

data:  ventoRC
W = 0.9142, p-value = 0.01897

> shapiro.test(ventoMLP)

      Shapiro-Wilk normality test

data:  ventoMLP
W = 0.9244, p-value = 0.03495

```

Figura 7: Teste de Shapiro-Wilk para as amostras de EPMA's com as topologias RC e MLP respectivamente.

[Fonte: elaboração própria]

```

> var.test(ventoRC,ventoMLP)

      F test to compare two variances

data:  ventoRC and ventoMLP
F = 7e-04, num df = 29, denom df = 29, p-value < 2.2e-16
alternative hypothesis: true ratio of variances is not equal to 1

```

Figura 8: Teste F para as amostras de EPMA's com as topologias RC e MLP respectivamente.

[Fonte: elaboração própria]

```
> t.test(ventoRC,ventoMLP,var.equal=TRUE,alternative="two.sided")  
  
Two Sample t-test  
  
data: ventoRC and ventoMLP  
t = -55.4554, df = 58, p-value < 2.2e-16  
alternative hypothesis: true difference in means is not equal to 0
```

Figura 9: Teste T-Student

[Fonte: elaboração própria]

```
> wilcox.test(ventoRC,ventoMLP)  
  
Wilcoxon rank sum test  
  
data: ventoRC and ventoMLP  
W = 0, p-value < 2.2e-16  
alternative hypothesis: true location shift is not equal to 0
```

Figura 10: Teste da Soma dos Postos de Wilcoxon.

[Fonte: elaboração própria]

APÊNDICE B - Script dos Testes Estatísticos

Segue abaixo o *script* utilizado no software R para a realização dos testes estatísticos.

```
dados<-read.csv("dados.csv")
ventoRC<-c(dados[,1])
ventoMLP<-c(dados[,2])

shapiro.test(ventoRC)
shapiro.test(ventoMLP)
var.test(ventoRC,ventoMLP)
t.test(ventoRC,ventoMLP,var.equal=TRUE,alternative="two.sided")
wilcox.test(ventoRC,ventoMLP)
```