



# **Comparação entre MLP e RC para Previsão de Geração de Potência Eólica Considerando Representações de Sazonalidade**

**Trabalho de Conclusão de Curso**

**Engenharia da Computação**

**Luis Gustavo Arcoverde Souto**  
**Orientador: Mêuser Jorge Silva Valença**



**UNIVERSIDADE  
DE PERNAMBUCO**

**Universidade de Pernambuco  
Escola Politécnica de Pernambuco  
Graduação em Engenharia de Computação**

**Luis Gustavo Arcoverde Souto**

**Comparação entre MLP e RC para  
Previsão de Geração de Potência  
Eólica Considerando Representações  
de Sazonalidade**

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

**Recife, Junho de 2017.**

### MONOGRAFIA DE FINAL DE CURSO

#### Avaliação Final (para o presidente da banca)\*

No dia 13 de julho de 2017, às 8:00 horas, reuniu-se para deliberar a defesa da monografia de conclusão de curso do discente **LUIZ GUSTAVO ARCOVERDE SOUTO**, orientado pelo professor **Mêuser Jorge Silva Valença**, sob título **Comparação entre MLP e RC para Previsão de Geração de Potência Eólica Considerando Representações de Sazonalidade**, a banca composta pelos professores:

**Sérgio Mario Lins Galdino**  
**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: 9,5 ( *nao e meio* )

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

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

*Sérgio Galdino*

SÉRGIO MARIO LINS GALDINO

*Mêuser Valença*

MÊUSER JORGE SILVA VALENÇA

\* Este documento deverá ser encadernado juntamente com a monografia em versão final.

*Dedico este trabalho a todos que me incentivaram e me apoiaram durante a minha jornada, principalmente meus pais e família.*

# Agradecimentos

Agradeço aos meus pais, Cláudio Feitoza Souto e Márcia Maria Cabral Arcoverde Souto, pelo apoio e educação.

Agradeço a minha namorada Heloisa Soares Jácomo de Araújo, que participou ativamente de todo meu processo de graduação, sempre me aconselhando e apoiando nas horas mais difíceis.

Agradeço aos meus familiares pelo auxílio e compreensão.

Agradeço a minha turma, por ser unida e ajudar uns aos outros nos momentos que mais precisávamos.

Agradeço aos amigos Pedro Queiroz, Renato Valença, Caio Vinicius e Everalda Soares por estarem sempre a disposição para me ajudar durante a construção do trabalho.

Finalmente agradeço ao meu orientador, Mêuser Valença, pelos seus ensinamentos, pela orientação e apoio.

# Resumo

O recurso eólico apresenta hoje um grande potencial como fonte alternativa energética. Uma busca por melhorias na previsão da geração eólica ocorre devido ao crescimento de sua capacidade. Um grande desafio desse tipo de previsão é prever a quantidade de energia gerada um dia à frente, uma vez que o vento é uma variável muito inconstante. Porém, prever essa geração eólica é de suma importância para determinar o quanto de energia será gerada no futuro, servindo de alerta, por exemplo, para o sistema elétrico de uma região ter seu fluxo energético otimizado. Uma maneira tradicional de previsão utilizando algoritmos de computação inteligente é a Rede Neural Artificial. Este trabalho está centrado na utilização de diferentes configurações de duas técnicas de redes neurais, o Reservoir Computing e o Multilayer Perceptron, para prever a potência gerada pelo vento. Como objetivo final, este trabalho apontará a possibilidade de melhorias e diminuição de erros nas previsões realizadas, baseados nos estudos comparativos e testes estatísticos aplicados aos resultados de cada rede.

**Palavras-chave:** Previsão de vento, Redes Neurais Artificiais, Multilayer Perceptron, Reservoir Computing

# Abstract

The wind resource today has great potential as an alternative energy source. A search for improvements in wind power forecasting is due to the growth in its capacity. A major challenge of this type of forecasting is to predict the amount of energy generated one day ahead, since the wind is a very inconsistent variable. However, predicting this wind generation is very important to determinate how much energy will be generated in the future, serving as an alert to electrical systems to have its energy flow best optimized. A traditional way of predicting using intelligent computing algorithms is the Artificial Neural Network. This work is centered on the use of different configurations of two types of neural networks, the Reservoir Computing and the Multilayer Perceptron, to calculate the power generated by the wind. As final objective, this work will point out the possibility of improvements and decreases of errors in the predictions made, based in comparative studies and methods of testing applied to the results of each network.

**Key-words:** Wind Power Prediction, Artificial Neural Network, Multilayer Perceptron, Reservoir Computing

# Sumário

<b>Capítulo 1 – Introdução</b>	<b>1</b>
1.1 Caracterização do Problema.....	1
1.2 Objetivos.....	2
1.2.1 Objetivos Gerais.....	2
1.2.2 Objetivos Específicos.....	3
1.3 Estrutura da Monografia.....	3
<b>Capítulo 2 - Fundamentação Teórica</b>	<b>4</b>
2.1 Energia Eólica.....	4
2.1.1 Conceitos.....	4
2.1.2 Turbina Eólica.....	5
2.1.3 Sazonalidade.....	8
2.1.4 Modelos de Previsão.....	9
2.2 Redes Neurais Artificiais.....	10
2.3 Multi-Layer Perceptron.....	12
2.3.1 Algoritmo Backpropagation.....	14
2.4 Reservoir Computing.....	16
<b>Capítulo 3 - Metodologia</b>	<b>19</b>
3.1 Base de dados.....	19
3.2 Pré- processamento dos dados.....	20
3.3 Entradas das Bases.....	24
3.3.1 Entradas da rede MLP.....	24
3.3.2 Entradas da rede RC.....	25
3.4 Configuração das Redes.....	26
3.4.1 Configurações das redes MLP.....	26

3.4.2	Configurações das redes RC.....	27
3.5	Testes estatísticos.....	28
3.5.1	Teste de Wilconxon.....	28
<b>Capítulo 4 - Resultados</b>		<b>30</b>
4.1	Experimentos com o MLP.....	30
4.1.1	Estudo de Caso MLP para o Parque A.....	30
4.1.2	Estudo de Caso MLP para o Parque B.....	31
4.1.3	Estudo de Caso MLP para o Parque C.....	31
4.2	Experimentos com o RC.....	32
4.2.1	Estudo de Caso RC para o Parque A.....	32
4.2.2	Estudo de Caso RC para o Parque B.....	33
4.2.3	Estudo de Caso RC para o Parque C.....	34
4.3	Comparação MLP x RC.....	34
4.3.1	Estudo comparativo para o Parque A.....	34
4.3.2	Estudo comparativo para o Parque B.....	35
4.3.3	Estudo comparativo para o Parque C.....	36
4.4	Box Plots.....	36
4.4.1	Box Plots para o Parque A.....	37
4.4.2	Box Plots para o Parque B.....	37
4.4.3	Box Plots para o Parque C.....	38
<b>Capítulo 5 Considerações Finais</b>		<b>40</b>
5.1	Conclusão.....	40
5.2	Trabalhos futuros.....	40
<b>Bibliografia</b>		<b>42</b>

# Índice de Figuras

<b>Figura 1.</b> Aerogeradores de Eixo Horizontal - Bahia.....	6
<b>Figura 2.</b> Aerogerador de Eixo Vertical.....	6
<b>Figura 3.</b> Componentes de uma turbina de eixo horizontal.....	7
<b>Figura 4.</b> Neurônio Biológico .....	10
<b>Figura 5.</b> Modelo de Neurônio de MacCulloch e Pitts .....	11
<b>Figura 6.</b> Arquitetura da Rede MLP com 2 camadas escondidas .....	13
<b>Figura 7.</b> Arquitetura da rede recorrente RC.....	16
<b>Figura 8.</b> Box Plots dos cenários para o Parque A .....	37
<b>Figura 9.</b> Box Plots dos cenários para o Parque B .....	38
<b>Figura 10.</b> Box Plots dos cenários para o Parque C .....	39

# Índice de Tabelas

<b>Tabela 1</b> Descrição da Equação de Alteração dos Neurônios do Reservatório .....	17
<b>Tabela 2</b> Normalização de uma amostra dos dados entre [0,15-0,85] .....	20
<b>Tabela 3</b> Descrição da Fórmula da Normalização .....	21
<b>Tabela 4</b> Base de Dados Defasada até (t-4).....	22
<b>Tabela 5</b> Correlação entre variáveis da Tabela 3.....	23
<b>Tabela 6</b> Parâmetros do Algoritmo RC .....	25
<b>Tabela 7</b> Principais parâmetros utilizados para a rede MLP .....	26
<b>Tabela 8</b> Descrição da equação do EPMA.....	27
<b>Tabela 9</b> Principais parâmetros utilizados para a rede MLP .....	27
<b>Tabela 10</b> Comparação do melhor cenário utilizando MLP para o Parque A .....	31
<b>Tabela 11</b> Comparação do melhor cenário utilizando MLP para o Parque B.....	31
<b>Tabela 12</b> Comparação do melhor cenário utilizando MLP para o Parque C .....	32
<b>Tabela 13</b> Comparação do melhor cenário utilizando RC para o Parque A .....	33
<b>Tabela 14</b> Comparação do melhor cenário utilizando RC para o Parque B .....	33
<b>Tabela 15</b> Comparação do melhor cenário utilizando RC para o Parque C.....	34
<b>Tabela 16</b> Comparação de melhor algoritmo para o Parque A .....	35
<b>Tabela 17</b> Comparação de melhor algoritmo para o Parque B .....	35
<b>Tabela 18</b> Comparação de melhor algoritmo para o Parque C .....	36

# Tabela de Símbolos e Siglas

EMQ – Erro Médio Quadrático

EPMA – Erro Percentual Médio Absoluto

MLP – Multilayer Perceptron

RC – Reservoir Computing

RNA - Redes Neurais Artificiais

---

# Capítulo 1 – Introdução

Este capítulo encontra-se dividido em três seções. A primeira seção trata do problema encontrado e os motivos de sua importância. Na segunda seção são apresentados os principais objetivos e metas. A última seção contém a estrutura da monografia.

## 1.1 Caracterização do Problema

A demanda por energia elétrica vem aumentando muito ao longo dos anos. Atualmente, existem diversas formas de obtenção de energia, porém as formas que não agredem o meio ambiente, como vento e sol, estão se tornando um foco bastante promissor.

Entre os anos de 2015 e 2024, a população brasileira aumentará em torno de 14 milhões de habitantes [6]. Quando a principal fonte de energia brasileira, a água, entra em escassez, o fornecimento se apara nas usinas termelétricas, que utilizam a queima de combustíveis fósseis para gerar energia.

Para contornar essa situação, o Brasil atualmente possui, como fontes alternativas de energia como as fazendas eólicas, que são a 3ª maior fonte de energia do país atualmente, usinas nucleares e estações solares [4]. E são nas fontes alternativas, consideradas energias limpas, que o Brasil precisa investir. A estimativa é de que a capacidade de geração eólica chegue a 24 mil MW, onde 21 mil serão oriundos do Nordeste [11].

A partir disso, os resultados são inúmeros, como: preservação do meio ambiente, gestão eficiente da geração de energia, aumento da produção energética no país e fim de aumento de custo por falta de água nas usinas hidrelétricas.

Atualmente, vários modelos tentam melhorar a previsão, isto é, diminuir o erro, utilizando series temporais. A maioria deles, como o de Box–Jenkins, possui uma grande desvantagem, pois eles só funcionam bem para problemas lineares e, geralmente, os modelos para previsão são não lineares [13].

---

Dessa forma, uma das maiores dificuldades encontradas acerca da precisão da previsão da energia eólica é fornecer uma previsão capaz de lidar com situações extremas, isto é, trabalhar situações inerentes ao usuário final, no qual não precise interferir no processo para obter um resultado final. Para que isso ocorra, é necessário o estudo e entendimento desses eventos extremos, os quais estão ligados a grandes desvios de geração de energia em relação ao que era esperado [10]. Além disso, a demanda de energia eólica é altamente incerta e quanto mais cedo esses eventos são previstos, mais eficazes são os procedimentos aplicáveis [4].

A solução trabalhada determinará previamente, e com melhor precisão possível, a quantidade de energia gerada em cada parque eólico. Além disso, a solução tem sua importância no planejamento, na melhor utilização dos recursos disponíveis de energia e ajuda a diminuir a incerteza na previsão da geração de energia futura. Para isto, serão comparados com diferentes estudos e métodos de testes já conhecidos na área estatística para garantir a relevância dos resultados obtidos pelas duas redes, Reservoir Computing e Multi Layer Perceptron.

## **1.2 Objetivos**

### **1.2.1 Objetivos Gerais**

Realizar a previsão de energia eólica gerada nos Parques A, B e C, utilizando o Reservoir Computing (RC) e a Multi Layer Perceptron (MLP). E contribuir para o estado da arte, proporcionando melhores resultados de previsões.

### **1.2.2 Objetivos Específicos**

1. Definir o melhor modelo de previsão
2. Selecionar as variáveis mais importantes no processo de previsão
3. Comparar os resultados entre as redes MLP e RC

---

## 1.3 Estrutura da Monografia

A monografia segue a seguinte estrutura:

- Capítulo 1 – Introdução: este capítulo contém as principais motivações para a pesquisa apontando o problema e objetivos a serem alcançados.
- Capítulo 2 – Fundamentação Teórica: este capítulo traz todo conhecimento relevante para o entendimento do problema e solução. Foram utilizadas as redes RC e MLP.
- Capítulo 3 – Metodologia: neste capítulo são encontrados os procedimentos utilizados para realização desse trabalho, isto é, aquisição da base de dados, tratamento dos dados e algoritmos utilizados.
- Capítulo 4 – Resultados: serão apresentados os resultados obtidos após as simulações do RC e MLP e os testes estatísticos utilizados.
- Capítulo 5 – Considerações finais: neste capítulo será apresentada a conclusão e junto a ela os possíveis trabalhos futuros.

---

# Capítulo 2 – Fundamentação

## Teórica

Este capítulo abrange o conteúdo necessário para a compreensão do problema e das soluções trabalhadas nesta monografia. Na seção 2.1 é descrita a produção de energia eólica brasileira. Na seção 2.2 estão os conceitos de redes neurais artificiais. Na seção 2.3 e 2.4 serão abordadas as arquiteturas MLP e RC, as quais foram utilizadas neste trabalho.

### 2.1 Energia Eólica

#### 2.1.1 Conceitos

Ventos são deslocamentos de ar causados pela diferença de pressões atmosféricas ao longo da superfície terrestre, migrando de zonas de alta pressão para zonas de baixa pressão. Turbinas eólicas têm como função aproveitar a energia existente no deslocamento de massas de ar e converter em energia cinética de rotação [2].

Uma vez que para calcular a energia cinética é necessário multiplicar a massa pela velocidade do vento ao quadrado e dividir o resultado por 2, fica claro que o principal fator dessa equação é a velocidade do vento.

$$K = \frac{mv^2}{2} \quad (2,1)$$

O movimento dos ventos na atmosfera ocorre de forma turbulenta, pois existem vários fatores que podem influenciar sua trajetória, tornando em algumas regiões o vento mais rápido, em outras mais forte ou as vezes constante. Por ser uma variável com grande irregularidade, não pode ser descrita de forma determinística, assim, faz-se necessário uma abordagem estatística [13].

Para suavizar o problema, a velocidade do vento é calculada através da variação de um valor médio das velocidades dos ventos em determinado período.

---

Outros fatores que influenciam na velocidade do vento:

- **Rugosidade:** É o atrito do vento com a superfície da Terra. Por isso, quanto mais obstáculos, maior será a rugosidade da região. Uma Região montanhosa, portanto, possui alta rugosidade, e os oceanos, por sua vez, possuem uma rugosidade muito baixa. Dessa forma, quanto maior a altura, menos rugosidade existirá no ambiente. É por isso que normalmente as medições são realizadas a partir de 25 metros e podem chegar aos 100 metros [17].
- **Outros:** Condições climáticas locais como monções, densidade do ar e relevo, pois podem causar efeito de aceleração ou desaceleração no escoamento do ar.

### 2.1.2 Turbina Eólica

Turbinas eólicas, ou aerogeradores, utilizam o vento para gerar energia elétrica. Para que isso ocorra, são necessárias grandes pás ligadas a turbina localizada na parte superior da torre de sustentação para transferir, através do movimento de rotação, parte da energia cinética do vento para o gerador.

As turbinas podem ser de eixo vertical [Figura 1], ou seja, as pás giram em um eixo paralelo a direção do vento. A principal vantagem desse modelo de turbina é o baixo custo de manutenção, uma vez que elas não trabalham com grande quantidade de energia cinética de vento, pois são instaladas próximas ao solo, e não precisam de sistemas que as façam girar seu eixo para conseguir capturar maiores quantidades de vento.

Existem também as turbinas de eixo horizontal [Figura 2], nesse caso as pás giram na direção perpendicular ao vento. Logo, esse modelo se faz mais eficaz quando se trata de conseguir acesso a maiores quantidades de vento e, assim, conseguindo gerar mais energia elétrica. Porém, existe uma desvantagem que se faz necessária para esse tipo de turbina: um sistema de rotação que ajusta a posição das pás com a direção do vento, ou seja, a existência desse sistema de rotação garante o desvio da turbina a depender da direção do vento, porém à custa de gasto de energia.

---

**Figura 1. Aerogeradores de Eixo Horizontal**



[Fonte: <http://atarde.uol.com.br/economia/noticias/1636493-bahia-tera-mais-30-novos-projetos-de-energia-solar-e-eolica-premium>]

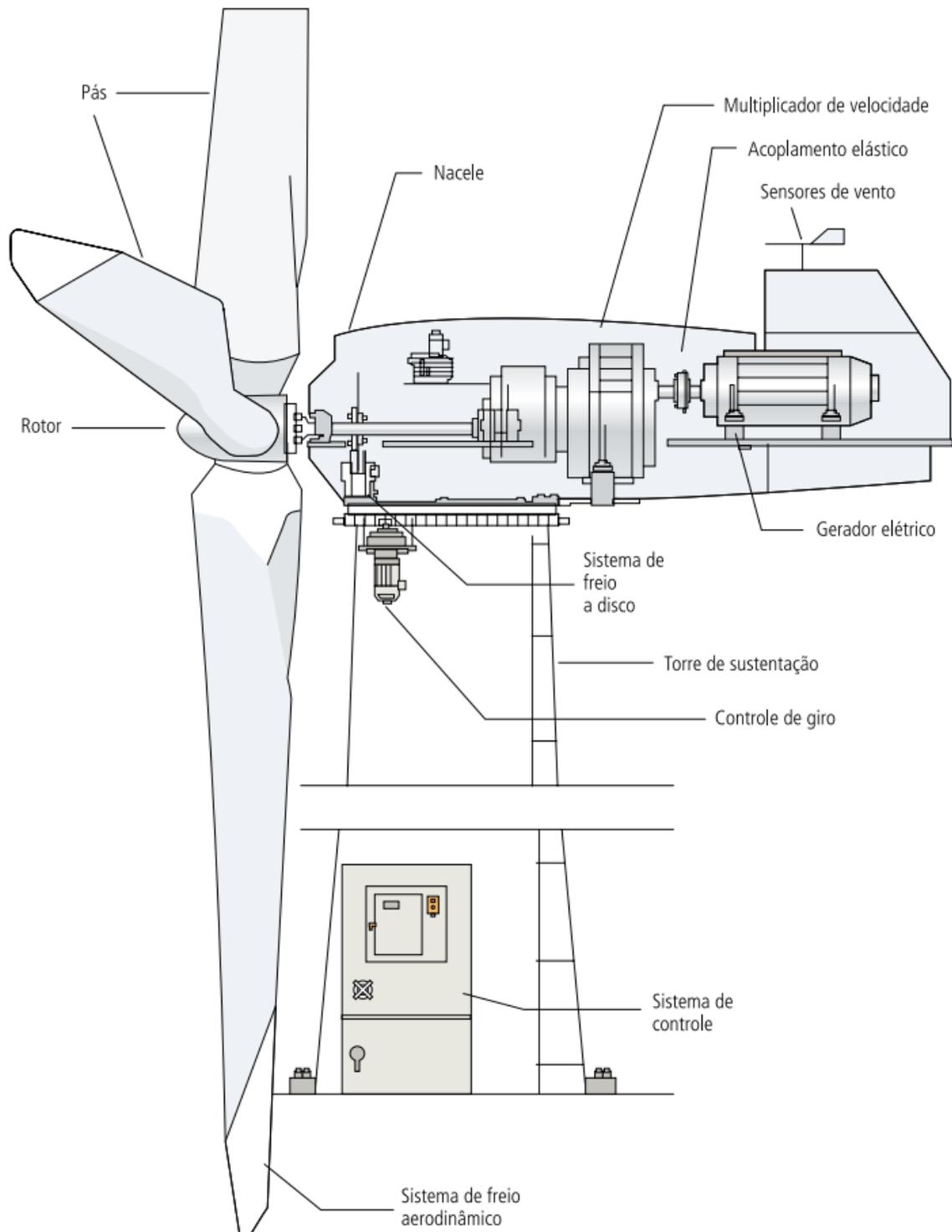
**Figura 2. Aerogerador de Eixo Vertical**



[Fonte: <https://evolucaoalp.wordpress.com/>]

Os aerogeradores de eixo horizontal, os mais usados e o foco deste trabalho, são compostos das seguintes partes:

**Figura 3.** Componentes de uma turbina de eixo horizontal



[Fonte: [http://www2.aneel.gov.br/aplicacoes/atlas/pdf/06-energia\\_eolica\(3\).pdf](http://www2.aneel.gov.br/aplicacoes/atlas/pdf/06-energia_eolica(3).pdf)]

---

Entretanto, o aproveitamento da relação energia cinética do vento e energia elétrica não é de 100%. As pás do aerogerador não ocupam toda a área por onde o vento passa, isto é, elas não impedem o vento, pois assim nunca iriam girar. Por outro lado, a intenção é deixar a massa de ar entrar em contato com as pás para que possam iniciar o processo rotacional responsável pela geração de energia.

Segundo o rendimento de Albert Betz [15], a geração de energia por um aerogerador, no melhor cenário, poderia chegar a 59,3% de aproveitamento da energia cinética do vento. Porém, a realidade mostra um aproveitamento em torno de 30% a 35% da energia cinética, que é convertida em energia mecânica.

Existe outra variável muito conhecida para a predição, que deve ser estudada antes das instalações das turbinas em um parque eólico. A direção do vento na região é um fator muito importante para identificar situações de rajadas de vento, mais conhecidas como rampas de vento, as quais podem causar turbulência. A partir disso, surge outro problema derivado da distribuição dos aerogerados em um parque eólico, a chamada sombra. Após atravessar a primeira turbina do parque, o vento sofre uma interferência que causa uma diminuição de seu fluxo, ocasionando a diminuição na geração de energia das demais turbinas. Portanto, é importante avaliar essas variáveis para obter o espaçamento ideal entre as turbinas, considerando o custo do terreno e das ligações à rede elétrica.

### 2.1.3 Sazonalidade

A variabilidade climática ou sazonalidade possui também uma grande influência na circulação geral e no comportamento do vento, isto é, a sazonalidade é uma das muitas variáveis que podem influenciar a velocidade do vento.

A sazonalidade representada utilizando as funções seno e cosseno, descritas nas equações abaixo, como duas novas entradas do sistema.

$$\cos\left(2\pi \times \frac{N}{T}\right) \quad (2,2)$$

$$\sin\left(2\pi \times \frac{N}{T}\right) \quad (2,3)$$

---

Em que:

- $N$  = varia de 1 até  $T$ ;
- $T$  = representa o número de pontos amostrais

#### **2.1.4 Modelos de previsão**

O crescimento da participação da energia eólica como principal fonte energética mundial, pede que ocorra uma alteração no planejamento, na forma de operação e no mercado de energia. Por isso, a previsão precisa ser confiável e segura. Como a força motriz desse sistema, o vento, é uma variável inconstante, faz-se necessário realizar a previsão para garantir o aviso prévio de sua escassez. Tudo isso aumenta o tempo e melhora o poder de decisão de parques eólicos tomarem devidas providencias para solucionar o problema, além de criar espaços para uma programação e um equilíbrio em geração de energia [3].

As previsões podem ser classificadas de acordo com a escala de tempo [12]:

- Curtíssimo prazo: São as previsões de poucos minutos à frente utilizadas principalmente para controlar a potência ativa do gerador.
- Curto prazo: De 1 a 72 horas. Utilizadas para planejar a operação e a distribuição da energia.
- Médio prazo: De 5 a 7 dias. Utilizadas na manutenção das plantas eólicas.
- Longo prazo: Acima de 7 dias. Utilizadas em planejamento de recursos energéticos e em leilões de energia

Os modelos de previsão mais usados são os modelos físicos e os modelos estatísticos/inteligência artificial. Os modelos físicos realizam as previsões com base nos dados meteorológicos e físicos e utilizam métodos numéricos para previsão. Demandam grande esforço computacional e requerem um número elevado de dados, sendo recomendados para previsões de médio e longo prazo. Os modelos estatísticos e de inteligência artificial são mais simples e mais rápidos na realização das previsões. Contudo, possuem tendência à diminuição na precisão da previsão conforme o aumento da escala de tempo, por essa razão são mais recomendados para previsões de curtíssimo e curto prazo [16]. Uma das técnicas mais comumente

---

usadas na construção dos modelos de inteligência artificial são as Redes Neurais Artificiais [6][10].

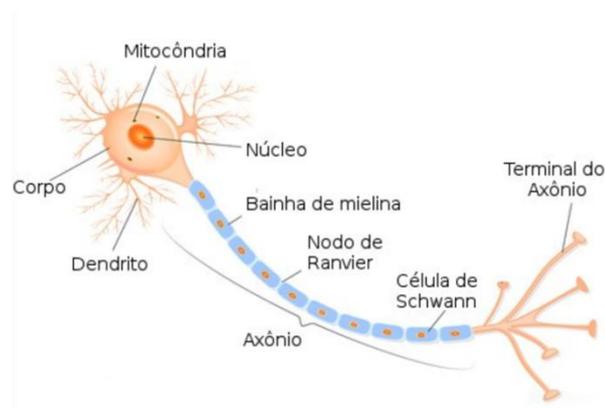
## 2.2 Redes Neurais Artificiais

As Redes Neurais Artificiais são uma das técnicas de aprendizagem inteligente de máquinas. Elas compartilham a propriedade de aprender baseadas em exemplos para então obter o poder de generalizar possíveis situações nas quais ainda não aprenderam. Para sua criação os cientistas se inspiraram no cérebro humano, que possui a capacidade de processamento paralelo e de obter conhecimento através de experiências. Portanto, esses modelos têm como base as redes neurais biológicas [18].

Diante disso, faz-se necessário um breve entendimento sobre o comportamento e características do neurônio biológico [Figura 4] para facilitar o aprendizado das técnicas e arquiteturas de RNA.

O neurônio é uma célula que faz parte do sistema nervoso humano e possui os dendritos, que conduzem os estímulos vindos de outras células, o corpo celular, o axônio, capaz de transmitir informações para outras células, e por fim o terminal do axônio, também conhecido como telodendros.

**Figura 4.** Neurônio Biológico

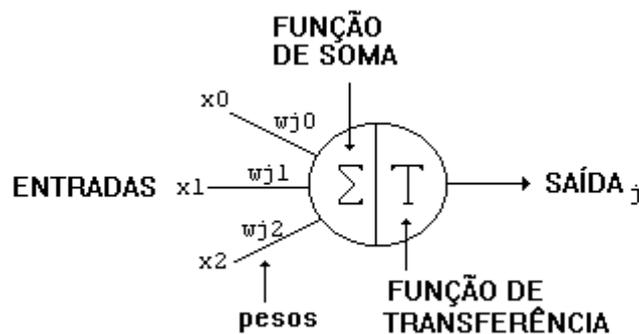


[Fonte: <http://amavc.com.br/conheca-o-funcionamento-de-um-neuronio/>]

Existe uma região denominada de sinapse, onde a comunicação entre os neurônios acontece. “Elas ocorrem no contato das terminações nervosas chamadas axônios, usualmente com os dendritos de outro neurônio” [5].

Segundo Valença (2012), o neurônio artificial seria um “conjunto de elementos de processamento interconectados que geralmente encontram-se em camadas”[16]. O primeiro modelo de neurônio artificial foi apresentado pelos cientistas McCulloch e Pitts [Figura 5] em 1943 [16]. Eles foram os difusores da lei do tudo ou nada, que determina a liberação do impulso nervoso entre neurônios. Além disso, para que essa transmissão ocorra o impulso deve ser maior que um valor mínimo, denominado limiar excitatório.

**Figura 5.** Modelo de Neurônio de MacCulloch e Pitts



[fonte: <http://redesneuraisartificiais.blogspot.com.br/2010/10/o-primeiro-modelo-de-um-neuronio-criado.html>]

O modelo proposto pela dupla é formado por um conjunto de entrada ( $x_0, x_1, x_2, \dots, x_n$ ), pelos pesos entre a camada de entrada e saída ( $W_{j0}, W_{j1}, W_{j2}, \dots, W_{jn}$ ), por uma regra de propagação ( $\Sigma$ ) (2,4) e uma função de ativação ou transferência  $f(net_i)$  (2,5) (2,6).

A regra de propagação é dada por:

$$net_i = \sum_{j=1}^n W_{ij} X_j - \theta \quad (2,4)$$

---

A função de ativação é a função degrau dada por:

$$f(net_i) = 1, \forall net_i \geq 0 \quad (2,5)$$

$$f(net_i) = 0, \forall net_i < 0 \quad (2,6)$$

Esse foi o início que as redes neurais precisavam. Logo em seguida surgiu a nova arquitetura chamada Perceptron, a qual apareceu em meados dos anos 50 e teve como principal diferencial o aprendizado supervisionado ou com professor, no qual a rede já conhece a sua saída, para ajustar os pesos. Logo depois, já no começo da década de 60, surgiram as novas funções de ativação que, segundo Valença (2012), são mais utilizadas dentro dos cenários das redes neurais: a função hiperbólica (2,7), a função sigmoide logística (2,8) e a função linear (2,9).

$$f(net_i) = \frac{e^{net_i} - e^{-net_i}}{e^{net_i} + e^{-net_i}} \quad (2,7)$$

$$f(net_i) = \frac{1}{1 + e^{-net_i}} \quad (2,8)$$

$$f(net_i) = net_i \quad (2,9)$$

Para a construção da rede neural artificial é aplicada uma ou mais camadas contendo essas informações. Com o estudo dessas novas técnicas foi possível, então, chegar ao entendimento de algumas das redes muito utilizadas atualmente que são a MLP e RC, as quais serão abordadas na sequência.

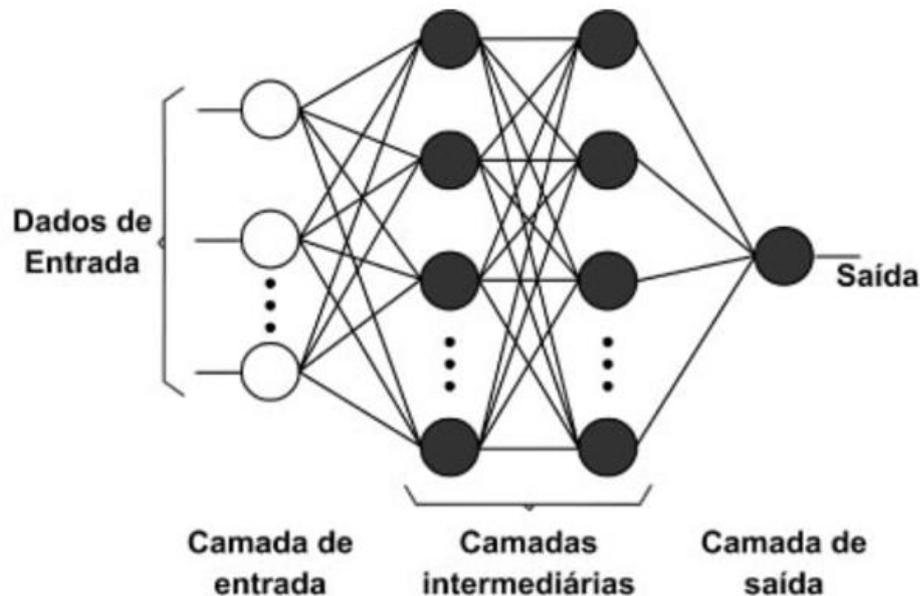
## 2.3 Multilayer Perceptron

As redes Multi-Layer Perceptron (MLP) são compostas por uma camada de entrada, uma ou mais camadas intermediárias, ou camadas escondidas, e uma camada de saída (Figura 6). O surgimento de múltiplas camadas concedeu a rede MLP o poder de solucionar problemas que não linearmente separáveis, isto é, a MLP é capaz de aproximar qualquer função, situação que antes era impossível de ser

---

realizada por redes de apenas uma camada escondida [1]. Está é uma das principais características dessa rede e o principal ponto de divergência entre seu “pai” o Perceptron e “mãe” a Adaline.

**Figura 6.** Arquitetura da Rede MLP com 2 camadas escondidas



[Fonte: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1806-11172011000100009](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1806-11172011000100009)]

Em cada camada é possível existir um ou mais neurônios e cada um deles possui uma função de ativação, sendo a sigmoide logística, equação descrita abaixo, a utilizada nesse trabalho. É uma função real cujo domínio é o conjunto dos números reais, a derivada é sempre positiva e sua faixa de valores possui uma maior variância longe das extremidades da função.

$$y = f(net) = \frac{1}{1 + e^{-net}} \quad (2,10)$$

A MLP tem a capacidade de aprender, fato que ocorre em toda rede neural. O treinamento da rede é o período em que ocorre o aprendizado sobre o problema. É nesse momento que os pesos dos neurônios são alterados, de acordo com algum algoritmo de treinamento, dando importância para as variáveis de maior valia para resolver o problema.

---

Na literatura, um algoritmo de treinamento muito bem sucedido e bastante difundido na atualidade é o *backpropagation*, desenvolvido por Paul Werbos em 1974 e apresentado na seção seguinte [19]. A grande dificuldade do treinamento de uma MLP está em calcular os pesos ótimos para a camada intermediária ou escondida.

### 2.3.1 Algoritmo *Backpropagation*

O *backpropagation* é o algoritmo mais utilizado em treinamento de redes neurais. Para treinar, o algoritmo ajusta os pesos da rede utilizando a ordem inversa da propagação do sinal. O algoritmo possui 2 fases: a primeira fase é chamada *forward*, caracterizada por propagar o sinal no sentido da camada de entrada, passando pela camada escondida até a camada de saída. No final, é calculado um erro a partir da resposta obtida pela propagação do sinal e outro erro que já era o esperado, o qual recebe o nome de erro médio quadrático (EMQ).

$$e_i(n) = \frac{1}{2} (d_i - y_i)^2 \quad (2,11)$$

- $d_i$ : O valor desejado
- $y_i$ : O valor estimado pela rede

Esse erro será calculado para cada neurônio da camada de saída e para cada conjunto de entrada oferecido para rede.

Já a segunda fase, mais conhecida como *backward*, ocorre a retropropagação dos erros, o que resultará no ajuste dos pesos, isto é, as alterações dos pesos levaram em conta o erro do ciclo anterior para minimizar o erro do próximo ciclo que a rede irá processar. Para calcular os pesos, a rede segue a seguinte formula (2,12):

$$w_{ij}^m(novo) = w_{ij}^m(antigo) + \alpha \delta_i^m f^{m-1}(net_i^{m-1}) \quad (2,12)$$

- $w_{ij}^m(novo)$ : é o valor do novo peso calculado no ciclo
- $w_{ij}^m(antigo)$ : É o valor do peso calculado no ciclo anterior

- 
- $\alpha$ : É a taxa de aprendizagem
  - $\delta_i^m$ : representa a sensibilidade na camada de saída
  - $f^{m-1}(net_i^{m-1})$ : foram os sinais emitidos da camada anterior

Para calcular a sensibilidade, é utilizada a seguinte equação:

$$\delta_i^m = f^{m'}(net_i^m) \cdot e_i(n) \quad (2,13)$$

- $f^{m'}(net_i^m)$ : É a derivada da função de ativação da camada de saída
- $e_i(n)$ : É o erro

Para a camada intermediária é utilizada a fórmula a seguir:

$$\delta_i^{m-1} = f^{m-1'}(net_i^{m-1}) \sum_{j=1}^n w_{ij}^m \delta_j^m \quad (2,14)$$

- $f^{m-1'}(net_i^{m-1})$ : É a derivada da função de ativação da camada anterior
- $n$ : É o número de neurônios
- $w_{ij}^m$ : São os pesos para a sinapse  $ij$  na camada  $m$
- $\delta_j^m$ : É o valor da sensibilidade da camada imediatamente a frente da camada atual.

O algoritmo *backpropagation* irá realizar esses passos pela quantidade de vezes definida na sua inicialização. Porém, a quantidade de repetições do algoritmo é sempre alta, uma vez que não se sabe a quantidade certa para uma base de entrada desconhecida. Dessa forma, o algoritmo pode chegar ao seu melhor resultado tanto no 10º quando no 100º, por exemplo. Isso pode gerar um fenômeno conhecido como *overfitting*, isto é, o treinamento excessivo da rede. Utilizando o exemplo citado acima, caso o melhor resultado da rede alcance o melhor resultado no 10º ciclo, a rede iria continuar em treinamento, mas sem obter melhores

---

resultados, isto pode levar a perda da capacidade de generalização, pois a rede iria ficar muito refinada para certo tipo de entrada. Por outro lado, se a rede não receber treinamento suficiente, ela pode não ter uma boa capacidade de generalização. Portanto, faz-se necessário um meio termo, que pode ser alcançado pela validação cruzada, a qual será utilizada como critério de parada do treinamento.

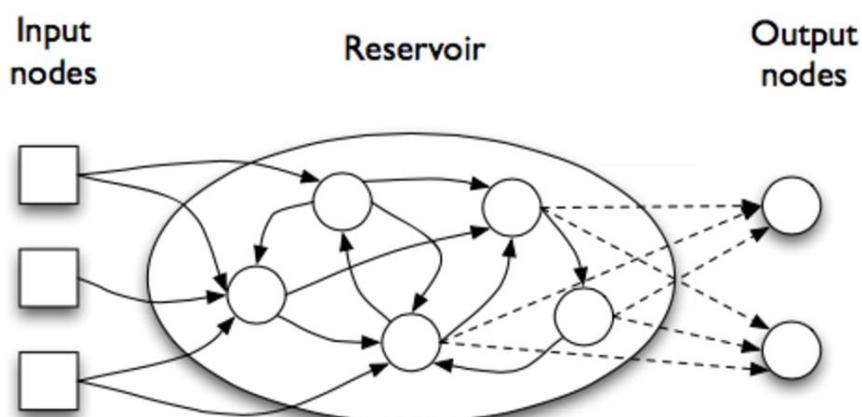
Para realizar os treinamentos a base de dados deve ser dividida em três partes, onde a primeira parte de 50% destinados para treinamento, à segunda parte de 25% para validação cruzada e os últimos 25% para realização dos testes.

Para definir a quantidade de ciclos de treinamento da base de dados e a validação cruzada utiliza-se a função erro que ao atingir o seu mínimo deve ser o ponto de parada do treinamento da rede. Para finalizar, o conjunto de testes é separado para validar o desempenho modelo construído.

## 2.4 Reservoir Computing

O RC faz parte das redes neurais que possuem retroalimentação, isto é, um *feedback* representado por ligações entre neurônios da camada escondida e/ou de saída. Dessa forma, essa rede é capaz de simular uma memória que tem o poder de analisar valores presentes e futuros juntos, gerando uma grande dificuldade em seu treinamento. Por outro lado, uma vez treinada, a RC se torna uma ótima ferramenta quando se trata de previsão utilizando dados temporais [1].

**Figura 7.** Arquitetura da rede recorrente RC



---

[Fonte:[https://www.researchgate.net/publication/307578398\\_Reservoir\\_Computing\\_com\\_Hierarquia\\_para\\_Previsao\\_de\\_Vazoes\\_Medias\\_Diarias](https://www.researchgate.net/publication/307578398_Reservoir_Computing_com_Hierarquia_para_Previsao_de_Vazoes_Medias_Diarias)]

A figura acima descreve a rede neural recorrente utilizando o novo paradigma chamado RC, o qual foi criado por David Verstraeten [16]. Nesse esquema, as linhas sólidas representam conexões com valores aleatórios. Já as linhas pontilhadas mostram as ligações que iram receber treinamento [Figura 7].

Ao analisar a Figura 7, pode ser determina a taxa de conectividade entre os neurônios, a qual é diretamente proporcional a quantidade de neurônios ligados entre si. Essa taxa é necessária por causa das ligações entre neurônios que esse tipo de modelo possui.

Para executar o *backpropagation* é necessário seguir os seguintes passos

- Determinar a quantidade de neurônios de entrada;
- Determinar a quantidade de neurônios no reservatório;
- Determinar os pesos da rede aleatoriamente entre -1 e 1;
- Dividir a base em 3 conjuntos: treinamento, validação cruzada e testes;
- Utilizar a técnica de Warm Up;
- Utilizar a validação cruzada (*cross validation*) como critério de parada do treinamento.

A primeira etapa do processo é a atualização dos neurônios do reservatório baseando-se nos dados de treinamento. Neste momento, os pesos ainda estão como no início em que foram preenchidos, ou seja, sem nenhuma alteração.

Para atualizar os neurônios do reservatório e dar continuidade ao processo, deve ser utilizada a seguinte equação:

$$x(t + 1) = f(W_{in} u(t) + W_{res} x(t)) \quad (2,15)$$

**Tabela 1.** Descrição da Equação de Alteração dos Neurônios do Reservatório

Variáveis da Formula	Significado
$t$	Representa o tempo

---

$W_{in} u(t)$	Matriz que armazena o produto de entradas versus pesos no instante $t$
$W_{res} x(t)$	Matriz de pesos do reservatório no instante $t$
$x(t + 1)$	Estado de um neurônio no reservatório no instante seguinte a $t$

[Fonte: elaboração própria]

Dessa forma, os únicos pesos que devem ser ajustados durante esse momento são os do reservatório e da camada de saída. Agora restando apenas aplicar um método de ajuste desses pesos. Uma boa técnica que realiza o mesmo processo de muitas outras, porém com uma velocidade muito maior, é a pseudo inversa de Moore-Penrose. Baseada no determinante da matriz formada pelos pesos. O fato dessa matriz de pesos não ser quadrada, faz com que o preenchimento com valores iguais a zero, até que seu estado mude e a matriz passe a ser quadrada. É desse fato curioso que surgiu o seu nome pseudo inversa, pois no começo ela não é quadrada, gerando uma falsa impressão de que aquela matriz não pode ser utilizada, porém pode-se reajustar este resultado da matriz e ela passa a ser quadrada.

---

# Capítulo 3 – Metodologia

Neste capítulo será abordado com detalhes o processo realizado na pesquisa para solucionar o problema. Na seção 3.1 serão descritas as bases de dados, na seção 3.2 será exposto todo o pré-processamento dos dados. Já na seção 3.3 e 3.4 serão analisadas as entradas das redes MLP e RC e por último a seção 3.5 aborda todos os testes estatísticos utilizados para inferir, com um teor estatístico, a melhor solução do problema.

## 3.1 Base de Dados

As três bases de dados utilizadas nos experimentos foram retiradas da ONS (Operador Nacional de Sistema Elétrico), que é o órgão responsável por gerir e controlar as operações das instalações de geração e transmissão de energia elétrica em todo o país.

Dentre toda a base foram selecionadas as seguintes variáveis: potência gerada, velocidade do vento e direção do vento. Porém, durante o processo de análise dos dados, foi aferido que os dados de velocidade e direção do vento não estavam em boa quantidade e qualidade, contendo diversos espaços vazios que poderiam comprometer o correto funcionamento da rede. Por isso, única variável elegível para ser utilizada nos experimentos foi a potência do vento.

Esses valores de potência foram utilizados no formato de 30 em 30 minutos e coletados no período de 01 de dezembro de 2011 até 31 de julho de 2012 para todas as três bases Parque A, Parque B e Parque C. Os parques receberam esses nomes em formato de letras do alfabeto, pois existe um sigilo de informação a respeito desses dados que foi acordado com o órgão cedente (ONS).

Para cada Parque ficou um total de 48 entradas possíveis, equivalentes a 1 dia passado para serem selecionadas com a correlação e preverem 48 saídas, as quais correspondem a 1 dia de previsão a frente, uma vez que são a previsão futura também é gerada de 30 em 30 minutos.

---

## 3.2 Pré-Processamento dos dados

A primeira etapa no pré-processamento dos dados é a sua normalização. A normalização é uma técnica que procura estabelecer um determinado limite de valores, isto é, dados com valores muito altos vão passar a ter uma discrepância bem menor em relação a valores menores.

**Tabela 2.** Normalização de uma amostra dos dados entre [0.15-0.85]

Valores Reais	Valores Normalizados
15.025	0.35
13,57	0.33
12.263	0.31
11.343	0.30
9.343	0.27
8.392	0.26
10.571	0.29
13.087	0.32
10.486	0.28
10.395	0.29
10.128	0.28

[Fonte: elaboração própria]

Segundo a literatura, a base de dados deve ser normalizada de certa forma, que seus valores sejam alterados para os limites da função de ativação. Neste trabalho será abordada a sigmoide como função de ativação, a qual está limitada entre 0 e 1. Porém, normalizar os dados entre 0 e 1 pode causar algumas complicações, como estagnar o treinamento caso a derivada da função de ativação chegue em um mínimo ou máximo local. Por esse motivo, os dados são normalizados entre valores próximos aos limites da função, geralmente entre 0.1 ou 0.15 até 0.9 ou 0.85. Para a função de ativação tangente hiperbólica, que é limitada pelo intervalo de -1 a 1, os dados devem ser normalizados, também, próximos aos seus limites.

Uma das opções de normalização mais empregadas para problemas que tratam de séries temporais é a transformação linear.

$$y = (b - a) \frac{(x_i - x_{min})}{(x_{max} - x_{min})} + a \quad (3,1)$$

**Tabela 3.** Descrição da Fórmula da Normalização

Variáveis da Formula	Significado
$y$	Valor da variável após a normalização
$a$	Limites mínimo desejado para a normalização da base [0,15]
$b$	Limite máximo desejado para a normalização da base [0,85]
$x_{min}$	Menor valor encontrado na base
$x_{max}$	Maior valor encontrado na base
$x_i$	Valor da base na posição $i$ que se deseja normalizar

[Fonte: elaboração própria]

Após a normalização será aplicada a técnica conhecida como defasagem, que simula valores passados, defasando os valores atuais, para determinar valores futuros. Segue um exemplo de defasagem a partir dos dados reais:

---

**Tabela 4.** Base de Dados Defasada até (t-4)

Dias	t-4	t-3	t-2	t-1	t
1					0.35
2				0.35	0.33
3			0.35	0.33	0.31
4		0.35	0.33	0.31	0.30
5	0.35	0.33	0.31	0.30	0.27
6	0.33	0.31	0.30	0.27	0.26
7	0.31	0.30	0.27	0.26	0.29
8	0.30	0.27	0.26	0.29	0.32
9	0.27	0.26	0.29	0.32	0.28
10	0.26	0.29	0.32	0.28	0.29
11	0.29	0.32	0.28	0.29	0.28
12	0.32	0.28	0.29	0.28	
13	0.28	0.29	0.28		
14	0.29	0.28			
15	0.28				

[Fonte: elaboração própria]

A tabela acima, representa a técnica de defasagem de dados com mais detalhes. Em primeiro momento só existe os valores da coluna t, porém para treinar a rede neural precisamos defasar essa coluna para a esquerda, simulando uma “volta no tempo”. Dessa forma, é possível obter (t-1), (t-2), (t-3) e (t-4). Assim, podemos utilizar essas 4 defasagens de t para representar a entrada da rede neural e o t (original) seria nossa saída esperada, fazendo com que o treinamento supervisionado seja bem sucedido.

Entretanto, um problema encontrado nesse processo é saber após a defasagem qual seria dentre as entradas as que teriam maior influencia sobre a previsão futura. Para isso, foi utilizada a correlação entre os dados como técnica de seleção das entradas com maior relevancia para o treinamento.

**Tabela 5.** Correlação entre variáveis da Tabela 3.

	<i>Coluna 1</i>	<i>Coluna 2</i>	<i>Coluna 3</i>	<i>Coluna 4</i>	<i>Coluna 5</i>
<i>Coluna 1</i>	1				
<i>Coluna 2</i>	0.65	1			
<i>Coluna 3</i>	7.55E-17	0.33	1		
<i>Coluna 4</i>	-0.23	-0.34	0.11	1	
<i>Coluna 5</i>	-0.39	-0.58	-0.61	-0.006	1

[Fonte: elaboração própria]

Analisando a tabela 5, percebe-se que na quinta linha estão localizados os valores de correlação das entradas em relação ao valor da saída. A diagonal dessa matriz é composta por valores iguais a 1, pois a correlação de uma variável em relação a ela mesma é sempre igual a 1, sendo este o maior valor de correlação possível.

Já a correlação entre a saída (linha 5) e o terceiro neurônio (coluna 3) é negativa, isso significa que ele é descartado do experimento, mas isso não significa dizer que por ser negativo o valor seria uma opção ruim. O sinal negativo na correlação não quer dizer que o valor encontrado seja ruim, apenas que quanto mais próximo do -1 o valor negativo for, melhor a sua correlação inversa será.

Para os valores positivos, que serão os utilizados nos experimentos, inicialmente foram escolhidos os valores de correlação acima de 0,6. Porém, foram obtidos melhores resultados ao aceitar valores acima de 0,3. Assim, todas as colunas com a correlação entre 0,3 e menores que 1, foram escolhidas para os experimentos.

Após obter as variáveis de entrada, devem ser realizados os experimentos tanto ta MLP quanto da RC.

---

## 3.3 Entradas das Redes

Nessa seção, serão apresentadas as entradas de ambas as redes MLP e RC. Definindo e explicando o motivo de cada escolha.

As redes neurais podem analisar os dados de entrada com diferentes configurações, basta alterar algumas de suas variáveis como a função de ativação ou a quantidade de neurônios na camada escondida que o resultado final da rede já será alterado para melhor ou pior. Por isso, faz-se necessário realizar diversos experimentos para encontrar o melhor resultado para suas variáveis.

Para este trabalho, foram analisados dados da média da potência eólica gerada de 30 em 30 minutos de 3 parques eólicos diferentes e também dados da sazonalidade, segue descrição na seção 3.3.1.

### 3.3.1 Entradas da Rede MLP

Para a rede MLP, o número de neurônios na camada de entrada depende diretamente da correlação das variáveis da entrada da base, a qual foi descrita na seção 3.2. Já a saída possui 48 valores, que correspondem a previsão de 1 dia a frente, visto que cada variável equivale a 30 minutos.

Para o Parque A, a correlação dos dados selecionou 19 entradas dentre as 48 analisadas, sendo aceitas correlações maiores que 0,3. Foram geradas 2 bases para rede MLP desse Parque, a primeira contendo as 19 variáveis e uma segunda contendo 21 variáveis, isto é, adicionando 2 variáveis para representação da sazonalidade, o seno e o cosseno.

O Parque B recebeu 20 entradas que também foram selecionadas pela correlação acima de 0,3. Foi criada uma segunda base com 22 entradas para a adição do seno e cosseno, respectivamente.

Já o Parque C ficou com 11 entradas cuja correlação foi maior que 0,3 e também foi criada uma segunda base para adicionar as variáveis de sazonalidade.

Ademais, a rede MLP depende de outros parâmetros como a função de ativação, quantidade de neurônios na camada escondida, validação cruzada, taxa de aprendizagem, momento e o número máximo de ciclos.

---

Diante disso, para o algoritmo de treinamento, foi utilizada a técnica do *backpropagation* - descrito na seção 2.3.1 - a função de ativação foi a sigmoide logística descrita pela equação (3,2).

$$f(net_i) = \frac{1}{1 + e^{-net_i}} \quad (3,2)$$

Por fim, foi escolhi a função linear, representada pela equação (3,3) para a camada de saída.

$$f(net_i) = net_i \quad (3,3)$$

### 3.3.2 Entradas da Rede RC

A RC depende, assim como a MLP, de um conjunto de dados para entrada. A quantidade de variáveis para cada experimento foi igual a MLP, porém para o RC foi necessário dividir cada base da MLP em outras 6. Das 6 subdivisões, 3 eram de entradas e 3 de saídas. Dentre as 3, a primeira era a entrada para o treinamento da RC, que possui 50% dos dados de entrada, já a segunda servia para calcular a validação cruzada e era equivalente a 25% dos outros 50% restantes. Então, os últimos 25% serviam como testes para validar o algoritmo depois de treinado.

Além disso, o algoritmo RC possui os seguintes campos como variáveis importantes na sua execução:

**Tabela 6.** Parâmetros do Algoritmo RC

Quantidade de neurônios no reservatório
Taxa de conectividade
Quantidade de ciclos de <i>Warm Up</i>
Função de ativação no reservatório
Função de ativação na camada de saída
Critério de parada de treinamento

[Fonte: elaboração própria]

Após definir essas variáveis, é possível iniciar o treinamento da rede RC. Como método de treinamento da RC para este trabalho foi utilizada a pseudo-inversa de Moore-Penrose, descrita na seção 2.4.

---

## 3.4 Configuração das redes

Nesta etapa, foram descritos as configurações das variáveis citadas na seção 3.3.

### 3.4.1 Configurações da Rede MLP

A rede neural MLP utilizada nos experimentos está implementada em Java. O algoritmo foi implementado por alunos da Escola Politécnica de Pernambuco no ambiente de desenvolvimento chamado Eclipse, a partir de estudos e pesquisas com orientação do professor Mêuser Jorge Silva Valença. Porém, foram geradas modificações ao código para se adaptar a realidade do problema em análise.

**Tabela 7.** Principais parâmetros utilizados para a rede MLP

PARÂMETROS	REDE MLP
Neurônios na camada de entrada	19 e 21 / 20 e 22 / 11 e 13
Neurônios na camada escondida	6
Neurônios na camada de saída	48
Algoritmo de treinamento	<i>backpropagation</i>
Forma de inicialização dos pesos da rede	Aleatória
Função de ativação	Sigmoide Logística
Máximo de ciclos	600
Experimentos	30
Taxa de aprendizado	0,8
Momentum	0,2

[Fonte: elaboração própria]

Como método de parada foi utilizado a validação cruzada, já descrito na seção 2.3. Após 30 simulações, já é possível considerar que as médias dos resultados pertencem a uma distribuição normal. Os neurônios da camada de entrada, descritos na Tabela 7, são 19, 20 e 11 que representam as melhores variáveis selecionadas de cada um dos parques A, B e C respectivamente. Já os valores 21, 22 e 13 são referentes às bases com a sazonalidade adicionada para serem feitas as comparações.

Como a base utilizada não possui dados faltantes e nem valores iguais à zero, a métrica de análise das técnicas foi o erro percentual médio absoluto (EPMA), que é calculado após cada rodada ou ciclo.

$$EPMA = \sum_{i=1}^n \frac{|d_i - y_i|}{d_i} * 100 \quad (3,4)$$

**Tabela 8.** Descrição da equação do EPMA

Variáveis da Formula	Significado
$d_i$	Saída desejada para a previsão $i$
$y_i$	Saída calculada
$n$	Número de Simulações

[Fonte: elaboração própria]

### 3.4.2 Configurações da Rede RC

O algoritmo do Reservoir Computing foi implementada em Java, também por alunos da Escola Politécnica de Pernambuco no ambiente de desenvolvimento chamado Eclipse, a partir de estudos e pesquisas com orientação do professor Mêuser Jorge Silva Valença.

Após estudo na literatura, foi escolhido um valor próximo para a quantidade de neurônios no reservatório, terminou sendo igual a 25. Isso não quer dizer que esta seja a melhor opção, portanto representa uma quantidade já utilizada em trabalhos passados, vide em TATYANA [20], e mostrou bons resultados.

**Tabela 9.** Principais parâmetros utilizados para a rede RC

PARÂMETROS	REDE RC
Número de neurônios na camada de entrada	19 e 21 / 20 e 22 / 11 e 13
Número de neurônios na camada de saída	48

Quantidade de neurônios no reservatório	25
Forma de inicialização dos pesos da rede	Aleatória
Taxa de conectividade	0,3 – 0,4
Quantidade de ciclos no <i>Warm up</i>	10
Função de ativação no Reservatório	Sigmoide Logística
Função de ativação na camada de saída	Linear
Taxa de aprendizado	0,8
Critério de parada de treinamento	<i>Cross validation</i>

[Fonte: elaboração própria]

Os dados se comportam da mesma forma que na MLP. A quantidade de ciclos para o *warm up* foi baseada em trabalhos anteriores que utilizaram RC para predição de situações cujos dados se retratavam de forma temporal [4][1].

A quantidade de experimentos é a mesma que na MLP, para assumir que os resultados fazem parte de uma distribuição normal e 48 saídas, pois o intuito é saber a previsão da potência 1 dia a frente.

## 3.5 Testes estatísticos

Após a finalização dos 30 ciclos de todos os experimentos, foram obtidos 30 erros para cada situação, que são os EPMA's. Esses erros servirão de entrada para testes estatísticos que confirmam, de forma matemática, os resultados encontrados.

Os testes foram realizados com o Software R [14], considerando o nível de significância igual a 5% (0,05), que é a porcentagem comumente utilizada na engenharia [15].

### 3.5.1 Teste de Wilcoxon

Este é um teste não paramétrico, isto é, não faz nenhuma hipótese em relação à distribuição de probabilidade da variável em análise. Este teste tem o

---

objetivo de verificar para amostras independentes, se as medianas podem ser consideradas originárias de certa população de dados.

O p-value, que já é a probabilidade dos 30 EPMS gerados pelos algoritmos, deve ser comparado com o nível de significância adotado ( $\alpha = 5\%$ ) e caso esse resultado seja menor que 0.05 ou 5%, significa dizer que o resultado encontra-se na região de rejeição, então pode-se considerar que as medianas são diferentes. Assim, sendo elas estatisticamente diferentes e o trabalho busca a melhor taxa de acerto, deve ser escolhido o resultado que possuir a menor mediana.

No entanto, se a comparação der maior do que 5% significa dizer que as medianas são consideradas estatisticamente iguais, assim tanto faz utilizar um ou outro modelo. Sendo assim, estatisticamente, não existe nenhum ganho significativo. Dessa forma, é feita a escolha do modelo mais simples, o menor custo computacional ou procurar outra vantagem para escolha, pois o erro nesse caso não é vantagem.

---

# Capítulo 4 – Resultados

Este capítulo tem como objetivo detalhar e expor todos os resultados encontrados durante os experimentos com as redes neurais MLP e RC na previsão da geração eólica um dia à frente. Acompanhando a metodologia descrita no capítulo 3, serão retratados os resultados para a base de dados obtida da ONS.

As seções 4.1 e 4.2 mostram os experimentos realizados com as respectivas redes MLP e RC para os respectivos parques A, B e C. Aplicando as comparações entre as melhores MLPs e RCs com sazonalidade e sem sazonalidade. Então após selecionar para cada parque a sua melhor MLP e melhor RC - baseando-se nas medianas geradas a partir dos 30 erros EPMA - a seção 4.3 realiza uma comparação dentre elas e apresenta qual obteve melhor eficiência para aquele parque em específico.

## 4.1 Experimentos com o MLP

Para os experimentos com a rede MLP foi comparado entre os três parques A, B e C qual possui o melhor cenário, sendo eles: o primeiro sem sazonalidade e o segundo com sazonalidade. Dessa forma, obtendo o melhor resultado para cada Parque.

Para a MLP, foi utilizada a melhor configuração encontrada na literatura (VALENÇA, 2009), a qual utiliza 6 neurônios na camada escondida.

### 4.1.1 Estudo de caso MLP para o Parque A

Para a previsão de um dia a frente de geração de potência eólica utilizando a MLP, neste parque, poderiam ter sido usados ambos os cenários, uma vez que a variável *p-value*, obtida pelo teste de Wilcoxon, foi superior a 5%, demonstrando que os modelos são estatisticamente iguais e não existe um ganho significativo. Portanto, a escolha do modelo deve levar em conta outros fatores como: o menor custo computacional, melhor taxa de acerto ou procurar outra vantagem. Então, para visar sempre uma melhoria dos resultados, foi utilizada como forma de vantagem a base cujo valor da mediana foi menor, isto é, o cenário 1 [Tabela 10].

**Tabela 10.** Comparação do melhor cenário utilizando MLP para o Parque A

Parque A			
Cenário	Sazonalidade	P-value	Mediana dos EPMA's
1	sem	0,4147	39,20656
2	com		39,2674

[Fonte: elaboração própria]

#### 4.1.2 Estudo de caso MLP para o Parque B

Para a previsão de um dia a frente de geração de potência eólica utilizando o MLP, neste parque, poderia escolher ambos os cenários, uma vez que o *p-value* foi superior a 5%, demonstrando que os resultados são estatisticamente iguais e não existe um ganho significativo.

Então, buscando sempre uma melhoria dos resultados, foi selecionado o cenário 2 cujo valor da mediana foi um pouco menor que o cenário 1 [Tabela 11].

**Tabela 11.** Comparação do melhor cenário utilizando MLP para o Parque B

Parque B			
Cenário	Sazonalidade	P-value	Mediana dos EPMA's
1	sem	0,4671	61,91609
2	com		61,84141

[Fonte: elaboração própria]

#### 4.1.3 Estudo de caso MLP para o Parque C

Para a previsão de um dia a frente de geração de potência eólica utilizando a MLP, neste parque, poderiam ter sido usados ambos os cenários, uma vez que o *p-*

---

*value* foi superior a 5%, demonstrando que os resultados são estatisticamente iguais e não existe um ganho significativo. Então, buscando sempre uma melhoria dos resultados, foi selecionada a base cujo valor da mediana foi menor, isto é, o cenário 1 cujo valor da mediana foi um pouco menor do que o cenário 2 [Tabela 12].

**Tabela 12.** Comparação do melhor cenário utilizando MLP para o Parque C

Parque C			
Cenário	Sazonalidade	P-value	Mediana dos EPMA's
1	sem	0,1922	3,717633
2	com		3,764222

[Fonte: elaboração própria]

## 4.2 Experimentos com o RC

Para os experimentos com o algoritmo RC foi comparado entre os três parques A, B e C qual deles possui o melhor cenário, sendo eles: o primeiro sem sazonalidade e o segundo com sazonalidade. Dessa forma, obtendo o melhor resultado para cada Parque.

Para a RC, foi utilizada a configuração descrita na seção 3.4.2

### 4.2.1 Estudo de caso RC para o Parque A

No caso da Tabela 13, percebe-se que o *p-value* gerado foi um valor muito pequeno. Por isso, o software R aplicou o seu valor *default* para números muito pequenos e próximos de zero. Dessa forma, sendo o *p-value* menor que 5%, o resultado encontra-se em região de rejeição, considerando então as medianas como diferentes. Portanto, deve ser escolhido o cenário com menor mediana [Tabela 13].

**Tabela 13.** Comparação do melhor cenário utilizando RC para o Parque A

Parque A			
Cenário	Sazonalidade	P-value	Mediana dos EPMA's
1	sem	2,2*10 <sup>-16</sup>	36,46915
2	com		25,36381

[Fonte: elaboração própria]

A partir disso, o melhor cenário e o escolhido da Tabela 13 foi o número 2.

#### 4.2.2 Estudo de caso RC para o Parque B

No caso da Tabela 14, percebe-se que o *p-value* gerado foi um valor muito pequeno e o software R também aplicou o seu valor *default* para números muito pequenos e próximos de zero. Dessa forma, sendo o *p-value* menor que 5%, o resultado encontra-se em região de rejeição, considerando então as medianas como diferentes. Portanto, deve ser escolhido o cenário com menor mediana. Logo, o melhor cenário para esse experimento é o de número 2 [Tabela 14].

**Tabela 14.** Comparação do melhor cenário utilizando RC para o Parque B

Parque B			
Cenário	Sazonalidade	P-value	Mediana dos EPMA's
1	sem	2,2*10 <sup>-16</sup>	36,51005
2	com		31,12505

[Fonte: elaboração própria]

---

### 4.2.3 Estudo de caso RC para o Parque C

No caso da Tabela 15, percebe-se que o *p-value* foi menor que 5%. Dessa forma, o resultado encontra-se em região de rejeição, considerando então as medianas como diferentes. Portanto, deve ser escolhido o cenário com menor mediana. Logo, o melhor cenário para esse experimento é o de número 2 [Tabela 15].

**Tabela 15.** Comparação do melhor cenário utilizando RC para o Parque C

Parque C			
Cenário	Sazonalidade	P-value	Mediana dos EPMA's
1	sem	0,0005762	6,430605
2	com		5,60636

[Fonte: elaboração própria]

## 4.3 Comparação MLP x RC

Nesta seção, serão comparados os melhores resultados de cara rede para cada um dos três parques, obtendo ao final a melhor opção de rede neural para cada parque eólico estudado.

### 4.3.1 Estudo comparativo para o Parque A

Ao observar os resultados da comparação entre as redes na tabela 16, pode-se afirmar que o RC obteve um resultado melhor que a MLP, pois a sua mediana foi de 25,36, sendo assim, a melhor escolha de rede neural para o Parque A [Tabela 16].

**Tabela 16.** Comparação de melhor algoritmo para o Parque A

<b>Parque A</b>			
<b>Cenário</b>	<b>Algoritmo</b>	<b>P-value</b>	<b>Mediana dos EPMA's</b>
1	MLP	2,20E-16	39,20656
2	RC		25,36381

[Fonte: elaboração própria]

#### 4.3.2 Estudo comparativo para o Parque B

Ao observar os resultados da comparação entre as redes na tabela 17, pode-se afirmar que o RC RC obteve um resultado melhor que a MLP com uma mediana de 31,12, sendo assim, a melhor escolha de rede neural para o Parque B [Tabela 17].

**Tabela 17.** Comparação de melhor algoritmo para o Parque B

<b>Parque B</b>			
<b>Cenário</b>	<b>Algoritmo</b>	<b>P-value</b>	<b>Mediana dos EPMA's</b>
1	MLP	2,20E-16	61,84141
2	RC		31,12505

[Fonte: elaboração própria]

#### 4.3.3 Estudo comparativo para o Parque C

Analisando a Tabela 18, percebe-se que ambas as redes obtiveram uma mediana muito baixa, comparando com os resultados dos outros parques.

---

Entretanto, a MLP foi melhor do que o RC, apresentando uma mediana de 3,71, sendo assim, a melhor escolha de rede neural para o Parque C [Tabela 18].

**Tabela 18.** Comparação de melhor algoritmo para o Parque A

Parque C			
Cenário	Algoritmo	P-value	Mediana dos EPMA's
1	MLP	2,20E-16	3,717633
2	RC		5,60636

[Fonte: elaboração própria]

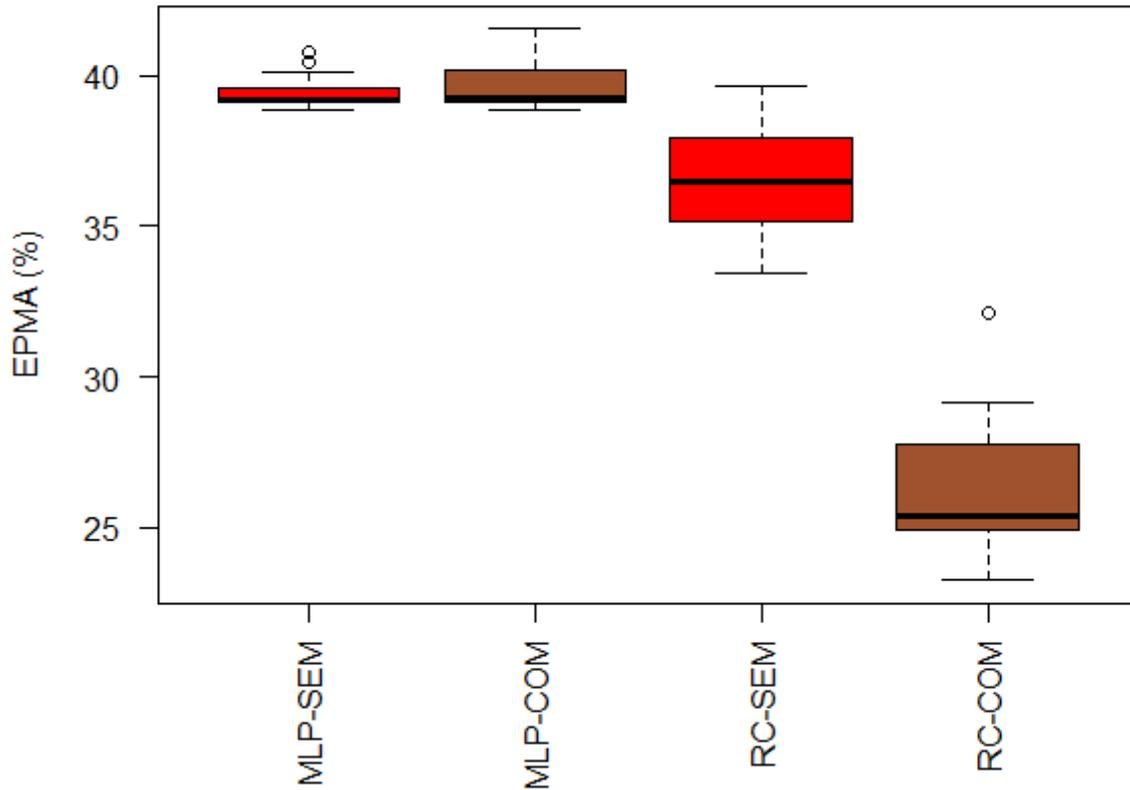
## 4.4 Box Plots

Nesta seção são abordados os gráficos do tipo *box plots* de modo que facilite a observação dos resultados obtidos. Para diferenciar entre os modelos de gráficos com e sem sazonalidade foram utilizadas cores diferentes: vermelho que representa a ausência da sazonalidade e um marrom para representar a adição da sazonalidade.

### 4.4.1 Box Plots para o Parque A

Para os experimentos realizados no parque C e representados na figura 10, observa-se que o RC com a adição da sazonalidade obteve um menor EPMA.

**Figura 8.** Box Plots dos cenários para o Parque A

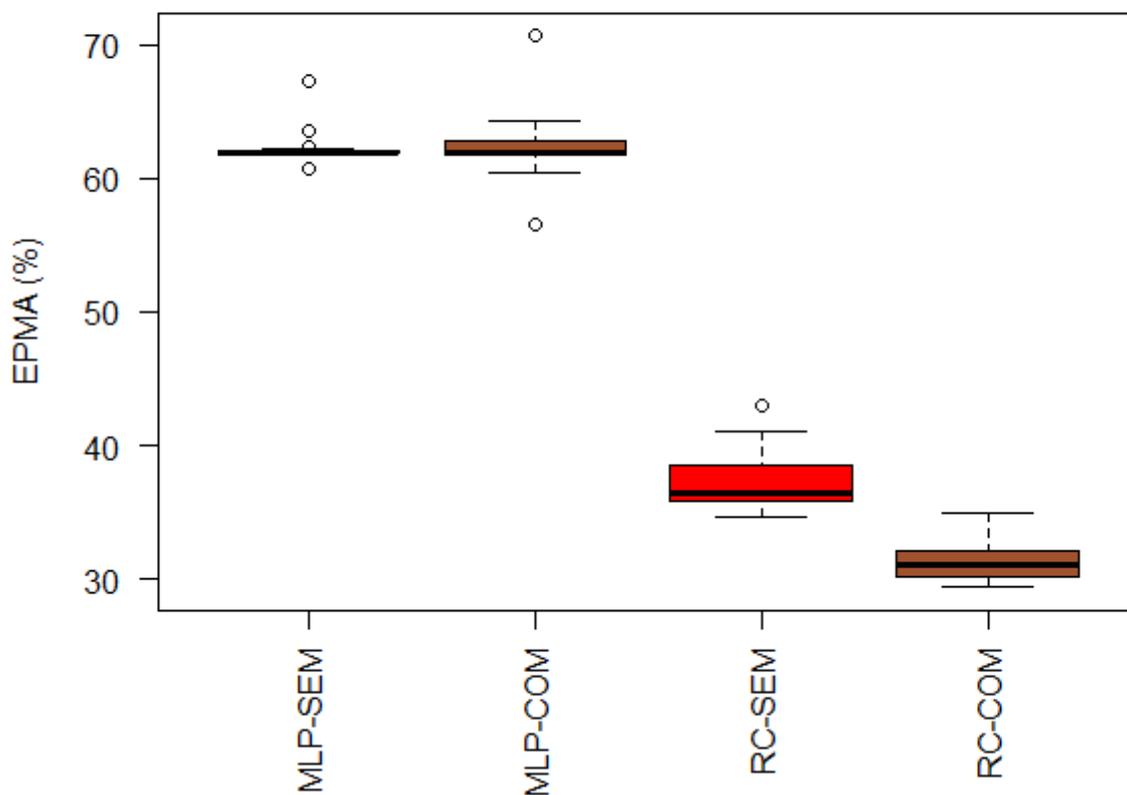


[Fonte: elaboração própria]

#### 4.4.2 Box Plots para o Parque B

Para os experimentos realizados no parque C e representados na figura 10, observa-se que o RC com a adição da sazonalidade obteve um menor EPMA.

**Figura 9.** Box Plots dos cenários para o Parque B

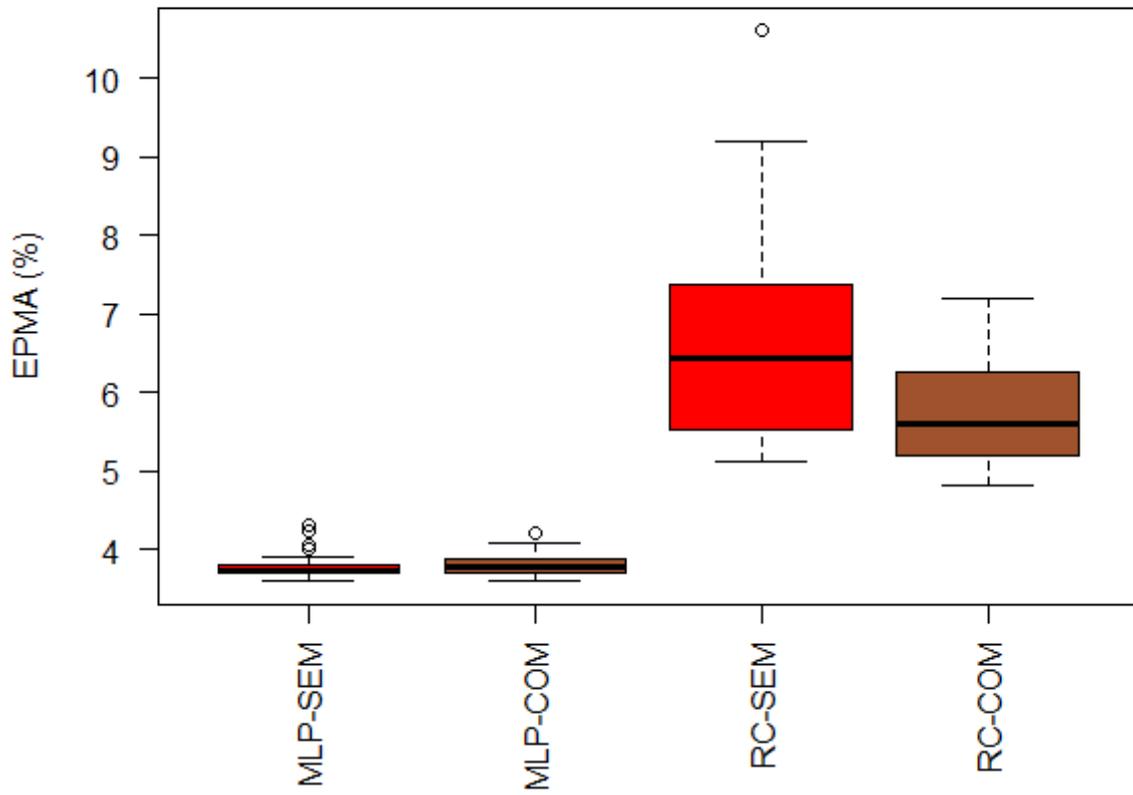


[Fonte: elaboração própria]

#### 4.4.3 Box Plots para o Parque C

Para os experimentos realizados no parque C e representados na figura 10, observa-se que o MLP sem a adição da sazonalidade obteve um menor EPMA.

Figura 10. Box Plots dos cenários para o Parque C



---

# Capítulo 5 – Considerações Finais

Este capítulo tem como objetivo descrever as considerações finais do trabalho subdividindo-a em 2 seções: a 5.1, onde será descrita a conclusão do trabalho e a 5.2, contendo possíveis trabalhos futuros.

## 5.1 Conclusão

Este trabalho teve como objetivo principal fazer um estudo comparativo entre as técnicas RC e MLP para prever a quantidade de potência gerada um dia a frente, considerando dados reais. Para validar e comparar essas técnicas, foram aplicados testes estatísticos.

Os resultados obtidos mostram que a influência da sazonalidade melhorou a maioria dos resultados durante a fase dos experimentos. Na fase final, descrita na seção 4.3, onde ocorreram as comparações entre os melhores resultados das redes MLP e RC, a RC obteve melhores resultados para os parques A e B, cujos valores de medianas foram 25,36 e 31,12 respectivamente. Porém, o MLP foi a melhor rede para o Parque C, conseguindo atingir uma mediana de 3,71.

## 5.2 Trabalhos Futuros

Como trabalhos futuros, podem ser listadas algumas configurações de redes para serem testadas:

- Utilizar algoritmos de busca em vez da correlação dos dados para determinação das variáveis de entrada das redes. PSO (Particle Swarm Optimization) é um dos algoritmos mais utilizados;

- Utilizar outros algoritmos de treinamento para a rede MLP, como o algoritmo colônia de formigas ou de enxames.

- Definir novas configurações para RC, por exemplo, alterar a quantidade de neurônios no reservatório, que possam causar alterações positivas no EPMA;

- 
- Utilizar diferentes funções de ativação;
  - Ampliar a margem dos testes utilizando dados ainda mais antigos para os mesmos parques eólicos e depois realizar comparações;
  - Utilizar outras bases de dados e realizar testes estatísticos.

# Bibliografia

- [1] AGUIAR, Bruna Cavalcanti Galle de. Utilizando Reservoir computing para previsão na geração de energia de Parques Elólicos brasileiros. 2015. 58 f. Dissertação (Mestrado) - Curso de Pós-graduação em Engenharia da Computação, Universidade de Pernambuco, Recife, 2015.
- [2] ANEEL. Aplicações - atlas em energia eólica. Disponível em: <[http://www2.aneel.gov.br/aplicacoes/atlas/pdf/06-energia\\_eolica\(3\).pdf](http://www2.aneel.gov.br/aplicacoes/atlas/pdf/06-energia_eolica(3).pdf)>. Acesso em: 29 abr. 2017.
- [3] ANEEL. Evolução da energia eólica no brasil. Disponível em: <[http://www.aneel.gov.br/aneel-essencial/-/asset\\_publisher/c4m6oiomklad/content/evolucao-da-energia-eolica-no-brasil?inheritredirect=false](http://www.aneel.gov.br/aneel-essencial/-/asset_publisher/c4m6oiomklad/content/evolucao-da-energia-eolica-no-brasil?inheritredirect=false)>. Acesso em: 16 mai. 2017.
- [4] BEZERRA, Sabrina G. T. A.; ANDRADE, Camila B. de; VALENÇA, Mêuser J. S. Using Reservoir Computing and Trend Information for Short-Term Streamflow Forecasting. Springer, p.308-316, ago. 2016.
- [5] ECOMP. Dissertação. Disponível em: <<http://tcc.ecomp.poli.br/20122/bruna%20cavalcanti%20galle%20de%20aguia%20r.pdf>>. Acesso em: 08 abr. 2017.
- [6] EMPRESA DE PESQUISA ENERGÉTICA. Projeção da demanda de energia elétrica para os próximos 10 anos (2015-2024). Disponível em: <<http://www.epe.gov.br/mercado/documents/dea%2003-2015-%20proje%20c3%a7%20c3%b5es%20da%20demanda%20de%20energia%20el%20c3%a9trica%202015-2024.pdf>>. Acesso em: 10 mai. 2017.
- [7] GREENPEACE and EREC. Conselho Europeu de Energia Renovável (Erec) revolução energética A caminho do desenvolvimento limpo, Dezembro 2010.
- [8] OLIVEIRA, Juliana Lima; COSTA, Alexandre Araújo. Estudo de variabilidade do vento em escala sazonal sobre o nordeste brasileiro utilizando o RAMS: os casos de 1973-1974 e 1982-1983. Bras. meteorol, São Paulo, v. 26, n. 1, mar. 2011.

- [9] ONS. Operador nacional do sistema elétrico. Disponível em: <<http://www.ons.org.br/home/>>. Acesso em: 11 abr. 2017.
- [10] PEREIRA, SALOMÃO Almeida; LOUREIRO, Lorena Drumond; MARTINS, Fernanda Roberta Gomes da Silva Pinto. Estudo comparativo entre modelos de previsão de demanda: ensaio em um produto classe a de uma empresa de perfumes e cosméticos. In: ENEGEP, Fortaleza, 2006.
- [11] PORTAL BRASIL. Brasil é o quarto país em que energia eólica mais cresce no mundo. Disponível em: <<http://www.brasil.gov.br/infraestrutura/2016/01/brasil-e-um-dos-principais-geradores-de-energia-eolica-do-mundo>>. Acesso em: 08 mai. 2017.
- [12] RODRIGUES, Guilherme. Utilização de Redes Neurais para Previsão de Ventos no Horizonte de 24 Horas. 2007. 119 f. Dissertação de Mestrado do Curso de Engenharia Elétrica, Universidade Federal do Rio de Janeiro, Rio de Janeiro.
- [13] SGT – Superintendência de Gestão Tarifária da Agência Nacional de Energia Elétrica (ANEEL). (2015) Bandeiras Tarifárias. Disponível em: <http://www.aneel.gov.br/bandeiras-tarifarias>. Acesso em 05 de setembro de 2016.
- [14] “The R Project for Statistical computing” [Online]. Available: <https://www.r-project.org/> Acessado em 15 de outubro de 2016.
- [15] VALENÇA, Mêuser Jorde Silva. Aplicando Redes Neurais. Olinda: Livro Rápido, 2005.
- [16] VALENÇA, Mêuser Jorde Silva. Fundamentos De Redes Neurais. Recife: Livro Rápido, 2012.
- [17] VERSTRAETEN, David. Reservoir Computing: Computation with dynamical systems. 2009
- [18] VALENÇA, J. M. S. Fundamentos das Redes Neurais, [S.l.]: Livro Rápido, 2011.
- [19] Werbos, J. P. Beyond Regression: New tools for prediction and analysis in behavioral sciences, Tese (Doutorado) Harvard University, 1974.

- [20] ECOMP. Uso de reservoir computing para classificação e previsão de eventos rampa. Disponível em: <[http://www.ecomp.poli.br/~posgraduacao/images/stories/convites/msc\\_ec\\_%20chamada\\_defesa\\_97.pdf](http://www.ecomp.poli.br/~posgraduacao/images/stories/convites/msc_ec_%20chamada_defesa_97.pdf)>. Acesso em: 08 abr. 2017.