



# **APLICAÇÃO DE MULTILAYER PERCEPTRON E RESERVOIR COMPUTING PARA PREVISÃO DE VAZÃO CONSIDERANDO SAZONALIDADE**

**Trabalho de Conclusão de Curso**

**Engenharia da Computação**

**Pedro Henrique de Queiroz Santos**  
**Orientador: Mêuser Jorge Silva Valença**



UNIVERSIDADE  
DE PERNAMBUCO

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

**Pedro Henrique de Queiroz Santos**

**APLICAÇÃO DE MULTILAYER PERCEPTRON E  
RESERVOIR COMPUTING PARA PREVISÃO DE  
VAZÃO CONSIDERANDO 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 9:00 horas, reuniu-se para deliberar a defesa da monografia de conclusão de curso do discente **PEDRO HENRIQUE DE QUEIROZ SANTOS**, orientado pelo professor **Mêuser Jorge Silva Valença**, sob título **APLICAÇÃO DE MULTI-LAYER PERCEPTRON E RESEVOIR COMPUTING PARA PREVISÃO DE VAZÃO CONSIDERANDO 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 (novo 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 MARIO LINS GALDINO

MÊUSER JORGE SILVA VALENÇA

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

*Para Verônica Queiroz, minha mãe, por sempre acreditar em mim e me incentivar a alcançar novos horizontes.*

# Agradecimentos

Agradeço a todos que fizeram parte da minha graduação direta ou indiretamente. A minha família que foi o meu porto seguro e incentivo para chegar até este ponto. Aos amigos da Poli por sempre estudarmos juntos e compartilharmos bons e maus momentos. Aos amigos do Ciência sem Fronteira que me mostraram que nem todo conhecimento se é adquirido nos muros de uma Universidade e que amizade transcende distâncias e culturas. Aos amigos do Colégio Imaculada por sempre estarem presentes, mesmo não estando perto. Aos diversos professores que participaram da minha formação. Ao professor Mêuser Valença por ter aceitado me orientar nesta monografia. A Everalda Solange por ser uma monitora exemplar do começo ao fim da graduação. A todos vocês: esta conquista é nossa, meus sinceros agradecimentos.

# Resumo

A sobrevivência na Terra só é possível devido a existência de água. Utilizada na agricultura, pecuária e no setor energético, seu uso deve ser eficiente. Por ser um recurso limitado, a melhor gestão da água é uma preocupação para líderes de diversos setores. Modelos estatísticos e hidrológicos apesar de funcionarem, têm se mostrado limitados e custosos na previsão de vazão em usinas hidroelétricas devido a características não-lineares presentes nos dados e escassez de informações geográficas sobre as bacias hidrográficas. A medida que a tecnologia avança, novas maneiras de se prever vazão surgem e redes neurais têm se provado uma solução viável e precisa para a previsão de vazões futuras. Este trabalho tem como objetivo comparar resultados de vazão mensal considerando representações de sazonalidade. Ao comparar os cenários possíveis das codificações de sazonalidade adicionadas aos dados e variando configurações das redes MLP e RC, espera-se ter como resultado qual arquitetura gera menores erros para cada representação de sazonalidade em diferentes usinas hidroelétricas.

# Abstract

Life on Earth is only possible due to the existence of water. Because it is demanded by agriculture, livestock and for generating energy, the use of water must be efficient. Since it is a limited resource, improving water management is a concern for leaders of several departments. Although statistical and hydrological models work for predicting watershed's flow, they have been proven limited and costly due to non-linear characteristics present on the data and lack of hydrographic basins' geographical information. As technology evolve, new ways of predicting watershed's flow appear and neural networks have proven to be a viable and precise solution to predict future watershed's flow. This research has the goal to compare results of monthly flow considering seasonality representations. By comparing the possible scenarios of seasonality codes added up to the data and by varying settings for RC and MLP networks, it is expected as a final result to point the architecture that generates smaller errors for each seasonality representation in different Brazilian hydroelectric plants.

# Sumário

<b>CAPÍTULO 1 INTRODUÇÃO</b> .....	<b>1</b>
1.1 CARACTERIZAÇÃO DO PROBLEMA E MOTIVAÇÕES .....	1
1.2 OBJETIVOS E METAS .....	2
1.2.1 <i>Objetivo Geral</i> .....	2
1.2.2 <i>Objetivos específicos</i> .....	2
1.3 ESTRUTURA DA MONOGRAFIA .....	2
<b>CAPÍTULO 2 FUNDAMENTAÇÃO TEÓRICA</b> .....	<b>3</b>
2.1 PLANEJAMENTO ENERGÉTICO NO BRASIL.....	3
2.2 REDES NEURAIS.....	4
2.3 MULTILAYER PERCEPTRON .....	6
2.3.1 <i>Treinamento da MLP</i> .....	7
2.4 RESERVOIR COMPUTING .....	9
2.4.1 <i>Treinamento da RC</i> .....	11
<b>CAPÍTULO 3 METODOLOGIA</b> .....	<b>13</b>
3.1 BASES DE DADOS .....	13
3.1.1 <i>Bento Munhoz</i> .....	13
3.1.2 <i>Furnas</i> .....	14
3.1.3 <i>Três Marias</i> .....	15
3.1.4 <i>Tucuruí</i> .....	15
3.2 PRÉ-PROCESSAMENTO DOS DADOS .....	16
3.2.1 <i>Normalização dos dados</i> .....	16
3.2.2 <i>Defasagem dos dados</i> .....	17
3.2.3 <i>Correlação dos dados</i> .....	18
3.3 REPRESENTANDO SAZONALIDADE .....	19
3.4 APRESENTANDO AS VARIÁVEIS DE ENTRADA À MLP .....	21
3.4.1 <i>Parâmetros utilizados no MLP</i> .....	22
3.5 APRESENTANDO AS VARIÁVEIS DE ENTRADA À RC .....	23
3.5.1 <i>Parâmetros utilizados no RC</i> .....	23
3.6 TESTES ESTATÍSTICOS .....	24
3.6.1 <i>Teste Shapiro-Wilk</i> .....	25
3.6.2 <i>Teste F</i> .....	25
3.6.3 <i>Teste t de Student</i> .....	25

3.6.4 Teste Wilcoxon.....	26
<b>CAPÍTULO 4 RESULTADOS.....</b>	<b>27</b>
4.1 EXPERIMENTOS COM A BASE BENTO MUNHOZ.....	27
4.1.1 Hidroelétrica Bento Munhoz: sem representação de sazonalidade.....	27
4.1.2 Hidroelétrica Bento Munhoz: sazonalidade de 12 bits .....	29
4.1.3 Hidroelétrica Bento Munhoz: sazonalidade de 4 bits .....	31
4.1.4 Hidroelétrica Bento Munhoz: sazonalidade de senos e cossenos .....	32
4.2 EXPERIMENTOS COM A BASE FURNAS .....	34
4.2.1 Hidroelétrica Furnas: sem representação de sazonalidade .....	34
4.2.2 Hidroelétrica Furnas: sazonalidade de 12 bits .....	36
4.2.3 Hidroelétrica Furnas: sazonalidade de 4 bits .....	38
4.2.4 Hidroelétrica Furnas: sazonalidade de senos e cossenos.....	39
4.3 EXPERIMENTOS COM A BASE TRÊS MARIAS .....	41
4.3.1 Hidroelétrica Três Marias: sem representação de sazonalidade .....	41
4.3.2 Hidroelétrica Três Marias: sazonalidade de 12 bits.....	43
4.3.3 Hidroelétrica Três Marias: sazonalidade de 4 bits .....	45
4.3.4 Hidroelétrica Três Marias: sazonalidade de senos e cossenos.....	47
4.4 EXPERIMENTOS COM A BASE TUCURUÍ.....	48
4.4.1 Hidroelétrica Tucuruí: sem representação de sazonalidade .....	49
4.4.2 Hidroelétrica Tucuruí: sazonalidade de 12 bits.....	50
4.4.3 Hidroelétrica Tucuruí: sazonalidade de 4 bits.....	52
4.4.4 Hidroelétrica Tucuruí: sazonalidade de senos e cossenos.....	54
4.5 COMPARATIVO DOS RESULTADOS .....	55
4.5.1 Hidroelétrica Bento Munhoz.....	55
4.5.2 Hidroelétrica de Furnas.....	58
4.5.3 Hidroelétrica Três Marias .....	61
4.5.4 Hidroelétrica Tucuruí .....	63
<b>CAPÍTULO 5 CONSIDERAÇÕES FINAIS.....</b>	<b>67</b>
5.1 CONCLUSÃO .....	67
5.2 TRABALHOS FUTUROS.....	67
<b>BIBLIOGRAFIA .....</b>	<b>69</b>

# Índice de Figuras

<b>Figura 1.</b> Representação de um neurônio biológico.....	5
<b>Figura 2.</b> Neurônio computacional de McCulloch e Pitts.....	5
<b>Figura 3.</b> Rede MLP.....	7
<b>Figura 4.</b> Rede RC.....	11
<b>Figura 5.</b> Usina Hidroelétrica Governador Bento Munhoz da Rocha Netto.....	14
<b>Figura 6.</b> Usina Hidroelétrica de Furnas.....	14
<b>Figura 7.</b> Usina Hidroelétrica de Três Marias.....	15
<b>Figura 8.</b> Usina Hidroelétrica de Tucuruí.....	16
<b>Figura 9.</b> Boxplot dos cenários da base de dados de Bento Munhoz.....	58
<b>Figura 10.</b> Boxplot dos cenários da base de dados de Furnas.....	60
<b>Figura 11.</b> Boxplot dos cenários da base de dados de Três Marias.....	63
<b>Figura 12.</b> Boxplot dos cenários da base de dados de Tucuruí.....	66

# Índice de Tabelas

<b>Tabela 1.</b> Dados pós normalização .....	17
<b>Tabela 2.</b> Dados defasados três vezes no passado .....	17
<b>Tabela 3.</b> Tabela onde as três primeiras colunas representam as variáveis de entrada e a última a variável de saída.....	18
<b>Tabela 4.</b> Correlação entre os dados da tabela 3.....	18
<b>Tabela 5.</b> Representação de sazonalidade com 12 bits .....	19
<b>Tabela 6.</b> Representação de sazonalidade com 4 bits .....	20
<b>Tabela 7.</b> Representação de sazonalidade com senos e cossenos, com os valores finais já calculados usando as equações (7) e (8) .....	21
<b>Tabela 8.</b> Quantidade de neurônios de entrada para cada cenário.....	22
<b>Tabela 9.</b> Média de EPMA's para cada cenário da hidrelétrica de Bento Munhoz (BM) sem representação de sazonalidade .....	28
<b>Tabela 10.</b> Resultados dos <i>p-value</i> para os cenários 1-BM e 2-BM em Bento Munhoz sem representação de sazonalidade .....	28
<b>Tabela 11.</b> Resultados dos <i>p-value</i> para os cenários 2-BM e 3-BM em Bento Munhoz sem representação de sazonalidade .....	29
<b>Tabela 12.</b> Média de EPMA's para cada cenário da hidrelétrica de Bento Munhoz (BM) com 12 bits de sazonalidade .....	29
<b>Tabela 13.</b> Resultados dos <i>p-value</i> para os cenários 1-BM e 2-BM em Bento Munhoz com 12 bits de sazonalidade .....	30
<b>Tabela 14.</b> Resultados dos <i>p-value</i> para os cenários 2-BM e 3-BM em Bento Munhoz com 12 bits de sazonalidade .....	31
<b>Tabela 15.</b> Média de EPMA's para cada cenário da hidrelétrica de Bento Munhoz (BM) com 4 bits de sazonalidade .....	31
<b>Tabela 16.</b> Resultados dos <i>p-value</i> para os cenários 1-BM e 2-BM em Bento Munhoz com 4 bits de sazonalidade .....	32
<b>Tabela 17.</b> Resultados dos <i>p-value</i> para os cenários 2-BM e 3-BM em Bento Munhoz com 4 bits de sazonalidade .....	32
<b>Tabela 18.</b> Média de EPMA's para cada cenário da hidrelétrica de Bento Munhoz (BM) utilizando senos e cossenos para representar sazonalidade.....	33

<b>Tabela 19.</b> Resultados dos <i>p-value</i> para os cenários 1-BM e 2-BM em Bento Munhoz utilizando senos e cossenos para representar sazonalidade.....	33
<b>Tabela 20.</b> Resultados dos <i>p-value</i> para os cenários 2-BM e 3-BM em Bento Munhoz utilizando senos e cossenos para representar sazonalidade.....	34
<b>Tabela 21.</b> Média de EPMA's para cada cenário da hidrelétrica de Furnas (F) sem representação de sazonalidade .....	35
<b>Tabela 22.</b> Resultados dos <i>p-value</i> para os cenários 1-F e 2-F em Furnas sem bits de sazonalidade .....	35
<b>Tabela 23.</b> Resultados dos <i>p-value</i> para os cenários 2-F e 3-F em Furnas sem bits sazonalidade .....	36
<b>Tabela 24.</b> Média de EPMA's para cada cenário da hidrelétrica de Furnas (F) com 12 bits de sazonalidade.....	36
<b>Tabela 25.</b> Resultados dos <i>p-value</i> para os cenários 1-F e 2-F em Furnas com 12 bits de sazonalidade .....	37
<b>Tabela 26.</b> Resultados dos <i>p-value</i> para os cenários 1-F e 3-F em Furnas com 4 bits de sazonalidade .....	37
<b>Tabela 27.</b> Média de EPMA's para cada cenário da hidrelétrica de Furnas (F) com 4 bits de sazonalidade.....	38
<b>Tabela 28.</b> Resultados dos <i>p-value</i> para os cenários 1-F e 2-F em Furnas com 4 bits de sazonalidade .....	39
<b>Tabela 29.</b> Resultados dos <i>p-value</i> para os cenários 2-F e 3-F em Furnas com 4 bits de sazonalidade .....	39
<b>Tabela 30.</b> Média de EPMA's para cada cenário da hidrelétrica de Furnas (F) utilizando senos e cossenos para representar sazonalidade.....	40
<b>Tabela 31.</b> Resultados dos <i>p-value</i> para os cenários 1-F e 2-F em Furnas utilizando senos e cossenos para representar sazonalidade.....	40
<b>Tabela 32.</b> Resultados dos <i>p-value</i> para os cenários 2-F e 3-F em Furnas usando senos e cossenos para representar sazonalidade.....	41
<b>Tabela 33.</b> Média de EPMA's para cada cenário da hidrelétrica de Três Marias (TM) sem representação de sazonalidade .....	42
<b>Tabela 34.</b> Resultados dos <i>p-value</i> para os cenários 1-TM e 2-TM em Três Marias sem representação de sazonalidade .....	42
<b>Tabela 35.</b> Resultados dos <i>p-value</i> para os cenários 2-TM e 3-TM em Três Marias sem representação de sazonalidade .....	43

<b>Tabela 36.</b> Média de EPMA's para cada cenário da hidrelétrica de Três Marias (TM) com 12 bits de sazonalidade .....	43
<b>Tabela 37.</b> Resultados dos <i>p-value</i> para os cenários 1-TM e 2-TM em Três Marias com 12 bits de sazonalidade .....	44
<b>Tabela 38.</b> Resultados dos <i>p-value</i> para os cenários 2-TM e 3-TM em Três Marias com 12 bits de sazonalidade .....	45
<b>Tabela 39.</b> Média de EPMA's para cada cenário da hidrelétrica de Três Marias (TM) com 4 bits de sazonalidade .....	45
<b>Tabela 40.</b> Resultados dos <i>p-value</i> para os cenários 1-TM e 2-TM em Três Marias com 4 bits de sazonalidade .....	46
<b>Tabela 41.</b> Resultados dos <i>p-value</i> para os cenários 2-TM e 3-TM em Três Marias com 4 bits de sazonalidade .....	46
<b>Tabela 42.</b> Média de EPMA's para cada cenário da hidrelétrica de Três Marias (TM) utilizando senos e cossenos para representar sazonalidade .....	47
<b>Tabela 43.</b> Resultados dos <i>p-value</i> para os cenários 1-TM e 2-TM em Três Marias usando senos e cossenos para representar sazonalidade .....	48
<b>Tabela 44.</b> Resultados dos <i>p-value</i> para os cenários 2-TM e 3-TM em Três Marias usando senos e cossenos para representar sazonalidade .....	48
<b>Tabela 45.</b> Média de EPMA's para cada cenário da hidrelétrica de Tucuruí (T) sem representação de sazonalidade .....	49
<b>Tabela 46.</b> Resultados dos <i>p-value</i> para os cenários 1-T e 2-T em Tucuruí sem bits representando sazonalidade .....	49
<b>Tabela 47.</b> Resultados dos <i>p-value</i> para os cenários 1-T e 3-T em Tucuruí sem bits para representar a sazonalidade .....	50
<b>Tabela 48.</b> Média de EPMA's para cada cenário da hidrelétrica de Tucuruí (T) com 12 bits de sazonalidade .....	50
<b>Tabela 49.</b> Resultados dos <i>p-value</i> para os cenários 1-T e 2-T em Tucuruí com 12 bits de sazonalidade .....	51
<b>Tabela 50.</b> Resultados dos <i>p-value</i> para os cenários 2-T e 3-T em Tucuruí com 12 bits de sazonalidade .....	52
<b>Tabela 51.</b> Média de EPMA's para cada cenário da hidrelétrica de Tucuruí (T) com 4 bits de sazonalidade .....	52
<b>Tabela 52.</b> Resultados dos <i>p-value</i> para os cenários 1-T e 2-T em Tucuruí com 4 bits de sazonalidade .....	53

<b>Tabela 53.</b> Resultados dos <i>p-value</i> para os cenários 2-T e 3-T em Tucuruí com 4 bits de sazonalidade .....	53
<b>Tabela 54.</b> Média de EPMA's para cada cenário da hidrelétrica de Tucuruí (T) utilizando senos e cossenos para representar sazonalidade.....	54
<b>Tabela 55.</b> Resultados dos <i>p-value</i> para os cenários 1-T e 2-T em Tucuruí usando senos e cossenos para representar sazonalidade.....	55
<b>Tabela 56.</b> Resultados dos <i>p-value</i> para os cenários 1-T e 3-T em Tucuruí usando senos e cossenos para representar a sazonalidade.....	55
<b>Tabela 57.</b> Média de EPMA's para cada cenário da hidrelétrica de Bento Munhoz, onde RC-25 significa rede RC com 25 neurônios no reservatório. ....	56
<b>Tabela 58.</b> Resultados dos <i>p-value</i> para os C enários 1 e 2 em Bento Munhoz.....	56
<b>Tabela 59.</b> Resultados dos <i>p-value</i> para os cenários 2 e 3 em Bento Munhoz.....	57
<b>Tabela 60.</b> Resultados dos <i>p-value</i> para os cenários 3 e 4 em Bento Munhoz.....	57
<b>Tabela 61.</b> Média de EPMA's para cada cenário da hidrelétrica de Furnas, onde RC-25 significa rede RC com 25 neurônios no reservatório e RC-35 significa 35 neurônios no reservatório.....	58
<b>Tabela 62.</b> Resultados dos <i>p-value</i> para os cenários 1 e 2 em Furnas.....	59
<b>Tabela 63.</b> Resultados dos <i>p-value</i> para os cenários 2 e 3 em Furnas.....	59
<b>Tabela 64.</b> Resultados dos <i>p-value</i> para os cenários 2 e 4 em Furnas.....	60
<b>Tabela 65.</b> Média de EPMA's para cada cenário da hidrelétrica de Três Marias, onde RC-25 significa rede RC com 25 neurônios no reservatório. ....	61
<b>Tabela 66.</b> Resultados dos <i>p-value</i> para os cenários 1 e 2 em Três Marias.....	61
<b>Tabela 67.</b> Resultados dos <i>p-value</i> para os cenários 2 e 3 em Três Marias.....	62
<b>Tabela 68.</b> Resultados dos <i>p-value</i> para os cenários 3 e 4 em Três Marias.....	62
<b>Tabela 69.</b> Média de EPMA's para cada cenário da hidrelétrica de Tucuruí.....	64
<b>Tabela 69.</b> Resultados dos <i>p-value</i> para os cenários 1 e 2 em Tucuruí.....	64
<b>Tabela 70.</b> Resultados dos <i>p-value</i> para os cenários 2 e 3 em Tucuruí.....	65
<b>Tabela 71.</b> Resultados dos <i>p-value</i> para os cenários 3 e 4 em Três Marias.....	65

# Tabela de Símbolos e Siglas

EPMA – Erro Percentual Médio Absoluto

MLP – Multi-layer perceptron

RC – Reservoir Computing

RNAs – Redes Neurais Artificiais

# Capítulo 1

## Introdução

Este capítulo encontra-se dividido em três seções. A primeira seção expõe as características do problema tratado nesta e as motivações para o desenvolvimento deste projeto. Na segunda seção são apresentados os principais objetivos e metas. Na terceira e última, apresenta-se a estrutura desta monografia.

### 1.1 Caracterização do problema e motivações

Um dos fatores que torna a sobrevivência no planeta Terra possível é a existência de água. Esta é essencial nos setores energéticos, agropecuários e nas comunidades. Apesar de 75 % da face da terra ser coberta por água, apenas 3% deste total é de água doce [8], ou seja, apenas este pequeno montante precisa ser dividido para as pessoas e suas necessidades.

Por ser um bem tão escasso, a gestão dos recursos hídricos e o seu melhor aproveitamento é uma preocupação [9]. No Brasil, modelos estatísticos e hidrológicos têm funcionado para auxiliar na gerência destes recursos, porém ambos possuem limitações [7]. Enquanto os modelos estatísticos não conseguem lidar com características não-lineares presentes nas séries temporais que representam as vazões nos rios, os modelos hidrológicos confiam fortemente na qualidade e quantidade dos dados das bacias hidrográficas, porém estes são escassos. [3]

Para auxiliar nas tomadas de decisão, a tecnologia tem se provado poderosa e menos custosa para efetuar previsões sobre as vazões futuras das bacias hidrográficas brasileiras. Pesquisas têm mostrado que o uso das Redes Neurais Artificiais é eficiente na previsão de séries temporais, mais especificamente para o escopo deste trabalho, em previsão de vazão [3][10][16][17]. Dentre elas, podemos citar as redes chamadas de Multilayer Perceptron e Reservoir Computing.

Com o ensejo de contribuir com a gama de soluções que resolvem problemas de previsão de vazão em bacias hidroelétricas, este trabalho visa explorar e comparar o desempenho das representações de sazonalidade utilizando MLP e RC.

## 1.2 Objetivos e Metas

### 1.2.1 Objetivo Geral

O principal objetivo desta monografia é comparar os resultados de previsão da vazão mensal considerando representações de sazonalidade usando as redes neurais MLP e RC.

### 1.2.2 Objetivos específicos

1. Estudar as redes neurais MLP e RC aplicadas à previsão de vazão;
2. Obter os erros na previsão de vazões mensais considerando diferentes representações de sazonalidade;
3. Analisar os dados estatisticamente para definir qual rede neural obtém melhores resultados para cada tipo de representação de sazonalidade;

## 1.3 Estrutura da Monografia

No **Capítulo 1**, o leitor vai ser apresentado às motivações por trás deste trabalho e qual o problema envolvido, além dos objetivos que se deseja alcançar nesta monografia.

Já no **Capítulo 2**, será apresentado um panorama da disponibilidade hídrica no Brasil, além de explorar os conceitos de redes neurais, biológicas e artificiais, com um foco maior nas arquiteturas utilizadas para resolver os problemas apontados no capítulo 1.

O **Capítulo 3** desenvolve a metodologia empregada para se atingir os objetivos do trabalho. Ele abrange desde a escolha das hidroelétricas, até o pré-processamento dos dados, a configuração das redes MLP e RC e os testes usados para avaliar os resultados obtidos.

Os resultados dos experimentos descritos no capítulo 3 podem ser vistos no **Capítulo 4**, juntamente com a comparação dos diversos cenários justificados pelos testes estatísticos

Por fim, o **Capítulo 5** expõe as conclusões sobre os resultados obtidos no capítulo 4, além de apontar ideias para trabalhos futuros.

# Capítulo 2

## Fundamentação Teórica

### 2.1 Planejamento energético no Brasil

O crescimento populacional no Brasil e o processo de industrialização do século XX fez que o país começasse a produzir energia, a princípio advindas do petróleo e do carvão [2]. Porém, nos primeiros anos do século XXI o mundo passou a caminhar para um futuro mais sustentável, com os vários acordos sobre aquecimento global [15], e o Brasil também adotou as novas políticas as quais diversos países se comprometeram.

Com isso, diferente do século anterior, novas fontes de energia renováveis são procuradas. Apesar de investir em energia eólica e solar, o Brasil conta com 1266 usinas hidroelétricas instaladas, o que corresponde a 61% da energia produzida no país, tendo o maior potencial hidroelétrico do mundo [1].

Com o objetivo de aproveitar ao máximo o potencial hidroelétrico do Brasil, diminuir as demais fontes poluentes e otimizar o custo de geração, é definido um Plano Mensal de Operações (PMO), onde as diversas empresas que produzem energia hidroelétrica devem determinar a capacidade de geração energética e o possível intercâmbio de energia [18].

Uma das métricas utilizadas para determinar o PMO é a vazão natural afluente, que é o volume de água que passa por uma determinada seção de um rio. Porém, sem a devida modelagem, esta é uma variável difícil de ser estimada. Os dados de vazão são séries temporais, ou seja, conjunto de observações ordenadas no tempo [13].

A importância de ser prever o comportamento e a vazão de um curso do rio no futuro é ajudar os gestores de energia a tomar melhores decisões. Para isso, diversos modelos tentam aproximar valores de vazão, seja anual, mensal ou diária. Existem

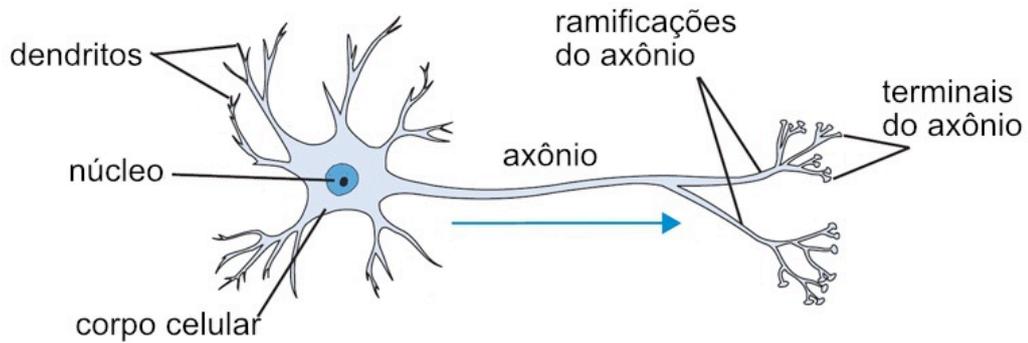
três abordagens comuns na área: modelos estatísticos, modelos hidrológicos e inteligência artificial [3].

Os modelos estatísticos, que lidam com regressão, funcionam e trazem boas previsões em alguns casos, porém não conseguem lidar muito bem com informações não-lineares contidas nas séries de vazão. Os modelos hidrológicos também fazem o seu papel, mas confiam fortemente na quantidade e qualidade dos dados hidrológicos, mesmo estes sendo escassos. Por fim, a inteligência artificial, mais especificamente redes neurais funcionam bem, lidam com não-linearidade e são mais baratas de serem implementadas [3].

## 2.2 Redes Neurais

Redes Neurais Artificiais fazem parte de uma subárea da Inteligência Artificial que toma o cérebro humano como inspiração. Assim como os humanos, as redes neurais adquirem conhecimento através da experiência. Há o fluxo de entrada, processamento e a geração de saídas, que são as informações que resolvem problemas do mundo prático.

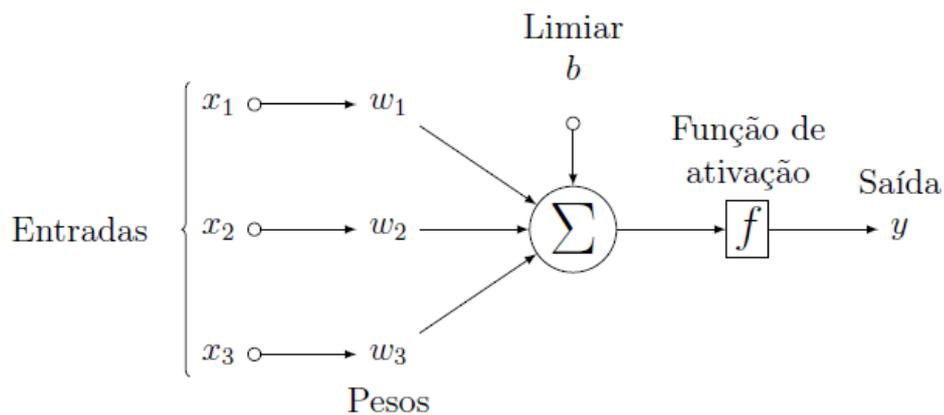
O cérebro humano é capaz de executar milhões de operações em paralelo. Apesar de ser formado por indivíduos lentos e pouco inteligentes, chamado de neurônios, estes em conjunto dão o poder de processamento que faz os seres humanos serem inteligentes. O neurônio biológico é composto por um núcleo, axônios e dendritos e é capaz de propagar pulsos elétricos de acordo com o estímulo recebido, como mostra a figura 1. Se o estímulo for maior do que dado limiar, o pulso é propagado, caso contrário não propaga [17].



**Figura 1.** Representação de um neurônio biológico.

Fonte: <http://cs231n.github.io/neural-networks-1/>

As redes neurais artificiais, por sua vez, funcionam de maneira análoga ao cérebro humano, também sendo constituídas de unidades de processamento mais simples interligadas entre si. Estas unidades são os neurônios artificiais, que ao receberem os dados de entrada, têm uma função de ativação processando a entrada líquida que vai fazer o neurônio disparar, ou não, de acordo com um limiar [18]. A figura 2 mostra o neurônio computacional de McCulloch e Pitts.



**Figura 2.** Neurônio computacional de McCulloch e Pitts.

Fonte: Extraída de [3]

O conhecimento adquirido por uma rede neural está localizado nos pesos das conexões entre os neurônios. Tais pesos são atualizados de acordo com o tipo de

aprendizado, que pode ser classificado em dois: aprendizado supervisionado ou não-supervisionado [3].

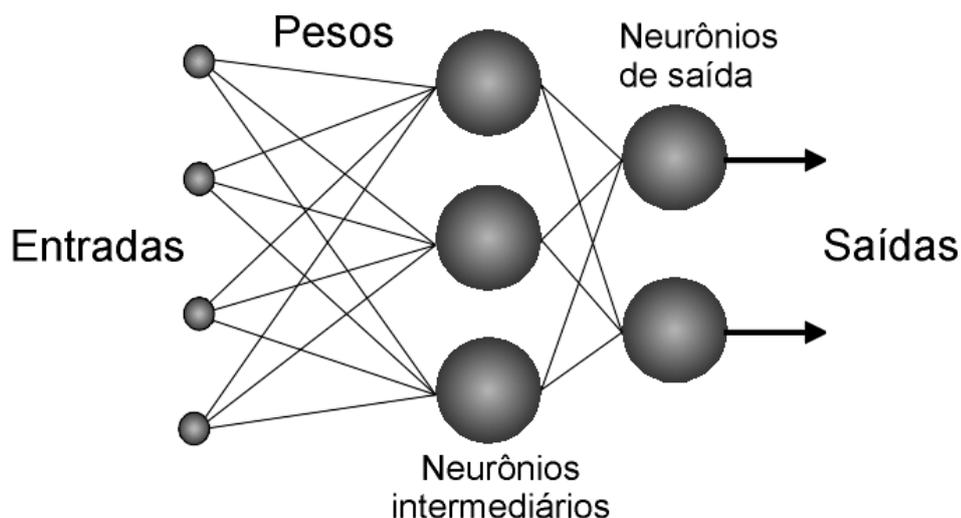
O treinamento supervisionado provê para a rede neural, além das entradas, a saída, ou seja, os pesos são atualizados de acordo com a diferença da saída real e a saída calculada pela rede. Por outro lado, o treinamento não-supervisionado só entrega os dados de entrada para a rede, fazendo com que os pesos sejam recalculados de acordo com as regras internas da rede, baseadas unicamente na relação entre as entradas [18].

As redes neurais também podem ser classificadas no sentido em que a informação percorre a rede. Utilizando este critério, elas podem ser classificadas em redes *feedforward* ou recorrentes. As redes *feedforward* têm a informação passando pela camada de entrada, uma ou mais camadas internas e a camada de saída, num único sentido. Já as redes recorrentes possuem pelo menos um laço de realimentação, dando a estas redes o poder de processar informações não-lineares [3].

## 2.3 Multilayer Perceptron

O MLP é uma generalização do Perceptron, que por sua vez foi a primeira rede *feedforward* a ser proposta. A arquitetura da rede MLP envolve uma camada de entrada, uma ou mais camadas intermediárias (também chamadas de camadas ocultas ou escondidas) e uma camada de saída.

O modo de uso mais tradicional é formado por uma camada de entrada, uma oculta e uma de saída, onde a camada de entrada contém os neurônios representantes das variáveis do problema, a camada escondida dá a rede a capacidade de resolver problemas não lineares e a camada de saída representa a resposta da rede [18]. A rede MLP com apenas uma camada escondida pode ser vista na figura 3.



**Figura 3.** Rede MLP.

Fonte: <http://www.cerebromente.org.br/n05/tecnologia/rna.htm>

De acordo com a definição do neurônio artificial, o sinal só é propagado quando um limiar é atingido, por isso, na camada oculta e na de saída devem haver funções de ativação. Sendo a função de ativação da camada oculta não-linear, para garantir que a rede resolve problemas não-lineares, e a da camada de saída linear ou não. Tipicamente as funções não-lineares mais usadas são a sigmoide logística e a tangente hiperbólica [17].

### 2.3.1 Treinamento da MLP

A fim de obter bons resultados, todas as redes neurais devem ajustar os seus pesos sinápticos até que se atinja um conjunto ótimo. Nas redes MLP, o treinamento era um problema complexo pela existência de uma camada escondida com erros desconhecidos. Em 1974 o algoritmo *backpropagation* foi proposto e acelerou a adoção das redes MLP [18].

O algoritmo é uma generalização da regra delta, que é a responsável por ajustar os pesos de acordo com os resultados fornecidos e os calculados. Com a nova proposta, os erros da camada de saída são facilmente determinados ao se comparar as saídas desejadas com as saídas calculadas, já os erros das camadas ocultas são determinados ao se propagar os erros da camada de saída recursivamente [18]. Há duas fases no *backpropagation*:

- Fase *forward*: o sinal é propagado no sentido progressivo, ou seja, da camada de entrada para a camada de saída;
- Fase *backward*: o sinal é retro propagado da camada de saída em direção a camada de entrada, passando pelas camadas escondidas, enquanto os pesos sinápticos são ajustados de acordo com a regra delta generalizada.

Os passos do algoritmo *backpropagation* são os seguintes:

1. Na fase *forward*, a entrada líquida para os neurônios da camada escondida é calculada, onde a entrada para cada neurônio da camada é a soma ponderada dos inputs da rede, usando os pesos entre as conexões), como descreve a fórmula:

$$net_i = \sum_{j=0}^{Ninputs} w_{ij}x_{jj} \quad (1)$$

2. A mesma entrada líquida,  $net_i$ , passa por uma função não-linear e gera as saídas da camada intermediária;
3. As saídas da camada oculta passam a ser as entradas da camada de saída e são usados para calcular a entrada líquida da camada de saída, conforme a equação 1;
4. A função de ativação de ativação da camada de saída é aplicada na entrada líquida, gerando os resultados da rede
5. O erro é calculado ao se subtrair o resultado desejado de cada variável de saída com seu respectivo resultado calculado, conforme a equação 2:

$$e_i(n) = (d_i - y_i) \quad (2)$$

6. Já na fase *backward* do algoritmo, os pesos sinápticos são reajustados de acordo com a equação 3:

$$w_{ij}^m = w_{ij}^m(\text{antigo}) + \alpha \beta \delta_i^m f^{m-1}(\text{net}_j^{m-1})_i \quad (3)$$

Sendo  $w_{ij}^m$  os pesos que passarão por ajuste;  $\alpha$  a taxa de aprendizagem;  $\beta$  o momento, que auxilia na convergência do modelo;  $f^{m-1}(\text{net}_j^{m-1})$  os sinais emitidos pelas camadas anteriores; e  $\delta_i^m$ , de acordo com Valença [17][18], a sensibilidade da camada, que é definida pela fórmula 4:

$$\delta_i^{m-1} = (f^{m-1})'(\text{net}_j^{m-1}) \sum_{j=1}^{\text{Neurônios}} w_{ij}^m \delta_i^m \quad (4)$$

## 2.4 Reservoir Computing

A Reservoir Computing está inserida no contexto de redes neurais recorrentes, ou seja, possui pelo menos um laço de realimentação, gerando um comportamento dinâmico não-linear. A realimentação significa que a saída de um elemento pode influenciar, em parte, as entradas aplicadas àquele elemento. Estes laços também permitem que a rede adquira representação de estados, resolvendo problemas complexos como reconhecimento e processamento de voz, previsão não-linear, entre outros [3].

O processo de treinamento de redes recorrentes é muito complexo, para facilitar surgiu o paradigma da Reservoir Computing. Esta rede possui uma espécie de memória que combina valores presentes e passados, o que a torna uma ferramenta interessante para solução de problemas que envolvem séries temporais.

O diferencial da arquitetura da RC em relação a MLP consiste em um reservatório de neurônios com não-linearidade e recorrência; uma propriedade que permite que as entradas ressoem pela rede por um período de tempo, chamada de *Echo State Property*; e uma camada de saída que pode ser treinada com uma função de regressão ou classificação linear [3][16].

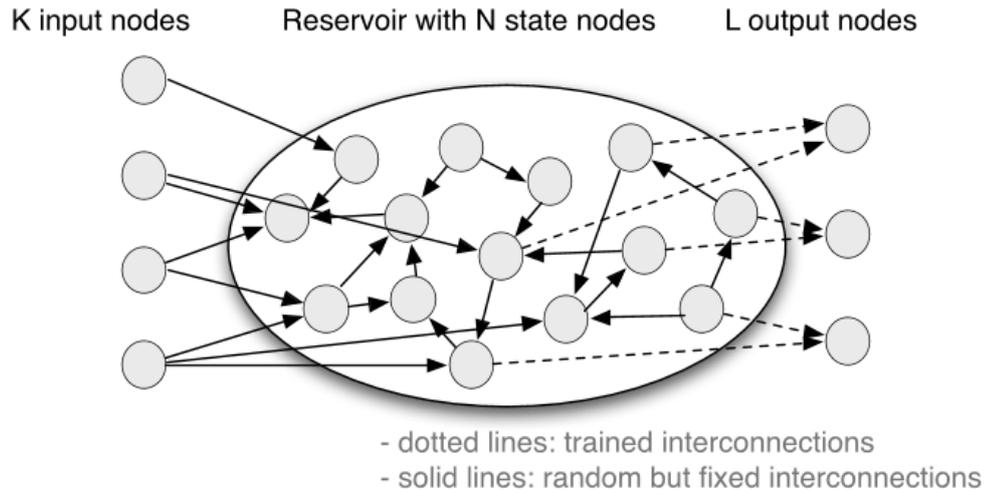
Detalhando melhor a arquitetura da RC, a camada de entrada é semelhante a MLP, ou seja, são as variáveis de entrada do problema. Conectado com os neurônios de entrada, há um reservatório de neurônios interconectados de maneira aleatória. O

reservatório possui a propriedade de eco e, graças a recorrência, este é capaz de manter um conjunto de transformações lineares além de valores atuais e passados da rede. Esta circulação de valores passados e presentes é que dá a rede a capacidade de resolver problemas temporais [16][3]. A arquitetura básica da rede RC pode ser vista na figura 4.

Inicialmente os pesos são gerados de maneira aleatória, para este trabalho, os valores estão no intervalo entre -1 e 1. Os estados dos neurônios do reservatório são inicializados com o valor 0. Para que o estado inicial não venha a interferir em estados futuros da rede, um processo de *Warm Up* [3] é realizado na rede antes do treinamento, este é definido pela seguinte equação:

$$x[k + 1] = f(W_{res}x[k] + W_{in}u[k]) \quad (5)$$

Onde  $W_{in}u[k]$  é a matriz que contém os valores da camada de entrada multiplicados pelos valores dos pesos entre a camada de entrada e o reservatório em um instante  $k$ .  $W_{res}x[k]$  é a matriz contendo os estados dos neurônios do reservatório no mesmo instante  $k$ . Após a soma das duas matrizes, a função de ativação de um neurônio é aplicada e o resultado é atribuído ao estado seguinte de um neurônio.



**Figura 4.** Rede RC. Linhas sólidas representam conexões que foram geradas aleatoriamente, mas não sofrem treinamento, diferente das linhas tracejadas que sofrem treinamento.

Fonte: <https://www.semanticscholar.org/paper/Band-pass-Reservoir-Computing-Wyffels-Schrauwen/3bf2a14ce4e9d2f17621c0ffb5e8544c9547>

Alguns parâmetros devem ser definidos para a rede RC antes do seu treinamento (vide seção 3.5.1). Após definida a configuração, o treinamento pode ser realizado.

#### 2.4.1 Treinamento da RC

O método de treinamento chamado de Pseudoinversa de Moore-Penrose foi escolhido para esta monografia.

Como os pesos e limiares são gerados de maneira aleatória em um momento inicial, o sistema passa de não-linear para linear, de acordo com a equação:

$$\mathbf{H}\beta = \mathbf{T} \quad (5)$$

Onde  $\mathbf{H}$  e  $\mathbf{T}$  são as matrizes das saídas do repositório e das saídas desejadas, respectivamente. [3][16] A determinação dos pesos entre a camada de saída e o reservatório é o mínimo quadrado para o sistema, definido a seguir:

$$\hat{\beta} = \mathbf{H}^+ \mathbf{T} \quad (5)$$

Onde  $H^+$  é a inversa generalizada da matriz  $H$ . O mínimo quadrado é o menor valor dentre todas as soluções e único. Este método permite que a rede tenha boa capacidade de generalização sem muito custo computacional [3].

Vale frisar que todos os nós da entrada e da saída estão conectados ao reservatório, mas nem todos os nós do reservatório estão conectados entre si [3].

# Capítulo 3

## Metodologia

Neste capítulo, a metodologia abordada para atingir os objetivos desta monografia é apresentada. A seção 3.1 descreve as bases de dados utilizadas. Na seção 3.2 consta o pré-processamento dos dados para uso nas redes neurais. A seção 3.3 aborda as representações de sazonalidade. As entradas e parâmetros das redes MLP e RC podem ser verificados nas seções 3.4 e 3.5, respectivamente. Por fim, os testes estatísticos que validaram os resultados dos algoritmos serão apresentados na seção 3.6.

### 3.1 Bases de dados

Para esta monografia utilizou-se dados de vazão mensal de quatro usinas hidroelétricas localizadas em bacias hidrográficas com características distintas distribuídas pelo Brasil. São elas: Bento Munhoz, Furnas, Três Marias e Tucuruí.

#### 3.1.1 Bento Munhoz

Inaugurada em 1980, a Hidroelétrica de Governador Bento Munhoz da Rocha Netto (figura 5), anteriormente chamada de Foz do Areia, está localizada no município de Pinhão no Paraná e tem capacidade de 1.676 MW de potência [6]. Dados de vazão mensal no rio Iguaçu foram coletados entre janeiro de 1970 e dezembro de 2003.



**Figura 5.** Usina Hidroelétrica Governador Bento Munhoz da Rocha Netto.

Fonte: Extraída de [6]

### 3.1.2 Furnas

Entrando em operação em 1963, a hidroelétrica de Furnas (figura 6) tem capacidade de 1.216 MW e está localizada entre os municípios de São José da Barra e São João Batista do Glória, em Minas Gerais [11]. Dados de vazão mensal no rio Grande foram coletados entre janeiro de 1931 e dezembro de 2003.



**Figura 6.** Usina Hidroelétrica de Furnas.

Fonte: Extraída de [11]

### 3.1.3 Três Marias

Localizada no município de Três Marias (figura 7), em Minas Gerais, a hidroelétrica de Três Marias teve o início de suas operações em 1962 e tem capacidade de 396MW. [4]. Dados de vazão mensal no rio São Francisco foram coletados entre janeiro de 1970 e dezembro de 2003.



**Figura 7.** Usina Hidroelétrica de Três Marias.

Fonte: Extraída de [4]

### 3.1.4 Tucuruí

A hidroelétrica de Tucuruí (figura 8), localizada no município de mesmo nome no estado do Pará, teve sua inauguração no ano de 1984 e possui capacidade de 8.370 MW [5]. Dados de vazão mensal no rio Tocantins foram coletados entre janeiro de 1970 e dezembro de 2003.



**Figura 8.** Usina Hidroelétrica de Tucuruí.

Fonte: Extraída de [5]

## 3.2 Pré-processamento dos dados

Após selecionar as bases de dados, é necessário transformar os dados brutos coletados em dados significativos para os algoritmos de redes neurais. Para isso, três passos foram necessários: normalização, defasagem e correlação dos dados.

### 3.2.1 Normalização dos dados

A normalização dos dados trata-se em transformar os dados de maneira que estes pertençam a um intervalo em comum e bem definido. Com esta transformação, as redes neurais darão a mesma importância para todos os dados, sem distinção.

A técnica de normalização mais utilizada em problemas de séries temporais é a transformação linear, descrita pela equação 6:

$$y_i = (b - a) \frac{x_i - x_{min}}{x_{max} - x_{min}} + a \quad (6)$$

Onde  $y_i$  é o valor da variável na posição  $i$  após a transformação,  $a$  e  $b$  são os extremos do intervalo que se deseja alcançar,  $x_{min}$  e  $x_{max}$  são os valores dos limites do intervalo original e  $x_i$  é o valor da variável original na posição  $i$  da base.

De acordo com Valença [18], os intervalos da normalização devem ser definidos de acordo com a função de ativação utilizadas nas técnicas de previsão.

Com isto, é possível evitar valores onde a derivada da função de ativação seja igual a zero. Como utilizamos a sigmoide logística como função de ativação, o intervalo de normalização utilizado nesta monografia foi de 0,15 a 0,85.

### 3.2.2 Defasagem dos dados

Após a normalização dos dados, realiza-se a defasagem dos dados. A defasagem é um processo auto regressivo consiste em gerar uma tabela onde os valores futuros são consequências dos valores passados [18].

A tabela 1 apresenta uma série temporal qualquer normalizada. Já a tabela 2 mostra os dados defasados e a tabela 3 faz a representação de uma rede neural com 3 neurônios de entrada e 1 de saída, onde o valor no instante  $t$  é a saída da rede e os valores  $(t-1)$ ,  $(t-2)$  e  $(t-3)$  são as suas entradas.

**Tabela 1.** Dados pós normalização

$t$
0.37
0.51
0.85
0.15
0.49
0.65

**Tabela 2.** Dados defasados três vezes no passado

$(t-3)$	$(t-2)$	$(t-1)$	$t$
			0.37
		0.37	0.51
	0.37	0.51	0.85
0.37	0.51	0.85	0.15
0.51	0.85	0.15	0.49
0.85	0.15	0.49	0.65
0.15	0.49	0.65	
0.49	0.65		
0.65			

Após eliminar as linhas que possuem dados faltantes, obteve-se a tabela 3.

**Tabela 3.** Tabela onde as três primeiras colunas representam as variáveis de entrada e a última a variável de saída

0.37	0.51	0.85	0.15
0.51	0.85	0.15	0.49
0.85	0.15	0.49	0.65

Para fins práticos desta monografia, optou-se por prever um mês baseado nos doze meses anteriores. Por este motivo, defasou-se cada uma das bases de dados em doze vezes no passado, de maneira que, neste ponto do pré-processamento dos dados, se tivesse doze entradas e uma saída na rede neural.

### 3.2.3 Correlação dos dados

Depois de defasar os dados, obteve-se doze entradas e uma saída. Porém, nem todas as doze entradas influenciam na previsão futura. Com isso, buscou-se um método para determinar quais das entradas são determinantes para o resultado final das redes neurais, este método foi a correlação entre as variáveis de entrada em relação a variável de saída. A correlação determina justamente o quão influente uma variável é em relação a outra. A tabela 4 foi gerada ao se usar a ferramenta de correlação do Microsoft Excel usando os dados da tabela 3 como entrada.

**Tabela 4.** Correlação entre os dados da tabela 3

	Coluna 1	Coluna 2	Coluna 3	Coluna 4
Coluna 1	1			
Coluna 2	-0.700551387	1		
Coluna 3	-0.299354557	-0.471164309	1	
Coluna 4	0.904323027	-0.32892464	-0.677987524	1

Observando-se a tabela 3, é verifica-se que as colunas 1, 2 e 3 são as variáveis de entrada e a coluna 4 é a de saída. Como almeja-se determinar as variáveis de entrada que mais influenciam na variável de saída, pode-se focar apenas na linha 4 da tabela 4 onde é apresentada a correlação da coluna 4 da tabela 3 em relação as colunas 1, 2 e 3.

O coeficiente de correlação pode assumir valores entre -1 e 1. Sendo:

$$\text{Correlação}(a, b) = \begin{cases} 0, & \text{se } a \text{ não possui qualquer relação com } b \\ 1, & \text{se } a \text{ possui relação totalmente proporcional com } b \\ -1, & \text{se } a \text{ possui relação inversamente proporcional com } b \end{cases}$$

Para cada base de dados escolheu-se um limiar para eliminar as variáveis de entrada que possuem baixo coeficiente de correlação em relação a saída, sendo esse limiar de  $\pm 0.30$  para a maioria das bases, ou seja, as variáveis que possuíam correlação entre -0.30 e +0.30 foram descartadas.

### 3.3 Representando sazonalidade

Para representar a sazonalidade no treinamento da rede, precisa-se adicionar uma certa quantidade de variáveis de entrada dependendo da representação escolhida. Para esta monografia, considerou-se três tipos de representação: 12 bits, 4 bits e representação com senos e cossenos.

Com 12 bits, adiciona-se 12 colunas onde cada posição representa o mês desejado na saída, como mostra a tabela 5.

**Tabela 5.** Representação de sazonalidade com 12 bits

<b>Janeiro</b>	1	0	0	0	0	0	0	0	0	0	0	0
<b>Fevereiro</b>	0	1	0	0	0	0	0	0	0	0	0	0
<b>Março</b>	0	0	1	0	0	0	0	0	0	0	0	0
<b>Abril</b>	0	0	0	1	0	0	0	0	0	0	0	0
<b>Maio</b>	0	0	0	0	1	0	0	0	0	0	0	0
<b>Junho</b>	0	0	0	0	0	1	0	0	0	0	0	0
<b>Julho</b>	0	0	0	0	0	0	1	0	0	0	0	0
<b>Agosto</b>	0	0	0	0	0	0	0	1	0	0	0	0
<b>Setembro</b>	0	0	0	0	0	0	0	0	1	0	0	0
<b>Outubro</b>	0	0	0	0	0	0	0	0	0	1	0	0
<b>Novembro</b>	0	0	0	0	0	0	0	0	0	0	1	0
<b>Dezembro</b>	0	0	0	0	0	0	0	0	0	0	0	1

Como 12 bits utiliza mais linhas de entrada, esta representação torna-se custosa para os algoritmos de redes neurais. Por esta razão, existe a representação

de 4 bits. Este valor é o mínimo necessário para representar os 12 meses, já que  $2^4 \text{ bits} = 16$  combinações, vide tabela 6.

**Tabela 6.** Representação de sazonalidade com 4 bits

<b>Janeiro</b>	0	0	0	0
<b>Fevereiro</b>	0	0	0	1
<b>Março</b>	0	0	1	0
<b>Abril</b>	0	0	1	1
<b>Mai</b>	0	1	0	0
<b>Junho</b>	0	1	0	1
<b>Julho</b>	0	1	1	0
<b>Agosto</b>	0	1	1	1
<b>Setembro</b>	1	0	0	0
<b>Outubro</b>	1	0	0	1
<b>Novembro</b>	1	0	1	0
<b>Dezembro</b>	1	0	1	1

Por fim, VALENÇA propôs a sazonalidade representada por uma combinação de funções periódicas, neste caso senos e cossenos [18]. Onde adiciona-se apenas duas entradas, na primeira constando os valores de seno e na segunda os valores de cosseno, sendo estes determinados pelas equações 7 e 8:

$$\text{variável 1} = \text{sen} \frac{2\pi N}{T} \quad (7)$$

$$\text{variável 2} = \text{cos} \frac{2\pi N}{T} \quad (8)$$

Sendo N o número do mês que se deseja representar (1, 2, 3 ... 12) e T igual ao tamanho do período, neste caso sempre vai corresponder a um ano, ou seja, T = 12 meses. A representação usando senos e cossenos encontra-se na tabela 7.

**Tabela 7.** Representação de sazonalidade com senos e cossenos, com os valores finais já calculados usando as equações (7) e (8)

<b>Janeiro</b>	0.5000	0.8660
<b>Fevereiro</b>	0.8660	0.5000
<b>Março</b>	1.0000	0.0000
<b>Abril</b>	0.8660	-0.5000
<b>Mai</b>	0.5000	-0.8660
<b>Junho</b>	0.0000	-1.0000
<b>Julho</b>	-0.5000	-0.8660
<b>Agosto</b>	-0.8660	-0.5000
<b>Setembro</b>	-1.0000	0.0000
<b>Outubro</b>	-0.8660	0.5000
<b>Novembro</b>	-0.5000	0.8660
<b>Dezembro</b>	0.0000	1.0000

Além das três representações de sazonalidade discutidas anteriormente, existe a representação de não-sazonalidade, ou seja, utiliza-se os dados como estão após a correlação.

### 3.4 Apresentando as variáveis de entrada à MLP

Após preparar os dados, é necessário ajustar o algoritmo de acordo com as peculiaridades de cada problema. Cada parâmetro pode causar efeitos diversos no resultado que a MLP prevê.

Antes da etapa de correlação das variáveis, cada base de dados apresentava 12 neurônios de entrada e 1 de saída. Após eliminar as entradas que influenciavam pouco na previsão de vazão com o processo de correlação e de adicionar as representações de sazonalidade para cada usina hidrelétrica, o número total de neurônios de entrada para cada cenário mudou, como pode ser visto na tabela 8.

**Tabela 8.** Quantidade de neurônios de entrada para cada cenário

	Sem Sazonalidade	12 bits	4 bits	Senos e cossenos
Bento Munhoz	6	18	10	8
Furnas	8	20	12	10
Três Marias	7	19	11	9
Tucuruí	10	22	14	12

### 3.4.1 Parâmetros utilizados no MLP

Além das entradas, o algoritmo da rede MLP precisa ter seus parâmetros ajustados para operar. Estes parâmetros foram escolhidos por apresentarem um resultado satisfatório em cenários parecidos de outros trabalhos [10][16], que também utilizaram o mesmo software desenvolvido na Linguagem Java e no ambiente de desenvolvimento do Eclipse por alunos da Escola Politécnica de Pernambuco. Os parâmetros foram definidos da seguinte maneira:

- Número de ciclos: 600;
- Número de simulações: 30;
- Algoritmo de treinamento: *backpropagation*;
- Taxa de aprendizagem: 0.85;
- Taxa de momento: 0.25;
- Quantidade de neurônios na camada escondida: 6;
- Quantidade de neurônios de saída: 1;
- Função de ativação da camada escondida: sigmoide logística;
- Função de ativação da camada de saída: linear;
- Divisão dos dados: 50% para treinamento, 25% para validação cruzada e 25% para testes

Vale ressaltar que a escolha do número de ciclos ser 600, não significa que o algoritmo vai rodar este número de ciclos. Se a rede atingir um ponto de convergência, os pesos ótimos serão guardados e a próxima simulação começará, o método de parada é a validação cruzada. O número de simulações é 30 porque com esta quantidade já se pode assumir que as médias obedecem a uma distribuição normal.

No total serão considerados quatro cenários por base de dados para a MLP, todos usando seis neurônios na camada escondida, mas variando na representação de sazonalidade, conforme a tabela 8.

Após cada experimento, o Erro Percentual Médio Absoluto (EPMA) é calculado, conforme a equação 7. Esta métrica pôde ser utilizada porque as variáveis de entrada da rede não possuem nenhum dado faltante ou nulo.

$$EPMA = \frac{\sum_{i=1}^n \left| \frac{d_i - y_i}{d_i} \right|}{n} \times 100\% \quad (7)$$

Onde  $d_i$  e  $y_i$  representam, respectivamente, a saída desejada e calculada da previsão e  $n$  representa o número de simulações. Os 30 erros percentuais médios de cada configuração das redes servem de entrada para os testes estatísticos que determinaram quais configurações obtiveram melhor desempenho.

## 3.5 Apresentando as variáveis de entrada à RC

Assim como a MLP, a RC também necessita que seus parâmetros sejam ajustados para operar. Dependendo dos valores dos parâmetros, resultados diferentes podem ser gerados.

Para a RC, os neurônios de entrada serão os mesmos utilizados na MLP, conforme a tabela 8.

### 3.5.1 Parâmetros utilizados no RC

O software da rede RC utilizado nesta monografia também foi desenvolvido na Linguagem Java e no ambiente de desenvolvimento do Eclipse por alunos da Escola Politécnica de Pernambuco. Além dos neurônios de entrada já definidos, os parâmetros foram utilizados da seguinte maneira:

- Número de simulações: 30;
- Método de treinamento: pseudo inversa de Moore-Penrose;
- Critério de parada do treinamento: *cross validation*;
- Taxa de conectividade: [0.3,0.4];

- Quantidade de ciclos de *Warm Up*: 10;
- Quantidade de neurônios no reservatório: 25 ou 35;
- Quantidade de neurônios de saída: 1 (previsão de 1 mês a frente);
- Função de ativação do reservatório: sigmoide logística;
- Função de ativação da camada de saída: linear;
- Divisão dos dados: 50% para treinamento, 25% para validação cruzada e 25% para testes

A quantidade de neurônios no reservatório vai ser testada para cada base, de maneira a se identificar qual funciona melhor para cada cenário. Já a taxa de conectividade aqui utilizada já foi testada em outros trabalhos [3][16] e retorna bons resultados com um menor tempo de processamento para previsão de vazão. Também em trabalhos passados [3][16] observou-se que 10 ciclos de *Warm Up* são suficientes para que o estado inicial da RC não afete no seu treinamento. O número de simulações é o mesmo utilizado no MLP pelo mesmo motivo, ou seja, os resultados das médias são normalmente distribuídos.

No total serão considerados oito cenários por base de dados para a RC, variando-se a quantidade de neurônios no reservatório (25 ou 35) e também a representação de sazonalidade, conforme a tabela 8. Assim como a MLP, o EPMA foi calculado no fim de cada experimento. Estes serão utilizados pelos testes estatísticos para comparar quais configurações geram melhores resultados.

### 3.6 Testes estatísticos

Para avaliar os resultados obtidos pelas duas redes neurais, ou seja, o conjunto de erros percentuais médios gerados ao fim de cada cenário, utilizou-se os seguintes testes estatísticos: Shapiro-Wilk, Teste F, Teste t de Student e Wilcoxon. Para automatizar tais testes, um script escrito no Software R foi utilizado. O nível de significância escolhido foi de 5%, já que é um valor usualmente adotado em problemas de engenharia [18]. Primeiro executa-se os testes de Shapiro-Wilk e F, sendo a hipótese zero dos dois atendidos, executa-se então o teste t de Student por ser paramétrico; caso contrário, o teste de não-paramétrico de Wilcoxon é executado.

### 3.6.1 Teste Shapiro-Wilk

Este é o primeiro teste a ser executado e tem a finalidade de verificar se os dados satisfazem uma distribuição normal [3][16]. Os seguintes passos são seguidos:

1. Determinar duas possibilidades:
  - a. H0: Amostra satisfaz uma distribuição normal.
  - b. H1: Amostra não satisfaz uma distribuição normal.
2. Caso o *p-value*, ou nível de significância, seja maior que 5% a amostra satisfaz H0, caso contrário satisfaz H1.

### 3.6.2 Teste F

Este é o segundo teste a ser executado e tem como objetivo verificar se a variância de duas amostras são estatisticamente iguais [3][16]. Os seguintes passos são seguidos:

1. Determinar duas possibilidades:
  - a. H0: Amostra apresenta variâncias estatisticamente idênticas.
  - b. H1: Amostra não apresenta variâncias estatisticamente idênticas.
2. Caso o *p-value* seja maior que 5% a amostra satisfaz H0, caso contrário satisfaz H1.

### 3.6.3 Teste t de Student

Dado que os testes de Shapiro-Wilk e F passaram, este é o terceiro teste a ser executado. Isso ocorre porque o teste T de Student é paramétrico, ou seja, as amostras precisam pertencer a uma distribuição normal e possuir variâncias estatisticamente idênticas [3][16]. Tendo isto determinado, este teste visa comparar a média entre duas amostras e segue os seguintes passos:

1. Determinar duas possibilidades:

a. H0: As médias entre as amostras são iguais.

b. H1: As médias entre as amostras são diferentes.

2. Caso o *p-value* seja maior que 5% a amostra satisfaz H0, caso contrário satisfaz H1.

### 3.6.4 Teste Wilcoxon

Dado que pelo menos um dos dois testes dentre o Shapiro-Wilk e F falham, o próximo teste a ser executado é o de Wilcoxon. Por ser um teste não-paramétrico, ele não faz nenhuma hipótese quanto a distribuição de probabilidade dos dados. O teste de Wilcoxon tem como objetivo verificar que, mesmo que as amostras sejam independentes, elas pertencem a uma população com mediana igual [3][16]. Os seguintes passos são seguidos:

1. Determinar duas possibilidades:

a. H0: As medianas entre as amostras são iguais.

b. H1: As medianas entre as amostras são diferentes.

2. Caso o *p-value* seja maior que 5% a amostra satisfaz H0, caso contrário satisfaz H1.

# Capítulo 4

## Resultados

Este capítulo visa apresentar os resultados obtidos aplicando as redes neurais MLP e RC na previsão de vazão mensal considerando critérios de sazonalidade. A metodologia descrita no capítulo 3 foi utilizada, abrangendo os cenários descritos para as quatro bases de dados: Bento Munhoz, Furnas, Três Marias e Tucuruí.

As seções 4.1 a 4.4 apresentam os experimentos realizados para cada uma das bases de dados, validando as representações de sazonalidade e utilizando tanto o cenário que usa a MLP quanto os dois cenários propostos com a RC. A seção 4.6 faz o comparativo entre os resultados para cada tipo de sazonalidade e determina qual a melhor representação para cada base de dado. Todas as sessões envolvem a previsão de vazão de um mês considerando doze meses passados, com os trinta erros percentuais médios gerados.

### 4.1 Experimentos com a base Bento Munhoz

Esta seção será dividida em quatro subseções, uma para cada representação de sazonalidade. Em cada subseção, será discutido qual configuração de rede neural obtém o melhor desempenho para a sazonalidade que está em questão em relação a Base de Bento Munhoz.

#### 4.1.1 Hidroelétrica Bento Munhoz: sem representação de sazonalidade

Para a hidrelétrica de Bento Munhoz sem nenhuma sazonalidade codificada, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 9.

**Tabela 9.** Média de EPMA para cada cenário da hidrelétrica de Bento Munhoz (BM) sem representação de sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA (%)	Desvio Padrão
<b>1-BM</b>	MLP	6	-	21.88	3.52
<b>2-BM</b>	RC	-	25	24.93	1.44
<b>3-BM</b>	RC	-	35	27.01	1.40

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

#### 4.1.1.1 Comparação entre o Cenário 1-BM e o Cenário 2-BM sem sazonalidade

Aplicando o teste de Shapiro-Wilk no Cenário 1-BM, obteve-se um *p-value* menor que 5%, o que indica que as amostras não são normalmente distribuídas. Aplicando o mesmo teste no Cenário 2-BM, obteve-se um *p-value* maior que 5%, indicando que as amostras têm uma distribuição normal.

Como o Cenário 1-BM falha no teste de Shapiro-Wilk, aplicou-se o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* menor que 5%, o que indica que a mediana entre as amostras não é igual.

O Cenário 1-BM foi o vencedor por ter um EPMA menor. Assim, o Cenário 2-BM foi descartado. A tabela 10 exhibe os resultados obtidos nos testes.

**Tabela 10.** Resultados dos *p-value* para os cenários 1-BM e 2-BM em Bento Munhoz sem representação de sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>1-BM</b>	0.0004	9.0e-07
<b>2-BM</b>	0.11	

#### 4.1.1.2 Comparação entre o Cenário 1-BM e o Cenário 3-BM sem sazonalidade

Como o Cenário 1-BM falha no teste de Shapiro-Wilk, aplicou-se o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* menor que 5%, o que indica que a mediana entre as amostras é estatisticamente diferente. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois, ou seja, o **Cenário 1-BM** foi o vencedor com um EPMA de 21.88%. A tabela 11 exibe os resultados dos testes.

**Tabela 11.** Resultados dos *p-value* para os cenários 1-BM e 3-BM em Bento Munhoz sem representação de sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>2-BM</b>	0.0004	3.574e-10
<b>3-BM</b>	0.07	

#### 4.1.2 Hidroelétrica Bento Munhoz: sazonalidade de 12 bits

Para a hidrelétrica de Bento Munhoz utilizando 12 bits de sazonalidade, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 12.

**Tabela 12.** Média de EPMA's para cada cenário da hidrelétrica de Bento Munhoz (BM) com 12 bits de sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA	Desvio Padrão
<b>1-BM</b>	MLP	6	-	22.42	3.16
<b>2-BM</b>	RC	-	25	21.16	1.02
<b>3-BM</b>	RC	-	35	22.16	1.36

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA's para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

#### 4.1.2.1 Comparação entre o Cenário 1-BM e o Cenário 2-BM com sazonalidade de 12 bits

Aplicando o teste de Shapiro-Wilk no Cenário 1-BM, obteve-se um *p-value* menor que 5%, o que indica que as amostras não são normalmente distribuídas. Aplicando o mesmo teste no Cenário 2-BM, obteve-se um *p-value* maior que 5%, provando que as amostras têm uma distribuição normal.

Como o Cenário 1-BM falha no teste de Shapiro-Wilk, aplicou-se o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* maior que 5%, o que indica que a mediana entre as amostras é estatisticamente igual. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O Cenário 2-BM foi o vencedor com um EPMA de 21.16%. A tabela 13 exhibe os resultados obtidos nos testes.

**Tabela 13.** Resultados dos *p-value* para os cenários 1-BM e 2-BM em Bento Munhoz com 12 bits de sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>1-BM</b>	0.04	0.15
<b>2-BM</b>	0.74	

#### 4.1.2.2 Comparação entre o Cenário 2-BM e o Cenário 3-BM com sazonalidade de 12 bits

Já se sabe que o teste de Shapiro-Wilk para o cenário 2-BM obteve um *p-value* maior que 5%. Ao repetir o processo para o cenário 3-BM, obtém-se também um valor maior que 5%, indicando que as amostras têm uma distribuição normal. Aplica-se, então, o teste F que resultou em um *p-value* maior que 5%, provando que as amostras possuem variâncias estatisticamente iguais.

Com as duas hipóteses anteriores atendidas, executa-se o teste t de Student, que por ter um *p-value* menor que 5%, indicou que as amostras não possuem médias estatisticamente iguais. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O **cenário 2-BM** foi o vencedor com um EPMA de 21.16%. A tabela 14 exhibe os resultados dos testes.

**Tabela 14.** Resultados dos *p-value* para os cenários 2-BM e 3-BM em Bento Munhoz com 12 bits de sazonalidade

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
<b>2-BM</b>	0.74	0.12	0.002
<b>3-BM</b>	0.95		

#### 4.1.3 Hidroelétrica Bento Munhoz: sazonalidade de 4 bits

Para a hidrelétrica de Bento Munhoz utilizando 4 bits de sazonalidade, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 15.

**Tabela 15.** Média de EPMA para cada cenário da hidrelétrica de Bento Munhoz (BM) com 4 bits de sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA	Desvio Padrão
<b>1-BM</b>	MLP	6	-	22.46	3.54
<b>2-BM</b>	RC	-	25	21.46	0.84
<b>3-BM</b>	RC	-	35	23.08	1.47

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

##### 4.1.3.1 Comparação entre o cenário 1-BM e o cenário 2-BM com sazonalidade de 4 bits

Aplicando-se o teste de Shapiro-Wilk tanto no cenário 1-BM quanto no cenário 2-BM, obteve-se um *p-value* maior que 5%. Aplica-se, então, o teste F que resultou em um *p-value* menor que 5%, indicando que as amostras não possuem variâncias estatisticamente iguais.

Como o teste F falha, aplicou-se, então, o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* maior que 5%, o que indica que a mediana entre as amostras é igual. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O cenário 2-BM foi o vencedor com um EPMA de 21.46%. A tabela 16 exibe os resultados obtidos nos testes.

**Tabela 16.** Resultados dos  $p$ -value para os cenários 1-BM e 2-BM em Bento Munhoz com 4 bits de sazonalidade

Cenário	Shapiro-Wilk	Teste F	Wilcoxon
<b>1-BM</b>	0.28	1.44e-11	0.50
<b>2-BM</b>	0.90		

#### 4.1.3.2 Comparação entre o cenário 2-BM e o cenário 3-BM com sazonalidade de 4 bits

Já se sabe que o teste de Shapiro-Wilk para o cenário 2-BM gerou um  $p$ -value maior que 5%. Ao repetir o processo para o cenário 3-BM, também se obtém um valor maior que 5%, indicando que as amostras têm uma distribuição normal. Aplica-se, então, o teste F que resultou em um  $p$ -value menor que 5%, indicando que as amostras não possuem variâncias estatisticamente iguais.

Como o teste F falha, aplicou-se, então, o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um  $p$ -value menor que 5%, o que indica que a mediana entre as amostras não é igual. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O **cenário 2-BM** foi o vencedor com um EPMA de 21.46%. A tabela 17 exibe os resultados obtidos nos testes.

**Tabela 17.** Resultados dos  $p$ -value para os cenários 2-BM e 3-BM em Bento Munhoz com 4 bits de sazonalidade

Cenário	Shapiro-Wilk	Teste F	Wilcoxon
<b>2-BM</b>	0.90	0.003	4.64e-07
<b>3-BM</b>	0.06		

#### 4.1.4 Hidroelétrica Bento Munhoz: sazonalidade de senos e cossenos

Para a hidrelétrica de Bento Munhoz utilizando representação de senos e cossenos para sazonalidade, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 18.

**Tabela 18.** Média de EPMA para cada cenário da hidrelétrica de Bento Munhoz (BM) utilizando senos e cossenos para representar sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA	Desvio Padrão
<b>1-BM</b>	MLP	6	-	22.03	2.90
<b>2-BM</b>	RC	-	25	21.63	0.86
<b>3-BM</b>	RC	-	35	23.09	1.09

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

#### 4.1.4.1 Comparação entre o cenário 1-BM e o cenário 2-BM com sazonalidade de senos e cossenos

Aplicando-se o teste de Shapiro-Wilk tanto no cenário 1-BM quanto no cenário 2-BM, obteve-se um *p-value* maior que 5%. Aplica-se, então, o teste F que resultou em um *p-value* menor que 5%, indicando que as amostras não possuem variâncias estatisticamente iguais.

Como o teste F falha, aplicou-se, então, o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* maior que 5%, o que indica que a mediana entre as amostras é igual. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O cenário 2-BM foi o vencedor com um EPMA de 21.63%. A tabela 19 exibe os resultados obtidos nos testes.

**Tabela 19.** Resultados dos *p-value* para os cenários 1-BM e 2-BM em Bento Munhoz utilizando senos e cossenos para representar sazonalidade

Cenário	Shapiro-Wilk	Teste F	Wilcoxon
<b>1-BM</b>	0.41	3.91e-09	0.44
<b>2-BM</b>	0.14		

#### 4.1.4.2 Comparação entre o cenário 2-BM e o cenário 3-BM com sazonalidade de senos e cossenos

Já se sabe que o teste de Shapiro-Wilk para o cenário 2-BM obteve um *p-value* maior que 5%. Ao repetir o processo para o cenário 3-BM, obtém-se também um valor maior que 5%, indicando que as amostras têm uma distribuição normal. Aplica-se, então, o teste F que resultou em um *p-value* maior que 5%, provando que as amostras possuem variâncias estatisticamente iguais. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois.

Com as duas hipóteses anteriores atendidas, executa-se então o teste t de Student, que por ter um *p-value* menor que 5%, indicou que as amostras não possuem médias estatisticamente iguais. O **cenário 2-BM** foi o vencedor com um EPMA de 21.63%. A tabela 20 exibe os resultados dos testes.

**Tabela 20.** Resultados dos *p-value* para os cenários 2-BM e 3-BM em Bento Munhoz utilizando senos e cossenos para representar sazonalidade

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
<b>2-BM</b>	0.14	0.19	4.51e-07
<b>3-BM</b>	0.51		

## 4.2 Experimentos com a base Furnas

Esta seção será dividida em quatro subseções, uma para cada representação de sazonalidade. Em cada subseção, será discutido qual configuração de rede neural obtém o melhor desempenho para a sazonalidade que está em questão em relação a Base de Furnas.

### 4.2.1 Hidroelétrica Furnas: sem representação de sazonalidade

Para a hidrelétrica de Furnas sem nenhuma sazonalidade codificada, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 21.

**Tabela 21.** Média de EPMA para cada cenário da hidrelétrica de Furnas (F) sem representação de sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA	Desvio Padrão
<b>1-F</b>	MLP	6	-	17.02	3.42
<b>2-F</b>	RC	-	25	14.83	0.35
<b>3-F</b>	RC	-	35	14.82	0.48

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

#### 4.2.1.1 Comparação entre o cenário 1-F e o cenário 2-F sem sazonalidade

Aplicando-se o teste de Shapiro-Wilk tanto no cenário 1-F quanto no cenário 2-F, obteve-se um *p-value* maior que 5%. Aplica-se, então, o teste F que resultou em um *p-value* menor que 5%, indicando que as amostras não possuem variâncias estatisticamente iguais.

Como o teste F falha, aplicou-se, então, o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* maior que 5%, o que indica que a mediana entre as amostras é igual. A tabela 22 exhibe os resultados obtidos nos testes. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O cenário 2-F foi o vencedor com um EPMA de 14.83%.

**Tabela 22.** Resultados dos *p-value* para os cenários 1-F e 2-F em Furnas sem bits de sazonalidade

Cenário	Shapiro-Wilk	Teste F	Wilcoxon
<b>1-F</b>	0.005	2.2e-16	0.006
<b>2-F</b>	0.48		

#### 4.2.1.2 Comparação entre o cenário 2-F e o cenário 3-F sem sazonalidade

Já se sabe que o teste de Shapiro-Wilk para o cenário 2-F obteve um *p-value* maior que 5%. Ao repetir o processo para o cenário 3-F, obtém-se também um valor maior que 5%, indicando que as amostras têm uma distribuição normal. Aplica-se,

então, o teste F que resultou em um *p-value* maior que 5%, provando que as amostras possuem variâncias estatisticamente iguais.

Com as duas hipóteses anteriores atendidas, executa-se então o teste t de Student, que por ter um *p-value* maior que 5%, indicou que as amostras possuem médias estatisticamente iguais. Como os valores dos EPMA são próximos e, de acordo com os testes, eles são estatisticamente equivalentes, o **cenário 2-F** é o vencedor por ser mais performático já que usa menos neurônios no reservatório. A tabela 23 exibe os resultados dos testes.

**Tabela 23.** Resultados dos *p-value* para os cenários 2-F e 3-F em Furnas sem bits sazonalidade

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
<b>2-F</b>	0.48	0.09	0.96
<b>3-F</b>	0.58		

#### 4.2.2 Hidroelétrica Furnas: sazonalidade de 12 bits

Para a hidrelétrica de Furnas utilizando 12 bits de sazonalidade, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 24.

**Tabela 24.** Média de EPMA para cada cenário da hidrelétrica de Furnas (F) com 12 bits de sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA	Desvio Padrão
<b>1-F</b>	MLP	6	-	14.03	1.77
<b>2-F</b>	RC	-	25	12.36	0.36
<b>3-F</b>	RC	-	35	12.60	0.30

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

#### 4.2.2.1 Comparação entre o cenário 1-F e o cenário 2-F com sazonalidade de 12 bits

Aplicando o teste de Shapiro-Wilk no cenário 1-F, obteve-se um *p-value* maior que 5%, o que indica que as amostras são normalmente distribuídas. Aplicando o mesmo teste no cenário 2-F, obteve-se um *p-value* menor que 5%, indicando que as amostras não têm uma distribuição normal.

Como o cenário 2-F falha no teste de Shapiro-Wilk, aplicou-se o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* menor que 5%, o que indica que a mediana entre as amostras não é igual.

O cenário 2-F foi o vencedor por possuir um menor EPMA. Assim, o cenário 1-F foi descartado. A tabela 25 exibe os resultados obtidos nos testes.

**Tabela 25.** Resultados dos *p-value* para os cenários 1-F e 2-F em Furnas com 12 bits de sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>1-F</b>	0.10	1.54e-07
<b>2-F</b>	0.0004	

#### 4.2.2.2 Comparação entre o cenário 2-F e o cenário 3-F com sazonalidade de 12 bits

Já se sabe que o teste de Shapiro-Wilk para o cenário 2-F gerou um *p-value* menor que 5%. Aplicou-se, então, o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* menor que 5%, o que indica que a mediana entre as amostras é estatisticamente diferente. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois, ou seja, o **Cenário 2-F** foi o vencedor com um EPMA de 12.36%. A tabela 26 exibe os resultados dos testes.

**Tabela 26.** Resultados dos *p-value* para os cenários 2-F e 3-F em Furnas com 4 bits de sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>2-F</b>	0.0004	0.0008
<b>3-F</b>	0.11	

#### 4.2.3 Hidroelétrica Furnas: sazonalidade de 4 bits

Para a hidrelétrica de Furnas utilizando 4 bits de sazonalidade, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 27.

**Tabela 27.** Média de EPMA para cada cenário da hidrelétrica de Furnas (F) com 4 bits de sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA	Desvio Padrão
<b>1-F</b>	MLP	6	-	15.23	2.32
<b>2-F</b>	RC	-	25	13.22	0.49
<b>3-F</b>	RC	-	35	12.95	0.31

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

##### 4.2.3.1 Comparação entre o cenário 1-F e o cenário 2-F com sazonalidade de 4 bits

Aplicando-se o teste de Shapiro-Wilk tanto no cenário 1-F quanto no cenário 2-F, obteve-se um *p-value* maior que 5%. Aplica-se, então, o teste F que resultou em um *p-value* menor que 5%, indicando que as amostras não possuem variâncias estatisticamente iguais.

Como o teste F falha, aplicou-se, então, o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* maior que 5%, o que indica que a mediana entre as amostras é igual. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O cenário 2-F foi o vencedor com um EPMA de 13.22%. A tabela 28 exibe os resultados obtidos nos testes.

**Tabela 28.** Resultados dos *p-value* para os cenários 1-F e 2-F em Furnas com 4 bits de sazonalidade

Cenário	Shapiro-Wilk	Teste F	Wilcoxon
<b>1-F</b>	0.13	8.35e-13	6.03e-05
<b>2-F</b>	0.07		

#### 4.2.3.2 Comparação entre o cenário 2-F e o cenário 3-F com sazonalidade de 4 bits

Já se sabe que o teste de Shapiro-Wilk para o cenário 2-F obteve um *p-value* maior que 5%. Ao repetir o processo para o cenário 3-F, obtém-se também um valor maior que 5%, indicando que as amostras têm uma distribuição normal. Aplica-se, então, o teste F que resultou em um *p-value* maior que 5%, provando que as amostras possuem variâncias estatisticamente iguais.

Com as duas hipóteses anteriores atendidas, executa-se então o teste t de Student, que por ter um *p-value* maior que 5%, indicou que as amostras possuem médias estatisticamente iguais. Como os valores dos EPMA são próximos e, de acordo com os testes, eles são estatisticamente equivalentes, o **cenário 2-F** é o vencedor por ser mais performático já que usa menos neurônios no reservatório. A tabela 29 exibe os resultados dos testes.

**Tabela 29.** Resultados dos *p-value* para os cenários 2-F e 3-F em Furnas com 4 bits de sazonalidade

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
<b>2-F</b>	0.07	0.01	0.01
<b>3-F</b>	0.87		

#### 4.2.4 Hidroelétrica Furnas: sazonalidade de senos e cossenos

Para a hidrelétrica de Furnas utilizando representação de senos e cossenos para sazonalidade, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 30.

**Tabela 30.** Média de EPMA's para cada cenário da hidrelétrica de Furnas (F) utilizando senos e cossenos para representar sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA	Desvio Padrão
<b>1-F</b>	MLP	6	-	14.22	2.28
<b>2-F</b>	RC	-	25	12.70	0.18
<b>3-F</b>	RC	-	35	12.88	0.34

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA's para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

#### 4.2.4.1 Comparação entre o cenário 1-F e o cenário 2-F com sazonalidade de senos e cossenos

Aplicando o teste de Shapiro-Wilk no cenário 1-F, obteve-se um *p-value* menor que 5%, o que indica que as amostras não são normalmente distribuídas. Aplicando o mesmo teste no cenário 2-F, obteve-se um *p-value* maior que 5%, provando que as amostras têm uma distribuição normal.

Como o cenário 1-F falha no teste de Shapiro-Wilk, aplicou-se o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* menor que 5%, o que indica que a mediana entre as amostras é diferente. Então vence o que apresenta menor EPMA, que neste caso foi o cenário 2-F. A tabela 31 exhibe os resultados obtidos nos testes.

**Tabela 31.** Resultados dos *p-value* para os cenários 1-F e 2-F em Furnas utilizando senos e cossenos para representar sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>1-F</b>	0.001	0.002
<b>2-F</b>	0.67	

#### 4.2.4.2 Comparação entre o cenário 2-F e o cenário 3-F com sazonalidade de senos e cossenos

Já se sabe que o teste de Shapiro-Wilk para o cenário 1-F gerou um *p-value* maior que 5%. Ao repetir o processo para o cenário 3-F, também se obtém um valor maior que 5%, indicando que as amostras têm uma distribuição normal. Aplica-se, então, o teste F que resultou em um *p-value* menor que 5%, indicando que as amostras não possuem variâncias estatisticamente iguais.

Como o teste F falha, aplicou-se, então, o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* maior que 5%, o que indica que a mediana entre as amostras é igual. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O **cenário 2-F** foi o vencedor com um EPMA de 12.70%. A tabela 32 exibe os resultados obtidos nos testes.

**Tabela 32.** Resultados dos *p-value* para os cenários 2-F e 3-F em Furnas usando senos e cossenos para representar sazonalidade

Cenário	Shapiro-Wilk	Teste F	Wilcoxon
<b>2-F</b>	0.67	0.001	0.06
<b>3-F</b>	0.30		

### 4.3 Experimentos com a base Três Marias

Esta seção será dividida em quatro subseções, uma para cada representação de sazonalidade. Em cada subseção, será discutido qual configuração de rede neural obtém o melhor desempenho para a sazonalidade que está em questão em relação a Base de Três Marias.

#### 4.3.1 Hidroelétrica Três Marias: sem representação de sazonalidade

Para a hidrelétrica de Três Marias sem nenhuma sazonalidade codificada, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 33.

**Tabela 33.** Média de EPMA para cada cenário da hidrelétrica de Três Marias (TM) sem representação de sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA	Desvio Padrão
<b>1-TM</b>	MLP	6	-	17.90	3.23
<b>2-TM</b>	RC	-	25	14.45	0.73
<b>3-TM</b>	RC	-	35	14.14	0.80

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

#### 4.3.1.1 Comparação entre o cenário 1-TM e o cenário 2-TM sem sazonalidade

Aplicando-se o teste de Shapiro-Wilk tanto no cenário 1-TM quanto no cenário 2-TM, obteve-se um *p-value* maior que 5%. Aplica-se, então, o teste F que resultou em um *p-value* menor que 5%, indicando que as amostras não possuem variâncias estatisticamente iguais.

Como o teste F falha, aplicou-se, então, o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* maior que 5%, o que indica que a mediana entre as amostras é igual. Então o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O cenário 2-TM foi o vencedor com um EPMA de 14.45%. A tabela 34 exibe os resultados obtidos nos testes.

**Tabela 34.** Resultados dos *p-value* para os cenários 1-TM e 2-TM em Três Marias sem representação de sazonalidade

Cenário	Shapiro-Wilk	Teste F	Wilcoxon
<b>1-TM</b>	0.13	4.37e-12	4.64e-07
<b>2-TM</b>	0.62		

#### 4.3.1.2 Comparação entre o cenário 2-TM e o cenário 3-TM sem sazonalidade

Já se sabe que o teste de Shapiro-Wilk para o cenário 2-TM obteve um *p-value* maior que 5%. Ao repetir o processo para o cenário 3-TM, obtém-se também um valor maior que 5%, indicando que as amostras têm uma distribuição normal. Aplica-se, então, o teste F que resultou em um *p-value* maior que 5%, provando que as amostras possuem variância estatisticamente iguais.

Com as duas hipóteses anteriores atendidas, executa-se então o teste t de Student, que por ter um *p-value* maior que 5%, indicou que as amostras possuem médias estatisticamente iguais. Como os valores dos EPMA são próximos e, de acordo com os testes, eles são estatisticamente equivalentes, o **cenário 2-TM** é o vencedor por ser mais performático já que usa menos neurônios no reservatório. A tabela 35 exhibe os resultados dos testes.

**Tabela 35.** Resultados dos *p-value* para os cenários 2-TM e 3-TM em Três Marias sem representação de sazonalidade

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
<b>2-TM</b>	0.62	0.61	0.12
<b>3-TM</b>	0.05		

#### 4.3.2 Hidroelétrica Três Marias: sazonalidade de 12 bits

Para a hidroelétrica de Três Marias utilizando 12 bits de sazonalidade, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 36.

**Tabela 36.** Média de EPMA para cada cenário da hidroelétrica de Três Marias (TM) com 12 bits de sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA	Desvio Padrão
<b>1-TM</b>	MLP	6	-	16.05	3.84
<b>2-TM</b>	RC	-	25	13.48	0.76
<b>3-TM</b>	RC	-	35	14.61	1.07

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA's para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

#### 4.3.2.1 Comparação entre o cenário 1-TM e o cenário 2-TM com sazonalidade de 12 bits

Aplicando o teste de Shapiro-Wilk no cenário 1-TM, obteve-se um *p-value* menor que 5%, o que indica que as amostras não são normalmente distribuídas. Aplicando o mesmo teste no cenário 2-TM, obteve-se um *p-value* maior que 5%, indicando que as amostras têm uma distribuição normal.

Como o cenário 1-TM falha no teste de Shapiro-Wilk, aplicou-se o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* menor que 5%, o que indica que a mediana entre as amostras não é igual.

O cenário 2-TM foi o vencedor pois possui um menor EPMA. Assim, o cenário 1-TM foi descartado. A tabela 37 exhibe os resultados obtidos nos testes.

**Tabela 37.** Resultados dos *p-value* para os cenários 1-TM e 2-TM em Três Marias com 12 bits de sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>1-TM</b>	1.24e-05	7.73e-05
<b>2-TM</b>	0.77	

#### 4.3.2.2 Comparação entre o cenário 2-TM e o cenário 3-TM com sazonalidade de 12 bits

Já se sabe que o teste de Shapiro-Wilk para o cenário 2-TM obteve um *p-value* maior que 5%. Ao repetir o processo para o cenário 3-TM, obtém-se também um valor maior que 5%, indicando que as amostras têm uma distribuição normal. Aplica-se, então, o teste F que resultou em um *p-value* maior que 5%, provando que as amostras possuem variâncias estatisticamente iguais.

Com as duas hipóteses anteriores atendidas, executa-se então o teste t de Student, que por ter um *p-value* maior que 5%, indicou que as amostras possuem

médias estatisticamente iguais. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O **cenário 2-TM** foi o vencedor com um EPMA de 13.48%. A tabela 38 exibe os resultados dos testes.

**Tabela 38.** Resultados dos *p-value* para os cenários 2-TM e 3-TM em Três Marias com 12 bits de sazonalidade

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
<b>2-TM</b>	0.77	0.07	1.85e-05
<b>3-TM</b>	0.88		

#### 4.3.3 Hidroelétrica Três Marias: sazonalidade de 4 bits

Para a hidrelétrica de Três Marias utilizando 4 bits de sazonalidade, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 39.

**Tabela 39.** Média de EPMA para cada cenário da hidrelétrica de Três Marias (TM) com 4 bits de sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA	Desvio Padrão
<b>1-TM</b>	MLP	6	-	19.15	5.37
<b>2-TM</b>	RC	-	25	13.88	0.84
<b>3-TM</b>	RC	-	35	14.18	0.75

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

##### 4.3.3.1 Comparação entre o cenário 1-TM e o cenário 2-TM com sazonalidade de 4 bits

Aplicando o teste de Shapiro-Wilk no cenário 1-TM, obteve-se um *p-value* menor que 5%, o que indica que as amostras não são normalmente distribuídas. Aplicando o mesmo teste no cenário 2-TM, obteve-se um *p-value* maior que 5%, indicando que as amostras têm uma distribuição normal.

Como o cenário 1-TM falha no teste de Shapiro-Wilk, aplicou-se o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* menor que 5%, o que indica que a mediana entre as amostras não é igual.

O cenário 2-TM foi o vencedor pois possui um menor EPMA. Assim, o cenário 1-TM foi descartado. A tabela 40 exibe os resultados obtidos nos testes.

**Tabela 40.** Resultados dos *p-value* para os cenários 1-TM e 2-TM em Três Marias com 4 bits de sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>1-TM</b>	0.0002	4.28e-08
<b>2-TM</b>	0.11	

#### 4.3.3.2 Comparação entre o cenário 2-TM e o cenário 3-TM com sazonalidade de 4 bits

Já se sabe que o teste de Shapiro-Wilk para o cenário 2-TM obteve um *p-value* maior que 5%. Ao repetir o processo para o cenário 3-TM, obtém-se também um valor maior que 5%, indicando que as amostras têm uma distribuição normal. Aplica-se, então, o teste F que resultou em um *p-value* maior que 5%, provando que as amostras possuem variâncias estatisticamente iguais.

Com as duas hipóteses anteriores atendidas, executa-se então o teste t de Student, que por ter um *p-value* maior que 5%, indicou que as amostras possuem médias estatisticamente iguais. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O **cenário 2-TM** foi o vencedor com um EPMA de 13.88%. A tabela 41 exibe os resultados dos testes.

**Tabela 41.** Resultados dos *p-value* para os cenários 2-TM e 3-TM em Três Marias com 4 bits de sazonalidade

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
<b>2-TM</b>	0.11	0.53	0.14
<b>3-TM</b>	0.58		

#### 4.3.4 Hidroelétrica Três Marias: sazonalidade de senos e cossenos

Para a hidrelétrica de Três Marias utilizando representação de senos e cossenos para sazonalidade, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 42.

**Tabela 42.** Média de EPMA para cada cenário da hidrelétrica de Três Marias (TM) utilizando senos e cossenos para representar sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA (%)	Desvio Padrão
<b>1-TM</b>	MLP	6	-	16.36	3.23
<b>2-TM</b>	RC	-	25	13.77	0.74
<b>3-TM</b>	RC	-	35	14.68	1.20

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

##### 4.3.4.1 Comparação entre o cenário 1-TM e o cenário 2-TM com sazonalidade de senos e cossenos

Aplicando o teste de Shapiro-Wilk no cenário 1-TM, obteve-se um *p-value* menor que 5%, o que indica que as amostras não são normalmente distribuídas. Aplicando o mesmo teste no cenário 2-TM, obteve-se um *p-value* maior que 5%, indicando que as amostras têm uma distribuição normal.

Como o cenário 1-TM falha no teste de Shapiro-Wilk, aplicou-se o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* menor que 5%, o que indica que a mediana entre as amostras não é igual.

O cenário 2-TM foi o vencedor pois um menor EPMA. Assim, o cenário 1-TM foi descartado. A tabela 43 exibe os resultados obtidos nos testes.

**Tabela 43.** Resultados dos  $p$ -value para os cenários 1-TM e 2-TM em Três Marias usando senos e cossenos para representar sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>1-TM</b>	0.001	3.77e-05
<b>2-TM</b>	0.62	

#### 4.3.4.2 Comparação entre o cenário 2-TM e o cenário 3-TM com sazonalidade de senos e cossenos

Já se sabe que o teste de Shapiro-Wilk para o cenário 2-TM obteve um  $p$ -value maior que 5%. Ao repetir o processo para o cenário 3-TM, obtém-se também um valor maior que 5%, indicando que as amostras têm uma distribuição normal. Aplica-se, então, o teste F que resultou em um  $p$ -value maior que 5%, provando que as amostras possuem variâncias estatisticamente iguais.

Com as duas hipóteses anteriores atendidas, executa-se então o teste t de Student, que por ter um  $p$ -value menor que 5%, indicou que as amostras não possuem médias estatisticamente iguais. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O **cenário 2-TM** foi o vencedor com um EPMA de 13.77%. A tabela 44 exibe os resultados dos testes.

**Tabela 44.** Resultados dos  $p$ -value para os cenários 2-TM e 3-TM em Três Marias usando senos e cossenos para representar sazonalidade

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
<b>2-TM</b>	0.62	0.01	0.0009
<b>3-TM</b>	0.31		

## 4.4 Experimentos com a base Tucuruí

Esta seção será dividida em quatro subseções, uma para cada representação de sazonalidade. Em cada subseção, será discutido qual configuração de rede neural obtém o melhor desempenho para a sazonalidade que está em questão em relação a Base de Tucuruí.

#### 4.4.1 Hidroelétrica Tucuruí: sem representação de sazonalidade

Para a hidroelétrica de Tucuruí sem nenhuma sazonalidade codificada, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 45.

**Tabela 45.** Média de EPMA para cada cenário da hidroelétrica de Tucuruí (T) sem representação de sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA (%)	Desvio Padrão
1-T	MLP	6	-	14.04	2.03
2-T	RC	-	25	10.73	0.58
3-T	RC	-	35	11.02	0.65

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

##### 4.4.1.1 Comparação entre o cenário 1-T e o cenário 2-T sem sazonalidade

Aplicando o teste de Shapiro-Wilk no cenário 1-T, obteve-se um *p-value* maior que 5%, o que indica que as amostras são normalmente distribuídas. Aplicando o mesmo teste no cenário 2-T, obteve-se um *p-value* menor que 5%, indicando que as amostras não têm uma distribuição normal.

Como o cenário 2-T falha no teste de Shapiro-Wilk, aplicou-se o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* menor que 5%, o que indica que a mediana entre as amostras não é igual.

O cenário 2-T foi o vencedor pois possui um menor EPMA. Assim, o cenário 1-T foi descartado. A tabela 46 exibe os resultados obtidos nos testes.

**Tabela 46.** Resultados dos *p-value* para os cenários 1-T e 2-T em Tucuruí sem bits representando sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
1-T	0.02	2.35e-15
2-T	0.008	

#### 4.4.1.2 Comparação entre o cenário 2-T e o cenário 3-T sem sazonalidade

Já se sabe que o teste de Shapiro-Wilk para o cenário 2-T gerou um *p-value* menor que 5%. Aplicou-se, então, o teste não-paramétrico de Wilcoxon entre os dois cenários e obteve-se um *p-value* maior que 5%, o que indica que a mediana entre as amostras é igual. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O **cenário 2-T** foi o vencedor com um EPMA de 10.73%. A tabela 47 exibe os resultados obtidos nos testes.

**Tabela 47.** Resultados dos *p-value* para os cenários 2-T e 3-T em Tucuruí sem bits para representar a sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>2-T</b>	0.008	0.51
<b>3-T</b>	0.66	

#### 4.4.2 Hidroelétrica Tucuruí: sazonalidade de 12 bits

Para a hidrelétrica de Tucuruí utilizando 12 bits de sazonalidade, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 48.

**Tabela 48.** Média de EPMA para cada cenário da hidrelétrica de Tucuruí (T) com 12 bits de sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA (%)	Desvio Padrão
<b>1-T</b>	MLP	6	-	11.25	2.43
<b>2-T</b>	RC	-	25	10.24	0.72
<b>3-T</b>	RC	-	35	9.90	0.52

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

#### 4.4.2.1 Comparação entre o cenário 1-T e o cenário 2-T com sazonalidade de 12 bits

Aplicando o teste de Shapiro-Wilk no cenário 1-T, obteve-se um *p-value* menor que 5%, o que indica que as amostras não são normalmente distribuídas. Aplicando o mesmo teste no cenário 2-T, obteve-se um *p-value* maior que 5%, provando que as amostras têm uma distribuição normal.

Como o cenário 1-T falha no teste de Shapiro-Wilk, aplicou-se o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* maior que 5%, o que indica que a mediana entre as amostras é estatisticamente igual. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O cenário 2-T foi o vencedor com um EPMA de 10.24%. A tabela 49 exibe os resultados obtidos nos testes.

**Tabela 49.** Resultados dos *p-value* para os cenários 1-T e 2-T em Tucuruí com 12 bits de sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>1-T</b>	0.0009	0.17
<b>2-T</b>	0.91	

#### 4.4.2.2 Comparação entre o cenário 2-T e o cenário 3-T com sazonalidade de 12 bits

Já se sabe que o teste de Shapiro-Wilk para o cenário 2-T obteve um *p-value* maior que 5%. Ao repetir o processo para o cenário 3-T, obtém-se também um valor maior que 5%, indicando que as amostras têm uma distribuição normal. Aplica-se, então, o teste F que resultou em um *p-value* maior que 5%, provando que as amostras possuem variâncias estatisticamente iguais.

Com as duas hipóteses anteriores atendidas, executa-se então o teste t de Student, que por ter um *p-value* maior que 5%, indicou que as amostras possuem médias estatisticamente iguais. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois. O **cenário 3-T** foi o vencedor com um EPMA de 9.90%. A tabela 50 exibe os resultados dos testes.

**Tabela 50.** Resultados dos *p-value* para os cenários 2-T e 3-T em Tucuruí com 12 bits de sazonalidade

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
<b>2-T</b>	0.91	0.08	0.04
<b>3-T</b>	0.71		

#### 4.4.3 Hidroelétrica Tucuruí: sazonalidade de 4 bits

Para a hidrelétrica de Tucuruí utilizando 4 bits de sazonalidade, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 51.

**Tabela 51.** Média de EPMA's para cada cenário da hidrelétrica de Tucuruí (T) com 4 bits de sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA (%)	Desvio Padrão
<b>1-T</b>	MLP	6	-	12.08	2.27
<b>2-T</b>	RC	-	25	10.11	0.73
<b>3-T</b>	RC	-	35	9.78	0.48

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA's para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

##### 4.4.3.1 Comparação entre o cenário 1-T e o cenário 2-T com sazonalidade de 4 bits

Aplicando o teste de Shapiro-Wilk no cenário 1-T, obteve-se um *p-value* menor que 5%, o que indica que as amostras não são normalmente distribuídas. Aplicando o mesmo teste no cenário 2-T, obteve-se um *p-value* maior que 5%, indicando que as amostras têm uma distribuição normal.

Como o cenário 2-T falha no teste de Shapiro-Wilk, aplicou-se o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* menor que 5%, o que indica que a mediana entre as amostras não é igual.

O cenário 2-T foi o vencedor pois possui um menor EPMA. Assim, o cenário 1-T foi descartado. A tabela 52 exibe os resultados obtidos nos testes.

**Tabela 52.** Resultados dos *p-value* para os cenários 1-T e 2-T em Tucuruí com 4 bits de sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>1-T</b>	5.34e-05	3.46e-07
<b>2-T</b>	0.24	

#### 4.4.3.2 Comparação entre o cenário 2-T e o cenário 3-T com sazonalidade de 4 bits

Já se sabe que o teste de Shapiro-Wilk para o cenário 2-T obteve um *p-value* maior que 5%. Ao repetir o processo para o cenário 3-T, obtém-se também um valor maior que 5%, indicando que as amostras têm uma distribuição normal. Aplica-se, então, o teste F que resultou em um *p-value* maior que 5%, provando que as amostras possuem variâncias estatisticamente iguais.

Com as duas hipóteses anteriores atendidas, executa-se então o teste t de Student, que por ter um *p-value* maior que 5%, indicou que as amostras possuem médias estatisticamente iguais. Como são estatisticamente equivalentes, o **cenário 2** foi escolhido por seu algoritmo obter uma melhor performance. O **cenário 3-T** foi o vencedor com um EPMA de 9.78%. A tabela 53 exibe os resultados dos testes.

**Tabela 53.** Resultados dos *p-value* para os cenários 2-T e 3-T em Tucuruí com 4 bits de sazonalidade

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
<b>2-T</b>	0.24	0.02	0.05
<b>3-T</b>	0.98		

#### 4.4.4 Hidroelétrica Tucuruí: sazonalidade de senos e cossenos

Para a hidrelétrica de Tucuruí utilizando representação de senos e cossenos para sazonalidade, executou-se tanto o MLP quanto o RC para gerar os cenários descritos na tabela 54.

**Tabela 54.** Média de EPMA's para cada cenário da hidrelétrica de Tucuruí (T) utilizando senos e cossenos para representar sazonalidade

Cenário	Rede Neural	Neurônios na camada escondida	Neurônios no reservatório	EPMA (%)	Desvio Padrão
<b>1-T</b>	MLP	6	-	11.68	1.80
<b>2-T</b>	RC	-	25	10.19	0.61
<b>3-T</b>	RC	-	35	10.38	0.76

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA's para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

##### 4.4.4.1 Comparação entre o cenário 1-T e o cenário 2-T com sazonalidade de senos e cossenos

Aplicando o teste de Shapiro-Wilk no cenário 1-T, obteve-se um *p-value* maior que 5%, o que indica que as amostras são normalmente distribuídas. Aplicando o mesmo teste no cenário 2-T, obteve-se um *p-value* menor que 5%, indicando que as amostras não têm uma distribuição normal.

Como o cenário 2-T falhas no teste de Shapiro-Wilk, aplicou-se o teste não-paramétrico de Wilcoxon entre os dois Cenários e obteve-se um *p-value* menor que 5%, o que indica que a mediana entre as amostras não é igual.

O cenário 2-T foi o vencedor um menor EPMA. Assim, o cenário 1-T foi descartado. A tabela 55 exhibe os resultados obtidos nos testes.

**Tabela 55.** Resultados dos  $p$ -value para os cenários 1-T e 2-T em Tucuruí usando senos e cossenos para representar sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>1-T</b>	0.78	0.0002
<b>2-T</b>	0.04	

#### 4.4.4.2 Comparação entre o cenário 2-T e o cenário 3-T com sazonalidade de senos e cossenos

Já se sabe que o teste de Shapiro-Wilk para o cenário 1-T gerou um  $p$ -value menor que 5%. Como o Cenário 2-T falha no teste de Shapiro-Wilk, aplicou-se o teste não-paramétrico de Wilcoxon entre os dois cenários e obteve-se um  $p$ -value maior que 5%, o que indica que a mediana entre as amostras é estatisticamente igual. Então, o cenário com melhor desempenho é o que possui o menor EPMA dentre os dois, ou seja, o **Cenário 2-T** foi o vencedor com um EPMA de 10.19%. A tabela 56 exhibe os resultados obtidos nos testes.

**Tabela 56.** Resultados dos  $p$ -value para os cenários 2-T e 3-T em Tucuruí usando senos e cossenos para representar a sazonalidade

Cenário	Shapiro-Wilk	Wilcoxon
<b>2-T</b>	0.04	0.3898
<b>3-T</b>	0.98	

## 4.5 Comparativo dos resultados

Após testar qual a melhor rede neural e configuração para cada representação de sazonalidade em todas as bases utilizadas nesta monografia, deseja-se verificar qual a melhor representação de sazonalidade dependendo da base de dados adotada.

### 4.5.1 Hidroelétrica Bento Munhoz

O vencedor para cada representação de sazonalidade foi o cenário 2-BM, com exceção para ausência de representação de sazonalidade que teve o cenário 1-BM como vencedor. Esses dados e a média de EPMA se encontram na tabela 57.

**Tabela 57.** Média de EPMA para cada cenário da hidrelétrica de Bento Munhoz, onde RC-25 significa rede RC com 25 neurônios no reservatório e MLP-6 significa MLP com 6 neurônios na camada escondida.

Cenário	Configuração	Sazonalidade	EPMA (%)	Shapiro-Wilk
1	MLP-6	Sem sazonalidade	21.88	0.0004
2	RC-25	12 bits	21.16	0.74
3	RC-25	4 bits	21.46	0.90
4	RC-25	Senos e cossenos	21.63	0.14

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

#### 4.5.1.1 Comparação entre o cenário 1 e o cenário 2

Para esta comparação, o cenário 1 apresenta um *p-value* menor que 5% no teste de Shapiro-Wilk. Aplica-se o teste de não-paramétrico de Wilcoxon que obteve um *p-value* maior que 5%. Então, o melhor desempenho é o cenário com o menor EPMA, ou seja, o cenário 2 com 21.16%. Então, o melhor desempenho é o cenário com o menor EPMA, ou seja, o cenário 2 com 21.16%. Os resultados dos testes podem ser vistos na tabela 58.

**Tabela 58.** Resultados dos *p-value* para os C cenários 1 e 2 em Bento Munhoz

Cenário	Shapiro-Wilk	Wilcoxon
1	0.0004	0.86
2	0.74	

#### 4.5.1.2 Comparação entre o cenário 2 e o cenário 3

Tanto o cenário 2 quanto o cenário 3 apresentaram valores maiores que 5% no teste Shapiro-Wilk. O teste F também gerou um valor maior que 5%. Com as duas hipóteses atendidas, o teste paramétrico t-Student foi executado e gerou um resultado maior que 5%. Como os cenários são estatisticamente equivalentes, o vencedor é o

que usa menos recursos ao executar o algoritmo, neste caso o cenário 3. Os resultados dos testes podem ser vistos na tabela 59.

**Tabela 59.** Resultados dos *p-value* para os cenários 2 e 3 em Bento Munhoz

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
2	0.74	0.30	0.22
3	0.90		

#### 4.5.1.3 Comparação entre o cenário 3 e o cenário 4

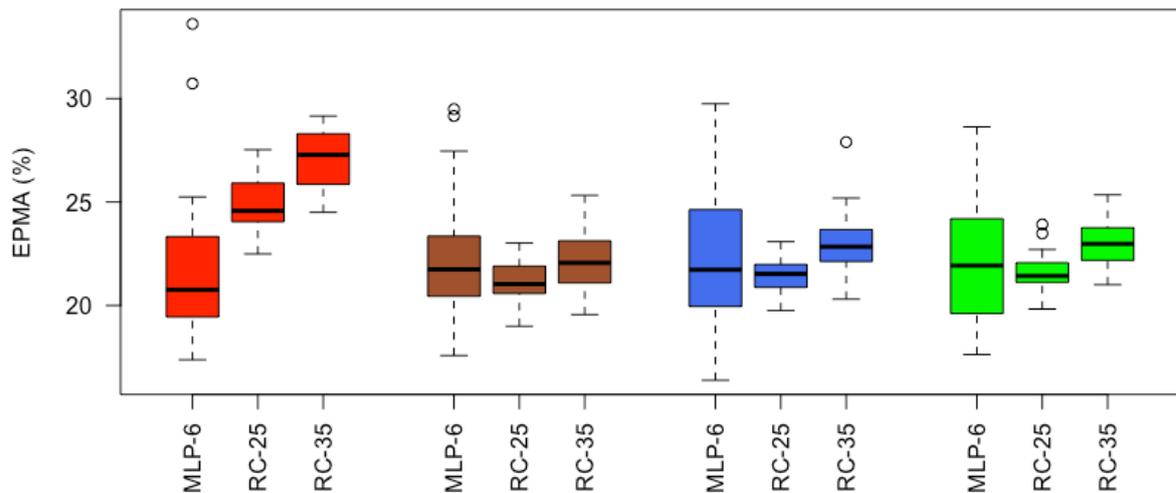
Tanto o cenário 3 quanto o cenário 4 apresentaram valores maiores que 5% no teste Shapiro-Wilk. O teste F também gerou um valor maior que 5%. Com as duas hipóteses atendidas, o teste paramétrico t-Student foi executado e gerou um resultado maior que 5%. Como os cenários são estatisticamente equivalentes, o vencedor é o que usa menos recursos ao executar o algoritmo, neste caso o **cenário 4**. Os resultados dos testes podem ser vistos na tabela 60.

**Tabela 60.** Resultados dos *p-value* para os cenários 3 e 4 em Bento Munhoz

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
3	0.90	0.90	0.43
4	0.14		

#### 4.5.1.4 Resultado

Após executados todos os testes estatísticos, os resultados podem ser visualizados no gráfico boxplot da figura 9.



**Figura 9.** Boxplot dos cenários da base de dados de Bento Munhoz. A cor vermelha representa o cenário sem sazonalidade, marrom com sazonalidade de 12 bits, azul com sazonalidade de 4 bits e verde com sazonalidade de senos e cossenos. MLP-6 significa a rede MLP com 6 neurônios na camada escondida, RC-25 e RC-35 significam a rede RC com 25 e 35 neurônios no reservatório, respectivamente.

Fonte: Elaboração própria

Com os testes executados e o boxplot, prova-se que a melhor representação de sazonalidade para a base de Bento Munhoz é a de senos e cossenos, quando usada a RC com 25 neurônios no reservatório.

#### 4.5.2 Hidroelétrica de Furnas

O cenário 2-F foi o vencedor para todos as representações de sazonalidade. A média de EPMA para cada tipo de sazonalidade se encontra na tabela 61.

**Tabela 61.** Média de EPMA para cada cenário da hidrelétrica de Furnas, onde RC-25 significa rede RC com 25 neurônios no reservatório.

Cenário	Configuração	Sazonalidade	EPMA (%)	Shapiro-Wilk
1	RC-25	Sem sazonalidade	14.83	0.48
2	RC-25	12 bits	12.36	0.0004
3	RC-25	4 bits	13.22	0.07
4	RC-25	Senos e cossenos	12.70	0.67

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

#### 4.5.2.1 Comparação entre o cenário 1 e o cenário 2

Para esta comparação, o cenário 2 apresentou um valor menor que 5% no teste Shapiro-Wilk. Com isso, o teste não-paramétrico de Wilcoxon foi executado e gerou um resultado menor que 5%. Então, o melhor desempenho é o cenário com o menor EPMA, ou seja, o cenário 2 com 12.36%. Os resultados dos testes podem ser vistos na tabela 58.

**Tabela 62.** Resultados dos *p-value* para os cenários 1 e 2 em Furnas

Cenário	Shapiro-Wilk	Wilcoxon
1	0.48	2.2e-16
2	0.11	

#### 4.5.2.2 Comparação entre o cenário 2 e o cenário 3

O cenário 2 apresentou um *p-value* menor que 5% no teste Shapiro-Wilk. Por este motivo, utilizou-se o teste não-paramétrico de Wilcoxon que gerou um resultado menor que 5%. Então, o melhor desempenho é o cenário com o menor EPMA, ou seja, o cenário 2 com 12.36%. Os resultados dos testes podem ser vistos na tabela 63.

**Tabela 63.** Resultados dos *p-value* para os cenários 2 e 3 em Furnas

Cenário	Shapiro-Wilk	Wilcoxon
2	0.0004	7.957e-11
3	0.07	

#### 4.5.2.3 Comparação entre o cenário 2 e o cenário 4

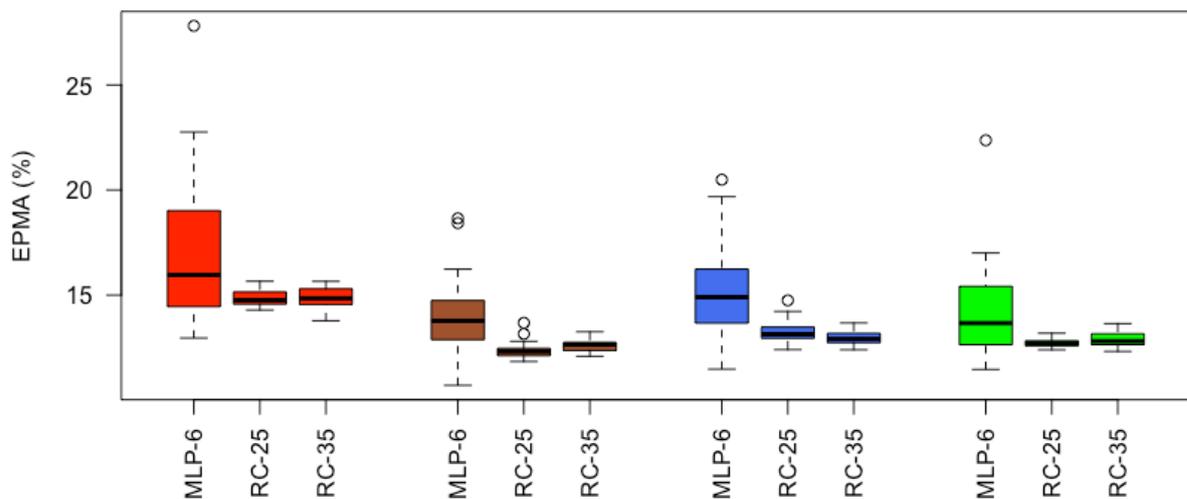
O cenário 2 apresentou um *p-value* menor que 5% no teste Shapiro-Wilk. Por este motivo, utilizou-se o teste não-paramétrico de Wilcoxon que gerou um resultado menor que 5%. O vencedor é, então, o **cenário 2** que possui o menor EPMA de 12.36%. Os resultados dos testes podem ser vistos na tabela 64.

**Tabela 64.** Resultados dos  $p$ -value para os cenários 2 e 4 em Furnas

Cenário	Shapiro-Wilk	Wilcoxon
2	0.0004	2.325e-07
4	0.67	

#### 4.5.2.4 Resultado

Após executados todos os testes estatísticos, os resultados podem ser visualizados no gráfico boxplot da figura 10.



**Figura 10.** Boxplot dos cenários da base de dados de Furnas. A cor vermelha representa o cenário sem sazonalidade, marrom com sazonalidade de 12 bits, azul com sazonalidade de 4 bits e verde com sazonalidade de senos e cossenos. MLP-6 significa a rede MLP com 6 neurônios na camada escondida, RC-25 e RC-35 significam a rede RC com 25 e 35 neurônios no reservatório, respectivamente.

Fonte: Elaboração própria

Com os testes executados e o boxplot, prova-se que a melhor representação de sazonalidade para a base de Furnas é a de 12 bits, quando usada a RC com 25 neurônios no reservatório.

### 4.5.3 Hidroelétrica Três Marias

Para todos os cenários da hidroelétrica de Três Marias, a rede RC com 25 neurônios no reservatório se mostrou a configuração mais eficiente. A média de EPMA para cada tipo de sazonalidade se encontra na tabela 65.

**Tabela 65.** Média de EPMA para cada cenário da hidrelétrica de Três Marias, onde RC-25 significa rede RC com 25 neurônios no reservatório.

Cenário	Configuração	Sazonalidade	EPMA (%)	Shapiro-Wilk
1	RC-25	Sem sazonalidade	14.45	0.62
2	RC-25	12 bits	13.48	0.77
3	RC-25	4 bits	13.88	0.11
4	RC-25	Senos e cossenos	13.77	0.62

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

#### 4.5.3.1 Comparação entre o cenário 1 e o cenário 2

Para esta comparação, tanto o cenário 1 quanto o cenário 2 apresentaram valores maiores que 5% no teste Shapiro-Wilk. O teste F também gerou um valor maior que 5%. Com as duas hipóteses atendidas, o teste paramétrico t-Student foi executado e gerou um resultado menor que 5%. Então, o melhor desempenho é o cenário com o menor EPMA, ou seja, o cenário 2 com 13.48%. Os resultados dos testes podem ser vistos na tabela 66.

**Tabela 66.** Resultados dos *p-value* para os cenários 1 e 2 em Três Marias

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
1	0.62	0.81	5.58e-06
2	0.77		

#### 4.5.3.2 Comparação entre o cenário 2 e o cenário 3

Tanto o cenário 2 quanto o cenário 3 apresentaram valores maiores que 5% no teste Shapiro-Wilk. O teste F também gerou um valor maior que 5%. Com as duas hipóteses atendidas, o teste paramétrico t-Student foi executado e gerou um resultado maior que 5%. Como os cenários são estatisticamente equivalentes, o vencedor é o que usa menos recursos ao executar o algoritmo, neste caso o cenário 3. Os resultados dos testes podem ser vistos na tabela 67.

**Tabela 67.** Resultados dos *p-value* para os cenários 2 e 3 em Três Marias

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
2	0.77	0.61	0.05
3	0.11		

#### 4.5.3.3 Comparação entre o cenário 3 e o cenário 4

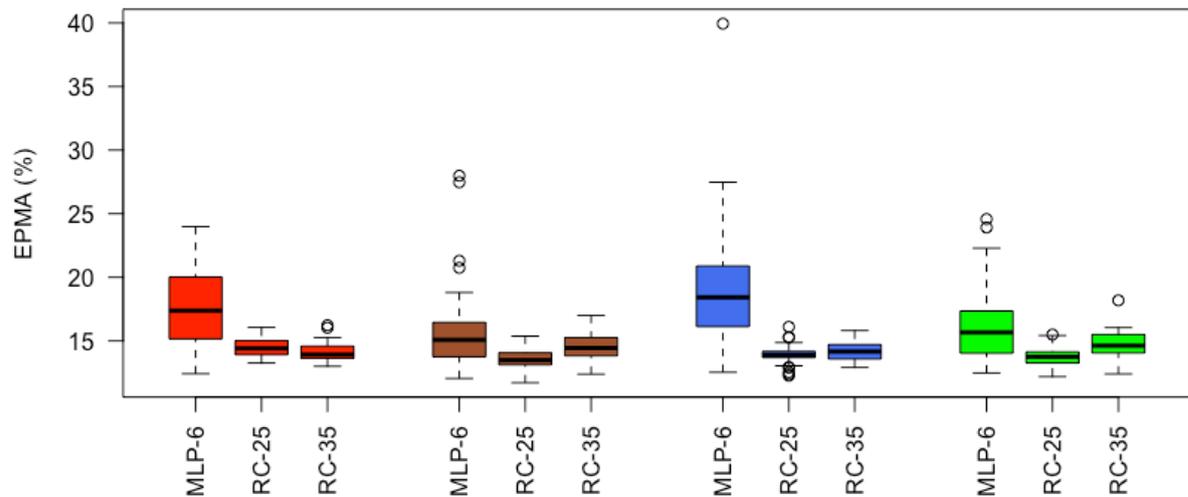
Tanto o cenário 3 quanto o cenário 4 apresentaram valores maiores que 5% no teste Shapiro-Wilk. O teste F, porém, gerou um resultado menor que 5%. Por este motivo, utilizou-se o teste não-paramétrico de Wilcoxon que gerou um resultado maior que 5%. Como os cenários são estatisticamente equivalentes, o vencedor é o que usa menos recursos ao executar o algoritmo, neste caso o **cenário 4**. Os resultados dos testes podem ser vistos na tabela 68.

**Tabela 68.** Resultados dos *p-value* para os cenários 3 e 4 em Três Marias

Cenário	Shapiro-Wilk	Teste F	Wilcoxon
3	0.11	0.52	0.60
4	0.62		

#### 4.5.3.4 Resultado

Após executados todos os testes estatísticos, os resultados podem ser visualizados no gráfico boxplot da figura 11.



**Figura 11.** Boxplot dos cenários da base de dados de Três Marias. A cor vermelha representa o cenário sem sazonalidade, marrom com sazonalidade de 12 bits, azul com sazonalidade de 4 bits e verde com sazonalidade de senos e cossenos. MLP-6 significa a rede MLP com 6 neurônios na camada escondida, RC-25 e RC-35 significam a rede RC com 25 e 35 neurônios no reservatório, respectivamente.

Fonte: Elaboração própria

Com os testes executados e o boxplot, prova-se que a melhor representação de sazonalidade para a base de Três Marias é a de senos e cossenos, quando usada a RC com 25 neurônios no reservatório.

#### 4.5.4 Hidroelétrica Tucuruí

O cenário 2-T obteve melhores resultados para todos os tipos de sazonalidade da base de Tucuruí, exceto na representação de sazonalidade de 12 bits que teve o cenário 3-T como vencedor. Estes dados juntamente com a média de EPMA se encontram na tabela 69.

**Tabela 69.** Média de EPMA para cada cenário da hidrelétrica de Tucuruí, onde RC-25 e RC-35 significam a rede RC com 25 e 35 neurônios no reservatório, respectivamente.

Cenário	Configuração	Sazonalidade	EPMA (%)	Shapiro-Wilk
1	RC-25	Sem sazonalidade	10.73	0.008
2	RC-35	12 bits	9.90	0.71
3	RC-25	4 bits	10.11	0.24
4	RC-25	Senos e cossenos	10.19	0.04

Para se analisar o cenário mais eficiente, os testes estatísticos foram executados utilizando os EPMA para realizar as comparações necessárias. As seções seguintes apresentam tais aplicações e resultados.

#### 4.5.4.1 Comparação entre o cenário 1 e o cenário 2

O cenário 1 apresenta um *p-value* menor que 5% no teste Shapiro-Wilk. O teste F também gerou um valor maior que 5%. O teste Wilcoxon foi executado e gerou um resultado menor que 5%. Então, o melhor desempenho é o cenário com o menor EPMA, ou seja, o cenário 2 com 9.90%. Os resultados dos testes podem ser vistos na tabela 69.

**Tabela 69.** Resultados dos *p-value* para os cenários 1 e 2 em Tucuruí

Cenário	Shapiro-Wilk	Wilcoxon
1	0.008	6.64e-08
2	0.71	

#### 4.5.4.2 Comparação entre o cenário 2 e o cenário 3

Tanto o cenário 2 quanto o cenário 3 apresentaram valores maiores que 5% no teste Shapiro-Wilk. O teste F também gerou um valor maior que 5%. Com as duas hipóteses atendidas, o teste paramétrico t-Student foi executado e gerou um resultado maior que 5%. Como os cenários são estatisticamente equivalentes, o vencedor é o que usa menos recursos ao executar o algoritmo, neste caso o cenário 3. Os resultados dos testes podem ser vistos na tabela 70.

**Tabela 70.** Resultados dos *p-value* para os cenários 2 e 3 em Tucuruí

Cenário	Shapiro-Wilk	Teste F	Teste t de Student
2	0.71	0.07	0.22
3	0.24		

#### 4.5.4.3 Comparação entre o cenário 3 e o cenário 4

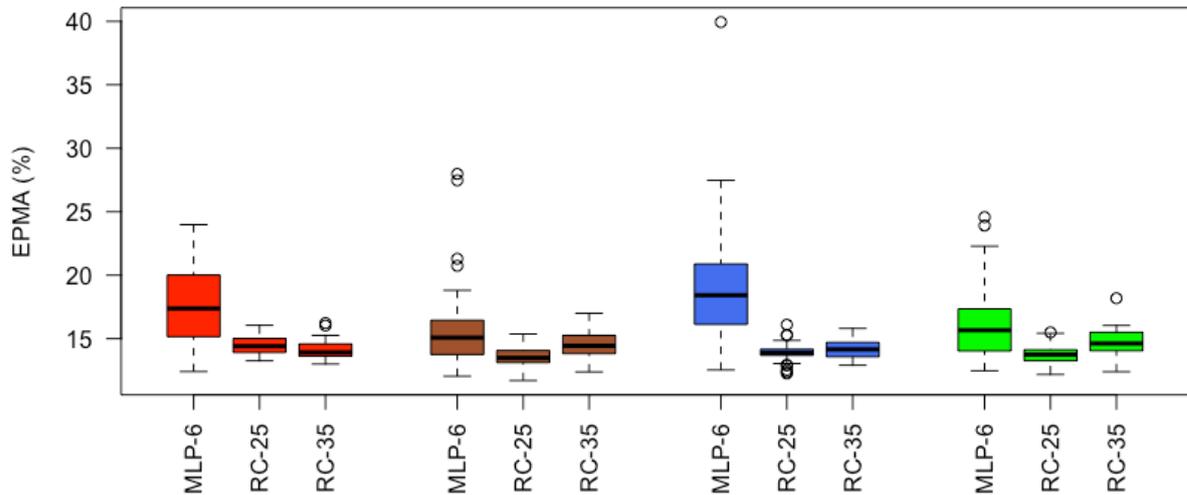
O cenário 4 apresentou um *p-value* menor que 5% no teste Shapiro-Wilk. Por este motivo, utilizou-se o teste não-paramétrico de Wilcoxon que gerou um resultado maior que 5%. Então, como as medianas são estatisticamente equivalentes, escolheu-se o **cenário 4** como vencedor. Os resultados dos testes podem ser vistos na tabela 71.

**Tabela 71.** Resultados dos *p-value* para os cenários 3 e 4 em Tucuruí

Cenário	Shapiro-Wilk	Wilcoxon
3	0.24	0.54
4	0.04	

#### 4.5.4.4 Resultado

Após executados todos os testes estatísticos, os resultados podem ser visualizados no gráfico boxplot da figura 12.



**Figura 12.** Boxplot dos cenários da base de dados de Tucuruí. A cor vermelha representa o cenário sem sazonalidade, marrom com sazonalidade de 12 bits, azul com sazonalidade de 4 bits e verde com sazonalidade de senos e cossenos. MLP-6 significa a rede MLP com 6 neurônios na camada escondida, RC-25 e RC-35 significam a rede RC com 25 e 35 neurônios no reservatório, respectivamente.

Fonte: Elaboração própria

Com os testes executados e o boxplot, prova-se que a melhor representação de sazonalidade para a base de Tucuruí é a de senos e cossenos, quando usada a RC com 25 neurônios no reservatório.

# Capítulo 5

## Considerações Finais

Este capítulo visa descrever as conclusões do autor na seção 5.1 e trabalhos futuros na seção 5.2.

### 5.1 Conclusão

A gestão de energia é uma preocupação significativa para líderes de governo e indústrias. Por causa disto, dados mais assertivos de como atuar no futuro dependendo da previsão de vazão são cada vez mais demandados. Esta monografia teve como principal objetivo comparar representações de sazonalidade com redes MLP e RC para verificar quais obtinham menores erros, quando se previa um mês dados os doze meses anteriores.

A rede RC com 25 neurônios no reservatório gerou melhores resultados para todas as bases, em detrimento das redes RC com 35 neurônios no reservatório e MLP com 6 neurônios camada escondida. Para as bases de Bento Munhoz, Três Marias e Tucuruí, a configuração utilizando senos e cossenos para representar sazonalidade se sobressaiu porque conseguiu efetuar previsões com menores EPMA's e utilizou menos neurônios de entrada, fazendo com que o algoritmo tenha uma melhor performance. Já a base de Furnas teve um melhor desempenho ao utilizar a representação de 12 bits.

### 5.2 Trabalhos futuros

Para possíveis trabalhos futuros, tem-se:

- Inferir a correlação da sazonalidade real das bacias hidrográficas com a representação de sazonalidade;
- Experimentar demais configurações da RC, como variar números de neurônios no reservatório, por exemplo;

- Realizar mais experimentos com dados de vazão de outras usinas hidroelétricas para verificar qual representação de sazonalidade apresenta melhor desempenho;
- Prever mais meses no futuro para verificar se os erros iriam se alterar para cada representação de sazonalidade.

# Bibliografia

- [1] ANEEL. **Atlas da energia elétrica no Brasil**, 3rd ed. Brasília
- [2] ANEEL. **Matriz de Energia Elétrica**. Disponível em: <<http://www2.aneel.gov.br/aplicacoes/capacidadebrasil/OperacaoCapacidadeBrasil.cfm>>. Acesso em 29 de junho de 2017.
- [3] BEZERRA, Sabrina Guimarães Tavares de Andrade. **Reservoir computinha com Hierarquia para Previsão de Vazão médias Diárias**. 2016. 73 f. Dissertação (Mestrado) - Curso de Pós-graduação em Engenharia da Computação, Universidade de Pernambuco, Recife.
- [4] CEMIG. **Usinas Hidrelétricas da Cemig**. Disponível em: <[http://www.cemig.com.br/pt-br/a\\_cemig/Nossa\\_Historia/Paginas/Usinas\\_Hidreletricas.aspx](http://www.cemig.com.br/pt-br/a_cemig/Nossa_Historia/Paginas/Usinas_Hidreletricas.aspx)>. Acesso em 29 de junho de 2017.
- [5] CIDADE DE TUCURUÍ. **Informações e Fotos da Usina Hidrelétrica Tucuruí**. Disponível em: <[http://cidadedetucurui.com/inicio/usina\\_hidreletrica\\_tucurui/USINA\\_HIDRELETRICA\\_TUCURUI.htm](http://cidadedetucurui.com/inicio/usina_hidreletrica_tucurui/USINA_HIDRELETRICA_TUCURUI.htm)>. Acesso em 29 de junho de 2017.
- [6] COPEL. **Usina Bento Munhoz da Rocha Netto**. Disponível em: <<http://www.copel.com/hpcopel/root/nivel2.jsp?endereco=%2Fhpcopel%2Froot%2Fpagcopel2.nsf%2F044b34faa7cc1143032570bd0059aa29%2Fe307f2c9b2edc56303257412004fdb91>>. Acesso em 29 de junho de 2017.
- [7] C. Tucci. **Modelos Hidrológicos**, 2, Ed. Porto Alegre: Editora da UFRGS, 2005.
- [8] DECICINO, Ronaldo, **Água potável: Apenas 3% das águas são doces**. Disponível em: <<https://educacao.uol.com.br/disciplinas/geografia/agu>>

- a-potavel- apenas-3-das-aguas-sao-doces.htm>, acesso em 26 de abril de 2017.
- [9] EASTER, K. William; DIXON, John A.; HUFSCHMIDT, Maynard M. **Watershed resources management: Studies from Asia and the Pacific**. 2. ed. Honolulu: Asian, 1991.
- [10] FIGUEIREDO, Bruno Monteiro. **Estudo Comparativo entre Técnicas de sazonalidade para previsão de vazões Utilizando Redes Neurais**. 2014. 68 f. TCC (Graduação) - Curso de Engenharia de Computação, Universidade de Pernambuco, Recife, 2014.
- [11] FURNAS. **Parque Gerador - Usina Hidrelétrica de Furnas**. Disponível em: <[http://www.furnas.com.br/hotsites/sistema-furnas/usina\\_hidr\\_furnas.asp](http://www.furnas.com.br/hotsites/sistema-furnas/usina_hidr_furnas.asp)>. Acesso em 29 de junho de 2017.
- [12] G. G. D. Oliveira, O. C. Pedrollo, e N. Maria. **O Desempenho das Redes Neurais Artificiais (RNAs) para Simulação Hidrológica Mensal**. Revista Brasileira de Recursos Hídricos, vol. 19, no. 2, pp. 251–265, 2014.
- [13] J. D. Hamilton. **Time Series Analysis**. Princeton University Press, 1994.
- [14] MAIA, James Lacerda. **Estabelecimento de Vazões de Outorga na Bacia hidrográfica do Alto Sapucaí, com a utilização de sazonalidade**. 2003. 116 f. Dissertação (Mestrado) - Curso de Pós-graduação em Engenharia da Energia, Universidade Federal de Itajubá, Itajubá.
- [15] PROTOCOLO DE KYOTO. **Aquecimento Global**. Disponível em: <<http://protocolo-de-kyoto.info/aquecimento-global.html>>. Acesso em 29 de junho de 2017.
- [16] SOARES, Everalda S. M. **Multilayer Perceptron e Reservoir Computing aplicadas em um processo Chuva x Vazão**. 2016. 56f. Monografia (Graduação) – Bacharelado em Engenharia da Computação, Universidade de Pernambuco, Pernambuco.

- [17] VALENÇA, Mêuser Jorge Silva. **Aplicando Redes Neurais**. Olinda: Livro Rápido, 2005.
- [18] VALENÇA, Mêuser Jorge Silva. **Fundamentos de Redes Neurais**. Recife: Livro Rápido, 2012.