



Utilização de redes neurais para previsão de situações críticas no Rio São Francisco

Trabalho de Conclusão de Curso

Engenharia da Computação

Fernando Baptistella de Lima
Orientador: Mêuser Jorge Silva Valença



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

Fernando Baptistella de Lima

**UTILIZAÇÃO DE REDES NEURAIS
PARA PREVISÃO DE SITUAÇÕES
CRÍTICAS NO RIO SÃO FRANCISCO**

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 2018.

MONOGRAFIA DE FINAL DE CURSO

Avaliação Final (para o presidente da banca)*

No dia 12 de julho de 2018, às 8:00 horas, reuniu-se para deliberar a defesa da monografia de conclusão de curso do discente **FERNANDO BAPTISTELLA DE LIMA**, orientado pelo professor **Mêuser Jorge Silva Valença**, sob título **Utilização de redes neurais para previsão de situações críticas no rio São Francisco**, a banca composta pelos professores:

Sérgio Mário 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: 10,0 (dez)

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

O discente terá 7 dias para entrega da versão final da monografia a contar da data deste documento.

Sérgio Galdino

SÉRGIO MÁRIO LINS GALDINO

Mêuser Jorge Silva Valença

MÊUSER JORGE SILVA VALENÇA

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

Dedico este trabalho para meus pais, Rogério Ferreira e Rose Baptistella, à minha irmã, Paula Baptistella e à minha namorada Michelle de Lima Meira por acreditarem em mim e me incentivarem a alcançar meus objetivos.

Agradecimentos

A realização deste trabalho ocorreu apenas pelo fato que durante minha jornada me deparei com pessoas excepcionais que contribuíram para o meu crescimento tanto pessoal quanto profissional. Desta forma, primeiramente eu agradeço em especial aos meus pais, Rogério e Rose, que sempre se esforçaram e mesmo enfrentando dificuldades me possibilitaram estudar e a me desenvolver como pessoa, sempre ajudando e me apoiando. Agradeço à minha irmã Paula, seu noivo Paulo e Dougie por estarem sempre ao meu lado e serem ótimos amigos.

Agradeço imensamente à minha namorada, grande amiga e companheira, Michelle, pela paciência, apoio e compreensão dos meus momentos de ausência dedicados ao estudo e por ter me ajudado em diversos momentos. Por estar ao meu lado, me abraçar quando preciso e fazer observar a vida com outra perspectiva.

Agradeço aos amigos que fiz que contribuíram direta ou indiretamente, em especial aos meus grandes amigos do grupo do Unidos do Facão: Arthur Hin, Gabriel Pontual (Pão), Giovanni Barbarini (Barba), Miguel Cruz, Rafael Cadena e Tiago Pereira, pelas risadas e momentos engraçados. Agradeço também à Elza Santos, Pedro Macedo e Pedro Pereira por serem pessoas maravilhosas comigo as quais eu as considero como segunda família.

Agradeço à Diego Pinheiro por ter sido o principal motivador durante meu ingresso na Universidade e por me despertar ao conhecimento do mundo acadêmico. Agradeço também aos amigos que fiz na Universidade, principalmente, Caio Albuquerque, Erik Vinícius, Gemerson Gerardo, Jonas Cordeiro, Matheus Henrique e Pedro Buarque. Além dos recentes amigos que fiz no C.E.S.A.R que me incentivaram durante os últimos meses, em especial Cristovão Rufino.

E por fim, ao meu orientador, professor Mêuser Valença, por ter acreditado em mim e no meu potencial com ensinamentos e paciência. E aos amigos de projeto de pesquisa Eva Solange e Breno Soares pela ajuda e contribuição.

Resumo

O gerenciamento de recursos hídricos tornou-se essencial para minimizar os impactos das situações críticas de secas que ocasionam diversos prejuízos na agricultura, pecuária, setor energético e na população. Modelos que utilizam redes neurais vem substituindo os estatísticos e hidrológicos na previsão de vazão em usinas hidroelétricas ao longo dos anos. Devido as características não-lineares serem mais precisas para a previsão de vazões futuras a utilização das redes neurais é mais recomendada. Este trabalho utilizou os modelos de redes neurais *Multilayer Perceptron* e *Reservoir Computing* na previsão da vazão diária para os volumes escoados no posto de São Francisco e dos volumes afluentes à própria barragem de Sobradinho, situada no estado da Bahia (BA). Foi realizado o teste estatístico de Wilcoxon para estudo comparativo entre os experimentos realizados onde variou o número de neurônios nas camadas intermediárias em ambas as técnicas. Os melhores resultados com os menores erros foi utilizando a rede *Reservoir Computing* utilizando 10 neurónios na camada do reservatório.

Palavras-chave: Previsão de vazão, Redes Neurais Artificiais, Multi-Layer Perceptron, Reservoir Computing.

Abstract

The management of water resources has become essential to minimize the impacts of critical drought situations that cause diverse damages in agriculture, livestock, energy sector and population. Models that use neural networks have been replacing the statistical and hydrological ones in the prediction of flow in hydroelectric plants over the years. Because the nonlinear characteristics are more accurate for predicting future flows, the use of neural networks is more recommended. This work used the models of the neural networks Multilayer Perceptron and Reservoir Computing in the forecast of the daily flow for the volumes drained in the San Francisco post and the volumes affluent to the dam of Sobradinho, located in the state of Bahia. Performing the Wilcoxon statistical test for a comparative study between experiments performed by varying the number of neurons in the intermediate layers. The best results obtained with the smallest errors were using the Reservoir Computing technique where the best scenario was with 10 neurons in the reservoir layer.

Keywords: Forecast of flow., Artificial Neural Network, Multilayer Perceptron, Reservoir Computing.

Sumário

Capítulo 1 Introdução	1
1.1 Caracterização do Problema	1
1.2 Objetivos	4
1.2.1 Objetivos Gerais	4
1.2.2 Objetivos Específicos	4
1.3 Estrutura da Monografia	4
Capítulo 2 Fundamentação Teórica	6
2.1 Potencial Hidrelétrico Brasileiro	6
2.2 Redes Neurais	8
2.2.1 Neurônio biológico	8
2.2.2 Neurônio Artificial	9
2.3 Multi-Layer Perceptron	11
2.3.1 Algoritmo de treinamento	13
2.3.2 Critério de parada	14
2.4 Reservoir Computing	15
2.4.1 Algoritmo de treinamento	16
Capítulo 3 Metodologia	18
3.1 Base de dados	18
3.2 Análise dos dados	20
3.3 Pré-processamento	23

3.4 Arquitetura da rede MLP	24
3.5 Arquitetura da rede RC	25
3.6 Testes estatísticos	27
Capítulo 4 Resultados	29
4.1 Experimento com duração de 60 dias	30
4.1.1 Comparação entre os cenários MLP e RC-10	31
4.1.2 Comparação entre os cenários MLP e RC-20	31
4.1.3 Comparação entre os cenários MLP e RC-30	32
4.2 Experimentos com duração de 30 dias	33
4.2.1 Comparação entre os cenários MLP e RC-10	34
4.2.2 Comparação entre os cenários MLP e RC-20	34
4.2.3 Comparação entre os cenários MLP e RC-30	34
4.3 Resumo dos experimentos	35
4.3.1 Resultado para previsão com duração de 60 dias	35
4.3.2 Resultado para previsão com duração de 30 dias	36
Capítulo 5 Conclusão	38
5.1 Considerações Finais	38
5.2 Trabalhos Futuros	39
Referências	40

Índice de Figuras

Figura 1.	Volume Útil dos reservatórios do sistema hídrico de São Francisco.....	2
Figura 2.	Estrutura de Oferta Interna de energia no Brasil.....	7
Figura 3.	Potencial Hidrelétrico Brasileiro em cada estágio por bacia hidrográfica ..	8
Figura 4.	Neurônio biológico.....	9
Figura 5.	Modelo computacional do neurônio de McCulloch e Pitts.....	10
Figura 6.	Modelo de uma rede Multi-Layer Perceptron com uma camada intermediária	12
Figura 7.	Diferença entre as topologias de uma rede alimentada adiante (esquerda) e de uma rede recorrente (direita)	15
Figura 8.	Representação do sistema RC com a camada respositório que contem conexões aleatórias indicadas por linhas sólidas.	16
Figura 9.	Sistema Hidráulico do rio São Francisco	19
Figura 10.	Usina Hidroelétrica de Sobradinho	19
Figura 11.	Exemplo do cálculo da média quinzenal para a base de dados da Usina Hidrelétrica de Sobradinho.	20
Figura 12.	Exemplo do cálculo da média quinzenal para a base de dados da Usina Hidrelétrica de São Francscico	21
Figura 13.	Volume máximo do afluente a sobradinho com duração de 60 dias	22
Figura 14.	<i>Boxplot</i> dos cenários com duração 60 dias	36
Figura 15.	<i>Boxplot</i> dos cenários com duração 30 dias	37

Índice de Tabelas

Tabela 1. Exemplo de normalização no intervalo [0-1].....	24
Tabela 2. Principais parâmetros utilizados e seus valores para a Rede Neural Artificial <i>Multi-layer Perceptron</i> (MLP)	25
Tabela 3. Principais parâmetros utilizados e seus valores para a a Rede Neural Artificial <i>Reservoir Computing</i> (RC).....	26
Tabela 4. Média de EPMAs para cada cenário utilizando média móvel de 60 dias com 12 variáveis.	30
Tabela 5. Média de EPMAs para cada cenário utilizando média móvel de 60 dias com 16 variáveis.	31
Tabela 6. Resultados dos p-value para cenários com cenário MLP e RC-10 com duração de 60 dias.....	31
Tabela 7. Resultados dos p-value para cenários com cenário MLP e RC-20 com duração de 60 dias.....	32
Tabela 8. Resultados dos p-value para cenários com cenário MLP e RC-30 com duração de 60 dias.....	32
Tabela 9. Média de EPMAs para cada cenário utilizando média móvel de 30 dias com 12 variáveis	33
Tabela 10. Média de EPMAs para cada cenário utilizando média móvel de 30 dias com 16 variáveis	33
Tabela 11. Resultados dos p-value para cenários com cenário MLP e RC-10 com duração de 30 dias.....	34
Tabela 12. Resultados dos p-value para cenários com cenário MLP e RC-20 com duração de 30 dias.....	34

Tabela 13. Resultados dos p-value para cenários com cenário MLP e RC-30 com duração de 30 dias.....	35
--	----

Tabela de Símbolos e Siglas

ANA - Agência Nacional de Águas

ANEEL - Agência Nacional de Energia Elétrica

BEM - Balanço Energético Nacional

Chesf - Companhia Hidroelétrica do São Francisco

CBHSF - Comitê da Bacia Hidrográfica do Rio São Francisco

EMQ - Erro Médio Quadrático

EPMAs - Erros Percentuais Médios Absolutos

GW - Gigawatt

Inmet - Instituto Nacional de Meteorologia

MLP - Multi-Layer Perceptron

OIEE - Oferta Interna de Energia Elétrica

RC - Reservoir Computing

RNNs - Redes Neurais Recorrentes

SIN - Sistema interligado nacional

SIPOT - Sistema de Informações do Potencial Hidrelétrico Brasileiro

POLI - Escola Politécnica de Pernambuco

RNAs - Redes Neurais Artificiais

UPE – Universidade de Pernambuco

Capítulo 1

Introdução

Este capítulo encontra-se dividido em três seções, a primeira seção abordará as características e conteúdo do problema com suas respectivas motivações e soluções para a realização deste projeto. Na segunda seção serão discutidos os objetivos gerais e específicos que este trabalho possuirá. Por fim, a terceira e última parte apresentará a estrutura que esta monografia possui.

1.1 Caracterização do Problema

O Brasil possui a maior bacia hidrográfica do mundo com um conjunto de mais de 100 usinas hidroelétricas. Esta quantidade representa uma participação de 68,1% na matriz elétrica brasileira, segundo o relatório consolidado do Balanço Energético Nacional (BEN) realizado pelo ministério de Minas e Energia em 2016 [20]. Assim os gerenciamentos de bacias hidrográficas tornaram-se um dos principais focos para controle do uso de recursos hídricos e energéticos em todo o país.

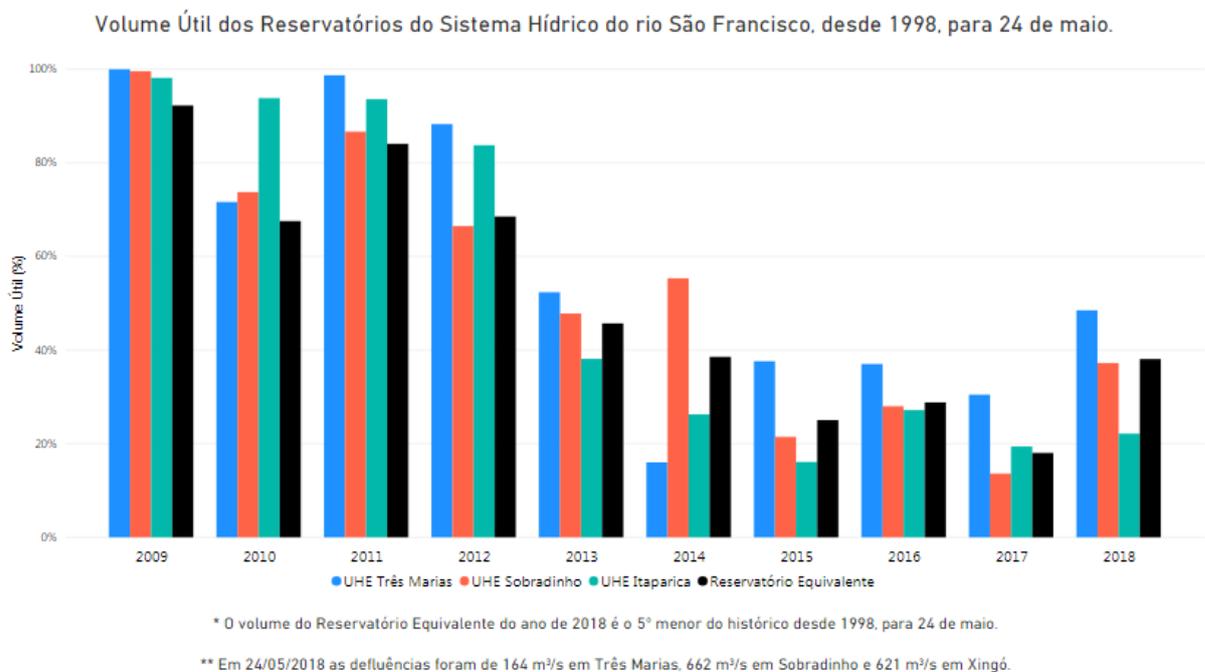
Uns dos principais e mais produtivos rios e usinas instaladas é bacia do São Francisco, totalizando seis usinas pertencente ao parque gerador da Companhia Hidroelétrica do São Francisco (Chesf). Sendo responsável por 70% da disponibilidade hídrica da região nordeste e abrangendo os seguintes estados: Bahia, Minas Gerais, Pernambuco, Alagoas, Sergipe, Goiás e o Distrito Federal e do norte de Minas Gerais, cobrindo 7,5% de todo o território nacional, segundo o Comitê da Bacia Hidrográfica do Rio São Francisco (CBHSF) [1].

Esta região apresenta prolongadas estiagens com baixa pluviosidade e alta evapotranspiração. A precipitação média anual na região é de 1.003 mm, muito abaixo da média nacional, de 1.761 mm, segundo dados do Instituto Nacional de Meteorologia (Inmet) [21].

A região hidrográfica do São Francisco possui uma vazão média é de 2.846 m³/s (1,58% da vazão média nacional, de 179.516 m³/s) [19]. Além disso, vale ressaltar um importante papel que o São Francisco realiza na geração de energia elétrica, representando à base de suprimento de energia do Nordeste. Em 2013 possuía um potencial instalado de 10.708 MW, responsável por 12% do total do país. Entre as usinas pertencentes à bacia destacam-se as usinas de Xingó (3.162 MW), Paulo Afonso IV (2.462 MW), Luiz Gonzaga (1.479 MW) e Sobradinho (1.050 MW) [21], localizadas nos estados de Sergipe, Bahia, Pernambuco e Bahia, respectivamente.

Desde 2012 vem ocorrendo uma das maiores secas nesta bacia considerada como uma das principais do Brasil, chegando a possuir apenas 6,99% de seu volume útil segundo a ANA (Agência Nacional de Águas) do governo federal [2]. Nota-se que é importante tomar decisões validadas e confiáveis para que se possa ter um bom gerenciamento do uso da bacia para a população.

Figura 1. Volume Útil dos reservatórios do sistema hídrico de São Francisco



[Fonte: <http://www3.ana.gov.br/porta/ANA/sala-de-situacao/sao-francisco>]

No lago de Sobradinho, maior reservatório de água do Nordeste e ponto central para a geração de energia elétrica da região viveu a pior seca em 83 anos em 2015[13]. Na época sua capacidade total caiu para menos de 6%, sua vazão correspondia em 900 m³/segundo, número bem menor que o volume normal.

Em 1992 Valença, Cavalcanti e Oliveira [16], realizaram o trabalho do caminho crítico de decisão para a operação hidráulica de Sobradinho cujo objetivo era realizar a previsão de volumes afluentes à Sobradinho, em função de informações do passado.

Foi utilizado à técnica de regressão múltipla que realizava a previsão com os valores dos volumes escoados no posto de São Francisco, situado à jusante da Barragem de Três Marias e dos volumes afluentes à própria barragem de Sobradinho.

A previsão da vazão nas usinas hidroelétricas é uma das práticas mais utilizadas para minimizar os impactos no gerenciamento e conseqüentemente racionalização do uso da energia e abastecimento de água [18]. Um dos meios utilizados é a abordagem dos métodos estatísticos e probabilísticos na solução de problemas de previsão de séries temporais.

Desta forma emerge à necessidade do estudo de sistemas que possam melhorar a previsão de maneira mais efetiva. Surge assim nos últimos anos o uso das técnicas de redes neurais. Abordagem que vem sendo atrativa para uma grande variedade de problemas que envolvem previsões [26]. Uma das principais vantagens destes modelos se deve ao fato de conseguir solucionar problemas que não são linearmente separáveis.

Estudos recentes demostram que o uso das Redes Neurais Artificiais (RNAs) para a previsão de séries temporais vem aumentando, principalmente ao uso de previsão da vazão e chuva [8] [5]. Entre essas redes estão a *Multi-Layer Perceptron* (MLP) e a *Reservoir Computing* (RC).

1.2 Objetivos

1.2.1 Objetivos Gerais

O objetivo principal deste projeto é realizar a previsão das situações críticas no rio São Francisco, utilizando dados das médias das vazões diárias da bacia e dos volumes afluentes de Sobradinho. Será comparado os resultados obtidos na previsão da vazão utilizando as redes neurais *Multi-Layer Perceptron* (MLP) e a *Reservoir Computing* (RC), verificando os resultados mais eficientes, ou seja, possuir um menor erro de acerto na previsão da vazão.

1.2.2 Objetivos Específicos

1. Estudar os algoritmos das redes MLP e RC.
2. Obter os erros nas previsões das vazões tanto para a rede MLP como para a RC.
3. Analisar os dados estatisticamente para saber qual rede obtêm o melhor resultado.

1.3 Estrutura da Monografia

Esta monografia possui a seguinte estrutura:

- No Capítulo 1 – Introdução
Será apresentado neste capítulo as principais motivações por trás desta pesquisa e qual o problema envolvido e os objetivos que se deseja alcançar nesta monografia.
- Capítulo 2 – Fundamentação Teórica
Este capítulo irá descrever a fundamentação teórica, que seria o conteúdo responsável por todo o conhecimento de base para o entendimento do trabalho pelo leitor, como o panorama da disponibilidade hídrica no Brasil, conceitos de redes neurais, biológicas e artificiais e as arquiteturas utilizadas no desenvolvimento deste trabalho

- O Capítulo 3 – Metodologia

Será o desenvolvimento da 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

- O Capítulo 4 – Resultados

Esse capítulo irá conter o detalhamento dos experimentos, cenários realizados e os resultados das comparações através dos testes estatísticos.

- Capítulo 5 – Conclusão

Este capítulo, abordará as conclusões a partir dos resultados do capítulo anterior e os trabalhos futuros.

Capítulo 2

Fundamentação Teórica

Este capítulo abordará os assuntos fundamentais para a compreensão do deste trabalho que visa a resolução do problema descrito no capítulo anterior. A seção 2.1 detalha a situação hidrelétrica brasileira e o potencial das bacias do Brasil. A seção 2.2 apresenta os conceitos relacionados as Redes Neurais Artificiais (RNAs) e em seguida é detalhada as arquiteturas *Multi-Layer Perceptron* e *Reservoir Computing* nas seções 2.3 e 2.4, respectivamente.

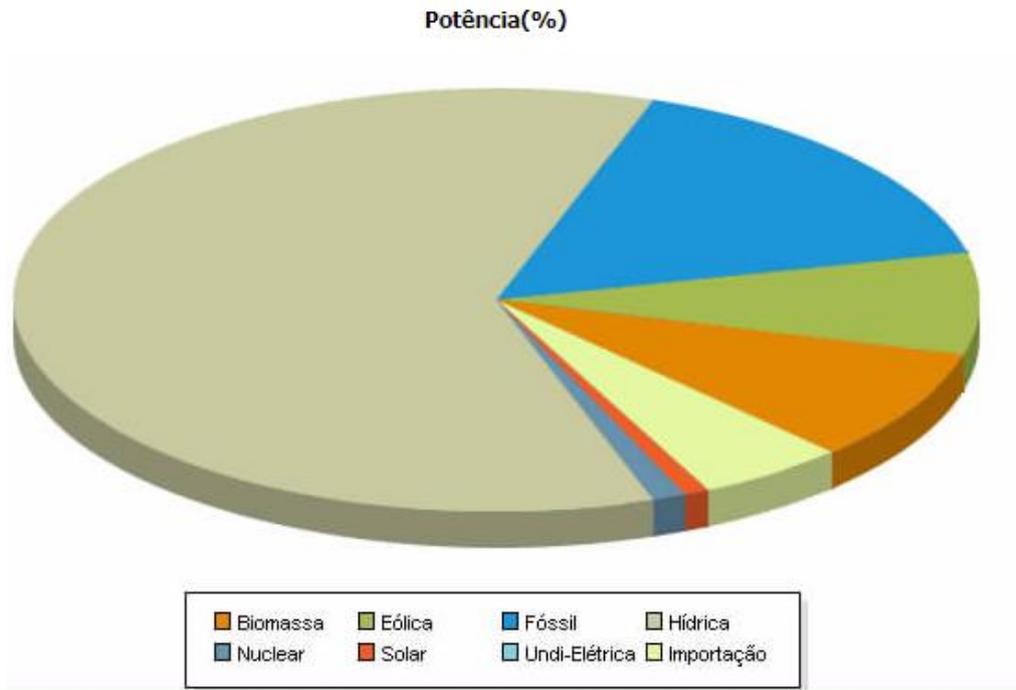
2.1 Potencial Hidrelétrico Brasileiro

A cada dia os estudos relacionados aos melhor uso dos recursos disponíveis, tais como energia solar, hidrelétrica, nuclear e eólica vem aumentando gradativamente. A energia elétrica hoje é um recurso indispensável para o desenvolvimento socioeconômico de muitos países e regiões. Ela é a principal fonte renovável de geração de eletricidade globalmente, fornecendo 71% de toda a eletricidade renovável em 2016 [27].

Nas últimas décadas vem sendo discutidos os impactos que o uso excessivo dos recursos energéticos podem trazer, essa discussão leva em conta fatores como sustentabilidade, poluição ambiental, custo social e segurança energética. Considerado uma das energias renováveis mais flexível à energia hidrelétrica vem ganhando vantagem para o seu desenvolvimento ao redor do mundo. Os principais países geradores de energia hidrelétrica em 2016 foram China, EUA, Brasil, Canadá, Índia e Rússia [10].

Segundo a Agência Nacional de Energia Elétrica (ANEEL), cerca de 60% da matriz energética brasileira é composta por hidroelétricas [6]. Sendo que a energia hidráulica ficou mais acentuada em 2016 ficando com 67,1% na estrutura da Oferta Interna de Energia Elétrica (OIEE) [14], isto é, total da energia disponibilizada para movimentar a economia de um país ou região, conforme mostra a Figura 2.

Figura 2. Estrutura de Oferta Interna de energia no Brasil



[Fonte:<http://www2.aneel.gov.br/aplicacoes/capacidadebrasil/OperacaoCapacidadeBrasil.cfm>]

Um reservatório de armazenamento, também conhecido como barragem, acumula o curso da água enquanto a afluência natural é maior ou menor em relação a demanda. A energia é produzida nas hidroelétricas através da vazão diária do rio que espera resultar em um valor ótimo de volume de água. Para seu uso modelos foram desenvolvidos para possibilitar uma boa administração desse recurso para a realização da previsão da vazão utilizando técnicas estatísticas, hidrológicas e de inteligência artificial [16].

De acordo com o SIPOT (Sistema de Informações do Potencial Hidrelétrico Brasileiro) o potencial hidrelétrico brasileiro é estimado em cerca de 260 GW, sendo que 40,5% está localizado na Bacia Hidrográfica do Amazonas [24].

Para obter o potencial hidrelétrico brasileiro é necessário realizar a soma da parcela estimada com a inventariada. A parcela estimada resulta da somatória do potencial remanescente, dados existentes sem qualquer levantamento detalhado que considera um trecho do curso d'água e do potencial individualizado.

O resultado da parcela estimada é composto com o valor da inventariada que resulta do estudo da bacia hidrográfica, viabilidade do projeto básico em construção ou em operação. O potencial hidráulico brasileiro por sub-bacia hidrográfica é apresentada na Figura 3.

Figura 3. Potencial Hidrelétrico Brasileiro em cada estágio por bacia hidrográfica

Potencial Hidrelétrico Brasileiro em cada Estágio por Bacia Hidrográfica (valores em MW) Dezembro de 2017 - Fonte: SIPOT Eletrobras							
Bacia	Estimado	Inventário	Viabilidade	Projeto Básico	Construção	Operação	Total Geral
Rio Amazonas	30.595	36.242	2.424	948	794	23.213	94.216
Rio Tocantins	1.875	8.077	3.738	120	-	13.250	27.060
Atlântico Norte e Nordeste	707	871	466	50	-	812	2.905
Rio São Francisco	1.561	3.895	6.140	234	-	10.786	22.616
Atlântico Leste	1.423	5.796	665	811	54	5.424	14.172
Rio Paraná	5.107	9.563	1.889	2.206	400	43.538	62.704
Rio Uruguai	342	4.097	292	432	152	6.406	11.720
Atlântico Sudeste	2.031	1.857	2.218	326	5	3.793	10.231
Totais	43.640	70.399	17.831	5.127	1.405	107.223	245.625

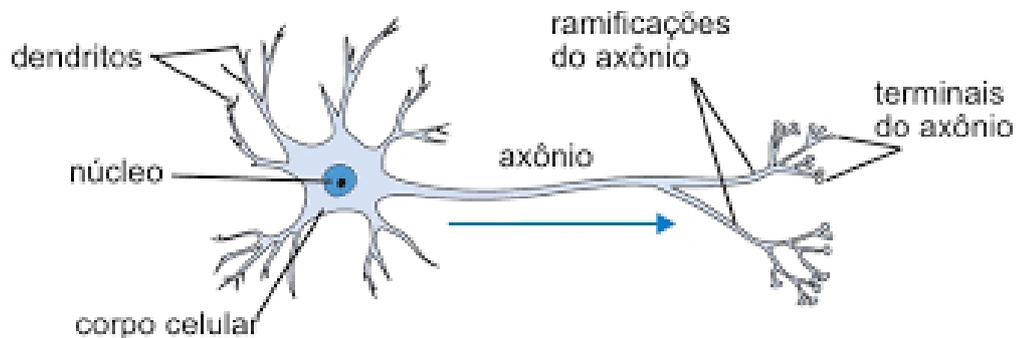
[Fonte: <http://eletrobras.com/pt/AreasdeAtuacao/geracao/sipot/PotencialHidre%C3%A9tricoBrasileiro-Bacia-Dezembro2017.pdf>]

2.2 Redes Neurais

2.2.1 Neurônio biológico

As redes neurais biológicas são compostas por aproximadamente 10 bilhões de células chamadas de neurônios, que processam e transmitem a propagação ou a inibição dos impulsos nervosos [17]. A resposta poderá ser propagada através da região do axônio, conforme ilustrada na figura 4, para outras células por meio de um estímulo nervoso. A região que ocorre a transição dos estímulos é chamada de sinapse, através da união entre um axônio e dendrito dos neurônios, conhecidos como receptores dos sinais.

Figura 4. Neurônio biológico



[Fonte: https://www.researchgate.net/figure/Figura-22-Representacao-de-um-neuronio-biologico-Fonte-extraido-e-adaptado-de-http_fig1_307578398]

O psicólogo canadense Donald Hebb estudou um mecanismo capaz de realizar as trocas de informações que pudesse armazenar e recuperá-las através dos neurônios [17]. Ele propôs que o aumento na eficiência sináptica surge através da estimulação repetida e persistente, propriedade de neurônios individuais.

Segundo a teoria hebbiana, "células que disparam juntas, permanecem conectadas", e "quando um neurônio recebe um estímulo de outro neurônio, e se ambos estão altamente ativos, o peso entre estes deve ser fortalecido, caso contrário enfraquecido" [15]. Com base no seu estudo surgiu o desenvolvimento das redes neurais artificiais por outros pesquisadores.

2.2.2 Neurônio Artificial

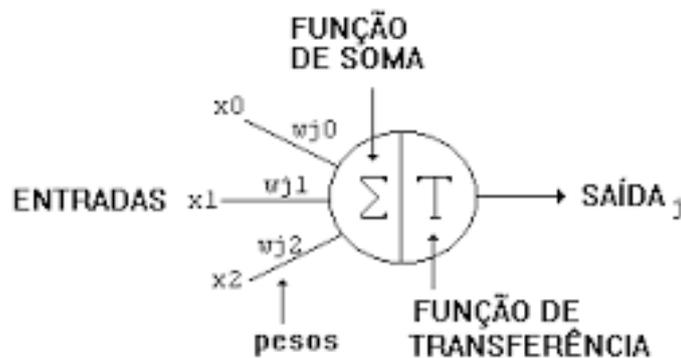
As Redes Neurais Artificiais (RNAs), subárea da Inteligência Artificial, são modelos computacionais complexos e poderosos que foram propostos por McCulloch e Pitts em 1943 [3], como uma metáfora do cérebro humano para realizar o aprendizado de máquina. De maneira geral, as redes biológicas e artificiais realizam a mesma função, ambas recebem informações de sinais, processam e emitem uma saída como resposta.

Nas redes artificiais os neurônios estão interconectados e dispostos em camadas. Segundo Valença [15], as redes artificiais são um "sistema constituído por elementos de pré-processamento interconectados chamados de neurônios, os quais

estão dispostos em camadas”. As particularidades que tornam essa técnica tão poderosa está relacionada a resolução de problemas que não são linearmente separáveis [15].

A imagem abaixo representa o modelo computacional do neurônio de Mcculloch e Pitts que constitui por um conjunto de valores de entrada x_i que irá gerar um sinal de saída j_0 .

Figura 5. Modelo computacional do neurônio de McCulloch e Pitts



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

O resultado é obtido pela multiplicação dos valores de entrada pelos seus respectivos pesos da camada de saída w_i . Caso as conexões sinápticas forem maiores que zero ela é dita como excitatória ($w_i > 0$), caso contrário é dita como inibitória. A regra de propagação é dada pela equação 2.1:

$$net_i = \sum_{j=1}^n W_{ij} X_j - \theta \quad (2.1)$$

A realização da transmissão do sinal poderá ou não ser realizada seguindo a Lei do Tudo ou Nada. Segundo Valença [15], o impulso pode ser realizada com base na intensidade do estímulo, chamado limiar excitatório. A função de ativação, também conhecida como função de transferência, mensura a influência de propagação. Portanto, representam abstração do esforço do disparo que os

neurônios possuem sobre a saída, sendo assim uma taxa de ativação do potencial de ação da célula.

Frank Rosenblatt desenvolveu a arquitetura *Perceptron* em 1958 [11], introduzindo o conceito de aprendizagem supervisionada. Nesta abordagem os pesos sinápticos da rede neural são atualizados de acordo com a diferença da saída real e a saída calculada pela rede, processo realizado por meio de aprendizado pela repetição.

Posteriormente foi apresentado a arquitetura de Adaline por Widrow e Hoff [11], a grande diferença deste modelo é o de possibilitar o uso da função contínua. Segundo Valença, as principais funções de ativação são a linear, sigmoide logística e tangente hiperbólica [15], representadas nas equações 2.2, 2.3 e 2.4, respectivamente.

$$y = f(net_i) = net_i \quad (2.2)$$

$$y = f(net_i) = \frac{1}{1 + e^{-net_i}} \quad (2.3)$$

$$y = f(net_i) = \frac{e^{net_i} - e^{-net_i}}{e^{net_i} + e^{-net_i}} \quad (2.4)$$

Estes modelos foram essenciais para o desenvolvimento de vários outros tipos de redes neurais artificiais, sendo que cada uma delas possui um paradigma e uma arquitetura de processamento específica. Neste trabalho foram utilizados dois tipos de redes neurais chamadas de *Multi-Layer Perceptron* (MLP) e a *Reservoir Computing* (RC) que serão abordadas nas próximas sessões.

2.3 Multi-Layer Perceptron

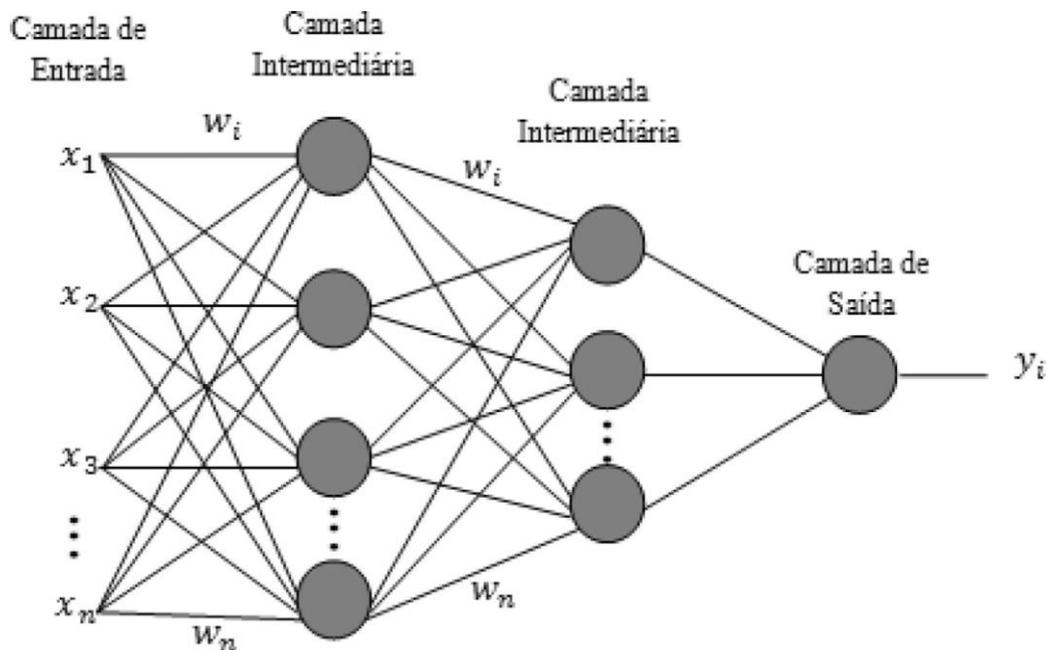
Os neurônios sozinhos não possuem grande capacidade de resolver problemas mais complexos. Por causa disso problemas do mundo real que são

considerados funções linearmente não separáveis não podem ser resolvidos através das arquiteturas de *Perceptron* e *Adaline*.

A rede *Multi-Layer Perceptron* (MLP) desenvolvido por Paul Werbos em 1974, também conhecida como *Gradient Descent Optimization* [11], é uma generalização da rede *Perceptron* [15] pelo simples fato de possuir uma ou várias camadas escondidas de neurônios que são os responsáveis pelas soluções de problemas não lineares.

A primeira camada representa a entrada (*input layer*) para o problema enquanto a última camada (*output layer*) prevê os valores para a rede. Entre essas duas camadas podem existir uma ou mais camadas intermediárias, também conhecida como camada escondida (*hidden layer*). A figura 6 mostra a estruturação de uma rede MLP com camadas escondidas.

Figura 6. Modelo de uma rede *Multi-Layer Perceptron* com uma camada intermediária



[Fonte: <http://www.scielo.br/img/revistas/rbmet/v31n1//0102-7786-rbmet-31-01-0024-gf02.jpg>]

2.3.1 Algoritmo de treinamento

A MLP realiza seu aprendizado de forma supervisionada utilizando o algoritmo *Backpropagation*, desenvolvido por Paul Werbos em 1974 [15]. A complexidade no treinamento aumentou comparada com os modelos anteriores, pelo fato do erro desta camada não ser possível de ser reconhecido.

Desta forma, o algoritmo possui duas fases onde a primeira fase *forward* representa a propagação da rede onde ocorre o treinamento dos dados no sentido progressivo até a camada de saída. Após o resultado, o erro é calculado a partir da resposta desejada de cada variável de saída com seu respectivo resultado utilizando o erro médio quadrático (EMQ), conforme a equação 2.5 a seguir:

$$e_i(n) = \frac{1}{2} (d_i - y_i)^2 \quad (2.5)$$

Onde:

- d_i é o valor desejado;
- y_i é o valor estimado pela rede neural;

Ao final desta etapa, caso o erro tenha como valor um resultado nulo ou muito próximo de zero a rede é dita como ter convergido ao valor esperado, encerrando o ciclo de treinamento. Caso não ocorra a convergência o algoritmo *backward* inicia a segunda fase, onde o sinal é retropropagado pelas camadas em direção a entrada para ajustar os erros para cada neurônio da rede.

Os pesos serão ajustados de acordo com a regra delta generalizada com o objetivo de minimizar o erro obtido na rede que essencialmente mensura o quanto divergente se encontra o valor esperado ao objetivo da rede. A equação 2.6 é responsável pelo ajuste dos pesos ou regra delta generalizada:

$$w_{ij}^m(t+1) = w_{ij}^m(t) + \alpha \beta \delta_i^m f^{m-1}(net_j^{m-1}) \quad (2.6)$$

Onde:

- $w_{ij}^m(t+1)$ é o novo valor atribuído ao i-ésimo peso do neurônio j;
- $w_{ij}^m(t)$ é o valor atual do i-ésimo peso do neurônio j no instante t;
- α é a taxa de aprendizagem;
- δ_i^m mede a sensibilidade do neurônio;
- β é a taxa de momento;
- $f^{m-1}(net_j^{m-1})$ sinais de entrada emitidos pelos neurônios da camada anterior;

O cálculo da sensibilidade para a camada de saída é obtida pela equação 2.7 a seguir:

$$\delta_i^{m-1} = (f^{m-1})'(net_j^{m-1}) \sum_{j=1}^{Neurônios} w_{ij}^m \delta_i^m \quad (2.7)$$

2.3.2 Critério de parada

A fim de minimizar os erro, os pesos da rede devem ser ajustados recursivamente até que um conjunto de pesos ótimos seja estabelecido. Durante o treinamento o momento de definir a parada é essencial pois caso seja treinada demais poderá ocorrer a generalização chamada de *overfitting*. O treinamento excessivo da rede ocasiona perda da capacidade de aprendizado, isto é, decorando padrões.

Ao mesmo tempo a rede pode não ser treinada o suficiente para gerar bons resultados, desta forma será aplicado nesta monografia à validação cruzada como critério de parada, citado por Valença [15].

Para empregar esta técnica a base de dados utilizada para o desenvolvimento da rede é dividida em três partes: treinamento, validação e teste. Os 50% dos dados contidos são direcionados para treinamento para ajustar os pesos durante o treinamento. Outros 25% da base original são utilizados para validação cruzada onde os erros são calculados.

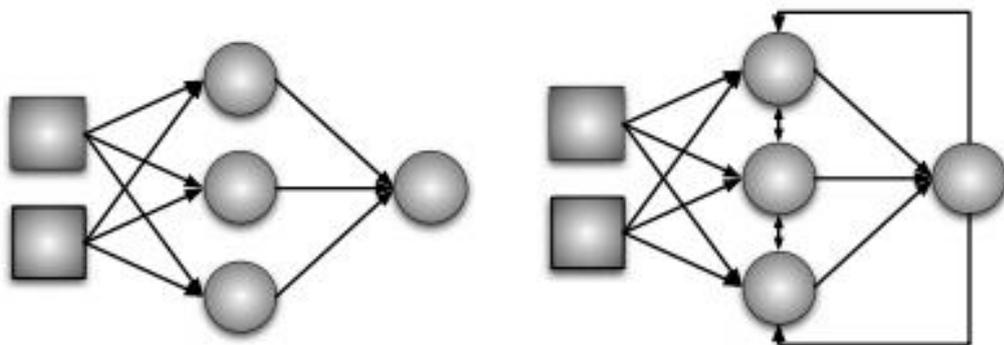
Os dados restante serão utilizados para testes. Durante o treinamento, por meio da validação cruzada é calculada a função erro que ao atingir o seu mínimo deve-se interromper o treinamento. Por ultimo o conjunto de testes que serve para verificar o desempenho do modelo da rede.

2.4 Reservoir Computing

As Redes Neurais Artificiais (RNNs) e as Redes Neurais Recorrentes (RNNs) possuem camada inicial para as variáveis de entrada e uma camada de saída para ser treinada. O diferencial entre as redes consiste que a recorrente possui na camada intermediária conexões de realimentação de informação gerando um comportamento dinâmico não-linear e capaz de armazenar o histórico dos padrões de entrada através das suas conexões recorrentes [11][12].

Desta forma esta rede vem sendo utilizada para resolução de problemas que envolvem séries temporais [18][22], devido que o uso do histórico dos valores é essencial para à relização das previsões que se deseja obter. A figura 7 diferencia a arquitetura entre as a rede MLP e RNN.

Figura 7. Diferença entre as topologias de uma rede alimentada adiante (esquerda) e de uma rede recorrente (direita)

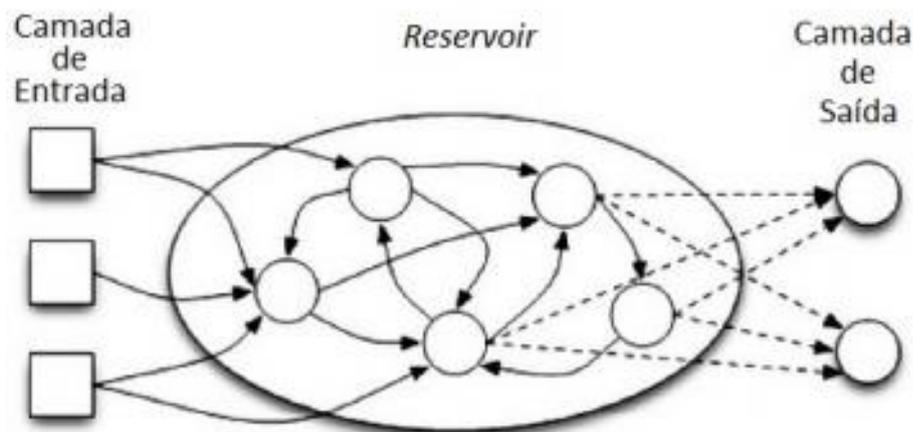


[Fonte: VERSTRAETEN, 2009, p. 13]

Um novo paradigma de treinamento de redes neurais recorrentes é conhecido como *Reservoir Computing* (RC), desenvolvida por David Verstraeten em 2009 [7]. O grande diferencial está relacionado à camada intermediária, chamada de reservatório, onde os neurônios estão interconectados de maneira aleatória.

As ligações entre os neurônios da camada do reservatório é determinada pela taxa de conectividade, quanto maior a taxa, maior a quantidade de neurônios ligados entre si. A rede neural incorpora um vasto conjunto de informações do passado e do presente, reforçando a sua aplicabilidade em sistemas dinâmicos. A Figura 8 demonstra a estrutura da rede RC.

Figura 8. Representação do sistema RC com a camada respositório que contém conexões aleatórias indicadas por linhas sólidas.



[Fonte: VERSTRAETEN, 2009, p. 28]

2.4.1 Algoritmo de treinamento

Conforme é realizada nas redes MLP, inicialmente é definida a quantidade de neurônios na camada inicial e a camada de saída com base nos dados que serão utilizados. A quantidade de neurônios no reservatório é escolhido empiricamente.

Logo em seguida, é definida as matrizes de pesos da camada de entrada para o reservatório W_{en} e para as conexões do reservatório W_{res} . Ambas foram geradas a partir de uma distribuição aleatória, entre o intervalo $[-1,1]$. Além disso, é

definida a taxa de conectividade, porcentagem de neurônios com pesos diferentes de zero sendo escolhida empiricamente, variando de acordo com a melhor performance possível para a aplicação.

A base de dados é dividida em três subconjuntos: treinamento, validação cruzada e testes, seguindo o padrão visto na rede de MLP. Antes de iniciar o treinamento os estados dos neurônios do reservatório são inicializados com valor igual a zero nos primeiros n ciclos utilizando a técnica *warm up* [5].

Esta abordagem possibilita que os estados iniciais não impactem a rede com a influência do valor zero atribuído ao estado inicial. Ao fim do *warm up*, acontece o início do treinamento com a atualização dos neurônios do reservatório. Os estados são atualizados segundo a equação 2.8 abaixo:

$$x(t + 1) = f(W_{in}u(t) + W_{res}x(t)) \quad (2.8)$$

Onde $W_{res}x(t)$ é a matriz da multiplicação dos valores vindos da camada de entrada pelos pesos que conectam esses neurônios ao *reservoir* em um tempo t e $W_{in}u(t)$ é a matriz com os estados dos neurônios do *reservoir* no mesmo tempo k . O resultado do neurônio da camada do reservatório será o cálculo da função de ativação a partir das duas parcelas descritas anteriormente.

Durante o processo do treinamento, as matrizes dos pesos de entrada e do repositório não são atualizada e possuem pesos fixos, apenas para a camada de saída W_{said} .

Nesta camada calculam-se os pesos das conexões utilizando o cálculo da matriz inversa de Moore–Penrose ou pseudoinversa, usada pelo fato do determinante da matriz poder se aproximar a zero. Ao fim deste processo, inicia-se a validação cruzada realizando a simulação com o conjunto de teste para calcular o Erro Percentual Médio Absoluto Normalizado (EPMA). Sendo repetido até que o critério de parada seja atingindo e o treinamento seja finalizado.

Capítulo 3

Metodologia

Neste capítulo será discutido a metodologia empregada com detalhes para a solução do problema. Na seção 3.1 serão descritos as bases de dados utilizadas. Na seção 3.2 e 3.3 é analisada a base e todo o pré-processamento realizado nos dados antes de serem apresentados nas redes neurais MLP e RC, abordadas nas seções 3.4 e 3.6, respectivamente. Por fim, a seção 3.6 aborda os testes estatísticos utilizados para validar os resultados.

3.1 Base de dados

Com base no trabalho de Valença, Cavalcanti e Oliveira realizado em 1992 [16], foi desenvolvido o caminho crítico da decisão para a operação hidráulica de Sobradinho este trabalho segue a mesma abordagem para a realização da previsão.

Foram utilizadas nos experimentos deste projeto duas bases de dados reais pertencentes a bacia do rio São Francisco. Os dados foram cedidos pela Companhia Hidroelétrica do São Francisco (Chesf), órgão responsável pela coordenação e controle da operação das instalações de geração e transmissão de energia elétrica no Sistema Interligado Nacional (SIN).

A primeira base de dados é referente a vazão diária em m³/segundo do posto de São Francisco, situado à jusante da Barragem de Três Marias. O rio São Francisco nasce em Minas Gerais, na Serra da Canastra, e chega a sua foz, no Oceano Atlântico, entre Alagoas e Sergipe, percorrendo cerca de 2.800 km de extensão. Os dados são referentes as coletadas diárias desde janeiro de 1933 à dezembro de 2015.

Figura 9. Sistema Hidráulico do rio São Francisco



[fonte: <http://www3.ana.gov.br/porta/ANA/sala-de-situacao/sao-francisco/sao-francisco-saiba-mais>]

A segunda base de dados é referente a Usina Hidroelétrica de Sobradinho, localizada no estado da Bahia, a cerca de 40 km das cidades de Juazeiro (BA) e Petrolina (PE). Sendo inaugurada em 1979 o reservatório tem cerca de 320 km de extensão com capacidade de armazenamento de 34,1 bilhões [9]. Foram utilizados dados diários da vazão em m^3 /segundo da Usina de Sobradinho a partir de janeiro de 1931 até dezembro de 2017.

Figura 10. Usina Hidroelétrica de Sobradinho



[fonte: <http://www.naynneto.com.br/reducao-da-vazao-de-agua-da-barragem-sobradinho-e-autorizada-pelo-ibama/>]

3.2 Análise dos dados

Foi realizado o cálculo da máxima média quinzenal das vazões diárias escoados no posto de São Francisco e dos volumes afluentes à própria barragem de Sobradinho. Esses dados foram obtidos a partir do histórico de vazões médias diárias entre os meses de outubro, novembro, dezembro e janeiro, conforme mostrado no exemplo das figuras 11 e 12, respectivamente.

Figura 11. Exemplo do cálculo da média quinzenal para a base de dados da Usina Hidrelétrica de Sobradinho.

	NOVQ1 - Sb	NOVQ - Sb2	DEZQ1 - Sb	DEZQ2 - Sb	JANQ1 - Sb	JANQ2 - Sb
2005	1095.333333	1938.000000	3410.000000	4617.333333	4789.333333	3405.333333
2006	2071.466667	3084.000000	2712.666667	4030.000000	5009.333333	5505.333333
2007	824.000000	1088.000000	1579.333333	1994.000000	1968.000000	1842.666667
2008	935.333333	1384.666667	2317.333333	2146.000000	3806.666667	4484.000000
2009	2592.666667	2250.000000	1788.666667	2892.666667	3108.000000	2781.333333
2010	1255.333333	1770.000000	2476.666667	2553.333333	2815.333333	4124.000000
2011	1368.666667	1432.666667	2704.000000	3385.333333	4375.333333	5175.333333
2012	1261.333333	2436.000000	2625.333333	1792.000000	1394.000000	1840.000000
2013	874.000000	970.666667	1520.666667	3239.333333	4322.000000	2142.666667
2014	364.666667	868.000000	1364.666667	2077.333333	1328.000000	607.333333

Fonte: Elaboração do autor (2018)

Figura 12. Exemplo do cálculo da média quinzenal para a base de dados da Usina Hidrelétrica de São Francisco.

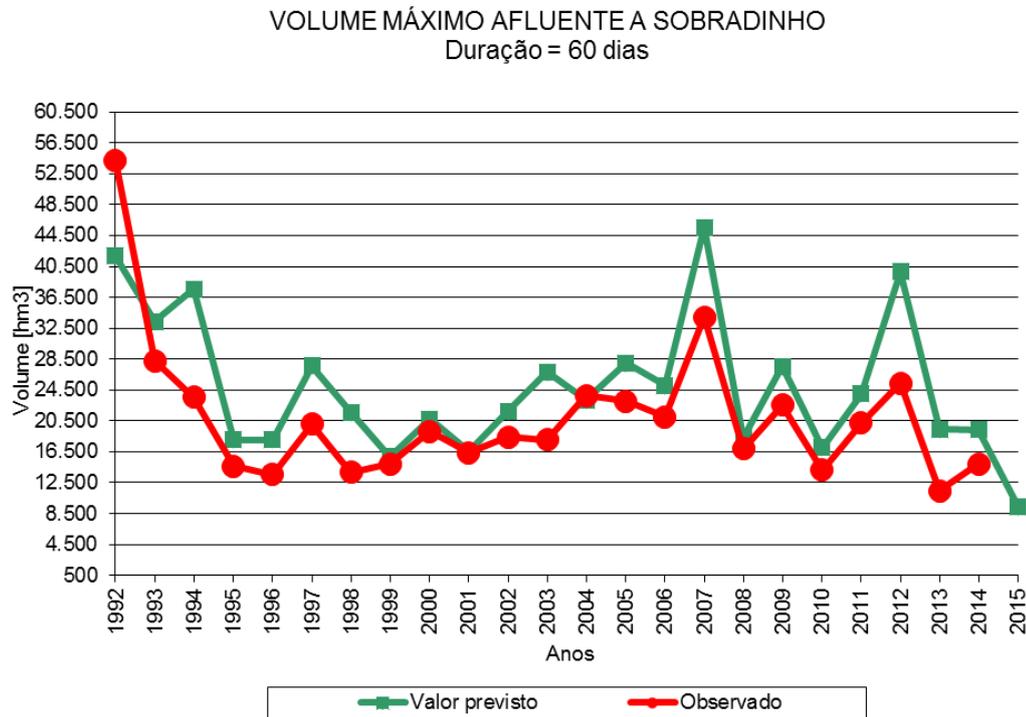
	NOVQ1 - SF	NOVQ2 - SF	DEZQ1 - SF	DEZQ2 - SF	JANQ1 - SF	JANQ2 - SF
2005	1408.866667	2521.600000	5052.666667	6222.066667	3608.666667	1872.800000
2006	2597.266667	2112.133333	4207.200000	7982.666667	7060.533333	6976.133333
2007	1172.933333	1227.133333	1616.866667	2281.200000	1781.266667	2024.200000
2008	1037.066667	1853.066667	1769.533333	3350.533333	5352.600000	3669.600000
2009	2754.400000	1620.866667	2501.733333	2271.866667	3529.666667	1935.266667
2010	1577.600000	1824.066667	2468.866667	1777.866667	4367.466667	3609.800000
2011	1188.466667	1881.533333	3095.000000	5420.600000	7396.133333	7123.066667
2012	1744.800000	2293.466667	1977.333333	1721.200000	1232.200000	2429.533333
2013	897.600000	1052.200000	2134.133333	5087.200000	2373.666667	1496.400000
2014	466.266667	1136.666667	1702.866667	1637.933333	684.666667	392.933333

Fonte: Elaboração do autor (2018)

Resultando em um total de 12 variáveis utilizadas como valores de entrada para a rede neural das médias quinzenais dentro o período de novembro e janeiro. Além disso, o outro cenário realizado neste trabalho utiliza as médias quinzenais entre os meses de outubro, novembro, dezembro e janeiro, resultando em 16 variáveis de entrada.

Foram usadas as vazões diárias para obter a previsão dos volumes máximos dos afluentes utilizando a média móvel simples do período de 60 e 30 dias dentre os meses de outubro e janeiro. Foi considerado esse meses pelo fato de corresponder ao período de maior estiagem da bacia do São Francisco.

Figura 13. Volume máximo do afluyente a sobradinho com duração de 60 dias.



Fonte: Elaboração do autor (2018)

A previsão é baseada na média pelo fato que os dados históricos conterem valores randômicos, que dificultam na previsão exata da demanda. Assim a média móvel remove as oscilações mais fortes contendo uma menor variabilidade. A equação da média móvel simples está descrita na equação 3.1 abaixo:

$$MMS_{t+1} = \frac{1}{n} \sum_{i=1-n}^t A_i \quad (3.1)$$

Onde:

- MMS_{t+1} é o valor final do período t ;
- A_i é o valor do período atual;
- n é o valor total de período em cada média;

3.3 Pré-processamento

Após a realização da média móvel simples é necessário realizar o processamento dos dados brutos coletados em dados significativos. Desta forma, a primeira etapa do pré-processamento dos dados é a normalização. Esta etapa visa garantir que o conjunto de dados façam parte de um mesmo intervalo. Evitará realizando este procedimento que valores altos possam ter uma maior influência e garantindo que as variáveis em intervalos diferentes recebam a mesma atenção com os valores mais baixos durante o treinamento.

As variáveis devem ter seus valores proporcionais aos limites da função de ativação usada na camada de saída, segundo Valença [15]. A técnica utilizada é a mais usada em problemas de séries temporais para realizar a normalização, a transformação linear descrita na equação 3.2 abaixo:

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

Onde:

- y_i = valor normalizado;
- x_i = valor original;
- x_{\min} = valor mínimo de x ;
- x_{\max} = valor máximo de x ;
- a e b = limites escolhidos.

De acordo com Valença [15], os intervalos da normalização devem ser definidos de acordo com a função de ativação utilizadas nas técnicas de previsão. Neste trabalho foi utilizadas a sigmoide logística como função de ativação, o intervalo de normalização utilizado nesta monografia foram limitados entre [0 e 1], logo os dados são geralmente normalizados entre [0,10 e 0,90] ou [0,15 e 0,85]. A tabela 1 abaixo mostra um exemplo do processo de normalização entre valores escolhidos aleatoriamente.

Tabela 1. Exemplo de normalização no intervalo [0-1]

Valores Originais	Normalização	Menor valor	Maior valor
18.24	0.78157	3.01	19.89
16.24	0.69863	3.01	19.89
15.00	0.64721	3.01	19.89
12.25	0.533175	3.01	19.89
10.00	0.439869	3.01	19.89
7.34	0.329561	3.01	19.89
6.34	0.288092	3.01	19.89
5.00	0.232523	3.01	19.89
4.96	0.230864	3.01	19.89

Fonte: Elaboração do autor (2018)

3.4 Arquitetura da rede MLP

Após a realização da análise dos dados e pré-processamento junto com a escolha da variáveis de entrada foi necessário definir os conjuntos dos parâmetros utilizados na rede neural. Foi utilizado o algoritmo da rede MLP desenvolvido em Java por alunos da Escola Politécnica de Pernambuco (POLI) utilizando a IDE de desenvolvimento Eclipse, a partir das pesquisas na aplicação de redes neurais na previsão da vazão [5][18].

A configuração da rede depende de parâmetros como a quantidade de neurônios na camada escondida, função de ativação, critério de parada de treinamento, taxa de aprendizagem, momento e o número máximo de ciclos.

O algoritmo de treinamento utilizado foi o *Backpropagation* e a função de ativação foi a sigmoide logística e na camada de saída, a função de ativação escolhida foi a sigmoide logística, descritas na seção 2.2. Os principais parâmetros utilizados e seus valores para a Rede Neural Artificial *Multi-layer Perceptron* (MLP) estão descrito na tabela 2 a seguir.

Tabela 2. Principais parâmetros utilizados e seus valores para a Rede Neural Artificial *Multi-layer Perceptron* (MLP)

Parâmetros	Rede MLP
Número de ciclos	600
Número de simulações	30
Algoritmo de treinamento	Backpropagation
Taxa de aprendizagem	0.85
Momentum	0.25
Neurônios na camada de entrada	12/16
Neurônios na camada escondida	10/20/30
Neurônios de saída	1
Função de ativação da camada escondida	Sigmoide logistica
Inicializacao dos pesos	Aleatória
Critério de parada	Cross Validation

Fonte: Elaboração do autor (2018)

A rede irá realizar o número total de ciclos treinamento, sendo que em cada ciclo será realizado até que seja atingido um ponto de convergência, e assim, a próxima simulação se iniciará. O método de parada utilizado é o de validação cruzada descrito descrito na seção 2.3

O experimento realizado para a técnica da rede MLP foi utilizar a melhor configuração encontrada na literatura (VALENÇA, 2009), com 6 neurônios na camada escondida. Este cenário irá ser comparado com os resultados obtidos pela rede RC.

3.5 Arquitetura da rede RC

Assim como o algoritmo MLP, a técnica RC também foi implementada em Java e foi desenvolvida por alunos da POLI no ambiente de desenvolvimento Eclipse. Conforme visto na seção anterior, a quantidade de experimentos é a mesma que arquitetura MLP, com 12 variáveis de entradas.

Os dados de entrada de vazão foram divididos em seis arquivos .csv: “*inputs*” e “*outputs*”, advindas do treino; “*crossinputs*” e “*crossoutputs*”, da validação cruzada; e “*testinputs*” e “*testoutputs*” para teste. Os nomes dos arquivos seguiam um padrão requisitado pelo desenvolvimento da rede neural. Nesta monografia os principais parâmetros adotados e seus valores para a Rede Neural Artificial *Reservoir Computing* (RC) estão descrito na tabela 3.

Tabela 3. Principais parâmetros utilizados e seus valores para a a Rede Neural Artificial *Reservoir Computing* (RC)

Parâmetros	Rede RC
Taxa de aprendizado	0.8
Número de simulações	30
Algoritmo de treinamento	Pseudoinversa de Moore-Penrose
Taxa de conectividade	0.3 - 0.4
Quantidade de ciclos de Warm Up	10
Neurônios no reservatório:	10/20/30
Neurônios de entrada	12/16
Neurônios de saída	1
Função de ativação do reservatório	Sigmoide logistica
Função de ativação da camada de saída	Linear
Inicializacao dos pesos	Aleatoria
Critério de parada	<i>Cross Validation</i>

Fonte: Elaboração do autor (2018)

Foram definidas a quantidade de neurônios no reservatório com base no desempenho para cada cenário, os cenários que foram desenvolvidos possuem 10, 20 e 30. Em relação a taxa de conectividade foi considerado pesquisas anteriores [5] [4] e retorna bons resultados com um menor tempo de processamento para previsão de vazão.

Além disso foi observado que nos trabalhos que 10 ciclos de *Warm Up* são suficientes para que o estado inicial da RC não afete no seu treinamento [5] [4]. Com as variáveis de entrada e as configurações definidas, a técnica de treinamento

proposta para a rede nesta monografia foi a pseudo-inversa de Moore-Penrose, descrita na seção 2.4.

3.6 Testes estatísticos

Em cada arquitetura foi realizada 30 experimentos gerando assim 30 erros percentuais médios absolutos (EPMA) após o término dos experimentos. A quantidade de amostras foi definida por poder assumir que as médias estariam em uma distribuição normal.

Os erros de cada configuração das redes servem de entrada para serem utilizados nos testes estatísticos que determinam quais configurações obtiveram melhor desempenho. Esta métrica de avaliação em cada experimento foi utilizada pois a base não possui dados faltantes e valores iguais a zero. A utilização deste tipo de erro, o EPMA pode ser expresso na equação 3.3.

$$EPMA = \sum_{i=1}^n \frac{\left| \frac{d_i - y_i}{d_i} \right| * 100}{n} \quad (3.3)$$

O EPMA é a média da diferença absoluta entre os valores previstos e os atuais onde d_i e y_i representam a saída desejada e a prevista para a i -ésima previsão i , respectivamente. O número n de previsão é a quantidade de previsões realizadas, também conhecida como rodadas.

Com o término dos experimentos, os erros percentuais médios são utilizados no comparativo dos testes estatísticos que determinam quais configurações obtiveram melhor desempenho. Nesta monografia é proposto o teste de Wilcoxon pelo fato de ser não-paramétrico, isto é, não se supõe nada a respeito da distribuição de seus dados, ou seja, se eles são normalmente distribuídos ou não. É apenas necessário que os dados sejam ordenáveis. O objetivo deste teste é verificar que as amostras apresentadas são estatisticamente iguais ou não. Para o desenvolvimento do teste estatístico os seguintes passos são seguidos:

1. Determinar duas possibilidades:
 - a. H_0 : As medianas entre as amostras são iguais.
 - b. H_1 : As medianas entre as amostras são diferentes.
2. Caso o p-value seja maior que 5% a amostra satisfaz H_0 , caso contrário satisfaz H_1 .

Para o desenvolvimento dos testes estatísticos foram realizados utilizando a linguagem de programação Python e o *software* RStudio que é um *software* livre para desenvolvimento da linguagem de programação R [25].

Capítulo 4

Resultados

Este capítulo tem como objetivo apresentar todos os resultados obtidos aplicados nas redes neurais MLP e RC utilizando dados de vazão diária da Usina Hidrelétrica de Sobradinho e de São Francisco. Nos experimentos, os dados disponíveis nas bases de dados correspondem ao período entre 1933 e 2015. A metodologia utilizada é a descrita no capítulo 3, esses experimentos visam demonstrar qual o melhor experimento para cada rede neural.

Foram realizados dois experimentos onde procura-se obter previsões com durações de 60 e 90 dias utilizando a máxima da média móvel simples, conforme descrito na seção 3.2. Em cada experimentos houve dois cenários para a utilização das arquiteturas das redes neurais. O primeiro utiliza as médias máximas quinzenais entre os meses de novembro, dezembro e janeiro totalizando em 12 variáveis de entrada.

O outro cenário utiliza as médias quinzenais entre os meses de outubro, novembro, dezembro e janeiro, resultando em 16 variáveis de entrada. Em cada cenário são realizadas utilizando a melhor configuração encontrada para a rede MLP, conforme descrito na seção 3.4. Os cenários com a arquitetura RC serão três, descritos na seção 3.5. Para a definição do número de neurônios utilizados na camada de reservatório precisou a realização de mais cenários para determinar a melhor arquitetura para as previsões da vazão.

As seções 4.1 e 4.2 apresentam esses experimentos realizados para a base de dados utilizando os cenários descritos anteriormente que foram realizadas 30 execuções, onde cada execução resultava em um erro percentual médio absoluto, discutido na seção 3.6. Por fim, na seção 4.3 é realizada à análise dos resultados obtidos entre os cenários.

4.1 Experimento com duração de 60 dias

Como explicado anteriormente, o experimento realizado utiliza as previsões com duração de 60 dias utilizando 12 e 16 variáveis de entrada. Esses valores de entrada variam de período, para o primeiro caso é utilizado entre os meses de novembro à janeiro, enquanto o segundo entre outubro e janeiro. Para cada ciclo executado foi obtido o erro percentual médio absoluto (EPMA). A tabela 4 indica a média dos erros percentuais nos experimentos para os cenários com 12 variáveis de entrada realizado para os experimentos com a duração da média móvel de 60 dias.

Tabela 4. Média de EPMA's para cada cenário utilizando média móvel de 60 dias com 12 variáveis.

Com 12 variáveis de entrada			
Cenário	Neurônios	Média do EPMA's	Desvio Padrão
MLP	6	22.98521	7.91641
RC-10	10	5.792784	0.759101
RC-20	20	6.30966	0.9240695
RC-30	30	11.27744	2.594474

Fonte: Elaboração do autor (2018)

A tabela 5 indica a média dos erros percentuais nos experimentos para os cenários com 16 variáveis de entrada também realizado com a duração da média móvel de 60 dias.

Tabela 5. Média de EPMA's para cada cenário utilizando média móvel de 60 dias com 16 variáveis.

Com 16 variáveis de entrada			
Cenário	Neurônios	Média do EPMA's	Desvio Padrão
MLP	6	20.63404	5.286927
RC-10	10	13.90685	1.526989
RC-20	20	18.73204	2.719519
RC-30	30	18.92088	3.08033

Fonte: Elaboração do autor (2018)

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.

4.1.1 Comparação entre os cenários MLP e RC-10

Aplicando o teste não-paramétrico de Wilcoxon obteve-se valor menor que 0.05 para o p-value nos dois cenários realizados, indicando que a mediana entre as amostras não é igual. Estes valores estão ilustrados na tabela abaixo. O cenário com 12 variáveis obteve o menor valor de EPMA com 5.792784%.

Tabela 6. Resultados dos p-value para cenários com cenário MLP e RC-10 com duração de 60 dias.

Cenário	Teste Estatístico	p-Value	p-Value > 5%
12 variáveis	Wilcoxon	2.2e-16	Não
16 variáveis	Wilcoxon	1.572e-13	Não

Fonte: Elaboração do autor (2018)

4.1.2 Comparação entre os cenários MLP e RC-20

O cenário com 12 variáveis apresentou um p-value menor que 5% no teste Wilcoxon, enquanto com 16 teve como resultado p-value maior que 5%, indicando

que a mediana entre as amostras é igual. Então, o cenário com 16 variáveis possui o melhor desempenho é com o menor EPMA. Os resultados dos testes podem ser visto na tabela 7.

Tabela 7. Resultados dos p-value para cenários com cenário MLP e RC-20 com duração de 60 dias.

Cenário	Teste Estatístico	p-Value	p-Value > 5%
12 variáveis	Wilcoxon	2.29e-15	Não
16 variáveis	Wilcoxon	0.0995	Sim

Fonte: Elaboração do autor (2018)

4.1.3 Comparação entre os cenários MLP e RC-30

Assim como nos outros cenários com 12 variáveis apresentou um p-value menor que 5% no teste Wilcoxon e com desempenho melhor com a rede RC. Para 16 variáveis o p-value foi maior que 5%, indicando que a mediana entre as amostras é igual. O melhor desempenho para este cenário é utilizando a arquitetura RC pois possui o menor EPMA, 18.92088%.

Tabela 8. Resultados dos p-value para cenários com cenário MLP e RC-30 com duração de 60 dias.

Cenário	Teste Estatístico	p-Value	p-Value > 5%
12 variáveis	Wilcoxon	5.998e-13	Não
16 variáveis	Wilcoxon	0.1421	Sim

Fonte: Elaboração do autor (2018)

Conforme visto para os experimentos com duração de 60 dias foi realizado os testes estatísticos utilizando os EPMA para realizar as comparações necessárias. A seção a seguir apresenta os resultados para os experimentos com duração de 30 dias.

4.2 Experimentos com duração de 30 dias

Conforme visto no experimento com previsão com média móvel de 30 dias, os experimentos a seguir foram divididos em 2 cenários com 12 e 16 variáveis de entrada utilizando a rede MLP e RC. Em cada cenário também foram realizadas 30 execuções resultando em um erro percentual médio absoluto calculado. As médias EPMA's nos experimentos para os cenários com 12 variáveis de entrada com a duração da média móvel de 30 dias são mostrados na tabela 9.

Tabela 9. Média de EPMA's para cada cenário utilizando média móvel de 30 dias com 12 variáveis

Com 12 variáveis de entrada			
Cenário	Neurônios	Média do EPMA's	Desvio Padrão
MLP	6	14.31481	4.279445
RC-10	10	6.163837	1.089593
RC-20	20	6.53111	0.9640175
RC-30	30	10.808779	3.66733

Fonte: Elaboração do autor (2018)

Assim como demonstrado na tabela anterior, a tabela 10 mostra os resultados média dos erros percentuais nos experimentos para os cenários com 16 variáveis de entrada também realizado com a duração da média móvel de 60 dias.

Tabela 10. Média de EPMA's para cada cenário utilizando média móvel de 30 dias com 16 variáveis

Com 16 variáveis de entrada			
Cenário	Neurônios	Média do EPMA's	Desvio Padrão
MLP	6	13.87759	2.94139
RC-10	10	6.146646	1.058894
RC-20	20	8.549325	1.309061
RC-30	30	13.59423	3.60351

Fonte: Elaboração do autor (2018)

4.2.1 Comparação entre os cenários MLP e RC-10

Realizando a comparação entre os cenários verifica-se que ambas obtiveram um p-value menor que 5% no teste Wilcoxon. O resultado com 16 variáveis apresentou a média do EPMA um pouco menor comparado com 12 variáveis, com 6.163837%. O resultado dos testes podem ser visto na tabela 11.

Tabela 11. Resultados dos p-value para cenários com cenário MLP e RC-10 com duração de 30 dias.

Cenário	Teste Estatístico	p-Value	p-Value > 5%
12 variáveis	Wilcoxon	2.2e-16	Não
16 variáveis	Wilcoxon	2.2e-16	Não

Fonte: Elaboração do autor (2018)

4.2.2 Comparação entre os cenários MLP e RC-20

Assim como os cenários anteriores os resultados do p-value foi menor que 5% no teste Wilcoxon. O resultado do teste podem ser visto na tabela 12.

Tabela 12. Resultados dos p-value para cenários com cenário MLP e RC-20 com duração de 30 dias

Cenário	Teste Estatístico	p-Value	p-Value > 5%
12 variáveis	Wilcoxon	2.25e-16	Não
16 variáveis	Wilcoxon	3.991e-15	Não

Fonte: Elaboração do autor (2018)

4.2.3 Comparação entre os cenários MLP e RC-30

Com 12 variáveis apresentou p-value menor que 5% no teste. Enquanto o cenário com 16 variáveis obteve-se um p-value maior que 5%, o que indica que a mediana entre as amostras é igual, conforme demonstrado na tabela 13. Logo, o

cenário que obteve o melhor desempenho é o que possui o menor EPMA dentre os dois, sendo o RC.

Tabela 13. Resultados dos p-value para cenários com cenário MLP e RC-30 com duração de 30 dias

Cenário	Teste Estatístico	p-Value	p-Value > 5%
12 variáveis	Wilcoxon	0.0001433	Não
16 variáveis	Wilcoxon	0.6973	Sim

Fonte: Elaboração do autor (2018)

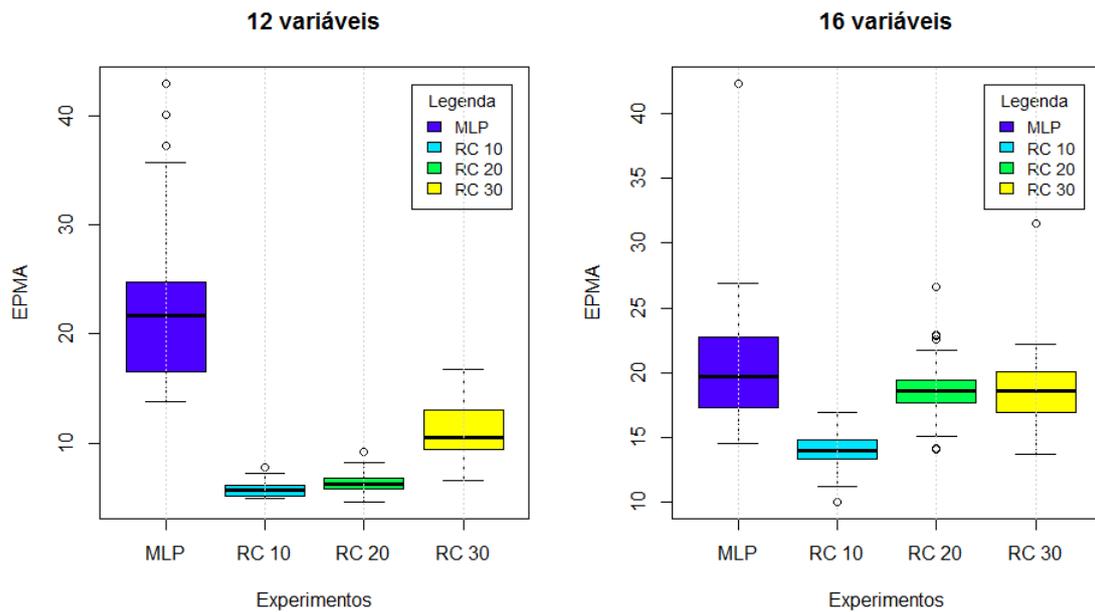
4.3 Resumo dos experimentos

Após executados todos os testes estatísticos para facilitar a compreensão dos resultados e determinar qual técnica é a melhor, foi feita a representação dos resultados em *Box Plots*. Verifica-se que o experimento que utiliza a rede RC possui melhor representação com menor quantidade de neurônios na camada do reservatório em cada experimento realizado.

4.3.1 Resultado para previsão com duração de 60 dias

Observa-se na figura 14 os resultados em *Box Plots* para os experimentos com duração de 60 dias com os cenários de 12 e 16 variáveis. A cor azul escura representa o cenário com a rede MLP, enquanto as cores azul claro, verde e amarelo representa o uso da RC. A sigla MLP inserida nos gráficos de *Boxplot* significa o uso da rede *Multi-layer Perceptron* utilizando 6 neurônios na camada escondida, enquanto RC 10, RC 20 e RC 30 significam a rede RC com 10, 20 e 30 neurônios no reservatório, respectivamente.

Figura 14. Boxplot dos cenários com duração 60 dias



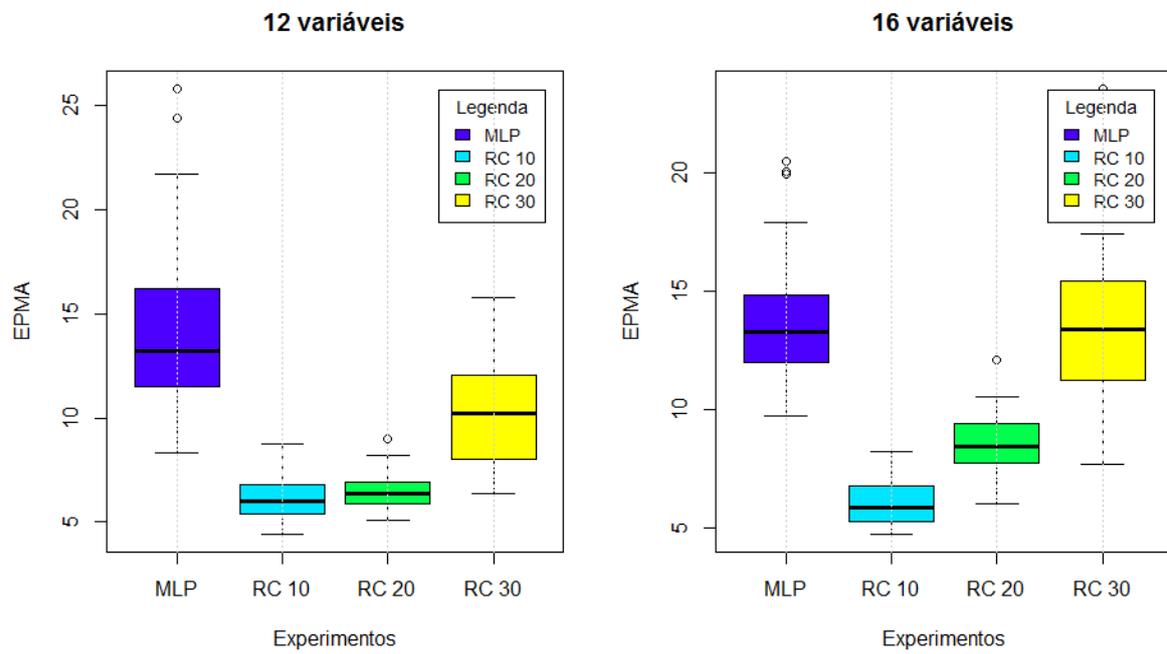
Fonte: Elaboração do autor (2018)

Em ambas situações o melhor resultado é para a rede RC com apenas 10 neurônios na camada do reservatório. O resultado que possui o menor valor de EPMA é com 12 variáveis que possui 5.79%. Os piores caso encontrados é referente a rede MLP, contudo o cenário com 12 variáveis de entrada é encontrado o pior com média de EMPAs 22.98%. Logo a rede que possui melhor desempenho é a RC.

4.3.2 Resultado para previsão com duração de 30 dias

Assim como o resultado do experimento de 60 dias os melhores resultados se encontram com a rede RC que utiliza 10 neurônios na camada de reservatório com 12 variáveis.

Figura 15. Boxplot dos cenários com duração 30 dias



Fonte: Elaboração do autor (2018)

Capítulo 5

Conclusão

Este capítulo visa descrever as considerações finais deste trabalho na seção 5.1. Por fim, os trabalhos futuros na seção 5.2.

5.1 Considerações Finais

A previsão de vazão das bacias hidrográficas é realizada diariamente para o planejamento energético do país e prevenção para o uso da usina. Assim, devido as recentes secas que vem aumentando especialmente na região nordeste do Brasil, conforme visto no reservatório de Sobradinho, vem crescendo a necessidade pela melhora da precisão das previsões. Desta forma, possibilitará em um melhor gerenciamento das bacias e prevendo os impactos que a poluição poderá sofrer em situações como casos críticos. Desta forma, esta monografia realizou uma análise comparativa das redes neurais artificiais MLP e RC para prevê a vazão utilizando médias modais entre os períodos de 30 e 60 dias considerando meses de dados de vazão passados.

Pelos resultados obtido nos experimentos realizados observa-se que a rede RC utilizando apenas 10 neurônios no reservatório gerou melhores resultados para todas os cenários e conseguiu efetuar previsões com menores EPMA's, assim o algoritmo obteve a melhor performance em cada cenário realizado nos experimentos. Em comparação com a rede MLP, obteve os piores resultados em todos os cenários.

5.2 Trabalhos Futuros

Para trabalhos futuros, alguns aspectos podem ser utilizados:

- Realizar mais experimentos na rede RC determinando outros parâmetros para sua configuração.
- Realizar mais experimentos nas redes e MLP e RC utilizando mais períodos de previsões e utilizando outros parâmetros para as suas configurações.
- Obter dados de outras bases e comparar os resultados da previsão diária da vazão utilizando as redes RC e MLP

Referências

- [1] **A bacia.** Disponível em: <<http://cbhsaofrancisco.org.br/2017/a-bacia/>>. Acessado em 13 abril, 2018.
- [2] **Acompanhamento da Bacia do Rio São Francisco.** Disponível em:<http://arquivos.ana.gov.br/saladesituacao/BoletinsDiarios/SF_19-10-2017.pdf>. Acessado em 20 abril, 2018
- [3] ALEKSANDER, I. e MORTON, H. **An Introduction to Neural Computing.** International Thomson Computer Press, 1995. 490 p.
- [4] AGUIAR, Bruna Cavalcanti Galle de. **Utilizando Reservoir computing para previsão na geração de energia de Parques Elólicos brasileiros.** 2015. 58 f. Dissertação (Mestrado) - Curso de Pós-graduação em Engenharia da Computação, Universidade de Pernambuco, Recife, 2015
- [5] BEZERRA, Sabrina Guimarães Tavares de Andrade. **Reservoir compotinha 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.
- [6] "BIG – Banco de Informações de Geração: **Capacidade de geração do Brasil**" [Online] available: <http://www2.aneel.gov.br/aplicacoes/capacidade-brasil/capacidadebrasil.cfm>. Acesso em 23 maio de 2018.
- [7] D. Verstraeten, **Reservoir computing: computation with dynamical systems**, Ph. D dissertation 2009.
- [8] DEBASTIANE, Aline bernarda; SILVA, Riardo Dal` Agnol; PADILHA, Victor Luís; SÁ, Eder Alexandre; RAFAELI, Silvio Luís, ANTUNES, Thiago Alves. **XXI Simpósio brasileiro de recursos hídricos**, 2015, Brasília. Modelagem.
- [9] "**Descrição do Aproveitamento de Sobradinho**" [Online] available: <https://www.chesf.gov.br/SistemaChesf/Pages/SistemaGeracao/Sobradinho.a.spx>. Acessado em 22 abril de 2018.

- [10] “**Energy Resources**” [Online] available: <https://www.worldenergy.org/data-resources/resource/hydropower/>. Acessado em 29 abril de 2018.
- [11] ENGELBRECHT, A. P. **Computational Intelligence An Introduction**. 1. ed. West Sussex, Inglaterra: Wiley, 2002.
- [12] FERREIRA, A. A. **Um método para design e treinamento de reservoir computing aplicado à previsão de séries temporais**, Tese de Doutorado, Universidade de Pernambuco, 2011.
- [13] “**Reservatório de Sobradinho, no Nordeste, vive pior seca em 83 anos**” [Online] available: <http://g1.globo.com/bom-dia-brasil/noticia/2015/10/reservatorio-de-sobradinho-no-nordeste-vive-pior-seca-em-83-anos.html/>. Acessado em 3 de abril de 2018.
- [14] “**Matriz de Energia Elétrica**” [Online] <http://www2.aneel.gov.br/aplicacoes/capacidadebrasil/OperacaoCapacidadeBrasil.cfm>. Acessado em 29 abril de 2018.
- [15] M. Valença, **Fundamento das Redes Neurais**, Olinda: Livro Rápido, 2009.
- [16] M. J. S. Valença, A. R. T. Cavalcanti e S. C. Oliveira. **Caminho Crítico da Decisão para a Operação Hidráulica de Sobradinho**. I Simpósio de recursos Hídricos do Nordeste, Vol 1, pages 23 – 30, Recife 1992.
- [17] Martinez. A, Allodi. S, Uziel D, **Neuroanatomia Essecial**, 1 edição, Editora: Guanabara Koogan, Ano 2014
- [18] NOVAES, R. V. **Análise comparativa de redes multi-layer perceptron (MLP) e reservoir computing (RC) para previsão de vazão**, Monografia, Universidade de Pernambuco, 2017.
- [19] “**Relatórios - CHESF**” [Online]: <http://www3.ana.gov.br/porta/ANA/sala-de-situacao/sao-francisco/relatorios-chesf/relatorios-chesf>. Acessado em 16 maio de 2018.

- [20] **“Relatório Final BEN 2017”**. Disponível em: <http://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-46/topico-82/Relatorio_Final_BEN_2017.pdf>. Acessado em 06 abril de 2018.
- [21] **“Sala de situação”** [Online] available: <http://www3.ana.gov.br/portal/ANA/sala-de-situacao/sao-francisco/sao-francisco-saiba-mais>. Acessado em 07 de maio de 2018.
- [22] SANTOS, P. H. Q. **Aplicação de Multilayer Perceptron e Reservoir Computing para previsão de vazão considerando sazonalidade**, Monografia, Universidade de Pernambuco, 2017.
- [23] SOARES, E. S. M. **Multi-layer perceptron e reservoir computing aplicadas em um processo chuva x vazão**, Monografia, Universidade de Pernambuco, 2016.
- [24] **“Potencial hidrelétrico brasileiro está entre os cinco maiores do mundo”** [Online] available: <http://www.brasil.gov.br/editoria/infraestrutura/2011/12/potencial-hidreletrico-brasileiro-esta-entre-os-cinco-maiores-do-mundo>. Acessado em 11 de abril de 2018.
- [25] **“The R Project for Statistical computing”** [Online]. Available: <https://www.rproject.org/>. Acessado em 15 de abril de 2018
- [26] VALENÇA, M. J. S. **Prevendo com redes neurais: o estado da arte**. Artigo apresentado no XVI Simpósio Brasileiro de Recursos Hídricos, João PessoaPB, 2005.
- [27] **“World Energy Resources”** [Online] <https://www.worldenergy.org/wp-content/uploads/2016/10/World-Energy-Resources-Full-report-2016.10.03.pdf>. Acessado em 04 de maio de 2018