



MULTI-LAYER PERCEPTRON E RESERVOIR COMPUTING APLICADAS EM UM PROCESSO CHUVA x VAZÃO

Trabalho de Conclusão de Curso

Engenharia de Computação

**Everalda Solange de Moraes Soares
Orientador: Mêuser Jorge Silva Valença**



**UNIVERSIDADE
DE PERNAMBUCO**

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

Everalda Solange de Moraes Soares

***MULTI-LAYER PERCEPTRON E RESERVOIR
COMPUTING APLICADAS EM UM PROCESSO
CHUVA x VAZÃO***

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

Recife, Novembro de 2016.

MONOGRAFIA DE FINAL DE CURSO

Avaliação Final (para o presidente da banca)*

No dia 22 de 12 de 2016, às 9:00 horas, reuniu-se para deliberar a defesa da monografia de conclusão de curso do discente **EVERALDA SOLANGE DE MORAES SOARES**, orientado pelo professor **Mêuser Jorge Silva Valença**, sob título **Multi-Layer Perceptron e Reservoir Computing Aplicadas em um Processo Chuva x Vazão**, a banca composta pelos professores:

Mêuser Jorge Silva Valença

Sérgio Mario Lins Galdino

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 (nove, cinco)

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

O discente terá 07 dias para entrega da versão final da monografia a contar da data deste documento.

MÊUSER JORGE SILVA VALENÇA

SÉRGIO MARIO LINS GALDINO

Este trabalho é dedicado a Escola de Referência em Ensino Médio Joaquim Olavo por ter me dado a oportunidade de chegar onde eu cheguei. Por todo o ensinamento provido pelos meus professores principalmente de matemática, física e química.

Agradecimentos

Agradeço a todos que nunca me deixaram desistir de viver, a todos os professores que contribuíram para a minha formação os quais foram como pais na minha caminhada e acreditaram que eu poderia ir além. Agradeço a minha mãe Solange Barbosa por toda a sua vontade e carinho em me ajudar a ter um futuro melhor. Agradecimento especial ao professor Mêuser Valença por ter aceitado ser meu orientador e por toda a sua paciência. A todos vocês meus sinceros agradecimentos.

Resumo

O uso de recursos hídricos pode ser observado desde a agricultura irrigada à produção de energia elétrica. Essas atividades são diretamente afetadas pelas mudanças nas condições climáticas, portanto, o gerenciamento de bacias hidrográficas tornou-se um dos principais focos para controle do uso de recursos hídricos e energéticos em todos os países. O desafio é achar maneiras de aproveitar adequadamente os recursos hídricos do planeta. No Brasil, a maior parcela da energia elétrica produzida provém de usinas hidroelétricas. Dependendo da vazão das bacias hidrográficas, a força das águas pode aumentar a quantidade de energia a ser gerada. Este trabalho propõe o uso de redes neurais na previsão da vazão de bacias hidrográficas em um processo chuva x vazão. Algumas configurações serão utilizadas em um estudo comparativo entre redes *Multi-layer Perceptron* e *Reservoir Computing*. Como resultado deste estudo comparativo será identificada a rede que possui melhor desempenho neste processo através de testes estatísticos.

Abstract

The use of water resources can be observed from irrigated agriculture to production of electricity. These activities are directly affected by changes in climatic conditions, so watershed management has become a major focus for the use of water and energy resources in all countries. The challenge is find ways to properly take advantage of the planet's water resources. In Brazil the largest part of the electricity produced comes from hydroelectric power plants. Depending on the watershed's flow, the water strength can increase the amount of electricity to be generated. This work proposes the use of neural networks in the prediction of the flow of watersheds in a rainfall x flow process. Some configurations will be used in a comparative experiment between Multilayer Perceptron and Reservoir Computing networks. As a result of this comparative study, the network that has higher performance is identified by statistical tests.

Sumário

Capítulo 1 Introdução	1
1.1 Caracterização do Problema e Motivações.....	1
1.2 Objetivos e Metas.....	2
1.2.1 Objetivo Geral.....	2
1.2.2 Objetivos Específicos.....	3
1.3 Estrutura da Monografia	3
Capítulo 2 Fundamentação Teórica	4
2.1 Produção Energética Brasileira	4
2.1.1 Geração de Energia Elétrica por Hidroelétricas	4
2.2 Redes Neurais.....	6
2.3 <i>Multi-Layer Perceptron</i>	9
2.4 <i>Reservoir Computing</i>	12
Capítulo 3 Metodologia	15
3.1 Base de dados.....	15
3.2 Pré- processamento dos dados	16
3.3 Apresentando as variáveis de entrada a MLP.....	19
3.3.1 Configurações utilizadas na rede MLP.....	20
3.4 Apresentando as variáveis de entrada a RC.....	21
3.4.1 Configurações utilizadas nas redes RC.....	21
3.5 Testes estatísticos	22
3.5.1 Teste Shapiro-wilk.....	23
3.5.2 Teste F.....	23
3.5.3 Teste t de Student.....	24

3.5.4 Teste de Wilcoxon.....	24
Capítulo 4 Resultados	25
4.1 Experimentos Base Serra do Facão	25
4.1.1 Hidroelétrica Serra do Facão: Estudo de caso com a rede MLP	25
4.1.2 Hidroelétrica Serra do facão: Estudo de caso com a rede RC	26
4.1.2.1 comparação do Cenário 1 com o Cenário 2.....	27
4.1.2.2 comparação do Cenário 2 com o Cenário 3.....	27
4.1.2.3 comparação do Cenário 2 com o Cenário 4.....	28
4.1.2.4 comparação do Cenário 2 com o Cenário 5.....	29
4.2 Experimentos Base Corumbá IV	30
4.2.1 Hidroelétrica Corumbá IV: Estudo de caso com a rede MLP.....	30
4.2.2 Hidroelétrica Corumbá IV: Estudo de caso com a rede RC.....	31
4.2.2.1 comparação do Cenário 1 com o Cenário 2.....	32
4.2.2.2 comparação do Cenário 1 com o Cenário 3.....	33
4.2.2.3 comparação do Cenário 1 com o Cenário 4.....	34
4.2.2.4 comparação do Cenário 1 com o Cenário 5.....	34
4.3 Comparativo dos resultados entre a MLP e RC	35
4.3.1 Hidroelétrica Serra do Facão MLP X RC.....	35
4.3.2 Hidroelétrica Corumbá IV MLP X RC.....	36
Capítulo 5 Considerações Finais	37
5.1 Conclusão.....	37
5.2 Trabalhos futuros.....	37
Referências	39

Índice de Figuras

Figura 1. Usina Itaipu.....	1
Figura 2. Desenho Esquemático de uma Usina Hidroelétrica.....	6
Figura 3. Neurônio Biológico.....	6
Figura 4. Comunicação entre neurônios.....	7
Figura 5. Modelo do Neurônio de Mcculloch e Pitts.....	8
Figura 6. Uma rede MLP.....	9
Figura 7. Arquitetura de uma rede recorrente RC.....	12
Figura 8. Usina Hidroelétrica Serra do Facão.....	15
Figura 9. Usina Hidroelétrica Corumbá IV.....	16

Índice de Tabelas

Tabela 1. Usinas Hidroelétricas Instaladas no Brasil.....	4
Tabela 2. Normalização no intervalo [0-1].....	17
Tabela 3. Base defasada até (t-2).....	18
Tabela 4. Correlação entre variáveis da Tabela 3.....	18
Tabela 5. Média das EPMA's para cada cenário na Serra do Facão.....	26
Tabela 6. Resultados do p <i>value</i> para os Cenário 1 e 2 na Serra do Facão.....	27
Tabela 7. Resultados do p <i>value</i> para os Cenário 2 e 3 na Serra do Facão.....	28
Tabela 8. Resultados do p <i>value</i> para os Cenário 2 e 4 na Serra do Facão.....	29
Tabela 9. Resultados do p <i>value</i> para os Cenário 2 e 5 na Serra do Facão.....	30
Tabela 10. Média das EPMA's para cada cenário na Corumbá IV.....	31
Tabela 11. Resultados do p <i>value</i> para os Cenário 1 e 2 na Corumbá IV	32
Tabela 12. Resultados do p <i>value</i> para os Cenário 1 e 3 na Corumbá IV	33
Tabela 13. Resultados do p <i>value</i> para os Cenário 1 e 4 na Corumbá IV.....	34
Tabela 14. Resultados do p <i>value</i> para os Cenário 1 e 5 na Corumbá IV	35
Tabela 15. Resultados do p <i>value</i> para MLP e RC na Serra do Facão	36
Tabela 16. Resultados do p <i>value</i> para MLP e RC na Corumbá IV	36

Tabela de Símbolos e Siglas

ANEEL – Agência Nacional de Energia Elétrica

EMQ – Erro Médio Quadrático

EPMA – Erro Percentual Médio Absoluto

RNAs – Redes Neurais Artificiais

MLP – *Multi-Layer Perceptron*

MME – Ministério de Minas e Energia

RC – *Reservoir Computing*

Capítulo 1

Introdução

Este capítulo encontra-se dividido em três seções. A primeira seção levanta as características do problema encontrado 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

A utilização de recursos hídricos pode ser observada na agricultura e práticas de engenharia. Essas atividades são diretamente afetadas pelas mudanças nas condições climáticas que adicionam incertezas futuras na sua gestão [1]. Com isso, o gerenciamento de bacias hidrográficas tornaram-se um dos principais focos para controle do uso de recursos hídricos e energéticos em todos países [2].

O desafio é achar maneiras de aproveitar adequadamente os recursos hídricos do planeta. Uma das principais utilizações desse recurso é na geração de energia elétrica. Segundo a Agência Nacional de Energia Elétrica (ANEEL, 2010) 75% da matriz energética do Brasil é composta por hidroelétricas, mas menos de 30% do parque hidroelétrico é aproveitado [7].

Figura 1. Usina Itaipu



[fonte:<http://mundoeducacao.bol.uol.com.br>]

Tendo essa informação como base, nota-se que é importante tomar decisões validadas com informações confiáveis como, por exemplo, valores baseados em dados estatísticos e probabilísticos para que se possa ter uma boa administração da bacia. A previsão da vazão nas usinas hidroelétricas é uma das práticas utilizadas que pode minimizar os impactos no seu gerenciamento e conseqüentemente a racionalização do uso da energia [3].

Existem vários métodos que tentam diminuir o erro ao fazer a previsão da vazão de uma hidroelétrica usando as séries temporais, alguns modelos estatísticos como, por exemplo, os métodos de previsão baseados na metodologia de Box-Jenkins são utilizados para fazer previsão desse tipo de série [4]. Uma das principais desvantagens desse método é que na prática a maioria dos problemas são não lineares e estes modelos são, de modo geral, lineares. A partir disto é notório que é preciso usar outros métodos para contornar esse tipo de problema.

Estudos recentes mostram que o uso das redes neurais artificiais (RNAs) para a previsão de series temporais vem aumentando. Principalmente na previsão da vazão e chuva [5] [6]. Entre essas redes estão a *Multi-Layer Perceptron* (MLP) e a *Reservoir Computing* (RC).

A fim de ajudar neste cenário, este trabalho pretende realizar um estudo comparativo dos resultados na previsão da vazão em um processo chuva x vazão usando as redes MLP e RC. Verificando-se desta forma a rede neural que obteve resultados mais eficientes, ou seja, a que possui um menor erro de acerto na previsão da vazão.

1.2 Objetivos e Metas

1.2.1 Objetivo Geral

O objetivo principal desta monografia é comparar os resultados na previsão da vazão em um processo chuva x vazão usando as redes neurais MLP e RC.

1.2.2 Objetivos Específicos

1. Estudo e uso dos algoritmos das redes MLP e RC que foram implementadas por alunos das Escola Politécnica de Pernambuco.
2. Obter os erros nas previsão das vazões diárias em um processo chuva x vazão tanto para a rede MLP como para a RC.
3. Analisar os dados estatisticamente para saber qual rede que obtém melhor resultado.

1.3 Estrutura da Monografia

Esta monografia foi estruturada da seguinte forma:

- **Capítulo 1- Introdução**
Neste capítulo é apresentada as principais motivações da pesquisa apontando o problema e objetivos a serem alcançados.
- **Capítulo 2 - Fundamentação Teórica**
Este capítulo traz todo conhecimento relevante para o entendimento do problema e sua solução, sendo apresentado a produção energética brasileira e as redes neurais artificiais tendo como foco as arquiteturas utilizadas nesta monografia, as redes MLP e RC.
- **Capítulo 3 - Metodologia**
Neste capítulo é apresentada a metodologia empregada nesta monografia, onde descreve as bases de dados utilizadas e todo o tratamento de seus dados, assim como as configurações empregadas na rede MLP e RC e os testes de avaliação desses algoritmos.
- **Capítulo 4 - Resultados**
Neste capítulo, têm-se o detalhamento dos experimentos realizados e o resultado da comparação dos cenários através dos testes estatísticos.
- **Capítulo 5 - Considerações Finais**
Este capítulo, aborda as conclusões a partir dos resultados do capítulo anterior e os trabalhos futuros.

Capítulo 2

Fundamentação Teórica

Este capítulo aborda o conteúdo necessário para compreendermos o problema descrito no capítulo anterior e a solução proposta. Na seção 2.1 é descrita a produção energética brasileira. Na seção 2.2 é conceituado o que são redes neurais artificiais e as duas sessões seguintes 2.3 e 2.4 são abordadas as redes empregadas neste trabalho, a rede MLP e RC respectivamente.

2.1 Produção Energética Brasileira

A produção de energia, principalmente da energia elétrica no Brasil, ingressou no século XXI, em busca do desenvolvimento sustentável [7]. No qual tenta-se associar o aumento da oferta de energia, o consumo adequado e a preservação do meio ambiente. Uma das principais fontes de geração energia sustentável é a energia provida por hidroelétricas.

2.1.1 Geração de Energia Elétrica por Hidroelétricas

Para a geração e transmissão de energia elétrica, o Brasil conta com um conjunto de mais de 100 usinas hidroelétricas. Os principais rios onde as usinas mais produtivas foram instaladas e sua respectiva capacidade de geração são listadas a seguir.

Tabela 1. Usinas Hidroelétricas Instaladas no Brasil

Rio	Usinas Instaladas	Capacidade(kW)
Tocantins – PA	1	8.535.000
São Francisco – AL	6	7.441.601
Grande – MG	12	7.203.500
Paraná – PR	2	7.000.085
São Francisco – BA	7	6.809.501
Iguaçu – PR	5	6.674.000

Paraná – MS	4	6.544.038
Paraná – SP	3	6.535.200
Madeira – RO	2	6.525.760
Paranaíba – GO	5	5.640.802
Grande – SP	7	5.217.400
Paranaíba – MG	4	4.982.800
São Francisco – SE	1	3.162.000
Tocantins – TO	4	2.731.450
Paranapanema – SP	11	2.398.156
Uruguai – SC	2	2.305.000
Uruguai – RS	2	2.305.000
Paranapanema – PR	8	2.184.700
Tietê – SP	11	1.906.224
Pelotas – SC	2	1.838.250

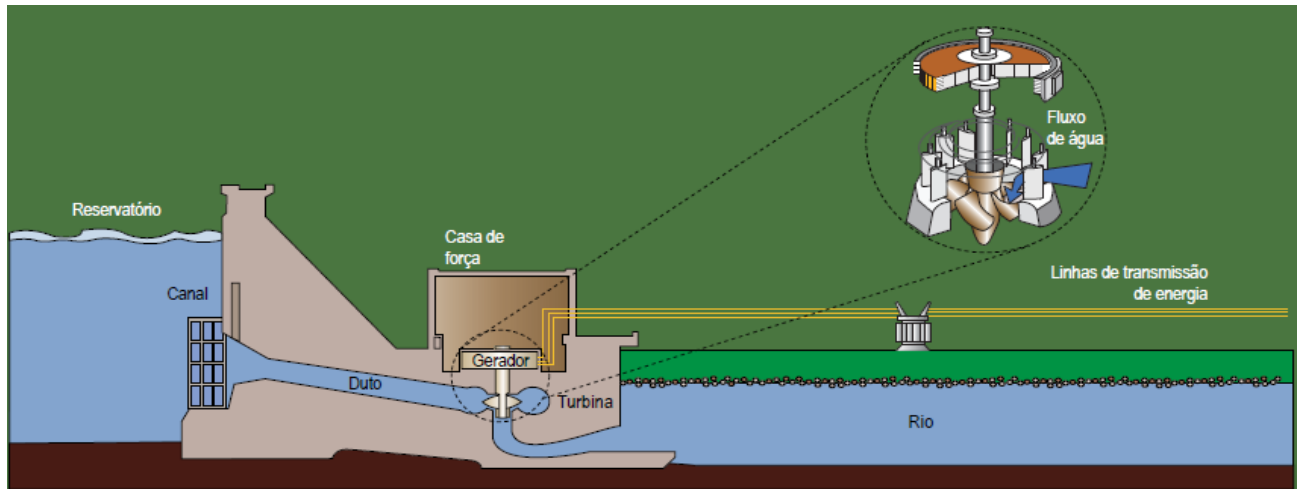
[fonte: ANEEL]

Todo o ano o Ministério de Minas e Energia (MME) faz um balanço energético a partir dos resultados preliminares do ano anterior com o intuito de expandir a oferta de energia e verificar a sua demanda no Brasil e em alguns outros países [7].

Para produzir energia providas por hidroelétricas é necessário checar a vazão do rio diariamente, ou seja a quantidade de água que passa pela barragem. A barragem tem a função de interromper o curso do rio com o intuito de formar um reservatório, o reservatório por sua vez forma uma espécie de desnível com o rio sendo capaz de captar a quantidade de água pertinente, tendo assim a função de regular a vazão do rio para o período de muita chuva ou estiagem.

Segundo a ANEEL uma usina hidroelétrica é esquematizada por possuir um canal por onde a água que vem do reservatório irá passar, e é neste canal onde se encontra as turbinas (que é uma estrutura formada por um conjunto de pás ligadas a um ponto central chamado de eixo, e este eixo esta conectado ao gerador). Com o movimento giratório promovido pelo movimento da água, ocorre a transformação da energia cinética em energia elétrica. Depois de passar pelas turbinas a água é devolvida ao rio.

Figura 2. Desenho Esquemático de uma Usina Hidroelétrica



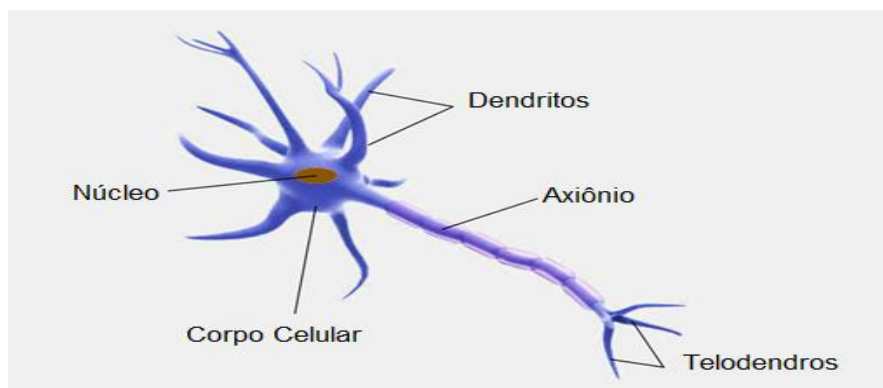
[fonte: Atlas de energia elétrica do Brasil – ANEEL]

2.2 Redes Neurais

Segundo Valença (2012) para se entender o comportamento de um neurônio artificial é necessário conhecer primeiro as características do neurônio biológico e como ele se comporta na transmissão de impulsos de um neurônio para outro.

Na literatura, de forma geral, o neurônio pode ser conceituado como uma célula do sistema nervoso de onde surgem ramificações (dendritos) e um prolongamento chamado de axônio, do axônico saem os terminais chamados de telodendros.

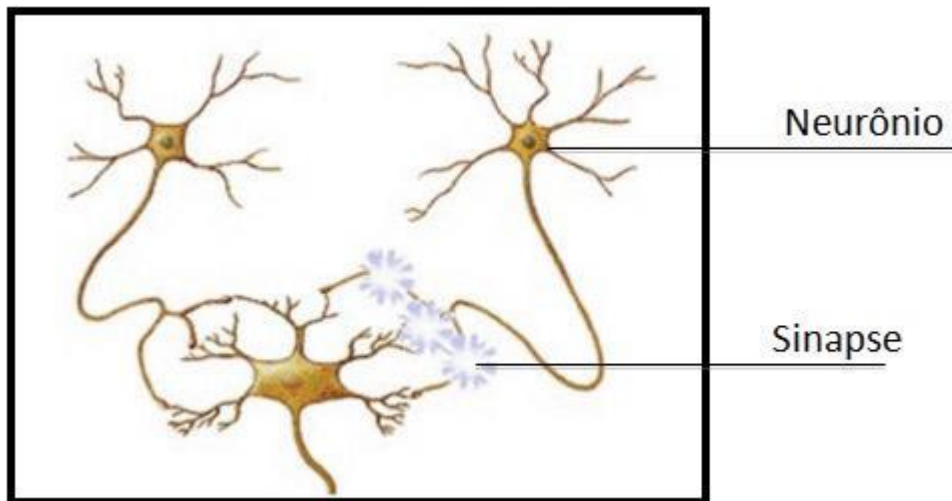
Figura 3. Neurônio Biológico



[fonte: <http://dicionariosaude.com/neuronio-celula-nervosa/>]

Na região chamada de sinapse é onde ocorre a passagem do impulso nervoso entre os neurônios, ou seja, a região onde acontece a comunicação dos neurônios. Esta região é constituída pelo axônio de um neurônio e pelos dendritos de um outro neurônio, sendo a transmissão do impulso transmitido neste sentido.

Figura 4. Comunicação entre neurônios

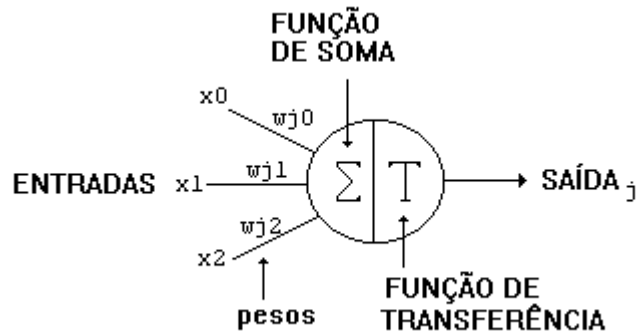


[fonte: http://www.neurovitta.com.br/ctd_noticias_detalhes.php?id=OQ]

De maneira simples o neurônio artificial segundo Valença é um “conjunto de elementos de processamento interconectados que geralmente encontram se em camadas”[8]. A primeira representação de um neurônio em forma matemática foi apresentada por McCulloch e Pitts em 1943 [8]. Neste modelo é onde foi introduzida a lei do tudo ou nada, esta lei é a que determina se o neurônio libera ou não o impulso nervoso para um outro neurônio. Para que esta transmissão ocorra, o impulso tem que ser maior que um dado limiar, chamado de limiar excitatório.

Este modelo é formado por um conjunto de valores de entrada ($x_0, x_1, x_2, \dots, x_n$), pelas sinapses que são os pesos entre a camada de entrada e camada de saída ($W_{j0}, W_{j1}, W_{j2}, \dots, W_{jn}$), por uma regra de propagação e uma função chamada de função de ativação ou de transferência.

Figura 5. Modelo do Neurônio de Mcculloch e Pitts



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

A regra de propagação é dada por:

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

A função de ativação é a função degrau dada por:

$$f(net_i) = 1, \forall net_i \geq 0 \quad (2.2)$$

$$f(net_i) = 0, \forall net_i < 0$$

Entre os anos 50 e 60 surgiram duas novas arquiteturas: a *Perceptron* em 1957 e a *Adaline* em 1960. Com a primeira foi introduzida o conceito de aprendizado supervisionado ou com professor para ajustar os pesos, já a segunda permitiu o uso de valores contínuos através a introdução de novas funções de ativação.

Segundo Valença (2012) as principais funções de ativação são: a função linear (2.3), a função sigmoide logística (2.4) e a tangente hiperbólica (2.5).

$$f(net_i) = net_i \quad (2.3)$$

$$f(net_i) = \frac{1}{1+e^{-net_i}} \quad (2.4)$$

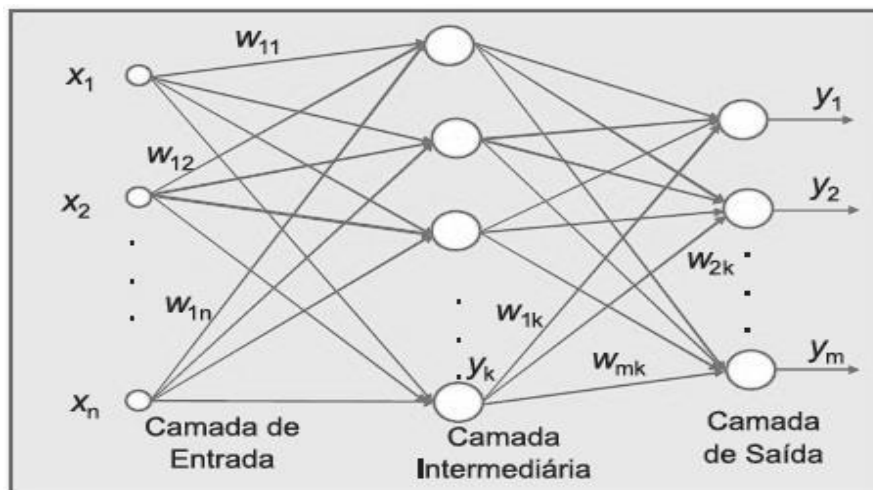
$$f(net_i) = \frac{e^{net_i} - e^{-net_i}}{e^{net_i} + e^{-net_i}} \quad (2.5)$$

O estudo destas redes foram essenciais para o entendimento e estudo das redes mais usadas recentemente que são a MLP e a RC que serão abordadas nas próximas sessões.

2.3 Multi-Layer Perceptron

Uma rede *Multi-layer Perceptron* ou simplesmente MLP é uma generalização da rede *Perceptron* [8]. A diferença entre elas é que a MLP possui além das camadas de entrada e de saída, uma camada intermediária também chamada de camada escondida.

Figura 6. Uma rede MLP



[fonte:http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0080-21072012000100008]

Segundo Valença uma rede MLP apresenta pelo menos três camadas, uma camada de entrada onde cada neurônio equivale a uma das variáveis de entrada ($x_1, x_2, x_3, \dots, x_n$) pelo menos uma camada intermediária, que é a camada que faz com que a MLP consiga resolver problemas não linearmente separáveis e uma camada de saída ($y_1, y_2, y_3, \dots, y_n$) a qual é responsável por dar a resposta do problema [7].

As principais funções de ativação de uma rede MLP são a sigmoide logística e a tangente hiperbólica, assim como a rede *Perceptron* esta rede também possui um treinamento supervisionado. O algoritmo mais utilizado para o treinamento desta rede é o *Backpropagation*. Este algoritmo foi o mesmo utilizado por Widrow & Hoff

para treinar a rede *Adaline*. *Backpropagation* foi desenvolvido por Paul Werbos em 1974, e é formado por duas fases:

A primeira é a fase *forward* que é caracterizado pela propagação do sinal em sentido progressivo, ou seja, da camada de entrada para a camada de saída. Ao se chegar na camada de saída o erro é calculado a partir da resposta obtida e a resposta esperada. Este erro é o erro médio quadrático (EMQ) dado por:

$$e_i(n) = \frac{1}{2}(d_i - y_i)^2 \quad (2.6)$$

Sendo d_i o valor desejado e y_i o valor estimado pela rede, este erro é calculado para cada neurônio da camada de saída e para cada conjunto de entradas apresentado a rede.

A segunda fase é chamada de *backward* é a fase onde acontece a retropropagação dos erros para ajustar os pesos, ou seja, os pesos serão alterados tomando em consideração os erros encontrados na fase anterior, com o intuito de minimizar o erro no próximo exemplo que a rede precisará interpretar. O reajuste dos pesos é calculado pela seguinte fórmula:

$$w_{ij}^m(novo) = w_{ij}^m(antigo) + \alpha \delta_i^m f^{m-1}(net_i^{m-1}) \quad (2.7)$$

Onde $w_{ij}^m(novo)$ é o valor do novo peso para a sinapse ij na camada m , $w_{ij}^m(antigo)$ é o valor do peso antigo para a sinapse ij na camada m , α é a taxa de aprendizagem, $f^{m-1}(net_i^{m-1})$ foram os sinais emitidos da camada anterior e por fim δ_i^m representa a sensibilidade na camada de saída que é dada por:

$$\delta_i^m = f^{m'}(net_i^m) \cdot e_i(n) \quad (2.8)$$

Onde $f^{m'}(net_i^m)$ é a derivada da função de ativação da camada de saída e $e_i(n)$ o erro. Para as camadas entre a camada de saída e entrada é usada a seguinte fórmula:

$$\delta_i^{m-1} = f^{m-1'}(net_i^{m-1}) \sum_{j=1}^n w_{ij}^m \delta_j^m \quad (2.9)$$

Onde $f^{m-1'}(net_i^{m-1})$ é a derivada da função de ativação da camada anterior “ $m-1$ ”, n é o número de neurônios, w_{ij}^m são os pesos para a sinapse ij na camada m e δ_j^m o

valor da sensibilidade da camada imediatamente a frente da camada atual. Assim pode ser verificado que a sensibilidade é calculada recursivamente da última camada para a primeira.

O critério de parada para treinamento da rede nesta monografia foi a validação cruzada, este tipo de critério é muito utilizado e evita os principais problemas no treinamento da base de dados.

Com a utilização da validação cruzada evitaremos o *overfitting*, ou seja, o treinamento excessivo da rede. Uma rede que é treinada mais do que deve começa a decorar padrões, perdendo assim sua capacidade de generalização. Ao mesmo tempo a rede pode não ser treinada o suficiente para gerar bons resultados, também perdendo sua capacidade de generalização. Para a determinação do critério de parada com a validação cruzada citado por Valença [7], geralmente os dados são divididos em três subconjuntos independentes:

- 1) 50% para treinamento;
- 2) 25% para validação cruzada;
- 3) 25% para testes.

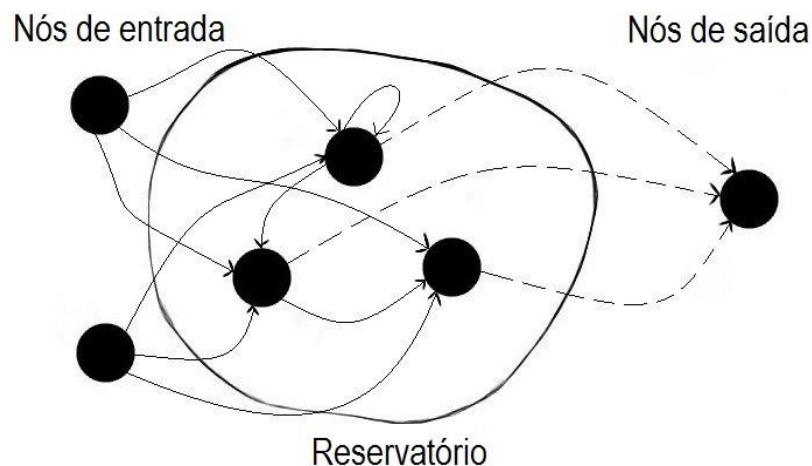
No treinamento da rede, os pesos serão ajustados, na validação cruzada é onde saberemos o número de ciclos que devemos parar de treinar, pois é calculada a função erro que ao atingir o seu mínimo deve-se interromper o treinamento. E por último o conjunto de testes que serve para verificar o desempenho do modelo.

2.4 Reservoir Computing

Redes neurais recorrentes são caracterizadas por possuírem um *feedback loop*, ou seja uma realimentação. Por possuírem este comportamento elas são capazes de conter um tipo de memória que combina valores presentes e passados, por isso é um tipo de rede muito difícil de treinar, porém também é uma ferramenta computacional muito poderosa, principalmente na resolução de problemas temporais [13].

Em 2009 surgiu um novo paradigma de treinamento desse tipo de rede chamado *Reservoir Computing* (RC) criado por David Verstraeten [6]. A rede RC consiste de uma camada de entrada (nós de entrada), um camada de neurônios recorrentes chamada de reservatório o qual é gerado aleatoriamente, e também por uma camada de saída (nós de saída) que são treinadas usando a técnica *one shot* (Verstraeten,2009).

Figura 7. Arquitetura de uma rede recorrente RC.



Na figura acima é representada a arquitetura de uma rede recorrente RC onde conexões com valores aleatórios estão indicadas pelas linhas sólidas, e as linhas pontilhadas indicam as ligações que passarão por treinamento.

Pela figura nota-se que todos os neurônios da camada de entrada estão conectados com os neurônios do reservatório e que todos os neurônios do reservatório estão ligados a camada de saída, mas nem todos os neurônios do reservatório estão ligados entre si. Por isso é determinado uma taxa de conectividade entre os neurônios do reservatório, quanto maior a taxa, maior a

quantidade de neurônios ligados entre si, este valor ainda é definido empiricamente, pois não existe estudos para comprovar que uma taxa maior ou menor ajudasse na previsão de problemas temporais.

Para construir esta rede, inicialmente é determinada a quantidade de neurônios da camada de entrada que, como na MLP, vai ser igual a quantidade de variáveis de entrada da rede mais um, que é o valor do limiar. Logo em seguida, deve-se determinar a quantidade de neurônios no reservatório que é escolhido empiricamente.

Após esses passos, os pesos da rede devem ser determinados, os quais serão inicializados de forma aleatória no intervalo de $[-1,1]$. Este valores também são determinados empiricamente e os estados dos neurônios do reservatório são inicializados com valor igual a zero. Porém, para que este estado inicial não influencie no treinamento da rede, uma técnica chamada de *warm up* é indicada [6]. *Warm up* ou esquentar é responsável por atualizar os estados dos neurônios no reservatório e deve ser executada por um número de ciclos necessários para que os estados iniciais iguais a 0 sejam esquecidos pela rede.

Assim como na MLP, o critério de parada indicado foi o *cross validation* (Valença, 2012), onde antes de enviar qualquer valor de entrada para a rede, os dados passam por um pré-processamento. O pré-processamento a base de dados é dividida em três subconjuntos: o conjunto de treinamento, validação cruzada e de testes. Finalizado todo este processo, o treinamento da rede pode ser inicializado. O primeiro passo é a atualização dos neurônios do reservatório a partir do conjunto de treinamento. É importante salientar que neste momento os pesos entre a camada de entrada e do reservatório e os pesos entre os neurônios do reservatório são gerados aleatoriamente, ou seja, não são atualizados.

A atualização dos neurônios do reservatório é feita segundo a equação abaixo:

$$x(t + 1) = f(W_{in} u(t) + W_{res} x(t)) \quad (2.10)$$

Onde t representa o tempo, $W_{in} u(t)$, a matriz que armazena o produto dos valores das variáveis de entrada pelos valores dos pesos que conectam a camada de entrada ao reservatório no instante t e $W_{res} x(t)$, a matriz de pesos do reservatório no mesmo instante t e por final $x(t + 1)$ representa o estado de

um neurônio no reservatório no instante seguinte a t , o qual é calculado pela função de ativação aplicada na soma de $W_{in} u(t)$ e $W_{res} x(t)$.

Os únicos pesos que são ajustados nesta fase são os pesos entre o reservatório e a camada de saída, para ajustá-los existem várias técnicas, uma delas é a pseudo-inversa de Moore-Penrose que é um tipo de técnica *one shot*. Esta técnica geralmente é usada porque o determinante da matriz (formada por esses pesos) pode se aproximar a zero. Um outro ponto a salientar é que esta matriz de pesos dificilmente é uma matriz quadrada, logo para contornar este problema, zeros devem ser adicionados até que a matriz se torne quadrada.

O objetivo do uso da pseudo-inversa de Moore-Penrose é converter este sistema não linear por uma aproximação linear, ou seja, solucionar problemas a partir da equação abaixo:

$$A \cdot x = B \tag{2.11}$$

Onde A para o nosso problema é a matriz de saída do reservatório e B a matriz de saídas desejadas. O x são os pesos entre o reservatório e a camada de saída. Neste caso, procura-se os valores x tais que a norma de $A \cdot x - B$ seja a menor possível (mínimos quadrados) e que representem, uma solução aproximada do sistema $A \cdot x = B$.

Capítulo 3

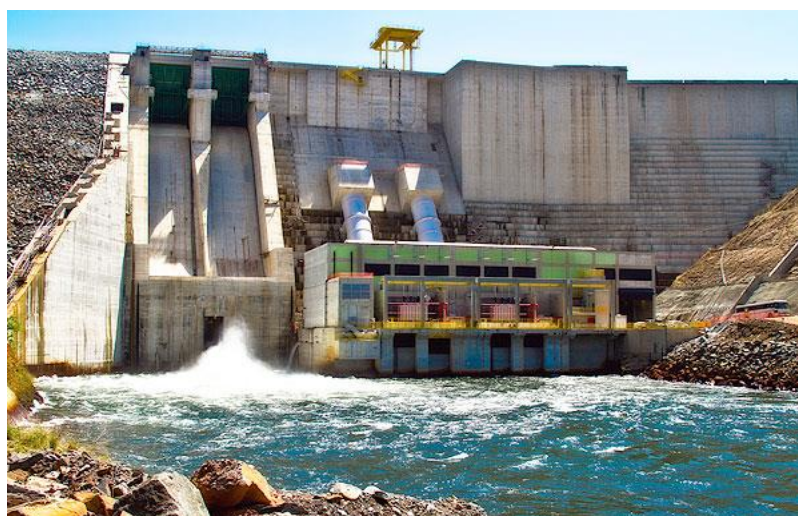
Metodologia

Neste capítulo será abordado toda a metodologia empregada para a solução do problema descrito. Na seção 3.1 as bases de dados utilizadas estão descritas, na seção 3.2 expõe todo pré-processamento dos dados antes de serem apresentados as redes neurais. Na seção 3.3 e 3.4 descrevem as entradas para a rede MLP e RC respectivamente e na seção 3.5 aborda todos os testes estatísticos utilizados para validar os resultados.

3.1 Base de dados

Foram utilizadas duas bases de dados. A primeira base de dados é da Usina Hidroelétrica Serra do Facão, a qual foi construída no rio São Marcos, no estado de Goiás. Inaugurada em 2010 o seu nível máximo é de 756 m acima do nível do mar com a capacidade de gerar 210 megawatts que é o bastante para atender mais de 1 milhão de habitantes [15]. Dados diários da vazão e da chuva da Usina Serra do Facão foram coletados a partir de janeiro de 1995 até dezembro de 2001 foram utilizados.

Figura 8 . Usina Hidroelétrica Serra do Facão



[Fonte: http://www.furnas.com.br/hotsites/sistemafurnas/usina_hidr_serrafacao.asp]

A segunda base é da Usina Hidroelétrica Corumbá IV, que também é localizada no estado de Goiás, só que no rio Corumbá. Esta usina foi inaugurada em 2006, seu reservatório tem um nível de 76 m de altura com a capacidade de gerar 129,6 megawatts [16]. Dados diários da vazão e da chuva desta usina foram coletados a partir de janeiro de 1995 até dezembro de 2001 foram utilizados.

Figura 9 . Usina Hidroelétrica Corumbá IV



[Fonte: <http://www.caldas.com.br/diversos/usinade.htm>]

3.2 Pré- processamento dos dados

O primeiro passo no pré processamento dos dados, é a sua normalização. A normalização é a técnica onde os dados passam a fazer parte de um mesmo intervalo, ou seja, valores muito altos agora passam a receber a mesma importância dos valores mais baixos. A tabela abaixo mostra um exemplo do processo de normalização.

Tabela 2. Normalização no intervalo [0-1]

Valores Reais	Valores normalizados
17.22	0.38
18.24	0.40
15.27	0.35
5.39	0.19
12.37	0.30
32.34	0.63
0.03	0.10
0.10	0.10
0.00	0.10
0.71	0.11
9.91	0.26
0.06	0.10
0.00	0.10

Segundo Valença [8] as variáveis da base de dados devem ser normalizadas de forma que seus valores variem entre os limites da função de ativação. Caso a função de ativação seja a sigmóide, função esta limitada entre 0 e 1, os dados deverão ser normalizados entre 0.1 e 0.9 ou 0.15 e 0.85 para evitar os extremos onde a derivada da função é aproximada de zero, caso a função de ativação seja a tangente hiperbólica, a qual é limitada entre -1 e 1, os dados deverão ser normalizados entre -0.9 e 0.9 ou -0.85 e 0.85.

O tipo de normalização mais usada para problemas de séries temporais é de transformação linear, dada por:

$$y = (b - a) \frac{(x_i - x_{min})}{(x_{max} - x_{min})} + a \quad (3.1)$$

Onde y é o valor da variável após a normalização, a e b são os limites que se quer ter para os valores da base, para este trabalho foi utilizada a função sigmóide com o limite entre 0.1 e 0.9 . O x_{min} é o menor valor encontrado na base e x_{max} é o maior valor. O x_i é um valor da base na posição i que se deseja normalizar.

Após os dados serem normalizados, uma técnica auto regressiva chamada defasagem é utilizada. A defasagem é o processo que tomando como base valores passados pode-se determinar os valores futuros.

Tabela 3. Base defasada até (t-2)

t-2	t-1	t
		0.38
	0.38	0.4
0.38	0.4	0.35
0.4	0.35	0.19
0.35	0.19	0.3
0.19	0.3	0.63
0.3	0.63	0.1
0.63	0.1	0.1
0.1	0.1	0.1
0.1	0.1	0.11
0.1	0.11	0.26
0.11	0.26	0.1
0.26	0.1	0.1
0.1	0.1	
0.1		

Na tabela acima, retrata um exemplo de uma rede neural com dois neurônios de entrada e um de saída, sendo (t-2) e (t-1) entradas e t seria o que nós busca-se usando a rede neural , ou seja, a saída esperada.

Um das carências do processo de defasagem é que nem todas as variáveis passadas influenciam na previsão futura, nesse caso para a previsão da vazão. Logo é necessário o uso de alguma técnica para selecionar as variáveis passadas. A técnica para seleção usada nesta monografia é a correlação entre as variáveis. A correlação indica o quanto uma variável tem influência perante outra. Na tabela 4 ilustra a correlação da tabela 3.

Tabela 4. Correlação entre variáveis da Tabela 3.

	<i>Coluna 1</i>	<i>Coluna 2</i>	<i>Coluna 3</i>
Coluna 1	1		
Coluna 2	0.181588	1	
Coluna 3	-0.03725	0.174038	1

Verificando esta tabela de correlação podemos fazer a seguinte análise. Na linha 3, que é onde apresentada-se a correlação do valor da saída esperada pelas

vaiáveis de entrada, pode-se dizer que a correlação de uma variável por ela mesma é sempre igual a 1 (coluna 3, coluna 3), e este é o maior valor que a correlação pode chegar, ou seja, é muito forte.

A correlação entre a saída e o segundo neurônio (coluna 3, coluna 2) já é menor que 1, mas não é muito forte, para esta monografia apenas valores maiores que 0.3 foram considerados. A correlação entre a saída e o primeiro neurônio (coluna 3, coluna 1) é negativa, isso não significa que ele é descartado mas sim que é de grandeza inversa a da saída, e este valor pode chegar a -1. Quanto mais próximo de -1, mais forte inversamente é a correlação. Valor igual a 0 significa que as variáveis não possuem nenhuma relação entre si.

Com as variáveis de entrada selecionadas através da correlação, pode-se apresenta-las as redes neurais e verificar o resultado obtido.

3.3 Apresentando as variáveis de entrada a MLP.

A MLP assim como as outras redes neurais é composta por um conjunto de variáveis, o ajuste dessas variáveis é chamada de configuração da rede, cada configuração pode trazer resultados similares entre si ou não.

O número de neurônios de entrada da rede depende da base de dados, o número de entradas é obtida através da correlação das variáveis, descrita na seção anterior (neste caso a base possui dados de chuva e vazão). A saída da rede corresponde ao valor da previsão futura da vazão do rio em 12 dias.

A correlação na base da Serra do Facão nos deu 24 entradas entre as 34 entradas possíveis, as correlações menores que 0.3 fizeram com que 10 possíveis entradas fossem descartadas. Já a base de Corumbá IV, de entrada foram 21 variáveis também das 34 possíveis. Também foi tomada o mesmo métodos de seleção, as variáveis com correlação maior que 0.3 permaneceram.

Esta rede também depende de outros parâmetros, a quantidade de neurônios na camada escondida, a função de ativação, o critério de parada de treinamento, a taxa de aprendizagem, o momento e o número máximo de ciclos.

O algoritmo de treinamento utilizado foi o *Backpropagation* descrito na seção 2.3, a função de ativação foi a sigmoide logística descrita pela equação (2.4).

$$f(net_i) = \frac{1}{1+e^{-net_i}} \quad (2.4)$$

Por fim na camada de saída, a função de ativação escolhida foi a linear representada na equação (2.3)

$$f(net_i) = net_i \quad (2.3)$$

3.3.1 Configurações utilizadas na rede MLP

A rede MLP esta implementada em java, e foi implemetada por alunos da Escola Politécnica de Pernambuco no ambiente de desenvolvimento eclipse, a partir das pesquisas na aplicação de redes neurais na previsão da vazão [10], nesta monografia um cenário foi escolhido, onde foram adotados os seguintes valores:

- 6 neurônios na camada intermediária;
- 30 rodadas ou simulações ;
- Taxa de aprendizagem foi de 0.85;
- Momento de 0.25;
- O número de ciclos igual a 600.

Com 30 simulações já pode-se considerar que a média dos resultados pertencem a uma distribuição normal, e que apesar do número de ciclos ser considerado alto, se a rede chegar em um ponto de convergência, ela vai parar e guardar os pesos ótimos. O método de parada utilizado é o de validação cruzada descrito descrito na seção 2.3.

Como a base utilizada não possui dados faltantes e nem valores iguais a zero, a métrica de avaliação das técnicas foi o erro percentual médio absoluto (EPMA), o qual é calculado após cada experimento. A utilização deste tipo de erro o EPMA pode ser expresso na seguinte equação.

$$EPMA = \sum_{i=1}^n \frac{|d_i - y_i| * 100}{n} \quad (3.2)$$

Onde, d_i representa a saída desejada para a previsão i . y_i representa a saída prevista na previsão i e n é a quantidade de rodadas, ou seja, a quantidade de previsões realizadas. Com todos os 30 erros percentuais médios dos 30 ciclos, pode-se fazer o comparativo com a outra arquitetura utilizada que é a Reservoir Computing usando os testes estatísticos descrito na seção 3.5.

3.4 Apresentando as variáveis de entrada a RC.

A RC assim como a MLP depende de um conjunto de variáveis, estas são as que caracterizam este tipo de rede, o ajuste delas também é chamada de configuração da rede e que também podem trazer resultados similares entre si ou não.

O número de neurônios de entrada da rede foram os mesmos apresentados a rede MLP. Para a base da Serra do Facão foram 24 entradas e para a base Corumbá IV foram 21.

Esta rede também depende de outros parâmetros, são eles:

- Quantidade de neurônios no reservatório;
- Taxa de conectividade;
- Quantidades de ciclos no *Warm up*;
- Função de ativação no Reservatório;
- Função de ativação na Camada de Saída;
- Critério de parada de treinamento.

Com as variáveis de entrada definidas, pode-se começar o treinamento do modelo. Dentre os métodos de treinamento deste tipo de rede, a utilizada nesta monografia foi a pseudo-inversa de Moore-Penrose, descrita na seção 2.4.

3.4.1 Configurações utilizadas nas redes RC

A rede RC esta implementada em java, também implementada por alunos da Escola Politécnica de Pernambuco no ambiente de desenvolvimento eclipse, a maioria dos parâmetros utilizados nas pesquisas foram escolhidos empiricamente,

nesta monografia foram adotados cinco cenários de teste variando a quantidade de neurônios no reservatório.

No primeiro cenário foram adotados os seguintes valores:

- Quantidade de neurônios no reservatório = 25;
- Taxa de conectividade = [0.3,0.4];
- Quantidades de ciclos no *Warm up* = 10;
- Função de ativação no reservatório = sigmoide;
- Função de ativação na camada de saída = linear;
- Critério de parada de treinamento = *cross validation*;
- Número de experimentos = 30;
- Número de saída = 12 (previsão de 12 dias a frente da vazão).

De forma geral, utiliza-se um intervalo de [0.3,0.4] como taxa conectividade pois considerando-se os testes realizados em outras pesquisas [13] [14], este intervalo traz menor processamento e bons resultado na previsão de vazão. A quantidade de ciclos para o *warm up* foi considerado 10 ciclos pois através de testes nos trabalhos [13][14] mostram que com 10 ciclos, os zeros inicializantes da rede já não faziam nenhuma interferência no treinamento da rede.

A quantidade de experimentos é a mesma que na MLP, para assumir que os resultados fazem parte de uma distribuição normal e 12 saídas pois o intuito é saber a previsão da vazão em doze dias a frente.

Nos cenários 2, 3, 4 e 5, foram utilizados 35, 50, 75 e 100 neurônios respectivamente no reservatório. Cada cenário assim como na MLP passaram por 30 experimentos, os quais geraram 30 EPMA's. Estes resultados serão comparados pelos testes estatísticos descritos na seção a seguir.

3.5 Testes estatísticos

Após o término dos 30 experimentos onde obtive-se 30 EPMA's, como métrica de avaliação destes resultados, foi-se utilizado testes estatísticos. Estes testes são para validar os resultados encontrados. Nesta monografia usaremos os seguintes: Shapiro -Wilk, teste F, t de Student e Wilcoxon.

Os testes foram realizados com o *Software R* [12] considerando o nível de significância igual a 5 %, que é a porcentagem comumente utilizada na engenharia [9].

O teste de Shapiro-Wilk propõe verificar se as amostras da base são normalmente distribuídas. O teste F é onde se verifica se as amostras possuem variâncias estatisticamente iguais. Se esses dois testes passarem, é possível aplicar os chamados testes paramétricos, neste caso utilizaremos o t de Student, caso contrário é necessário o uso de testes não paramétricos como o Wilcoxon.

3.5.1 Teste Shapiro-Wilk

O teste Shapiro-Wilk é usado para verificar se os dados fazem parte de uma população normalmente distribuídas. Para realizar o teste deve-se considerar os seguintes passos:

1. Determinar duas possibilidades:
 - a. H0: Amostra é normalmente distribuída.
 - b. H1: Amostra não é normalmente distribuída.
2. Caso o nível de significância, chamado de *p value*, seja menor que 0.05 a amostra satisfaz a H1, caso contrário satisfaz a H0.

Este teste é o primeira hipótese para aplicar um teste paramétrico, sabendo-se que a distribuição seja normal.

3.5.2 Teste F

O teste F tem como finalidade verificar se a variância de duas amostras são estatisticamente iguais. Para realizar o teste deve-se considerar os seguintes passos:

1. Determinar duas possibilidades
 - a. H0: Amostra possui variâncias estatisticamente iguais.
 - b. H1: Amostra não possui variâncias estatisticamente iguais.
2. Caso o *p value* seja menor que 0.05 a amostra satisfaz a H1, caso contrário satisfaz a H0.

Este teste é o segundo princípio para aplicar um teste paramétrico. Assumindo que a variância da distribuição são estatisticamente iguais.

3.5.3 Teste t de Student

O teste t de Student é um teste paramétrico, ou seja, para aplicá-lo temos que garantir que as duas hipóteses anteriores sejam respeitadas: as amostras devem ser normalmente distribuídas e possuir variância estatisticamente iguais. Este teste compara a média entre duas amostras e para realizar este teste deve-se seguir os seguintes passos:

1. Determinar duas possibilidades
 - a. H0: As médias entre as amostras são iguais.
 - b. H1: As médias entre as amostras não são iguais.
2. Caso o p *value* seja menor que 0.05 a amostra satisfaz a H1, caso contrário satisfaz a H0.

Se uma das hipóteses anteriores não são satisfeitas deve-se aplicar um testes não paramétricos.

3.5.4 Test Wilcoxon

O teste Wilcoxon é um teste não paramétrico, ou seja, não faz nenhuma hipótese em relação a distribuição das probabilidades dos dados. Este teste tem o objetivo de verificar que mesmo sendo de amostras independentes, elas fazem parte de uma população onde as medianas são iguais. Para realizar o teste deve-se considerar os seguintes passos:

1. Determinar duas possibilidades
 - a. H0: as medianas entre as amostras são iguais.
 - b. H1: as medianas entre as amostras não são iguais.
2. Caso o p *value* seja menor que 0.05 a amostra satisfaz a H1, caso contrário satisfaz a H0.

Capítulo 4

Resultados

Este capítulo tem como objetivo apresentar todos os resultados obtidos aplicando as redes neurais MLP e RC utilizando dados de chuva e vazão na previsão da vazão diária. Seguindo a metodologia descrita no capítulo anterior, ilustrando todos os cenários descritos para as duas bases de dados: Serra do Facão e Corumbá IV.

A seção 4.1 apresenta os experimentos realizados usando a base de dados da hidroelétrica Serra do Facão, aplicando o cenário com a rede MLP e os cinco cenários com a rede RC. Logo após esta seção, a seção 4.2 demonstra todos estes experimentos usando a base dados da hidroelétrica Corumbá IV. Por fim na seção 4.3 faz o comparativos entre as técnicas utilizadas e indica qual delas possui a melhor eficiência.

4.1 Experimentos com a base Serra do Facão

Como explicado anteriormente, todos os experimentos foram divididos em 6 cenários, o primeiro cenário utilizando a rede MLP pois segundo pesquisas a configuração utilizada nesta rede é a melhor na previsão da vazão e os outros cinco cenários foram utilizando a rede RC pois alguns dados eram escolhidos empiricamente, logo havia necessidade de buscar cenários que obtivesse o melhor resultado utilizando esta rede.

Em cada cenário foram realizadas 30 execuções, em cada execução era obtido o erro percentual médio absoluto.

4.1.1 Hidroelétrica Serra do Facão: Estudo de caso com a rede MLP

Para a base de dados da Serra do Facão foram apresentadas a rede 24 entradas das 34 possíveis, sendo estas escolhidas pela correção dos dados. Das variáveis de entrada foram consideradas os 12 últimos dias da vazão e últimos 10 dias de chuva, da chuva seguinte ou chuva futura apenas os 2 primeiros dias

foram introduzidos, os outros 8 dias de chuvas futura foram descartadas já que a correlação era menor que 0.3.

Para a configuração descrita na seção 3.3.1, na previsão de 12 dias seguintes de vazão, a média entre os 30 EPMA's foi gerada obtendo uma porcentagem de erro igual a 21,51%. Este resultado será comparado com o melhor cenário da rede RC.

4.1.2 Hidroelétrica Serra do facão: Estudo de caso com a rede RC

Na RC também foram apresentadas as mesmas 24 entradas que foram introduzidas na MLP. Alguns valores nesta rede ainda são escolhido empiricamente, a quantidade de neurônios no reservatório em uma rede RC é um valor que precisa de mais estudos para ser determinado. Para se determinar este valor foram criados cinco cenários que variassem a quantidade de neurônios e que trouxesse a melhor previsão da vazão em 12 dias seguintes.

Assim como na rede MLP, para cada cenário foram obtidos 30 EPMA's, cada cenário será comparado entre si utilizando as técnicas descritas na seção 3.5. O melhor cenário na RC também foi comparado com o resultado da MLP através dos testes estatísticos.

O primeiro cenário apresenta 25 neurônios no reservatório, o segundo, terceiro, quarto e quinto apresentam 35, 50, 75 e 100 respectivamente. A tabela 5 exibe a média do erro percentual médio absoluto para cada cenário.

Tabela 5. Média das EPMA's para cada cenário na Serra do Facão

Cenários	Quantidade de neurônios	EPMA	Desvio padrão
1	25	15.36%	0.5720558
2	35	14.91%	0.241211
3	50	14.93%	0.2903761
4	75	15.45%	0.2842297
5	100	15.94%	0.2799488

Para analisar qual cenário seria o mais eficiente foram feitas comparações utilizando os testes estatísticos. Nas seções seguintes são apresentadas todas estas aplicações e resultados.

4.1.2.1 Comparação do Cenário 1 com o Cenário 2

Aplicando o primeiro teste Shapiro-Wilk no Cenário 1, obtivemos o valor do *p value* menor que 0.05, mostrando que o conjunto de dados satisfaz H1 que são amostras não normalmente distribuída.

Aplicando este mesmo teste no Cenário 2, obtivemos o valor do *p value* maior que 0.05, mostrando que o conjunto de dados satisfaz a H0 onde as amostras são normalmente distribuídas.

Como este primeiro teste trouxe um valor de *p value* < 0.05, não podemos aplicar um teste paramétrico, deve-se aplicar um teste não paramétrico. Neste caso, aplica-se o teste Wilcoxon onde obtive-se o *p value* < 0.05.

Neste último teste verificamos que para os cenários 1 e 2, as medianas entre as amostras não são iguais, logo consideramos apenas o conjunto que possui amostras normalmente distribuídas. Podemos dizer então que o Cenário 2 tem melhor desempenho que o Cenário 1 na previsão da vazão. Sendo o Cenário 1 então descartado. A tabela 6 exibe os resultados obtidos nos testes.

Tabela 6. Resultados do *p value* para os Cenário 1 e 2 na Serra do Facão

Cenários	<i>p value</i> (Shapiro- Wilk)	<i>p-value</i> (Wilcoxon)
1	0.02016	0.0002127
2	0.6825	

4.1.2.2 Comparação do Cenário 2 com o Cenário 3

Nesta fase já sabemos que o Cenário 2 é considerada uma amostra normalmente distribuída pelo resultado do teste Shapiro-Wilk. Aplicando este mesmo teste ao Cenário 3, obtivemos um valor de *p value* maior que 0.05.

Quando as duas amostras são considerados normalmente distribuídas aplicamos testes paramétricos. O primeiro teste utilizado foi o teste F no qual o resultado obtido para o *p value* foi maior que 0.05, mostrando que as duas amostras satisfaz H0, ou seja, as amostra possuem variâncias estatisticamente iguais.

Com as amostras que satisfazem H_0 no teste F e que são normalmente distribuídas, pode-se aplicar o teste t de Student, onde o valor obtido para o *p value* foi maior que 0.05. Sendo assim os cenários 2 e 3 tem amostras onde as médias são estatisticamente iguais. A tabela abaixo exhibe os valores encontrados nesta comparação dos cenários 2 e 3.

Tabela 7. Resultados do *p value* para os Cenário 2 e 3 na Serra do Facão

Cenários	<i>p value</i> (Shapiro- Wilk)	Teste F	Teste t de Student
2	0.6825	0.3233	0.8162
3	0.2572		

Para este cenário verifica-se as médias dos dois cenários para poder determinar que cenário que tem melhor desempenho. Vericando os resultados com o *Software R*, encontramos os seguintes valores.

```
> mean(rc35)
[1] 14.91394
> mean(rc50)
[1] 14.93003
```

Onde *mean(rc35)* é a média para o Cenário 2, e *mean(rc50)* a média para o Cenário 3, pode-se verificar que a média do Cenário 2 é menor que do Cenário 3. Logo considera-se que o Cenário 2 é o vencedor, ou seja, continua-se com o cenário que apresenta 35 neurônios no reservatório.

4.1.2.3 Comparação do Cenário 2 com o Cenário 4

Sabe-se que o Cenário 2 passou no teste de Shapiro-wilk, aplicamos este teste para o Cenário 4, e verifica-se que o *p value* encontrado é maior que 0.05. logo pode ser aplicado o teste F.

No teste F, obtive-se o *p value* maior que 0.05. Os dois testes acima foram satisfeitos, pode-se então aplicar o teste t de Studente, no qual foi obtido valor para o *p value* < 0.05, ou seja, as amostras não possuem médias estatisticamente iguais. Na tabela abaixo é exibido os resultados encontrados.

Tabela 8. Resultados do *p value* para os cenários 2 e 4 na Serra do Facão

Cenários	<i>p value</i> (Shapiro- Wilk)	Teste F	Teste t de Student
2	0.6825	0.382	8.542e-11
4	0.89		

A definição do cenário com melhor desempenho então é definida pela menor média encontrada. Verificando-se os resultados com o *Software R*, foi encontrado os seguintes valores.

```
> mean(rc35)
[1] 14.91394
> mean(rc75)
[1] 15.45557
```

Onde *mean(rc35)* é a média para o Cenário 2, e *mean(rc75)* a média para o Cenário 4, pode-se verificar que a média do Cenário 2 é menor que do Cenário 4. Logo considera-se que o Cenário 2 é o vencedor, ou seja, continua-se com o cenário que apresenta 35 neurônios no reservatório.

4.1.2.4 Comparação do Cenário 2 com o Cenário 5

Já sabe-se que o Cenário 2 continua a ser o que tem melhor desempenho entre os cenários comparados e que passou no teste de Shapiro- Wilk, aplica-se este teste para o Cenário 5, e verifica-se que o *p value* encontrado é maior que 0.05. Portanto, podere-se aplicar o teste F.

Para o teste F, o *p value* encontrado também foi maior que 0.05. Se os dois testes acima foram satisfeitos, pode-se então aplicar o teste t de Student, no qual o *p value* foi menor que 0.05, ou seja, as amostras do cenários 2 e 5 não possuem médias estatisticamente iguais. Na tabela abaixo é exibido os resultados encontrados.

Tabela 9. Resultados do *p value* para os cenários 2 e 5 na Serra do Facão

Cenários	<i>p value</i> (Shapiro- Wilk)	Teste F	Teste t de Student
2	0.6825	0.4274	2.2e-16
4	0.1743		

Assim como na comparação entre os cenários 2 e 4, a definição do cenário com melhor desempenho é estabelecido pela média dos EPMA's de cada cenário. Verificando os resultados no *Software R*, encontra-se os seguintes valores.

```
> mean(rc35)
[1] 14.91394
> mean(rc100)
[1] 15.94132
```

Onde $mean(rc35)$ é a média para o Cenário 2, e $mean(rc100)$ a média para o Cenário 5, pode-se verificar que a média do Cenário 2 é menor que do Cenário 5. Logo considera-se que o Cenário 2 é o vencedor, ou seja, o cenário que possui 35 neurônios no reservatório tem melhor desempenho na determinação da vazão diária para a hidroelétrica Serra do Facão. Além disso, ter 35 neurônios no reservatório é preferível pois diminui tempo de execução.

4.2 Experimentos Base Corumbá IV

Assim como na base da Hidroelétrica Serra do Facão, todos os experimentos foram divididos em 6 cenários, um cenário para a rede MLP pois segundo pesquisas anteriores a configuração utilizada nesta rede já é a melhor na previsão da vazão e os outros cinco cenários para a rede RC pois alguns parâmetros foram escolhidos de forma empírica e houve a necessidade de testá-los.

Em cada cenário também foram realizadas 30 execuções e em cada execução o erro percentual médio absoluto foi calculado.

4.2.1 Hidroelétrica Corumbá IV: Estudo de caso com a rede MLP

Para a base Corumbá IV foram apresentada a rede 21 entradas das 34 possíveis, sendo estas escolhidas através correção das amostras de dados. Das variáveis de entrada foram consideradas os 12 últimos dias da vazão e últimos 9 dias de chuva passada, nenhum dia da chuva seguinte ou chuva futura foram considerados pois a correlação era menor que 0.3.

Para a configuração descrita na seção 3.3.1, a média entre os 30 EPMA's calculada foi de 21.25%. Este valor será comparado com o cenário mais eficiente da rede RC.

4.2.2 Hidroelétrica Corumbá IV: Estudo de caso com a rede RC

Assim como na MLP, para a rede RC também foram apresentadas as mesmas 21 entradas. Sabendo-se que ainda não existem muitas pesquisas na determinação da quantidade de neurônios no reservatório, se fez necessários criar cenários para testes. Para descobrir a quantidade de neurônios que desse a melhor previsão dos 12 dias seguintes de vazão para esta base foram criados os cinco cenários como foi feito na base da Hidroelétrica Serra do Facão.

Para cada cenário com seus 30 EPMA's, se fez a comparação entre eles utilizando as técnicas estatísticas descritas na seção 3.5. O cenário mais eficiente utilizando a rede RC é o que foi comparado com o resultado da MLP.

Para a base Corumbá IV, no primeiro, segundo, terceiro, quarto e quinto cenários foram definidos 25, 35, 50, 75 e 100 neurônios respectivamente. A tabela abaixo exhibe para cada cenário a média do erro percentual médio absoluto.

Tabela 10. Média das EPMA's para cada cenário na Corumbá IV

Cenários	Quantidade de neurônios	EPMA	Desvio padrão
1	25	20.34%	0.3951667
2	35	20.27%	0.335409
3	50	20.66%	0.2969692
4	75	21.37%	0.4723669
5	100	22.10%	0.4466163

Para analisar estes cenários, comparações utilizando os testes estatísticos foram realizados. As aplicações dos testes e resultados serão exibidos nas seções seguintes.

4.2.2.1 Comparação do Cenário 1 com o Cenário 2

Para a comparação dos cenários, o primeiro teste é o Shapiro- Wilk, o qual mostra se o conjunto apresentado é normalmente distribuído ou não. Nos cenários 1 e 2, obtive-se o valor do *p value* > 0.05, mostrando que com 25 neurônios ou 35 as amostras pertencem a uma população normalmente distribuída.

Como este primeiro teste apresentou *p value* > 0.05, pode-se aplicar o teste F. Para ambos cenários utilizando o teste F o *p value* foi maior que 0.05, mostrando que as duas amostras possuem variâncias estatisticamente iguais.

Sabendo que as amostras são normalmente distribuídas e que possuem variâncias estatisticamente iguais, aplica-se o teste t de Student. O resultado na comparação dos cenários 1 e 2 utilizando este teste foi o *p value* maior que 0.05, indicando que as médias entre estes valores são estatisticamente iguais.

A tabela abaixo exhibe os valores encontrados nesta comparação dos cenários 1 e 2.

Tabela 11. Resultados do *p value* para os cenários 1 e 2 na Corumbá IV

Cenários	<i>p value</i> (Shapiro- Wilk)	Teste F	Teste t de Student
1	0.6389	0.3824	0.4626
2	0.2543		

Comparando os cenários 1 e 2, verifica-se que as suas médias para poder determinar o melhor cenário. Vericando os resultados no *Software R*, encontra-se os seguintes valores.

```
> mean(rc25)
[1] 20.3436
> mean(r35)
[1] 20.27361
```

Onde *mean(rc25)* é a média para o Cenário 1, e *mean(rc35)* a média para o Cenário 2, podemos verificar que a média do Cenário 2 é menor que do Cenário 1. Mas são tão próximas entre si, que neste caso a determinação da quantidade de neurônios no reservatório dará-se pelo cenário que trouxe menor trabalho de processamento. Quanto menos neurônio no reservatório, menor será o processamento para fazer a previsão, logo o vencedor nesta comparação é o Cenário 1.

4.2.2.2 Comparação do Cenário 1 com o Cenário 3

Nesta segunda comparação já sabe-se que o Cenário 1 pertence a uma população normalmente distribuída através do teste Shapiro-Wilk. Aplicando este mesmo teste no Cenário 3, obtive-se o *p value* maior que 0.05.

Quando as duas amostras são normalmente distribuídas pode-se aplicar testes paramétricos. O primeiro teste foi o teste F, obtendo como resultado um p

$value > 0.05$, mostrando que as duas amostras possuem variâncias estatisticamente iguais.

Passando por estes dois testes, pode-se aplicar o teste t de Student, onde o valor obtido para o $p\ value$ foi menor que 0.05. A tabela abaixo indica o valor do $p\ value$ para os cenários 2 e 3.

Tabela 12. Resultados do $p\ value$ para os cenários 1 e 3 na Corumbá IV

Cenários	$p\ value$ (Shapiro- Wilk)	Teste F	Teste t de Student
1	0.6389	0.1297	0.0007406
3	0.6935		

Para o caso onde o t de Student teste possui valor $p\ value < 0.05$, é verificada a menor média nos cenários 1 e 3. As médias foram calculadas com o auxílio do *software* R listadas a seguir.

```
> mean(rc25)
[1] 20.3436
> mean(rc50)
[1] 20.66657
```

Onde $mean(rc25)$ é a média para o Cenário 1, e $mean(rc50)$ a média para o Cenário 3, pode-se verificar que a menor média ocorre no Cenário 1. Para esta comparação o Cenário 1 é o escolhido.

4.2.2.3 Comparação do Cenário 1 com o Cenário 4

Como o Cenário 1 é normalmente distribuído, podemos já aplicar o teste Shapiro- Wilk no Cenário 4 e como resultado o $p\ value$ foi maior que 0.05. Pode-se aplicar o teste F.

Aplicando-se o teste F, obteve-se o $p\ value$ também maior que 0.05. Como estes dois testes acima foram satisfeitos, pode-se aplicar o teste t de Student, no qual o valor calculado para o $p\ value$ foi menor que 0.05. Na tabela a seguir é ilustrado os resultados encontrados.

Tabela 13. Resultados do $p\ value$ para os cenários 1 e 4 na Corumbá IV

Cenários	$p\ value$ (Shapiro- Wilk)	Teste F	Teste t de Student
1	0.6389	0.342	4.906e-13
4	0.8527		

As amostras não possuem médias estatisticamente iguais. A escolha do melhor cenário é diferenciada pela menor média encontrada. Vericando os resultados no *Software R*, encontra-se os seguintes valores.

```
> mean(rc35)
[1] 20.3436
> mean(rc75)
[1] 21.39378
```

Onde *mean(rc35)* é a média para o Cenário 1, e *mean(rc75)* a média para o Cenário 4, pode-se verificar que a média do Cenário 1 é menor que do Cenário 4 e constata-se que continua-se com o cenário que apresenta 25 neurônios no reservatório.

4.2.2.4 Comparação do Cenário 1 com o Cenário 5

Até este momento o Cenário 1 é o indicado para ser utilizado na compração com a MLP, mas ainda tem-se a necessidade de aplicar os testes para o Cenário 5, aplicando Shapiro- Wilk no Cenário 5, o *p value* encontrado é maior que 0.05. Então pode-se aplicar os testes paramétricos.

Para o teste F, o *p value* encontrado também foi maior que 0.05. Satisfazendo assim os pré requisitos para a aplicação do teste t de Student, este teste foi aplicado e obteve de resultado o *p value* menor que 0.05. Na tabela abaixo é ilustrado os resultados encontrados para cada cenário.

Tabela 14. Resultados do *p value* para os cenários 1 e 5 na Corumbá IV

Cenários	<i>p value</i> (Shapiro- wilk)	Teste F	Teste t de Student
2	0.6389	0.514	2.2e-16
4	0.5313		

Para verificar o cenário com melhor desempenho, a média foi calculada utilizando o *Software R*, estas médias são listadas a seguir.

```
> mean(rc3)
[1] 20.3436
> mean(rc7)
[1] 22.11458
```

Onde *mean(rc25)* é a média para o Cenário 1, e *mean(rc100)* a média para o Cenário 5, pode-se verificar que a média do Cenário 1 é menor que do Cenário 5.

Logo os resultados da rede com 25 neurônios no reservatório é a que vai ser compara com a rede MLP.

4.3 Comparativo dos resultados: MLP x RC

Após obtenção da melhor configuração utilizando a rede *Reservoir Computing*, nesta seção este melhor valor será comparado com o resultado da rede MLP, visando confirmar qual rede que tem melhor desempenho na previsão de 12 dias de vazão em um processo chuva x vazão.

4.3.1 Hidroelétrica Serra do Facão: MLP x RC

Aplicando o teste Shapiro- Wilk para os resultados da MLP obteve-se um valor menor que 0.05 para o *p value*. Desta maneira deve-se aplicar o teste não paramétrico de Wilcoxon. Este teste foi aplicado e obteve-se novamente um valor menor que 0.05 para o *p value*. Estes valores estão ilustrados na tabela abaixo para as duas técnicas.

Tabela 15. Resultados do *p value* para MLP e RC na Serra do Facão

Técnicas	<i>p value</i> (Shapiro- Wilk)	<i>p-value</i> (Wilcoxon)
RC	0.6825	2.2e-16
MLP	1.818e-5	

A técnica que possui melhor desempenho é considerada a possui menor erro percentual. Para a MLP a média dos 30 EMPAs foi de 21.51%, já a RC foi de 14.91%. Logo a melhor rede é a RC.

4.3.2 Hidroelétrica Corumbá IV: MLP x RC

Para a amostra na MLP, o *p value* indica que os dados não podem ser considerados normalmente distribuídos, deve-se usar nesta comparação o teste não paramétrico Wilcoxon, aplicando este teste foi verificado que suas medianas são estatisticamente iguais, pois o *p value* retornado foi maior que 0.05. Este valor para as duas técnicas estão representados na tabela abaixo.

Tabela 16. Resultados do *p value* para MLP e RC na Corumbá IV

Técnicas	<i>p value</i> (Shapiro- wilk)	<i>p-value</i> (Wilcoxon)
RC	0.6389	0.1188
MLP	0.01857	

Para determinar qual técnica é a melhor, verifica-se a que possui a menor taxa de erro. Para a MLP a média dos 30 EMPAs foi de 21.25%, já a RC foi de 20.34%. Logo a rede que possui melhor desempenho é a RC.

Capítulo 5

Considerações Finais

Este capítulo tem como objetivo descrever as considerações finais na seção 5.1, assim como uma lista com trabalhos futuros na seção 5.2.

5.1 Conclusão

A previsão da vazão diária de bacias hidrográficas são realizadas continuamente para o planejamento energético do país. Diante deste fato, ao longo dos anos a preocupação com a precisão da previsão só tem aumentado. Este trabalho teve como principal objetivo fazer um estudo comparativo entre técnicas RC e MLP para prever a vazão diária nos 12 dias seguintes considerando dados de chuva e vazão. Para que fosse possível comparar estas técnicas testes estatísticos foram aplicados.

Os resultados obtidos mostram que a utilização da rede RC na previsão da vazão diária deu um ganho significativo tanto quanto aplicadas para a base da Serra do Facão quanto para a base de corumbá IV. Para a primeira, a rede RC obteve uma diferença de 6.6% e para a segunda uma diferença de quase 1%. Comprovando que a rede RC é mais precisa que a rede MLP para este estudo de caso.

5.2 Trabalhos futuros

Para trabalhos futuros, alguns aspectos podem ser utilizados.

- Utilizar algoritmos de busca em vez da correlação dos dados para determinação das variáveis de entrada das redes. PSO (Particle swarm optimization) é um dos algoritmos mais utilizados.
- Uso de outros algoritmos de treinamento de rede para a MLP como colônia de formigas.

- Realizar mais experimentos na rede RC determinando outros parâmetros para sua configuração.
- Verificar se com um maior número de ciclos as redes apresentam uma menor taxa de erro.
- Obter dados de outras bases e verificar se também haverá ganho na previsão diária da vazão utilizando a rede RC.

Referências

- [1] BROOKS, Kenneth N.; FFOLLIOTT, Peter F.; MAGNER, Joseph A. Hydrology and Management of Watersheds. 4. ed. Iowa: John Wiley & Sons Inc, 2013. 552 p. Disponível em: <<http://www.wiley.com/WileyCDA/WileyTitle/productCd-0470963050.html>>. Acesso em: 02 set. 2016.
- [2] DIXON, John A.; EASTER, K. William; HUFSCHMIDT, Maynard M. Integrated Watershed Management: An Integrated framework with studies from Asia and the Pacific. Singapura e Honolulu: Institute Of Southeast Asian Studies Com Colaboração do East Coast Center do Policy Institute, 1991. 215 p.
- [3] LOUREIRO, Igor Bittencourt. Laboatório remoto para ensino de eficiência energética em um sistema de bombeamento de água. 2009. 121 f. Dissertação (Mestrado) - Curso de Engenharia Elétrica, Instituto de Tecnologia, Universidade Federal do Pará, Belém do Pará, 2009.
- [4] PEREIRA, Salomão Almeida; LOUREIRO, Lorena Drumond; MARTINS, Fernanda Roberta Gomes da Silva Pinto; SANTOS, André Clementino Oliveira; TUBINO, Dalvio Ferrari. XXVI Encontro nacional de engenharia de produção, 2006, Fortaleza. Estudo comparativo entre modelos de previsão de demanda: ensaio em um produto classe a de uma empresa de perfumes e cosméticos. Fortaleza, 2006. 10 p.
- [5] 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 chuva-vazão mensal utilizando redes neurais artificiais. Brasília, 2015. 8 p.
- [6] D. Verstraeten, *Reservoir computing: computation with dynamical systems*, Ph. D *dissertation* 2009.
- [7] ANEEL, Atlas da energia eletrica no Brasil, 3rd ed. Brasilia.
- [8] M. Valença, Fundamento das Redes Neurais, Olinda: Livro Rápido, 2009.
- [9] M. Valença, Aplicando Redes Neurais. Olinda: Livro Rápido, 2005.

-
- [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] Ministério de Minas e Energia, Resenha Energética, 2015. 31f.
- [12] “*The R Project for Statistical computing*” [Online]. Available: <https://www.r-project.org/> Acessado em 15 de outubro de 2016.
- [13] BEZERRA, Sabrina Guimarães Tavares de Andrade. *Reservoir computing 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, 2016.
- [14] 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
- [15] “Usina Hidroelétrica Serra do Facão” [Online] available: <http://www.sefac.com.br/> Acessado em 2 de setembro de 2016.
- [16] “Usina Hidroelétrica Corumbá IV” [Online] available: <http://www.corumbaconcessoes.com.br/uhe-corumba-iv/visao-geral/> Acessado em 2 de setembro de 2016.