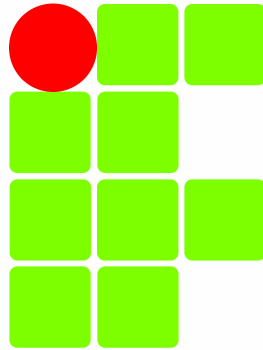


INSTITUTO FEDERAL DE GOIÁS – IFG  
CÂMPUS GOIÂNIA  
PROGRAMA DO MESTRADO EM TECNOLOGIAS DE PROCESSOS SUSTENTÁVEIS



INSTITUTO FEDERAL DE GOIÁS

# Modelagem de Oxigênio Dissolvido utilizando Redes Neurais Artificiais

Saulo Rodrigues e Silva

Goiânia  
2014

**Saulo Rodrigues e Silva**

**Modelagem de Oxigênio Dissolvido  
utilizando Redes Neurais Artificiais**

Dissertação apresentada ao Programa do Mestrado em Tecnologias de Processos Sustentáveis do Instituto Federal de Goiás – IFG, como requisito parcial para a obtenção do título de Mestre em Ciências.

Área de concentração: Modelagem Ambiental, Sistemas de Produção Limpa

Orientador: Fernando Schimidt, Dr.

Coorientador: José Luis Domingos, Dr.

Goiânia

2014

Catálogo na fonte

S586 Silva, Saulo Rodrigues e.  
Modelagem de oxigênio dissolvido utilizando redes neurais artificiais [manuscrito] / Saulo Rodrigues e Silva. -- Goiânia, 2014. 120 f. : il.

Dissertação (Mestrado) – Instituto Federal de Goiás, Câmpus Goiânia, Programa do Mestrado Profissional em Tecnologias de Processos Sustentáveis, 2014.

“Orientador: Prof. Dr. Fernando Schmidt”.

1. Redes neurais artificiais. 2. Retropropagação de erro. 3. Oxigênio dissolvido. I. Título.

CDD 026.0221

Ficha catalográfica elaborada pela Bibliotecária Maria Aparecida Rodrigues de Souza, CRB-1/1497.  
Biblioteca Atena, Câmpus Inhumas  
Instituto Federal de Educação, Ciência e Tecnologia de Goiás

**Saulo Rodrigues e Silva**

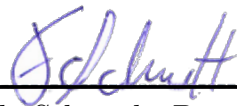
**Modelagem de Oxigênio Dissolvido  
utilizando Redes Neurais Artificiais**

Dissertação apresentada ao Programa do Mestrado em Tecnologias de Processos Sustentáveis do Instituto Federal de Goiás – IFG, como requisito parcial para a obtenção do título de Mestre em Ciências.

Área de concentração: Modelagem Ambiental, Sistemas de Produção Limpa

Goiânia, 28 de Junho de 2014

Banca Examinadora:



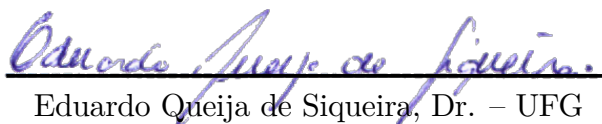
---

Fernando Schmidt, Dr. – IFG



---

Wesley Pacheco Calixto, Dr. – IFG



---

Eduardo Queija de Siqueira, Dr. – UFG

*Dedico este trabalho às minhas filhas, Laura Beatriz e Elisa Bianca: por vocês, tudo vale a pena!*

---

# Agradecimentos

À Deus, pelo respirar.

À minha família, que cedeu gentilmente muito do seu tempo comigo para a conclusão deste projeto.

Ao amigo Leonardo Garcia, pelo apoio e presença fundamentais ao trabalho.

Ao professor Fernando Schmidt, pela orientação, incentivos e amizade que construímos.

Ao professor José Luis Domingos, pelo tempo dedicado na co-orientação, revisões e valiosas colaborações no texto.

Ao professor Wesley Pacheco Calixto, pelas participações nas bancas de qualificação e defesa, que resultaram em importantes apontamentos na direção de explorar metodologias de sensibilidade de dados, além de incontáveis melhorias no trabalho e até mesmo ajuda com Latex.

Ao professor Eduardo Queija de Siqueira, pela participação na banca de defesa, que terminou em inestimáveis considerações e apontamentos para a melhoria deste trabalho.

À professora Warde Antonieta da Fonseca Zang, coordenadora deste programa de mestrado, que sempre incentivou e inspirou de maneira muito particular esta primeira turma.

À todos os professores do programa, que abraçaram o desafio de ensinar dentro de um programa multidisciplinar, acreditando no potencial dos alunos.

À colega Leila Ninon, secretária do programa de mestrado e companheira que sempre manteve a porta aberta aos alunos.

Aos desenvolvedores responsáveis pelos projetos de código aberto (*open source software*) a seguir, cujos programas foram responsáveis pelo trabalho computacional desse projeto de mestrado:

❑ Gnu Linux<sup>1</sup>: sistema operacional seguro e robusto;

❑ KDE<sup>2</sup>: meu sistema de gerenciamento de janelas;

---

<sup>1</sup> <http://www.gnu.org>

<sup>2</sup> <http://www.kde.org>

- JAVA<sup>3</sup>: uma linguagem de programação aberta, universal e poderosa;
- Neuroph<sup>4</sup>: *framework* desenvolvido em JAVA para criar e treinar redes neurais, utilizado neste trabalho;
- L<sup>A</sup>T<sub>E</sub>X<sup>5</sup> e abnT<sub>E</sub>X<sup>6</sup>: o sistema de formatação de texto utilizado nos textos e até mesmo nas apresentações;
- Inkscape<sup>7</sup>: utilizado para elaborar as imagens vetoriais;
- Gimp<sup>8</sup>: software para para correção e ajustes de imagens do tipo *bitmap*;
- Scilab<sup>9</sup>: ferramenta de matemática computacional.

Certamente existem outros nomes entre amigo(a)s, colegas, professores, pessoas que acabam nos apoiando ou inspirando de alguma forma, mesmo quando não sabem.

A todos, simplesmente obrigado.

---

<sup>3</sup> <http://www.java.com>

<sup>4</sup> <http://neuroph.sourceforge.net/>

<sup>5</sup> <http://www.latex-project.org>

<sup>6</sup> <https://code.google.com/p/abntex2/>

<sup>7</sup> <http://www.inkscape.org>

<sup>8</sup> <http://www.gimp.org>

<sup>9</sup> <http://www.scilab.org>

*“We forget that the water cycle and the life cycle are one.”*

*Jacques Yves Cousteau*



---

# Resumo

Silva, S. R. **Modelagem de Oxigênio Dissolvido utilizando Redes Neurais Artificiais**. 117 p. Dissertação – Câmpus Goiânia, Instituto Federal de Goiás, 2014 .

Este trabalho descreve a utilização da técnica de Redes Neurais Artificiais (RNA), com algoritmo de retropropagação de erro, para modelar o valor do Oxigênio Dissolvido (OD), utilizando-se como entrada dez parâmetros físico químicos provenientes de amostras de água do Rio Meia Ponte, cuja bacia se situa na região metropolitana de Goiânia, GO, e Rio Paraíba do Sul, cuja bacia está localizada no sudeste do estado de São Paulo. A arquitetura da RNA é discutida, assim como os resultados. O objetivo não é apenas encontrar a melhor arquitetura para o problema dado, mas também avaliar a acuidade da predição do OD. Outra contribuição é encontrar uma técnica que permita treinar as RNA com menor quantidade de dados de entrada. Para tanto foram conduzidos experimentos com a técnica de Análise de Componentes Principais (PCA), que se mostrou eficiente ao representar os dados originais através de sua variância acumulada. Para os experimentos, os dados são anteriormente tratados, sendo divididos em dois conjuntos utilizados no treinamento supervisionado e nos testes de avaliação. Para os parâmetros físico químicos selecionados, a RNA apresentou erros pequenos na comparação do OD observado contra o OD calculado, em vários pontos amostrados. O estudo do OD é importante porque sua quantidade no corpo hídrico está diretamente relacionada à capacidade daquele corpo sustentar vida. Modelar com precisão o OD significa que a ferramenta pode ser utilizada para pesquisas futuras prevendo Oxigênio Dissolvido baseadas na extrapolação de um dos dez parâmetros de entrada, que podem variar de acordo com mudanças no ambiente ecológico do rio. É apresentada uma aplicação que implementa a técnica avaliada, programada em linguagem Java, multi plataforma e de código aberto.

**Palavras-chave:** Redes Neurais Artificiais. Retropropagação de erro. Oxigênio dissolvido.

---

# Abstract

Silva, S. R. **Dissolved Oxigen modeling using Artificial Neural Networks.** 117 p. Master Dissertation – Goiânia Campus, Federal Institute of Goiás, 2014 .

This paper describes the utilization of the Artificial Neural Networks (ANN) technique, with the error backpropagation training algorithm, for modeling Dissolved Oxygen (DO) by using ten physical and chemical parameters as network input proceeding from water samples of Meia Ponte river, whose drainage basin is located in metropolitan region of Goiania (GO), and Paraiba do Sul river, whose basin is located on southwest of Sao Paulo state. The ANN architecture is discussed, as well as the results. The objective is not only to find the best architecture for the given problem but also avaliate the acuracy of prediction for the dissolved oxygen values. Another contribution is to find a tecnique that allow to train ANN with less input data. For such it was conducted experiments with the Principal Component Analysis (PCA), which has shown efficient by representing the original data throught their accumulated variance. For the experiments, the data are briefly processed and then divided in two data sets that are used for supervised training and avaliation tests. For the chosen physical and chemical parameters, the ANN has presented small errors in comparison with the real dissolved oxygen and the one calculated by the net, for many sampling points. The study of dissolved oxygem is important because it's quantity on the hidric body is directly connected with the capacity of that body to sustain live. A acuracy of dissolved oxygen means that the tool can be used for future researches predicting dissolved oxygen based in the extrapolation of one of the ten input parameters. These parameters may vary due to changes in the ecological river environment. It is presented an application that implements the Artificial Neural Network, programmed using Java language, multi platform, open source code, that can be used by any interested person.

**Keywords:** Artificial Neural Networks. Backpropagation error. Dissolved oxygen.

---

## Lista de ilustrações

Figura 1 – Localização da bacia do Rio Meia Ponte no Estado de Goiás . . . . .	21
Figura 2 – Municípios que fazem parte da bacia do Rio Meia Ponte no Estado de Goiás . . . . .	21
Figura 3 – Localização da bacia do Rio Paraíba do Sul no Estado de São Paulo (SP)	22
Figura 4 – Decomposição da matriz de dados originais para PCA . . . . .	35
Figura 5 – Estrutura genérica de RNA multicamadas de arquitetura Entrada-Processamento- Saída . . . . .	39
Figura 6 – Modelo de neurônio biológico. . . . .	40
Figura 7 – Modelo de Neurônio Artificial, segundo McCulloch-Pits. . . . .	41
Figura 8 – Diagrama de blocos de aprendizagem com professor. . . . .	45
Figura 9 – Esquema gráfico do algoritmo de retropropagação de erro. . . . .	46
Figura 10 – Sigmoid bipolar, no intervalo $[-1, 1]$ . . . . .	47
Figura 11 – Localização dos pontos de amostragem do Rio Meia Ponte utilizados nos experimentos. . . . .	52
Figura 12 – Exemplo do <i>software</i> NEUROPH de rede MLP na arquitetura (10-11-1). . . . .	61
Figura 13 – Visualização gráfica do <i>software</i> NEUROPH de rede MLP na arquitetura (10-11-1). . . . .	61
Figura 14 – Erro médio quadrático da rede (10- $x$ -1) para o Ponto 1. . . . .	64
Figura 15 – Número de iterações em função da quantidade de neurônios na camada intermediária para o Ponto 1. . . . .	66
Figura 16 – Dado original contra dado modelado pela RNA para o Ponto 1. . . . .	67
Figura 17 – Dispersão de dado original contra dado modelado pela RNA para o Ponto 1. . . . .	67
Figura 18 – Erro médio quadrático da rede (10- $x$ -1) para o ponto 3. . . . .	69
Figura 19 – Número de iterações em função da quantidade de neurônios na camada intermediária para o ponto 3. . . . .	70
Figura 20 – Dado original contra dado modelado pela RNA para o ponto 3. . . . .	70

Figura 21 – Dispersão de dado original contra dado modelado pela RNA para o ponto 3. . . . .	71
Figura 22 – Erro médio quadrático da rede (10- $x$ -1) para o ponto 4. . . . .	72
Figura 23 – Número de iterações em função da quantidade de neurônios na camada intermediária para o ponto 4. . . . .	73
Figura 24 – Dado original contra dado modelado pela RNA para o ponto 4. . . . .	73
Figura 25 – Dispersão de dado original contra dado modelado pela RNA para o ponto 4. . . . .	74
Figura 26 – Erro médio quadrático da rede (10- $x$ -1) para o ponto 5. . . . .	75
Figura 27 – Número de iterações em função da quantidade de neurônios na camada intermediária para o ponto 5. . . . .	76
Figura 28 – Dado original contra dado modelado pela RNA para o ponto 5. . . . .	76
Figura 29 – Dispersão de dado original contra dado modelado pela RNA para o ponto 5. . . . .	77
Figura 30 – Erro médio quadrático da rede (10- $x$ -1) para o ponto 11. . . . .	78
Figura 31 – Número de iterações em função da quantidade de neurônios na camada intermediária para o ponto 11. . . . .	79
Figura 32 – Dado original contra dado modelado pela RNA para o ponto 11. . . . .	79
Figura 33 – Dado original contra dado modelado pela RNA para o ponto 11. . . . .	80
Figura 34 – Erro Total da Rede e Erro Médio Quadrático da rede (10- $x$ -1) para o Rio Paraíba do Sul. . . . .	83
Figura 35 – Dado original contra dado modelado pela RNA para o Rio Paraíba do Sul. . . . .	84
Figura 36 – Dispersão do dado original contra dado modelado pela RNA para o Rio Paraíba do Sul. . . . .	84
Figura 37 – Erro Total da Rede e Erro Médio Quadrático da rede (6- $x$ -1) para o ponto 3 - PCA. . . . .	87
Figura 38 – Dado original contra dado modelado pela RNA para o Ponto 3 - PCA. . . . .	88
Figura 39 – Dispersão do dado original contra dado modelado pela RNA para o Ponto 3 - PCA. . . . .	89
Figura 40 – Erro Total da Rede e Erro Médio Quadrático da rede (5- $x$ -1) em função do número de neurônios na camada intermediária para o ponto 11 - PCA. . . . .	91
Figura 41 – Dado original contra dado modelado pela RNA para o Ponto 11 - PCA. . . . .	92
Figura 42 – Dispersão do dado original contra dado modelado pela RNA para o Ponto 11 - PCA. . . . .	92
Figura 43 – Erro Total da Rede e Erro Médio Quadrático da rede (8- $x$ -1) em função do número de neurônios na camada intermediária para o Rio Paraíba do Sul - PCA. . . . .	95

Figura 44 – Dado original contra dado modelado pela RNA para o Rio Paraíba do Sul - PCA. . . . .	96
Figura 45 – Dispersão do dado original contra dado modelado pela RNA para o Rio Paraíba do Sul - PCA. . . . .	96
Figura 46 – Comparação entre Modelo com dados teóricos e PCA - Ponto 3. . . . .	97
Figura 47 – Comparação entre Modelo com dados teóricos e PCA - Ponto 11. . . . .	98
Figura 48 – Comparação entre Modelo com dados reais e PCA - Rio Paraíba do Sul. . . . .	98
Figura 49 – Esquema de neurônio na camada de saída. . . . .	114
Figura 50 – Esquema de neurônio na camada escondida. . . . .	115

---

## Lista de tabelas

Tabela 1 – Valores de Oxigênio Dissolvido para Rios de Água Doce segundo resolução CONAMA 357. . . . .	23
Tabela 2 – Concentração máxima de fósforo total em mg/L para três classes de água doce segundo a Resolução CONAMA 357/05. . . . .	33
Tabela 3 – Parâmetros orgânicos e inorgânicos segundo a resolução CONAMA 357 para rios de classe 2. . . . .	38
Tabela 4 – Amostras do ponto de coleta 4 situado no município de Inhumas - Jusante 500m do Lançamento da ETE - antes da aplicação do processo de interpolação linear. . . . .	53
Tabela 5 – Tempo para execução dos treinamentos no <i>software</i> NEUROPH. Experimento 1 - Ponto 3. . . . .	58
Tabela 6 – Configuração para treinamento da RNA do Experimento 1 . . . . .	62
Tabela 7 – Dados médios obtidos com o treinamento da rede (10- $x$ -1) para o Ponto 1. . . . .	64
Tabela 8 – Resultado do teste da rede (10-4-1) para o Ponto 1. . . . .	65
Tabela 9 – Dados médios obtidos com o treinamento da rede (10- $x$ -1) para o Ponto 3. . . . .	69
Tabela 10 – Dados médios obtidos com o treinamento da rede (10- $x$ -1) para o Ponto 4. . . . .	72
Tabela 11 – Dados médios obtidos com o treinamento da rede (10- $x$ -1) para o Ponto 5. . . . .	75
Tabela 12 – Dados médios obtidos com o treinamento da rede (10- $x$ -1) para o Ponto 11. . . . .	78
Tabela 13 – Configuração para treinamento da RNA do Experimento 2 . . . . .	81
Tabela 14 – Dados médios obtidos com o treinamento da rede (10- $x$ -1) para o Rio Paraíba do Sul. . . . .	82
Tabela 15 – Configuração para treinamento da RNA do Experimento 3 . . . . .	85
Tabela 16 – Porcentagem de variância do ponto 3. . . . .	86

Tabela 17 – Dados médios obtidos com o treinamento da rede (6- $x$ -1) para o Ponto 3 - PCA. . . . .	87
Tabela 18 – Porcentagem de variância para o ponto 11. . . . .	90
Tabela 19 – Dados médios obtidos com o treinamento da rede (5- $x$ -1) para o Ponto 11 - PCA. . . . .	91
Tabela 20 – Porcentagem de variância para o Rio Paraíba do Sul. . . . .	93
Tabela 21 – Dados médios obtidos com o treinamento da rede (8- $x$ -1) para o Rio Paraíba do Sul - PCA. . . . .	95

---

## Lista de siglas

**ACP** Análise de Componentes Principais

**API** *Application Programming Interface*

**CETESB** Companhia Ambiental do Estado de São Paulo

**CONAMA** Conselho Nacional do Meio Ambiente

**CSV** *Comma-separated Values*

**DBO** Demanda Bioquímica de Oxigênio

**EMBRAER** Empresa Brasileira de Aeronáutica

**IAP** Índice de Qualidade da Água Bruta para fins de Abastecimento Público

**IBGE** Instituto Brasileiro de Geografia e Estatística

**IQA** Índice de Qualidade de Águas

**MLP** *Multi Layer Perceptron*

**OO** Orientado a Objeto

**OD** Oxigênio Dissolvido

**PCA** *Principal Component Analysis*

**pH** Potencial Hidrogeniônico

**RNA** Redes Neurais Artificiais

**SANEAGO** Empresa de Saneamento do Estado de Goiás

**SEPLAN** Secretaria de Planejamento do Estado de Goiás

**UGRHI** Unidade Hidrográfica de Gerenciamento de Recursos Hídricos



---

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b> . . . . .	<b>18</b>
<b>1.1</b>	<b>Área de estudo</b> . . . . .	<b>20</b>
<b>1.2</b>	<b>Objetivos</b> . . . . .	<b>24</b>
<b>1.3</b>	<b>Justificativa</b> . . . . .	<b>24</b>
<b>1.4</b>	<b>Premissa</b> . . . . .	<b>25</b>
<b>1.5</b>	<b>Estrutura do Trabalho</b> . . . . .	<b>25</b>
<b>1.6</b>	<b>Revisão Bibliográfica</b> . . . . .	<b>25</b>
1.6.1	Modelagem de Oxigênio Dissolvido . . . . .	25
1.6.2	Aplicações de Redes Neurais na Área Ambiental . . . . .	27
<b>2</b>	<b>REFERENCIAL TEÓRICO</b> . . . . .	<b>29</b>
<b>2.1</b>	<b>Parâmetros Ambientais e Físico Químicos</b> . . . . .	<b>29</b>
2.1.1	Parâmetros Físicos da Água . . . . .	29
<b>2.2</b>	<b>Quimiometria</b> . . . . .	<b>34</b>
2.2.1	Análise dos Componentes Principais . . . . .	35
<b>2.3</b>	<b>Legislação Formal e Material</b> . . . . .	<b>36</b>
<b>2.4</b>	<b>Redes Neurais Artificiais</b> . . . . .	<b>38</b>
2.4.1	Redes Neurais Multicamadas . . . . .	39
2.4.2	Neurônios Biológicos . . . . .	40
2.4.3	Neurônios Artificiais . . . . .	41
2.4.4	Definição da Arquitetura da RNA . . . . .	42
2.4.5	Paradigmas de Aprendizagem . . . . .	44
2.4.6	Algoritmo de Retropropagação de Erro . . . . .	46
2.4.7	Desafio - melhorar a generalização da RNA . . . . .	49
<b>3</b>	<b>METODOLOGIA</b> . . . . .	<b>50</b>
<b>3.1</b>	<b>Análise dos dados</b> . . . . .	<b>50</b>
3.1.1	Amostras de dados . . . . .	50

3.1.2	Condicionamento dos Dados para Processamento . . . . .	54
<b>3.2</b>	<b>Recursos computacionais . . . . .</b>	<b>54</b>
3.2.1	NEUROPH . . . . .	55
3.2.2	SCILAB . . . . .	56
3.2.3	Computador utilizado nos experimentos . . . . .	57
<b>4</b>	<b>RESULTADOS E DISCUSSÃO . . . . .</b>	<b>59</b>
<b>4.1</b>	<b>Treinamento/Testes de verificação . . . . .</b>	<b>59</b>
<b>4.2</b>	<b>Experimento 1 - dados teóricos . . . . .</b>	<b>60</b>
4.2.1	Configuração para o treinamento . . . . .	62
4.2.2	Ponto de amostragem 1 . . . . .	63
4.2.3	Ponto de amostragem 3 . . . . .	68
4.2.4	Ponto de amostragem 4 . . . . .	71
4.2.5	Ponto de amostragem 5 . . . . .	74
4.2.6	Ponto de amostragem 11 . . . . .	77
<b>4.3</b>	<b>Experimento 2 - dados reais . . . . .</b>	<b>80</b>
4.3.1	Configuração para o treinamento . . . . .	81
4.3.2	Ponto de amostragem PARB 02310 . . . . .	81
<b>4.4</b>	<b>Experimento 3 - PCA . . . . .</b>	<b>84</b>
4.4.1	Ponto de amostragem 3 . . . . .	85
4.4.2	Ponto de amostragem 11 . . . . .	89
4.4.3	Rio Paraíba do Sul . . . . .	93
<b>4.5</b>	<b>Comparação entre os Experimentos 1, 2 e 3 . . . . .</b>	<b>96</b>
<b>5</b>	<b>CONCLUSÃO . . . . .</b>	<b>99</b>
<b>5.1</b>	<b>Objetivos alcançados . . . . .</b>	<b>101</b>
<b>5.2</b>	<b>Trabalhos futuros . . . . .</b>	<b>102</b>
	<b>Referências . . . . .</b>	<b>103</b>

## APÊNDICES 110

<b>APÊNDICE A</b>	<b>– ALGORITMO DE TREINAMENTO <i>BACKPROPAGATION</i> . . . . .</b>	<b>111</b>
<b>APÊNDICE B</b>	<b>– DEMONSTRAÇÃO MATEMÁTICA DO ALGORITMO <i>BACKPROPAGATION</i> . . . . .</b>	<b>113</b>
<b>B.1</b>	<i>j</i> pertence à camada de saída . . . . .	114
<b>B.2</b>	<i>j</i> é neurônio pertencente à camada escondida . . . . .	115

APÊNDICE C	–	<i>SCRIPT</i> DE NORMALIZAÇÃO . . . . .	117
C.1		normaliz.sci . . . . .	117

---

## Introdução

*“Water is a unique liquid; without it, life as we know it is impossible.”*

A afirmação de Daniel B. Botkin e Edward A. Keller, pesquisadores e autores da obra *“Environmental Science: Earth as a Living Planet”*, com a qual iniciam seu capítulo sobre *“Water Supply, Use and Management”*, dá o tom do papel deste precioso componente da vida, completamente atrelado à própria existência da vida na terra. Desde as primeiras civilizações, basicamente nômades, datadas de 200.000 anos a.C., até os dias atuais, passando pelo aparecimento dos primeiros povoados há 10.000 anos, a água pura sempre teve importância vital no abastecimento e saneamento da civilização (JUUTI et al., 2007).

Mensurar a importância deste componente vital para os seres vivos está além do escopo deste trabalho. É necessário, no entanto, trazer à memória a responsabilidade para com a preservação da qualidade deste valioso líquido, principalmente em função das mudanças ambientais registradas nas últimas décadas.

Apesar de ser abundante no ecossistema, é preciso fazer algumas considerações sobre os números que cercam o líquido da vida. O planeta Terra possui 1,4 bilhão de km<sup>3</sup> de água (10,4 bilhões de litros). Destes, 97,5% encontra-se nos oceanos e não pode ser aproveitada para consumo, por conta da salinidade ou dificuldade de extração. Dos 2,5% de água doce restantes, a grande maioria (77,2%) está em lugares de difícil acesso, como geleiras e calotas polares. Outros 22,4% está armazenada no subsolo, 0,399% encontra-se disponível em rios e lagos e apenas 0,001% da água existente no planeta está na atmosfera e disponível para uso humano através da precipitação. Com tempo de residência média de nove dias, o menor dentre todas as fontes deste recurso, a água encontrada na atmosfera é o ator importante no ciclo hidrológico, garantindo o abastecimento de rios, lagos, além da irrigação do solo, através das trocas entre superfície e o ambiente atmosférico (BOTKIN; KELLER, 2011).

Se a quantidade do precioso líquido potável é pequena, projeções da Organização das Nações Unidas (ONU) indicam aumento em seu uso para fins não agrícolas de aproximadamente 897 km<sup>3</sup> no ano 2000 para 2.216 km<sup>3</sup> em 2050. Outras estimativas apontam para o risco de cerca de 4 bilhões de habitantes, que vivem no sul do continente

asiático, nos estados árabes e na África sub-sahariana viverem no meio do século em situação de *stress* de água, dispondo de menos de 1.700 m<sup>3</sup> por pessoa por ano. Ainda mais drástica é a situação de 1,4 bilhão de pessoas da mesma região, que viverão em situação de absoluta escassez, com menos de 100 m<sup>3</sup> do líquido *per capita*, por ano, para fins de agricultura, indústria, energia e ambiente (UNESCO, 2009)(WWAP, 2014). Outros dados registram aumento na demanda atual, pois enquanto o uso de água em 1975 era de 700 m<sup>3</sup> ao ano, *per capita*, e toda a humanidade consumia 3.850 km<sup>3</sup> ao ano, números atualizados demonstram que a demanda humana total de água é de cerca de 6.000 km<sup>3</sup> ao ano, uma fração significativa do total de água potável disponível (BOTKIN; KELLER, 2011).

Enquanto o aumento da demanda parece inevitável, estudos apontam para outra variável que merece consideração. Segundo dados do Painel Intergovernamental de Mudanças Climáticas (*Intergovernmental Panel on Climate Change - IPCC*) da ONU, formado por cientistas de todo o mundo, no século XX as atividades humanas aumentaram a presença do gás de efeito estufa na atmosfera em cerca de 30%, se comparado com a era pré-industrial. O impacto destas emissões já começam a ficar aparentes: aumento da temperatura média da superfície do planeta em 0,7°C. Segundo o estudo, o dióxido de carbono, principal gás do efeito estufa, pode ter seus lançamentos aumentados em até 63% até 2030, ultrapassando as 450 partes por milhão (ppm), o que pode provocar o aquecimento da terra entre 1,4°C e 5,8°C até 2100. Em um cenário positivo, se estes lançamentos se estabilizarem em 450ppm, a temperatura ainda deve subir em torno de 2°C (UN, 2006).

Ainda segundo o Painel da ONU, o aquecimento do planeta tende a provocar severas mudanças climáticas, como o efeito do derretimento de geleiras, que causa mudanças no ciclo hidrológico de muitos rios asiáticos e seus fluxos, impactando diretamente na precipitação e portanto na disponibilidade de água potável. O continente americano pode sofrer os mesmos efeitos se o cenário se confirmar, com diminuição de mais de 20% do escoamento de rios importantes. Países em desenvolvimento podem sofrer em até 70% de redução da disponibilidade de água *per capita*, neste cenário, tendo intensificados problemas como a educação de mulheres e crianças, uma vez que pesquisas em 45 países em desenvolvimento apontam que atualmente, em lugares onde a fonte de água potável se encontra distante das residências, as mulheres são responsáveis por dois terços da mão de obra para coletá-la, enquanto que em 12% dos casos, crianças são empregadas para fazer o transporte do recurso para seus lares, o que lhes impede de frequentar regularmente as aulas (UNICEF; WHO, 2010)(UNESCO, 2009).

A partir do histórico apresentado, cabe o cuidado com o tratamento da água disponível com vistas a promover sua qualidade e consequente utilização. Existem registros do conhecimento da influência dos diferentes tipos de água na saúde de seres humanos desde 400 anos a.C.. Mesmo naquela época, testes sensoriais já eram realizados para medir a

qualidade do líquido, através do paladar, cheiro, temperatura e aparência, e mais tarde, utilizou-se alguns métodos para melhorar a qualidade da água, como tanques de sedimentação, coadores e filtros e dão assim, uma ideia da importância destes cuidados para a qualidade de vida dos seres humanos (JUUTI et al., 2007).

## 1.1 Área de estudo

A nação brasileira possui uma das maiores áreas territoriais do planeta e a maior bacia hidrográfica do mundo ocupa grande parte da Amazônia nacional. Pode-se destacar a presença de diversos aquíferos e uma extensa capilaridade fluvial em grande parte da sua superfície.

Estes recursos hídricos em abundância, são um privilégio. Mas não podem, entretanto, ser confundidos como garantia de estoque infindável de água potável. Antes, constituem-se em oportunidade para o exercício da responsabilidade para com esta e as futuras gerações, no sentido de proteger os mananciais hídricos existentes e promover a descontaminação de rios e lagos, além do reaproveitamento de águas residuais industriais, visando promover a sustentabilidade ambiental.

Em plano local, o Estado de Goiás é o sétimo em tamanho territorial e possui a população estimada em 6.080.588 habitantes. A economia estadual, baseada em atividades agropecuárias, industriais e de serviços, cresceu em 67% nos últimos 10 anos, com índices acima da média nacional, conforme dados da Secretaria de Planejamento do Estado de Goiás (SEPLAN), gerando Produto Interno Bruto de R\$ 103,44 bilhões, o nono maior do país (SEPLAN, 2012).

De acordo com o Instituto Brasileiro de Geografia e Estatística (IBGE), a região metropolitana de Goiânia, capital do Estado, possui 20 municípios dentro de suas fronteiras e importante participação no quantitativo populacional total, com aproximadamente 2 milhões de habitantes (IBGE, 2010). Esta região está situada na bacia hidrográfica do Meia Ponte, localizada inteiramente no Estado de Goiás, como pode ser vista na Figura 1, integrante da grande Bacia do Rio Paraná. O rio homônimo, caracterizado como de Classe 2 pela resolução 357 do CONAMA, banha 38 municípios, ilustrados na Figura 2 e situados na região de maior densidade populacional do Estado, totalizando cerca de 2.590.972 habitantes (SEPLAN, 2012) e geradores de 49% da riqueza estadual (DO CERRADO, 2008a). O curso principal do rio nasce na Serra dos Brandões, no município de Itauçu, a aproximadamente 50km a noroeste da capital, de onde parte para a trajetória de 437 km rumo ao deságue no Rio Paranaíba (MELLO et al., 2011). O Rio Meia Ponte é de vital importância para a vida sócio-econômica local pois 82% dos recursos deste manancial são utilizados para abastecimento público ou irrigação (produção de alimentos) (DO CERRADO, 2008b). No entanto, sua localização na região mais populosa e economicamente ativa do Estado traz sérias consequências à sua saúde, especialmente para o

trecho que atravessa a Cidade de Goiânia, uma vez que a diluição de efluentes domésticos e industriais sem tratamento deterioram a qualidade das suas águas (DO CERRADO, 2008a)(BRASIL, 2005)(CARVALHO; SIQUEIRA, 2011).



Figura 1 – Localização da bacia do Rio Meia Ponte no Estado de Goiás



Figura 2 – Municípios que fazem parte da bacia do Rio Meia Ponte no Estado de Goiás

O Rio Paraíba do Sul, cujos dados de um ponto de amostragem também alimentam este estudo, faz parte da Unidade Hidrográfica de Gerenciamento de Recursos Hídricos (UGRHI) homônima, localizada no sudeste do Estado de São Paulo (SP). A bacia Paraíba do Sul, vista na Figura 3, banha 36 municípios em três estados, servindo uma população total de 2.904.650 habitantes (HÍDRICOS et al., 2008). O Rio abastece uma das regiões mais economicamente ativas do estado paulista, o Vale do Paraíba, incluindo o município de São José dos Campos, importante expoente econômico local, que possui população de 638.990 habitantes e uma base consolidada de indústrias da área de aeronáutica, como a Empresa Brasileira de Aeronáutica (EMBRAER) e o Instituto Tecnológico de Aeronáutica (HENRIQUE, 2012).

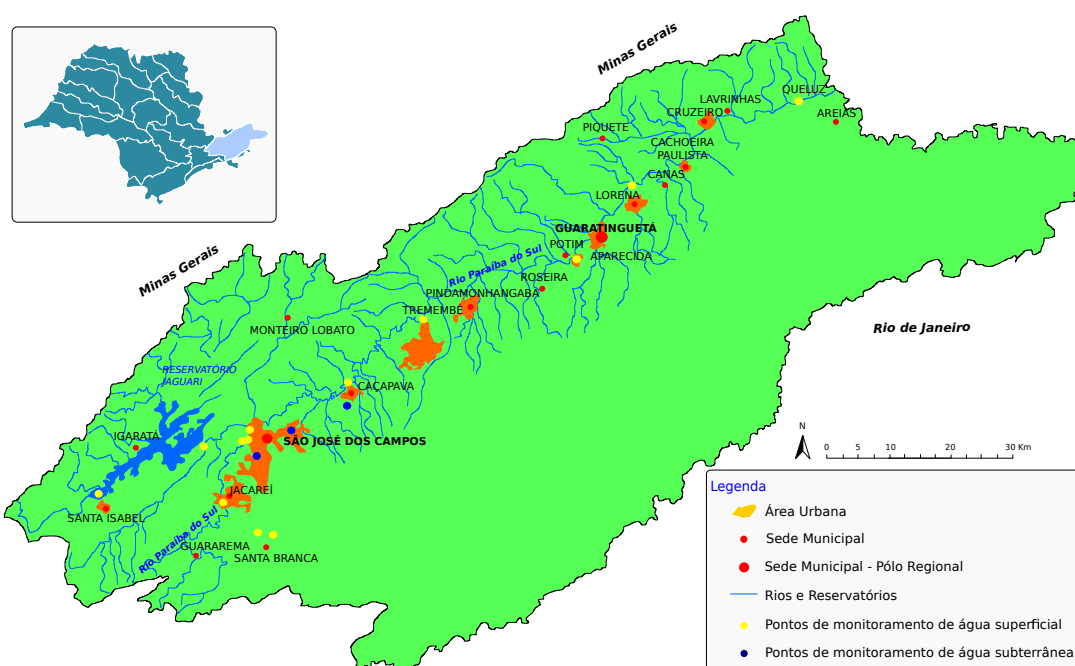


Figura 3 – Localização da bacia do Rio Paraíba do Sul no Estado de São Paulo (SP)

A legislação brasileira caracteriza os recursos hídricos de Água Doce utilizados para abastecimento e produção de alimentos, após tratamento convencional, como recursos de Classe 2, através da resolução nº. 357 do Conselho Nacional do Meio Ambiente (CONAMA). Estes recursos são inestimáveis para a subsistência da população. A Tabela 1 exibe o valor de Oxigênio Dissolvido para esta classe, assim como para os corpos hídricos das demais classificações. Assim, a análise de qualidade da água destes mananciais é essencial para sua longevidade e consequentemente a proteção da população. Vários métodos e ferramentas podem ser utilizadas para este fim, na tentativa de proteger os recursos hídricos de água doce, tão necessários e fundamentais à manutenção da vida (UN, 2006). Algumas análises são baseadas na coleta de amostras de água para posterior análise laboratorial. Em seguida, os quantitativos de cada componente físico-químico e biológico que compõem o manancial são comparados com as especificações da legislação



vigente. Frequentemente este trabalho de análise dos parâmetros é apoiado por índices ou indicadores da qualidade de água, construídos por agências, institutos de pesquisa e universidades, como é o caso do Índice de Qualidade da Água Bruta para fins de Abastecimento Público (IAP) e do Índice de Qualidade de Águas (IQA) (BRASIL, 2013).

Tabela 1 – Valores de Oxigênio Dissolvido para Rios de Água Doce segundo resolução CONAMA 357.

Classificação	Destinação	Tipo de tratamento	Valor de OD (mg/L)
Classe 1	abastecimento para consumo humano	simplificado	não inferior a 6,0
Classe 2	abastecimento para consumo humano	convencional	não inferior a 5,0
Classe 3	abastecimento para consumo humano	convencional ou avançado	não inferior a 4,0
Classe 4	navegação ou paisagismo	nd	superior a 2,0

Enquanto o Oxigênio Dissolvido (OD) costuma constar nos parâmetros avaliados pelos índices, este trabalho procura, através da avaliação de outros parâmetros, com a ajuda do ferramental das redes neurais, modelar seu valor, ou seja, através do quantitativo das variáveis de entrada, avaliar a resposta do sistema para a variável de saída.

Conforme Cunha (2013), o processo de modelagem de sistemas ecológicos, hidrodinâmicos e bioquímicos, que consiste em estabelecer hipóteses sobre o funcionamento do sistema, vem sendo utilizado há muito tempo na ciência. O modelo pode estimar, ainda que de forma relativa, o impacto da descarga de resíduos no caso da composição de recursos hídricos. Para tal, a escolha das variáveis de entrada é estratégica e se dá em função do propósito de cada estudo particular, visando estabelecer relações qualitativas e quantitativas entre as características do rio (SIQUEIRA, 1996).

Segundo WARWICK (1991 apud SIQUEIRA, 1996), modelos com menos variáveis de entrada costumam ser mais flexíveis e garantem melhor desempenho que modelos sofisticados, mesmo que as complexidades do sistema sejam reais. Isto porque as interações entre variáveis de sistemas ecológicos apresentam alta complexidade de análise por parte de sistemas analíticos convencionais. Por esta razão, a modelagem destes sistemas exige programas de computador altamente especializados em análise não linear dos dados (CUNHA, 2013), agindo como especialistas, quando corretamente utilizados. DEMETRACOPoulos e STEFAN (1983 apud SIQUEIRA, 1996) que modelaram o reservatório nº. 2 do Rio Mississípi, observaram vários modelos surgir para o mesmo reservatório desde a década de 60. Eles reiteram que, além de complexo, o processo de modelação é interativo e necessita ser continuamente aprimorado para o mesmo recurso hídrico.

A modelagem de sistemas hídricos é estratégica para tomada de decisão, visando garantir a qualidade e quantidade de água que satisfaça as necessidades humanas e ecológicas diante dos crescentes usos. Scariot (2008) lembra que garantir o acesso à água significa mais do que apenas atendimento a necessidades básicas, pois “a garantia do acesso à qualidade da água assegura acesso justo à saúde, desenvolvimento econômico e ambientes ecologicamente saudáveis para uma nação”.

Desta forma, com o risco de redução da quantidade de água no futuro e a qualidade

das águas disponíveis atualmente nos rios cada vez mais comprometida, torna-se estratégico entender o OD, com vistas à preservação dos corpos hídricos, que vêm sofrendo com pressões do aumento populacional, uso por indústrias, localizadas nas regiões mais populosas, além da utilização para irrigação e produção de alimentos, responsável por grande uso e poluição dos recursos hídricos. Modelar este parâmetro é importante em função de seu próprio valor para os cursos de água superficiais. A respiração de organismos vivos aeróbios e a manutenção da vida aquática deve-se a valores saudáveis deste parâmetro. Baixas concentrações de OD, abaixo de 4 a 5 mg/L, ou mesmo condições anaeróbias conduzem a desequilíbrio no ecossistema em geral, com mortalidade de peixes, liberação de gás sulfídrico, odores e efeitos estéticos, interferindo na saúde do ambiente aquático (SIQUEIRA, 1996)(MONTEIRO, 1997).

## 1.2 Objetivos

Este trabalho tem por objetivo principal modelar o parâmetro Oxigênio Dissolvido nos Rios Meia Ponte (GO) e Paraíba do Sul (SP) com base na técnica de Redes Neurais Artificiais (RNA).

Para atingir o objetivo principal, outros objetivos secundários fizeram-se necessários:

- ❑ Avaliar os parâmetros de qualidade de água, baseados na resolução 357 e 430 do CONAMA;
- ❑ Aplicar os conceitos de redes neurais artificiais, seus algoritmos de treinamento e suas potencialidades;
- ❑ Disponibilizar um sistema computacional que implemente as técnicas de redes neurais, programado utilizando a linguagem Java, multiplataforma e aberto, de forma a permitir sua posterior utilização por qualquer interessado;
- ❑ Avaliar técnicas que possibilitem treinar as Redes Neurais Artificiais utilizando menor quantidade de dados, como a matriz de componentes principais gerada através da técnica de Análise de Componentes Principais.
- ❑ Avaliar qual é a quantidade de neurônios na camada intermediária que define a melhor arquitetura da rede para cada experimento.

## 1.3 Justificativa

Buscar um processo computacional que use o menor número de parâmetros para encontrar o valor de Oxigênio Dissolvido (OD) mais próximo do real, frente aos vários parâmetros disponíveis na resolução 357 do CONAMA. Esta proposta objetiva ajudar na

estimativa de qualidade dos mananciais de água potável de uso da população, provendo as autoridades de gestão de recursos hídricos com ferramental científico para tomada de decisão. O modelo é mais simples, pois possui menos parâmetros que outros existentes, além de ser de fácil implementação, a partir de recursos facilmente encontrados e disponíveis em qualquer instituição/centro de pesquisa. Como motivação pretende-se encontrar um modelo a partir de uma abordagem local que possa ser generalizada e aplicada para qualquer corpo hídrico.

## 1.4 Premissa

É possível estimar os valores de OD em um ponto a partir de outros parâmetros de qualidade de água neste mesmo ponto.

## 1.5 Estrutura do Trabalho

Este trabalho está organizado em quatro capítulos, incluindo o Capítulo 1, que introduz o leitor aos temas centrais que motivam e justificam esta pesquisa, incluindo objetivos, organização do texto e revisão bibliográfica sobre modelagem de OD e aplicação de RNA na área ambiental.

O Capítulo 2 faz breve levantamento teórico sobre os Parâmetros Ambientais utilizados na modelagem e citados na resolução CONAMA 357/05 como sendo importantes para a classificação dos mananciais, além de Índices de Qualidade de Água e sua importância, finalizando o Capítulo 2 com Redes Neurais Artificiais, Aspectos de Paradigmas de Aprendizagem e o Algoritmo de Retropropagação de Erro.

O Capítulo 3 trata da metodologia do trabalho, fazendo relato da análise dos dados utilizados no treinamento das redes neurais e dos recursos computacionais necessários para o desenvolvimento dos experimentos.

O Capítulo 4 conduz aos experimentos propriamente ditos, através dos quais buscou-se avaliar a capacidade de modelagem do OD através dos dados de pontos de amostragem de água do Rio Meia Ponte e Rio Paraíba do Sul, enquanto é feita a análise dos resultados.

O Capítulo 5 avalia os objetivos alcançados e sugere trabalhos futuros, concluindo o trabalho.

## 1.6 Revisão Bibliográfica

### 1.6.1 Modelagem de Oxigênio Dissolvido

Areerachakul et al. (2011) descrevem a utilização de uma RNA com algoritmo *Levenberg-Marquardt* para predição de oxigênio dissolvido no distrito de Dusit, em Bangkok, Tai-

lândia a partir de dez parâmetros de entrada. A arquitetura que encontrou o menor erro foi com 16 neurônios na camada intermediária.

O trabalho de Kisi e AY (2011) descreve a utilização de dois tipos diferentes de RNA para modelagem de OD no estado norte-americano do Colorado. A rede *perceptron* multicamadas foi utilizada com algoritmo de retropropagação de erro, enquanto a rede neural de base radial faz correção dos pesos por ajustamento linear.

O texto dissertativo de KNAPIK (2009) descreve a implementação e calibração de um modelo QUAL2E para um estudo mais detalhado da dinâmica da matéria orgânica em um corpo de água, com estudo de caso desenvolvido na Bacia do Alto Iguaçu, englobando Curitiba e Região Metropolitana. A calibração e simulação das variáveis OD, DBO, Nitrogênio e Fósforo foram utilizadas no modelo QUAL2E com o objetivo de consolidar o diagnóstico.

Mwegoha et al. (2010) descrevem o desenvolvimento de um modelo matemático para predição dos efeitos dos parâmetros Velocidade do Vento, Luz, pH, Temperatura, Dióxido de Carbono Dissolvido e Demanda Bioquímica de Oxigênio no Oxigênio Dissolvido em tanques de peixe.

Nunes et al. (2009) apresentam um sistema *web* para cálculo de OD a partir de características hidráulicas, químicas e físicas oriundas da Bacia do Rio Turvo Sujo, situada na Bacia do Rio Doce, abrangendo parte dos municípios de Viçosa, Coimbra, Cajuri, Teixeiras e Guaraciaba, a partir do modelo de Streeter-Phelps (1925).

Palani et al. (2009) discorrem sobre a modelagem de OD de água do mar, retirada da região do Estreito East Johor em Singapura, utilizando para este estudo dois tipos de RNA, uma *Multi Layer Perceptron* (MLP) e outra rede neural de Regressão Geral (GRNN), a partir da utilização de quatorze parâmetros de entrada para as redes. Os resultados alcançados mostraram boa performance, com erros de menos de 1,55 mg/L para a rede GRNN e de 2,25 mg/L para a rede MLP na comparação entre os valores reais e modelados pelas redes.

A dissertação de Williams (2007) apresenta a modelagem de OD no Lago Powell, situado nos estados norte-americanos de Utah e Arizona, a partir da simulação hidrodinâmica e de qualidade de água CE-QUA-W2. Em função da existência de poucos dados de qualidade de água para o referido lago, processos limnológicos foram observados para o desenvolvimento de métodos empíricos que descrevam a dinâmica do OD.

A tese de Li (2006) apresenta a elaboração de um modelo de Oxigênio Dissolvido e Temperatura (DDOT) mono-dimensional de corrente dinâmica para avaliar possíveis efeitos de modificação humana para o Rio Lower Flint, no estado norte-americano da Georgia.

Heathcote (2009) avalia cinco modelos de OD para o Rio Grand, no estado canadense de Ontario. Os quatro modelos analisados QUAL2K, WASP7, HSPF/BASINS (todos disponíveis como ferramentas de código aberto) e MIKE11 têm os resultados comparados

ao modelo GRSM, desenvolvido especificamente pelas autoridades de conservação do manancial para avaliação de impactos de estratégias alternativas de gerenciamento de água em nutrientes e concentração de OD. Os resultados da avaliação indicam que o modelo GRSM possui a melhor capacidade de modelagem do OD, assim como de crescimento de algas e limitação de luz.

O trabalho de Dutra (2014) apresenta a modelagem via modelo matemático QUAL2K dos parâmetros OD e Demanda Bioquímica de Oxigênio (DBO) em um trecho do Rio Paraibuna situado entre os Córregos Tapera e Independência, situados em Juiz de Fora, MG.

### 1.6.2 Aplicações de Redes Neurais na Área Ambiental

GOMES et al. (2007) em seu “Modelo Baseado em Técnica de Redes Neurais para Previsão de Vazões na Bacia do Rio São Francisco” reforçam que a escolha de redes neurais artificiais para trabalhos com séries temporais se dá principalmente “devido à sua capacidade de tratar com relações de entrada-saída de dados não-lineares, destacando sua habilidade de aprendizado e capacidade de generalização, associação e busca paralela” através de informações “muitas vezes camufladas por ruídos.”

WEIGANG et al. (1998), em seu trabalho “*Prediction of the Level of Paraguay River using Neural Networks*”, implementam uma rede neural com algoritmo de retropropagação para predição dos níveis do Rio Paraguai, em Ladário (MS). A justificativa do autor para a escolha das RNA é a de que análise de séries temporais necessitam de representações de funções não-lineares mais robustas.

Um aspecto importante para o uso das RNA em processos hidrológicos é que “possibilitam a resolução de sistemas mal-definidos ou não claramente compreendidos”, conforme definido por SANTOS e FILHO (2003) em seu trabalho “Modelagem Hidrológica Urbana por meio de Redes Neurais Artificiais: uma Aplicação para a Bacia do rio Tamanduateí em São Paulo, SP”.

O texto dissertativo de RIBEIRO (2003), “Rede Neural com Retropropagação: uma Aplicação na Classificação de Imagens de Satélite”, descreve e implementa RNA multica-madas com algoritmo de retropropagação para verificar a utilização do solo com base em imagens de satélite.

Nazareno et al. (2003) abordam redes MLP com algoritmo de retropropagação para reconhecimento de caracteres, demonstrando a boa atuação deste modelo e arquitetura na solução de problemas não-lineares, em seu artigo intitulado “*Neural Trainer - Software Genérico para o Treinamento de Redes Neurais Multilayer Perceptron com aprendizado supervisionado por retropropagação de erros, um caso de uso: OCR (Optical Character Recognition)*”.

O trabalho de SOUSA e SOUSA (2010) denominado “Rede neural artificial aplicada à previsão de vazão da Bacia Hidrográfica do Rio Piancó” demonstra o treinamento de

uma RNA utilizando-se de 120 amostras, enquanto que ELSHORBAGY et al. (2000) demonstram no trabalho “*Performance Evaluation Of Artificial Neural Networks For Runoff Prediction*” o treinamento da rede com 50 amostras de dados, demonstrando a viabilidade do treinamento utilizando bases de dados amostrais pequenas.

BORSATO et al. (2009) reforçam o valor do algoritmo de retropropagação de erro juntamente com a regra de aprendizagem supervisionada em seu artigo “Aplicação de redes neurais artificiais na identificação de gasolinas adulteradas comercializadas na região de Londrina (PR)”, ao concluir que “a utilização de redes neurais mostrou ser uma ferramenta útil para identificar gasolinas adulteradas já que foi possível reconhecer 100% das amostras analisadas” com este tipo de rede mesmo utilizando número reduzido de 40 amostras para o experimento.

No trabalho *Can Neural Networks be used in Data-Poor Situations?*, Silvert e Baptist (1998) analisam o uso de redes neurais para aplicações ecológicas, reiterando que estas aplicações raramente tem grandes quantidades de dados disponíveis, em função da dificuldade, alto custo e falta de automatização dos processos de coleta de dados (salvo raras exceções como sensoreamento remoto ou dados de telemetria automatizada), citando o caso de uma rede treinada com 14 amostras e testada com 5. No entanto, destaca que estas aplicações exigem métodos cada vez mais sofisticados de análise, o que incentiva o uso de RNA. O autor defende o uso de pré-processamento de dados e descreve o uso com sucesso de Transformada de Fourier para escolher os componentes mais sensíveis como entrada no treinamento de uma RNA.

Weckman et al. (2010) utilizam redes neurais para estimativa de custo em seu artigo *Using Neural Networks with Limited Data to Estimate Manufacturing Cost*, aplicando técnicas de expansão de dados para aumentar com sucesso o conjunto de dados de treinamento para sua rede neural, demonstrando bons resultados diante de outras técnicas de estimativa de custo.

Cunha et al. (2010) utilizaram uma rede neural com algoritmo de retropropagação de erro para avaliar os dados da série histórica da malária no Município de Cantá, Estado de Roraima, através de um estudo de previsão, feito a partir de um número de 72 amostras. O resultado comprovou um maior poder de previsão da RNA quando os resultados são comparados ao modelo de regressão logística.

O texto intitulado *Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique* demonstra a utilização de RNA para a criação de um modelo de DBO a partir de 8 parâmetros de entrada para o Rio Melen, na Turquia. A partir da utilização de técnicas de sensibilidade de dados, os mais representativos foram escolhidos para servir como parâmetros de entrada (DOGAN et al., 2009).

---

## Referencial Teórico

### 2.1 Parâmetros Ambientais e Físico Químicos

#### 2.1.1 Parâmetros Físicos da Água

A água pode ser modificada por sólidos presentes na mesma, o que pode causar variações em sua aparência. Estas variações estão relacionadas aos conteúdos físicos da mesma, podendo-se destacar alguns:

##### 2.1.1.1 Temperatura da Água

A ocorrência de variação na temperatura dos corpos hídricos é fenômeno natural e pode variar em função de diversos fatores, como as estações do ano, período do dia, profundidade, latitude e longitude. Apesar disso, o homem pode contribuir com estas variações através de despejos industriais (termelétricas, usinas nucleares) ou sistemas de resfriamento. Araujo e Oliveira (2011) indicam a importância deste parâmetro na medida em que ele influencia diretamente no metabolismo dos microrganismos, acarretando maior gasto energético. Segundo eles, 10°C de aumento na temperatura são capazes de dobrar a taxa metabólica, aumentando a degradação de matéria orgânica e conseqüentemente o consumo de oxigênio dissolvido. Isto denuncia a clara governança da temperatura nas cinéticas das reações químicas envolvidas no equilíbrio da água.

##### 2.1.1.2 Condutividade

Este parâmetro indica, através de expressão numérica, a capacidade da água conduzir corrente elétrica. Está diretamente relacionado com a presença de íons dissolvidos em seu meio, ou seja, é função da quantidade de sais e da temperatura da amostra e mede indiretamente a concentração de poluentes (MIZUTORI, 2009). Este parâmetro não determina quais íons estão presentes em determinada amostra de água, mas está diretamente relacionado a impactos ambientais ocorridos em função de lançamentos de resíduos industriais,

mineração, esgotos, etc, influenciando diretamente na OD presente no corpo hídrico. O valor deste parâmetro não está fixado na resolução CONAMA 357, porém a Companhia Ambiental do Estado de São Paulo (CETESB) determina valores acima de 100mS/cm para ambientes impactados negativamente (BAGGIO; HORN, 2011).

### 2.1.1.3 Turbidez

Representa o grau de atenuação de intensidade que um feixe de luz sofre ao atravessar a amostra de água, devido a absorção e espalhamento da luz, ocasionado pela presença de sólidos em suspensão como partículas inorgânicas (areia, argila, silte) e de detritos orgânicos (algas, bacterianas, plânctons) (MIZUTORI, 2009). Várias atividades provocam elevação da turbidez, como erosão das margens dos rios, ou mesmo despejo de efluentes sanitários. Em altas intensidades, ela pode reduzir a fotossíntese, que diminui a atividade primária do rio e conseqüentemente a concentração de Oxigênio Dissolvido.

### 2.1.1.4 Sólidos Dissolvidos Totais

A quantidade e natureza dos sólidos dissolvidos totais na água é grande, podendo variar de cerca de 20 a 1000mg/L. Sua presença na amostra significa aumento de turbidez e conseqüente diminuição da transparência que ocasionará, mais tarde, redução na atividade de fotossíntese. Os sólidos dissolvidos são resíduos filtráveis e segundo Mizutori (2009) águas com alta concentração desta variável geralmente apresentam potabilidade inferior e podem conduzir a reações fisiológicas adversas no consumidor. Por este motivo é fixado pela resolução CONAMA 357 (CONAMA, 2005), o limite máximo de 500 mg/L, como disposto na Tabela 3. Nos rios analisados, os valores máximos encontrados foram de 198 mg/L e 106,86 mg/L para os rios de SP e GO, respectivamente. FONSECA e SALVADOR (2005) apontam uma relação direta entre o aumento deste parâmetro e da condutividade. Assim, entende-se que o aumento dos sólidos dissolvidos totais em um corpo hídrico é inversamente proporcional à presença de oxigênio dissolvido neste meio.

### 2.1.1.5 Oxigênio Dissolvido

A quantidade de OD dentro dos corpos de água naturais é variável, dependendo de atividade fotossintética (como das algas) e de fatores como temperatura, turbulência, pressão e salinidade das águas. Assim, é comum variações de concentração de OD dentro de estações do ano, meses, dias ou mesmo ao longo do dia. Pequenas variações não são prejudiciais, mas a pouca disponibilidade de oxigênio atmosférico dentro do meio aquático pode interferir na capacidade de manutenção da vida dos seres aeróbios, que o utilizam na respiração, na qual ocorre a oxidação de compostos orgânicos gerando energia para a formação de Trifosfato de Adenosina (ATP)<sup>1</sup>, fonte de energia para todas as reações

<sup>1</sup> Do termo inglês *Adenosine Triphosphate*.



metabólicas. Para o abastecimento humano, a baixa quantidade de OD presente na água pode indicar grande quantidade de matéria orgânica, prejudicando elementos de filtração e tratamento. Ademais, este é um dos principais parâmetros de caracterização dos efeitos de poluição das águas pois o seu nível indica a capacidade de um corpo hídrico se autodepurar e manter a vida aquática (MIZUTORI, 2009). A solubilidade do oxigênio é inversamente proporcional à temperatura e salinidade da água. Scariot (2008) destaca que a variação temporal e espacial do OD no corpo hídrico pode ser afetada positiva ou negativamente por vários processos físico e bioquímicos, como por exemplo, pela temperatura, salinidade, turbulência, atividade fotossintética das algas e plantas e pressão atmosférica, aumentando com a queda da pressão atmosférica e diminuindo com o aumento da temperatura e da salinidade. As fontes ou sumidouros de OD representam as variações de concentrações de oxigênio no corpo hídrico. MONTEIRO (1997) e Fiorucci e Filho (2005) destacam como principais fontes de OD a difusão de oxigênio atmosférico através da interface ar-água, como no rearejamento proveniente de afluentes e a fotossíntese de plantas aquáticas. Ainda segundo Scariot (2008), a demanda de oxigênio pelo sedimento é um dos maiores sumidouros de OD no corpo hídrico. Esta demanda incorpora a demanda biológica da respiração dos seres vivos, além da demanda química da oxidação de substâncias carbonadas, substâncias azotadas e sedimentos do curso de água. A produção de algas atua tanto como fonte e sumidouro de OD nos rios, uma vez que durante o dia a fotossíntese das algas produz oxigênio em excesso à demanda. Durante a noite, com a impossibilidade da fotossíntese, o consumo de oxigênio pela respiração das algas pode reduzir drasticamente a concentração de OD.

#### 2.1.1.6 Demanda Bioquímica de Oxigênio (DBO)

O Oxigênio Dissolvido nos corpos de água é de vital importância para os organismos aeróbios, que dependem dele para sua sobrevivência. A redução de sua concentração pode ocasionar não apenas a redução deste tipo de vida mas também o aparecimento de outras formas de vida no meio anaeróbio resultante, que produzem resíduos metabólicos nocivos para determinados usos da água. Por este motivo, a DBO é um dos mais importantes parâmetros para avaliação da qualidade da água, uma vez que define a quantidade necessária de oxigênio para metabolização da matéria orgânica biodegradável existente no meio aquático (MIZUTORI, 2009)(CETESB, 2004). A  $DBO_{5,20}$  é o teste padrão realizado a temperatura constante de 20°C durante o período também fixo de 5 dias (VALENTE et al., 1997). Despejos de efluentes industriais ou domésticos no recurso hídrico podem criar uma região de baixo oxigênio dissolvido e alta DBO, restringindo a vida de espécies que toleram muita matéria orgânica e pouco oxigênio (BOTKIN; KELLER, 2011). Segundo Ceretta (2004), a demanda de oxigênio na água é exercida por três classes de materiais:

- Matéria orgânica carbonácea utilizada como fonte de alimentos para organismos

aeróbios, resultando  $\text{CO}_2$ ,  $\text{H}_2$  e  $\text{NH}_3$ ;

- Matéria orgânica nitrogenada oxidável usada como fonte de alimentos para bactérias específicas resultando  $\text{NH}_3$ ,  $\text{CO}_2$ ,  $\text{H}_2\text{O}$  e  $\text{SO}_2$ ;
- Compostos químicos redutores como íon ferroso ( $\text{Fe}^{+2}$ ), sulfito ( $\text{SO}_2^{-}$ ) e sulfeto ( $\text{S}^{2-}$ ), que são oxidados pelo oxigênio dissolvido.

#### 2.1.1.7 pH

O Potencial Hidrogeniônico (pH) representa a atividade hidrogeniônica, ou seja, a concentração de íons  $\text{H}^+$  e  $\text{OH}^-$  que representa a intensidade de ácidos ou álcalis presentes na solução (MIZUTORI, 2009). Ceretta (2004) complementa que em condições de pH elevado, o ferro tende a tornar-se menos disponível para algumas plantas, afetando a produção de toda comunidade aquática e trazendo efeito letal para peixes e outras vidas aquáticas. O valor do pH é estratégico pois denuncia a eficiência do processo de tratamento, em função do contato do líquido com partes metálicas sujeitas a corrosão, além de afetar o sentido do paladar da água. Sua relação com o OD é direta, ou seja, baixos teores deste componente indicam baixos valores de pH no mesmo corpo hídrico.

#### 2.1.1.8 Nitrato

Este elemento é encontrado em águas originárias de esgotos domésticos, industriais ou ainda quando resultante de drenagem de áreas fertilizadas. Em grandes concentrações pode ser responsável por doença infantil denominada metemoglobinemia (ou cianose), que é provocadora da despigmentação da pele, e letal. Por este motivo, o padrão de potabilidade definido pela portaria n° 2.914 de 2011 do Ministério da Saúde estabelece que o valor máximo de nitrato permitido é de 10 mg/L (BRASIL, 2011). Sua quantidade no corpo hídrico é inversamente proporcional à concentração de OD, uma vez que este é consumido durante o processo de nitrificação e por algas que tem neste componente um elemento vital para seu crescimento.

#### 2.1.1.9 Fósforo Total

O Fósforo é o elemento não metálico, que pode ocorrer em diferentes formas: orgânica, inorgânica e como espécies dissolvidas ou particuladas. A resolução CONAMA 357/05 considera o Fósforo como elemento a ser monitorado para se medir o padrão de qualidade da água, e estabelece valores diferentes para os variados ambientes naturais. Estes ambientes são divididos em dois grupos: ambientes lênticos e lóticos. Os primeiros são caracterizados por tempo de residência superior a 40 dias e são representados por lagos e lagoas. Já os ambientes lóticos possuem tempo de residência inferior a 2 dias e são representados por aqueles ambientes com fluxo de água contínuo, como rios. A construção de

reservatórios caracteriza a formação do terceiro grupo, com tempo de residência entre 2 e 40 dias (FERREIRA et al., 2009). A Tabela 2 informa as concentrações de Fósforo Total para os diferentes ambientes (BAIO, 2009). O Fósforo é o nutriente ativamente utilizado pelas plantas, por este motivo raramente encontrado em concentrações significativas em águas naturais. Concentrações acima de 0,1 mg/L podem favorecer o crescimento de algas e desenvolvimento de lodo, afetando o uso da água para abastecimento público (MIZUTORI, 2009).

Tabela 2 – Concentração máxima de fósforo total em mg/L para três classes de água doce segundo a Resolução CONAMA 357/05.

	<b>Lêntico</b>	<b>Intermediário</b>	<b>Lótico</b>
Classe 1	0,02	0,025	0,1
Classe 2	0,03	0,05	0,1
Classe 3	0,05	0,075	0,15

#### 2.1.1.10 Nitrogênio Amoniacal

O Nitrogênio pode ser encontrado em diversas formas nas águas naturais. Depois do Carbono, é o elemento mais exigido pelas células vivas, por este motivo tornando-se fertilizante dos meios aquáticos onde possibilitam o crescimento em maior extensão dos seres vivos que os utilizam, especialmente as algas, em um efeito chamado de “eutrofização”. O Nitrogênio Amoniacal, juntamente com o nitrogênio orgânico, é encontrado nas águas em sua forma reduzida. Nitrito e nitrato são encontrados nas formas oxidadas. Através de testes de laboratório, pode-se avaliar a localização da fonte de poluição através da verificação das formas de nitrogênio encontradas em amostras; as oxidadas indicam fontes distantes, enquanto que as reduzidas indicam proximidade da fonte poluidora. Sua composição amoniacal constitui-se na somatória da amônia ( $\text{NH}_3$ ) com o íon amônio ( $\text{NH}_4^+$ ), encontrada naturalmente nas águas superficiais ou subterrâneas resultante da decomposição da matéria orgânica, de forma que altos níveis indicam poluição doméstica ou industrial, derivada principalmente da hidrólise sofrida pela uréia na água, vinda de lançamentos de esgotos (BRIGANTE; ESPÍNDOLA, 2003 apud GASPAROTTO, 2011)(MIZUTORI, 2009).

A amônia é bastante restritiva à vida dos peixes, sendo que muitas espécies não suportam mais que 5 mg/L deste elemento. Além disso, provoca consumo de Oxigênio Dissolvido das águas naturais ao ser oxidada biologicamente, a chamada DBO de segundo estágio.

#### 2.1.1.11 Nitrito

Outro composto do Nitrogênio, pode ocorrer naturalmente na metabolização de organismos anaeróbios, no entanto, o aumento desse íon indica que os processos biológicos

estão sendo influenciados por poluição orgânica. É coadjuvante no aparecimento da metemoglobinemia, uma doença em que o nitrito se liga à molécula de hemoglobina, a qual torna-se incapaz de transportar oxigênio para as células do organismo, podendo resultar em cianose, anóxia cerebral e morte. O desenvolvimento desta doença a partir do nitrato nas águas potáveis depende da sua conversão bacteriana para nitrito durante a digestão, o que pode ocorrer na saliva e no trato gastrointestinal (GASPAROTTO, 2011).

Por este motivo, os limites considerados pela resolução federal CONAMA 357/05 para este elemento é de 1,0 mg/L N, como apresentado na Tabela 3 (p. 38). A concentração de Nitrito no corpo hídrico tem relação inversamente proporcional ao OD, uma vez que o consumo deste elemento é afetado pelo processo de nitrificação, do qual este composto de Nitrogênio também faz parte.

## 2.2 Quimiometria

De acordo com Otto (2007), o desenvolvimento da disciplina de Quimiometria está estritamente relacionado com o surgimento do computador. Na década de 70, muitos especialistas já trabalhavam com métodos matemáticos e estatísticos que eram submetidos a computadores *mainframe*<sup>2</sup> para cálculo e depois viriam a se transformar em métodos quimiométricos.

Ele define quimiometria, portanto, como:

a disciplina química que usa métodos matemáticos e estatísticos, i) para desenhar ou selecionar os mais adequados procedimentos e experimentos e ii) para prover a máxima informação química através da análise de dados químicos.

Portanto, a partir da análise de dados químicos, advindos dos mais variados tipos de análises de laboratório, objetiva-se obter informações relevantes para tomada de decisão. Para tanto, é também importante que os dados sejam pré-processados antes de qualquer tipo de análise exploratória, buscando remover fontes de variação indesejáveis que não são removidas naturalmente durante a análise de dados. Alguns métodos podem ser destacados:

- ❑ Centragem dos dados na média;
- ❑ Escalonamento pela variância;
- ❑ Auto-escalonamento;
- ❑ Escalonamento pela amplitude.

<sup>2</sup> Grandes computadores conectados a vários terminais que davam acesso aos seus recursos.

A presença de amostras anômalas (*outliers*), que tenham comportamento diferente das demais amostras, podem interferir no resultado.

Outra ferramenta importante para buscar informações nos dados é a verificação da presença de correlações, ou seja, explicações para o nível de associação entre duas variáveis. Dentre as técnicas existentes para este fim, as mais utilizadas são as correlações de *Pearson* e a de *Spearman*. Para análise exploratória, que têm por objetivo extrair o máximo de informações possíveis dos resultados, pode-se utilizar: Análise de Componentes Principais (ACP), Análise Hierárquica de Agrupamentos (HCA) e Análise de Fatores (FA) (BAIO, 2009).

### 2.2.1 Análise dos Componentes Principais

SCHIMIDT et al. (2008) definem a Análise de Componentes Principais (ACP), do termo inglês *Principal Component Analysis* (PCA), como a ferramenta multivariada de análise de componentes principais que consiste na “mudança de base dos dados multivariados, representados pela matriz  $X$ ”, contendo os dados das amostras. A ideia principal é a de transformar o espaço multidimensional em subespaço menor, ou seja, aproximar a matriz  $X$  pelo produto de duas matrizes pequenas - as matrizes *score* e *loading*, de acordo com a Expressão 1.

$$\mathbf{X} = \mathbf{T} \mathbf{L}^T \quad (1)$$

onde  $X$  é a matriz de dados originais, consistindo de  $n$  linhas (amostras) e  $p$  colunas (variáveis);  $T$  é a matriz de *scores* com  $n$  linhas e  $d$  colunas número de componentes principais (CP);  $L$  é a matriz de *loadings* com  $d$  colunas e  $p$  linhas; e  $T$  é a transposta da matriz de *loadings*, gerada pelo PCA, que troca as linhas pelas colunas, como presente na Expressão 1 e Figura 4.

O diagrama mostra a equação  $\mathbf{X} = \mathbf{T} \mathbf{L}^T$  representada por retângulos. O retângulo  $\mathbf{X}$  tem altura  $n$  e largura  $p$ . O retângulo  $\mathbf{T}$  tem altura  $n$  e largura  $d$ . O retângulo  $\mathbf{L}^T$  tem altura  $d$  e largura  $p$ . Um sinal de igualdade está entre  $\mathbf{X}$  e  $\mathbf{T}$ .

Figura 4 – Decomposição da matriz de dados originais para PCA

A Expressão 2 exhibe as combinações lineares das variáveis originais que formam as novas coordenadas.

$$X = t_1 l'_1 + t_2 l'_2 + \dots + t_d l'_d + E \quad (2)$$

Os CP são determinados de acordo com o critério de variância - diferentemente do modelo original dos dados - onde o primeiro CP concentra a maioria da variância do modelo,

o segundo CP mais que o terceiro, e assim sucessivamente (OTTO, 2007). Ou seja, “ $X$  (multivariável) representada por muitos vetores, passa a ter as informações relevantes, representadas por um ou mais autovetores. Este subsistema deve reter as informações relevantes e representar, de modo mais adequado, cada variável individualmente.” (SCHMIDT et al., 2008) ainda reforçam que:

“PCA descreve, em termos geométricos das componentes principais, a covariância das variáveis usando o menor número de autovetores. O novo subsistema consiste, então, dos eixos das componentes principais, sendo as coordenadas das amostras nesses novos eixos denominados de *scores*. Como cada componente principal é uma combinação linear das variáveis originais, os coeficientes dessas variáveis (ou seja, os autovetores) são chamados de *loadings*.”

Uma das propostas deste trabalho é verificar se a matriz de componentes principais gerada através da técnica de PCA pode ser um dos substitutos eficientes no treinamento da rede neural, uma vez que com número menor de variáveis a RNA necessitaria de menos neurônios na camada de entrada e poderia exigir menor custo computacional. Assim, a técnica será utilizada na preparação de dados para treinamento da rede, para fins de comparação.

## 2.3 Legislação Formal e Material

O Brasil tem uma legislação extensa, no que tange à questão da água, seus usos e tratamentos. O **Sistema de avaliação de qualidade de água, saúde e saneamento**, é alicerçado em Leis, Decretos, Portarias e Resoluções Federais, que legislam em todas as áreas globais envolvendo saúde e saneamento. De acordo com (BRASIL, 2010), é composta por um extenso conjunto de letras, entre leis, decretos, portarias e resoluções federais. As principais utilizadas neste trabalho são:

### □ MINISTÉRIO DA SAÚDE

- Portaria nº. 2.914, de 12 de dezembro de 2011 - Dispõe sobre os procedimentos de controle e de vigilância da qualidade da água para consumo humano e seu padrão de potabilidade;

### □ CONAMA

- Resolução CONAMA nº. 357, de 17 de março de 2005 - Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes, e dá outras providências;

- Resolução CONAMA n.º. 430, de 13 de maio de 2011 - Dispõe sobre as condições e padrões de lançamento de efluentes, complementa e altera (mas não revoga) a Resolução n.º. 357 do CONAMA.

A resolução normativa CONAMA n.º. 357, publicada em 17 de março de 2005, faz parte do marco regulatório brasileiro que estabelece o enquadramento e classificação das águas doces, salinas e salobras, assim como os componentes físico-químicos e a variação destes com o intuito de criar níveis de qualidade de água associados com seus respectivos usos. Este texto normativo ampliou e revogou a resolução CONAMA n.º. 20, de 1986, que tinha a mesma função.

Nesse sentido, em seu Capítulo I, intitulado “Das Definições”, o texto cria o entendimento que define os corpos de água com base em sua salinidade, como verifica-se no Art. 2º, item I:

“CAPÍTULO I  
DAS DEFINIÇÕES

Art. 2º.

Para efeito desta Resolução são adotadas as seguintes definições:

I - águas doces: águas com salinidade igual ou inferior a 0,5 ‰;

II - águas salobras: águas com salinidade superior a 0,5 ‰ e inferior a 30 ‰;

III - águas salinas: águas com salinidade igual ou superior a 30 ‰;”

De acordo com o texto da Resolução, o Rio Meia Ponte e o Rio Paraíba do Sul se enquadram como rios de Classe 2, de água doce. Já no Capítulo II, intitulado, “Da classificação dos Corpos de Água”, os rios de água doce são classificados como águas destinadas a vários outros usos, dentre os quais o abastecimento público após tratamento, conforme se vê em sua Seção I, Art. 4º:

“CAPÍTULO II  
DA CLASSIFICAÇÃO DOS CORPOS DE ÁGUA

Seção I

Das Águas Doces

Art. 4º.

As águas doces são classificadas em:

III - classe 2: águas que podem ser destinadas:

a) ao abastecimento para consumo humano, após tratamento convencional;

b) à proteção das comunidades aquáticas;

c) à recreação de contato primário, tais como natação, esqui aquático e mergulho, conforme Resolução CONAMA n.º. 274, de 2000;

d) à irrigação de hortaliças, plantas frutíferas e de parques, jardins, campos de esporte e lazer, com os quais o público possa vir a ter contato direto; e

e) à aquicultura e à atividade de pesca.”

Assim, uma vez definidas e classificadas as águas dos rios que fazem parte deste trabalho, resta verificar a tabela de parâmetros orgânicos e inorgânicos, utilizados para quantificar os valores máximos destes parâmetros de forma que os mananciais se enquadrem na

classificação normativa. Conforme o Art. 15 do Capítulo II, observadas algumas exceções, a tabela que determina os parâmetros para os mananciais de classe 1, também se aplica aos mananciais de classe 2, conforme a Tabela 3. As únicas variáveis utilizadas neste trabalho e que não constam nesta tabela são a Temperatura da água e a Condutividade.

Tabela 3 – Parâmetros orgânicos e inorgânicos segundo a resolução CONAMA 357 para rios de classe 2.

<b>Resolução CONAMA 357</b>	
Classe 2 - Águas doce	
Parâmetros	Valor Máximo
pH	6,0 a 9,0
Demanda Bioquímica de Oxigênio	5 dias a 20 °C até 5 mg/L O <sub>2</sub>
Nitrogênio Amoniacal	3,7 mg/L N, para pH ≤ 7,5 2,0 mg/L N, para 7,5 < pH ≤ 8,0 1,0 mg/L N, para 8,0 < pH ≤ 8,5 0,5 mg/L N, para pH < 8,5
Nitrato	10,0 mg/L N
Nitrito	1,0 mg/L N
Fósforo	a) até 0,030 mg/L, em ambientes lânticos; e, b) até 0,050 mg/L, em ambientes intermediários
Turbidez	até 100 unidades nefelométrica de turbidez (UNT)
Sólidos Dissolvidos Totais	500 mg/L
Oxigênio Dissolvido	não inferior a 5 mg/L O <sub>2</sub>

Uma observação importante se faz necessária a respeito da resolução CONAMA N°. 430, de 13 de maio de 2011; este novo texto altera e complementa a resolução n°. 357 de 2005, especificamente no que tange à criação de dois padrões de emissão de efluentes, um destinado a efluentes de modo geral e outro para esgoto sanitário. No entanto, a norma não muda as classificações ou valores empregados para os enquadramentos dos mananciais (CONAMA, 2011).

## 2.4 Redes Neurais Artificiais

Para HAYKIN (2001), RNA são sistemas de processamento da informação que executam suas tarefas com técnicas que tentam se assemelhar com o funcionamento da mente humana. Uma definição mais qualificada é dada por BRAGA et al. (2000), que as definem como modelos matemáticos que se assemelham às estruturas neurais biológicas e que têm capacidade computacional adquirida por meio de aprendizado e generalização. Seu aprendizado está definido como o processo pelos quais os parâmetros da rede são estimulados e adaptados pelo ambiente onde a rede está inserida. Este processo de adaptação às características particulares de determinado problema tem o objetivo de levar as RNA a capacidade de generalização, ou seja, gradativamente dar respostas coerentes para dados não apresentados à rede durante o treinamento.



RODRIGUES et al. (2007) completam:

“A RNA procura construir um computador (ou um programa) que tenha um comportamento similar ao dos circuitos cerebrais, o que implica em que espera-se ver um comportamento inteligente emergindo, aprendendo novas tarefas, errando, fazendo generalizações e descobertas.”

As RNA se diferenciam pela arquitetura e pela forma como os pesos associados às conexões são ajustados durante o processo de aprendizado. A arquitetura é definida pelo número de camadas (camada única ou múltiplas camadas), pelo número de nós em cada camada, tipo de conexão entre os nós e por sua topologia. Elas podem ser utilizadas para variedade de propósitos, dentre os quais se destacam a classificação, controle, modelagem, predição e previsão (BAPTISTA; MORGADO-DIAS, 2013).

### 2.4.1 Redes Neurais Multicamadas

As redes multicamadas, também conhecidas como MLP, são definidas como um modelo de rede que mapeia um conjunto de entradas em uma ou mais saídas, tendo a capacidade de aprender e obter respostas para problemas que não apresentam separabilidade linear (HAYKIN, 2001) e são exemplificadas de maneira genérica na Figura 5, onde é possível destacar dois tipos de arquitetura; RNA com apenas uma camada intermediária serão representadas na figura como tendo apenas a primeira camada escondida, além das demais. As redes com duas camadas escondidas seguem o modelo da figura, onde verifica-se a camada de entrada dos dados, a primeira e segunda camada escondidas, além da camada de saída da RNA.

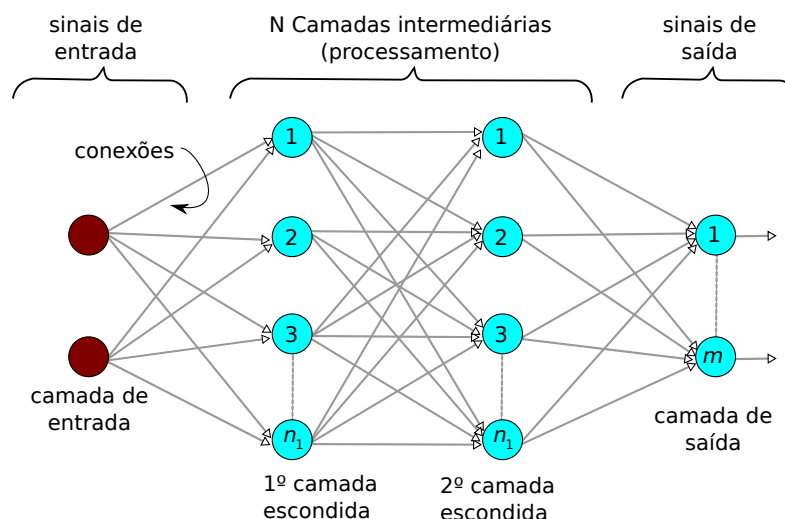


Figura 5 – Estrutura genérica de RNA multicamadas de arquitetura Entrada-Processamento-Saída

Assim, as RNA possuem arquiteturas (a maneira como os vários neurônios na rede são conectados) extremamente paralelas, operando com algoritmos voltados ao processo

de aprendizagem e necessitando de alta capacidade de processamento (os algoritmos de aprendizagem geralmente exigem computadores de última geração para maximizar sua performance) (LAWRENCE et al., 1996)(BRAGA et al., 2000).

Dentre suas principais aplicações, pode-se citar o aprendizado e os reconhecimentos, associação e classificação de padrões. As RNA são extremamente tolerantes a falhas, possuem alta capacidade de auto-adaptação e podem gerenciar aplicações que necessitam de atualizações ou respostas em tempo real. A principal diferença entre a computação *convencional* e as RNA reside na sua capacidade de resolver problemas práticos sem a necessidade da definição de listas de regras ou de modelos precisos.

## 2.4.2 Neurônios Biológicos

O funcionamento do cérebro humano há muito tempo é tema de inúmeras linhas de pesquisa em diversas áreas. As observações deste órgão fundamental ao funcionamento do sistema nervoso animal, renderam importantes descobertas em variadas áreas do conhecimento. Seus 10 bilhões de neurônios e sua capacidade de 60 trilhões de sinapses o tornam o mais eficiente sistema de processamento não-linear existente, com impressionante capacidade de execução eficiente de tarefas em paralelo. Mesmo com tanta capacidade, o cérebro humano possui eficiência energética de  $10^{-16}$  joules por operação por segundo, se comparado aos computadores atuais ( $10^{-6}$  joules por operação por segundo) (HAYKIN, 2001).

A Figura 6 ilustra a anatomia de um neurônio biológico. Nela podem ser verificados os três componentes de particular interesse para a compreensão do neurônio biológico: dendritos, soma e axônios. Os muitos “Dendritos” que um neurônio possui são responsáveis por receber o sinal de outros neurônios. FAUSETT (1994) destaca que os sinais são impulsos elétricos que são transmitidos através do “gap sináptico” por meio de processo químico, os neurotransmissores. A “Soma”, ou “Corpo celular”, destaca-se como região responsável pelo *processamento* das informações, pela soma dos sinais recebidos; os “Axônios” promovem a saída da informação neural para a próxima célula, que a receberá através dos “Dendritos”. Os neurônios se agrupam no cérebro biológico e se comunicam entre si, para proporcionar aprendizado, memória, raciocínio. A comunicação intraneural se dá através de sinais químicos, denominados Pré-sinápticos. No entanto, quando necessitam estabelecer sua *rede* de comunicação, se utilizam de sinais elétricos, também denominados de sinais Pós-sinápticos (BRAGA et al., 2000)(HAYKIN, 2001).

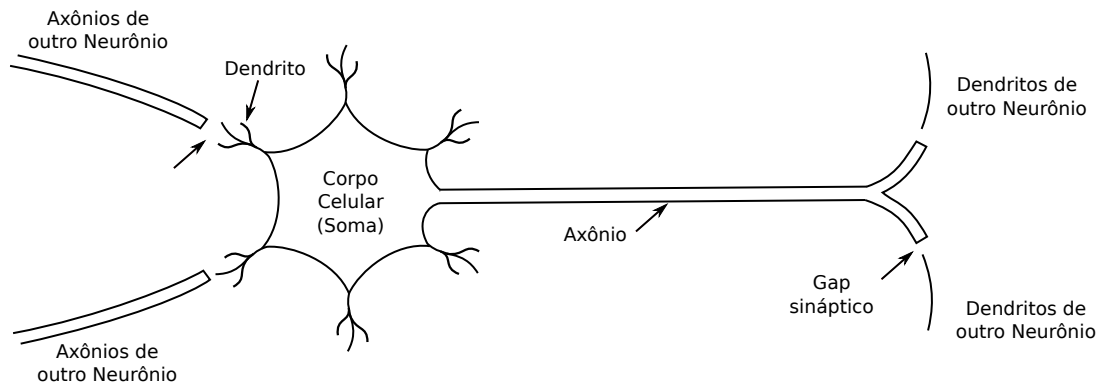


Figura 6 – Modelo de neurônio biológico.

### 2.4.3 Neurônios Artificiais

As primeiras pesquisas com neurônios e Redes Neurais Artificiais (RNA) apareceram em função da observação de suas contrapartes biológicas no início do século passado.

Para fins de definição, divide-se as RNA em:

1. neurônios (elementos de processamento);
2. conexão;
3. padrões: dados de entrada da rede;
4. funções: modelos matemáticos utilizados no reconhecimento de padrões.

De acordo com HAYKIN (2001) e FAUSETT (1994) os neurônios artificiais implementam modelos de axônios e dendritos de sua contraparte biológica através de conexões e as sinapses utilizando ponderações ou pesos de ajuste. A Figura 7 exhibe o modelo inicial para um neurônio artificial, proposto por MACCULLOCH e PITTS (1943) e aprimorado por ROSENBLATT (1958), que o nomeou de Perceptron.

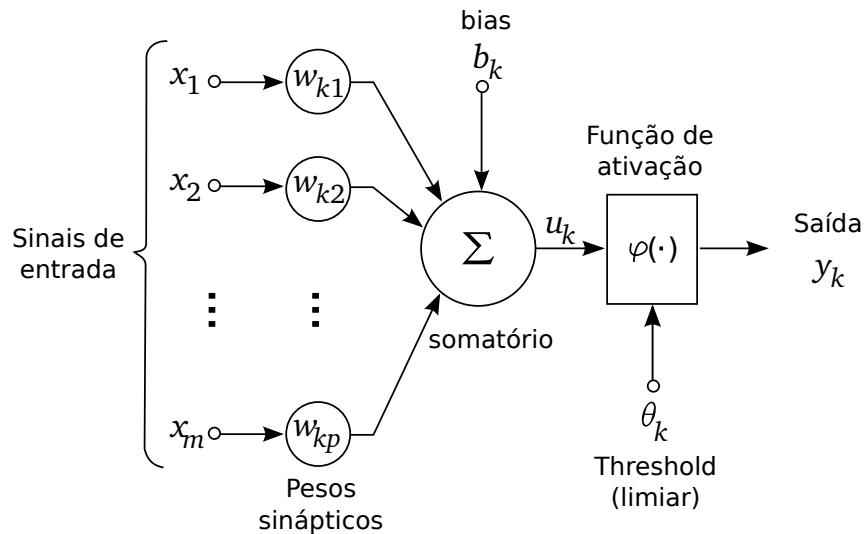


Figura 7 – Modelo de Neurônio Artificial, segundo McCulloch-Pits.

O modelo de neurônio artificial de McCulloch-Pits (1943) implementa:

- $x_m$  são sinais de entrada da rede;
- $w_{km}$  são pesos ou pesos sinápticos, associados a cada entrada;
- $b_k$  é o termo bias;
- $u_k$  é a combinação linear dos sinais de entrada;
- Função de somatório;
- $\varphi(\cdot)$  é função de ativação;
- $y_k$  é a saída (única, em direção a outros neurônios).

Cada neurônio artificial possui número de entradas e um único sinal de saída que se propaga através das conexões sinápticas com outros neurônios (COSTA, 2006). Os pesos são os parâmetros ajustáveis da rede e devem ser reajustados à medida que o conjunto de treinamento é apresentado à rede. O processo de aprendizado supervisionado da RNA resulta no ajuste dos pesos de tal forma que o valor de saída (estimado) da rede seja o mais próximo possível do valor desejado. O modelo de neurônio também inclui o termo “bias”, que é aplicado externamente e serve para aumentar os graus de liberdade, permitindo melhor adaptação, por parte da RNA, ao conhecimento fornecido a ela (SOARES; SILVA, 2011).

#### 2.4.4 Definição da Arquitetura da RNA

A configuração da rede neural depende da particularidade do problema que se quer resolver (ZURADA, 1995), restringindo sua aplicação para outros problemas sem a devida

reconfiguração. A rede precisa levar em conta se os dados são lineares ou não-lineares. Alguns dos parâmetros que modificam a configuração de uma RNA são: i) o número de neurônios da camada de entrada; ii) o número de camadas intermediárias, situadas entre a camada de entrada e a de saída; iii) número de neurônios de cada camada intermediária (SANTOS; FILHO, 2003)(LAWRENCE et al., 1996)(STATHAKIS, 2009). O resultado desta definição chama-se arquitetura, ou topologia, que define a maneira com que os vários neurônios constituintes da rede neural são interconectados (GOMES et al., 2007).

De acordo com STATHAKIS (2009), a melhor topologia é julgada pela capacidade de generalizar dados não “vistos” pela rede. Desta forma, em uma primeira fase, o processo de treinamento é conduzido com o objetivo de apresentar um padrão à RNA. Durante a execução desta fase, a cada novo ciclo, o algoritmo de retropropagação de erro corrige os pesos sinápticos de forma que o erro de saída da rede seja o mais próximo possível do erro configurado nas variáveis de treinamento. Uma vez concluído o treinamento e alcançado o erro mínimo esperado, uma nova fase é iniciada, através da qual um novo conjunto de dados, não apresentados à RNA durante a fase de treinamento, são utilizados para a verificação do aprendizado, ou seja, da capacidade de generalização da rede através de dados novos. A RNA deve apresentar resposta (erro de verificação) compatível com o nível de erro obtido durante a fase de treinamento.

Os ajustes de topologia definem a facilidade que a rede terá de concluir o treinamento com o nível de erro esperado, a partir dos parâmetros de configuração escolhidos. Topologias ruins (número incorreto de camadas ou neurônios nestas camadas) podem conduzir a dificuldades no aprendizado e posterior generalização, além de desnecessário aumento do esforço computacional (STATHAKIS, 2009)(FANUCCHI et al., 2014). Este elemento é especialmente importante para o trabalho com grandes redes, configuradas com centenas de neurônios de entrada e milhares de amostras de dados para treinamento. Neste caso, a topologia define a economia computacional e pode economizar horas de treinamento, mesmo em computadores de última geração.

Na grande maioria dos casos, redes com poucos neurônios (também denominados nós) na camada intermediária e poucas camadas oferecem melhores resultados. No entanto, mesmo ajustes para estes casos são particulares e muito esforço e testes são necessários para verificar a melhor solução.

De acordo com Soares e Silva (2011), alguns pesquisadores afirmam que com uma camada intermediária já é possível calcular uma função arbitrária qualquer, a partir de dados fornecidos. Esta camada deve ter  $(2i+1)$  neurônios, onde  $i$  é o número de variáveis de entrada. Outros defendem a necessidade de duas camadas intermediárias, suficientes para aproximar qualquer função matemática. Neste caso, a segunda camada teria o dobro dos neurônios da camada de saída.

Assim, STATHAKIS (2009) sugere alguns métodos e equações para ajudar a definição da quantidade máxima de nós e camadas intermediárias em busca do menor esforço

computacional no treinamento.

#### 2.4.4.1 Cálculo para redes de uma camada intermediária

Para rede com dez entradas, uma saída e apenas uma camada intermediária, utiliza-se a Expressão 3 para definir o número máximo de nós da camada intermediária, onde  $N$  = entrada e  $m$  = saída da rede. No caso da rede com  $N = 10$  e  $m = 1$ , o resultado da equação é 10,9544, ou 11 nós máximos na camada intermediária.

$$nós^{máx} = 2 \cdot \sqrt{(m + 2)N} \quad (3)$$

#### 2.4.4.2 Cálculo para redes de duas camadas intermediárias

Para rede com dez entradas, uma saída e duas camadas intermediárias, as expressões a seguir podem ser utilizadas para definição do número máximo de nós da primeira e segunda camadas intermediárias, onde  $N$  = número de parâmetros de entrada e  $m$  = valor de saída da rede. Portanto, para  $N = 10$  e  $m = 1$ .

A Expressão (4) define o número de nós da camada 1. Para a rede de arquitetura 10- $x$ -1, o resultado é 9,1287, ou 9 nós máximos na primeira camada.

$$nós^{máx}(camada1) = \sqrt{(m + 2) \cdot N} + 2 \cdot \sqrt{\left(\frac{N}{m + 2}\right)} \quad (4)$$

A Expressão (5) define o número de nós da camada 2. Para a mesma rede 10- $x$ -1 o resultado é 1,8257, ou 2 nós máximos nesta camada.

$$nós^{máx}(camada2) = m \cdot \sqrt{\left(\frac{N}{m + 2}\right)} \quad (5)$$

#### 2.4.4.3 Testes

STATHAKIS (2009) ainda sugere que a utilização deste ferramental não é consenso na literatura. Em muitos casos existe a necessidade de testes com outras topologias para assegurar o melhor desempenho e generalização. Métodos heurísticos também são utilizados com o objetivo de determinar a expressão que consiga relacionar de forma mais eficiente o número de entradas da rede com o de saída e aproximar melhor a topologia. Outra vertente é a utilização de algoritmos de construção e poda. Os primeiros visam à construção gradual da rede por meio da adição de neurônios, até conseguir a arquitetura que assegure o melhor treinamento e generalização. Os algoritmos de poda, por sua vez, partem do caminho oposto, inicializando a rede neural com número razoável de neurônios e cortando aqueles que têm pouca influência no erro.

## 2.4.5 Paradigmas de Aprendizagem

BRAGA et al. (2000 apud SOARES; SILVA, 2011) definem aprendizagem como o processo pelo qual os parâmetros livres da rede (pesos sinápticos e bias) são adaptados através do processo de estimulação pelo ambiente onde a mesma está inserida. Esta característica é fundamental à RNA, pois concede a este programa de computador a capacidade de aprender do ambiente e de melhorar seu desempenho através da aprendizagem. São dois os métodos principais para o treinamento de redes neurais: o aprendizado supervisionado e o não supervisionado.

### 2.4.5.1 Aprendizado Supervisionado

Soares e Silva (2011) definem aprendizado supervisionado como o método em que o “professor” detêm todo o conhecimento sobre o ambiente, representado pelo conjunto de exemplos de entrada-saída. O objetivo deste método é fazer com que a RNA, que desconhece o ambiente, se iguale ao professor. Assim, quando a rede neural inicia o treinamento e está utilizando este paradigma de aprendizagem, conta com o auxílio deste professor. Logo que o vetor de entrada é apresentado, o professor orienta sobre qual a saída ótima para aquele vetor de treinamento. Esta resposta, que representa o conhecimento disponível do professor sobre o ambiente, orienta o ajuste dos pesos sinápticos através da combinação do vetor de treinamento e do sinal de erro indicado pelo professor. Este ajuste é realizado passo a passo e interativamente, até que o sinal de saída da rede seja parecido com o indicado. Uma vez alcançado este objetivo, a rede neural continua seu trabalho sem a assistência do professor. O diagrama ilustrativo deste processo, adaptado de Soares e Silva (2011), pode ser observado na Figura 8, que ilustra o ambiente, ao qual apenas o professor tem acesso e sobre o qual tem pleno conhecimento, através dos exemplos de entrada-saída. Este professor informa à rede neural a resposta desejada. Após iniciado o processo de treinamento, a saída da rede se torna a resposta calculada, e a cada novo ciclo, esta saída deve gerar um sinal de erro que, através do sistema de aprendizagem, fará os ajustes nos parâmetros de aprendizagem para que a resposta calculada se torne cada vez mais próxima do exemplo do professor, ou seja, da resposta desejada.

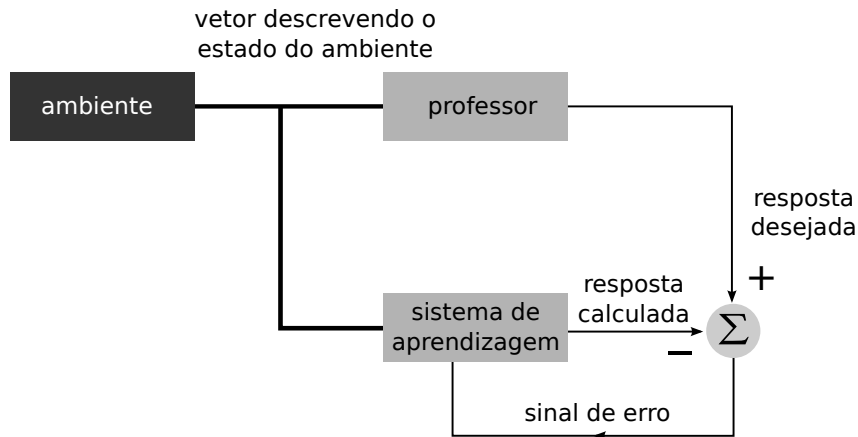


Figura 8 – Diagrama de blocos de aprendizagem com professor.

Este método de aprendizado é utilizado pelo algoritmo de retropropagação de erro e portanto para este trabalho. Neste sentido, durante a apresentação de cada vetor de entrada na rede, são introduzidos os valores dos dez parâmetros de entrada (Temperatura da água, pH, DBO, Condutividade, Nitrogênio Amoniacal, Nitrato, Nitrito, Fósforo, Turbidez, Sólidos Dissolvidos Totais) juntamente com o valor do parâmetro de saída (OD), que representam todo o conhecimento do professor sobre o ambiente (dados de qualidade de água do rio). O valor OD representa a saída ótima para a rede, que a cada iteração do algoritmo de retropropagação, recebe instruções do processo de aprendizado supervisionado de como ajustar os pesos sinápticos de modo a buscar aproximar a resposta (saída) da rede àquela exibida como ótima.

Como medida de desempenho da rede é utilizado o erro médio quadrático ou a soma de erros quadráticos sobre a amostra de treinamento. A Expressão 6 representa o sinal de erro deste sistema de aprendizado (a diferença entre a resposta desejada,  $d$ , e a saída da rede, ou resposta observada  $z$ ):

$$e(n) = d(n) - z(n) \quad (6)$$

#### 2.4.5.2 Aprendizado Não-Supervisionado

A aprendizagem não supervisionada é definida de modo geral pelo estado em que não há o “professor” externo para supervisionar o processo de aprendizado e a rede deve se “familiarizar” às regularidades estatísticas dos dados de entrada. Silva (2008) reforça que, nestes casos, a rede neural deve, por si só, decidir quais são as classes pertinentes, simplesmente agrupando padrões semelhantes.



## 2.4.6 Algoritmo de Retropropagação de Erro

O algoritmo de retropropagação de erro (*backpropagation*) é um algoritmo de treinamento que opera em redes neurais Multicamadas (MLP) e se comporta bem com problemas/dados não-lineares, motivo pelo qual foi escolhido para este trabalho.

O treinamento da rede neural multicamadas por meio deste algoritmo envolve três etapas: a propagação do sinal de entrada pela rede (para frente), a retropropagação do erro associado (para trás) e o ajuste dos pesos sinápticos, conforme ilustra a Figura 9, adaptada de (SILVA et al., 2010).

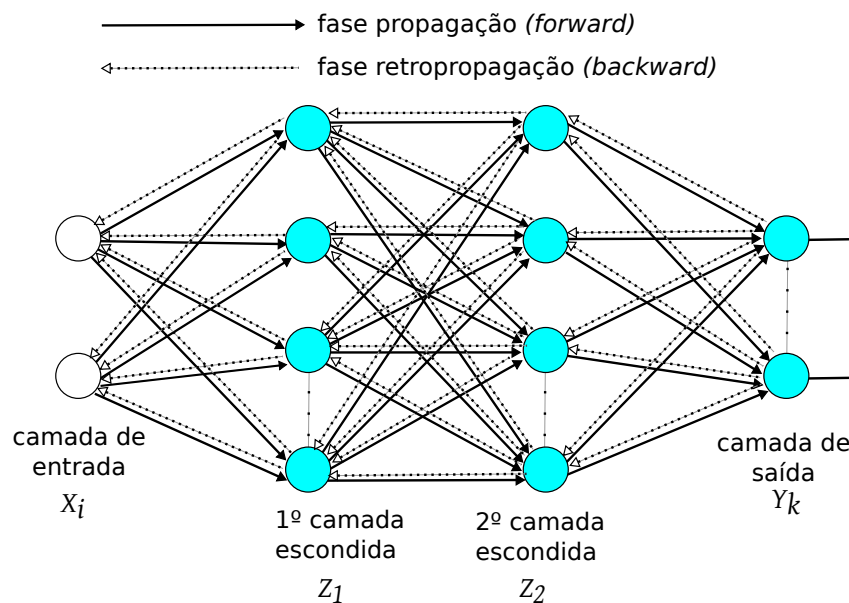


Figura 9 – Esquema gráfico do algoritmo de retropropagação de erro.

Durante o treinamento, cada unidade de entrada  $X_i$  recebe um sinal que é difundido para cada unidade escondida  $Z_1, \dots, Z_p$ . Cada unidade calcula seu sinal de ativação e envia um sinal  $z_i$  para cada unidade de saída  $Y_k$ . Estas unidades de saída, então, calculam seus sinais de ativação  $y_k$ , que correspondem à resposta da rede para o padrão de entrada dado.

Durante o treinamento, cada saída  $y_k$  determinada pela rede é comparada ao valor alvo (*target*)  $t_k$ , para determinar o erro associado ao padrão de entrada a cada unidade. Com base no fator de erro  $\delta_k$  com  $k \in \mathbb{N}$  com  $k = 1, 2, \dots, m$  é calculado e é usado para distribuir o erro das unidades escondidas  $Z_j$  com  $j \in \mathbb{N}$  com  $j = 1, 2, \dots, m$ ) conectadas a  $Y_k$ .

A função de ativação utilizada na rede treinada com algoritmo *backpropagation* dever ser contínua e diferenciável em todos os seus pontos. Quando se treina padrões cujos *targets* estão no intervalo  $[-1, 1]$ , a função *sigmoid* (Figura 10) pode ser usada:

$$f(x) = \frac{2}{1 + e^{-x}} - 1 \quad (7)$$

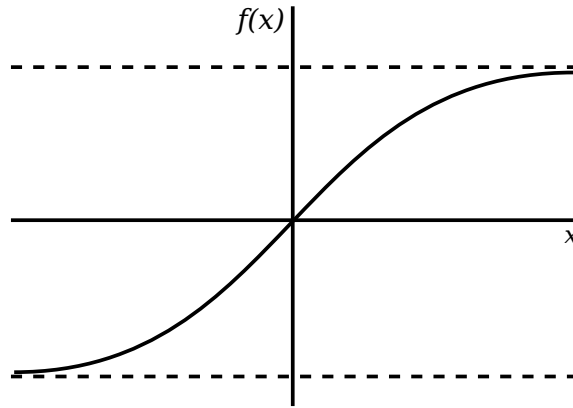


Figura 10 – Sigmoid bipolar, no intervalo  $[-1, 1]$ .

Sua derivada é dada por:

$$f'(x) = \frac{1}{2}[1 + f(x)][1 - f(x)] \quad (8)$$

De acordo com SCHIMIDT et al. (2002) a Equação 9 resume o cálculo da correção de pesos deste algoritmo de treinamento:

$$\Delta w^k = \eta \cdot D + m \cdot \Delta w^{k-1} \quad (9)$$

Onde:

$\Delta w^k$  representa a diferença entre os pesos inicial e final para a iteração  $k$ ;

$\eta$  é chamado de “taxa de aprendizagem”;

$D$  é a derivada do erro calculado através da função de transferência;

$m$  é chamado de “momento” e evita a possibilidade de se encontrar um mínimo local;

$\Delta w^{k-1}$  representa a diferença entre os pesos inicial e final para a iteração  $k - 1$ .

O algoritmo disposto no Apêndice A apresenta o procedimento de treino da rede neural multicamada por meio da metodologia *backpropagation* (FAUSETT, 1994).

#### 2.4.6.1 Formas de inicialização dos pesos

Algumas considerações devem ser realizadas a respeito deste algoritmo. Os seus primeiros passos consiste na inicialização dos pesos e bias dos neurônios. O processo de atualização dos pesos é dependente do valor que a função de ativação e sua derivada assumem durante o treinamento. Por esta razão, é importante que os pesos iniciais sejam tais que não se tenham valores nulos como respostas a estas funções. Escolher valores iniciais aleatórios e pequenos (entre  $-0,5$  e  $0,5$ , por exemplo) é considerada boa estratégia.

### 2.4.6.2 Quanto treinar a rede

No que se refere ao critério de parada do algoritmo, de acordo com HAYKIN (2001), não se pode demonstrar que o algoritmo de retropropagação convergiu e não existem critérios bem definidos para encerrar a sua operação. A prática comum é considerar que sua convergência tenha ocorrido quando o valor absoluto do Erro Quadrático Médio (MSE) estiver suficientemente pequeno.

Entretanto, ao treinar a rede por meio do algoritmo *backpropagation*, é necessário ter em mente que dois objetivos são buscados: i) fazer com que a rede aprenda determinado padrão; ii) fazer com que a rede tenha capacidade de generalização, de forma a dar respostas adequadas para novos padrões de entrada. Neste sentido, continuar o treinamento da rede até que o erro quadrático médio seja mínimo pode não ser vantajoso, uma vez que ela pode ficar especializada a tal ponto que perca a sensibilidade para novos padrões de entrada.

O padrão de parada encontrado na literatura consiste em dividir o conjunto de treinamento da rede em dois subconjuntos disjuntos  $A$  e  $B$ ; convencionou-se distribuir 70% dos dados para o subconjunto  $A$  e 30% para o subconjunto  $B$ . Os valores de  $A$  são utilizados no treinamento da rede e, regularmente, a rede é testada para elementos do conjunto  $B$ , fazendo neste momento o cálculo do erro. No momento em que a rede for capaz de reconhecer padrões de ambos os conjuntos com baixa taxa de erro, o treinamento terá acabado.

### 2.4.6.3 O número de camadas escondidas

O treinamento da rede com mais de uma camada escondida pode ser conseguido por meio de pequenas modificações no algoritmo disposto no Apêndice A, num processo similar de retropropagação dos erros  $\delta$ .

Entretanto, resultados teóricos mostram que uma camada escondida é o suficiente para a rede do tipo *backpropagation* aproximar qualquer mapeamento contínuo de padrões de entrada para padrões de saída (FAUSETT, 1994). Por este motivo todas as RNA deste trabalho foram treinadas utilizando-se uma camada intermediária entre a camada de entrada e a de saída da rede.

### 2.4.6.4 Momento ( $\mu$ )

A atualização dos pesos no treinamento da rede com algoritmo *backpropagation* com momento é feita numa direção dada pela combinação do gradiente anterior com o gradiente atual. Esta modificação é vantajosa quando alguns dados do treinamento são muito diferentes da maioria. É desejável usar taxa de aprendizado pequena para impedir grande perturbação na direção de decréscimo do gradiente quando um padrão muito incomum for

apresentado. Porém, é também preferível manter o ritmo rápido no treinamento quando os padrões de treinamento forem relativamente similares.

A convergência pode se tornar mais rápida se o termo “momento” for adicionado na fórmula de atualização de pesos. Para que o momento possa ser utilizado, é necessário ter um ou mais valores de pesos atualizados nas etapas anteriores. Por exemplo, na mais simples implementação do *backpropagation* com momento, a nova atualização de pesos no passo  $t + 1$  é feita tendo por base os valores dos pesos na etapa  $t$  e  $t - 1$ .

O momento permite que a rede faça ajustes razoavelmente grandes de pesos na direção geral de crescimento da maioria dos padrões, ao mesmo tempo que a utilização de taxa de aprendizado pequena previne o efeito de grandes erros vindos de qualquer padrão de treinamento. Ele também reduz a probabilidade da rede encontrar pesos que são de mínimo local, mas não global. Quando utiliza-se o momento, a rede não segue na direção do gradiente e sim na direção da combinação do gradiente atual com vários outros gradientes calculados anteriormente nas correções dos pesos, isto promove a corrida em direção aos pesos mínimos. Ver o Apêndice B.

### 2.4.7 Desafio - melhorar a generalização da RNA

ELSHORBAGY et al. (2000) e BORSATO et al. (2009) confirmam ser possível treinar RNA com pequenos bancos de dados. No entanto, é comum a utilização de grandes bases de amostras para treinamento e verificação da rede, com a finalidade de melhorar a generalização da mesma.

Métodos numéricos como interpolação de dados, ajuste de dados e transformada de Fourier podem ser utilizados com o objetivo de incrementar o número de amostras de uma base de dados existente (WECKMAN et al., 2010)(SILVERT; BAPTIST, 1998).

A interpolação de dados é definida como o método que permite a obtenção de um conjunto de dados novos a partir de um conjunto previamente conhecido. De acordo com Urroz (2001), “a interpolação simples linear é tipicamente aplicada a uma tabela de valores  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  de modo a obter o valor  $y_k$  correspondente ao valor  $x_k$  dado, localizado dentro da faixa de valores de  $x$ .”

É importante dizer que a técnica de interpolação linear pode ser um procedimento perigoso se não adotado corretamente. Quando se utiliza tal técnica, é importante separar os dados das variáveis e aplicar a técnica uma variável por vez, não estabelecendo uma relação entre elas, mas apenas com a própria variável.

---

## Metodologia

Nesta seção serão detalhados os aspectos metodológicos concernentes ao presente trabalho. Aspectos como a origem dos dados utilizados e os recursos computacionais empregados nos experimentos são detalhados.

### 3.1 Análise dos dados

#### 3.1.1 Amostras de dados

A correta escolha dos dados é importante para a representação dos fenômenos estudados, por refletir o comportamento das variáveis analisadas (RODRIGUES et al., 2007).

Para este trabalho são utilizados dados de duas origens distintas:

□ Rio Meia Ponte (GO):

- Ponto de amostragem<sup>1</sup> 1, localizado no município de Itauçu;
- Ponto de amostragem<sup>1</sup> 3, 4 e 5, localizados no município de Inhumas;
- Ponto de amostragem<sup>1</sup> 11, localizado no município de Goiânia.

□ Rio Paraíba do Sul (SP):

- Ponto de monitoramento<sup>2</sup> PARB 02300, localizado no município de São José dos Campos;

##### 3.1.1.1 Amostras do Rio Meia Ponte (GO)

Os dados utilizados nos experimentos resultam de análise efetuada pela Empresa de Saneamento do Estado de Goiás (SANEAGO), que retira amostras bimestralmente de pontos de coleta no Rio Meia Ponte destinados a análise de qualidade de água. A Figura 11

---

<sup>1</sup> Nomenclatura da SANEAGO.

<sup>2</sup> Nomenclatura da CETESB.

exibe a localização dos cinco pontos de coleta situados nos municípios de Itauçu, Inhumas e Goiânia. Para cada ponto, são coletadas informações de onze parâmetros distintos, num total de aproximadamente dezesseis amostras, registradas entre 07/05/2008 e 06/07/2011, cujos limites admissíveis e as unidades métricas estão de acordo com os níveis exigidos na Resolução CONAMA 357, para águas de classe 2, como pode ser observado na Tabela 3 (p. 38). Os parâmetros são:

- Temperatura da água;
- pH;
- Demanda Bioquímica de Oxigênio (DBO);
- Condutividade;
- Nitrogênio Amoniacal;
- Nitrato;
- Nitrito;
- Fósforo;
- Turbidez;
- Sólidos Dissolvidos Totais;
- Oxigênio Dissolvido (OD).

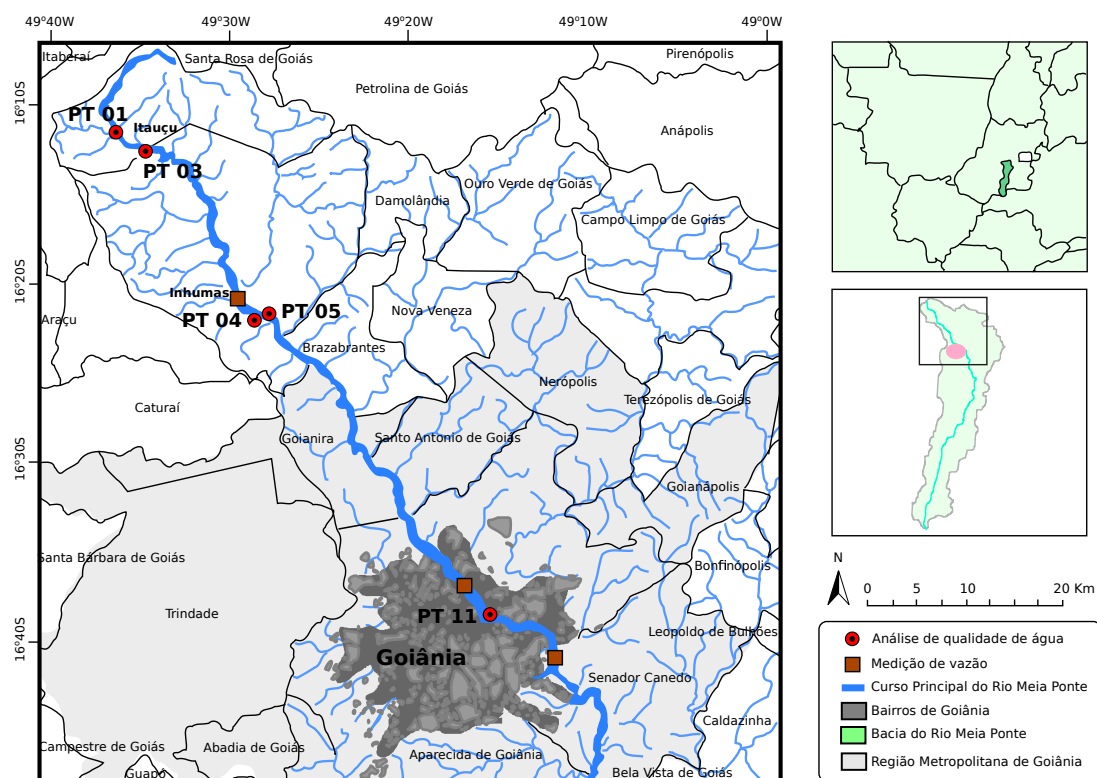


Figura 11 – Localização dos pontos de amostragem do Rio Meia Ponte utilizados nos experimentos.

A Tabela 4 dispõe o arquivo de dados do ponto de amostragem 4, situado no município de Inhumas. Como o número de amostras é relativamente pequeno, os dados são submetidos ao processo matemático de interpolação linear. Assim, uma vez separadas as variáveis (neste caso, colunas), o processo é aplicado, observando-se a relação temporal, ou seja, o processo resulta em novas amostras entre duas existentes, simulando uma maior frequência de coleta das amostras no respectivo ponto de amostragem. Esta operação, no entanto, empresta a estes conjuntos de dados caráter teórico, mesmo advindo de amostras reais.

### 3.1.1.2 Amostras do Rio Paraíba do Sul (SP)

Através do sítio na internet<sup>3</sup>, pode-se ter acesso ao sistema INFOÁGUAS, da CETESB, a partir do qual estão disponíveis dados de análises de águas dos diversos mananciais do Estado. Assim, com o objetivo de se fazer um experimento com o treinamento de redes neurais com dados reais, em menor quantidade do que os de Goiás, são reunidas informações sobre o Rio Paraíba do Sul. Os dados são oriundos do ponto de coleta PARB 02310, situado na ponte de acesso ao loteamento Urbanova no município de São José dos Campos (SP), nas coordenadas

<sup>3</sup> <http://www.cetesb.sp.gov.br/>

Tabela 4 – Amostras do ponto de coleta 4 situado no município de Inhumas - Jusante 500m do Lançamento da ETE - antes da aplicação do processo de interpolação linear.

Data Coleta	1	2	3	4	5	6	7	8	9	10	11
	Temp. Água (°C)	pH	D.B.O. (mg/L)	Condutividade (mS/cm)	Nitrogênio amoniacal (mg/L N)	Nitrato (mg/L N)	Nitrito (mg/L N)	Fósforo (mg/L)	Turbidez (UNT)	Sólidos Diss. Totais (mg/L)	O.D. (mg/L)
7/5/2008	21	7,29	2,6	94,9	0,1	0,06	0,071	0,105	21,5	52,19	6,3
2/7/2008	19	7,64	1,6	105,6	0,23	0,06	0,121	0,127	25,3	58,08	7,6
3/9/2008	22	7,72	3,6	119,6	0,26	0,22	0,284	0,101	13,4	65,78	6,8
5/11/2008	25	7,34	3,9	160,3	0,08	0,3	0,324	0,13	33,1	88,16	6,5
7/1/2009	23	7,41	2,5	101,3	0,01	0,04	0,013	0,117	52,5	55,71	6,6
13/5/2009	22	7,61	1,5	115,3	0,01	0,17	0,056	0,048	45	63,41	7,5
8/7/2009	20	7,7	1,2	112,1	0,1	0,12	0,34	0,15	15,1	61,65	7,6
2/9/2009	23	7,51	1,5	155,3	0,65	0,39	0,491	0,09	13,3	85,41	6,7
8/9/2010	21	7,69	1	108,3	0,07	0,05	0,019	0,3	13,7	59,56	7
3/11/2010	23	7,22	3,7	121,1	0,1	0,06	0,055	0,075	63,3	66,6	7,5
5/1/2011	23	6,75	3	84,6	0,18	0,02	0,116	0,004	436	46,5	5
2/3/2011	24	7,09	4,2	94,9	0,01	0,08	0,016	0,061	141	52,19	7
4/5/2011	23	7,1	1,5	96,4	0,12	0,16	0,038	0,012	389	53,02	5,7
6/7/2011	19	7,38	1,5	130,5	0,06	0,05	0,01	0,083	37,9	71,77	8,2

✓ Latitude: 23°11'45.0348"S

✓ Longitude: 45°55'49.3032"O

Buscou-se dados dos onze parâmetros estudados, coletados e analisados entre 18/01/90 e 05/12/12:

- Temperatura da água;
- pH;
- Demanda Bioquímica de Oxigênio (DBO);
- Condutividade;
- Nitrogênio Amoniacal;
- Nitrato;
- Nitrito;
- Fósforo;
- Turbidez;
- Sólidos Dissolvidos Totais;
- Oxigênio Dissolvido (OD).



### 3.1.2 Condicionamento dos Dados para Processamento

Para cada ponto de amostragem existe um arquivo de dados em texto puro (TXT)<sup>4</sup> contendo as informações de cada parâmetro físico-químico e os valores de suas leituras, oriundos das análises das amostras de água. Para minimizar problemas oriundos do uso de unidades distintas entre os parâmetros faz-se necessário que dados multivariados sejam normalizados, padronizados em uma mesma escala. A utilização de redes neurais com função de transferência sigmoideal exige que estes dados estejam normalizados no intervalo [0,1]. Esta função é que melhor resultados apresenta para dados não-lineares, por isto é escolhida para este trabalho (GOMES et al., 2007).

Uma vez que o programa *Scilab* não apresenta função interna de normalização de dados, é desenvolvido um *script* com esta finalidade, disposto no Apêndice C. As funções devem ser aplicadas a vetores de dados de cada variável. O *script* está preparado para receber a matriz contendo todas as variáveis, porém necessita que as matrizes de entrada tenham os vetores das variáveis dispostos em colunas, com as amostras de cada variável apresentadas nas linhas.

O *script* em questão implementa a técnica de normalização linear no intervalo [0,1], obedecendo à seguinte Expressão:

$$X_n = \frac{X_o - X_{min}}{X_{max} - X_{min}} \quad (10)$$

A função de normalização inversa obedece à Expressão:

$$X_o = X_n(X_{max} - X_{min}) + X_{min} \quad (11)$$

onde:

$X_o = x$  original;

$X_n = x$  normalizado;

$X_{max}$  = maior valor de  $x$  por parâmetro;

$X_{min}$  = menor valor de  $x$  por parâmetro.

## 3.2 Recursos computacionais

O computador é um recurso cada vez mais necessário às pesquisas de qualquer área do conhecimento, uma vez que ferramentas implementam algoritmos e técnicas que possibilitam tratamento de dados cada vez mais eficientes. Neste trabalho, destacam-se do uso

<sup>4</sup> Arquivos em texto puro são caracterizados por não incorporarem, em seu conteúdo, informações sobre formatação dos dados (nem mesmo formatação de caracteres, como negrito, itálico, fonte, ou tamanho de folha, margens, etc). Por este motivo, são os mais indicados para transportar dados entre plataformas (sistemas operacionais) e programas diferentes, pois, além de ter tamanho sensivelmente menor que os demais arquivos, os dados estão limpos para leitura, não sendo necessária a aplicação de nenhum filtro para separar a informação principal das informações de formatação.

habitual, dois programas utilizados para o tratamento dos dados e configuração/testes com as redes neurais, como verifica-se a seguir.

### 3.2.1 NEUROPH

Neuroph *Studio*<sup>5</sup> é um *framework*<sup>6</sup> JAVA para redes neurais artificiais criado para desenvolver arquiteturas comuns de redes neurais. O programa contém bibliotecas de código aberto (*open source*) com número de classes que corresponde aos conceitos básicos de RNA. É Orientado a Objeto (OO)<sup>7</sup> e sua *Application Programming Interface* (API) é rica em detalhes, com camadas de abstração que permitem manipulação direta das características da rede durante seu uso, além de redes, neurônios e modelos de aprendizagem pré-definidos.

O programa conta ainda com editor gráfico que permite criar, treinar, armazenar e carregar redes neurais JAVA, facilitando o seu uso para quem não deseja acessar o *framework* diretamente, através da criação de aplicação JAVA. Este editor conta com interface gráfica, criada para permitir sua utilização a partir de menus e ícones, buscando facilitar o trabalho dos usuários interessados em configurar uma rede neural. A partir da interface é possível gerenciar i) projetos; ii) redes neurais; iii) conjuntos de treinamentos; configurando detalhes em cada um dos elementos anteriores.

Por ser desenvolvido em JAVA, o *framework* “herda” muitas características desta linguagem de programação, que há muito tempo já a tornam excelente escolha para o desenvolvimento de programas. É importante ressaltar duas delas i) seu código aberto favorece o constante desenvolvimento da linguagem, que se moderniza constantemente e permite que suas ferramentas de desenvolvimento e documentação sejam também copiadas e utilizadas livremente por qualquer pessoa; ii) os programas JAVA rodam nas mais variadas arquiteturas e plataformas, ou seja, nos mais diferentes tipos de processadores e sistemas operacionais. Assim, o programa desenvolvido para determinada plataforma pode ser preparado para ser executado em outras, com pouco ou nenhum esforço, permitindo que não apenas os programas desenvolvidos naquela plataforma, mas por conseguinte o próprio NEUROPH seja utilizado por número muito maior de usuários.

Em função de ser licenciado como código aberto sob a licença Apache 2.0<sup>8</sup>, o código fonte da aplicação pode ser copiado e o usuário pode ter acesso aos aspectos mais internos do desenvolvimento do programa, modificá-lo, usá-lo e (re)distribuí-lo livremente. No entanto, a maioria dos usuários preferem copiar a versão executável do programa, que

<sup>5</sup> A versão 2.8 foi copiada do sítio do projeto <<http://neuroph.sourceforge.net>>

<sup>6</sup> Conjunto de bibliotecas e/ou classes que podem ser utilizadas no desenvolvimento de aplicações.

<sup>7</sup> Orientação a Objeto é um conceito moderno para desenvolvimento de componentes de *software* que procura criar programas de computador que refletem melhor os objetos do mundo real que necessitem ser neles representados. Seu objetivo é tornar os componentes reusáveis, facilitando a manutenção pela equipe de desenvolvimento, melhorando assim sua qualidade.

<sup>8</sup> <http://www.apache.org/licenses/>

está disponível e pronta para uso em computadores utilizando sistemas operacionais Windows, Linux ou OS X (Macintosh), que juntos compreendem mais de 99% dos sistemas operacionais instalados em computadores *desktop* e *laptop*.

Alguns tipos de redes suportados pelo *framework* são:

- ❑ *Adaline*;
- ❑ *Perceptron*;
- ❑ MLP com algoritmo *Backpropagation*;
- ❑ *Hopfield Network*;
- ❑ *Bidirectional Associative Memory*;
- ❑ *Neuro Fuzzy Reasoner*.

A literatura demonstra que este *framework* tem sido utilizado para desenvolvimento de diversos experimentos com redes neurais, conforme apresentado em (ELVIRA, 2013), (BAPTISTA; MORGADO-DIAS, 2013), (ANDERSEN, 2010), (ERASMUS et al., 2010) e (SEVARAC, 2006).

Neste trabalho, a versão do *framework* utilizada para os experimentos foi a 2.8.

### 3.2.2 SCILAB

Scilab<sup>9</sup> é uma linguagem de programação associada com rica coleção de algoritmos numéricos cobrindo muitos aspectos de problemas de computação científica. Do ponto de vista do *software*, Scilab é uma linguagem interpretada. Isso significa normalmente processos de desenvolvimento mais rápidos, porque o usuário acessa diretamente uma linguagem de alto nível, com um rico conjunto de características providas pelas bibliotecas. A linguagem do Scilab permite compilar dinamicamente e “associar” outras linguagens como Fortran e C, de modo que bibliotecas externas possam ser usadas como se elas fossem parte das características internas do programa. Usuários do Scilab podem desenvolver seus próprios módulos de modo a resolver seus problemas particulares (BAUDIN, 2010). O programa também faz interface com *LabVIEW*, ambiente de desenvolvimento e plataforma para linguagem de programação visual da *National Instruments*.

Do ponto de vista da licença, Scilab é um *software* livre no sentido de que o usuário não paga por ele, além do seu código fonte ser aberto, de acordo com a licença Cecill<sup>10</sup>. Assim, o programa é distribuído juntamente com seu código fonte e o usuário tem acesso aos seus aspectos mais internos, podendo modificá-lo e usá-lo livremente.

<sup>9</sup> <http://www.scilab.org>

<sup>10</sup> Cecill and free software. <http://www.cecill.info>.

O sistema de ajuda do Scilab documenta suas funções e bibliotecas. Ajuda adicional pode ser conseguida i) em seu sítio de internet; ii) através de manuais, como o de (BAUDIN, 2010) *Introduction to Scilab*; iii) e livros publicados especificamente sobre o assunto, como o de (CAMPBELL et al., 2006) *Modeling and Simulation in Scilab/Scicos*.

Scilab é dedicado principalmente a computação científica e provê acesso a grande quantidade de bibliotecas de áreas como álgebra linear, integração numérica e otimização. Os dados podem ser importados de vários tipos de arquivos de dados (CSV, TXT, DAT, dentre outros) e exportados para formato *Comma-separated Values* (CSV) livremente através de comandos internos da aplicação ou mesmo da interface gráfica. Há também funcionalidades de visualização de dados, incluindo gráficos 2D/3D e animação. Os gráficos podem ser exportados como imagens GIF - (*Graphics Interchange Format*), Postscript, Postscript-Latex, PNG e Xfig.

Algumas das funções científicas que estão disponíveis na aplicação:

- ❑ Álgebra linear, matrizes;
- ❑ Polinômios e funções racionais;
- ❑ Interpolação, aproximação;
- ❑ Otimização linear, quadrática e não-linear;
- ❑ Resolvedor de Equações Diferenciais Ordinárias e Equações Algébricas Diferenciais;
- ❑ Otimização de Inequações Matriciais Lineares, Controle clássico e robusto;
- ❑ Otimizações diferenciáveis e não-diferenciáveis;
- ❑ Processamento de sinal;
- ❑ Tradutor MATLAB-para-Scilab;
- ❑ Estatística.

Neste trabalho, a versão utilizada para os cálculos computacionais foi a 5.4, para o sistema operacional Gnu Linux.

Este *software* foi utilizado para, através da função de interpolação linear, obter valores intermediários dentro da faixa de valores de cada variável das amostras de água do Rio Meia Ponte (GO), visando aumentar o banco de dados original de 16 amostras para 141. Procura-se avaliar a diferença de generalização utilizando a base original e os arquivos com dados teóricos obtidos por interpolação.

### 3.2.3 Computador utilizado nos experimentos

Nos experimentos realizados neste trabalho, todas as RNA foram treinadas no mesmo computador, cuja configuração consiste de processador Intel Quad-Core i5 de 2.50 GHz, 8 GB de memória RAM DDR3 e sistema operacional GNU/Linux 64 bits.

De modo a quantificar o custo computacional para o treinamento das RNA deste experimento, é mostrada na Tabela 5 o tempo aproximado de treinamento para as redes neurais do Experimento 1, Ponto de Amostragem 3, que conta com 141 amostras, utilizando-se o computador descrito acima. As RNA deste treinamento contam com 10 entradas. O tempo de treinamento para as RNA do Experimento 3, que contam com 6 entradas, para o mesmo ponto de amostragem teve variação muito pequena se comparado com os do Experimento 1. A fase de treinamento é a que toma mais tempo/recursos de processamento, uma vez que é nela que o modelo é desenvolvido, através de erros e correções recorrentes. A fase posterior, de testes com amostras não vistas pela rede, toma uma fração pequena do tempo de treinamento.

Tabela 5 – Tempo para execução dos treinamentos no *software* NEUROPH. Experimento 1 - Ponto 3.

Neurônios na camada intermediária ( $x$ )	Iteração	Tempo (segundos)
1	10000	104
2	10000	104
3	2169	23
4	2591	27
5	1581	16
6	1757	18
7	2683	28
8	2254	23
9	1661	17
10	1292	13
11	1548	16
12	1230	13
13	1591	17
14	1917	20
15	1602	17
16	2098	22
17	2124	22
18	1305	14
19	1525	16
20	1621	17
21	1739	18
22	1451	15
23	2240	23

---

## Resultados e Discussão

Este capítulo descreve, em suas várias seções, uma sequência de experimentos que consistem no treinamento de redes neurais, com arquiteturas diferentes, utilizando dados de parâmetros físico-químicos oriundos de cinco pontos de amostragem do Rio Meia Ponte (GO) e de um ponto de amostragem do Rio Paraíba do Sul (SP). Os primeiros são dados teóricos, advindos de pré-processamento matemático com a finalidade de melhorar os conjuntos de treinamento e testes do Rio Meia Ponte; os dados do Rio Paraíba do Sul são amostras reais retiradas de um ponto de amostragem do manancial. Objetiva-se avaliar qual é a quantidade de neurônios na camada intermediária que define a melhor arquitetura da rede.

### 4.1 Treinamento/Testes de verificação

Em geral, o desenvolvimento de aplicação de RNA segue as seguintes etapas (SANTOS; FILHO, 2003):

1. obtenção de dados;
2. separação de conjuntos para treinamento e verificação;
3. configuração da rede;
4. treinamento;
5. verificação.

Assim, dos conjuntos de amostras referentes a cada ponto de amostragem, são separados dois conjuntos de dados, sendo:

- ❑ 70% destinados ao treinamento da rede neural;
- ❑ 30% para testar a eficiência da rede no reconhecimento do padrão.

Esta separação dos dados pode ser feita manualmente, ou seja, a partir do *software* de matemática computacional, retiram-se os primeiros 70% da matriz de dados original, na sequência em que aparecem, que farão parte do subconjunto  $A$  de treinamento, sendo então armazenados em um novo arquivo. Os 30% de dados remanescentes compõem o subconjunto  $B$ , composto de dados de verificação/testes. Uma outra forma de proceder a esta separação é utilizando-se uma função interna do próprio Neuroph, que separa automaticamente o conjunto original de dados em dois subconjuntos,  $A$  de treinamento e  $B$  de verificação/testes. Esta função, no entanto, utiliza um algoritmo aleatório que embute um certo grau de aleatoriedade na escolha das amostras que farão parte dos subconjuntos  $A$  e  $B$ . Este método foi empregado para os experimentos deste trabalho.

## 4.2 Experimento 1 - dados teóricos

Esta seção descreve o treinamento de diversas RNA utilizando-se dados do Rio Meia Ponte, interpolados linearmente, a partir de várias configurações de arquitetura; variou-se o número de nós na camada intermediária entre 1 e 23. Para simplificar a notação, adota-se a codificação, de acordo com o seguinte exemplo: (10-1-1), ou seja, dez sinais na camada de entrada, um neurônio na camada intermediária e um sinal na camada de saída. Na segunda configuração, acrescentou-se um neurônio na camada intermediária (10-2-1) e assim sucessivamente, até a configuração (10-23-1). A Figura 12 apresenta um exemplo de rede neural com dez neurônios na camada de entrada, onde neurônios na camada intermediária e um neurônio na camada de saída, conforme exibição do *software* Neuroph. A Figura 13, oriunda do mesmo programa, apresenta a visualização gráfica da mesma arquitetura de rede, onde percebe-se que cada neurônio da camada de entrada relaciona-se com todos os neurônios da camada intermediária. Esta figura ainda apresenta as entradas da rede associadas aos respectivos neurônios desta camada (In 1 até In 10) e o neurônio de saída associado ao valor de saída da rede (Out 1). Assim, em ambas as figuras, *Layer 1* refere-se à primeira camada, ou a camada de entrada; *Layer 2* refere-se à primeira camada intermediária (e neste caso única), o que nos leva à camada de saída (e última), ou *Layer 3*. As representações dos neurônios das referidas figuras estão com valores nulos (simbolizados pelo número 0) pois nenhuma informação lhes é atribuída, ou seja, nenhum conjunto de dados para treinamento da RNA é inserido na mesma. Os únicos neurônios com valores atribuídos são os da extrema direita, chamados *bias* e sempre portadores do valor (1). RIBEIRO (2003) destaca que o *bias* tem o efeito de aumentar ou diminuir a entrada líquida da função de ativação, dependendo se este neurônio é positivo ou negativo, respectivamente. Ou seja, permite à rede neural melhor adaptação ao conhecimento a ela fornecido.

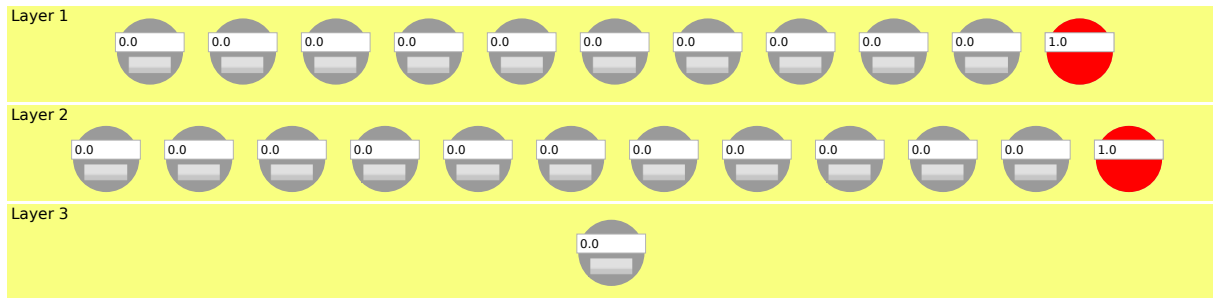


Figura 12 – Exemplo do *software* NEUROPH de rede MLP na arquitetura (10-11-1).

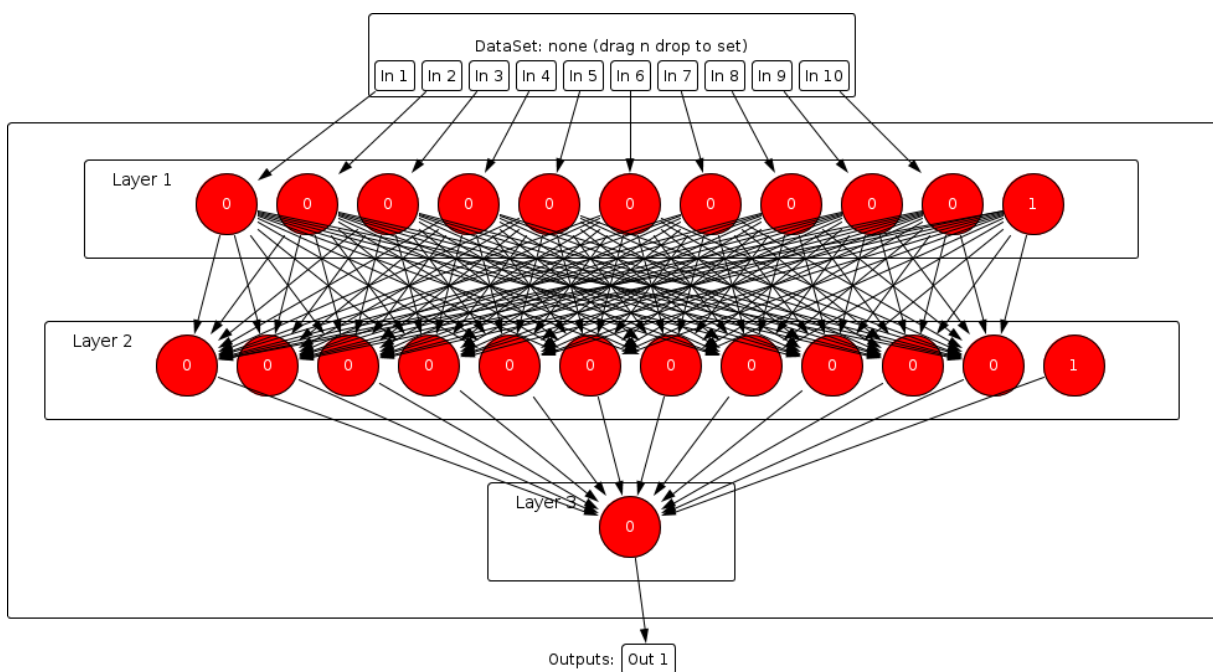


Figura 13 – Visualização gráfica do *software* NEUROPH de rede MLP na arquitetura (10-11-1).

Utilizou-se na entrada da RNA dez sinais (parâmetros) normalizados de entrada:

- Temperatura da água (in 1);
- pH (in 2);
- Demanda Bioquímica de Oxigênio - DBO (in 3);
- Condutividade (in 4);
- Nitrogênio Amoniacal (in 5);



- ❑ Nitrato (in 6);
- ❑ Nitrito (in 7);
- ❑ Fósforo (in 8);
- ❑ Turbidez (in 9);
- ❑ Sólidos Dissolvidos Totais (in 10);

E apenas um sinal (parâmetro) normalizado de saída:

- ❑ Oxigênio Dissolvido (out 1).

### 4.2.1 Configuração para o treinamento

Para cada experimento, a rede é treinada três vezes para aumentar a probabilidade de produzir bom conjunto inicial de pesos. Uma vez obtidos os resultados, estes valores são utilizados para gerar média dos três experimentos; os parâmetros de inicialização do algoritmo são padronizados para todas as redes, com vistas a permitir melhor comparação entre as arquiteturas, como pode ser observado na Tabela 6.

Tabela 6 – Configuração para treinamento da RNA do Experimento 1

Variáveis de treinamento	Limite
Erro máximo ( <i>Max error</i> )	0,0001
Limite de iteração ( <i>Iteration</i> ) ou Época	10.000
Taxa de aprendizagem ( <i>Learning rate</i> )	0,2
Momento ( <i>Momentum</i> )	0,7

Para o propósito deste projeto, o número máximo de épocas é configurado em  $10^4$ . O parâmetro “erro máximo” é configurado para  $10^{-4}$ , pois pretende-se alcançar erro baixo o suficiente para que o modelo da rede seja o mais próximo do real. As variáveis “taxa de aprendizagem” e “momento” são configuradas com os valores padrão da aplicação, 0,2 e 0,7, respectivamente, pois valores menores nestes parâmetros produzem variação muito pequena nos pesos sinápticos, corrigidos a cada ciclo do algoritmo de retropropagação, elevando o número de iterações (e conseqüentemente o esforço computacional) necessários para se aproximar da convergência.

Cada um dos experimentos de treinamento da rede deve parar quando **uma** das seguintes condições for atingida:

- ❑ Valor configurado da variável “Erro Máximo” for alcançado.
- ❑ Valor configurado da variável “Limite de iteração” for alcançado;

## 4.2.2 Ponto de amostragem 1

Os dados originais são oriundos de análise de água coletada neste ponto de amostragem do rio, submetidos à técnica de interpolação linear (URROZ, 2001) para a produção de mais valores para aumentar o conjunto de treinamento. Antes do treinamento, o conjunto é padronizado (normalizados entre  $[0,1]$ ), sendo posteriormente separados em dois conjuntos através do arquivo contendo os dados de treinamento, que visam promover o aprendizado da RNA, e de outro contendo os dados de teste ou verificação, responsáveis por verificar a capacidade de generalização da rede:

□ Ponto de amostragem 1 - 131 amostras de 11 parâmetros, conforme a seção 3.1.1:

- Dados para treinamento: 91 amostras (70%);
- Dados para verificação: 40 amostras (30%).

### 4.2.2.1 Resultados

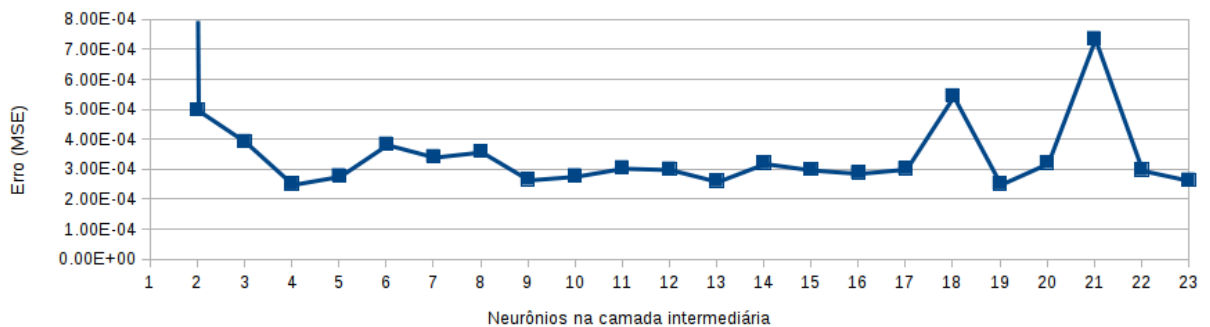
A Tabela 7 traz os dados dos resultados de treinamento das RNA com dez entradas, uma saída e uma camada intermediária, variando o número de neurônios nesta camada entre 1 e 23. Os valores de Erro Total da Rede, Iteração e Erro Quadrático Médio<sup>1</sup> dos Testes são médias das triplicatas. O cálculo do desvio padrão para os resultados de Erro Quadrático dos testes efetuados em todas as arquiteturas é de 4,08-E03.

O gráfico da Figura 14 apresenta a dinâmica do Erro Quadrático Médio em função do número de neurônios da camada intermediária. O treinamento da rede na arquitetura (10-1-1) produziu erro quadrático muito alto, motivo pelo qual este valor é mantido fora do gráfico, para comportar melhor a escala dos demais valores. O menor erro é encontrado com a utilização de quatro neurônios na camada intermediária (10-4-1), arquitetura que também convergiu com o menor número de iterações possível (3011), conforme verificado na Tabela 7. Portanto, pode-se dizer que esta seria a configuração da rede com menor custo computacional para o problema proposto.

<sup>1</sup> A diferença entre a estimativa e o valor real elevada ao quadrado.

Tabela 7 – Dados médios obtidos com o treinamento da rede (10- $x$ -1) para o Ponto 1.

Neurônios na camada intermediária ( $x$ )	Erro Máx.	Limite máx. iteração	Erro Total da Rede (70%)	Iteração	Erro Quadrático dos Testes (30%)
1	0,0001	10000	4.05E-03	10000	1.99E-02
2	0,0001	10000	1.86E-04	10000	4.99E-04
3	0,0001	10000	1.15E-04	10000	3.93E-04
4	0,0001	10000	9.99E-05	3011	2.51E-04
5	0,0001	10000	9.99E-05	4961	2.79E-04
6	0,0001	10000	9.99E-05	3465	3.85E-04
7	0,0001	10000	9.99E-05	3084	3.43E-04
8	0,0001	10000	9.99E-05	7254	3.60E-04
9	0,0001	10000	9.99E-05	4175	2.67E-04
10	0,0001	10000	9.99E-05	3860	2.79E-04
11	0,0001	10000	9.99E-05	5109	3.06E-04
12	0,0001	10000	9.99E-05	5905	3.01E-04
13	0,0001	10000	9.99E-05	3375	2.61E-04
14	0,0001	10000	9.99E-05	5311	3.22E-04
15	0,0001	10000	9.99E-05	5723	3.00E-04
16	0,0001	10000	9.99E-05	5753	2.89E-04
17	0,0001	10000	9.99E-05	5021	3.03E-04
18	0,0001	10000	9.99E-05	4218	5.46E-04
19	0,0001	10000	9.99E-05	4735	2.52E-04
20	0,0001	10000	9.99E-05	4911	3.22E-04
21	0,0001	10000	9.99E-05	4840	7.37E-04
22	0,0001	10000	9.99E-05	5092	2.99E-04
23	0,0001	10000	9.99E-05	3555	2.64E-04

Figura 14 – Erro médio quadrático da rede (10- $x$ -1) para o Ponto 1.

A Tabela 8 exibe como exemplo, os resultados obtidos no teste da arquitetura (10-4-1). As colunas numeradas como IN1 a IN10 representam o vetor com dez valores normalizados de entrada para o treinamento da RNA, oriundos dos parâmetros a partir dos quais se pretende modelar a saída, de acordo com a Seção 4.2; a coluna “Saída da Rede” representa o valor estimado pela rede neural; a coluna “Valor desejado” informa o valor do dado real, ou seja, o próprio O.D. que faz parte do conjunto de dados de treinamento; a coluna “Erro” representa a diferença entre o valor desejado e o valor

Tabela 8 – Resultado do teste da rede (10-4-1) para o Ponto 1.

Amostras	Parâmetros										Saída da rede	Valor desejado	Erro
	IN1	IN2	IN3	IN4	IN5	IN6	IN7	IN8	IN9	IN10			
1	0,6857	0,4283	0,1425	0,7800	0,4886	0,0585	0,1429	0,1045	0,5544	0,1390	0,3423	0,3714	-0,0292
2	0,0571	0,7735	0,2349	0,5850	0,5426	0,1829	0,9571	0,2075	0,9463	0,1250	0,6784	0,6786	-0,0002
3	0,3571	0,6080	0,0678	0,9550	0,4316	0,0878	0,0143	0,2179	0,4557	0,0253	0,7725	0,7857	-0,0132
4	0,4571	0,6655	0,2071	0,6200	0,5265	0,0951	0,6429	0,2507	0,3242	0,1472	0,6140	0,6143	-0,0003
5	0,3429	0,7912	0,6496	0,2600	0,7856	0,1220	0,7429	0,2746	0,6779	0,0961	0,5131	0,5000	0,0131
6	0,1429	0,5619	0,0141	0,7000	0,4061	0,3293	0,0000	0,1791	0,3591	0,0130	0,9351	0,9286	0,0065
7	0,7143	0,2850	0,4835	0,5650	0,6884	0,3000	0,1857	0,4358	0,1376	0,7246	0,2985	0,3000	-0,0015
8	0,7143	0,1221	0,5998	0,5850	0,7566	0,7000	0,2429	0,7582	0,1242	0,8820	0,1058	0,1286	-0,0228
9	0,7143	0,3487	0,5857	0,4500	0,7519	0,1610	0,0857	0,0448	0,1826	0,1381	0,5473	0,5571	-0,0098
10	0,0857	0,5354	0,0084	0,6000	0,4057	0,4220	0,0000	0,1672	0,3228	0,0097	0,9464	0,9571	-0,0108
11	0,2571	0,6150	0,0253	0,9000	0,4068	0,1439	0,0000	0,2030	0,4315	0,0197	0,8620	0,8714	-0,0094
12	0,3286	0,9000	0,3726	0,3300	0,0808	0,1000	0,0286	0,2478	0,0658	0,1000	0,6341	0,6429	-0,0088
13	0,4143	0,7000	0,4425	0,3900	0,2423	0,3000	0,0857	0,4149	0,0765	0,3000	0,4889	0,5000	-0,0111
14	0,7143	0,4195	0,3788	0,5300	0,6271	0,0341	0,1429	0,1701	0,1181	0,5296	0,4078	0,4143	-0,0065
15	0,1714	0,7327	0,1945	0,5600	0,5190	0,1220	0,9143	0,2657	0,5208	0,1507	0,6283	0,6286	-0,0003
16	0,7857	0,3805	0,2741	0,8500	0,5657	0,0000	0,1429	0,2015	0,5134	0,1102	0,1481	0,1429	0,0052
17	0,7143	0,4088	0,4476	0,4500	0,6692	0,1659	0,1143	0,0597	0,0913	0,1799	0,4441	0,4571	-0,0131
18	0,6571	0,3540	0,1839	0,7600	0,5129	0,0000	0,1429	0,1433	0,5154	0,1399	0,1550	0,1429	0,0121
19	0,0000	0,4956	0,0000	0,4500	0,4052	0,5610	0,0000	0,1493	0,2685	0,0047	0,9478	1,0000	-0,0522
20	0,3000	0,3035	0,0024	0,5900	0,4064	0,1854	0,0000	0,2537	0,2544	0,0245	0,2599	0,3000	-0,0401
21	0,5143	0,6513	0,2019	0,6600	0,5234	0,1024	0,5714	0,2299	0,3651	0,1416	0,6279	0,6286	-0,0006
22	0,7143	0,1628	0,5707	0,5800	0,7396	0,6000	0,2286	0,6776	0,1275	0,8426	0,1462	0,1714	-0,0253
23	0,3429	0,9611	0,4664	0,2400	0,1960	0,0146	0,1143	0,1970	0,1409	0,0146	0,6416	0,6429	-0,0012
24	0,2857	0,7876	0,5667	0,3250	0,7370	0,1341	0,7857	0,2612	0,7315	0,1019	0,5431	0,5357	0,0074
25	0,5714	0,5469	0,1866	0,9700	0,5056	0,0585	0,0571	0,2448	0,4738	0,0371	0,5843	0,5714	0,0129
26	0,7143	0,0407	0,6580	0,5950	0,7907	0,9000	0,2714	0,9194	0,1174	0,9607	0,5006	0,0429	0,0077
27	0,1429	0,7389	0,1874	0,5750	0,5149	0,1341	0,9286	0,2537	0,6007	0,1474	0,6366	0,6429	-0,0062
28	0,7143	0,4133	0,3875	0,5400	0,6322	0,0171	0,1429	0,1821	0,1329	0,5681	0,4085	0,4214	-0,0129
29	0,6714	0,1000	0,6522	0,5700	0,7270	0,9000	0,2571	0,9164	0,1087	0,9000	0,0700	0,0714	-0,0015
30	0,7143	0,4628	0,3182	0,4600	0,5916	0,1537	0,1429	0,0866	0,0148	0,2602	0,3655	0,3643	0,0012
31	0,5429	0,4000	0,5473	0,4800	0,4847	0,6000	0,1714	0,6657	0,0926	0,6000	0,2789	0,2857	-0,0068
32	0,2857	0,7080	0,2228	0,5000	0,5357	0,0732	0,8571	0,3134	0,2013	0,1640	0,5544	0,5714	-0,0170
33	0,6429	0,5265	0,2262	0,9750	0,5303	0,0488	0,0714	0,2537	0,4799	0,0410	0,5103	0,5000	0,0103
34	0,4000	0,7947	0,7326	0,1950	0,8342	0,1098	0,7000	0,2881	0,6242	0,0903	0,4807	0,4643	0,0164
35	0,2571	0,3310	0,0020	0,5700	0,4062	0,2390	0,0000	0,2388	0,2564	0,0217	0,3743	0,4000	-0,0257
36	0,5143	0,8018	0,8985	0,0650	0,9314	0,0854	0,6143	0,3149	0,5168	0,0787	0,4010	0,3929	0,0081
37	0,4571	0,7982	0,8155	0,1300	0,8828	0,0976	0,6571	0,3015	0,5705	0,0845	0,4437	0,4286	0,0151
38	0,7143	0,3664	0,4253	0,5550	0,6543	0,1000	0,1571	0,2746	0,1443	0,6459	0,3841	0,3857	-0,0016
39	0,0857	0,7513	0,1732	0,6050	0,5065	0,1585	0,9571	0,2299	0,7604	0,1407	0,6544	0,6714	-0,0170
40	0,9143	0,4071	0,3642	0,9400	0,6184	0,0000	0,1429	0,2597	0,5114	0,0805	0,1400	0,1429	-0,0028

estimado pelo treinamento da rede. As linhas exibidas, numeradas de 1 a 40, representam as amostras separadas especificamente para verificar a acuidade do treinamento, nunca tendo sido apresentadas à rede em fase anterior e totalizam 30% dos dados disponíveis para desenvolver o experimento.

O valor do erro quadrático médio (*Total Mean Square Error*), calculado a partir dos erros obtidos nos testes com a rede de arquitetura (10-4-1), é  $2,51E-4$ . Este erro representa a quantia pela qual o resultado estimado se distancia do valor esperado (neste caso, o valor de O.D. coletado a partir de análises de amostras de água). Quanto menor este erro, mais próxima a estimativa do valor real. Assim, quando confrontada com dados nunca vistos, a capacidade de generalização da rede treinada permitiu que ela estimasse os valores de Oxigênio Dissolvido com margens pequenas de erro. A Figura 15 demonstra o número de neurônios necessários para que a rede alcance o valor de erro desejado e o treinamento cesse, ou, quando não alcançado o treinamento, atinja o número máximo de iterações

permitido.

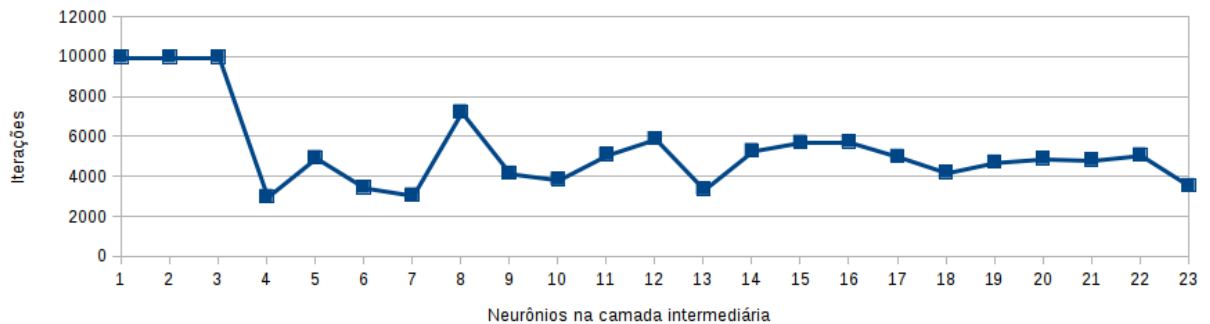


Figura 15 – Número de iterações em função da quantidade de neurônios na camada intermediária para o Ponto 1.

#### 4.2.2.2 Comparação do modelo com dados reais

A partir do treinamento da rede de arquitetura (10-4-1) é possível normalizar inversamente os dados de saída (*output*) da rede com a finalidade de verificar a distância destes dados com o alvo do treinamento, o valor real de Oxigênio Dissolvido. A Figura 16 apresenta estes valores iniciais da variável O.D., antes de serem padronizados e utilizados no conjunto de teste, contra o O.D. modelado pela rede, depois da efetuada a normalização reversa. O eixo *Y* apresenta a escala do OD e o eixo *X* as amostras utilizadas no teste na ordem em que aparecem no conjunto de verificação. Estes dados não estão classificados em ordem de data pois as novas amostras, geradas para todo o Experimento 1 a partir da técnica de interpolação linear, não têm data. A Figura 17 exhibe o gráfico de dispersão para o mesmo ponto. Verifica-se que com erro quadrático médio pequeno desta arquitetura as diferenças entre os valores originais e modelados são pequenas, o que pode ser confirmado pelo valor de coeficiente de correlação, calculado em 0,996, validando o método de modelagem através de RNA.

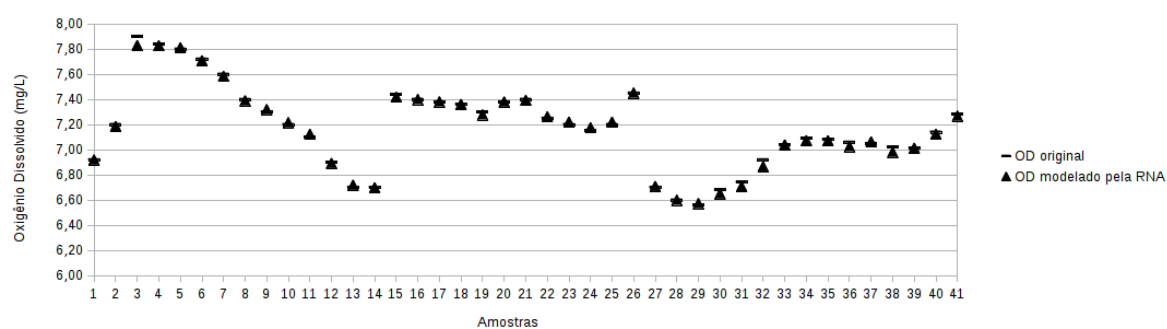


Figura 16 – Dado original contra dado modelado pela RNA para o Ponto 1.

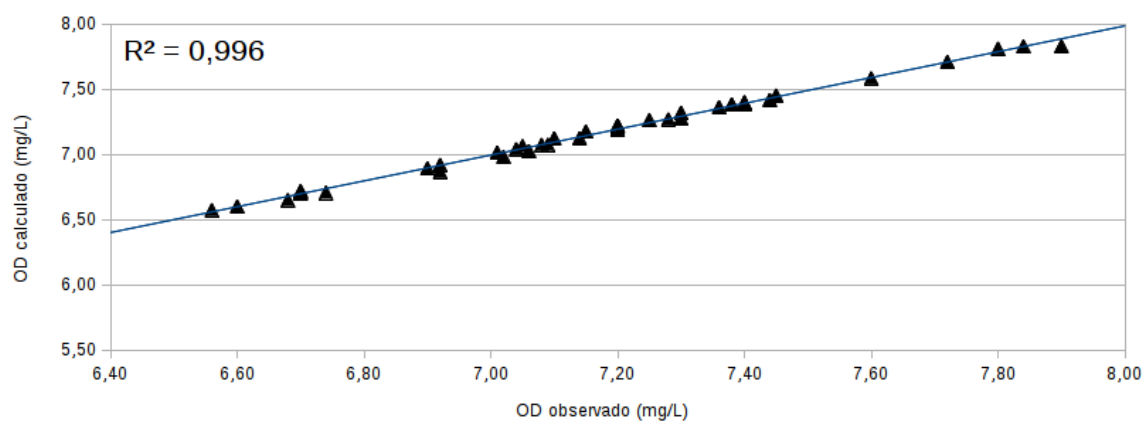


Figura 17 – Dispersão de dado original contra dado modelado pela RNA para o Ponto 1.

### 4.2.3 Ponto de amostragem 3

Os dados originais são oriundos de análise de água coletada neste ponto de amostragem do rio, submetidos à técnica de interpolação linear para a produção de mais valores para aumentar o conjunto de treinamento. Antes do treinamento, o conjunto é padronizado (normalizados entre  $[0,1]$ ), sendo posteriormente separados em dois conjuntos através do arquivo contendo os dados de treinamento, que visam promover o aprendizado da RNA, e outro contendo os dados de teste ou verificação, responsáveis por verificar a capacidade de generalização da rede, e outro os dados de teste:

□ Ponto de amostragem 3 - 141 amostras de 11 parâmetros, conforme a seção 3.1.1:

- Dados para treinamento: 98 amostras (70%);
- Dados para verificação: 43 amostras (30%).

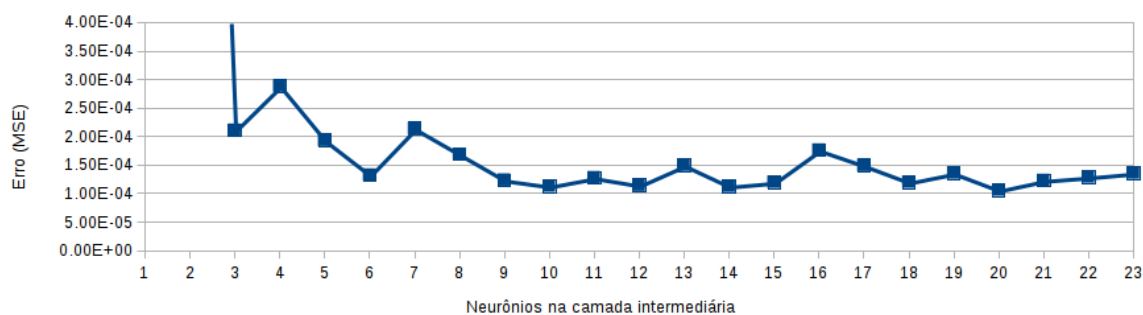
#### 4.2.3.1 Resultados

A Tabela 9 traz os dados dos resultados de treinamento das RNA com dez entradas, uma saída e uma camada intermediária. Após o treinamento das RNA com dez entradas, uma saída e uma camada intermediária, variando o número de neurônios nesta camada entre 1 e 23, é calculado o desvio padrão para os resultados de Erro Quadrático dos testes efetuados. O resultado obtido é  $8,75E-04$ .

O gráfico da Figura 18 apresenta a dinâmica do Erro Quadrático Médio em função do número de neurônios da camada intermediária. Os treinamentos das redes nas arquiteturas (10-1-1) e (10-2-1) produziram valores altos de erro quadrático, motivo pelo qual são mantidos fora do gráfico, para comportar melhor a escala dos demais valores. O menor erro é encontrado com a utilização de vinte neurônios na camada intermediária (10-20-1). Esta arquitetura não é a que convergiu com o menor número de iterações (1621), mas a que resultou em menor diferença entre o valor real e o modelo da RNA. De fato, a arquitetura (10-10-1) convergiu com o número mais baixo de iterações (1292), mas com valor de Erro Quadrático Médio ligeiramente maior que a arquitetura (10-20-1), motivo pelo qual não é escolhida. No entanto, cabe salientar que esta diferença pequena de erro não deve mudar drasticamente o resultado da comparação entre os dados do modelo e os dados reais; sendo assim, pode-se validar o uso da Expressão 3 (p. 43) para definição da quantidade máxima de neurônios na camada intermediária. Segundo esta expressão, o número de neurônios na camada intermediária é 11, valor que resultaria em pequena diferença de erro dos testes e igualmente pequena diferença entre OD modelado e real.

Tabela 9 – Dados médios obtidos com o treinamento da rede (10- $x$ -1) para o Ponto 3.

Neurônios na camada intermediária ( $x$ )	Erro Máx.	Limite máx. iteração	Erro Total da Rede (70%)	Iteração	Erro Quadrático dos Testes (30%)
1	0,0001	10000	2,64E-03	10000	3.94E-03
2	0,0001	10000	6,67E-04	10000	2.12E-03
3	0,0001	10000	9,99E-05	2169	2.11E-04
4	0,0001	10000	9,99E-05	2591	2.89E-04
5	0,0001	10000	9,99E-05	1581	1.93E-04
6	0,0001	10000	9,99E-05	1757	1.32E-04
7	0,0001	10000	9,99E-05	2683	2.14E-04
8	0,0001	10000	9,99E-05	2254	1.69E-04
9	0,0001	10000	9,99E-05	1661	1.24E-04
10	0,0001	10000	9,99E-05	1292	1.13E-04
11	0,0001	10000	9,99E-05	1548	1.28E-04
12	0,0001	10000	9,99E-05	1230	1.15E-04
13	0,0001	10000	9,99E-05	1591	1.50E-04
14	0,0001	10000	9,99E-05	1917	1.13E-04
15	0,0001	10000	9,99E-05	1602	1.20E-04
16	0,0001	10000	9,99E-05	2098	1.76E-04
17	0,0001	10000	9,99E-05	2124	1.49E-04
18	0,0001	10000	9,99E-05	1305	1.20E-04
19	0,0001	10000	9,99E-05	1525	1.36E-04
20	0,0001	10000	9,99E-05	1621	1.06E-04
21	0,0001	10000	9,99E-05	1739	1.23E-04
22	0,0001	10000	9,99E-05	1451	1.29E-04
23	0,0001	10000	9,99E-05	2240	1.36E-04

Figura 18 – Erro médio quadrático da rede (10- $x$ -1) para o ponto 3.

O valor do erro quadrático médio (*Total Mean Square Error*) da rede de arquitetura (10-20-1), é 1,06E-4. Este erro representa a quantia pela qual o resultado estimado se distancia do valor esperado (neste caso, o valor de O.D. coletado a partir de análises de amostras de água). A Figura 19 demonstra o número de neurônios necessários para que a rede alcance o valor de erro desejado e o treinamento cesse, ou, quando não alcançado o treinamento, atinja o número máximo de iterações permitido.



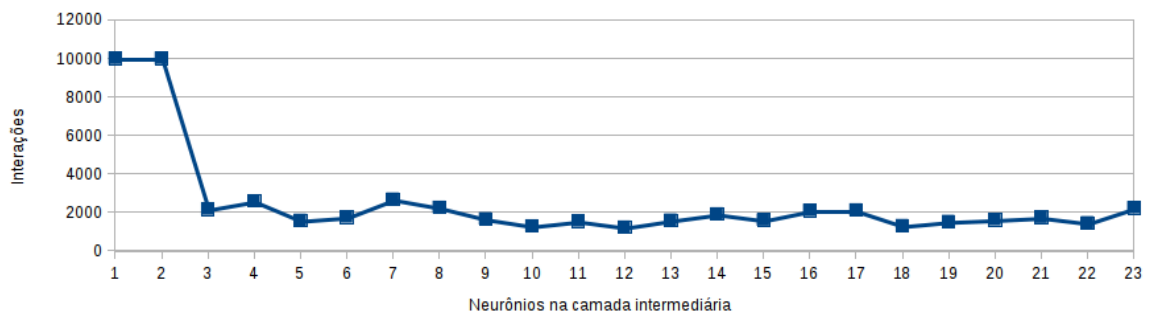


Figura 19 – Número de iterações em função da quantidade de neurônios na camada intermediária para o ponto 3.

#### 4.2.3.2 Comparação do modelo com dados reais

A partir do treinamento da rede de arquitetura (10-20-1), é possível normalizar inversamente os dados de saída (*output*) da rede com a finalidade de verificar a distância destes dados com o alvo do treinamento, o valor real de Oxigênio Dissolvido. A Figura 20 apresenta estes valores iniciais da variável O.D., antes de serem padronizados e utilizados no conjunto de teste, contra o O.D. modelado pela rede, depois da efetuada a normalização reversa. A Figura 21 exige a informação de dispersão dos mesmos dados. Verifica-se através do erro quadrático médio pequeno desta arquitetura que as diferenças entre os valores originais e modelados são pequenas, o que pode ser confirmado pelo coeficiente de correlação calculado em 0,998, validando o método de modelagem através de RNA.

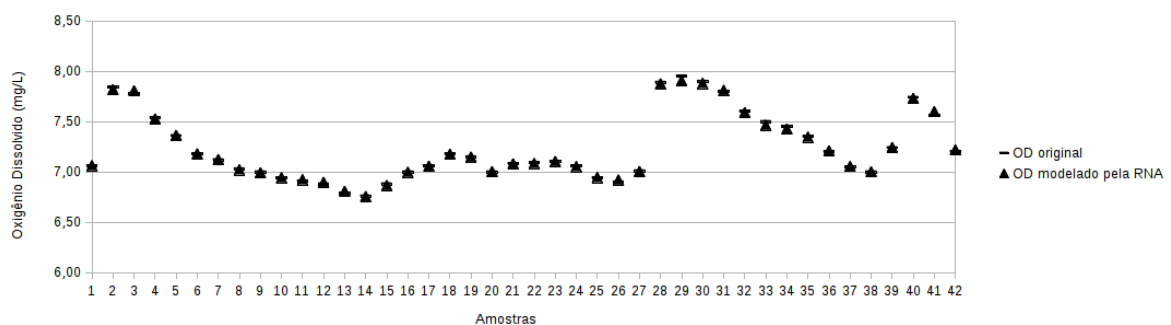


Figura 20 – Dado original contra dado modelado pela RNA para o ponto 3.

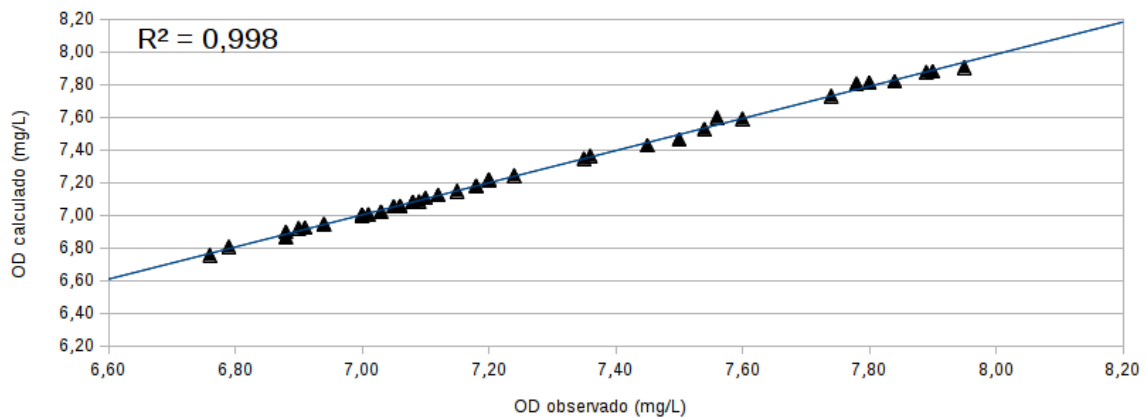


Figura 21 – Dispersão de dado original contra dado modelado pela RNA para o ponto 3.

#### 4.2.4 Ponto de amostragem 4

Os dados originais são oriundos de análise de água coletada neste ponto de amostragem do rio, submetidos à técnica de interpolação linear para a produção de mais valores para aumentar o conjunto de treinamento. Antes do treinamento, o conjunto é padronizado (normalizados entre  $[0,1]$ ), sendo posteriormente separados em dois conjuntos através do arquivo contendo os dados de treinamento, que visam promover o aprendizado da RNA, e de outro contendo os dados de teste ou verificação, responsáveis por verificar a capacidade de generalização da rede, e outro os dados de teste:

- Ponto de amostragem 4 - 131 amostras de 11 variáveis, conforme a seção 3.1.1:
  - Dados para treinamento: 91 amostras (70%);
  - Dados para verificação: 40 amostras (30%).

##### 4.2.4.1 Resultados

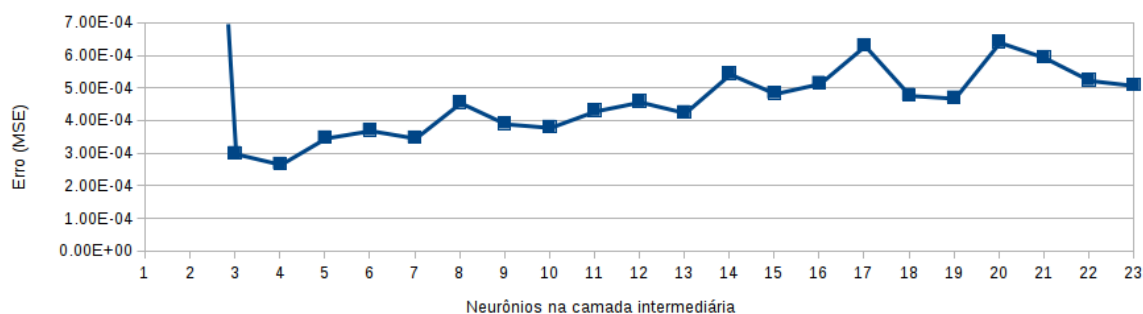
Após o treinamento das RNA com dez entradas, uma saída e uma camada intermediária, variando o número de neurônios nesta camada entre 1 e 23, os resultados são apresentados na Tabela 10. É calculado o desvio padrão para os resultados de Erro Quadrático dos testes efetuados. O resultado obtido é  $3,02E-03$ .

O gráfico da Figura 22 apresenta a dinâmica do Erro Quadrático Médio em função do número de neurônios da camada intermediária. Os treinamentos das redes nas arquiteturas (10-1-1) e (10-2-1) produziram valores altos de erro quadrático, motivo pelo qual são mantidos fora do gráfico, para comportar melhor a escala dos demais valores. O menor erro é encontrado com a utilização de quatro neurônios na camada intermediária (10-4-1).

Tabela 10 – Dados médios obtidos com o treinamento da rede (10- $x$ -1) para o Ponto 4.

Neurônios na camada intermediária ( $x$ )	Erro Máx.	Limite máx. iteração	Erro Total da Rede (70%)	Iteração	Erro Quadrático dos Testes (30%)
1	0,0001	10000	8,77E-03	10000	1,49E-02
2	0,0001	10000	4,79E-04	10000	2,56E-03
3	0,0001	10000	1,00E-04	2797	3,00E-04
4	0,0001	10000	1,00E-04	2353	2,68E-04
5	0,0001	10000	9,99E-05	1105	3,49E-04
6	0,0001	10000	9,99E-05	1123	3,72E-04
7	0,0001	10000	9,99E-05	1143	3,48E-04
8	0,0001	10000	9,99E-05	1460	4,57E-04
9	0,0001	10000	1,00E-04	1576	3,92E-04
10	0,0001	10000	9,99E-05	1476	3,81E-04
11	0,0001	10000	1,00E-04	1032	4,32E-04
12	0,0001	10000	9,99E-05	1266	4,60E-04
13	0,0001	10000	9,99E-05	1189	4,26E-04
14	0,0001	10000	1,00E-04	1703	5,45E-04
15	0,0001	10000	1,00E-04	1780	4,85E-04
16	0,0001	10000	1,00E-04	1413	5,16E-04
17	0,0001	10000	1,00E-04	1661	6,32E-04
18	0,0001	10000	9,99E-05	1393	4,80E-04
19	0,0001	10000	1,00E-04	1761	4,71E-04
20	0,0001	10000	9,99E-05	1519	6,43E-04
21	0,0001	10000	1,00E-04	1711	5,95E-04
22	0,0001	10000	1,00E-04	1522	5,26E-04
23	0,0001	10000	1,00E-04	1903	5,10E-04

Esta arquitetura não é a que convergiu com o menor número de iterações, mas a que resultou em menor diferença entre o valor real e o modelo da RNA.

Figura 22 – Erro médio quadrático da rede (10- $x$ -1) para o ponto 4.

O valor do erro quadrático médio (*Total Mean Square Error*) para a rede de arquitetura (10-4-1), é 2,67E-4. Este erro representa a quantia pela qual o resultado estimado se distancia do valor esperado (neste caso, o valor de O.D. coletado a partir de análises de amostras de água). A Figura 23 demonstra o número de neurônios necessários para que a rede alcance o valor de erro desejado e o treinamento cesse, ou, quando não alcançado

o treinamento, atinja o número máximo de iterações permitido.

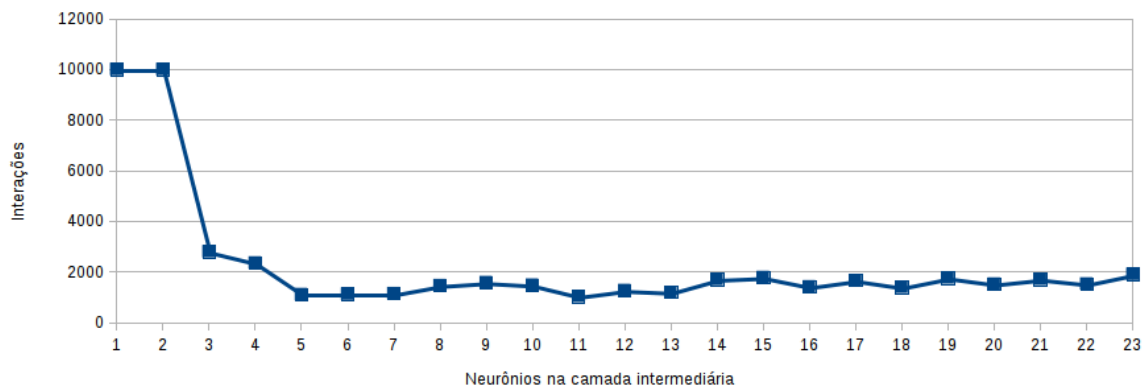


Figura 23 – Número de iterações em função da quantidade de neurônios na camada intermediária para o ponto 4.

#### 4.2.4.2 Comparação do modelo com dados reais

A partir do treinamento da rede de arquitetura (10-4-1), é possível normalizar inversamente os dados de saída (*output*) da rede com a finalidade de verificar a distância destes dados com o alvo do treinamento, o valor real de Oxigênio Dissolvido. A Figura 24 apresenta estes valores iniciais da variável O.D., antes de serem padronizados e utilizados no conjunto de teste, contra o O.D. modelado pela rede, depois da efetuada a normalização reversa. A Figura 25 ilustra a dispersão para os mesmos dados. Verifica-se que com erro quadrático médio pequeno desta arquitetura as diferenças entre os valores originais e modelados são pequenas, o que pode ser verificado pelo valor calculado de coeficiente de correlação de 0,990, validando o método de modelagem através de RNA.

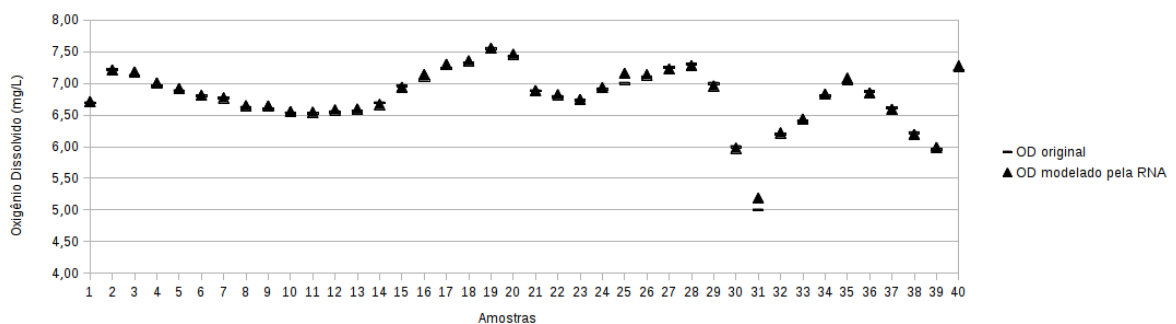


Figura 24 – Dado original contra dado modelado pela RNA para o ponto 4.

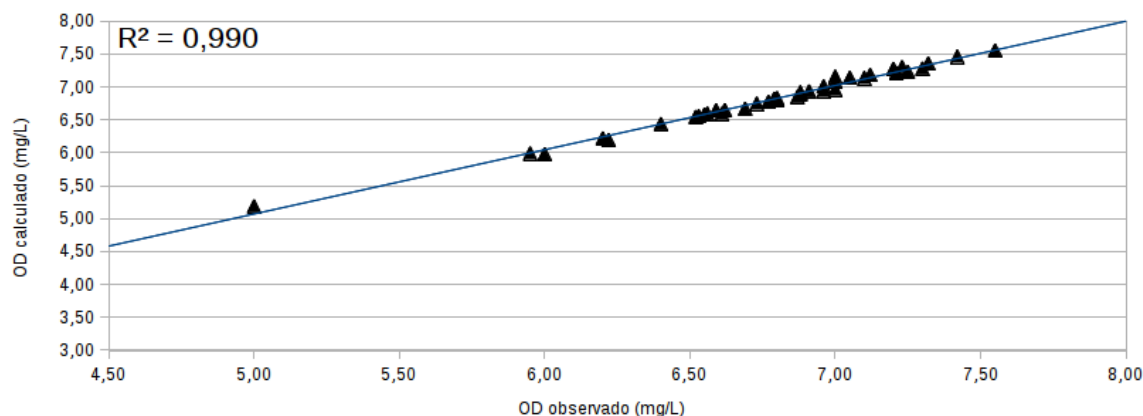


Figura 25 – Dispersão de dado original contra dado modelado pela RNA para o ponto 4.

## 4.2.5 Ponto de amostragem 5

Os dados originais são oriundos de análise de água coletada neste ponto de amostragem do rio, submetidos à técnica de interpolação linear para a produção de mais valores para aumentar o conjunto de treinamento. Antes do treinamento, o conjunto é padronizado (normalizados entre  $[0,1]$ ), sendo posteriormente separados em dois conjuntos através do arquivo contendo os dados de treinamento, que visam promover o aprendizado da RNA, e de outro contendo os dados de teste ou verificação, responsáveis por verificar a capacidade de generalização da rede, e outro os dados de teste:

- Ponto de amostragem 5 - 141 amostras de 11 variáveis, conforme a seção 3.1.1:
  - Dados para treinamento: 98 amostras (70%);
  - Dados para verificação: 43 amostras (30%).

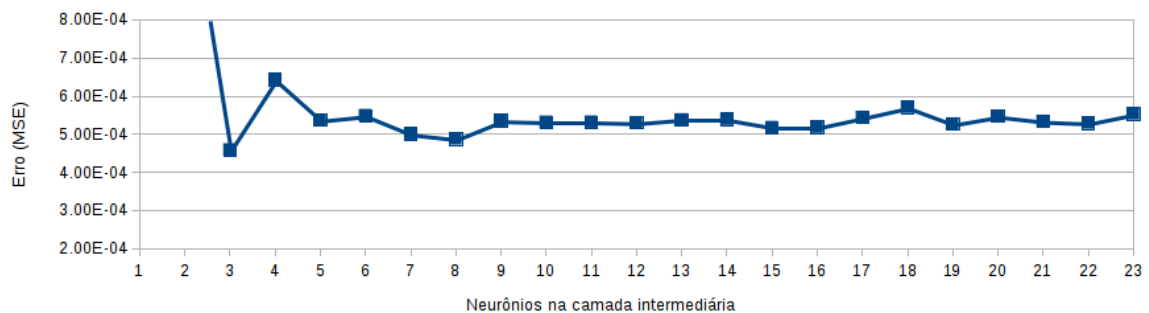
### 4.2.5.1 Resultados

A Tabela 11 apresenta os dados resultantes do treinamento das RNA com dez entradas, uma saída e uma camada intermediária, variando o número de neurônios nesta camada entre 1 e 23. O cálculo do desvio padrão para os resultados de Erro Quadrático dos testes efetuados é  $1,70E-03$ .

O gráfico da Figura 26 apresenta a dinâmica do Erro Quadrático Médio em função do número de neurônios da camada intermediária. Os treinamentos das redes nas arquiteturas (10-1-1) e (10-2-1) produziram valores altos de erro quadrático, motivo pelo qual são mantidos fora do gráfico, para comportar melhor a escala dos demais valores. A arquitetura (10-8-1) produziu um dos menores erros quadráticos dos testes, tendo finalizado o treinamento com sucesso, sem que alcançasse o máximo de iterações.

Tabela 11 – Dados médios obtidos com o treinamento da rede (10- $x$ -1) para o Ponto 5.

Neurônios na camada intermediária ( $x$ )	Erro Máx.	Limite máx. iteração	Erro Total da Rede (70%)	Iteração	Erro Quadrático dos Testes (30%)
1	0,0001	10000	2,98E-03	10000	8,74E-03
2	0,0001	10000	2,59E-04	10000	1,19E-03
3	0,0001	10000	1,43E-04	10000	4,58E-04
4	0,0001	10000	1,00E-04	2320	6,43E-04
5	0,0001	10000	1,00E-04	1379	5,36E-04
6	0,0001	10000	1,00E-04	2168	5,49E-04
7	0,0001	10000	1,00E-04	2219	5,00E-04
8	0,0001	10000	1,00E-04	1837	4,87E-04
9	0,0001	10000	1,00E-04	2198	5,35E-04
10	0,0001	10000	1,00E-04	2700	5,31E-04
11	0,0001	10000	1,00E-04	1966	5,31E-04
12	0,0001	10000	1,00E-04	2623	5,30E-04
13	0,0001	10000	1,00E-04	1665	5,37E-04
14	0,0001	10000	1,00E-04	1806	5,39E-04
15	0,0001	10000	1,00E-04	1815	5,17E-04
16	0,0001	10000	1,00E-04	2039	5,18E-04
17	0,0001	10000	1,00E-04	1943	5,44E-04
18	0,0001	10000	1,00E-04	2087	5,70E-04
19	0,0001	10000	1,00E-04	3591	5,26E-04
20	0,0001	10000	1,00E-04	1764	5,47E-04
21	0,0001	10000	1,00E-04	3779	5,34E-04
22	0,0001	10000	1,00E-04	2127	5,29E-04
23	0,0001	10000	1,00E-04	2119	5,53E-04

Figura 26 – Erro médio quadrático da rede (10- $x$ -1) para o ponto 5.

O valor do erro quadrático médio (*Total Mean Square Error*), calculado a partir dos erros obtidos nos testes com a rede de arquitetura (10-8-1), é 4,87E-4. A Figura 27 demonstra o número de neurônios necessários para que a rede alcance o valor de erro desejado e o treinamento cesse, ou, quando não alcançado o treinamento, atinja o número máximo de iterações permitido.

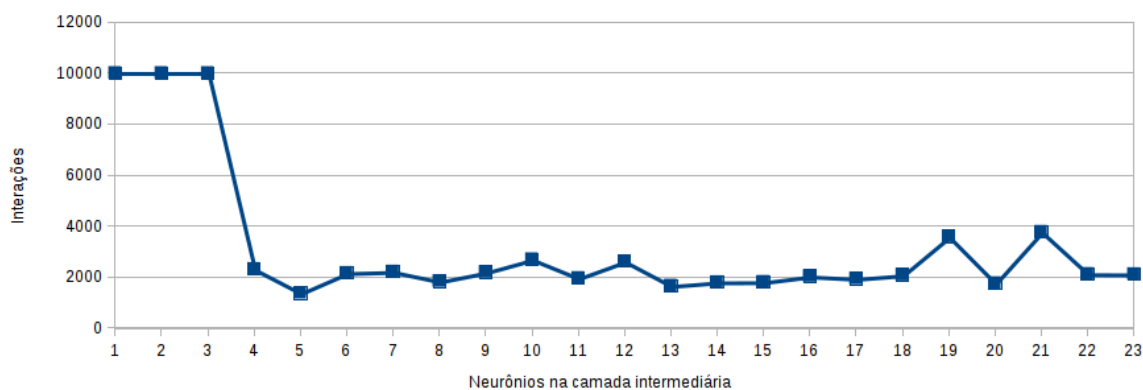


Figura 27 – Número de iterações em função da quantidade de neurônios na camada intermediária para o ponto 5.

#### 4.2.5.2 Comparação do modelo com dados reais

A partir do treinamento da rede de arquitetura (10-8-1), é possível normalizar inversamente os dados de saída (*output*) da rede com a finalidade de verificar a distância destes dados com o alvo do treinamento, o valor real de Oxigênio Dissolvido. A Figura 28 apresenta estes valores iniciais da variável O.D., antes de serem padronizados e utilizados no conjunto de teste, contra o O.D. modelado pela rede, depois da efetuada a normalização reversa. A Figura 29 ilustra a dispersão dos mesmos dados. Verifica-se que com erro quadrático médio pequeno desta arquitetura as diferenças entre os valores originais e modelados são pequenas, o que pode ser confirmado pelo valor de coeficiente de correlação calculado em 0,992, validando o método de modelagem através de RNA.

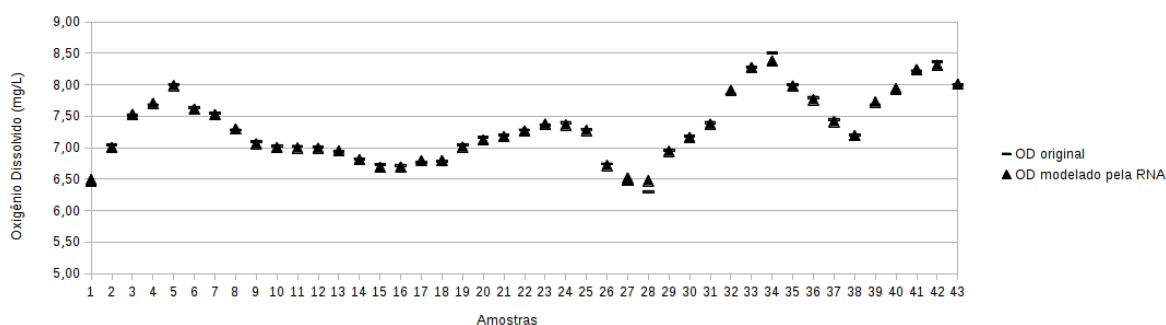


Figura 28 – Dado original contra dado modelado pela RNA para o ponto 5.

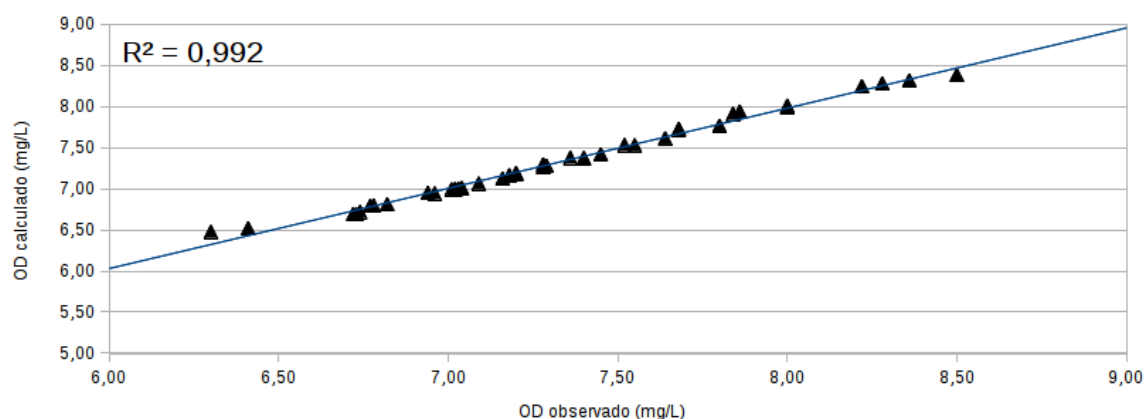


Figura 29 – Dispersão de dado original contra dado modelado pela RNA para o ponto 5.

## 4.2.6 Ponto de amostragem 11

Os dados originais são oriundos de análise de água coletada neste ponto de amostragem do rio, submetidos à técnica de interpolação linear para a produção de mais valores para aumentar o conjunto de treinamento. Antes do treinamento, o conjunto é padronizado (normalizados entre  $[0,1]$ ), sendo posteriormente separados em dois conjuntos através do arquivo contendo os dados de treinamento, que visam promover o aprendizado da RNA, e de outro contendo os dados de teste ou verificação, responsáveis por verificar a capacidade de generalização da rede, e outro os dados de teste:

- Ponto de amostragem 11 - 141 amostras de 11 variáveis, conforme a seção 3.1.1:
  - Dados para treinamento: 98 amostras (70%);
  - Dados para verificação: 43 amostras (30%).

### 4.2.6.1 Resultados

A Tabela 12 apresenta os dados resultantes do treinamento das RNA com dez entradas, uma saída e uma camada intermediária, variando o número de neurônios nesta camada entre 1 e 23. É calculado o desvio padrão para os resultados de Erro Quadrático dos testes efetuados, cujo resultado é  $2,94E-03$ .

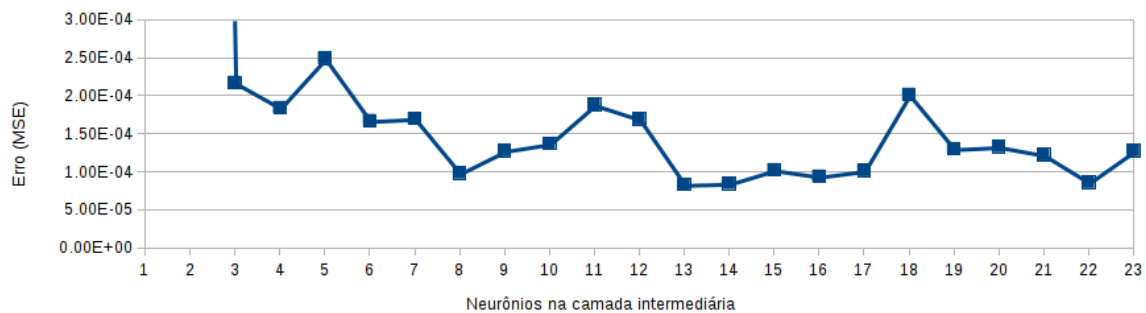
O gráfico da Figura 30 apresenta a dinâmica do Erro Quadrático Médio em função do número de neurônios da camada intermediária. Os treinamentos das redes nas arquiteturas (10-1-1) e (10-2-1) produziram valores altos de erro quadrático, motivo pelo qual são mantidos fora do gráfico, para comportar melhor a escala dos demais valores. O menor erro é encontrado com a utilização de treze neurônios na camada intermediária (10-13-1).



Tabela 12 – Dados médios obtidos com o treinamento da rede (10- $x$ -1) para o Ponto 11.

Neurônios na camada intermediária ( $x$ )	Erro Máx.	Limite máx. iteração	Erro Total da Rede (70%)	Iteração	Erro Quadrático dos Testes (30%)
1	0,0001	10000	9,12E-03	10000	1,41E-02
2	0,0001	10000	1,92E-03	10000	3,26E-03
3	0,0001	10000	1,91E-04	10000	2,17E-04
4	0,0001	10000	2,88E-04	10000	1,84E-04
5	0,0001	10000	1,00E-04	3085	2,49E-04
6	0,0001	10000	1,25E-04	10000	1,67E-04
7	0,0001	10000	1,17E-04	10000	1,70E-04
8	0,0001	10000	1,00E-04	2316	9,83E-05
9	0,0001	10000	1,01E-04	10000	1,28E-04
10	0,0001	10000	1,00E-04	3444	1,37E-04
11	0,0001	10000	1,00E-04	3297	1,88E-04
12	0,0001	10000	1,00E-04	4826	1,69E-04
13	0,0001	10000	1,00E-04	3631	8,32E-05
14	0,0001	10000	1,00E-04	6889	8,46E-05
15	0,0001	10000	1,00E-04	2671	1,02E-04
16	0,0001	10000	1,00E-04	2716	9,40E-05
17	0,0001	10000	1,03E-04	10000	1,01E-04
18	0,0001	10000	1,50E-04	10000	2,02E-04
19	0,0001	10000	1,00E-04	6803	1,30E-04
20	0,0001	10000	1,16E-04	10000	1,33E-04
21	0,0001	10000	1,00E-04	4314	1,22E-04
22	0,0001	10000	1,00E-04	2872	8,60E-05
23	0,0001	10000	1,00E-04	7225	1,28E-04

Esta arquitetura não é a que convergiu com o menor número de iterações, mas a que resultou em menor diferença entre o valor real e o modelo da RNA.

Figura 30 – Erro médio quadrático da rede (10- $x$ -1) para o ponto 11.

Ao final do teste com a arquitetura (10-13-1), a aplicação disponibiliza o valor do erro quadrático médio (*Total Mean Square Error*) calculado a partir dos erros obtidos do teste. O resultado obtido é 8,32E-5. A Figura 31 demonstra o número de neurônios necessários para que a rede alcance o valor de erro desejado e o treinamento cesse, ou, quando não alcançado o treinamento, atinja o número máximo de iterações permitido.

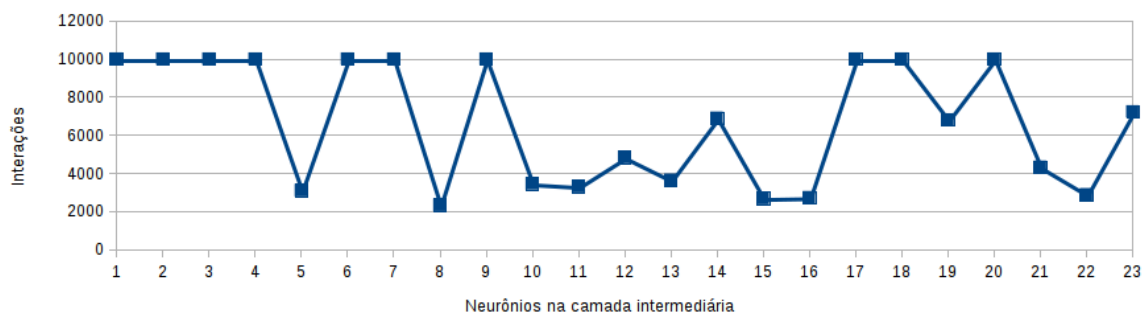


Figura 31 – Número de iterações em função da quantidade de neurônios na camada intermediária para o ponto 11.

#### 4.2.6.2 Comparação do modelo com dados reais

A partir do treinamento da rede de arquitetura (10-13-1), é possível normalizar inversamente os dados de saída (*output*) da rede com a finalidade de verificar a distância destes dados com o alvo do treinamento, o valor real de Oxigênio Dissolvido. A Figura 32 apresenta estes valores iniciais da variável O.D., antes de serem padronizados e utilizados no conjunto de teste, contra o O.D. modelado pela rede, depois da efetuada a normalização reversa. Pode-se verificar nesta figura que a amostra de número 6 apresenta um comportamento anômalo, sendo considerada um *outlier*, presente no conjunto original de dados que não foi modificado para a aplicação da RNA. Optou-se por não remover esta amostra. A Figura 33 ilustra a dispersão dos mesmos dados. Verifica-se que com erro quadrático médio baixo desta arquitetura, as diferenças entre os valores originais e modelados são pequenas, o que pode ser confirmado pelo valor de coeficiente de correlação calculado em 0,920, validando o método de modelagem através de RNA.

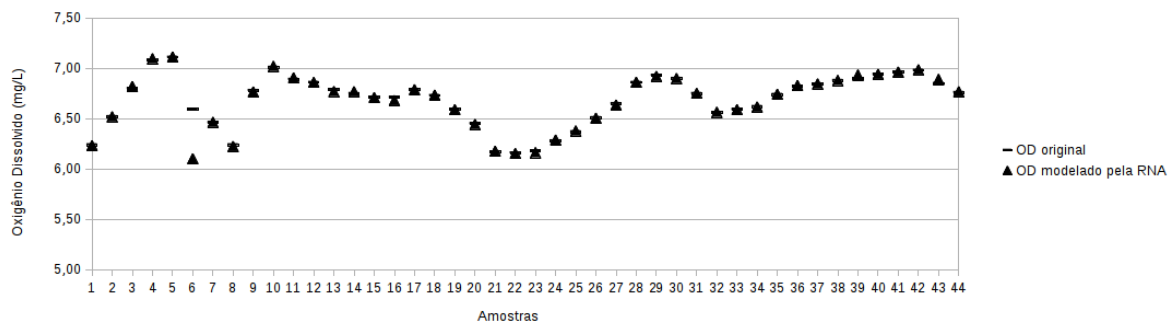


Figura 32 – Dado original contra dado modelado pela RNA para o ponto 11.

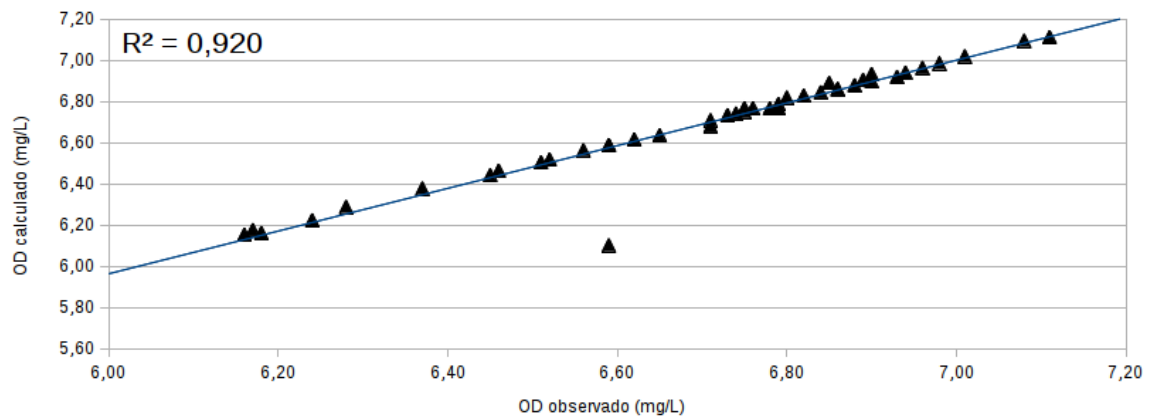


Figura 33 – Dado original contra dado modelado pela RNA para o ponto 11.

### 4.3 Experimento 2 - dados reais

Esta seção descreve a utilização do conjunto de dados do Rio Paraíba do Sul (SP), sendo este mais completo e abrangente com maior escala de variação dos parâmetros, para o treinamento de diversas arquiteturas de RNA com várias configurações de arquitetura; variou-se o número de neurônios na camada intermediária entre 1 e 23.

São utilizados dez sinais (variáveis) normalizados na entrada da RNA:

- Temperatura da água;
- pH;
- Demanda Bioquímica de Oxigênio - DBO;
- Condutividade;
- Nitrogênio Amoniacal;
- Nitrato;
- Nitrito;
- Fósforo;
- Turbidez;
- Sólidos Dissolvidos Totais;

E apenas um sinal (variável) normalizado de saída:

- Oxigênio Dissolvido.

### 4.3.1 Configuração para o treinamento

Para cada experimento, a rede é treinada três vezes para aumentar a probabilidade de produzir um bom conjunto inicial de pesos. Uma vez obtidos os resultados, estes valores são utilizados para gerar uma média dos três experimentos; os parâmetros de inicialização do algoritmo são padronizados para todas as redes, com vistas a permitir melhor comparação entre as arquiteturas, conforme apresentado na Tabela 13.

Tabela 13 – Configuração para treinamento da RNA do Experimento 2

Variáveis de treinamento	Limite
Erro máximo ( <i>Max error</i> )	0,0001
Limite de iteração ( <i>Iteration</i> ) ou época	10.000
Taxa de aprendizagem ( <i>Learning rate</i> )	0,2
Momento ( <i>Momentum</i> )	0,7

Para o propósito deste projeto, o número máximo de épocas é configurado em  $10^4$ . O parâmetro “erro máximo” é configurado para  $10^{-4}$ , pois pretende-se alcançar erro baixo o suficiente para que o modelo da rede seja o mais próximo do valor real. As variáveis “taxa de aprendizagem” e “momento” são configuradas com os valores padrão da aplicação, 0,2 e 0,7, respectivamente, pois valores menores nestes parâmetros produzem uma variação muito pequena nos pesos sinápticos, corrigidos a cada ciclo do algoritmo de retropropagação, elevando o número de iterações necessárias para se aproximar da convergência.

Cada um dos experimentos devem parar quando **uma** das seguintes condições for atingida:

- ❑ Valor configurado da variável “Erro Máximo” for alcançado.
- ❑ Valor configurado da variável “Limite de iteração” for alcançado;

### 4.3.2 Ponto de amostragem PARB 02310

Os dados utilizados neste experimento correspondem a um conjunto de 155 amostras de onde variáveis diferentes, conforme a seção 3.1.1. Antes de sua utilização no treinamento das redes neurais, o conjunto é padronizado (normalizados entre  $[0,1]$ ), posteriormente separados em dois conjuntos através do arquivo contendo os dados de treinamento, que visam promover o aprendizado da RNA, e outro contendo os dados de teste ou verificação, responsáveis por verificar a capacidade de generalização da rede, e outro os dados de teste:

- ❑ Dados para treinamento: 109 amostras (70%);
- ❑ Dados para verificação: 46 amostras (30%).

### 4.3.2.1 Resultados

A Tabela 14 apresenta os dados resultantes do treinamento das RNA com dez entradas, uma saída e uma camada intermediária, variando o número de neurônios nesta camada entre 1 e 23. O resultado obtido no cálculo do desvio padrão para os resultados de Erro Quadrático dos testes efetuados é  $1,58E-02$ .

Tabela 14 – Dados médios obtidos com o treinamento da rede (10- $x$ -1) para o Rio Paraíba do Sul.

Neurônios na camada intermediária ( $x$ )	Erro Máx.	Limite máx. iteração	Erro Total da Rede (70%)	Iteração	Erro Quadrático dos Testes (30%)
1	0,0001	10000	7,59E-03	10000	3,05E-02
2	0,0001	10000	7,78E-03	10000	3,57E-02
3	0,0001	10000	5,32E-03	10000	2,78E-02
4	0,0001	10000	3,61E-03	10000	3,48E-02
5	0,0001	10000	3,16E-03	10000	5,20E-02
6	0,0001	10000	2,17E-03	10000	5,42E-02
7	0,0001	10000	1,77E-03	10000	5,30E-02
8	0,0001	10000	1,23E-03	10000	4,21E-02
9	0,0001	10000	1,00E-03	10000	5,38E-02
10	0,0001	10000	8,65E-04	10000	7,35E-02
11	0,0001	10000	3,54E-04	10000	6,94E-02
12	0,0001	10000	6,54E-04	10000	5,64E-02
13	0,0001	10000	8,44E-04	10000	6,69E-02
14	0,0001	10000	2,60E-04	10000	5,94E-02
15	0,0001	10000	5,30E-04	10000	5,18E-02
16	0,0001	10000	4,01E-04	10000	7,18E-02
17	0,0001	10000	5,77E-04	10000	8,22E-02
18	0,0001	10000	6,77E-04	10000	8,50E-02
19	0,0001	10000	6,86E-04	10000	4,95E-02
20	0,0001	10000	2,55E-04	10000	5,28E-02
21	0,0001	10000	7,48E-04	10000	4,57E-02
22	0,0001	10000	5,70E-04	10000	5,64E-02
23	0,0001	10000	2,17E-04	10000	7,65E-02

O gráfico da Figura 34 apresenta a dinâmica do Erro Total da Rede, registrado durante o treinamento, e do Erro Quadrático Médio<sup>2</sup>, registrado durante os testes, em função do número de neurônios da camada intermediária. A arquitetura (10-8-1) é a escolhida para a realização do teste e normalização inversa do resultado. A arquitetura não é a que apresentou o menor erro quadrático dos testes, mas a que apresentou um dos menores erros e tem melhores condições para finalizar o treinamento que as arquiteturas com poucos neurônios na camada intermediária, uma vez que, de acordo com a Seção 2.4.4, a literatura confirma que a melhor solução terá a quantidade de neurônios próxima do número de entradas na rede. As redes testadas não convergiram para o menor erro configurado, com as configurações utilizadas no treinamento. Todos os testes terminaram por atingir o limite máximo de iterações (10.000).

<sup>2</sup> a diferença entre a estimativa e o valor real elevada ao quadrado.

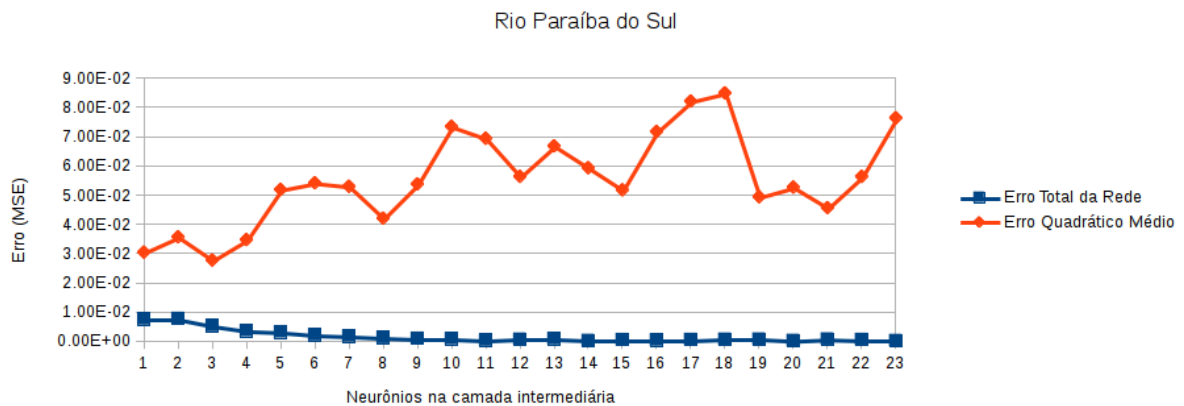


Figura 34 – Erro Total da Rede e Erro Médio Quadrático da rede (10- $x$ -1) para o Rio Paraíba do Sul.

Ao final do teste, a aplicação disponibiliza o valor do erro quadrático médio (*Total Mean Square Error*), calculado a partir dos erros obtidos do teste. No caso da arquitetura (10-8-1) para o Rio Paraíba do Sul, este erro é de 4,21E-02.

#### 4.3.2.2 Comparação do modelo com dados reais

A partir do treinamento da rede de arquitetura (10-8-1), é possível normalizar inversamente os dados de saída (*output*) da rede com a finalidade de verificar a distância destes dados com o alvo do treinamento, o valor real de Oxigênio Dissolvido. A Figura 35 apresenta a plotagem dos valores iniciais da variável O.D., antes de serem padronizados e utilizados no conjunto de teste, contra o O.D. modelado pela rede, depois da efetuada a normalização reversa. A Figura 36 ilustra a dispersão dos mesmos dados. Verifica-se que com erro quadrático médio alto desta arquitetura, as diferenças entre os valores originais e modelados são consideráveis, o que pode ser confirmado pelo valor do coeficiente de correlação, calculado em 0,428, suscitando a necessidade de mudanças nos parâmetros de treinamento em busca de valor de erro menor.

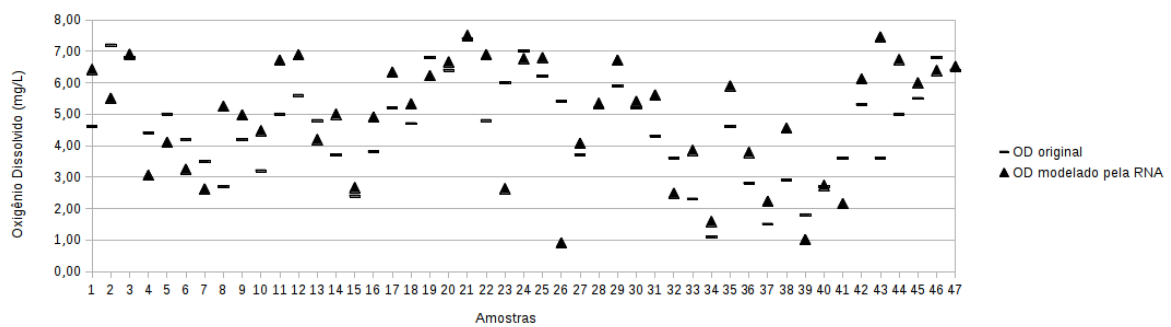


Figura 35 – Dado original contra dado modelado pela RNA para o Rio Paraíba do Sul.

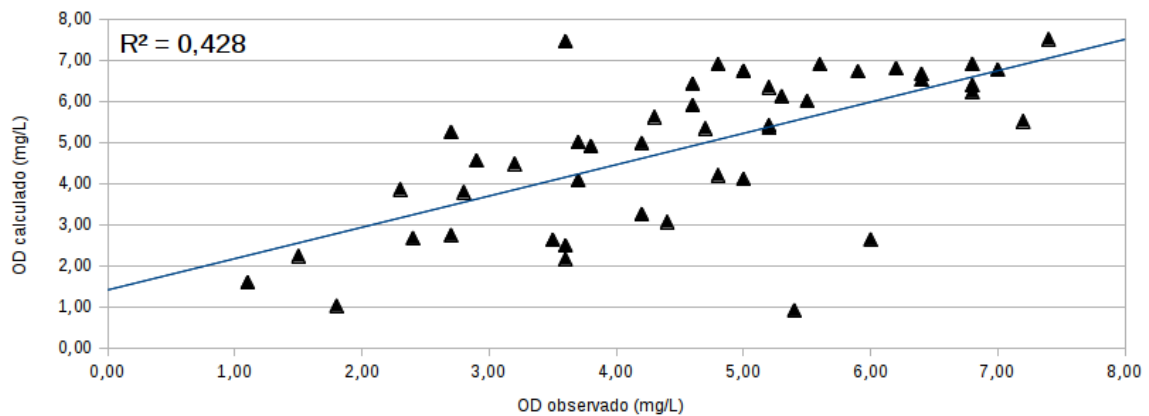


Figura 36 – Dispersão do dado original contra dado modelado pela RNA para o Rio Paraíba do Sul.

## 4.4 Experimento 3 - PCA

De acordo com a seção 2.2.1, criou-se um experimento buscando verificar se, após submetidos os dados das amostras originais à técnica PCA, o resultado poderia reduzir o tempo ou custo computacional de treinamento da rede, tendo em vista a redução do número de entradas da RNA. Isto se dá em função de que, originalmente as entradas da RNA são os próprios parâmetros físico-químicos analisados nas amostras de água; porém a técnica de PCA permite descrever os mesmos dados em um número de Componentes Principais menor que a quantidade de parâmetros de entrada e que contenha maior variância dos dados, como pode ser observado neste resultado. Desta forma, o resultado da técnica de PCA aplicada aos dados oriundos das análises de água para os pontos de amos-

tragem 3 e 11 do Rio Meia Ponte (GO) e o único ponto de coleta de amostras avaliado no Rio Paraíba do Sul (SP), forneceram os dados para o treinamento das redes.

Para cada experimento, a rede é treinada três vezes para aumentar a probabilidade de produzir um bom conjunto inicial de pesos, obtidos os valores dos resultados, que são utilizados para gerar uma média; os parâmetros de inicialização do algoritmo são padronizados para permitir melhor comparação entre as arquiteturas, como verifica-se na Tabela 15.

Tabela 15 – Configuração para treinamento da RNA do Experimento 3

Variáveis de treinamento	Limite
Erro máximo ( <i>Max error</i> )	0,0001
Limite de iteração ( <i>Iteration</i> ) ou época	10000
Taxa de aprendizagem ( <i>Learning rate</i> )	0,2
Momento ( <i>Momentum</i> )	0,7

Cada um dos experimentos devem parar quando **uma** das seguintes condições for atingida:

- Valor configurado da variável “Erro Máximo” for alcançado.
- Valor configurado da variável “Limite de iteração” for alcançado;

#### 4.4.1 Ponto de amostragem 3

As operações que se seguiram no software *SciLab* são:

- Carregamento dos dados originais;
- Interpolação dos dados, gerando 141 amostras;
- Aplicação do PCA;
- Avaliação da matriz de variância dos CP;
- Avaliação da matriz de dados.

A Tabela 16 mostra o percentual acumulado de variância para cada número de CP, obtida através da matriz lambda, que representa respectivamente os autovalores e sua variância (dividida por 100%). O número de linhas da matriz representa o componente principal (CP), ou seja, primeira linha, primeira CP; segunda linha, segunda CP, respectivamente.



Tabela 16 – Porcentagem de variância do ponto 3.

Nome	Autovalores	Variância acumulada (%)
CP Número 1	2,824E-01	28,24
CP Número 2	2,149E-01	49,74
CP Número 3	1,741E-01	67,15
CP Número 4	1,043E-01	77,59
CP Número 5	9,051E-02	86,64
CP Número 6	5,802E-02	92,44
CP Número 7	3,666E-02	96,11
CP Número 8	3,191E-02	99,30
CP Número 9	5,814E-03	99,88
CP Número 10	1,213E-03	100,00
CP Número 11	1,108E-09	100,00

#### 4.4.1.1 Configuração para o treinamento

Como se pode observar na Tabela 16, os seis primeiros componentes principais da matriz  $\lambda$ , resultado da aplicação da técnica de PCA, representam uma variância acumulada total de 92,44% de dez parâmetros originais. Deste modo, as RNA são treinadas e verificadas para várias configurações de arquitetura; variou-se o número de neurônios na camada intermediária entre 1 e 15. Para simplificar a notação, adota-se a codificação de acordo com o seguinte exemplo: (6-1-1), ou seja, seis sinais na camada de entrada, um neurônio na camada intermediária e um sinal na camada de saída. Na segunda configuração, acrescentou-se um neurônio na camada intermediária (6-2-1) e assim sucessivamente, até a configuração (6-15-1). Os parâmetros de inicialização do algoritmo são padronizados para este experimento, como verifica-se na Tabela 15.

Utilizou-se na entrada da RNA seis sinais (variáveis) normalizadas de entrada.

□ *Scores* de 1 a 6 de acordo com a Tabela 16.

E apenas um sinal (variável) normalizado de saída:

□ Oxigênio Dissolvido.

#### 4.4.1.2 Amostras de dados para PCA

Desta forma, a partir da análise da matriz  $\lambda$ , são separadas as seis primeiras colunas da matriz de *scores* gerada no PCA; em seguida, adicionou-se a coluna do oxigênio dissolvido, que serve como referência para o aprendizado supervisionado da rede (saída), normalizou-se os dados, para serem utilizados no treinamento, em dois conjuntos:

□ Ponto de amostragem 3 - 141 amostras de 6 *scores* para 6 CP:

- Dados para treinamento: 99 amostras (70%);
- Dados para verificação: 42 amostras (30%).

#### 4.4.1.3 Resultados

A Tabela 17 e a Figura 37 apresentam os dados obtidos no treinamento das RNA com seis entradas e uma camada intermediária, variando o número de neurônios desta camada entre 1 e 15, dispostos na forma de informações sobre o erro da rede e o erro do treinamento em função do número de neurônios da camada intermediária. Nota-se que apesar da rede convergir para o treinamento em algumas arquiteturas, o erro quadrático médio dos testes é sempre mais alto do que o pretendido para o experimento ( $1,0E-04$ ).

Tabela 17 – Dados médios obtidos com o treinamento da rede (6- $x$ -1) para o Ponto 3 - PCA.

Neurônios na camada intermediária ( $x$ )	Erro Máx.	Limite máx. iteração	Erro Total da Rede (70%)	Iteração	Erro Quadrático dos Testes (30%)
1	0,0001	10000	6,29E-03	10000	1,56E-02
2	0,0001	10000	1,91E-03	10000	6,23E-03
3	0,0001	10000	7,50E-04	10000	2,66E-03
4	0,0001	10000	1,65E-04	7888	1,22E-03
5	0,0001	10000	1,27E-04	5977	1,09E-03
6	0,0001	10000	2,58E-04	7260	2,22E-03
7	0,0001	10000	1,75E-04	7424	1,33E-03
8	0,0001	10000	1,24E-04	7916	1,20E-03
9	0,0001	10000	1,07E-04	5998	6,98E-04
10	0,0001	10000	1,00E-04	4931	7,82E-04
11	0,0001	10000	1,06E-04	6848	5,27E-04
12	0,0001	10000	1,00E-04	5318	8,56E-04
13	0,0001	10000	1,00E-04	3892	7,75E-04
14	0,0001	10000	1,14E-04	5337	6,60E-04
15	0,0001	10000	1,15E-04	7993	7,89E-04

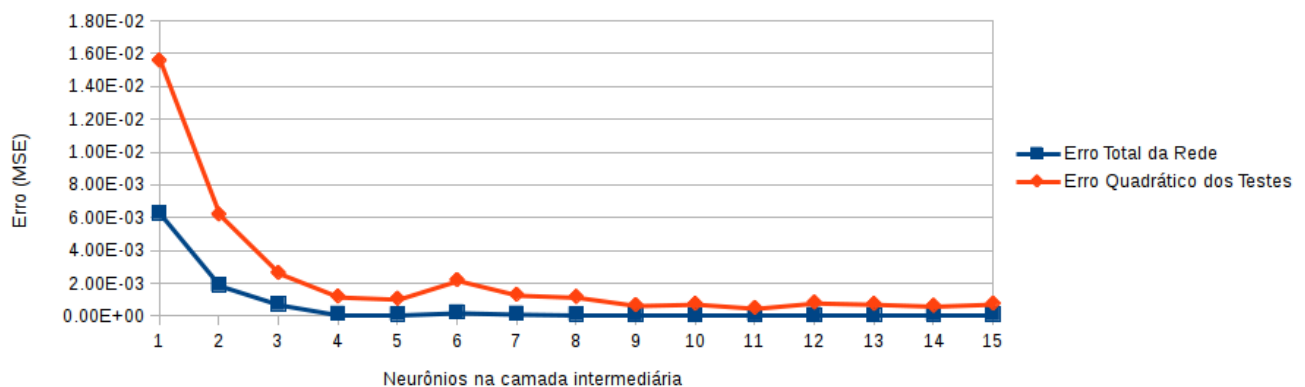


Figura 37 – Erro Total da Rede e Erro Médio Quadrático da rede (6- $x$ -1) para o ponto 3 - PCA.

#### 4.4.1.4 Comparação do modelo com dados reais

A partir do treinamento da rede de arquitetura (6-11-1), é possível normalizar inversamente os dados de saída da rede com a finalidade de verificar a distância destes dados com o alvo do treinamento, o valor real de Oxigênio Dissolvido. A Figura 38 apresenta estes valores iniciais da variável O.D., antes de serem padronizados e utilizados no conjunto de teste, contra o O.D. modelado pela rede, depois da efetuada a normalização reversa. Nota-se que a linha que representa os dados originais mostra-se mais linear que os gráficos anteriores. Isto se dá em função de que os dados para teste deste experimento não são separados pelo programa Neuroph, mas manualmente. Quando se utiliza a função de separação de dados para treinamento/teste deste *software*, o mesmo faz a separação de modo aleatório dentro do conjunto total de dados, evitando a linearidade aqui constatada. A Figura 39 ilustra a dispersão dos mesmos dados. Verifica-se que apesar do erro quadrático médio desta arquitetura, apresentado na Figura 37, não ter atingido valores tão baixos quanto o erro total de treinamento da rede, as diferenças entre os valores originais e modelados são baixas, o que pode ser confirmado pelo valor de coeficiente de correlação calculado em 0,989, validando o método de modelagem através de RNA usando dados advindos da técnica de PCA.

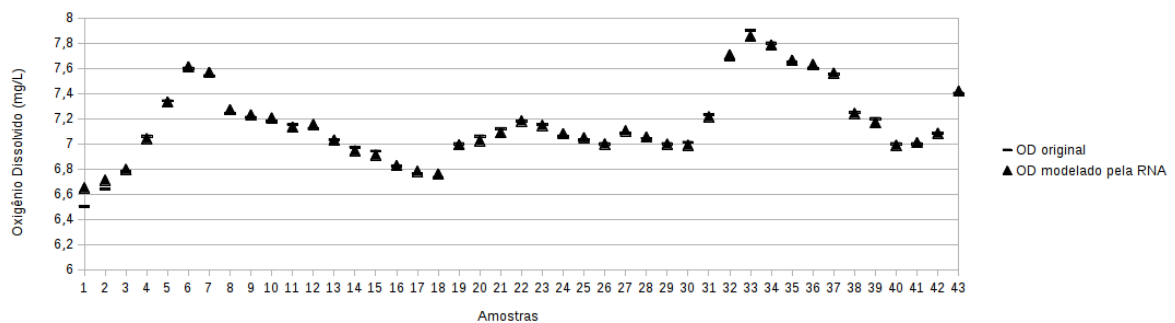


Figura 38 – Dado original contra dado modelado pela RNA para o Ponto 3 - PCA.

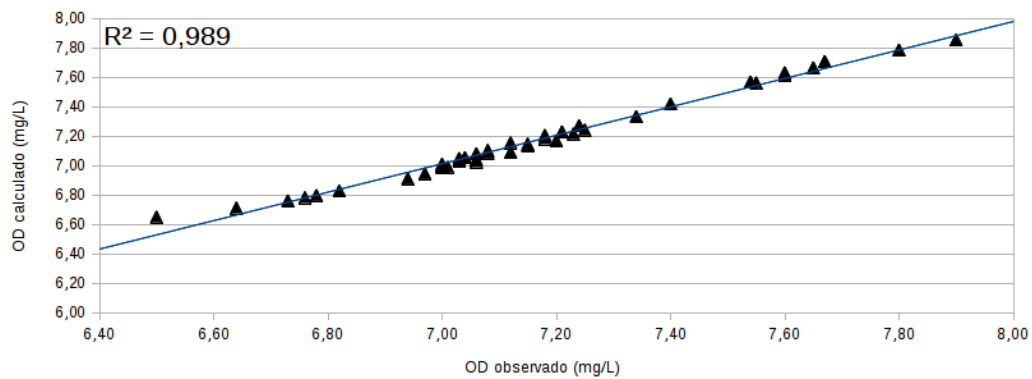


Figura 39 – Dispersão do dado original contra dado modelado pela RNA para o Ponto 3 - PCA.

#### 4.4.2 Ponto de amostragem 11

As operações que se seguiram no software *SciLab* são:

- Carregamento dos dados originais;
- Interpolação dos dados, gerando 141 amostras;
- Aplicação do PCA;
- Avaliação da matriz de variância dos CP;
- Avaliação da matriz de dados.

A Tabela 18 mostra o percentual acumulado de variância para cada número de CP, obtida através da matriz *lambda*, que representa respectivamente os autovalores e sua variância (dividida por 100%). O número de linhas da matriz representa o componente principal (CP), ou seja, primeira linha, primeira CP; segunda linha, segunda CP, respectivamente.

##### 4.4.2.1 Configuração para o treinamento

Como se pode observar na Tabela 18, os cinco primeiros componentes principais da matriz *lambda*, resultado do PCA, representam uma variância acumulada total de 92,14% dos dados originais. Deste modo, as RNA são treinadas e verificadas para várias configurações de arquitetura; variou-se o número de neurônios na camada intermediária entre 1 e 15. Para simplificar a notação, adota-se a codificação de acordo com o seguinte exemplo: (5-1-1), ou seja, cinco sinais na camada de entrada, um neurônio na camada intermediária e um sinal na camada de saída. Na segunda configuração, acrescentou-se um neurônio na camada intermediária (5-2-1) e assim sucessivamente, até a configuração (5-15-1). Os

Tabela 18 – Porcentagem de variância para o ponto 11.

Nome	Autovalores	Variância acumulada (%)
CP Número 1	4,388E-01	43,88
CP Número 2	1,860E-01	62,49
CP Número 3	1,369E-01	76,18
CP Número 4	9,359E-02	85,54
CP Número 5	6,594E-02	92,14
CP Número 6	3,809E-02	95,94
CP Número 7	1,895E-02	97,84
CP Número 8	1,061E-02	98,90
CP Número 9	8,609E-03	99,76
CP Número 10	2,364E-03	100,00
CP Número 11	1,394E-09	100,00

parâmetros de inicialização do algoritmo são padronizados para este experimento, como verifica-se na Tabela 15:

Utilizou-se na entrada da RNA seis sinais (variáveis) normalizadas de entrada.

□ *Scores* de 1 a 5, de acordo com a Tabela 18.

E apenas um sinal (variável) normalizado de saída.

□ Oxigênio Dissolvido.

Cada um dos experimentos devem parar quando **uma** das seguintes condições for atingida:

□ Valor configurado da variável “Erro Máximo” for alcançado.

□ Valor configurado da variável “Limite de iteração” for alcançado;

#### 4.4.2.2 Amostras de dados para PCA

Desta forma, a partir da análise da matriz *lambda*, são separadas as cinco primeiras colunas da matriz de *scores* gerada no PCA; em seguida, adicionou-se a coluna do oxigênio dissolvido, que serve como referência para o aprendizado supervisionado da rede (saída), normalizou-se os dados, para serem utilizados no treinamento, em dois conjuntos:

□ Ponto de amostragem 11 - 141 amostras de 5 *scores* para 5 CP:

– Dados para treinamento: 99 amostras (70%);

– Dados para verificação: 42 amostras (30%).

#### 4.4.2.3 Resultados

A Figura 40 apresenta o erro total da rede e o erro do treinamento em função do número de neurônios da camada intermediária. Os dados são obtidos no treinamento das RNA com seis entradas e uma camada intermediária, variando o número de neurônios desta camada entre 1 e 15, conforme apresentado na Tabela 19. Nota-se que na medida em que são acrescentados neurônios na camada intermediária, a rede aumenta o poder de processamento, o que produz a redução de ambos os erros. O menor erro quadrático médio é encontrado na arquitetura (5-15-1), com um dos menores números de iterações do conjunto.

Tabela 19 – Dados médios obtidos com o treinamento da rede (5- $x$ -1) para o Ponto 11 - PCA.

Neurônios na camada intermediária ( $x$ )	Erro Máx.	Limite máx. iteração	Erro Total da Rede (70%)	Iteração	Erro Quadrático dos Testes (30%)
1	0,0001	10000	3,21E-03	10000	4,56E-03
2	0,0001	10000	1,86E-03	10000	3,05E-03
3	0,0001	10000	6,79E-04	10000	2,34E-03
4	0,0001	10000	2,75E-04	10000	1,06E-03
5	0,0001	10000	2,09E-04	10000	8,90E-04
6	0,0001	10000	1,34E-04	7808	7,19E-04
7	0,0001	10000	1,43E-04	8803	5,51E-04
8	0,0001	10000	1,12E-04	7407	4,48E-04
9	0,0001	10000	1,38E-04	8065	6,18E-04
10	0,0001	10000	1,55E-04	8191	6,59E-04
11	0,0001	10000	1,48E-04	8743	5,83E-04
12	0,0001	10000	1,31E-04	8824	5,24E-04
13	0,0001	10000	1,14E-04	9190	4,77E-04
14	0,0001	10000	1,24E-04	9370	4,42E-04
15	0,0001	10000	1,24E-04	7924	4,40E-04

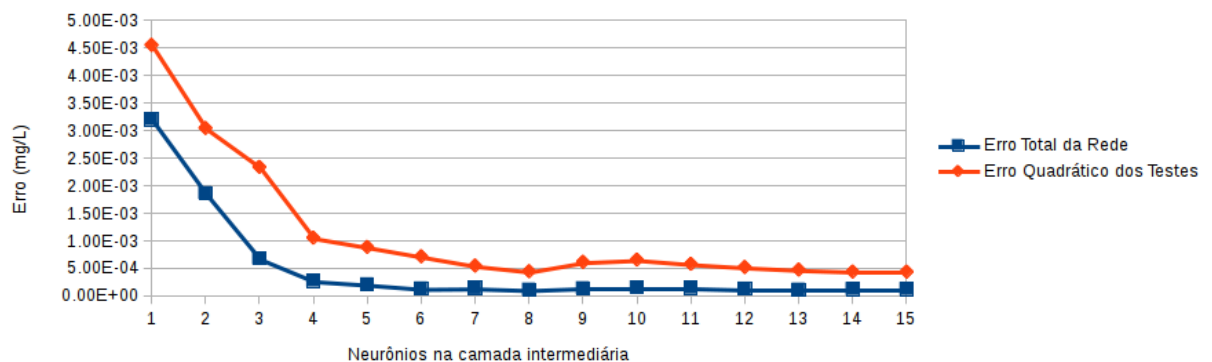


Figura 40 – Erro Total da Rede e Erro Médio Quadrático da rede (5- $x$ -1) em função do número de neurônios na camada intermediária para o ponto 11 - PCA.

#### 4.4.2.4 Comparação do modelo com dados reais

A partir do treinamento da rede de arquitetura (5-15-1), é possível normalizar inversamente os dados de saída da rede com a finalidade de verificar a distância destes dados com o alvo do treinamento, o valor real de Oxigênio Dissolvido. A Figura 41 apresenta estes valores iniciais da variável O.D., antes de serem padronizados e utilizados no conjunto de teste, contra o O.D. modelado pela rede, depois da efetuada a normalização reversa. A Figura 42 ilustra a dispersão dos mesmos dados. Assim como no experimento com PCA do Ponto de amostragem 3, é possível afirmar que a RNA conseguiu correlacionar as cinco entradas de Componentes Principais com a saída, modelando corretamente e com nível de erro baixo o valor de O.D., o que pode ser confirmado pelo valor de coeficiente de correlação calculado em 0,993. O Erro Quadrático dos Testes deste experimento, apresentado na Figura 40, é fruto da maior variação dos parâmetros físico-químicos deste ponto de amostragem, situado dentro da capital goiana.

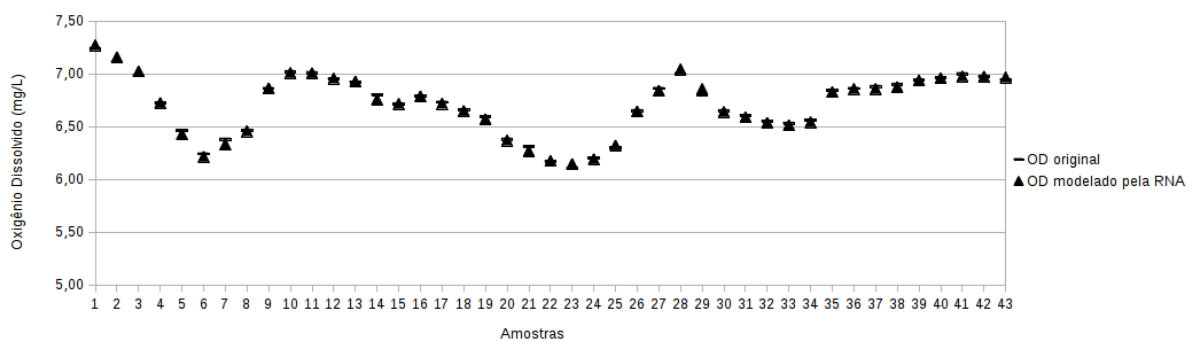


Figura 41 – Dado original contra dado modelado pela RNA para o Ponto 11 - PCA.

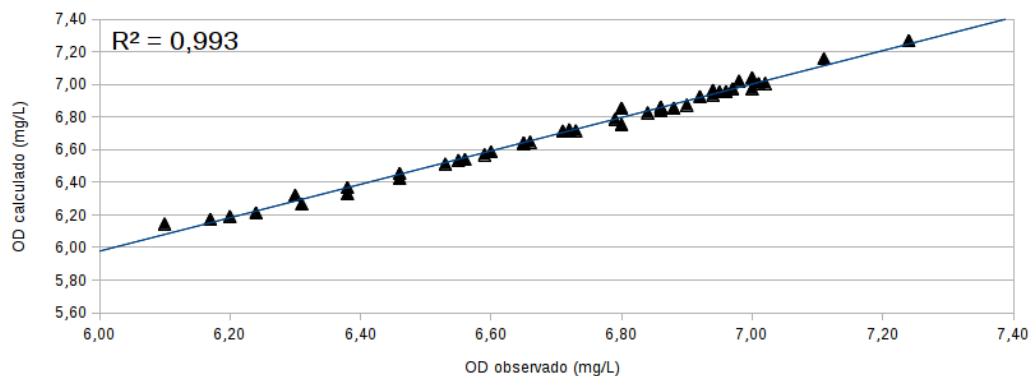


Figura 42 – Dispersão do dado original contra dado modelado pela RNA para o Ponto 11 - PCA.

### 4.4.3 Rio Paraíba do Sul

As operações que se seguiram no software *SciLab* são:

- ❑ Carregamento dos dados originais;
- ❑ Aplicação do PCA;
- ❑ Avaliação da matriz de variância dos CP;
- ❑ Avaliação da matriz de dados.

A Tabela 20 mostra o percentual acumulado de variância para cada número de CP, obtida através da matriz *lambda*, que representa respectivamente os autovalores e sua variância (dividida por 100%). O número de linhas da matriz representa o componente principal (CP), ou seja, primeira linha, primeira CP; segunda linha, segunda CP, respectivamente.

Tabela 20 – Porcentagem de variância para o Rio Paraíba do Sul.

Nome	Autovalores	Variância acumulada (%)
CP Número 1	2,730E-01	27,31
CP Número 2	1,422E-01	41,53
CP Número 3	1,204E-01	53,58
CP Número 4	9,083E-02	62,67
CP Número 5	8,881E-02	71,55
CP Número 6	7,473E-02	79,02
CP Número 7	5,872E-02	84,89
CP Número 8	5,124E-02	90,02
CP Número 9	4,662E-02	94,68
CP Número 10	3,243E-02	97,92
CP Número 11	2,077E-02	100,00

#### 4.4.3.1 Configuração para o treinamento

Como se pode observar na Tabela 20, os oito primeiros componentes principais da matriz *lambda*, resultado do PCA, representam uma variância acumulada total de 90,02% dos dados originais. Deste modo, as RNA são treinadas e verificadas para várias configurações de arquitetura; variou-se o número de neurônios na camada intermediária entre 1 e 15. Para simplificar a notação, adota-se a codificação de acordo com o seguinte exemplo: (8-1-1), ou seja, oito sinais na camada de entrada, um nó na camada intermediária e um sinal na camada de saída. Na segunda configuração, acrescentou-se um nó na camada intermediária (8-2-1) e assim sucessivamente, até a configuração (8-15-1). Os parâmetros de inicialização do algoritmo são padronizados para este experimento, como apresentado na Tabela 15:

Utilizou-se na entrada da RNA seis sinais (variáveis) normalizadas de entrada.



- *Scores* de 1 a 8.

E apenas um sinal (variável) normalizado de saída.

- Oxigênio Dissolvido.

Cada um dos experimentos devem parar quando uma das seguintes condições for atingida:

- Valor configurado da variável “Erro Máximo” for alcançado.
- Valor configurado da variável “Limite de iteração” for alcançado;

#### 4.4.3.2 Amostras de dados para PCA

Desta forma, a partir da análise da matriz lambda, são separadas as oito primeiras colunas da matriz de *scores* gerada no PCA; em seguida, adicionou-se a coluna do oxigênio dissolvido, que serve como referência para o aprendizado supervisionado da rede (saída), normalizou-se os dados, para serem utilizados no treinamento, em dois conjuntos:

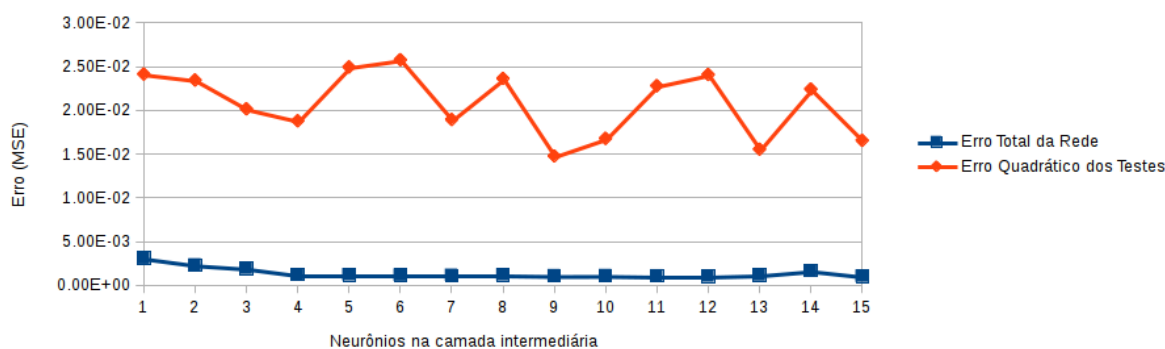
- Ponto de amostragem 11 - 155 amostras de 8 *scores* para 8 CP:
  - Dados para treinamento: 109 amostras (70%);
  - Dados para verificação: 46 amostras (30%).

#### 4.4.3.3 Resultados

A Figura 43 apresenta o erro total da rede e o erro do treinamento em função do número de neurônios da camada intermediária. Os dados são obtidos no treinamento das RNA com seis entradas e uma camada intermediária, variando o número de neurônios desta camada entre 1 e 15, conforme apresenta a Tabela 21. Nota-se que apesar da rede convergir para o treinamento em algumas arquiteturas, o erro quadrático médio dos testes é sempre muito alto. O menor erro quadrático médio é encontrado na arquitetura (8-9-1).

Tabela 21 – Dados médios obtidos com o treinamento da rede (8- $x$ -1) para o Rio Paraíba do Sul - PCA.

Neurônios na camada intermediária ( $x$ )	Erro Máx.	Limite máx. iteração	Erro Total da Rede (70%)	Iteração	Erro Quadrático dos Testes (30%)
1	0,0001	10000	3,11E-03	10000	2,41E-02
2	0,0001	10000	2,33E-03	10000	2,35E-02
3	0,0001	10000	1,94E-03	10000	2,02E-02
4	0,0001	10000	1,23E-03	10000	1,88E-02
5	0,0001	10000	1,21E-03	10000	2,49E-02
6	0,0001	10000	1,16E-03	10000	2,58E-02
7	0,0001	10000	1,15E-03	10000	1,89E-02
8	0,0001	10000	1,20E-03	10000	2,36E-02
9	0,0001	10000	1,10E-03	10000	1,48E-02
10	0,0001	10000	1,12E-03	10000	1,68E-02
11	0,0001	10000	1,08E-03	10000	2,28E-02
12	0,0001	10000	1,05E-03	10000	2,41E-02
13	0,0001	10000	1,19E-03	10000	1,56E-02
14	0,0001	10000	1,67E-03	10000	2,24E-02
15	0,0001	10000	1,04E-03	10000	1,66E-02

Figura 43 – Erro Total da Rede e Erro Médio Quadrático da rede (8- $x$ -1) em função do número de neurônios na camada intermediária para o Rio Paraíba do Sul - PCA.

#### 4.4.3.4 Comparação do modelo com dados reais

A partir do treinamento da rede de arquitetura (8-9-1), é possível normalizar inversamente os dados de saída da rede com a finalidade de verificar a distância destes dados com o alvo do treinamento, o valor real de Oxigênio Dissolvido. A Figura 44 apresenta estes valores iniciais da variável O.D., antes de serem padronizados e utilizados no conjunto de teste, contra o O.D. modelado pela rede, depois da efetuada a normalização reversa. A Figura 45 ilustra a dispersão dos mesmos dados. A maior amplitude e menor linearidade dos dados reais advindos das amostras do rio Paraíba do Sul, terminam por produzir maiores diferenças entre os valores originais e modelados pela RNA. No

entanto, a técnica de PCA produziu resultados melhores para este ponto de amostragem se comparados com o Experimento 2, o que pode ser confirmado pelo valor de coeficiente de correlação calculado em 0,761.

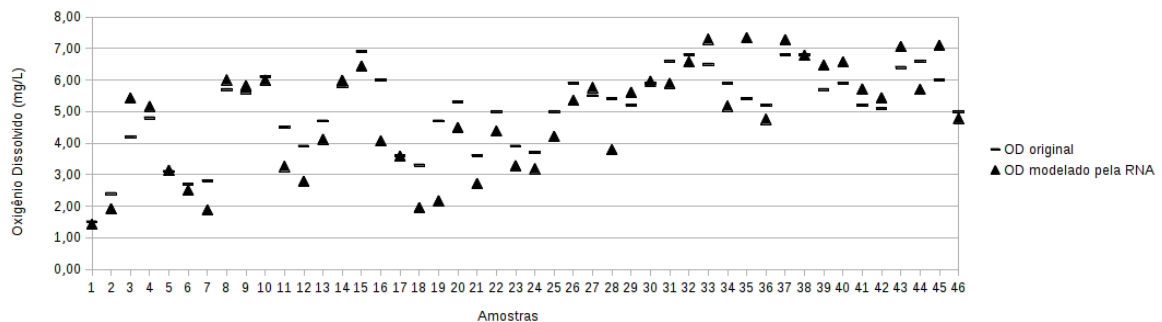


Figura 44 – Dado original contra dado modelado pela RNA para o Rio Paraíba do Sul - PCA.

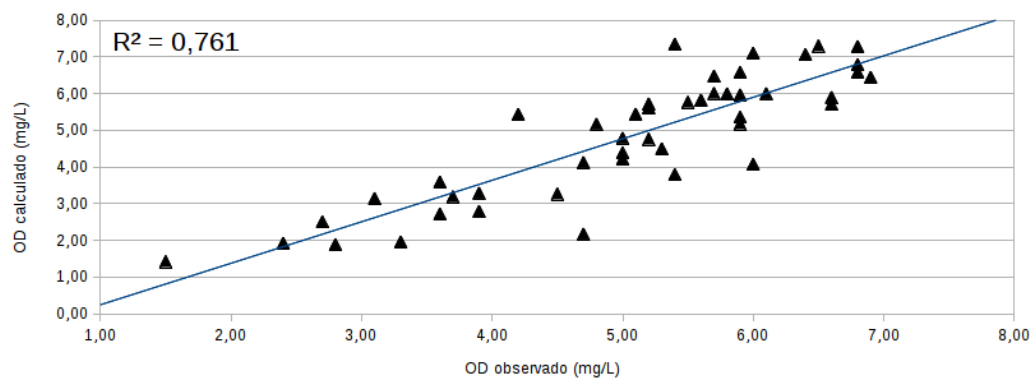


Figura 45 – Dispersão do dado original contra dado modelado pela RNA para o Rio Paraíba do Sul - PCA.

## 4.5 Comparação entre os Experimentos 1, 2 e 3

A Seção 4.2 descreve o treinamento de diversas redes neurais utilizando-se de dados teóricos do Rio Meia Ponte (GO). A Seção 4.3 descreve este mesmo treinamento a partir de dados reais advindos de amostras de água do Rio Paraíba do Sul (SP). Por fim, a Seção 4.4 descreve o treinamento a partir de dados oriundos da técnica de Análise de Componentes Principais. Assim, os dados colhidos de amostras de água de ambos os rios são também submetidos à técnica de PCA, permitindo uma comparação entre os resultados da modelagem de Oxigênio Dissolvido a partir do treinamento das redes utilizando-se de ambos os tipos de dados.

A Figura 46 apresenta a plotagem dos pares ordenados que representam os valores de Oxigênio Dissolvido do Ponto de amostragem 3 do Rio Meia Ponte (GO), modelados a partir de dados teóricos e dados advindos de PCA. Como verifica-se, o experimento com dados teóricos produziu um resultado bem pequeno de erro quando compara-se o OD original e OD modelado, utilizando-se de 10 entradas na rede.

A plotagem dos dados advindos da técnica de PCA demonstra que a rede não teve dificuldades para correlacionar as entradas (6 Componentes Principais) com o valor esperado do OD, gerando resultados que podem ser comparados em termos de qualidade com o experimento conduzido através de dados teóricos.

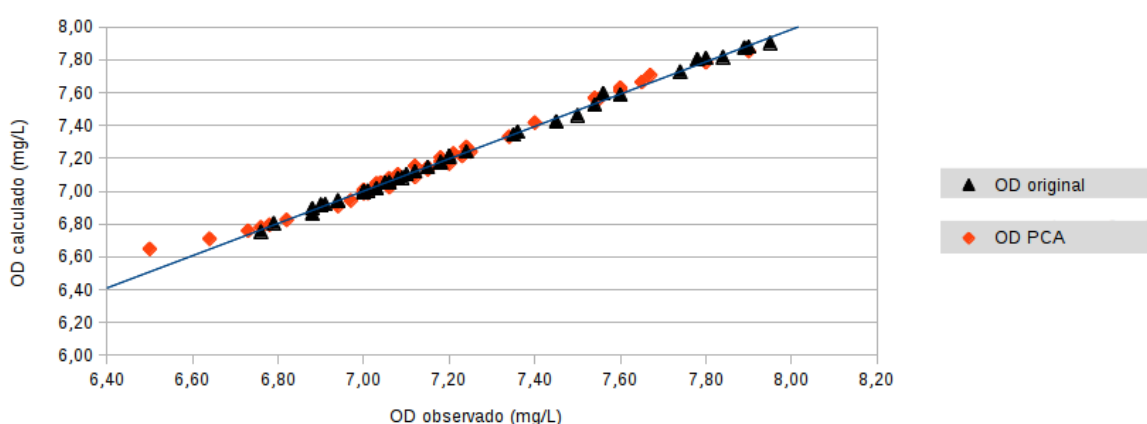


Figura 46 – Comparação entre Modelo com dados teóricos e PCA - Ponto 3.

A Figura 47 apresenta a plotagem dos pares ordenados que representam os valores de OD do Ponto de amostragem 11 do mesmo rio goiano, modelados utilizando-se das duas técnicas. Verifica-se que no caso deste ponto coletor de amostras, a modelagem a partir de ambas as técnicas mostra-se bastante eficiente, com valores próximos entre o dado modelado e o real.

No caso da técnica de PCA, pode ser verificado pequena dispersão dos pontos, que a redução do número de entradas na rede neural para 5 (o número de Componentes Principais que descrevem 92,14% de variância) não chega a reduzir a acuidade do treinamento, permitindo convergência no experimento com diferença pequena dos valores originais e obtidos através desta técnica.

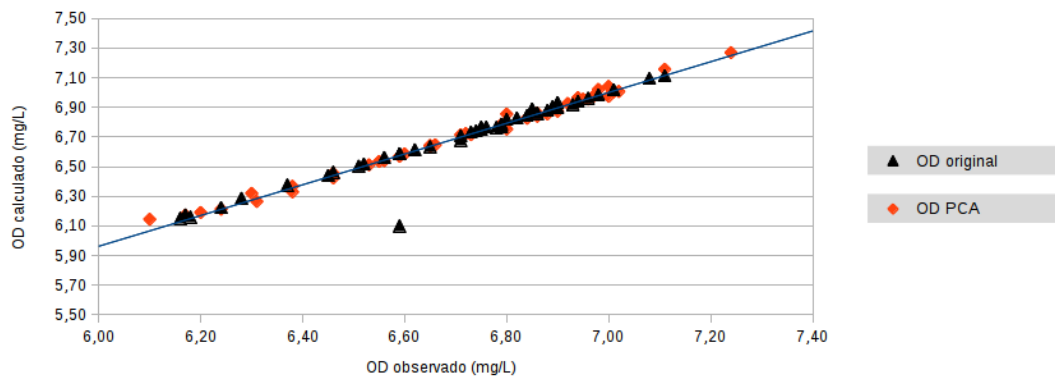


Figura 47 – Comparação entre Modelo com dados teóricos e PCA - Ponto 11.

A Figura 48 apresenta a plotagem dos pares ordenados representando os valores de OD para o Rio Paraíba do Sul (SP), modelados a partir de dados reais e advindos da técnica de PCA. Neste caso específico, o maior distanciamento dos pontos que representam os dados originais da linha de tendência demonstra que as maiores amplitude e grau de não-linearidade dos dados reais deste manancial aumentaram o erro de treinamento da RNA, o que provoca automaticamente o distanciamento entre o valor modelado e original.

A partir da análise dos dados advindos da técnica de PCA, verifica-se que a utilização de oito Componentes Principais (que representam 90,02% de variância) como entrada da RNA proporcionou um resultado melhor se comparado com o encontrado com a utilização dos dados originais, com pontos mais próximos da linha de tendência.

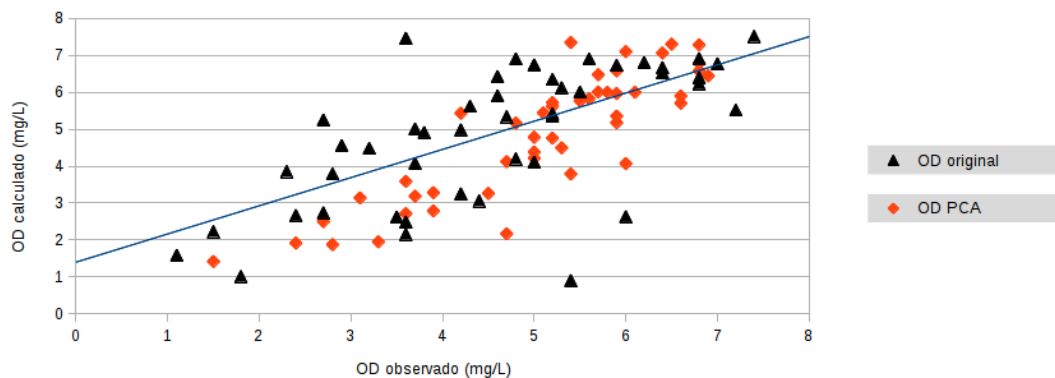


Figura 48 – Comparação entre Modelo com dados reais e PCA - Rio Paraíba do Sul.

As comparações demonstram que a utilização de dados advindos da técnica de PCA para alimentar o treinamento de redes neurais pode ser utilizada como técnica de sensibilidade de dados objetivando reduzir o número de entradas da rede, poupando esforço computacional e tempo de treinamento, principalmente para grandes bases de dados.

---

## Conclusão

Neste trabalho pode-se concluir de maneira geral, que a pesquisa com Redes Neurais Artificiais oferece vantagens, como a capacidade de aproximar funções não lineares, podendo, através de treinamento, modelar o sistema através de relações de entrada e saída, conhecendo apenas os valores de entrada e saída. No entanto, a maior vantagem do uso de modelos baseados em redes neurais é a sua simplicidade, o pouco tempo e esforço computacional requeridos após a fase de treinamento.

A contribuição deste trabalho inclui a discussão da técnica de Redes Neurais Artificiais, buscando ajudar na avaliação da qualidade da água potável para uso da população, através do modelo de Oxigênio Dissolvido, que visa prover as autoridades de gestão de recursos hídricos com ferramental científico para tomada de decisão, uma vez que este modelo é simples e de fácil implementação por parte de autoridades e pesquisadores. Através dos vários experimentos realizados, verifica-se o treinamento de RNA com menor número de parâmetros de entrada e de pequena quantidade de dados dos parâmetros, além da comparação entre os métodos utilizados visando determinar qual conjunto de dados de treinamento e arquitetura de rede produziu um valor modelado de OD mais próximo do valor observado na natureza.

Verificou-se, através do Experimento 1, Seção 4.2 (p. 60), realizado a partir de amostras do Rio Meia Ponte (GO), que mesmo com uma quantidade pequena de dados ecológicos é possível treinar uma RNA para produzir um modelo com valor de erro baixo. Foi possível determinar a melhor arquitetura de rede para completar a tarefa. O experimento gera números baixos de iteração e de erros nos treinamentos, o que proporciona baixo erro do modelo. A acuidade é próxima a de alguns oxímetros, ou medidores portáteis de oxigênio dissolvido, encontrados no mercado.

Por outro lado, o Experimento 2, Seção 4.3 (p. 80), conduzido a partir de dados reais do Rio Paraíba do Sul (SP), fornecidos pelo sítio na *internet* da CETESB, mostrou um nível de erro entre OD calculado e observado muito acima do esperado. Isto se dá em função da RNA não ter alcançado o erro de treinamento baixo o suficiente, o que dá a ideia do alto grau de não-linearidade das amostras e conseqüente dificuldade

para a convergência. Como observado na Seção 3.1.1.2 (p. 52), os 22 anos de dados disponíveis para este manancial resultaram em número de amostras insuficientes para representar o sistema hídrico em todos os meses do ano, caracterizando mudanças de estação, clima, etc. Desta forma, os dados coletados para este experimento contêm média de 51 dias para cada amostra, aproximadamente. Outro importante fator que contribui para o aumento do erro do treinamento é a amplitude maior nos dados do rio paulista, se comparado com a maioria das informações coletadas do rio goiano, com exceção do Ponto de Coleta 11, localizado dentro da capital do Estado. Esta maior amplitude é visível em parâmetros como o OD, que tem variação maior que todos os pontos de amostragem do rio goiano, com valores entre 0,9 e 8,0 mg/L. A Condutividade, assim como quase todos os outros parâmetros, exibe limite inferior mais baixo que os encontrados no rio de Goiás. Apesar do erro relativamente baixo nesta fase do experimento, todo este cenário trouxe dificuldade no treinamento, contribuindo para erros altos na fase de testes de avaliação de dados não vistos pela rede. Em sistemas apoiados por aprendizagem supervisionada, este erro, que representa a diferença entre a resposta desejada e a resposta observada, como apresentado na Equação 6 (p. 45), é medido pela função chamada Função Objetivo, que tem o propósito de corrigir este erro observado para todos os neurônios. Problemas como este não podem ser caracterizados como fruto de *overfitting*<sup>1</sup>. Este problema é descartado, pois é necessária quantidade maior de neurônios na camada intermediária para que ocorra, condição não presente na arquitetura escolhida para o experimento do Rio Paraíba do Sul, ou em qualquer outro experimento. Conduziram-se experimentos com outros dois tipos de algoritmo de treinamento de redes neurais com os dados deste ponto de amostragem para validar este resultado. Realizou-se experimento com o algoritmo de Retropropagação Resiliente (Rprop) e com o algoritmo *Levenberg-Marquardt*, havendo todos demonstrado resultados semelhantes de valores de erros de treinamento.

Já o Experimento 3, Seção 4.4 (p. 84), conduzido a partir da utilização da técnica de PCA aplicada nos dados de dois pontos de amostragem do Rio Meia Ponte e um ponto de amostragem do Rio Paraíba do Sul, mostrou-se bastante eficiente ao reduzir a quantidade de entradas para a RNA, atingindo o objetivo proposto, que era o de permitir a redução do número de entradas em uma rede neural, a partir da substituição dos dados originais por um conjunto menor descrevendo maior variância. Para o ponto de amostragem 3 do Rio Meia Ponte, a rede foi treinada com 6 entradas neste experimento, contra 10 do primeiro. No caso do ponto de amostragem 11 do mesmo manancial, a RNA foi treinada com 5 entradas, sendo esta a metade das entradas utilizadas no primeiro experimento, obtendo praticamente o mesmo resultado. O grande expoente, no entanto, foi o resultado para o ponto de amostragem do Rio Paraíba do Sul, onde o treinamento da RNA a partir de 8 entradas obteve melhor resultado que aqueles resultantes do treinamento da RNA com

<sup>1</sup> Nesta condição, a rede neural memoriza os dados apresentados na fase de treinamento; com isto, ela torna-se incapaz de generalizar e reconhecer padrões não vistos durante o treinamento.

10 entradas advindas de dados reais (Experimento 2), o que pode ser comprovado pelo cálculo do coeficiente de correlação, que resultou em valor de 0,761 para o Experimento 2, contra 0,428 quando a RNA foi treinada com dados de PCA.

Os resultados mostrados através dos experimentos fornecem ferramentas para a tomada de decisão por parte de autoridades, objetivando a manutenção da qualidade e a garantia dos múltiplos usos dos recursos hídricos, além de poderem servir de referência para trabalhos futuros de modelagem utilizando redes neurais com algoritmo de retropropagação. Apesar disso, esta estratégia de modelagem oferece limitações, como a dependência da qualidade da água a montante do ponto onde se pretende modelar o valor de OD.

A ferramenta *Neuroph Studio* foi apresentada e mostra-se importante recurso computacional para conduzir experimentos com redes neurais artificiais, não apenas pela eficiência apresentada, mas vale destacar que sua natureza aberta permite sua utilização por qualquer interessado neste assunto, em diversas plataformas computacionais, sem o pagamento de licenças.

Ressalta-se que a precisão dos resultados do modelo é dependente da quantidade e qualidade dos dados de entrada. Por este motivo, recomenda-se o armazenamento sistemático de dados, por parte das autoridades do nosso país, que representem os sistemas ambientais como todo, principalmente daqueles que se situam em áreas de grande densidade populacional. Quanto maior o tempo disponível de dados coletados, maiores as chances de desenvolver modelos analíticos que descrevam a realidade e sirvam como ferramenta para conhecer o futuro. A pobreza destes dados coloca em risco o estudo destes sistemas, em busca de informações que possam ajudar a mitigar os impactos provocados pelo desenvolvimento e ocupação não planejados, podendo contribuir para piorar a qualidade de vida de milhões de pessoas.

## 5.1 Objetivos alcançados

### □ Objetivo Principal

- Neste trabalho foi utilizada ferramenta baseada em Redes Neurais Artificiais para a modelagem do parâmetro químico Oxigênio Dissolvido, considerado importante para a avaliação da qualidade de água pela legislação brasileira. Assim, considera-se o objetivo principal alcançado.

### □ Objetivos Secundários

- Avaliar a legislação brasileira que caracteriza os recursos hídricos e os parâmetros de qualidade da água e a importância deste parâmetro para planejar o sistema a ser modelado.



- A partir das RNA treinadas, avaliou-se a quantidade de neurônios na camada intermediária que define a melhor arquitetura de rede para cada experimento;
- Em função da pequena quantidade de dados para os sistemas hídricos que fizeram parte deste estudo, avaliaram-se também técnicas que possibilitem o treinamento de redes neurais com quantidades menores de dados. Assim, conduziram-se três experimentos:
  - \* treinamento com dados teóricos, pré-processados para permitir a expansão do banco de dados através da técnica de interpolação linear;
  - \* utilização de dados reais para treinamento e avaliação dos parâmetros que interferem na utilização de dados com alto grau de não-linearidade;
  - \* treinamento com dados advindos da técnica estatística chamada PCA com a finalidade de reduzir o número de entradas na rede.

## 5.2 Trabalhos futuros

Entre as perspectivas de trabalhos futuros, está a possibilidade de avaliar outras abordagens de redes neurais, com os mesmos dados utilizados neste trabalho. Podem ser comparados o número de iterações para convergência, o erro de treinamento e o erro dos testes da rede.

A avaliação de outros algoritmos de treinamento semelhantes ao de retropropagação de erro também pode ser interessante com vistas a melhorar a capacidade de generalização com dados não-lineares. Alguns trabalhos sugerem a utilização de técnicas conhecidas como construção e poda, ou mesmo de algoritmos genéticos, para determinação automática da melhor arquitetura da rede neural, reduzindo etapas ao eliminar a necessidade de testes com variadas arquiteturas.

A utilização de outras técnicas de sensibilidade de dados pode ajudar a definir uma melhor relação entre os componentes de entrada e a saída da rede, permitindo a eliminação de entradas que causam aumento de custo computacional à tarefa de treinar uma rede neural, ou mesmo a escolha de outros parâmetros que descrevam melhor a dinâmica envolvida na modelagem do OD.

Avaliar a existência de relação entre os dez parâmetros de entrada e o oxigênio dissolvido é também uma continuação importante para este trabalho. Tal estudo permite estabelecer as relações entre os parâmetros, conhecendo melhor o papel destes componentes na dinâmica do modelo e, através da rede neural treinada, extrapolar um (ou vários) destes parâmetros, simulando uma disfunção no ambiente com vistas a visualizar, neste mesmo ambiente, o resultado do Oxigênio Dissolvido.

---

## Referências

- ANDERSEN, A. C. Autonomous Neural Development and Pruning. 2010. Disponível em <<http://blog.andersen.im/2010/12/autonomous-neural-development-and-pruning/>> Acesso em 01/14.
- ARAÚJO, R. C. d. S.; OLIVEIRA, F. H. P. C. de. Análise de parâmetros limnológicos da água dos açudes do prata e do meio (Pernambuco-Brasil). **XIX Simpósio Brasileiro de Recursos Hídricos**, 2011.
- AREERACHAKUL, S.; JUNSAWANG, P.; POMSATHIT, A. Prediction of Dissolved Oxygen Using Artificial Neural Network. **2011 International Conference on Computer Communication and Management**, v. 5, 2011.
- BAGGIO, H.; HORN, A. H. Avaliação geoquímica ambiental da água superficial do Rio do Formoso, Buritizeiro - MG. **Geografias artigos científicos**, v. 07, n. 2, p. 56–67, julho-dezembro 2011.
- BAIO, J. A. F. **Avaliação da contaminação nos principais corpos d'água do município de São Carlos/SP**. 111 f. p. Dissertação (Mestrado) — Universidade de São Paulo, 2009. Dissertação (Mestrado em Química Analítica).
- BAPTISTA, D.; MORGADO-DIAS, F. A survey of artificial neural network training tools. **Neural Computing and Applications**, Springer, p. 609–615, September 2013. ISSN 0941-0643.
- BAUDIN, M. Introduction to Scilab. Scilab Consortium - Digiteo, january 2010.
- BORSATO, D.; MOREIRA, I.; NOBREGA, M. M.; MOREIRA, M. B.; DIAS, G. H.; DA SILVA, R. S. D. S. F.; BONA, E. Aplicação de redes neurais artificiais na identificação de gasolinas adulteradas comercializadas na região de Londrina - Paraná. **Quim. Nova**, v. 32, n. 9, p. 2328–2332, 2009.
- BOTKIN, D.; KELLER, E. **Environmental Science: Earth as a Living Planet**. 8. ed. [S.l.]: Wiley, 2011. ISBN 9780470118559.
- BRAGA, A. P.; CARVALHO, A. C. P. L. F.; LUDEMIR, T. B. **Redes neurais artificiais: teoria e aplicações**. [S.l.]: LTC, 2000.
- BRASIL, A. N. D. Á. A. **Panorama da Qualidade das Águas Superficiais no Brasil**. [S.l.], 2005. Disponível em: <http://www.ana.gov.br>. Acesso em 06/2012.

- \_\_\_\_\_. **Indicadores de Qualidade - Introdução**. [S.l.], 2013. Disponível em: <<http://pnqa.ana.gov.br/IndicadoresQA/introdu%C3%A7%C3%A3o.aspx>>. Acesso em 11/2013.
- BRASIL, M. D. S. **Água Brasil - Legislação Federal**. [S.l.]: Fundação Oswaldo Cruz, 2010. Disponível em: <<http://www.aguabrasil.iciet.fiocruz.br/index.php?pag=leg>>. Acesso em 09/2013.
- \_\_\_\_\_. **Portaria 2914**. [S.l.], 2011. Disponível em: [http://www.comitepcj.sp.gov.br/download/Portaria\\_MS\\_2914-11.pdf](http://www.comitepcj.sp.gov.br/download/Portaria_MS_2914-11.pdf). Acesso em 11/2013.
- BRIGANTE, J.; ESPÍNDOLA, E. Limnologia fluvial: um estudo no rio Mogi-Guaçu. p. 278, 2003.
- CAMPBELL, S. L.; CHANCELIER, J.-P.; NIKOUKHAH, R. **Modeling and Simulation in Scilab/Scicos**. [S.l.]: Springer, 2006.
- CARVALHO, G. L.; SIQUEIRA, E. Q. Qualidade da Água do Rio Meia Ponte no Perímetro Urbano do Município de Goiânia Goiás. **REEC. Revista Eletrônica de Engenharia Civil**, v. 1, n. 2, p. 19–33, 2011.
- CERETTA, M. C. **Avaliação dos aspectos da qualidade da água na sub-bacia hidrográfica do Arroio Cadena - município de Santa Maria - RS**. Dissertação (Mestrado), 2004.
- CETESB, C. D. T. D. S. A. **Relatório de Qualidade das Águas Interiores do Estado de São Paulo**. [S.l.], 2004. Disponível em <<http://www.cetesb.sp.gov.br>> Acesso em 06/2012.
- CONAMA, C. N. D. M. A. **Resolução CONAMA 357**. [S.l.], 2005. Disponível em <<http://www.mma.gov.br/port/conama/res/res05/res35705.pdf>> Acesso em 06/2012.
- \_\_\_\_\_. **Resolução CONAMA 430**. [S.l.], 2011. Disponível em: <<http://www.mma.gov.br/port/conama/legiabre.cfm?codlegi=646>> Acesso em 06/2012.
- COSTA, H. R. d. N. **Aplicação de técnicas de inteligência artificial em processos de fabricação de vidro**. . Tese (Doutorado) — Escola Politécnica da Universidade de São Paulo, 2006.
- CUNHA, A. C. d. Revisão descritiva sobre qualidade da água, parâmetros e modelagem de ecossistemas aquáticos tropicais. **Biota Amazônia**, v. 3, n. 1, p. 124–143, 2013. ISSN 2179-5746.
- CUNHA, G. B. da; LUITGARDS-MOURA, J. F.; NAVES, E. L. M.; ANDRADE, A. O.; PEREIRA, A. A.; TEREZINHAMILAGRE, S. A utilização de uma rede neural artificial para previsão da incidência da malária no Município de Cantá, Estado de Roraima. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 43, n. 5, p. 567–570, Sep-oct 2010.
- DEMETRACOPOULOS, A. C.; STEFAN, H. G. Model of the Mississippi River Pool: Mass Transport. **Journal of Environmental Engineering**, v.109, n. 5, p. 1006–19, 1983.

DO CERRADO, A. A. P.-Á. Situação Ambiental das Águas da Bacia do Rio Meia Ponte, Goiás. 2008. Disponível em: [www.proaguasdocerrado.org.br/bacia.php](http://www.proaguasdocerrado.org.br/bacia.php) Acesso em 11/2012.

\_\_\_\_\_. Situação Ambiental das Águas da Bacia do Rio Meia Ponte, Goiás. 2008. Disponível em: [www.proaguasdocerrado.org.br/bacia4.php](http://www.proaguasdocerrado.org.br/bacia4.php) Acesso em 11/2012.

DOGAN, E.; SENGORUR, B.; KOKLU, R. Modeling biological oxygen demand of the Melen River in Turkey using an artificial neural network technique. **Journal of Environmental Management**, Elsevier, v. 90, p. 1229–1235, 2009.

DUTRA, W. C. P. **MODELAGEM DOS PARÂMETROS DE QUALIDADE DE ÁGUA EM TRECHO URBANIZADO DO RIO PARAIBUNA EM JUIZ DE FORA (MG)**. 2014.

ELSHORBAGY, A.; SIMONOVIC, S. P.; PANU, U. S. **Journal of Hydrologic Engineering**, v. 5, n. 4, p. 424–427, October 2000.

ELVIRA, M. G. Analysis of Intrusion Detection System Using Multilayer Perceptron Neural Networks. **Reporte Técnico RT-0002-2013**, Dezembro 2013. In *Sistemas Inteligentes: Reportes Finales Ago-Dic 2013*.

ERASMUS, J. B.; ERASMUS, J. S.; CSALL, P. O. C.; CSALL, M. J. C. Backgammon AI Player. April 2010.

FANUCCHI, R. Z.; OLESKOVICZ, M.; BARBOSA, D. Uma Análise da Distorção Harmônica para a Detecção de Faltas de Alta Impedância via Redes Neurais Artificiais. **Simpósio Brasileiro de Sistemas Elétricos**, Anais, April 2014. ISSN 2177-6164. Disponível em <<http://www.sbse.org.br/anais/PDF/SBSE2014-0219.pdf>> Acesso em 04/2014.

FAUSETT, L. **Fundamentals of neural networks: architectures, algorithms, and applications**. [S.l.]: Prentice-Hall, Inc., 1994.

FERREIRA, L. R.; BORTONI, M. A.; CAMARGO, L.; SEBASTIANI, M.; OLIVEIRA, F. A. de; LOURENÇO, F. L. Heterogeneidade espacial de fitoplâncton entre ambientes lóticos a montante e a jusante do reservatório do Monjolinho, São Carlos, SP. Rodovia Washington Luís, km 235 - SP-310, 2009. VI Congresso de Meio Ambiente da AUGM.

FIORUCCI, A. R.; FILHO, E. B. A importância do Oxigênio Dissolvido em Ecossistemas Aquáticos. **QUÍMICA NOVA NA ESCOLA**, n. 22, p. 10–16, November 2005.

FONSECA, H. S.; SALVADOR, N. N. B. Estudo integrado da qualidade das águas da bacia hidrográfica do rio Bonito em Descalvado - SP. **Revista Inst. Adolfo Lutz**, v. 64, n. 2, p. 179–185, 2005. Disponível em <<http://pesquisa.bvs.br/brasil/resource/pt/ses-6070>>. Acesso em 06/14.

GASPAROTTO, F. A. **Avaliação ecotoxicológica e microbiológica da água de nascentes urbanas no município de Piracicaba-SP**. 89 p. Dissertação (Mestrado) — Centro de Energia Nuclear na Agricultura, 2011.

GOMES, L. F. C.; MONTENEGRO, S. M. G. L.; VALENCA, M. J. D. S. Modelo Baseado na Técnica de Redes Neurais para Previsão de Vazões na Bacia do Rio São Francisco. **RBRH: Revista Brasileira de Recursos Hídricos**, Associação Brasileira de Recursos Hídricos, fevereiro 2007.

HAYKIN, S. **Redes Neurais: princípios e prática**. [S.l.]: Bookman, 2001.

HEATHCOTE, I. W. **A Review and Comparison of Dissolved Oxygen Models Appropriate for the Grand River Watershed**. [S.l.], 2009.

HENRIQUE, M. A. **A Industrialização do Município de São José dos Campos-SP: Uma Abordagem a Partir da História Econômica Local**. 2012. 57 p. Monografia de Especialização em Gestão Pública Municipal. Disponível em <<http://repositorio.roca.utfpr.edu.br/jspui/handle/1/1795>>. Acesso em 03/2014.

HÍDRICOS, S. P. E. C. d. R.; PALÁCIOS, A. C.; RODRIGUES, F. B. **Guia do sistema paulista de recursos hídricos**. 2. ed. [S.l.]: Coordenadoria de Recursos Hídricos, Secretaria do Meio Ambiente, Governo do Estado de São Paulo, 2008. 96 p. Disponível em <[http://www.sigrh.sp.gov.br/sigrh/index/publicacoes\\_files/guia/guia.pdf](http://www.sigrh.sp.gov.br/sigrh/index/publicacoes_files/guia/guia.pdf)> Acesso em 01/14. ISBN 9788586624575.

IBGE. **Censo Demográfico Brasil 2010**. [S.l.], 2010. Disponível em <<http://www.ibge.gov.br/estadosat/perfil.php?sigla=go>> Acesso em 06/12.

JUUTI, P. S.; KATKO, T. S.; S. Vuorinen, H. **Environmental History of Water**. [S.l.]: IWA Publishing, 2007. 640 p. ISBN 9781843391104.

KISI, O.; AY, M. Modeling Dissolved Oxygen (DO) Concentration Using Different Neural Network Techniques. **International Balkans Conference on Challenges of Civil Engineering**, p. 19–21, Maio 2011.

KNAPIK, H. G. **REFLEXÕES SOBRE MONITORAMENTO, MODELAGEM E CALIBRAÇÃO NA GESTÃO DE RECURSOS HÍDRICOS: ESTUDO DE CASO DA QUALIDADE DA ÁGUA DA BACIA DO ALTO IGUAÇU**. 197 p. Dissertação (Mestrado), 2009.

LAWRENCE, S.; GILES, C. L.; TSOI, A. C. What Size Neural Network Gives Optimal Generalization? Convergence Properties of Backpropagation. University of Maryland, 1996.

LI, G. **Stream Temperature and Dissolved Oxygen Modeling in the Lower Flint River Basin, GA**. 240 p. Tese (Doutorado), 2006.

MACCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **Bulletin of Mathematical Biophysics**, v. 5, p. 115–133, 1943.

MELLO, L. T. A. d.; TRINDADE, M. C.; MARCUZZO, F. F. N. Caracterização da qualidade das águas do principal rio da região metropolitana de Goiânia. **XIX Simpósio Brasileiro de Recursos Hídricos**, 2011.

MIZUTORI, I. S. **Caracterização da qualidade das águas fluviais em meios peri-urbanos: o caso da bacia hidrográfica do Rio Morto – RJ**. . Dissertação (Mestrado) — Centro de Tecnologia e Ciências, Março 2009.

- MONTEIRO, A. J. Modelação do Oxigénio Dissolvido em rios. 1997. Disponível em <<https://fenix.tecnico.ulisboa.pt/downloadFile/3779571269514/Oxig?nio%20Dissolvido.pdf>>. Acesso em 01/2014.
- MWEGOHA, W. J. S.; KASEVA, M. E.; SABAI, S. M. M. Mathematical modeling of dissolved oxygen in fish ponds. **African Journal of Environmental Science and Technology**, v. 4, n. 9, p. 625–638, September 2010.
- NAZARENO, G. S.; SILVA, A. T. R. da; SCHNEIDER, A. M. NeuralTrainer – Software Genérico para o Treinamento de Redes Neurais Multilayer Perceptron com Aprendizado Supervisionado por Retropropagação de Erros, Um Caso de Uso: OCR (Optical Character Recognition). **Anais do V Encontro de Estudantes de Informática do Tocantins.**, p. pp. 243–250, 2003.
- NUNES, D. G.; SILVA, A. d. J.; SILVA, D. D. da; Mota Ramos, M.; MATOS, A. T. de. DEPURA: SISTEMA WEB DE MODELAGEM DE AUTODEPURAÇÃO DO RIO TURVO SUJO. **XVIII Simpósio Brasileiro de Recursos Hídricos**, November 2009.
- OTTO, M. **Chemometrics - Statistics and Computer Application in Analytical Chemistry**. Second edition. [S.l.]: Wiley-VCH, 2007. ISBN 3-527-31418-0.
- PALANI, S.; LIONG, S.-Y.; TKALICH, P.; PALANICHAMY, J. Development of a neural network model for dissolved oxygen in seawater. **Indian Journal of Marine Sciences**, v. 38, n. 2, p. 151–159, June 2009.
- RIBEIRO, L. N. **Rede Neural com Retropropagação: uma Aplicação na Classificação de Imagens de Satélite**. Dissertação (Mestrado), 2003.
- RODRIGUES, L. C.; SILVA, P. P. C. D.; LINDEN, R. Séries Temporais no Consumo de Energia Elétrica no Estado do Rio de Janeiro. 2007.
- ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological Review**, v. 65, n. 6, p. 386–408, 1958.
- SANTOS, C. C. D.; FILHO, A. J. P. Modelagem Hidrológica Urbana por meio de Redes Neurais Artificiais: uma Aplicação para a Bacia do rio Tamanduateí em São Paulo, SP. **Revista Brasileira de Meteorologia**, v. 18, n. 2, p. 149–159, 2003.
- SCARIOT, M. R. **Modelagem e Simulação Sistêmica de Rios: Avaliação dos Impactos Ambientais no Rio Mogi-Guaçu/SP**. Tese (Doutorado) — FACULDADE DE ENGENHARIA DE ALIMENTOS, 2008.
- SCHIMIDT, F.; BUENO, M. I.; POPPI, R. J. et al. Aplicação de modelos quimiométricos à Espectroscopia de Fluorescência de Raios- X de energia dispersiva. **Quím. Nova**, v. 25, n. 6, p. 946–956, 2002.
- SCHIMIDT, F.; GODINHO, M. D. S.; PEREIRA, R. O.; RIBEIRO, K. D. O.; DE OLIVEIRA, A. E.; DE OLIVEIRA, S. B. Classificação de Refrigerantes Através de Análise de Imagens e Análise de Componentes Principais (PCA). **Quim. Nova**, v. 31, n. 6, p. 1485–1489, 2008.
- SEPLAN. **Goiás, Visão Geral**. [S.l.], 2012. Disponível em: [http://www.seplan.go.gov.br/sepin/goias.asp?id\\_cad=6000](http://www.seplan.go.gov.br/sepin/goias.asp?id_cad=6000). Acesso em 11/2012.

- SEVARAC, Z. Neuro Fuzzy Reasoner for Student Modeling . Department of Information Systems, p. 740–744, July 2006.
- SILVA, I. N. da; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais para engenharia e ciências aplicadas**. [S.l.]: ARTLIBER, 2010. ISBN 8588098539.
- SILVA, M. M. **Uma Abordagem Evolucionária Para o Aprendizado Semi-Supervisionado em Máquinas de Vetores de Suporte**. Dissertação (Mestrado) — Escola de Engenharia, November 2008.
- SILVERT, W.; BAPTIST, M. Can Neural Networks be used in Data-Poor Situations? **Int. Workshop on Applications of Artificial Neural Networks to Ecological Modelling**, 1998.
- SIQUEIRA, E. Q. d. **Aplicação do modelo de qualidade de água (QUAL2E) na modelação de oxigênio dissolvido no rio Meia Ponte (GO)**. Dissertação (Mestrado) — Escola de Engenharia de São Carlos-EESC/Departamento de Hidráulica e Saneamento, 1996. Dissertação de Mestrado.
- SOARES, P. L. B.; SILVA, J. P. da. Aplicação de Redes Neurais Artificiais em Conjunto com o Método Vetorial da Propagação de Feixes na Análise de um Acoplador Direcional Baseado em Fibra Ótica . **Revista Brasileira de Computação Aplicada**, v. 3, n. 2, p. 58–72, set. 2011. ISSN 2176-6649.
- SOUSA, W. D. S.; SOUSA, F. D. A. S. D. Rede neural artificial aplicada à previsão de vazão da Bacia Hidrográfica do Rio Piancó. **Revista Brasileira de Engenharia Agrícola e Ambiental**, UAEA/UFCG, v. 14, n. 2, p. p.173–180, 2010.
- STATHAKIS, D. How many hidden layers and nodes? **International Journal of Remote Sensing**, Taylor & Francis, v. 30, n. 8, p. 2133–2147, April 2009. DOI: 10.1080/01431160802549278.
- UN, U. N. **Beyond scarcity: Power, poverty and the global water crisis**. [S.l.], 2006.
- UNESCO, W. W. A. P. **The United Nations World Water Development Report 3: Water in a Changing World**. [S.l.], 2009. Disponível em <<http://webworld.unesco.org/water/wwap/wwdr/wwdr3/tableofcontents.shtml>> Acesso em 01/14.
- UNICEF; WHO, W. H. O. **Progress on Sanitation and Drinking-water: 2010 Update**. [S.l.], 2010.
- URROZ, G. E. **DataFittingand Interpolation with SCILAB**. [S.l.], 2001.
- VALENTE, J. P. S.; PADILHA, P. M.; SILVA, A. M. M. Oxigênio dissolvido (OD), demanda bioquímica de oxigênio (DBO) e demanda química de oxigênio (DQO) como parâmetros de poluição no ribeirão Lavapés/Botucatu - SP. **Eclética Química**, scielo, v. 22, p. 49–66, 00 1997. ISSN 0100-4670.
- WARWICK, J. J. Optimizing Water Quality Modeling in an Uncertain Environment. **Water Science and Technology**, v.24, n. n. 6, p. p.41–48, 1991.

- WECKMAN, G. R.; PASCHOLD, H. W.; DOWLER, J. D.; WHITING, H. S.; A. Young, W. Using Neural Networks with Limited Data to Estimate Manufacturing Cost. **Journal of Industrial and Systems Engineering**, v. 3, n. 4, p. pp 257–274, 2010.
- WEIGANG, L.; SÁ, L. D. D. A.; GALVÃO, G. P.; BEVILAQUA, R. M. Prediction of the Level of Paraguay River using Neural Networks. **Pesq. agropec. bras.**, v. 33, p. p.1791–1797, 1998.
- WILLIAMS, N. T. **MODELING DISSOLVED OXYGEN IN LAKE POWELL USING CE-QUAL-W2**. Dissertação (Mestrado) — Department of Civil and Environmental Engineering, April 2007.
- WWAP, U. N. W. W. A. P. **The United Nations World Water Development Report 2014: Water and Energy**. [S.l.], 2014.
- ZURADA, J. M. **Introduction to Artificial Neural Systems**. [S.l.]: PWS Publishing Company, 1995.



# Apêndices

---

# Algoritmo de treinamento

## *Backpropagation*

Este apêndice lista o algoritmo de treinamento pela regra de retropropagação, para redes MLP.

---

**Algoritmo 1** Algoritmo de treinamento *Backpropagation*

---

- 1: Inicialize os pesos (valores pequenos e aleatórios)
- 2: **enquanto** condição de parada não é satisfeita **faça**
- 3:   **para** cada par de treinamento **faça**
- 4:     **para** cada unidade de entrada  $X_i, i = 1, \dots, n$  **faça**
- 5:       receba o sinal  $x_i$  e difunda esse sinal para todas as unidades da camada acima (camada escondida)
- 6:     **fim para**
- 7:     **para** cada unidade escondida  $Z_j, j = 1, \dots, p$  **faça**
- 8:       faça a soma ponderada dos sinais de entrada

$$z_{in_j} = v_{0j} + \sum_{i=1}^n x_i v_{ij}$$

depois aplique sua função de ativação para calcular o sinal de saída

$$z_j = f(z_{in_j})$$

e envie esse sinal para todas as unidades da camada acima (unidades de saída)

- 9:   **fim para**
- 10:   **para** cada unidade de saída  $Y_k, k = 1, \dots, m$  **faça**
- 11:     faça sua soma ponderada dos sinais de entrada

$$y_{in_k} = w_{0k} + \sum_{j=1}^p z_j w_{jk}$$

e aplique sua função de ativação para calcular o sinal de saída

$$y_k = f(y_{in_k})$$

- 12:   **fim para**
- 13:   **para** cada unidade de saída  $Y_k, k = 1, \dots, m$  **faça**
- 14:     receba o *target* correspondente ao padrão de treinamento e calcule o erro

$$\delta_k = (t_k - y_k) f'(y_{in_k})$$

calcule a sua correção do peso (usando o  $w_{jk}$  anterior),

$$\Delta w_{jk} = \alpha \delta_k z_j$$

e calcule o bias

$$\Delta w_{0k} = \alpha \delta_k$$

- 15:   **fim para**
- 16:   **para** cada unidade de escondida  $Z_j, j = 1, \dots, p$  **faça**
- 17:     encontre o  $\delta$  dos neurônios da camada escondida

$$\delta_{in_j} = \sum_{k=1}^m \delta_k w_{jk}$$

e multiplique a derivada de sua função de ativação para calcular o erro

$$\delta_j = \delta_{in_j} f'(z_{in_j})$$

e calcule o termo de correção de peso

$$\Delta v_{ij} = \alpha \delta_j x_i$$

A seguir, calcule a correção do bias

$$\Delta v_{0j} = \alpha \delta_j$$

- 18:   **fim para**
  - 19:   **para** cada unidade  $Y_k, k = 1, \dots, m$ , atualize os pesos e os bias ( $j = 0, \dots, p$ ) **faça**
  - 20:      $w_{jk}(\text{novo}) = w_{jk}(\text{antigo}) + \Delta w_{jk}$
  - 21:   **fim para**
  - 22:   **para** cada unidade  $Z_j, j = 1, \dots, p$ , atualize os pesos e os bias ( $i = 0, \dots, n$ ) **faça**
  - 23:      $v_{ij}(\text{novo}) = v_{ij}(\text{antigo}) + \Delta v_{ij}$
  - 24:   **fim para**
  - 25:   **fim para**
  - 26: **fim enquanto**
-

## Demonstração matemática do algoritmo *Backpropagation*

A regra de propagação de um neurônio  $j$  é dada por

$$net_j = \sum w_{ij} + b_j \quad (12)$$

onde  $w_{ij}$  é o peso sináptico de  $j$  para a entrada  $i$  e  $b_j$  é seu *bias*. O algoritmo de aprendizado baseia-se na minimização do erro quadrático por meio do método do gradiente decrescente para o ajuste de peso. Esse processo pode ser descrito pela equação:

$$\Delta w_{ij} = -\alpha \frac{\partial E}{\partial w_{ij}} \quad (13)$$

O erro quadrático médio para um padrão  $P$  é definido por

$$E^P = \frac{1}{2} \sum_j (t_j^P - y_j^P)^2 \quad (14)$$

onde  $t_j^P$  é o  $j$ -ésimo *target* do padrão  $P$  e  $y_j^P$  é a saída do  $j$ -ésimo neurônio para o mesmo padrão.

O cálculo de  $\Delta w_{ij}$  é dado por

$$\Delta w_{ij} = -\alpha \frac{\partial E^P}{\partial w_{ij}} = -\alpha \underbrace{\frac{\partial E^P}{\partial net_j^P}}_{\delta_j^P} \cdot \underbrace{\frac{\partial net_j^P}{\partial w_{ij}}}_{x_i} = \alpha \delta_j^P x_i \quad (15)$$

O valor  $\delta_j^P$  é o erro encontrado no neurônio  $j$  para o padrão  $P$ . Para a sua obtenção, deve-se calcular a derivada parcial de

$$\delta_j^P = -\frac{\partial E^P}{\partial net_j^P} \quad (16)$$

Para tanto, é necessário encontrar uma função que descreva o erro em função de  $net$ , ou seja,  $E^P = f(net_j^P)$ . Como  $f(net_j^P) = y_j^P$ , obtém-se

$$\delta_j^P = -\frac{\partial E^P}{\partial y_j^P} \cdot \frac{\partial y_j^P}{\partial net_j^P} \quad (17)$$

A expressão do erro  $\delta_j^P$ , utilizada na correção dos pesos, assume configurações diferentes, dependendo do local onde o neurônio em análise se encontra – na camada escondida ou na camada de saída. Assim, temos duas situações

## B.1 $j$ pertence à camada de saída

A Figura 49 mostra o esquema de um neurônio  $j$  que se encontra na camada de saída.

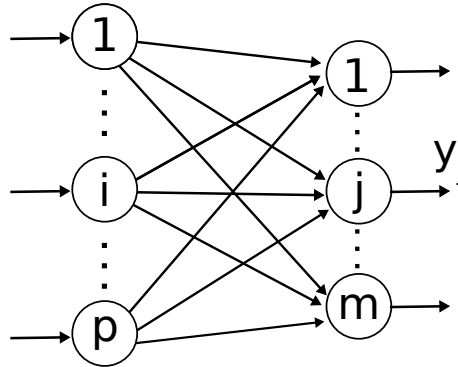


Figura 49 – Esquema de neurônio na camada de saída.

Para o cálculo de  $\delta_j^P$  nesse caso, considere a expressão do erro quadrático total do neurônio  $j$ , dada por:

$$E^P = \sum_j \frac{1}{2} (t_j^P - y_j^P)^2 \quad (18)$$

Sua derivada parcial em relação a  $y$  (primeira parte da equação procurada) corresponde a:

$$\frac{\partial E^P}{\partial y_j^P} = -(t_j^P - y_j^P) \quad (19)$$

Por outro lado, temos que:

$$\frac{\partial y_j^P}{\partial net_j^P} = \frac{\partial f(net_j^P)}{\partial net_j^P} = f'(net_j^P) \quad (20)$$

Os resultados acima permitem que a expressão de erro para um neurônio de saída  $j$  possa ser descrita por

$$\delta_j^P = (t_j^P - y_j^P) \cdot f'(net_j^P) \quad (21)$$

Como o ajuste de pesos é dado por  $\Delta w_{ij} = \alpha \delta_j^P x_i$ , seu valor na camada de saída é definido por:

$$\Delta w_{ij} = \alpha \cdot (t_j^P - y_j^P) \cdot f'(net_j^P) \cdot x_i \quad (22)$$

## B.2 $j$ é neurônio pertencente à camada escondida

Nesse caso, o neurônio  $j$  encontra-se na camada escondida e se comunicam com neurônios  $k$ , da camada de saída, conforme pode ser visto na Figura 50.

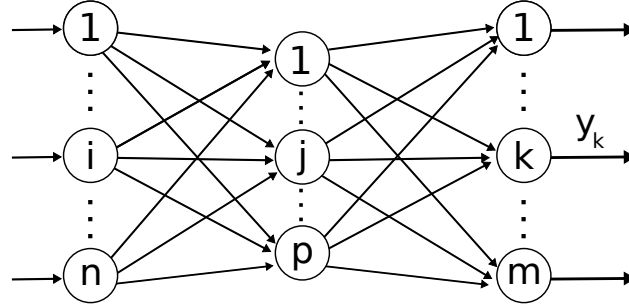


Figura 50 – Esquema de neurônio na camada escondida.

A determinação do  $\delta_j^P$  consiste na adequação de suas derivadas parciais:

$$\delta_j^P = \underbrace{-\frac{\partial E^P}{\partial y_j^P}}_I \cdot \underbrace{\frac{\partial y_j^P}{\partial net_j^P}}_{II} \quad (23)$$

O cálculo de  $I$  é simples, uma vez que

$$\frac{\partial E^P}{\partial y_j^P} = \frac{\partial f(net_j^P)}{\partial net_j^P} = f'(net_j^P) \quad (24)$$

Já o cálculo de  $II$  exige uma análise mais aprofundada, já que ela envolve o erro da camada de saída e a saída de um neurônio da camada escondida. Inicialmente, considere a expressão do erro quadrático médio, que é calculado na camada de saída da rede neural:

$$E^P = \frac{1}{2} \sum_k (t_k^P - y_k^P)^2 \quad (25)$$

A expressão  $II$  pode, por tanto se escrita da seguinte forma:

$$\frac{\partial E^P}{\partial y_j^P} = \frac{\partial}{\partial y_j^P} \left[ \frac{1}{2} \sum_k (t_k^P - y_k^P)^2 \right] = - \sum_k (t_k^P - y_k^P) \frac{\partial y_k^P}{\partial y_j^P} \quad (26)$$

Dessa maneira é preciso determinar  $\frac{\partial y_k^P}{\partial y_j^P}$ , o que consiste em encontrar uma função que envolva neurônios da camada de saída e da camada escondida ( $y_k = f(y_i)$ ). Essa relação pode ser facilmente obtida utilizando-se a função  $net$ , uma vez que  $net_k = \sum_k w_{jk} y_j + b_k$ . Ou seja,  $y_k = f(net_k)$ .

$$\frac{\partial y_k^P}{\partial y_j^P} = \frac{\partial net_k^P}{\partial y_j^P} \cdot \frac{\partial y_k^P}{\partial net_k^P} \Rightarrow \begin{cases} \frac{\partial net_k^P}{\partial y_j^P} = w_{jk} \\ \frac{\partial y_k^P}{\partial net_k^P} = f'(net_k^P) \end{cases} \quad (27)$$

Logo,

$$\frac{\partial y_k^P}{\partial y_j^P} = f'(net_k^P) w_{jk} \quad (28)$$

Substituindo esse resultado na expressão

$$\frac{\partial E^P}{\partial y_j^P} = - \sum_k (t_k^P - y_k^P) \frac{\partial y_k^P}{\partial y_j^P} \quad (29)$$

obtemos:

$$\frac{\partial E^P}{\partial y_j^P} = - \sum_k \left[ \underbrace{(t_k^P - y_k^P) f'(net_k^P)}_{\delta_k^P} w_{jk} \right] \quad (30)$$

que nos leva a uma expressão que é capaz de envolver a saída do neurônio  $j$  da camada escondida com o erro dos neurônios  $k$  da camada de saída.

$$\frac{\partial E^P}{\partial y_j^P} = - \sum_k \delta_k^P w_{jk} \quad (31)$$

Tais resultados nos permitem, finalmente, encontrar o erro associado ao neurônio  $j$ :

$$\delta_j^P = \frac{\partial E^P}{\partial net_j^P} = - \frac{\partial E^P}{\partial y_j^P} \cdot \frac{\partial y_j^P}{\partial net_j^P} = \left( - \sum_k \delta_k^P w_{jk} \right) f'(net_j^P) \quad (32)$$

Por fim, obtemos a fórmula geral para o treinamento de neurônios  $j$  que se encontram na camada escondida de uma rede neural:

$$\Delta w_{ij} = \alpha \left( \sum_k \delta_k^P w_{jk} \right) f'(net_j^P) x_i \quad (33)$$

---

## Script de Normalização

### C.1 normaliz.sci

O programa *Scilab* pode utilizar novos módulos, contendo novas rotinas ou operações, para expandir suas funcionalidades, como visto na seção 3.2.2. Estes módulos, chamados *scripts*, são desenvolvidos como arquivos contendo trechos de código que serão utilizados para operações que as rotinas internas do programa não suportam.

O código seguinte apresenta o conteúdo do arquivo “normaliz.sci”, que implementa a rotina “normaliz”, responsável pela normalização dos dados dos arquivos originais de treinamento entre [0,1]. Este arquivo, desenvolvido pelo professor Dr. Fernando Schmidt, deve ser carregado como *script* dentro do programa *Scilab*. O *script* exige que as matrizes de dados originais tenham as variáveis distribuídas em colunas e as amostras destas nas linhas.

```
1 function [y] = normaliz(x)
2 // Normalização da matriz X entre 0 e 1
3 // y : matriz de saída com os dados normalizados de x
4 y=[];
5
6 // Modo de exibição
7 mode(0);
8
9 // Exibir aviso para a exceção de ponto flutuante
10 ieee(1);
11
12 // I/O formato: [y] = normaliz(x);
13 [k,n] = size(x);
14
15 mm = max(x,'r'); //valor maximo de x, para cada coluna
16
17 mi = min(x,'r'); //valor minimo de x, para cada coluna
18
19 for i = 1:k
20     for j = 1:n
21         y(i,j) = (x(i,j) - mi(j))/(mm(j) - mi(j));
22     end;
23 end
24
25 endfunction
26
```