

REDUÇÃO DE VARIÁVEIS DE ENTRADA DE REDES NEURAS ARTIFICIAIS A PARTIR DE DADOS DE ANÁLISE DE COMPONENTES PRINCIPAIS NA MODELAGEM DE OXIGÊNIO DISSOLVIDO**Saulo Rodrigues e Silva^{a,*} e Fernando Schmidt^{a,b}**^aInstituto Federal de Educação, Ciência e Tecnologia de Goiás, Campus Inhumas, 75400-000 Inhumas – GO, Brasil^bInstituto Federal de Educação, Ciência e Tecnologia de Goiás, Campus Goiânia, 74055-110 Goiânia – GO, Brasil

Recebido em 23/07/2015; aceito em 09/11/2015; publicado na web em 18/02/2015

REDUCTION OF INPUT VARIABLES IN ARTIFICIAL NEURAL NETWORKS AS FROM PRINCIPAL COMPONENTS ANALYSIS DATA IN THE MODELING OF DISSOLVED OXYGEN. The objective of this work is to demonstrate the efficient utilization of the Principal Components Analysis (PCA) as a method to pre-process the original multivariate data, that is rewrite in a new matrix with principal components sorted by its accumulated variance. The Artificial Neural Network (ANN) with backpropagation algorithm is trained, using this pre-processed data set derived from the PCA method, representing 90.02% of accumulated variance of the original data, as input. The training goal is modeling Dissolved Oxygen using information of other physical and chemical parameters. The water samples used in the experiments are gathered from the Paraíba do Sul River in São Paulo State, Brazil. The smallest Mean Square Errors (MSE) is used to compare the results of the different architectures and choose the best. The utilization of this method allowed the reduction of more than 20% of the input data, which contributed directly for the shorting time and computational effort in the ANN training.

Keywords: Principal Components Analysis; Artificial Neural Networks; dissolved oxygen; modeling.

INTRODUÇÃO

A água é um dos bens mais importantes para a manutenção da vida em nosso planeta. Como componente essencial da riqueza de qualquer país, governos têm dedicado atenção ao seu controle e preservação, na forma de órgãos reguladores e um complexo arcabouço legal. Dada sua importância estratégica para a saúde de uma população, instituições não governamentais também têm dedicado atenção ao tema. Assim, estudos regulatórios e relatórios não governamentais têm sugerido um grande aumento na demanda deste recurso nas próximas décadas, especialmente para uso agricultável. Em face deste cenário, faz-se necessário o investimento em ferramentas de apoio à decisão que possibilitem a avaliação dos recursos existentes em termos não apenas quantitativos, mas principalmente qualitativos. Dentre as ferramentas existentes, a utilização de modelos matemáticos para avaliar o comportamento do sistema é comum. O modelo utilizado deve ser capaz de realizar simulações para diferentes cenários de desenvolvimento na bacia, sendo necessário ajustá-lo às condições naturais através de dados de monitoramento.¹ O processo de modelagem de sistemas ecológicos, hidrodinâmicos ou bioquímicos busca conhecer seu comportamento a partir do estabelecimento de hipóteses sobre seu funcionamento, e vem sendo utilizado há muito tempo na ciência.² 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.³ 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.^{4,5} 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,² agindo como especialistas, quando corretamente utilizados. Pesquisadores que modelaram o reservatório número 2 do Rio Mississipi observaram vários modelos surgirem para o mesmo reservatório desde a década de 60.^{6,7} 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. 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”.⁸ 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 Oxigênio Dissolvido (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.⁹ 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.^{7,3} A modelagem através de 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. Mesmo se tratando de um método não linear, a construção de modelos baseados em redes neurais requer pouco tempo, além de mínimo esforço computacional após a fase de treinamento.

*e-mail: saulors@yahoo.com

REDES NEURAIS ARTIFICIAIS

A neurociência vem, há muitas décadas, estudando o funcionamento do cérebro biológico. A partir destes conhecimentos, foi proposto o modelo de neurônio artificial, visto na Figura 1, por MacCulloch e Pitts, posteriormente aprimorado por Rosenblatt, que o nomeou de *Perceptron*^{10,11}

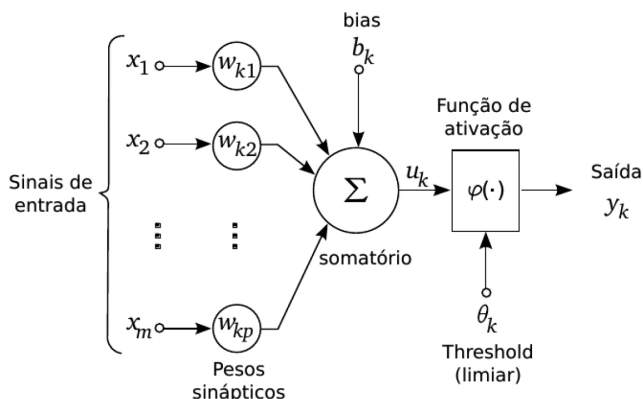


Figura 1. Modelo de Neurônio Artificial, segundo McCulloch-Pitts¹⁰

O neurônio artificial proposto implementa: x_m são sinais de entrada da rede; w_{kp} 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; Σ é a função de somatório; $\phi(\cdot)$ é função de ativação; y_k é a saída (única, em direção a outros neurônios). Um conjunto de neurônios artificiais conectados entre si representam uma Rede Neural Artificial (RNA), que tem como objetivo processar um conjunto de dados de calibração, inseridos na camada de entrada da rede, propagando-se por cada neurônio, o qual terá um número de entradas e uma saída. Os resultados, medidos através do sinal de saída, constituem-se no cálculo do menor erro RMS (“*Root Mean Square*”) ou erro quadrático médio, segundo a Equação (1), dirão se a rede conseguiu “aprender” o padrão que busque descrever o sistema estudado. Os pesos são os parâmetros ajustáveis da rede e devem ser reajustados à medida que o conjunto de calibração 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.

$$RMS = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (1)$$

em que y_i é a saída calculada e \hat{y}_i é a saída da calibração. Uma vez aprendido o padrão, dados novos são utilizados para determinar a acuidade do treinamento. É esperado que os valores de erro RMS oriundos dos testes sejam tão baixos quanto do conjunto de calibração.

Arquitetura (otimização)

Um importante aspecto das RNA é a sua arquitetura, pois a configuração da rede neural depende da particularidade do problema que se quer resolver,¹² restringindo sua aplicação para outros problemas sem a sua devida reconfiguração. Deste modo, antes de ser utilizada, a arquitetura da RNA necessita representar os dados com os quais se pretende trabalhar, para que seja possível mapear as relações não-lineares entre sua(s) entrada(s) e a saída pretendida. Alguns dos parâmetros que modificam esta arquitetura 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 (para mais de uma).

Algoritmo de treinamento *Backpropagation*

O algoritmo de correção dos pesos sinápticos utilizado neste experimento é o *Backpropagation (BP)*, ou Retropropagação, cuja principal aplicação é a modelagem de dados não-lineares, ou seja, que não apresentam uma função de ordem definida para o comportamento matemático das variáveis.¹³ Na Equação 2, é resumido o cálculo da correção de pesos deste algoritmo de treinamento:

$$\Delta w^k = \eta \cdot D + m \cdot \Delta w^{k-1} \quad (2)$$

onde: Δw^k representa a diferença entre os pesos inicial e final para a iteração k ; η 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; Δw^{k-1} representa a diferença entre os pesos inicial e final para a iteração $k - 1$.

ANÁLISE DOS COMPONENTES PRINCIPAIS

Análise de Componentes Principais, do termo inglês *Principal Component Analysis (PCA)*, é definida como a ferramenta multivariada de reconhecimento de padrões não supervisionado que consiste na mudança de base dos dados multivariados, representados pela matriz X , contendo os dados das amostras.¹⁴ 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 de “*scores*” e “*loadings*”, de acordo com a Equação 3.

$$X = S L^T \quad (3)$$

onde X é a matriz de dados originais, consistindo de n linhas (amostras) e p colunas (variáveis); S 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 o símbolo para identificar a transposta da matriz de *loadings* L , gerada pelo PCA, que troca as linhas pelas colunas, como presente na Equação 3.

MATERIAL E MÉTODOS

Este trabalho foi feito utilizando o software livre NEUROPH v2.8, disponível para *download*.¹⁵ Pode-se ter acesso ao sistema INFOÁGUAS, da CETESB, através do sítio na internet,¹⁶ 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 utilizando-se dados reais, foram reunidas informações sobre o Rio Paraíba do Sul. Os dados, totalizando 155 amostras, são oriundos do ponto de coleta PARB 02310, localizado na ponte de acesso ao loteamento Urbanova no município de São José dos Campos -SP, nas coordenadas 23° 11' 45.0348" sul de latitude e 45° 55' 49.3032" oeste de longitude. Buscaram-se dados de onze parâmetros físico-químicos, coletados e analisados entre 18/01/90 e 05/12/12: Temperatura da água, pH, Demanda Bioquímica de Oxigênio (DBO), Condutividade, Nitrogênio Amônia, Nitrato, Nitrito, Fósforo, Turbidez, Sólidos Dissolvidos Totais e Oxigênio Dissolvido. Os parâmetros de inicialização da RNA para todos os experimentos foram padronizados e estão apresentados na Tabela 1:

Os experimentos de treinamento devem parar em duas condições específicas *i*) quando o valor configurado da variável “Erro

Tabela 1. Configuração para treinamento da RNA

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

Máximo” for alcançado; *ii*) ou quando o valor configurado da variável “Limite de iteração” for alcançado. Antes de sua utilização na rede neural, o conjunto total de amostras é dividido em dois conjuntos, contendo os dados de treinamento (ou calibração), que visa promover o aprendizado da RNA, e outro contendo os dados de teste ou verificação, responsável por verificar a capacidade de generalização da rede, sendo que neste trabalho foram utilizadas 109 amostras (70%) como dados para treinamento e 46 amostras (30%) como dados para verificação.

Condicionalmento dos dados para processamento

Os dados do referido ponto de amostragem estão gravados em arquivo com formato de texto puro (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., estando limpos e por este motivo, são os mais indicados para transportar dados entre plataformas (sistemas operacionais) e programas diferentes) contendo as informações de cada parâmetro físico-químico e os valores de suas leituras, oriundos das análises das amostras de água bruta. Para minimizar problemas oriundos do uso de unidades e faixas de valores distintos 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 *sigmoidal* exige que estes dados estejam normalizados no intervalo [0,1]. Esta função é a que melhor apresenta resultados para aplicações em dados não lineares, por isto foi escolhida para este trabalho.¹⁷ Foi desenvolvido um *script* no ambiente computacional *Scilab* para implementar esta técnica de normalização linear no intervalo [0,1], de acordo com a seguinte equação 4:

$$X_n = \frac{X_o - X_{min}}{X_{max} - X_{min}} \quad (4)$$

A função de normalização inversa obedece à equação:

$$X_o = X_n(X_{max} - X_{min}) + X_{min} \quad (5)$$

sendo: X_o : x original; X_n : x normalizado; X_{max} : maior valor de x por parâmetro; X_{min} : menor valor de X por parâmetro.

RESULTADOS E DISCUSSÃO

Em geral, o desenvolvimento de aplicação de RNA segue as seguintes etapas:⁸ (1) obtenção de dados; (2) separação de conjuntos para treinamento e verificação; (3) configuração da rede; (4) treinamento; (5) verificação.

Modelagem com dados reais

O treinamento da RNA com várias configurações de arquitetura foi feito com dados de análise de água bruta do Rio Paraíba do Sul através da variação do número de neurônios 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 neurônios 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). O gráfico da Figura 2 apresenta a dinâmica do Erro Total da Rede, registrado durante o treinamento, e do Erro Quadrático Médio (a diferença entre a estimativa e o valor real elevada ao quadrado), 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 a literatura¹⁸ 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 (10000).

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 $4,21 \times 10^{-2}$. 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 3 apresenta os valores iniciais da variável OD, antes de serem padronizados e utilizados no conjunto de teste, contra o OD modelado pela rede, depois da efetuada a normalização reversa. Verifica-se que, com erro quadrático médio alto desta arquitetura, as diferenças entre os valores originais e modelados são consideráveis,



Figura 2. Erro Total da Rede e Erro Médio Quadrático da rede (10-x-1) para o Rio Paraíba do Sul

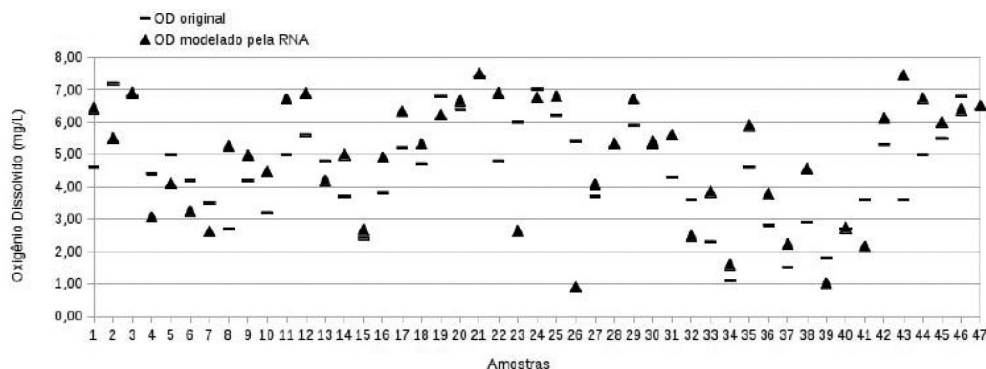


Figura 3. Dado original contra dado modelado pela RNA para o Rio Paraíba do Sul

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.

Modelagem com dados de PCA

Desenvolveu-se uma metodologia buscando verificar se, após submetidos os dados das amostras originais ao método PCA, existe redução no 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 análise de componentes principais tem a função de reduzir sistematicamente o número de variáveis independentes para um conjunto menor, conceitualmente mais coerente de variáveis.¹⁹ A Tabela 2 mostra o percentual acumulado de variância para cada número de CP, obtida através do cálculo do PCA no ambiente *Scilab*, que representa respectivamente os autovalores e sua variância total.

Como se pode observar na Tabela 2, os oito primeiros componentes principais representam uma variância acumulada total de 90,02% dos dados originais. Por mais que as RNA sejam conhecidas pela sua tolerância a dados ruidosos ou incompletos, buscou-se para este experimento uma representação mínima de 90% de variância acumulada, o que justifica o número escolhido de CP, muito próximo do original. Deste modo, as RNA são treinadas e verificadas para várias configurações de arquitetura, variando-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 neurônio na camada intermediária e um sinal na camada de saída. Na segunda configuração, acrescentou-se um neurônio na camada intermediária (8-2-1) e assim sucessivamente, até a configuração (8-15-1). Utilizou-se na entrada da RNA oito sinais (variáveis) normalizadas de entrada:

Tabela 2. Resultados para a PCA aplicada aos dados do Rio Paraíba do Sul

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

conjunto de *Scores* de 1 a 8, e apenas um sinal (variável) normalizado de saída: Oxigênio Dissolvido. A Figura 4 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 oito entradas e uma camada intermediária, variando o número de neurônios desta camada entre 1 e 15. Nota-se que apesar de a 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).

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 5 apresenta os valores iniciais da variável OD, antes de serem padronizados e utilizados no conjunto de teste, contra o OD modelado pela rede, depois da efetuada a normalização reversa.

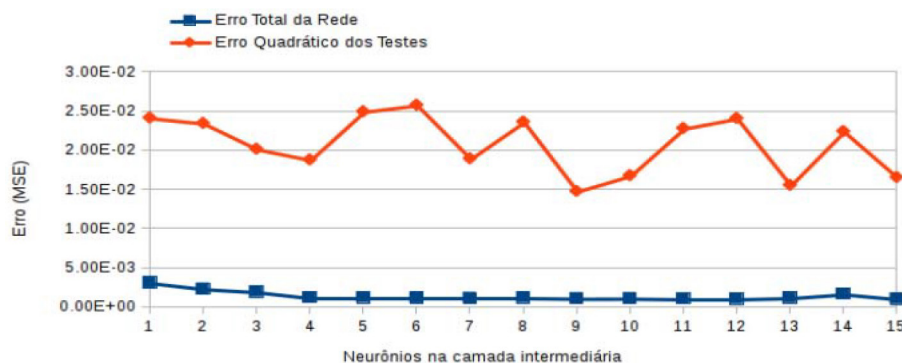


Figura 4. Erro Total da Rede e Erro Médio Quadrático da rede (8-x-1) em função do número de neurônios na camada intermediária para o Rio Paraíba do Sul - PCA

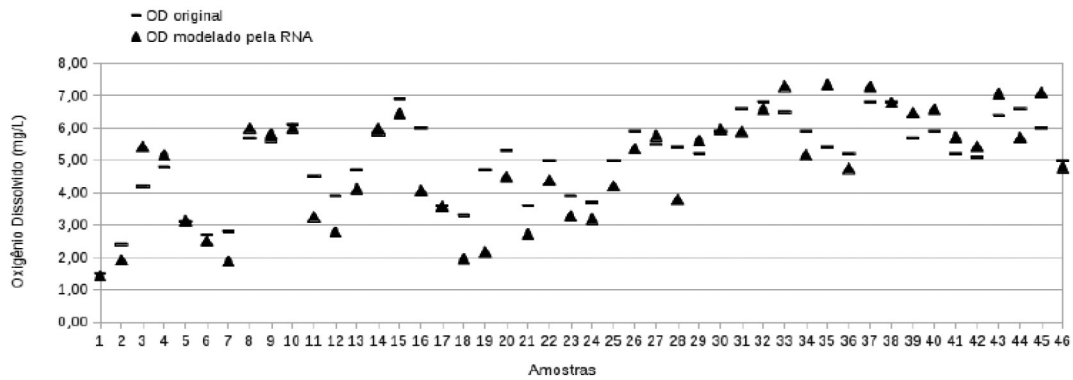


Figura 5. Dado original contra dado modelado pela RNA para o Rio Paraíba do Sul - PCA

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. Para confirmar estas diferenças foi feito um teste t pareado entre os pares de dados dos parâmetros originais e os obtidos pelo método de PCA, buscando verificar se há diferenças estatisticamente significativas (Tabela 3). O resultado do teste mostra que existem diferenças entre os modelos com 95% de confiança ($\alpha = 0,05$). O modelo obtido com o método PCA produziu resultados melhores para este ponto de amostragem quando comparados com o modelo utilizando dados reais (10 entradas), o que pode ser também confirmado pelo valor de coeficiente de correlação calculado em 0,761.

Tabela 3. Parâmetros do teste t pareado aplicado aos modelos

	Média	Variância
Dados originais	4,20768	2,11106
Dados do PCA	5,28022	0,71113

$t = -4,57996$
 $p = 3,67135 \cdot 10^{-5}$
 $\alpha = 0,05 \quad n = 46$

A Figura 6 apresenta a comparação conjunta entre os pares ordenados representando os valores de OD para o Rio Paraíba do Sul, modelados a partir de dados reais e os advindos do método de PCA. Como se verifica, 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 do método 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. Testes posteriores mostraram que um número menor ou maior do que 8 componentes principais não proporcionaram resultados melhores para o coeficiente de correlação R.

CONCLUSÃO

Neste trabalho buscou-se verificar o impacto da utilização de PCA como método de redução de dados no treinamento de RNA. Inicialmente a RNA foi treinada com dados originais de 11 variáveis do Rio Paraíba do Sul. Verificou-se o erro RMS e um coeficiente de correlação calculado de 0,428 para este experimento. Posteriormente, a matriz de componentes principais gerada através do método de PCA, que permite descrever os dados originais em um número menor de CP, foi utilizada para alimentar a entrada da RNA. Deste modo, a rede foi treinada com oito entradas, representando os oito componentes principais, que se mostraram mais eficientes como substitutos dos dados originais. Este resultado pôde ser comprovado através da verificação de menor erro RMS e coeficiente de correlação calculado de 0,761. Portanto, verifica-se que a rede treinada com dados provenientes de redução de variáveis por PCA apresentou menor erro que a rede treinada com dados originais. Mesmo que a RNA treinada com estes dados não tenha convergindo para o menor erro esperado, ou seja, aquele configurado na variável “Erro máximo” antes do treinamento, foi demonstrado que a técnica aumentou a variância dos dados que puderam ser encontrados em número menor de CP. Adicionalmente, verifica-se que o menor número de neurônios

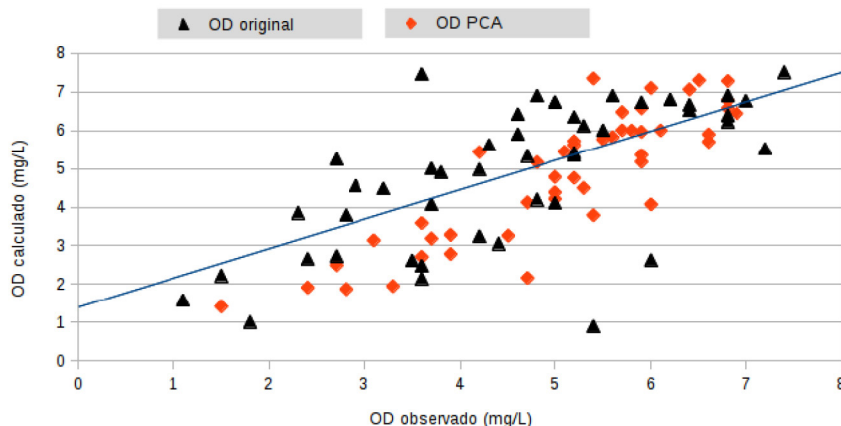


Figura 6. Comparação entre Modelo com dados reais (10 variáveis) e PCA (8 “scores”) para o Rio Paraíba do Sul

na camada de entrada exige menor custo computacional para o reconhecimento do padrão, poupando esforço computacional e tempo de treinamento, principalmente para grandes bases de dados. Este resultado se mostra importante na modelagem do OD para um corpo hídrico, na medida em que determinar este parâmetro é fundamental para se criar um modelo preditivo que ajude no entendimento sobre os impactos dos diversos tipos de uso da água para o sistema hídrico em questão. Outros estudos são importantes para o aperfeiçoamento deste resultado. A avaliação de diferentes 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,^{20,21} ou mesmo de algoritmos genéticos,²² para determinação automática da melhor arquitetura da rede neural, reduzindo etapas ao eliminar a necessidade de testes com variadas arquiteturas.

REFERÊNCIAS

1. Teodoro, A.; Ide, C. N.; Ribeiro, M. L.; Broch, S. A. O.; Silva, J. B.; *Eng. Sanit. Ambiental* **2013**, *18*, 275.
2. Cunha, A. C. D.; *Biota Amazônia* **2013**, *3*, 124.
3. <https://fenix.tecnico.ulisboa.pt/downloadFile/3779571269514/Oxig?nio%20Dissolvido.pdf>, Acessado em Janeiro 2016.
4. Warwick, J. J.; *Water Sci. Technol.* **1991**, *24*, 41.
5. Siqueira, E. Q.; *Dissertação de Mestrado*, Escola de Engenharia de São Carlos, Departamento de Hidráulica e Saneamento, USP, São Carlos, Brasil, 1996.
6. Demetraco-Poulos, A. C.; Stefan, H. G.; *J. Environ. Eng.* **1983**, *109*, 1006.
7. Scariot, M. R.; *Tese de Doutorado*, Faculdade de Engenharia de Alimentos, Universidade Estadual de Campinas, Campinas, Brasil, 2008.
8. Santos, C. C. D.; Filho, A. J. P.; *Revista Brasileira de Meteorologia* **2003**, *18*, 149.
9. Nunes, D. G.; *Dissertação de Mestrado*, Universidade Federal de Viçosa, Viçosa, Brasil, 2008.
10. MacCulloch, W.S.; Pitts, W.; *Bull. Math. Biophys.* **1943**, *5*, 115.
11. Rosenblatt, F.; *Psychological Review* **1958**, *65*, 386.
12. Zurada, J. M.; *Introduction to Artificial Neural Systems*, PWS Publishing Company, EUA, 1995.
13. Schimidt, F.; Bueno, M. I. M. S.; Poppi, R. J.; *Quim. Nova* **2002**, *25*, 946.
14. Godinho, M. D. S.; Pereira, R. O.; Ribeiro, K. D. O.; Schimidt, F.; Oliveira, S. B.; Oliveira, A. E.; *Quim. Nova* **2008**, *31*, 1485.
15. <http://neuroph.sf.net>. Acessado em Janeiro de 2016.
16. <http://www.cetesb.sp.gov.br>. Acessado em Janeiro 2016.
17. Gomes, L. F. C.; Montenegro, S. M. G. L.; Valença, M. J. D. S.; *Revista Brasileira de Recursos Hídricos* **2010**, *15*, 5.
18. e Silva, S. R.; *Dissertação de Mestrado*, Instituto Federal de Educação, Ciência e Tecnologia de Goiás, Goiânia, Brasil, 2014.
19. Dunteman, G. H.; *Principal Components Analysis*, Sage University Paper, v. 69, EUA, 1989.
20. Parpinelli, R. S.; *Dissertação de Mestrado*, Centro Federal De Educação Tecnológica do Paraná, Curitiba, Brasil, 2001.
21. e Souza, S. B. S.; *Tese de Doutorado*, Universidade Federal de Goiás, Goiânia, Brasil, 2014.
22. Fernandes Junior, F. E.; *Dissertação de mestrado*, Universidade Estadual de Campinas, Campinas, Brasil, 2014.