

UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

YURI LIMA CAVALCANTE

APLICAÇÃO DAS TÉCNICAS ANÁLISE  
MULTIVARIADA E REDES NEURAIS  
ARTIFICIAIS NA CLASSIFICAÇÃO DAS  
ÁGUAS DE RESERVATÓRIOS DE  
HIDRELÉTRICAS: UM ESTUDO DE  
CASO NA REGIÃO AMAZÔNICA

UFPA / ICEN / PPGCC  
Campus Universitário do Guamá  
Belém-Pará-Brasil  
2013



UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

YURI LIMA CAVALCANTE

APLICAÇÃO DAS TÉCNICAS ANÁLISE  
MULTIVARIADA E REDES NEURAS  
ARTIFICIAIS NA CLASSIFICAÇÃO DAS  
ÁGUAS DE RESERVATÓRIOS DE  
HIDRELÉTRICAS: UM ESTUDO DE  
CASO NA REGIÃO AMAZÔNICA

UFPA / ICEN / PPGCC  
Campus Universitário do Guamá  
Belém-Pará-Brasil  
2013

UNIVERSIDADE FEDERAL DO PARÁ  
INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

YURI LIMA CAVALCANTE

APLICAÇÃO DAS TÉCNICAS ANÁLISE  
MULTIVARIADA E REDES NEURAIS  
ARTIFICIAIS NA CLASSIFICAÇÃO DAS  
ÁGUAS DE RESERVATÓRIOS DE  
HIDRELÉTRICAS: UM ESTUDO DE  
CASO NA REGIÃO AMAZÔNICA

Dissertação submetida à Banca  
Examinadora do Programa de Pós-  
Graduação em Ciência da Computação da  
UFPA para a obtenção do Grau de Mestre  
em Ciência da Computação.

UFPA / ICEN / PPGCC  
Campus Universitário do Guamá  
Belém-Pará-Brasil  
2013



## **AGRADECIMENTOS**

À todos que me auxiliaram durante o desenvolvimento deste trabalho, em especial à minha mãe, à minha co-orientadora prof. Dr. Terezinha Ferreira de Oliveira e ao meu orientador prof. Dr. Antonio Morais da Silveira.

## RESUMO

Esta pesquisa apresenta uma metodologia para classificação de águas de reservatórios de hidrelétricas da Amazônia utilizando Análise Multivariada de Dados e Redes Neurais Artificiais (RNA). As amostras de água dos reservatórios das hidrelétricas de Coaracy Nunes e Tucuruí foram coletadas durante o período de Junho de 2008 à Setembro de 2009. Os parâmetros físico-químicos analisados foram: temperatura, condutividade, pH, turbidez, sólidos totais em suspensão, alcalinidade, amônia, demanda química de oxigênio, dureza e oxigênio dissolvido. Os metais analisados foram: Alumínio (Al), Bário (Ba), Cálcio (Ca), Ferro (Fe), Potássio (K), Magnésio (Mg), Manganês (Mn) e Sódio (Na). Para cada reservatório foi aplicada a Análise Fatorial para redução de variáveis. As variáveis constituintes dos fatores resultantes foram utilizadas como variáveis de entrada das RNAs propostas para classificar as amostras de água como pertencentes ao período seco ou chuvoso, a classificação resultante do uso das RNAs é comparada com os a classificação realizada pela Análise Discriminante. Os resultados obtidos demonstram que a combinação das técnicas estatísticas com Redes Neurais Artificiais pode ser utilizada para geração de soluções robustas em problemas de classificação, sendo estas aplicações estatísticas também relevantes em casos de gestão ambiental e política de tomada de decisão, para identificar quais fatores devem ser estudados e/ou modificados para recuperar corpos de água degradados ou contaminados.

**Palavras-chaves: Reservatório de Hidrelétrica. Análise Multivariada de Dados. Redes Neurais Artificiais. Parâmetros Físico-químicos. Metais.**

## ABSTRACT

This research presents a methodology to classify Amazon's hydroelectric reservoirs waters using Multivariate Data Analysis and Artificial Neural Networks (ANN). Coaracy Nunes and Tucuruí hydroelectric reservoir's water samples were collected between June of 2008 and September of 2009. The physicochemical parameters analyzed were: temperature, conductivity, pH, turbidity, total suspended solids, alkalinity, ammonia, chemical oxygen demand, hardness and dissolved oxygen. The metals analyzed were: Aluminum (Al), Barium (Ba), Calcium (Ca), Iron (Fe), Potassium (K), Magnesium (Mg), Manganese (Mn) and Sodium (Na). For each reservoir were applied the Factorial Analysis to reduce the number of variables. The variables constituting the resulting factors were used as input variables of the proposed ANNs to classify the water samples as belonging to dry or rainy seasons and the resulting classification using ANNs were compared to Discriminant Analysis classification. The results obtained shows that the combination of statistical techniques with Artificial Neural Networks can be used to generate robust solutions in classification problems, being this techniques applications relevant to environmental management and police decision-making process, to identify which factor should be further studied and/or modified to recover degraded or contained water bodies.

**Keywords: Hydroelectric Reservoir. Multivariate Data Analysis. Artificial Neural Networks. Physicochemical Parameters. Metals.**

# SUMÁRIO

<b>INTRODUÇÃO .....</b>	<b>14</b>
<b>1. ESTADO DA ARTE .....</b>	<b>17</b>
<b>2. MATERIAIS E MÉTODOS.....</b>	<b>23</b>
2.1. Área de estudo .....	23
2.2. Coleta de dados.....	25
2.3. Análise dos parâmetros físico-químicos e metais da água .....	27
2.4. Métodos estatísticos e computacionais.....	32
2.4.1. Análise Multivariada.....	32
2.4.2. Redes Neurais Artificiais.....	35
<b>3. APLICAÇÃO DA METODOLOGIA E RESULTADOS .....</b>	<b>39</b>
3.1. Reservatório Coaracy Nunes .....	39
3.2. Reservatório de Tucuruí .....	46
<b>4. CONCLUSÕES .....</b>	<b>53</b>
<b>5. TRABALHOS FUTUROS .....</b>	<b>55</b>
<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>56</b>

## LISTA DE ABREVIações

AC	Análise de Cluster
ACP	Análise de Componentes Principais
AD	Análise Discriminante
AF	Análise Fatorial
AFP	Análise de Fatores Principais
AG	Algoritmo Genético
ANOVA	Análise de Variância
LCS	Learning Classifier System
OD	Oxigênio dissolvido
pH	Potencial de Hidrogênio
RNA	Rede Neural Artificial
SD	Disco de Secchi
STS	Sólidos Totais em Suspensão

## LISTA DE FIGURAS

Figura 2.1: Localização da usina hidrelétrica de Coaracy Nunes .....	24
Figura 2.2: Localização da usina hidrelétrica de Tucuruí .....	25
Figura 2.3: Pontos de coleta distribuídos pelo reservatório de Coaracy Nunes .....	26
Figura 2.4: Pontos de coleta distribuídos pelo reservatório de Tucuruí .....	27
Figura 2.5: Arquitetura típica de uma RNA .....	35
Figura 2.6: (a) função degrau; (b) função sinal; (c) função sigmoide; (d) função linear; (e) função tangente hiperbólica .....	36
Figura 2.7: Perceptron multicamada .....	36
Figura 2.8: Fase de propagação (LNCC, 2012).....	37
Figura 2.9: Fase de retropropagação (LNCC, 2012) .....	37
Figura 3.1: RNA proposta para classificação das amostras de água do reservatório de Coaracy Nunes	40
Figura 3.2: Box-plots do Ca, K, Mg e Na analisados das amostras de água da superfície, meio e fundo .....	44
Figura 3.3: Box-plot do DQO na superfície, meio e fundo .....	44
Figura 3.4: RNA proposta para classificação das amostras de água do reservatório de Tucuruí.....	47
Figura 3.5: Intervalo de confiança das médias das concentrações (mgL <sup>-1</sup> ) do Ba, Fe, K e Na.....	50
Figura 3.6: Box –plots com os intervalos de confiança das médias dos parâmetros do fator 3 .....	51

## LISTA DE QUADROS

Quadro 2.1: Campanhas de coleta de Coaracy Nunes e de Tucuruí com os respectivos períodos de coleta .....	25
Quadro 2.2: Valores recomendados por diferentes instituições governamentais para os parâmetros físico-químicos .....	30
Quadro 2.3: Condições operacionais do ICP-OES.....	31
Quadro 2.4: Valores recomendados por diferentes instituições governamentais para os metais .....	31

## LISTA DE TABELAS

Tabela 3.1: Resultado dos parâmetros físico-químicos medidos em Coaracy Nunes .....	40
Tabela 3.2: Concentrações dos metais medidas no reservatório de Coaracy Nunes .....	42
Tabela 3.3: Fatores, variância total e comunalidades obtidos pela análise de componente principal e análise fatorial de Coaracy Nunes .....	44
Tabela 3.4: Resultado dos parâmetros físico-químicos medidos em Tucuruí.....	48
Tabela 3.5: Concentrações dos metais em $\text{mg/L}^{-1}$ medidas no reservatório de Tucuruí .....	48
Tabela 3.6: Fatores, variância total e comunalidades obtidos pela análise fatorial com extração por análise de componente principal e rotação varimax de Tucuruí .....	49

## INTRODUÇÃO

A água é um dos constituintes mais essenciais do ambiente humano. Este recurso natural gera desenvolvimento socioeconômico para a sociedade em geral, mais especificamente para a indústria, agricultura e uso público. Qualidade da água se refere às características físicas, químicas e biológicas necessárias à existência da vida, além de ser um dos principais fatores envolvidos no controle da saúde e dos estados patológicos nos organismos vivos. Esta, em uma região, é determinada tanto por processos naturais, como intemperismo e erosão do solo, quanto pela ação antropogênica (SINGH, *et al.*, 2004; KAZI, *et al.*, 2009). O monitoramento das variáveis da qualidade da água é de fundamental importância para a sociedade. Seus valores devem estar dentro de padrões definidos por legislações, já que a presença de certos poluentes em determinadas concentrações pode representar uma ameaça tanto à saúde humana quanto aos ecossistemas (VRANA, *et al.*, 2005).

Os parâmetros físico-químicos (demanda química de oxigênio, pH, oxigênio dissolvido, etc) de corpos de água podem apresentar características próprias que variam de acordo com o ambiente no qual estão inseridos. Estas variações podem afetar diretamente a biota local, a qual pode ou não estar preparada para lidar com estas variações. As concentrações iônicas das águas da Amazônia ilustram muito bem as diferenças regionais que ocorrem nesta área (TUNDISI e TUNDISI, 2008).

A presença de metais em corpos de água é importante para o ciclo geoquímico e para saúde ambiental, levando a preocupações sobre suas influências em diversos ecossistemas quando em concentrações elevadas (KABATA-PENDIAS, 2011). Nestas concentrações os metais estão entre os poluentes ambientais mais comuns nas águas e em biotas associadas, sendo introduzidos no sistema a partir de fontes naturais ou antropogênicas (PAPAFILIPPAKI, KOTTI e STAVROULAKIS, 2008).

Devido ao desenvolvimento e crescimento econômico e populacional, países tropicais vêm construindo cada vez mais reservatórios, aumentando, portanto, a demanda por água potável e poder hidrelétrico (ATO BATELE e UGWUMBA, 2008), tornando os reservatórios o tipo de

lago predominante em muitas destas regiões no mundo (LEWIS, 2000), como é o caso do Brasil que vem implantando várias hidrelétricas nas últimas décadas. Reservatórios estão entre os rios e lagos com a hidrodinâmica grandemente influenciada pelos seus regimes de gerência operacional. Embora alguns reservatórios possuam características comuns com os lagos, outras características são específicas destes (CHAPMAN, 1996).

Diversos estudos foram conduzidos com o intuito de identificar e caracterizar variações sazonais nos parâmetros físico-químicos e nas concentrações dos metais em ambientes aquáticos (WENNER, *et al.*, 2004; OUYANG, 2005; KAZI, *et al.*, 2009; WANG, *et al.*, 2009; ZAHRAIE e HOSSEINI, 2009; COLETTI, *et al.*, 2010; HAUSER-DAVIS, *et al.*, 2012; SONG, *et al.*, 2011; YENILMEZ, KESKIN e AKSOY, 2011). Dentre as técnicas mais utilizadas para avaliação destes parâmetros destacam-se análise multivariada (OUYANG, 2005; KAZI, *et al.*, 2009; COLETTI, *et al.*, 2010; SONG, *et al.*, 2011; YENILMEZ, KESKIN e AKSOY, 2011) e, mais recentemente, técnicas de inteligência computacional (WANG, *et al.*, 2009; ZAHRAIE e HOSSEINI, 2009; HAUSER-DAVIS, *et al.*, 2010).

Neste contexto, este estudo tem como objetivo propor um método qualitativo utilizando técnicas estatísticas e da inteligência computacional para classificação de águas de reservatórios de hidrelétricas da região amazônica quanto à sazonalidade. Para atingir este objetivo foi necessário identificar um reservatório para estudo de caso, em seguida realizar uma análise exploratória na base de dados do reservatório utilizando Análise Exploratória de Dados e Análise Multivariada, identificar parâmetros relevantes para classificação das águas de reservatórios de hidrelétricas quanto à sazonalidade, desenvolver um modelo de Rede Neural Artificial (RNA) para classificar as amostras de água, classificar estas mesmas amostras utilizando Análise Discriminante, comparar e validar os resultados obtidos pela RNA com os resultados obtidos pela Análise Discriminante e por fim aplicar a metodologia proposta a outro estudo de caso.

O primeiro reservatório escolhido como estudo de caso foi o reservatório da usina hidrelétrica de Coaracy Nunes, por ser considerado um reservatório pouco profundo, e o segundo reservatório foi o da hidrelétrica de Tucuruí, por ser considerado um reservatório profundo. As coletas para a caracterização dos parâmetros físico-químicos e dos metais presentes na água foram realizadas entre os anos de 2008 e 2009. Nesse período foram realizadas quatro campanhas trimestrais de coleta de amostras de água nos reservatórios das usinas citadas.

Espera-se que a abordagem proposta auxilie no processo de tomada de decisão em relação à poluição de lagos, servindo como uma alternativa para classificação de águas desses ambientes na Amazônia, contribuindo, portanto com o debate sobre este importante e vital recurso natural, problema este que interessa toda a sociedade.

No capítulo 1 é apresentado o levantamento do estado da arte sobre o tema. No capítulo 2 são detalhados a área de estudo, o processo de coleta de dados, o processo de análise dos parâmetros físico-químicos e metais, bem como os métodos estatísticos e computacionais utilizados. O capítulo 3 aborda a aplicação da metodologia e análise dos resultados. No capítulo 4 estão presentes as conclusões do trabalho. No capítulo 5 estão presentes sugestões de trabalhos futuros.

## 1. ESTADO DA ARTE

Diversos estudos foram conduzidos com o intuito de identificar e caracterizar variações sazonais nos parâmetros físico-químicos e nas concentrações dos metais em ambientes aquáticos. Dentre as técnicas mais utilizadas destacam-se análise multivariada e técnicas de inteligência computacional.

Simeonov, *et al.* (2003) apresentam a necessidade e utilização da análise estatística multivariada em bancos de dados grandes e complexos, a fim de obter melhores informações sobre a qualidade das águas superficiais, o projeto de amostragem, protocolos analíticos e a poluição e efetivo controle/gestão das águas de superfície. A aplicação de técnicas estatísticas de análise multivariada foi feita no monitoramento das águas superficiais no norte da Grécia. O conjunto de dados consistiu em resultados analíticos de uma pesquisa de três anos realizada em grandes sistemas fluviais (Aliakmon, Axios, Gallikos, Loudias e Strymon), bem como córregos, afluentes e valas. Vinte e sete parâmetros foram monitorados mensalmente em 25 locais de amostragem totalizando 22.350 observações. O conjunto de dados foi tratado por Análise de Agrupamento, Análise de Componentes Principais e Análise de Regressão Múltipla em Componentes Principais. A Análise de Agrupamento mostrou quatro diferentes grupos de similaridade entre os locais de amostragem refletindo diferentes características físico-químicas e os níveis de poluição dos sistemas de água estudados. Seis fatores latentes foram identificados e explicam 90% da variância total do conjunto de dados que se referem à orgânica de nutrientes, à análise físico-química, ao intemperismo, a lixiviação do solo e à toxidez antropogênica.

Wenner, *et al.* (2004) Analizou dados coletados por sondas automáticas, de 1995 a 2000, em 55 pontos dentro de 22 NERR (Reservas de Pesquisa de Estuários dos Estados Unidos), para caracterizar a variabilidade sazonal e comparar os diversos sistemas estuarinos e regiões. Modelos de análise de variância (ANOVA) revelaram várias diferenças significativas nos parâmetros de qualidade da água em muitos pontos amostrados dentro de uma Reserva; entretanto essa tendência não foi observada nas comparações interanuais dentro de uma

Reserva ou região. Salinidade e eventos hipóxicos (28% de saturação no oxigênio dissolvido) foram maiores no verão e no outono. Evapotranspiração e precipitação demonstraram influência nos padrões de salinidade nas Reservas. Diminuições abruptas em curto prazo, e mantidas em longo prazo, na salinidade associada com eventos de precipitação foram documentadas durante a passagem de sistemas tropicais. Uma diminuição abrupta na temperatura da água durante a passagem de uma tempestade foi observada, com aumento de efeitos de resfriamento relacionado ao aumento da intensidade da tempestade. A hipoxia foi fortemente influenciada pela latitude e pelo clima. A síntese dos parâmetros de qualidade da água indica que Reservas não tem experimentado muitos dos problemas encontrados em outros locais mais populosos.

Ouyang (2005) utilizou as técnicas Análise de Componentes Principais (ACP) e Análise de Fatores Principais (AFP) para avaliar a eficiência de uma rede de monitoramento da qualidade das águas de superfície onde as variáveis avaliadas foram estações de monitoramento, visando identificar estações de monitoramento mais importantes para análise da qualidade da água do rio. Foram analisadas 22 estações de monitoramento de parâmetros físico-químicos e biológicos localizadas no rio St. Johns na Flórida, EUA. O estudo identificou três estações de monitoramento como tendo pouca importância para explicar a variância anual dos dados e a AF revelou que o carbono orgânico total, carbono orgânico dissolvido, nitrogênio total, nitrato e nitrito dissolvido, ortofosfato, alcalinidade, salinidade, Magnésio (Mg) e Cálcio (Ca) foram os parâmetros mais importantes para explicar as variações nos parâmetros da qualidade da água do rio.

Allan, Vrana e Greenwood (2006) consideram as limitações associadas ao monitoramento da qualidade da água (resolução temporal e espacial; informações sobre a biodisponibilidade) e apresentam, na forma de uma revisão, ferramentas de monitoramentos biológicos e de produtos químicos que podem se tornar parte de uma "caixa de ferramentas" de técnicas para utilização por aqueles encarregados de avaliar a qualidade da água em decorrência da implantação da diretiva (2000/60/CE) que exige que os membros da União Européia utilizem ferramentas de baixo custo como parte do monitoramento. As técnicas de monitoramentos biológicos incluem: biomarcadores, biossensores, sistemas de alerta precoce e bioensaios. A metodologia de amostragem para avaliação química compreende: biossensores, imunoenaios, amostragem passiva, e sensores. São feitas a descrição dos dispositivos, bem como uma discussão da adequação desses para diferentes tipos de monitoramento. Os autores resumem a necessidade da garantia de qualidade dos dados bem como a validação dos

métodos e acrescentam que essas técnicas podem complementar os monitoramentos já existentes, fornecendo informações adicionais com o objetivo de se obter uma imagem mais representativa da qualidade de um corpo de água.

Hao, Huang e Zhu (2008) propuseram um modelo de Rede Neural Artificial (RNA) com treinamento back-propagation para avaliar a qualidade da água de Yellow River, China, levando em consideração o oxigênio dissolvido (OD), a demanda biológica de oxigênio (DBO), índice de permanganato e NH<sub>3</sub>-NH<sub>4</sub>. A rede conseguiu atingir uma alta precisão e os resultados podem fornecer referências para o planejamento e proteção de ambientes aquáticos.

Kazi, *et al.* (2009) aplicaram as técnicas análise multivariada, Análise de Cluster (AC) e Análise de Componentes Principais (ACP) para avaliar e interpretar dados complexos de qualidade da água e fontes de poluição para obter informações melhores sobre a qualidade da água e desenvolver um sistema de monitoramento. Foram analisados 36 parâmetros de amostras de água do lago Manchar, Paquistão, entre os anos de 2005 e 2006. As correlações químicas identificadas pela ACP foram utilizadas pela AC para classificar as amostras, identificando três pontos de coleta significantes. Os resultados revelaram que a maior causa de deterioração da qualidade da água está relacionada com a entrada de poluentes de origem industrial, doméstica e agrícola no lago.

Wang, *et al.* (2009) propuseram um sistema de classificação das regras de operação com base no Learning Classifier System (LCS), que aprende através de cessão de crédito (o algoritmo de brigada) e descoberta de regras (o Algoritmo Genético) para extrair de regras de funcionamento do reservatório de abastecimento de água. O sistema proposto adquiriu uma taxa de identificação on-line de 95%, para amostras de treinamento off-line a taxa foi 85%. São feitas comparações entre o sistema de classificação regra e a RNA. Os resultados indicam que o sistema classificador de aprendizado é prático e eficaz para se obter as regras operacionais do fornecimento de reservatório.

Coletti, *et al.* (2010) utilizaram a análise fatorial para propor um índice de qualidade da água que pudesse indicar a influência de atividades agrícolas na qualidade da água. As águas de uma bacia hidrográfica predominantemente agrícola foram monitoradas no período de setembro de 2003 a setembro de 2004, onde foram realizadas coletas mensais de amostras de água de seis pontos de coleta e oito parâmetros foram analisados: nitrato, nitrogênio, amônia, fósforo total (FT), condutividade elétrica, pH, sólidos totais em suspensão (STS) e turbidez.

Os resultados indicaram que a amônia, o nitrogênio e o nitrato foram os parâmetros que mais contribuíram para a determinação do índice.

Hauser-davis, *et al.* (2010) aplicaram a análise discriminante e RNA na classificação de três espécies de peixes no estado do Rio de Janeiro, Brasil. A RNA apresentou resultados satisfatórios, mesmo com grupos compostos de animais de diversos tamanhos, indicando esta técnica como uma excelente alternativa para problemas de classificação com dados desbalanceados.

Matthews, Bernard e Winter (2010) utilizaram a Medium Resolution Imaging Spectrometer (MERIS) para avaliar o impacto hipertrófico em Zeekoevlei, um pequeno lago de água doce situado em Cape Flats, em Cape Town, África do Sul, dominada por microcystinas de cianobactéria. O objetivo do estudo foi avaliar as propriedades ópticas do lago, os procedimentos de correção atmosférica, e comparar o desempenho de algoritmos empíricos e semi-analíticos em água hipertrófica. Os parâmetros de qualidade da água foram obtidos *in situ* e foram feitas medições radiométricas simultânea no MERIS. As medições de radiância em profundidade 0,66 m foram corrigidas e processadas e estimativas verticais foram atenuadas pelo coeficiente de radiação para cima. A reflectância da água foi normalizada e algoritmos empíricos foram utilizados para estimar clorofila-a (Chl a), sólidos totais em suspensão (STS), disco de Secchi (SD) e absorção por CDOM, os quais foram simultaneamente coletados no MERIS e medidos *in situ*. Os algoritmos empíricos forneceram valores do coeficiente de correlação altos, embora tenham uma capacidade limitada para separar as covariantes dos sinais da água. O algoritmo de Rede Neural Artificial (RNA) utilizada no MERIS também foi utilizado para obter as concentrações de constituintes da água. No entanto, estes não conseguiram produzir comparações razoáveis com medições *in situ*, devido à falha de correção atmosférica e divergência entre as propriedades ópticas e faixas usadas para treinamento dos algoritmos. Mapas produzidos que utilizavam os algoritmos empíricos efetivamente mostraram a variabilidade espacial e temporal dos parâmetros de qualidade da água durante abril de 2008. Com base nos resultados argumenta-se que MERIS é o sensor de corrente ideal para aplicações de detecção de mudanças frequentes nas águas

Odermatt, Giardino e Heege (2010) desenvolveram uma rede neural para ser aplicada juntamente com as técnicas semi-analítica de sensoriamento remoto para águas eutrofizadas de lagos oligomesotróficos na área perialpina, já que eles são insensíveis às variações de concentração de clorofila entre 1 e 10 mg/m<sup>3</sup>. A rede neural foi baseada no algoritmo Case-2-

Regional para MERIS, juntamente com o algoritmo de correção ICOL. Os algoritmos foram aplicados a um conjunto de 239 imagens de satélite obtidas no período de 2003 à 2008, e os resultados experimentais foram comparados com os dados oficiais de qualidade de água coletadas em 6 lagos perialpinos no mesmo período. As estimativas da detecção remota podem fornecer uma adequada fonte de dados para complementar séries de dados *in situ* da camada superior de água de 5 m.

Song, *et al.* (2011) analisaram a qualidade da água do rio Beijiang, China, para determinar as influências das áreas agrícolas e urbanas no reservatório. Foram utilizados 18 parâmetros da qualidade da água a partir de coletas mensais de amostras de água de 13 pontos de coleta durante os anos de 2005 e 2006. Uma análise de correlação bivariada foi utilizada para avaliar as correlações dos parâmetros, enquanto a ACP foi aplicada para extrair as variáveis que mais influenciam nas variações regionais da qualidade da água do rio. Seis componentes foram extraídos, os quais juntos explicam mais que 78% e 84% da variância total agrícola e urbana respectivamente. A regressão linear multivariada foi utilizada para estimar a contribuição de todas as fontes de poluição identificadas. Altos coeficientes de determinação das equações de regressão sugerem que a técnica é aplicável para estimativa das fontes da maioria dos parâmetros físico-químicos do rio Beijiang.

No estudo de Yenilmez, Keskin e Aksoy (2011) foram analisadas as tendências dos parâmetros de qualidade da água no lago Eymir durante um período de 10 anos utilizando o teste de Mann-Kendall. Os parâmetros estudados foram: oxigênio dissolvido (OD), fósforo total (FT), sólidos totais em suspensão (STS) e profundidade Secchi (SD). Além disso, foram obtidas as médias de precipitação, o volume do lago, e temperatura ambiente anuais. De acordo com os resultados do teste Mann-Kendall os valores de temperatura ambiente, precipitação e volume do lago apresentam tendências decrescentes de 1998 a 2008. O OD e os STS apresentam tendências crescente, enquanto FT e SD apresentaram tendências decrescentes. A mudança no volume do lago tem um impacto significativo sobre as tendências de OD, STS, e SD. Estes resultados indicam que, além das condições de eutrofização, o balanço hídrico em condições de seca tem um impacto significativo na qualidade da água do lago Eymir.

Xue e chen (2011) desenvolveram um modelo de avaliação de qualidade da água baseado em RNA com back-propagation, tendo como parâmetros de qualidade da água o oxigênio dissolvido (OD), demanda química de oxigênio (DQO), fenol volátil, cianeto, mercúrio,

arsênio e cromo ( $\text{Cr}^{6+}$ ). Exemplos práticos indicaram que este tipo de método de avaliação é bastante confiável e preciso quando comparado com métodos de avaliação convencionais.

O estudo de Khalil, *et al.* (2012) visa investigar o potencial de Redes Neurais Artificiais (RNA) na modelagem da relação entre diferentes variáveis da qualidade da água, com o propósito de reduzir o número de variáveis a serem monitoradas. A pesquisa utilizou dados de 33 registros mensais dos parâmetros da qualidade da água monitorados em 102 pontos do sistema de drenagem do delta do rio Nilo, Egito, obtidos desde agosto de 1997, fornecidos pelo Drainage Research Institute (DRI). Os pesquisadores analisaram o impacto de três fatores na modelagem da RNA: utilização de entradas diferentes, tamanho das amostras de treinamento e teste e o número de neurônios na camada intermediária. As RNAs propostas utilizaram combinações das concentrações do Oxigênio Dissolvido, Demanda Química de Oxigênio e Demanda Biológica de Oxigênio em 3 pontos de coleta como variáveis de entrada e saída das RNAs. Os resultados obtidos pelas redes foram comparados com os resultados obtidos com Regressão Logística. A pesquisa concluiu que RNAs são bastante úteis quando aplicadas para predição dos parâmetros da qualidade da água e que a escolha das variáveis a serem utilizadas para treinamento e teste podem influenciar no erro da rede.

Apesar de serem pesquisas relacionadas com a qualidade da água, os trabalhos anteriores não abordam a classificação de águas de reservatório quanto à sazonalidade e não foram projetados para reservatórios da Amazônia, fazendo desta uma pesquisa original. A quantidade de pesquisas que utilizam análise multivariada e RNAs para solucionar diversos problemas associados à qualidade da água indica que é possível atingir o objetivo da pesquisa utilizando-se estas técnicas.

## 2. MATERIAIS E MÉTODOS

### 2.1. Área de estudo

Para este estudo foram escolhidos o reservatório da usina hidrelétrica de Coaracy Nunes, por ser considerado um reservatório raso e o reservatório da usina hidrelétrica de Tucuruí, por ser considerado um reservatório profundo, com profundidade máxima de 77 m e profundidade média de 19,8 m (ELETRONORTE, 1988).

A usina hidrelétrica de Coaracy Nunes está localizada no estado do Amapá, nas coordenadas  $0^{\circ} 54'$  N latitude e  $51^{\circ} 15'$  W longitude (Figura 2.1). A usina foi construída no meio do rio Araguari e pertence às Centrais Elétricas do Norte do Brasil S/A. Sua topografia é caracterizada pela transição de relevo de planície para áreas de serras. O clima quente e úmido apresenta uma estação seca no período de Setembro a Novembro. O nível de precipitação médio anual nesta área é de aproximadamente 2571,5 mm, com os maiores níveis observados de Fevereiro a Abril, apresentando médias de 1138,5 mm (45% do total anual), enquanto que em Setembro, Outubro e Novembro estes valores diminuem consideravelmente, apresentando médias de 45,5 mm.

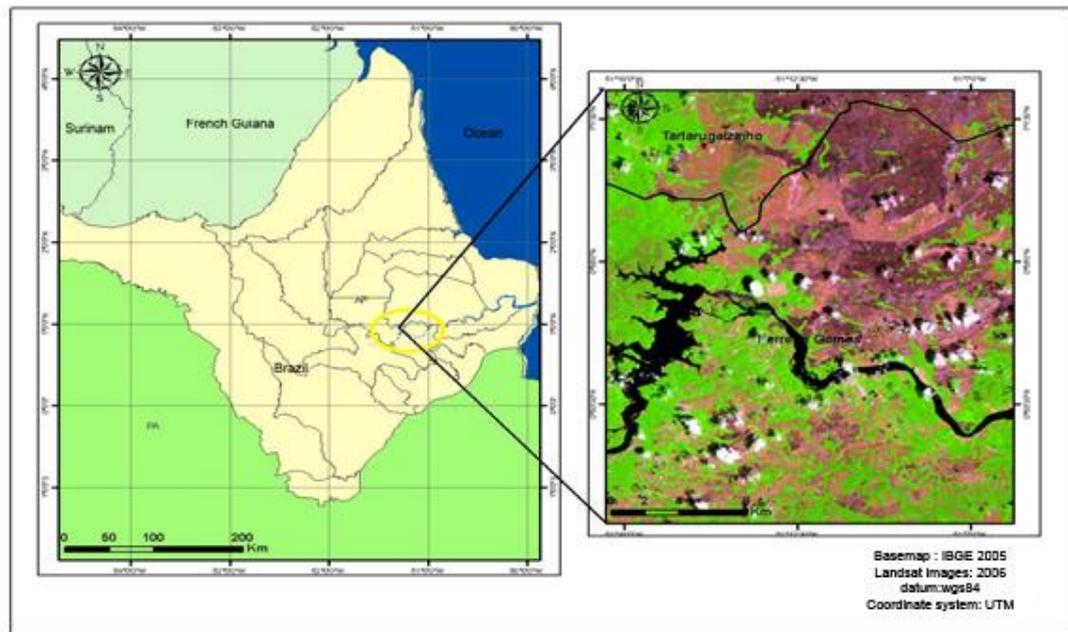


Figura 2.1: Localização da usina hidrelétrica de Coaracy Nunes

A usina hidrelétrica de Tucuruí está localizada no estado do Pará, nas coordenadas: latitude  $03^{\circ} 43' - 05^{\circ} 15'$ , longitude  $49^{\circ} 12' - 50^{\circ} 00' W$  (Figura 2.2). A usina foi construída no rio Tocantins, a cerca de 7 km da cidade de Tucuruí e a 300 km da cidade de Belém. O reservatório possui um total de área inundada de aproximadamente  $2.850 \text{ m}^2$ , com aproximadamente 50,8 milhões de  $\text{m}^3$  de água (ELETRONORTE, 1988). O período de Dezembro a Maio é considerado chuvoso na região e o período de Junho a Novembro é considerado seco, com precipitação anual entre 2250 e 2500 mm, sendo Março o mês mais chuvoso e Setembro o mais seco (FISCH, JANUÁRIO e SENNA, 1990).

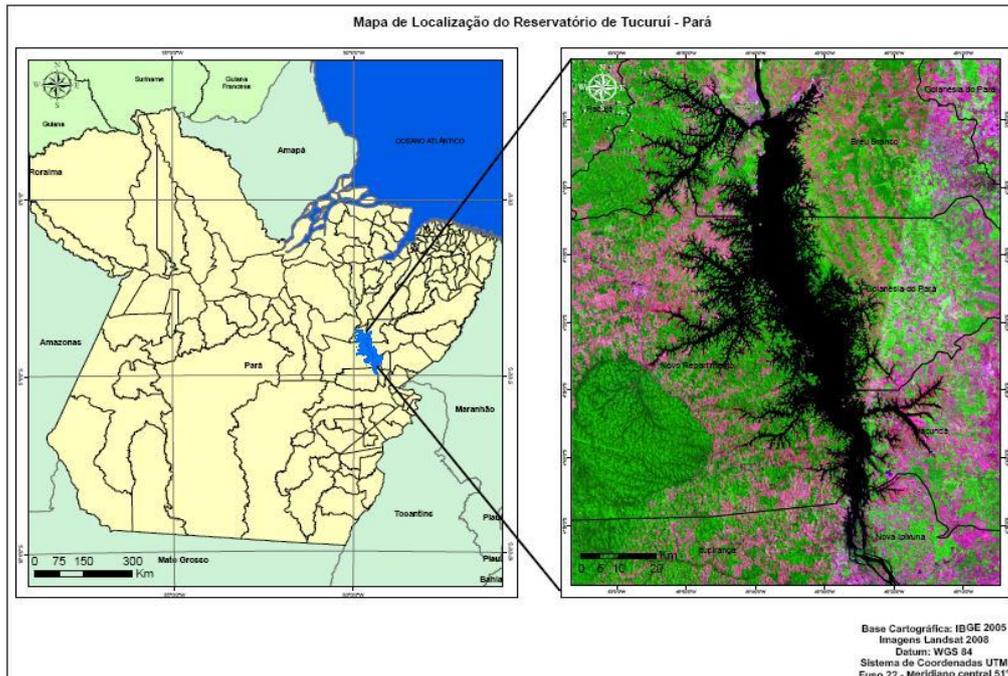


Figura 2.2: Localização da usina hidrelétrica de Tucuruí

## 2.2. Coleta de dados

Nos dois reservatórios foram realizadas quatro campanhas de coleta de amostras de água (Quadro 2.1) em pontos de coleta distribuídos nas montantes dos dois reservatórios. Estes foram escolhidos com o intuito de obter amostras representativas dos reservatórios como um todo (Figura 2.3 e Figura 2.4).

As amostras de água do reservatório de Coaracy Nunes foram coletadas entre outubro de 2008 e setembro de 2009 e as amostras de Tucuruí foram coletadas entre junho/julho 2008 e março/abril 2009.

Quadro 2.1: Campanhas de coleta de Coaracy Nunes e de Tucuruí com os respectivos períodos de coleta

Campanha	Coaracy Nunes		Tucuruí	
	Período	Clima	Período	Clima
1	23/10/2008 a 26/10/2008	Seco	28/05/2008 a 07/06/2008	Chuvoso
2	11/02/2009 a 16/02/2009	Chuvoso	04/09/2008 a 10/09/2008	Seco

3	05/05/2008 a 10/05/2009	Chuvoso	09/12/2008 a 18/12/2008	Seco
4	09/09/2009 a 16/09/2009	Seco	27/03/2009 a 06/04/2009	Chuvoso

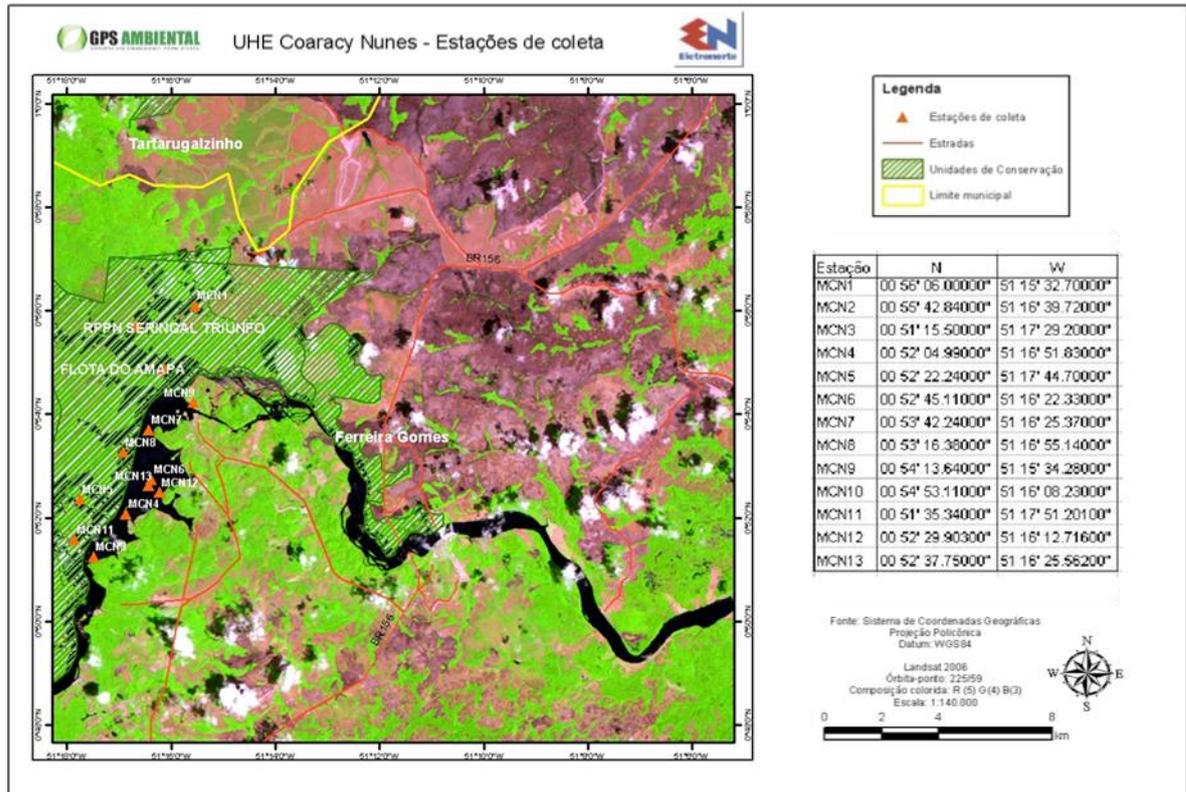


Figura 2.3: Pontos de coleta distribuídos pelo reservatório de Coaracy Nunes

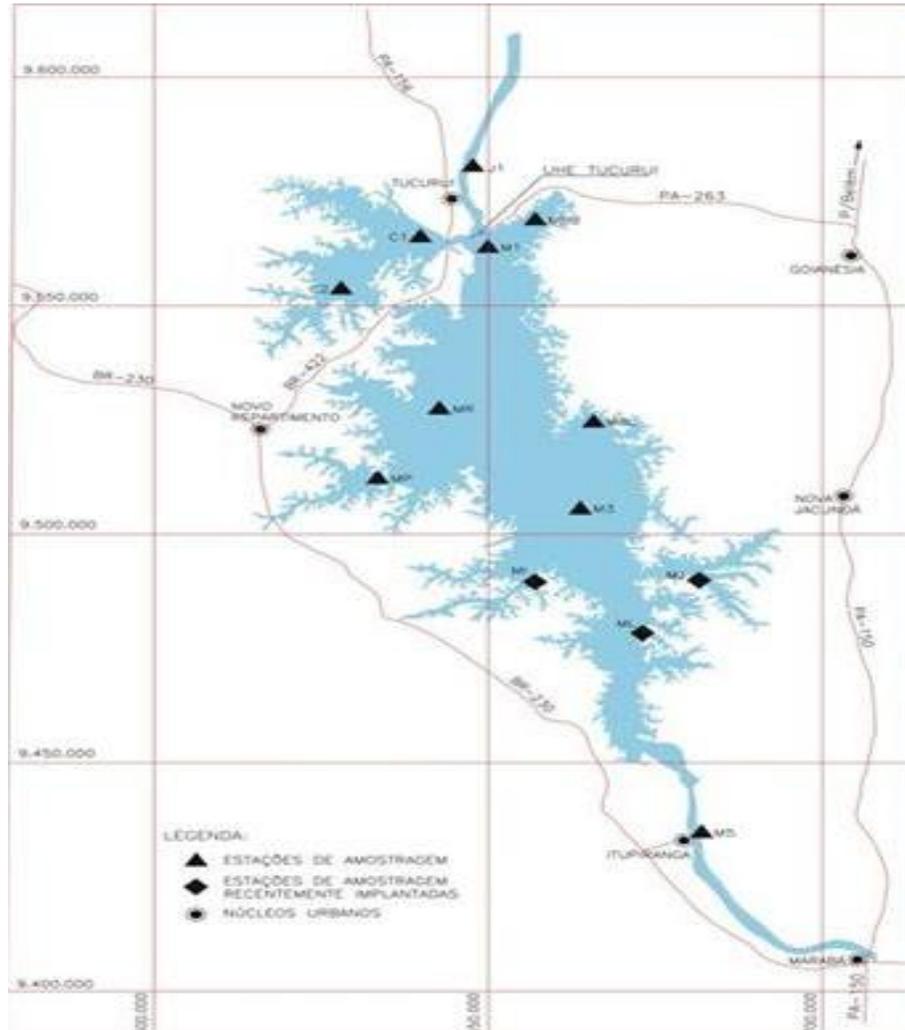


Figura 2.4: Pontos de coleta distribuídos pelo reservatório de Tucuruí

Foram escolhidos treze pontos de coleta de amostras em cada reservatório e cada ponto foi dividido em três profundidades, visando obter o perfil vertical da coluna de água: superfície, meio e fundo. As amostras foram coletadas com o auxílio de uma voadeira e armazenadas em frascos eppendorf de 50 ml para metais e acidificadas com ácido nítrico. A preservação das amostras foi realizada de acordo com USA EPA Standard Method 1060B (EPA, 1996).

### 2.3. Análise dos parâmetros físico-químicos e metais da água

Cada corpo de água doce tem um padrão individual de características físico-químicas que são largamente influenciadas pelo clima, geomorfologia, e condições geoquímicas que prevalecem na base de drenagem e lençol aquíferos. A qualidade físico-química da água ocorrerá normalmente em condições sem sinais de impactos antropogênicos (CHAPMAN, 1996). Na prática, essa condição é difícil de encontrar devido aos transportes atmosférico de contaminante e subseqüente depósito em locais distantes da origem. Oduwole (1997)

considera que diferenças significativas desses parâmetros entre os períodos seco e chuvoso podem ser indicadores de mudanças da qualidade da água.

Os parâmetros físico-químicos para análise da qualidade da água considerados para este estudo foram: **temperatura** da água, **condutividade**, **pH**, **turbidez**, **sólidos totais em suspensão**, **alcalinidade**, **amônia**, **demanda química de oxigênio**, **dureza** e **oxigênio dissolvido**, sendo o pH, condutividade elétrica e temperatura analisados *in situ*. As amostras foram coletadas utilizando a garrafa de Von Dorn de 2,5 L de capacidade em três profundidades. Estes foram coletados e analisados utilizando material de referência certificados, quando pertinentes, técnicas e equipamentos adequados e analistas treinados para tal.

O principal interesse do estudo da **temperatura** é seu papel na estruturação térmica dos corpos d'água já que influencia no metabolismo, na reprodução e comportamento migratório de organismos aquáticos (CHAPMAN, 1996; ATOBATELE e UGWUMBA, 2008; DAVIES, ABOLUDE e UGWUMBA, 2008; DEEKAE, ABOWEI e ALFRED-OCKIYA, 2010).

A **condutividade elétrica** é a medida da capacidade que a água possui de conduzir correntes elétricas. Este parâmetro proporciona uma boa indicação de mudanças na composição da água como partículas poluentes (ATOBATELE e UGWUMBA, 2008) e, depende da concentração iônica e da temperatura, podendo variar se houver alguma mudança nas concentrações de sólidos dissolvidos na água. (DEEKAE, ABOWEI e ALFRED-OCKIYA, 2010).

Segundo Deekae, Abowei e Alfred-Ockiya (2010) a concentração do íon hidrogênio na água é usualmente medida em termos de **pH**, que é definido como o logaritmo negativo da concentração do íon hidrogênio, sendo o valor sete neutro, valores acima de sete tendem para básico, enquanto valores abaixo de sete tendem para ácido. Valores de pH maiores que 7 e menores que 8,5 são ideais para produtividade biológica, enquanto valores de pH inferiores a 4 são prejudiciais a vida aquática. A maioria dos organismos não tolera grandes variações do pH durante o tempo e se tais condições persistirem podem ocorrer mortes. O pH de águas naturais é fortemente influenciado pela concentração do dióxido de carbono, que é considerado um gás ácido (ATOBATELE e UGWUMBA, 2008; DEEKAE, ABOWEI e ALFRED-OCKIYA, 2010).

A **turbidez** é a redução da transparência de fluidos devido à presença de materiais em suspensão. De acordo com Atobatele e Ugwumba (2008) a biomassa de fitoplâncton, a

disponibilidade de nutrientes e o aumento de sólidos em suspensão influenciam na transparência da água e conseqüentemente na turbidez. Os autores também observaram que a turbidez reduz a penetração da luz, diminuindo a produção primária e secundária.

Os **sólidos totais em suspensão** (STS) é a porção dos sólidos totais que fica retida em um filtro que propicia a retenção de partículas de diâmetro maior ou igual a 1,2  $\mu\text{m}$ . São caracterizados através de amostras líquidas (mg/L) ou sólidas (%), com a finalidade de se verificar a possibilidade de degradação aeróbia/anaeróbia dos sólidos em suspensão, tendo sido empregado também quando em sólidos totais sedimentados, para estimar o conteúdo orgânico do sedimento, em corpos d'água (CHAPMAN, 1996).

A **alcalinidade** ( $\text{HCO}_3^-$ ,  $\text{CO}_3^{2-}$  e  $\text{OH}^-$ ) dos corpos de água se refere à quantidade e tipos de íons dissolvidos (ânions) que alteram o pH para o lado alcalino. É causada por ou atribuída à presença de bicarbonatos, carbonatos, hidróxidos, boratos, silicatos e fosfatos. Estes íons são derivados de rochas dissolvidas, sais, solos, lixos industriais e atividades de plantas (DEEKAE, ABOWEI e ALFRED-OCKIYA, 2010).

A **amônia** ( $\text{NH}_3$ ) ocorre naturalmente em corpo de água onde há uma quebra de nitrogênio orgânico e matéria inorgânica no solo e água, excreção pela biota, redução do gás nitrogênio no corpo de água por microrganismo e da troca de gases com a atmosfera, sendo um indicador de descarga de processos industriais, esgotos fertilização e lixos comunitários (CHAPMAN, 1996). A amônia é formada pela decomposição da matéria orgânica, durante o processo de amonificação. Nesse processo, a amônia resulta tanto da decomposição aeróbica quanto anaeróbica da parte nitrogenada da matéria orgânica.

A **demanda química de oxigênio** (DQO) é um parâmetro relacionado à quantidade de oxigênio consumido por materiais e por substâncias orgânicas e minerais, que se oxidam sobre condições definidas. O parâmetro torna-se importante por estimar o potencial poluidor de efluentes domésticos e industriais, assim como por estimar um impacto dos mesmos sobre os ecossistemas aquáticos (CHAPMAN, 1996).

A **dureza** da água está associada com a concentração de íons de determinados minerais dissolvidos nela, principalmente cálcio e magnésio (MUSTAPHA, 2008). Águas com valores de dureza acima de 15 ppm são consideradas satisfatórias para o crescimento de peixe, enquanto valores abaixo de 5 ppm de carbonato de cálcio ( $\text{CaCO}_3$ ) provocam uma diminuição da população e eventual morte de peixes (ATOBATELE e UGWUMBA, 2008).

A quantidade de **oxigênio dissolvido** (OD) na água é importante para organismos aquáticos, pois afeta o crescimento, sobrevivência, distribuição, comportamento e fisiologia dos organismos aquáticos (DEEKAE, ABOWEI e ALFRED-OCKIYA, 2010). A medida de sua concentração é usada para indicar o grau de poluição por matéria orgânica e o nível de auto-purificação da água (ESTEVES, 1998).

O Quadro 2.2 apresenta os valores limites estabelecido pela legislação CONAMA 357/05 e Portaria do Ministério da Saúde 518/04, dos parâmetros físico-químicos analisados.

Quadro 2.2: Valores recomendados por diferentes instituições governamentais para os parâmetros físico-químicos

Padrão de Portabilidade			Padrão de Qualidade Ambiental
Parâmetro	Portaria nº 1469/00 - MS	Portaria nº 518/04 - MS	CONAMA nº 357/05 Classe II
<b>pH</b>	6,0 - 9,5	6,0 - 9,5	6,0 - 9,0
<b>Condutividade</b>	-	-	-
<b>Turbidez</b>	5 UT	5 UT	100 UT
<b>STS</b>	1.000 mg/L	1.000 mg/L	-
<b>Alcalinidade</b>	-	-	-
<b>Dureza</b>	500 mg/L	500 mg/L	500 mg/L
<b>DQO</b>	-	-	-
<b>OD</b>	-	-	Superior 5 mg/L

Os metais têm contribuído de forma significativa para a poluição do ar, da água e do solo, interferindo na manutenção da biota terrestre e aquática em decorrência principalmente de fontes naturais ou antropogênicas (PAPAFILIPPAKI, KOTTI e STAVROULAKIS, 2008; LEITE, DORNFELD e ESPÍNDOLA, 2013).

Apesar da presença de metais em corpos de água ser importante para o ciclo geoquímico e para saúde ambiental, em concentrações altas o acúmulo de metais nesses ecossistemas aquáticos tem despertado interesse sob vários aspectos, principalmente em relação ao destino e possíveis efeitos desses contaminantes, sua ciclagem biogeoquímica, comportamento e distribuição na cadeia alimentar (KABATA-PENDIAS, 2011; LEITE, DORNFELD e ESPÍNDOLA, 2013).

A técnica multielementar com espectrometria de emissão óptica em plasma por acoplamento indutivo (ICP-OES - Varian) foi utilizada para determinar a concentração dos metais. Os metais analisados foram: Alumínio (Al), Bário (Ba), Cálcio (Ca), Ferro (Fe), Potássio (K), Magnésio (Mg), Manganês (Mn) e Sódio (Na). Para o reservatório de Coaracy Nunes o Bário foi desconsiderado da análise por ter apresentado valores abaixo do limite de detecção. As condições operacionais do ICP-OES estão dispostas no Quadro 2.3. O padrão de referência utilizado para os testes de recuperação foi o SLRS-4, material de referência para traços de metais em águas de rios (National Research Council, Canada).

Quadro 2.3: Condições operacionais do ICP-OES

Potência da radiofrequência (KW)	1,4
Pressão do nebulizador (kPa)	180
Fluxo de argônio do plasma (L min <sup>-1</sup> )	15
Fluxo de argônio auxiliar (L min <sup>-1</sup> )	0,75
Introdução da amostra (s)	40
Tempo de lavagem (s)	30
Tempo de estabilização (s)	15
Tempo de leitura (s)	1
Velocidade da bomba peristáltica (rpm)	15

O Quadro 2.4 apresenta os valores limites estabelecido pela legislação CONAMA 357/05 e Portaria do Ministério da Saúde 518/04, dos metais analisados.

Quadro 2.4: Valores recomendados por diferentes instituições governamentais para os metais

Padrão de Portabilidade			Padrão de Qualidade Ambiental
Metal	Portaria nº 1469/00 - MS	Portaria nº 518/04 - MS	CONAMA nº 357/05 Classe II
Al	0,2 mg/L	0,2 mg/L	0,1 mg/L
Ca	-	-	-
Fe	0,3 mg/L	0,3 mg/L	0,3 mg/L
K	-	-	-
Mg	-	-	-
Mn	0,1 mg/L	0,1 mg/L	0,1 mg/L
Na	200 mg/L	200 mg/L	-

## 2.4. Métodos estatísticos e computacionais

### 2.4.1. Análise Multivariada

A análise multivariada se refere aos métodos estatísticos que analisam simultaneamente mais de duas variáveis. Algumas destas técnicas executam em uma única análise o que antes exigia várias análises univariadas, outras foram projetadas para lidar somente com questões multivariadas, como Análise Fatorial e Análise Discriminante (HAIR *et al.*, 1998).

Nesta pesquisa a Análise Fatorial (AF) foi aplicada para identificar as variáveis mais representativas, reduzindo o número de variáveis para utilizar os fatores resultantes decorrelacionados entre si na Análise Discriminante (AD). Esta foi utilizada para classificar as amostras de água em dois grupos que representam os períodos seco e chuvoso dos reservatórios, bem como avaliar o desempenho das RNAs propostas.

#### 2.4.1.1. *Análise Fatorial*

A Análise Fatorial (AF) é uma técnica estatística multivariada que objetiva definir a estrutura subjacente em uma matriz de dados. Pode ser aplicada para analisar as correlações entre um grande número de variáveis e explicá-las em termos de suas dimensões inerentes comuns (fatores). Esta técnica visa resumir e reduzir um determinado conjunto de dados. É indicada quando há a necessidade de identificar as variáveis mais representativas em um determinado conjunto de dados ou de representar melhor um número menor de variáveis. A análise fatorial pode auxiliar na seleção de um subconjunto representativo de variáveis ou na criação de novas variáveis como substitutas das variáveis originais (HAIR *et al.*, 1988).

A matriz de correlação dos dados de entrada pode ser calculada a partir da computação das correlações entre as variáveis ou entre os respondentes individuais. As variáveis para AF devem ser métricas e deve-se minimizar o número de variáveis, mas manter um número razoável de variáveis por fator (HAIR *et al.*, 1998)

Neste trabalho, os pré-requisitos para adequação da análise fatorial utilizados como critérios para escolher os fatores que melhor representam a estrutura dos dados, foram: o Kaiser-Meyer-Olkin Measure of Adequacy (KMO) maior que 0,50 e o teste de Bartlett. O KMO é uma medida que serve para quantificar o grau de correlação entre variáveis. Este índice varia de 0 a 1, alcançando 1 quando cada variável é perfeitamente prevista sem erro pelas outras variáveis, sendo inaceitáveis valores a baixo de 0,50. O teste de Bartlett de esfericidade é um método que examina a matriz de correlação inteira, fornecendo a probabilidade estatística de que a estas correlações significantes entre pelo menos algumas das variáveis (JOHNSON e WICHERN, 2001; HAIR *et al.*, 1998).

A Análise de Componentes Principais (ACP) é um método de extração de fatores apropriado quando a preocupação principal é a previsão ou o número mínimo de fatores necessários para explicar a parte máxima da variância representada no conjunto original de variáveis. O método primeiro extrai as combinações de variáveis que explicam o maior montante de variância e então segue para combinações que explicam montantes cada vez menores de variância. A quantidade de fatores a extrair pode ser escolhida com base em critérios pré-determinados, normalmente o percentual de variância ou o critério da raiz latente. No critério de percentagem de variância os fatores são escolhidos de forma que juntos expliquem pelo menos 95% da variância total. No critério da raiz latente, um dos critérios mais utilizados, que também foi utilizado neste trabalho, qualquer fator individual deve

explicar a variância de pelo menos uma variável. Cada variável contribui com um valor 1 do auto valor total. Logo, apenas os fatores que tenham raízes latentes ou autovalores maiores que 1 são considerados significantes, os demais fatores são descartados (JOHNSON e WICHERN, 2001; HAIR *et al.*, 1998).

A matriz fatorial não-rotacionada é computada para auxiliar na obtenção de uma indicação preliminar do número de fatores a extrair. Soluções com fatores não-rotacionados atingem o objetivo de redução de dados, porém, podem não fornecer a interpretação mais adequada das variáveis. Em geral, a rotação é desejável porque simplifica a estrutura fatorial. A rotação fatorial pode ser ortogonal, na qual os eixos são mantidos a 90 graus, ou oblíqua, quando não há restrição de os eixos serem ortogonais. A rotação fatorial ortogonal é a mais comum e entre as principais abordagens ortogonais destacam-se a *quadrimax* (simplifica as linhas da matriz fatorial), *varimax* (simplifica as colunas da matriz fatorial) utilizada neste trabalho e *equimax* (simplifica as linhas e colunas da matriz fatorial, porém não apresenta bons resultados) (JOHNSON e WICHERN, 2001; HAIR *et al.*, 1998).

As cargas fatoriais e as comunalidades (valor que representa a quantia de variância explicada pela solução fatorial para cada variável) devem ser analisadas. Caso não apresentem valores satisfatórios é preciso remover alguma variável e executar a análise de componentes ou de fatores comum novamente.

#### 2.4.1.2. *Análise Discriminante*

A Análise Discriminante (AD) é uma técnica multivariada que permite discriminar uma variável dependente categórica a partir de variáveis independentes métricas (HAIR *et al.*, 1998). Neste trabalho as categorias ou grupos representaram os períodos seco ( $Y=0$ ) e chuvoso ( $Y=1$ ) e os escores descorrelacionados (independentes) da AF entraram com as variáveis independentes (JOHNSON e WICHERN, 2001; HAIR *et al.*, 1998).

A AD pode ser linear (Fisher) ou não linear (quadrática, cúbica, dentre outras), sendo a linear adequada quando não existem diferenças significativas entre as matizes de covariância dos grupos e a não linear adequada quando estas diferenças são significativas (JOHNSON e WICHERN, 2001; HAIR *et al.*, 1998).

A discriminação é obtida ao estabelecer pesos às variáveis para maximizar a variância entre os grupos (função discriminante), resultando em um escore Z discriminante. A média dos escores de cada grupo é chamada de centroide e a variável dependente é discriminada de

acordo com a distância em relação aos centróides de cada grupo (JOHNSON e WICHERN, 2001; HAIR *et al.*, 1998).

Durante a AD as amostras devem ser divididas para desenvolver a função discriminante (amostra de análise) e para testar a função discriminante (amostra de teste). Depois que a função discriminante for computada é preciso avaliar o seu nível de significância, sendo o  $\lambda$  de Wilks um dos critérios mais comuns. Esta estatística determina o critério de otimalidade (relacionado com a probabilidade de erro de classificação), o qual avalia a separação entre os grupos para cada variável a partir da razão entre a dispersão das médias dos grupos e a variância total (JOHNSON e WICHERN, 2001; HAIR *et al.*, 1998).

#### 2.4.2. Redes Neurais Artificiais

Redes Neurais Artificiais (RNAs) é uma técnica de inteligência computacional que pode ser definida como um modelo de raciocínio baseado no cérebro humano (NEGNEVITSKY, 2005). As RNAs consistem em um conjunto de unidades básicas de processamento (neurônios) ligadas entre si por links que transmitem sinais de um neurônio para outro. Estes links possuem um peso numérico que representa a importância de cada entrada do neurônio. Cada neurônio recebe uma quantidade de sinais de entrada e sempre produz um único sinal de saída que pode ser transmitido para outros neurônios. (NEGNEVITSKY, 2005; RUSSELL e NORVIG, 2010). Um exemplo de RNA pode ser visto na Figura 2.5.

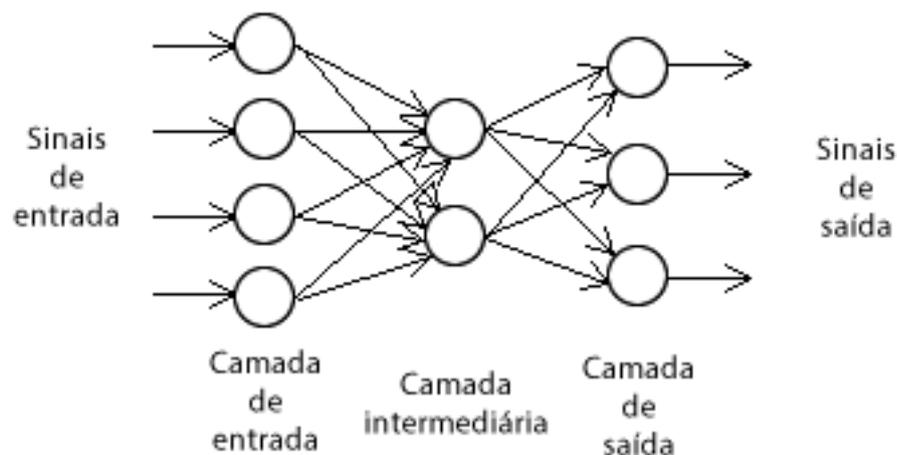


Figura 2.5: Arquitetura típica de uma RNA

A saída de cada neurônio é computada a partir de uma função de ativação. Existem diversas funções de ativação, porém apenas poucas apresentaram aplicação prática (NEGNEVITSKY, 2005). Dentre as escolhas mais comuns destacam-se as funções degrau,

sinal, linear, sigmoide e tangente hiperbólica (Figura 2.6) (GOMES e LUDERMIR, 2008; NEGNEVITSKY, 2005).

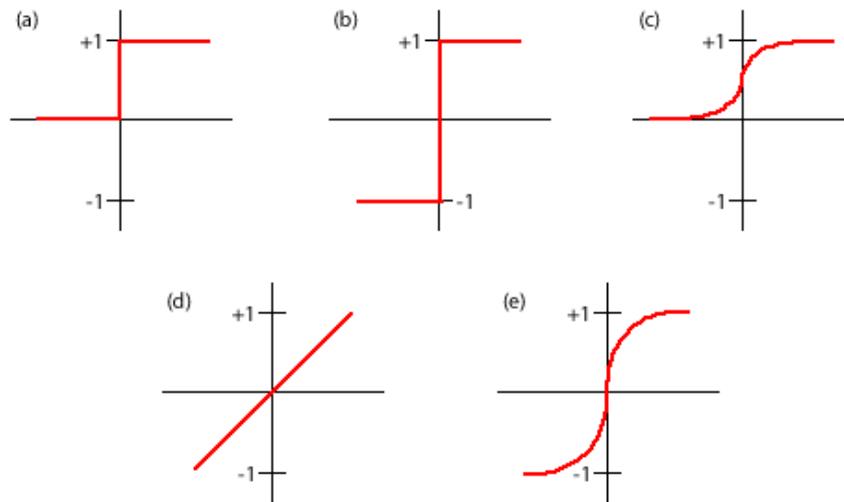


Figura 2.6: (a) função degrau; (b) função sinal; (c) função sigmoide; (d) função linear; (e) função tangente hiperbólica

O aprendizado das RNAs ocorre a partir de repetitivos ajustes nos pesos de todos os links da rede. O responsável por estes ajustes é o algoritmo de treinamento utilizado e sua escolha é baseada na arquitetura da RNA. A RNA mais simples é a perceptron, uma rede composta de um único neurônio, a qual é capaz de aprender funções linearmente separáveis. Problemas mais complexos podem ser resolvidos com a utilização da arquitetura perceptron multicamada (Figura 2.7). Este modelo de RNA consiste em vários neurônios interligados entre si com uma ou mais camadas escondidas (NEGNEVITSKY, 2005).

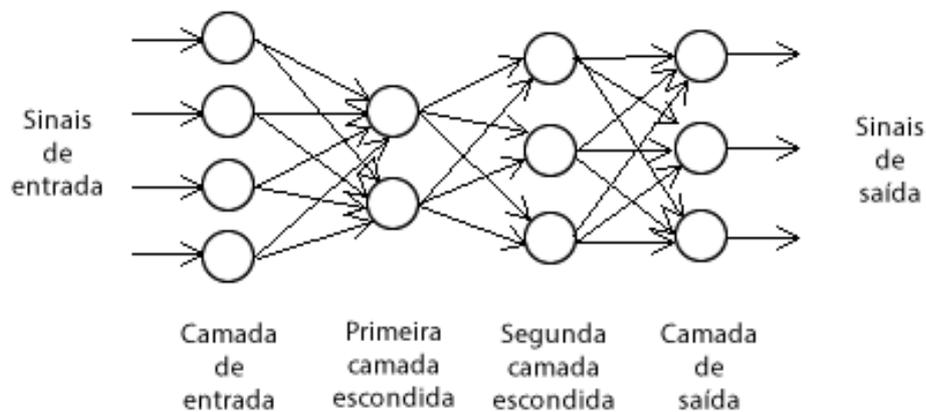


Figura 2.7: Perceptron multicamada

Existem diversos algoritmos que ajustam os pesos dos links (pesos sinápticos) da RNA chamados algoritmos de treinamento. O algoritmo Backpropagation e suas derivações é um dos algoritmos para treinamento de RNAs multicamadas, mais difundido na comunidade científica. Baseia-se no aprendizado supervisionado por correção de erros. Sua utilização compreende duas fases de propagação:

**1° - Propagação:** Um padrão é apresentado à camada da entrada da rede. A atividade resultante flui através da rede, camada por camada, até que a resposta seja produzida pela camada de saída, onde é obtido a resposta da rede e o erro é calculado, como visto na Figura 2.8;

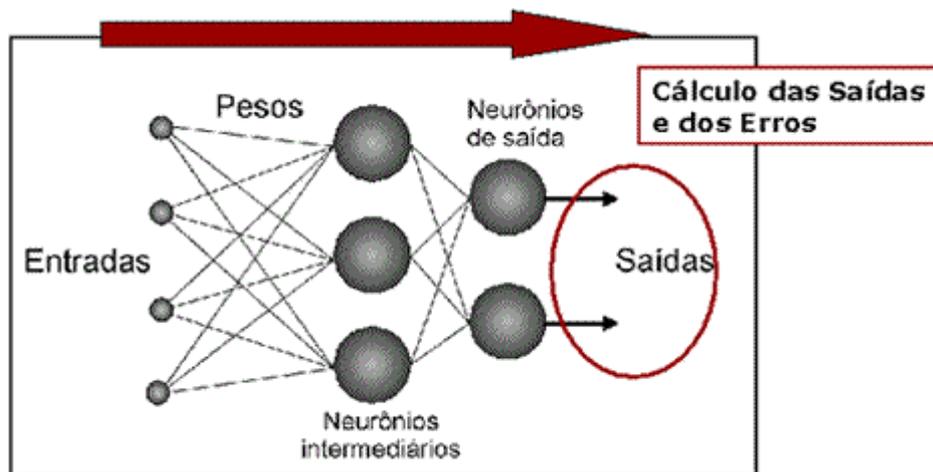


Figura 2.8: Fase de propagação (LNCC, 2012)

**2° - Retropropagação ("backpropagation"):** Desde a camada de saída até a camada de entrada, são feitas alterações nos pesos sinápticos, como visto na Figura 2.9.

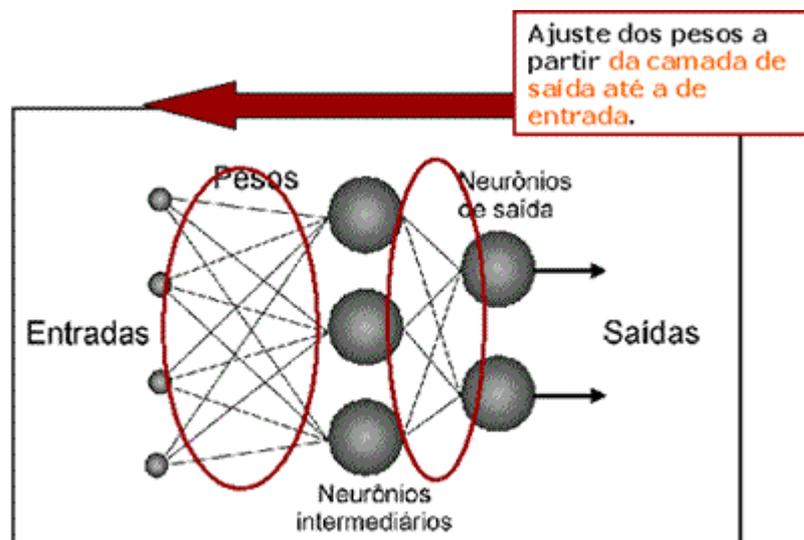


Figura 2.9: Fase de retropropagação (LNCC, 2012)

Depois que a rede estiver treinada e o erro estiver em um nível satisfatório, ela poderá ser utilizada como uma ferramenta para classificação de novos dados. Para isto, a rede deverá ser utilizada apenas no modo progressivo (feed-forward). Ou seja, novas entradas são apresentadas à camada de entrada, são processadas nas camadas intermediárias e os resultados são apresentados na camada de saída, como no treinamento, mas sem a retropropagação do erro. A saída apresentada é o modelo dos dados, na interpretação da RNA.

### 3. APLICAÇÃO DA METODOLOGIA E RESULTADOS

As análises estatísticas foram conduzidas com auxílio dos softwares Statistica® e SPSS. As redes neurais artificiais foram criadas com auxílio dos softwares Matlab e SPSS.

#### 3.1. Reservatório Coaracy Nunes

A Análise Fatorial por Componentes Principais com rotação ortogonal dos eixos via varimax foi utilizada para reduzir o número de variáveis e para obtenção de fatores independentes constituídos pelos parâmetros físico-químicos e pelos metais analisados da água do reservatório de Coaracy Nunes. Os pré-requisitos para adequação da Análise Fatorial foram utilizados como critérios para escolher os fatores que melhor representam a estrutura dos dados. Estes critérios foram: autovalores maiores que 1, comunalidades e KMO maiores que 0,50 e teste de Bartlett significativo.

Os escores resultantes da análise fatorial foram utilizados como variáveis de entrada para análise discriminante, com o intuito de classificar as amostras em dois grupos que correspondem aos períodos de coleta das amostras de água do reservatório (seco ou chuvoso). De um total de 140 amostras, foram utilizadas 98 para obtenção da função discriminante (amostra de análise) e 42 para testar a função discriminante (amostra de teste).

O resultado do teste Teste M de Box foi não significativo ( $p=0,329$ ), indicando equivalência na matriz de covariância dos dois grupos (adequação da análise discriminante linear) e o valor  $\lambda$  de Wilks resultou em 0,121 com teste significativo ( $p=0,000$ ).

A quantidade de variáveis constituintes dos fatores resultantes da análise fatorial e o valor do  $\lambda$  de Wilks próximo a 0 indicam que a RNA necessária para realizar a mesma classificação possuirá uma estrutura simples e que não será difícil obter boas classificações.

Foi criada uma Rede Neural Artificial, com auxílio do software Matlab6p1, para classificar as amostras de água do reservatório nos dois grupos (período de coleta seco ou chuvoso), com a mesma repartição da amostra sendo 98 para treinamento e 42 para validação

da rede. A RNA proposta (Figura 3.1) utiliza as variáveis Ca, K, Mg, Na e DQO, constituintes dos fatores resultantes da análise fatorial, como variáveis de entrada e os grupos como variável de saída. A RNA foi treinada com os seguintes parâmetros:

- RNA: Perceptron multicamada com uma camada escondida com cinco neurônios.
- Algoritmo de treinamento: Levenberg-Marquardt backpropagation.
- Taxa de treinamento: 0,001.
- Função de transferência: *tansig* em todos os neurônios da rede.
- Erro de aproximação da rede:  $10^{-14}$ .

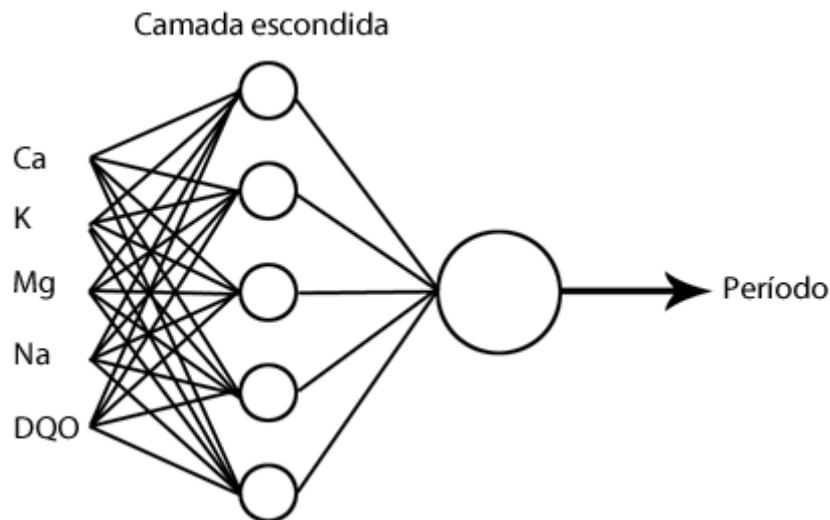


Figura 3.1: RNA proposta para classificação das amostras de água do reservatório de Coaray Nunes

A Tabela 3.1 apresenta o intervalo de confiança de 95% da média dos parâmetros físico-químicos medidos no reservatório de Coaracy Nunes durante as quatro campanhas de coleta e para os períodos seco e chuvoso.

Tabela 3.1: Resultado dos parâmetros físico-químicos medidos em Coaracy Nunes

Parâmetro	Campanha de coleta				IC95% (período seco)	IC95% (período chuvoso)
	1 (seco)	2 (chuvoso)	3 (chuvoso)	4 (seco)		

Temperatura (°C)	29,58±0,17	26,81±0,24	27,86±0,30	29,61±0,30	29,59±0,13	27,21±0,23
Condutividade ( $\mu\text{s cm}^{-1}$ )	45,55±0,69	20,34±0,30	16,65±0,18	20,05±0,37	33,34±3,06	18,55±0,47
pH	6,79±0,06	6,45±0,09	6,61±0,04	6,20±0,13	6,51±0,10	6,53±0,05
STS	1,54±0,55	7,06±2,36	2,72±1,75	2,87±0,86	2,29±0,56	4,14±1,48
Alcalinidade (mg $\text{CaCO}_3 \text{ L}^{-1}$ )	34,59±13,75	6,47±0,39	8,62±0,63	10,16±0,48	22,37±7,33	7,51±0,44
Amônia ( $\mu\text{g L}^{-1}$ )	31,38±8,00	8,81±2,85	22,13±17,69	80,10±188,32	35,12±10,20	13,42±6,26
DQO ( $\text{mg O}_2 \text{ L}^{-1}$ )	10,30±0,53	15,90±0,36	10,19±0,39	6,10±1,00	8,29±0,73	13,05±0,74
Dureza ( $\text{mg CaCO}_3$ )	4,92±0,18	6,11±0,26	6,18±0,36	6,32±0,20	5,59±0,21	6,14±0,21
OD ( $\text{mg L}^{-1}$ )	6,33±0,26	5,31±0,27	5,83±0,29	6,10±0,38	6,20±0,23	5,57±0,20

O valores do pH da água não são estatisticamente diferentes entre os períodos seco e chuvoso, enquanto a temperatura, condutividade, alcalinidade, amônia e oxigênio dissolvido apresentaram valores elevados durante o período seco. Já sólido total em suspensão, dureza e demanda química de oxigênio apresentaram valores elevados durante o período chuvoso.

Os valores da alcalinidade são reflexos dos perfis de carbonatos e bicarbonatos (WETZEL, 2001), com alguma contribuição de silicato e fosfato, e ajudam a estabilizar os valores de pH do reservatório (DEEKAE *et al.*, 2010). Altos valores de alcalinidade durante o período seco foram observados em outros reservatórios (ADEBISI, 1981; HART e ZABBEY, 2005; DAVIES *et al.*, 2008; DEEKAE *et al.*, 2010,) sendo atribuídos a grande evaporação durante este período, concentrando íons e dióxido de carbono na água e aumentando a liberação de íon bicarbonato pelo sedimento (DEEKAE *et al.*, 2010; MUSTAPHA, 2008). Essa grande evaporação pode ser decorrente do aumento da temperatura e pode influenciar o aumento da condutividade e amônia durante o período seco (Tabela 3.1).

Aumentos na dureza da água foram observados no presente estudo durante o período chuvoso, assim como em outros reservatórios tropicais (ATOBATELE e UGWUMBA, 2008). Os níveis de DQO, por outro lado, foram muito altos durante os dois períodos, variando de  $8,29\pm 0,73 \text{ mgO}_2\text{L}^{-1}$  no período seco a  $13,05\pm 0,74 \text{ mgO}_2\text{L}^{-1}$  no período chuvoso.

Em climas tropicais, a água possui pouca habilidade de manter oxigênio em altas temperaturas quando comparada a baixas (LEWIS, 2000). Assim, costuma-se observar a

relação inversa entre temperatura e oxigênio dissolvido, como no caso do reservatório na Turquia (TEPE e MUTLU, 2005). Este não é o caso no presente estudo, pois altos valores de oxigênio dissolvido foram observados durante o período seco, quando temperatura um pouco mais alta da água foi observada. Alguns estudos também relatam esta exceção, como no caso de Oben (2000) que realizou um levantamento limnológico do impacto de efluentes agrícolas e domésticos em 3 lagos artificiais em Ibadan, Nigéria.

Fontes de oxigênio dissolvido no ambiente aquático incluem a atmosfera e fotossíntese, enquanto a perda de oxigênio inclui respiração, decaimento de bactérias aeróbias e decomposição de sedimento (GUPTA e GUPTA, 2006). O decréscimo nos valores do oxigênio dissolvido no período chuvoso poderia como postulado por outros autores, ser decorrente de *blooms* de fitoplâncton e da decomposição de materiais orgânicos que normalmente ocorre durante este período (ATOBATELE e UGWUMBA, 2008; OKAYI, 2003).

A Tabela 3.2 apresenta o intervalo de confiança de 95% da média das concentrações dos metais medidas no reservatório de Coaracy Nunes durante as quatro campanhas de coleta para os períodos seco e chuvoso.

Tabela 3.2: Concentrações dos metais medidas no reservatório de Coaracy Nunes

Metal (mg L <sup>-1</sup> )	Campanha de coleta				IC95% da Média (período seco)	IC95% da Média (período chuvoso)
	1 (seco)	2 (chuvoso)	3 (chuvoso)	4 (seco)		
Al	0,07±0,01	0,33±0,05	0,17±0,09	0,09±0,01	0,08±0,01	0,25±0,06
Ca	0,83±0,02	0,70±0,03	0,43±0,01	0,63±0,01	0,74±0,03	0,57±0,04
Fe	0,32±0,02	0,43±0,02	0,34±0,06	0,33±0,08	0,33±0,04	0,39±0,03
K	1,03±0,02	0,71±0,03	0,48±0,01	0,73±0,02	0,89±0,04	0,60±0,03
Mg	0,57±0,00	0,38±0,02	0,32±0,00	0,48±0,01	0,52±0,01	0,35±0,01
Mn	0,09±0,02	0,06±0,01	0,05±0,02	0,10±0,04	0,10±0,02	0,06±0,01
Na	2,04±0,02	1,25±0,08	0,99±0,01	1,51±0,18	1,79±0,11	1,12±0,05

Pode-se observar na Tabela 3.2 que os metais, com exceção do Fe e Al, apresentaram maiores concentrações durante o período seco. Isto pode estar diretamente relacionado com a diminuição do volume de água do reservatório devido à evaporação durante o período e conseqüentemente ao aumento da concentração dos metais na coluna de água. Sabe-se que os padrões de distribuição vertical da maioria dos cátodos e ânodos estão associados com a profundidade da massa de água, portanto, em áreas rasas (profundidade máxima de 20m), como no caso do reservatório de Coaracy Nunes, sem estratificação térmica duradoura, estes elementos nas suas formas dissolvidas estão distribuídos uniformemente ao longo da coluna de água.

Uma vez que os fenômenos de estratificação térmica e química (relacionados ao oxigênio dissolvido) não influenciam na distribuição do perfil vertical dos metais, a profundidade não foi considerada um fator.

Os parâmetros físico-químicos analisados apresentaram valores médios abaixo dos limites estabelecidos pelas portarias do ministério da saúde e pelo CONAMA. O alumínio (Al) apresentou valores médios acima do limite estabelecido pelas portarias do ministério da saúde (0,2 mg/L) durante a segunda campanha e valores médios acima do limite estabelecido pelo CONAMA (0,1 mg/L) durante a segunda e terceira campanhas. O Ferro (Fe) apresentou valores médios acima do limite estabelecido pelas portarias do ministério da saúde (0,3 mg/L) e pelo CONAMA (0,3 mg/L) durante as quatro campanhas. O Manganês (Mn) apresentou valores médios acima do limite estabelecido pelas portarias do ministério da saúde (0,1 mg/L) e pelo CONAMA (0,1 mg/L) durante a quarta campanha.

Pode-se observar na Figura 3.2 e na Figura 3.3 os box-plots dos metais e do DQO nas três profundidades utilizadas na análise.

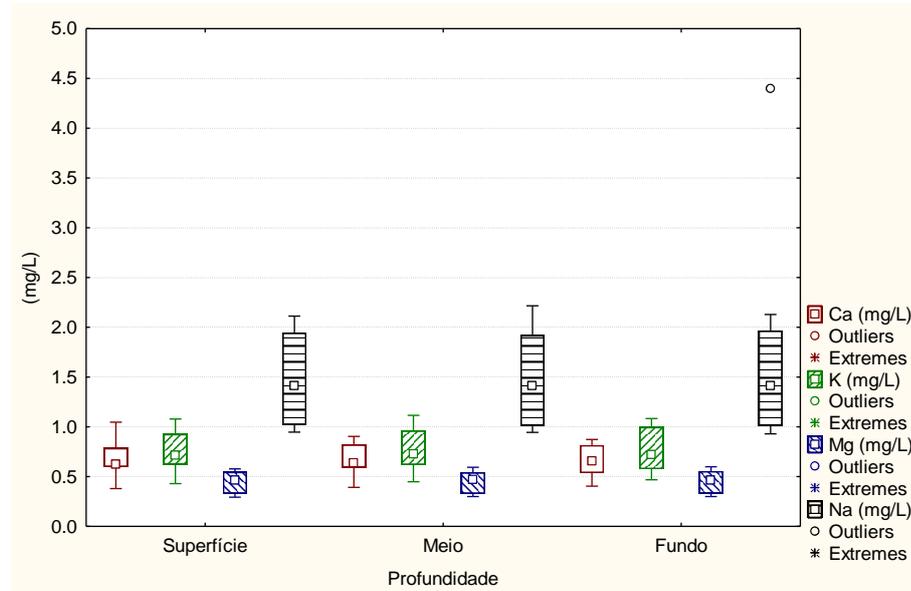


Figura 3.2: Box-plots do Ca, K, Mg e Na analisados das amostras de água da superfície, meio e fundo

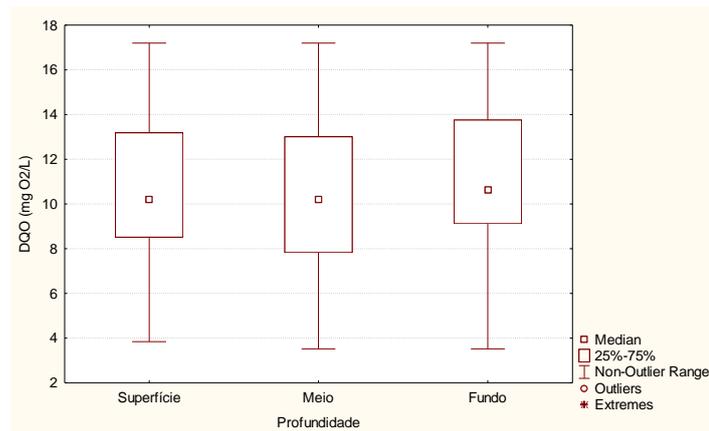


Figura 3.3: Box-plot do DQO na superfície, meio e fundo

Depois da análise descritiva dos dados foi feito o estudo das correlações entre os metais e os parâmetros físico-químicos, seguida pela Análise Fatorial. As variáveis Ca, K, Mg, Na e Demanda química de oxigênio (DQO) apresentaram comunalidades apropriadas para o uso da técnica, resultando em dois fatores representando 94,30% da variância total, todos com autovalores maiores que 1 (Tabela 3.3). O KMO foi 0,70, maior que 0,50 para toda a matriz de correlação e o teste de Bartlett apresentou um valor significativo.

Tabela 3.3: Fatores, variância total e comunalidades obtidos pela análise de Componente Principal e Análise Fatorial de Coaracy Nunes

Variável	Comunalidade	Fator
----------	--------------	-------

		1	2
Ca ( mgL <sup>-1</sup> )	0,93	<b>0,94</b>	-0,20
K ( mgL <sup>-1</sup> )	0,99	<b>0,99</b>	-0,01
Mg ( mgL <sup>-1</sup> )	0,96	<b>0,94</b>	-0,29
Na ( mgL <sup>-1</sup> )	0,85	<b>0,91</b>	-0,11
DQO(mg O <sub>2</sub> /L)	0,99	-0,03	<b>-0,99</b>
Autovalor		3,59	1,13
% Variância		71,76	22,57
Cumulativo %		71,76	94,33

O fator 1 é composto pelas variáveis: Ca, K, Mg e Na, que juntas representam 71,76% da variância total dos dados. Como esperado, as variáveis que constituem este fator apresentaram uma correlação significativa. O fator 2 é composto pela demanda química de oxigênio (DQO), que representa 22,57% da variância total dos dados. O intervalo de confiança de 95% da média do DQO foi respectivamente [7,55, 9,20] e [12,41, 13,88] durante o período seco e chuvoso respectivamente. Assim, este parâmetro foi confirmado como uma importante variável anual para este reservatório.

Os elementos Ca, K, Mg e Na são os principais envolvidos na regulação da dureza, alcalinidade e pH da água, pois eles afetam diretamente estes parâmetros dependendo de suas concentrações. Por causa disto, é lógico interpretar o fator resultante da análise multivariada e da rede neural artificial como uma conexão direta entre estes elementos e o DQO, devido outra conexão direta entre estes elementos e os parâmetros compartilhados: dureza, alcalinidade e pH da água, que também estão conectados diretamente ao DQO.

Altos valores de DQO foram associados à poluição (TEPE e MUTLU, 2005). É recomendável que os valores de DQO sejam inferiores a 2 mgL<sup>-1</sup> para água potável (APHA, 1995). Uma possível explicação para estes valores elevados é o alto índice de decomposição orgânica resultante de atividades antrópicas próximas ao reservatório, as quais incluem fazendas e terras utilizadas para criação de búfalo e cultivo de eucalipto, que produzem esgoto

doméstico e escoamentos agrícolas. Como altos níveis de Ca, K, Mg e Na podem, em muitos casos, serem resultantes de esgoto doméstico e escoamentos agrícolas, postula-se que eles podem ter sido introduzidos no sistema, apresentando uma influência direta nos níveis de DQO, como foi confirmado pela análise multivariada e pela rede neural artificial.

Com a Análise Discriminante foi obtido 96,40% e 95,70% de classificações corretas para a função discriminante e amostras de teste respectivamente, enquanto que, utilizando as variáveis correlacionadas Ca, K, Mg, Na e DQO como variáveis de entrada da RNA foi possível classificar corretamente 100% das amostras tanto no treinamento quanto na validação da rede, portanto, ambas as metodologias são adequadas para análise dos dados apresentados neste estudo.

### **3.2. Reservatório de Tucuruí**

Da mesma forma que o reservatório de Coroacy Nunes, foi conduzido a Análise Fatorial por Componentes Principais com rotação ortogonal dos eixos via varimax, utilizando os mesmos critérios para escolha dos fatores que melhor representam a estrutura dos dados.

Os escores resultantes da AF foram utilizados como variáveis de entrada para Análise Discriminante, com o intuito de classificação quanto ao período de coleta das amostras de água do reservatório (seco ou chuvoso). De um total de 164 amostras, foram utilizadas 117 para obtenção da função discriminante (amostra de análise) e 32 para testar a função discriminante (amostra de teste).

O resultado do teste M de Box foi significativo ( $p=0,70$ ) indicando a adequação da AD não linear (não equivalência na matriz de covariância dos dois grupos), e o valor  $\lambda$  de Wilks resultou em 0,625 com teste significativo ( $p=0,000$ ).

A quantidade de variáveis constituintes dos fatores resultantes da análise fatorial e o valor do  $\lambda$  de Wilks afastado de 0 indicam que a RNA necessária para realizar a mesma classificação possuirá uma estrutura mais complexa e que será difícil obter boas classificações.

Foi criada uma Rede Neural Artificial, com auxílio do software SPSS, para classificar as amostras de água do reservatório quanto ao período de coleta (seco ou chuvoso). As mesmas amostras utilizadas na análise da função discriminante e de teste na AD foram utilizadas respectivamente para treinamento e validação da rede. A RNA proposta (Figura 3.4) utiliza as variáveis Ba, Ca, Fe, K, Mg, Na, profundidade, condutividade, STS, turbidez, NH<sub>3</sub>, OD e

DQO, constituintes dos fatores resultantes da análise fatorial, como variáveis de entrada e o período seco ou chuvoso como variável de saída. A RNA foi treinada com os seguintes parâmetros:

- RNA: Perceptron multicamada com uma camada escondida com seis neurônios e dois neurônios na camada de saída.
- Algoritmo de treinamento: Scaled Conjugate Gradient.
- Lambda inicial: 0,0000005.
- Sigma inicial: 0,00005.
- Função de transferência: *tansig* na camada intermediária e *logsig* na camada de saída.
- Erro de aproximação da rede:  $10^{-3}$ .

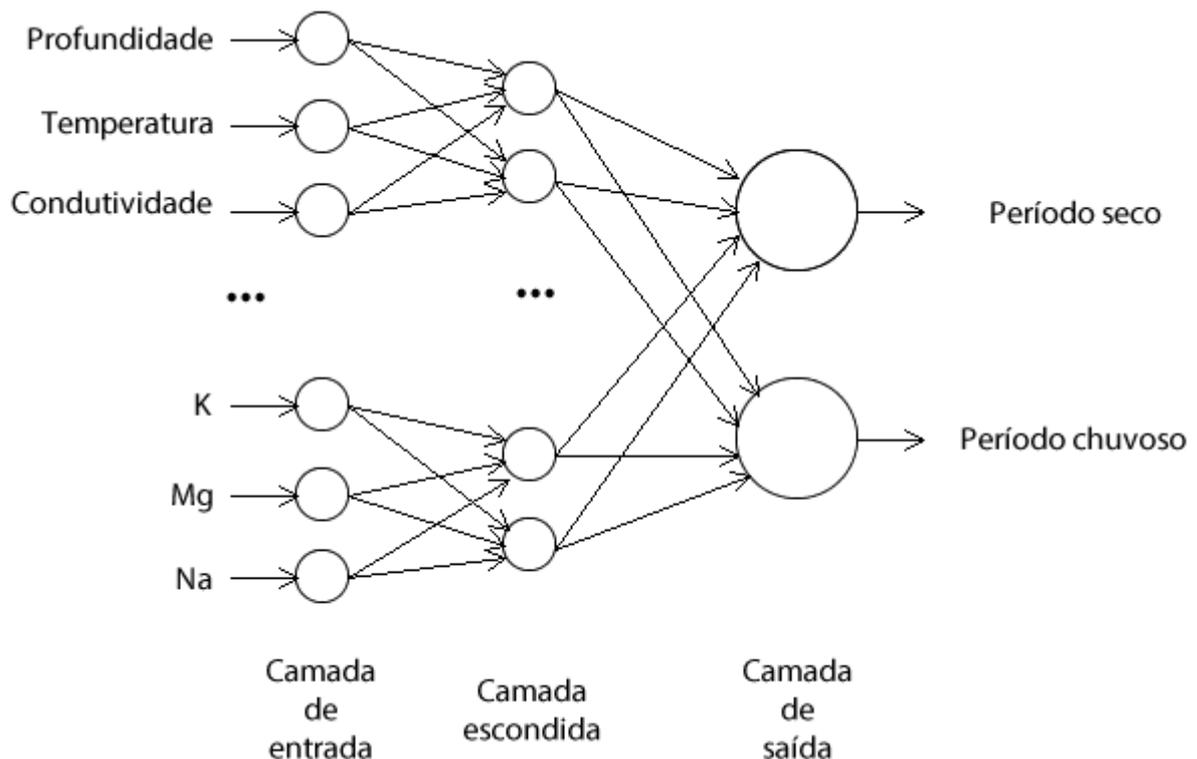


Figura 3.4: RNA proposta para classificação das amostras de água do reservatório de Tucuruí

A Tabela 3.4 apresenta o intervalo de confiança de 95% dos parâmetros físico-químicos medidos no reservatório de Tucuruí durante as quatro campanhas de coleta para os períodos seco e chuvoso.

Tabela 3.4: Resultado dos parâmetros físico-químicos medidos em Tucuruí

Parâmetro	Campanha				IC 95% (período seco)	IC 95% (período chuvoso)
	1 (chuvoso)	2 (seco)	3 (seco)	4 (chuvoso)		
Profundidade (m)	10,11±2,82	9,00±2,38	7,50±2,21	8,89±2,58	8,24±1,61	9,50±1,89
Temperatura (°C)	29,25±0,20	29,98±0,13	30,10±0,17	29,37±0,12	30,04±0,10	29,31±0,12
Condutividade ( $\mu\text{s cm}^{-1}$ )	56,17±1,75	56,60±3,30	62,83±2,36	54,72±0,98	59,76±2,06	55,44±0,99
pH	7,51±0,08	7,03±0,07	7,36±0,05	7,38±0,08	7,19±0,05	7,44±0,03
Turbidez (NTU)	6,51±1,35	3,28±0,97	9,13±2,17	9,28±0,99	6,54±1,44	7,92±0,86
DQO ( $\text{mg O}_2 \text{ L}^{-1}$ )	6,86±0,44	6,15±0,40	8,08±0,88	9,13±0,51	7,13±0,51	8,15±0,39
OD ( $\text{mg L}^{-1}$ )	4,89±0,41	4,94±0,41	5,37±0,31	4,91±0,36	5,16±0,25	4,90±0,27

A Tabela 3.5 apresenta o intervalo de confiança de 95% da média das concentrações dos metais medidas no reservatório de Tucuruí durante as quatro campanhas de coleta para os períodos seco e chuvoso.

Tabela 3.5: Concentrações dos metais em  $\text{mg/L}^{-1}$  medidas no reservatório de Tucuruí

Parâmetro	Campanha				IC 95% da Média (período seco)	IC 95% da Média (período chuvoso)
	1 (chuvoso)	2 (seco)	3 (seco)	4 (chuvoso)		
Al	0,09±0,02	0,03±0,01	0,09±0,04	0,08±0,07	0,07±0,03	0,08±0,03
Fe	0,91±0,27	0,77±0,45	0,47±0,12	0,48±0,10	0,62±0,23	0,69±0,15
Mn	0,07±0,02	0,08±0,03	0,04±0,01	0,07±0,04	0,05±0,01	0,07±0,02
Ba	0,04±0,00	0,03±0,01	0,04±0,01	0,04±0,00	0,03±0,00	0,03±0,00
Ca	3,42±0,07	2,78±0,05	3,24±0,11	2,76±0,07	3,01±0,07	3,09±0,07
Mg	1,76±0,02	1,36±0,03	1,44±0,05	0,99±0,02	1,40±0,03	1,37±0,06
Na	2,12±0,08	2,26±0,13	1,75±0,23	1,13±0,05	2,00±0,14	1,62±0,09

A turbidez apresentou valores médios acima do limite estabelecido pelas portarias do ministério da saúde (5 UT) durante a primeira, terceira e quarta campanha. O oxigênio dissolvido (OD) apresentou valores médios bem próximos do limite mínimo estabelecido pelo

CONAMA (5 mg/L) durante as quatro campanhas. O alumínio (Al) apresentou valores próximos do limite máximo estabelecido pelo CONAMA (0,1 mg/L). O Ferro (Fe) apresentou valores acima do limite máximo estabelecido pelas portarias do ministério da saúde (0,3 mg/L) e pelo CONAMA (0,3 mg/L).

Os parâmetros pH, alcalinidade, Al e Mn não apresentaram comunalidades adequadas para utilização da técnica, portanto, somente profundidade, temperatura, condutividade, STS, turbidez, NH<sub>3</sub>, OD, DQO, Ba, Ca, Fe, K, Mg e Na foram analisados, resultando em quatro fatores que juntos representam 77,06% da variância total dos dados, com autovalores maiores que 1 (Tabela 3.6). O KMO foi 0,70, maior que 0,50 para toda a matriz de correlação e o teste de Bartlett apresentou um valor significativo.

Tabela 3.6: Fatores, variância total e comunalidades obtidos pela análise fatorial com extração por análise de componente principal e rotação varimax de Tucuruí

Variável	Comunalidade	Fator			
		1	2	3	4
Condutividade (µs/cm)	0,73	<b>0,72</b>	0,38	0,25	-0,06
STS (mgL <sup>-1</sup> )	0,76	<b>0,71</b>	0,47	0,12	-0,09
Ba (mgL <sup>-1</sup> )	0,83	<b>0,87</b>	0,23	0,16	-0,08
Fe (mgL <sup>-1</sup> )	0,82	<b>0,72</b>	0,53	0,12	-0,03
K (mgL <sup>-1</sup> )	0,65	<b>0,72</b>	-0,15	0,08	0,31
Na(mgL <sup>-1</sup> )	0,82	<b>0,84</b>	-0,27	-0,16	0,09
Profundidade (m)	0,61	0,07	<b>0,77</b>	-0,10	0,04
Temperatura(°C)	0,67	0,06	<b>-0,79</b>	-0,08	-0,17
OD (mg O <sub>2</sub> /L)	0,81	-0,47	<b>-0,73</b>	0,03	0,24
Turbidez (NTU)	0,83	-0,08	-0,05	<b>0,87</b>	0,25
NH <sub>3</sub> (µg/L)	0,73	0,12	0,10	<b>0,82</b>	0,20
DQO (mg O <sub>2</sub> /L)	0,79	0,38	0,06	<b>0,72</b>	-0,35
Ca (mgL <sup>-1</sup> )	0,83	-0,04	0,14	0,25	<b>0,86</b>
Mg (mgL <sup>-1</sup> )	0,91	0,13	-0,02	0,02	<b>0,95</b>
% Variância		28,17	18,39	15,67	14,83
Cumulativo %		28,17	46,56	62,23	77,06

O fator 1 é composto pelas variáveis: Condutividade, STS, Ba, Fe, K e Na, que juntas representam 28,17% da variabilidade total dos dados. As variáveis que constituem este fator

estão fortemente correlacionadas. A Figura 3.5 apresenta o intervalo de confiança de 95% das médias das concentrações dos metais que constituem este fator.

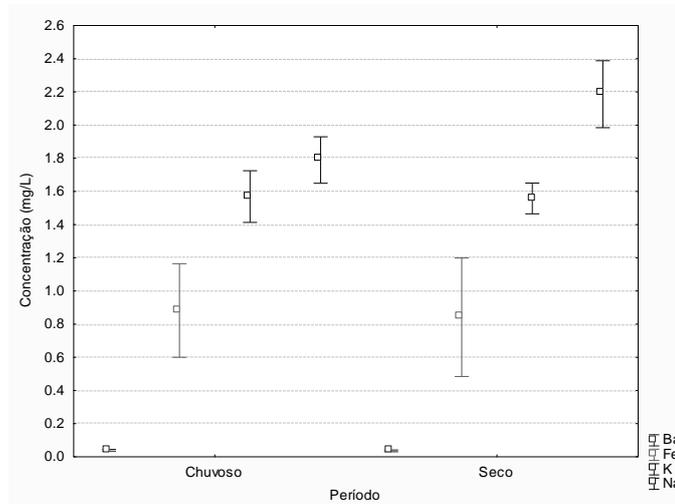


Figura 3.5: Intervalo de confiança das médias das concentrações (mgL<sup>-1</sup>) do Ba, Fe, K e Na

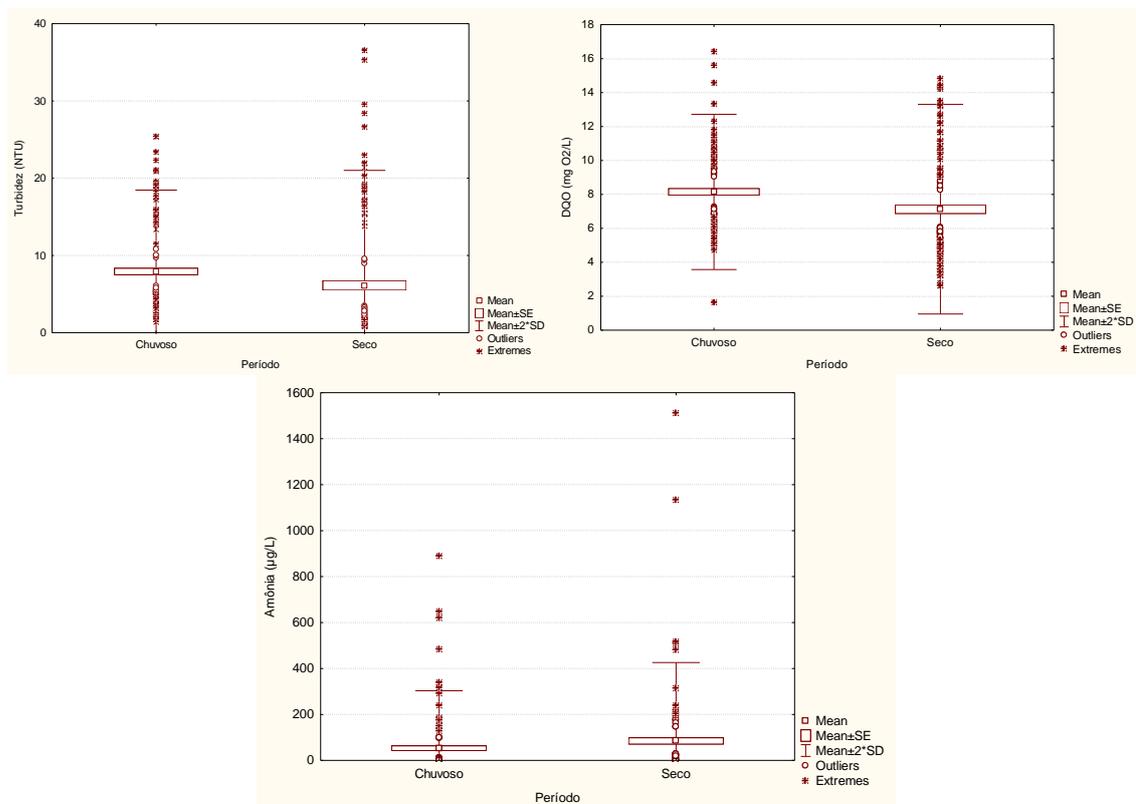
Os valores da condutividade no reservatório variaram de 57,70 a 61,82  $\mu\text{S cm}^{-1}$  no período seco e de 54,45 a 56,43  $\mu\text{S cm}^{-1}$  no período chuvoso. A condutividade é sensível a campos elétricos de íons dissolvidos.

A energia do reservatório vem não apenas do rio Tocantins, mas também de seus afluentes, o que explica a presença de Na<sup>+</sup> e K<sup>+</sup> em seis pontos de coleta durante o período chuvoso. Estes pontos estão localizados na área de recarga, sob o domínio de rochas alcalinas, contribuindo para a formação de água de composição mista, tais como bicarbonato de cálcio, magnésio, sódio ou águas de sódio. Nestas áreas de recarga, a lixiviação de cátions é facilitada pela morfologia do terreno, composto por colinas que dominam o lado oeste do reservatório. A água intrusiva pode ter densidades diferentes em relação à água do reservatório, e pode penetrar todos os perfis verticais da coluna de água. Esta intrusão resulta em um transporte de materiais suspensos e carga de nutrientes ou organismos transportados para várias profundidades (TUNDISI e TUNDISI, 2008).

O fator 2 é composto pelas variáveis: profundidade, temperatura e oxigênio dissolvido, que juntas representam 18,19% da variabilidade total dos dados. Os valores médios de oxigênio dissolvido variaram de 4,91 a 5,41 mgL<sup>-1</sup> no período seco e de 4,63 a 5,17 mgL<sup>-1</sup> no período chuvoso. Valores próximos ou iguais a 0,00 mgL<sup>-1</sup> foram obtidos dentro do reservatório, uma característica de lagos estratificados onde a acumulação de material em

decomposição é alta no fundo e em baixas temperaturas e o consumo de oxigênio é alto (TUNDISI e TUNDISI, 2008), indicando um estado eutrófico ou quase eutrófico.

O fator 3 é composto pelas variáveis: turbidez, DQO e  $\text{NH}_3$ , que juntas representam 15,67% da variabilidade total dos dados. A principal forma de nitrogênio analisada neste estudo foi o íon amônio ( $\text{NH}_4^+$ ) ou amônia, o qual é importante para os processos de produção primários (ESTEVES, 1988). Os maiores valores foram encontrados durante o período seco, alcançando aproximadamente  $1513,05 \text{ mgL}^{-1}$  formado principalmente em grandes profundidades do reservatório, em um perfil de concentração crescente a partir da superfície, onde o fitoplâncton cresce. Durante o período seco, o nível da água do reservatório é menor, assim como o volume de água intrusiva, STS e turbidez e com isto é possível atingir um equilíbrio químico aquoso, no entanto, durante este período foi observada uma pequena variação na profundidade, com profundidades máximas no período seco e chuvoso de 60 e 67 m respectivamente. A liberação de íons em grandes profundidades também parece ser acompanhada por um aumento da turbidez. A Figura 3.6 apresenta o intervalo de confiança de 95% das variações das médias deste fator.



O fator 4 é composto pelas variáveis: Ca e Mg, que juntas representam 14,83% da variabilidade total dos dados. O caráter bicarbonatado da água é característico de águas continentais e reflete no equilíbrio  $\text{CO}_2\text{-H}_2\text{O}$ , bem como na hidrólise de silicatos (dissolução de feldspatos), embora a sílica não tenha sido analisada. Os solos da área são caracterizados pela baixa fertilidade e alta acidez (ESTEVEES, 1998; TUNDISI *et al.*, 2003). As águas da região drenam grandes extensões de solos ácidos (Latossolos) e pobres em nutrientes, portanto, é razoável esperar concentrações extremamente baixas dos íons em solução (KRONBERG *et al.* 1979; STALLARD e EDMOND 1983; KRONBERG, 1987; STALLARD e EDMOND 1987; KONHAUSER *et al.* 1994). Ainda assim, as análises químicas da água de Tucuruí distinguem-se de águas de terrenos sedimentares pela presença marcante do  $\text{Ca}^{2+}$  com amplo domínio do  $\text{Na}^+$  ou como segundo elemento mais abundante, aproximando-as de águas que drenam terrenos cristalinos dominados por rochas básicas. As concentrações de  $\text{Ca}^{2+}$  e  $\text{Na}^+$  podem resultar da dissolução de feldspatos, principalmente do plagioclásio cálcico na ausência da calcita ( $\text{CaCO}_3$ ) enquanto que o  $\text{Na}^+$  pode vir dos feldspatos alcalinos.

Os teores dos cátions principais ( $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Na}^+$  e  $\text{K}^+$ ) obedecem à ordem  $\text{Ca}^{2+} > \text{Na}^+ > \text{Mg}^{2+} \leq \text{K}^+$ . Esses constituintes possuem um caráter conservativo no meio aquoso, com pequenas diminuições na concentração do  $\text{Ca}^{2+}$  dentro do reservatório. A composição química geral, com o domínio do íon  $\text{HCO}_3^-$ , indica águas mistas, de bicarbonatadas cálcio-magnesianas-sódicas no período seco a cálcicas no período chuvoso; algumas amostras tendem ao campo sódico no período seco.

Com Análise Discriminante não linear foi possível classificar corretamente 91,00% e 88,40% das amostras de análise e teste respectivamente. Sem a pressuposição de independências das variáveis, com a rede neural proposta foi possível classificar corretamente 100% das amostras de análise e validação respectivamente. Análogo às metodologias utilizadas na classificação das amostras de águas do reservatório de Coroacy Nunes, as técnicas estatística multivariada e a neural são adequadas para análise dos dados apresentados neste estudo.

## 4. CONCLUSÕES

Este estudo utilizou técnicas exploratórias de dados (Análise Fatorial, Discriminante e Redes Neurais Artificiais) para interpretar informação e aquisição do conhecimento resultante necessário para aplicação no estudo das concentrações de metais e dos parâmetros físico-químicos nos reservatórios das usinas hidrelétricas de Coaracy Nunes e Tucuruí, localizadas no norte do Brasil.

O reservatório de Coaracy Nunes não apresentou diferenças significativas nos valores do pH da água durante o período seco e chuvoso, enquanto que a temperatura, condutividade, alcalinidade, amônia e oxigênio dissolvido apresentaram valores elevados no período seco e os sólidos totais em suspensão, dureza e demanda química de oxigênio apresentaram valores elevados durante o período chuvoso. Neste reservatório, os resultados da aplicação das técnicas estatísticas mostraram que Ca, K, Mg e Na não possuem ligação direta com a demanda química de oxigênio, o que indica a possibilidade de terem sido introduzidos no sistema, provavelmente a partir de esgotos domésticos e escoamentos agrícolas nas proximidades do reservatório. Foram observados valores de Al, Fe e Mn acima dos limites estabelecidos pelas portarias do ministério da saúde e pelo CONAMA.

As águas do reservatório de Tucuruí foram caracterizadas pela forte presença de  $\text{Ca}^{2+}$  com predominância de  $\text{Na}^+$ , características que se aproxima de águas que drenaram terrenos cristalinos dominados por pedras básicas e feldspatos alcalinos presentes na área e o estado trófico desta área foi indicado como eutrófico ou quase eutrófico. Foram observados valores de Al e Fe acima dos limites estabelecidos pelas portarias do ministério da saúde e pelo CONAMA, e os resultados da aplicação das técnicas estatísticas não indicou nenhum impacto antrópico no reservatório.

No presente estudo foi possível verificar que a quantidade de variáveis constituintes dos fatores resultantes da análise fatorial, bem como o valor do  $\lambda$  de Wilks podem indicar o grau de complexidade de RNAs em problemas de classificação. Este estudo também demonstrou

que a combinação das técnicas estatísticas com redes neurais artificiais pode ser utilizada para geração de soluções mais precisas em problemas de classificação. As aplicações estatísticas também são relevantes em casos de gestão ambiental e política de tomada de decisão, pois auxiliam na identificação de fatores que devem ser estudados e/ou modificados para recuperar corpos de água degradados ou contaminados.

## 5. TRABALHOS FUTUROS

Neste estudo foram analisados os reservatórios de duas usinas hidrelétricas, um considerado profundo e outro pouco profundo, e as amostras de água de cada reservatório foram classificadas quanto à sazonalidade do período no qual foram coletados. Como trabalho futuro é possível aplicar a mesma metodologia em reservatórios diferentes e classificar as amostras de água em relação a critérios diferentes da sazonalidade, como por exemplo, a profundidade ou a distribuição espacial dos pontos de coleta.

Outra possibilidade de trabalho futuro é o desenvolvimento de um software que utilize técnicas da estatística e da inteligência computacional na geração de diagnósticos automatizados de corpos d'água de reservatórios de hidrelétrica, indicando possíveis fontes de poluição e soluções para recuperar estes corpos d'água.

## REFERÊNCIAS BIBLIOGRÁFICAS

ADEBISI, A. A. The physico-chemical hydrology of a tropical seasonal river upper Ogun river. **Hydrobiologia**, 79, Março 1981. 157–65.

ALLAN, I. J.; VRANA, B.; GREENWOOD, R. A “toolbox” for biological and chemical monitoring requirements for the European Union’s Water Framework Directive. **Talanta**, 69, 2006. 302-322.

APHA. **American Public Health Association Standard methods for the examination of water and wastewater**. American Public Health Association Inc. Nova York. 1995.

ATOBATELE, O. E.; UGWUMBA, O. A. Seasonal variation in the physicochemistry of a small tropical reservoir (Aiba Reservoir, Iwo, Osun, Nigeria). **African Journal of Biotechnology**, 7, Junho 2008. 1962-1971.

CHAPMAN, D. (Ed.). **Water Quality Assessments: A guide to the use of biota, sediments and water in environmental monitoring**. 2. ed. London: UNESCO/WHO/UNEP, 1996. 651 p.

COLETTI, C.; TESTEZLAF, R.; RIBEIRO, T. A. P.; SOUZA, R. T. G.; PEREIRA, D.A. Water quality index using multivariate factorial analysis. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, 14, Março 2010. 517–522.

DAVIES, O. A.; ABOLUDE, D. S.; UGWUMBA, A. A. A. Phytoplankton of the lower reaches of Okpoka creek, Port Harcourt, Nigeria. **Journal of Fisheries International**, 3, 2008. 83-90.

DEEKAE, S. N.; ABOWEI, J. F. N.; ALFRED-OCKIYA, J. F. Seasonal Variation of Some Physical and Chemical Parameters of Luubara Creek, Ogoni Land, Niger Delta, Nigeria. **Research Journal of Environmental and Earth Sciences**, 2, Outubro 2010. 208-215.

ELETRONORTE. **Manual do sistema de gestão ambiental - UHE Tucuruí**. ELETRONORTE. Tocantins. 1988.

EPA. **US EPA STANDARD METHODS 1060B. Collection and preservation of samples**. US EPA. [S.l.]. 1996.

ESTEVES, F. A. **Fundamentos de Limnologia**. 2. ed. Rio de Janeiro: Interciência, 1998. 602 p.

FISCH, G. F.; JANUÁRIO, M.; SENNA, R. C. Impacto ecológico em Tucuruí (PA): Climatologia. **Acta Amazonica**, 20, 1990. 49-60.

GOMES, G. S. S.; LUDERMIR, T. B. Redes neurais artificiais com funções de ativação complemento loglog e probit para aproximar funções na presença de observações extremas. Learning and nonlinear models. **Revista da Sociedade Brasileira de Redes Neurais (SBRN)**, 6, 2008. 142-153.

GUPTA, S. K.; GUPTA, R. C. **General and Applied Ichthyology: Fish and Fisheries**. Nova Délhi: Chand (S.) & Co Ltd, 2006. 1130 p.

HAIR, J. F.; TATHAM, R. L.; ANDERSON, R. E.; BLACK, W. **Multivariate data analysis**. 5. ed. New Jersey: Prentice Hall, 1998.

HAO, Z.; HUANG, X.; ZHU, C. Water quality evaluation through application of neural network at Yellow River. **International Conference on Natural Computation**, Washington, DC, USA, 3, 2008. 394-397.

HART, A. I.; ZABBEY, N. Physico-chemistry and benthic fauna of Woji creek in the lower Niger Delta. **Nigeria Environ Ecol**, 23, 2005. 361–368.

HAUSER-DAVIS, R. A.; OLIVEIRA, T. F.; SILVEIRA, A. M.; SILVA, T. B.; ZIOLLI, R. L. Case study: comparing the use of nonlinear discriminating analysis and Artificial Neural Networks in the classification of three fish species: acaras (*Geophagus brasiliensis*), tilapias (*Tilapia rendalli*) and mullets (*Mugil liza*). **Ecological Informatics**, 5, 2010. 474–478.

HAUSER-DAVIS, R. A.; OLIVEIRA, T. F.; SILVEIRA, A. M.; PROTÁZIO, J. M. B.; ZIOLLI, R. L. Logistic regression and fuzzy logic as a classification method for feral fish sampling sites. **Environ Ecol Stat**, 19, 2012. 473–483.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. 5. ed. [S.l.]: Prentice Hall, 2001. 767 p.

KABATA-PENDIAS, A. **Trace elements in soils and plants**. 4. ed. Boca Raton, Florida: Taylor and Francis Group, 2011.

KAZI, T. G.; ARAIN, M. B.; JAMALI, M. K.; JALBANI, N.; AFRIDI, H. I.; SARFRAZ, R. A.; BAIG, J. A.; SHAH, A. Q. Assessment of water quality of polluted lake using multivariate statistical techniques: A case study. **Ecotoxicology and Environmental Safety**, 72, 2009. 301–309.

KHALIL, B. M.; AWADALLAH, A. G.; KARAMAN, H.; EL-SAYED, A. Application of Artificial Neural Networks for the Prediction of Water Quality Variables in the Nile Delta. **Journal of Water Resource and Protectio**, 4, 2012, 388-394.

KONHAUSER, K. O.; FYFE, W. S.; KRONBERG, B. I. Multielement Chemistry of Some Amazonian Waters and Soils. **Chemical Geology**, 111, Janeiro 1994. 155–175.

KRONBERG, B. I.; FYFE, W. S.; LEONARDOS JR., O. H.; SANTOS, A. M. The chemistry of some Brazilian soils: Element mobility during intense weathering. **Chemical Geology**, 24, 1979. 211–229.

KRONBERG, B. I.; NESBITT, H. W.; FYFE, W. S. Mobilities of Alkalis, Alkaline-Earths and Halogens during Weathering. **Chemical Geology**, 60, Março 1987. 41–49.

LAMB, J. C. **Water quality and its control**. Nova York: John Wiley & Sons, 1985. 384 p.

LEITE, M. A.; DORNFELD, C. B.; ESPÍNDOLA, E. L. G. Análise da concentração de metais na água do reservatório de Salto Grande: diagnóstico ambiental e de saúde pública. Disponível em: <[HTTP://www.bvsde.paho.org/bvsAIDIS/PuertoRico29/busos.pdf](http://www.bvsde.paho.org/bvsAIDIS/PuertoRico29/busos.pdf)>. Acesso em: 23 Abril 2013.

LEWIS, W. M. Basis for the protection and management of tropical lakes. **Lakes and Reservoirs: Research and Management**, 5, 2000. 35-48.

LNCC. O Algoritmo Backpropagation. **Laboratório Nacional de Computação Científica - Ministério da Ciência e Tecnologia**. Disponível em: <[http://www.lncc.br/~labinfo/tutorialRN/frm4\\_backpropagation.htm](http://www.lncc.br/~labinfo/tutorialRN/frm4_backpropagation.htm)>. Acesso em: Outubro 2012.

MACHIWA, P. K. Water quality management and sustainability: the experience of Lake Victoria Environmental Management Project (LVEMP)—Tanzania. **Physics and Chemistry of the Earth, Parts A/B/C**, 28, 2003. 1111–1115.

MATTHEWS, M. W.; BERNARD, S.; WINTER, K. Remote sensing of cyanobacteria-dominant algal blooms and water quality parameters in Zeekoevlei, a small hypertrophic lake, using MERIS. **Remote Sensing of Environment**, 114, 2010. 2070–2087.

MUSTAPHA, M. K. Assessment of the water quality of Oyun Reservoir, Offa, Nigeria, using selected physico-chemical parameters. **Turkish Journal of Fisheries and Aquatic Sciences**, 8, 2008. 309-319.

NEGNEVITSKY, M. **Artificial Intelligence: A guide to intelligent systems**. 2. ed. [S.l.]: Pearson Education, 2005. 415 p.

OBEN, B. O. **Limnological assessment of the impact of agricultural and domestic effluent on three man-made lakes in Ibadan, Nigeria**. [S.l.]: University of Ibadan, 2000. 334 p. PhD degree.

ODERMATT, D.; GIARDINO, C.; HEEGE, T. Chlorophyll retrieval with MERIS Case-2-Regional in perialpine lakes. **Remote Sensing of Environment**, 114, 2010. 607-617.

ODUWOLE, G. A. **Indices of pollution in Ogunpa and Ona Rivers, Ibadan: physico-chemical, trace metal and plankton studies**. [S.l.]: University of Ibadan, 1997. 293 p. PhD degree.

OKAYI, R. G. **Effect of effluent discharge on water quality, distribution and abundance of plankton and fish species of River Benue**. Ibadan: University of Ibadan, 2003.

OUYANG, Y. Evaluation of river water quality monitoring stations by principal component analysis. **Water Research**, n. 39, 2005. 2621–2635.

PAPAFILIPPAKI, A. K.; KOTTI, M. E.; STAVROULAKIS, G. G. Seasonal variations in dissolved heavy metals in the Keritis River, Chania, Greece. **Global NEST Journal**, 10, 2008. 320-325.

RUSSELL, S.; NORVIG, P. **Artificial Intelligence: A modern approach**. 3. ed. [S.l.]: Pearson Education, 2010. 1132 p.

SIMEONOV, V.; STRATIS, J. A.; SAMARA, C.; ZACHARIADIS, G.; VOUTSA, D.; ANTHEMIDIS, A.; SOFONIOU, M.; KOUIMTZIS, T. Assessment of the surface water quality in Northern Greece. **Water Research**, 37, 2003. 4119–4124.

SINGH, K. P.; MALIK, A.; MOHAN, D.; SINHA, S. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) - a case study. **Water Research**, 38, Novembro 2004. 3980–3992.

SONG, M. W.; HUANG, P.; LI, F.; ZHANG, H.; XIE, K. Z.; WANG, Z. H.; HE, G. X. Water quality of a tributary of the Pearl River, the Beijiang, Southern China: implications from multivariate statistical analyses. **Environ Monit Assess**, 172, 2011. 589–603.

STALLARD, R. F.; EDMOND, J. M. Geochemistry of the Amazon.2. The Influence of Geology and Weathering Environment on the Dissolved-Load. **Journal of Geophysical Research: Oceans**, 88, Setembro 1983. 9671–9688.

STALLARD, R. F.; EDMOND, J. M. Geochemistry of the Amazon.3. Weathering Chemistry and Limits to Dissolved Inputs. **Journal of Geophysical Research:Oceans**, 92, 1987. 8293–8302.

TEPE, Y.; MUTLU, E. Physico-chemical characteristics of Hatay Harbiye Spring water, Turkey. **Journal of the Institute of Science and Technology of Dumlupınar University**, 6, 2005. 77-88.

TUNDISI, J. G.; SANTON, M. A.; MENEZES, C. F. S. **Tucuruí reservoir and hydroelectric Power plant**. [S.l.]. 2003.

TUNDISI, J. G.; TUNDISI, T. M. **Limnologia**. São Paulo: Oficina de textos, 2008.

VRANA, B.; ALLAN, I. J.; GREENWOOD, R.; MILLS, G. A.; DOMINIAK, E.; SVENSSON, K.; KNUTSSON, J.; MORRISON, G. Passive sampling techniques for monitoring pollutants in water. **Trends in Analytical Chemistry**, 24, Novembro 2005. 845–868.

WANG, X. L.; YIN, Z. J.; LV, Y. B.; LI, S. F. Operating rules classification system of water supply reservoir based on Learning Classifier System. **Expert Systems with Applications**, 36, Abril 2009. 5654–5659.

WENNER, E.; SANGER, D.; ARENDT, M.; HOLLAND, A. F.; CHEN, Y. Variability in dissolved oxygen and other water-quality variables within the national estuarine research reserve system. **Journal of Coastal Research**, 45, 2004. 17-38.

WETZEL, R. G. **Limnology, lake and river ecosystems**. Nova York: Academic Press, 2001. 1006 p.

XUE, X.; CHEN, Y. Water quality evaluation of surface water based on back propagation neural networks. **International Symposium on Water Resource and Environmental Protection**, Xi'an, China, Maio 2011.

YENILMEZ, F.; KESKIN, F.; AKSOY, A. Water quality trend analysis in Eymir Lake, Ankara. **Physics and Chemistry of the Earth, Parts A/B/C**, 36, 2011. 135-140.

ZAHRAIE, B.; HOSSEINI, S. M. Development of reservoir operation policies considering variable agricultural water demands. **Expert Systems with Applications**, 36, Abril 2009. 4980–4987.