

**ISAIAS DA SILVA PEREIRA**

**ÍNDICES DE QUALIDADE DA ÁGUA PARA CONSUMO HUMANO:  
MODELAGEM E INFLUÊNCIA DO USO E OCUPAÇÃO DO SOLO**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Civil, para obtenção do título de *Doctor Scientiae*.

**VIÇOSA**  
**MINAS GERAIS – BRASIL**  
**2014**

**Ficha catalográfica preparada pela Biblioteca Central da Universidade  
Federal de Viçosa - Câmpus Viçosa**

T

P436i  
2014  
Pereira, Isaias da Silva, 1965-  
Índices de qualidade da água para consumo humano :  
modelagem e influência do uso e ocupação do solo / Isaias da  
Silva Pereira. – Viçosa, MG, 2014.  
xiv, 104f. : il. (algumas color.) ; 29 cm.

Orientador: Rafael Kopschitz Xavier Bastos.  
Tese (doutorado) - Universidade Federal de Viçosa.  
Referências bibliográficas: f.95-104.

1. Água - Qualidade. 2. Recursos hídricos - Análise.  
3. Solos - Uso. 4. Solos - Composição. 5. Redes neurais  
artificiais. I. Universidade Federal de Viçosa. Departamento de  
Engenharia Civil. Programa de Pós-graduação em Engenharia  
Civil. II. Título.

CDD 22. ed. 553.7

**ISAIAS DA SILVA PEREIRA**

**ÍNDICES DE QUALIDADE DA ÁGUA PARA CONSUMO HUMANO:  
MODELAGEM E INFLUÊNCIA DO USO E OCUPAÇÃO DO SOLO**

Tese apresentada à Universidade Federal de Viçosa, como parte das exigências do Programa de Pós-Graduação em Engenharia Civil, para obtenção do título de *Doctor Scientiae*.

APROVADA: 18 de fevereiro de 2014.

---

Prof<sup>a</sup> Maria Lúcia Calijuri

---

Prof<sup>a</sup> Ann Honor Munteer

---

Prof. Elpídio Inácio Fernandes Filho

---

Prof. Carlos Antônio Oliveira Vieira

---

Prof. Rafael Kopschitz Xavier Bastos  
(Orientador)

## AGRADECIMENTOS

*A Deus pela paciência e cumplicidade.*

*Em especial, ao amigo professor e orientador Rafael K. X. Bastos, obrigado por tudo, minha admiração e respeito pela lisura com o povo brasileiro.*

*À Prof<sup>a</sup>. Paula Dias Bevilacqua, pela amizade, respeito e contribuição na execução do trabalho.*

*Aos Professores do DEC: Carlos Vieira, Eduardo Marques, Lúcia Calijuri e Ann Mounteer pelo acolhimento e amizade.*

*À dona Benvenida, minha mãe, ao seu Mundico, meu pai e à mãe Francisca pela vinda ao mundo.*

*Em especial, à minha irmã Mariquinha pelo amor incondicional que carregamos em nós.*

*À Luíza pelo cuidado, admiração, carinho e respeito.*

*Aos meus filhos Fábio, Lia, Rafael e Gabriel pelo carinho, amizade, cuidado e apoio nos momentos em que a vida nos fez simplesmente amigos.*

*Ao Dito e ao Compadre Nonato pela cumplicidade, pela verdadeira e sólida amizade forjada no tempo, caminhada que nos permite sermos sempre os meninos de nossas mães.*

*Às minhas irmãs Francismar, Viliani, Isailda e Isanilda e Chico por tudo.*

*Aos sobrinhos(as): Jeferson, Jackson, Jeciane, Joyce, Kassandra, Lais, Juliana, Robson, Jadson, Maicon, Tarucauê, Anahi, Yuêberi, Raimundo, Cecília, Isabela, Miguel, Ana Clara, Junior e Juan.*

*Ao meu amigo e cunhado Zé Carvalho.*

*À Liliane, Adriana, Victor Luis Eduardo, Raimundão, Tião, Silvério, Japonês, Niel, Antonio Carlos e Arlindo.*

*A João Pimenta, Cocó, Bjei, Adriana, Zenga, Gentileza, Lardotto, Chatisse, Tiago, Daniel, Grazi, Léo, Anibal, Mateus, Beleiza, Soneca, Felipe, Luísa, Aline e Pedro pela amizade e pronta colaboração.*

*Aos Amigos do Romão dos Reis: Eduardinho I e II, Eduardão, Rodrigo I e II, Chico, Biro, Dinho, Zezé, Guilherme e Pedrinho.*

*Aos amigos da Padoca: Carlão, Fernando, Ivens, Josi, Hugo, Cícero, Francês, Débora, Aglaê, Valério, Consuelo e Comadre Rosa e Ana Clara.*

*Ao amigo Álvaro pelos ensinamentos e prosas mineiradas vida afora.*

*Ao CNPq pela disponibilização da bolsa e financiamento da pesquisa.*

*À Universidade Federal de Viçosa, à Divisão de Água e Esgotos, ao Departamento de Engenharia Civil e ao Programa de Pós-Graduação em Engenharia Civil, pela formação e condições de trabalho proporcionadas.*

## **BIOGRAFIA**

**ISAIAS DA SILVA PEREIRA**, filho de Raimundo da Silva Pereira e Benvinda Alves de Moura, nasceu em Teresina – PI, no bairro Lourival Parente, em 06 de agosto de 1965.

Concluiu o primeiro grau no colégio Unidade Escolar Monsenhor Cícero Portela Nunes e o segundo no colégio Cipreve.

Graduou-se em Engenharia Agrônômica pela Universidade Federal de Viçosa em janeiro de 1992.

Concluiu o mestrado em Ciências Agrárias, área de concentração Manejo e Conservação do Solo e da Água, pela Universidade de Brasília em julho de 1997.

Professor da rede pública do Distrito Federal de Ciências Físicas e Biológicas 1996.

Professor Substituto (FAV/UNB) de Edafologia, Manejo e Conservação, Adubos e Adubações, Estágio Supervisionado e Solos Florestais de 1997 a 1999.

Servidor da Secretaria de Meio Ambiente e Recursos Hídricos do Distrito Federal e no Ministério do Meio Ambiente de 1999 a 2001.

Consultor e gerente técnico do Ministério da Saúde/FUNASA na área de Vigilância Ambiental - VIGIAGUA - de 2001 a 2009 e 2012/13.

Professor de Fertilidade do solo, Manejo e Conservação e Fundamentos de Ciência do Solo da FAV/UNB - 2013.

Em março de 2009 iniciou o curso de Doutorado em Engenharia Civil na Universidade Federal de Viçosa, na área de concentração Saneamento e Ambiental.

## SUMÁRIO

<b>LISTA DE QUADROS .....</b>	<b>vii</b>
<b>LISTA DE FIGURAS .....</b>	<b>viii</b>
<b>LISTA DE TABELAS .....</b>	<b>x</b>
<b>RESUMO .....</b>	<b>xiii</b>
<b>ABSTRACT .....</b>	<b>xiv</b>
<b>1. INTRODUÇÃO.....</b>	<b>1</b>
<b>1.1. Gerais.....</b>	<b>3</b>
<b>1.2. Específicos .....</b>	<b>3</b>
<b>2. REVISÃO DE LITERATURA .....</b>	<b>4</b>
<b>2.1. Índices de Qualidade da Água.....</b>	<b>4</b>
<b>2.2. Análise de Componentes Principais - ACP .....</b>	<b>17</b>
<b>2.3. Redes neurais artificiais - RNA .....</b>	<b>24</b>
<b>2.4. Uso e ocupação do solo e variáveis de qualidade da água .....</b>	<b>31</b>
<b>3. MATERIAL E MÉTODOS .....</b>	<b>34</b>
<b>3.1. Área de estudo .....</b>	<b>34</b>
<b>3.2. Avaliação da qualidade da água e determinação do IQA e do IAP tradicionais.....</b>	<b>36</b>
<b>3.3. Prerrogativas para análise fatorial de componentes principais.....</b>	<b>39</b>
<b>3.4. Elaboração da matriz de coeficientes de correlação .....</b>	<b>39</b>
<b>3.5. Extração de fatores.....</b>	<b>39</b>
<b>3.6. Rotação ortogonal dos eixos .....</b>	<b>40</b>
<b>3.7. Determinação dos escores fatoriais.....</b>	<b>40</b>
<b>3.8. Análises de correlações entre o índice Cetesb e obtidos pelos escores fatoriais.....</b>	<b>40</b>
<b>3.9. Redes neurais artificiais na modelagem dos índices da qualidade da água .....</b>	<b>41</b>
<b>3.9.1. Construção dos modelos de RNAs .....</b>	<b>41</b>
<b>3.9.2. Avaliação de desempenho das RNAs.....</b>	<b>42</b>
<b>3.10. Uso e ocupação do solo e variáveis de qualidade da água.....</b>	<b>44</b>

3.11.	Analises estatísticas e nomenclatura dos índices modelados .....	46
4.	RESULTADOS E DISCUSSÃO.....	47
4.1.	Condições para realização da análise fatorial de componentes principais	47
4.1.1.	Teste de normalidade.....	47
4.1.2.	Teste de adequacidade da amostra .....	47
4.1.2.1.	Teste de adequacidade das variáveis do IQA <sub>Sb</sub> .....	47
4.1.2.2.	Teste de adequacidade das variáveis do IAP <sub>Sb</sub> .....	48
4.1.3.	Teste de Esfericidade .....	49
4.1.4.	Aplicação da Análise de Componentes Principais para IQA e IAP <sub>Sb</sub> .....	49
4.1.4.1.	Estatística descritiva das variáveis do IQA <sub>Sb</sub> .....	49
4.1.4.2.	Estatística descritiva das variáveis do IAP <sub>Sb</sub> .....	50
4.1.4.3.	Matriz de correlação do IQA <sub>Sb</sub> .....	51
4.1.4.4.	Matriz de correlação do IAP <sub>Sb</sub> .....	52
4.1.4.5.	Extração de fatores e rotação de eixos do IQA <sub>Sb</sub> .....	54
4.1.4.6.	Extração de fatores e rotação de eixos do IAP <sub>Sb</sub> .....	58
4.1.4.7.	Modelagem do IQA <sub>Sb</sub> pelo método de Bartlett .....	63
4.1.4.8.	Modelagem do IAP <sub>Sb</sub> pelo método de Bartlett.....	68
4.2.	Uso de redes neurais na modelagem do IQA <sub>Sb</sub> e IAP <sub>Sb</sub> .....	71
4.2.1.	Arquiteturas de redes neurais selecionadas para IQA <sub>Sb</sub> e IAP <sub>Sb</sub> .....	71
4.2.2.	Variáveis selecionadas para IQA <sub>Sb</sub> e IAPSb .....	73
4.2.3.	Modelos de predição para IQA <sub>Sb</sub> pela redes neurais e IAPSb e avaliação de desempenho.....	74
4.3.	Uso e ocupação do solo e variáveis de qualidade da água .....	79
4.3.1.	Teste de normalidade.....	79
4.3.2.	Diagnóstico do uso e ocupação do solo na bacia do ribeirão são Bartolomeu.....	79
4.3.3.	Correlações entre usos ocupações e variáveis de qualidade da água .....	86
5.	Conclusões .....	92
6.	Referências Bibliográficas.....	94

## LISTA DE QUADROS

Quadro 1- Classificação da água bruta com base no IQA .....	8
Quadro 2 - Limites inferiores e superiores para metais e potencial de formação de trihalometanos para cálculo do ISTO.....	13
Quadro 3- Faixas de número de células de cianobactérias e respectiva taxaço para cálculo do ISTO. ....	14
Quadro 4- Faixas de validade da aplicação da AF/ACP pelo teste de KMO.....	19
Quadro 5- Síntese do primeiro estágio do planejamento da AF. ....	20

## LISTA DE FIGURAS

Figura 1 - Curvas de variação de qualidade água das variáveis do IQA.....	9
Figura 2 - Curva de qualidade padrão para as variáveis incluídas no ISTO.....	12
Figura 3 - Representação de um neurônio artificial. ....	25
Figura 4 - Função de ativação sigmóidal. ....	26
Figura 5 - Localização do município de Viçosa e da bacia do ribeirão São Bartolomeu	34
Figura 6 - Delimitação das sub-bacias, localização das propriedades rurais e dos pontos de coleta de amostras água (círculos vermelhos vazados), bacia de captação do ribeirão São Bartolomeu, Viçosa–MG .....	38
Figura 7 - Mapa de uso e ocupação do na bacia ribeirão São Bartolomeu, Viçosa – MG. Fonte: Silva (2010).....	45
Figura 8 - Resultados médios de alumínio dissolvido e pH nos doze meses de monitoramento. ....	54
Figura 9 - Número de fatores extraídos com autovalor maior que a unidade.....	55
Figura 10 - Número de fatores extraídos com autovalor maior que a unidade.....	58
Figura 11 - Granja de criação de suínos na sub bacia do Córrego do Engenho.....	61
Figura 12 - Lançamento de efluentes da suinocultura no Córrego do Engenho. ....	61
Figura 13 - Distribuição dos valores médios do IQA Cetesb comparados com os modelados pelo método de Bartlett ( $IQAS_bF1$ ) nas oito sub bacias durante doze meses. ....	65
Figura 14 - Distribuição dos valores médios do IQA Cetesb comparados com os modelados pelo método de Bartlett ( $IQAS_bF2$ ) nas oito sub bacias durante doze meses. ....	66
Figura 15 - Distribuição dos valores médios do IQA Cetesb comparados com os modelados pelo método de Bartlett ( $IQAS_bF3$ ) nas oito sub bacias durante doze meses. ....	67
Figura 16 - Distribuição dos valores médios do IAP Cetesb comparados com os modelados pelo método de Bartlett ( $IAPS_bF1$ ) nas oito sub bacias durante doze meses. ....	70
Figura 17 - Arquitetura da rede de melhor desempenho para o $IQAS_b$ .....	72
Figura 18 - Arquitetura da rede de melhor desempenho para o $IAP_{Sb}$ .....	72
Figura 19 - Valores do IQA Cetesb versus $IQAS_b$ preditos pela rede neural R8. ....	76
Figura 20 - Valores do IQA Cetesb versus $IQAS_b$ preditos pela rede neural R7. ....	76

Figura 21 - Valores do IQA Cetesb versus IQA <sub>Sb</sub> preditos pela rede neural R6. ....	77
Figura 22 - Valores do IQA Cetesb versus IQA <sub>Sb</sub> preditos pela rede neural R5. ....	77
Figura 23 - Valores do IAP Cetesb versus IAP <sub>Sb</sub> preditos pela rede neural R5.....	78
Figura 24 – Usos e ocupação do solo por sub bacias no ribeirão São Bartolomeu.....	83
Figura 25 – Índices da qualidade da água por sub bacias do ribeirão São Bartolomeu..	84

## LISTA DE TABELAS

Tabela 1 - Comunalidades extraídas e medidas de adequação das variáveis do IQA <sub>Sb</sub> .	48
Tabela 2 - Comunalidades extraídas e medidas de adequação das variáveis do IAP <sub>Sb</sub> .	48
Tabela 3 – Estatística descritiva das variáveis do IQA <sub>Sb</sub>	50
Tabela 4 – Estatística descritiva das variáveis do IAP <sub>Sb</sub>	50
Tabela 5 – Coeficientes de correlação das variáveis físico, químicas e biológica do IQA <sub>Sb</sub> .	52
Tabela 6 – Coeficientes de correlação das variáveis físico químicas do IAP <sub>Sb</sub> .	53
Tabela 7 – Matriz de cargas fatoriais não rotacionadas para o IQA.	56
Tabela 8 – Matriz de cargas fatoriais rotacionadas para o IQA.	58
Tabela 9 – Matriz de cargas fatoriais não rotacionadas para o IAP.	59
Tabela 10 – Matriz de cargas rotacionadas ortogonalmente para o IAP.	62
Tabela 11 – Estatística descritiva dos índices calculados pelo método de Bartlett.	63
Tabela 12 – Matriz de correlação do IAP Cetesb e IQA modelado pela ACP.	64
Tabela 13 – Estatística descritiva dos índices calculados pelo método de Bartlett.	68
Tabela 14 – Matriz de correlação do IAP Cetesb e IAP modelado pela ACP.	68
Tabela 15 - Redes neurais selecionadas para o IQA <sub>Sb</sub>	71
Tabela 16 - Redes neurais selecionadas para o IAP <sub>Sb</sub>	72
Tabela 17 - Variáveis selecionadas pelas redes neurais para o IQA <sub>Sb</sub>	73
Tabela 18 - Avaliação de desempenho das redes neurais utilizadas na predição do IQA <sub>Sb</sub>	75
Tabela 19 - Avaliação de desempenho da rede neural 10 utilizada na predição do IAP <sub>Sb</sub>	75
Tabela 20 - Uso e ocupação do solo na bacia do ribeirão São Bartolomeu.	82
Tabela 21 - Uso e ocupação do solo por sub bacias no ribeirão São Bartolomeu.	85
Tabela 22 - Coeficientes de correlações entre as variáveis de qualidade da água e uso e ocupação do solo	87

## LISTAS DE SIGLAS

ACP - Análise fatorial de componentes principais.

$Al^{3+}$  - Alumínio

Cálcio -  $Ca^{2+}$ ,

CETESB - Companhia Ambiental do Estado de São Paulo

CONAMA - Conselho Nacional de Meio Ambiente

Ctot - Coliformes totais

DBO - Demanda biológica de oxigênio

ETA: Estação de Tratamento de Água

$Fe^{2+}$  - Ferro dissolvido

IAP - Índice de Qualidade de Água Bruta para Fins de Abastecimento Público

$IAP_{sb}$  - Índice de abastecimento público aferido para a bacia do São Bartolomeu.

$IAP_{sbF}$  - Índice de abastecimento público fatorial aferido para a bacia do São Bartolomeu

Índice de substâncias tóxicas e organolépticas - ISTO

IQA - Índice de Qualidade da Água.

IQAB: Índice de Qualidade de Água Bruta Afluente a Estações de Tratamento de Água Convencionais.

$IQA_{sb}$  - Índice da qualidade da água aferido para a bacia do São Bartolomeu

$IQA_{sbF}$  - Índice da qualidade da água fatorial aferido para a bacia do São Bartolomeu

IQUAS: Índice de Qualidade de Águas Subterrâneas Destinadas ao Uso na Produção de Água Potável.

ISTO: Índice de Substâncias Tóxicas e Organolépticas

$Mg^{2+}$  - Magnésio -

MLP - Multilayer Perceptrons

$Mn^{2+}$  - Manganês dissolvido

MS - Ministério da Saúde

NSF - National Sanitation Foundation Institution

Ntot - Nitrogênio total

OD - Oxigênio dissolvido

PFTHM - Potencial de formação de trihalometanos

pH – Potencial Hidrogeniônico

Ptot - Fósforo total

Ptot - Fósforo Total

RNAs - Redes neurais artificiais.

SAAE: Serviço Autônomo de Água e Esgoto de Viçosa

SO<sub>4</sub><sup>2-</sup> - Sulfato

ST - Sólidos totais

Tur - Turbidez

## RESUMO

PEREIRA, Isaias da Silva. D.Sc., Universidade Federal de Viçosa, fevereiro de 2014. **Índices de qualidade da água para consumo humano: modelagem e influência do uso e ocupação do solo.** Orientador: Rafael Kopschitz Xavier Bastos. Coorientadora: Paula Dias Bevilacqua.

Este trabalho teve como objetivo a modelagem de índices de qualidade da água pelo uso de análise de componentes principais e redes neurais artificiais para bacia do ribeirão São Bartolomeu no município de Viçosa, Minas Gerais, Brasil e avaliar a influência do uso e ocupação do solo nas variáveis de qualidade da água. Foram monitoradas, no período de outubro de 2009 a setembro de 2010, as variáveis oxigênio dissolvido, demanda bioquímica de oxigênio, fósforo, nitrogênio total, cor aparente, *E. Coli*, coliformes totais, turbidez, sólidos totais, temperatura, alumínio, cádmio, cromo, zinco, ferro, manganês, níquel e (oo) cistos de protozoários (*Giardia e Cryptosporidium*). Testes de adequacidade Kayser-Mayer-Olkin (KMO), de esfericidade de Bartlett e de normalidade pelo teste Kolmogorov-Smirnov indicaram que as análises poderiam ser realizadas. Os índices modelados pelo uso de análise de componentes principais (ACP) e de redes neurais artificiais (RNAs) apresentaram resultados semelhantes aos calculados pelo método desenvolvido pela Cetesb (Companhia Ambiental do Estado de São Paulo, Brasil). As variáveis oxigênio dissolvido, turbidez, nitrogênio total, cor e pH foram as que mais apareceram nos modelos, devendo ser mantidas, impreterivelmente, nos programas de monitoramentos da qualidade da água. Agricultura, café, pastagem degradada e piscicultura foram os usos do solo mais impactantes na degradação da qualidade da água. Áreas ocupadas com pastagem não degradada e formações florestais favoreceram uma melhor proteção dos solos e melhor qualidade da água. Os índices modelados selecionaram variáveis ajustadas às condições socioeconômicas da bacia, no entanto, novos estudos precisam ser desenvolvidos para avaliar a capacidade de generalização desses modelos em outras condições socioeconômicas e ambientais diferentes das encontradas na bacia do ribeirão São Bartolomeu.

## ABSTRACT

PEREIRA, Isaias da Silva. D.Sc., Universidade Federal de Viçosa, February, 2014. **Indices of water quality for human consumption. Modeling and influence of the use and occupation of land.** Adviser: Rafael Kopschitz Xavier Bastos. Co-adviser: Paula Dias Bevilacqua.

This study aimed the modeling of water quality indexes by using principal component analysis and artificial neural networks to São Bartolomeu watershed, in Viçosa, Minas Gerais, Brazil. It also concerned the evaluation of the influence of soil use and occupation on water quality variables. The following variables were monitored in the period from October 2009 to September 2010: dissolved oxygen, biochemical oxygen demand, phosphorus, total nitrogen, apparent color, *E. coli*, total coliform, turbidity, total solids, temperature, aluminum, cadmium, chromium, zinc, iron, manganese, nickel, and (oo) cysts, protozoan (*Cryptosporidium and Giardia*). Tests of adequacy (Kayser-Mayer-Olkin (KMO) and Bartlett's sphericity) and normality (Kolmogorov Smirnov) indicated the analyses could be performed. The indices modeled by using principal component analysis (PCA) and artificial neural networks (ANN) showed similar results to those calculated by the method developed by Cetesb (The Environmental Company of the State of São Paulo, Brazil). The variables dissolved oxygen, turbidity, total nitrogen, pH and color were the most apparent in the models and should necessarily be maintained in water quality monitoring programs. Agriculture, coffee, degraded pasture and fish-farming were the uses of soil that most strongly deteriorated water quality. Areas occupied by forest and non degraded grassland favored soil protection and a better water quality. The modeled indices selected variables which were adjusted to the socioeconomic conditions in the studied area, however, new studies are needed to evaluate the generalization of these models in other areas which hold distinct socioeconomic and environmental conditions than those from São Bartolomeu watershed.

## 1.INTRODUÇÃO

A qualidade da água, que varia em função das condições naturais e do uso e ocupação do solo na bacia hidrográfica, pode ser analisada por variáveis que representam suas características físicas, químicas e biológicas e que determinam a compatibilidade com diferentes usos (MERTEN e MINELLA, 2002).

O monitoramento da qualidade da água geralmente utiliza um grande número de variáveis, produzindo uma quantidade de dados que dificulta a análise e interpretação dos resultados (SIMEONOV et al., 2003). Cada sistema hídrico apresenta características próprias e por isso, a definição de variáveis indicadoras de sua qualidade se torna uma tarefa desafiadora para a pesquisa em qualidade da água (TOLEDO e NICOLELLA, 2002). Dessa maneira, o uso de índices de qualidade da água tem se mostrado como um indicador capaz resumir em um número adimensional informações referentes a um grupo de múltiplas variáveis, com informações que facilitam a comunicação com o público leigo; todavia apresentam como principal desvantagem a perda de informações das variáveis individualizadas e de sua capacidade de interação (TOLEDO e NICOLELLA, 2002; COLETTI et al, 2010).

No campo específico do abastecimento de água para consumo humano, um dos índices mais difundidos, no mundo e no Brasil, é o IQA (Índice de Qualidade da Água), desenvolvido na década de 1970 pela *National Sanitation Foundation Institution* (EUA), ajustado pela CETESB para as condições ambientais das bacias hidrográficas do Estado de São Paulo em 1975. Posteriormente, a CETESB desenvolveu o Índice de Qualidade de Água Bruta para fins de Abastecimento Público (IAP), que representou um avanço em relação ao IQA, por incorporar variáveis relacionadas aos riscos à saúde da população e à aceitação da água para consumo humano (CETESB, 2008).

A maioria dos programas estaduais de monitoramento da qualidade da água no Brasil utiliza o IQA como índice de referência. O uso do IAP torna-se mais restrito, devido à maior complexidade das variáveis envolvidas e ao maior custo das análises.

A motivação para o desenvolvimento desse trabalho veio da necessidade de se modelar índices e selecionar variáveis que sejam mais ajustadas as condições socioeconômicas e ambientais da bacia do ribeirão São Bartolomeu que é caracterizada pela presença de pequenas propriedades rurais dedicadas, em sua maioria, a produção leiteira, sem nenhuma atividade industrial e com poucas propriedades com criação

intensiva de animais, condições socioeconômicas bem diferenciadas das encontradas na maioria das bacias do Estado de São Paulo.

Com o intuito de selecionar variáveis mais representativas autores como Toledo e Nicolella, 2002; Meireles, 2007; Coletti et al, 2010; Al-Badaii (2013), utilizaram técnicas de análise multivariada de dados na extração de informações partindo de um conjunto grande de variáveis para um grupo de menor dimensão, com maior representatividade, melhor interpretação dos resultados e sem perda significativa das informações.

Outro ponto que motivou o estudo da associação entre uso e ocupação do solo e variáveis da qualidade da água é a importância desse tema para o planejamento e gestão dos recursos hídrico, pois a maioria dos trabalhos como os de Arcova e Cicco (1999); Guedes et al (2012) e Al-Badaii (2013) fazem inferência sobre a influência dos diferentes tipos de uso e ocupação do solo com as variáveis de qualidade da água, sem o estabelecimento de qualquer relação estatística capaz de comprovar essas relações.

Dessa forma, esse trabalho tem como uma de suas hipótese que a modelagem do índice de qualidade da água (IQA) e do índice de qualidade de água bruta para fins de abastecimento público (IAP) pelo uso de análise fatorial de componentes principais e de redes neurais artificiais possibilitará a seleção de um número menor de variáveis representativas e ajustadas as condições socioeconômicas e ambientais da bacia ribeirão São Bartolomeu. A bacia hidrográfica será a escala de análise, uma vez que é considerada como a unidade territorial de planejamento das ações e das políticas de recursos hídricos (BRASIL, 1997).

A outra, é que os tipos de usos e ocupação do solo que oferecem menor proteção da superfície do solo e conseqüentemente maior exposição aos processos de erosão hídrica estarão associados positivamente com as variáveis que contribuem para perda de qualidade da água dos mananciais superficiais na bacia.

## **OBJETIVOS**

### **1.1. Gerais**

- Propor modelagem do índice de qualidade da água (IQA) e do índice de qualidade de água bruta para fins de abastecimento público (IAP) ajustados às condições socioeconômicas e ambientais para a bacia do ribeirão São Bartolomeu.
- Averiguar a influência do uso e ocupação do solo nas variáveis de qualidade da água.

### **1.2. Específicos**

- Usar análise fatorial de componentes principais e redes neurais artificiais na modelagem de índices de qualidade da água para consumo humano.
- Comparar os resultados de qualidade da água calculados pelo índice desenvolvido pela Companhia Ambiental do Estado de São Paulo (CETESB) com os modelados pela análise fatorial de componentes principais e redes neurais artificiais.
- Verificar a influência dos diferentes tipos de uso e ocupação do solo e nas variáveis de qualidade da água através da análise dos resultados das correlações de Pearson.

## **2.REVISÃO DE LITERATURA**

### **2.1. Índices de Qualidade da Água**

Um índice de qualidade da água tem como finalidades agregar, classificar e disponibilizar informações simplificadas sobre a qualidade da água de um determinado corpo hídrico. Pode ser utilizado para verificar o cumprimento da legislação, a definição de prioridades de usos (abastecimento de água para consumo humano, definição do estado trófico, balneabilidade, preservação da vida aquática, etc.), comparação de condições ambientais em diferentes locais, avaliar degradação ou melhora na qualidade do corpo hídrico. Deve, sempre que possível, ser desenvolvido para situações específicas que considerem as relações ambientais, econômicas e sociais existentes na bacia hidrográfica (LEITE e FONSECA, 1994).

A vantagem da utilização de índices da qualidade da água está na sua capacidade de resumir informações de um grupo grande de variáveis em um número adimensional. Esse procedimento permite a classificação da água em diferentes classes de qualidade. Facilita a comunicação à população sobre as condições da qualidade da água, identifica problemas em trechos do curso de água que necessitem de cuidados especiais, além de atuar como ferramenta para a tomada de decisão por técnicos e gestores de recursos hídricos (SANTOS et al., 2001). No entanto, sua principal desvantagem consiste na perda de informação das variáveis individuais e da possibilidade de interação entre elas. A rigor, um índice, apesar de fornecer uma avaliação integrada, jamais substituirá uma avaliação detalhada da qualidade das águas de uma determinada bacia hidrográfica (LEITE e FONSECA, 1994).

Os primeiros estudos que se tem notícia relacionando o nível de pureza com a poluição da água foram realizados na Alemanha em 1848 (OTT, 1978) e procuravam sintetizar dados de qualidade da água com a ocorrência de determinadas comunidades de organismos aquáticos; em lugar de um valor numérico, a qualidade da água era categorizada por classes de poluição (DERÍSIO, 2007). Posteriormente, e ao longo de mais de um século, vários países desenvolveram e aplicaram diferentes sistemas para classificar a qualidade da água, usualmente de dois tipos: aqueles relativos à quantidade de poluição detectada e aqueles relativos à vida de comunidades de organismos macro e microscópicos, como por exemplo, peixes, organismos bentônicos e plantas (MEIRELES, 2007).

Em 1965, Horton, pesquisador alemão, fez a primeira proposição formal de um índice da qualidade da água, agregando variáveis quantitativas e qualitativas, e apresentado como uma ferramenta para avaliação de programas de redução da poluição e para informação pública. O Índice de Horton consistia, basicamente, em um somatório ponderado de subíndices (“notas” atribuídas a cada variável que compõe o índice, de acordo com a respectiva concentração ou valor), divididos pelo somatório dos pesos (fixados para cada variável) multiplicado por dois coeficientes, que consideram a temperatura e a poluição evidente de um curso d’água (Equação 1) (DERÍSIO, 2007). As variáveis e os respectivos pesos que compunham o Índice de Horton são: oxigênio dissolvido – OD (4), pH (4), tratamento de esgoto (4), coliformes fecais (2), alcalinidade (1), cloretos (1), condutividade elétrica (1), carbono (extraído por clorofórmio) Equação (1).

$$I = \frac{\sum_{i=1}^n W_i l_i}{\sum_{i=1}^n W_i} M_1 M_2 \quad (1)$$

onde:

$W_i$  = peso de cada variável (i) que entra no cálculo;

$l_i$  = subíndice da variável (i);

$M_1$  = coeficiente que reflete a temperatura; = 1 se  $T < 34^\circ\text{C}$ ; = 0,5 se  $T > 34^\circ\text{C}$ ;

$M_2$  = coeficiente que reflete a poluição aparente; = 1 se sólidos sedimentáveis, 0,1 mL/L; = 0,5 se sólidos sedimentáveis  $> 0,1$  mL/L.

Brown et al., (1970) apresentaram um índice bastante similar em sua estrutura ao Índice de Horton, consistindo também em um somatório ponderado de subíndices e pesos. Esses pesquisadores contaram com o apoio da *National Sanitation Foundation Institution* (NSF) e, por este motivo, esse índice ficou conhecido como o IQA-NSF. A construção do índice foi baseada no Método Delphi de construção de consenso a partir da consulta a 142 especialistas, no caso sobre que variáveis de qualidade de água deveriam entrar no cálculo do IQA, o peso relativo das mesmas e uma “nota” que representasse a condição de qualidade da água em função do valor com que cada variável se apresente na água. Inicialmente, foram selecionadas 35 variáveis indicadoras de qualidade da água e, destas, nove foram selecionadas para compor o IQA-NSF; para

cada variável, além do peso relativo á sua importância, foi estabelecida uma curva de variação da qualidade da água, de acordo com sua concentração ou medida. (Figura 1).

Em sua formulação original, o IQA-NSF é uma combinação linear (somatória) dos produtos entre cada subíndice e seu respectivo peso, conforme a Equação 2. Ao final, o IQA é expresso por uma nota em escala de 0 a 100, à qual é associada uma valoração de qualidade da água, por exemplo, de “ótima” à “péssima” (Tabela 1). Entretanto, de acordo com Ott (1978), *apud* Leite e Fonseca (1994), a formulação aditiva apresenta um problema matemático denominado de eclipsamento, que ocorre quando pelo menos uma variável apresenta um valor muito ruim, mas o IQA não reflete esta situação, ocasionando com isso a superestimação da qualidade da água. Para resolver esse problema de resultados mascarados, Landwehr e Deininger (1976) propuseram a forma multiplicativa, onde os pesos tornam-se potência dos subíndices; com isso, quando qualquer dos subíndices se aproxima de zero, o IQA também se aproximará. Assim, o IQA é calculado pelo produtório ponderado das notas de qualidades ( $q_i$ ) das variáveis que integram o índice, elevadas ao peso ( $w_i$ ), conforme descrito na Equação 2. O índice, final, é resultante do somatório individual de cada variável associada ao peso relativo desse, refletindo sua importância na variabilidade total (Equação 3).

$$IQA = \prod_{i=1}^N q_i^{w_i} \quad (2)$$

$$IQA = \sum_{i=1}^n q_i^{w_i} \quad (3)$$

Onde:

IQA: Índice de Qualidade da Água (um número entre 0 e 100);

$q_i$ : qualidade da  $i$ -ésima variável extraída da curva média de variação de qualidade, em função da concentração ou medida (Figura 1).

$w_i$ : peso correspondente a  $i$ -ésima variável, atribuído em função da importância de cada variável na composição do índice, varia de 0 e 1 (Equação 4).

$$\sum_{i=1}^n w_i = 1 \quad (4)$$

Onde:

$n$ : número de variáveis que entram no cálculo do IQA.

O IQA-NSF foi originalmente concebido para avaliar a qualidade geral de águas superficiais, ou seja, efeitos da poluição nos cursos d'água (OTT, 1978), sendo composto pelas seguintes variáveis e respectivos pesos: OD (oxigênio dissolvido) (17%), coliformes fecais (15%), pH (12%), DBO (demanda bioquímica de oxigênio) (10%), nitrato (10%), fosfato (10%), temperatura (10%), turbidez (8%) e sólidos totais (8%). Posteriormente, assumiu a função mais específica de indicador de qualidade da água tendo em vista a tratabilidade para abastecimento para consumo humano e, no Brasil, foi adaptado pela Cetesb, com nitrato e fosfato sendo substituídos, respectivamente, por nitrogênio total e fósforo total, mantendo-se os mesmos pesos ( $w = 0,10$ ) e curvas de qualidade estabelecidos no IQA-NSF (CETESB, 2008).

Na Figura 1 são apresentadas as curvas médias de qualidade das variáveis que compõem o IQA e os respectivos pesos relativos (indicados no canto superior direito de cada gráfico). No Quadro 1 são apresentados os critérios de classificação da água com base no IQA-NSF e no IQA CETESB, com indicação dos estados brasileiros que adotam cada um desses critérios (ANA, 2014).

Quadro 1- Classificação da água bruta com base no IQA

Faixas do IQA-NSF, utilizadas nos seguintes estados: AL, MG, MT, PR, RJ, RN, RS	Faixas de IQA CETESB, utilizadas nos seguintes estados: BA, CE, ES, GO, MS, PB, PE, SP	Avaliação da Qualidade da Água
91-100	80-100	Ótima
71-90	52-79	Boa
51-70	37-51	Regular
26-50	20-36	Ruim
0-25	0-19	Péssima

Fonte: Adaptado de ANA (2014)

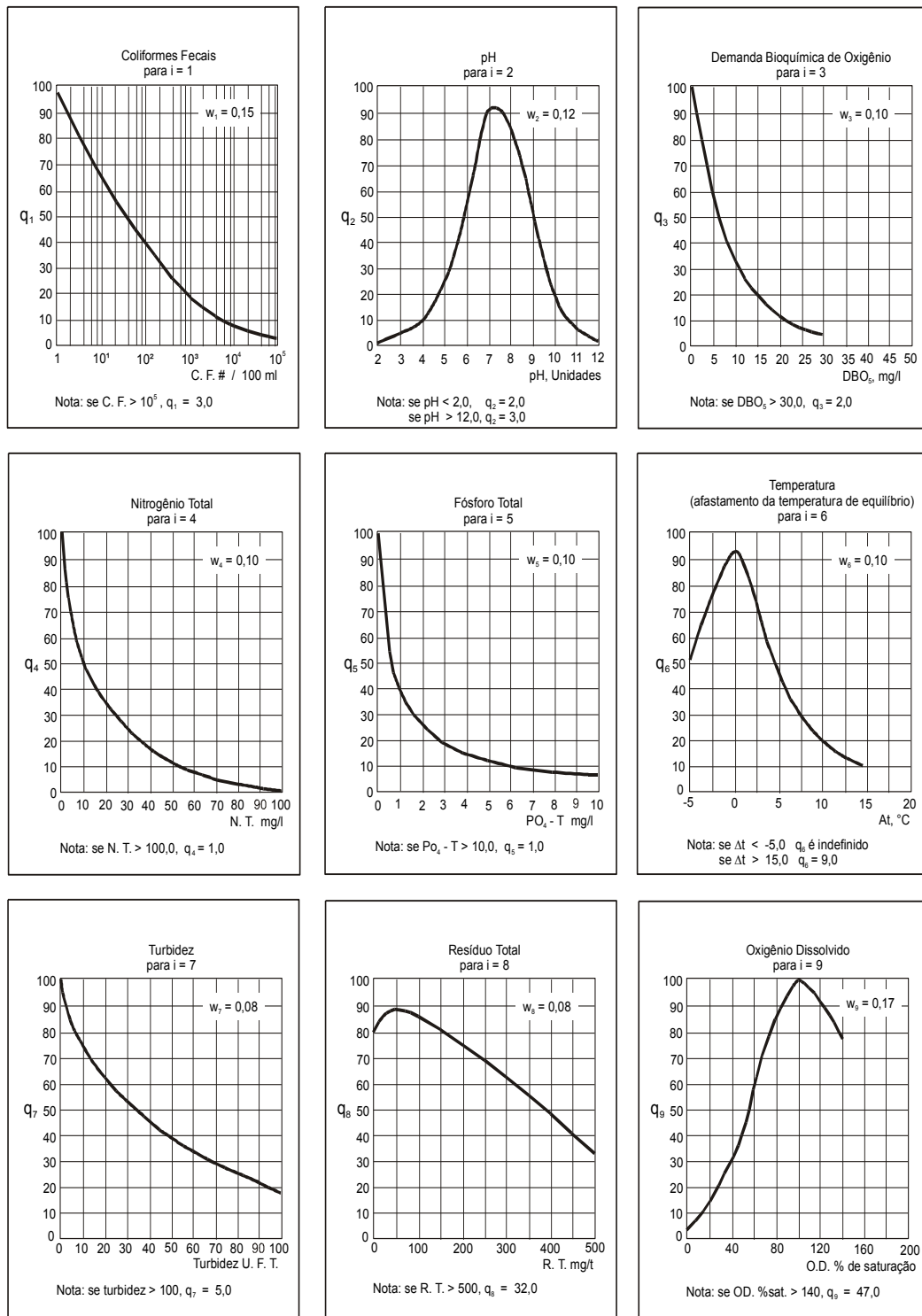


Figura 1- Curvas de variação de qualidade água das variáveis do IQA.

Fonte: CETESB (2008)

Desde sua formulação, o IQA-NSF tem tido ampla aplicação em várias regiões do mundo. No Brasil, o interesse pelos índices de qualidade da água aumentou a partir do início da década de 1970, quando o Conselho Nacional de Meio Ambiente em seu relatório anual manifestou a necessidade da utilização desses recursos em programas de gerenciamento ambiental (MEIRELES, 2007). A partir de 1975 o IQA começou a ser utilizado pela Cetesb, com as modificações anteriormente referidas. Nas décadas seguintes, outros estados brasileiros adotaram o IQA (em sua formulação original ou com as adaptações da Cetesb – ver Quadro 1), sendo hoje o principal índice de qualidade da água utilizado no país (ANA, 2014).

Também como já referido, o IQA-NSF foi originalmente concebido para avaliar a qualidade geral de águas superficiais e, como destacado pela Cetesb (2013), as variáveis que fazem parte do cálculo desse índice refletem, principalmente, a poluição dos corpos hídricos ocasionada pelo lançamento de esgotos domésticos, ou mesmo por contribuições de efluentes industriais, desde que sejam de natureza orgânica biodegradável. Por sinal, segundo PNMA (2008), essa foi exatamente uma das motivações para as adaptações promovidas pela IQA-Cetesb, pois sendo os rios do estado de São Paulo fortemente impactados por esgotos domésticos, é de se esperar a presença e de outras formas de nitrogênio (ex.: nitrogênio orgânico e amoniacal) e fósforo (ex.: fósforo orgânico), além de nitratos e fosfatos. Entretanto, como destacado por Souza e Libânio (2009), como qualquer índice, o IQA apresenta limitações, como, por exemplo, a sobreposição de variáveis que fornecem informações semelhantes, tais como oxigênio OD e DBO, turbidez e sólidos totais.

Posteriormente, tal como incorporado pela Cetesb em 1975, o IQA passou a ter como “determinante principal a utilização para abastecimento público” (CETESB, 2008). Nesse caso, as limitações se vêm acentuadas, pois o IQA não leva em consideração variáveis relevantes para o consumo humano, tais como: substâncias e propriedades potencialmente tóxicas (ex: metais pesados, compostos orgânicos com potencial mutagênico e potencial de formação de trihalometanos), cianobactérias (e cianotoxinas), substâncias que afetam as propriedades organolépticas da água, além de organismos patogênicos (ANA, 2014; CETESB, 2013). Dito isso, há que se considerar que o IQA-NSF vinha acompanhado do Índice de Toxidez (IT), também proposto por Brown *et al.* (1970) e utilizado pela Cetesb desde . O IT assume valor igual a 0 ou 1, indicando, respectivamente, a presença de substâncias tóxicas acima do limite máximo

admissível ou a ausência destas. O IT deve ser multiplicado pelo IQA, confirmando seu valor (IT=1), ou anulando-o (IT=0).

Tendo essas limitações em conta, a literatura reporta várias iniciativas de adaptar o IQA, incorporando alguns desses aspectos, como, por exemplo: (i) Índice Qualidade de Água Bruta (IQAB), proposto por Souza e Libânio (2009), composto pelas variáveis turbidez, cor verdadeira, cianobactérias, algas, *E.coli*, ferro e manganês, formulado também de acordo com a técnica de consulta Delphi para a seleção de variáveis a comporem o índice, mas com as respectivas curvas de qualidade arbitradas pelos autores; (ii) Índice de Qualidade de Água Bruta para Abastecimento Público (IQABP), proposto por Silva et al., (2012), criado a partir da inclusão da variável densidade de cianobactérias na fórmula do IQA – CETESB, utilizando as faixas de números de células de cianobactérias, e a sua respectiva taxação, usadas no cálculo do Índice de Substâncias Tóxicas e Organolépticas (ISTO) – CETESB (descrito a seguir).

No entanto, a mais bem consolidada iniciativa de contornar limitações do IQA é da própria Cetesb, que em 1998 propôs e em 2002 adotou o Índice de Qualidade de Água Bruta para Fins de Abastecimento Público (IAP), composto pelo próprio IQA e pelo ISTO (Índice de Substâncias Tóxicas e Organolépticas) (CETESB, 2008). Assim, em 2002, o IQA, isoladamente, retoma seu sentido original de indicador geral sobre as condições de qualidade das águas superficiais (mais particularmente de impactos do lançamento de cargas orgânicas), mas, juntamente com o ISTO, compõe um novo indicador mais específico de tratabilidade da água para consumo humano (CETESB, 2013).

Como relatado em Cetesb (2013), o ISTO é composto por um grupo de variáveis que indicam a presença de substâncias tóxicas, - (potencial de formação de trihalometanos - PFTHM, número de células de cianobactérias, cádmio, chumbo, cromo total, mercúrio e níquel) e, por outro, de variáveis que afetam a qualidade organoléptica da água (ferro, manganês, alumínio, cobre e zinco), isto é, estão relacionadas à aceitação da água para consumo humano, não a problemas de saúde.

Para cada variável incluída no ISTO foram estabelecidas curvas de qualidade que atribuem ponderações variando de 0 a 1. As curvas de PFTHM e de metais foram construídas utilizando-se dois níveis de qualidade ( $q_i$ ), que associam os valores numéricos 1 e 0,5, respectivamente, ao limite inferior (LI) e ao limite superior (LS), conforme apresentado na Figura 2. As faixas de variação de qualidade ( $q_i$ ) refletem as

seguintes condições de qualidade da água bruta destinada ao abastecimento para consumo humano:

- Valor medido  $\leq$  LI: águas adequadas para o consumo humano; atendem ao padrão de potabilidade estabelecido na Portaria N° 2914 / 2011 do Ministério da Saúde (BRASIL, 2011) em relação às variáveis analisadas.
- $LI < \text{Valor medido} \leq LS$ : águas adequadas para tratamento convencional ou avançado; atendem aos padrões de qualidade da classe 3 da Resolução CONAMA N° 357 / 2005 (BRASIL, 2005) em relação às variáveis determinadas.
- Valor medido  $> LS$ : águas que não devem ser submetidas apenas a tratamento convencional; não atendem aos padrões de qualidade da classe 3 da Resolução CONAMA N° 357/2005 em relação às variáveis avaliadas

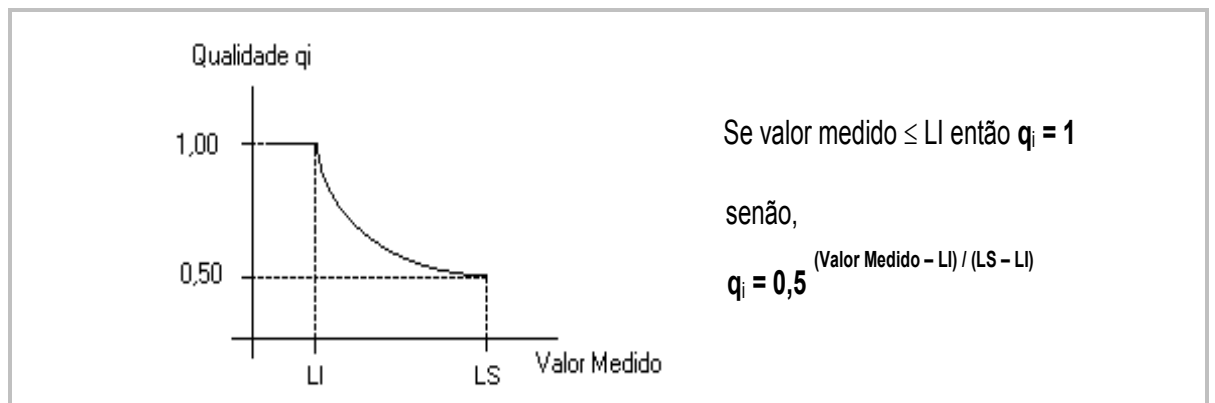


Figura 2 - Curva de qualidade padrão para as variáveis incluídas no ISTO.

Os valores constantes no padrão de potabilidade estabelecido pela Portaria MS N° 2914 /2011 foram considerados como sendo o limite inferior e os padrões de qualidade de água doce da Classe 3 da Resolução CONAMA N° 357 / 2005 como o limite superior (CETESB, 2013).

No caso do PFTHM, foi estabelecida uma equação de regressão linear entre essa variável na água bruta e trihalometanos na água tratada; para isso foram utilizados valores médios de séries históricas de ambas as variáveis em mananciais da Região Metropolitana de São Paulo. O limite superior do PFTHM foi estimado para a concentração de THM estabelecida para água de consumo humano pela Portaria MS N° 2914/2011 (100 µg/L), enquanto que o limite inferior foi estimado a partir do nível de THM estabelecido na legislação norte americana (80µg/L) (CETESB, 2013) (Quadro 2)

Quadro 2 - Limites inferiores e superiores para metais e potencial de formação de trihalometanos para cálculo do ISTO

Grupo	Variáveis	Unidade	Limite Inferior	Limite Superior
Tóxicos	Cádmio	mg/L	0,005	0,01
	Chumbo	mg/L	0,01	0,033
	Cromo Total	mg/L	0,05	0,059
	Níquel	mg/L	0,02	0,025
	Mercúrio	mg/L	0,001	0,002
	PFTHM	µg/L	373	461
Organolépticos	Alumínio Dissolvido	mg/L	0,2	2
	Cobre Dissolvido	mg/L	2	8
	Ferro Dissolvido	mg/L	0,3	5
	Manganês	mg/L	0,1	5
	Zinco	mg/L	5	5,9

Ainda de acordo com Cetesb (2013), devido à importância da presença de cianobactérias e do potencial de liberação de cianotoxinas para a avaliação da qualidade da água para consumo humano, em 2005 a contagem do número de células de cianobactérias passou a fazer parte do ISTO. A taxa associada ao número de células de cianobactérias (Quadro 3), foi baseada em dispositivos da Portaria N° 2914/2011 e da Resolução Conama N° 357 / 2005 em dados existentes da rede de monitoramento da

CETESB A Portaria N° 2914/2011 determina o monitoramento de cianotoxinas no manancial, no ponto de captação, apenas quando a densidade de cianobactérias for superior a 20.000 células /mL, com frequência semanal. Por sua vez, a Resolução Conama N° 357 / 2005, estabelece limites de 20.000 células / mL para águas de classe 1 (que podem ser destinadas ao abastecimento para consumo humano, após tratamento simplificado - clarificação por meio de filtração e desinfecção e correção de pH quando necessário) e de classe 2 (águas que podem ser destinadas ao abastecimento para consumo humano, após tratamento convencional - clarificação com utilização de coagulação e floculação, seguida de desinfecção e correção de pH), e de 100.000 células / mL para águas de classe 3 (águas que podem ser destinadas ao abastecimento para consumo humano, após tratamento convencional ou avançado - técnicas de remoção e/ou inativação de constituintes refratários aos processos convencionais de tratamento, os quais podem conferir à água características, tais como: cor, odor, sabor, atividade tóxica ou patogênica).

Nos pontos de amostragem, situados em ambientes lênticos e utilizados para abastecimento público, o número de células de cianobactérias é variável obrigatória para o cálculo do IAP; nos demais pontos é opcional (CETESB, 2013).

Quadro 3- Faixas de número de células de cianobactérias e respectiva taxaço para cálculo do ISTO.

N° de células por mL (N)	Taxação (q <sub>NCC</sub> )
$N \leq 20.000$	1.00
$20.000 < N \leq 50.000$	0.80
$50.000 < N \leq 100.000$	0.70
$100.000 < N \leq 200.000$	0.60
$200.000 < N \leq 500.000$	0.50
$N > 500.000$	0.35

O componente de substâncias tóxicas (ST) do ISTO é obtido pela Equação 5, ou seja, da multiplicação dos dois valores mínimos mais críticos do grupo de variáveis que indicam a presença dessas substâncias na água:

$$ST = \text{Mín-1} (q_{\text{THMFP}}; q_{\text{Cd}}; q_{\text{Cr}}; q_{\text{Pb}}; q_{\text{Ni}}; q_{\text{Hg}}; q_{\text{NCC}}) \times \text{Mín-2} (q_{\text{THMFP}}; q_{\text{Cd}}; q_{\text{Cr}}; q_{\text{Pb}}; q_{\text{Ni}}; q_{\text{Hg}}; q_{\text{NCC}}) \quad (5)$$

O componente relativo ao grupo de substâncias que afetam a qualidade organoléptica da água (SO) pode ser obtido por meio da Equação 6, ou seja, a média aritmética dos valores de qualidade padronizados ( $q_i$ ) desse grupo

$$SO = \text{Média Aritmética} (q_{\text{Al}}; q_{\text{Cu}}; q_{\text{Zn}}; q_{\text{Fe}}; q_{\text{Mn}}) \quad (6)$$

O ISTO é resultado do produto dos grupos de substâncias tóxicas e as que alteram a qualidade organoléptica da água, como descrito na Equação 7. Por fim, o IAP é obtido pelo produto do IQA pelo ISTO (Equação 8) e a classificação do IAP é feita tal qual para o IQA CETESB (Quadro 1).

$$ISTO = ST \times SO \quad (7)$$

$$IAP = IQA \times ISTO \quad (8)$$

Embora o ISTO seja considerado um avanço importante como índice qualidade (tratabilidade) da água bruta para abastecimento para consumo humano, outra limitação do IQA permanece e é trazida para o ISTO: a determinação de todas as nove variáveis que compõem o IQA é indispensável, pois em sua formulação matemática a falta de qualquer variável inviabiliza o cálculo do IQA e, por conseguinte do IAP (mesmo que isso não se aplique ao ISTO).

Com o intuito de selecionar variáveis mais representativas autores como Toledo e Nicolella, 2002; Meireles, 2007; Coletti et al, 2010; Al-Badaii (2013), utilizaram técnicas de análise multivariada de dados na extração de informações partindo de um conjunto grande de variáveis para um grupo de menor dimensão, com maior representatividade, melhor interpretação dos resultados e sem perda significativa das informações.

No Brasil, técnicas de estatística multivariada vem sendo utilizadas por diversos pesquisadores no desenvolvimento de índices da qualidade da água. Haase et al., (1989); Haase e Possoli (1993); Carvalho et al., (2000), Toledo e Nicollela (2002), Coletti et al. e (2010) Guedes et al. (2012) são alguns exemplos de pesquisas que utilizaram essa técnica.

Outro ponto que motivou o estudo da associação entre uso e ocupação do solo e variáveis da qualidade da água é a importância desse tema para o planejamento e gestão dos recursos hídricos, pois a maioria dos trabalhos como os de Arcova e Cicco (1999); Guedes et al (2012) e Al-Badaii (2013) fazem inferência sobre a influência dos diferentes tipos de uso e ocupação do solo com as variáveis de qualidade da água, sem o estabelecimento de qualquer relação estatística capaz de comprovar essas relações.

## 2.2. Análise de Componentes Principais - ACP

A análise de componentes principais (ACP) é uma das diversas técnicas da estatística multivariada, mais especificamente de análise fatorial, que permite que um número grande de variáveis observadas sejam agrupadas em fatores (combinações lineares de variáveis correlacionadas entre si), ou seja, condensa a informação contida nas variáveis em um pequeno grupo de novas dimensões que resumem ou explicam o conjunto de variáveis observadas com o mínimo de perda de informação (HAIR et al., 2009).

A ACP consiste, essencialmente, em reescrever as coordenadas das amostras em outro sistema de eixos mais conveniente para a análise dos dados. As combinações lineares entre as variáveis originais, nesse caso, são chamadas componentes principais. A análise agrupa os indivíduos de acordo com sua variação, isto é, os indivíduos são agrupados segundo suas variâncias, ou seja, segundo seu comportamento dentro da população. O primeiro fator tende a ser o que explica a maior parte da variância dos dados; o segundo fator e os seguintes são então baseados na quantia residual da variância, explicando porções sucessivamente menores da variância (HAIR et al., 2009).

Como sintetizado por Figueiredo Filho e Silva Júnior (2010), cabe aqui destacar a diferença entre análise de componentes principais (ACP) e análise fatorial (AF): ambas as técnicas procuram produzir combinações lineares de variáveis que capturem o máximo possível a variância das variáveis observadas, mas enquanto na ACP *toda* a variância é utilizada, na AF apenas a variância *compartilhada*. Ainda segundo esses autores, citando Garson (2009), “a análise de componentes principais é em geral preferida para fins de redução de dados (traduzindo o espaço das variáveis num espaço ótimo de fatores), enquanto a análise fatorial é em geral preferida quando o objetivo da pesquisa é detectar a estrutura dos dados ou a modelagem causal”. Porém, de acordo Hair *et al.* (2009), na maioria dos casos, tanto a ACP quanto a AF chegam aos mesmos resultados se o número de variáveis superar 30 ou se as comunalidades excederem 0,60 para a maior parte das variáveis (as comunalidades representam a proporção da variância para cada variável incluída na análise que é explicada pelos componentes extraídos - Schawb (2012))

Cabe aqui também destacar outra diferenciação feita em Figueiredo Filho e Silva Júnior (2010), entre (i) análise fatorial exploratória (AFE), geralmente utilizada nos estágios mais embrionários da pesquisa, procurando explorar a relação entre um

conjunto de variáveis, identificando padrões de correlação, ou ainda para criar variáveis independentes ou dependentes que podem ser utilizadas posteriormente em modelos de regressão; e (ii), a análise fatorial confirmatória (AFC), utilizada para testar hipóteses.

A ACP demanda três etapas, as quais embutem requisitos que precisam ser satisfeitos para a utilização da análise: (i) a preparação e análise da matriz de correlação; (ii) a extração dos fatores comuns e a possível redução de variáveis explicativas; e (iii) a rotação dos eixos relativos aos fatores comuns, visando uma solução mais simples e facilmente interpretável.

Na primeira etapa, como em outras técnicas, é preciso inicialmente verificar a adequabilidade da base de dados, por exemplo, tamanho da amostra, distribuição dos dados e verificação de existência de correlações.

Em relação ao número de casos, Hair et al. (2009) sugerem que a amostra deve ser superior a 50 observações e ainda que a razão entre o número de casos e a quantidade de variáveis deve exceder cinco para um.

No que diz respeito ao padrão de correlação entre as variáveis, a matriz de correlações deve exibir a maior parte dos coeficientes com valor acima de 0,30. A adequabilidade do conjunto de dados e das variáveis pode ser verificada por meio do teste de Kayser-Mayer-Olkin (KMO), comparando-se a magnitude dos coeficientes de correlação observados com a magnitude dos coeficientes de correlações parciais (NORUSIS, 1990), conforme descrito na Equação 9.

$$KMO = \frac{\sum \sum_{k \neq j} r_{kj}^2}{\sum \sum_{k \neq j} r_{kj}^2 + \sum \sum_{k \neq j} a_{kj}^2} \quad (9)$$

onde:

$r_{kj}$ : Coeficiente de correlação simples entre as variáveis k e j.

$a_{kj}$ : coeficientes de correlação parcial entre as variáveis k e j.

Se a soma dos coeficientes de correlação parciais ao quadrado entre os pares de variáveis for pequena comparada à soma dos coeficientes de correlação simples ao quadrado, as medidas de KMO serão próximas a um, indicando que não há restrição ao uso do modelo da análise dos fatores. Valores baixos de KMO indicam que a AF/ACP não pode ser aplicada ao conjunto de dados (NORUSIS, 1990; HAIR et al., 2009). Pallant (2007) sugere 0,6 como um limite razoável. Friel (2009) sugere a seguinte escala para interpretar o valor da estatística KMO: entre 0,90 e 1 excelente; entre 0,80 e 0,89 bom; entre 0,70 e 0,79 mediano; entre 0,60 e 0,69 medíocre; entre 0,50 e 0,59 ruim e entre 0 e 0,49 inadequado. Já Hair et al. (2009) sugerem 0,50 como patamar aceitável. O Quadro 4 apresenta as faixas de aceitação propostas por para aplicação ao conjunto de dados da ACP.

Quadro 4- Faixas de validade da aplicação da AF/ACP pelo teste de KMO.

Valor de KMO	Aplicabilidade da AF/ACP
$KMO \geq 0,9$	Excelente
$0,8 \leq KMO < 0,9$	Ótima
$0,7 \leq KMO < 0,8$	Boa
$0,6 \leq KMO < 0,7$	Regular
$0,5 \leq KMO < 0,6$	Minimante Aceitável
$0,5 < KMO$	Inaceitável

Adicionalmente, o teste de esfericidade de Bartlett indica se a matriz de correlação é uma matriz identidade e, caso positivo, mostra que suas variáveis são desassociadas, ou seja, não apresentam qualquer relação e a análise fatorial não pode ser aplicada. A estatística *Bartlett Test of Sphericity* (BTS) deve ser estatisticamente significativa ( $p < 0,05$ ), indicando a existência de correlações suficientes entre as variáveis para realização da análise (HAIR et al., 2009).

O Quadro 5, resume esses requerimentos da primeira etapa do planejamento da AF (Figueiredo Filho e Silva Júnior, 2010).

Quadro 5- Síntese do primeiro estágio do planejamento da AF.

Nível de mensuração	Variáveis contínuas e discretas
Amostra	Amostras mínimas entre 50 e 100 casos; razão entre o número de observações e a quantidade de variáveis igual ou superior a cinco.
Correlação	Maior parte dos coeficientes de correlação deve apresentar valores acima de 0,30.
KMO	Quanto maior melhor, tendo 0,50 como o patamar mínimo de adequabilidade (Hair et al 2006)
BTS	$p < 0,05$

Na sequência se deve determinar a técnica de extração dos fatores: componentes principais há outras: fatores principais, fatoração por imagem; fatoração por verossimilhança máxima; fatoração alfa; mínimos quadrados não ponderados; mínimos quadrados. ACP é a técnica mais utilizada e a diferença principal entre ACP e AF foi destacada anteriormente.

Depois, na definição do número de fatores que serão extraídos, como relatam Figueiredo Filho e Silva Júnior (2010), o pesquisador deve ponderar entre parcimônia e explicação. Quanto mais fatores forem extraídos, menor é o grau de parcimônia, no entanto, maior é a quantidade total de variância pelos fatores. Por outro lado, quanto menos fatores forem extraídos, maior é o grau de parcimônia, todavia, menor será a quantidade total de variância carregada pelos fatores. Dessa forma, na visão desses autores, a solução ótima é identificar o número mínimo de fatores que maximiza a quantidade de variância total explicada. De acordo com Hair et al. (2009), a variância acumulada dos dados no conjunto de fatores deve ser no mínimo de 60% .

Por fim, como apresentado por Figueiredo Filho e Silva Júnior (2010), o pesquisador deve decidir o tipo de rotação dos fatores, que se refere ao método matemático que rotaciona os eixos no espaço geométrico. Sobre os dois principais tipos de rotação (ortogonal - *Varimax*, *Quartimax*, *Equamax* e oblíqua - *direct oblimin*, *Promax*), esses autores explicam que, de acordo com Tabachnick e Fidell (2007), as rotações ortogonais são mais fáceis de reportar e de interpretar, porém os fatores devem ser independentes (pressuposto que, segundo os referidos autores, na prática, é mais difícil de ser respeitado); por sua vez, as rotações oblíquas permitem que os fatores sejam correlacionados, mas são mais difíceis de descrever e interpretar. Ainda de acordo com Figueiredo Filho e Silva Júnior (2010), citando Hair et al. (2009) e Pallant (2007), em geral, essas duas formas de rotação produzem resultados bastante semelhantes,

principalmente quando o padrão de correlação entre as variáveis utilizadas é claro; citam ainda que o tipo de rotação ortogonal *Varimax* é o mais comumente utilizado e que esse método procura minimizar o número de variáveis que apresentam altas cargas em cada fator.

Segundo Toledo e Nicolella (2002) após análise dos resultados da análise fatorial é possível a construção de um ou mais fatores de escala (escores fatoriais) a partir da constatação da existência de determinadas dimensões subjacentes aos dados. O interesse é examinar os dados amostrais em termos dessas dimensões, ao invés de examinarmos cada variável separadamente. Foi utilizado o método de Bartlett para a criação dos fatores de escala, que minimiza a variância do erro, sendo os escores preditos pelo modelo considerados como índice de qualidade de água.

Feitas essas breves considerações sobre a fundamentação teórica da AF / ACP, a seguir são citados alguns estudos que utilizaram essa técnica no desenvolvimento de índices da qualidade da água e, ou na investigação dos principais impactos de poluição, e respectivas fontes, em cursos d'água.

Haase et al. (1989) avaliaram a viabilidade do uso da AF como um instrumento para interpretação da qualidade da água na bacia do Guaíba - R. Foram selecionados 24 pontos de amostragens, com a análise de 40 variáveis; oito foram selecionadas para compor o índice da qualidade da água - IQA. Os resultados mostraram que a aplicação foi consistente e representativa do ambiente em estudo.

Haase e Possoli (1993) utilizaram AF na elaboração de um IQA, avaliando seu emprego no Rio Grande do Sul, em dois regimes hidrólogos diferentes (Arroio Velhaco e Rio Gravataí). Foram utilizadas seis variáveis relacionadas à qualidade da água: pH, condutividade elétrica (CE), saturação de OD, DBO, DQO e alcalinidade. Foram obtidos dois fatores que explicaram 79,1% da variância total das variáveis originais. O fator 1, como componente que represente a parte comum mais importante das variáveis analisadas, foi denominado IQA, explicando 50,4% da da variância total das variáveis originais. A comparação da variação do IQA ao longo dos cursos d'água com uma classificação segundo o déficit de OD e com um critério de zoneamento ecológico baseado em características morfológicas, físicas, químicas e biológicas mostrou congruência nos resultados.

Toledo e Nicolella (2002) usaram a técnica de AF e o método de Bartlett com o propósito de desenvolvimento de um IQA para a bacia hidrográfica do ribeirão Jardim,

Guaíra - SP. Dez variáveis de qualidade da água foram analisadas: fósforo total e dissolvido, pH, OD, amônia, nitrato, CE, turbidez, sólidos totais (ST) em suspensão (SS) e clorofila. OD, fósforo, amônia e CE foram que mais contribuíram para a composição do índice. Os três primeiros fatores explicaram 71% da variância dos dados. O primeiro fator da ACP explicou 45% da variância total dos dados e foi utilizado na construção do IQA; o segundo 15% e o terceiro 12%.

Coletti et al. (2010) utilizaram ACP para a formulação de um IQA na bacia do rio das Pedras em Mogi Guaçu - SP, de uso predominantemente agrícola. Foram analisadas oito variáveis (nitrato, nitrogênio amoniacal, fósforo total, condutividade elétrica, pH e turbidez). Os três primeiros fatores da ACP explicaram 72% da variância total dos dados. O primeiro, com 38%, foi adotado como IQA, sendo que pela análise fatorial, amônia, nitrogênio amoniacal e nitrato foram as variáveis que mais contribuíram na determinação do IQA. Os autores consideraram que o IQA proporcionou boa caracterização da qualidade da água ao longo da bacia do rio das Pedras.

Guedes et al. (2012) utilizaram ACP para avaliar a qualidade da água na bacia do médio Rio Pomba, na Zona da Mata de Minas Gerais. Foram realizados dois monitoramentos, o primeiro em outubro de 2008 e outro em janeiro de 2009, com nove pontos de amostragem. Dezoito variáveis foram analisadas. Os três primeiros componentes explicaram 74,3% da variância total dos dados. As variações na qualidade da água no médio Rio Pomba foram definidas por um grupo de nutrientes, associado à influência de fontes de poluição difusa, como drenagem das áreas agrícolas e esgoto doméstico; por um grupo orgânico (OD, DBO e DQO), causado pelo lançamento de esgoto doméstico no curso de água e por um grupo de sólidos em suspensão, expressando o processo de erosão hídrica na bacia.

Netto et al. (2013), em estudo na bacia do Rio Grande - RJ compararam o IQA CETESB com o IQA calculado com base em ACP e no método de Bartlett. Foram monitorados onze pontos trimestralmente, durante dois anos, com determinação das variáveis necessárias ao cálculo do IQA-CETESB: pH, temperatura, OD, fósforo total, nitrogênio total, turbidez, coliformes termotolerantes, DBO, resíduo total, tendo sido obtidos 68 resultados para cada variável. Foram obtidos quatro fatores que explicam 28,7% da variância total das variáveis originais. O fator 4, com menor variabilidade dos dados (11,99%), representou a parte comum mais abrangente das variáveis observadas, e foi associado ao índice de qualidade de água (IQA). Adotou-se o procedimento “no rotate”, já

que outros procedimentos de rotação dos eixos (varimax, equimax e quartimax) não produziram resultados adequados para estarem associados a um índice de qualidade água. Os autores concluíram os IQAs calculados pelo método estatístico Bartlett e pelo modelo proposto pela CETESB apresentaram coerência nos valores numéricos encontrados (respectivamente 66,9 e 65,8).

Na Turquia, Akbal et al. (2011) avaliaram a qualidade da água na costa do Mar Negro,. Seleccionaram dez pontos de amostragens, de água doce (rios) e do mar,. e vinte e cinco variáveis foram analisadas: carbono total, carbono inorgânico total, carbono orgânico total, cromo, cádmio, cobre, chumbo, ferro, níquel, manganês, fenóis, surfactantes, amônia, nitratos, nitritos, fósforo total, halogênicos orgânicos absorvíveis (AOX), sulfato, dureza, OD, pH, temperatura, sólidos totais dissolvidos, CE e salinidade. Técnicas de estatística multivariada – análise de cluster e AF/ ACP – foram aplicadas na investigação das fontes de aportes dos poluentes. Os resultados indicaram que 82,24% da variância da água do rio e 89,65% da variância total da água do mar foram explicados por sete fatores. Carbono orgânico total, nitrato, nitrito, Cu, Pb, Fe, Mn e Cr, sulfato, dureza e sólidos dissolvidos totais, salinidade e temperatura foram as variáveis que mais contribuíram para o modelo de qualidade da água dos rios e fenóis, amônia ionizada, nitrato, nitrito, Cd, Pb, Cr, Ni, Cu e Mn, dureza, condutividade elétrica, sólidos dissolvidos totais, salinidade, oxigênio dissolvido e temperatura foram as mais relevantes para a água do mar.

Na Espanha, Mendiguchía *et al.* (2004) estudaram os efeitos causados por atividades antrópicas sobre a qualidade da água do rio Guadalquivir.. Os resultados de análise de cluster e AF / ACP, permitiram a identificação de quatro zonas de qualidade da água no rio, as quais foram associadas ao uso e ocupação do solo. Três fatores explicaram cerca de 80% da variância dos dados. O primeiro (47% da variância) foi associado a nitrato, amônia e manganês; o segundo componente (22,5% da variância) aos sólidos em suspensão e fosfato e o terceiro (9,7% da variância) à nitrato e cobre.

### **2.3. Redes neurais artificiais - RNA**

As redes neurais artificiais (RNAs) constituem uma ferramenta computacional que vem sendo utilizada em diversas áreas do conhecimento, na modelagem de problemas de predição, classificação e controle de dados (HAYKIN, 2001).

São compostas por unidades de processamento denominadas de neurônios, conectadas por sinapses, associados aos pesos que representam a importância de cada sinal transmitido. Diferem de modelos computacionais tradicionais por apresentarem adaptabilidade ou aprendizagem, capacidade de generalização, agrupamento, organização de dados e tolerância a falhas (RABUNAL e DORRADO, 2006).

De acordo com Wasserman (1989) podem ser treinadas pela exposição repetida de um conjunto de dados, reconhecendo características comuns entre eles e agrupando-os ordenadamente. Também podem ser programadas para associar um conjunto de dados de entrada e suas respectivas saídas. Elas podem generalizar através de exemplos imperfeitos (ruídos) e extrair informações essenciais das entradas contendo tanto dados relevantes como irrelevantes.

Assemelha-se ao cérebro humano pela arquitetura, operações, conhecimentos adquiridos no processo de aprendizagem e por forças de conexão entre neurônios, conhecidas como pesos sinápticos, que são utilizadas para armazenar o conhecimento adquirido (HAYKIN, 2001).

Quando ativado, o neurônio dispara um sinal eletroquímico ao longo do axônio, na forma de um pulso. Esse sinal cruza as sinapses para outros neurônios, que podem responder ao sinal. Um neurônio só dispara um sinal se o sinal recebido pelos dendritos exceder um certo nível, denominado limiar de disparo ou limiar de excitação. Se o estímulo recebido estiver abaixo do limiar de excitação, o neurônio não emitirá um impulso nervoso do estímulo. Por outro lado, se o limiar for ultrapassado, o potencial de ação do neurônio será sempre o mesmo, qualquer que seja a intensidade do estímulo (HEBB, 2002).

A força do sinal recebido por um neurônio depende da eficácia das sinapses, que são uma lacuna, entre duas ramificações de neurônios distintos, preenchida com um fluido. O sinal não pode ultrapassar eletricamente esse espaço. Assim, substâncias químicas especiais, chamadas neurotransmissores, desempenham esse papel. Elas são liberadas pela membrana emissora pré-sináptica e se difundem através do espaço para

os receptores da membrana do neurônio receptor pós-sináptico (CERQUEIRA et al, 2001).

O neurônio artificial é a unidade de processamento que é fundamental para operação de uma rede neural, formado pelo conjunto de sinapses ou elos de conexão, cada uma caracterizada por um peso ( $w_{k1}, w_{k2}, \dots, w_{kn}$ ). Um sinal  $x_j$  na entrada da sinapse  $j$  conectada ao neurônio  $k$  é multiplicada pelo peso sináptico  $w_{kj}$ . O peso sináptico de um neurônio está em um intervalo que inclui valores positivos e negativos, os primeiros são excitatórios e os segundos inibitórios do neurônio (McCulloch e PITTS, 1943).

Os sinais de entrada são somados e ponderados pelas sinapses dos neurônios. A função de ativação é referida como uma função restritiva e que limita a amplitude do sinal de saída de um neurônio a um intervalo unitário fechado  $[0,1]$  ou alternativo  $[-1,1]$ . O modelo neural inclui o bias  $b_k$  aplicado externamente, que tem efeito de aumentar ou diminuir a entrada da função de ativação (HAYKIN, 2001) (Figura 3).

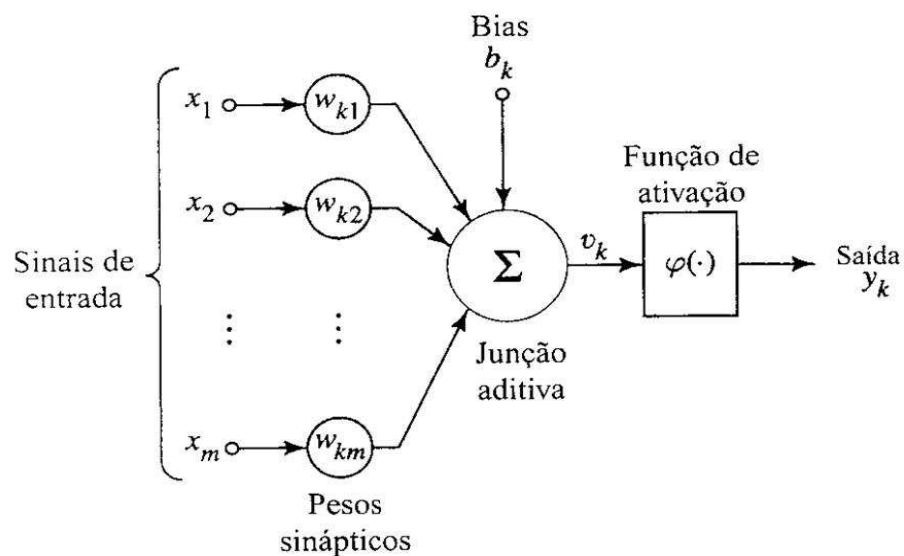


Figura 3 - Representação de um neurônio artificial.

Fonte: Haykin (2001)

A função sigmoidal (Figura 4) é a função de ativação mais utilizada na construção de redes neurais, definida como uma função estritamente crescente que exibe um balanceamento adequado entre o comportamento linear e não linear (HAYKIN, 2001). Tem sido apontada como a mais empregada na predição e previsão de variáveis de recursos hídricos (LINGIREDDY e BRION, 2005).

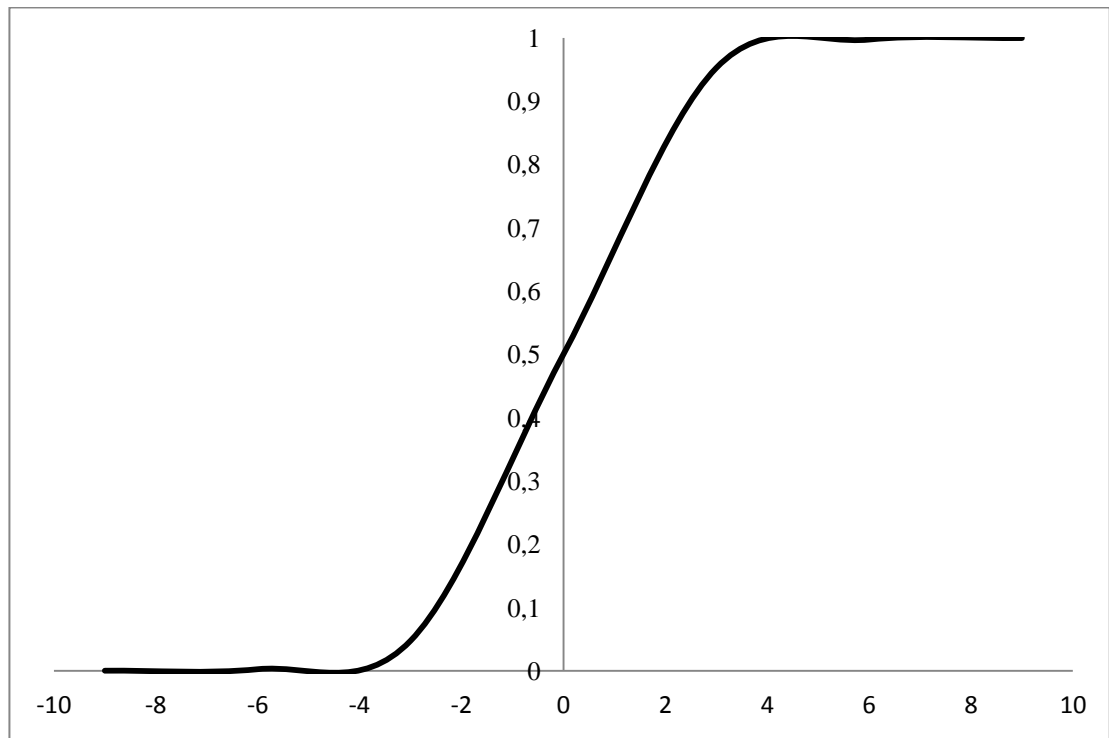


Figura 4 - Função de ativação sigmoidal.

Fonte: Adaptado de Haykin (2001)

Os primeiros estudos sobre redes neurais ocorreram na década de 1940 com o trabalho de McCulloch e Pitts (1943). Esses autores descreveram um cálculo lógico das redes neurais que unificava os estudos de neurofisiologia e da lógica matemática. Em 1949 o psicólogo Donald Hebb escreveu livro a "*The Organization of Behaviour*", no qual foi apresentada pela primeira vez a formulação de uma regra de aprendizagem fisiológica para modificação sináptica.

Em 1958 o psicólogo Frank Rosenblatt introduziu o conceito de aprendizado em redes neurais. O modelo proposto por Rosenblatt ficou conhecido como Perceptron e consistia de uma rede neural formada por uma camada de entrada e outra de saída.

Minsky e Papert (1969) fizeram um estudo desses algoritmos e publicaram o livro *Perceptrons*, onde mostraram que uma rede formada de uma única camada de neurônios, independente do algoritmo de aprendizagem, é capaz da resolução de problemas apenas quando os conjuntos são linearmente separáveis. Esses resultados foram devastadores para o desenvolvimento das RNAs que ficaram praticamente paradas até o início da década de 1980.

O reaparecimento efetivo da computação neural ocorreu em 1982 com a publicação do artigo de Hopfield que tratou pela primeira vez do princípio do armazenamento de informações em redes dinamicamente estáveis (HAYKIN, 2001).

Em 1986 foi relado por Rumelhart, Hinton e Williams o desenvolvimento do algoritmo de retropropagação (backpropagation). No mesmo ano foi publicado o livro *Parallel Distributed Processing: Explorations in the Microstructures of Cognition* editado por Rumelhart e McClelland. Este livro exerceu grande influência na utilização da aprendizagem por retropropagação, que emergiu com o algoritmo mais popular para o treinamento de Perceptrons de Múltiplas camadas (HAYKIN, 2001).

As redes neurais Multilayer Perceptrons (MLP) são estruturadas com uma camada de entrada, uma ou mais camadas ocultas (intermediárias) e uma camada de saída. A introdução de camadas intermediárias forneceu a estrutura necessária para que as redes neurais pudessem aprender diferentes tipos de agrupamentos e possibilidade de sua aplicação em sistemas não lineares. Os neurônios ocultos capacitam a rede a aprender tarefas complexas extraindo progressivamente as características mais significativas dos padrões de entrada. A definição do número de camadas ocultas e de neurônios em cada uma delas depende da complexidade do problema. Estudos apontam que uma MLP com uma única camada oculta é capaz de modelar a maioria dos problemas (CERQUEIRA, 2001).

Um requisito necessário para o uso de redes neurais é saber, ou intuir, que existe uma relação entre os dados de entrada e os de saída. Essa relação não precisa ser exata. Quando se usam redes neurais, o objetivo não é conhecer a natureza exata da relação entre os dados de entrada e os de saída, se essa relação fosse conhecida, seria possível modelá-la diretamente (VERÍSSIMO, 2007).

Os tipos de treinamento mais usados em redes neurais são treinamento supervisionado e treinamento não supervisionado. No aprendizado supervisionado o usuário reúne um conjunto de casos para treinamento, que contêm dados de entrada e suas respectivas saídas, e a rede aprende a prever a relação entre os dois (HAYKIN, 2001). No aprendizado não-supervisionado não é fornecido um valor de saída desejado para cada padrão de entrada, cabendo à rede neural criar uma classe ou grupo representativo das características relevantes da entrada que é fornecida (VERÍSSIMO, 2007).

O algoritmo de treinamento mais utilizado é o backpropagation, conhecido como de retropropagação do erro que foi apresentado por Rumelhart et al. (1986). Tem como característica o ajuste dos pesos e limiares da rede, de forma a minimizar o erro em suas previsões sobre o conjunto de dados de treinamento (HAYKIN, 2001).

O procedimento de aprendizagem utiliza vetores que mapeiam um conjunto de entradas para um conjunto de saídas. O aprendizado é realizado por ajuste dos pesos da rede para minimizar as diferenças entre a saída atual e a saída desejada. A rede é inicializada com pesos aleatórios. Durante o processo de aprendizagem, um vetor de entrada é apresentado para a rede e propagado para determinar os valores da saída. A seguir, o vetor de saída produzido pela rede é comparado com o vetor de saída desejado, resultando num sinal de erro, que é retropropagado através da rede para permitir o ajuste dos pesos. Este processo é repetido até que a rede responda, para cada vetor de entrada, com um vetor de saída com valores suficientemente próximos dos valores desejados (RUMELHART et al., 1986).

No modelo backpropagation o cálculo do vetor do gradiente da superfície de erro aponta na direção do declive mais íngreme de forma que um pequeno movimento nessa direção implica na diminuição do erro. Uma dificuldade é decidir o tamanho dos passos que deve ser dados. Passos largos podem convergir mais rapidamente, mas também podem sobrepasar a solução ou andar na direção errada. Passos muito pequenos podem ir na direção correta, mas requerem um grande número de iterações. O tamanho dos passos é proporcional à inclinação da superfície e a uma constante especial chamada taxa de aprendizagem ( $\eta$ ) (VIEIRA, 2000).

A determinação da taxa de aprendizagem ( $\eta$ ) para um problema específico não é uma tarefa simples, valores típicos estão na faixa de 0,1 a 1,0. Há vários modos sugeridos para ajustar a taxa de aprendizagem. A estratégia habitual é observar os erros

enquanto são atualizados os pesos das conexões. Se a mudança num peso específico não resulta em redução do erro global, então o algoritmo passou o mínimo global e a taxa de aprendizagem ( $\eta$ ) deve ser diminuída. Por outro lado, se mudanças sucessivas para um peso particular reduzem o erro global, então a taxa de aprendizagem ( $\eta$ ) deve ser aumentada (VIEIRA, 2000).

Valores pequenos na taxa de aprendizagem ( $\eta$ ) tornam o treinamento lento para obtenção de um resultado aceitável. Porém, valores grandes colocam a rede em oscilação sem alcançar o mínimo global. Isto normalmente acontece quando a superfície de erro possui uma área muito íngreme e estreita (como um vale). A adição de um impulso o termo momentum ( $\mu$ ) é um modo de apressar a aprendizagem sem introduzir oscilação na rede (VIEIRA, 2000).

O efeito deste melhoramento é que áreas planas na superfície de erro são atravessadas rapidamente com passos grandes, em áreas rugosas, ocorre à redução do tamanho do passo. Essa alteração do tamanho do passo aumenta a velocidade de aprendizagem significativamente (ROJAS, 1996).

As redes neurais possuem uma boa capacidade de generalização ou seja, apresentam bons resultados quando aplicadas a um conjunto de dados que não participaram do treinamento da rede. No entanto, em algumas situações as redes neurais mesmo depois de treinadas apresentam pequenos erros de calibração e elevados erros de predição devido ao número excessivo de neurônios utilizados na camada intermediária o que pode comprometer sua capacidade de generalização (BRAGA et al., 2007).

A capacidade de aprender através de exemplos e de generalizar a informação aprendida é, sem dúvida, o atrativo principal da solução de problemas através de RNAs. A generalização, que está associada à capacidade de a rede aprender através de um conjunto de exemplos e posteriormente dar respostas coerentes para dados não conhecidos, é uma demonstração de que a capacidade das redes neurais vai muito além do que simplesmente mapear relações de entrada e saída (BRAGA et al., 2007).

A seguir são apresentados alguns, dos muitos exemplos disponíveis na literatura de aplicação de redes neurais à predição da qualidade da água, seja de variáveis isoladamente ou de índices.

Gazzaz et al. (2012), utilizaram RNA para a predição do IQA no rio Kinta, na Malásia. A arquitetura ótima da rede encontrada foi 23-34-1 e a melhor predição do

IQA se deu com o algoritmo de treinamento de *quick propagation* (QP), com taxa de aprendizagem de 0,06 e coeficiente de QP de 1,75. Os valores estimados e calculados de IQA apresentaram correlação positiva, elevada e significativa ( $r = 0,977$ ,  $p < 0,01$ ).

Juahir et al. (2004) utilizaram, RNA no desenvolvimento e validação de um IQA para o rio Langat na Malásia, dados de 30 estações de monitoramento, sendo dividida em duas subséries. A primeira, com seis variáveis de qualidade da água independentes, foi utilizada para treinar, testar e validar RNAs. A seguir, foi aplicado teste de regressão múltipla linear para exclusão das variáveis que respondessem pelas menores contribuições para a variância no IQA, tendo sido obtidos os seguintes resultados: OD, DBO, SS e amônia contribuíram com aproximadamente 71% da variância, enquanto DQO e pH com apenas 8 e 2% , respectivamente. Na sequência, uma segunda série de dados, com apenas as quatro primeiras variáveis foi utilizada também para treinar, testar e validar RNAs. Encontrou-se que os coeficiente de correlação obtido com seis e quatro variáveis foram muito próximos (0,92 e 0,91, respectivamente), sugerindo que a RNA treinada com apenas quatro variáveis (sem DQO e pH) foi capaz de estimar o IQA com “precisão aceitável”.

No Brasil Sousa e Sousa (2010) utilizaram RNA para modelar a relação não linear entre chuva e vazão na bacia hidrográfica do rio Piancó, no semiárido paraibano. Avaliaram a capacidade modelagem do processo chuva-vazão em base mensal e se considerou, durante o seu treinamento, a influência da arquitetura da rede e da inicialização dos pesos. Das 120 arquiteturas, apenas as 24 melhores foram selecionados para análise, destas apenas uma foi escolhida para modelar vazões médias mensais na bacia estudada. Produziu o melhor resultado com valores para o coeficiente de determinação, de eficiência e erro padrão da estimativa de 92,0, 77,0% e 8,29, respectivamente.

Nasir et al. (2011) aplicaram RNA para modelagem e predição da qualidade da água no rio Juru. Dados secundários de 12 pontos de monitoramento 30 variáveis de qualidade da água foram utilizados como variáveis de entrada o modelo e um IQA como variável de saída. Três modelos de predição do IQA foram propostos. O desempenho das RNAs foi determinado pelos coeficientes de determinação ( $R^2$ ) e de correlação ( $R$ ) e pela raiz quadrada do erro quadrático médio (RMSE). A RNA1 fpoi a que apresentou o melhor desempenho com valores de  $R^2$  (0,99) e de RMSE (2,90) em relação aos outros

modelos. Esses autores sugerem que a utilização de RNA substituiria com vantagens o cálculo do IQA pelo método tradicional.

#### **2.4. Uso e ocupação do solo e variáveis de qualidade da água**

A qualidade da água varia de acordo com as condições ambientais (clima, relevo, geologia e pedologia), o uso e ocupação do solo. Em bacias hidrográficas onde há predomínio de cobertura natural do solo, ocorre uma maior proteção contra a erosão, aporte de sedimentos e carreamento de nutrientes para a água (BERTONI e NETO, 1999), sendo portanto, a análise desses elementos importante na avaliação da qualidade da água para o abastecimento de consumo humano.

No Brasil, a principal forma de erosão é a hídrica, que é provocada pela ação das chuvas, acarretando na desagregação e no transporte do solo pela enxurrada. Além das partículas de solo em suspensão, o escoamento superficial transporta nutrientes, matéria orgânica, sementes e agrotóxicos, acarretando o empobrecimento gradativo dos solos agrícolas, a elevação do custo de produção e, às vezes, resultando no abandono de áreas anteriormente produtivas (BERTONI e NETO, 1999).

Bacias hidrográficas exploradas por agricultura intensiva (culturas anuais e perenes) ou que apresentem percentuais elevados de pastagens em avançado de estado de degradação favorecem o processo erosivo do solo e o aumento do escoamento superficial (MERTEN e MINELLA, 2002).

A principal consequência deste impacto é o assoreamento que, além de modificar ou degradar a qualidade da água, afeta a fauna e a flora, provoca o decréscimo da velocidade de infiltração e armazenamento de água no solo, além da redução da disponibilidade hídrica nos mananciais (VANZELA et al., 2010).

Estudos da relação entre diferentes tipos de uso e ocupação do solo e variáveis da qualidade da água têm sido realizados como o objetivo de avaliar o potencial impacto dessas atividades sobre as variáveis da qualidade da água (VANZELA et al., 2010).

Trabalhos de Arcova e Cicco (1999) avaliaram a qualidade da água de microbacias com diferentes usos do solo na região de Cunha, Estado de São Paulo; Ranzini e Lima (2002) avaliaram o comportamento hidrológico do balanço de nutrientes e perdas de solo em duas microbacias reflorestadas com *Eucalyptus*, no Vale do Paraíba em São Paulo; Bertol et al., (2007) estudaram a perda de solo e, qualidade da água e

recentemente, Vanzela et al. (2010) observaram a influência do uso e ocupação do solo nos recursos hídricos do Córrego Três Barras, Marinópolis, São Paulo; e Huang et al. (2013) avaliaram os impactos do uso da terra na qualidade da água no Lago Basin na China.

Resultados obtidos por Boyacioglu (2007) na bacia do rio Menderes Buyuk na Turquia, mostraram que a qualidade da água foi mais afetada pelo uso agrícola e ocupação urbana. Sendo OD, DBO, Nitrogênio,  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$  e  $\text{SO}_4^{2-}$  as variáveis que apresentaram correlações significativas com esses tipos de uso e ocupação do solo.

De acordo com Queiroz et al. (2010) o uso de terraceamento e plantio direto, favorecem a reposição, manutenção da matéria orgânica, proteção, cobertura e melhoria da estrutura do solo, além da infiltração da água e diminuição do escoamento superficial.

A manutenção da cobertura do solo com vegetação natural influenciou positivamente as variáveis de qualidade da água das nascentes, enquanto as atividades agrícolas com preparo convencional do solo, aração e gradagem, contribuem para a perda de qualidade (DONADIO et al., 2005).

A retirada da vegetação contribui para o assoreamento e o aumento da turbidez das águas (OLIVEIRA et al., 1994); ARCOVA e CICCIO (1999) mostraram que nas bacias hidrográficas de uso agrícola, quando comparadas às de uso florestal, o transporte de sedimentos e a perda de nutrientes são maiores.

Os sistemas aquáticos são receptores das descargas resultantes das várias atividades humanas existentes nas bacias hidrográficas (MARGALEF, 1983). O uso e ocupação do solo na bacia do ribeirão São Bartolomeu é caracterizado por atividade agropecuária desenvolvida por pequenos produtores rurais e, mais recentemente, expansão urbana, com a implantação de condomínios e loteamentos; existem ainda áreas florestais, mas nos locais com maior declividade há predominância de pastagens e pastagens degradadas.

O uso de corretivos, fertilizantes e agrotóxicos de forma aleatória é prática corriqueira. A maioria das propriedades dedica-se à bovinocultura leiteira. Registra-se ainda criação de equinos, caprinos, aves e peixes (BASTOS, 2009).

O estabelecimento de associações entre as variáveis de qualidade da água e as alterações ocorridas na bacia hidrográfica, sejam essas de origem natural ou antrópica, constitui-se numa importante informação para avaliação da qualidade da água e definição dos usos dos recursos hídricos na bacia hidrográfica (ARCOVA e CICCIO, 1999).

### 3. MATERIAL E MÉTODOS

#### 3.1. Área de estudo

O trabalho foi realizado no município de Viçosa Minas Gerais, localizado na Zona da Mata, Estado de Minas Gerais, que se encontra a altitude de 649 m e tem como coordenadas geográficas o paralelo de 20°45'14'', latitude S, e o meridiano de 42°52'54'', longitude W Gr.; apresente área de cerca 300 km<sup>2</sup> e população estimada em 2013 de aproximadamente 76.147 habitantes (IBGE, 2014).

Um dos principais mananciais de abastecimento de água da cidade é o ribeirão São Bartolomeu (SB), cuja (micro)bacia hidrográfica se encontra inteiramente do município de Viçosa, pertencendo à sub-bacia do rio Piranga e, por extensão, à bacia do rio Doce (Figura 5).

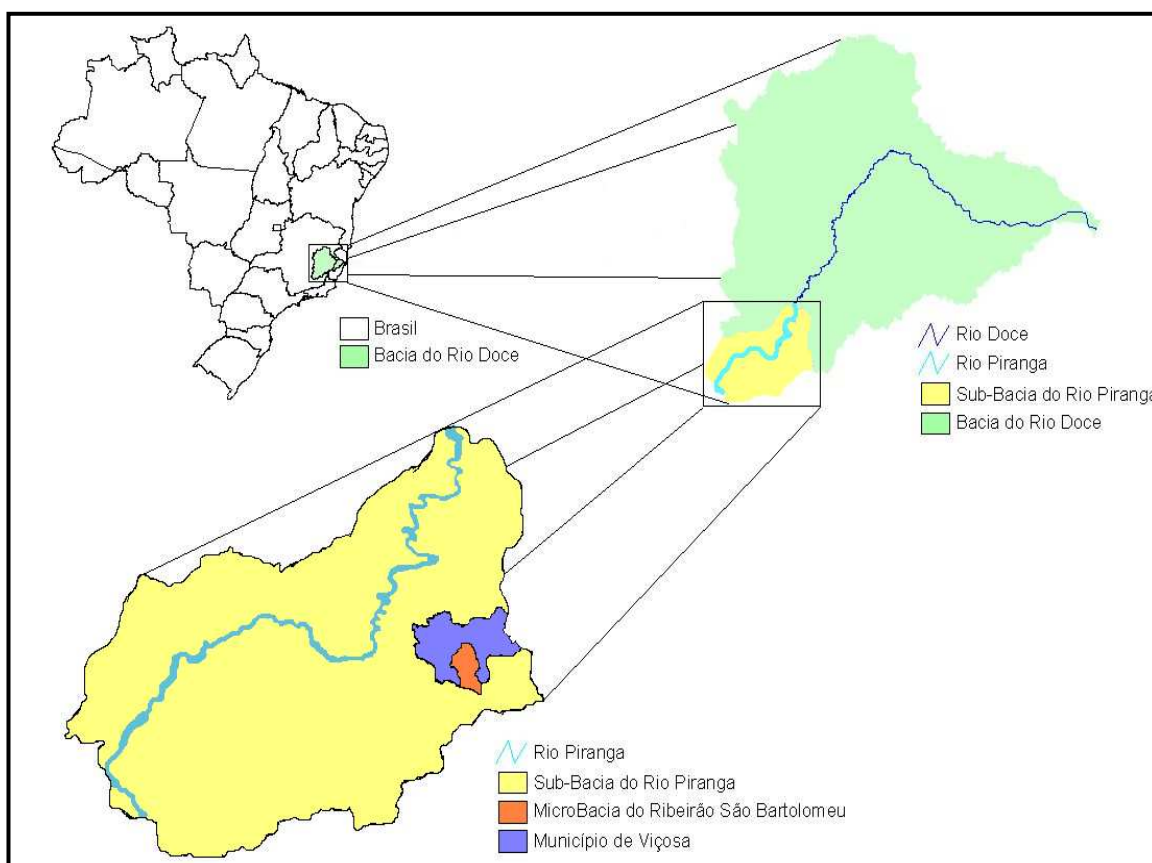


Figura 5 - Localização do município de Viçosa e da bacia do ribeirão São Bartolomeu  
Fonte: Silva (2010).

A bacia do São Bartolomeu, com densidade de drenagem de 4,5 km/km<sup>2</sup>, está à montante do campus da Universidade Federal de Viçosa (UFV), orientada no sentido sul-norte, apresentando na porção sul, limites com os municípios de Paula Cândido e Coimbra; ao norte, ainda na área urbana do município de Viçosa, no bairro Barrinha, o ribeirão deságua no rio Turvo Sujo que é tributário do rio Piranga, (BASTOS, 2009).

No campus da UFV, o São Bartolomeu recebe cinco barramentos consecutivos. As duas primeiras represas cumprem a função de reservatórios de acumulação, sendo que na segunda encontram-se as estruturas de captação de água para abastecimento do campus universitário da UFV e de parte da cidade.

A UFV capta cerca de 50 L/s para tratamento (em ETA de ciclo completo) e abastecimento do campus universitário, onde se encontram quase 20.000 usuários de um complexo tão diverso quanto: pavilhões de aula, alojamentos, restaurantes, hotel, posto de saúde, unidades administrativas, centro desportivo, piscinas, hospital veterinário, supermercado e indústria de laticínios.

No mesmo ponto, o Serviço Autônomo de Água e Esgoto de Viçosa (SAAE) capta 100 L/s e aduz água para a ETA I, também em ciclo completo. Esse sistema abastece cerca de 70% da população urbana de Viçosa no período de chuvas, sendo os restantes 30% atendidos pela ETA II, cujo manancial é o rio Turvo Sujo. Esses dois sistemas são interligados e em época de estiagem a situação se inverte, a ETA II passa a abastecer aproximadamente 70% da população urbana (BASTOS, 2009).

O uso e ocupação do solo na bacia de captação são caracterizados por atividades agropecuárias em pequenas propriedades rurais, com predominância de criação de gado de leite e mais recentemente, expansão urbana, com a implantação de condomínios e loteamentos (BASTOS, 2009).

### 3.2. Avaliação da qualidade da água e determinação do IQA e do IAP tradicionais

Informações sobre a qualidade da água foram obtidas do banco de dados da Divisão de Água e Esgotos (DAG) da UFV e de um programa de monitoramento executado, no âmbito do presente trabalho, mensalmente, de outubro de 2009 a setembro de 2010. Foram selecionados sete pontos de coleta localizados nas saídas das sub-bacias do Córrego do Engenho, Antuérpia, Palmital, Paraíso, Santa Catarina, Machados, São Lucas, e um oitavo ponto (Acamari) que representa a totalidade da área de drenagem da bacia do SB à montante do ponto de captação de água para abastecimento para consumo humano (chamada aqui de bacia de captação) (Figura 6). As coletas de amostras de água seguiram os procedimentos recomendados pela CETESB (1987)

As variáveis de controle incluíram aquelas utilizadas na composição do IQA: oxigênio dissolvido (OD), demanda bioquímica de oxigênio (DBO), nitrogênio total (Ntot), fósforo total (Ptot), potencial hidrogênico (pH), temperatura (Temp), turbidez (Turb), sólidos suspensos totais (SST), *Escherichia coli* (E. coli), além de coliformes totais (CT) e cor aparente, analisadas de acordo com os procedimentos do *Standard Methods for the Examination of Water and Wastewater* (APHA, 2005).

Foram também determinados (por espectrofotometria de chama) os metais cádmio (Cd), chumbo (Pb), cromo (Cr), níquel (Ni) usados no Índice de Substâncias Tóxicas (ST); e alumínio ( $Al^{3+}$ ), ferro ( $Fe^{2+}$ ), cobre (Cu), manganês ( $Mn^{2+}$ ) e zinco (Zn) usados no cálculo Índice de Substâncias Organolépticas (SO)

A concentração dos (oo)cistos foi realizada por meio das técnicas de floculação com carbonato de cálcio (VESEY et al., 1993) e filtração em membranas (FRANCO et al., 2001). Para amostras com turbidez inferior a 20 uT, volumes de 2 L de amostra eram filtrados e em seguida o material retido na membrana, incluindo os (oo)cistos, era ressuspenso por extração mecânica, procedendo-se à lavagem da membrana conforme descrito por Franco et al. (2001). Já para amostras com turbidez elevada, o que inviabiliza a filtração, volumes de 10 L eram concentrados por floculação com carbonato de cálcio e posterior sedimentação, de acordo com a técnica preconizada por Vesey et al. (1993). A enumeração dos (oo)cistos foi realizada com uso do kit de diagnóstico MERIFLUOR<sup>®</sup> C/G (Meridian Diagnostics, Inc.) pela técnica de imunofluorescência direta.

Os Índices de Qualidade da Água (IQA), –de Substâncias Tóxicas e Organolépticas (ISTO) e de Qualidade de Água Bruta para fins de Abastecimento Público (IAP) foram calculados de acordo com as equações 3, 5, 6, 7 e 8, apresentadas no item 3.1 deste trabalho. Esses índices, calculados da maneira tradicional, em referência à sua formulação original e conceitual, foram aqui denominados  $IAQ_{Cetesb}$  e  $IAP_{Cetesb}$ .

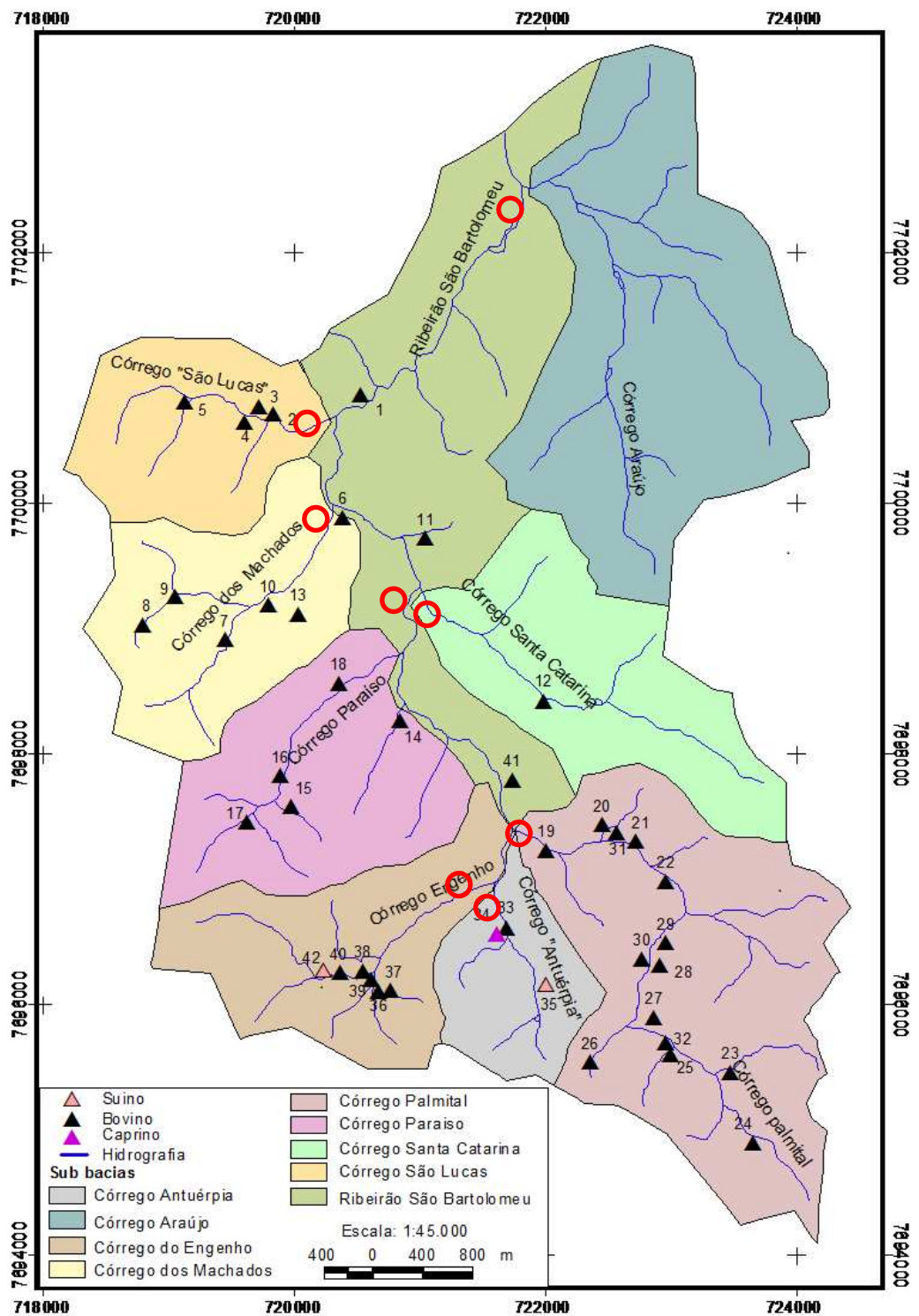


Figura 6 - Delimitação das sub-bacias, localização das propriedades rurais e dos pontos de coleta de amostras água (círculos vermelhos vazados), bacia de captação do ribeirão São Bartolomeu, Viçosa-MG

Fonte: Adaptado de Dias (2007).

### **3.3.Prerrogativas para análise fatorial de componentes principais**

Antes de aplicar a análise fatorial de componentes principais, faz-se necessário verificar a distribuição da normalidade, da adequabilidade do conjunto de dados e das variáveis individualmente e a verificação de existência de correlações pela esfericidade de Bartlett.

### **3.4. Elaboração da matriz de coeficientes de correlação**

Atendidas as prerrogativas para realização da ACP foi elaborada uma matriz de correlação e estatística descritiva das variáveis selecionadas para composição do IQA e IAP São Bartolomeu, considerando-se as que apresentaram correlações significativas ( $p < 0,05$ ). Essa matriz foi tomada como a base para a determinação das variáveis que foram analisadas nos fatores.

### **3.5. Extração de fatores**

O número de fatores extraídos foi definido pelo critério da raiz latente e pelo percentual de variância acumulada. Pelo método da raiz latente qualquer fator individual deve explicar a variância de pelo menos uma variável que contribui com um valor igual a um do autovalor total. Logo, apenas os fatores com raiz latentes ou autovalores maiores que um são considerados significativos e mantidos na análise e fatores, os de raiz latente menores que um são insignificantes e são descartados do modelo (NORUSIS, 1990). Já a variância acumulada dos dados no conjunto de fatores deve ser de no mínimo de 60% (HAIR et al., 2009).

### **3.6. Rotação ortogonal dos eixos**

Foi realizada uma rotação ortogonal ou simplesmente rotação dos eixos, pelos métodos quartimax, equimax e varimax para os dados do IQA e IAP. O processo maximiza a variância entre os fatores, alterando a raiz característica sem afetar a proporção da variância total explicada pelo conjunto.

### **3.7. Determinação dos escores fatoriais**

Para cada fator foi determinado o valor de escores fatoriais, pelo método de Bartlett, que corresponderam aos índices da qualidade da água, considerando-se o grupo de variáveis correlacionadas em cada fator. Os índices obtidos por este método possuem média zero e variância igual a um.

Valores extremos dos escores fatoriais (positivos ou negativos) representam classes de pior ou de melhor qualidade da água. Sendo assim, quanto mais negativo for o valor do escore fatorial (índice), melhor será a qualidade da água e quanto mais positiva, pior a qualidade da água.

### **3.8. Análises de correlações entre o índice Cetesb e obtidos pelos escores fatoriais**

Realizou-se análise de correlação, utilizando-se coeficiente de correlação de Pearson com análise de variância a nível de 5% (\*) e 1% (\*\*) de probabilidade, entre o IQA calculado pelo método Cetesb com os obtidos pelo método de Bartlett para AF/ACP.

### 3.9. Redes neurais artificiais na modelagem dos índices da qualidade da água

#### 3.9.1. Construção dos modelos de RNAs

Foram utilizadas redes neurais do tipo Multilayer Perceptron (MLP) definidas com três camadas: a de entrada, a intermediária ou oculta e a de saída. O número de neurônios na camada de entrada, utilizados na fase de treinamento, foi definido pela quantidade de variáveis de qualidade da água utilizadas na composição dos índices da qualidade da água (IQA) e índice de qualidade de água bruta para fins de abastecimento público (IAP) desenvolvidos pela Cetesb.

No software *statistica neural networks* não há a necessidade de normalização dos dados. A interface permite que a melhor rede seja selecionada pelo próprio software ou que o usuário faça a sua escolha. O número de camadas intermediárias e o número de neurônios nessas camadas foram encontrados de tentativa e erro. Considerou-se como resultados de saída valores das classes de qualidade da água variando de 0 a 100.

Para o ajuste da rede utilizou-se a função de transferência sigmoideal, conforme definida matematicamente na Equação 10, sendo o Backpropagation o algoritmo de treinamento.

$$f(x) = \frac{1}{1 + e^{(-x)}} \quad (10)$$

A taxa de aprendizagem ( $\eta$ ) e o momento ( $\mu$ ) foram definidos pelo software *Statistic Neural Network 4.0* e a duração e a parada do processamento da rede determinado pela convergência para o menor erro.

Para o desenvolvimento da arquitetura da RNA foram usados 350 conjuntos dados, distribuídos aleatoriamente nos subgrupos treinamento 176 (50%), validação 87 (25%) e 87 (25%) na fase de teste da rede.

Nas fase de treinamento e validação foram ajustados os pesos das conexões, determinado o momento de parada do treinamento e definição da arquitetura das redes, sendo selecionadas as que apresentaram os menores erros. Na fase de teste foi avaliada a capacidade de generalização a partir a partir de um conjunto de dados desconhecidos pelas redes neurais.

### 3.9.2. Avaliação de desempenho das RNAs

Para avaliação de desempenho das RNAs utilizou-se o coeficiente de determinação ( $R^2$ ), coeficiente de eficiência (E) e o erro padrão de estimativa (EPE).

O coeficiente de determinação ( $R^2$ ) é definido como o quadrado do coeficiente de correlação de Pearson, sendo uma medida da proporção da variabilidade de uma variável, que é explicada pela variabilidade de outra. É pouco comum que se tenha uma correlação perfeita ( $R^2 = 1$ ), porque existem muitos fatores que determinam as relações entre variáveis (SOUSA e SOUSA, 2010). A Equação 11 apresenta a fórmula de cálculo do coeficiente de correlação de Pearson.

$$R^2 = \frac{n \sum (I_o \times I_e) - \sum I_o \times \sum I_e}{[\sqrt{n \sum I_o^2 + (o)^2}] [n \sum I_e^2 - (\sum I_e)^2]} \quad (11)$$

Em que:

$R^2$  - Coeficiente de determinação

$I_e$  – índice estimado

$I_o$  – índice observado

$n$  – número de observações

O Coeficiente de eficiência foi definido por Nash e Sutcliffeem (1970). Segundo Moriasi et al., (2007) este coeficiente tem sido utilizado em diversos estudos na avaliação de modelos hidrológicos na definição de variáveis e na predição de modelos de qualidade da água (Equação 12).

$$E = 1 - \frac{\sum (I_o - I_e)}{\sum (I_o - I_m)} \quad (12)$$

Em que:

E- Coeficiente de eficiência

Ie – índice estimado

Io – índice observado

Im – Média do índice observado

O erro padrão da estimativa estipula o desvio médio entre os valores observados e os estimados. Informa quanto menor o valor melhor a predição do modelo, expresso pela Equação 13.

$$EPE = \sqrt{\frac{\sum (I_o - I_e)^2}{n-1}} \quad (13)$$

Em que:

EPE - Erro padrão de estimativa

Ie – índice estimado

Io – índice observado

n – número de observações

### 3.10. Uso e ocupação do solo e variáveis de qualidade da água

O mapa de uso e ocupação do solo foi elaborado a partir de um mosaico de toda a área de estudos, composta por imagens IKONOS 2007 com resolução espacial ou geométrica e reamostrada para 1m x 1m já corrigida radiométrica e geometricamente (adquirida pelo Projeto Plano de Segurança da Água financiado pelo Ministério da Saúde) (BASTOS, 2009).

Foram identificados e vetorizados, manualmente, polígonos contendo as diferentes classes de uso e ocupação do solo. Dentre as classes de interesse destacam-se: Agricultura (culturas anuais de milho, feijão e mandioca), café, áreas urbanizadas, pastagem degradada, pastagem não degradada, formações florestais (mata, capoeira e reflorestamentos) e piscicultura (hidrografia) (Figura 7).

Para cada tipo de uso e ocupação do solo foi calculado um índice de ocupação determinado pela relação da área ocupada (ha), pela área total da sub-bacia (ha), conforme apresentado na equação 14.

$$IC = \frac{\text{Área Ocupada (ha)}}{\text{Área Total da Sub bacia (ha)}} \quad (14)$$

Em que:

IC = Índice de ocupação

A relação entre os diferentes tipos de usos e ocupação do solo e as variáveis de qualidade foi mensurada pela aplicação de análise de correlação, utilizando-se coeficiente de correlação de Pearson, com análise de variância ao nível de 1% e 5% de probabilidade. Foi realizado teste de normalidade dos dados pela aplicação do teste de Teste de Kolmogorov-Smirnov ao nível de 5% de significância.

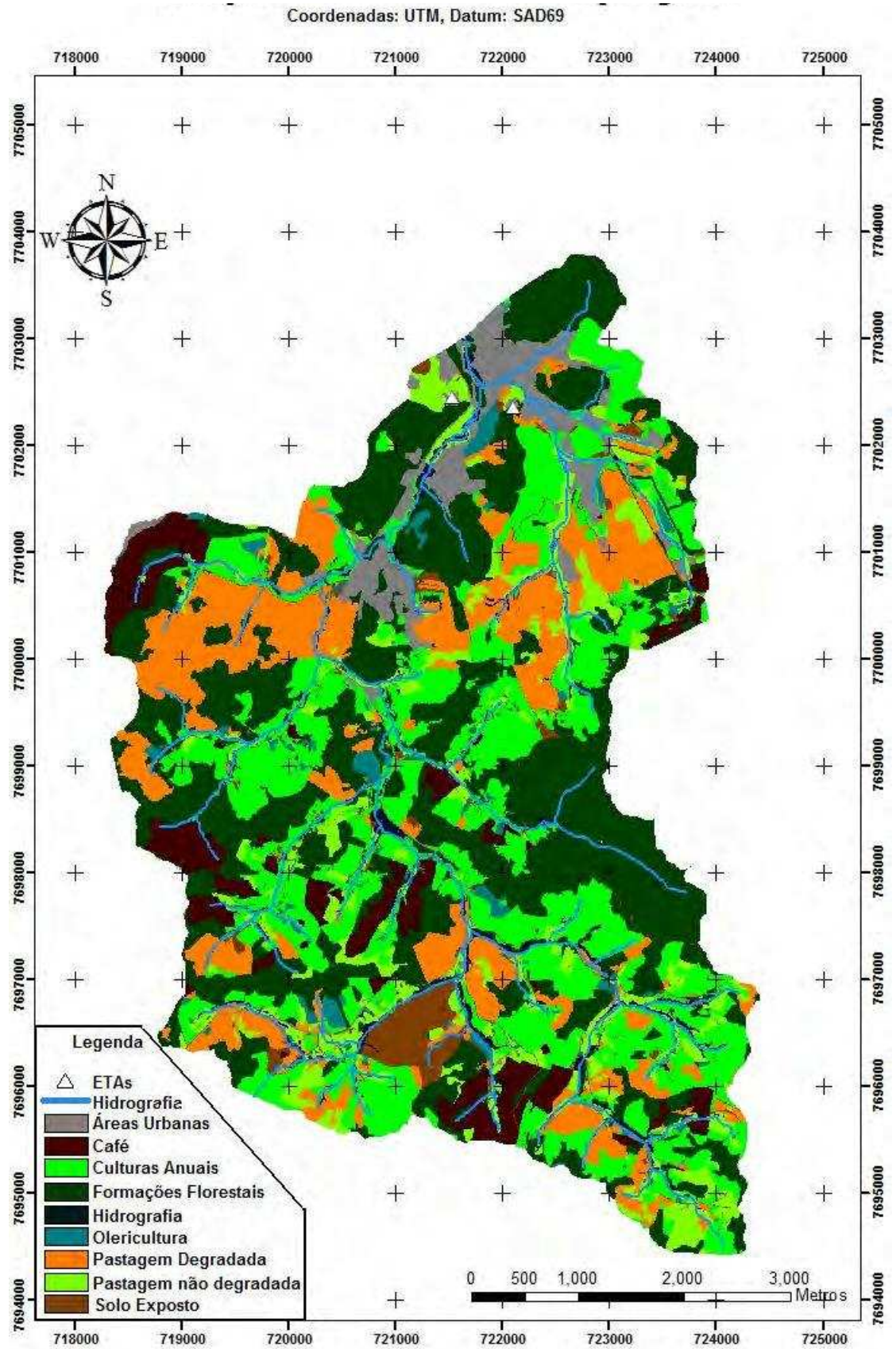


Figura 7 - Mapa de uso e ocupação do na bacia ribeirão São Bartolomeu, Viçosa – MG.  
 Fonte: Silva (2010).

### **3.11. Análises estatísticas e nomenclatura dos índices modelados**

As análises estatísticas foram realizadas pelo software SPSS – Statistical Package for the Social Sciences 18.0, gráficos e tabelas foram elaborados no Excel (Microsoft, 2007).

Os índices modelados pela análise de componentes principais e redes neurais artificiais receberam a nomenclatura de índice da qualidade da água da bacia do ribeirão São Bartolomeu (IQAS<sub>b</sub>) e índice da qualidade da água para fins de abastecimento público da bacia do ribeirão São Bartolomeu (IAPS<sub>b</sub>).

## **4. RESULTADOS E DISCUSSÃO**

### **4.1. Condições para realização da análise fatorial de componentes principais**

A realização da análise de componentes principais demanda que o conjunto de dados e as variáveis constituintes do  $IQA_{Sb}$  e  $IAP_{Sb}$  atendam as exigências das condições de distribuição de normalidade, adequabilidade e esfericidade.

#### **4.1.1. Teste de normalidade**

Os resultados do teste de normalidade de Kolmogorov-Smirnov, para  $IQA_{Sb}$  e  $IAP_{Sb}$ , mostram que os dados apresentam uma distribuição normal ao nível de 5% de significância.

#### **4.1.2. Teste de adequabilidade da amostra**

O teste da adequabilidade (Kaiser-Meyer-Olkin, KMO) aplicado para comparar a magnitude dos coeficientes, apresentou valor de 0,70 e 0,68 para o  $IQA_{Sb}$  e  $IAP_{Sb}$ , respectivamente, demonstrando que ACP pode ser aplicada, uma vez que a limitação de uso ocorre quando esse valor de KMO é inferior a 0,50 (Tabelas 5 e 6).

##### **4.1.2.1. Teste de adequabilidade das variáveis do $IQA_{Sb}$**

De onze variáveis analisadas, apenas sete apresentaram valores de KMO maiores ou iguais a 0,5, sendo portanto mantidas na análise de componentes principais. Das sete foram selecionadas três variáveis físicas (turbidez, cor e sólidos totais); três químicas (oxigênio dissolvido, nitrogênio (total e pH) e uma biológica coliformes totais (Tabela 1).

Tabela 1 - Comunalidades extraídas e medidas de adequação das variáveis do IQA<sub>Sb</sub>

Variáveis	Comunalidades Extraídas	Medidas de Adequação (KMO)	
		Variáveis	Amostra
OD	0,60	0,78	
Ntot	0,73	0,73	
Turb	0,71	0,69	
Cor	0,74	0,67	0,70
SST	0,50	0,65	
pH	0,74	0,64	
Ctot	0,61	0,71	

#### 4.1.2.2. Teste de adequacidade das variáveis do IAP<sub>Sb</sub>

Dezoito variáveis foram analisadas, somente nove cumpriram os requisitos exigidos pelo teste de adequabilidade para realização da ACP ( $KMO \geq 0,5$ ). Foram selecionadas turbidez, cor, oxigênio dissolvido, nitrogênio total, fósforo total, pH, alumínio, ferro dissolvido e manganês dissolvido (Tabela 2).

Tabela 2 - Comunalidades extraídas e medidas de adequação das variáveis do IAP<sub>Sb</sub>.

Variáveis	Comunalidades Extraídas	Medidas de Adequação	
		Variáveis	Amostra
OD	0,60	0,50	
Ntot	0,68	0,67	
Ftot	0,55	0,51	
Turb	0,73	0,73	
pH	0,56	0,60	0,68
Al	0,50	0,78	
Fe	0,60	0,62	
Mn	0,61	0,59	
Cor	0,60	0,70	

### **4.1.3. Teste de Esfericidade**

O teste de esfericidade de Bartlett foi significativo ( $p < 0,05$ ) para o IQA<sub>Sb</sub> e IAP<sub>Sb</sub>, indicando a existência de correlações suficientes para a realização da análise de componentes principais.

Resultados semelhantes foram observados nos trabalhos de Rodrigues e Rodriguez (2003); Meireles (2007); Mahmood et al.(2011); Al-Badaii et al., (2013).

Esses resultados indicam que análise fatorial de componentes principais pode ser aplicada ao conjunto de dados na avaliação da qualidade da água e modelagem do índice de Qualidade das Águas Brutas para Fins de Abastecimento Público da bacia do ribeirão São Bartolomeu IAP<sub>Sb</sub>.

### **4.1.4. Aplicação da Análise de Componentes Principais para IQA e IAP<sub>Sb</sub>**

A discussão dos resultados da análise de componentes principais foi dividida em cinco etapas: estatística básica descritiva; matriz de correlação; extração dos fatores, rotação dos eixos e modelagem dos índices da qualidade da água pelo método de Bartlett para o IQASb e IAPSb.

#### **4.1.4.1. Estatística descritiva das variáveis do IQA<sub>Sb</sub>**

As variáveis sólidos totais, nitrogênio e cor apresentaram os maiores valores de coeficiente de variação, enquanto pH teve pouca variabilidade ao longo dos oito pontos de monitoramento (Tabela 3). Os coeficientes de variação são indicativos da maior ou menor influencia das atividades antrópicas (agricultura, lançamentos de efluentes e de excretos de animais) nos valores dessas variáveis nos diferentes trechos da bacia do ribeirão São Bartolomeu. Toledo e Nicoletta (2002), Coletti et al. (2010) encontraram uma maior variabilidade de nitrogênio e uma menor variação do pH nos pontos de monitoramento.

Tabela 3 – Estatística descritiva das variáveis do IQA<sub>Sb</sub>

Variáveis	Observações	Mínimo	Máximo	Média	Desvio Padrão	Coef. de Variação
OD	96	0,70	9,20	5,57	2,01	36,08
Ntot	96	0,30	7,35	1,63	1,49	91,33
Turb	96	1,35	46,20	14,52	8,75	60,28
Cor	96	7,50	337,00	80,46	64,22	79,82
ST	96	5,00	740,00	182,83	171,74	93,93
pH	96	5,94	7,70	6,72	0,36	5,39
Ctot	96	214,00	45500,00	12675	9128	72,02

#### 4.1.4.2. Estatística descritiva das variáveis do IAP<sub>Sb</sub>

Alumínio, manganês, fósforo total e ferro foram as variáveis que apresentaram os maiores valores de coeficiente de variação (Tabela 4). Maiores concentrações de ferro, manganês e alumínio podem estar associadas a ambientes hidromórficos com baixa saturação de oxigênio e presença de ferro e manganês na forma reduzida, sendo mais solúveis em água.

Tabela 4 – Estatística descritiva das variáveis do IAP<sub>Sb</sub>

Variáveis	Observações	Mínimo	Máximo	Média	Desvio Padrão	Coef. de Variação
OD	96	0,70	9,20	5,57	2,01	36,08
Ntot	96	0,30	7,35	1,63	1,49	91,33
Ptot	96	0,00	0,26	0,06	0,06	100,00
Turb	96	1,35	46,20	14,52	8,75	60,28
Cor	96	7,50	337,00	80,46	64,22	79,82
pH	96	5,94	7,70	6,72	0,36	5,39
Al	96	0,00	0,92	0,10	0,23	222,30
Fe	96	0,25	10,20	2,61	2,40	91,95
Mn	96	0,01	1,17	0,23	0,23	100,00

#### 4.1.4.3. Matriz de correlação do IQA<sub>SB</sub>

A interação entre as variáveis pode ser considerada como um ponto de partida para avaliação da qualidade da água e a seleção das variáveis de maior importância para composição dos modelos. Oxigênio dissolvido, nitrogênio e turbidez foram as variáveis que apresentaram os maiores valores de coeficientes de correlação, significativos ao nível de 5% de probabilidade, com um maior número de variáveis (Tabela 5).

De acordo com essa tabela, a turbidez e a cor (0,66) apresentaram o maior valor de coeficiente de correlação, seguidas por nitrogênio e cor (0,57), nitrogênio e turbidez (0,50), oxigênio dissolvido e turbidez (-0,50), oxigênio e coliformes totais (-0,47) e coliformes e sólidos suspensos totais (0,44).

O valor elevado do coeficiente de correlação entre turbidez e cor pode ser explicado pela interpretação da cor aparente, que inclui em seu valor final uma parcela da turbidez da água. Quando a cor é determinada após centrifugação obtém-se a cor verdadeira que não sofre interferência de materiais coloidais e em suspensão, diluindo a relação com a turbidez. A cor é um padrão de natureza estética e componente do padrão de aceitação para consumo humano. Porém se a cor for devido a substâncias húmicas pode ser um bom indicativo da presença de precursores de trihalometanos.

A medida que aumenta a concentração da turbidez na água aumenta o consumo de oxigênio dissolvido, provavelmente pelo processo de decomposição da matéria orgânica. A mesma tendência de consumo de oxigênio foi observada para os coliformes totais, possivelmente pelo aumento da população microbiana na água. Os coliformes incluem uma gama variada de bactérias que vai desde as que habitam o intestino de animais e humanos, aos de vida livre, sem contar que ainda podem se multiplicarem na água

O nitrogênio pode ser de origem mineral ou orgânica. Valores altos e significativos ( $p < 0,05$ ) do coeficiente de correlação entre turbidez e nitrogênio, é um indicativo do aumento do escoamento superficial nas áreas de agricultura e de pastagem degradada. Outra possibilidade é que tenha sido originado dos lançamentos difusos de esgotos domésticos e de efluentes resultantes da criação de animais.

Tabela 5 – Coeficientes de correlação das variáveis físico, químicas e biológica do IQA<sub>Sb</sub>.

Variáveis	OD	Ntot	Turb	Cor	SST	pH
Ntot	-0,17					
Turb	<b>-0,50</b>	<b>0,50</b>				
Cor	<b>-0,31</b>	<b>0,57</b>	<b>0,66</b>			
ST	<b>-0,30</b>	<b>0,35</b>	0,24	0,10		
pH	<b>0,36</b>	0,18	<b>-0,36</b>	-0,20	0,04	
Ctot	<b>-0,47</b>	0,15	<b>0,35</b>	0,35	<b>0,44</b>	-0,20

Valores em negrito expressam as correlações significativas ao nível de 5% de probabilidade.

#### 4.1.4.4. Matriz de correlação do IAP<sub>Sb</sub>

A extração da matriz de correlação possibilita a identificação da associação entre as variáveis, além de mostrar a tendência geral dos dados, bem como as variáveis de maior potencialidade para a modelagem do índice.

Cor e turbidez apresentaram os maiores valores de coeficiente de correlação (0,66), seguidas por nitrogênio e cor (0,57), fósforo total e ferro (0,53), Fósforo total e turbidez (0,49), nitrogênio e turbidez (0,50), ferro e manganês (0,50) e pH e alumínio (-0,46) (Tabela 6).

A correlação significativa entre oxigênio dissolvido e turbidez pode estar associada ao lançamento de efluentes domésticos nas propriedades rurais. De acordo com Bastos (2009) cerca de um terço dos esgotos produzidos na bacia são lançados nos cursos d'água sem nenhum tipo de tratamento. O aumento do consumo de oxigênio indica que está sendo utilizado pelas bactérias decompositoras na estabilização da matéria orgânica.

Tabela 6 – Coeficientes de correlação das variáveis físico químicas do IAP<sub>Sb</sub>.

Variáveis	OD	Ntot	Ptot	Turb	Cor	pH	Al	Fe
Ntot	-0,17							
Ptot	0,04	0,17						
Turb	<b>-0,50</b>	<b>0,50</b>	<b>0,49</b>					
Cor	<b>-0,31</b>	<b>0,57</b>	0,23	<b>0,66</b>				
pH	<b>0,36</b>	0,18	<b>-0,30</b>	<b>-0,36</b>	-0,20			
Al	<b>-0,33</b>	0,19	-0,16	0,04	0,04	<b>-0,46</b>		
Fe	<b>-0,40</b>	0,04	<b>0,53</b>	<b>0,40</b>	0,18	<b>-0,45</b>	0,04	
Mn	<b>-0,43</b>	0,03	<b>0,50</b>	0,23	0,01	-0,18	-0,28	<b>0,53</b>

Valores em negrito expressam as correlações significativas ao nível de 5% de probabilidade.

De acordo com a Cetesb (2009) as concentrações podem ser influenciadas pelo pH e matéria orgânica e outros ligantes.

As concentrações de alumínio dissolvido na água são fortemente influenciadas pelo pH. Nos meses em que o valor de pH foram inferiores 6,50, as concentrações de alumínio dissolvida na água ficaram abaixo de 0,3 mg/L e nulas quando o pH foi superior a 6,50 (Figura 8).

A presença de ferro e manganês nas águas pode ser atribuída as áreas alagadas com baixa concentração de oxigênio que aumentam a solubilidade do ferro e manganês e conseqüentemente uma maior concentração na água. Segundo Resende et al (2007) em condições de excesso de água o ambiente é de redução com predomínio de ferro e manganês reduzidos.

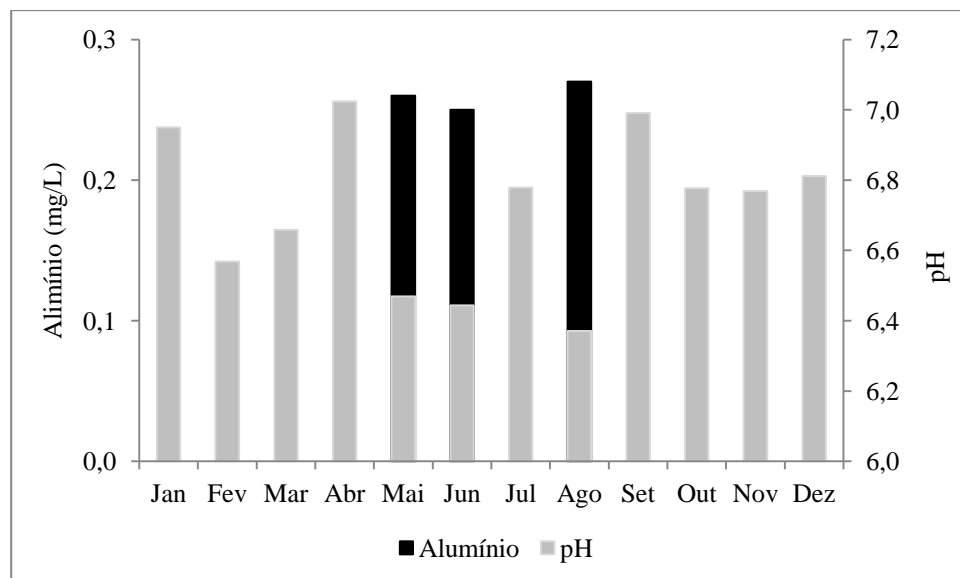


Figura 8 - Resultados médios de alumínio dissolvido e pH nos doze meses de monitoramento.

#### 4.1.4.5. Extração de fatores e rotação de eixos do IQA<sub>Sb</sub>

Após a aplicação da ACP os fatores com autovalor menor que um e com menos de 60% de explicação da variância total das variáveis foram excluídos do modelo. Foram portanto, obtidos três fatores (F1, F2 e F3) para as matrizes fatoriais rotacionadas e não rotacionadas, que apresentaram autovalor maior que um (Figura 9) e percentual de explicação da variância total das variáveis originais da matriz de cargas fatoriais não rotacionadas para o IQA foi de 65,75 (Tabela 7).

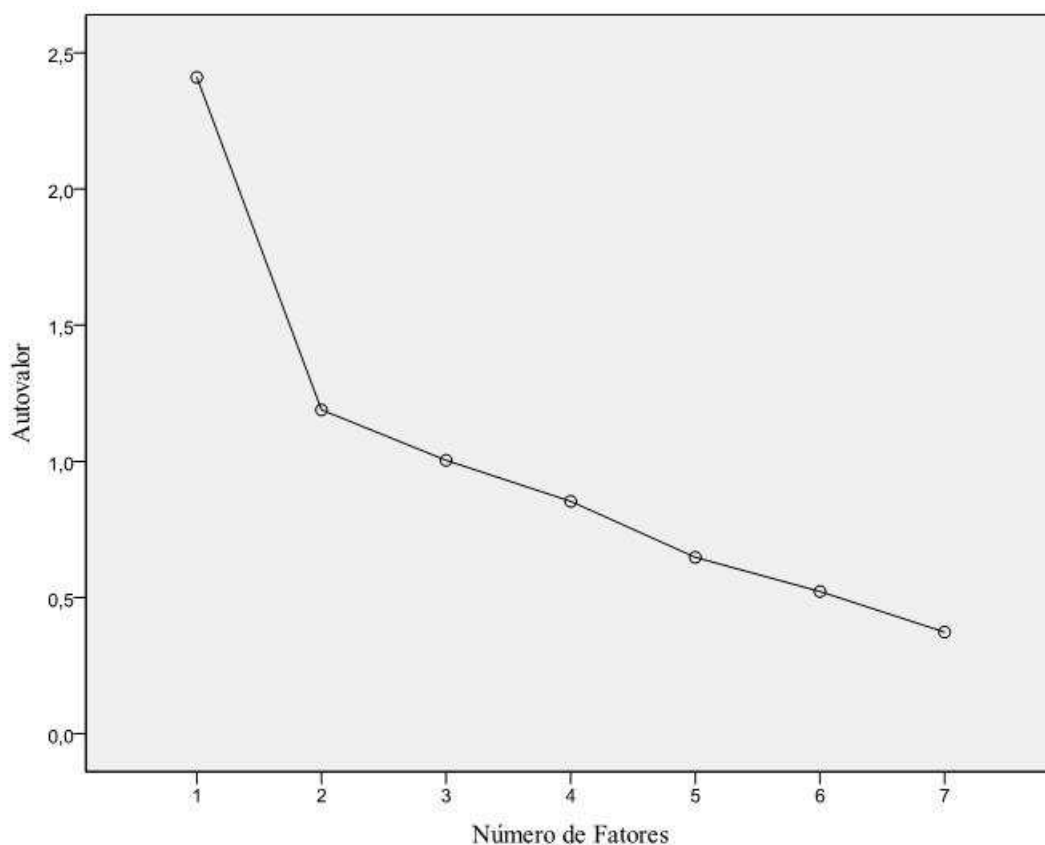


Figura 9 - Número de fatores extraídos com autovalor maior que a unidade.

Para a exclusão de fatores da AF / CP levou-se em consideração o critério da raiz latente, que considera que qualquer fator individual deve explicar a variância de no mínimo uma variável para ser mantido na análise. Destes, apenas três permaneceram na análise (F1, F2 e F3). Observa-se que o primeiro fator explicou o maior percentual da variância de nitrogênio, turbidez e cor. O segundo de coliformes totais, oxigênio dissolvido e sólidos suspensos, ficando o pH isolado no último. (Tabela 7).

Trabalhos de Helena et al., (2000); Candia, (2007), Coletti et al., (2010) e Al-Badai et al., (2012) e Li (2013), utilizaram essa mesma metodologia na análise da qualidade de água em bacias hidrográficas e encontraram resultados similares aos deste trabalho.

Os fatores F1, F2 e F3 não rotacionados apresentaram variâncias de 34,42%, 16,98% e 14,35%, contra 27,18%, 20,31% e 18,26% dos fatores rotacionados. Observa-se que na segunda situação ocorreu uma melhor redistribuição da variância nos três fatores (Tabela 7).

A variação entre os fatores F1 e F2 na matriz não rotacionada, foi de 17,44% e após a aplicação da rotação, essa diferença caiu para 6,87%. A relação entre os fatores F2 e F3 ficou na casa do 2% para as duas situações (Tabela 7).

Tabela 7 – Matriz de cargas fatoriais não rotacionadas para o IQA.

Variáveis	Fatores não Rotacionados		
	F1	F2	F3
OD	-0,63	-0,60	0,41
Ntot	0,61	0,13	0,48
Turb	0,83	0,32	0,07
Cor	0,73	0,00	0,40
SST	0,40	0,59	-0,18
pH	-0,33	0,05	0,23
Ctot	0,40	0,78	-0,60
% Variância não rotacionada explicada	34,42	16,98	14,35
% Variância acumulada dos fatores não rotacionados	34,42	51,14	65,75
Autovalor	2,40	1,19	1,00

Nem sempre os fatores obtidos sem a rotação dos eixos são de fácil interpretação, devido aos valores das cargas fatoriais estarem próximos da média. Nesse caso, foi necessária a geração de uma nova matriz fatorial rotacionada ortogonalmente, que redistribuiu a variância entre os três fatores criando um padrão fatorial simples e de fácil interpretação. Foram aplicadas as rotações ortogonais quartimax, varimax e equimax. Destas, a varimax foi a que apresentou melhores resultados interpretativos, principalmente entre os fatores F1 e F2 e a capacidade de análise e interpretação dos resultados das variáveis estudadas (Tabela 8).

Segundo Toledo e Nicolella (2002) e Coletti et al. (2010) para que o primeiro fator possa ser considerado como o índice da qualidade da água é necessário que além de explicar a maior parte da variabilidade dos dados, todas as variáveis devem apresentar os maiores valores de cargas fatoriais. Por não ter ocorridos essa situação optou-se nesse estudo a análise individualizada de cada fator, como será demonstrado abaixo.

O primeiro fator rotacionado explicou 27,18% da variância total dos dados. Nitrogênio total (0,79), cor (0,82) e turbidez (0,70) apresentaram as maiores cargas fatoriais nesse fator (Tabela 12).

Este fator expressa o efeito do escoamento superficial de áreas agrícolas, esgotos domésticos e excretos de animais que provavelmente, favoreceram o aumento da concentração de nitrogênio na água. Guedes et al., (2012) encontrou resultados semelhantes.

Na bacia do ribeirão São Bartolomeu há uma predominância de áreas de agricultura e pastagem (65%), sendo que desse montante, aproximadamente 16% da pastagem está em processo acentuado de degradação, sem contar que existem áreas agrícolas com práticas de manejo e conservação inadequadas para as condições de solo, do clima e declividade da bacia (Bastos, 2009).

No segundo fator de 20,31% da variância acumulada foi explicada por sete variáveis. Destas, chama-se atenção para coliformes totais (0,78), oxigênio dissolvido (-0,60) e sólidos suspensos totais (0,59) (Tabela 8).

O segundo fator expressa a alteração da qualidade da água pelo aumento da concentração de matéria orgânica proveniente do lançamento de esgotos domésticos e de efluentes oriundos de atividades de criação de animais.

Os coliformes totais por incluírem bactérias que não são de origem exclusivamente fecais, não foram considerados como indicador de poluição pelo lançamento de esgotos e efluentes.

Nesse caso considerou-se os sólidos totais como os responsáveis pelo aporte de material orgânico na água. Esse aporte de material orgânico pode ser explicado pela condição sanitária da bacia, cerca de 30% do esgoto produzido nas propriedades rurais são lançados diretamente nos cursos d'água sem nenhum tipo de tratamento, contribuindo, assim, para o aumento da carga orgânica nos sólidos suspensos totais (Bastos, 2009).

O terceiro fator explicou 18,26% da variância dados, sendo o pH a variável que apresentou o maior percentual de explicação sua variância (0,86) (Tabela 12). É considerada uma variável importante para o bom desempenho do processo de tratamento e manutenção da rede de distribuição de água, não sendo portanto, importante para análise da qualidade da água na bacia, por apresentar pouca variabilidade ao longo dos pontos de monitoramento.

Tabela 8 – Matriz de cargas fatoriais rotacionadas para o IQA.

Variáveis	Fatores Rotacionados		
	F1	F2	F3
OD	-0,25	-0,60	0,39
Ntot	0,79	0,13	0,30
Turb	0,70	0,32	-0,34
Cor	0,82	0,002	-0,25
SST	0,22	0,59	0,32
pH	-0,09	0,05	0,86
Ctot	-0,02	0,78	-0,04
% Variância rotacionada explicada	27,18	20,31	18,26
% Variância acumulada dos fatores	27,18	47,49	65,75
Autovalor	2,40	1,19	1,00

#### 4.1.4.6.Extração de fatores e rotação de eixos do IAP<sub>Sb</sub>

Procedimento semelhantes aos do IQAS<sub>b</sub> foram utilizados na análise e interpretação dos resultados da ACP do IAP<sub>Sb</sub>. Foram extraídos os fatores F1, F2 e F3 que explicaram 63,20% da variância total dos dados originais (Tabela 9), também apresentaram auto valor maior que um, ilustrado na Figura 10.

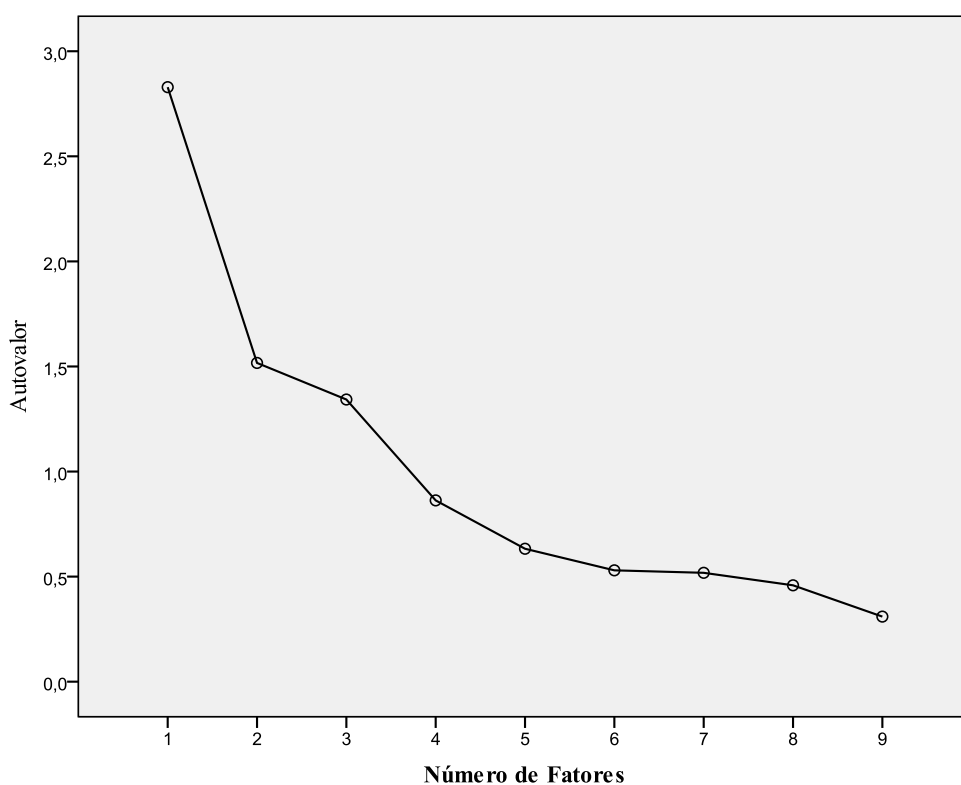


Figura 10 - Número de fatores extraídos com autovalor maior que a unidade.

Os fatores não rotacionados F1, F2 e F3 explicaram 31,43%, 16,85% e 14,92% das variâncias, respectivamente, contra 24,85%, 22,40% e 15,95% dos fatores rotacionados (Tabela 9).

A variação da variância acumulada do primeiro fator para o segundo da matriz não rotacionada, foi de 14,58% e de 1,93% do segundo para o terceiro. Foram aplicadas as rotações ortogonais quartimax, varimax e equimax.

Após a aplicação da rotação dos eixos, essa diferença caiu para 2,45% entre os dois primeiros, passando para 6,45 entre os dois últimos (Tabela 10). Andrade et al., (2007) encontraram resultados semelhantes no estudo da seleção dos indicadores da qualidade das águas superficiais pelo emprego da análise multivariada.

A rotação varimax permitiu uma melhor interpretação dos fatores (F1, F2 e F3) obtidos pela matriz de cargas fatoriais não rotacionadas. Esses fatores, obtidos sem a rotação dos eixos, nem sempre são de fácil interpretação, o que de certa forma, dificulta uma avaliação da qualidade da água (Tabela 10).

Tabela 9 – Matriz de cargas fatoriais não rotacionadas para o IAP.

Variáveis	Fatores não Rotacionados		
	F1	F2	F3
OD	-0,64	-0,11	0,26
Ntot	0,41	0,69	0,16
Ftot	0,66	-0,19	-0,21
Turb	0,78	0,15	0,32
pH	-0,44	0,49	-0,41
Al	-0,15	-0,48	0,73
Fe	0,59	-0,50	-0,14
Mn	0,54	-0,29	-0,46
Cor	0,60	0,41	0,41
% Variância explicada	31,43	16,85	14,92
% Variância acumulada dos fatores	31,43	48,28	63,20
Autovalor	2,83	1,51	1,34

O primeiro fator apresentou cargas fatoriais positivas para manganês (0,76), ferro (0,74) e fósforo total (0,70) e negativa para oxigênio dissolvido (-0,60). As outras variáveis foram mais bem explicadas no segundo e terceiro fatores (Tabela 10). Esse fator pode está associado à variáveis indicadoras da alteração da qualidade da água pelo aumento de matéria orgânica na água proveniente do lançamento de efluentes e ou de excretos de animais.

Acarga fatorial negativa (-0,60) do oxigênio dissolvido pode ser um sinal de seu consumo no processo de decomposição da matéria orgânica, rica em fosforo, originada de dejetos de animais ou de esgotos domésticos lançados nos cursos d'água sem nenhum tipo de tratamento.

Toledo e Nicolella (2002) mostraram que o aporte de fósforo na água tinha como principal fonte os esgotos domésticos, sendo portanto, considerado como o principal agente causador da degradação da qualidade da água em um trecho da bacia do ribeirão Jardim no município de Guaira em São Paulo.

Na bacia do ribeirão São Bartolomeu 30% dos esgotos domésticos produzidos nas propriedades rurais são lançados sem tratamento nos cursos d'água. Nesta bacia, existem duas grandes suinoculturas, uma caprinocultura e diversas propriedades rurais com criação extensiva de bovinos (Figura 11), nestas, os animais possuem acesso livre aos cursos d'água (BASTOS, 2009).

No córrego do Engenho foi observado, no período de menor vazão do ano, o lançamento de efluentes, por uma suinocultura, sem nenhum tipo de tratamento, sendo portanto, uma fonte importante de poluição da água por matéria orgânica (Figuras 11 e 12).

Os efluentes produzidos nas atividades de criação de animais e o lançamentos de esgotos domésticos contribuem para o aumento da carga orgânica na água diminuindo a concentração de oxigênio dissolvido e qualidade da água.

De acordo com Cetesb (2009) a presença de ferro e manganês nas águas pode ser atribuída ao intemperismo de minerais e rochas, principalmente quando óxidos, sulfetos, carbonatos e silicatos que contêm estes metais.

Na bacia do ribeirão São Bartolomeu predominam solos muito intemperizados, que podem ser fontes potenciais de Ferro e manganês pela solubilização na solução da solo e lixiviação para os curso d'águas.

Ferro e manganês não estão relacionados a riscos à saúde da população, mas com a perda de qualidade pelo padrão organoléptico da água para consumo humano (VON SPERLING, 2005).

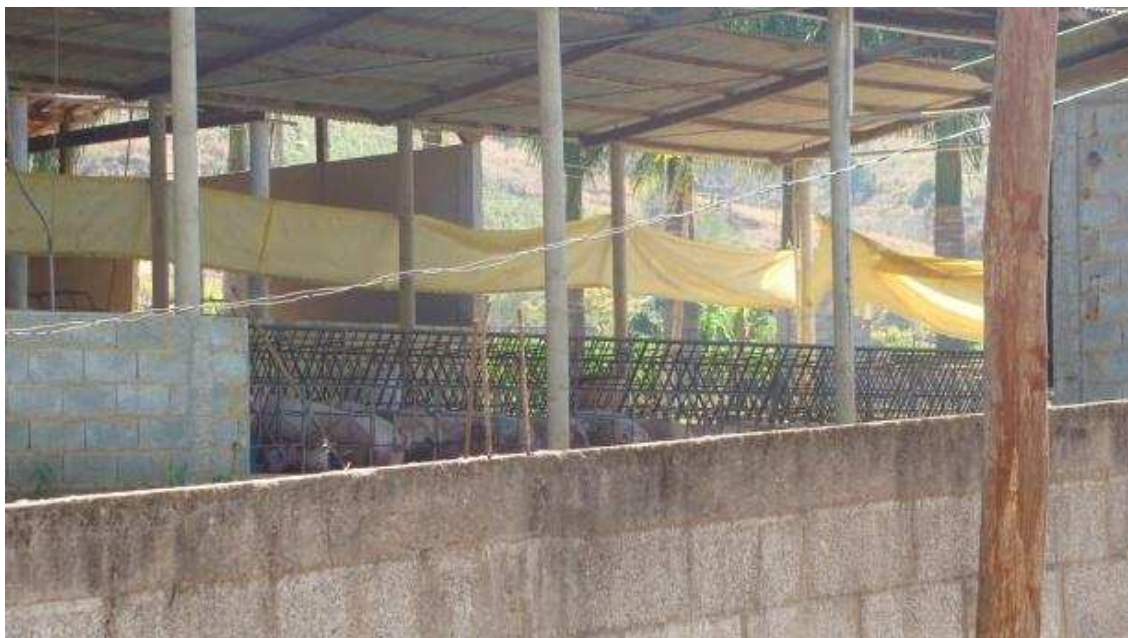


Figura 11 - Granja de criação de suínos na sub bacia do Córrego do Engenho.



Figura 12 - Lançamento de efluentes da suinocultura no Córrego do Engenho.

No segundo fator os valores foram positivos para cor (0,82), nitrogênio total (0,77) e turbidez (0,73) (Tabela 10). Turbidez e cor são variáveis indicadoras do aumento do processo erosivo na bacia.

O nitrogênio é o nutriente exigido em maior quantidade pelas plantas. Sua utilização na agricultura pode ser pela aplicação de fertilizantes ou de compostos orgânicos nitrogenados. Nessa bacia, os agricultores são de baixo poder aquisitivo e por isso esse nutriente pode está relacionado a um maior uso de adubos orgânicos nas propriedades agrícolas em relação aos fertilizantes minerais.

O coeficiente de correlação entre áreas agrícolas e turbidez foi positivo ( $r=0,24$ ,  $p < 0,05$ ), sendo portanto, um bom indicativo dessa situação. Esses resultados são indicativos do carreamento de partículas do solo das áreas agrícolas pela erosão contribuindo para o aumento da concentração de nitrogênio na água.

Resultados similares foram encontrados por Simeonov et al., (2003) e Andrade et al., (2007) que observaram a ocorrência de alterações na qualidade das águas superficiais, causadas pelo escoamento superficial de áreas agrícolas com elevada concentração de nitrogênio.

No terceiro e último fator, as variáveis com maiores cargas fatoriais foram pH (-0,68) e alumínio (0,84) (Tabela 10). De acordo com a Cetesb (2009), na água, o alumínio pode ocorrer em diferentes formas, sendo diretamente influenciado pelas variações do pH, temperatura e presença de fluoretos, sulfatos, matéria orgânica e outros ligantes, sendo sua solubilidade baixa em pH entre 5,5 e 6,0.

Tabela 10 – Matriz de cargas rotacionadas ortogonalmente para o IAP.

Variáveis	Fatores Rotacionados		
	F1	F2	F3
OD	-0,60	-0,36	0,19
Ntot	-0,06	0,77	-0,27
Ftot	0,70	0,21	0,04
Turb	0,37	0,73	0,24
pH	-0,35	-0,15	-0,68
Al	-0,26	-0,09	0,84
Fe	0,74	-0,01	0,28
Mn	0,76	-0,04	-0,10
Cor	0,08	0,82	0,13
% Variância explicada	24,85	22,40	15,95
% Variância acumulada no fatores	24,85	47,25	63,20
Autovalor	2,83	1,51	1,34

#### 4.1.4.7. Modelagem do IQA<sub>Sb</sub> pelo método de Bartlett

Na tabela 11 são apresentados os resultados da estatística descritiva para o IQA São Bartolomeu calculado pelo escore fatorial pelo método de Bartlett. De acordo com Haase et al., (1989); Toledo e Nicolella, (2002) o índice predito por esse método deve ter média zero e variância igual a um e quanto trabalho e quanto mais negativos os valores dos escores fatoriais melhor será a qualidade da água. Os valores dos escores fatoriais calculados pela AF / CP pelo método de Bartlett foram considerados como índices da qualidade da água, sendo interpretados individualmente em cada fator.

Tabela 11 – Estatística descritiva dos índices calculados pelo método de Bartlett.

Índices	Observações	Mínimo	Máximo	Média	Variância
IQA <sub>Sb</sub> F1	96	-1,47	6,86	0,00	1,00
IQA <sub>Sb</sub> F2	96	-1,80	2,26	0,00	1,00
IQA <sub>Sb</sub> F3	96	-2,10	3,07	0,00	1,00

Os valores dos escores fatoriais , pelo método de Bartlett, que representam o índice da qualidade da água para o primeiro fator (IQA<sub>Sb</sub>F1) variaram de -0,42 a 0,70, de -0,79 a 0,73 para o segundo (IQA<sub>Sb</sub>F2) e de -1,03 a 1,01 para o terceiro (IQA<sub>Sb</sub>F3) (Figuras 13, 14 e 15).

Observa-se que o modelo do IQA ficou bem ajustado para os dois primeiros fatores, no terceiro, houve uma inversão, onde os valores mais positivos foram os de melhor qualidade, possivelmente por apresentar apenas duas variáveis com maior valor de cargas fatoriais Alumínio e pH.

Tomando-se com exemplo, comparamos os valores dos escores fatoriais com os do IQA calculado pelo método Cetesb, observa-se que na sub bacia do córrego do Engenho IQA<sub>Sb</sub> para o primeiro fator apresentou o maior valor positivo (0,70) e um dos piores desempenhos de qualidade da água com IQA igual a 50. Já na bacia do São Lucas onde o escore fatorial (-0,42) foi o mais negativo, o IQA calculado pela Cetesb apresentou o melhor valor de qualidade da água (62) (Figura 13). Resultados semelhantes foram observados no segundo fator (Figura 14).

Avnalizando a matriz de correlação entre IQA Cetesb e IQA calculado pelos escores fatoriais (Tabela 12), nota-se a existência de correlações negativas, de maior valor de coeficiente de correlação, entre o IQA Cetesb e o IQASbF1 (-0,42,  $p < 0,01$ ) e com o IQASbF2 (-0,42,  $p < 0,01$ ). Já com o IQASbF3 ( $r = 0,27$ ,  $p < 0,01$ ) os valores foram menores e positivos.

O terceiro fator apresentou comportamento inverso aos dois anteriores, possivelmete por ter apenas duas das sete variáveis de qualidade da água explicada nesse fator (Figura 15), o que pode indicar não ter ocorrido um bom ajuste do modelo de cálculo do IQA pelo método de Bartlett para esse fator.

Tabela 12 – Matriz de correlação do IAP Cetesb e IQA modelado pela ACP.

Variáveis	IQA Cetesb	IQASbF1	IQASbF2
IQASbF1	<b>-0,42</b>		
IQASbF2	<b>-0,74</b>	0,00	
IQASbF3	<b>0,27</b>	0,00	0,00

Valores em negrito expressam as correlações significativas ao nível de 1% de probabilidade

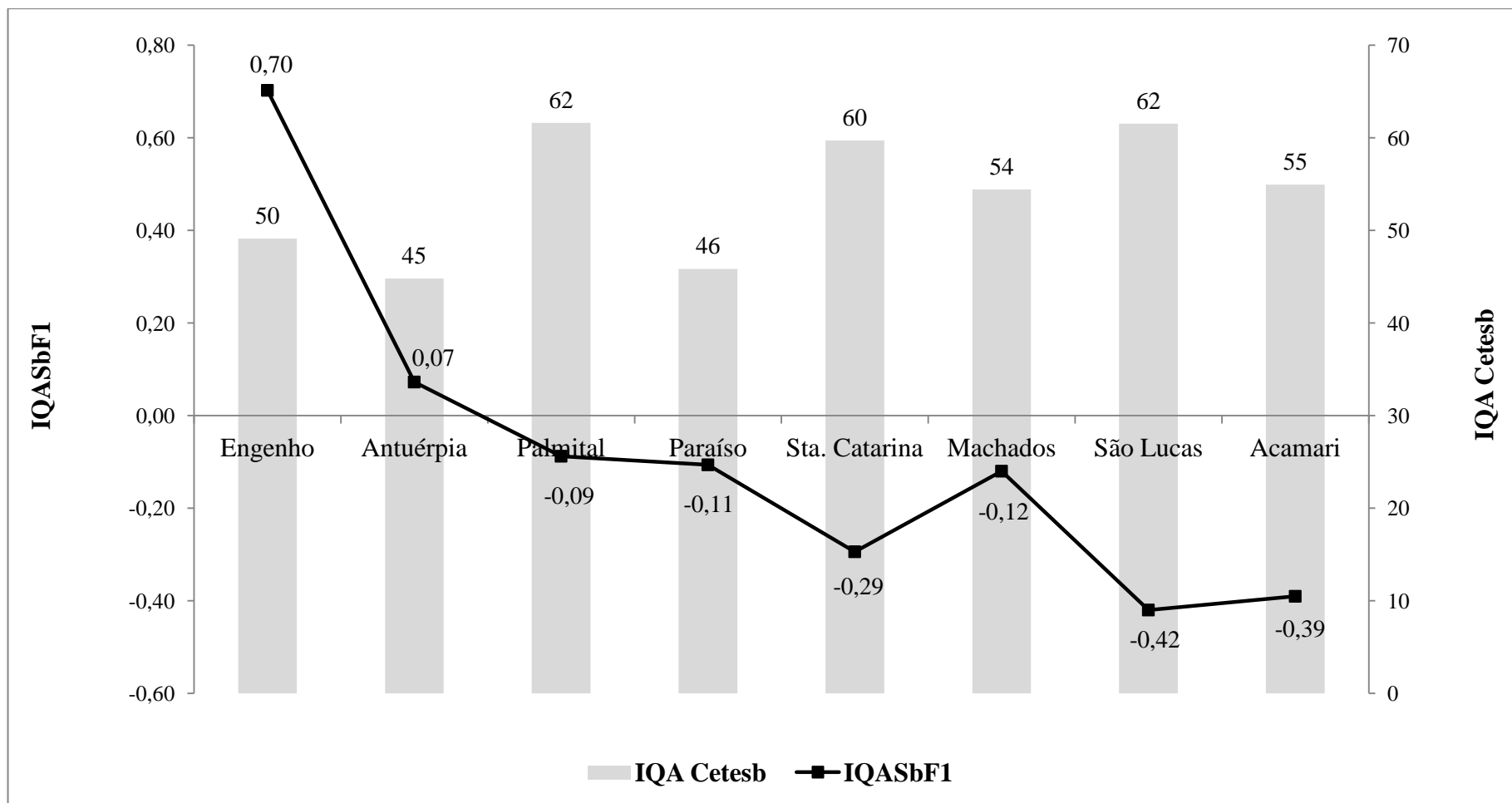


Figura 13 - Distribuição dos valores médios do IQA Cetesb comparados com os modelados pelo método de Bartlett (IQAS<sub>b</sub>F1) nas oito sub-bacias durante doze meses.

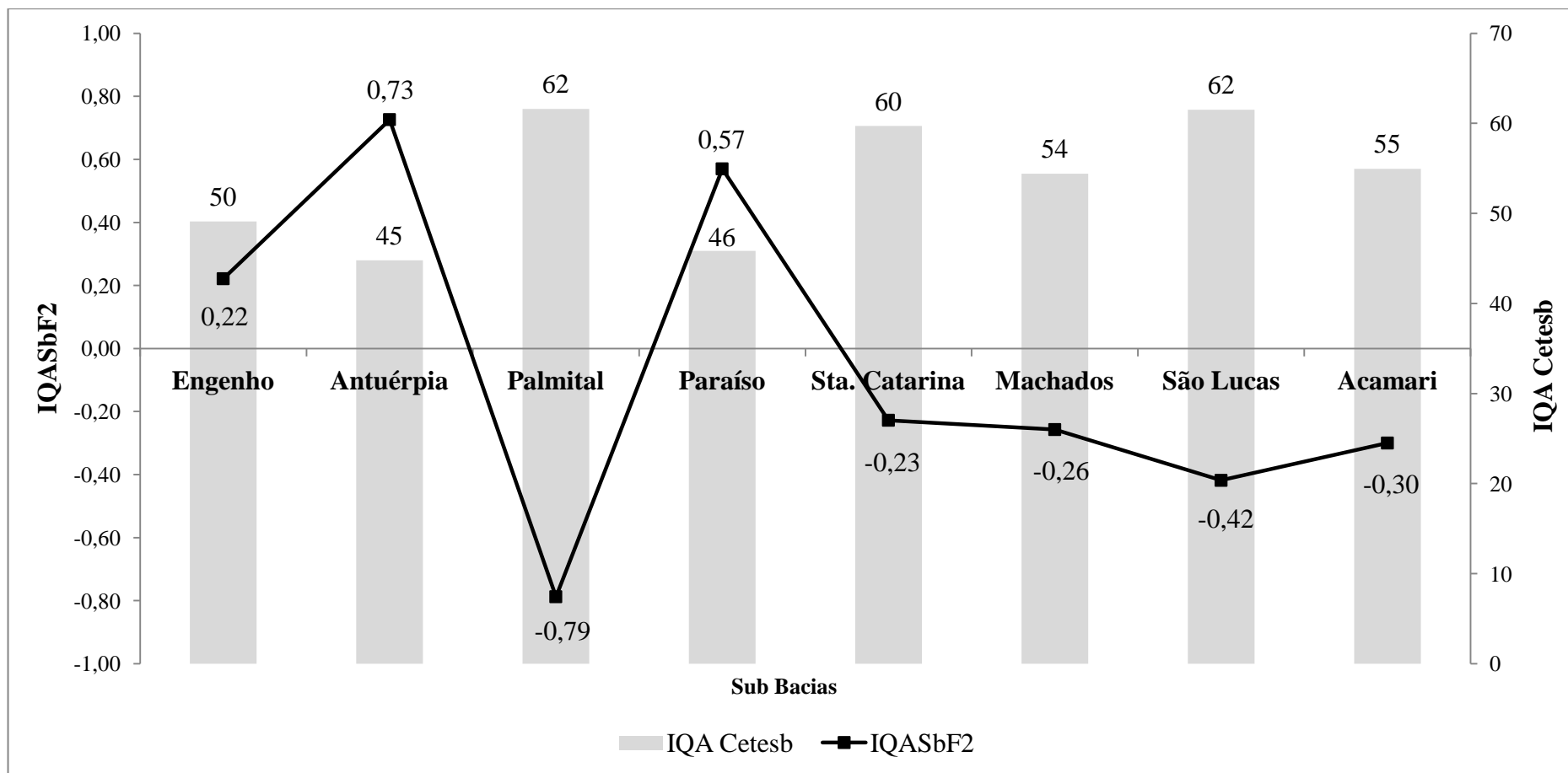


Figura 14 - Distribuição dos valores médios do IQA Cetesb comparados com os modelados pelo método de Bartlett (IQAS<sub>b</sub>F<sub>2</sub>) nas oito sub bacias durante doze meses.

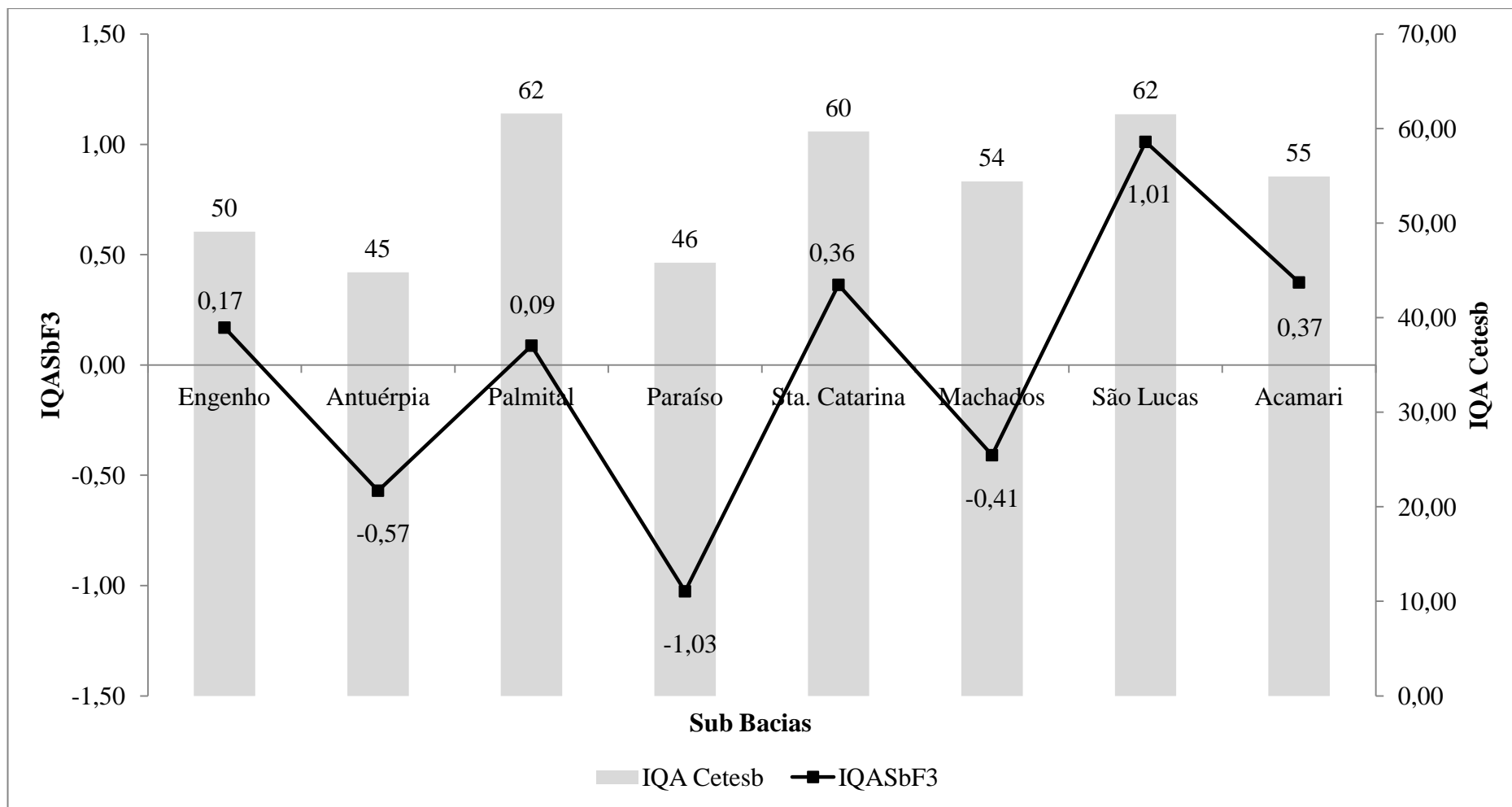


Figura 15 - Distribuição dos valores médios do IQA Cetesb comparados com os modelados pelo método de Bartlett (IQAS<sub>b</sub>F3) nas oito sub bacias durante doze meses.

#### 4.1.4.8. Modelagem do IAP<sub>Sb</sub> pelo método de Bartlett

Na tabela 13 são apresentados os resultados da estatística descritiva para o IAP São Bartolomeu calculado pelo escore fatorial pelo método de Bartlett.

Tabela 13 – Estatística descritiva dos índices calculados pelo método de Bartlett.

Índices	Observações	Mínimo	Máximo	Média	Variância
IAP <sub>Sb</sub> F1	96	-0,81	0,92	0,00	1
IAP <sub>Sb</sub> F2	96	-0,55	0,55	0,00	0,16
IAP <sub>Sb</sub> F3	96	-0,42	0,25	0,00	0,05

Os modelos de cálculo do IQA pelo método de Bartlett não ficaram bem ajustados para os três fatores, apenas o primeiro fator apresentou coeficientes de correlações significativas ao nível de 5% de probabilidade entre o IQA Cetesb e IQA modelado ( $r = -0,31$ ,  $p < 0,01$ ) (Tabela 14).

Tabela 14 – Matriz de correlação do IAP Cetesb e IAP modelado pela ACP.

Variáveis	IAPCetesb	IAPSbF1	IAPSbF2
IAPSbF1	<b>-0,31</b>		
IAPSbF2	-0,16	0,02	
IAPSbF3	-0,10	0,03	0,02

Valores em negrito expressam correlações significativas ao nível de 1% de probabilidade.

Os índices calculados, pelo método de Bartlett, para o primeiro fator (IQASbF1) variaram de -0,81 a 0,92 e de acordo com Haase et al., (1989) quanto mais negativo for o valor do escore fatorial, melhor será a qualidade da água (Figuras 16).

Tomando-se com exemplo, comparamos os valores dos escores fatoriais com os do IAP calculado pelo método Cetesb, observa-se que na sub bacia do córrego do São Lucas o valor escore fatorial foi o mais negativo (-0,81) e o IQA calculado pela Cetesb apresentou o melhor de qualidade da água (62). Equanto, na bacia do córrego do Engenho IQASb para o primeiro fator apresentou o maior valor positivo (0,92) e um dos piores desempenhos de qualidade da água com IQA Cetesb igual a 50 (Figura 16).

Por essa análise, observa-se que as piores qualidades da água nas bacias do córrego do Engenho, Antuérpia e Paraíso e a melhor qualidade nas sub bacias do Palmital, Santa Catarina, São Lucas e Acamari e conseqüentemente os valores mais negativos de escores fatoriais foram observados nessas sub bacias (Figura 16).

Nas duas primeiras, existem suinoculturas que lançam efluentes no curso d'água sem tratamento, enquanto, na bacia do São Lucas os esgotos do bairro Romão do Reis são coletados e tratados, o que contribui significativamente para melhoria da qualidade da água.

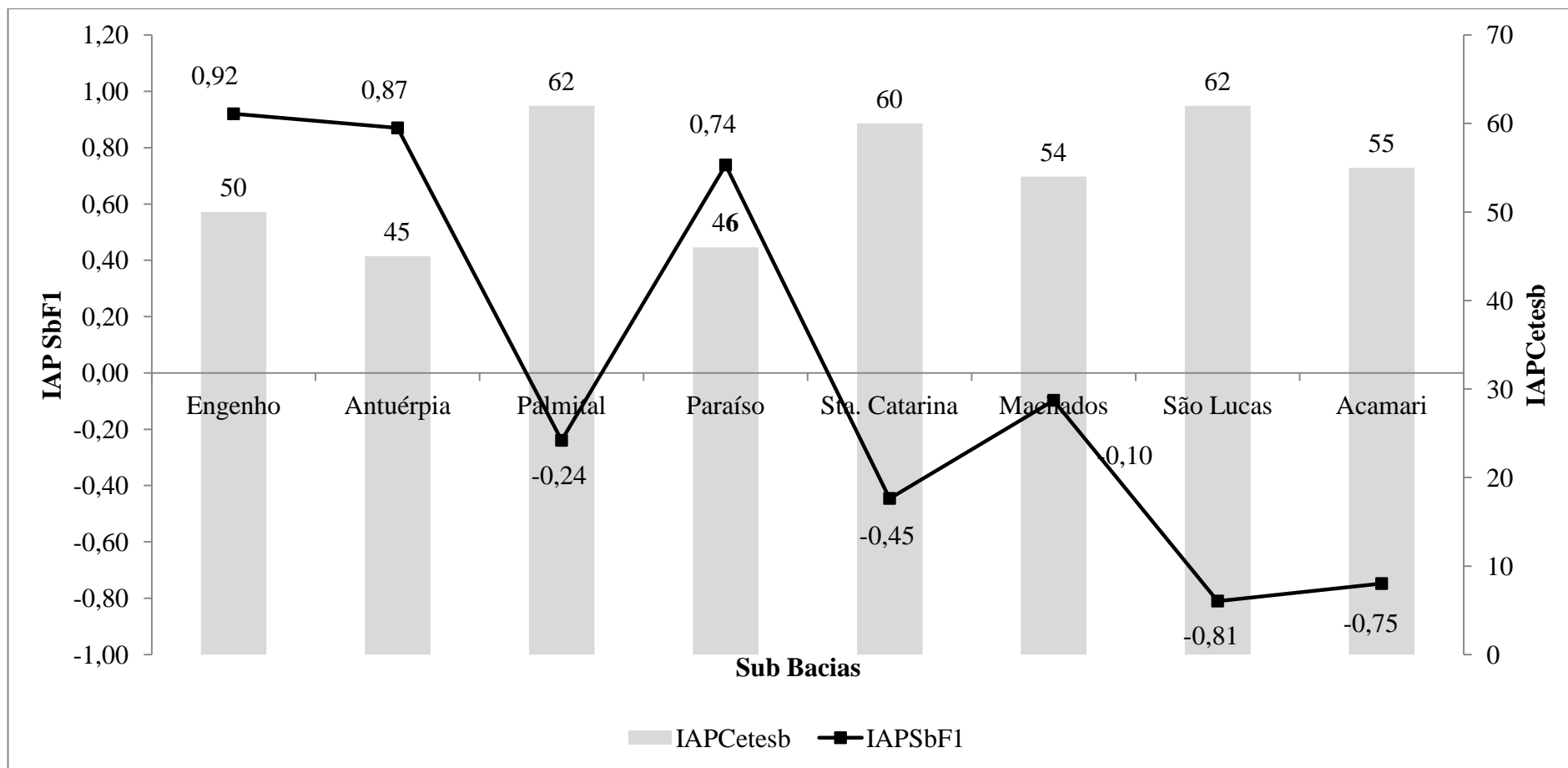


Figura 16 - Distribuição dos valores médios do IAP Cetesb comparados com os modelados pelo método de Bartlett ( $IAP_{SbF1}$ ) nas oito sub-bacias durante doze meses.

## 4.2. Uso de redes neurais na modelagem do IQAS<sub>b</sub> e IAPS<sub>b</sub>

O uso de redes neurais artificiais foi sistematizado em três etapas: escolha das melhores arquiteturas, seleção de variáveis e desenvolvimento de modelos de predição e de avaliação dos índices.

### 4.2.1. Arquiteturas de redes neurais selecionadas para IQAS<sub>b</sub> e IAPS<sub>b</sub>

Foram testadas inicialmente varias arquiteturas de redes, das quais foram selecionadas, oito para o IQAS<sub>b</sub> e dez para IAPS<sub>b</sub>, por apresentaram os menores erros quadráticos médios nas fases de treinamento e validação (Tabelas 15 e 16).

A rede sete foi a de melhor desempenho para IQAS<sub>b</sub> (Figura 17), com erro de quadrático médio de 1,77, 1,63 e 1,63 nas fases de treinamento, validação e teste, respectivamente. Para a rede 10 do IAPS<sub>b</sub> (Figura 18) os resultados foram de 3,33, 3,01 e 3,01 para treinamento, validação e teste (Tabelas 15 e 16).

Além destas, na composição dos modelos de predição do IQAS<sub>b</sub> foram selecionadas as redes cinco, seis e oito. Nesta fase foram consideradas apenas as arquiteturas que apresentaram menores erros de calibração, ou seja, a menor diferença entre a validação e o teste e um número de pelo menos três variáveis na camada de entrada.

Tabela 15 - Redes neurais selecionadas para o IQAS<sub>b</sub>

Redes	Erro Médios Quadráticos			Arquitetura		
	Treinamento	Validação	Teste	Camadas		
				Entrada	Intermediária	Saída
R1	8,72	8,81	9,15	1	1	1
R2	8,70	8,79	9,15	1	2	1
R3	6,69	6,30	6,30	1	3	1
R4	6,67	6,27	6,27	1	12	1
R5	4,82	2,93	2,95	3	4	1
R6	2,54	2,27	2,28	4	6	1
R7	1,77	1,63	1,63	5	7	1
R8	1,77	1,74	1,74	7	10	1

Tabela 16 - Redes neurais selecionadas para o IAP<sub>Sb</sub>

Redes	Erros Quadráticos Médios			Arquiteturas		
	Treinamento	Validação	Teste	Camadas		
				Entrada	Intermediária	Saída
R1	8,19	6,95	7,30	6	2	1
R2	4,78	4,95	4,47	7	10	1
R3	6,15	4,57	4,30	7	19	1
R4	5,33	4,12	4,37	8	4	1
R5	4,81	4,02	3,71	9	18	1
R6	3,59	3,82	2,79	11	8	1
R7	4,08	3,63	4,47	11	14	1
R8	3,77	3,28	3,30	12	16	1
R9	3,96	3,05	3,12	13	14	1
R10	3,33	3,01	3,01	16	12	1

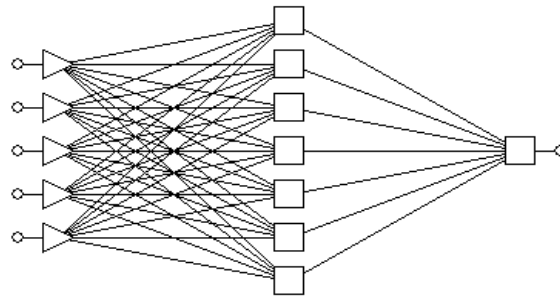


Figura 17 - Arquitetura da rede de melhor desempenho para o IQA<sub>Sb</sub>.

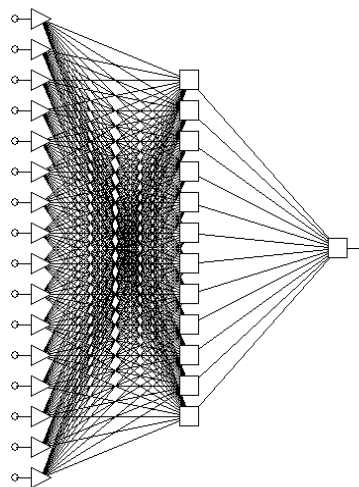


Figura 18 - Arquitetura da rede de melhor desempenho para o IAP<sub>Sb</sub>.

#### 4.2.2. Variáveis selecionadas para IQASb e IAPSb

As redes selecionadas para a predição do IQASb e IAPSb apresentaram o número de variáveis, neurônios, na camada de entrada inferior ao número de variáveis definidas pela Cetesb para a composição do IQA e IAP (Tabela 17).

Tabela 17 - Variáveis selecionadas pelas redes neurais para o IQASb

Índices	Redes	Variáveis Selecionadas
IQASb	R5	OD, ST e CT
	R6	OD, Ntot, Turb e CT
	R7	OD, DBO, Ntot, Turb e CT
	R8	OD, DBO, Ntot, Turb, ST, pH e CT
IAPSb	R10	OD, DBO, Ntot, Ptot, Turb, pH, Ctot, Cd, Cr, Giard, Crypto, Al, Cu, Fe, Mn e Zn

No software *statisca neural networks* não há a necessidade de normalização dos dados, mesmo assim foram testadas a análise com dados normalizados e não normalizados tendo sido obtidos resultados semelhantes. O software foi capaz de realizar a predição dos resultados de saída com apenas algumas das variáveis de entrada, selecionadas para o modelo por apresentaram os menores erros quadráticos médios individuais. Esse procedimento pode reduzir tempo de análise, de publicação resultados, dos custos operacionais e de tomada de decisão pelos gestores responsáveis pelo gerenciamento dos recursos hídricos.

As RNAs possuem capacidade de aprender por exemplos, baseadas nos padrões fornecidos, de classificar, de estimar e de selecionar características específicas dentro de um conjunto de dados (HAYKIN, 2001; SINGH et al., 2009).

Esses resultados mostram que as RNAs selecionaram as variáveis mais representativas para a composição dos índices da qualidade da água na bacia do ribeirão São Bartolomeu, além de incluírem novas variáveis como giárdia e *Cryptosporidium* que não fazem parte das variáveis adotadas nos programas de monitoramento da qualidade da água.

A adoção desses índices modelados pelo uso de redes neurais artificiais ainda carece de ajustes e sua utilização na bacia do ribeirão São Bartolomeu depende de mais

estudos e comparações com os resultados obtidos pelo método convencional desenvolvido pela Cetesb. No entanto, os resultados mostram que as redes neurais artificiais é uma ferramenta que bem ajustada as condições da bacia pode ser adotada pelo programa de monitoramento da qualidade da água.

#### **4.2.3. Modelos de predição para IQA<sub>Sb</sub> pela redes neurais e IAPSb e avaliação de desempenho.**

Foram utilizados os 350 conjuntos dados, divididos em três sub conjuntos: de treinamento 176 (50%), validação 87 (25%) e teste 87 (25%). Esses sub conjuntos foram obtidos aleatoriamente. O número de neurônios de entrada foram as variáveis de qualidade da água e o de saída um único número que variou de acordo com as classes de qualidade da água numa escala de 0 a 100.

Nesta etapa avaliou-se os resultados dos índices calculados pelo método Cetesb com os obtidos pelos modelos de predição das redes neurais artificiais. A rede neural R8 foi a que apresentou o melhor desempenho para do IQA<sub>Sb</sub>, com valor de coeficiente de determinação ( $R^2$ ) igual a 0,99, do coeficiente de eficiência (CE) de 0,99 e do erro padrão de estimativa (EPE) igual a 2,36. Já a rede R5 foi considerada como a de pior desempenho entre as quatro selecionadas para o modelo de predição do IQA<sub>Sb</sub> (Tabela 18). No entanto, mesmo apresentando esse baixo desempenho, teve a capacidade de predição do IQA com apenas três variáveis (OD, ST e CT), situação que não deve ser testada para averiguação de sua aplicabilidade as condições da bacia.

Segundo Braga (2007), um número excessivo de neurônios na camada intermediária ou oculta pode levar à memorização dos resultados pela rede, enquanto que poucos neurônios podem comprometer a capacidade de generalização. A rede 8 possui 10 neurônios na camada oculta e a rede 5 quatro, essa configuração pode ter influenciado para uma melhor capacidade de generalização da rede 8 quando comparada com a rede 5.

Tabela 18 - Avaliação de desempenho das redes neurais utilizadas na predição do IQA<sub>Sb</sub>

Redes	R <sup>2</sup>	Erro Padrão da Estimativa (EPE)	Coefficiente de Eficência (E)
R5	0,95	5,01	0,95
R6	0,97	4,13	0,98
R7	0,98	2,45	0,99
R8	0,99	2,36	0,99

Para o IAPSb a rede neural dez apresentou valores de coeficiente de determinação (R<sup>2</sup>) igual a 0,97, do coeficiente de eficiência (CE) de 0,97 e do erro padrão de estimativa (EPE) igual a 5,17 (Tabela 19).

Tabela 19 - Avaliação de desempenho da rede neural 10 utilizada na predição do IAP<sub>Sb</sub>

Redes	R <sup>2</sup>	Erro Padrão da Estimativa (EPE)	Coefficiente de Eficência (E)
R10	0,97	5,17	0,97

Utilizou-se o método estratificado por amostragem, sendo o conjunto de testes dados divididos em extratos aleatórios de classes de qualidade da água. Procurou-se extrair na amostra grupos que representassem minimamente as cinco classes de qualidade da água. No entanto, essa condição não foi plenamente atendida, sendo que em algumas situações obtiveram grupos maiores dentro uma classe ou faixa de classificação em relação a outras, o que comprometeu uma distribuição uniforme, levando a presença de espaços vazios. Comportamentos semelhantes também foram observados nos trabalhos de Cerqueira et al. (2001); Menezes et al. (2011); Najah et al. (2009).

Observa-se a inexistência de resultados preditos para IQASb e IAPSb e de IQA calculados pelo método Cetesb para as faixas entre 12 e 25, 35 e 40, 45 e 50 e 70 e 80.

Nas figuras 19, 20, 21 e 22 são apresentadas as comparações entre os resultados do IQA calculados pela Cetesb com os preditos pelos modelos das redes neurais.

Observa-se que os dados, do IQAS<sub>b</sub>, calibrados pelos modelos desenvolvidos pelas redes neurais, R7 e R8, oito estão próximos aos calculados pela Cetesb e distribuídos junto da linha de 45°. Já em relação as redes R5 e R6 observa-se uma maior dispersão desses dados, mostrando a maior eficiência de predição do IQAS<sub>b</sub> pelas redes sete e oito.

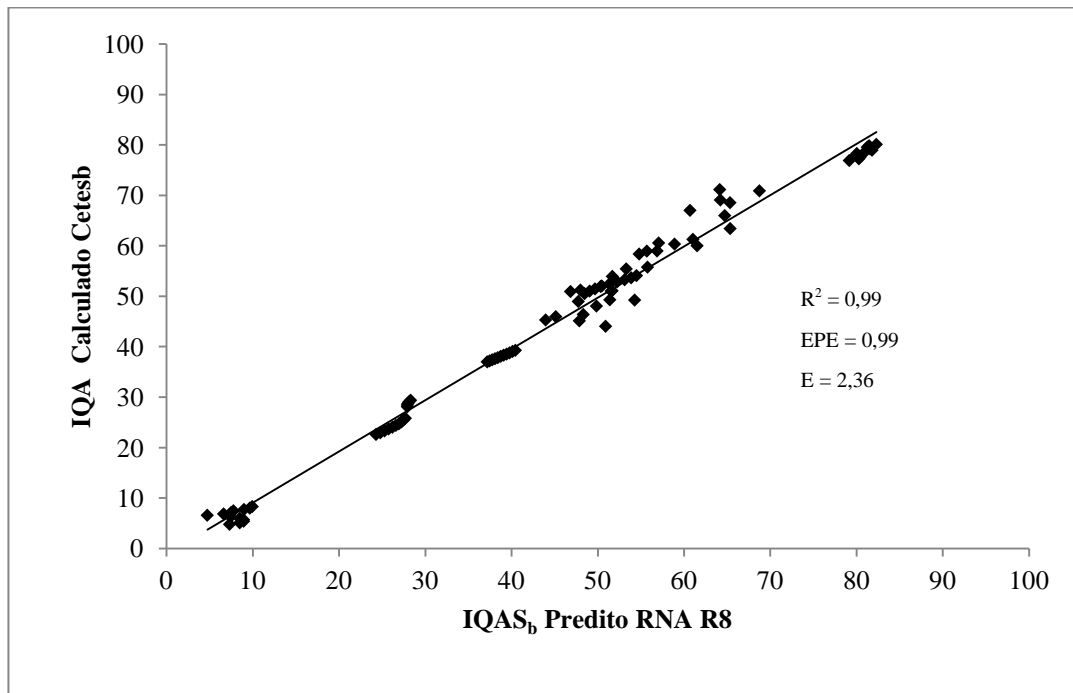


Figura 19 - Valores do IQA Cetesb versus IQAS<sub>b</sub> preditos pela rede neural R8.

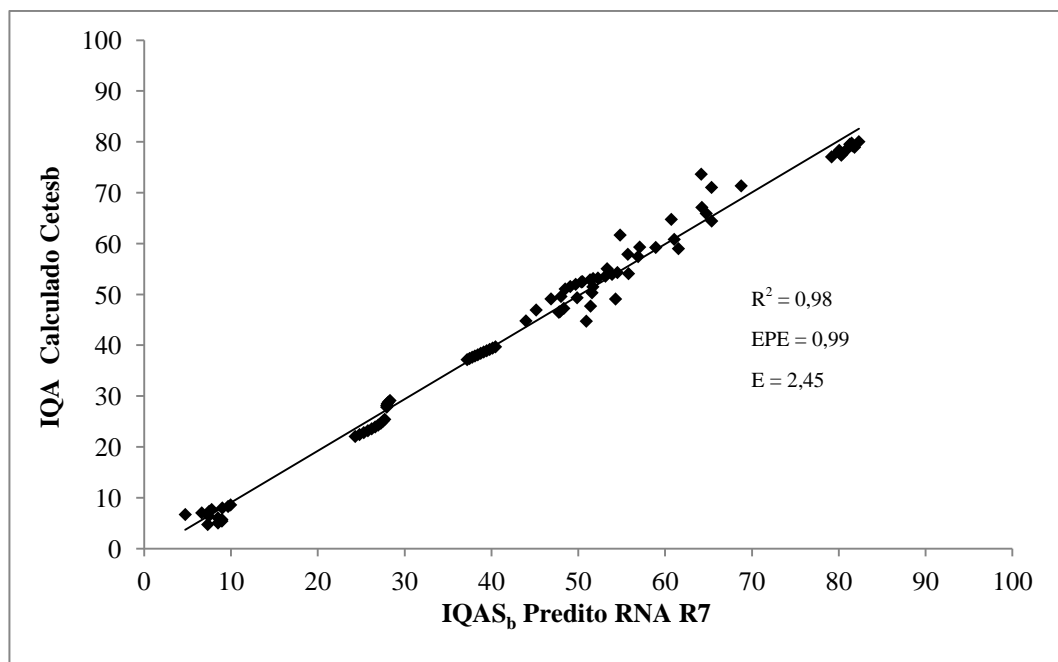


Figura 20 - Valores do IQA Cetesb versus IQAS<sub>b</sub> preditos pela rede neural R7.

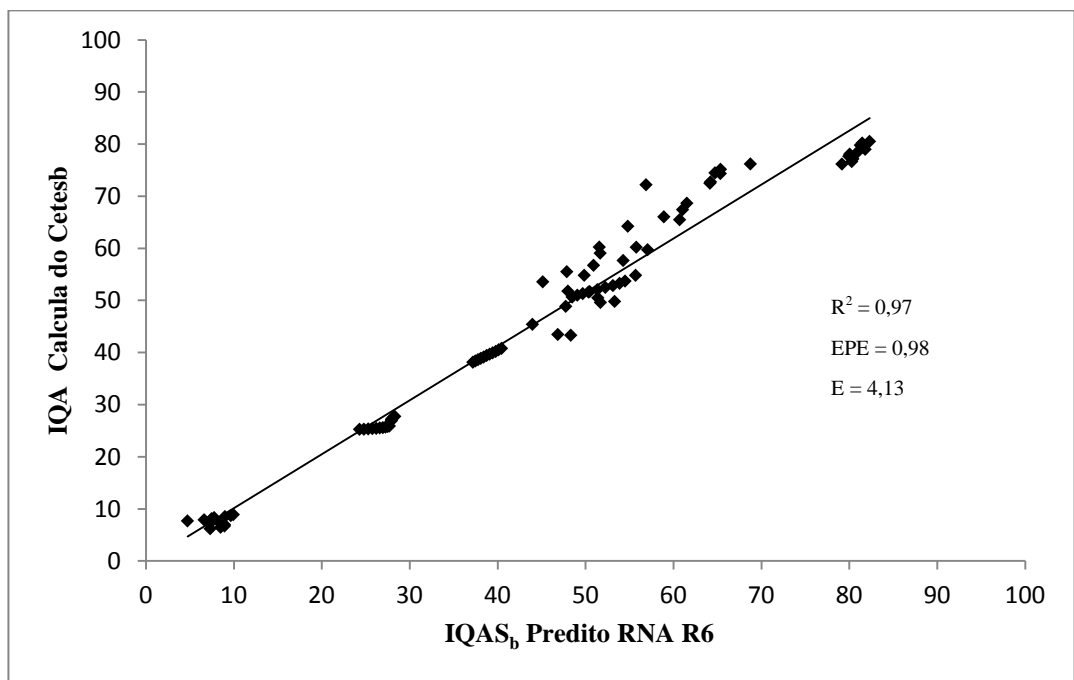


Figura 21 - Valores do IQA Cetesb versus IQA<sub>Sb</sub> preditos pela rede neural R6.

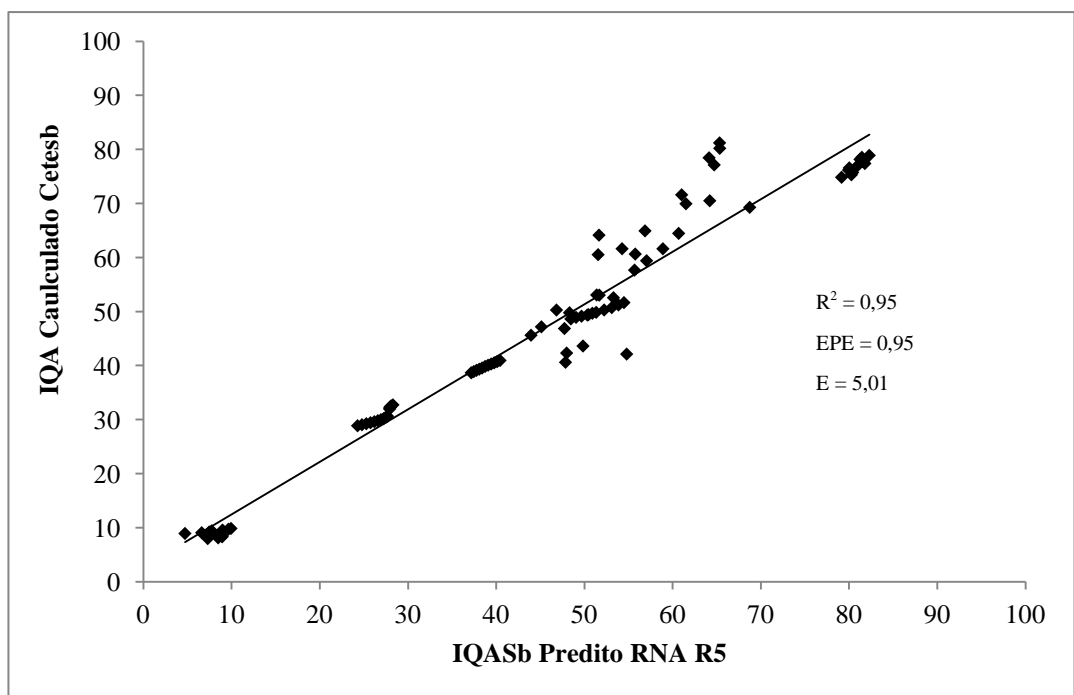


Figura 22 - Valores do IQA Cetesb versus IQA<sub>Sb</sub> preditos pela rede neural R5.

Com relação ao  $IAP_{Sb}$  os resultados calculados pela Cetesb versus os preditos pelo modelo desenvolvido pela rede neural R10 estão próximos e distribuídos ao longo da linha de  $45^\circ$  (Figura 23), mostrando uma boa capacidade de predição do modelo desenvolvido pela rede neural dez.

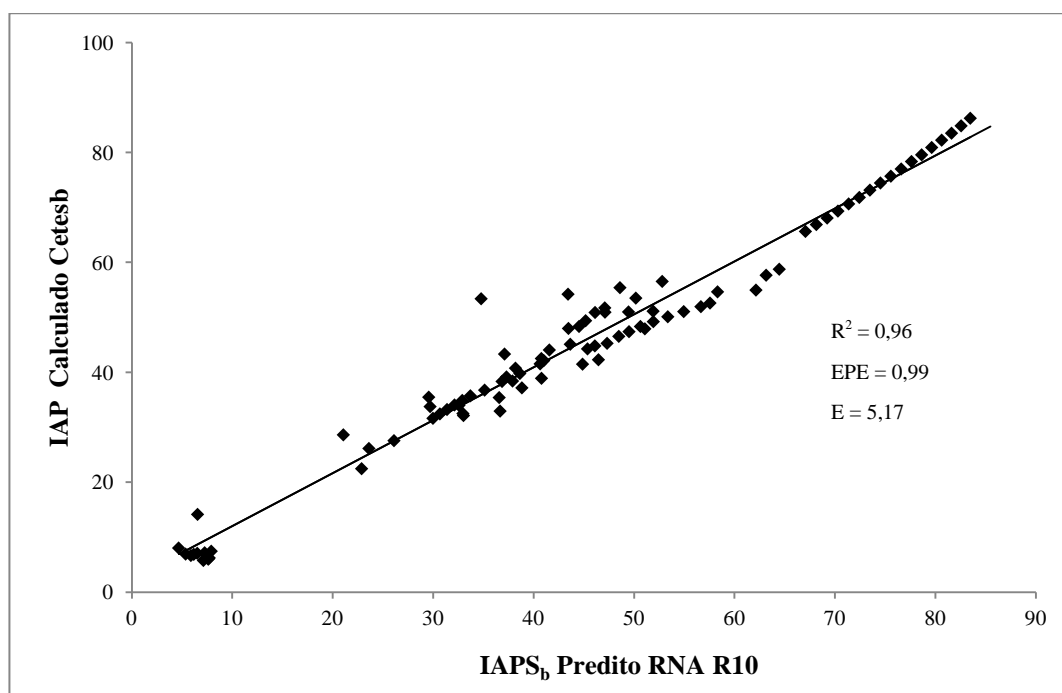


Figura 23 - Valores do IAP Cetesb versus  $IAP_{Sb}$  preditos pela rede neural R5.

Os resultados mostram que a modelagem desses índices pelo emprego de redes neurais apresentou uma boa capacidade de predição para  $IQA_{Sb}$  e  $IAP_{Sb}$ , a partir de um conjunto de dados desconhecidos pela rede, neural e da possibilidade de utilização dessa ferramenta na rotina dos programas de monitoramento.

### **4.3. Uso e ocupação do solo e variáveis de qualidade da água**

Esse tópico foi organizado com a discussão dos resultados do teste de normalidade, análise do uso e ocupação do solo e das associações destes com as variáveis de qualidade da água. Para avaliar a influência do uso e ocupação do solo nas variáveis de qualidade da água na bacia de São Bartolomeu efetuou-se um diagnóstico do uso e ocupação do solo, aplicou-se o teste de normalidade de Kolmogorov-Smirnov, e, por fim, foi realizada uma correlações entre usos ocupações e variáveis de qualidade da água.

#### **4.3.1. Teste de normalidade**

Os resultados do teste de normalidade de Kolmogorov-Smirnov, para os dados de uso e ocupação do solo e variáveis de qualidade da água, apresentaram distribuição normalidade ao nível de 5% de significância.

#### **4.3.2. Diagnóstico do uso e ocupação do solo na bacia do ribeirão São Bartolomeu**

De acordo com os resultados apresentados na tabela 20, observa-se que formações florestais (32,88%) e pastagem não degradadas (33,89%) são os dois usos e ocupação do solo com o maior percentual de ocupação de área na bacia do ribeirão São Bartolomeu, sendo os melhor contribuem para uma melhor proteção do solo.

Segundo Adamoli et al. (1986) os solos ocupados por pastagens quando manejados adequadamente, sem excesso de animais por hectare, com manutenção da fertilidade e melhoria das propriedades físicas do solo contribuem para apresentam diminuição do aporte de sedimentos e melhorias da qualidade da água.

O avanço das áreas de agricultura, áreas urbanas, café, pastagem degradada e piscicultura apresentam efeito adverso ao que foi apresentado anteriormente, intensificam o processo erosivo, aumentam o escoamento superficial, a deposição de material orgânico e mineral e conseqüentemente contribuem para perda de água em quantidade e qualidade.

A bacia do ribeirão São Bartolomeu apresenta estrutura fundiária caracterizada por pequenas propriedades rurais que desenvolvem agricultura de subsistência e criação de bovinos para produção de leite (BEZERRA, 2011). A baixa capacidade de investimento em uso de sistemas extensivos de criação de animais, baixa aquisição de

fertilizantes para reposição dos nutrientes reirados da área e a falta de orientação técnica, podem estar contribuindo para que cerca de 16%% das áreas da bacia esteja ocupadas por pastagens degradadas (20)

O restante da área da bacia está ocupada pelo cultivo de café com 11,14%, agricultura 3,88%, uso urbano com 1,93% e piscicultura com 0,34% (Tabela 20). A Figura 23 apresenta a distribuição dos usos e ocupação do solo e a Figura 24 mostra os diferentes valores de os índices da qualidade da água por sub bacias do ribeirão São Bartolomeu, destes três sub bacias apresentaram qualidade regular ( $36 \leq IQA \leq 51$ ) e cinco foram classificadas como boa ( $51 \leq IQA \leq 79$ ).

As sub bacias do córrego Palmital, São Lucas e Santa Catarina, e apresentaram os melhores valores de índice de qualidade da água, ao passo que, Antuérpia, Paraíso e Córrego do Engenho os piores valores de IQA (Figura 24).

As condições naturais e atividades antrópicas são os principais fatores que influenciam a qualidade da água em uma bacia hidrográfica. Considerando-se o segundo fator, observa-se que no córrego Santa Catarina cerca de 72% (223/309) de sua área estão ocupadas por formações florestais, já nos córregos São Lucas e Palmital predominam a ocupação por pastagem não degradadas, que são da ordem de 44,99% (101/225) e 34,56% (193/559), respectivamente (Tabela 21). Esse comportamento está favorecendo uma melhor da qualidade da água, como pode ser avaliado pelos maiores valores de IQA de 62 (Palmital e São Lucas) e 60 para a sub bacia do córrego Santa Catarina (Figura 24).

Do lado oposto, estão as sub bacias do Antuérpia, Paraíso e córrego do Engenho com os piores resultados para o índice da qualidade distribuídos da seguinte maneira: 45, 46 e 50, respectivamente.

Os piores indicadores de qualidade da água estão sendo influenciados pelos maiores percentuais de pastagem degradada 33,85% (46/155) na sub bacia do Antuérpia e no córrego do Engenho 47,77% (60/126). Já na sub bacia do Paraíso, café 31% (60/190) foi o uso de maior influência na perda de qualidade da água (Tabela 21).

A piscicultura, que mesmo sendo o uso de menor percentual de ocupação de área na bacia 0,34% (8,51/2464), tem um papel importante na perda de qualidade da água, pois ocupa cerca de 1,0% (1,25/126) da área do córrego do Engenho, 0,34% (0,46/135) do antuérpia e 0,60% (1,12/190) do córrego do Paraíso (Tabela 21).

Sendo uma importante fonte de aporte de material orgânico, correlação negativa com oxigênio dissolvido ( $r = -0,24$ ,  $p < 0,05$ ), fósforo total ( $r = 0,50$ ,  $p < 0,01$ ) e nitrogênio ( $r = 0,0$ ,  $p < 0,01$ ).

Em praticamente todas as propriedades, da bacia, existem pequenos criatórios de peixes que são abastecidos por nascentes ou minas e que despejam seus efluentes diretamente nos cursos d'água sem nenhum tipo de tratamento. Os efluentes produzidos por essa atividade estão aumentando o aporte de material orgânico, nutrientes e o consumo de oxigênio dissolvido pelas bactérias responsáveis pela mineralização da matéria orgânica.

Tabela 20 - Uso e ocupação do solo na bacia do ribeirão São Bartolomeu.

Uso e Ocupação do Solo	Área (ha)	Área (%)
Agricultura (culturas anuais)	96,06	3,88
Área Urbana	47,90	1,93
Café	275,78	11,14
Formações Florestais	814,00	32,88
Pastagem Degradada	383,03	15,47
Pastagem não Degradada	839,07	33,89
Pisicultura (hidrografia)	8,51	0,34
Total	2464,35	99,54

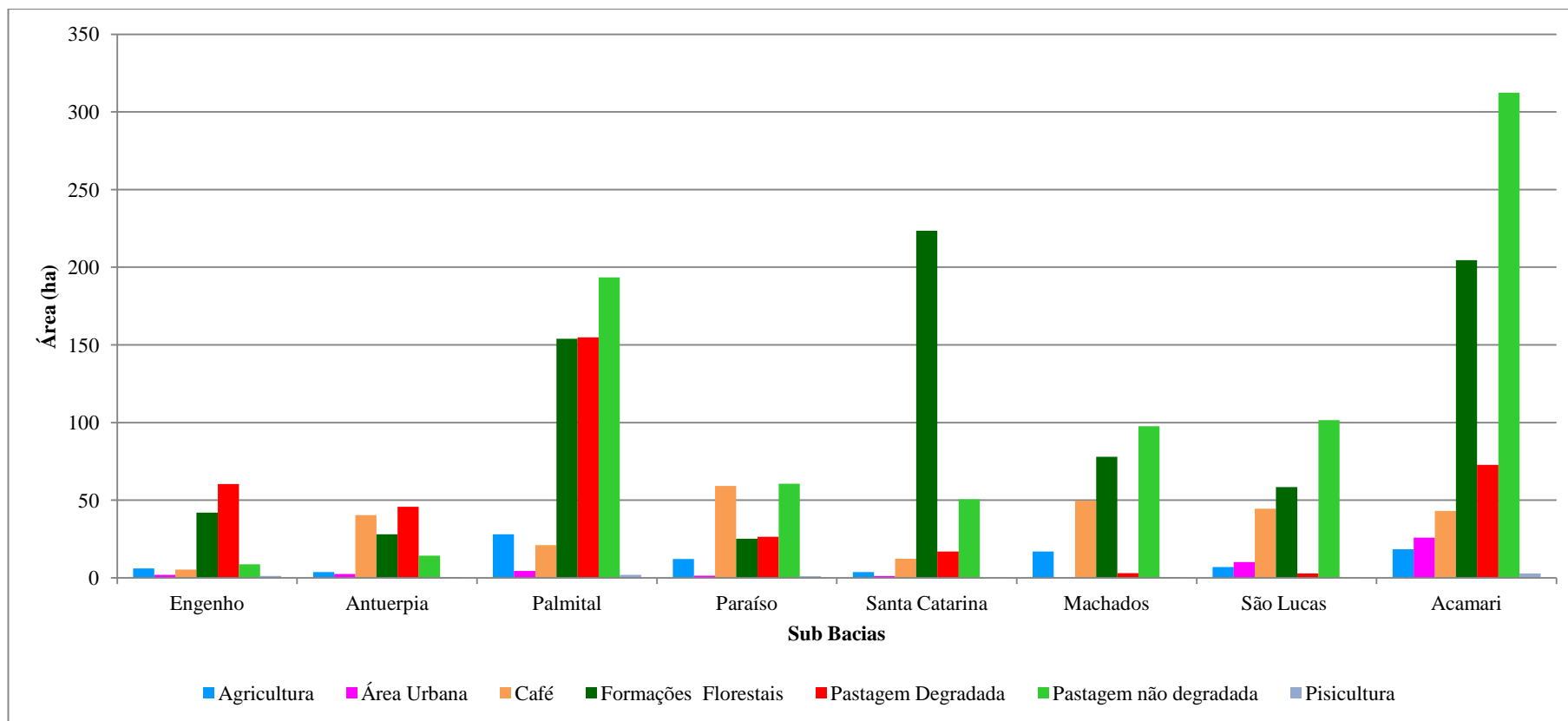


Figura 24 – Usos e ocupação do solo por sub bacias no ribeirão São Bartolomeu

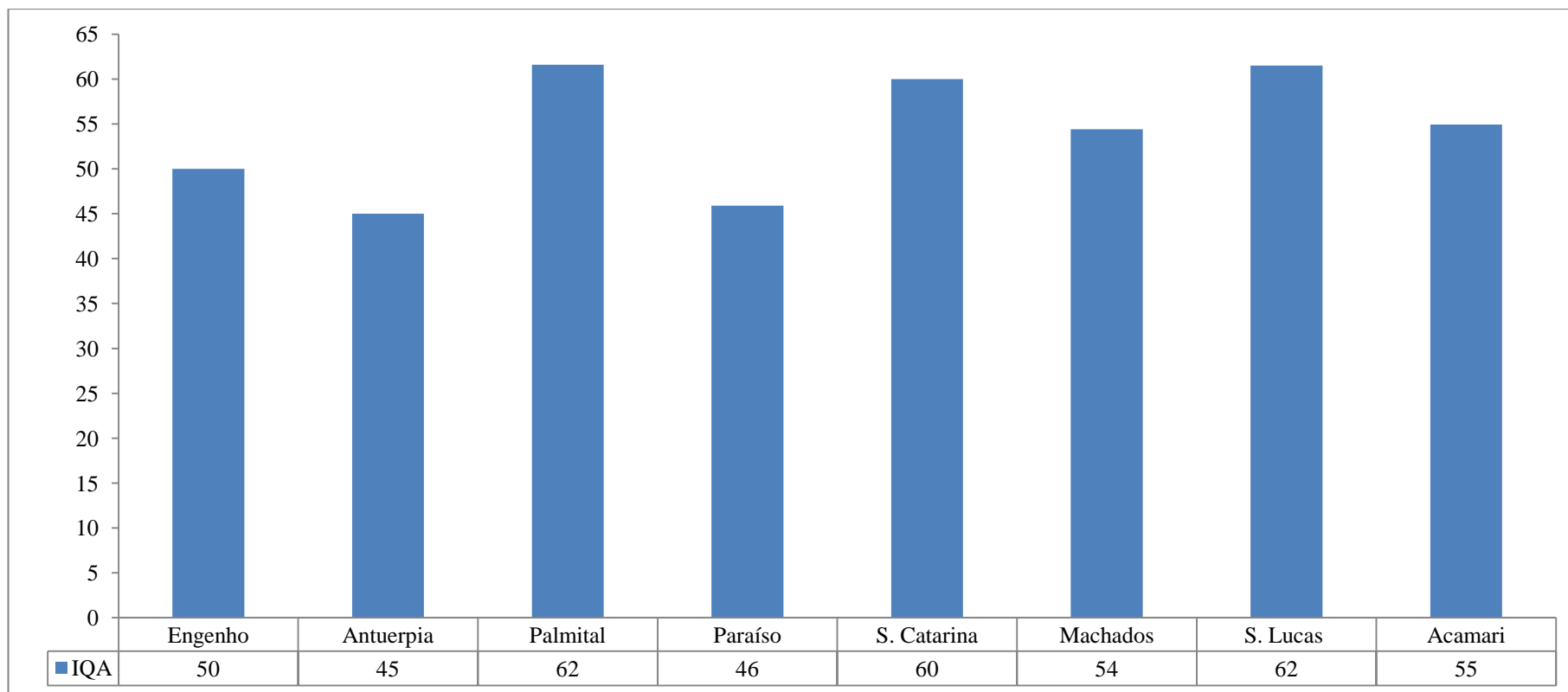


Figura 25 – Índices da qualidade da água por sub bacias do ribeirão São Bartolomeu.

Tabela 21 - Uso e ocupação do solo por sub bacias no ribeirão São Bartolomeu.

Usos e Ocupação do Solo	Engenho	Antuerpia	Palmital	Paraíso	Santa Catarina	Machados	São Lucas	Acamari
	Área (ha)							
Agricultura (culturas anuais)	6,06	3,76	27,99	12,15	3,78	16,92	7,04	18,37
Área Urbana	2,01	2,55	4,38	1,37	1,26	0,23	10,23	25,88
Café	5,37	40,42	21,12	59,25	12,31	49,72	44,55	43,05
Formaçõe Florestais	42,08	28,13	154,04	25,21	223,53	77,97	58,55	204,50
Pastagem Degradada	60,34	45,83	154,82	26,44	17,00	3,04	2,82	72,74
Pastagem não Degradada	8,68	14,24	193,53	60,54	50,61	97,62	101,48	312,38
Pisicultura (Hidrografia)	1,25	0,46	1,95	1,12	0,24	0,58	0,12	2,78
Solo Exposto	0,53	0,00	2,12	4,28	0,47	0,43	0,77	2,72
<b>Total</b>	<b>126,32</b>	<b>135,38</b>	<b>559,95</b>	<b>190,37</b>	<b>309,20</b>	<b>246,49</b>	<b>225,55</b>	<b>682,42</b>

### 4.3.3. Correlações entre usos ocupações e variáveis de qualidade da água

Na Tabela 26 são apresentados os resultados da análise de correlação entre as variáveis de qualidade da água e os diferentes usos e ocupação dos solo na bacia hidrográfica do ribeirão São Bartolomeu.

Inicialmente, observa-se que pH e oxigênio dissolvido (OD) foram as variáveis de qualidade da água que apresentaram os maiores números de coeficientes de correlações significativos ( $p < 0,05$ ) com usos e ocupação do solo (Tabela 22).

Com relação ao OD verificou-se uma tendência de seu maior consumo a medida que áreas como café ( $r = -0,60$ ,  $p < 0,01$ ), agricultura ( $r = -0,35$ ,  $p < 0,01$ ), piscicultura ( $r = -0,24$ ,  $p < 0,01$ ) e pastagem degradada ( $r = -0,22$ ,  $p < 0,05$ ) passaram a ocupar maiores áreas nas sub bacias e sejam manejadas de maneira inadequada.

No entanto, o aumento de áreas que promovam uma melhor cobertura do solo como formações florestais ( $r = 0,32$ ,  $p < 0,01$ ) e pastagem ( $r = 0,26$ ,  $p < 0,05$ ) favorecem a existência de correlações positivas, sendo portanto, um bom indicador da melhoria da qualidade da água pela redução do consumo de oxigênio dissolvido nos processos de estabilização da matéria pela biomassa microbiana presente na água (22).

Era de se esperar um aumento da concentração dos sólidos nas águas dos mananciais, principalmente os de origem orgânica, oriundo das áreas urbanas. No entanto, a correlação entre OD e esse uso do solo não foi significativa ( $r = 0,20$ ,  $p < 0,05$ ) (Tabela 22), resultado similar foi observado por Vanzela et al. (2010) ( $r = 0,10$ ,  $p < 0,05$ ). Esses mesmos autores também encontraram resultados próximos aos observados neste trabalho para a correlação entre oxigênio dissolvido e formações florestais ( $r = 0,22$ ,  $p < 0,05$ ).

Nos núcleos rurais de maior adensamento populacional existem coleta e tratamento de esgotos em fossas sépticas e nas comunidades do Paraíso e Romão dos Reis há unidades estações de tratamento de esgotos. Essa política desenvolvida pelo SAAE Viçosa tem contribuído para a melhoria da qualidade da água pela diminuição do aporte de matéria orgânica proveniente de esgotos domésticos nos cursos d'água.

Tabela 22 - Coeficientes de correlações entre as variáveis de qualidade da água e uso e ocupação do solo

Usos e Ocupação do Solo	Variáveis de Qualidade da Água								
	Oxigênio Dissolvido	Nitrogênio Total	Fósforo Total	Turbidez	pH	<i>Escherichia Coli</i>	Ferro Total	Manganês	Zinco
Agricultura	-0,35**	-0,04	-0,01	0,24*	-0,30**	-0,19	0,10	0,44**	0,24*
Área Urbana	0,20	-0,03	-0,02	-0,33**	-0,32**	0,32**	-0,25*	-0,31**	-0,11
Café	-0,60**	-0,05	0,02	0,07	-0,31**	-0,16	0,11	0,15	-0,06
Formações Florestais	0,32**	-0,003	-0,01	-0,19	0,27**	0,02	-0,06	-0,13	-0,08
Pastagem Degradada.	-0,22*	0,31**	0,53**	0,25*	-0,26*	-0,12	0,50**	0,19	0,33**
Pastagem não Degradada	0,26*	-0,29**	-0,50**	-0,40**	0,30**	0,27**	-0,53**	-0,15	-0,19
Piscicultura	-0,24*	0,30**	0,50**	0,34**	-0,25*	0,02	0,41**	0,32**	0,40**

\*\*Correlações significativas ao nível de 1%

\*Correlações significativas ao nível de 5%

A turbidez apresentou coeficientes de correlação significativos ( $p < 0,05$ ) com cinco dos setes tipos de uso e ocupação do solo. Destes, agricultura ( $r = 0,24$ ,  $p < 0,05$ ), pastagem degradada ( $r = 0,25$ ,  $p < 0,05$ ) e piscicultura ( $r = 0,34$ ,  $p < 0,01$ ) foram positivos (Tabela 22).

As principais causas da degradação das pastagens são o excesso de animais por hectare, a não utilização de adução e reposição de nutrientes necessários para manutenção da fertilidade do solo e principalmente o aumento da densidade, redução da porosidade e diminuição da infiltração de água no solo (MACEDO, 2009).

Os resultados ajudam na explicação da contribuição das pastagens degradadas para a perda de qualidade da água, assim como apontam para necessidade da adoção de sistema de manejo que busque a recuperação das áreas de pastagens degradadas, a criação de animais dentro da capacidade de suporte de cada pasto, o uso adequado de ração na alimentação dos peixes, o tratamento dos efluentes antes do lançamento nos cursos d'água e o uso de sistemas de manejo e conservação do solo e da água. Dessa forma, pode-se vislumbrar a melhoria das propriedades físicas, químicas e biológicas do solo, proporcionando o aumento da infiltração de água, reduzindo, assim, o escoamento superficial e o aporte de material de origem mineral e orgânica lançados na bacia.

As correlações negativas entre turbidez e pastagem ( $r = -0,40$ ,  $p < 0,01$ ) e turbidez e área urbana consolidada ( $r = -0,33$ ,  $p < 0,01$ ) (Tabela 22), mostram que as pastagens quando bem manejadas são eficientes na proteção do solo contra a erosão hídrica. O outro resultado indica que a maior parte do esgoto produzido está recebendo algum tipo de tratamento, o que vai de encontro com os resultados encontrados no trabalho de Bezerra (2011) onde apenas cerca de 33% dos esgotos são lançados diretamente nos cursos d'água.

De acordo com Bastos (2009), recentemente foi observado um aumento da expansão urbana, na bacia do ribeirão São Bartolomeu, com a implantação de condomínios e loteamentos.

A correlação positiva entre *Escherichia coli* e áreas urbanas ( $r = 0,32$ ,  $p < 0,01$ ) (Tabela 26) pode ser um bom indicador de que se áreas continuarem expandindo, terá como consequência o aumento da concentração de *Escherichia coli* na água.

Os resultados mostram que não só a expansão urbana tem contribuído para o aumento da concentração de *Escherichia coli* na água, mas sim o tipo de manejo que está sendo dados aos resíduos sólidos, efluentes e excretos de animais.

Nesta bacia 55,3% dos esgotos domésticos produzidos são coletados e recebem algum tipo de tratamento, 32,9% são lançados diretamente nos cursos d'água, 10,2% coletados em redes de esgotos e destinados a estações de tratamentos e apenas 1,6% é jogado à céu aberto. Com relação aos resíduos sólidos 91,8% são coletados pela Prefeitura Municipal de Viçosa e 8,2% são queimados (BEZERRA, 2011).

Nesta bacia 55,3% dos esgotos domésticos produzidos são coletados e recebem algum tipo de tratamento, 32,9% são lançados diretamente nos cursos d'água, 10,2% coletados em redes de esgotos e destinados a estações de tratamentos e apenas 1,6% é jogado a céu aberto. Com relação aos resíduos sólidos 91,8% são coletados pela Prefeitura Municipal de Viçosa e 8,2% são queimados (BEZERRA, 2011). Portanto, faz-se necessário a implantação de uma política que tenha como finalidade a melhoria dos serviços de saneamento e que dê a essa bacia a importância que ela tem para o abastecimento de água da cidade e da Universidade Federal de Viçosa

Outro ponto que chama atenção é a correlação positiva entre *Escherichia coli* e pastagem não degradada ( $r = -0,27, p < 0,01$ ) (226). Essa relação pode ser explicada por ser os bovinos potenciais reservatórios de *Escherichia coli*, eliminando-as por meio de suas fezes nos pastos (MERCADO, 2007) e por uma pastagem não degradada quando bem manejada apresentar vigor de produtividade e capacidade de sustentação dos níveis de produção e de qualidade exigida pelos animais, o que de certa forma contribui para o aumento da concentração de *Escherichia coli* na água.

Pelos coeficientes de correlações positivos entre nitrogênio total e pastagem degradada ( $r = 0,31, p < 0,01$ ) e fósforo total e pastagem degradada ( $r = 0,53, p < 0,01$ ) (Tabela 22).

Segundo Vanzela et al., (2005) a perda de solo em uma área de pastagem degradada foi de 0,28 t/ha e Bertoni e Neto (1999) encontraram um valor de perda para uma área de mata nativa de 0,004 t/ha. Esses resultados mostram que a perda de solo nas áreas de pastagem degradada chega a ser 70 vezes maior que uma área em condições naturais, mostrando que estas áreas são importantes fontes desses nutrientes para água, provavelmente originados da matéria orgânica, já que devido ao baixo poder econômico dos agricultores a capacidade de investimento em fertilizantes para

reposição desses nutrientes é muito baixa, sendo essa bacia caracterizada por pequenas propriedades rurais de subsistência.

Pelas correlações observadas entre nitrogênio total e piscicultura ( $r = 0,30$ ,  $p < 0,01$ ) e fósforo total e piscicultura ( $r = 0,30$ ,  $p < 0,01$ ) (Tabela 22), pode-se inferir que a atividade de piscicultura, que ocupa uma área muito pequena na bacia, inferior a 1%, sendo uma importante fonte de matéria orgânica e desses nutrientes para a água.

Segundo Galli e Torloni (1992), a ração principal fonte de energia utilizada na alimentação dos peixes, é rica em nitrogênio, fósforo e potássio. A adubação dos tanques com fertilizantes orgânicos (esterco de animais) e minerais, a base de nitrogênio, fósforo e potássio, é uma prática comum (BASTOS, 2003).

Silva et al. (2005) observaram, na entrada do curso d'água, que os efluentes oriundos de tanques de piscicultura apresentavam uma concentração de material em suspensão 1,65 vezes maior do que a observada na entrada do tanque. Para nitrogênio os valores foram um pouco menores, porém, seguiram a mesma tendência, saíram de 1,31 mg/L para 1,90 mg/L.

Por ser tratar de uma produção de subsistência, geralmente, não há controle sobre a dosagem de ração utilizada na alimentação dos peixes, limpeza dos tanques e, principalmente, o tratamento dos efluentes antes do lançamento nos cursos d'água. De certa forma, esse manejo inadequado tem favorecido o aumento da concentração de nutrientes, de material orgânico e diminuição da concentração de oxigênio dissolvido na água.

Uma pastagem bem manejada aumenta a capacidade de infiltração, diminui o escoamento superficial, o aporte de excrementos dos animais e nutrientes para os cursos d'água. O efeitos positivos das pastagens sobre a qualidade da água foram observados por Zigomar e Alves (2003), Silva et al., (2005), Vanzela et al. (2010) e Huang et al., (2013).

As correlações negativas entre nitrogênio total e pastagem não degradada ( $r = -0,29$ ,  $p < 0,01$ ) e fósforo total e pastagem não degradada ( $r = -0,50$ ,  $p < 0,01$ ) (226) comprovam a influência dessas áreas na diminuição da concentração desses nutrientes na água. Sendo portanto, necessário a implantação de ações visando a ampliação dessas áreas em detrimento das áreas de pastagens degradadas que atualmente ocupam cerca de

15% da área da bacia e causam efeitos adversos a qualidade da água por favorecer positivamente essas variáveis.

O pH apresentou correlações positivas apenas com formações florestais ( $r = 0,27$ ,  $p < 0,01$ ) e pastagem não degradadas ( $r = 0,30$ ,  $p < 0,01$ ), que são usos do solo que favorecem a melhoria da qualidade da água, pela diminuição do escoamento superficial, deposição de material de origem orgânica e mineral nos cursos d'água (Tabela 22).

Verificou-se uma tendência de redução do pH provocada pelas áreas de ocupação urbana ( $r = -0,32$ ,  $p < 0,01$ ), café ( $r = -0,31$ ,  $p < 0,01$ ), agricultura ( $r = -0,30$ ,  $p < 0,01$ ), pastagem degradada agricultura ( $r = -0,26$ ,  $p < 0,05$ ) e piscicultura ( $r = -0,25$ ,  $p < 0,05$ ), que contribuem para o aumento da concentração de ácido carbônico na água, produzidos durante o processo de decomposição da matéria orgânica pela biomassa microbiana (Tabela 22).

Quanto o ferro, observa-se a existência de correlações positivas apenas com pastagens degradadas ( $r = 0,50$ ,  $p < 0,01$ ) e piscicultura ( $r = 0,41$ ,  $p < 0,01$ ) (Tabela 24). Essa relação pode ser atribuída aos sedimentos com presença de ferro provenientes de áreas de pastagens degradadas ou originários nas áreas de piscicultura, que encontram nas áreas alagadas, piscicultura ou margens dos cursos d'água, ambiente propício para a redução do ferro, devido a baixa concentração de oxigênio, aumentando a solubilidade e disponibilidade desse elemento nas águas.

As correlações observadas para o manganês foram positivas para agricultura ( $r = 0,44$ ,  $p < 0,01$ ) e piscicultura ( $r = 0,32$ ,  $p < 0,01$ ). O zinco apresentou correlações positivas com agricultura ( $r = 0,24$ ,  $p < 0,05$ ), pastagem degradada ( $r = 0,33$ ,  $p < 0,01$ ) e piscicultura ( $r = 0,40$ ,  $p < 0,01$ ). O manganês e zinco são considerados micronutrientes em razão das pequenas quantidades exigidas pelas culturas agrícolas (Van Raij, 2011). Mesmo sendo os agricultores classificados como de subsistência, porém em algum momento foi realizada uma adubação com esses micronutrientes, refletindo a relação deles com a atividade agrícola.

As correlações desses elementos com a piscicultura podem ser atribuídas a uma maior solubilidade devido a uma baixa saturação de oxigênio. De acordo com Cetesb (2009) a concentração de zinco é aproximadamente quatro vezes maior em águas subterrâneas, quando comparada com as observadas em águas superficiais. Von Sperling (2005) afirma que na ausência de oxigênio dissolvido suas concentrações de manganês tendem a aumentar por se apresentar na forma solúvel.

## 5. Conclusões

A modelagem dos índices da qualidade da água para bacia do ribeirão São Bartolomeu (IQAS<sub>b</sub> e IAPS<sub>b</sub>), que foram desenvolvidos pelo uso de análise de componentes principais (ACP) e de redes neurais artificiais, apresentou resultados semelhantes aos do modelo desenvolvido pela Cetesb para o Estado de São Paulo.

Com essa técnica foi possível constatar que nem sempre o uso do primeiro fator da ACP, que tem predileção por uma parte significativa de autores, é suficiente para avaliar a qualidade das propriedades químicas, físicas e biológicas das amostras de água analisadas. No caso estudado, mesmo o primeiro fator da ACP tendo apresentado a maior variância acumulada dos dados do IQAS<sub>b</sub> e IAPS<sub>b</sub>, 27,18% e 24,85%, respectivamente, não pôde ser considerado como índice da qualidade, porque nem todas as variáveis selecionadas para composição dos índices apresentaram os maiores valores de cargas fatoriais nesse fator

Os resultados dos índices de qualidade da água, calculados pelo método estatístico Bartlett e pelo modelo proposto pela CETESB apresentaram comportamentos coerentes nos valores numéricos encontrados em cada sub-bacia hidrográfica.

Dessa forma, as variáveis oxigênio dissolvido, turbidez, nitrogênio total, cor e pH foram as que apareceram com maior frequência nos modelos de composição dos índices devendo, portanto, ser mantidas, impreterivelmente, nos programas de monitoramentos da qualidade da água na bacia do ribeirão São Bartolomeu.

O estudo demonstrou que a ocupação do solo com agricultura, café, pastagem degradada e piscicultura foram os mais impactantes na degradação da qualidade da água na bacia do ribeirão São Bartolomeu, por terem contribuído para o aumento do escoamento superficial, aporte de sedimentos e consumo de oxigênio dissolvido no processo de decomposição da matéria orgânica pela microrganismos aeróbicos.

As áreas ocupadas por formações florestais e pastagem não degradadas apresentaram correlações positivas com oxigênio dissolvido ( $r = 0,32$ ,  $p < 0,01$ ) e ( $r = 0,26$ ,  $p < 0,05$ ), respectivamente, por isso promoveram uma melhor cobertura do solo, proteção contra o processo erosivo, maior estabilidade física e melhoria da qualidade da água pelo aumento da concentração de oxigênio dissolvido na água.

As limitações da modelagem de índices da qualidade pelo emprego de análise fatorial de componentes principais está na disponibilidade de um banco de dados robusto, na complexidade dessas técnicas e na existência um corpo técnico habilitado para o desenvolvimento dos modelos, o que na maioria das vezes não são encontrados nos órgãos e empresas responsáveis pela gestão dos recursos hídricos.

O uso de índices da qualidade da água modelados pelo emprego ACP e redes neurais favoreceram a seleção de variáveis ajustadas para as condições socioeconômicas da bacia do ribeirão São Bartolomeu, no entanto, novos estudos precisam ser desenvolvidos para avaliar a capacidade de generalização desses modelos em outras condições socioeconômicas e ambientais diferentes das encontradas na bacia do ribeirão São Bartolomeu.

## 6. Referências Bibliográficas

AGÊNCIA NACIONAL DE ÁGUAS -ANA. Indicadores de Qualidade. Disponível em: <http://pnqa.ana.gov.br/IndicadoresQA/IndiceQA.aspx>. Acesso em: 14/08/2014.

ADAMOLI, J.; MACEDO, J.; AZEVEDO, J. G.; NETTO, J. M. Caracterização da região dos Cerrados. In: EMBRAPA/CPAC Solos dos Cerrados: tecnologias e estratégia de manejo. São Paulo: Nobel, 1986. p.33-74.

AKBAL, F.; GUREL, L.; BAHADIR, T.; GULER, I.; BAKAN, G.; BUYUKGUNGOR, H. Multivariate statistical techniques for the assessment of surface water quality at the mid-Black Sea Coast of Turkey. *Water, Air e Soil Pollution*, v. 216, n 1-4, p. 21-37, 2011.

AL-BADAIL, F.; SHUHAIMI-OTHTMAN, M.; GASIM, M. B. Water Quality Assessment of the Semenyih River, Selangor, Malaysia. *Journal of Chemistry*, 2013.

ALMEIDA, R. A. S. Índice de qualidade de águas subterrâneas destinadas ao uso na produção de água potável (IQUAS). 2007. 130f. Dissertação (Mestrado em Engenharia Ambiental) - Escola Politécnica, Universidade Federal da Bahia, Salvador, 2007.

ANDRADE, E. M. de; Araújo, L. de F. P.; Rosa; M. F.; Gomes, R. B.; Lobato, F. A. de O. Seleção dos indicadores da qualidade das águas superficiais pelo emprego da análise multivariada. *Engenharia Agrícola*, v. 27, n. 3, p. 683-690, 2007.

APHA-AWWA-WPCF; Standard Methods for Examination of Water and Wastewater, 20th ed.; Washington, 2005.

ARCOVA, F. C. S.; CICCIO, V. Qualidade da água de microbacias com diferentes usos do solo na região de Cunha, Estado de São Paulo. *Scientia Forestalis*, Piracicaba, n.56, p.125-134, 1999.

BASTOS, R. K. X.; Pereira, C. M.; Pivelli, R. P.; Lapolli, F. R., Lanna, E. A. T Utilização de esgotos sanitários em piscicultura. Utilização de esgotos tratados em fertirrigação, hidroponia e piscicultura. Rio de Janeiro: ABES, p. 163-223, 2003.

BASTOS, R.K.X. (Coord.). Plano de Segurança da Água. relatório final. Viçosa, 2009.

BERTOL, O. J.; RIZZI, N. E.; BERTOL, I.; ROLOFF, G. Perdas de solo e água e qualidade do escoamento superficial associadas à erosão entre sulcos em área cultivada sob semeadura direta e submetida às adubações mineral e orgânica. Revista Brasileira de Ciência do Solo, v.31, p. 781-792, 2007.

BERTONI, J.; NETO, F. L. Conservação do solo. São Paulo: 4ª ed. Ícone, 1999. 355p.

BEZERRA, N. R. Aplicação de redes bayesianas na identificação de perigos em sistemas de abastecimento de água para consumo humano: estudo de caso no município de Viçosa-Minas Gerais. 2011. 163f. Tese (Programa de Pós-graduação em Engenharia Civil), Universidade Federal de Viçosa. Viçosa, 2011.

BOYACIOGLU, Hülya. Surface water quality assessment using factor analysis. Water SA V. 32, n.3, p. 389-393, 2007.

BRAGA, A. P.; CARVALHO, A. C.; LUDERMIR, T. B. Redes Neurais Artificiais: teoria e aplicações. São Carlos: 2ª ed. LTC, 2007. 248p.

BRASIL. Resolução N° 357 março de 2005, Brasília, Ministério do Meio Ambiente, 2005. 27p (Ana).

BRASIL. Lei N° 9433, de 08 de janeiro de 1997, Brasília, Ministério do Meio Ambiente. 17p.

BRASIL. Portaria n° 2914 dezembro 2011, Brasília, Ministério da Saúde, 2011. 32p.

BROWN, R. M.; Mc CLEALLAND, N.; DEININGER, R.A.; TOZER, R.G. A water quality index do we dare. Water Sewage Works, v. 117, p. 339-343, 1970.

CANDIA, C. I. A construção de um indicador de qualidade de água para gestão de bacia hidrográfica usando análise fatorial. 2007. 113f. Dissertação (Programa de Pós Graduação em Tecnologias Ambientais), Universidade Federal de Mato Grosso do Sul, Campo Grande, 2007.

CARVALHO, A. R., SCHLITTLER, F. H. M., TORNISIELO, V. L. Relações da atividade agropecuária com parâmetros físicos químicos da água. Revista Química Nova. v. 23. n. 5, p. 618 - 622, 2000.

CERQUEIRA, E. O., ANDRADE, J. C., POPPI, R. J.; MELLO, C. Redes neurais e suas aplicações em calibração multivariada. Química Nova, v. 24, n. 6, p. 864-873, 2001.

COLETTI, C.; TESTEZLAF, R.; RIBEIRO, T. A. P.; DE SOUZA, R. T. G.; PEREIRA, D. DE A. Water quality index using multivariate factorial analysis. Revista Brasileira de Engenharia Agrícola e Ambiental, v.14. n.5. p.517-522, 2010.

COMPANHIA DE TECNOLOGIA DE SANEAMENTO AMBIENTAL - CETESB. Qualidade das águas superficiais no estado de São Paulo 2013. Série relatórios. São Paulo, 2014, 434p.

COMPANHIA DE TECNOLOGIA DE SANEAMENTO AMBIENTAL - CETESB. Guia de coleta e preservação de amostras de água. 1ª ed. São Paulo, 1987, 155p.

COMPANHIA DE TECNOLOGIA DE SANEAMENTO AMBIENTAL - CETESB. Qualidade das Águas Interiores de São Paulo. Série relatórios. São Paulo, 2008, 29p.

COMPANHIA DE TECNOLOGIA DE SANEAMENTO AMBIENTAL - CETESB. Significado ambiental e sanitário das variáveis de qualidade das águas e dos sedimentos e metodologias analíticas e de amostragem. Série relatórios. São Paulo, 2009, 45p.

DERÍSIO, J. C. Introdução ao controle da poluição ambiental. São Paulo: Editora Signus. 2007. 224p.

DIAS, G. M. F. Estudo integrado da qualidade de água da bacia do São Bartolomeu: análise epidemiológica, ambiental e espacial. Viçosa, 2007, 189p. Dissertação (Mestrado em Medicina Veterinária) – Departamento de Veterinária, Universidade Federal de Viçosa, Viçosa-MG, 2007.

DONADIO, N. M. M.; GALBIATTI, J. A.; PAULA, R. C. de. Qualidade da água de nascentes com diferentes uso do solo na bacia hidrográfica do Córrego Rico, São Paulo, Brasil. Engenharia Agrícola, v.25, n.1, p.115-125, 2005.

FIGUEIREDO FILHO, D. B.; SILVA JÚNIOR, J. A.. Visão além do alcance: uma introdução à análise fatorial. Opinião Pública, v. 16, n. 1, p. 160-185, 2010.

FRANCO, R. M. B.; CANTUSIO NETO, R.; BRANCO, N. Detecção de *Cryptosporidium* sp e *Giardia* sp em água pela técnica de filtração em membrana: estudo comparativo entre diferentes técnicas de diluição. Jornal Brasileiro de Patologia, v 37, n 4, p. 205, 2001.

FRIEL, C. M. 2009. Notes on Factor Analysis. Criminal Justice Centre, Sam Houston. <http://www.bama.ua.edu/~jcsenkbeil/gy523/Factor%20>. Acesso em: 14/04/2014.

GALLI, L. F.; TORLONI, C. E. C. Criação de Peixes. São Paulo, Editora Nobel, 1992. 119p.

GARSON, G. D. Regression analysis, from Statnotes: Topics in multivariate analysis. 2010.

GAZZAZ, N. M.; YUSOFF, M. K.; ARIS, A. Z.; JUAHIR, H.; RAMLI, M. F. Artificial neural network modeling of the water quality index for Kinta River (Malaysia) using water quality variables as predictors. Marine Pollution Bulletin, v. 64, n. 11, p. 2409–2420, 2012.

GUEDES, H. A. S.; SILVA D. D.; ELESBON, A. A. A.; RIBEIRO, C. B. M.; De MATOS, A. T.; SOARES, J. H. P. Aplicação da análise estatística multivariada no estudo da qualidade da água do Rio Pomba, MG. *Revista Brasileira de Engenharia Agrícola e Ambiental*. v.16. n.5. p.558-563, 2012.

HAASE, J.; KRIEGER, J. A. H.; POSSOLI, S. Estudo da viabilidade do uso da técnica de análise fatorial como um instrumento na interpretação da qualidade da água da bacia do Guaíba, RS, Brasil. *Ciência e Cultura*, v.41, n.6, p.576-582, 1989.

HAASE, J.; POSSOLI, S. Estudo da utilização da técnica de análise fatorial na elaboração de um índice de qualidade de água: comparação entre dois regimes hidrológicos diferentes, RS. *Acta Limnologica Brasiliensia*, v.6, p.245 - 255, 1993.

HAIR, J. F. Jr.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. *Análise multivariada de dados*. Tradução Adonai Schlup Sant'Anna. Porto Alegre: Bookman. 6 ed. 2009. 686p.

HAYKIN, S. *Redes Neurais - princípios e prática*, Porto Alegre, Brasil, Editora Bookman, 2a ed, . 2001. 900p.

HEBB, DONALD OLDING. *The organization of behavior: A neuropsychological theory*. Psychology Press, 2002.

HELENA, B.; PARDO, R.; VEGA, M.; BARRADO, E.; FERNANDEZ J. M.; FERNANDEZ, L. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga river, Spain) by principal component analysis. *Water Research*, v.34, n.3, p. 807-816, 2000.

HUANG, J.; ZHAN, J.; YAN, H.; WU, F.; DENG, X. Evaluation of the Impacts of Land Use on Water Quality: A Case Study in The Chaohu Lake Basin. *The Scientific World Journal*, 2013.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA - IBGE. Estimativa da população 2013. Disponível em: [http://www.ibge.gov.br/home/estatistica/populacao/estimativa2013/estimativa\\_dou.shtm](http://www.ibge.gov.br/home/estatistica/populacao/estimativa2013/estimativa_dou.shtm). Acesso em: 14/08/2014.

JUAHIR, H.; ZAIN, S. M.; TORIMAN, E. M.; MOKHTAR, M.; MAN, H. C. Application of artificial neural network models for predicting water quality index. *Jurnal Kejuruteraan Awam*, v. 16, n. 2, p 42-55, 2004.

LANDWEHR, J.M.; DEININGER, R.A. A comparison of several water quality indices. *Journal of Water Pollution Control Federation*, v. 48, n. 5, p 954 - 958. 1976.

LEITE, F.; FONSECA, O. Aplicação de índices de qualidade das águas na lagoa Caconde, Osório, RS. IN: Seminário de Qualidade das Águas Continentais no Mercosul, 1. Anais. Organizador: David M. Marques- ABRH. 1994.

LI, S.; Xia, X.; Tan, X.; Zhang, Q. Effects of catchment and riparian landscape setting on water chemistry and seasonal evolution of water quality in the upper Han River basin, China, v. 8, n. 1, p. 1 - 14, 2013.

LINGIREDDY, S.; BRION, G. M. Artificial Neural Networks in Water Supply Engineering. USA: ASCE, American Society of Civil Engineers Press, 2005. 192p.

MACEDO, M. C. M. Integração lavoura e pecuária: o estado da arte e inovações tecnológicas. *Revista Brasileira de Zootecnia*, v. 38, 2009.

MAHMOOD, A.; MUQBOOL, W.; MUMTAZ, M. W.; AHMAD, F. Application of Multivariate Statistical Techniques for the Characterization of Ground Water Quality of Lahore, Gujranwala and Sialkot (Pakistan). *Pakistan Journal of Analytical and Environmental Chemistry*, v. 12, 2011.

MARGALEF, R. *Limnologia*. Barcelona: Omega, 1983. 100 p.

McCULLOCH, W.; PITTS, W. A Logical calculus of the ideas immanent in nervous activity. *Bulletin of mathematical biophysics*. v.5, n. 4, p. 113-133, 1943.

MEIRELES, A. C. M. Dinâmica qualitativa das águas superficiais da bacia do Acaraú e uma proposta de classificação para fins de irrigação. 2007, 180f. Tese (Curso de Pós-graduação em Engenharia Civil), Universidade Federal do Ceará. Fortaleza, 2007.

MENDIGUCHÍA, C.; MORENO, C.; GALINDO-RIANO, M. D.; GARCÍA-VARGAS, M. Using chemometric tools to assess antropogenic effects in river water a case study: Guadalquivir river (Spain). *Analytica Chimica Acta*, v. 515, n. 1, p. 143-149, 2004.

MENEZES, F. C. Uso de redes neurais artificiais na predição da dosagem do sulfato de alumínio e do hidróxido de sódio no processo de coagulação. 2011. 93f. Dissertação (Programa de Pós-Graduação em Engenharia Industrial da Escola Politécnica), Universidade Federal da Bahia, Salvador, 2011.

MERCADO, E. C. Síndrome Urémico Hemolítico. *Revista Argentina de Microbiologia*, v.39, p.191-192, 2007.

MERTEN, G. H e MINELLA, J. P. Qualidade da água em bacias hidrográficas rurais: um desafio atual para sobrevivência futura. *Agroecologia e Desenvolvimento Rural Sustentável*, v.3, n.4, p.33-38, 2002.

MINSKY, M. L.; PAPERT, S.A. *Perceptrons: An introduction to computational geometry*, MIT Press, Cambridge, Massachusetts. 1969.

MORIASI, D. N.; ARNOLD, J. G.; LIEW, M. W. VAN; BINGNER, R. L.; HARMEL, R. D.; VEITH, T. L. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of American Society of Agricultural and Biological Engineers*, v.50, n.3, p.885-900, 2007.

NAJAH, A.; ELSHAFIE, A.; KARIM, O. A.; JAFFAR, O. Prediction of Johor River Water Quality Parameters Using Artificial Neural Networks *European Journal of Scientific Research*, v. 28, n. 3. p.422-435. 2009.

NASH, J. E.; SUTCLIFFE, J. V. River flow forecasting through conceptual models part 1 – A discussion of principles. *Journal of Hydrology*, v.10, n.3, p.282-290, 1970.

NASIR, M. F. M.; JUAHIR, H.; ROSLAN, N.; MOHD. I.; SHAFIE, N. A.; ALIAA, N.; RAMLI, N. Artificial neural networks combined with sensitivity analysis as a prediction model for water quality index in Juru River, Malaysia. *International Journal of Environmental Protection*, v.1, n. 3, p 1-8, 2011.

NETTO F.; CECILIO. A.; GOMES. A.;; PIRES L. Comparativo entre cálculos de índice de qualidade da Água: análise estatística multivariada e método cestesb. Disponível em: <  
[https://www.abrh.org.br/sgcv3/UserFiles/Sumarios/d8b1562d6b78282ce56a50640803d128\\_3086a99abc4dbc0506d9fa6190010fb1.pdf](https://www.abrh.org.br/sgcv3/UserFiles/Sumarios/d8b1562d6b78282ce56a50640803d128_3086a99abc4dbc0506d9fa6190010fb1.pdf) > Acesso em: 12/08/2014.

NORUSIS, M. J. *SPSS Base System User's Guide*. Chicago: SPSS Inc, 1990. 520p.

OLIVEIRA, A.T.; ALMEIDA, R.J. de; MELLO, J.M. de; GAVILANES, M.L. Estrutura fitossociológica e variáveis ambientais em um trecho de mata ciliar do córrego Vilas Boas, Reserva Biológica do Poço Bonito, Lavras (MG). *Revista Brasileira de Botânica*, São Paulo, v.17, n.1, p.67-85, 1994.

OTT, Wayne R. *Environmental indices: theory and practice*. 1978. 371p.

PALLANT, J. *SPSS Survival Manual*. Open University Press, 2007.

PROGRAMA NACIONAL DO MEIO AMBIENTE - PNMA. Índice e indicadores de qualidade da água - Revisão da Literatura. 2008. 114p.

QUEIROZ, M. M. F.; IOST.; C. GOMES, S. D.; VILAS BOAS, M. A. Influência do uso do solo na qualidade da água de uma microbacia hidrográfica rural. *Revista Verde*, v.5, n.4, p. 200 - 210, 2010.

RABUNAL, R. J.; DORRADO, J. *Artificial neural networks in real-life applications*. Idea Group. EUA. 2006. 395p.

RANZINI, M.; LIMA, W.P. Comportamento hidrológico, balanço de nutrientes e perdas de solo em duas microbacias reflorestadas com Eucaliptus, no Vale do Paraíba, SP. Scientia Forestalis, Piracicaba, n.61, p.144-159, 2002.

RESENDE, M.; CURI, N.; REZENDE, S. B.; CORRÊA, G. F. PEDOLOGIA: base para distinção de ambientes. Editora UFLA. 2007, 322p.

RODRIGUES, M. L. K.; RODRIGUEZ, M. T. R.. Statistical factor analysis of heavy metal pollution in the Caí river basin sediments (Brazil). In: 2ND International Symposium on Contaminated Sediments, Quebec City, Canadá. Pôster. 26-28 de M 2003.

ROJAS, R. The Backpropagation Algorithm. Neural Networks. Springer-Verlag, Berlin. 1996. 34p. Disponível em: <<http://page.mi.fu-berlin.de/rojas/neural/chapter/K7.pdf>>. Acesso em: 09/10/2011.

ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychol. n. 65, p. 386-408, 1958.

RUMELHART, D.; HILTON, G.; WILLIAMS, R. Learning Interval Representation by Error Propagation. Parallel Distributed Processing: Explorations in the Microstructure of Cognition, 1: Foundations. MIT Press, Cambridge, Massachusetts. p. 218-362. 1986.

SANTOS, I.; FILL, H. D.; SUGAI, M. R. V. B; BUBA, H.; KISHI, R. T.; LAUTERT, L. F. Hidrometria Aplicada. LACTEC- Instituto de Tecnologia para o Desenvolvimento. Curitiba, PR. 372p, 2001.

SCHAWB, A. J. Eletronic classroom. 2012. Disponível em: <<http://www.utexas.edu/ssw/eclassroom/schwab.html>>. Acesso em: 14/10/2013.

SILVA, D. D.; PRUSKI, F. F.; SCHAEFER, C. E. G. R.; AMORIM, R. S. S.; PAIVA, K. W. N. Efeito da cobertura nas perdas de solo em um Argissolo Vermelho-Amarelo utilizando simulador de chuva. Engenharia Agrícola, v.25, n.2, p. 409-419, 2005.

SILVA, C. H. C. Identificação de fragilidades ambientais na bacia do ribeirão São Bartolomeu, Viçosa-MG utilizando análise multicritério. 2010. 120f. Dissertação. (Programa de Pós-Graduação em Engenharia Civil), Universidade Federal de Viçosa, Viçosa, 2010.

SILVA, G. L.; AURELIANO, J. T.; LUCENA, S. V. D. O., Proposição de um Índice de Qualidade de Água Bruta para Abastecimento Público. Revista de Gestão de Água da América Latina. v. 1. n. 9. p. 17-24. 2012.

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, Amsterdam, v.37, n.17, p. 4. 119 - 4124, 2003.

SINGH, K. P.; BASANT, A.; MALIK, A.; JAIN, G. Artificial neural network modeling of the river water quality: A case study. Ecol. Model., v. 220, p. 888-895, 2009.

SOUSA, W. S.; SOUSA, F. A. S. Rede neural artificial aplicada à previsão de vazão da Bacia Hidrográfica do Rio Piancó. Revista Brasileira de Engenharia Agrícola e Ambiental, v.14, n.2, p.173-180, 2010.

SOUZA, M. E. T. A.; LIBÂNIO, M. Proposta de índice de qualidade para água bruta afluente a estações convencionais de tratamento. Engenharia Sanitária e Ambiental, v.14, n.4, p.471- 478, 2009.

TABACHNICK, B.; FIDELL, L. Using multivariate analysis. Needham Heights: Allyn e Bacon, 2007.

TOLEDO, L. G.; NICOLELLA, G. Índice de qualidade de água em microbacia sob uso agrícola e urbano. Scientia Agrícola, v.59, n.1, p.181-186, 2002.

VANZELA, L. S.; HERNANDEZ, F. B. T.; LIMA, R. C.; GARGANTINI, P. E. Influência antrópica no transporte de sedimentos em microbacia degradada. In: Congresso Nacional de Irrigação e Drenagem, XV, Teresina, 16 a 21 de outubro de 2005. 6p. CD-ROM.

VANZELA, L.; S.; HERNANDEZ, F.; B.; T.; FRANCO, R.; A.; M. Influência do uso e ocupação do solo nos recursos hídricos do Córrego Três Barras, Marinópolis. R. Bras. Eng. Agríc. e Ambiental, v. 14, n. 1, p. 55-64, 2010.

VAN RAIJ, B. Fertilidade do solo e manejo de nutrientes. Piracicaba: Esalq/USP, 2011, 420 p.

VERÍSSIMO, G. S. Desenvolvimento de um conector de cisalhamento em chapa dentada para estruturas mistas de aço e concreto e estudo do seu comportamento. 2007, 316f. Tese (Curso de Pós-graduação em Engenharia de Estruturas da Escola de Engenharia), Universidade Federal de Minas Gerais. Belo Horizonte, 2007.

VESEY, G.; SLADE, J.S.; BYRNE, M.; SHEPHERD, K.; FRICKER, C.R. A new method for the concentration of *Cryptosporidium* oocysts from water. *Journal of Applied Bacteriology*, v.75, p.82-86, 1993.

VIEIRA, C. A. O. Accuracy of remotely sensing classification of agricultural crops: a comparative study. 2000, 327f. Tese de Doutorado. University of Nottingham. 2000.

VON SPERLING, M.. Introdução à qualidade das águas e ao tratamento de esgotos. 3ª ed. Belo Horizonte: DESA / UFMG, 2005, 452p.

WASSERMAN, P. D. Neural computing: theory and practice. New York: Van Nostrand Reinhold, 1989. 230p.

ZIGOMAR, M. S.; ALVES, M. C. Movimento de água e resistência à penetração em um Latossolo Vermelho distrófico de cerrado, sob diferentes usos e manejos. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.7, n.1, p.18-23, 2003.