

UFMG – ICEx/DQ. 651^a
T. 254^a

**USOS DE MÉTODOS ESTATÍSTICOS
PARA DADOS DE QUALIDADE DE
ÁGUAS: ESTUDO DE CASO, RIO DAS
VELHAS**

Aluna: Eliana Aparecida Nonato Knupp

Orientadora: Heloiza Helena Ribeiro Schor

Co-orientadora: Cláudia de Vilhena Schayer Sabino

UFMG – ICEx/DQ. 651^a
T. 254^a

Eliana Aparecida Nonato Knupp

**USOS DE MÉTODOS ESTATÍSTICOS PARA DADOS DE
QUALIDADE DE ÁGUAS: ESTUDO DE CASO, RIO DAS
VELHAS**

**Tese apresentada ao Curso de Pós-
graduação do Departamento de Química
do Instituto de Ciências Exatas da
Universidade Federal de Minas Gerais,
como requisito parcial à obtenção do
título de doutor em ciências - química**

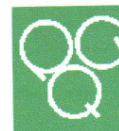
Universidade Federal de Minas Gerais

Departamento de Química

Fevereiro de 2007



CURSO DE PÓS-GRADUAÇÃO EM QUÍMICA
DEPARTAMENTO DE QUÍMICA -ICEX
31270-901 - BELO HORIZONTE - MG
TEL.: 031 - 3499-5732
FAX: 031 - 3499-5700
E-MAIL: pgquimic@zeus.qui.ufmg.br



**"Usos de Métodos Estatísticos para Dados de Qualidade de Águas: Estudo de
Caso Rio das Velhas"**

Eliana Aparecida Nonato

Tese aprovada pela banca examinadora constituída pelos Professores:

Profa. Heloiza Helena Ribeiro Schor - Orientadora
UFMG

Prof. José Cláudio Junqueira Ribeiro
FUMEC

Prof. Sérgio Duarte Segall
UNI-BH

Profa. Maria das Graças Lins Brandão
UFMG

Profa. Cláudia Carvalhinho Windmüller
UFMG

Belo Horizonte, 14 de fevereiro de 2007.

*Aos meus pais, Célia Silveira Nonato e Valdemar Nonato,
sempre presentes.*

AGRADECIMENTOS

Às Professoras Heloiza Helena Ribeiro Schor e Cláudia de Vilhena Schayer Sabino pela orientação, apoio, estímulo e liberdade de trabalho.

Ao Professor Eustáquio Vínicius Ribeiro de Castro agradeço às sugestões fundamentais na fase inicial do trabalho.

A Zenilde das Graças Guimarães Viola, coordenadora do Projeto Águas de Minas, agradeço pela cordialidade e informações sobre todo programa de monitoramento.

A equipe técnica do IGAM, em especial a Katiane Cristina de Brito Almeida e Leonardo Corradi, pela disponibilidade e atenção no fornecimento de relatórios e edições dos mapas.

Ao José Antônio Cardoso, responsável pelas análises físico-químicas no CETEC, agradeço a atenção no fornecimento de informações sobre as análises.

Aos colegas do Centro de Desenvolvimento da Tecnologia Nuclear, CDTN/CNEN, em especial a Milton Batista Franco, Adilson de Castro de Assis, Roberto Pellacani e João Batista Barbosa pelo apoio e liberações que possibilitaram a realização desse trabalho.

As bibliotecárias, Lenira Ferreira, Nívia Lima e Virgínia Rodrigues agradeço pela ajuda na aquisição de material bibliográfico e correção das referências bibliográficas.

A secretaria de pós-graduação do departamento de química pelas orientações administrativas.

Aos estatísticos, Otaviano Neves, Dalila Rodrigues e Naila Moura, agradeço pelo esclarecimento de dúvidas e por me despertar para o estudo de estatística.

Aos colegas do Departamento de Química, Ivanildo e Jorge, parceiros de orientação, agradeço a agradável convivência.

Aos meus familiares e amigos pelo carinho e incentivo nos momentos mais críticos.

Em especial, agradeço ao meu marido Wagner Fernandes Knupp, sempre presente com carinho, compreensão, serenidade e estímulo.

Sempre há o que aprender, ouvindo...vivendo... e sobretudo trabalhando.

Mas só aprende, quem se dispõe rever suas certezas.

Darcy Ribeiro

SUMÁRIO

Capítulo I – Introdução	1
1.1 Monitoramento da qualidade das águas superficiais no Estado de Minas Gerais ..	1
1.2 A sub – bacia do rio das Velhas	5
1.3 Requisitos básicos de qualidade	10
1.3.1 Legislação	10
1.3.2 Classificação de águas	11
1.4 Parâmetros indicativos da qualidade das águas	17
1.5 Coletas e análises	19
Capítulo II – Análise multivariada.....	22
2.1 Estatística multivariada.....	22
2.2 Pré – processamento dos dados	24
2.2.1 Dados censurados e discrepantes.....	24
2.2.2 Padronização dos dados.....	25
2.2.3 Matriz de covariância	27
2.2.4 Matriz de correlação	29
2.3 Análise de agrupamentos.....	30
2.3.1 Função similaridade.....	32
2.3.2 Métodos de associação entre agrupamentos	34
2.3.2.1 Critério de associação simples:	35
2.3.2.2 Critério de associação completa	35
2.3.2.3 Critério de associação média	36
2.3.2.4 Critério de associação média ponderada	36
2.3.2.5 Critério associação centróide.....	37
2.3.2.6 Critério de associação mediana	37
2.3.2.7 Método de Ward	37
2.4 Análise de componentes principais (PCA).....	40
2.4.1 Componentes principais exatas extraídas da matriz de covariância.....	41
2.4.2 Componentes principais exatas extraídas da matriz de correlação.....	46
2.4.3 Interpretação dos resultados na análise de componentes principais.....	47
2.5 Análise fatorial	48
2.6 Análise de regressão não linear	57
2.6.1 Formulação dos modelos de regressão não-lineares.....	57
2.6.2 O método de estimação por mínimos quadrados.....	58
2.6.3 A obtenção das estimativas de mínimos quadrados	59
2.6.4 Aproximação linear	61
2.6.5 O método de Gauss-Newton.....	63
Capítulo III – Análise de séries temporais.....	66
3.1 Modelos simples de previsão.....	68

3.1.1	Séries constantes.....	68
3.1.1.1	Média móvel.....	68
3.1.1.2	Alisamento exponencial simples	69
3.1.2	Séries com tendência	70
3.1.2.1	Alisamento exponencial duplo ou linear de Brown.....	70
3.1.2.2	Alisamento exponencial de Holt.....	71
3.1.3	Séries com tendência e sazonalidade.....	71
3.1.3.1	Alisamento exponencial de Holt Winter	71
3.2	Modelos auto-regressivos ou método Box e Jenkins	72
3.3	Critério de avaliação de modelos.....	74
Capítulo IV – Descrição e pré-tratamento dos dados		75
4.1	Descrição da base de dados	75
4.2	Metodologia de análises físico-química e estatística.....	77
Capítulo V – Análise de regressão.....		93
5.1	Descrição do conjunto de dados	93
5.2	Análise fatorial	95
5.3	Análise de variância aplicada ao fator 1	97
5.4	Regressão não linear - fator 1	99
5.5	Análise de variância aplicada ao fator 2	101
5.6	Regressão não linear - fator 2.....	103
5.7	Conclusões do capítulo.....	104
Capítulo VI – Análise exploratória da sub-bacia do rio das Velhas		106
6.1	Descrição do conjunto de dados	106
6.2	Análise de agrupamentos e componentes principais	108
6.3	Conclusões do capítulo.....	121
Capítulo VII - Análise exploratória do alto curso da sub-bacia do rio das Velhas.....		124
7.1	Descrição do conjunto de dados do alto curso do rio das Velhas.....	125
7.2	Análise de agrupamentos.....	126
7.3	Análise de componentes principais	130
7.4	Conclusões do capítulo.....	139
Capítulo VIII – Análise temporal dos dados		141
8.1	Escolha dos parâmetros e estações	141
8.2	Metodologia de análise.....	142
8.3	Avaliação da concentração de fosfatos total.....	143
8.3.1	Estação ribeirão do Onça próximo de sua foz no rio das Velhas (BV154)....	143
8.3.2	Rio das Velhas logo a jusante do ribeirão do Onça (BV105).....	146
8.3.3	Estação ribeirão Arrudas próximo de sua foz no rio das Velhas (BV155).....	150

8.3.4	Rio das Velhas logo a jusante do ribeirão Arrudas (BV083)	153
8.4	Avaliação dos teores de nitrogênio amoniacal	156
8.4.1	Estações ribeirão do Onça próximo de sua foz no rio das Velhas (BV154) e rio das Velhas logo a jusante do ribeirão do Onça (BV105).....	156
8.4.2	Estações no ribeirão Arrudas próxima de sua foz no rio das Velhas (BV155) e no rio das Velhas logo a jusante do ribeirão Arrudas (BV083).....	161
8.5	Avaliação dos teores de manganês	166
8.5.1	Estação rio das Velhas logo a jusante do ribeirão Arrudas (BV083)	166
8.5.2	Rio Itabirito a jusante da cidade de Itabirito (BV035/ AV110)	170
8.6	Avaliação dos teores de oxigênio dissolvido.....	173
8.6.1	Estações ribeirão do Onça próximo de sua foz no rio das Velhas (BV154) e rio das Velhas logo a jusante do ribeirão do Onça (BV105).....	173
8.6.2	Ribeirão Jequitibá próximo de sua foz no rio das Velhas (BV140)	178
8.6.3	Rio Cipó a montante da foz do rio Paraúna (BV162).....	180
8.7	Conclusões do capítulo.....	184
	Capítulo IX – Conclusões finais	186
	Capítulo X – Recomendações.....	190
	Capítulo XI – Referências bibliográficas.....	194

LISTA DE FIGURAS

FIGURA 1 – Localização das estações de amostragem	7
FIGURA 2 – Determinação das distâncias Euclidiana, Paineis A e B para 2 medidas e 2 variáveis. Figura adaptada da referência 34.....	33
FIGURA 3 - Agrupamento de pontos de amostragem em função das similaridades utilizando o método de Ward. Figura adaptada da referência 34.....	39
FIGURA 4 – Interpretação geométrica da análise de componentes principais.....	45
FIGURA 5 – Avaliação das componentes mais importantes a partir dos autovalores. Método de Cattell. Figura adaptada da referência 29	45
FIGURA 6 - <i>Biplot</i> para simultânea caracterização de escores e loadings em dados referentes a análise química de cabelo ³⁰	48
FIGURA 7 – Importância dos principais fatores avaliada pelo método de Cattell para a base de dados do monitoramento do rio das Velhas realizado no período de 1997 a 2004 em campanhas completas	95
FIGURA 8 – ANOVA aplicada ao fator 1 em relação a latitude das estações de amostragem localizadas na sub-bacia do rio das Velhas	98
FIGURA 9 – Valores reais e estimados para ln[fosfatos] no período de 1997 a 2004 para as estações compreendidas entre as latitudes 19°21' a 19°59' na sub-bacia do rio das Velhas	100
FIGURA 10 – Concentração de fosfatos totais real e estimada para as estações compreendidas entre as latitudes 19°21' a 19°59' na sub-bacia do rio das Velhas.....	101
FIGURA 11 - ANOVA aplicada ao fator 2 em relação a latitude das estações de amostragem localizadas na sub-bacia do rio das Velhas	102
FIGURA 12 - Concentração de manganês real e estimada para as estações compreendidas entre as latitudes 19°33' e 20°14' na sub-bacia do rio das Velhas	104
FIGURA 13 - Análise de agrupamento na matriz padronizada das medianas dos 41 parâmetros medidos nas 29 estações de amostragem do rio das Velhas, utilizando o método de Ward.....	108
FIGURA 14 – Valores médios, mínimos e máximos de IQA para estações de amostragem no período de 1997 a 2005 ¹⁻⁷	109
FIGURA 15 - Gráfico dos autovalores da matriz de covariância para as dez componentes principais da base de dados. As cinco primeiras componentes principais descrevem 82,2% da variabilidade dos dados.....	111
FIGURA 16 – Análise de agrupamento na matriz padronizada das medianas dos 32 parâmetros medidos nas 29 estações de amostragem do rio das Velhas, utilizando o método de Ward.....	113
FIGURA 17 – Gráfico dos autovalores da matriz de covariância para as dez componentes principais da matriz composta pelo 32 parâmetros mais importantes e OD. As cinco primeiras componentes principais descrevem 83,8% da variabilidade dos dados.....	116
FIGURA 18 - Gráfico <i>biplot</i> de escores e cargas dos parâmetros para as duas primeiras componentes principais para todas as estações da rede da macro-rede de monitoramento da sub-bacia do rio das Velhas. Escala de escores no intervalo de -0,3 a +0,5 e escala de pesos no intervalo de -4,5 a +4,5.....	118

- FIGURA 19** - Gráfico *biplot* de escores e cargas dos parâmetros para as componentes principais 1 e 3 para todas as estações da macro-rede de monitoramento da sub-bacia do rio das Velhas. Escala de escores no intervalo de -0,3 a +0,5 e escala de pesos no intervalo de -4,5 a +6,0. 121
- FIGURA 20** - Dendrograma indicando as estações separadas em grupos e o isolamento das estações AV310 e AV320. 128
- FIGURA 21** - Dispersões das variáveis de pares de estações próximas no dendrograma: painel A, estações AV220 e AV230; painel B, estações AV240, AV290; painel C, AV240 e AV280; painel D, estações AV280 e AV290. Os valores de r correspondem ao coeficiente de correlação de Pearson 129
- FIGURA 22** - Gráfico dos autovalores da matriz de covariância para as dez componentes principais da base de dados. As quatro primeiras componentes principais descrevem 79,6% da variabilidade dos dados. 131
- FIGURA 23**- Gráfico *biplot* de escores e cargas para as duas primeiras componentes principais para todas as estações da rede de monitoramento do alto curso do rio das Velhas. Escala de escores no intervalo de -1 a +1 e escala de pesos no intervalo de -6 a +6. 133
- FIGURA 24**- Gráfico *biplot* de escores e cargas para as duas primeiras componentes principais para as estações da rede básica, constituída das estações AV005, AV010, AV030, AV040, AV050, AV080, AV110, AV130, AV160, AV180, AV200, AV210, AV330, AV340, AV345 e AV350. Escala de escores no intervalo de -0,4 a +0,6 e escala de pesos no intervalo de -4 a +5. 134
- FIGURA 25**- Gráfico *biplot* de escores e cargas para as duas primeiras componentes principais para as estações da rede dirigida para controle da poluição, constituída das estações AV020, AV060, AV070, AV090, AV100, AV120, AV140, AV150, AV170, AV190, AV250, AV300, AV310 e AV320. Escala de escores no intervalo de -1 a +1 e escala de pesos no intervalo de -6 a +6. 135
- FIGURA 26**- Gráfico *biplot* de escores e cargas escalonados para as duas primeiras componentes principais para as estações da rede dirigida para controle da poluição excluindo as estações AV310 e AV320. Escala de escores no intervalo de -1 a +1 e escala de pesos no intervalo de -5 a +4. 136
- FIGURA 27**- Gráfico *biplot* de escores e cargas para as duas primeiras componentes principais para as estações AV220, AV230, AV240, AV260, AV270, AV280 e AV290, que constituem a rede dirigida para abastecimento doméstico. Escala de escores no intervalo de -0,6 a +0,8 e escala de pesos no intervalo de -2,5 a +3,5. 137
- FIGURA 28**- Dendrograma obtido na análise com os 22 principais parâmetros de cada uma das quatro componentes principais 138
- FIGURA 29** - Comparação dos dados históricos da concentração de fosfatos total no ribeirão do Onça, estação BV154, com os ajustes exponenciais dos modelos de Holt-Winter e Winter período de janeiro de 1994 a outubro de 2004 e as previsões em 2005 144
- FIGURA 30** – Comparação dos dados históricos da concentração de fosfatos total no rio das Velhas a jusante do ribeirão do Onça, estação BV105, com os ajustes exponenciais dos modelos de Holt- Winter e Winter período de janeiro de 1993 a outubro de 2004 e as previsões em 2005 148

FIGURA 31 – Comparação dos dados históricos da concentração de fosfatos total no ribeirão Arrudas, estação BV155, com os ajustes exponenciais dos modelos de Holt-Winter e Winter período de janeiro de 1994 a outubro de 2004 e as previsões em 2005	151
FIGURA 32 – Comparação dos dados históricos da concentração de fosfatos total no rio das Velhas a jusante do ribeirão Arrudas ,estação BV083, com os ajustes exponenciais dos modelos de Holt- Winter e Winter período de janeiro de 1994 a outubro de 2004 e as previsões em 2005	154
FIGURA 33 - Comparação dos dados históricos da concentração de nitrogênio amoniacal no ribeirão do Onça, PAINEL (A), e no rio da Velhas a jusante do ribeirão do Onça, PAINEL (B), com os ajustes exponenciais dos modelos de Winter e Holt- Winter ...	158
FIGURA 34 - Comparação dos dados históricos da concentração de nitrogênio amoniacal ribeirão Arrudas, PAINEL (A), e rio das Velhas a jusante do ribeirão Arrudas, PAINEL (B), com os ajustes exponenciais dos modelos de Holt - Winter e Holt- Winter com tendência amortecida	163
FIGURA 35 - Comparação dos dados históricos da concentração de manganês no rio das Velhas a jusante do ribeirão Arrudas, BV083, com os ajustes exponenciais dos modelos de Winter e Holt – Winter com tendência amortecida período de janeiro de 1993 a outubro de 2004 e as previsões em 2005	168
FIGURA 36 - Comparação dos dados históricos da concentração de manganês no rio Itabirito com os ajustes exponenciais dos modelos de Winter e Holt – Winter período de janeiro de 1993 a outubro de 2004 e as previsões em 2005	171
FIGURA 37 - Comparação dos dados históricos da concentração de oxigênio dissolvido no ribeirão do Onça, PAINEL (A), com os ajustes exponenciais pelos modelos de Winter e Holt – Winter com tendência amortecida e Winter e Holt – Winter no rio das Velhas a jusante do ribeirão do Onça, PAINEL (B).....	175
FIGURA 38 - Comparação dos dados históricos da concentração de oxigênio dissolvido no ribeirão Jequitibá com os ajustes exponenciais dos modelos de Winter e Holt – Winter período de janeiro de 1994 a outubro de 2004 e as previsões em 2005	179
FIGURA 39 - Comparação dos dados históricos da concentração de oxigênio dissolvido no rio Cipó com os ajustes exponenciais dos modelos de Winter e Holt – Winter período de janeiro de 2000 a outubro de 2004 e as previsões em 2005	182

LISTA DE TABELAS

TABELA 1 - Densidade de pontos de amostragem em cada bacia hidrográfica ⁷	4
TABELA 2 – Descrição das estações de amostragem da sub- bacia do rio das Velhas ⁸	8
TABELA 3 - Teores máximos de substâncias potencialmente prejudiciais em águas de classe 1, 2 e 3 ¹³	15
TABELA 4 - Atividades industriais e minerárias das regiões do alto, médio e baixo curso do rio das Velhas ⁸	76
TABELA 5 – Metodologia de análises dos principais parâmetros analisados em águas da sub-bacia do rio das Velhas ⁹⁰	78
TABELA 6 - Mediana dos parâmetros na macro-rede de monitoramento no período de setembro de 1997 a outubro de 2005	82
TABELA 7 - Mediana dos parâmetros nas estações de monitoramento do alto curso do rio das velhas no período de novembro de 2003 a dezembro de 2004.....	87
TABELA 8 – Classificação dos parâmetros monitorados em ordem decrescente de percentual de violação de classe de enquadramento em toda bacia do rio das Velhas, no período de 1997 a 2003 ⁸	92
TABELA 9 – Autovalores e cargas dos parâmetros nos fatores comuns	96
TABELA 10 – Equações de regressão para os parâmetros de maior carga no fator 1 para as estações BV062, BV063, BV067, BV076, BV083, BV105, BV130, BV135, BV137, BV153, BV154, BV155, BV156 e BV160	99
TABELA 11 – Equações de regressão para os parâmetros de maior carga no fator 2 para as estações BV013, BV035, BV037, BV062, BV063, BV067, BV076, BV083, BV105, BV130, BV135, BV137, BV139 e BV153	103
TABELA 12 – Autovalores, porcentagem de variância e variância acumulada para dez primeiras componentes principais da matriz da padronizada da mediana dos 41 parâmetros para 29 estações de amostragem ao longo da sub – bacia do rio das Velhas	110
TABELA 13 – Cinco primeiras componentes principais e respectivos componentes de cargas dos parâmetros com valor absoluto superior a 0,200 – matriz 29 x 41	112
TABELA 14 – Autovalores, porcentagem de variância e variância acumulada para dez primeiras componentes principais da matriz da padronizada da mediana dos 32 parâmetros mais importantes e oxigênio dissolvido para 29 estações de amostragem ao longo da sub – bacia do rio das Velhas.....	115
TABELA 15 - Cinco primeiras componentes principais e respectivos componentes de cargas dos parâmetros com valor absoluto superior a 0,200 - matriz 29 x 33.....	116
TABELA 16 – Autovalores, porcentagem de variância e variância acumulada para dez primeiras componentes principais da matriz da padronizada da mediana dos 39 parâmetros 40 estações de amostragem do alto curso do rio das Velhas.....	130
TABELA 17 - Cinco primeiras componentes principais e respectivos componentes de cargas dos parâmetros com valor absoluto superior a 0,200 - matriz 40 x 39.....	131
TABELA 18 – Comparação de modelos de previsão para teores para fosfatos total no ribeirão do Onça, estação BV154*	144

TABELA 19 – Valores reais e previstos para a concentração de fosfatos total em mg/L no ribeirão do Onça, estação BV154, no período de 2005 a 2007 pelo modelo de Winter	146
TABELA 20 – Comparação de modelos de previsão para teores para fosfatos total no rio das Velhas a jusante do ribeirão Onça, estação BV105*.....	147
TABELA 21 – Valores reais e previstos para a concentração de fosfato total em mg/L no rio das Velhas a jusante do ribeirão do Onça, estação BV105, no período de 2005 a 2007 pelo modelo de Winter.....	149
TABELA 22 – Comparação de modelos de previsão para teores para fosfatos total no ribeirão Arrudas, estação BV155*	150
TABELA 23 - Valores reais e previstos para a concentração de fosfatos total em mg/L no ribeirão Arrudas, estação BV155, no período de 2005 a 2007 pelo modelo de Holt com tendência amortecida	152
TABELA 24 – Comparação de modelos de previsão para teores para fosfatos total no rio das Velhas a jusante do ribeirão Arrudas, estação BV083*	153
TABELA 25 – Valores reais e previstos para a concentração de fosfatos total em mg/L no rio das Velhas a jusante do ribeirão Arrudas, estação BV083, em 2005 a 2007 pelo modelo de Winter.....	155
TABELA 26 – Comparação de modelos de previsão para teores para nitrogênio amoniacal nas estações BV154 e BV105	157
TABELA 27 – Valores reais e previstos para concentração de nitrogênio amoniacal em mg/L no ribeirão do Onça e no rio das Velhas a jusante do ribeirão do Onça de 2005 a 2007 pelo modelo de Holt–Winter.....	160
TABELA 28 – Comparação de modelos de previsão para teores para nitrogênio amoniacal no ribeirão Arrudas, BV155 e no rio das Velhas a jusante do ribeirão Arrudas, BV083	162
TABELA 29 – Valores reais e previstos para a concentração de nitrogênio amoniacal em mg/L na estação BV155 e BV083 de 2005 a 2007 pelo modelo de Holt – Winter com tendência amortecida	165
TABELA 30 - Comparação de modelos de previsão para teores de manganês total no rio das Velhas a jusante do ribeirão Arrudas, BV083*	167
TABELA 31 - Valores reais e previstos para a concentração de manganês em mg/L no rio das Velhas a jusante do ribeirão Arrudas no período de 2005 a 2007 pelo modelo de Winter	169
TABELA 32 - Comparação de modelos de previsão para teores de manganês total no rio Itabirito, BV035*	170
TABELA 33 – Valores reais e previstos para manganês no rio Itabirito, BV035, no período de 2005 a 2007 pelo modelo de Winter	172
TABELA 34 - Comparação de modelos de previsão para a concentração de oxigênio dissolvido em mg/L na estação ribeirão do Onça e rio das Velhas a jusante do ribeirão do Onça, BV154 e BV105	174
TABELA 35 – Valores reais e previstos para concentração de oxigênio dissolvido em mg/L nas estações ribeirão do Onça e no rio das Velhas a jusante do ribeirão do Onça no período de 2005 a 2007 pelos modelos de Holt – Winter com tendência amortecida e Winter respectivamente	177
TABELA 36 - Comparação de modelos de previsão para teores de oxigênio dissolvido no ribeirão Jequitibá, estação BV140*	178

TABELA 37 - Valores reais e previstos para concentração de oxigênio dissolvido em mg/L na estação no ribeirão Jequitibá no período de 2005 a 2007 pelo modelo de Winter .	180
TABELA 38 - Comparação de modelos de previsão para concentração de oxigênio dissolvido no rio Cipó, BV162*	181
TABELA 39 - Valores reais e previstos para concentração de oxigênio dissolvido em mg/L na estação BV162 de 2005 a 2007 pelo modelo de Winter	183

RESUMO

Um programa de monitoramento da qualidade das águas superficiais da sub-bacia do Rio Das Velhas, o principal afluente do Rio São Francisco que constitui a segunda maior bacia hidrográfica do Brasil, vem sendo realizado pelo Governo do Estado de Minas Gerais desde 1977. Este monitoramento gera uma base de dados de valores de mais de 50 parâmetros de qualidade de água medidos em 29 estações de monitoramento ao longo do rio e seus afluentes com frequência de leituras variando de mensal a anual, que é mantida pelo Instituto Mineiro de Gestão das Águas, IGAM. Assim, são detectadas as tendências de variação da qualidade das águas que são empregadas para estabelecer políticas públicas de controle ambiental tais como metas a serem atingidas para adequar os usos da água com níveis apropriados ao seu padrão de qualidade.

Este trabalho tem como objetivo a proposição de uma metodologia para o tratamento estatístico da base de dados da sub-bacia do Rio das Velhas, relativa ao período de 1997 a 2005, que emprega métodos estatísticos de análise multivariada e análise de séries temporais. A análise multivariada é aplicada à base de dados visando uma avaliação espacial das estações monitoradas destacando os parâmetros mais importantes e as regiões mais críticas. A avaliação temporal é proposta para investigar tendências e comportamento sazonal dos parâmetros de qualidade de água e a possibilidade de identificar características espaciais e temporais dos dados que produziram modelos de dependência espacial e temporal.

A análise fatorial associada à análise de variância permitiu a caracterização de áreas com maiores impactos de contaminações associadas a efluentes domésticos, industriais e de mineração, na região metropolitana de Belo Horizonte. Essas análises possibilitaram a elaboração de modelos de regressão não lineares preliminares de qualidade de água empregando os parâmetros mais correlacionados.

As análises de agrupamento hierárquico e de componentes principais foram empregadas para agrupar as estações de monitoramento em toda a sub-bacia, de acordo com similaridades e índices de qualidade de água, identificando os principais poluentes e as fontes de contaminação associadas e selecionando aqueles mais importantes na descrição da variabilidade da base de dados. A análise de componentes principais possibilitou uma importante redução no número estações de monitoramento e no número de parâmetros físico-químicos empregados no monitoramento, mantendo os mesmos agrupamentos das estações bem como a mesma identificação dos parâmetros mais importantes responsáveis pela grande variação na qualidade das águas.

A análise de séries temporais dos dados permitiu avaliar o comportamento dessas estações em relação às variações das concentrações dos poluentes mais críticos, destacados pela análise multivariada; detectando tendências, comportamento sazonal, correlações espaciais e temporais e gerando modelos de previsão para propor ações de controle mais efetivas.

A situação crítica da sub-bacia do rio das Velhas se deve principalmente ao impacto causado pelo transporte de efluentes de esgoto doméstico e industrial da região metropolitana de Belo Horizonte e pelas intensivas atividades minerárias de longa duração, que correm principalmente na região do seu alto curso, responsáveis pelo carregamento de rejeitos de mineração no seu leito. Os efeitos dessas fontes constantes de poluição se estendem por todo o Rio das Velhas e continuam sem a aplicação de medidas efetivas de controle.

A metodologia para tratamento dos dados aqui proposta ilustra a grande utilidade dos métodos estatísticos no tratamento de grandes bases de dados ambientais e poderá ser utilizada em outros estudos, permitindo uma diminuição do custo de programas de monitoração ambiental e um uso mais racional de pessoal e equipamentos, no laboratório e no campo.

ABSTRACT

A monitoring program of water quality of Rio Das Velhas sub-basin, the major tributary of the São Francisco River, the second larger river in Brazil, has been established by Government of the State of Minas Gerais since 1977, generating a large data base with values of more than 50 water quality parameters measured in 29 monitoring stations along the river and its tributaries, with monthly or yearly reading frequency, which is maintained by Instituto Mineiro de Gestão das Águas, IGAM. This effective, long-term monitoring of water quality establishes environmental policies and indicates tendencies of seasonal variation of water quality and goals to be achieved in order to adapt the use of the water with its quality standards.

The statistical treatment of the generated Rio Das Velhas data base in the period of 1997-2005 was performed employing methods of multivariate and time series analysis. The multivariate analysis was applied to the data base in order to perform a spatial evaluation of the sampling sites assigning the more important parameters and more critical areas. The temporal evaluation was performed in order to investigate seasonal behavior pattern of the water quality parameters and the possibility of spatial and temporal patterns in water quality that lead to space-time dependencies.

The factor analysis associated to the variance analysis allowed the characterization of the geographical areas with larger impacts of contaminations related to domestic, industrial and mining effluents in Belo Horizonte metropolitan area. Such analyses allowed the elaboration of preliminary non-linear regression models of water quality with the more correlated parameters.

Hierarchical cluster analysis and principal component analysis were employed to group the monitoring sites along the sub-basin according to similarities and water quality index, identifying the more important pollutants and polluting sources and selecting the most important ones which should describe the variability of the data base. The principal

components analyses provided an important reduction in the number of sampling stations and in the number of analyzed parameters maintaining the same clustering pattern of stations and the same identification of the more important parameters responsible for the large variation in water quality.

The time series data analysis allowed evaluating the behavior of some stations with respect to the variations of concentration of the most critical pollutants, assigned by the multivariate analysis, detecting tendencies, seasonality, space-time correlations and generating forecast models and means to propose more effective control actions.

The critical situation of the sub-basin of Das Velhas River is mainly due to the impact caused by the transport of domestic and industrial waste water discharge of Belo Horizonte metropolitan area, and the long term intensive mining activities mostly located on its high course, responsible for rejected materials carried into its waters. The effect of these constant sources of pollution is widespread and will continue without a suitable control activity.

The employed methodology of data treatment illustrates the usefulness of statistical methods in the treatment of large environmental data sets and could be used in any water quality study, allowing a reduction of the cost of environmental monitoring program and a more rational use of personnel and equipment, in laboratory and *in situ* as well.

ABREVIATURAS

IGAM – Instituto Mineiro de Gestão das Águas

FEAM - Fundação Estadual do Meio Ambiente

RV – Rio das Velhas

AV – Alto Velhas

MMA – Ministério do Meio Ambiente

SRH – Secretaria de Recursos Hídricos

SEMAD – Secretaria de Estado de Meio Ambiente e Desenvolvimento Sustentável

CONAMA – Conselho Nacional do Meio Ambiente

COPAM – Conselho Estadual de Política Ambiental

GPS – Global Position System

PROSAM – Programa de Saneamento Ambiental

PNMA II – Programa Nacional do Meio Ambiente II

ETE – Estação de Tratamento de Esgoto

Capítulo I – Introdução

1.1 Monitoramento da qualidade das águas superficiais no Estado de Minas Gerais

Os trabalhos de monitoramento e enquadramento dos cursos d'água de bacias hidrográficas indicam respectivamente as tendências de variação da qualidade das águas e estabelecem metas a serem atingidas para adequar os usos da água com os níveis apropriados à sua qualidade. A avaliação da qualidade dessas águas, sob a ótica das atividades desenvolvidas na bacia hidrográfica, direciona a definição de ações prioritárias para a melhoria contínua da qualidade ambiental da região.¹⁻⁸

O monitoramento das águas em Minas Gerais teve seu início em 1977, com a rede de amostragem operada pela Fundação Centro Tecnológico de Minas Gerais - CETEC, visando à avaliação da qualidade das águas das bacias do rio das Velhas, rio Paraopeba e rio Paraíba do Sul até o ano de 1988. A FEAM monitorou a bacia hidrográfica do rio Verde de 1987 a 1995 utilizando os serviços do CETEC. Nos anos seguintes, utilizando os

serviços da GEOSOL - Geologia e Sondagens – e, posteriormente, do CETEC, as bacias hidrográficas do rio das Velhas e do rio Paraopeba foram monitoradas de 1993 a 1997.⁸

Com a valorização dos recursos hídricos refletida na promulgação da Lei 9.433/97⁹ e a conseqüente criação de órgãos federais e estaduais dirigidos ao gerenciamento racional das águas, o trabalho de monitoramento foi reforçado pela FEAM, em 1997, desta vez com um monitoramento mais amplo e completo, estendido às oito principais bacias hidrográficas mineiras por meio de convênio com o Ministério do Meio Ambiente - MMA. No final de 1999, o Governo do Estado de Minas Gerais, por intermédio do Conselho Estadual de Recursos Hídricos - CERH, também destinou recursos para o Projeto Águas de Minas, passando o IGAM a integrar a coordenação do mesmo. O Projeto Águas de Minas tem por objetivo implantar o sistema de monitoramento físico – químico e bacteriológico da qualidade das águas superficiais das principais bacias hidrográficas de Estado de Minas Gerais. Em 2002, por estar mais adequadamente inserido nas competências da Agenda Azul do que nas da Agenda Marrom, a coordenação geral deste projeto passou para o IGAM, com participação da FEAM.⁸

A rede de monitoramento consiste de 242 estações de amostragem que abrangem as bacias do Rio São Francisco, Rio Grande, Rio Doce, Rio Paranaíba, Rio Jequitinhonha, Rio Paraíba do sul, Rio Mucuri e Rio Pardo 577.015 km², o que representa 98,3% de sua área total.¹⁻⁷

A definição dos locais de coleta é feita identificando áreas que caracterizam as condições naturais das águas de cada bacia hidrográfica e as principais interferências antrópicas, especialmente relacionadas à ocupação urbana e às atividades industriais e minerárias, além da agropecuária e silvicultura. Além disso, são considerados redes de qualidade de água anteriormente operadas em Minas Gerais e dados dos processos de licenciamento ambiental da FEAM/COPAM.¹⁻⁷

A rede em operação vem sendo adequada ao longo da execução dos trabalhos, adotando-se como referência a experiência desenvolvida pelos países membros da União Européia. Assim sendo, é estabelecida como meta a razão de uma estação de monitoramento por 1.000 km², que é a densidade média adotada nos mencionados países. Os números de pontos de coleta por bacia e sub-bacia contempladas, com as respectivas densidades, são apresentados na TABELA 1.¹⁻⁷

A sub-bacia do rio das Velhas que é foco deste trabalho é pertencente à bacia do Rio São Francisco, juntamente com as sub-bacias do rio São Francisco Sul, rio Pará, rio Paraopeba, rio das Velhas e rio São Francisco Norte. Em função da grande área da bacia do Rio São Francisco, da diversidade das condições naturais e econômicas da região e visando uma melhor descrição das diferentes características da mesma, a sua avaliação foi feita em cinco sub-bacias distintas, TABELA 1.¹⁻⁷

A escolha da sub-bacia do rio das Velhas se deve a sua grande importância econômica, social e histórica, aliadas ao seu grande interesse atual na recuperação das suas águas, foco de importantes projetos, tal como o Projeto Manuelzão. A questão da revitalização da sub-bacia do rio das Velhas está incluída nas principais políticas do Estado para recursos hídricos. Há projetos de recuperação dos seus efluentes que passam pela região metropolitana de Belo Horizonte com instalação de estações de tratamento de esgotos através do Programa de Saneamento Ambiental, PROSAM.⁸

TABELA 1 - Densidade de pontos de amostragem em cada bacia hidrográfica ⁷

BACIA HIDROGRÁFICA SUB-BACIA	NÚMERO DE PONTOS DE AMOSTRAGEM	DENSIDADE PONTOS/1000 km²
RIO SÃO FRANCISCO	97	0,41
São Francisco Sul	12	0,37
Rio Pará	13	1,06
Rio Paraopeba	18	1,49
Rio das Velhas	29	0,98
São Francisco Norte	25	0,17
RIO GRANDE	42	0,48
Rio das Mortes	7	1,06
Rio Verde	12	1,74
Restante da Bacia	23	0,31
RIO DOCE	32	0,45
Rio Piracicaba	8	1,49
Restante da Bacia	24	0,37
RIO PARANAÍBA	18	0,25
RIO JEQUITINHONHA	13	0,20
RIO PARAÍBA DO SUL	29	1,38
Rio Paraíbauna	8	1,18
Restante da Bacia	21	1,48
RIO MUCURI	8	0,55
RIO PARDO	3	0,24
TOTAL	242	0,42

1.2 A sub – bacia do rio das Velhas

O rio das Velhas tem sua nascente na Serra do Veloso, proximidade de Ouro Preto, e deságua no Rio São Francisco, em Guaiçuí, distrito de Várzea da Palma, próximo a Pirapora, com uma vazão média de 265 m³/s, percorrendo uma extensão de 716 km e drenando uma bacia de 29.173 km².¹⁻⁷ Para facilitar a descrição da bacia, ela é dividida em alto, médio e baixo curso. O alto curso bacia do rio das Velhas compreende uma área de 1.943 km², englobando desde a proximidade de sua nascente entre a Serra de Antônio Pereira, município de Ouro Preto, distrito de São Bartolomeu, até os limites dos municípios de Raposos e Sabará englobando os municípios de Belo Horizonte, Contagem e Caeté. O médio curso compreende a região do Rio Paraúna, o principal afluente do rio das Velhas, do córrego Salobinho ao norte, até os municípios de Curvelo e Corinto. O baixo curso compreende os municípios de Corinto, Diamantina, Buenópolis e Pirapora.¹⁰ Em toda extensão da bacia, estão localizados 56 municípios, com um total e 4,1 milhões de habitantes nas áreas urbanas e 186,5 mil na zona rural.⁸

Na sub – bacia do rio das Velhas são avaliados, atualmente, vinte e nove estações de amostragem no âmbito da macro – rede de monitoramento pelo projeto Águas de Minas. Dessas estações, dezesseis são ao longo do rio das Velhas.⁵⁻⁸ Além dessas estações, o IGAM vem realizando desde novembro de 2002, paralelamente ao monitoramento de toda a bacia do rio das Velhas, o aperfeiçoamento do monitoramento da qualidade das águas da bacia do alto curso do rio das Velhas, pelo Programa Nacional do Meio Ambiente II – PNMA II ¹¹⁻¹², em uma rede dirigida que objetiva o aumento do nível de conhecimento das condições de qualidade das águas de uma área específica de grande importância econômica e social do rio das Velhas e de seus formadores e em particular dos pontos críticos para o abastecimento de água, permitindo avaliações mais precisas da efetividade dos sistemas de controle ambiental implantados. A rede dirigida contempla trinta e sete estações de monitoramento de qualidade das águas, das quais, sete estações são coincidentes com a macro–rede operadas pelo Projeto Águas de Minas, sete estações dirigidas

ao abastecimento doméstico e operadas pela Companhia de Saneamento de Minas Gerais, COPASA, e mais 23 novas estações de amostragem. Entre as estações monitoradas na região do alto curso nove estações estão localizadas ao longo do rio das Velhas.

A descrição das estações de amostragem está apresentada na TABELA 2. A FIGURA 1 apresenta a localização dessas estações em um mapa. A região do alto curso do rio das Velhas é destacada a esquerda do mapa por possuir uma rede de monitoramento específica e mais abrangente. Para distinção, as estações pertencentes à macro-rede possuem a sigla BV, Bacia Velhas, e as estações pertencentes à rede dirigida à região do alto curso do rio das Velhas possuem a sigla AV, Alto Velhas.

Entre as vinte e nove estações que compõem a macro-rede de monitoramento as estações BV154, ribeirão do Onça a montante do rio das Velhas e BV155, ribeirão Arrudas a montante do rio das Velhas foram inseridas no programa em maio de 1994 e as estações BV160, ribeirão das Neves próximo de sua foz no ribeirão da Mata; BV161, ribeirão Santo Antônio próximo de sua foz no rio das Velhas e BV162, Rio Cipó a montante da foz do rio Paraúna, em maio de 2000.

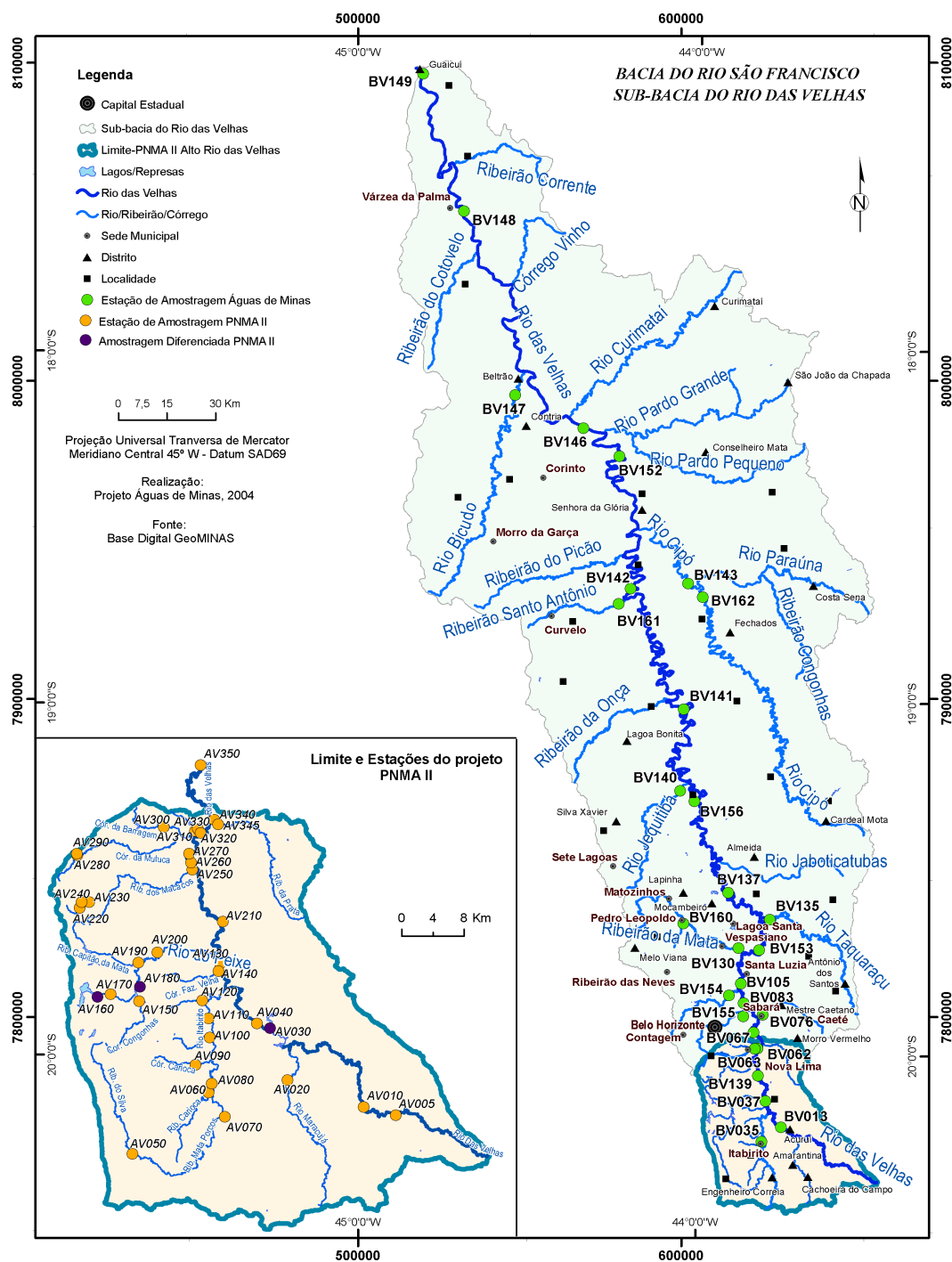


FIGURA 1 – Localização das estações de amostragem

TABELA 2 – Descrição das estações de amostragem da sub- bacia do rio das Velhas⁸

CÓDIGO	DESCRIÇÃO	LATITUDE	LONGITUDE
AV005	Rio das Velhas a montante de São Bartolomeu	20°17'	43°35'
AV010	Rio das Velhas a jusante do ribeirão do Funil	20°17'	43°37'
AV020	Rio Maracujá a montante da confluência do Córrego dos Padres	20°16'	43°43'
AV030	Represa Rio de Pedras – Corpo da barragem	20°13'	43°44'
BV013/AV040	Rio das Velhas a montante do Rio Itabirito	20°10'	43°47'
AV050	Ribeirão do Silva a montante do Córrego das Almas	20°21'	43°54'
AV060	Ribeirão Carioca a montante de sua confluência com o ribeirão Mata Porcos	20°17'	43°49'
AV070	Ribeirão Mata Porcos próximo de sua confluência com o ribeirão Sardinha	20°19'	43°47'
AV080	Rio Itabirito a montante de Itabirito	20°17'	43°48'
AV090	Córrego Carioca a montante da cidade de Itabirito	20°15'	43°49'
AV100	Córrego da Onça a montante do rio Itabirito	20°13'	43°48'
BV035/AV110	Rio Itabirito a jusante do Córrego da Onça	20°14'	43°48'
AV120	Córrego Moleque a montante do rio Itabirito	20°11'	43°49'
BV037/AV130	Rio das Velhas a jusante do rio Itabirito	20°08'	43°48'
AV140	Córrego Fazenda Velha a montante do rio das Velhas.	20°08'	43°47'
AV150	Ribeirão Congonhas a montante da represa das Codornas	20°11'	43°53'
AV160	Corpo da barragem Lagoa Grande	20°11'	43°57'
AV170	Córrego Lagoa Grande a montante da represa das Codornas	20°10'	43°56'
AV180	Lagoa das Codornas	20°10'	43°53'
AV190	Ribeirão Capitão da Mata a montante do rio do Peixe	20°08'	43°54'
AV200	Rio do Peixe a montante do rio das Velhas	20°07'	43°52'
BV139/AV210	Rio das Velhas a montante da ETA/ COPASA, em Bela Fama	20°04'	43°49'
AV220	Córrego Fechos na barragem principal	20°04'	43°58'
AV230	Córrego Fechos na barragem auxiliar	20°04'	43°57'
AV240	Córrego Seco na captação Fechos	20°04'	43°58'
AV250	Ribeirão Macacos a montante do rio das Velhas	20°02'	43°50'
AV260	Rio das Velhas em Bela Fama, alça direita	20°01'	43°50'
AV270	Rio das Velhas em Bela Fama - alça esquerda	20°01'	43°50'
AV280	Córrego Mutuca na captação Mutuca, barragem principal	20°01'	43°58'
AV290	Córrego Mutuca na captação Mutuca, barragem auxiliar	20°01'	43°58'
AV300	Córrego da Barragem em Nova Lima	19°59'	43°51'

TABELA 2 – Descrição das estações de amostragem da sub- bacia do rio das Velhas⁸ - CONTINUAÇÃO

CÓDIGO	DESCRIÇÃO	LATITUDE	LONGITUDE
BV062/AV310	Ribeirão Água Suja a montante do rio das Velhas	19°59'	43°50'
AV320	Córrego da Mina a montante do rio das Velhas	19°59'	43°49'
BV063/AV330	Rio das Velhas a jusante do Córrego da Mina	19°58'	43°49'
AV340	Ribeirão da Prata a montante do rio das Velhas	19°58'	43°48'
AV345	Ribeirão da Prata a montante da cidade de Raposos	19°58'	43°48'
BV067/AV350	Rio das Velhas a montante do ribeirão Sabará	19°55'	43°50'
BV076	Ribeirão Sabará a montante do rio das Velhas	19°53'	43°49'
BV155	Ribeirão Arrudas a montante do rio das Velhas	19°54'	43°51'
BV083	RV a jusante do ribeirão Arrudas	19°51'	43°52'
BV154	Ribeirão do Onça a montante do rio das Velhas	19°50'	43°51'
BV105	Rio das Velhas a jusante do ribeirão do Onça	19°49'	43°53'
BV160	Ribeirão das Neves próximo de sua foz no ribeirão da Mata	19°37'	44°02'
BV130	Ribeirão da Mata a montante do rio das Velhas	19°42'	43°53'
BV153	Rio das Velhas a jusante do ribeirão da Mata	19°42'	43°49'
BV135	Ribeirão Taquaraçu a montante do rio das Velhas	19°37'	43°48'
BV137	Rio das Velhas na ponte Raul Soares	19°33'	43°55'
BV156	Rio das Velhas a jusante do ribeirão Jaboticatubas	19°21'	44°00'
BV140	Ribeirão Jequitibá a montante do rio das Velhas	19°14'	44°01'
BV141	Rio das Velhas a jusante do ribeirão Jequitibá na cidade de Santana do Pirapama	19°01'	44°02'
BV161	Ribeirão Santo Antônio próximo de sua foz no rio das Velhas	18°43'	44°13'
BV142	Rio das Velhas a montante da foz do rio Paraúna	18°40'	44°12'
BV162	Rio Cipó a montante da foz do rio Paraúna	18°41'	43°59'
BV143	Rio Paraúna próximo de sua foz no rio das Velhas	18°38'	44°03'
BV152	Rio das Velhas em Santo Hipólito entre os rios Paraúna e Pardo Grande	18°18'	44°14'
BV146	Rio das Velhas a jusante do Rio Pardo Grande	18°13'	44°21'
BV147	Rio Bicudo próximo de sua foz no do rio das Velhas	18°08'	44°32'
BV148	Rio das Velhas em Várzea da Palma	17°36'	44°43'
BV149	RV em Guaicuí, estação que antecede o deságue no Rio São Francisco	17°12'	44°49'

1.3 Requisitos básicos de qualidade

1.3.1 Legislação

As principais normas legais que estabelecem os requisitos básicos de qualidade para água são: ¹⁻⁷

- Lei Nº 6938/81, de 02 de setembro de 1981 - Dispõe sobre a Política Nacional do Meio Ambiente, seus fins e mecanismos de formulação e aplicação, e dá outras providências.
- Resolução CONAMA Nº 357 de 17 de março de 2005. Estabelece a classificação das águas doces, salobras e salinas do Território Nacional e respectivas normas e padrões para qualidade das águas e lançamento de efluentes nas coleções de água, e dá outras providências.¹³
- Deliberação Normativa COPAM Nº 010/86, de 16 de dezembro de 1986 - Estabelece normas e padrões para qualidade das águas, lançamento de efluentes nas coleções de águas, e dá outras providências.
- Lei Nº 13.199 de 29 de janeiro de 1999 - Dispõe sobre a Política Estadual de Recursos Hídricos e dá outras providências.
- Lei Nº 9.433, de 08 de janeiro de 1997 - Institui a Política Nacional de Recursos Hídricos.⁹

1.3.2 Classificação de águas

A necessidade de qualidade da água está intrinsecamente relacionada ao uso que será feito desta água. A legislação em vigor, RESOLUÇÃO CONAMA Nº 357, de 17 de março de 2005, classifica as águas doces, salobras e salinas do Território Nacional, em treze classes:¹³

Águas doces

As águas doces são classificadas em:

I - Classe especial - águas destinadas:

- ao abastecimento para consumo humano, com desinfecção;
- à preservação do equilíbrio natural das comunidades aquáticas; e,
- à preservação dos ambientes aquáticos em unidades de conservação de proteção integral.

II - Classe 1 - águas que podem ser destinadas:

- ao abastecimento doméstico após tratamento simplificado;
- à proteção das comunidades aquáticas;
- à recreação de contato primário, tais como natação, esqui aquático e mergulho, conforme Resolução CONAMA nº 274, de 2000;
- à irrigação de hortaliças que são consumidas cruas e de frutas que se desenvolvam rentes ao solo e que sejam ingeridas cruas sem remoção de película; e,
- à proteção das comunidades aquáticas em Terras Indígenas.

III - Classe 2 - águas que podem ser destinadas:

- ao abastecimento para consumo humano, após tratamento convencional;
- à proteção das comunidades aquáticas;
- à recreação de contato primário;

- à irrigação de hortaliças e plantas frutíferas e de parques, jardins, campos de esporte e lazer, com os quais o público possa vir a ter contato direto; e;
- à aquicultura e à atividade de pesca.

IV - Classe 3 - águas que podem ser destinadas:

- ao abastecimento para consumo humano, após tratamento convencional ou avançado;
- à irrigação de culturas arbóreas, cerealíferas e forrageiras;
- à pesca amadora;
- à recreação de contato secundário, , referente a atividades em que o contato com a água é esporádico ou acidental e a possibilidade de ingerir água é pequena, como na pesca e na navegação (iatismo); e
- à dessedentação de animais.

V - Classe 4 - águas que podem ser destinadas:

- à navegação; e
- à harmonia paisagística;

Águas salinas

As águas salinas são classificadas em:

I - Classe especial - águas destinadas:

- à preservação dos ambientes aquáticos em unidades de conservação de proteção integral;
- à preservação do equilíbrio natural das comunidades aquáticas.

II - Classe 1 - águas que podem ser destinadas:

- à recreação de contato primário;
- à proteção das comunidades aquáticas; e
- à aquicultura e à atividade de pesca.

III - Classe 2 - águas que podem ser destinadas:

- à pesca amadora, e
- à recreação de contato secundário.

IV - Classe 3 – águas que podem ser destinadas:

- à navegação; e
- à harmonia paisagística.

Águas salobras

As águas salobras são assim classificadas:

I - Classe especial - águas destinadas:

- à preservação dos ambientes aquáticos em unidades de conservação de proteção integral;
- à preservação do equilíbrio natural das comunidades aquáticas.

II - Classe 1 - águas que podem ser destinadas:

- à recreação de contato primário;
- à proteção das comunidades aquáticas; e
- à aqüicultura e à atividade de pesca;
- ao abastecimento para consumo humano após tratamento convencional ou avançado; e
- à irrigação de hortaliças que são consumidas cruas e de frutas que se desenvolvam rentes ao solo e que sejam ingeridas cruas sem remoção de película, e à irrigação de parques, jardins, campos de esportes e lazer, com os quais o público possa vir a ter contato direto.

III - Classe 2 – águas que podem ser destinadas:

- à pesca amadora; e
- à recreação de contato secundário.

IV - Classe 3 – águas que podem ser destinadas:

- à navegação; e
- à harmonia paisagística.

Os teores máximos de substâncias potencialmente prejudiciais em águas de classe 1, 2 e 3 são listados na TABELA 3. O termo ambiente lêntico é referente a ambientes com água parada, com movimento lento ou estagnado; e ambiente lótico é relativo a águas continentais moventes.¹³

TABELA 3- Teores máximos de substâncias potencialmente prejudiciais em águas de classe 1, 2 e 3¹³

PARÂMETROS	TEORES MÁXIMOS		PARÂMETROS	TEORES MÁXIMOS	
	CLASSE 1 (mg/L)*	CLASSE 3 (mg/L)*		CLASSE 1 (mg/L)	CLASSE 3 (mg/L)
Alumínio	0,1 Al	0,2 Al	Fenóis totais	0,003 C ₆ H ₅ OH	0,01 C ₆ H ₅ OH
Antimônio	0,005 Sb	-	Ferro dissolvido	0,3 Fe	5,0 Fe
Arsênio total	0,01 As	0,033 As	Fluoreto total	1,4 F	1,4 F
Bário total	0,7 Ba	1,0 Ba	Fósforo total ¹	0,020 P	0,05 P
Berílio total	0,04 Be	0,1 Be	Fósforo total ²	0,025 P	0,075 P
Boro total	0,5 B	0,75 B	Fósforo total ³	0,1 P	0,15 P
Cádmio total	0,001 Cd	0,01 Cd	Lítio total	2,5 Li	2,5 Li
Chumbo total	0,01 Pb	0,033 Pb	Manganês total	0,1 Mn	0,5 Mn
Cianeto livre	0,005 CN	0,022 CN	Mercúrio total	0,0002 Hg	0,002 Hg
Cloreto total	250 Cl	250 Cl	Níquel total	0,025 Ni	0,025 Ni
Cloro residual total	0,01 Cl	-	Nitrato	10 N	10 N
Cobalto total	0,05 Co	0,2 Co	Nitrito	1,0 N	1,0 N
Cobre dissolvido	0,0009 Cu	0,013 Cu	pH	6,0 a 9,0	6,0 a 9,0
Cor	30 UPt	75 UPt	Prata total	0,01 Ag	0,05 Ag
Cromo total	0,05 Cr	0,05 Cr	Selênio total	0,01 Se	0,05 Se
DBO (5 dias a 20 ⁰ C)	3 de O ₂	10 de O ₂			

TABELA 3- Teores máximos de substâncias potencialmente prejudiciais em águas de classe 1, 2 e 3¹³
CONTINUAÇÃO

PARÂMETROS	TEORES MÁXIMOS		PARÂMETROS	TEORES MÁXIMOS	
	CLASSE 1 (mg/L)*	CLASSE 3 (mg/L)*		CLASSE 1 (mg/L)	CLASSE 3 (mg/L)
Nitrogênio amoniacal	3,7 N, para pH ≤ 7,5	13,3 N, para pH ≤ 7,5	Nitrogênio amoniacal	1,0 N, para 8,0 < pH ≤ 8,5	2,2 N, para 8,0 < pH ≤ 8,5
Nitrogênio amoniacal	2,0 N, para 7,5 < pH ≤ 8,0	5,6 N, para 7,5 < pH ≤ 8,0	Nitrogênio amoniacal	0,5 N, para pH > 7,5	1,0 N, para pH > 7,5
Oxigênio dissolvido	> 6,0 O ₂	> 4,0 O ₂	Sulfato total	250 SO ₄	250 SO ₄
Urânio total	0,02 U	0,02 U	Sulfeto **	0,002 S	0,3 S
Vanádio total	0,1 V	0,1 V	Sólidos dissolvidos totais	500	500
Zinco total	0,18 Zn	5,0 Zn	Turbidez	< 40 UNT***	< 100 UNT***

* exceto cor, pH e turbidez

** H₂S não dissociado

***UNT - unidades nefelométricas de turbidez

1 - ambiente lântico

2 - ambiente intermediário

3 - ambiente lótico

Para as águas doces de classe 2, estabelece - se os valores de até 5 mg/L O₂ para DBO, 0,030 mg/L de fósforo total em ambientes lânticos, 0,050 mg/L de fósforo total em ambientes intermediários, oxigênio dissolvido não inferior a 5 mg/L , 75 UPt para cor e turbidez até 100 UNT. Aos demais parâmetros aplicam-se as mesmas condições e padrões de águas de classe 1.¹³

1.4 Parâmetros indicativos da qualidade das águas

A poluição das águas tem como origem fontes diversas, associadas ao tipo de uso e ocupação do solo. Entre as principais fontes de poluição destacam-se: efluentes domésticos, efluentes industriais, carga difusa urbana, poluentes decorrentes de atividades ligadas à agricultura, silvicultura e pecuária, mineração, natural e acidental.⁵⁻⁷

Cada uma das fontes citadas possui acima características próprias quanto aos poluentes que carregam. Os esgotos domésticos apresentam compostos orgânicos biodegradáveis, nutrientes e microrganismos patogênicos. Nos efluentes industriais, há uma maior diversificação nos poluentes devido aos tipos de matérias – primas e processos industriais utilizados. Quanto à atividade agrícola, seus efeitos dependem muito das práticas utilizadas em cada região e da época do ano em que se realizam as preparações do terreno para o plantio, assim como, do uso intensivo dos defensivos agrícolas.⁵⁻⁷

A poluição natural está associada às chuvas e ao escoamento superficial, a salinização, à decomposição de animais e vegetais e a acidental é proveniente de derramamentos acidentais na linha de produção ou transporte.⁵⁻⁷

Foram adotados pelo IGAM parâmetros de monitoramento físicos, químicos, microbiológicos e bioensaios ecotoxicológicos que permitem caracterizar a qualidade da água e o grau de contaminação dos cursos d'água do estado de Minas Gerais. Entre os parâmetros físicos estão: temperatura do ar, temperatura da água, condutividade elétrica, sólidos totais, sólidos em suspensão, sólidos dissolvidos, cor, turbidez, alcalinidade total, alcalinidade bicarbonato, dureza cálcio, dureza magnésio e dureza total. Entre os parâmetros químicos: pH, oxigênio dissolvido (OD), demanda bioquímica de oxigênio (DBO), demanda química de oxigênio (DQO), nitrogênio orgânico, nitrogênio amoniacal, nitratos, nitritos, fosfatos total, surfactantes aniônicos, óleos e graxas, cianetos, fenóis, cloretos, ferro, potássio, sódio, sulfetos, sulfatos, cálcio, magnésio, manganês, zinco, bário, cádmio, boro, arsênio, níquel, chumbo, cobre, cromo (III), cromo (VI), selênio e mercúrio.

Entre os parâmetros microbiológicos: coliformes fecais, coliformes totais e estreptococos totais. Para avaliação da contaminação em cada estação são adotados pelo IGAM dois indicadores: IQA, Índice de qualidade das águas e CT, Contaminação por tóxicos.⁵⁻⁷

Os bioensaios ecotoxicológicos correspondem a testes de toxicidade que consistem na determinação do potencial tóxico de um agente químico ou de uma mistura complexa, sendo os efeitos desses poluentes detectados através da resposta de organismos vivos. No ensaio de toxicidade adotado pelo IGAM, o organismo aquático utilizado é a *Ceriodaphnia dubia*. São utilizadas as denominações agudo, crônico e não tóxico. O efeito agudo é caracterizado por uma resposta severa e rápida a um estímulo, a qual se manifesta nos organismos aquáticos em tempos relativamente curtos, 0 a 4 dias. O efeito crônico caracteriza-se pela resposta a um estímulo que continua por longos períodos de exposição do organismo aquático ao poluente que podem ser expressas através de mudanças comportamentais, alterações fisiológicas, genéticas e de reprodução. Os corpos de água não apresentam condições adequadas para a manutenção da vida aquática quando há efeitos de toxicidade aguda e crônica nas amostras de águas analisadas.⁵⁻⁷

O IQA, desenvolvido pela *National Sanitation Foundation*, reflete a interferência por esgotos sanitários, nutrientes e sólidos, é calculado a partir dos parâmetros oxigênio dissolvido, coliformes fecais, pH, demanda bioquímica de oxigênio, nitratos, fosfatos, variação na temperatura, turbidez e sólidos totais atribuindo pesos específicos a cada um³. Os valores de IQA variam de 0 a 100 e o nível de qualidade da água é avaliado como excelente ($90 < \text{IQA} \leq 100$), bom ($70 < \text{IQA} \leq 90$), médio ($50 < \text{IQA} \leq 70$), ruim ($25 < \text{IQA} \leq 50$) e muito ruim ($0 < \text{IQA} \leq 25$).^{8,14}

A CT é avaliada considerando-se a concentração de amônia, arsênio, bário, cádmio, chumbo, cianetos, cobre, cromo hexavalente, índice de fenóis, mercúrio, nitritos, nitratos e zinco. A caracterização por tóxicos é classificada como baixa ($\text{CT} < 1,2\text{P}$), média ($1,2\text{P} < \text{CT} < 2\text{P}$) ou alta ($\text{CT} > 2\text{P}$). Onde P corresponde ao limite definido na classe de enquadramento do curso d'água pelo Conselho Estadual de Política Ambiental – COPAM

na Deliberação Normativa Nº 10/86 para o trecho onde é localizada a estação de amostragem.⁸

1.5 Coletas e análises

Os serviços de amostragem e análises laboratoriais são contratados à Fundação Centro Tecnológico de Minas Gerais – CETEC – MG que possui certificação do IMETRO para os ensaios.

As atividades de coleta e análise laboratoriais são executadas com uma frequência trimestral, sendo, portanto, realizadas quatro campanhas durante um ano. Nos períodos críticos de fluxo das águas, cheias e estiagem, as campanhas de amostragem são completas.

Nessas campanhas são avaliados, por estação de coleta, 51 parâmetros físico-químicos e bacteriológicos. Entre esses parâmetros estão: condutividade elétrica, oxigênio dissolvido, pH *in loco*, pH laboratorial, temperatura da água, temperatura do ar, alcalinidade de bicarbonato, alcalinidade total, arsênio total, bário total, boro solúvel, cádmio total, cálcio total, chumbo total, cianeto total, cloretos, cobre total, coliformes fecais, coliformes totais, cor real, cromo hexavalente, cromo trivalente, demanda bioquímica de oxigênio, demanda química de oxigênio, dureza de cálcio, dureza de magnésio, dureza total, estreptococos fecais, ferro solúvel, fosfato total, índice de fenóis, magnésio total, manganês total, mercúrio total, níquel total, nitrogênio amoniacal, nitrogênio nítrico, nitrogênio nitroso, nitrogênio orgânico, óleos e graxas, potássio solúvel, selênio total, sódio solúvel, sólidos dissolvidos, sólidos suspensos, sólidos totais, sulfatos, sulfetos, surfactantes aniônicos, turbidez e zinco total.¹⁻⁸

Nas campanhas intermediárias são avaliados os principais indicadores da qualidade da água e parâmetros específicos, definidos com base no conhecimento da situação de qualidade de água e das atividades econômicas desenvolvidas na área de contribuição direta de cada estação. Entre os parâmetros físicos – químicos e bacteriológicos monitorados nessas

campanhas estão: cloretos, coliformes fecais, condutividade elétrica, demanda bioquímica de oxigênio, demanda química de oxigênio, fosfatos, nitrogênio orgânico, nitrogênio amoniacal, nitratos, nitritos, oxigênio dissolvido, pH, sólidos dissolvidos, sólidos em suspensão, sólidos totais, temperatura da água, temperatura do ar e turbidez. ¹⁻⁸

Os resultados de qualidade de água gerados são armazenados em uma base de dados, que contém informações atuais e históricas, permitindo observar a evolução da qualidade das águas nas últimas duas décadas. ¹⁻⁸

Este trabalho, desenvolvido a partir de um acordo de cooperação científica entre o Instituto de Ciências Exatas da Universidade Federal de Minas Gerais e o Instituto Mineiro de Gestão das Águas, IGAM, teve como objetivo a proposição de uma metodologia para o tratamento estatístico da base de dados mantida pelo IGAM relativa à bacia do rio das Velhas através de métodos estatísticos de análise multivariada e temporal; especificamente, objetivou-se:

- Proceder a uma análise inicial da base de dados do rio das Velhas buscando caracterizar as regiões mais críticas em termos de poluição, caracterizando os parâmetros mais importantes e seus respectivos agentes de contaminação.
- Proceder a uma análise exploratória de toda bacia do rio das Velhas, buscando correlacionar as estações de amostragem ao longo da sub-bacia, os principais parâmetros e seus respectivos agentes de contaminação. Paralelamente, pretende-se distinguir as estações com as mesmas características.
- Proceder a uma análise da base de dados da rede dirigida ao alto curso do rio das Velhas, buscando atender os mesmos objetivos acima e também propor uma nova rede de monitoramento reduzindo se possível, parâmetros e estações.

- Diagnosticar as estações mais críticas e os principais poluentes e assim proceder a uma análise temporal desses parâmetros buscando estabelecer um modelo de previsão para essas estações.

A análise multivariada aplicada à base foi proposta para uma avaliação espacial das estações monitoradas destacando parâmetros e regiões mais críticas. A avaliação temporal foi proposta para estudo do comportamento de parâmetros e previsões em curto prazo.

Através dos objetivos específicos foi possível mapear a sub-bacia do rio das Velhas, explorando ao máximo os dados coletados ao longo das campanhas de monitoramento dos anos de 1993 a 2005.

Esta tese é dividida em dez capítulos. Nos capítulos II e III estão descritos os métodos estatísticos empregados, ou seja, os métodos de estatística multivariada no capítulo II e os métodos de análise de séries temporais no capítulo III. No capítulo IV é feita a descrição e pré- tratamento da base de dados, os programas estatísticos utilizados e as metodologias de análise físico-química. No capítulo V são apresentados os resultados das análises de regressão e no capítulo VI os resultados de análise exploratória da base formada pela estações monitorados em toda sub-bacia do rio das Velhas.

No capítulo VII são apresentados os resultados de análises exploratórias da base formada pelas estações monitoradas no alto curso do rio das Velhas, no capítulo VIII os resultados da análise de séries temporais e no capítulo IX as conclusões finais.

A metodologia para tratamento dos dados aqui proposta poderá ser utilizada em outros estudos, permitindo uma diminuição do custo de programas de monitoração ambiental e um uso mais racional de pessoal e equipamentos, no laboratório e no campo.

Capítulo II – Análise multivariada

2.1 Estatística multivariada

Dados de poluição ambiental são caracterizados por sua alta variabilidade. Esta variabilidade é originada por influências geogênicas, hidrológicas, sazonais, além de possíveis fontes antrópicas. Na análise de um conjunto de dados, tais informações são perdidas usando somente métodos estatísticos univariados para interpretação dos dados.¹⁶⁻¹⁸

Na interpretação de dados de poluição ambiental em rios, a análise univariada pode ser útil na avaliação da conformidade das amostras hidrológicas com padrões legais. Neste caso, o exame e comparação devem ser realizados para cada parâmetro individualmente, com a necessidade de vários gráficos para avaliação de cada parâmetro, podendo levar a conclusões inexatas e qualitativas.

O modelo estatístico dos métodos multivariados considera a correlação entre muitas variáveis analisadas simultaneamente, permitindo a extração de uma quantidade maior de informações.¹⁵⁻²⁸

A estatística multivariada consiste em um conjunto de métodos estatísticos utilizados em situações nas quais várias variáveis são medidas simultaneamente, em cada elemento amostral. Embora a princípio o uso de métodos multivariados tenha sido aplicado a trabalhos na área das ciências humanas e biológicas, recentemente eles têm sido aplicados em trabalhos em áreas diversas e são mais evidentes na química, física, engenharias, geologia, ergonomia, etc. Esta expansão na aplicação dessa técnicas somente foi possível graças ao grande avanço da tecnologia computacional e ao grande número de softwares estatísticos com módulos de análise multivariada implementados.²⁹

A aparente facilidade esbarra em dois problemas de ordem prática que são a diversidade de métodos estatísticos que confundem o usuário na escolha de uma melhor solução para o problema e o desconhecimento por parte do usuário do sistema de trabalho que dificulta a escolha de uma inferência adequada. Quando não há conhecimento teórico prévio do sistema, as dificuldades começam logo na amostragem, no início do trabalho.

Basicamente, a estatística multivariada se divide em dois grupos: um primeiro, consistindo em técnicas exploratórias de simplificação da estrutura de variabilidade dos dados, e um segundo, consistindo em técnicas de inferência estatística. Fazem parte do primeiro grupo métodos como análise de componentes principais, análise fatorial, análise de correlações canônicas, análise de agrupamentos, análise discriminante e análise de correspondência. Esses métodos, na sua grande maioria, independem do conhecimento matemático da forma matemática da distribuição de probabilidades geradora dos dados amostrais. No segundo grupo, encontram-se os métodos de estimação de parâmetros, testes de hipóteses, análises de variância, de covariância e de regressão multivariada.²⁹

Em linhas gerais, os métodos de estatística multivariada são utilizados com o propósito de simplificar ou facilitar a interpretação do fenômeno que está sendo estudado através da construção de índices ou variáveis alternativas que sintetizem a informação original dos dados; construir grupos de elementos amostrais que apresentem similaridades entre si, possibilitando a segmentação do conjunto de dados original; investigar as relações de dependência entre as variáveis resposta associadas ao fenômeno e outros fatores (variáveis

explicativas), muitas vezes, com objetivos de predição; comparar populações ou validar suposições através de testes de hipóteses.²⁹

A análise multivariada proporciona uma avaliação mais racional de uma coleção de dados através da sua capacidade de avaliação das variáveis.¹⁷ Recentes trabalhos, também têm sido publicados com utilização da análise multivariada na interpretação dados espectrais de fluorescência³⁰, discriminação e classificação de adulterantes em mel³¹, classificação de gasolina³², aguardentes de cana e cafés.³³

2.2 Pré – processamento dos dados

2.2.1 Dados censurados e discrepantes

Um conjunto de dados antes de ser analisado deve ser revisado completamente. A presença de **dados censurados**, *missing data*, impede uma análise matemática. Os métodos clássicos da estatística descritiva aplicam-se a conjuntos de dados em que todos resultados sejam valores numéricos; a presença de resultados censurados, como é usual acontecer em estudos ambientais, torna-se não – trivial o tratamento estatístico destes conjuntos. Dados censurados não podem ser substituídos por zeros, em vez disso, esses dados devem ser preenchidos pela média das colunas ou linhas, ou na pior das hipóteses, por um número aleatório na mesma ordem de grandeza dos valores considerados nas colunas e linhas.

Dados numéricos altamente correlacionados, redundantes ou constantes, devem ser removidos com uso de técnicas estatísticas.^{32,33} Os dados são ditos correlacionados quando as variáveis estão relacionadas através de um equação matemática simples. O cálculo do **coeficiente de correlação de Pearson**, utilizado neste trabalho para avaliação das variáveis correlacionadas é descrito na seção 2.2.4. Dados são ditos redundantes ou constantes quando o mesmo valor de medida é obtido em diferentes observações. A presença desses dados costuma inviabilizar o emprego de algumas técnicas estatísticas, devido a isso, esses dados devem ser eliminados após serem identificados.^{34, 36}

Os **dados discrepantes** são resultados consideravelmente mais altos ou mais baixos em relação ao conjunto de dados. Em conjuntos de dados ambientais, os dados podem ser originados por acidentes ambientais, erros de transcrição de resultados e até problemas de calibração ou mau funcionamento de equipamentos empregados na obtenção da medida. O tratamento de dados e discrepantes também requer o uso de métodos estatísticos que não interfiram na distribuição original dos dados.^{36,37} Neste trabalho, dados discrepantes foram substituídos pela média das medidas no período de observação.

2.2.2 Padronização dos dados

Dado um conjunto de dados representado pela matriz \mathbf{X} , constituída de n de medidas de p variáveis:

$$\mathbf{X} = \begin{pmatrix} x_{11} & x_{12} & \Lambda & x_{1p} \\ x_{21} & x_{22} & \Lambda & x_{2p} \\ \mathbf{M} & & & \\ x_{n1} & x_{n2} & \Lambda & x_{np} \end{pmatrix} \quad (1)$$

Para eliminar uma constante, os dados podem ser modificados ao longo da origem das coordenadas.

Um procedimento comum é a média centrada, onde cada variável, x_{ik} , é centrada pela subtração da média das colunas:

$$x_{ik}^* = x_{ik} - \overline{x_k} \quad (2)$$

onde i é o índice das linhas, e k do das colunas, e $\overline{x_k}$ é a média das colunas calculada conforme a equação 3:

$$\bar{x}_k = \frac{1}{n} \sum_{i=1}^n x_{ik} \quad (3)$$

Freqüentemente os dados numéricos representam diferentes propriedades de amostras e de um objeto, podendo levar a diferenças na métrica das variáveis de coluna a coluna. Este comportamento pode levar a diferenças nos valores absolutos das variáveis, como também a diferença nas suas variâncias. Ambos tipos de distorção podem afetar a análise estatística por métodos multivariados. Essas diferenças podem ser eliminadas através do escalonamento dos dados para similares extensões e variâncias. Dois métodos de escalonamento importantes são, os que escalam os dados por sua amplitude:

$$x_{ik}^* = \frac{x_{ik} - x_k(\min)}{x_k(\max) - x_k(\min)} \quad 0 \leq x_{ik}^* \leq 1 \quad (4)$$

E os que escalam pelo desvio padrão:

$$x_{ik}^* = \frac{x_{ik} - \bar{x}_k}{s_k} \quad (5)$$

onde s_k :

$$s_k = \sqrt{\frac{\sum_{i=1}^n (x_{ik} - \bar{x}_k)^2}{n-1}} \quad (6)$$

onde n é o número de objetos. O escalonamento produz dados com média zero e variância unitária. O comprimento dos vetores é escalado para $\sqrt{n-1}$.

Em alguns casos, a normalização do comprimento do vetor dos dados é necessária:

$$x_{ik}^* = \frac{x_{ik}}{\|x_k\|} \quad (7)$$

onde,

$$\text{onde } \|x_k\| = \sqrt{x_{1k}^2 + x_{2k}^2 + \dots + x_{nk}^2} \quad (8)$$

A transformação dos dados em um novo sistema de coordenadas é uma outra possibilidade no pré-tratamento dos dados. Estes métodos são baseados na análise de componentes principais ou análise de fatores. O primeiro passo para esta transformação é a formação de uma matriz de dados que é derivada da matriz de dados originais e que reflete a relação entre os dados. Essas matrizes de dados derivadas são as matrizes de variância – covariância e a matriz de correlação.³⁴

2.2.3 Matriz de covariância

A **matriz de variância – covariância** ou simplesmente matriz de covariância, *Cov*, é obtida a partir da matriz de dados **X** pelo cálculo da variância, s^2 , de todas p e suas covariâncias.

$$s_k^2 = \frac{1}{n-1} \sum_{i=1}^n \left(x_{ik} - \bar{x}_k \right)^2 \text{ para } k = 1, 2, 3, \dots, p \quad (9)$$

A covariância entre os valores da j -ésima e k -ésima variáveis da matriz **X** é dada por:

$$\text{cov}(j, k) = \frac{1}{n-1} \sum_{i=1}^n \left(x_{ij} - \bar{x}_j \right) \left(x_{ik} - \bar{x}_k \right) \text{ para } j, k = 1 \dots p; j \neq k \quad (10)$$

A matriz de covariância, cujos elementos da diagonal são constituídos pelos valores de variância, s_{ik} , é dada por:

$$\mathbf{Cov} = \begin{pmatrix} s_{11}^2 & \text{cov}_{12} & \Lambda & \text{cov}_{1p} \\ \text{cov}_{21} & s_{22}^2 & & \text{cov}_{2p} \\ \mathbf{M} & \mathbf{M} & & \\ \text{cov}_{n1} & \text{cov}_{n2} & & \text{cov}_{np} \end{pmatrix} \quad (11)$$

A matriz de covariância é usada nos casos onde a métrica das variáveis é comparável. Se a métrica das variáveis é diferente, para análise simultânea dos constituintes principais e traços, as variáveis devem ser escaladas, pela equação 12.

$$x_{ik}^* = \frac{x_{ik} - \bar{x}_k}{s_k} \quad (12)$$

Se as variáveis são escaladas, a nova variância fica:

$$\begin{aligned} s_k^{2*} &= \frac{1}{n-1} \sum_{i=1}^n (x_{ik}^* - \bar{x}_k^*)^2 = \frac{1}{n-1} \sum_{i=1}^n \frac{1}{s_k^2} (x_{ik} - \bar{x}_k)^2 = \\ s_k^{2*} &= \frac{1}{n-1} \sum_{i=1}^n \frac{(n-1)s_k^2}{s_k^2} = 1 \end{aligned} \quad (13)$$

e a nova covariância será

$$\text{Cov}_{jk}^* = \frac{1}{n-1} \sum_{i=1}^n (x_{ij}^* - \bar{x}_j^*)(x_{ik}^* - \bar{x}_k^*) \quad (14)$$

e os elementos:

$$(x_{ik}^* - \bar{x}_k^*) = \frac{x_{ik}}{s_k} - \frac{1}{ns_k} \sum_{i=1}^n x_{ik} = \frac{x_{ik} - \bar{x}_k}{s_k} \quad (15)$$

$$Cov_{jk}^* = \frac{1}{n-1} \sum_{i=1}^n \left(\frac{x_{ij} - \bar{x}_j}{s_j} \right) \left(\frac{x_{ik} - \bar{x}_k}{s_k} \right) = \frac{Cov_{jk}}{s_j s_k} \quad (16)$$

e a nova matriz de covariância é igual a matriz de correlação, equação 19.

2.2.4 Matriz de correlação

Os **coeficientes de correlação**, r_{jk} , necessários para composição da **matriz de correlação** são calculados através da equação:

$$r_{jk} = \frac{Cov_{jk}}{s_j s_k} = \frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k)}{\left[\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2 \sum_{i=1}^n (x_{ik} - \bar{x}_k)^2 \right]^{1/2}} \quad j \neq k \quad (17)$$

r_{jk} conhecido como coeficiente de correlação de Pearson é empregado na identificação de variáveis correlacionadas, considerando o seguinte critério:

Se r_{jk} estiver muito próximo de 0, não há uma correlação significativa entre as variáveis, mas se r estiver próximo de -1 e +1 há uma correlação linear significativa.

O desvio padrão, s_j , é calculado pela equação:

$$s_j = \sqrt{\frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n-1}} \quad (18)$$

E a matriz de correlação, \mathbf{R} , é dada por:

$$\mathbf{R} = \begin{pmatrix} 1 & r_{12} & \Lambda & r_{1p} \\ r_{12} & 1 & & r_{2p} \\ \mathbf{M} & & & \mathbf{M} \\ r_{n1} & r_{n2} & \mathbf{K} & r_{np} \end{pmatrix} \quad (19)$$

Os cálculos das matrizes de variância e correlação são pré-requisitos para análise de componentes principais, análise fatorial, projeção de autovalores, entre outros.

2.3 Análise de agrupamentos

Na análise de agrupamentos, um conjunto de x_{ik} elementos amostrais, equação 1, constituído de $i = 1, 2, 3, \dots, n$ medidas de $k = 1, 2, 3, \dots, p$ variáveis aleatórias é agrupado em g grupos. Para cada variável k tem-se o vetor de medidas X_k definido por:

$$X_k = [X_{1k}, X_{2k}, \dots, X_{1p}]' \quad k = 1, 2, 3, \dots, p \quad (20)$$

Para que se possa proceder ao agrupamento de elementos, é necessário que se decida *a priori* a medida de similaridade ou dissimilaridade que será utilizada. Existem várias medidas diferentes e cada uma delas produz um determinado tipo de agrupamento.²⁹

Os diversos métodos para a análise de agrupamento podem ser enquadrados em quatro tipos gerais.³⁴

- **Métodos de partição:** procuram classificar as observações em grupos, definidos em função de variáveis, que sejam densamente ocupados em termos de observações daqueles com ocupação menos densa.
- **Métodos com origem arbitrária:** procuram classificar as observações segundo "g" conjuntos previamente definidos. Neste caso "g" pontos arbitrários servirão como centróides iniciais e as observações irão se agrupando, por similaridade, em torno desses centróides para formar agrupamentos.
- **Métodos por similaridade mútua:** procuram agrupar observações que tenham uma similaridade comum com outras observações. Inicialmente uma matriz $n \times n$ de similaridades entre todos os pares da observação é calculada. Em seguida, as similaridades entre colunas são repetidamente recalculadas. Colunas representando membros de um único agrupamento tenderão apresentar intercorrelações próximas a 1 e valores menores com não membros.
- **Métodos por agrupamentos hierárquicos:** para o seu desenvolvimento parte-se de uma matriz simétrica de coeficientes de associação entre itens e para a combinação dos mesmos, segundo níveis hierárquicos de similaridade, utiliza-se de um procedimento aglomerativo de tal modo que cada ciclo de agrupamento obedeça a uma ordem sucessiva no sentido do decréscimo de similaridade. Embora diversas medidas de similaridade tenham sido propostas, somente duas são geralmente usadas: coeficiente de correlação e coeficiente de variância. Se as variáveis forem padronizadas a partir da matriz inicial de dados, dando o mesmo peso a cada uma delas, qualquer um desses coeficientes poderá ser diretamente transformado no outro.

Qualquer método de agrupamento hierárquico é caracterizado pela determinação da similaridade e pelo modo que os agrupamentos presentes são ligados.

2.3.1 Função similaridade

Para variáveis contínuas, a **similaridade** é medida através das distâncias Euclidiana e Manhattan. Estas distâncias são casos especiais da distância Minkowski ou L_p - métrica, definida como:³⁴

$$d_{il} = \left[\sum_{k=1}^p (x_{ik} - x_{lk})^q \right]^{1/q} \quad (21)$$

onde, p é igual ao número de variáveis $k = 1, 2, 3, \dots; p$; e i, l são os índices das medidas $i = 1, 2, 3, \dots, l, \dots, n$.

Na maioria dos casos a distância Euclidiana é aplicada, considerando $q = 2$. A distância Euclidiana (Painel A, FIGURA 2) entre duas medidas 1 e 2 e duas variáveis é definida como:

$$d_{12} = \left[(x_{11} - x_{21})^2 + (x_{12} - x_{22})^2 \right]^{1/2} \quad (22)$$

Quando $q = 1$ a distância Manhattan (Painel B, FIGURA 2) é obtida por:

$$d_{il} = \sum_{k=1}^p |x_{ik} - x_{lk}| \quad (23)$$

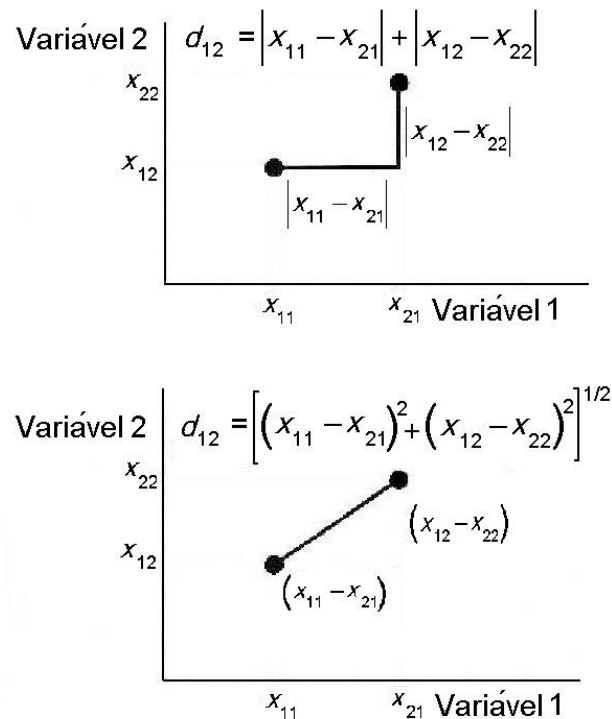


FIGURA 2 – Determinação das distâncias Euclidiana, Painei A e Manhattam, Painei B para 2 medidas e 2 variáveis. Figura adaptada da referência 34

Uma desvantagem das funções baseadas na L_p – métrica é a sua dependência com a métrica usada. Por isso, o escalonamento dos dados é inevitável se estas medidas são aplicadas.³⁴

Uma função que avalia as diferentes escalas das variáveis e ainda, suas correlações é a distância de Mahalanobis. Esta função invariante definida como:

$$D_{il}^2 = (X_i - \bar{X}) Cov^{-1} (X_l - \bar{X}) \quad (24)$$

onde Cov é a matriz de covariância e X_i e X_l os vetores coluna de vetores para cada medida i, l, \dots, n , definido como:

$$X_i = [x_{i1}, x_{i2}, \dots, x_{il}, \dots, x_{ip}] \quad i = 1, 2, 3, \dots, 1, \dots, n \quad (25)$$

O escalonamento dos dados não é necessário se a distância Mahanobis é usada. No entanto, distorções nos agrupamentos surgem se existem correlações entre as variáveis. A distância Euclidiana pode ser aplicada para variáveis altamente correlacionadas.

Complementarmente funções de distância, a função de similaridade, S_{il} também é utilizada. Essa função é baseada na distância de Minkowski, e definida como:^{30,32}

$$S_{il} = 1 - \frac{d_{il}}{d_{il}(\max)} \quad (26)$$

onde $d_{il}(\max)$ representa a distância máxima das medidas i e l , encontrados no conjunto de dados. As medidas completamente similares têm valores de $S_{il} = 1$ e para medidas completamente não similares valores de $S_{ij} = 0$ são esperados.

No presente trabalho, a distância Euclidiana é empregada como medida de similaridade.

2.3.2 Métodos de associação entre agrupamentos

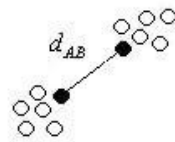
Na análise de agrupamentos por métodos hierárquicos, para decisão dos agrupamentos, diferentes métodos de associação podem ser aplicados para avaliação da similaridade entre as medidas.

Inicialmente, todas as medidas são separadas e cada uma formará um pequeno grupo. A fase inicial dos agrupamentos consiste da associação dos pequenos grupos de cada medida com base nas distâncias de Euclidiana ou de Manhattam. Os grupos maiores, formados por mais de uma medida, são unidos por critérios de associação entre as medidas que o compõem. A apresentação dos resultados de um agrupamento hierárquico é feita usualmente através de um **dendrograma**, que é um gráfico que representa a similaridade

entre as medidas como critério de agrupamento. A seguir serão discutidos alguns critérios de associação de grupos.^{34,39}

2.3.2.1 Critério de associação simples:

Neste caso, a similaridade entre dois agrupamentos é definida pelos dois elementos mais parecidos nos agrupamentos, ou seja, entre os “vizinhos” mais próximos, e utilizada na criação de um novo agrupamento g :

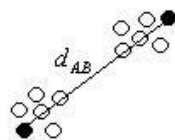


$$d_{gi} = \frac{d_{Ai} + d_{Bi}}{2} - \frac{|d_{Ai} - d_{Bi}|}{2} = \min(d_{Ai}, d_{Bi}) \quad (27)$$

Neste caso os resultados dos agrupamentos formados são negligentemente associados. Os agrupamentos são freqüentemente linearmente alongados em contraste com o usual agrupamento esférico.³⁴

2.3.2.2 Critério de associação completa

Este método baseia a agregação na maior distância entre elementos de agrupamentos opostos:



$$d_{gi} = \frac{d_{Ai} + d_{Bi}}{2} + \frac{|d_{Ai} - d_{Bi}|}{2} = \max(d_{Ai}, d_{Bi}) \quad (28)$$

Em geral bem separado, pequenos e compactos agrupamentos esféricos são formados.

2.3.2.3 Critério de associação média

Neste caso, a distância de um novo agrupamento g é calculada pela distância média entre todas as distâncias entre os elementos de dois agrupamentos A e B:

$$d_{gi} = \frac{d_{Ai} + d_{Bi}}{2} \quad (29)$$

O tamanho dos agrupamentos e seus pesos são supostamente iguais.

2.3.2.4 Critério de associação média ponderada

Diferente do critério anterior, nesse caso, o número de elementos em um agrupamento é usado para ponderar as distâncias nos agrupamentos e criação do agrupamento g :

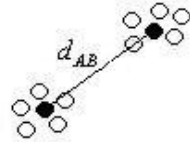
$$d_{gi} = \frac{n_A}{n} d_{Ai} + \frac{n_B}{n} d_{Bi} \quad (30)$$

com $n = n_A + n_B$

Onde n_A e n_B consistem no número de objetos dos agrupamentos A e B, respectivamente. Normalmente, nenhuma deformação do agrupamento é observada. Em alguma extensão pequenos agrupamentos consistindo de resultados discrepantes, *outliers*, podem surgir.

2.3.2.5 Critério associação centróide

Neste caso a centróide é calculada como a média de um agrupamento e aplicada como base para agregação sem distorção do espaço do agrupamento:



$$d_{gi} = \frac{n_A}{n} d_{Ai} + \frac{n_B}{n} d_{Bi} - \left(\frac{n_A - n_B}{n^2} \right) d_{AB} \quad (31)$$

2.3.2.6 Critério de associação mediana

Neste caso a centróide é calculada como a mediana de um agrupamento e aplicada como base para agregação:

$$d_{gi} = \frac{d_{Ai}}{2} + \frac{d_{Bi}}{2} - \frac{d_{AB}}{4} \quad (32)$$

Uma vantagem deste método é que a importância de um pequeno agrupamento é preservada depois da agregação com um grande agrupamento.

2.3.2.7 Método de Ward

Em 1963, Ward propôs um método de agrupamento que é fundamentado justamente nesta “mudança de variação” entre os grupos e dentro dos grupos que estão sendo formados em cada do agrupamento. Seu procedimento é também chamado de “mínima variância” e fundamenta-se nos seguintes princípios:

Inicialmente cada elemento é considerado como um único conglomerado. Em cada passo do algoritmo de agrupamento calcula-se a soma de quadrados dentro de cada aglomerado.

Esta soma é o quadrado da distância Euclidiana de cada elemento amostral pertencente ao conglomerado em relação ao correspondente vetor de médias do aglomerado, isto é:

$$SS_i = \sum_{k=1}^{n_i} (X_{ik} - \bar{X}_i)'(X_{ik} - \bar{X}_i) \quad (33)$$

onde n_i é o número de elementos no aglomerado A_i quando se está no passo s do processo de agrupamentos, X_{ik} é o vetor de observações do k –ésimo elemento amostral que pertence ao i –ésimo conglomerado, \bar{X}_i é o centróide do conglomerado A_i , e SS_i representa a soma de quadrados correspondente ao aglomerado A_i . No passo s , a soma de quadrados total dentro dos grupos, SSR , é definido como:

$$SSR = \sum_{i=1}^{g_s} SS_i \quad (34)$$

onde g_s é o número de grupos existentes quando se está no passo s .

A distância entre os aglomerados A_i e B_l é , então, definida como:

$$d(A_i, B_l) = \left[\frac{n_i n_l}{n_i + n_l} \right] (\bar{X}_i - \bar{X}_l)' (\bar{X}_i - \bar{X}_l) \quad (35)$$

que é a soma de quadrados entre os clusters A_i e B_l . Em cada passo do algoritmo de agrupamento, os dois conglomerados que minimizam a distância, equação 35, são combinados.

É possível demonstrar que a medida de distância, equação 35, é a diferença entre o valor de SSR depois e antes de se combinar os conglomerados A_i e B_l num único conglomerado.

Portanto, em cada passo do agrupamento, o método de Ward combina os dois conglomerados que resultam no menor valor SSR .²⁹

Este método de agregação conduz a um dendrograma bem estruturado e é o mais freqüentemente usado.

A escolha de um método de agrupamento depende dos dados estudados e do particular propósito da análise. Para um propósito explanatório, o mesmo grupo de dados pode ser estudado usando diferentes métodos de agrupamento sendo feita as comparações dos resultados de classificação.^{34,39}

O resultado final do agrupamento hierárquico é representado em forma de um dendrograma. No eixo x do dendrograma, os índices dos objetos agrupados são exibidos, enquanto no eixo y representa a correspondente distância ligação (ou uma adequada medida da similaridade) entre objetos ou agrupamentos, os quais são fundidos.

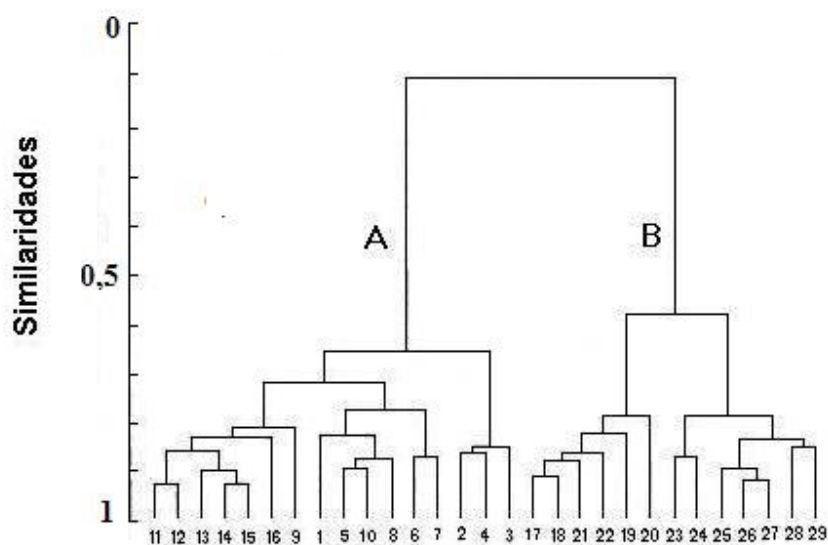


FIGURA 3 - Agrupamento de pontos de amostragem em função das similaridades utilizando o método de Ward. Figura adaptada da referência 34

2.4 Análise de componentes principais (PCA)

A técnica denominada de **análise de componentes principais**, popularmente chamada PCA, *Principal Component Analysis*, foi introduzida por Karl Person em 1901.²⁹ Seu objetivo principal é explicar a estrutura de variância e covariância de um vetor aleatório composto de p -variáveis aleatórias, através da construção de combinações lineares, linearmente independentes das variáveis originais. Essas combinações lineares são chamadas de componentes principais e não são correlacionadas entre si. Através dessa análise a informação contida nas p -variáveis originais é substituída pela informação contida em k ($k \leq p$) componentes principais não correlacionadas. A qualidade da aproximação depende do número de componentes mantidas no sistema e pode ser medida através da avaliação da proporção de variância total explicada por cada componente.^{29,34}

Quando a distribuição de probabilidades do vetor aleatório em estudo é normal, as componentes principais, são independentes e têm distribuição normal. Entretanto, a normalidade não é requisito necessário para que a técnica de análise de componentes principais possa ser utilizada. A obtenção das componentes principais envolve a decomposição da matriz de covariância do vetor aleatório de interesse. Uma transformação muito usual é a padronização das variáveis do vetor aleatório pelas respectivas médias e desvios padrões, gerando novas variáveis centradas em zero e com variâncias iguais a 1. Neste caso, as componentes principais são determinadas a partir da matriz de covariâncias das variáveis originais padronizadas, o que é equivalente a extrair as componentes principais utilizando a matriz de correlação das variáveis originais.²⁹

Uma vez determinadas as componentes principais, os seus valores numéricos, denominados de **escores**, podem ser calculados para cada elemento amostral. Deste modo, os valores de cada componentes podem ser analisados, usando-se técnicas estatísticas usuais como análise de variância e análise de regressão, dentre outras.

As análises de componentes principais juntamente com análise de agrupamentos são amplamente utilizadas em análise exploratória de dados.⁴⁰⁻⁵³

2.4.1 Componentes principais exatas extraídas da matriz de covariância

Seja $\mathbf{X} = [X_1, X_2, \dots, X_p]$ um vetor aleatório com vetor de médias $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_p]$ e matriz de covariâncias \mathbf{Cov} , dada pela equação 11.

Sejam $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p$ os autovalores da matriz \mathbf{Cov} , com os respectivos autovetores normalizados e_1, e_2, \dots, e_p , isto é, os autovetores e_k satisfazem as seguintes condições:

$$\mathbf{Cov} e_k = \lambda_k e_k, \text{ para todo } k = 1, 2, \dots, p \quad (36)$$

$$e_k e_j = \delta_{jk} \begin{cases} \delta_{jk} = 0 & \text{se } j \neq k \\ \delta_{jk} = 1 & \text{se } j = k \end{cases}$$

Considere o vetor aleatório $\mathbf{Y} = \mathbf{O}\mathbf{X}$, onde \mathbf{O} é a matriz ortogonal de dimensão $(p \times p)$ constituída dos autovetores, e_k , normalizados da matriz $\mathbf{Cov}_{n \times p}$, isto é,

$$\mathbf{O}_{p \times p} = \begin{bmatrix} e_{11} & e_{12} & \dots & e_{1p} \\ e_{21} & e_{22} & \dots & e_{2p} \\ \dots & \dots & \dots & \dots \\ e_{n1} & e_{n2} & \dots & e_{np} \end{bmatrix} = [e_1, e_2, \dots, e_k, \dots, e_p] \quad (36)$$

sendo o autovetor e_k denotado por $e_k = [e_{1k}, e_{2k}, e_{3k}, \dots, e_{nk}]'$.

O vetor \mathbf{Y} é composto de p combinações lineares das variáveis aleatórias do vetor de dados originais \mathbf{X} , tem vetor de médias igual a $\mathbf{O}\boldsymbol{\mu}$ e matriz de covariâncias $\boldsymbol{\Lambda}_{n \times p}$, que é uma matriz diagonal, cujos elementos são iguais a $a_{kk} = \lambda_k$, $k = 1, 2, \dots, p$ isto é,

$$\Lambda_{n \times p} = \begin{bmatrix} \lambda_1 & 0 & \Lambda & 0 \\ 0 & \lambda_2 & & 0 \\ \mathbf{M} & & \mathbf{O} & \mathbf{M} \\ 0 & 0 & \Lambda & \lambda_p \end{bmatrix} \quad (37)$$

Como $\lambda_1 > \lambda_2 > \dots > \lambda_p$ é possível escrever uma aproximação da matriz Λ , truncada nas K primeiras componentes principais.

Deste modo, surge à idéia de utilizar as combinações lineares em \mathbf{X} que definem \mathbf{Y} , como forma alternativa de se representar a estrutura de covariâncias do vetor \mathbf{X} tentando obter uma redução do espaço de variáveis, passando da dimensão p , para uma dimensão K menor que p .

Portanto, ao invés de se utilizar o vetor aleatório original na análise de dados, utiliza-se K combinações lineares principais. Os vetores aleatórios, \mathbf{X} e \mathbf{Y} , têm a mesma variância total e a mesma variância generalizada, sendo que o vetor \mathbf{Y} tem a vantagem de ser composto por variáveis aleatórias não correlacionadas, facilitando a interpretação conjunta dessas.²⁹

A K -ésima componente principal da matriz \mathbf{Cov} é definida como:

$$\mathbf{Y}_k = \mathbf{e}_k \mathbf{X} = [e_k X] = [e_{1k} X_1 + e_{2k} X_2 + \dots + e_{pk} X_p] \quad (38)$$

A **esperança** $E(Y_k)$ da componente \mathbf{Y}_k é definida pelo vetor de médias de \mathbf{X} :

$$E[Y_k] = \mathbf{e}_k \boldsymbol{\mu} = [e_{1k} \mu_1 + e_{2k} \mu_2 + \dots + e_{pk} \mu_p] \quad (39)$$

e a variância $\text{Var}(Y_k)$ por:

$$\text{Var}[Y_k] = \mathbf{e}_k' \mathbf{Cov} \mathbf{e}_k = \lambda_j \quad (40)$$

Lembrando que $\text{Cov}[Y_j, Y_k] = 0 \quad j \neq k$, ou seja, Λ é diagonal.

Cada autovalor λ_k representa a variância de uma componente principal Y_k . Como os autovalores estão ordenados em ordem decrescente, a primeira componente tem maior variabilidade e a p -ésima menor, como será demonstrado a seguir.

A **proporção da variância** de \mathbf{X} que é explicada pela K -ésima componente principal é definida como:

$$\frac{\text{Var}[\mathbf{Y}_k]}{\text{Variância total de } \mathbf{X}} = \frac{\lambda_k}{\text{Traço Cov}} = \frac{\lambda_k}{\sum_{j=1}^p \lambda_j} \quad (41)$$

onde:

$$\text{Traço Cov} = \sum_{k=1}^p s_{kk}^2 = \sum_{i=1}^p \lambda_k \quad \text{e} \quad s_{kk}^2 = \text{Var}[\mathbf{X}_k], \quad k=1,2,\dots,p$$

e os autovalores, λ_k , da matriz de covariância Λ representam a variância da k -ésima componente principal.

Em geral, a razão apresentada na equação 41 é multiplicada por 100, indicando o resultado em porcentagem.

A **proporção da variância total** que é explicada pela K primeiras componentes principais é definida como:

$$\frac{\sum_{j=1}^K \text{Var}[\mathbf{Y}_k]}{\text{Variância total de } \mathbf{X}} = \frac{\sum_{j=1}^K \lambda_j}{\text{Traço Cov}} = \frac{\sum_{j=1}^K \lambda_j}{\sum_{j=1}^p \lambda_j} \quad (42)$$

Se as K primeiras componentes principais explicam uma grande parte da variância total do vetor X , pode-se utilizar na análise dos dados apenas o vetor aleatório $[Y_1, Y_2, \dots, Y_k]$. Assim, matriz de covariância é aproximada pelas K primeiras componentes principais pela relação:

$$\mathbf{Cov} \approx \sum_{j=1}^K \lambda_j e_j e_j' \quad (43)$$

Considerando que $\sum_{j=1}^K e_j e_j' = \mathbf{I}_k$.

Cada parcela da equação 42 envolve uma matriz de dimensão correspondente apenas à informação de K -ésima componente principal $j = 1, 2, \dots, K$. Assim sendo, a variabilidade original do vetor X estará sendo aproximada pela soma de K matrizes, cada uma representando a variabilidade relacionada a uma componente, j . Quando $K = p$, tem-se que a matriz de covariâncias, \mathbf{Cov} , é reproduzida com exatidão pela soma de matrizes relacionadas às componentes principais, isto é,

$$\mathbf{Cov} \approx \sum_{j=1}^p \lambda_j e_j e_j' \quad (44)$$

Devido à própria forma de construção, a primeira componente principal é sempre mais representativa em termos da variância total e a p -ésima componente é sempre a de menor representatividade.

A FIGURA 4 mostra uma ilustração gráfica das componentes principais no caso de duas variáveis. Como pode ser observado no gráfico, cada ponto no sistema de coordenadas X_1 e X_2 é projetado ortogonalmente no novo sistema de coordenadas Y_1 e Y_2 .

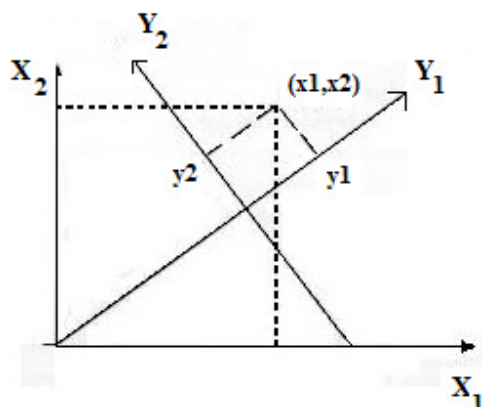


FIGURA 4 – Interpretação geométrica da análise de componentes principais.
Figura adaptada da referência 29

Alguns critérios são sugeridos na escolha das componentes mais importantes, como o método de Cattell e o método de Kaiser. O método de Cattell propõe uma separação visual das componentes principais através do gráfico de autovalores, λ_k , em função das componentes Y_k , como ilustrado na FIGURA 5.³⁴

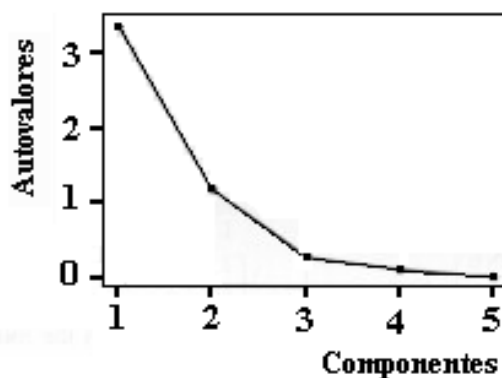


FIGURA 5 – Avaliação das componentes mais importantes a partir dos autovalores.
Método de Cattell. Figura adaptada da referência 29

Nessa ilustração é possível constatar que a partir da 4ª componente os autovalores tendem a um valor constante e próximo de zero. Assim, para este sistema de dados, as três primeiras componentes explicam quase totalidade da variabilidade dos dados.

O método de Kaiser exclui componentes com autovalores abaixo da média.³⁴

Os dois métodos foram empregados análise de componentes principais no conjunto de dados do presente trabalho.

2.4.2 Componentes principais exatas extraídas da matriz de correlação

A técnica de componentes principais aplicada a uma matriz de covariâncias de variáveis padronizadas por suas médias e desvio padrão é equivalente à obtenção de componentes principais da matriz de correlação \mathbf{R} das variáveis X_i originais, $i = 1, 2, \dots, p$, definida na equação 19.

Seja,

$$Z_j = \frac{(X_k - \mu_k)}{s_k} \quad (45)$$

O novo vetor das variáveis padronizadas em termos das médias, μ_k , e desvio padrão s_k .

Então, a k -ésima componente principal da matriz \mathbf{R} , $k = 1, 2, \dots, p$ é definida por:

$$\mathbf{Y}_k = \mathbf{e}_k \mathbf{Z} = [e_k Z] = [e_{1k} Z_1 + e_{2k} Z_2 + \dots + e_{pk} Z_p] \quad (46)$$

Uma vez que:

A variância de Y_k é igual a λ_k , $j = 1, 2, \dots, p$, e a covariância entre Y_j e Y_k é igual a zero, para qualquer $j \neq k$. A correlação entre a componente Y_k e a variável padronizada Z_j é igual a:

$$r_{Y_k, Z_j} = r_{Y_k, X_k} = e_{kj} \sqrt{\lambda_k} \quad (47)$$

e, conseqüentemente, as variáveis Z_j com os maiores coeficientes na componente principal Y_k são as mais correlacionadas com a componente.

A variância total do vetor aleatório $Z = [Z_1, Z_2, \dots, Z_p]$ é igual ao traço da matriz \mathbf{R} , que é igual ao valor p , ou seja, o número de variáveis medidas em cada elemento amostral. Portanto, a proporção da variância total explicada pela k -ésima componente principal, de acordo com a equação 41, é igual a $\frac{\lambda_k}{p}$, $k = 1, 2, \dots, p$.

Os coeficientes das componentes principais obtidos através da decomposição da matriz de correlação, \mathbf{R} , não são numericamente iguais aos coeficientes obtidos através da decomposição da matriz de covariâncias, \mathbf{Cov} . Em geral, em termos de explicação a variância total, determinada quando a matriz \mathbf{R} é utilizada, necessita de um número maior de componentes para explicar a mesma quantidade de variância total quando comparada a matriz \mathbf{Cov} .²⁹

2.4.3 Interpretação dos resultados na análise de componentes principais

A interpretação dos resultados na análise de componentes principais é conduzida pela análise dos gráficos de escores e cargas (*loadings*), como indicado na FIGURA 6.

Primeiramente o agrupamento entre os objetos deve ser reconhecido e depois a importância entre os parâmetros avaliados na discriminação entre os objetos deve ser discutida.

O gráfico de escores descreve a projeção a projeção linear das medidas representando a maior variância total dos dados. A correlação e a importância dos parâmetros é decidida pelo gráfico de cargas, que representa a projeção dos parâmetros nas componentes principais. A interpretação conjunta dos escores e cargas é possível se as cargas são escalonadas e superpostas ao gráfico de escores em um gráfico chamado *biplot*, FIGURA 6.³⁵

A correlação entre os parâmetros é descrita pelo cosseno do ângulo entre os vetores dos *loadings*. Quanto menor o ângulo maior é a correlação entre os parâmetros. Parâmetros não correlacionados são ortogonais entre si. A projeção do vetor de carga de um determinado parâmetro a componente principal considerada é uma medida da importância de um parâmetro no modelo de componentes principais. A proximidade dos objetos ao vetor dos *loadings* reflete a importância daquele parâmetro na construção do modelo.

Na FIGURA 6, há maiores correlações entre os metais (Cu e Mn) e entre os halogênios (Br, Cl, e I). Os elementos (Cu e I) não são correlacionados. Todos os parâmetros são importantes para descrever a primeira componente principal. Já a segunda componente é principalmente caracterizada pelos parâmetros I e Br. Os parâmetros Cu, Mn, I e Cl separam os agrupamentos (2,8,5) e (3,6,7) enquanto Br separa o agrupamento (2,5,8) do restante dos pontos.

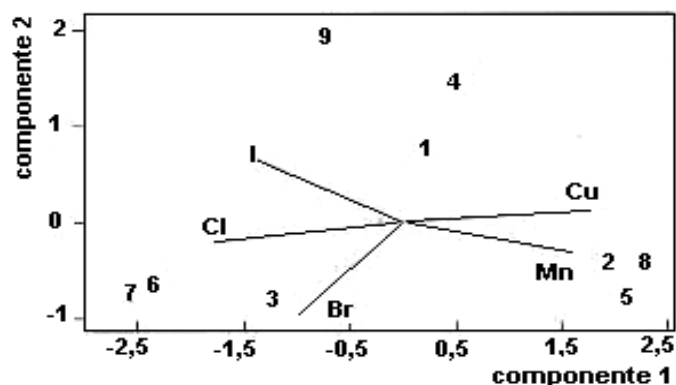


FIGURA 6- Biplot para simultânea caracterização de escores e loadings em dados referentes a análise química de cabelo³⁰

2.5 Análise fatorial

Assim como a análise de componentes principais, a **análise fatorial** tem como objetivo principal descrever a variabilidade original do vetor aleatório X , em termos de um número menor de m variáveis aleatórias, chamadas de **fatores comuns** e que estão linearmente relacionados com o vetor original X . Nesse modelo, parte da variabilidade de X é atribuída

aos fatores comuns, sendo o restante da variabilidade de X atribuído às variáveis que não foram incluídas no modelo, ou seja, ao erro aleatório.²⁹

Em linhas gerais, o que se espera é que as variáveis originais $X_k = k = 1, 2, \dots, p$ estejam agrupadas em sub-conjuntos de novas variáveis mutuamente não correlacionadas, sendo que a análise fatorial teria como objetivo a determinação desses fatores de agrupamento. Desse modo, em casos nos quais se tem um número grande de variáveis correlacionadas entre si, seria possível, a partir da análise fatorial, identificar um número menor de novas variáveis alternativas, não correlacionadas que de algum modo resumissem as informações principais das variáveis originais.

Seja, $X = [X_1, X_2, \dots, X_p]$ um vetor aleatório com vetor de médias μ , onde $\mu = (\mu_1, \mu_2, \dots, \mu_p)$, matriz de covariâncias \mathbf{Cov} e matriz de correlação \mathbf{R} . Sejam $Z_k = \frac{(X_k - \mu_k)}{s_k}$ as variáveis originais padronizadas, onde μ_k e s_k representam respectivamente a média e o desvio padrão da variável X_k , $k = 1, 2, \dots, p$. Neste caso, a matriz \mathbf{R} é a matriz de covariâncias do novo vetor aleatório $Z = [Z_1, Z_2, \dots, Z_p]$, como mostrado na seção 2.4.2.

O modelo de análise fatorial construído a partir da matriz de correlação teórica \mathbf{R} é um modelo que relaciona linearmente as variáveis padronizadas e os m fatores comuns que, a princípio, são desconhecidos.

Em notação matricial, o modelo pode ser expresso por:

$$\mathbf{Z} = \mathbf{D}[\mathbf{X} - \boldsymbol{\mu}] = \mathbf{L}\mathbf{F} + \boldsymbol{\varepsilon} \quad (48)$$

Onde,

$$\mathbf{D} = \begin{bmatrix} 1/s_1 & 0 & 0\Lambda & 0 \\ 0 & 1/s_2 & 0\Lambda & 0 \\ \mathbf{M} & \mathbf{M} & \mathbf{M} & \mathbf{M} \\ 0 & 0 & \Lambda & 1/s_p \end{bmatrix}$$

$$\mathbf{D}[\mathbf{X} - \boldsymbol{\mu}] = \begin{bmatrix} l_{11} & l_{12} & \Lambda & l_{1m} \\ l_{21} & l_{22} & \Lambda & l_{2m} \\ \mathbf{M} & \mathbf{M} & & \mathbf{M} \\ l_{p1} & l_{p2} & \Lambda & l_{pm} \end{bmatrix} \times \begin{bmatrix} F_1 \\ F_2 \\ \mathbf{M} \\ F_m \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \mathbf{M} \\ \varepsilon_m \end{bmatrix}$$

Onde cada elemento da equação matricial é expresso por:

$$\begin{aligned} Z_1 &= l_{11}F_1 + l_{12}F_2 + \Lambda + l_{1m}F_m + \varepsilon_1 \\ Z_2 &= l_{21}F_1 + l_{22}F_2 + \Lambda + l_{2m}F_m + \varepsilon_2 \\ &\quad \mathbf{M} \quad \mathbf{M} \quad \mathbf{M} \quad \mathbf{M} \\ Z_p &= l_{p1}F_1 + l_{p2}F_2 + \Lambda + l_{pm}F_m + \varepsilon_p \end{aligned} \quad (49)$$

Nesse modelo, F é um vetor aleatório contendo m fatores, também chamados de **variáveis latentes**, que descrevem os elementos da população em estudo e não são observáveis, $1 \leq m \leq p$, ou seja, não podem ser medidos *a priori*. Portanto, o modelo de análise fatorial assume que as variáveis Z_k estão relacionadas linearmente com novas variáveis aleatórias F_k , $k = 1, 2, \dots, m$, que precisarão ser identificadas de algum modo. O vetor ε é um vetor de erros aleatórios e corresponde aos erros de medida e à variação de Z_k , que não é explicada pelos fatores comuns F_i , $i = 1, 2, \dots, m$, incluídos no modelo.

O coeficiente l_{ik} , comumente chamado de carga, é o coeficiente da k -ésima variável padronizada Z_k no i -ésimo fator F_k e representa o grau de dependência linear entre Z_k e F_k , $k = 1, 2, \dots, m$. Os coeficientes l_{ik} ficam armazenados na matriz \mathbf{L} , que é uma matriz de parâmetros a serem determinados.

Assim, de acordo com o modelo apresentado na equação 49, a informação das p – variáveis originais padronizadas (Z_1, Z_2, \dots, Z_p) estará sendo representada por $(p+m)$ variáveis aleatórias não observáveis, isto é $(\epsilon_1, \epsilon_2, \dots, \epsilon_p)$ e (F_1, F_2, \dots, F_m) . O objetivo é identificar as novas m – variáveis, interpretá-las e calcular seus escores assim como foi feito na análise de componentes principais.

Algumas suposições são necessárias para que se possa operacionalizar a determinação dos fatores comuns:

- A esperança dos fatores é igual a zero, $E(F) = 0$, o que implica que $E[F_k] = 0$, $k = 1, 2, \dots, m$, ou seja, os fatores têm média igual a zero.
- A variância dos fatores é dada por:

$$\text{Var}(F) = \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & \Lambda & 0 \\ 0 & 1 & 0 & \Lambda & 0 \\ M & M & M & M & M \\ 0 & 0 & \Lambda & 1 & 0 \end{bmatrix}$$

Ou seja, todos os fatores F_k não são correlacionados e têm variâncias iguais a 1;

- A esperança dos erros é também nula, $E(\epsilon) = 0$, o que implica que $E(\epsilon_k) = 0$, $k = 1, 2, \dots, m$, ou seja, todos os erros têm médias iguais a zero;
- A variância dos erros é dada por:

$$\text{Var}(\epsilon) = \mathbf{\Psi} = \begin{bmatrix} \psi_1 & 0 & 0 & \Lambda & 0 \\ 0 & \psi_2 & 0 & \Lambda & 0 \\ M & M & M & M & M \\ 0 & 0 & \Lambda & \psi_p & 0 \end{bmatrix}$$

Ou seja,

$\text{Var}[\varepsilon_k] = \Psi_k$ e $\text{Cov}(\varepsilon_i, \varepsilon_k) = 0, \forall i \neq j$, o que significa dizer que os erros não são correlacionados entre si e não necessariamente têm a mesma variância;

- Os vetores ε e F são independentes. Portanto, $\text{Cov}(\varepsilon, F) = E(\varepsilon, F') = 0$

A suposição acima implica dizer que os vetores ε e F representam duas fontes de variação distintas, relacionadas às variáveis padronizadas Z_k , não havendo qualquer dependência entre estas fontes de informação. Um modelo fatorial que atende as suposições relacionadas acima é chamado de **ortogonal**, onde a ortogonalidade refere-se ao fato dos m fatores serem independentes entre si.

Uma consequência imediata das suposições acima está relacionada com a estrutura da matriz de correlação teórica \mathbf{R} . Quando o modelo ortogonal é assumido, a matriz \mathbf{R} pode ser reparametrizada na forma:

$$\mathbf{R} = \mathbf{L}\mathbf{L}' + \boldsymbol{\Psi} \quad (50)$$

O objetivo da análise fatorial é encontrar as matrizes \mathbf{L} e $\boldsymbol{\Psi}$ que possam representar a matriz \mathbf{R} para um dado valor de m , menor que o número de variáveis originais p .

As implicações da decomposição da equação acima são representadas como:

$$\text{Var}(Z_k) = l_{1k}^2 + l_{2k}^2 + \Lambda l_{mk}^2 + \psi_k = h_k^2 + \psi_k \quad (51)$$

onde,

$$h_k^2 = l_{1k}^2 + l_{2k}^2 + \Lambda l_{mk}^2, \quad j = 1, 2, \Lambda, p \quad (52)$$

o que significa que a variância de Z_k é decomposta em duas partes. A primeira, denotada por h_k^2 , é a variância de Z_k explicada pelos m fatores incluídos no modelo fatorial. Esta parte da variabilidade é chamada de **comunalidade**, nome dado em ao fato dos fatores F_k , $k = 1, 2, \dots, m$ aparecerem em todas as equações 49 e as variáveis de Z_k terem fonte de variação comum. A parte denotada por ψ_k , é a parte da variabilidade de Z_k associada apenas ao erro aleatório ϵ_k , específico de cada variável Z_k . Esta parte da variabilidade é chamada de **unicidade** ou **variância específica**.

A proporção da variância total explicada pelo fator F_i é dada por:

$$\text{Proporção da variância total explicada}_{F_k} = \frac{\sum_{k=1}^p 1_{ik}^2}{p} \quad (53)$$

Os fatores mais representativos no modelo são aqueles com maiores valores de proporção de variância. A determinação dos m fatores é feita através dos autovalores da matriz de correlação amostral \mathbf{R}^* . A matriz de correlação amostral $R_{p \times p}$ é estimada da matriz de correlação teórica R .

Para escolha dos autovalores mais importantes são utilizados os seguintes critérios:

- Permanecem aqueles autovalores, λ_j , que representam maiores proporções da variância total, e, portanto, o valor de m será igual ao número de autovalores retidos. A proporção da variância é relacionada com cada autovalor pela seguinte equação:

$$\text{Proporção da variância explicada} = \frac{\lambda_j}{p}, \quad j = 1, 2, \dots, p \quad (54)$$

- O uso do método de Kaiser, que utiliza a comparação do valor numérico de cada autovalor, λ_j , com o valor médio. O valor proposto de m será igual ao número de fatores com autovalores maiores ou iguais a média.

- O uso do método de Cattell, que utiliza a observação do gráfico onde os autovalores são plotados em função do número de fatores. Por este critério, procura –se no gráfico um “ponto de salto”, que estaria representando um decréscimo de importância em relação à variância total. O valor de m seria, então, igual ao número de autovalores anteriores ao “ponto de salto”.

Finalmente, uma escolha adequada do valor de m deve, levar em consideração a interpretabilidade dos fatores e o princípio da parcimônia, ou seja, a descrição da estrutura de variabilidade do vetor aleatório Z com um número pequeno de fatores. Uma vez escolhido o valor de m , é possível estimar as matrizes \mathbf{L} e Ψ .

O método de componentes principais, comumente utilizado como análise exploratória de dados, não exige informações ou suposições sobre a distribuição de probabilidades do vetor aleatório Z . Os métodos de fatores principais e da máxima verossimilhança também podem ser utilizados, mas são indicados apenas quando o vetor aleatório Z tem distribuição normal multivariada.

No método das componentes principais para cada autovalor λ_j , $j = 1, 2, \dots, m$ encontrado na estimação do valor m , encontra-se o autovetor normalizado correspondente e_j , onde $e_j = [e_{1j}, e_{2j}, e_{3j}, \dots, e_{pj}]$. As matrizes \mathbf{L} e Ψ são determinadas respectivamente por:

$$L = [\sqrt{\lambda_1} e_1 \sqrt{\lambda_2} e_2 \dots \sqrt{\lambda_m} e_m] \quad (55)$$

$$\Psi = \text{diag}(\mathbf{R} - \mathbf{L}\mathbf{L}') \quad (56)$$

A matriz $\Psi_{p \times p}$ tem a diagonal principal igual aos elementos da diagonal principal da matriz $(\mathbf{R} - \mathbf{L}\mathbf{L}')$.

A idéia básica deste procedimento reside na aplicação do **teorema de decomposição espectral** à matriz **R**. Por este teorema, a matriz de correlação amostral pode ser decomposta como uma soma de p matrizes cada uma relacionada com um autovalor da matriz **R**. Para um valor m fixo tem-se que:

$$\mathbf{R} = \sum_{j=1}^p \lambda_j \mathbf{e}_j \mathbf{e}_j' = \sum_{i=1}^m \lambda_i \mathbf{e}_i \mathbf{e}_i' + \sum_{i=m+1}^p \lambda_i \mathbf{e}_i \mathbf{e}_i' \quad (57)$$

Assim, uma aproximação para a matriz **LL'** será dada por:

$$\mathbf{LL}' = \left[\sqrt{\lambda_1} \mathbf{e}_1 \sqrt{\lambda_2} \mathbf{e}_2 \Lambda \sqrt{\lambda_m} \mathbf{e}_m \right] \left[\sqrt{\lambda_1} \mathbf{e}_1 \sqrt{\lambda_2} \mathbf{e}_2 \Lambda \sqrt{\lambda_m} \mathbf{e}_m \right]' = \sum_{j=1}^m \lambda_j \mathbf{e}_j \mathbf{e}_j' \quad (58)$$

Para construir a matriz ψ , pode-se considerar a utilização da seguintes matriz:

$$\mathbf{R} - \mathbf{LL}' = \sum_{j=m+1}^p \lambda_j \mathbf{e}_j \mathbf{e}_j' \quad (59)$$

Como a matriz acima não é diagonal, uma estimação de ψ é feita considerando apenas a sua diagonal. De acordo com a equação 59.

Assim, a matriz de correlação amostral original **R** está sendo aproximada por:

$$\mathbf{R} \approx \mathbf{LL}' + \psi \quad (60)$$

e a matriz residual proveniente do ajuste do modelo fatorial é dada por:

$$\text{Matriz residual} = \mathbf{R} - (\mathbf{LL}' + \psi) \quad (61)$$

A matriz residual pode servir como um critério para avaliação da qualidade de ajuste do modelo fatorial. Idealmente, seus valores deveriam ser próximos de zero. Entretanto, essa

matriz somente é nula quando o valor de m é igual a p , o que na prática não é a solução desejada.²⁵

A análise de fatores é um tratamento análogo à análise de componentes principais. A diferença entre os dois métodos reside principalmente na introdução da matriz de erros, equação 61. Como o ajuste do modelo fatorial é orientado pela determinação de uma matriz residual com valores mais próximos de zero, os fatores discriminam melhor o conjunto de variáveis aleatórias da base de dados original, o que nem sempre acontece na análise de componentes principais.

Diversos exemplos são encontrados na literatura na aplicação da análise fatorial na avaliação de dados ambientais; como na avaliação da qualidade de amostras de águas subterrâneas^{25,58} e de rios^{27,57}, no estudo arqueométrico de cerâmicas⁵⁹ e no estudo da biodisponibilidade de metais em sedimentos⁶⁰.

2.6 Análise de regressão não linear

Avaliar a possível relação entre uma variável dependente e uma ou mais variáveis independentes é uma das tarefas mais comuns em estatística. Pode-se atingir este objetivo por meio dos bem conhecidos modelos de regressão, os quais se dividem em duas classes distintas: **os lineares** e **os não-lineares**.^{61, 94-98}

Dentre as muitas diferenças existentes entre essas duas classes de modelos, a principal está relacionada as suas formulações. No caso linear, a partir de um conjunto de observações, busca-se o modelo que melhor explique a relação, se existir alguma, entre as variáveis inerentes a um dado fenômeno. Por exemplo, se a resposta de interesse, usualmente representada por y , depender de uma única variável independente, x , a partir da representação gráfica de y versus x , pode-se sugerir possíveis modelos. Na presença de várias variáveis independentes, uma alternativa para o ajuste de um possível modelo é partir, inicialmente, de um modelo completo e avaliar a qualidade do ajuste por meio de diagnósticos de regressão.^{62,63} A classe de modelos lineares é bastante flexível, uma vez que muitos modelos podem ser formulados.^{64,65} Já no caso não-linear, na maioria das vezes, as formulações de possíveis modelos são baseadas em condições teóricas inerentes ao fenômeno que se tem interesse modelar. Modelos formulados desta forma são chamados de **modelos mecanísticos**.^{66,67}

2.6.1 Formulação dos modelos de regressão não-lineares

Por definição, um modelo de regressão é não linear se pelo menos um de seus parâmetros aparecem na forma não linear.

Um modelo de regressão não-linear é considerado **intrinsecamente linear** se este pode ser reduzido a um modelo linear por meio de uma reparametrização apropriada. Em geral, usar-se o termo intrinsecamente linear para referir-se a modelos que podem ser linearizados via alguma transformação.

Em geral, na prática, um modelo não-linear é linearizado para facilitar a obtenção das estimativas dos parâmetros. O inconveniente do uso de uma transformação é que, além do parâmetro perder sua interpretação intrínseca, pode-se alterar a estrutura e distribuição dos erros, ou seja, se os erros do modelo original satisfizerem as suposições usuais de normalidade, independência e homogeneidade da variância, os erros do novo modelo, em geral, não satisfarão tais suposições.⁶⁸ Caso não seja possível obter uma reparametrização ou uma transformação apropriada que reduza o modelo a forma linear, tem-se os chamados modelos **intrinsecamente não-lineares**.

2.6.2 O método de estimação por mínimos quadrados

O método de estimação por mínimos quadrados é usado na análise de dados em que as observações são constituídas por variáveis resposta Y_i obtidas em diferentes níveis da variável independente X_i , $i=1, 2, 3, \dots$. Assume-se que a relação entre a variável resposta e a variável independente pode ser adequadamente representada por uma equação da forma:

$$Y = f(X, \theta) + \varepsilon \quad (62)$$

na qual $\mathbf{Y} = [Y_1, \dots, Y_n]'$ e $\mathbf{X} = [X_1, \dots, X_n]'$ são os vetores das variáveis dependentes e independentes respectivamente, $\theta = [\theta_1, \dots, \theta_n]'$ é o vetor de parâmetros desconhecidos, $f(X, \theta) = [f(X_1, \theta), \dots, f(X_n, \theta)]'$ é uma função das variáveis independentes e dos parâmetros chamada de **função esperança** ou **função de regressão** e $\varepsilon = [\varepsilon_1, \dots, \varepsilon_n]'$ é o vetor de erros aleatórios. Em geral, por motivos de inferência, assume-se que os erros são variáveis aleatórias independentes e identicamente distribuídas, normais com média zero e variâncias constantes $s^2 \mathbf{I}_n$, em que \mathbf{I}_n representa a matriz identidade de ordem n .

Considerando, função de regressão contínua e admitindo derivadas de primeira e segunda ordem com relação aos parâmetros, define-se a soma de quadrados dos erros da equação 62 por:

$$S(\theta) = \sum_{i=1}^n [y_i - f(X_i; \theta)]^2 \quad (63)$$

Pode-se notar que após a realização de um dado experimento (X_i, Y_i) , são observações fixas e conhecidas, portanto $S(\theta)$ depende exclusivamente de θ . Denota-se por $\hat{\theta}$ os estimadores de mínimos quadrados de θ , ou seja, os valores de $\hat{\theta}$ que minimizam $S(\theta)$. Para determinação dos estimadores de mínimos quadrados $\hat{\theta}$ deve-se derivar a equação 63 com relação a cada θ_j ($j = 1, \dots, p$). Procedendo desta forma, determinam-se p equações, que são especificadas por:

$$\sum_{i=1}^n [y_i - f(X_i; \theta)] \left[\frac{\partial f(X_i; \theta)}{\partial \theta_j} \right]_{\theta=\hat{\theta}} = 0 \quad (64)$$

Quando $\partial f(X_i, \theta) / \partial \theta_j$ não depende de θ , ou $\partial^2 f(X_i, \theta) / \partial \theta_j^2 = 0$, tem-se as **equações normais** do modelo de regressão linear. Para funções de regressão não-lineares, a dependência em θ estará presente em pelo menos uma das derivadas parciais de $f(X_i; \theta)$. Em modelos multiparamétricos, as soluções das equações normais podem ser extremamente difíceis de serem obtidas e algum método iterativo de resolução de equações normais não-lineares deve ser utilizado na maioria dos casos.^{67,69}

2.6.3 A obtenção das estimativas de mínimos quadrados

Vários métodos iterativos são propostos na literatura para obtenção das estimativas de mínimos quadrados dos parâmetros de um modelo de regressão não-linear. Os mais utilizados são o método de Gauss-Newton ou método da linearização, o método *steepest-descent* ou método do gradiente e o método de Marquardt⁶⁷. Esses métodos fazem uso das derivadas parciais da função de regressão, $f(X_i; \theta)$, com relação a cada parâmetro. Essa característica pode restringir suas aplicações, uma vez que, em geral, a função de regressão é bastante complexa.

Uma alternativa é calcular as derivadas numericamente, como, por exemplo, através de diferenças finitas ou usar o método chamado D.U.D. *Doesn't Use Derivatives*⁷¹. O método D.U.D. é bastante similar ao método de Gauss-Newton, exceto pelo fato de não exigir a especificação das derivadas parciais das funções de regressão. A grande maioria dos programas estatísticos possui rotinas com a implementação destes métodos, sendo que os mesmos produzem estimativas bastante similares e, em geral, são de rápida convergência⁶⁷. A rapidez na convergência depende da complexidade do modelo em estudo e, principalmente, da qualidade dos valores iniciais, necessários em qualquer método iterativo.⁶¹

Considerando como verdadeiras as suposições de que os erros na equação 62, são variáveis aleatórias independentes e identicamente distribuídas, normais com média zero e variância constante σ^2 , pode-se fazer certas considerações a cerca dos estimadores de mínimos quadrados. A validade dessas suposições pode ser avaliada através dos chamados diagnósticos de regressão.

Em modelos de regressão lineares, os estimadores de mínimos quadrados são não viciados, normalmente distribuídos, e ainda possuem variância mínima possível entre qualquer outra classe de estimadores. Essas propriedades são aceitas como as melhores propriedades que uma classe pode apresentar⁷². Podem existir outros estimadores, também não viciados, mas esses são menos eficientes devido às variâncias excederem a variância dos estimadores de mínimos quadrados. Já no caso não-linear, essas propriedades somente são válidas assintoticamente, isto é, quando o tamanho da amostra é suficientemente grande. Em geral, em pequenas amostras, essas propriedades são desconhecidas⁷³.

O tamanho da amostra necessário para os resultados assintóticos terem validade depende fundamentalmente do modelo em estudo⁷³. Podem existir modelos que, mesmo com amostras consideradas muito grandes, os resultados assintóticos não são válidos. Entretanto à medida que o tamanho da amostra aumenta, os resultados assintóticos vão se tornando mais aplicáveis. Quando os estimadores de mínimos quadrados são pouco viciados,

distribuição próxima da normal e variâncias próximas daquelas dadas pela matriz de variâncias-covariâncias assintóticas⁶⁶, pode-se afirmar que os estimadores de mínimos quadrados exibem um comportamento próximo do comportamento linear. Quanto mais próximo do linear for o comportamento do modelo, mais precisos serão os resultados assintóticos e, conseqüentemente, mais confiáveis serão as inferências. Avalia-se a extensão do comportamento não-linear através das medidas de não-linearidade.⁶¹

2.6.4 Aproximação linear

Para facilitar a compreensão dos resultados apresentados a seguir, considera-se a seguinte notação:

$$f(X, \theta) = [f(X_1, \theta), \dots, f(X_n, \theta)]'$$

$$F(\theta) = \frac{\partial f(\theta)}{\partial \theta'} = \left[\left[\frac{\partial f_i(X_i; \theta)}{\partial \theta_j} \right] \right]$$

$$F = F(\theta^*).$$

Como comentado anteriormente, os resultados em regressão não-linear somente são válidos assintoticamente. É feita, uma aproximação linear, de primeira ordem, em série de Taylor da função esperança em torno de uma vizinhança de θ^* , onde θ^* é considerado como sendo o verdadeiro valor do parâmetro θ . Essa aproximação linear é dada por:

$$f(X_i; \theta) \approx f(X_i; \theta^*) + \sum_{j=1}^p \left[\frac{\partial f(X_i; \theta)}{\partial \theta_j} \right]_{\theta=\theta^*} \times (\theta_j - \theta_j^*) \quad (65)$$

ou, em notação matricial:

$$F(\theta) \approx F(\theta^*) + F(\theta - \theta^*) \quad (66)$$

onde,

$$f(\theta) = [f(X_1, \theta), \dots, f(X_n, \theta)]'$$

$$F(\theta - \theta^*) = \frac{\partial f(\theta - \theta^*)}{\partial \theta'} = \left[\left(\frac{\partial f_i(X_i; \theta)}{\partial \theta_j} \right) \right] \times (\theta - \theta^*)$$

Usando a equação 65, pode-se escrever a soma de quadrados dos erros, definida na equação 62, na forma:

$$\begin{aligned} S(\theta) &= [Y - f(\theta)]' [Y - f(\theta)] = \|Y - f(\theta)\|^2 \\ S(\theta) &\approx \|Y - f(\theta^*) - F(\theta - \theta^*)\|^2 = \|z - F(\beta)\|^2 \end{aligned} \quad (67)$$

onde $z = Y - f(\theta^*)$ e $\beta = (\theta - \theta^*)$.

Utilizando propriedades de modelos lineares, a equação 67, será mínima quando β for dada por:

$$\hat{\beta} = (F'F)^{-1} F'z \quad (68)$$

Quando o tamanho da amostra n for suficientemente grande e sob certas condições de regularidade⁶⁶, $\hat{\theta}$ estará quase certamente em uma vizinhança em torno de θ^* . Assim, $\hat{\theta} - \theta^* \approx \beta$ e $\theta - \theta^* \approx \hat{\beta} = (F'F)^{-1} F'\varepsilon$. Além disso, na equação 67, quando $\theta = \hat{\theta}$, tem-se:

$$f(\hat{\theta}) - f(\theta^*) = F(\hat{\theta} - \theta^*) \approx (F'F)^{-1} F'\varepsilon \approx \mathbf{P}_F \varepsilon \quad (69)$$

e,

$$Y - f(\hat{\theta}) \approx Y - f(\theta^*) - F(\hat{\theta} - \theta^*) \approx \varepsilon - \mathbf{P}_F \varepsilon = (\mathbf{I}_n - \mathbf{P}_F) \varepsilon \quad (70)$$

onde $\mathbf{P}_F = F(F'F)^{-1} F'$ e $(\mathbf{I}_n - \mathbf{P}_F)$ é uma matriz simétrica e independente⁷⁵. Conseqüentemente, usando as equações 69 e 70 tem-se:

$$(n-p)s^2 = S(\theta) = \|Y - f(\hat{\theta})\|^2 \approx \|(\mathbf{I}_n - \mathbf{P}_F)\varepsilon\|^2 = \varepsilon'(\mathbf{I}_n - \mathbf{P}_F)\varepsilon \quad (71)$$

e

$$\|f(\hat{\theta}) - f(\theta^*)\| \approx \|F(\hat{\theta} - \theta^*)\| = (\hat{\theta} - \theta^*)' F' F (\hat{\theta} - \theta^*) \approx \|\mathbf{P}_F \varepsilon\|^2 = \varepsilon' \mathbf{P}_F \varepsilon \quad (72)$$

Portanto, usando 71 e 72 obtém-se:

$$S(\theta^*) - S(\hat{\theta}) \approx \theta' \theta - \theta' (\mathbf{I}_n - \mathbf{P}_F) \theta = \varepsilon' \mathbf{P}_F \varepsilon$$

$$S(\theta^*) - S(\hat{\theta}) \approx (\hat{\theta} - \theta^*)' F' F (\hat{\theta} - \theta^*) \quad (73)$$

Dado que $\varepsilon \sim N(0, s^2 \mathbf{I}_n)$, sob certas condições de regularidade e n suficientemente grande, pode-se demonstrar que:

- $\hat{\theta}$ e $s^2 = \frac{S(\hat{\theta})}{(n-p)}$ são estimadores consistentes de θ^* e σ^2 respectivamente;
- $\hat{\theta} - \theta \sim Np(0, \sigma^2 C^{-1})$, em que $C = F' F = F'(\theta^*) F(\theta)$;
- $\frac{(n-p)s^2}{\sigma^2} = \frac{\varepsilon'(\mathbf{I}_n - \mathbf{P}_F)\varepsilon}{\sigma^2} \sim \chi_{n-p}^2$;
- $\hat{\theta}$ e s^2 são estatisticamente independentes.

2.6.5 O método de Gauss-Newton

Suponha que $\theta^{(a)}$ é uma aproximação da estimativa de mínimos quadrados, $\hat{\theta}$ do modelo descrito pela equação 62. Para θ próximo de $\theta^{(a)}$ e considerando novamente uma expansão de Taylor de primeira ordem, da função de regressão como na equação 65, tem-se:

$$f(X_i, \theta) \approx f(X_i, \theta^{(a)}) + \sum_{j=1}^p \left[\frac{\partial f(X_i, \theta)}{\partial \theta_j} \right]_{\theta=\theta^{(a)}} (\theta_j - \theta_j^{(a)}) \quad (74)$$

ou, em notação matricial:

$$F(\theta) \approx F(\theta^{(a)}) + F'(\theta - \theta^{(a)}) \quad (75)$$

Definindo $r(\theta)$ como sendo um vetor de resíduos pode-se escrever:

$$r(\theta) = Y - f(\theta) \approx Y - f(\theta) - F^{(a)}(\theta - \theta^{(a)})$$

$$r(\theta) = r(\theta^{(a)}) - F^{(a)}(\theta - \theta^{(a)}) \quad (76)$$

Substituindo $r'(\theta)r(\theta)$ em $S(\theta)$, definido na equação 63, obtém-se:

$$S(\theta) \cong r'(\theta)r(\theta) - 2r(\theta)F^{(a)}(\theta - \theta^{(a)}) + (\theta - \theta^{(a)})'F'^{(a)}F^{(a)}(\theta - \theta^{(a)}) \quad (77)$$

Portanto, $S(\theta)$ será mínima quando:

$$\theta - \theta^{(a)} = [F'^{(a)} - F^{(a)}]^{-1} F'^{(a)} r(\theta^{(a)}) \quad (78)$$

Assim, devido à aproximação $\theta^{(a)}$, a próxima aproximação é dada por:

$$\theta^{(a+1)} = \theta^{(a)} + [F'^{(a)} - F^{(a)}]^{-1} F'^{(a)} r(\theta^{(a)}) \quad (79)$$

resultando no processo iterativo conhecido como *Método de Gauss-Newton* ou método da linearização.

Iniciando a determinação de $\theta^{(a)}$ com algum valor $\theta^{(a)}$, o processo continua até a convergência, que ocorre quando

$$|\theta^{(a+1)} - \theta^{(a)}| < \delta$$

onde δ é algum valor fixo, por exemplo, $\delta = e^{-10}$.

Apesar do método de Gauss-Newton ser numericamente estável, ele pode apresentar problemas, como, por exemplo: a convergência pode ser lenta se uma grande precisão for exigida; a matriz $F^{(a)}$ pode ser singular ou tornar-se singular durante o processo iterativo; e ainda a convergência pode ser para um mínimo local e não para o mínimo global.

Uma característica interessante do método de *Gauss-Newton* ocorre quando a função de regressão é linear. Neste caso, o processo converge em uma única interação, partindo-se de qualquer valor inicial.⁷⁴

Além do método *Gauss-Newton*, que é empregado no presente trabalho e dos já citados métodos de Marquardt e D.U.D., as estimativas de mínimos quadrados podem ser obtidas ainda por meio de outros métodos de otimização, como, por exemplo, o Método Simplex Nelder-Mead, Método de Otimização Quadrática, Método do gradiente Conjugado, Método Newton-Raphson, Método Newton-Raphson Rodge.

Capítulo III – Análise de séries temporais

Uma série temporal é definida como o conjunto de observações, N , de um fenômeno geradas sequencialmente no tempo.⁷⁵ São exemplos de series temporais, registros em intervalos regulares de tempo de dados econômicos, demográficos e ambientais; tais como os valores diários, semanais ou mensais da concentração de um poluente ambiental em uma localidade⁷⁶⁻⁸⁵. A análise de séries temporais é o processo de identificação das características, padrões e propriedades do conjunto de observações. Dentre os objetivos da análise de séries temporais, o maior deles é a previsão.

A previsão de séries temporais tem sido usada em diversos problemas reais, eliminando riscos gerados por incerteza e auxiliando o planejamento e a tomada de decisões, uma vez que a eficácia de uma decisão depende obviamente dos eventos anteriores à mesma. Diversos modelos foram desenvolvidos para prever séries temporais e selecionar o modelo mais adequado para prever uma dada série pode ser uma tarefa difícil, dependendo dos modelos avaliados e das características da série em questão.

Os modelos utilizados na análise de séries temporais, possuem propósito de identificar o padrão básico de **ruído**, a_t , que possa estar presente nas observações da série e usar esse padrão para prever valores futuros com a **série de saída**, Z_t . Os passos principais nessa análise consistem na estimação de uma **função de transferência**, F_t , conhecendo-se as séries de entrada, a_t , e saída, Z_t ; e no estudo o comportamento do sistema, ajustar os valores de entrada, a_t , e das previsões, Z_t . Os modelos que se aplicam na descrição de séries temporais são processos estocásticos, isto é, processos controlados por leis probabilísticas: $Z = \{Z(t), t \in T\}$, tal que, para cada $t \in T$, $Z(t)$ é uma variável aleatória.⁷⁵

Basicamente a análise temporal envolve a construção do gráfico da variável estudada em função do tempo para visualizar a existência ou não de componentes de tendência e sazonalidade, seleção de modelos mais adequados para descrever o comportamento da série e escolha do melhor modelo através da análise dos erros.

Os modelos de previsão se dividem em métodos simples e avançados. Entre os modelos simples, existem os **métodos de média móvel** e **alisamento exponencial simples** para séries localmente constantes; os **métodos de alisamento exponencial duplo** ou de **Brown** e **alisamento exponencial de Holt**, para séries com tendência e o **alisamento exponencial de Holt Winter** para séries com tendência e sazonalidade. Os modelos avançados fazem parte dos **modelos auto-regressivos** ou **método de Box e Jenkins**⁷⁵. Estes modelos serão descritos a seguir.

3.1 Modelos simples de previsão

3.1.1 Séries constantes

As séries localmente constantes são descritas por uma média μ_t somada a uma função a_t para o ruído:

$$Z_t = \mu_t + a_t \quad t= 1, 2, 3, \dots, N \quad (80)$$

Para essas séries são aplicáveis os métodos de média móvel e de alisamento exponencial simples.

3.1.1.1 Média móvel

O método de média móvel consiste em calcular a média aritmética de r observações mais recentes:

$$\hat{Z}_t = \frac{Z_t + Z_{t-1} + \dots + Z_{t-r+1}}{r} \quad (81)$$

e fazer a previsão do valor \hat{Z}_t por:

$$\hat{Z}_t = \hat{Z}_{t-1} + \frac{Z_t - Z_{t-r}}{r} \quad (82)$$

onde \hat{Z}_t é o valor estimado e Z_t o valor real.

As propriedades desse método dependem do valor de r escolhido e sua escolha será definida pelo menor erro quadrático médio, EQM , de previsão:

$$EQM = \frac{\sum_{t=1}^N (Z_t - \hat{Z}_t)^2}{N} \quad (83)$$

O método de média móvel é de simples aplicação; é aplicável quando se tem um número pequeno de observações e permite uma grande flexibilidade devido à variação de r de acordo com o padrão da série. Suas desvantagens são, a restrição de ser utilizado somente em séries estacionárias, a necessidade de armazenar pelo menos $(r-1)$ observações e a dificuldade de determinar o valor de r ⁷⁵.

3.1.1.2 Alisamento exponencial simples

Esse método extrai das observações da série temporal o comportamento para previsões atribuindo maior peso às observações mais recentes da série,

$$\hat{Z}_t = \alpha Z_t + \alpha(1-\alpha)Z_{t-1} + \alpha(1-\alpha)^2 Z_{t-2} + \dots \quad 0 \leq \alpha \leq 1 \quad (84)$$

$t = 1, 2, 3, \dots, N$

Assim, assumindo que o primeiro valor previsto será igual ao primeiro valor real da série, as previsões são determinadas por:

$$\begin{aligned} \hat{Z}_1 &= Z_1 \\ \hat{Z}_2 &= \alpha Z_2 + (1-\alpha)\hat{Z}_1 \\ \hat{Z}_3 &= \alpha Z_3 + (1-\alpha)\hat{Z}_2 \end{aligned}$$

A escolha de α será definida pelo menor erro quadrático médio, EQM , de previsão, de acordo com a equação 83.

O método de alisamento exponencial simples é de fácil entendimento, aplicação não dispendiosa e de grande flexibilidade permitida pela variação de α . Como a desvantagem do método está na dificuldade de encontrar o valor mais apropriado para α ⁷⁵.

3.1.2 Séries com tendência

As séries com tendência são descritas por uma média, μ , somada a uma função para tendência, T_t , e outra para o ruído, a_t :

$$Z_t = \mu + T_t + a_t \quad (85)$$

$$Z_t = \mu + \beta_0 + \beta_1 t + K + \beta_m t^m + a_t \quad (86)$$

$$\text{onde, } T_t = \beta_0 + \beta_1 t + K + \beta_m t$$

Para essas séries, são aplicáveis os métodos de alisamento exponencial duplo e o método de alisamento exponencial de Holt.

3.1.2.1 Alisamento exponencial duplo ou linear de Brown

A base do método é calcular um segundo valor exponencialmente alisado, de acordo com:

$$\bar{\bar{Z}}_t = \alpha \bar{Z}_t + (1 - \alpha) \bar{\bar{Z}}_{t-1} \quad (87)$$

As previsões são determinadas por:

$$\begin{aligned} \bar{\bar{Z}}_1 &= Z_1 \\ \bar{\bar{Z}}_2 &= \alpha \bar{Z}_2 + (1 - \alpha) \bar{\bar{Z}}_1 \\ \bar{\bar{Z}}_3 &= \alpha \bar{Z}_3 + (1 - \alpha) \bar{\bar{Z}}_2 \end{aligned}$$

A escolha de α será definida pelo menor erro quadrático médio, EQM , de previsão, de acordo com a equação 83⁷⁵.

3.1.2.2 Alisamento exponencial de Holt

Este método utiliza uma nova constante de suavização para modelar a tendência da série

$$\begin{aligned}\hat{Z}_t &= AZ_t + (1-A)(\hat{Z}_{t-1} + \hat{T}_{t-1}) & 0 < A < 1 & \quad e \quad t = 2,3,\Lambda, N \\ \hat{T}_t &= C(\hat{Z}_t - \hat{Z}_{t-1}) + (1+C)\hat{T}_{t-1} & 0 < C < 1 & \quad e \quad t = 2,3,\Lambda, N\end{aligned}\quad (88)$$

Neste caso, as constantes, A e C, são escolhidas de forma minimizar a somas dos erros quadráticos médio de previsão, calculado de acordo com a equação 83⁷⁵.

3.1.3 Séries com tendência e sazonalidade

3.1.3.1 Alisamento exponencial de Holt Winter

O alisamento exponencial de Holt Winter é utilizado para séries com tendência e sazonalidade. Para séries que apresentam uma variação sazonal constante, séries aditivas são utilizadas, na forma::

$$Z_t = \mu_t + F_t + T_t + a_t \quad (89)$$

As funções de sazonalidade e tendência são modeladas respectivamente por:

$$\hat{F}_t = D(Z_t - \hat{Z}_t) + (1-D)\hat{F}_{t-s}, \quad 0 < D < 1$$

$$\hat{T}_t = C(\hat{Z}_t - \hat{Z}_{t-1}) + (1-C)(\hat{T}_{t-1}), \quad 0 < C < 1$$

E as previsões do valor \hat{Z}_t são determinadas por;

$$\hat{Z}_t = A(Z_t - \hat{F}_{t-s}) + (1-A)(\hat{Z}_{t-1} + \hat{T}_{t-1}), \quad 0 < A < 1 \quad (90)$$

Para séries que apresentam irregularidade na variação sazonal, séries multiplicativas são utilizadas na descrição, como mostra a equação 91:

$$Z_t = \mu_t \times F_t + T_t + a_t \quad (91)$$

As funções de sazonalidade e tendência são modeladas respectivamente por:

$$\hat{F}_t = D \left\{ \frac{Z_t}{\hat{Z}_t} \right\} + (1-D)\hat{F}_{t-s} \quad 0 < D < 1, \quad t = s+1, \dots, N$$

$$\hat{T}_t = C(\hat{Z}_t - \hat{Z}_{t-1}) + (1-C)\hat{T}_{t-1} \quad 0 < C < 1, \quad t = s+1, \dots, N$$

E as previsões do valor \hat{Z}_t são determinadas por;

$$\hat{Z}_t = A \left\{ \frac{Z_t}{\hat{F}_{t-s}} \right\} + (1-A)(\hat{Z}_{t-1} + \hat{T}_{t-1}) \quad 0 < A < 1, \quad t = s+1, \dots, N \quad (92)$$

A determinação das constantes de alisamento A, C e D é realizada de modo a tornar mínimo o erro quadrático médio calculado de acordo com a equação 83.⁷⁵

3.2 Modelos auto-regressivos ou método Box e Jenkins

Os modelos avançados de previsão fazem parte da abordagem Box e Jenkins. Essa abordagem consiste em ajustar modelos auto – regressivos integrados de médias móveis a um conjunto de dados estacionários. A abordagem Box e Jenkins é sugerida em séries para séries com mais de 60 observações ($N > 60$).

Os parâmetros abordados nesse modelo são: auto-regressivos (p), média móvel (q) e diferenciação (d). A diferenciação é necessária para séries não estacionárias, onde diferenças são tomadas até que a série se torne estacionária.

Nessa metodologia os modelos podem ser: AR(p), MA (q), ARMA (p,q) e ARIMA (p,d,q).

A modelagem proposta por Box e Jenkins é da forma:

$$\varphi(B)Z_t = \theta(B)a_t \quad (93)$$

onde B é um operador de retardo tal que,

$$(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p)Z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)a_t \quad (94)$$

A função $\varphi(B)$ é de ordem p e corresponde a parte auto – regressiva (AR) e a função $\theta(B)$ é de ordem q e corresponde a parte de média móvel. Aplicando o operador B em ambos os lados, tem-se:

$$Z_t - \varphi_1 Z_{t-1} - \varphi_2 Z_{t-2} - \dots - \varphi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (95)$$

A identificação da ordem p, q é feita pela comparação das funções de autocorrelação, *FAC*, e as funções de autocorrelação parciais, *FACP*, da serie em estudo com modelos teóricos^{75,88}.

3.3 Critério de avaliação de modelos

O coeficiente de informação Bayesiana, BIC, é um critério de comparação entre vários modelos estimados para uma série. Quanto menor o valor desse critério, melhor é o modelo. O cálculo do BIC é definido pela equação:

$$BIC = N \times \log(EQM) + (\text{número de parâmetros}) \times \log N \quad (96)$$

onde, N = número de dados, EQM = erro quadrático médio de previsão e p e q = ordem do polinômio.⁷⁵

Capítulo IV – Descrição e pré-tratamento dos dados

4.1 Descrição da base de dados

Os corpos de água do Estado de Minas Gerais vêm sendo monitorados com objetivo de levantar, conhecer e divulgar as condições da qualidade das águas superficiais do Estado e propor ações de planejamento e controle de uso das mesmas.

O programa de monitoramento da sub - bacia do rio das Velhas, é constituído de 29 estações de amostragem que abrangem toda sua extensão. Este programa denominado de macro-rede de monitoramento vem sendo continuamente avaliado, visando aprimorar a localização dos pontos de amostragem, os parâmetros analisados e a frequência do monitoramento. Os relatórios anuais vêm alertando sobre a situação crítica na qual encontram-se suas águas, acusando problemas relacionados tanto a fontes pontuais de poluição, representadas pelos esgotos sanitários, atividades de mineração e efluentes de pequenas indústrias, quanto a fontes difusas, representadas pela drenagem pluvial da área urbana e rural e por rejeitos de mineração.¹⁻⁷

A TABELA 4 apresenta as principais atividades industriais e minerárias de cada região. Para facilitar a descrição da sub-bacia, ela é dividida em alto, médio e baixo curso.¹⁰

TABELA 4 - Atividades industriais e minerárias das regiões do alto, médio e baixo curso do rio das Velhas⁸

REGIÕES / ESTAÇÕES	DESCRIÇÃO DAS ATIVIDADES DE MINERAÇÃO	ATIVIDADES INDUSTRIAIS*
Alto curso BV013, BV035, BV037, BV067, BV076, BV139, BV062, BV063, BV155, BV083, BV154, BV105	A atividade extrativista na região do alto curso é intensa, as mineradoras utilizam a extração a céu aberto, o tipo de minério extraído é principalmente o de ferro, as técnicas empregadas são a de bancada única e escavação com equipamento mecânico. Podem ser encontradas extrações de ouro, topázio, argila, pedras e areia.	metalúrgicas >48 químicas > 24 alimentícias >22 produtos minerais não metálicos >20 mecânicas >10 madeira e mobiliário >9 borracha > 9 têxteis > 9 plástico > 6 material elétrico e comunicações > 4
Médio curso BV 130, BV153, BV135, BV137, BV160, BV156, BV140, BV141, BV142, BV143, BV161, BV162	A atividade minerária na região do médio curso é também intensa. Podem ser encontradas: extração de argila, de calcáreo, de pedras e areia, de minério de ferro, de ouro, de diamante.	indústria de produtos minerais não metálicos > 37 metalúrgicas > 30 alimentícias > 20 químicas > 12 bebidas > 7 perfumaria, sabões e vela > 5 têxteis > 3 madeira e mobiliário > 3 plásticos > 3 curtumes > 3
Baixo curso BV 152, BV146, BV147, BV148, BV149	A atividade minerária na região de baixo curso tem uma queda acentuada, comparada com as outras regiões, sendo uma de extração de minério de ferro, uma de extração de minério de metais não ferrosos, etc.	metalúrgicas > 3 bebidas > 3 química > 1 produtos minerais não metálicos > 2

*Valor aproximado do número de empresas

A região do alto curso do rio das Velhas abriga um dos principais mananciais de abastecimento urbano de água da Região Metropolitana de Belo Horizonte, atendendo a cerca de 50% de sua população, equivalente a dois milhões de habitantes.

Devido à relevância da região do alto curso, o IGAM vem realizando desde novembro de 2002, paralelamente ao monitoramento de toda a bacia do rio das Velhas, o aperfeiçoamento do monitoramento da qualidade das águas da bacia do alto curso do rio das Velhas, pelo Programa Nacional do Meio Ambiente II – PNMA II. Este programa é constituído de 37 estações na região do alto curso. As estações da região do alto curso estão descritas na TABELA 2 do capítulo 1.¹¹⁻¹²

Na região do alto curso, a rede dirigida ao abastecimento doméstico envolve as estações AV220 e AV230 Córrego dos Fechos nas barragens principal e auxiliar; AV240, Córrego Seco na captação dos Fechos; AV260 e AV270 rio das Velhas em Bela Fama, alça direita e esquerda; AV280 e AV290, Córrego da Mutuca, barragens principal e auxiliar, respectivamente¹¹⁻¹². Essas estações podem ser localizadas no mapa da FIGURA 1, capítulo 1.

Ao mesmo tempo, as águas do alto curso do rio das Velhas recebem efluentes de esgotos municipais e industriais de grande parte da região metropolitana de Belo Horizonte¹¹⁻¹². A importância econômica do alto curso do rio das Velhas advém do intenso desenvolvimento de atividades de mineração representadas, principalmente, por explorações de minério de ferro, ouro e gemas e da crescente ocupação urbana associada à forte expansão imobiliária.³

4.2 Metodologia de análises físico-química e estatística

A metodologia de análise dos parâmetros físico-químicos e microbiológicos,⁹⁰ utilizada pelo CETEC é apresentada na TABELA 5.

TABELA 5 – Metodologia de análises dos principais parâmetros analisados em águas da sub-bacia do rio das Velhas⁹⁰

PARÂMETRO	TIPO DE ENSAIO	REFERÊNCIA	LIMITE DE QUANTIFICAÇÃO
Alcalinidade total e alcalinidade bicarbonato	Titulometria	APHA ¹ 2320B	1,0 mg/L
Arsênio	Espectrometria de AA– gerador de hidretos	APHA 3114 B	0,0003 mg/L
Bário	Espectrometria de AA – plasma	APHA 3120B	0,005 mg/L
Boro	Espectrometria de AA – plasma	APHA 3120B	0,07 mg/L
Cádmio	Espectrometria de AA – forno de grafite	APHA 3113 B	0,0005 mg/L
Cálcio	Espectrometria de AA – plasma	APHA 3120B	0,10 mg/L
Chumbo	Espectrometria de AA – forno de grafite	APHA 3113 B	0,005 mg/L
Cianeto	Titulometria	APHA 4500-CN-F	0,01 mg/L
Cloreto	Colorimetria	USGS ² – I – 1187 78	0,3 mg/L
Cobre	Espectrometria de AA – plasma	APHA 3120B	0,004 mg/L
Coliformes fecais	Tubos múltiplos	APHA 9221 E	2 NMP/100 mL
Coliformes totais	Tubos múltiplos	APHA 9221 B	2 NMP/100 mL
Condutividade elétrica	Condutimetria	APHA 2510 B	1,00 µmho/cm
Cor	Colorimetria	APHA 2120B	5 UPt
Cromo hexavalente	Colorimetria	APHA 3500-Cr D	0,01 mg/L
Cromo trivalente	Espectrometria de AA – plasma	APHA 3120	0,04 mg/L
Demanda bioquímica de oxigênio	Winkler/incubação	ABNT/NBR ³ 12614/1192	2 mg/L
Demanda química de oxigênio	Titulometria	ABNT NBR 12614/1192	5 mg/L
Dureza cálcio	Titulometria	APHA 3500 – Ca D	1,0 mg/L
Dureza magnésio	Titulometria	APHA 3500 – Mg E	1,0 mg/L
Dureza total	Titulometria	APHA 2340	1,0 mg/L
Estreptococos fecais	Tubos múltiplos	APHA 9230 B	2 NMP/100 mL
Ferro solúvel	Espectrometria de AA – plasma	APHA 3120 B	0,03 mg/L
Fluoretos	Colorimetria	APHA 4500C	0,25 mg/L
Fosfatos	Colorimetria	APHA 4500- P C	0,01 mg/L P
Índice de fenóis	Colorimetria	ABNT NBR 10740/1989	0,001 mg/L
Manganês	Espectrometria de AA – plasma	APHA 3120B	0,003 mg/L
Magnésio	Espectrometria de AA – plasma	APHA 3120B	0,10 mg/L

TABELA 5 – Metodologia de análises dos principais parâmetros analisados em águas da sub-bacia do rio das Velhas⁹⁰

PARÂMETRO	TIPO DE ENSAIO	REFERÊNCIA	LIMITE DE QUANTIFICAÇÃO
Mercúrio	Espectrometria de AA – vapor frio	APHA 3112 B	0,20 µg/L
Níquel	Espectrometria de AA – grafite	APHA 3113B	0,004 mg/L
Nitrogênio amoniacal	Colorimetria	ABNT NBR10560/1988	0,10 mg/L
N Nitrato	Colorimetria	APHA 4500 - NO ₃ - E	0,01 mg/L
N Nitrito	Colorimetria	APHA NBR 12619	0,001 mg/L
N orgânico	Colorimetria	APHA 4500 – N orgânico B	0,10 mg/L
Oxigênio dissolvido	Titulometria	APHA NBR10559/1988	0,5 mg/L
Óleos e graxas	Gravimetria	APHA 5520	1 mg/L
pH	Potenciometria	APHA 5520 B	Não se aplica
Potássio	Espectrometria de AA – plasma	APHA 3120 B	0,10 mg/L
Selênio	Espectrometria de AA – gerador de hidretos	APHA 3114 B	0,0005 mg/L
Sódio	Espectrometria de AA – plasma	APHA 3120 B	0,10 mg/L
Sólidos dissolvidos	Gravimetria	ABNT NBR 10664/1989	1 mg/L
Sólidos em suspensão	Gravimetria	ABNT NBR 10664/1989	1 mg/L
Sólidos totais	Gravimetria	ABNT NBR 10664/1989	1 mg/L
Sulfatos	Gravimetria	APHA 4500 – SO ₄ 2E	1 mg/L
Sulfetos	Titulometria	APHA 4500 – S ₂ - F	0,5 mg/L
Surfactantes aniônicos	Colorimetria	ABNT NBR 10738/1989/1989	0,05 mg/L
Temperatura da água	Termometria	APHA 2550B	Não se aplica
Temperatura do ar	Termometria	APHA 2550B	Não se aplica
Turbidez	Turbidimetria	APHA 2130 B	1 mg/L
Zinco	Espectrometria de AA – plasma	APHA 3120 B	0,02 mg/L

*AA – Absorção atômica–As análises laboratoriais atendem às normas aprovadas pelo instituto nacional de metrologia, normalização e qualidade industrial ou, na sua ausência, aos métodos indicados no *Standard methods for the examination of water and wastewater*, 20^a edição, 1998. Ensaio realizado de acordo com procedimentos da 1= APHA- American Public Health Association; 2 = USGS-United States Geological Survey; 3 = ABNT-NBR (Associação Brasileira de Normas Técnicas).

Os limites de quantificação dos métodos de análise apresentados na TABELA 5, são na maioria dos métodos, 10 vezes inferiores aos padrões definidos para águas de classe 1.¹³

No banco de dados, crescente a cada campanha, o IGAM realiza um tratamento prévio dos dados brutos de análise publicando em relatórios anuais os resultados dos principais poluentes em cada região de coleta, buscando associar as principais fontes de impacto a cada estação^{5,8}. A implementação de uma metodologia de análise estatística sistemática do banco de dados foi o objetivo desse trabalho.

A primeira etapa teve como objetivo explorar a base de dados de toda sub-bacia buscando investigar através das análises fatorial e de variância, modelos de regressão em regiões da sub-bacia com características similares. Para essa etapa foi empregado o programa *SYSTAT 7.0 for Windows*⁸⁸. A base de dados utilizada nessa etapa é formada pelas estações que compõe a macro-rede de monitoramento no período compreendido entre janeiro de 1997 a julho de 2004.¹⁻⁷

A segunda etapa do tratamento dos dados teve como objetivo integrar os dados existentes através de análises exploratórias com análise de agrupamentos e componentes principais nas estações que compõem a macro-rede de monitoramento e nas estações da rede dirigida a alto curso do rio das Velhas. Para essa etapa do trabalho foi utilizado o pacote de programas: *Advanced Data Analysis Software S-PLUS*^{39,54, 55}. A mediana da concentração de cada parâmetro foi utilizada nesse estudo por ser mais representativa da variação de cada parâmetro durante o período e mais empregada na avaliação de dados ambientais²⁰. A TABELA 6 apresenta as medianas dos parâmetros em cada estação de amostragem da macro-rede de monitoramento no período de setembro de 1997 a outubro de 2005 e a TABELA 7 as estações que compõem a rede de monitoramento do alto curso do rio das Velhas no período de novembro de 2003 a dezembro de 2004.

Na terceira etapa do trabalho, modelos temporais para parâmetros relevantes em algumas estações de monitoramento foram avaliados. O objetivo dessa análise foi o diagnóstico das estações mais críticas e estimativas do comportamento futuro das mesmas.

A identificação dos parâmetros e estações mais importantes para esse estudo foi conseqüência das análises anteriores. O programa utilizado nessa etapa do trabalho foi *Forest Pro for Windows 3.5*⁸⁹. A base de dados utilizada nessa etapa envolve dados compreendidos no período entre janeiro de 1993 a novembro de 2005.^{1-7, 91-93}

A escolha dos programas ideais, para cada análise dos dados em cada etapa do trabalho foi fundamentada na conveniência na forma de organização dos dados e nos recursos disponíveis em cada um.

TABELA 6 - Mediana dos parâmetros na macro-rede de monitoramento no período de setembro de 1997 a outubro de 2005

	ESTAÇÃO	BV013	BV035	BV037	BV062	BV063	BV067
1	Alcalinidade de bicarbonato/mg/L	20,0	25,0	20,7	45,0	23,3	22,5
2	Alcalinidade total / mg/L CaCO ₃	19,3	25,9	21,2	44,0	22,8	21,9
3	Arsênio total / mg/L	1,00x10 ⁻³	7,00x10 ⁻⁴	9,00x10 ⁻⁴	9,14x10 ⁻²	1,41 x10 ⁻²	1,21x10 ⁻²
4	Bário total / mg/L	0,0945	0,200	0,174	0,114	0,133	0,155
5	Boro solúvel / mg/L	0,0700	0,0700	0,0700	0,0700	0,0700	0,0700
6	Cádmio total / mg/L	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴
7	Cálcio total / mg/L	4,50	7,80	5,70	12,8	8,10	8,20
8	Chumbo total / mg/L	9,00x10 ⁻³	9,00x10 ⁻³	9,00x10 ⁻³	5,00x10 ⁻³	5,00x10 ⁻³	8,00x10 ⁻³
9	Cianeto total / mg/L	0,0100	0,0100	0,0100	0,0100	0,0100	0,0100
10	Cloretos / mg/L	0,870	1,52	0,910	6,82	1,65	1,61
11	Cobre total / mg/L	7,00x10 ⁻³	7,00x10 ⁻³	7,00x10 ⁻³	7,00x10 ⁻³	7,00x10 ⁻³	7,00x10 ⁻³
12	Coliformes fecais / NMP/100mL	900	160x10 ³	700x 10	160x10 ³	30,0x10 ³	8,00x10 ³
13	Coliformes totais / NMP/100mL	2,70x10 ³	160x10 ³	40,0x10 ³	160x10 ³	90,0x10 ³	20,5x10 ³
14	Condutividade elétrica, in loco, 25°C / µmho/cm ⁻¹	44,8	73,8	57,3	149	105	94,0
15	Cor real / UPt	11,0	10,0	10,0	8,0	10,0	5,00
16	Cromo hexavalente / mg/L	0,0100	0,0100	0,0100	0,0100	0,0100	0,0100
17	Cromo trivalente / mg/L	0,0450	0,0400	0,0500	0,0400	0,0400	0,0500
18	DBO/ mg/L	2,00	4,00	2,00	10,9	3,00	2,00
19	DQO/ mg/L	8,40	16,0	10,5	30,5	12,5	10,0
20	Dureza de cálcio / mg/L CaCO ₃	11,1	18,9	13,6	31,5	21,2	20,8
21	Dureza de magnésio / mg/L CaCO ₃	8,20	9,10	8,30	21,3	14,9	12,0
22	Dureza total / mg/L CaCO ₃	17,5	24,4	21,4	48,1	36,1	32,9
23	Estreptococos fecais NMP/100mL	1,00x10 ³	40,0x10 ³	2,70x10 ³	160x10 ³	13,0x10 ³	5,00x10 ³
24	Ferro solúvel / mg/L	0,125	0,140	0,105	0,286	0,0950	0,0800
25	Fosfato total / mg/L P	0,0500	0,115	0,0500	0,333	0,0900	0,0700
26	Índice de fenóis/ mg/L	1,00x10 ⁻³	1,00x10 ⁻³	1,00x10 ⁻³	2,00x10 ⁻³	1,00x10 ⁻³	1,00x10 ⁻³
27	Magnésio total / mg/L	2,00	2,67	1,78	5,30	3,90	3,00
28	Manganês total / mg/L	0,210	0,340	0,28	0,45	0,46	0,37
29	Mercúrio total / µg/L	0,200	0,200	0,200	0,200	0,200	0,200
30	Níquel total / mg/L	5,00x10 ⁻³	4,00x10 ⁻³	4,00x10 ⁻³	4,00x10 ⁻³	0,0100	7,00x10 ⁻³
31	Nitrogênio amoniacal / mg/L	0,100	0,380	0,100	1,70	0,300	0,200
32	Nitrogênio nítrico / mg/L	0,140	0,130	0,130	0,360	0,230	0,260
33	Nitrogênio nitroso / mg/L	0,00370	0,00920	0,00700	0,0770	0,0290	0,0230
34	Nitrogênio orgânico / mg/L	0,200	0,570	0,300	0,950	0,400	0,350
35	Óleos e graxas / mg/L	1,00	1,00	1,00	1,03	1,00	1,00
36	Oxigênio dissolvido / mg/L	7,40	7,30	8,00	5,90	7,40	7,20
37	pH, in loco	7,00	7,00	7,20	7,10	7,10	7,10
38	pH, Lab	7,30	7,20	7,30	7,30	7,30	7,30
39	Potássio solúvel / mg/L	0,840	0,940	0,790	1,47	0,870	0,810
40	Selênio total / mg/L	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴
41	Sódio solúvel / mg/L	1,87	2,93	2,15	6,09	2,82	2,66
42	Sólidos dissolvidos, mg/L	35,0	54,0	47,0	96,0	80,0	75,0
43	Sólidos suspensos, mg/L	25,0	56,0	23,0	28,0	47,0	56,0
44	Sólidos totais / mg/L	59,0	88,0	75,0	116	159	120
45	Sulfatos / mg/L	1,10	5,15	3,40	11,5	14,2	10,4
46	Sulfetos / mg/L	0,500	0,500	0,500	0,500	0,500	0,500
47	Surfactantes aniônicos / mg/L	0,0500	0,0500	0,0500	0,0500	0,0500	0,0500
48	Temperatura da água / °C	22,0	21,8	22,2	21,8	22,3	22,8
49	Temperatura do ar / °C	23,0	25,0	24,0	25,0	26,0	26,0
50	Turbidez / NTU	36,5	33,5	33,7	24,1	61,5	41,9
51	Zinco total / mg/L	0,0200	0,0400	0,0300	0,0500	0,0500	0,0400

TABELA 6 -CONTINUAÇÃO

ESTAÇÃO	BV076	BV083	BV105	BV130	BV135	BV137
Alcalinidade de bicarbonato/mg/L	27,8	47,4	75,4	94,8	22,2	78,6
Alcalinidade total / mg/L CaCO ₃	27,8	45,5	73,2	95,7	21,8	75,9
Arsênio total / mg/L	2,20x10 ⁻³	1,12x10 ⁻²	1,73x10 ⁻²	1,10x10 ⁻³	6,00x10 ⁻⁴	3,78 x10 ⁻²
Bário total / mg/L	0,103	0,155	0,177	0,188	0,0665	0,172
Boro solúvel / mg/L	0,0200	0,0700	0,0700	0,0700	0,0700	0,0700
Cádmio total / mg/L	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴
Cálcio total / mg/L	8,50	15,9	17,5	34,7	5,30	18,9
Chumbo total / mg/L	6,00x10 ⁻³	8,00x10 ⁻³	8,00x10 ⁻³	7,00x10 ⁻³	5,00x10 ⁻³	0,0120
Cianeto total / mg/L	0,0100	0,0100	0,0100	0,0100	0,0100	0,0100
Cloretos / mg/L	3,88	11,9	16,0	11,8	1,23	11,2
Cobre total / mg/L	7,00x10 ⁻³	1,20 x10 ⁻²	1,70 x10 ⁻²	7,00x10 ⁻³	5,50x10 ⁻³	1,60 x10 ⁻²
Coliformes fecais / NMP/100mL	50,0x10 ³	160x10 ³	160x10 ³	13,0x10 ³	800	90,0x10 ³
Coliformes totais / NMP/100mL	160x10 ³	160x10 ³	160x10 ³	29,0x10 ³	2,30x10 ³	160x10 ³
Condutividade elétrica, in loco, 25°C / µmho/cm ⁻¹	98,2	200	244	260	50,0	222
Cor real / UPt	10,0	10,0	30,0	25,0	41,0	25,0
Cromo hexavalente / mg/L	0,0100	0,0100	0,0100	0,0100	0,0100	0,0100
Cromo trivalente / mg/L	0,0400	0,0500	0,0500	0,0400	0,0400	0,0450
DBO/ mg/L	4,00	17,7	21,0	5,50	2,00	12,0
DQO/ mg/L	18,1	45,5	58,5	26,5	7,4	42,2
Dureza de cálcio / mg/L CaCO ₃	20,7	37,2	42,3	86,0	13,1	47,1
Dureza de magnésio/mg/L CaCO ₃	15,7	15,1	15,8	14,3	6,0	15,6
Dureza total / mg/L CaCO ₃	34,8	46,6	56,2	99,7	19,1	61,1
Estrepto. fecais NMP/100mL	50,0x10 ³	90,0x10 ³	160x10 ³	30,0x10 ³	1,70x10 ³	30,0x10 ³
Ferro solúvel / mg/L	0,208	0,120	0,180	0,0670	0,355	0,126
Fosfato total / mg/L P	0,106	0,430	0,450	0,155	0,0450	0,345
Índice de fenóis / mg/L	1,00x10 ⁻³	4,00x10 ⁻³	4,00x10 ⁻³	1,90x10 ⁻³	1,00x10 ⁻³	3,00x10 ⁻³
Magnésio total / mg/L	3,90	3,60	3,80	3,50	1,40	3,90
Manganês total / mg/L	0,110	0,430	0,490	0,160	0,0900	0,600
Mercurio total / µg/L	0,20	0,20	0,20	0,20	0,20	0,20
Níquel total / mg/L	4,00x10 ⁻³	6,50x10 ⁻³	6,80x10 ⁻³	5,00x10 ⁻³	4,00x10 ⁻³	8,00x10 ⁻³
Nitrogênio amoniacal / mg/L	0,400	3,10	3,85	1,10	0,100	3,39
Nitrogênio nítrico / mg/L	0,740	0,330	0,110	0,770	0,110	0,150
Nitrogênio nitroso / mg/L	0,0320	0,0540	0,0510	0,143	0,00600	0,0190
Nitrogênio orgânico / mg/L	0,500	0,900	1,28	0,700	0,300	0,870
Óleos e graxas / mg/L	1,00	1,00	1,00	1,00	1,00	1,00
Oxigênio dissolvido / mg/L	7,20	5,20	1,60	4,90	7,20	1,90
pH, in loco	7,30	7,10	7,00	7,30	7,00	7,00
pH, Lab	7,40	7,40	7,20	7,60	7,30	7,30
Potássio solúvel / mg/L	1,21	2,05	4,15	4,01	1,14	3,53
Selênio total / mg/L	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴
Sódio solúvel / mg/L	4,02	7,82	15,7	14,1	3,42	13,0
Sólidos dissolvidos, mg/L	68,0	121	139	172	50,0	128
Sólidos suspensos, mg/L	10,0	80,0	84,0	94,0	25,0	103
Sólidos totais / mg/L	86,0	208	239	274	81,0	250
Sulfatos / mg/L	6,70	12,4	14,1	8,00	1,30	13,1
Sulfetos / mg/L	0,500	0,500	0,500	0,500	0,500	0,500
Surfactantes aniônicos / mg/L	0,0500	0,0500	0,0900	0,0500	0,0500	0,0500
Temperatura da água / °C	22,3	23,2	22,7	24,2	24,0	23,7
Temperatura do ar / °C	26,0	26,0	24,0	26,0	27,0	26,0
Turbidez / NTU	8,5	43,3	56,5	84,0	23,9	87,2
Zinco total / mg/L	0,0400	0,0850	0,0900	0,0700	0,0300	0,0800

TABELA 6 -CONTINUAÇÃO

ESTAÇÃO	BV139	BV140	BV141	BV142	BV143	BV146
Alcalinidade de bicarbonato/mg/L	20,3	98,7	64,5	57,8	13,0	44,3
Alcalinidade total / mg/L CaCO ₃	20,5	99,2	63,6	58,5	14,8	48,3
Arsênio total / mg/L	9,00x10 ⁻⁴	8,00x10 ⁻⁴	3,69x10 ⁻²	3,25 x10 ⁻²	1,10x10 ⁻³	1,78 x10 ⁻³
Bário total / mg/L	0,115	0,106	0,188	0,196	0,0325	0,163
Boro solúvel / mg/L	0,0700	0,0700	0,0700	0,0700	0,0700	0,0700
Cádmio total / mg/L	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴
Cálcio total / mg/L	5,50	31,2	25,1	23,9	7,50	19,5
Chumbo total / mg/L	5,00x10 ⁻³	5,00x10 ⁻³	6,00x10 ⁻³	7,00x10 ⁻³	5,00x10 ⁻³	8,00x10 ⁻³
Cianeto total / mg/L	0,0100	0,0100	0,0100	0,0100	0,0100	0,0100
Cloretos / mg/L	1,02	8,58	10,5	9,15	0,900	5,22
Cobre total / mg/L	7,00x10 ⁻³	7,00x10 ⁻³	7,00x10 ⁻³	8,00x10 ⁻³	7,00x10 ⁻³	7,00x10 ⁻³
Coliformes fecais / NMP/100mL	17,0x10 ³	300	1,10 x10 ³	1,10 x10 ³	280	230
Coliformes totais / NMP/100mL	42,5x10 ³	1,30x10 ³	8,00x10 ³	2,05x10 ³	1,20x10 ³	950
Condutividade elétrica, in loco, 25°C / µmho/cm ⁻¹	55,3	280	205	199	40,3	143
Cor real / UPt	5,00	25,0	15,0	17,0	50,0	25,0
Cromo hexavalente / mg/L	0,0100	0,0100	0,0100	0,0100	0,0100	0,0100
Cromo trivalente / mg/L	0,0400	0,0400	0,0500	0,0500	0,0400	0,0450
DBO/ mg/L	2,00	2,00	5,00	4,00	2,00	3,00
DQO/ mg/L	8,00	12,0	20,0	16,5	12,5	22,0
Dureza de cálcio / mg/L CaCO ₃	13,7	78,3	61,6	58,8	11,4	46,7
Dureza de magnésio/mg/L CaCO ₃	8,00	11,2	15,1	13,9	4,0	10,2
Dureza total / mg/L CaCO ₃	19,4	88,8	72,1	75,2	16,7	57,3
Estrepto. fecais NMP/100mL	8,00x10 ³	500	400	600	470	465
Ferro solúvel / mg/L	0,0850	0,0920	0,0550	0,0500	0,145	0,0700
Fosfato total / mg/L P	0,0500	0,180	0,190	0,135	0,0350	0,0900
Índice de fenóis / mg/L	1,00x10 ⁻³	1,50x10 ⁻³	2,00x10 ⁻³	1,00x10 ⁻³	2,00x10 ⁻³	1,00x10 ⁻³
Magnésio total / mg/L	1,93	2,80	3,60	3,50	1,00	2,40
Manganês total / mg/L	0,310	0,0500	0,500	0,280	0,0500	0,150
Mercurio total / µg/L	0,200	0,200	0,200	0,200	0,200	0,200
Níquel total / mg/L	0,0040	0,0040	0,0050	0,0055	0,0040	0,0045
Nitrogênio amoniacal / mg/L	0,100	0,160	1,20	0,120	0,100	0,170
Nitrogênio nítrico / mg/L	0,140	0,990	0,980	1,35	0,110	0,740
Nitrogênio nitroso / mg/L	8,00x10 ⁻³	0,0330	0,190	0,0877	4,00x10 ⁻³	0,0190
Nitrogênio orgânico / mg/L	0,400	0,650	0,750	0,700	0,350	0,750
Óleos e graxas / mg/L	1,00	1,00	1,00	1,00	1,00	1,00
Oxigênio dissolvido / mg/L	7,50	6,80	5,20	5,50	7,20	6,70
pH, in loco	7,20	7,60	7,20	7,20	7,10	7,50
pH, Lab	7,40	7,80	7,50	7,60	7,20	7,70
Potássio solúvel / mg/L	0,710	3,05	3,15	2,98	0,730	2,37
Selênio total / mg/L	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴
Sódio solúvel / mg/L	2,20	21,4	11,8	11,5	1,49	7,69
Sólidos dissolvidos, mg/L	41,0	173	134	130	39,0	101
Sólidos suspensos, mg/L	35,0	38,0	76,0	61,0	26,0	60,0
Sólidos totais / mg/L	83,0	217	198	197	66,0	144
Sulfatos / mg/L	2,90	11,9	10,1	9,25	1,05	7,80
Sulfetos / mg/L	0,500	0,500	0,500	0,500	0,500	0,500
Surfactantes aniônicos / mg/L	0,0500	0,0500	0,0500	0,0500	0,0500	0,0500
Temperatura da água / °C	22,5	24,1	26,0	25,1	24,0	27,0
Temperatura do ar / °C	26,0	27,0	29,0	25,0	26,0	31,0
Turbidez / NTU	35,2	28,4	30,5	37,2	38,4	40,8
Zinco total / mg/L	0,0300	0,0350	0,0600	0,0350	0,0200	0,0300

TABELA 6 -CONTINUAÇÃO

ESTAÇÃO	BV147	BV148	BV149	BV152	BV153	BV154
Alcalinidade de bicarbonato/mg/L	34,3	47,8	45,5	47,5	77,9	147,9
Alcalinidade total / mg/L CaCO ₃	37,9	49,4	46,8	48,8	79,9	141,3
Arsênio total / mg/L	4,00x10 ⁻⁴	1,43x10 ⁻²	1,16x10 ⁻²	1,97x10 ⁻²	1,52x10 ⁻²	1,20 x10 ⁻³
Bário total / mg/L	0,0750	0,1115	0,1215	0,1620	0,1770	0,1390
Boro solúvel / mg/L	0,0700	0,0700	0,0700	0,0700	0,0700	0,0700
Cádmio total / mg/L	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	0,0005	5,00x10 ⁻⁴
Cálcio total / mg/L	10,9	17,7	16,0	18,7	21,0	31,8
Chumbo total / mg/L	5,00x10 ⁻³	5,00x10 ⁻³	5,00x10 ⁻³	5,00x10 ⁻³	7,00x10 ⁻³	8,00x10 ⁻³
Cianeto total / mg/L	0,0100	0,0100	0,0100	0,0100	0,0100	0,0100
Cloretos / mg/L	0,730	5,38	4,76	5,93	15,8	25,4
Cobre total / mg/L	7,00x10 ⁻³	7,00x10 ⁻³	6,00x10 ⁻³	7,00x10 ⁻³	0,0110	8,00x10 ⁻³
Coliformes fecais / NMP/100mL	110	170	120	220	160x10 ³	160x10 ³
Coliformes totais / NMP/100mL	1200	1200	700	2,00x10 ³	160x10 ³	160x10 ³
Condutividade elétrica, in loco, 25°C / µmho/cm ⁻¹	69,4	123	119	151	265	418
Cor real / UPt	40,0	15,0	17,0	16,0	20,0	25,0
Cromo hexavalente / mg/L	0,0100	0,0100	0,0100	0,0100	0,0100	0,0100
Cromo trivalente / mg/L	0,0400	0,0450	0,0400	0,0500	0,0500	0,0400
DBO/ mg/L	2,00	2,00	2,00	3,00	19,0	32,5
DQO/ mg/L	13,5	14,5	12,5	16,0	57,4	86,5
Dureza de cálcio / mg/L CaCO ₃	28,6	44,6	40,8	46,5	52,3	79,0
Dureza de magnésio/mg/L CaCO ₃	8,50	8,60	9,90	9,20	15,4	15,5
Dureza total / mg/L CaCO ₃	37,8	57,3	51,8	56,4	65,9	93,2
Estrepto. fecais NMP/100mL	500	650	130	300	90,0x10 ³	160x10 ³
Ferro solúvel / mg/L	0,190	0,0700	0,0570	0,0860	0,150	0,0850
Fosfato total / mg/L P	0,0400	0,0700	0,0700	0,0950	0,380	0,850
Índice de fenóis / mg/L	1,00x10 ⁻³	1,80x10 ⁻³	1,00x10 ⁻³	1,00x10 ⁻³	4,00x10 ⁻³	4,00x10 ⁻³
Magnésio total / mg/L	2,20	2,00	2,30	2,00	3,90	3,90
Manganês total / mg/L	0,0500	0,0800	0,0700	0,200	0,520	0,210
Mercurio total / µg/L	0,200	0,200	0,200	0,200	0,200	0,200
Níquel total / mg/L	4,00x10 ⁻³	4,00x10 ⁻³	4,00x10 ⁻³	8,00x10 ⁻³	7,50x10 ⁻³	7,00x10 ⁻³
Nitrogênio amoniacal / mg/L	0,100	0,100	0,100	0,100	3,70	9,35
Nitrogênio nítrico / mg/L	0,130	0,680	0,790	0,740	0,140	0,150
Nitrogênio nitroso / mg/L	5,00x10 ⁻³	9,00x10 ⁻³	8,00x10 ⁻³	0,0140	0,0530	0,0760
Nitrogênio orgânico / mg/L	0,400	0,550	0,500	0,650	1,08	2,23
Óleos e graxas / mg/L	1,00	1,00	1,00	1,09	2,00	4,00
Oxigênio dissolvido / mg/L	6,80	6,70	6,30	6,60	1,10	3,10
pH, in loco	7,10	7,50	7,30	7,50	7,10	7,50
pH, Lab	7,40	7,60	7,50	7,70	7,20	7,50
Potássio solúvel / mg/L	1,22	1,98	1,88	2,33	4,38	8,75
Selênio total / mg/L	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴
Sódio solúvel / mg/L	2,64	6,61	6,94	8,27	17,2	35,2
Sólidos dissolvidos, mg/L	72,0	89,0	83,0	102,0	158,0	230,0
Sólidos suspensos, mg/L	23,0	30,0	26,0	65,0	96,0	90,0
Sólidos totais / mg/L	89,0	126	113	162	246	308
Sulfatos / mg/L	1,30	6,70	6,30	7,50	13,5	12,3
Sulfetos / mg/L	0,500	0,500	0,500	0,500	0,500	0,500
Surfactantes aniônicos / mg/L	0,0500	0,0500	0,0500	0,0500	0,0770	0,320
Temperatura da água / °C	23,7	25,8	27,1	26,9	23,7	24,0
Temperatura do ar / °C	24,0	27,0	30,0	30,0	26,0	25,0
Turbidez / NTU	53,3	25,9	21,5	43,5	53,5	43,5
Zinco total / mg/L	0,0200	0,0350	0,0380	0,0300	0,0760	0,0850

TABELA 6 -CONTINUAÇÃO

ESTAÇÃO	BV155	BV156	BV160	BV161	BV162
Alcalinidade de bicarbonato/mg/L	124	84,6	73,3	73,9	34,0
Alcalinidade total / mg/L CaCO ₃	125	77,5	73,4	71,9	47,1
Arsênio total / mg/L	1,50x10 ⁻³	0,0400	8,00x10 ⁻⁴	3,00x10 ⁻⁴	4,00x10 ⁻⁴
Bário total / mg/L	0,166	0,183	0,169	0,0270	0,0100
Boro solúvel / mg/L	0,0700	0,0700	0,0700	0,0700	0,0700
Cádmio total / mg/L	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴
Cálcio total / mg/L	30,5	25,3	19,1	27,0	11,8
Chumbo total / mg/L	6,00x10 ⁻³	7,00x10 ⁻³	6,00x10 ⁻³	5,00x10 ⁻³	5,00x10 ⁻³
Cianeto total / mg/L	0,0100	0,0100	0,0100	0,0100	0,0100
Cloretos / mg/L	30,0	11,9	8,90	5,05	1,02
Cobre total / mg/L	0,0130	0,00950	0,00500	0,00600	0,00400
Coliformes fecais / NMP/100mL	160x10 ³	8,00x10 ³	160x10 ³	235	140
Coliformes totais / NMP/100mL	160x10 ³	27,0x10 ³	160x10 ³	950	1,05x10 ³
Condutividade elétrica, in loco, 25°C / µmho/cm ⁻¹	430	225	191	176	81,2
Cor real / UPt	25,0	10,0	60,0	52,5	20,0
Cromo hexavalente / mg/L	0,0100	0,0100	0,0100	0,0100	0,0100
Cromo trivalente / mg/L	0,0400	0,0500	0,0400	0,0400	0,0400
DBO/ mg/L	52,5	7,50	7,00	2,00	2,00
DQO/ mg/L	123	20,0	29,0	13,0	10,0
Dureza de cálcio / mg/L CaCO ₃	74,5	61,7	48,4	57,2	31,0
Dureza de magnésio/mg/L CaCO ₃	19,8	14,2	12,4	9,80	6,50
Dureza total / mg/L CaCO ₃	95,6	75,2	60,8	67,0	38,4
Estrepto. fecais NMP/100mL	160x10 ³	1,95x10 ³	125x10 ³	800	130
Ferro solúvel / mg/L	0,130	0,0950	0,285	0,180	0,160
Fosfato total / mg/L P	0,860	0,250	0,160	0,0600	0,0200
Índice de fenóis / mg/L	0,0110	0,00300	0,00300	0,00100	0,00100
Magnésio total / mg/L	4,80	3,40	3,00	2,45	1,30
Manganês total / mg/L	0,640	0,680	0,150	0,0700	0,0300
Mercurio total / µg/L	0,200	0,200	0,200	0,200	0,200
Níquel total / mg/L	9,00x10 ⁻³	0,0115	7,00x10 ⁻³	4,00x10 ⁻³	4,00x10 ⁻³
Nitrogênio amoniacal / mg/L	9,30	2,90	1,27	0,100	0,100
Nitrogênio nítrico / mg/L	0,230	0,410	0,700	0,810	0,0800
Nitrogênio nitroso / mg/L	0,152	0,0610	0,0995	0,0240	0,00350
Nitrogênio orgânico / mg/L	2,73	0,700	0,800	0,500	0,300
Óleos e graxas / mg/L	3,00	1,00	1,00	1,00	1,00
Oxigênio dissolvido / mg/L	2,50	2,90	5,30	5,70	6,90
pH, in loco	7,40	7,10	7,20	7,30	7,30
pH, Lab	7,50	7,40	7,30	7,60	7,50
Potássio solúvel / mg/L	7,32	3,16	3,95	2,56	0,620
Selênio total / mg/L	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴	5,00x10 ⁻⁴
Sódio solúvel / mg/L	27,7	11,6	13,1	9,31	1,78
Sólidos dissolvidos, mg/L	223	133	141	131	51,5
Sólidos suspensos, mg/L	72,0	94,4	56,0	12,5	11,0
Sólidos totais / mg/L	300	222	220	159	67,5
Sulfatos / mg/L	16,9	9,95	4,70	3,40	2,40
Sulfetos / mg/L	0,500	0,500	0,500	0,500	0,500
Surfactantes aniônicos / mg/L	0,870	0,0500	0,0650	0,0500	0,0500
Temperatura da água / °C	23,7	24,9	23,7	24,3	25,0
Temperatura do ar / °C	26,0	28,0	27,0	27,0	26,0
Turbidez / NTU	41,0	67,3	62,9	11,7	7,3
Zinco total / mg/L	0,130	0,0650	0,0600	0,0400	0,0300

TABELA 7 - Mediana dos parâmetros nas estações de monitoramento do alto curso do rio das velhas no período de novembro de 2003 a dezembro de 2004

ESTAÇÕES	ALCALINIDADE TOTAL (mg/L)	ARSENIO TOTAL (mg/L)	BÁRIO TOTAL (mg/L)	CHUMBO TOTAL (mg/L)	CLORETOS (mg/L)	COBRE TOTAL (mg/L)	COLIFORMES FECALIS NMP/100mL	CONDUTIV. ELÉTRICA $\mu\text{mho}/\text{cm}^{-1}$
AV005	10,7	$3,50 \times 10^{-4}$	$9,00 \times 10^{-4}$	$5,00 \times 10^{-3}$	0,630	$4,00 \times 10^{-3}$	110	26,9
AV010	20,6	$1,70 \times 10^{-3}$	$5,30 \times 10^{-2}$	$5,00 \times 10^{-3}$	0,710	$1,10 \times 10^{-3}$	790	46,0
AV020	19,2	$1,00 \times 10^{-3}$	$6,10 \times 10^{-2}$	$5,00 \times 10^{-3}$	1,65	$6,00 \times 10^{-3}$	$13,5 \times 10^3$	50,0
AV030E	15,9	$7,10 \times 10^{-4}$	$1,70 \times 10^{-2}$	$5,00 \times 10^{-3}$	0,950	$6,00 \times 10^{-3}$	121	40,7
AV030F	17,3	$4,50 \times 10^{-4}$	$3,60 \times 10^{-2}$	$5,00 \times 10^{-3}$	0,930	$9,00 \times 10^{-3}$	195	44,3
AV040	15,1	$3,00 \times 10^{-4}$	$2,90 \times 10^{-2}$	$5,00 \times 10^{-3}$	0,870	$4,00 \times 10^{-3}$	115	41,4
AV050	26,2	$3,00 \times 10^{-4}$	$3,20 \times 10^{-2}$	$5,00 \times 10^{-3}$	0,340	$6,00 \times 10^{-3}$	690	79,6
AV060	7,70	$1,44 \times 10^{-3}$	$2,20 \times 10^{-2}$	$6,00 \times 10^{-3}$	0,580	$6,00 \times 10^{-3}$	305	19,6
AV070	20,0	$3,00 \times 10^{-4}$	$1,20 \times 10^{-1}$	$8,50 \times 10^{-3}$	0,370	$2,60 \times 10^{-2}$	170	70,2
AV080	18,1	$3,00 \times 10^{-4}$	$1,40 \times 10^{-1}$	$5,00 \times 10^{-3}$	0,500	$5,00 \times 10^{-3}$	400	58,4
AV090	7,70	$3,00 \times 10^{-4}$	$7,00 \times 10^{-3}$	$5,00 \times 10^{-3}$	0,350	$4,00 \times 10^{-3}$	30,0	25,9
AV100	6,40	$3,00 \times 10^{-4}$	$7,00 \times 10^{-3}$	$5,00 \times 10^{-3}$	0,530	$4,00 \times 10^{-3}$	170×10^3	28,2
AV110	19,2	$3,00 \times 10^{-4}$	$6,80 \times 10^{-2}$	$5,00 \times 10^{-3}$	1,38	$4,00 \times 10^{-3}$	$1,15 \times 10^3$	72,4
AV120	5,20	$3,00 \times 10^{-4}$	$2,00 \times 10^{-2}$	$5,00 \times 10^{-3}$	0,510	$4,00 \times 10^{-3}$	195	25,2
AV130	16,8	$4,00 \times 10^{-4}$	$4,70 \times 10^{-2}$	$5,00 \times 10^{-3}$	1,05	$4,00 \times 10^{-3}$	$5,15 \times 10^3$	53,4
AV140	11,5	$7,70 \times 10^{-4}$	$1,90 \times 10^{-2}$	$5,00 \times 10^{-3}$	0,380	$4,00 \times 10^{-3}$	50,0	50,1
AV150	9,10	$3,00 \times 10^{-4}$	$3,40 \times 10^{-2}$	$5,00 \times 10^{-3}$	0,620	$4,00 \times 10^{-3}$	90,0	68,7
AV160E	14,1	$3,00 \times 10^{-4}$	$6,00 \times 10^{-3}$	$5,00 \times 10^{-3}$	0,580	$4,00 \times 10^{-3}$	2,00	32,5
AV160F	14,6	$3,00 \times 10^{-4}$	$6,00 \times 10^{-3}$	$5,00 \times 10^{-3}$	0,510	$4,00 \times 10^{-3}$	195	32,6
AV170	12,2	$3,00 \times 10^{-4}$	$5,00 \times 10^{-3}$	$5,00 \times 10^{-3}$	0,470	$4,00 \times 10^{-3}$	190	34,0
AV180E	13,9	$3,00 \times 10^{-4}$	$2,20 \times 10^{-2}$	$5,00 \times 10^{-3}$	0,530	$4,00 \times 10^{-3}$	5,50	50,3
AV180F	12,6	$3,00 \times 10^{-4}$	$2,60 \times 10^{-2}$	$5,00 \times 10^{-3}$	0,640	$4,00 \times 10^{-3}$	195	51,5
AV190	16,7	$4,00 \times 10^{-4}$	$2,40 \times 10^{-2}$	$5,00 \times 10^{-3}$	0,560	$4,00 \times 10^{-3}$	65,0	49,5
AV200	11,1	$1,05 \times 10^{-3}$	$1,70 \times 10^{-2}$	$5,00 \times 10^{-3}$	0,360	$5,00 \times 10^{-3}$	50,0	27,5
AV210	17,6	$3,00 \times 10^{-4}$	$3,60 \times 10^{-2}$	$5,00 \times 10^{-3}$	1,11	$4,00 \times 10^{-3}$	$7,50 \times 10^3$	47,9
AV220	27,6	$3,00 \times 10^{-4}$	$5,00 \times 10^{-3}$	$5,00 \times 10^{-3}$	0,400	$4,00 \times 10^{-3}$	11,0	59,7
AV230	32,6	$3,00 \times 10^{-4}$	$5,00 \times 10^{-3}$	$5,00 \times 10^{-3}$	0,470	$4,00 \times 10^{-3}$	11,0	61,5
AV240	1,40	$3,00 \times 10^{-4}$	$5,00 \times 10^{-3}$	$5,00 \times 10^{-3}$	0,320	$4,00 \times 10^{-3}$	2,00	6,60
AV250	16,8	$3,00 \times 10^{-4}$	$1,10 \times 10^{-2}$	$5,00 \times 10^{-3}$	0,690	$4,00 \times 10^{-3}$	$1,33 \times 10^3$	42,1
AV260	18,1	$8,50 \times 10^{-4}$	$3,00 \times 10^{-2}$	$5,00 \times 10^{-3}$	1,08	$4,00 \times 10^{-3}$	$8,00 \times 10^3$	50,2
AV270	18,0	$3,00 \times 10^{-4}$	$2,70 \times 10^{-2}$	$5,00 \times 10^{-3}$	1,05	$5,00 \times 10^{-3}$	$6,50 \times 10^3$	51,3
AV280	5,10	$3,00 \times 10^{-4}$	$5,00 \times 10^{-3}$	$5,00 \times 10^{-3}$	0,300	$4,00 \times 10^{-3}$	23,0	13,2
AV290	5,70	$3,00 \times 10^{-4}$	$5,00 \times 10^{-3}$	$5,00 \times 10^{-3}$	0,400	$4,00 \times 10^{-3}$	2,00	15,9
AV300	31,5	$5,50 \times 10^{-4}$	$1,70 \times 10^{-2}$	$5,00 \times 10^{-3}$	5,14	$4,00 \times 10^{-3}$	125×10^3	117
AV310	38,6	$6,88 \times 10^{-2}$	$1,80 \times 10^{-2}$	$5,00 \times 10^{-3}$	6,28	$4,00 \times 10^{-3}$	125×10^3	141
AV320	27,1	$3,32 \times 10^{-2}$	$3,00 \times 10^{-2}$	$5,00 \times 10^{-3}$	8,12	$5,80 \times 10^{-2}$	$2,90 \times 10^3$	$1,82 \times 10^3$
AV330	21,4	$1,36 \times 10^{-2}$	$6,60 \times 10^{-2}$	$5,00 \times 10^{-3}$	1,75	$5,00 \times 10^{-3}$	$50,0 \times 10^3$	78,7
AV340	10,8	$3,00 \times 10^{-4}$	$6,00 \times 10^{-3}$	$5,00 \times 10^{-3}$	1,35	$5,00 \times 10^{-3}$	$11,7 \times 10^3$	31,3
AV345	9,00	$3,00 \times 10^{-4}$	$5,00 \times 10^{-3}$	$5,00 \times 10^{-3}$	0,530	$4,00 \times 10^{-3}$	110	24,0
AV350	19,7	$8,70 \times 10^{-3}$	$6,30 \times 10^{-2}$	$6,00 \times 10^{-3}$	1,40	$6,00 \times 10^{-3}$	$12,0 \times 10^3$	75,4

TABELA 7 - CONTINUAÇÃO

ESTAÇÕES	COR REAL (Upt)	DEMANDA BIOQUÍMICA DE OXIGÊNIO (mg/L)	DEMANDA QUÍMICA DE OXIGÊNIO (mg/L)	DUREZA DE CÁLCIO (mg/L)	DUREZA DE MAGNÉSIO (mg/L)	ESTREPTOCOCOS FECAIS (NMP/100mL)	FERRO SOLÚVEL (mg/L)	FERRO TOTAL (mg/L)
AV005	21,5	2,00	5,00	6,15	4,90	55,0	0,190	0,660
AV010	26,0	2,00	5,00	10,9	9,85	720	0,190	1,76
AV020	51,0	2,00	11,5	8,50	6,85	6,50x10 ³	0,210	2,18
AV030E	30,0	2,00	5,00	8,30	7,05	13,5	0,100	0,690
AV030F	26,5	2,00	7,00	9,50	6,75	305	0,120	1,77
AV040	10,0	2,00	11,50	7,55	7,00	135	0,140	1,13
AV050	9,00	2,00	6,00	21,3	12,2	900	3,00x10 ⁻²	0,420
AV060	44,0	2,00	10,0	3,30	1,85	14,0	0,340	1,40
AV070	24,0	2,00	6,50	16,6	10,3	95,0	5,00x10 ⁻²	1,07
AV080	23,5	2,00	11,5	13,6	8,45	95,0	0,110	3,49
AV090	14,0	2,00	5,00	6,55	5,40	170	0,110	0,520
AV100	14,5	2,00	5,00	4,25	3,80	2,25x10 ³	0,140	0,740
AV110	12,5	3,00	15,5	12,6	6,40	30,0x10 ³	0,150	2,15
AV120	21,5	2,00	7,00	4,15	3,95	650	0,160	1,05
AV130	12,5	2,00	12,5	9,85	7,60	6,15x10 ³	0,130	0,880
AV140	26,0	2,00	8,50	7,55	5,60	600	0,200	0,920
AV150	12,5	2,00	5,00	4,45	4,00	150	7,00x10 ⁻²	0,770
AV160E	6,00	2,00	8,00	9,25	7,00	2,00	3,00x10 ⁻²	0,110
AV160F	6,50	2,00	6,00	10,1	6,15	305	3,00x10 ⁻²	0,190
AV170	9,50	2,00	5,00	8,95	6,40	305	0,240	0,780
AV180E	14,0	2,00	8,00	7,00	5,80	2,00	3,00x10 ⁻²	0,200
AV180F	13,0	2,00	6,50	6,80	5,25	305	3,00x10 ⁻²	0,350
AV190	9,50	2,00	5,00	9,70	7,30	150	3,00x10 ⁻²	0,580
AV200	14,5	2,00	5,00	6,65	5,95	75,0	3,00x10 ⁻²	0,250
AV210	18,5	2,00	8,00	9,25	8,40	6,0x10 ³	0,140	2,44
AV220	5,00	2,00	5,00	21,5	12,5	8,00	3,00x10 ⁻²	0,150
AV230	7,00	2,00	5,00	20,4	12,6	8,00	3,00x10 ⁻²	0,190
AV240	5,00	2,00	5,00	2,60	2,20	2,00	3,00x10 ⁻²	0,170
AV250	16,5	2,00	5,00	9,25	9,15	800,0	9,00x10 ⁻²	1,52
AV260	13,5	2,00	5,00	10,0	8,85	6,50x10 ³	0,120	2,44
AV270	17,5	2,00	6,00	10,7	8,55	2,65x10 ³	0,110	3,20
AV280	10,0	2,00	5,00	3,90	3,00	34,0	4,00x10 ⁻²	0,210
AV290	6,00	2,00	5,00	5,20	3,10	4,00	4,00x10 ⁻²	0,130
AV300	14,5	4,00	15,0	19,1	14,9	70,0x10 ³	0,120	0,760
AV310	14,5	10,5	36,5	22,5	19,1	125x10 ³	0,290	2,52
AV320	11,0	4,50	19,0	739	462	2,30x10 ³	3,00x10 ⁻²	0,960
AV330	6,00	2,00	12,0	18,4	13,6	24,0x10 ³	9,00x10 ⁻²	5,34
AV340	18,0	3,00	11,5	5,90	5,45	40,0x10 ³	0,170	0,880
AV345	14,0	2,00	5,00	5,10	4,70	30,0	0,150	0,400
AV350	12,0	2,00	11,5	15,3	11,5	5,00x10 ³	0,100	3,48

TABELA 7 - CONTINUAÇÃO

ESTAÇÕES	FLUORETOS (mg/L)	FÓSFORO TOTAL (mg/L)	ÍNDICE DE FENÓIS (mg/L)	MANGANÊS TOTAL (mg/L)	NÍQUEL TOTAL (mg/L)	NITROGÊNIO AMONÍACAL (mg/L)	NITROGÊNIO NÍTRICO (mg/L)	NITROGÊNIO NITROSO (mg/L)
AV005	0,250	2,00x10 ⁻²	1,00x10 ⁻³	3,00x10 ⁻²	4,00x10 ⁻³	0,100	8,00x10 ⁻²	3,00x10 ⁻³
AV010	0,250	5,00x10 ⁻²	1,00x10 ⁻³	0,270	1,70x10 ⁻²	0,150	9,00x10 ⁻²	2,50x10 ⁻³
AV020	0,250	8,00x10 ⁻²	1,00x10 ⁻³	1,08	9,00x10 ⁻³	0,150	0,270	1,40x10 ⁻²
AV030E	0,250	3,50x10 ⁻²	1,50x10 ⁻³	0,100	8,00x10 ⁻³	0,150	0,120	4,00x10 ⁻³
AV030F	0,250	4,50x10 ⁻²	1,00x10 ⁻³	0,290	1,40x10 ⁻²	0,100	0,120	4,00x10 ⁻³
AV040	0,250	4,50x10 ⁻²	1,00x10 ⁻³	0,180	4,00x10 ⁻³	0,150	0,110	4,50x10 ⁻³
AV050	0,250	3,50x10 ⁻²	1,00x10 ⁻³	0,150	7,00x10 ⁻³	0,100	6,00x10 ⁻²	2,00x10 ⁻³
AV060	0,250	2,00x10 ⁻²	1,00x10 ⁻³	0,110	4,00x10 ⁻³	0,100	5,00x10 ⁻²	5,00x10 ⁻³
AV070	0,250	3,00x10 ⁻²	1,00x10 ⁻³	0,150	2,50x10 ⁻²	0,100	5,00x10 ⁻²	2,00x10 ⁻³
AV080	0,250	5,00x10 ⁻²	1,00x10 ⁻³	0,460	5,00x10 ⁻³	0,100	5,00x10 ⁻²	2,00x10 ⁻³
AV090	0,250	2,00x10 ⁻²	1,00x10 ⁻³	8,00x10 ⁻²	4,00x10 ⁻³	0,100	9,00x10 ⁻²	1,50x10 ⁻³
AV100	0,250	2,50x10 ⁻²	1,00x10 ⁻³	6,00x10 ⁻²	4,00x10 ⁻³	0,100	9,00x10 ⁻²	2,00x10 ⁻³
AV110	0,250	0,125	1,00x10 ⁻³	0,330	4,00x10 ⁻³	0,300	0,150	1,75x10 ⁻²
AV120	0,250	3,50x10 ⁻²	1,00x10 ⁻³	9,00x10 ⁻²	4,00x10 ⁻³	0,100	0,170	2,00x10 ⁻³
AV130	0,250	5,50x10 ⁻²	1,00x10 ⁻³	0,180	4,00x10 ⁻³	0,100	0,140	7,50x10 ⁻³
AV140	0,250	2,00x10 ⁻²	1,00x10 ⁻³	0,110	4,00x10 ⁻³	0,100	0,120	2,00x10 ⁻³
AV150	0,250	2,00x10 ⁻²	1,00x10 ⁻³	0,260	4,00x10 ⁻³	0,100	0,190	4,00x10 ⁻³
AV160E	0,250	1,50x10 ⁻²	1,00x10 ⁻³	2,00x10 ⁻²	4,00x10 ⁻³	0,100	3,00x10 ⁻²	2,00x10 ⁻³
AV160F	0,250	2,00x10 ⁻²	1,50x10 ⁻³	2,00x10 ⁻²	4,00x10 ⁻³	0,100	2,00x10 ⁻²	4,00x10 ⁻³
AV170	0,250	2,50x10 ⁻²	1,00x10 ⁻³	0,200	4,00x10 ⁻³	0,100	6,00x10 ⁻²	3,50x10 ⁻³
AV180E	0,250	1,50x10 ⁻²	1,50x10 ⁻³	4,00x10 ⁻²	4,00x10 ⁻³	0,100	0,110	2,50x10 ⁻³
AV180F	0,250	2,00x10 ⁻²	1,00x10 ⁻³	5,00x10 ⁻²	4,00x10 ⁻³	0,100	0,120	3,00x10 ⁻³
AV190	0,250	2,50x10 ⁻²	1,00x10 ⁻³	0,120	4,00x10 ⁻³	0,100	0,130	2,00x10 ⁻³
AV200	0,250	2,50x10 ⁻²	1,00x10 ⁻³	4,00x10 ⁻²	4,00x10 ⁻³	0,100	0,100	2,00x10 ⁻³
AV210	0,250	5,50x10 ⁻²	1,00x10 ⁻³	0,280	4,00x10 ⁻³	0,100	0,150	8,00x10 ⁻³
AV220	0,250	2,00x10 ⁻²	1,00x10 ⁻³	3,00x10 ⁻²	4,00x10 ⁻³	0,100	7,00x10 ⁻²	2,00x10 ⁻³
AV230	0,250	2,00x10 ⁻²	1,00x10 ⁻³	4,00x10 ⁻²	4,00x10 ⁻³	0,100	7,00x10 ⁻²	2,00x10 ⁻³
AV240	0,250	2,00x10 ⁻²	1,00x10 ⁻³	1,00x10 ⁻²	4,00x10 ⁻³	0,100	0,150	1,00x10 ⁻³
AV250	0,250	3,00x10 ⁻²	1,00x10 ⁻³	0,120	5,00x10 ⁻³	0,100	0,160	2,50x10 ⁻³
AV260	0,250	6,50x10 ⁻²	1,00x10 ⁻³	0,320	4,00x10 ⁻³	0,250	0,130	8,50x10 ⁻³
AV270	0,250	7,00x10 ⁻²	1,00x10 ⁻³	0,370	6,00x10 ⁻³	0,100	0,130	8,50x10 ⁻³
AV280	0,250	2,00x10 ⁻²	1,00x10 ⁻³	2,00x10 ⁻²	4,00x10 ⁻³	0,100	5,00x10 ⁻²	2,00x10 ⁻³
AV290	0,250	2,00x10 ⁻²	1,00x10 ⁻³	2,00x10 ⁻²	4,00x10 ⁻³	0,100	0,150	2,00x10 ⁻³
AV300	0,250	1,75x10	1,00x10 ⁻³	0,230	4,00x10 ⁻³	0,800	0,380	5,05x10 ⁻²
AV310	0,250	4,05x10	2,50x10 ⁻³	0,420	4,00x10 ⁻³	1,05	0,390	6,35x10 ⁻²
AV320	0,330	1,00x10 ⁻²	1,00x10 ⁻³	1,24	5,20x10 ⁻²	2,00	0,790	0,400
AV330	0,250	1,20x10	2,00x10 ⁻³	0,540	1,00x10 ⁻²	0,300	0,220	2,00x10 ⁻²
AV340	0,250	4,50x10 ⁻²	1,00x10 ⁻³	6,00x10 ⁻²	4,00x10 ⁻³	0,200	0,180	6,00x10 ⁻³
AV345	0,250	2,00x10 ⁻²	2,00x10 ⁻³	3,00x10 ⁻²	4,00x10 ⁻³	0,100	7,00x10 ⁻²	2,00x10 ⁻³
AV350	0,250	8,50x10 ⁻²	1,00x10 ⁻³	0,320	6,00x10 ⁻³	0,200	0,220	2,00x10 ⁻²

TABELA 7 - CONTINUAÇÃO

ESTAÇÕES	NITROGÊNIO ORGÂNICO (mg/L)	OXIGÊNIO DISSOLVIDO (mg/L)	pH IN LOCO	pH, LAB	POTÁSSIO SOLÚVEL (mg/L)	SELÊNIO TOTAL (mg/L)	SÓDIO SOLÚVEL (mg/L)	SÓLIDOS DISSOLVIDOS (mg/L)
AV005	0,150	7,80	6,80	6,85	0,470	5,00x10 ⁻⁴	1,01	23,0
AV010	0,200	7,50	7,10	7,25	0,590	5,00x10 ⁻⁴	1,28	35,0
AV020	0,250	7,00	6,80	7,00	1,39	5,00x10 ⁻⁴	3,19	42,5
AV030E	0,200	6,95	7,20	7,15	0,910	5,00x10 ⁻⁴	1,77	33,0
AV030F	0,300	6,35	7,20	7,10	0,920	5,00x10 ⁻⁴	1,79	36,0
AV040	0,250	7,55	6,90	7,40	0,780	5,00x10 ⁻⁴	1,62	33,5
AV050	0,350	7,50	7,40	7,40	0,290	5,00x10 ⁻⁴	0,360	55,5
AV060	0,300	7,75	6,60	6,65	0,820	5,00x10 ⁻⁴	1,95	26,0
AV070	0,300	7,60	7,20	7,25	0,380	5,00x10 ⁻⁴	0,590	48,5
AV080	0,200	7,70	7,00	7,30	0,530	5,00x10 ⁻⁴	1,11	45,0
AV090	0,200	7,85	6,50	6,85	0,520	5,00x10 ⁻⁴	0,750	23,0
AV100	0,300	7,65	6,70	7,00	0,550	5,00x10 ⁻⁴	0,530	23,0
AV110	0,250	7,50	7,10	7,20	0,950	5,00x10 ⁻⁴	2,58	52,5
AV120	0,300	7,80	6,80	7,10	0,490	5,00x10 ⁻⁴	0,870	23,0
AV130	0,300	8,00	7,20	7,20	0,810	5,00x10 ⁻⁴	1,95	43,5
AV140	0,250	7,80	7,15	7,20	0,510	5,00x10 ⁻⁴	0,620	35,5
AV150	0,250	7,30	6,90	7,20	0,320	5,00x10 ⁻⁴	3,44	43,5
AV160E	0,200	7,05	7,30	7,25	0,300	5,00x10 ⁻⁴	0,470	21,5
AV160F	0,300	6,35	7,20	7,30	0,360	5,00x10 ⁻⁴	0,420	21,0
AV170	0,200	7,00	6,90	6,95	0,300	5,00x10 ⁻⁴	0,420	22,0
AV180E	0,250	6,85	7,30	7,20	0,420	5,00x10 ⁻⁴	4,99	32,5
AV180F	0,300	6,90	7,00	7,00	0,430	5,00x10 ⁻⁴	4,98	35,5
AV190	0,200	7,50	7,60	7,40	0,310	5,00x10 ⁻⁴	2,65	35,0
AV200	0,200	7,80	7,50	7,55	0,210	5,00x10 ⁻⁴	1,01	24,0
AV210	0,250	7,40	7,50	7,45	0,710	5,00x10 ⁻⁴	2,15	39,0
AV220	0,200	7,40	7,20	7,20	0,110	5,00x10 ⁻⁴	0,300	39,0
AV230	0,200	7,40	7,40	7,40	0,150	5,00x10 ⁻⁴	0,940	38,0
AV240	0,200	7,50	6,10	6,40	0,100	5,00x10 ⁻⁴	0,270	12,0
AV250	0,250	7,65	7,60	7,55	0,430	5,00x10 ⁻⁴	1,07	32,5
AV260	0,350	7,50	7,40	7,40	0,690	5,00x10 ⁻⁴	2,07	37,5
AV270	0,300	7,60	7,50	7,50	0,700	5,00x10 ⁻⁴	2,16	39,0
AV280	0,200	7,90	6,50	7,00	0,100	5,00x10 ⁻⁴	0,290	14,0
AV290	0,200	7,30	6,50	6,80	0,140	5,00x10 ⁻⁴	0,290	17,5
AV300	0,400	6,35	7,15	7,20	0,950	5,00x10 ⁻⁴	3,10	70,5
AV310	0,450	6,30	7,30	7,30	1,26	5,00x10 ⁻⁴	5,21	97,0
AV320	0,550	6,40	7,30	7,30	4,28	9,40x10 ⁻³	19,1	1,62x10 ³
AV330	0,300	7,40	7,10	7,10	0,770	5,00x10 ⁻⁴	2,55	104
AV340	0,300	7,40	6,92	7,00	0,310	5,00x10 ⁻⁴	1,39	26,0
AV345	0,300	7,90	7,20	7,20	0,240	5,00x10 ⁻⁴	0,830	20,0
AV350	0,300	7,10	7,15	7,30	0,710	5,00x10 ⁻⁴	2,40	57,0

TABELA 7 - CONTINUAÇÃO

ESTAÇÕES	SÓLIDOS EM SUSPENSÃO (mg/L)	SULFATOS (mg/L)	SURFACTANTES ANIÔNICOS (mg/L LAS)	TEMPERATURA DA ÁGUA (°C)	TEMPERATURA DO AR (°C)	TURBIDEZ (NTU)	ZINCO TOTAL (mg/L)
AV005	9,50	1,00	5,00x10 ⁻²	20,5	23,5	3,88	2,00x10 ⁻²
AV010	14,5	1,30	5,00x10 ⁻²	22,4	27,0	23,6	2,00x10 ⁻²
AV020	15,5	1,25	5,00x10 ⁻²	22,6	27,5	14,9	2,50x10 ⁻²
AV030E	11,5	1,30	5,00x10 ⁻²	24,7	27,0	13,7	2,50x10 ⁻²
AV030F	17,5	1,60	5,00x10 ⁻²	22,2	27,0	27,7	3,00x10 ⁻²
AV040	12,0	1,10	5,00x10 ⁻²	22,7	21,5	21,7	3,50x10 ⁻²
AV050	13,0	6,30	5,00x10 ⁻²	21,3	23,0	8,26	2,00x10 ⁻²
AV060	4,00	1,00	5,00x10 ⁻²	22,1	25,0	6,51	3,00x10 ⁻²
AV070	24,5	4,75	5,00x10 ⁻²	21,8	25,0	25,0	6,50x10 ⁻²
AV080	57,5	2,60	5,00x10 ⁻²	22,4	26,5	45,3	5,00x10 ⁻²
AV090	7,50	3,10	5,00x10 ⁻²	20,5	24,0	3,06	2,00x10 ⁻²
AV100	6,00	2,30	5,00x10 ⁻²	21,6	24,0	4,12	2,00x10 ⁻²
AV110	33,5	3,30	5,00x10 ⁻²	22,1	23,5	22,2	4,00x10 ⁻²
AV120	10,5	3,30	5,00x10 ⁻²	21,6	23,5	11,1	3,50x10 ⁻²
AV130	17,5	2,65	5,00x10 ⁻²	22,5	23,5	14,0	4,00x10 ⁻²
AV140	7,50	2,20	5,00x10 ⁻²	21,1	25,0	7,16	2,50x10 ⁻²
AV150	8,00	4,75	5,00x10 ⁻²	20,7	22,0	9,70	2,00x10 ⁻²
AV160E	2,50	1,50	5,00x10 ⁻²	22,4	23,0	1,80	2,00x10 ⁻²
AV160F	3,50	1,35	5,00x10 ⁻²	21,1	23,0	1,63	2,00x10 ⁻²
AV170	7,50	1,00	5,00x10 ⁻²	20,5	25,5	5,06	2,00x10 ⁻²
AV180E	4,00	7,40	5,00x10 ⁻²	22,2	25,0	3,95	2,00x10 ⁻²
AV180F	4,00	6,95	5,00x10 ⁻²	21,1	25,0	4,45	3,50x10 ⁻²
AV190	8,50	3,75	5,00x10 ⁻²	20,8	21,0	6,75	2,00x10 ⁻²
AV200	6,50	1,80	5,00x10 ⁻²	22,0	24,0	5,44	2,00x10 ⁻²
AV210	61,5	2,40	5,00x10 ⁻²	21,8	22,0	100	2,00x10 ⁻²
AV220	8,00	2,00	5,00x10 ⁻²	20,8	21,0	1,54	2,00x10 ⁻²
AV230	8,00	2,40	5,00x10 ⁻²	20,6	21,0	2,09	2,00x10 ⁻²
AV240	4,00	1,00	5,00x10 ⁻²	20,7	22,0	1,29	2,00x10 ⁻²
AV250	20,00	1,65	5,00x10 ⁻²	21,6	26,0	38,2	2,00x10 ⁻²
AV260	79,0	2,75	5,00x10 ⁻²	22,5	24,5	81,5	2,00x10 ⁻²
AV270	46,0	3,00	5,00x10 ⁻²	21,7	25,5	75,5	2,00x10 ⁻²
AV280	2,00	1,00	5,00x10 ⁻²	18,8	21,0	1,55	2,00x10 ⁻²
AV290	4,00	1,00	5,00x10 ⁻²	20,6	21,0	1,24	2,00x10 ⁻²
AV300	10,0	3,70	5,00x10 ⁻²	22,0	24,0	5,70	2,00x10 ⁻²
AV310	37,0	12,6	6,00x10 ⁻²	22,3	25,0	32,7	5,50x10 ⁻²
AV320	48,5	920	5,00x10 ⁻²	24,1	25,5	5,81	1,00x10 ⁻¹
AV330	49,0	14,0	5,00x10 ⁻²	23,3	27,0	62,9	5,00x10 ⁻²
AV340	10,5	2,05	5,00x10 ⁻²	22,9	26,0	5,03	2,00x10 ⁻²
AV345	3,00	1,30	5,00x10 ⁻²	21,9	26,0	4,44	2,00x10 ⁻²
AV350	50,5	11,7	5,00x10 ⁻²	22,6	24,5	65,9	2,00x10 ⁻²

Entre os parâmetros analisados no programa de monitoramento da sub-bacia do rio das Velhas, alguns violaram em 20% os limites legais da DN COPAM 10/86, no período de 1997 a 2003. A TABELA 8 apresenta o percentual de violações em ordem decrescente do valor obtido para cada parâmetro. ⁸

TABELA 8 – Classificação dos parâmetros monitorados em ordem decrescente de percentual de violação de classe de enquadramento em toda bacia do rio das Velhas, no período de 1997 a 2003⁸

PARÂMETRO	VIOLAÇÕES (%)	PARÂMETRO	VIOLAÇÕES (%)
Fosfato total	83,5	Cobre	11,3
Coliformes fecais	64,9	Níquel	10,3
Coliformes totais	64,4	Ferro solúvel	7,7
Manganês	53,0	Zinco total	5,9
Índice de fenóis	51,5	Mercúrio	5,5
Óleos de graxas	38,4	Chumbo	3,3
DBO	31,9	Cádmio	2,1
Turbidez	29,7	Cromo III	0,8
Oxigênio dissolvido	21,1		
Nitrogênio amoniacal	19,6		
Cor	16,6		
Substâncias tensoativas	12,2		
Arsênio	12,0		

Os parâmetros que apresentaram os maiores percentuais de violação em toda bacia do rio das Velhas estão associados principalmente a efluentes de esgotos domésticos, tais como fosfato total, coliformes fecais, índice de fenóis, DBO, oxigênio dissolvido, nitrogênio amoniacal; e efluentes de mineração, como manganês, turbidez, arsênio, cobre e níquel.

Capítulo V – Análise de regressão

Este estudo teve como finalidade a modelagem da qualidade das águas na sub-bacia do rio das Velhas visando identificar os principais poluentes e as regiões mais impactadas da macro-rede de monitoramento do rio das Velhas, empregando análise fatorial e de variância. Dentro desse objetivo, a análise fatorial foi utilizada para descrever a variabilidade dos dados originais em termos fatores. Aos fatores mais importantes foi realizada a análise de variâncias em relação à latitude das estações de amostragem para identificação das estações mais críticas. Às estações identificadas pela análise de variâncias foram propostos modelos de regressão não linear entre os principais parâmetros presentes nos fatores.

5.1 Descrição do conjunto de dados

A utilização da estatística multivariada se justifica porque a extração de informações a partir de um banco de dados envolve a análise de grande número de parâmetros. Os dados utilizados fazem parte do programa de monitoramento do IGAM¹⁻⁷ no período compreendido entre janeiro de 1997 a julho de 2004.

A matriz de dados utilizada nessa análise é composta pelas 29 estações de amostragem que compõem a macro-rede de monitoramento da sub-bacia do rio das Velhas. Para cada estação foram avaliados 50 parâmetros, listados na TABELA 6 do capítulo IV.

Foram empregados os dados gerados em campanhas realizadas nos meses de janeiro e julho para análise fatorial, pois apenas nesses meses foram realizadas medidas de todos os 50 parâmetros, nas chamadas **campanhas completas**. As análises de todos os parâmetros nesses meses são feitas para caracterizar os períodos de chuva e de estiagem.

Entre esses parâmetros avaliados estão: condutividade elétrica, oxigênio dissolvido, pH *in loco*, temperatura da água, temperatura do ar, alcalinidade de bicarbonato, alcalinidade total, arsênio total, bário total, boro solúvel, cádmio total, cálcio total, chumbo total, cianeto total, cloretos, cobre total, coliformes fecais, coliformes totais, cor real, cromo hexavalente, cromo trivalente, demanda bioquímica de oxigênio, demanda química de oxigênio, dureza de cálcio, dureza de magnésio, dureza total, estreptococos fecais, ferro solúvel, fosfato total, índice de fenóis, magnésio total, manganês total, mercúrio total, níquel total, nitrogênio amoniacal, nitrogênio nítrico, nitrogênio nitroso, nitrogênio orgânico, óleos e graxas, potássio solúvel, selênio total, sódio solúvel, sólidos dissolvidos, sólidos suspensos sólidos totais, sulfatos, sulfetos, surfactantes aniônicos, turbidez e zinco total.¹⁻⁸ As medidas de pH realizadas em laboratório não foram incluídas na base de dados analisada por não terem sido realizadas medidas em algumas estações. A latitude e longitude de cada estação de amostragem também foram consideradas.

Devido às diferenças nas escalas entre as medidas, em todo conjunto de dados foi primeiramente aplicado o logaritmo natural. Sendo cada elemento, ${}^*x_{ik}$, da matriz X analisada igual ao logaritmo da concentração do parâmetro k:

$${}^*x_{ik} = \ln x_{ik} \quad (97)$$

5.2 Análise fatorial

Inicialmente foi utilizada análise fatorial pelo método das componentes principais visando à redução de parâmetros e que de algum modo sumarizassem as informações principais das variáveis originais em fatores comuns. Nesse modelo, parte da variabilidade de X é atribuída aos fatores comuns, sendo o restante da variabilidade de X atribuído às variáveis que não foram incluídas no modelo, ou seja, ao erro aleatório.²⁹

O método de Cattell^{29,34} foi utilizado para identificação dos principais fatores, FIGURA 7. Pode-se observar pela figura que a partir do sexto fator os autovalores tendem a valores constantes e pouco significativos.

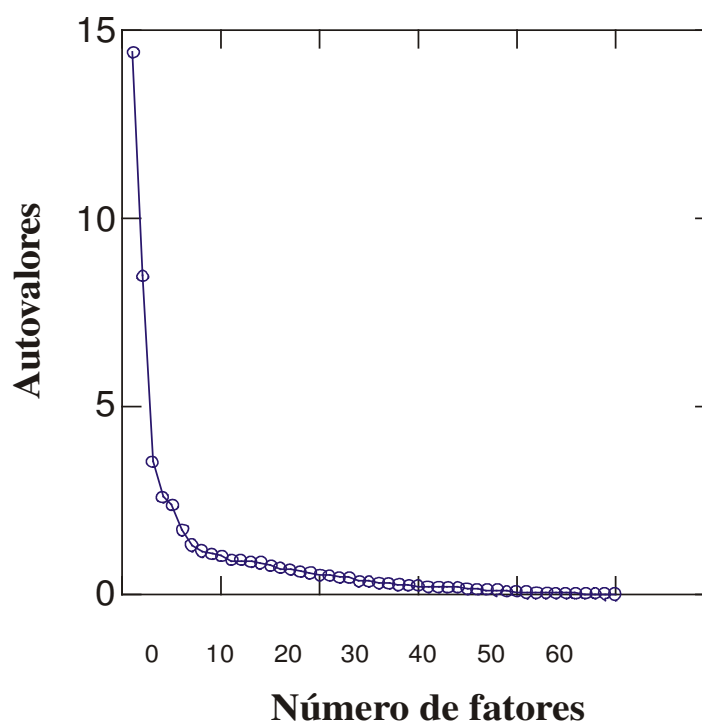


FIGURA 7– Importância dos principais fatores avaliada pelo método de Cattell para a base de dados do monitoramento do rio das Velhas realizado no período de 1997 a 2004 em campanhas completas

A TABELA 9 apresenta a composição e carga dos parâmetros para os fatores comuns e seus respectivos autovalores.

TABELA 9 – Autovalores e cargas dos parâmetros nos fatores comuns

PARÂMETROS	FATOR 1	FATOR 2	FATOR 3	FATOR 4	FATOR 5	FATOR 6
Proporção da variância total: 67,2%	AUTOVALOR (14,4) Prop. var*: 29,4%	AUTOVALOR (8,48) Prop. var: 17,2%	AUTOVALOR (3,53) Prop. var: 7,16%	AUTOVALOR (2,59) Prop. var: 5,25%	AUTOVALOR (2,37) Prop. var: 4,77%	AUTOVALOR (1,72) Prop. var: 3,46%
Alcanidade bicarbonato	0,867	-0,371	0,123	-0,132	-0,0190	0,0740
Alcalinidade total	0,864	-0,377	0,126	-0,135	-0,0240	0,0810
Boro	-0,0710	-0,105	-0,253	0,337	-0,814	-0,0580
Cálcio	0,801	-0,378	0,322	-0,118	-0,0260	0,0340
Cloretos	0,876	-0,213	0,0780	-0,0850	-0,0840	0,100
Cobre	0,196	0,816	0,244	0,0480	0,100	0,198
Condutividade elétrica	0,910	-0,347	0,144	-0,0380	0,0250	0,0430
Cor	-0,0310	0,452	-0,115	0,0130	-0,703	0,108
Cromo trivalente	-0,0390	0,623	0,343	0,207	0,261	0,314
DQO	0,694	0,410	-0,235	-0,0900	-0,116	0,106
DBO	0,830	0,168	-0,308	-0,0910	0,0720	0,0150
Dureza de cálcio	0,814	-0,379	0,326	-0,120	-0,0150	0,0220
Dureza de magnésio	0,658	-0,237	-0,104	-0,524	0,0350	0,138
Dureza total	0,854	-0,386	0,260	0,00100	0,0140	0,0360
Estreptococos fecais	0,453	0,531	-0,361	0,0630	0,188	-0,330
Fosfatos totais	0,626	0,231	-0,223	0,0590	0,0600	-0,0510
Índice de fenóis	0,597	0,0810	-0,284	-0,224	-0,0280	0,107
Magnésio	0,659	-0,240	-0,114	-0,508	0,0400	0,122
Manganês	0,257	0,804	0,0900	0,214	0,0940	-0,102
Níquel	0,0910	0,678	0,252	0,297	0,0960	0,204
Nitrogênio amoniacal	0,765	0,117	-0,364	-0,0610	0,107	-0,0200
Nitrogênio orgânico	0,669	0,236	-0,121	-0,203	0,00700	0,149
Nitratos	0,0810	-0,245	0,510	0,302	-0,333	-0,270
Nitritos	0,613	-0,241	0,209	0,303	0,0130	-0,467
Oxigênio dissolvido	-0,709	-0,162	0,345	0,235	0,064	-0,159
Potássio	0,908	-0,134	0,0820	-0,184	-0,155	0,027
Sódio	0,897	-0,339	0,0780	-0,084	-0,019	0,023
Sólidos dissolvidos	0,859	-0,0390	0,177	0,017	-0,128	0,061
Sólidos susp.	0,174	0,826	0,175	-0,016	-0,243	-0,172
Sólidos totais	0,388	0,749	0,347	0,0760	-0,191	-0,0210
Sulfatos	0,738	-0,130	0,0530	0,360	0,0550	-0,0630
Surfactantes	0,538	0,0730	-0,405	-0,0380	-0,205	0,176
Turbidez	0,0820	0,876	0,208	0,00300	-0,221	-0,129

* Prop. var* = proporção da variância

Identificamos dessa maneira, que os parâmetros analisados podem ser reduzidos a seis fatores, que explicam 67,2% da variância total dos dados. A porcentagem da variância total explicada pelos fatores é de 29,4%, para o primeiro fator, 17,2%, para o segundo, 7,16%, para o terceiro, 5,25% para o quarto, 4,77% para o quinto e 3,46 % para o sexto 3,4%.

Foram considerados como mais significativos os parâmetros com componentes de carga com valores absolutos superiores a 0,600 nos fatores. Utilizando esse critério, os parâmetros mais importantes no primeiro fator foram: alcalinidade bicarbonato, alcalinidade total, cálcio, cloretos, DQO, dureza de cálcio, dureza de magnésio, dureza total, fosfatos, magnésio, nitritos, nitrogênio amoniacal, nitrogênio orgânico, potássio, sódio, condutividade elétrica, oxigênio dissolvido, DBO, sólidos dissolvidos e sulfatos. A presença desses parâmetros é característica de efluentes de esgotos domésticos e industriais.

No segundo fator foram considerados os parâmetros cobre, cromo trivalente, níquel, manganês, sólidos em suspensão, sólidos totais e turbidez como os mais importantes. Os parâmetros presentes no segundo fator são característicos de contaminação de efluentes de mineração. Os terceiro, quarto e sexto fatores não apresentaram componentes de carga com valores absolutos superior a 0,600 e o quinto apresentou boro e cor como componentes com carga superior a 0,600.

5.3 Análise de variância aplicada ao fator 1

A análise de variância (ANOVA) é um método de se testar a igualdade de três ou mais médias populacionais através da análise das variâncias amostrais. Nesta análise é realizado um teste hipótese onde a hipótese nula considera médias iguais para todas as amostras testadas. A probabilidade dessa hipótese estar correta é avaliada pelo valor de **probabilidade** ou **valor p** dentro de um nível de significância estabelecido.⁵⁶ Neste trabalho foi utilizando um nível de significância de 0,05 para os testes; assim, valores de p iguais ou menores que 0,05 rejeitam a probabilidade de igualdade entre as amostras.

A ANOVA⁹⁵ aplicada ao fator 1 e as latitudes das estações ao longo da sub-bacia do rio das Velhas, demonstra haver diferenças em relação parâmetros que compõem esse fator e as estações com $p = 0$, FIGURA 8.

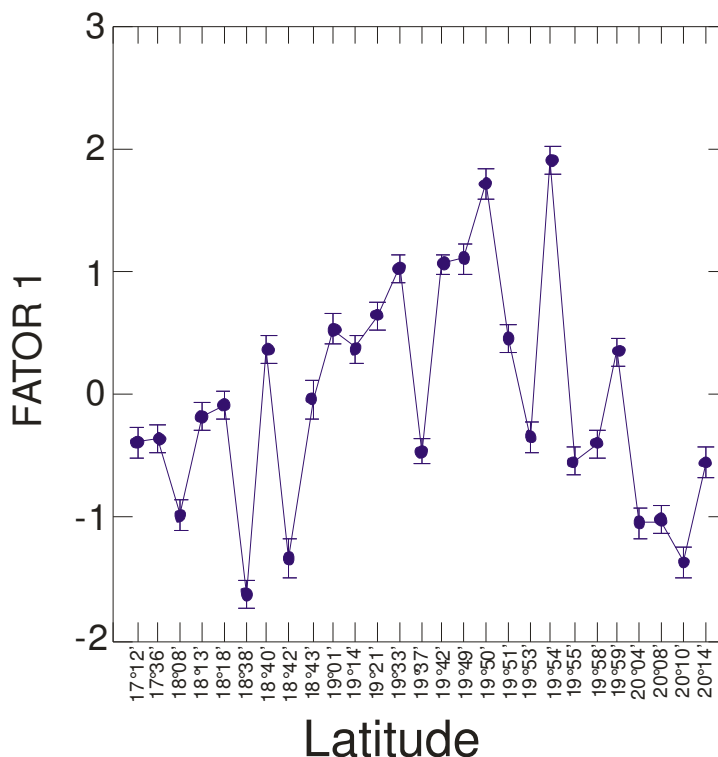


FIGURA 8 – ANOVA aplicada ao fator 1 em relação a latitude das estações de amostragem localizadas na sub-bacia do rio das Velhas

As maiores diferenças em relação ao fator 1 e a latitude das estações são observadas nas latitudes 19°50' e 19°54' correspondentes às estações BV154, ribeirão do Onça a montante do rio das Velhas e BV155, ribeirão Arrudas a montante do rio das Velhas, respectivamente. A localização dos máximos do fator em relação à latitude das estações sugere impacto de efluentes de esgotos sanitários e industriais na qualidade das águas principalmente nas estações localizadas na região metropolitana de Belo Horizonte. Para obtenção do modelo foram escolhidas regiões próximas a esses máximos, compreendidas entre as latitudes 19°21' a 19°59' correspondendo às estações BV062, BV063, BV067, BV076, BV083, BV105, BV130, BV135, BV137, BV153, BV154, BV155, BV156 e BV160 pertencentes a região metropolitana de Belo Horizonte, compreendida entre os municípios de Nova Lima até Jaboticatubas, TABELA 2.

5.4 Regressão não linear - fator 1

Os parâmetros escolhidos para obtenção das equações de regressão foram os que apresentaram carga relevante no fator 1 ou alto percentual de violação de classe de enquadramento no período de 1997 a 2004, TABELA 8.

A TABELA 10 apresenta as equações de regressão propostas para os parâmetros cloretos, DBO, fosfatos e DQO.

TABELA 10 – Equações de regressão para os parâmetros de maior carga no fator 1 para as estações BV062, BV063, BV067, BV076, BV083, BV105, BV130, BV135, BV 137, BV153, BV154, BV155, BV156 e BV160

PARÂMETROS	EQUAÇÕES DE REGRESSÃO
Cloretos	$\ln[\text{Cl}]^* = -76,2 + 0,359 \cdot \ln[\text{condutividade}] +$ $0,0690 \cdot \ln[\text{DQO}] + 0,0670 \cdot \ln[\text{DBO}] + 0,0310 \cdot \ln[\text{PO}_4^{-3}] +$ $0,329 \cdot \ln[\text{K}] + 0,496 \cdot \ln[\text{Na}] - 0,109 \cdot \ln[\text{Zn}] + 0,346 \cdot \text{Latitude} +$ $1,54 \cdot \text{Longitude}$ $r = 0,910$
DBO	$\ln[\text{DBO}] = -0,411 + 0,475 \cdot \ln[\text{DQO}] + 0,698 \cdot \ln[\text{Na}] - 0,202 \cdot \ln[\text{alcalinidade}$ $\text{total}] + -0,156 \cdot \ln[\text{O}_2 \text{ dissolvido}] + 0,140 \cdot \ln[\text{fenóis}] + 0,697 \cdot \text{Latitude} +$ $-0,283 \cdot \text{Longitude}$ $r = 0,870$
Fosfatos	$\ln[\text{PO}_4^{-3}] = -5,34 - 2,10 \cdot \ln[\text{dureza Ca}] - 0,398 \cdot \ln[\text{K}] + 1,86 \cdot \ln[\text{condutividade}]$ $+ 0,552 \cdot \ln[\text{alcalinidade total}] + 0,198 \cdot \ln[\text{DBO}] + 0,400 \cdot \ln[\text{sólidos totais}]$ $+ 0,297 \cdot \ln[\text{Cl}] + 2,00 \times 10^3 \cdot \text{Latitude} +$ $-0,0680 \cdot \text{Longitude}$ $r = 0,661$
DQO	$\ln[\text{DQO}] = -28,7 + 0,514 \cdot \ln[\text{DBO}] - 0,453 \cdot \ln[\text{Na}]$ $+ 0,121 \cdot \ln[\text{surfactantes}] + 0,319 \cdot \ln[\text{nitrogênio orgânico}] +$ $0,461 \cdot \ln[\text{sólidos dissolvidos}] + 0,191 \cdot \ln[\text{Cl}] + -0,129 \cdot \ln[\text{dureza total}] +$ $-0,232 \cdot \text{Latitude} + 0,795 \cdot \text{Longitude}$ $r = 0,825$

*As concentrações dos parâmetros correspondem a valores totais

A FIGURA 9 apresenta os valores reais e estimados para o modelo proposto para o parâmetro fosfato no período de 1997 a 2004.

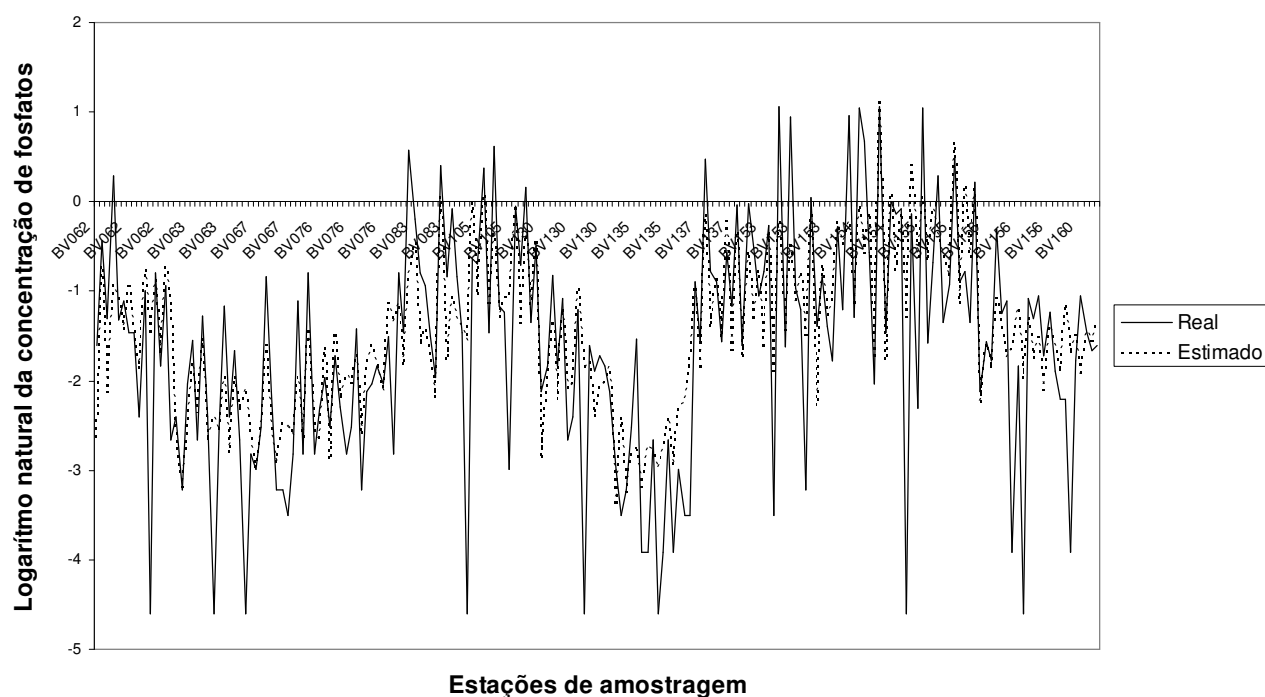


FIGURA 9– Valores reais e estimados para $\ln[\text{fosfatos}]$ no período de 1997 a 2004 para as estações compreendidas entre as latitudes $19^{\circ}21'$ a $19^{\circ}59'$ na sub-bacia do rio das Velhas

Avaliando a FIGURA 9 pode-se observar que o principal fator que contribui para coeficiente de correlação, $r = 0,661$, são os valores entre $-4,6$ a $-3,0$. Como o gráfico representa o logaritmo natural da concentração de fosfatos podemos dizer que a concentração real nesses pontos está entre $0,01$ a $0,05$ mg/L de fosfatos. Esta faixa de concentração é baixa e próxima do limite de detecção do método, como as dosagens próximas aos limites de detecção são mais sujeitas a erros, esse fator também pode ter influenciado o valor de r .

A validade dos modelos foi testada aplicando a equação de regressão aos resultados de monitoramento no ano de 2005. A FIGURA 10 apresenta uma comparação entre os valores reais e estimados para a concentração de fosfatos no ano de 2005.

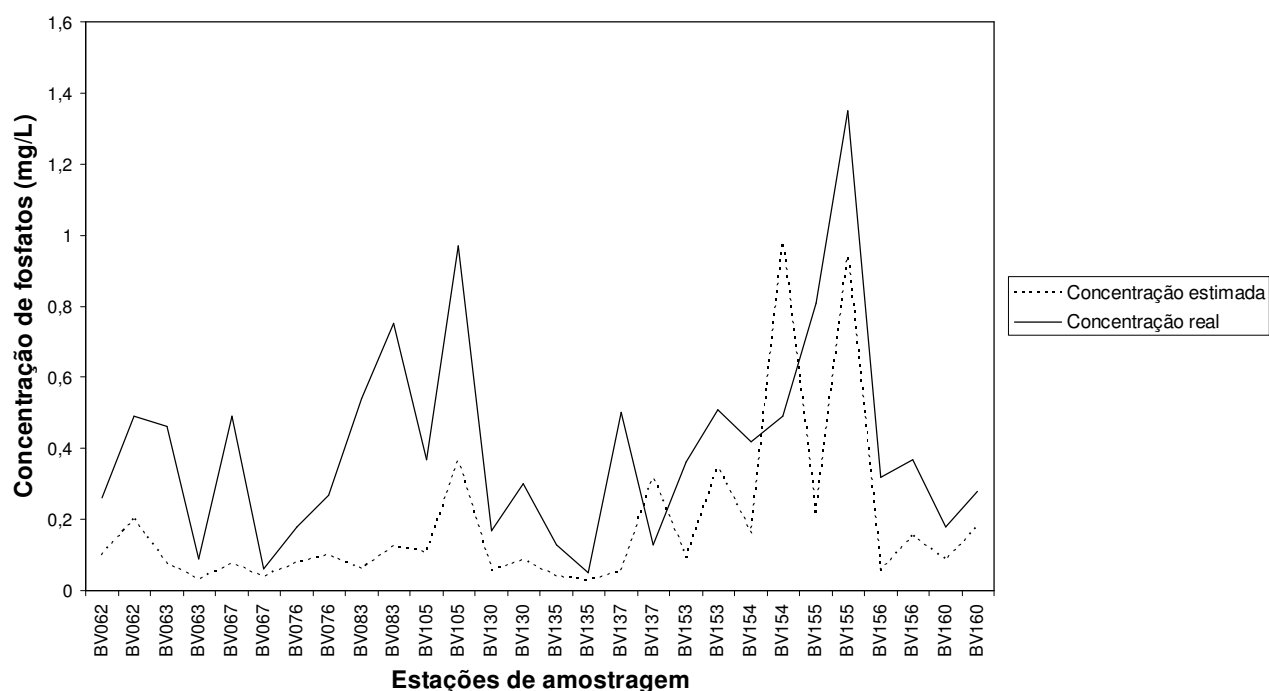


FIGURA 10 – Concentração de fosfatos totais real e estimada para as estações compreendidas entre as latitudes 19°21' a 19°59' na sub-bacia do rio das Velhas

Embora as tendências de crescimento e decréscimo da concentração de fosfatos tenham sido reproduzidas pela estimativa na maior parte das estações, não há uma boa aproximação aos valores reais. Esta dificuldade está associada a variações irregulares na concentração de fosfatos e dos outros parâmetros utilizados na sua estimativa. O uso dos dados das campanhas realizadas nos meses de janeiro e julho, ou seja, chuva e estiagem, apenas pode não ter acompanhado completamente as variações na concentração de fosfatos para essas estações.

5.5 Análise de variância aplicada ao fator 2

A ANOVA⁹⁵ aplicada ao fator 2 e as latitudes das estações ao longo da sub-bacia do rio das Velhas, demonstra haver diferenças entre os parâmetros que compõem esse fator e as estações ao longo da sub-bacia ($p = 0$), FIGURA 11.

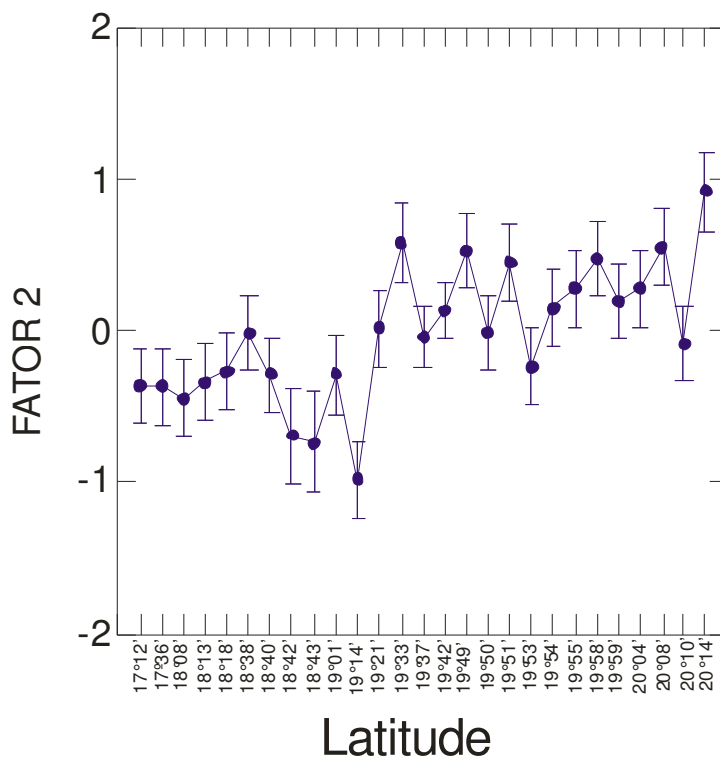


FIGURA 11- ANOVA aplicada ao fator 2 em relação a latitude das estações de amostragem localizadas na sub-bacia do rio das Velhas

As maiores diferenças em relação ao fator 2 e a latitude das estações são observadas nas latitudes $19^{\circ}33'$ e $20^{\circ}14'$ correspondentes às estações BV137, rio das Velhas na ponte Raul Soares no município de Lagoa Santa e BV035, rio Itabirito a jusante da cidade de Itabirito, respectivamente. A localização dos máximos do fator em relação à latitude das estações sugere impacto de efluentes de mineração, erosão e assoreamento para as estações compreendidas entre as latitudes $19^{\circ}33'$ e $20^{\circ}14'$. As estações de amostragem que correspondem a essa faixa de latitude são: BV013, BV035, BV037, BV062, BV063, BV067, BV076, BV083, BV105, BV130, BV135, BV137, BV139 e BV153, TABELA 2.

5.6 Regressão não linear - fator 2

Os parâmetros escolhidos para obtenção das equações de regressão foram os que apresentaram maior carga no fator 2.

A TABELA 11 apresenta as equações de regressão propostas para os parâmetros turbidez, manganês, cobre e sólidos em suspensão.

TABELA 11 – Equações de regressão para os parâmetros de maior carga no fator 2 para as estações BV013, BV035, BV037, BV062, BV063, BV067, BV076, BV083, BV105, BV130, BV135, BV137, BV139 e BV153

PARÂMETROS	EQUAÇÕES DE REGRESSÃO
Turbidez	$\ln[\text{turbidez}] = -1,95 + 0,535 \cdot \ln[\text{sólidos totais}] + 0,543 \cdot \ln[\text{sólidos em suspensão}] + 0,0480 \cdot \ln[\text{coliformes totais}] - 0,435 \cdot \ln[\text{condutividade elétrica}] + 0,154 \cdot \ln[\text{Mn}] - 0,230 \cdot \text{Latitude} + 0,162 \cdot \text{Longitude}$ $r = 0,903$
Manganês	$\ln[\text{Mn}] = -115 + 0,0290 \cdot \ln[\text{sólidos dissolvidos}] + 0,280 \cdot \ln[\text{sólidos em suspensão}] + 0,147 \cdot \ln[\text{turbidez}] + 0,243 \cdot \ln[\text{Ni}] + 0,273 \cdot \ln[\text{Cr}^{\text{III}}] + 1,60 \cdot \text{Latitude} + 1,87 \cdot \text{Longitude}$ $r = 0,797$
Cobre	$\ln[\text{Cu}] = -79,4 + 0,355 \cdot \ln[\text{Cr}^{\text{III}}] + 0,237 \cdot \ln[\text{Ni}] + 0,142 \cdot \ln[\text{surfactantes}] + 0,151 \cdot \ln[\text{Mn}] + 0,200 \cdot \text{Latitude} + 1,65 \cdot \text{Longitude}$ $r = 0,770$
Sólidos em suspensão	$\ln[\text{sólidos em suspensão}] = 50 + 0,779 \cdot \ln[\text{sólidos totais}] + 0,553 \cdot \ln[\text{turbidez}] + 0,140 \cdot \ln[\text{pH}] - 0,248 \cdot \text{Latitude} - 1,08 \cdot \text{Longitude}$ $r = 0,906$

* As concentrações dos parâmetros correspondem a valores totais

A presença de concentrações mais baixas e próximas aos limites de cada método também influenciou os valores dos coeficientes de correlação, $r = 0,770$ e $0,797$, nos modelos de cobre e manganês respectivamente.

A validade dos modelos foi testada aplicando a equação de regressão aos resultados de monitoramento no ano de 2005. A FIGURA 12 apresenta uma comparação entre os valores reais e estimados para a concentração de manganês no ano de 2005.

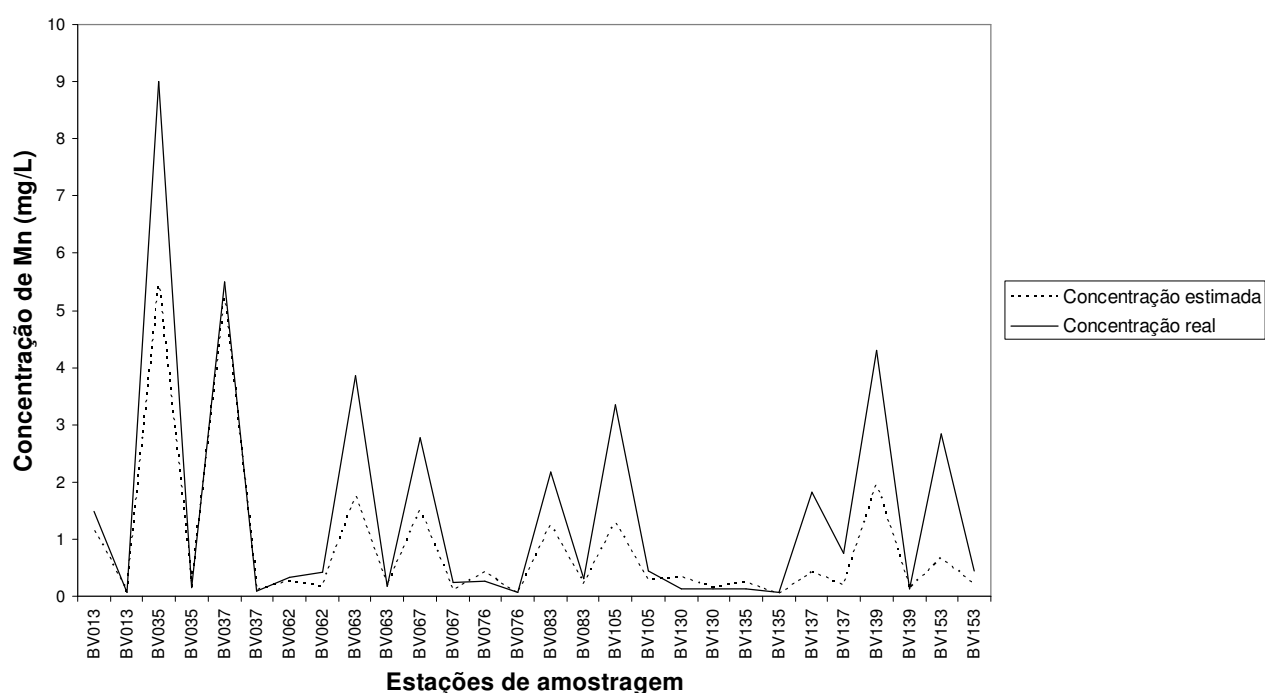


FIGURA 12 - Concentração de manganês real e estimada para as estações compreendidas entre as latitudes $19^{\circ}33'$ e $20^{\circ}14'$ na sub-bacia do rio das Velhas

As tendências de crescimento e decréscimo da concentração de manganês foram melhores estimadas pela equação de regressão para esse modelo. Novamente, o uso dos dados das campanhas realizadas apenas nos meses de janeiro e julho pode ter prejudicado o ajuste das concentrações estimadas às concentrações reais.

5.7 Conclusões do capítulo

A análise fatorial aplicada a matriz dos dados de monitoramento da sub-bacia do rio das Velhas no período compreendido entre janeiro de 1997 a julho de 2004 permitiu a caracterização dos parâmetros mais importantes destacados como os componentes de maior carga nos fatores 1 e 2.

Os parâmetros destacados pela carga no fator 1 são característicos de efluentes domésticos e industriais e a análise de variância mostrou que esse fator diferencia algumas estações do alto e baixo curso do rio das Velhas situadas na região metropolitana de Belo Horizonte como as mais atingidas. As estações BV 154, ribeirão do Onça a montante do rio das Velhas e BV155, ribeirão Arrudas a montante do rio das Velhas são as que apresentam as piores condições.

Para o fator 2, os parâmetros de maior componente de carga são característicos de efluentes de mineração como cobre, manganês, níquel e também erosão e assoreamento como sólidos em suspensão, sólidos totais e turbidez. A análise de variância destaca principalmente as estações BV035, rio Itabirito a jusante da cidade de Itabirito, situada no alto curso do rio das Velhas e a estação BV137, rio das Velhas na ponte Raul Soares no município de Lagoa Santa situada no médio curso.

Este estudo permitiu desenvolver um modelo preliminar de avaliação dos principais poluentes, destacados pela análise fatorial, em regiões específicas, destacadas pela análise de variância entre os fatores e a latitude das estações. Os coeficientes de correlação determinados nas equações de regressão para os parâmetros fosfatos totais e manganês podem ter sido influenciados pelo número de medidas utilizados na sua elaboração, ou seja, o uso apenas dos dados das campanhas realizadas apenas nos meses de chuva e estiagem. A existência de estações com teores mais baixos do poluente e próximo do limite de detecção dentro das regiões específicas também pode ter influenciado nos ajustes dos modelos aos valores reais em 2005. A variabilidade dos parâmetros nas regiões onde as equações de regressão foram determinadas também pode ter influenciado sua acurácia.

Capítulo VI – Análise exploratória da sub-bacia do rio das Velhas

Este estudo foi realizado na macro-rede de monitoramento da sub-bacia do rio das Velhas, descrita no capítulo I, visando identificar as principais regiões poluídas e os principais poluentes. Com essa finalidade foram empregadas as técnicas de análise estatística multivariada e de análise exploratória, especificamente análise de agrupamentos e análise de componentes principais. A análise de agrupamentos foi utilizada como técnica de classificação das estações em relação às suas similaridades. A análise de componentes, assim como a análise fatorial, tem como objetivo a descrição da variabilidade original dos dados em termos de um número menor de componentes e foi utilizada também de maneira comparativa à análise fatorial.

6.1 Descrição do conjunto de dados

O tratamento estatístico foi realizado em 41 parâmetros físico-químicos monitorados pelo IGAM em 29 estações de amostragem ao longo do rio das Velhas, TABELA 2. O conjunto dos dados foi organizado em forma de uma matriz composta de 29 linhas

correspondentes às estações de amostragem e 41 colunas correspondentes aos parâmetros de qualidade das águas.

Cada elemento da matriz de dados, x_{ik} , corresponde à mediana das medidas realizadas entre setembro de 1997 a outubro de 2005 de 41 dos 51 parâmetros monitorados, TABELA 6, nas 29 estações de amostragem.

A mediana da concentração do parâmetro foi utilizada nesse estudo por ser mais representativa da variação de cada parâmetro durante o período, não influenciada por *outlier* como a média e mais empregada na avaliação de dados ambientais.³⁸

Parâmetros como cádmio, cianetos, cromo hexavalente, mercúrio, selênio e sulfetos foram excluídos por apresentarem mediana constante e igual que os limites de quantificação para todas as estações. Parâmetros como alcalinidade bicarbonato, cálcio, magnésio e dureza total são muito correlacionados com os parâmetros alcalinidade total, dureza de cálcio, dureza de magnésio e dureza de cálcio respectivamente. Assim, os últimos foram mantidos na base de dados analisada e os primeiros foram eliminados. Os valores do *coeficiente r de correlação de Pearson*^{34,55} desses parâmetros em relação aos parâmetros mantidos são respectivamente: $r = 0,995$ de alcalinidade bicarbonato com relação à alcalinidade total, $r = 0,994$ de alcalinidade cálcio com relação à dureza de cálcio, $r = 0,991$ de alcalinidade magnésio com relação à dureza de magnésio, e $r = 0,992$ de dureza total com relação à dureza de cálcio.

Devido às diferenças nas escalas entre as medidas, os parâmetros foram padronizados através da equação:

$$x_{ik} = \frac{x_{ik} - \bar{x}_k}{s_k} \quad (98)$$

onde, \bar{x}_k é a média da k-ésima coluna e s_k seu desvio padrão. Essa padronização foi utilizada de forma alternativa ao cálculo de logaritmo natural, utilizado no capítulo anterior, e aplicada à mediana dos parâmetros durante o período em estudo não envolveu a eliminação dos dados de monitoramento das campanhas intermediárias.

A TABELA 5 do capítulo IV apresenta a metodologia de ensaio de cada avaliado e a TABELA 6 do mesmo capítulo as medianas dos dados originais.

6.2 Análise de agrupamentos e componentes principais

Inicialmente a análise de agrupamentos foi aplicada à matriz das medianas padronizadas buscando avaliar a classificação das estações por similaridade considerando os parâmetros avaliados, o dendrograma obtido é mostrado na FIGURA 13. Para esta análise, a distância Euclidiana foi calculada na matriz das medianas padronizadas das estações e como técnica hierárquica aglomerativa foi utilizado o método de Ward, descritos no capítulo II.

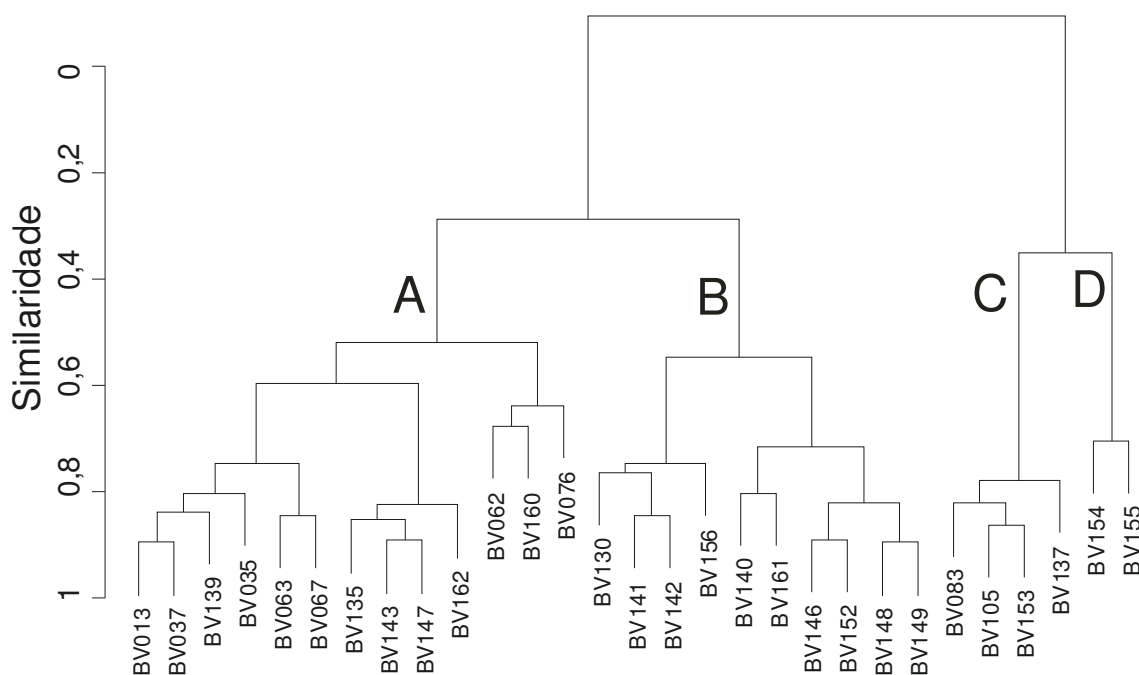


FIGURA 13 - Análise de agrupamento na matriz padronizada das medianas dos 41 parâmetros medidos nas 29 estações de amostragem do rio das Velhas, utilizando o método de Ward

Os resultados da análise mostram que as estações podem ser classificadas em 4 grupos com níveis de similaridade acima de 0,4. É possível estabelecer uma relação entre a ordem das estações no dendrograma e o índice de qualidade das águas, IQA, para cada estação. Na FIGURA 14 estão apresentados os valores médios de IQA e os limites mínimos e máximos de todas as estações no período de 1997 a 2005.

A primeira estação à esquerda, BV013, rio das Velhas a montante do rio Itabira próximo da sua nascente, apresentou melhores índices de qualidade e as estações à direita BV154, ribeirão do Onça a montante do rio das Velhas e BV155, ribeirão Arrudas a montante do rio das Velhas apresentou os piores índices de qualidade de água, IQA. Levando-se em consideração a variação do IQA durante o período em estudo para as estações entre os dois extremos do dendrograma observa-se que a qualidade das águas piora no sentido da esquerda para direita dessa classificação. A maioria das estações localizadas nos grupos A e B apresentaram IQA variando de bom a médio e as estações dos C e D apresentaram IQA médio a ruim.

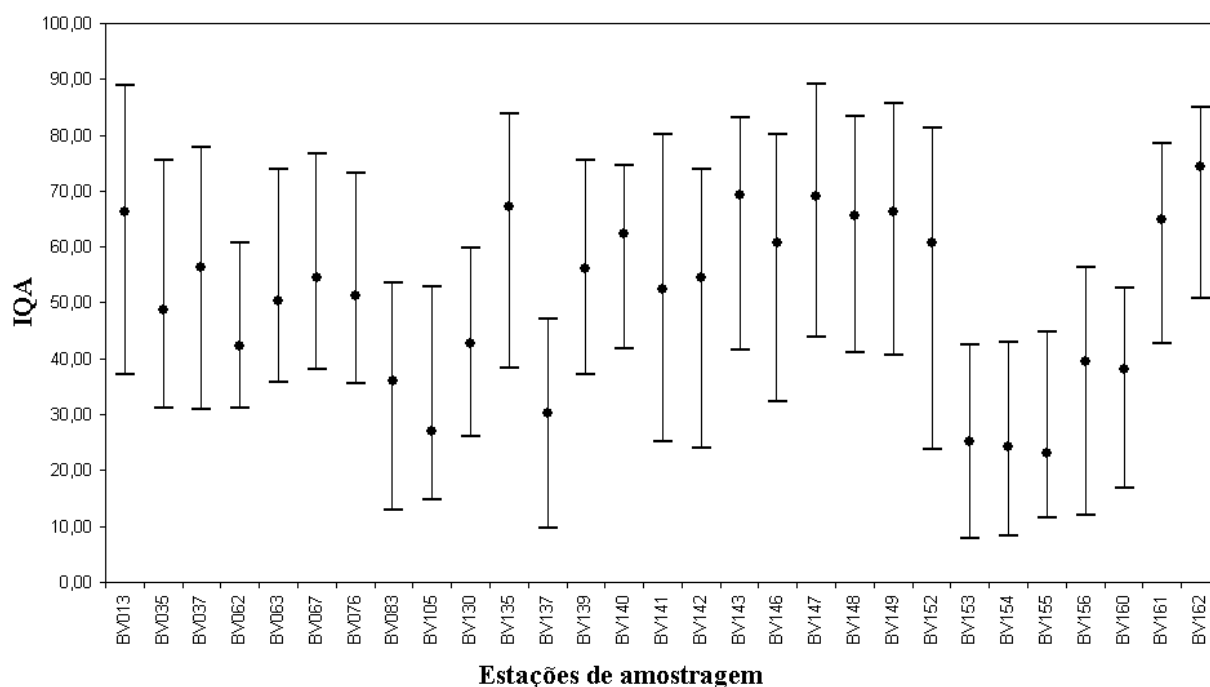


FIGURA 14 – Valores médios, mínimos e máximos de IQA para estações de amostragem no período de 1997 a 2005¹⁻⁷

A partir desses dados foi realizada a análise de componentes principais. O propósito da análise de componentes principais é a redução da complexidade dos dados multivariados e a detecção de estrutura na relação entre os dados, transformando-os em um novo conjunto de variáveis, as chamadas componentes principais, escritas como uma combinação linear das variáveis originais.

Na TABELA 12 estão relacionados os autovalores, a porcentagem da variância explicada para dez primeiras componentes principais e a variância acumulada por essas componentes.

TABELA 12 – Autovalores, porcentagem de variância e variância acumulada para dez primeiras componentes principais da matriz da padronizada da mediana dos 41 parâmetros para 29 estações de amostragem ao longo da sub – bacia do rio das Velhas

COMPONENTES	AUTOVALORES	PORCENTAGEM DA VARIÂNCIA (%)	VARIÂNCIA ACUMULADA (%)
Componente 1	18,9	47,7	47,7
Componente 2	5,75	14,5	62,2
Componente 3	4,15	10,5	72,7
Componente 4	2,03	5,12	77,8
Componente 5	1,75	4,41	82,2
Componente 6	1,31	3,32	85,6
Componente 7	0,936	2,36	87,9
Componente 8	0,874	2,21	90,1
Componente 9	0,749	1,89	92,0
Componente 10	0,716	1,81	93,8

O método de Cattell, aplicado aos autovalores da matriz de covariância para as dez primeiras componentes principais, apresentado na FIGURA 15, mostra que a partir da quinta componente os autovalores tendem para valores constantes e menores que 1,5; valor considerado pouco significativo, mostrando que as 5 primeiras componentes principais são as mais importantes.

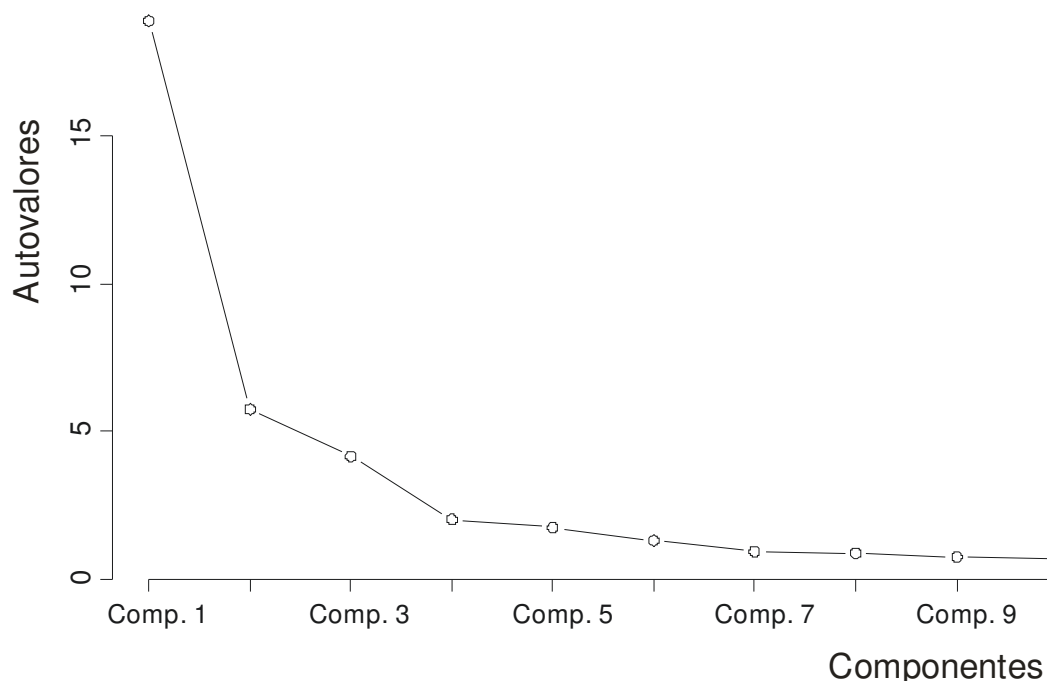


FIGURA 15 - Gráfico dos autovalores da matriz de covariância para as dez componentes principais da base de dados. As cinco primeiras componentes principais descrevem 82,2% da variabilidade dos dados

As componentes 1 (47,7 %), 2 (14,5%), 3 (10,5 %), 4 (5,12 %) e 5 (4,41 %) explicam 82,2% da variabilidade dos dados. Na TABELA 13 estão indicadas a composição e a carga dos parâmetros nas cinco primeiras componentes principais.

As cinco componentes principais apresentam em conjunto 33 parâmetros com valores absolutos de cargas superiores à 0,200 e que são considerados os mais importantes. São eles: cloretos, fosfatos, condutividade elétrica, zinco, nitrogênio orgânico, DQO, nitrogênio amoniacal, potássio, DBO, sólidos totais, sólidos dissolvidos, sódio, bário, coliformes totais, arsênio, coliformes fecais, dureza cálcio, temperatura do ar, nitratos, temperatura do água, pH in loco, pH laboratorial, cromo III, sólidos em suspensão, turbidez, manganês, chumbo, oleos e graxas, surfactantes, cor, ferro, dureza magnésio e boro.

TABELA 13 – Cinco primeiras componentes principais e respectivos componentes de cargas dos parâmetros com valor absoluto superior a 0,200 – matriz 29 x 41

Parâmetros	Comp. 1	Parâmetros	Comp. 2	Parâmetros	Comp. 3	Parâmetros	Comp. 4	Parâmetros	Comp. 5
Cloretos	0,221	Bário	0,311	Cromo III	0,355	Dureza magnésio	0,361	Cor	0,491
Fosfatos	0,214	Coliformes totais	0,235	Bário	0,311	Arsênio	0,357	Ferro	0,426
Cond. elétrica	0,213	Arsênio	0,223	Sólidos em suspensão	0,258	Nitratos	0,278	Turbidez	0,215
Zinco	0,213	Coliformes fecais	0,201	Turbidez	0,239	Cor	-0,303	Óleos e graxas	-0,230
Nitrogênio orgânico	0,212	Dureza cálcio	-0,257	Manganês	0,237	Turbidez	-0,314	Surfactantes	-0,257
DQO	0,211	Temp. do ar	-0,270	Arsênio	0,223	Boro	-0,386		
Nitrogênio amoniacal	0,210	Nitratos	-0,301	Chumbo	0,203				
Potássio	0,208	Temp. da água	-0,331	Óleos e graxas	-0,205				
DBO	0,206	pH in loco	-0,342	Surfactantes	-0,208				
Sólidos totais	0,206	pH laboratorial	-0,384	Cor	-0,220				
Sólidos dissolvidos	0,202								
Sódio	0,201								

Dos 33 parâmetros mais importantes na classificação das estações através da sua variabilidade, foi eliminado parâmetro pH laboratorial por expressar a mesma medida do pH in loco, confirmada pelo valor da correlação de *Pearson* igual a $r = 0,876$. Com os 32 parâmetros selecionados pela análise de componentes principais foi realizada uma nova análise de agrupamentos, visando avaliar a classificação das estações com o conjunto reduzido de parâmetros. Na FIGURA 16 é apresentada a classificação das 29 estações de amostragem levando-se em consideração os 32 parâmetros selecionados na análise de componentes principais.

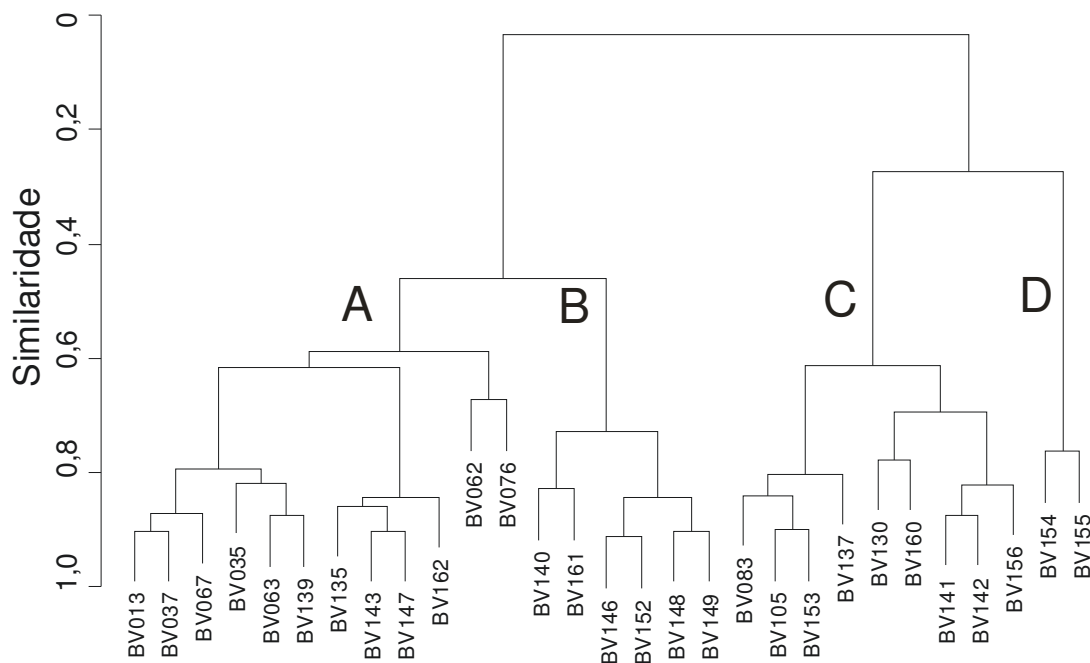


FIGURA 16 – Análise de agrupamento na matriz padronizada das medianas dos 32 parâmetros medidos nas 29 estações de amostragem do rio das Velhas, utilizando o método de Ward

A redução dos parâmetros introduziu pequenas alterações na nova classificação pela análise de agrupamento quando comparada à análise de agrupamentos realizada com os 41 parâmetros, mostrada na FIGURA 13. A principal alteração é o deslocamento das estações BV130, ribeirão da Mata a montante do rio da Velhas, BV141; rio das Velhas a jusante do ribeirão Jequitibá na cidade de Santana do Pirapama; BV142, rio das Velhas a montante da foz do rio Paraúna; BV156, rio das Velhas a jusante do rio Jaboticatubas e BV160, ribeirão das Neves próximo de sua foz no ribeirão da Mata dos grupos A e B para o grupo C. Todas estas estações apresentaram IQA médio e $CT > 2P$ como as estações do grupo C.

Os resultados da análise de agrupamentos com a base de parâmetros reduzida mostram que as estações são ainda classificadas em 4 grupos, mas com níveis de similaridade acima de 0,25. As estações agrupadas mais à esquerda do dendrograma, correspondentes ao grupo A apresentaram melhores índices de qualidade das águas em comparação as estações localizadas nos demais grupos.

Comparando a classificação das estações em relação à média do índice de qualidade da água, IQA, durante o período de 1997-2005, apresentada na FIGURA 14, pode-se observar que para as estações do grupo A, a estação BV162, Rio Cipó a montante da foz do Rio Paraúna apresentou IQA bom, as estações BV035, Rio Itabirito a jusante da cidade de Itabirito e BV062, ribeirão Água Suja próximo de sua foz no rio das Velhas apresentaram IQA ruim e as demais estações apresentaram IQA médio. A piora da qualidade das águas das estações BV035 e BV062 se deve à contaminação provocada por minerações, característica dessa região, como mostrado na TABELA 4 do capítulo IV e evidenciado pelo índice de contaminação por tóxicos alto, CT^{1-7} , superior a 2. As estações do grupo B, apresentaram IQA médio. Com exceção apenas da estação BV141, rio das Velhas a jusante do ribeirão Jequitibá e BV142, rio das Velhas a montante do rio Paraúna todas as estações dos grupos C apresentaram IQA ruim e $CT > 2$. Grande parte das estações agrupadas nos grupos C e D são estações localizadas na região metropolitana de Belo Horizonte que frequentemente são sujeitas a descargas de efluentes domésticos e industriais sem tratamento.

A partir desses resultados, foi feita uma nova análise de componentes principais com o objetivo de classificar as estações em relação aos seus principais parâmetros poluentes. Nessa nova análise, o parâmetro oxigênio dissolvido, característico de águas de boa qualidade, foi acrescentado com a finalidade de classificar as estações com maiores concentrações desse parâmetro.

Na TABELA 14 estão relacionados os autovalores, a porcentagem da variância explicada para dez primeiras componentes principais e a variância acumulada para análise componentes principais realizada na matriz composta pelos 32 parâmetros mais importantes e oxigênio dissolvido nas 29 estações de amostragem.

TABELA 14 – Autovalores, porcentagem de variância e variância acumulada para dez primeiras componentes principais da matriz da padronizada da mediana dos 32 parâmetros mais importantes e oxigênio dissolvido para 29 estações de amostragem ao longo da sub – bacia do rio das Velhas

COMPONENTES	AUTOVALORES	PORCENTAGEM DA VARIÂNCIA (%)	VARIÂNCIA ACUMULADA (%)
Componente 1	15,2	47,6	47,6
Componente 2	4,44	13,9	61,5
Componente 3	3,70	11,5	73,0
Componente 4	1,81	5,69	78,7
Componente 5	1,62	5,08	83,8
Componente 6	1,19	3,74	87,5
Componente 7	0,799	2,51	90,0
Componente 8	0,659	2,07	92,1
Componente 9	0,575	1,80	93,9
Componente 10	0,481	1,51	95,4

Avaliando os dados da TABELA 14, observa-se que as componentes nessa nova análise, as componentes 1 (47,6 %), 2 (13,9%), 3 (11,5 %), 4 (5,69 %) e 5 (5,08 %) apresentam maior porcentagem da variância. Essas componentes somadas explicam 83,8% da variabilidade dos dados.

O método de Cattell, aplicado aos autovalores da matriz de covariância para as dez primeiras componentes principais apresentado na FIGURA 17, mostra que a partir da quinta componente os autovalores tendem para valores constantes e menores que 1,5; valor considerado pouco significativo, mostrando que as cinco primeiras componentes principais são as mais importantes.

As cinco principais componentes apresentam em conjunto trinta e três parâmetros com valores absolutos de cargas superiores à 0,200 e considerados os mais importantes. São eles: cloretos, fosfatos, condutividade elétrica, nitrogênio orgânico, DQO, nitrogênio amoniacal, potássio, zinco, DBO, sólidos totais, sódio, sólidos dissolvidos, oxigênio dissolvido, coliformes totais, coliformes fecais, manganês, chumbo, dureza de cálcio, temperatura do ar, nitratos, pH *in loco*, temperatura do água, cromo III, bário, sólidos em suspensão, turbidez, arsênio, óleos e graxas, surfactantes, cor, ferro, dureza de magnésio e boro. A segunda análise de componentes principais não permitiu uma nova redução de parâmetros e resultou cargas próximas às encontradas na primeira análise.

Representando graficamente os escores e as cargas de cada parâmetro nas componentes principais é possível avaliar a importância desses parâmetros e das estações na base de dados. Na FIGURA 18 estão representados por vetores os pesos de cada parâmetro nas componentes principais 1 e 2 e simultaneamente os escores de cada parâmetro em um gráfico *biplot*, onde as cargas e os escores foram escalonados por métodos convencionais^{17, 35, 54}. Quanto mais paralelo for o vetor de carga ao eixo da componente principal maior é a importância do parâmetro correspondente. A proximidade das estações ao vetor de carga de cada parâmetro monitorado reflete a importância do parâmetro na separação das estações.

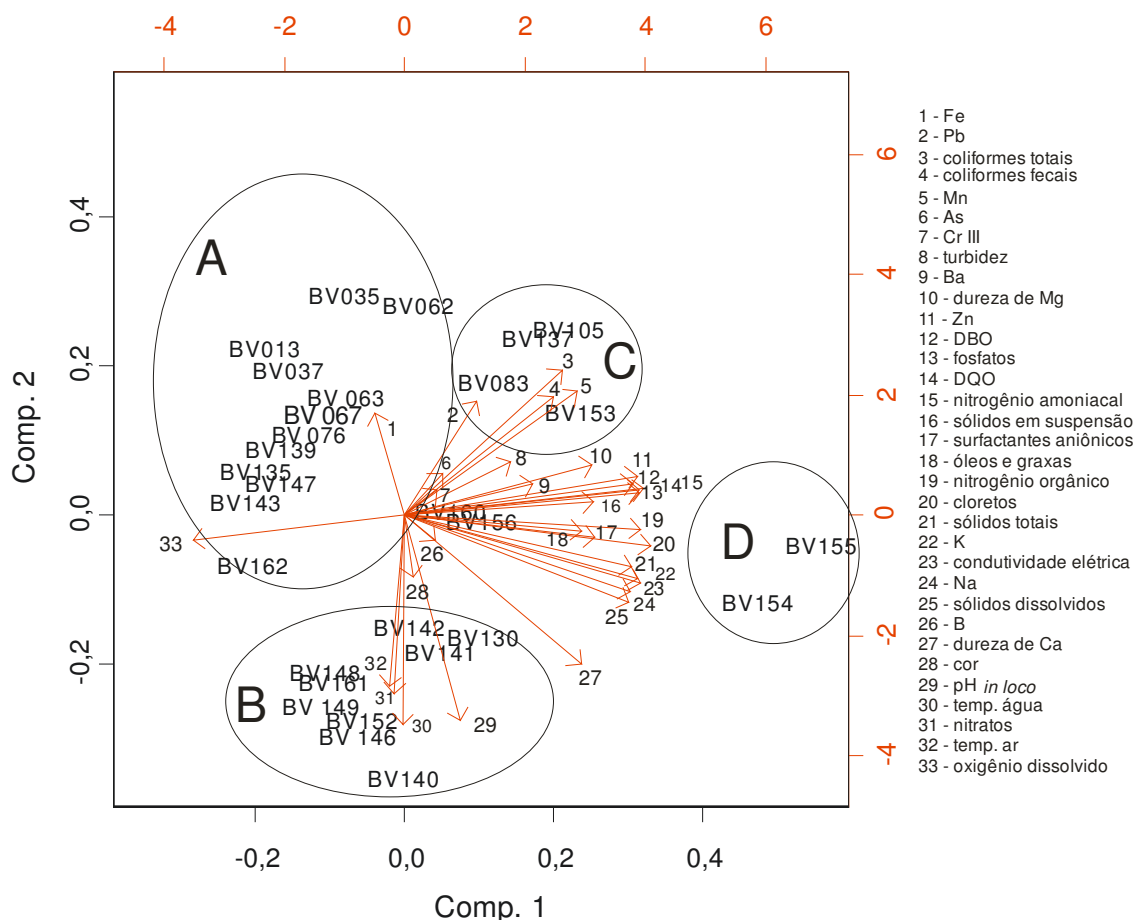


FIGURA 18 - Gráfico *biplot* de escores e cargas dos parâmetros para as duas primeiras componentes principais para todas as estações da rede da macro-rede de monitoramento da sub-bacia do rio das Velhas. Escala de escores no intervalo de -0,3 a +0,5 e escala de pesos no intervalo de -4,5 a +4,5.

O gráfico *biplot* apresenta uma separação das estações em quatro grupos:

As estações do grupo A: BV162, BV143, BV147, BV135, BV139, BV076, BV067, BV063, BV013, BV037, BV035 e BV062 foram caracterizadas pelos parâmetros oxigênio dissolvido e ferro. A proximidade da estação em relação ao parâmetro oxigênio dissolvido reflete uma melhor qualidade da água, assim, a estação BV162, rio Cipó a montante do rio Paraúna no município de Presidente Juscelino pode ser caracterizada como a estação com águas de melhor qualidade como confirmado pelo IQA, FIGURA 14. A presença de ferro é uma característica do solo. Por outro lado, as estações BV035, rio Itabirito a jusante do Córrego Cata Branca e BV062, ribeirão Água Suja próximo de sua foz no rio das Velhas mais próximas ao primeiro grupo apresentaram pior qualidade devido aos impactos de mineração. As estações que

compõem esse grupo também aparecem associadas no grupo A da análise de agrupamentos, FIGURAS 13 e 16.

O grupo B é composto pelas estações BV142, BV141 e BV130 caracterizadas pelo parâmetro cor e mais próximas às estações do segundo grupo e pelas estações BV148, BV161, BV149, BV152, BV146 e BV140 caracterizadas pelos parâmetros temperatura da água, temperatura do ar, pH e nitratos. Os piores índices de qualidade para esse grupo foram apresentados para estações BV130, ribeirão da Mata próximo de sua foz no rio das Velhas; BV161, ribeirão Santo Antônio próximo de sua foz no rio das Velhas; BV141, rio das Velhas na cidade de Santana do Pirapama e BV142, rio das Velhas a montante da foz do rio Paraúna; e as demais apresentaram índices médios, FIGURA 14.

O grupo C composto pelas estações BV083, rio das Velhas a jusante do ribeirão Arrudas; BV105, rio das Velhas a jusante do ribeirão do Onça; BV137, rio das Velhas na ponte Raul Soares próximo ao município de Lagoa Santa e BV153, rio das Velhas a jusante do ribeirão da Mata município de Santa Luzia foram caracterizadas pelos parâmetros coliformes fecais¹⁻⁸, coliformes totais¹⁻⁸, chumbo, manganês, turbidez, dureza de magnésio e bário. A presença desses poluentes caracteriza o lançamento de esgoto sanitário e de efluentes industriais. A presença de manganês é bastante marcante no rio das Velhas e a ocorrência deste elemento está relacionada, principalmente, ao fato da sub-bacia do rio das Velhas, sobretudo seu alto curso, estar inserida na região do Quadrilátero Ferrífero do Estado de Minas Gerais. A presença de manganês pode também estar relacionada aos lançamentos das indústrias do ramo siderúrgico e têxtil, localizadas próximas ao ribeirão Arruda. A presença de chumbo e a turbidez estão correlacionadas com a presença de manganês. As estações que compõem esse grupo também aparecem associadas no grupo C da análise de agrupamentos, FIGURAS 13 e 16.

O grupo D composto pelas estações BV154, ribeirão do Onça a montante do rio das Velhas e BV155, ribeirão Arrudas a montante do rio das Velhas foi caracterizado pelos parâmetros DBO, DQO, fosfatos, sólidos em suspensão, nitrogênio orgânico, surfactantes, cloretos, óleos e graxas, sólidos totais, condutividade elétrica, sólidos totais e sólidos dissolvidos. Por serem os vetores desses parâmetros paralelos ao eixo da componente 1, eles são os principais na sua composição. A presença desses poluentes

caracteriza os impactos causados por lançamentos de esgoto sem tratamento, efluentes industriais, erosão e assoreamento. As estações que compõem esse grupo também aparecem associadas no grupo D da análise de agrupamentos, FIGURAS 13 e 16. As estações BV160 e BV156 ficaram indefinidas entre dois grupos.

A proximidade das estações caracterizadas por águas de pior qualidade aos vetores dos parâmetros característicos de esgotos também reflete a contribuição desses na piora da qualidade da águas. Os parâmetros temperatura da água e do ar refletem características climáticas das estações onde foram importantes; o pH reflete as características do solo e a presença de nitratos reflete a depuração da carga orgânica e de nutrientes recebidos ao longo do trecho pela existência de níveis não críticos de oxigênio dissolvido.¹⁻⁷ As estações que compõem esse grupo também aparecem associadas no grupo B da análise de agrupamentos, FIGURAS 13 e 16

Na FIGURA 19 os pesos de cada parâmetro nas componentes principais 1 e 3 estão representados por vetores e simultaneamente os escores de cada parâmetro em um gráfico *biplot*. Nesta figura as estações impactadas por resíduos industriais e efluentes de mineração ficam mais evidentes. As estações BV156, rio das Velhas a jusante do rio Jaboticatubas; BV141, rio das Velhas na cidade de Santana do Pirapama e BV142, rio das Velhas a montante da foz do rio Paraúna; foram caracterizadas pela presença de cromo III, associado ao lançamento de efluentes industriais.

As estações BV083, rio das Velhas a jusante do ribeirão Arrudas; BV130, ribeirão da Mata próximo de sua foz no rio das Velhas e BV137, rio das Velhas na ponte Raul Soares próximo ao município de Lagoa Santa foram caracterizadas pela presença de chumbo, arsênio, turbidez, manganês, sólidos em suspensão e bário; provenientes de atividades minerárias que impactam a estação BV130, erosão e assoreamento.

As estações BV153, rio das Velhas a jusante do ribeirão da Mata município de Santa Luzia e BV105, rio das Velhas logo a jusante do ribeirão do Onça são caracterizadas por dureza de cálcio, dureza de magnésio e sólidos totais. Os parâmetros que caracterizam essas estações estão associados à erosão, assoreamento e resíduos sólidos urbanos.

As águas das estações BV013, rio das Velhas logo a montante da foz do rio Itabirito e BV139, rio das Velhas a montante da ETA/COPASA, em Bela Fama; destacaram-se como de melhor qualidade devido à associação ao parâmetro oxigênio dissolvido.

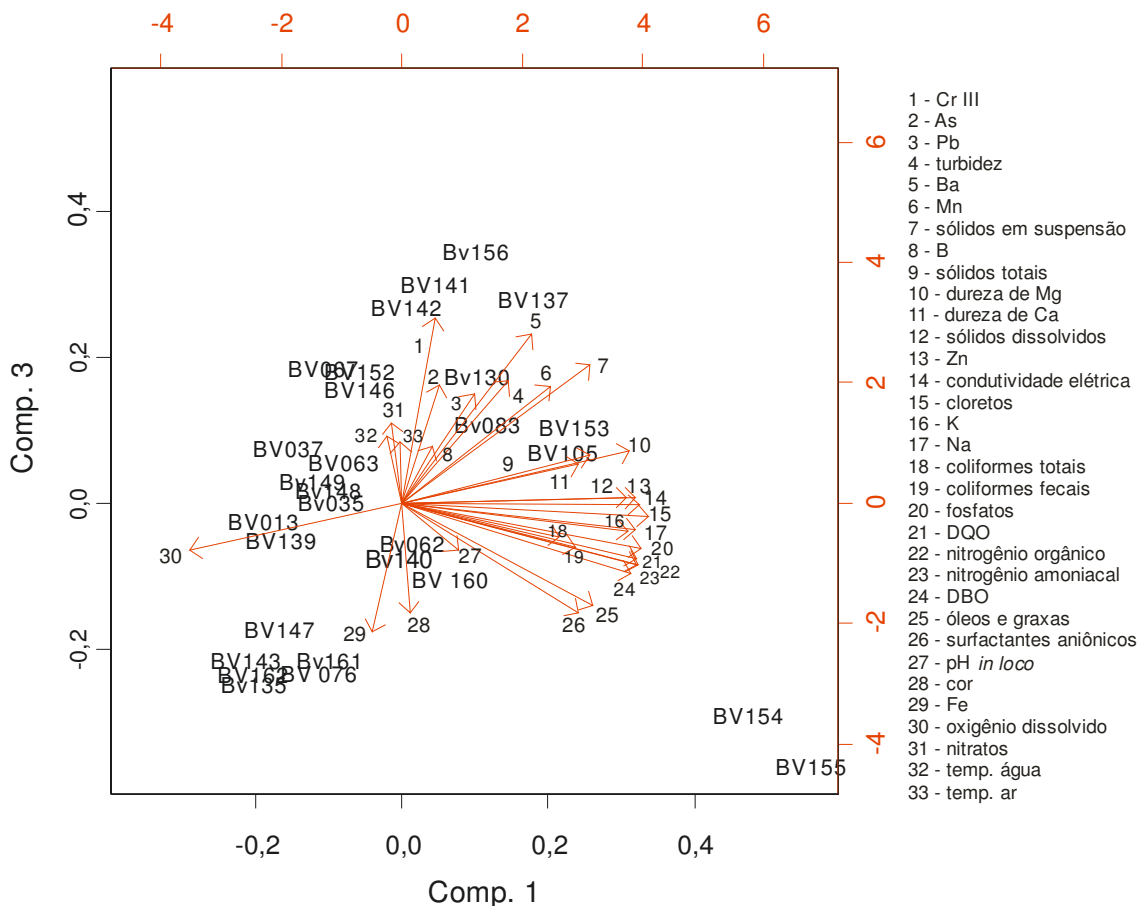


FIGURA 19 - Gráfico *biplot* de escores e cargas dos parâmetros para as componentes principais 1 e 3 para todas as estações da macro-rede de monitoramento da sub-bacia do rio das Velhas. Escala de escores no intervalo de -0,3 a +0,5 e escala de pesos no intervalo de -4,5 a +6,0.

6.3 Conclusões do capítulo

É possível pela análise de agrupamentos avaliar a qualidade das águas de acordo com sua classificação. Levando-se em consideração a variação do IQA durante o período em estudo para as estações entre os dois extremos do dendrograma, observa-se que a qualidade das águas piora no sentido da esquerda para direita dessa classificação. A maioria das estações localizadas nos grupos A e B apresentaram IQA variando de bom a médio e as estações dos C e D apresentaram IQA médio a ruim.

A análise de componentes principais mostrou que é possível uma redução dos parâmetros utilizados para avaliar a qualidade das águas das estações. São eles: cloretos, fosfatos, condutividade elétrica, nitrogênio orgânico, DQO, nitrogênio amoniacal, potássio, zinco, DBO, sólidos totais, sódio, sólidos dissolvidos, oxigênio dissolvido, coliformes totais¹⁻⁸, coliformes fecais¹⁻⁸, manganês, chumbo, dureza de cálcio, temperatura do ar, nitratos, pH *in loco*, temperatura do água, cromo III, bário, sólidos em suspensão, turbidez, arsênio, óleos e graxas, surfactantes, cor, ferro, dureza de magnésio e boro.

Os gráficos *biplot* evidenciaram quais são os principais parâmetros responsáveis pelos agrupamentos das estações nas análises de agrupamentos e de componentes principais. Assim, as estações do grupo A e B apresentam águas de melhor qualidade e foram associadas por parâmetros como oxigênio dissolvido, temperatura da água e do ar e ferro proveniente do solo. A presença de nitratos em algumas estações do grupo B é provocada pela depuração da carga orgânica pelo oxigênio dissolvido presente e confirma a melhoria de qualidade de suas águas. Entre as estações desses dois grupos podem ser destacadas: BV162, rio Cipó a montante do rio Paraúna no município de Presidente Juscelino; BV013, rio das Velhas logo a montante da foz do rio Itabirito e BV139, rio das Velhas a montante da ETA/COPASA, em Bela Fama.

Por outro lado, as estações dos grupos C e D apresentam águas de pior qualidade, pois estão associadas por parâmetros como coliformes fecais, nitrogênio amoniacal, fosfatos, turbidez, cloretos, nitrogênio orgânico provenientes de lançamento efluentes de esgotos. A contaminação causada pelas atividades de mineração em áreas próximas às estações de amostragem também é destacada pelas cargas atribuídas a parâmetros como sólidos em suspensão, turbidez, ferro total, manganês, chumbo, arsênio, cromo trivalente e bário, que estão entre os mais importantes, evidenciando o impacto ambiental causado por essa atividade em regiões específicas. Entre as estações desse grupo destacaram-se: BV 154, ribeirão do Onça próximo de sua foz no rio das Velhas; BV105, rio das Velhas a jusante do ribeirão do Onça; BV155, ribeirão Arrudas próximo de sua foz no rio das Velhas; BV083, rio das Velhas a jusante do ribeirão Arrudas; BV137, rio das Velhas na Ponte Raul Soares próximo ao município de Lagoa Santa e BV153, rio das Velhas a jusante do ribeirão da Mata.

A presença de poluentes característicos de efluentes esgotos entre a maioria dos parâmetros evidenciados na análise de componentes principais nos permite concluir que o principal impacto causado nas águas da sub-bacia do rio das Velhas é atribuído a deficiências nos processos de tratamento e controle desses efluentes durante o período estudado.

Os parâmetros destacados por suas cargas nas primeiras componentes principais também foram destacados na análise fatorial, apresentada no capítulo IV, por suas respectivas cargas nos principais fatores. Entre esses parâmetros estão presentes: cloretos, DBO, condutividade elétrica, potássio, sódio, sólidos totais, sólidos dissolvidos, cobre, fosfatos, manganês e turbidez. A redução de dados e identificação dos parâmetros mais importantes apresentou basicamente os mesmos resultados com as duas técnicas. As pequenas diferenças observadas nos resultados estão associadas ao pré-tratamento da base de dados e ao período avaliado em cada caso.

Capítulo VII - Análise exploratória do alto curso da sub-bacia do rio das Velhas

Este estudo foi realizado na rede amostragem definida pelo “Programa de Aperfeiçoamento do Monitoramento da Qualidade das Águas da Bacia do Alto Curso do rio das Velhas”. Esta rede é monitorada paralelamente ao monitoramento de toda bacia do rio das Velhas, objetivando o aumento do nível de conhecimento das condições de qualidade das águas de uma área específica de grande importância econômica e social do rio das Velhas e de seus efluentes e em particular dos pontos críticos para abastecimento de água.^{8, 11-12}. A área em estudo está localizada na porção centro sul do Estado de Minas Gerais, região em que está inserida uma das maiores províncias do minério de ferro no Brasil, o Quadrilátero Ferrífero, que possui uma área aproximada de 7200 km². A parcela correspondente ao presente estudo situa-se na porção central do Quadrilátero Ferrífero possuindo, aproximadamente, 2000 km², cobrindo a maior parte da região do alto curso do rio das Velhas.¹⁰⁰

Esse programa de monitoramento é dispendioso e por isso há necessidade de otimização da rede de amostragem proposta bem como da frequência de amostragem e do número de parâmetros analisados, sem perda da informação. Com essa finalidade foram empregadas as técnicas de análise estatística multivariada e exploratória;

especificamente a análise de agrupamentos e análise de componentes principais no tratamento da base de dados constituída de medidas de parâmetros químicos e físico-químicos feitas no período de 2003-2004 em águas da sub-bacia do alto curso do rio das Velhas.

A rede de monitoramento do alto curso do rio das Velhas é composta por 7 estações operadas pela Companhia de Saneamento de Minas Gerais - COPASA (AV220, AV230, AV240, AV260, AV270, AV280 e AV290), 7 estações (AV040, AV110, AV130, AV210, AV310, AV330 e AV350) monitoradas pelo Programa Águas de Minas e mais 23 novas estações de amostragem, que permitem uma melhor caracterização da situação de qualidade e dos efeitos de fontes pontuais e difusas de poluição, a partir da análise de parâmetros físico-químicos e biológicos. O “Programa de Aperfeiçoamento do Monitoramento da Qualidade das Águas da Bacia do Alto Curso do Rio das Velhas” engloba no total 37 estações de monitoramento divididas em redes básicas AV005, AV010, AV030, AV040, AV050, AV080, AV110, AV130, AV160, AV180, AV200, AV210, AV330, AV340, AV345 e AV350, *redes dirigidas para controle da poluição* AV020, AV060, AV070, AV090, AV100, AV120, AV140, AV150, AV170, AV190, AV250, AV300, AV310 e AV320 e em rede dirigidas para *o abastecimento doméstico*, AV220, AV230, AV240, AV260, AV270, AV280 e AV290, descritas na TABELA 2 e FIGURA 1 do capítulo I. Foram adotadas as frequências de amostragem mensal, para córregos que drenam áreas reconhecidamente impactadas⁸, estações AV090, AV100, AV120, AV140, AV150, AV300 e AV320, e trimestral para as demais estações.¹¹⁻¹²

7.1 Descrição do conjunto de dados do alto curso do rio das Velhas

Os dados analisados representam 39 parâmetros físico-químicos monitorados em água, listados na TABELA 6 do capítulo IV, em 37 estações de amostragem ao longo da sub-bacia do Alto Curso do Rio das Velhas listadas na TABELA 2 e FIGURA 1 do capítulo I. Entre as trinta e sete estações de monitoramento, as estações AV030, AV160 e AV180 foram avaliadas com coletas em diferentes profundidades por se tratarem de represas. Assim, nos pontos AV030E, AV160E e AV180E, as amostras foram coletadas respectivamente a 0,15m, 1,0m e 0,125m da superfície e nos pontos AV030F, AV160F

e AV180F as amostras foram coletadas na metade da zona fótica, ou seja, a 4,5m, 12,5m e 7,5m de profundidade, respectivamente. A zona fótica corresponde à faixa que vai até o ponto de máxima profundidade atingida pela luz e onde ainda é possível ocorrer a fotossíntese¹¹⁻¹². Os parâmetros analisados de acordo com as normas padronizadas⁹⁰ estão listados na TABELA 5 do capítulo IV.

A frequência das análises dos parâmetros foi igual à frequência de amostragem, que variou entre mensal, trimensal e semestral, ou menor para alguns parâmetros de ocorrência pouco comum ou que apresentaram pequena variabilidade anual.

Inicialmente foram calculadas as medianas no período de 2003 a 2004 de cada parâmetro em cada estação de amostragem. Considerando que nas estações AV030, AV160 e AV180, as amostras foram coletadas em profundidades diferentes, os dados foram organizados em uma matriz **X** de elementos x_{ik} , onde i corresponde a cada uma das 40 linhas compostas pelas 37 estações de amostragem e de três pontos coletados abaixo da superfície (AV030F, AV160F e AV180F) e k aos 39 parâmetros. Alguns parâmetros listados na TABELA 6 que apresentaram mediana constante foram excluídos desse tratamento, os 39 parâmetros considerados nesse tratamento são descritos na TABELA 7, capítulo IV.

7.2 Análise de agrupamentos

Devido às diferenças nas escalas dos parâmetros monitorados eles foram padronizados através da equação 98. A técnica de agrupamento hierárquica, combinada com a análise de componentes principais e a avaliação de gráficos de dispersão entre estações foram utilizadas com a finalidade de avaliar as similaridades das estações de amostragem e parâmetros monitorados.

O dendrograma exibido na FIGURA 20 mostra as estações agrupadas e o isolamento das estações AV310, rio das Velhas a montante do ribeirão de Água Suja, e AV320, rio das Velhas a montante do Córrego da Mina. A separação das estações no dendrograma reflete os resultados obtidos no monitoramento das águas dos cursos do Alto rio das Velhas⁸. Eles apontam melhores condições, segundo o índice de qualidade das águas,

IQA, e contaminação por tóxicos, CT, para estação AV005, rio das Velhas a montante de São Bartolomeu, enquanto as estações isoladas no dendrograma, AV310 e AV320, apresentaram os piores valores de IQA e CT. A qualidade das águas do ribeirão Água Suja, AV310, apresentou nível ruim nas coletas, segundo os índices CT e IQA e a estação AV320, rio das Velhas a montante do Córrego da Mina, apresentou IQA médio, e alto índice de CT. As demais estações apresentam valores intermediários. O isolamento das duas últimas estações mais contaminadas no dendrograma sugere que a proximidade entre as estações reflete o grau de contaminação das águas. A proximidade da estação AV005 no dendrograma às estações AV090, Córrego Carioca a montante da cidade de Itabirito AV170, Córrego Lagoa Grande a montante da represa das Codornas, AV100, Córrego da Onça a montante do Rio Itabirito, AV120, Córrego Moleque a montante do Rio Itabirito, e AV140, Córrego Fazenda Velha a montante do rio das Velhas, AV150, ribeirão Congonhas a montante da Represa das Codornas, AV240 Córrego Seco na captação Fechos, AV290 Córrego Mutuca na captação Mutuca, barragem auxiliar e AV280 Córrego Mutuca na captação Mutuca, barragem principal, sugere contaminações mais baixas ou pouco significantes nessas estações, confirmadas pelo bom IQA e baixa CT. A estação AV340, ribeirão da Prata a montante do rio das Velhas apresentou IQA médio e baixa CT.

As águas das estações AV070, AV080, AV300, AV330 e AV350, agrupadas proximamente às estações AV310 e AV320, foram avaliadas como de baixa qualidade em relação aos índices IQA e CT. A proximidade de estações no dendrograma de agrupamento indica também uma possível correlação entre elas.

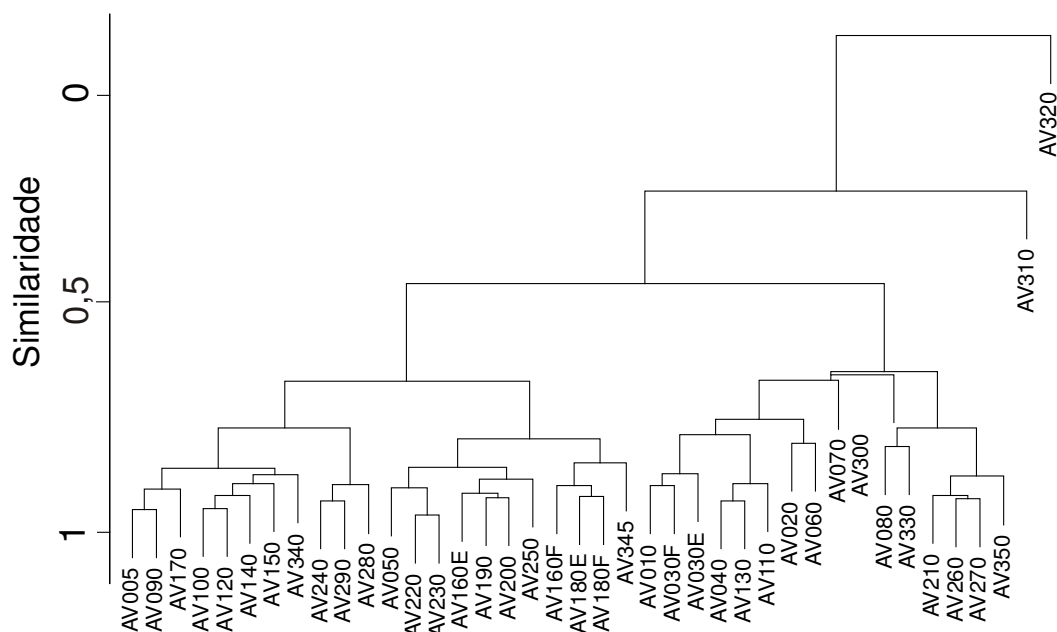


FIGURA 20 - Dendrograma indicando as estações separadas em grupos e o isolamento das estações AV310 e AV320.

Com base na análise do dendrograma de agrupamento foi investigada a correlação entre as estações mais próximas nos agrupamentos através dos gráficos de dispersão das variáveis de cada par de estações e cálculo do *coeficiente de correlação de Pearson*. A linearidade indicativa da correlação entre as estações pode ser empregada para redução da base de dados, eliminando estações correlacionadas.

Na FIGURA 21 estão representadas as dispersões das variáveis de alguns pares de estações próximas no dendrograma: no painel A, as estações AV220 e AV230, localizadas no Córrego Fechos na barragem principal e na barragem auxiliar respectivamente, com coeficiente de correlação $r = 0,960$; no painel B as estações AV240, localizada no Córrego Seco na captação Fechos e AV290, localizada no Córrego Mutuca na captação Mutuca, barragem auxiliar, com $r = 0,900$; no painel C as estações AV240 e AV280 localizada na barragem principal do Córrego Mutuca na captação Mutuca com $r = 0,610$; e no painel D as estações AV280 e AV290 com $r = 0,750$. As estações AV260 e AV270 localizadas no rio das Velhas em Bela Fama na alça direita e na alça esquerda respectivamente também estão correlacionadas, com coeficiente de correlação $r = 0,860$. É possível eliminar da amostragem algumas das estações correlacionadas empregando um critério de seleção como, os valores de IQA, CT e o valor de r acima do mínimo 0,8, por exemplo. Sugere-se assim a eliminação das estações AV220, AV240 por apresentarem melhores índices de qualidade das águas em

relação às respectivas estações correlacionadas, reduzindo a rede de monitoramento do Alto Velhas. A estação AV280 apesar de correlacionada com a estação AV290, apresentou valor de $r = 0,610$ com a estação AV240 e por isso deve ser mantida no monitoramento da rede de abastecimento doméstico. As estações AV260 e AV270, com $r = 0,860$ e IQA de médio a ruim, devem também ser mantidas.

As estações AV030, AV160 e AV180 avaliadas em diferentes profundidades apresentaram coeficientes de correlação menores que 0,8. As estações AV030E e AV030F apresentaram coeficiente de correlação, $r = 0,581$; as estações AV160E e AV160F, $r = 0,0691$ e as estações AV180E e AV180F, $r = 0,136$. Esses valores sugerem diferenças em relação à qualidade das águas nas diferentes profundidades avaliadas que podem estar associadas à mobilidade das águas e também à contaminação por sedimentos.

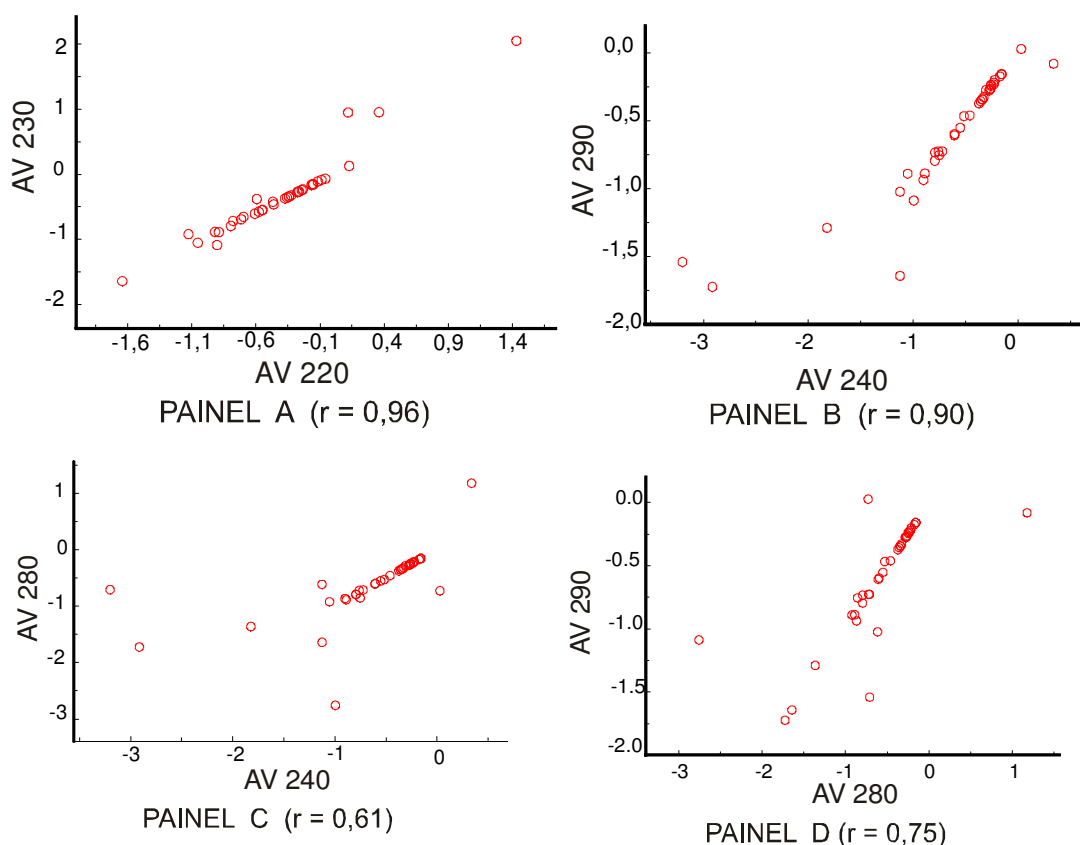


FIGURA 21 - Dispersões das variáveis de pares de estações próximas no dendrograma: painel A, estações AV220 e AV230; painel B, estações AV240, AV290; painel C, AV240 e AV280; painel D, estações AV280 e AV290. Os valores de r correspondem ao coeficiente de correlação de Pearson

7.3 Análise de componentes principais

Na TABELA 16 estão relacionados os autovalores, a porcentagem da variância explicada para dez primeiras componentes principais e a variância acumulada para análise componentes principais realizada na matriz composta pelos 39 parâmetros e 40 estações de amostragem.

TABELA 16– Autovalores, porcentagem de variância e variância acumulada para dez primeiras componentes principais da matriz da padronizada da mediana dos 39 parâmetros 40 estações de amostragem do alto curso do rio das Velhas

COMPONENTES	AUTOVALORES	PORCENTAGEM DA VARIÂNCIA (%)	VARIÂNCIA ACUMULADA (%)
Componente 1	16,8	44,1	44,1
Componente 2	7,13	18,7	62,8
Componente 3	3,86	10,1	72,9
Componente 4	2,55	6,70	79,6
Componente 5	1,53	4,03	83,6
Componente 6	1,35	3,55	87,2
Componente 7	0,929	2,44	89,6
Componente 8	0,852	2,24	91,9
Componente 9	0,600	1,58	93,4
Componente 10	0,499	1,31	94,8

O método de Cattell, aplicado aos autovalores da matriz de covariância para as dez primeiras componentes principais apresentados na FIGURA 22, mostra que a partir da quarta componente os autovalores tendem para valores menores que 1,5, valor considerado pouco significativo, mostrando que as quatro primeiras componentes principais são as mais importantes. As proporções das variâncias totais dessas quatro componentes principais são 44,1 %, 18,7%, 10,1% e 6,70% respectivamente. Essas componentes somadas explicam 79,6% da variabilidade dos dados.

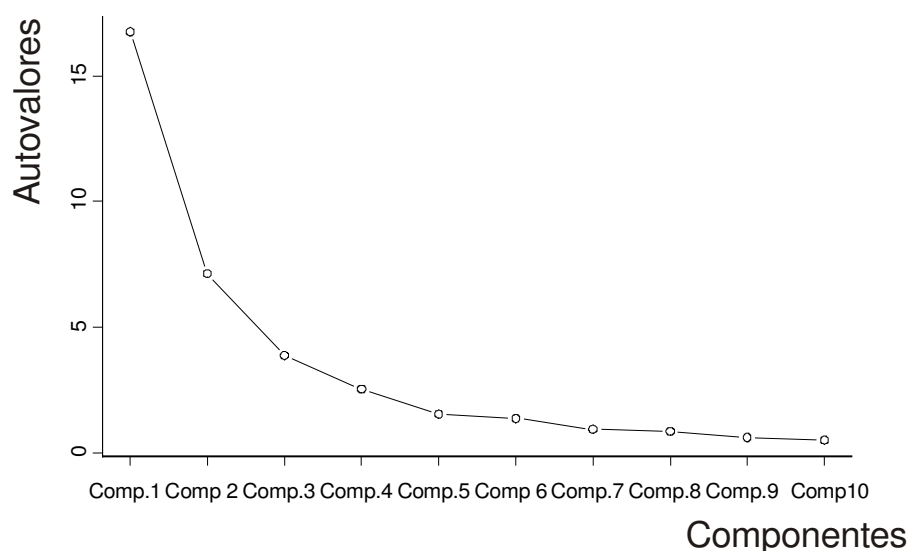


FIGURA 22 - Gráfico dos autovalores da matriz de covariância para as dez componentes principais da base de dados. As quatro primeiras componentes principais descrevem 79,6% da variabilidade dos dados

As contribuições dos seis parâmetros mais importantes nas quatro primeiras componentes principais estão representadas na TABELA 17.

TABELA 17 - Cinco primeiras componentes principais e respectivos componentes de cargas dos parâmetros com valor absoluto superior a 0,200 - matriz 40 x 39

Parâmetro	Comp.1	Parâmetro	Comp.2	Parâmetro	Comp.3	Parâmetro	Comp.4
Nitrogênio amoniacal	0,233	Fósforo	-0,339	Bário	0,383	pH lab.	0,470
Nitritos	0,229	Estreptococos	-0,314	Turbidez	0,363	pH in loco	0,443
Potássio	0,227	Coliformes fecais	-0,293	Ferro total	0,361	Cor	-0,411
Sódio	0,223	Surfactantes	-0,292	Sólidos em suspensão	0,331	Ferro solúvel	-0,389
Nitratos	0,221	DBO	-0,256	Temp. ar	0,257	Temp. ar	-0,263
Condutividade elétrica	0,221	DQO	-0,243	Temp. água	0,223		
Cloretos	0,221	Fenóis	-0,234	Cor	0,222		
Sólidos dissolvidos	0,220	Arsênio	-0,208				
Dureza de magnésio	0,217						
Dureza de cálcio	0,216						
Sulfatos	0,215						
Fluoretos/ Selênio	0,213						
Cobre	0,201						

Na componente 1 os parâmetros com maior peso são nitrogênio amoniacal, nitritos, K, Na, nitratos e condutividade; na componente principal 2 são fósforo, estreptococos, coliformes fecais, surfactantes DBO e DQO; na componente principal 3 são bário, turbidez, ferro total, sólidos em suspensão, temperatura do ar e cor e na componente 4 são pH lab, pH loco, cor, ferro solúvel, temperatura do ar e alcalinidade. Os demais parâmetros estão correlacionados com aqueles presentes nas componentes principais mais importantes. Por exemplo, os teores de arsênio estão correlacionados com DBO, DQO, surfactantes, fósforo, a nitrogênio amoniacal, alcalinidade, nitratos, coliformes fecais, Na e K, presentes nas componentes principais 1 e 2. Os parâmetros mais importantes nas componentes 1 e 2 são característicos de efluentes domésticos e industriais, indicando contaminação por estes efluentes nas águas do Alto Curso do Rio das Velhas.

Representando graficamente os escores e as cargas de cada parâmetro nas componentes principais é possível avaliar a importância desses parâmetros e das estações na base de dados. Na FIGURA 23 estão representados por vetores os pesos de cada parâmetro nas componentes principais 1 e 2 e simultaneamente os escores de cada parâmetro em um gráfico *biplot*. Novamente é observada a nítida separação das estações AV310 e AV320 das demais, como na análise de agrupamentos. Observa-se também uma distinção das estações AV300, Córrego da Barragem Nova Lima, e AV330, rio das Velhas a jusante do Córrego da Mina das estações restantes.

Podemos observar que a separação da estação AV320 se deve principalmente aos parâmetros fluoretos, dureza Mg, sulfatos, condutividade, sólidos dissolvidos, dureza Ca, Ni, Cu, Se, nitritos, Na, K, Zn, Mn, nitratos e N amoniacal, correlacionados com oxigênio dissolvido que apresenta correlação negativa. Estes parâmetros caracterizam os impactos causados por lançamento de esgoto sem tratamento e atividades minerárias. A separação da estação e AV310 se deve principalmente aos parâmetros alcalinidade, As, DBO, DQO, fenóis, surfactantes, coliformes fecais, estreptococos, fósforo, turbidez, ferro solúvel, ferro total, correlacionados com oxigênio dissolvido, OD, que apresenta correlação negativa. Os parâmetros que distinguem a estação AV310 são característicos de poluentes e justificados pelos impactos dos lançamentos, tanto de origem doméstica quanto industrial e minerária, provenientes do município de Nova Lima.

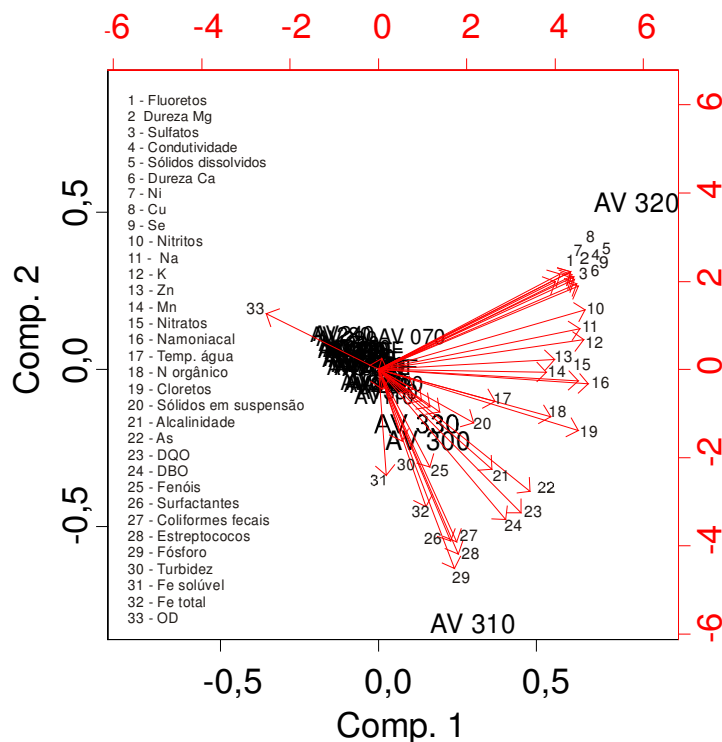


FIGURA 23- Gráfico *biplot* de escores e cargas para as duas primeiras componentes principais para todas as estações da rede de monitoramento do alto curso do rio das Velhas. Escala de escores no intervalo de -1 a +1 e escala de pesos no intervalo de -6 a +6.

A análise de componentes principais também foi realizada separadamente nas matrizes formadas pelas estações que compõem cada um dos três grupos de estações da rede de monitoramento, a rede básica, a rede dirigida para controle da poluição e a rede dirigida para o abastecimento doméstico buscando caracterizá-las nos seus respectivos grupos. O gráfico *biplot* de escores e cargas de cada parâmetro para as duas primeiras componentes principais para as estações da *rede básica*, AV005, AV010, AV030, AV040, AV050, AV080, AV110, AV130, AV160, AV180, AV200, AV210, AV330, AV340, AV345 e AV350 é mostrado na FIGURA 24.

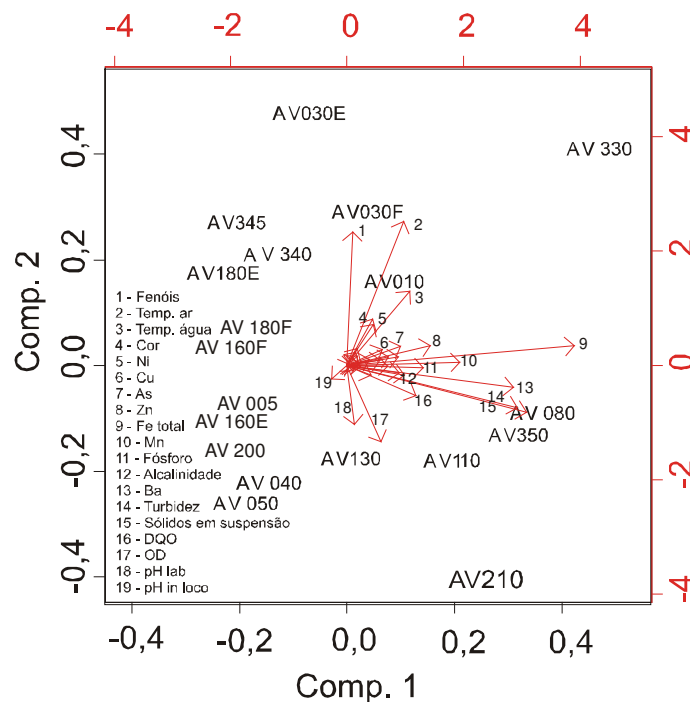


FIGURA 24- Gráfico *biplot* de escores e cargas para as duas primeiras componentes principais para as estações da rede básica, constituída das estações AV005, AV010, AV030, AV040, AV050, AV080, AV110, AV130, AV160, AV180, AV200, AV210, AV330, AV340, AV345 e AV350. Escala de escores no intervalo de -0,4 a +0,6 e escala de pesos no intervalo de -4 a +5.

Entre as estações que compõem a rede básica, a estação AV080, Rio Itabirito a montante de Itabirito, AV350, rio das Velhas a montante do ribeirão Sabará, AV110, Rio Itabirito a jusante do Córrego do Onça, AV210, rio das Velhas a jusante de Rio Acima, e AV330 rio das Velhas a jusante do Córrego da Mina foram caracterizados pelos parâmetros Zn, ferro total, Mn, Ba, sólidos em suspensão e turbidez, parâmetros que caracterizam atividades de mineração. As estações AV030E e AV030F foram discriminadas por parâmetros como fenóis e temperatura da água e do ar e cor. As inconformidades verificadas em relação aos índices de fenóis estão relacionadas à descarga de esgotos sanitários e efluentes industriais nas proximidades dessa estação. A estação AV130, rio das Velhas a jusante do Rio Itabirito é discriminada pelo parâmetro oxigênio dissolvido, que reflete na melhor qualidade da água nessa estação. As demais estações que compõem a rede básica não apresentaram parâmetros que as discriminasse.

O gráfico *biplot* de escores e cargas de cada parâmetro para as duas primeiras componentes principais para as estações da rede dirigida para controle da poluição, constituída das estações AV020, AV060, AV070, AV090, AV100, AV120, AV140, AV150, AV170, AV190, AV250, AV300, AV310 e AV320 é mostrado na FIGURA 25.

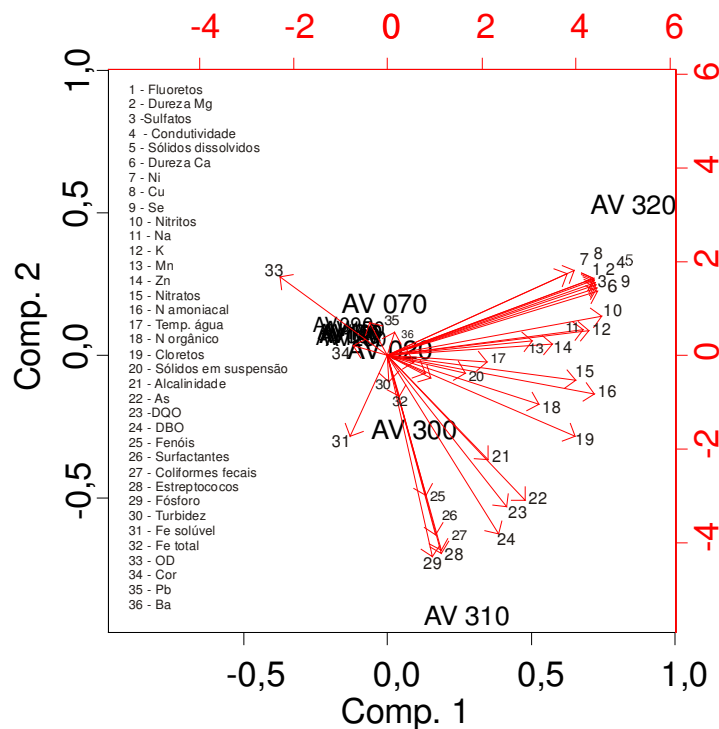


FIGURA 25- Gráfico biplot de escores e cargas para as duas primeiras componentes principais para as estações da rede dirigida para controle da poluição, constituída das estações AV020, AV060, AV070, AV090, AV100, AV120, AV140, AV150, AV170, AV190, AV250, AV300, AV310 e AV320. Escala de escores no intervalo de -1 a +1 e escala de pesos no intervalo de -6 a +6.

A análise de componentes principais realizada com as estações que compõem a rede dirigida para o controle de poluição indica a distinta separação das estações poluídas AV310 e AV320. Essas estações são discriminadas por parâmetros que caracterizam rejeitos de mineração e também de poluição doméstica e industrial. Como os índices de contaminação dessas duas estações são bem superiores aos das demais, não é possível distinguir os parâmetros que caracterizam as demais estações da rede de controle de poluição. Para essa finalidade uma nova análise foi realizada excluindo as estações AV310 e AV320. O novo gráfico *biplot* de escores e cargas para esse subgrupo de estações é mostrado na FIGURA 26.

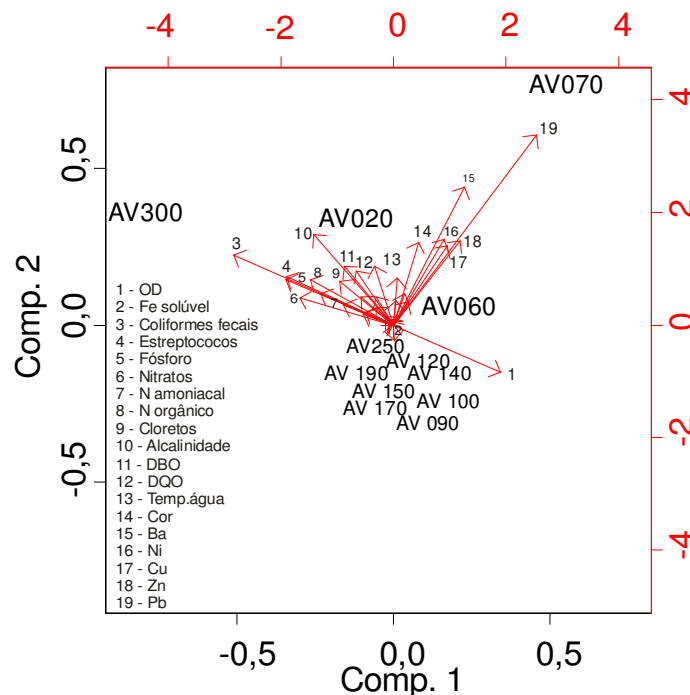


FIGURA 26- Gráfico *biplot* de escores e cargas escalonados para as duas primeiras componentes principais para as estações da rede dirigida para controle da poluição excluindo as estações AV310 e AV320. Escala de escores no intervalo de -1 a +1 e escala de pesos no intervalo de -5 a +4

Nessa análise é observada a discriminação das estações AV070, ribeirão Mata Porcos, próximo de sua confluência com o ribeirão Sardinha, caracterizada pelos parâmetros Pb, Ba, Ni, Zn e cor, característicos de atividades de mineração. Próximo a este ponto, a estação AV060, ribeirão Carioca a montante da confluência do ribeirão Mata Porcos também é distinguida pelos mesmos parâmetros evidenciando também a influência de rejeitos de mineração. A estação AV300, Córrego da Barragem em Nova Lima, é discriminada pela presença de coliformes fecais, estreptococos, fósforo, nitratos, nitrogênio amoniacal, nitrogênio orgânico e cloretos provenientes de efluentes sanitários. A estação AV020, Rio Maracujá a montante da confluência do Córrego dos Padres foi caracterizada pelos parâmetros alcalinidade, DBO e DQO, também indicativos da contaminação de efluentes sanitários. As demais estações são discriminadas pelo parâmetro oxigênio dissolvido, característico de água de boa qualidade.

O gráfico *biplot* de escores e cargas de cada parâmetro para as duas primeiras componentes principais para as estações da rede dirigida para abastecimento doméstico, constituída das estações AV220, AV230, AV240, AV260, AV270, AV280 e AV290 é mostrado na FIGURA 27.

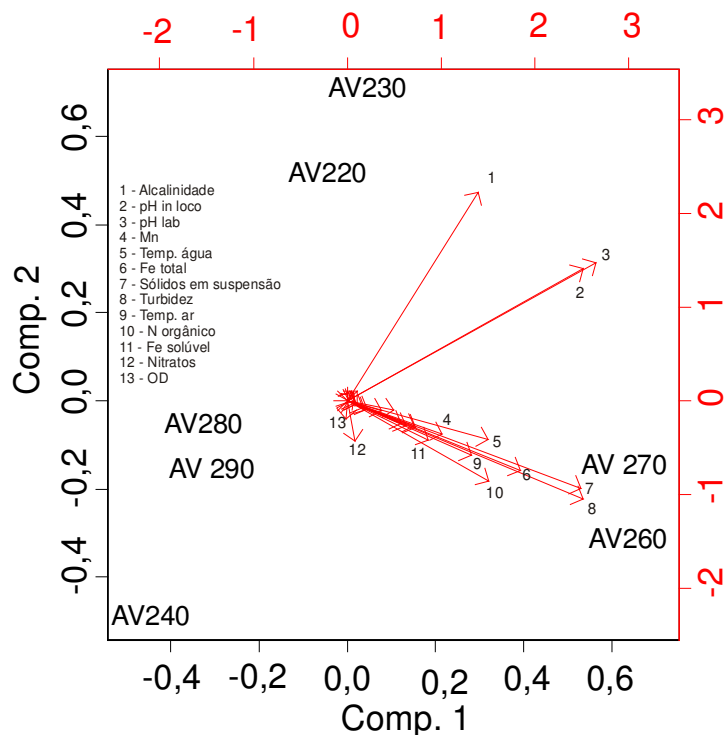


FIGURA 27- Gráfico *biplot* de escores e cargas para as duas primeiras componentes principais para as estações AV220, AV230, AV240, AV260, AV270, AV280 e AV290, que constituem a rede dirigida para abastecimento doméstico. Escala de escores no intervalo de -0,6 a +0,8 e escala de pesos no intervalo de -2,5 a +3,5.

Para as estações que compõem a rede dirigida ao abastecimento doméstico, as estações AV260, rio das Velhas em Bela Fama (alça direita), e AV270, rio das Velhas em Bela Fama (alça esquerda), são discriminadas pela pelos parâmetros sólidos em suspensão, turbidez e Fe total, indicativos da contaminação por efluentes de mineração. As estações AV220, Córrego dos Fechos na barragem principal, e AV230 Córrego dos Fechos na barragem auxiliar são discriminadas pelos parâmetros alcalinidade e pH. A discriminação por esses parâmetros não é considerada relevante, uma vez que essas estações apresentaram excelente IQA e baixa contaminação por tóxicos⁸. As estações AV240, AV280 e AV290 não são discriminadas por nenhum parâmetro indicativo de poluição, estando entre as menos poluídas da rede de abastecimento.

Com base na variabilidade explicada pelas componentes principais, foram avaliados os agrupamentos das estações considerando apenas os principais parâmetros de cada uma das quatro componentes principais mais importantes, ou seja, nitrogênio amoniacal, nitritos, K, Na, nitratos, condutividade; fósforo, estreptococos, coliformes fecais, surfactantes, DBO, DQO, bário, turbidez, ferro total, sólidos em suspensão, temperatura

do ar, cor, pH lab, pH in loco, ferro solúvel e alcalinidade, somando 22 parâmetros. O dendrograma obtido é mostrado na FIGURA 28.

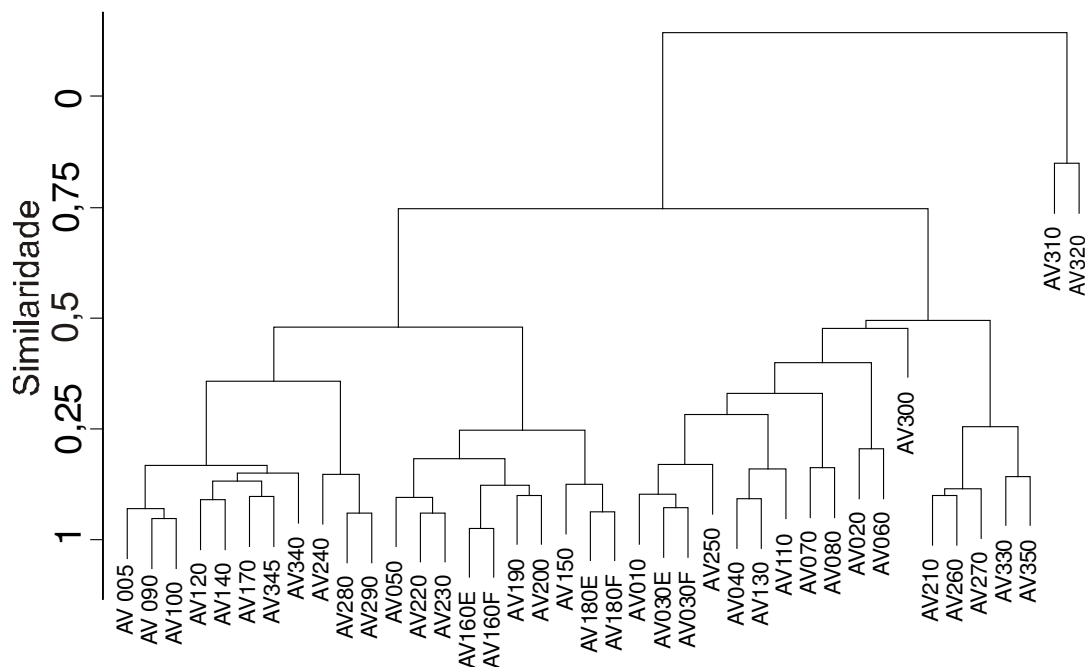


FIGURA 28- Dendrograma obtido na análise com os 22 principais parâmetros de cada uma das quatro componentes principais

A redução de parâmetros não produziu alterações significativas nos agrupamentos das estações por similaridade. Esta observação permite concluir que a caracterização e avaliação das estações poderiam ser realizadas apenas com o monitoramento destes parâmetros mais importantes.

Uma nova tabela de parâmetros a serem analisados pode ser proposta com base nos parâmetros mais importantes na discriminação das estações, selecionados a partir da análise de componentes principais e dos seus respectivos teores medianos. Os parâmetros considerados são nitrogênio amoniacal, nitritos, K, Na, nitratos, condutividade; fósforo, estreptococos, coliformes fecais, surfactantes, DBO, DQO, bário, turbidez, ferro total, sólidos em suspensão, temperatura do ar, cor, pH laboratorial, pH in loco, cor, ferro solúvel e alcalinidade. Os parâmetros pH laboratorial e pH *in loco* são empregados para controle e evidência de conservação das amostras. Eles contem a mesma informação o que fica evidente no tratamento estatístico.

7.4 Conclusões do capítulo

O tratamento estatístico dos dados de monitoramento da qualidade da água na região do Alto Curso do Rio das Velhas mostrou que apenas 22 parâmetros de qualidade descrevem as quatro mais importantes componentes principais, explicando aproximadamente 80% da variabilidade dos dados. Entre os principais parâmetros estão presentes variáveis que caracterizam o impacto causado pelo lançamento de esgoto sanitário, como nitrogênio amoniacal, nitratos, nitritos, fósforo, estreptococos, coliformes fecais, DBO e DQO. Esta observação nos permite concluir que o principal impacto causado no Alto Curso do Rio das Velhas também é atribuído a deficiências nos processos de tratamento e controle desses efluentes durante o período estudado.

Alguns parâmetros característicos de efluentes de mineração, sólidos em suspensão, turbidez, Fe total, estão entre os mais importantes, evidenciando o impacto ambiental causado por essa atividade na região. A presença de arsênio na estação AV310, ribeirão Água Suja a montante do rio das Velhas, que atingiu no passado teores alarmantes, chegando a alcançar o valor correspondente a 58 vezes o limite de classe 3¹³ em julho de 1995, vem sendo reduzida a partir de 2002 com fiscalizações efetivas da FEAM. Esta redução teve grande contribuição das obras de reabilitação ambiental do antigo depósito do Morro do Galo, localizado as margens do ribeirão Água Suja, e que constituía a fonte potencial de arsênio para o rio das Velhas. Não obstante, a contaminação por arsênio no alto curso do Rio das Velhas ainda não está totalmente eliminada.

Foi observada uma correlação entre pares de algumas estações da rede dirigida para o abastecimento, o que recomenda a eliminação de uma das estações correlacionadas, reduzindo o número de estações nesse grupo em futuro monitoramento. Um critério foi estabelecido, eliminando as estações que apresentam melhores valores de IQA e CT em relação às suas respectivas estações correlacionadas.

As estações AV240, AV280 e AV290 não são discriminadas por nenhum parâmetro, estando entre as de águas de melhor qualidade da rede dirigida para o abastecimento. As estações AV260 e AV270, localizadas no rio das Velhas em Bela Fama na alça direita e na alça esquerda respectivamente são distinguidas pelos parâmetros sólidos em

suspensão e turbidez, característicos de atividades minerárias e são as mais contaminadas desse grupo. Essas estações são importantes no abastecimento da cidade de Belo Horizonte e sua contaminação aumenta o custo no tratamento da água.

Nas estações da rede dirigida para o controle da poluição destacam-se as estações AV310, rio das Velhas a montante do ribeirão de Água Suja, e AV320, rio das Velhas a montante do Córrego da Mina, como as águas de pior qualidade de toda a rede de monitoramento. Por outro lado, a análise estatística destacou as estações onde a qualidade da água é melhor, particularmente a estação AV005, rio das Velhas a montante de São Bartolomeu.

A avaliação das principais variáveis responsáveis pelo agrupamento das estações permitiu a preparação de um novo mapa de monitoramento de parâmetros de controle de poluição a ser empregado em futuro monitoramento do alto curso da bacia do rio das Velhas.

Capítulo VIII – Análise temporal dos dados

Neste estudo foram selecionadas para análise temporal as estações mais críticas reveladas pelas análises de componentes principais, fatorial e agrupamentos. Uma vez realizada a análise exploratória dos dados monitorados na sub-bacia do rio das Velhas, a avaliação temporal dos dados tem como objetivo o estudo do comportamento de alguns parâmetros que apresentaram maiores índices de violação dos limites de classe de enquadramento^{1-8, 13}. Pretende-se dessa forma avaliar o comportamento dessas estações, detectando tendências e propondo ações mais efetivas.

8.1 Escolha dos parâmetros e estações

Os parâmetros que apresentaram os maiores percentuais de violação em toda bacia do rio das Velhas estão associados principalmente a efluentes de esgotos domésticos, tais como fosfato total, coliformes fecais, índice de fenóis, DBO, oxigênio dissolvido, nitrogênio amoniacal e efluentes de mineração, como manganês, turbidez, arsênio, cobre, níquel, como mostra a TABELA 8 do capítulo IV.

As análises de fatores e componentes principais realizadas nos capítulos V e VI, caracterizam a maioria dos parâmetros violados entre os fatores e as primeiras componentes principais. A caracterização das estações em relação aos parâmetros mais críticos nos gráficos *biplot*, FIGURAS 18 e 19 do capítulo VI, foi o critério empregado na escolha das estações submetidas à avaliação temporal.

8.2 Metodologia de análise

Os dados analisados correspondem às concentrações em mg/L de fosfatos, nitrogênio amoniacal nas estações BV154, ribeirão do Onça próximo de sua foz no rio das Velhas; BV155, ribeirão Arrudas próximo de sua foz no rio das Velhas; BV105, rio das Velhas logo a jusante do ribeirão do Onça e BV083, rio das Velhas a jusante do ribeirão Arrudas, manganês BV035/AV110, rio Itabirito a jusante do Córrego Cata Branca e BV083; e oxigênio dissolvido, OD, nas estações BV140, rio Jequitibá próximo de sua foz no rio das Velhas; BV162, rio Cipó a montante da foz do rio Paraúna; BV105 e BV154 nos meses de janeiro, abril, julho e dezembro.

A análise temporal compreende o período de janeiro de 1993 a novembro de 2005 para as estações BV083, BV105 e BV140; janeiro de 1994 a novembro de 2005 para as estações BV154 e BV155 e janeiro de 2000 a novembro de 2005 para a estação BV162. Como as estações BV154 e BV155 foram inseridas na macro-rede de monitoramento em maio de 1994 e a estação BV162 em maio de 2000; os dados censurados nas campanhas de janeiro e abril foram substituídos pela média anual da concentração no parâmetro no mês correspondente. Assim, nas análises temporais de fosfatos e nitrogênio amoniacal nas estações dos ribeirões do Onça, BV154, e Arrudas, BV155, as medidas de concentração nos meses de janeiro e abril de 1994, correspondem às médias nos respectivos meses nos anos de 1995 a 2004. Na estação BV162, as medidas de oxigênio dissolvido, nos meses de janeiro e abril do ano de 2000 correspondem às médias nos respectivos meses nos anos de 2001 a 2004.^{1-7, 91-93}

Inicialmente os dados foram ordenados em uma matriz, de elementos x_{ik} , onde i correspondente ao mês e ano da coleta e análise e k corresponde ao parâmetro analisado naquela estação.

Basicamente, como descrito no capítulo III, a análise temporal envolve a construção do gráfico da variável estudada em função do tempo para visualizar a existência ou não de componentes de tendência e sazonalidade, seleção de modelos mais adequados para descrever o comportamento da série e escolha do melhor modelo através da análise dos erros. É fundamental que a variação do parâmetro em estudo seja analisada em intervalos regulares de tempo.

Os métodos de alisamento exponencial podem ser aplicados em séries com um número de informações em relação ao tempo, N , inferior a 60; já os modelos autoregressivos ou método de Box e Jenkins requerem séries maiores, com $N > 60$. Como o número de medidas não ultrapassa a 48 nas estações analisadas, foram empregados nesse trabalho os métodos de alisamento exponencial.

8.3 Avaliação da concentração de fosfatos total

O fósforo é originado naturalmente da dissolução de compostos do solo e da decomposição da matéria orgânica. A origem antrópica é associada aos despejos domésticos e industriais, aos detergentes, excrementos de animais e fertilizantes. A presença de fósforo nos corpos d'água desencadeia o desenvolvimento de algas ou outras plantas aquáticas desagradáveis, principalmente em reservatórios ou águas paradas, podendo conduzir ao processo de eutrofização.¹⁻⁷

8.3.1 Estação ribeirão do Onça próximo de sua foz no rio das Velhas (BV154)

A avaliação temporal de fosfatos total nessa estação foi realizada empregando os dados coletados com periodicidade trimestral no período compreendido entre janeiro de 1994 a novembro de 2004, totalizando 44 medidas. Os resultados das campanhas em 2005 foram utilizados na comparação dos valores previstos por cada modelo.

A TABELA 18 apresenta a comparação dos valores das constantes de alisamento, coeficiente bayesiano, BIC, e erro quadrático médio de previsão, EQM, para os modelos avaliados para série.

TABELA 18 – Comparação de modelos de previsão para teores para fosfatos total no ribeirão do Onça, estação BV154*

MODELO	NÍVEL	TENDÊNCIA	SAZONALIDADE	BIC	EQM
Holt	0,0466	0,0596	-	1,68	1,54
Winter	0,168	-	0,132	1,27	1,16
Holt-Winter	0,127	0,0246	0,146	1,30	1,14

* As constantes de nível, tendência e sazonalidade são calculadas de forma a minimizar a soma do erro quadrático médio (EQM) de acordo com as equações 88 e 89. O EQM e o BIC são calculados pelas equações 83 e 96, respectivamente.

O modelo de Winter apresenta menor coeficiente bayesiano e o modelo de Holt – Winter apresenta menor erro quadrático de previsão.

A FIGURA 29 mostra os ajustes dos modelos de Winter e o modelo de Holt–Winter em relação aos dados históricos.

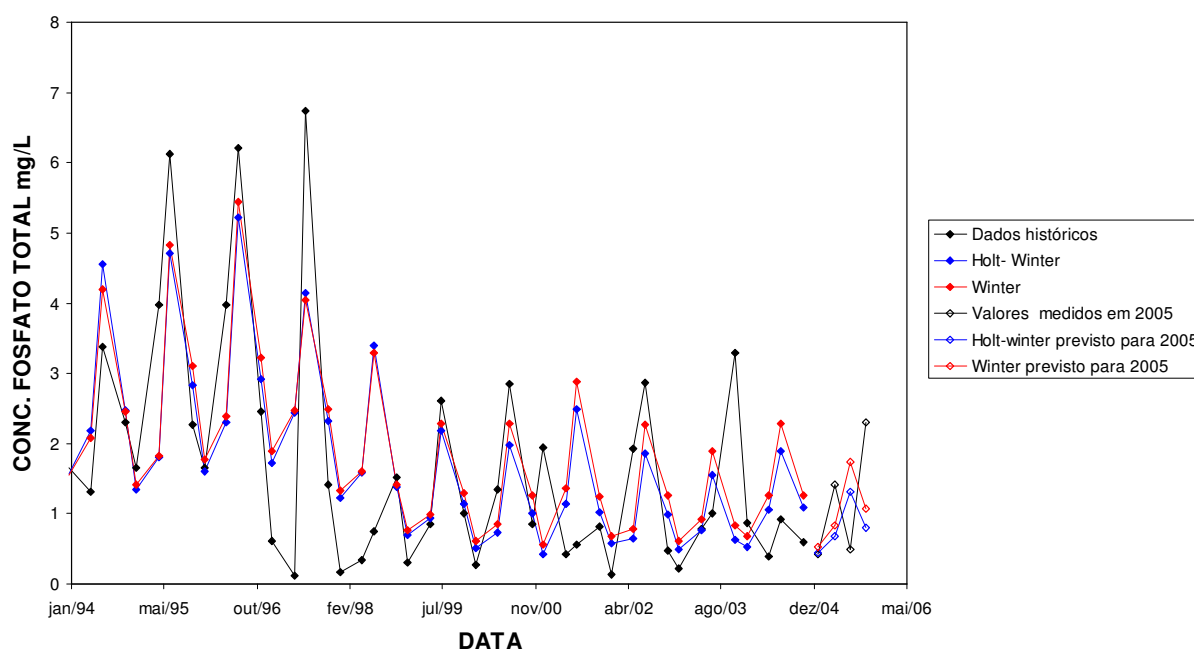


FIGURA 29 - Comparação dos dados históricos da concentração de fosfatos total no ribeirão do Onça, estação BV154, com os ajustes exponenciais dos modelos de Holt-Winter e Winter período de janeiro de 1994 a outubro de 2004 e as previsões em 2005

Os dados históricos, simbolizados com linha preta, apresentam variações sazonais, com concentrações mais baixas nos meses de janeiro e mais altas no mês de julho. É

observada uma tendência de decréscimo da concentração de fosfatos nessa estação, esse decréscimo foi mais significativo nas campanhas realizadas em 1998 e 2001.

Os modelos de Winter, simbolizado com linha vermelha e Holt–Winter, simbolizado com linha azul, apresentaram pequenas diferenças no ajuste com os dados históricos da concentração de fosfatos total na estação BV154; reproduzindo as variações sazonais durante quase todo período; com exceção das campanhas nos anos de 1998 e 2001.

O modelo de Winter apresenta valores mais próximos dos reais nos picos sazonais do parâmetro. Como as diferenças em cada modelo estão no fato do modelo de Winter, ser aplicado para séries sazonais e o de Holt-Winter para séries com tendência e sazonalidade, é observado que a tendência de decréscimo não é muito representativa na definição do modelo. Por essa razão, e também devido à pequena diferença em relação ao erro quadrático médio de previsão em relação ao modelo de Holt-Winter, o modelo de Winter pode ser escolhido como mais próximo no ajuste dos dados históricos.

A TABELA 19 apresenta os valores reais e previstos e seus limites mínimos e máximos para concentração de fosfatos total nos anos de 2005 a 2007 pelo modelo de Winter.

TABELA 19– Valores reais e previstos para a concentração de fosfatos total em mg/L no ribeirão do Onça, estação BV154, no período de 2005 a 2007 pelo modelo de Winter

DATA	PREVISÃO			VALOR REAL EM 2005
	Mínimo	Previsão	Máximo	
jan/05	0,0	0,437	1,21	0,420
abr/05	0,0	0,676	1,92	1,41
jul/05	0,0	1,33	3,87	0,490
nov/05	0,0	0,813	2,48	2,31
jan/06	0,0	0,367	1,25	
abr/06	0,0	0,563	1,88	
jul/06	0,0	1,09	3,68	
nov/06	0,0	0,664	2,39	
jan/07	0,0	0,297	1,28	
abr/07	0,0	0,450	1,84	
jul/07	0,0	0,862	3,48	
nov/07	0,0	0,516	2,30	

Pode-se observar pela TABELA 19, que a concentração de fosfatos no ribeirão do Onça apresentou em 2005 um comportamento diferente do observado na série nos anos de 1994 a 2004. Há um aumento da concentração nos meses de abril e novembro e uma diminuição nos meses de janeiro e julho. Por essa razão, os valores médios previstos para 2005 não são tão próximos dos valores reais. A concentração de fosfatos nessa estação está entre os limites mínimos e máximos de previsão.

As previsões para os anos de 2006 e 2007 indicam uma diminuição na concentração dos fosfatos totais nessa estação; indicando uma melhoria na qualidade das águas comparada aos últimos dez anos.

8.3.2 Rio das Velhas logo a jusante do ribeirão do Onça (BV105)

A estação BV105 é a primeira estação de amostragem no rio das Velhas localizada após o ribeirão do Onça. Os impactos da contaminação do ribeirão do Onça no rio das Velhas são avaliados pelos parâmetros monitorados nessa estação. Os teores de fosfatos totais foram avaliados com periodicidade trimestral no período compreendido entre janeiro de

1993 a novembro de 2004, totalizando 48 medidas. Os resultados das campanhas em 2005 foram utilizados na comparação dos valores previstos por cada modelo. A TABELA 20 apresenta a comparação dos valores das constantes de alisamento, coeficiente bayesiano, BIC, e erro quadrático médio de previsão, EQM, para os modelos avaliados.

TABELA 20 – Comparação de modelos de previsão para teores para fosfatos total no rio das Velhas a jusante do ribeirão Onça, estação BV105*

MODELO	NÍVEL	TENDÊNCIA	SAZONALIDADE	BIC	EQM
Holt	0,0675	0,0529	-	0,998	0,915
Winter	0,169	-	0,0683	0,921	0,845
Holt-Winter	0,0778	0,0297	0,0682	0,954	0,839

* As constantes de nível, tendência e sazonalidade são calculadas de forma a minimizar a soma do erro quadrático médio (EQM) de acordo com as equações 88 e 89. O EQM e o BIC são calculados pelas equações 83 e 96, respectivamente.

O modelo de Winter apresentou o menor coeficiente bayesiano e o modelo de Holt-Winter apresentou o menor erro quadrático de previsão.

A FIGURA 30 apresenta os ajustes dos modelos de Winter e o modelo de Holt – Winter em relação aos dados reais da concentração de fosfatos total.

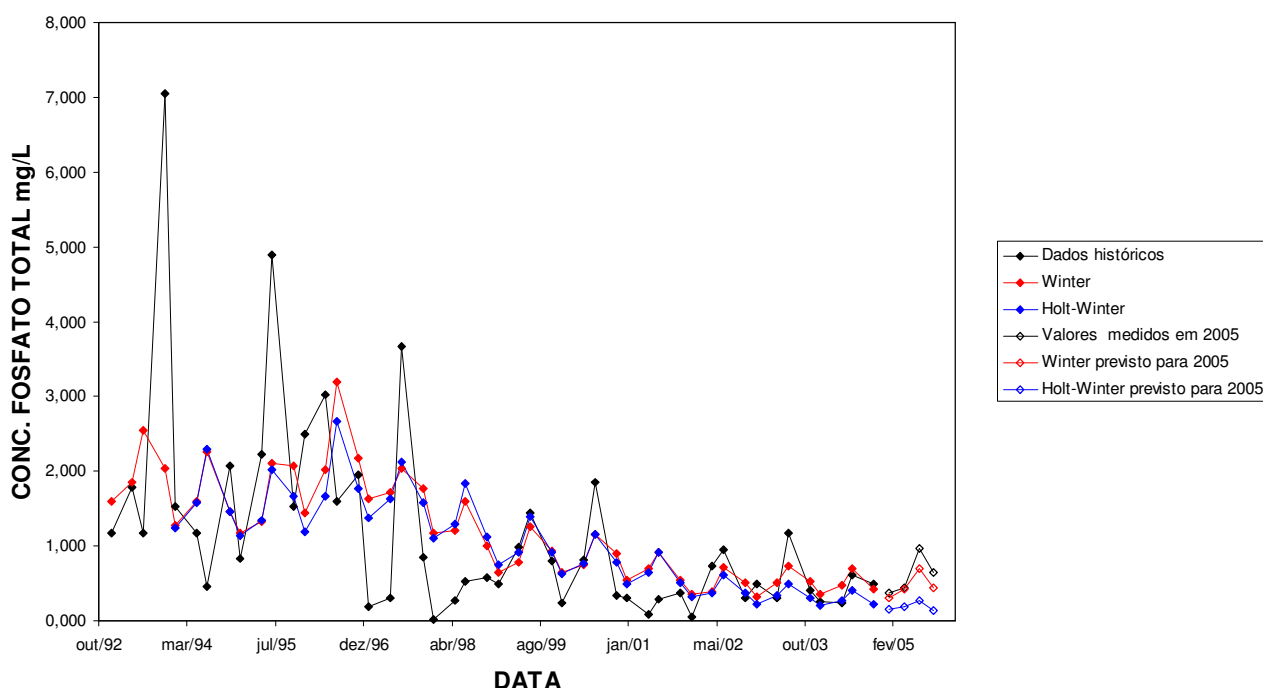


FIGURA 30 – Comparação dos dados históricos da concentração de fosfatos total no rio das Velhas a jusante do ribeirão do Onça, estação BV105, com os ajustes exponenciais dos modelos de Holt- Winter e Winter período de janeiro de 1993 a outubro de 2004 e as previsões em 2005

Os dados históricos, como na série anterior, apresentaram um decréscimo nos teores de fosfatos total a partir do ano de 1998. É observada uma tendência de decréscimo da concentração de fosfatos nessa estação, esse decréscimo foi mais significativo nas campanhas realizadas em 1998 e 2001. Os máximos na concentração de fosfatos acontecem nos meses de julho e as concentrações mais baixas nos meses de janeiro. Embora a concentração de fosfatos nessa estação seja mais baixa que no ribeirão do Onça, pode-se perceber a influência desse ribeirão na qualidade das águas dessa estação.

No ajuste aos dados históricos da série, o modelo de Winter apresenta valores mais próximos dos reais nos picos sazonais do parâmetro. Por essa razão, e também pela pequena diferença em relação ao erro quadrático médio de previsão em relação ao modelo de Holt- Winter; o modelo de Winter foi escolhido como mais próximo no ajuste dos dados históricos.

A TABELA 21 apresenta os valores reais e previstos e seus limites mínimos e máximos para concentração de fosfatos total nos anos de 2005 a 2007 pelo modelo de Winter.

TABELA 21 – Valores reais e previstos para a concentração de fosfato total em mg/L no rio das Velhas a jusante do ribeirão do Onça, estação BV105, no período de 2005 a 2007 pelo modelo de Winter

DATA	PREVISÃO			VALOR REAL EM 2005
	Mínimo	Previsão	Máximo	
jan/05	0,0	0,319	0,761	0,370
abr/05	0,0	0,426	1,02	0,450
jul/05	0,0	0,664	1,59	0,970
nov/05	0,0	0,509	1,23	0,650
jan/06	0,0	0,319	0,799	
abr/06	0,0	0,426	1,05	
jul/06	0,0	0,664	1,61	
nov/06	0,0	0,509	1,25	
jan/07	0,0	0,319	0,835	
abr/07	0,0	0,426	1,08	
jul/07	0,0	0,664	1,63	
nov/07	0,0	0,509	1,28	

A concentração de fosfatos medida em 2005 apresenta mínimo no mês de janeiro e máximo no mês de julho como observado na série temporal da concentração de fosfatos total nos anos de 1993 a 2004; por isso há um bom ajuste dos valores médios previstos pelo modelo de Winter e os valores reais. As previsões para os de 2006 e 2007 variam em relação a 2005 apenas nos máximos de previsão, indicando concentrações mais baixas que no ribeirão do Onça.

É importante observar que, embora as concentrações de fosfatos totais estejam mais baixas nessa estação que na anterior, nos meses de julho e novembro os limites estão superiores aos limites de fosfatos em ambiente lótico que é de 0,15 mg/L de fósforo e 0,46 mg/L de fosfatos para águas de classe 3.

8.3.3 Estação ribeirão Arrudas próximo de sua foz no rio das Velhas (BV155)

A estação BV155, ribeirão Arrudas próximo de sua foz no rio das Velhas, juntamente com o ribeirão do Onça, BV154, apresentaram as piores condições de qualidade do rio das Velhas. A avaliação temporal de fosfatos total nessa estação foi realizada empregando os dados coletados com periodicidade trimestral no período compreendido entre janeiro de 1994 a novembro de 2004, totalizando 44 medidas. Os resultados das campanhas em 2005 foram utilizados na comparação dos valores previstos por cada modelo.

A TABELA 22 apresenta a comparação dos valores das constantes de alisamento, coeficiente bayesiano, BIC, e erro quadrático médio de previsão, EQM, para os modelos avaliados para série.

TABELA 22 – Comparação de modelos de previsão para teores para fosfatos total no ribeirão Arrudas, estação BV155*

MODELO	NÍVEL	TENDÊNCIA	SAZONALIDADE	BIC	EQM
Holt	0,266	0,0204	-	2,12	1,94
Holt <i>Tendência amortecida</i>	0,251	0,334 0,986	-	2,21	1,94
Holt-Winter	0,265	0,0200	0,0245	2,20	1,94

* As constantes de nível, tendência e sazonalidade são calculadas de forma a minimizar a soma do erro quadrático médio (EQM) de acordo com as equações 88 e 89. O EQM e o BIC são calculados pelas equações 83 e 96, respectivamente.

A constante de amortecimento, introduzida no modelo Holt, tem a finalidade de suavizar as diferenças de tendência apresentadas pela série. O modelo temporal de Holt – Winter para essa estação apresenta menor erro de previsão e coeficiente bayesiano que os modelos de Holt com tendência amortecida e Holt.

A FIGURA 31 apresenta a comparação dos ajustes dos modelos de Holt com tendência amortecida e o modelo de Holt – Winter em relação aos dados reais.

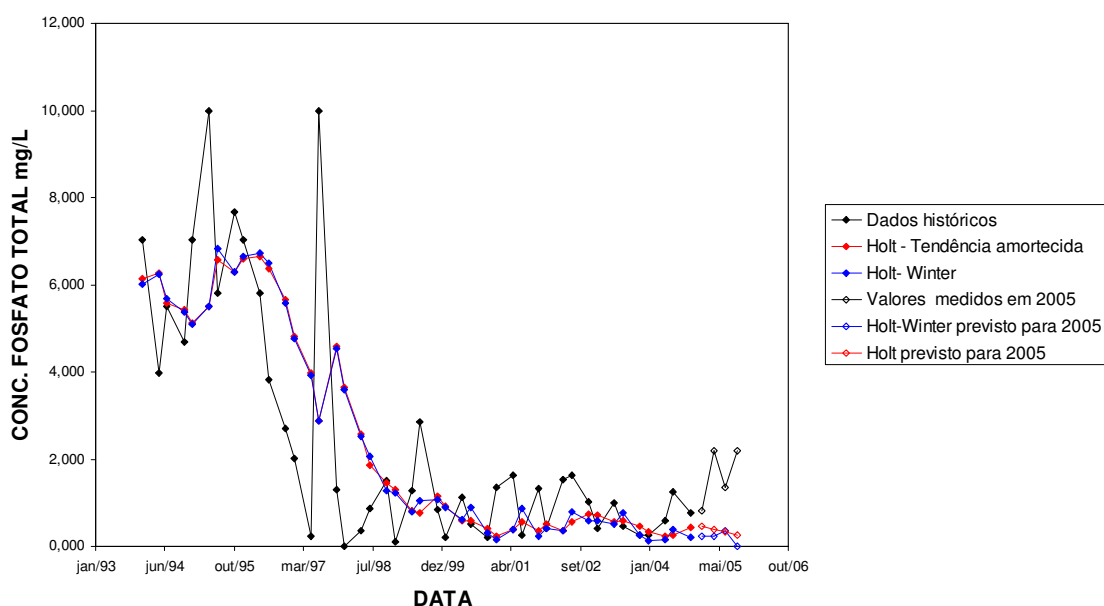


FIGURA 31– Comparação dos dados históricos da concentração de fosfatos total no ribeirão Arrudas, estação BV155, com os ajustes exponenciais dos modelos de Holt- Winter e Winter período de janeiro de 1994 a outubro de 2004 e as previsões em 2005

Pode-se observar, pelo comportamento dos dados históricos, que a partir das campanhas de 1998, as concentrações de fosfatos total no ribeirão Arrudas, BV155, apresentaram uma tendência de decréscimo como no ribeirão do Onça até as campanhas de 2005. As maiores concentrações foram observadas nos meses de julho e as menores em janeiro. As concentrações de fosfatos no ribeirão Arrudas foram mais altas que no ribeirão do Onça, BV154.

Embora, os valores dos coeficientes BIC e EQM para o modelo de Holt com constante de amortecimento para tendência não sejam menores que os coeficientes para o modelo de Holt–Winter, o modelo de Holt com tendência amortecida apresenta melhor ajuste os teores de fosfatos totais a partir de 2004.

A TABELA 23 apresenta os valores reais e previstos e seus limites mínimos e máximos para concentração de fosfatos total nos anos de 2005 a 2007 pelo modelo de Winter.

TABELA 23 - Valores reais e previstos para a concentração de fosfatos total em mg/L no ribeirão Arrudas, estação BV155, no período de 2005 a 2007 pelo modelo de Holt com tendência amortecida

DATA	PREVISÃO			VALOR REAL EM 2005
	Mínimo	Previsão	Máximo	
jan/05	0,0	0,449	1,53	0,810
abr/05	0,0	0,386	1,50	2,18
jul/05	0,0	0,323	1,47	1,35
nov/05	0,0	0,262	1,44	2,19
jan/06	0,0	0,201	1,41	
abr/06	0,0	0,142	1,39	
jul/06	0,0	0,0830	1,36	
nov/06	0,0	0,0250	1,33	
jan/07	0,0	0,000	1,30	
abr/07	0,0	0,000	1,28	
jul/07	0,0	0,000	1,25	
nov/07	0,0	0,000	1,22	

Os valores medidos de concentração de fosfatos nas campanhas de 2005 apresentaram um crescimento não observado nos anos anteriores, excedendo os limites máximos de previsão nos meses de abril e novembro. Por essa razão, as previsões para os anos 2006 e 2007 não refletem as condições reais da qualidade das águas nessa estação. Uma melhor previsão pode ser conseguida com a inclusão das campanhas do 2005 no modelo utilizado.

O aumento da concentração de fosfatos afeta diretamente a qualidade das águas no rio das Velhas, pois as águas do ribeirão Arrudas contaminam a estação BV083, rio das Velhas a jusante do ribeirão Arrudas. O aumento da concentração de fosfatos favorece o crescimento de algas que podem causar eutroficação. O aumento da matéria orgânica provocado pelo crescimento das algas provoca um aumento da demanda química de oxigênio necessária para sua oxidação e conseqüentemente provoca morte da fauna aquática.

8.3.4 Rio das Velhas logo a jusante do ribeirão Arrudas (BV083)

A estação BV083 é a primeira estação de amostragem no rio das Velhas localizada após o ribeirão Arrudas. Os impactos da contaminação do ribeirão Arrudas no rio das Velhas são avaliados pelos parâmetros monitorados nessa estação. Os teores de fosfatos totais foram avaliados com periodicidade trimestral no período compreendido entre janeiro de 1993 a novembro de 2004, totalizando 48 medidas. Os resultados das campanhas em 2005 foram utilizados na comparação dos valores previstos por cada modelo.

A TABELA 24 apresenta a comparação dos valores das constantes de alisamento, coeficiente bayesiano, BIC, e erro quadrático médio de previsão, EQM, para os modelos avaliados para série.

TABELA 24 – Comparação de modelos de previsão para teores para fosfatos total no rio das Velhas a jusante do ribeirão Arrudas, estação BV083*

MODELO	NÍVEL	TENDÊNCIA	SAZONALIDADE	BIC	EQM
Winter	0,0706	-	0,0433	0,755	0,697
Holt – Winter	0,0131	0,330	0,0416	0,790	0,672
<i>Tendência amortecida</i>		0,995			
Holt-Winter	0,0311	0,0414	0,0461	0,758	0,672

* As constantes de nível, tendência e sazonalidade são calculadas de forma a minimizar a soma do erro quadrático médio (EQM) de acordo com as equações 88 e 89. O EQM e o BIC são calculados pelas equações 83 e 96, respectivamente.

O modelo de Winter proposto para os dados históricos dessa estação apresenta menor coeficiente BIC e maior EQM que o modelo Holt – Winter.

A FIGURA 32 apresenta a comparação dos ajustes dos modelos de Winter e o modelo de Holt – Winter em relação aos dados reais.

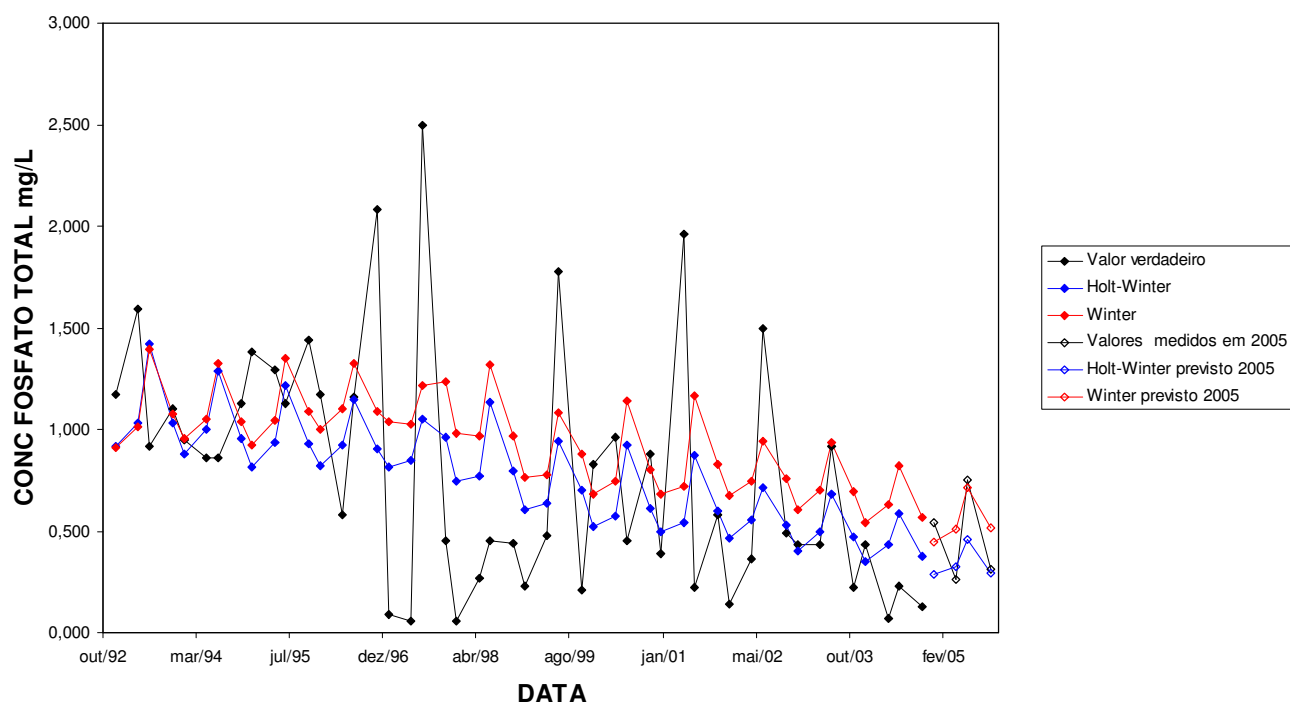


FIGURA 32 – Comparação dos dados históricos da concentração de fosfatos total no rio das Velhas a jusante do ribeirão Arrudas ,estação BV083, com os ajustes exponenciais dos modelos de Holt- Winter e Winter período de janeiro de 1994 a outubro de 2004 e as previsões em 2005

Os valores históricos para concentração de fosfatos nessa estação mantiveram-se mais baixos que na estação anterior e com concentrações máximas principalmente nos meses de julho e mínimas nos meses de janeiro. Há uma tendência de decrescimento nas concentrações de fosfatos até o ano de 2004 e o ano de 1998 apresentou concentrações mais baixas em todas as campanhas.

O modelo de Winter apresenta melhor ajuste aos dados reais, mas não acompanhou as concentrações máximas dos dados reais nos meses de julho.

A TABELA 25 apresenta os valores reais e previstos e seus limites mínimos e máximos para concentração de fosfatos total nos anos de 2005 a 2007 pelo modelo de Winter.

TABELA 25– Valores reais e previstos para a concentração de fosfatos total em mg/L no rio das Velhas a jusante do ribeirão Arrudas, estação BV083, em 2005 a 2007 pelo modelo de Winter

DATA	PREVISÃO			VALOR REAL EM 2005
	Mínimo	Previsão	Máximo	
jan/05	0,0	0,285	0,770	0,540
abr/05	0,0	0,327	0,911	0,260
jul/05	0,0	0,456	1,31	0,750
nov/05	0,0	0,293	0,874	0,310
jan/06	0,0	0,232	0,719	
abr/06	0,0	0,263	0,849	
jul/06	0,0	0,363	1,22	
nov/06	0,0	0,230	0,813	
jan/07	0,0	0,179	0,668	
abr/07	0,0	0,200	0,787	
jul/07	0,0	0,270	1,13	
nov/07	0,0	0,167	0,751	

A concentração de fosfatos em 2005 apresentou níveis mais altos nos meses de janeiro e julho e mais próximos aos valores máximos previstos pelo modelo de Winter. Esse comportamento é preocupante uma vez que ações de controle vêm sendo realizadas pela COPASA como a Estação de Tratamento de Esgoto, ETE, implantada no ribeirão Arrudas em 2002 com a finalidade de tratar grande parte dos esgotos dos municípios de Belo Horizonte e Contagem. Caso esse comportamento atípico seja mantido, as previsões para os anos 2006 e 2007 não refletem as condições reais da qualidade das águas nessa estação.

As concentrações máximas de fosfatos totais observadas em 2005 e nas campanhas de 1993, 1995 e 1997 no rio das Velhas logo a jusante do ribeirão do Onça são maiores que no rio das Velhas logo a jusante do ribeirão Arrudas evidenciando maiores impactos no rio das Velhas pelo ribeirão Onça por despejos com maior concentração de fosfatos totais.

8.4 Avaliação dos teores de nitrogênio amoniacal

O nitrogênio na forma amoniacal é uma substância tóxica não persistente e, em baixas concentrações, não causa nenhum dano fisiológico aos seres humanos e animais. Grandes quantidades de amônia podem causar sufocamento aos peixes. A concentração total de nitrogênio é altamente importante considerando-se os aspectos tópicos do corpo d'água. Aumentos súbitos indicam contaminação recente por esgotos domésticos e efluentes orgânicos.¹⁻⁷

8.4.1 Estações ribeirão do Onça próximo de sua foz no rio das Velhas (BV154) e rio das Velhas logo a jusante do ribeirão do Onça (BV105)

Como mencionado anteriormente, as qualidades das águas nas estações BV154 e BV105 estão diretamente relacionada. As concentrações de nitrogênio amoniacal foram avaliadas com periodicidade trimestral, no período compreendido entre janeiro de 1994 a novembro de 2004 para o ribeirão do Onça, BV154, e janeiro de 1993 a novembro de 2004 para o rio das Velhas a jusante do ribeirão do Onça, BV105. Os resultados das campanhas em 2005 foram utilizados na comparação dos valores previstos por cada modelo.

A TABELA 26 apresenta a comparação dos valores das constantes de alisamento, coeficiente bayesiano, BIC, e erro quadrático médio de previsão, EQM, para os modelos avaliados para concentração de nitrogênio amoniacal em cada estação.

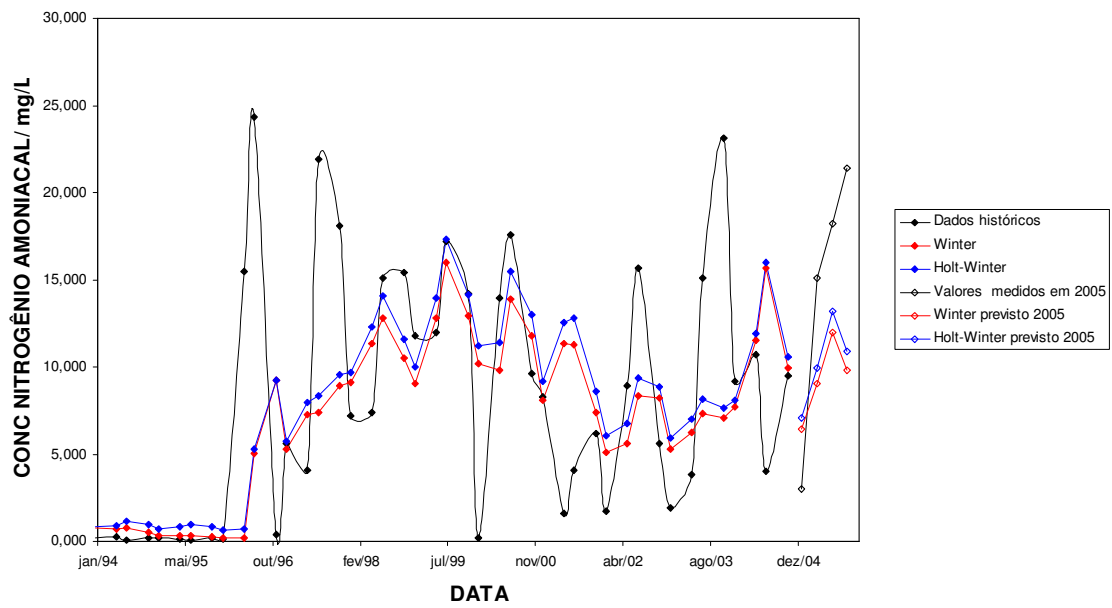
TABELA 26 – Comparação de modelos de previsão para teores para nitrogênio amoniacal nas estações BV154 e BV105

ESTAÇÃO	MODELO	NÍVEL	TENDÊNCIA	SAZONALIDADE	BIC	EQM
BV154	Winter	0,254	-	0,0623	7,15	6,56
	Holt – Winter <i>Tendência amortecida</i>	0,256	$3,00 \times 10^{-5}$ 0,703	0,0625	7,79	6,56
	Holt-Winter	0,233	0,0247	0,0542	7,42	6,52
BV105	Winter	0,981	-	0,689	2,34	2,16
	Holt – Winter	0,319	0,00548	0,125	2,70	2,39
	Holt	0,275	0,00838	-	2,86	2,64

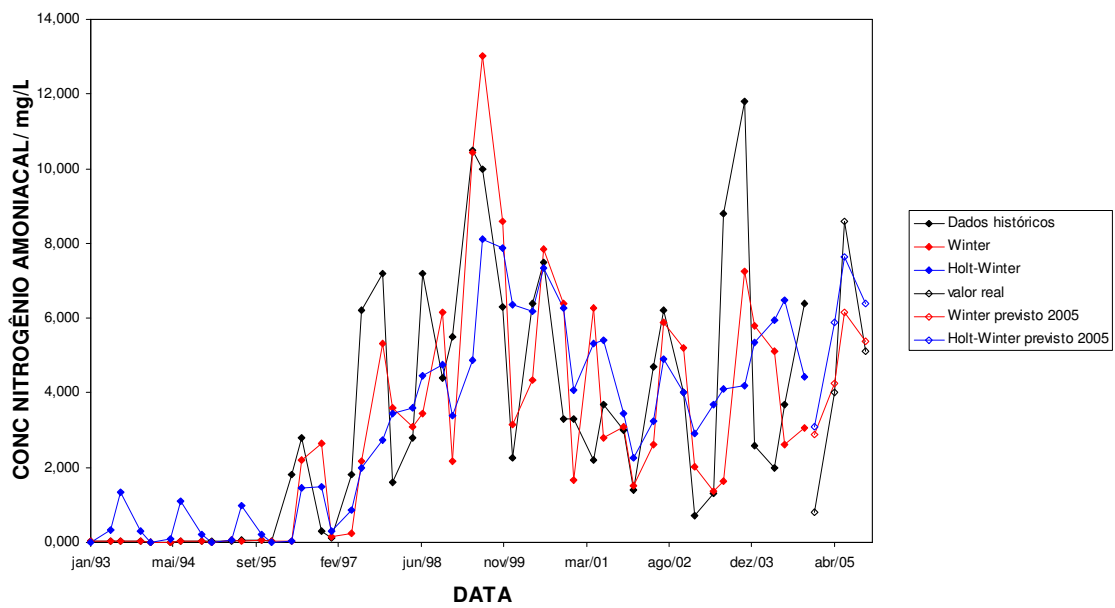
* As constantes de nível, tendência e sazonalidade são calculadas de forma a minimizar a soma do erro quadrático médio (EQM) de acordo com as equações 88 e 89. O EQM e o BIC são calculados pelas equações 83 e 96, respectivamente.

O modelo de Winter proposto para os dados históricos da estação BV154 apresenta menor coeficiente BIC e o modelo Holt – Winter menor EQM. Para estação BV105, o modelo de Winter série apresenta menores coeficientes BIC e EQM.

A FIGURA 33 apresenta a comparação entre os ajustes do modelo de Winter e o modelo de Holt – Winter em relação aos dados reais para cada estação.



PAINEL (A)



PAINEL (B)

FIGURA 33 - Comparação dos dados históricos da concentração de nitrogênio amoniacal no ribeirão do Onça, PAINEL (A), e no rio da Velhas a jusante do ribeirão do Onça, PAINEL (B), com os ajustes exponenciais dos modelos de Winter e Holt- Winter

Os dados históricos mostram que as variações da concentração de nitrogênio amoniacal na estação do ribeirão do Onça, PAINEL A, apresentam comportamento sazonal e tendência de crescimento. As concentrações são maiores nos meses de julho e menores nos meses de janeiro. As baixas concentrações observadas até abril de 1996 podem estar associadas à metodologia de análise empregada pelo laboratório até essa época, uma vez que, esse crescimento repentino não foi observado nas séries de concentrações de fosfatos totais. Para a estação BV105, PAINEL B, que apresentou concentrações menores que o ribeirão do Onça, também é observado comportamento sazonal com concentrações mais altas nos meses de julho e novembro. Há tendência de crescimento e as baixas concentrações até abril de 1996 apresentam a mesma justificativa.

O modelo de Holt – Winter apresenta melhor ajuste aos dados históricos no período de janeiro de 1994 a outubro de 2004 e também uma melhor previsão para o ano de 2005 no ribeirão do Onça. Na estação do rio das Velhas a jusante do ribeirão do Onça, embora o modelo de Winter tenha melhor ajuste aos valores reais de janeiro de 1993 a outubro de 2004, a previsão feita em 2005 pelo modelo de Holt – Winter é mais próxima dos valores reais.

A TABELA 27 apresenta os valores previstos e limites mínimos e máximos pelo modelo de Holt - Winter para o ano de 2005 para as duas estações.

TABELA 27– Valores reais e previstos para concentração de nitrogênio amoniacal em mg/L no ribeirão do Onça e no rio das Velhas a jusante do ribeirão do Onça de 2005 a 2007 pelo modelo de Holt–Winter

DATA	PREVISÃO								VALOR REAL EM 2005	
	Mínimo		Previsão				Máximo			
	BV154	BV105	BV154	BV154*	BV105	BV105*	BV154	BV105	BV154	BV105
jan/05	0,0	0,0	6,41	6,29	3,08	2,51	18,4	11,3	3,0	0,800
abr/05	0,0	0,0	9,05	8,08	5,89,	3,92	26,3	23,5	15,1	4,00
jul/05	0,0	0,0	12,0	12,8	7,65	6,38	35,1	34,4	18,2	8,60
nov/05	0,0	0,0	9,82	12,1	6,38	5,89	29,8	37,9	21,4	5,10
jan/06	0,0	0,0	6,41	6,29	3,08	2,51	21,5	35,4		
abr/06	0,0	0,0	9,05	8,08	5,89,	3,92	28,5	41,8		
jul/06	0,0	0,0	12,0	12,8	7,65	6,38	36,8	48,8		
nov/06	0,0	0,0	9,82	12,1	6,38	5,89	31,8	50,8		
jan/07	0,0	0,0	6,41	6,29	3,08	2,51	24,0	48,1		
abr/07	0,0	0,0	9,05	8,08	5,89,	3,92	30,6	53,5		
jul/07	0,0	0,0	12,0	12,8	7,65	6,38	38,5	59,4		
nov/07	0,0	0,0	9,82	12,1	6,38	5,89	33,6	60,7		

*Valores previstos com séries iniciadas em janeiro de 1997

A concentração de nitrogênio amoniacal foi aumentando em cada campanha na estação do ribeirão do Onça não apresentado o mesmo comportamento sazonal apresentado nos anos anteriores. As concentrações medidas nas campanhas de 2005 estão entre os valores médios e máximos previstos pelo modelo de Holt – Winter.

Na estação do rio das Velhas a jusante do ribeirão do Onça, a variação na concentração de nitrogênio amoniacal apresenta comportamento sazonal, mas seus valores são mais altos que nas campanhas de 2004. As concentrações medidas nas campanhas de 2005 são menores que no ribeirão do Onça e mais próximas aos valores médios de previsão.

As previsões médias para os anos de 2006 e 2007 apresentam os mesmos valores que em 2005 e diferem apenas em relação às concentrações máximas.

Para mostrar de forma inequívoca a qualidade das previsões do método de séries temporais, foram apresentadas na TABELA 27 as previsões para os teores de nitrogênio amoniacal nas estações BV105 e BV154 com e sem a inclusão das campanhas anteriores a janeiro de 2007. Nas estações destacadas com asterisco na referida tabela os teores de nitrogênio amoniacal foram previstos sem a inclusão das campanhas anteriores a janeiro de 1997 pelo modelo de Holt–Winter. Os resultados das mesmas previsões utilizando as séries iniciadas em janeiro de 1994 para estação BV154 e janeiro 1993 para estação BV105 também estão representados na referida tabela. As pequenas diferenças nas previsões para os anos de 2005 a 2007 demonstram a propriedade do método de séries temporais, confirmando a expectativa que as baixas concentrações observadas até dezembro de 1996 não prejudicariam as estimativas das previsões. Assim demonstra-se que as variações nas campanhas mais recentes são mais importantes na estimativa das previsões das concentrações que as das primeiras campanhas e que, mesmo sendo aparentemente discrepantes, os dados das primeiras campanhas não devem ser eliminados *a priori* da base de dados.

O aumento na concentração de nitrogênio amoniacal nas campanhas de 2005 no ribeirão do Onça, BV154, que influencia diretamente a qualidade das águas do rio das Velhas, estação BV105, é preocupante uma vez que as maiores concentrações medidas no mês de julho na estação BV105 são superiores ao limite de classe 3 que corresponde a 13,3 mg/L de N para $\text{pH} \leq 7,5$. O acompanhamento das várias formas de nitrogênio ao longo de um determinado trecho de um rio indica qual é a capacidade desse corpo d'água para degradar e transformar a carga orgânica nitrogenada e, principalmente, qual é a sua capacidade de assimilar determinadas classes de resíduos. A presença de nitrogênio na forma amoniacal indica que a matéria orgânica presente no rio ainda não foi oxidada.

8.4.2 Estações no ribeirão Arrudas próxima de sua foz no rio das Velhas (BV155) e no rio das Velhas logo a jusante do ribeirão Arrudas (BV083)

Devido a sua localização, a qualidade das águas nessas estações está diretamente relacionada. As concentrações de nitrogênio amoniacal foram avaliadas com periodicidade trimestral, no período compreendido entre janeiro de 1994 a novembro de 2004 para o ribeirão Arrudas, BV155, e janeiro de 1993 a novembro de 2004 no rio das Velhas a jusante do ribeirão Arrudas, BV083. Os resultados das campanhas em 2005

foram utilizados na comparação dos valores previstos das concentrações de nitrogênio amoniacal por cada modelo.

A TABELA 28 apresenta a comparação dos valores das constantes de alisamento, coeficiente bayesiano, BIC e erro quadrático médio de previsão, EQM, para os modelos avaliados para concentração de nitrogênio amoniacal em cada estação.

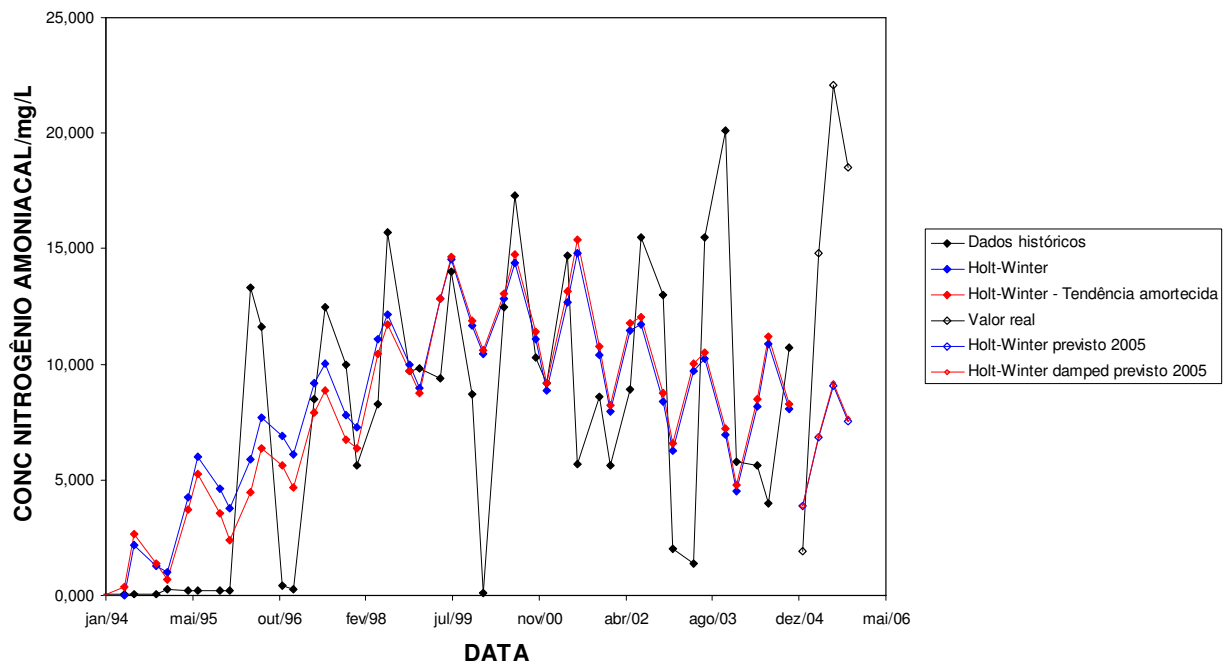
TABELA 28 – Comparação de modelos de previsão para teores para nitrogênio amoniacal no ribeirão Arrudas, BV155 e no rio das Velhas a jusante do ribeirão Arrudas, BV083

ESTAÇÃO	MODELO	NÍVEL	TENDÊNCIA	SAZONALIDADE	BIC	EQM
BV155	Holt	0,0423	0,990	-	5,40	4,96
	Holt – Winter <i>Tendência amortecida</i>	0,0483	0,735 0,983	0,107	5,41	4,55
	Holt-Winter	0,0371	0,997	0,100	5,19	4,56
BV083	Winter	0,161	-	0,0674	3,28	3,03
	Holt – Winter <i>Tendência amortecida</i>	0,0366	0,722 0,982	0,0506	3,52	2,99
	Holt-Winter	0,137	$8,00 \times 10^{-5}$	0,0590	3,47	3,07

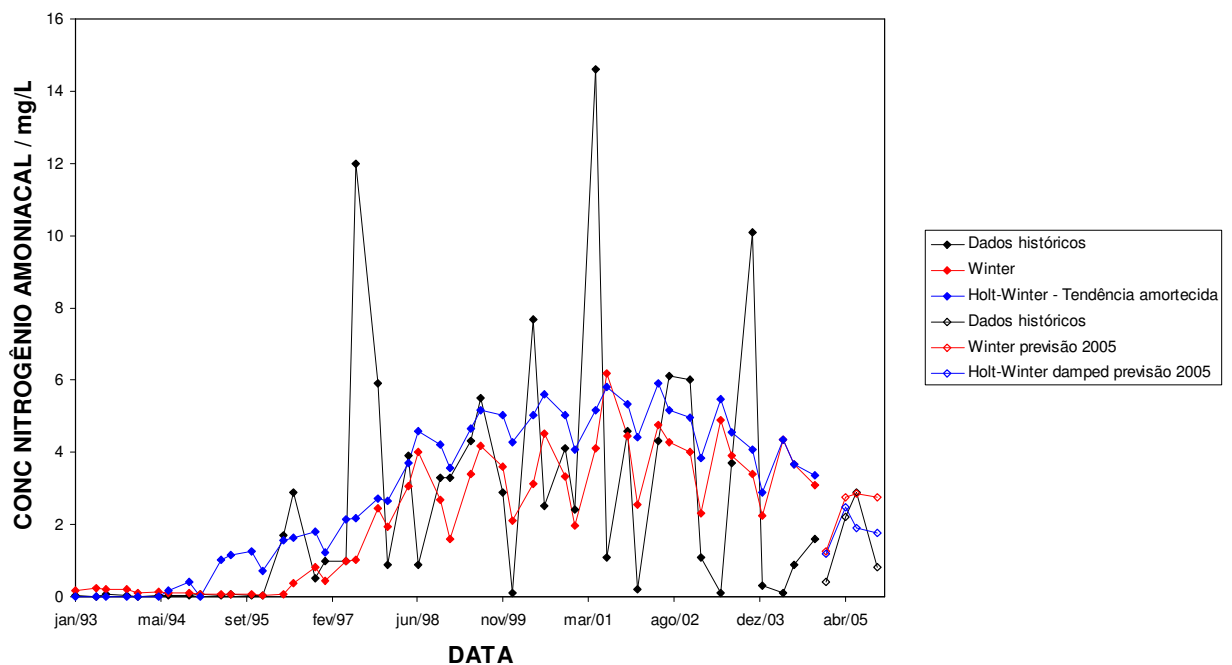
* As constantes de nível, tendência e sazonalidade são calculadas de forma a minimizar a soma do erro quadrático médio (EQM) de acordo com as equações 88 e 89. O EQM e o BIC são calculados pelas equações 83 e 96, respectivamente.

O modelo de Holt – Winter apresenta menor coeficiente BIC e o Holt – Winter com tendência amortecida menor EQM para os dados históricos da estação BV155. Para a estação BV083, o modelo de Winter proposto para os dados históricos apresenta menores coeficientes BIC e o modelo Holt – Winter com tendência amortecida menor EQM.

A FIGURA 34 apresenta a comparação dos ajustes exponenciais desses modelos com os dados históricos para teores de nitrogênio amoniacal para cada estação.



PAINEL (A)



PAINEL (B)

FIGURA 34 - Comparação dos dados históricos da concentração de nitrogênio amoniacal ribeirão Arrudas, PAINEL (A), e rio das Velhas a jusante do ribeirão Arrudas, PAINEL (B), com os ajustes exponenciais dos modelos de Holt - Winter e Holt- Winter com tendência amortecida

Os dados históricos mostram que as variações da concentração de nitrogênio amoniacal na estação do ribeirão Arrudas, PAINEL A, apresentam comportamento sazonal com máximo nos meses de julho e novembro e tendência de crescimento. As concentrações são maiores nos meses de julho e menores nos meses de janeiro. A estação no rio das Velhas a jusante do ribeirão Arrudas, PAINEL B, apresenta concentrações mais baixas de nitrogênio amoniacal que o ribeirão Arrudas, também é observado o comportamento sazonal com concentrações mais altas nos meses de julho e novembro e tendência de decréscimo. As baixas concentrações registradas até abril de 1996 devem estar associadas à metodologia de análise então empregada.

As previsões dos teores de nitrogênio amoniacal em 2005 são semelhantes nos dois modelos, com o modelo de Holt – Winter com tendência amortecida apresentando ajustes e maiores limites máximos dos valores previstos para a estação BV155. Para estação BV083 o modelo de Holt – Winter com tendência amortecida também produziu melhor ajuste aos dados históricos.

A TABELA 29 apresenta os valores previstos e limites mínimos e máximos dos teores de nitrogênio amoniacal pelo modelo de Holt – Winter com tendência amortecida para as duas estações.

TABELA 29 – Valores reais e previstos para a concentração de nitrogênio amoniacal em mg/L na estação BV155 e BV083 de 2005 a 2007 pelo modelo de Holt – Winter com tendência amortecida

DATA	PREVISÃO								VALOR REAL EM 2005	
	Mínimo		Previsão				Máximo			
	BV155	BV083	BV155	BV155*	BV083	BV083*	BV155	BV083	BV155	BV083
jan/05	0,0	0,0	3,87	5,41	1,25	1,51	14,0	4,282	1,90	0,400
abr/05	0,0	0,0	6,87	8,55	2,76	3,77	17,1	9,51	14,8	2,20
jul/05	0,0	0,0	9,09	12,0	2,84	3,56	19,3	9,84	22,1	2,90
nov/05	0,0	0,0	7,55	11,6	2,77	4,28	17,8	9,69	18,5	0,800
jan/06	0,0	0,0	3,02	5,41	1,25	1,51	13,3	4,86		
abr/06	0,0	0,0	6,01	8,55	2,76	3,77	16,3	9,79		
jul/06	0,0	0,0	8,23	12,0	2,84	3,56	18,6	10,1		
nov/06	0,0	0,0	6,70	11,6	2,77	4,28	17,1	10,0		
jan/07	0,0	0,0	2,16	5,41	1,25	1,51	12,6	5,36		
abr/07	0,0	0,0	5,16	8,55	2,76	3,77	15,6	10,1		
jul/07	0,0	0,0	7,38	12,0	2,84	3,56	17,8	10,4		
nov/07	0,0	0,0	5,85	11,6	2,77	4,28	16,3	10,2		

*Valores previstos com séries iniciadas em janeiro de 1997

O significativo aumento da concentração de nitrogênio amoniacal no ribeirão Arrudas observado em 2005 excedeu ao comportamento da série nos anos anteriores e por isso não foram verificadas boas previsões. As concentrações nos meses de julho e novembro de 2005 são superiores aos limites máximos de previsão. Por essa razão, as previsões para os anos 2006 e 2007 não refletem as condições reais da qualidade das águas nessa estação.

Apesar da concentração de nitrogênio amoniacal no rio das Velhas a jusante do ribeirão Arrudas ter aumentado no ano de 2004, em relação ao ano de 2003, esse crescimento não refletiu no comportamento da série e as concentrações medidas nas campanhas de 2005 estão entre os limites mínimos e médios de previsão. As concentrações previstas de nitrogênio amoniacal para os anos de 2006 e 2007 também estão abaixo do limite estabelecido para águas de classe 3, que é de 13,3 mg/L N.

Para testar a influência das baixas concentrações de nitrogênio amoniacal observadas até janeiro de 1997, foram comparados os resultados das análises temporais utilizando o modelo Holt – Winter com tendência amortecida empregando as séries iniciadas em janeiro de 1997 e as séries iniciadas em janeiro de 1994 para a estação BV155 e iniciadas em janeiro 1993 para estação BV083. Observam-se previsões de concentrações de nitrogênio amoniacal maiores quando são utilizadas as séries menores iniciadas em janeiro de 1997. Essa diferença está associada a maior influência do aumento atípico na concentração de nitrogênio amoniacal observada no ano de 2004 na série menor.

8.5 Avaliação dos teores de manganês

A presença de manganês é bastante marcante no rio das Velhas e a ocorrência deste elemento está relacionada principalmente ao fato da sub-bacia do rio das Velhas, sobretudo seu alto curso, estar inserida na região do Quadrilátero Ferrífero do Estado de Minas Gerais. A presença de manganês pode também estar relacionada aos lançamentos das indústrias do ramo siderúrgico e têxtil. O manganês destaca-se pelo maior número de violações aos limites da legislação observadas na maioria dos trechos monitorados. O limite máximo de manganês em águas de classe 3 é de 0,5 mg/L.¹⁻⁷

8.5.1 Estação rio das Velhas logo a jusante do ribeirão Arrudas (BV083)

Os teores de manganês foram avaliados na estação rio das Velhas logo a jusante do ribeirão Arrudas com periodicidade trimestral no período compreendido entre janeiro de 1993 a novembro de 2004, totalizando 48 medidas. Os resultados das campanhas em 2005 foram utilizados na comparação dos valores previstos por cada modelo.

A TABELA 30 apresenta a comparação dos valores das constantes de alisamento, coeficiente bayesiano, BIC, e erro quadrático médio de previsão, EQM, para os modelos avaliados para série de medidas de teores de manganês total no rio das Velhas a jusante do ribeirão Arrudas.

TABELA 30 - Comparação de modelos de previsão para teores de manganês total no rio das Velhas a jusante do ribeirão Arrudas, BV083*

MODELO	NÍVEL	TENDÊNCIA	SAZONALIDADE	BIC	EQM
Winter	0	-	0,204	0,709	0,655
Holt – Winter	$2,00 \times 10^{-5}$	0,0972	0,202	0,738	0,655
Holt-Winter <i>Tendência amortecida</i>	0	0,0993 0,985	0,202	0,767	0,655

* As constantes de nível, tendência e sazonalidade são calculadas de forma a minimizar a soma do erro quadrático médio (EQM) de acordo com as equações 88 e 89. O EQM e o BIC são calculados pelas equações 83 e 96, respectivamente.

O modelo de Winter apresenta menor coeficiente BIC em relação aos demais. Os três modelos de ajuste apresentam o mesmo valor de erro quadrático médio em relação aos dados reais de teores de manganês total.

A FIGURA 35 apresenta uma comparação dos ajustes dos modelos de Winter e Holt – Winter com tendência amortecida em relação aos valores reais de teores de manganês total e previsão para o ano de 2005.

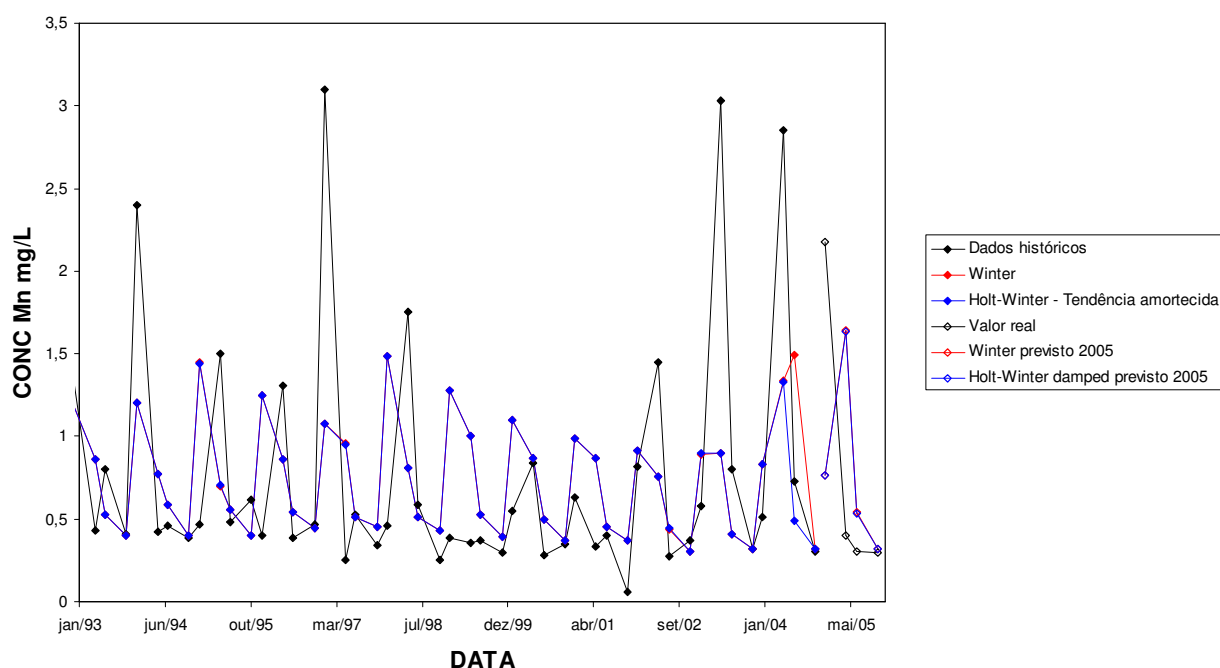


FIGURA 35 - Comparação dos dados históricos da concentração de manganês no rio das Velhas a jusante do ribeirão Arrudas, BV083, com os ajustes exponenciais dos modelos de Winter e Holt – Winter com tendência amortecida período de janeiro de 1993 a outubro de 2004 e as previsões em 2005

Os dados históricos de concentração de manganês total apresentam comportamento sazonal, com maiores concentrações nos meses de abril e janeiro. Não é observada nenhuma tendência de crescimento ou decrescimento da concentração de manganês nesse período. Os ajustes de cada modelo apresentaram valores coincidentes durante o período de janeiro de 1993 a janeiro de 2004. As diferenças de ajustes embora pequenas são mais destacadas a partir de julho de 2004. Os valores previstos para 2005 são também coincidentes para os dois modelos.

A TABELA 31 apresenta os valores previstos de teores de manganês total e limites mínimos e máximos pelo modelo de Winter para o ano de 2005.

TABELA 31 - Valores reais e previstos para a concentração de manganês em mg/L no rio das Velhas a jusante do ribeirão Arrudas no período de 2005 a 2007 pelo modelo de Winter

DATA	PREVISÃO			VALOR REAL EM 2005
	Mínimo	Previsão	Máximo	
jan/05	0,0	0,764	3,91	2,18
abr/05	0,0	1,64	3,75	0,402
jul/05	0,0	0,539	1,23	0,303
nov/05	0,0	0,316	0,712	0,297
jan/06	0,0	0,698	1,59	
abr/06	0,0	1,64	3,75	
jul/06	0,0	0,539	1,23	
nov/06	0,0	0,316	0,721	
jan/07	0,0	0,698	1,59	
abr/07	0,0	1,64	3,75	
jul/07	0,0	0,54	1,23	
nov/07	0,0	0,316	0,721	

No mês de janeiro de 2005 a concentração de manganês no rio das Velhas a jusante do ribeirão Arrudas foi maior que nos anos anteriores e manteve-se mais baixa nos meses seguintes. O valor real em janeiro de 2005 está entre os limites médio e máximo de previsão e nos meses seguintes os valores medidos estão entre os limites mínimos e médios.

A previsão para os anos de 2006 e 2007 estima concentrações de manganês total acima no limite de classe 3 para os meses de janeiro, abril e maio.

As concentrações mais altas de manganês total nos meses mais chuvosos como janeiro e menores em meses de estiagem como julho e novembro indica o carreamento desse metal do solo para o rio em períodos chuvosos

8.5.2 Rio Itabirito a jusante da cidade de Itabirito (BV035/ AV110)

A presença de manganês nessa estação está associada às atividades de mineração. Os teores de manganês foram avaliados com periodicidade trimestral no período compreendido entre janeiro de 1993 a novembro de 2004, totalizando 48 medidas. Os resultados das campanhas em 2005 foram utilizados na comparação dos valores previstos por cada modelo.

A TABELA 32 apresenta a comparação dos valores das constantes de alisamento, coeficiente bayesiano, BIC, e erro quadrático médio de previsão, EQM, para os modelos avaliados para série de teores de manganês total.

TABELA 32 - Comparação de modelos de previsão para teores de manganês total no rio Itabirito, BV035*

MODELO	NÍVEL	TENDÊNCIA	SAZONALIDADE	BIC	EQM
Winter	0,00262	-	0,120	1,62	1,50
Holt – Winter	0	0,0204	0,120	1,69	1,50
Holt-Winter <i>Tendência amortecida</i>	0,00213	0,180 0,922	0,116	1,76	1,50

* As constantes de nível, tendência e sazonalidade são calculadas de forma a minimizar a soma do erro quadrático médio (EQM) de acordo com as equações 88 e 89. O EQM e o BIC são calculados pelas equações 83 e 96, respectivamente.

O modelo de Winter apresenta menor coeficiente BIC em relação aos demais. O mesmo valor de erro quadrático médio em relação aos dados reais foi determinado nos três modelos de ajuste.

A FIGURA 36 apresenta uma comparação dos ajustes com os modelos de Winter e Holt – Winter com tendência amortecida em relação aos valores reais de concentração de manganês total e suas previsões para o ano de 2005.

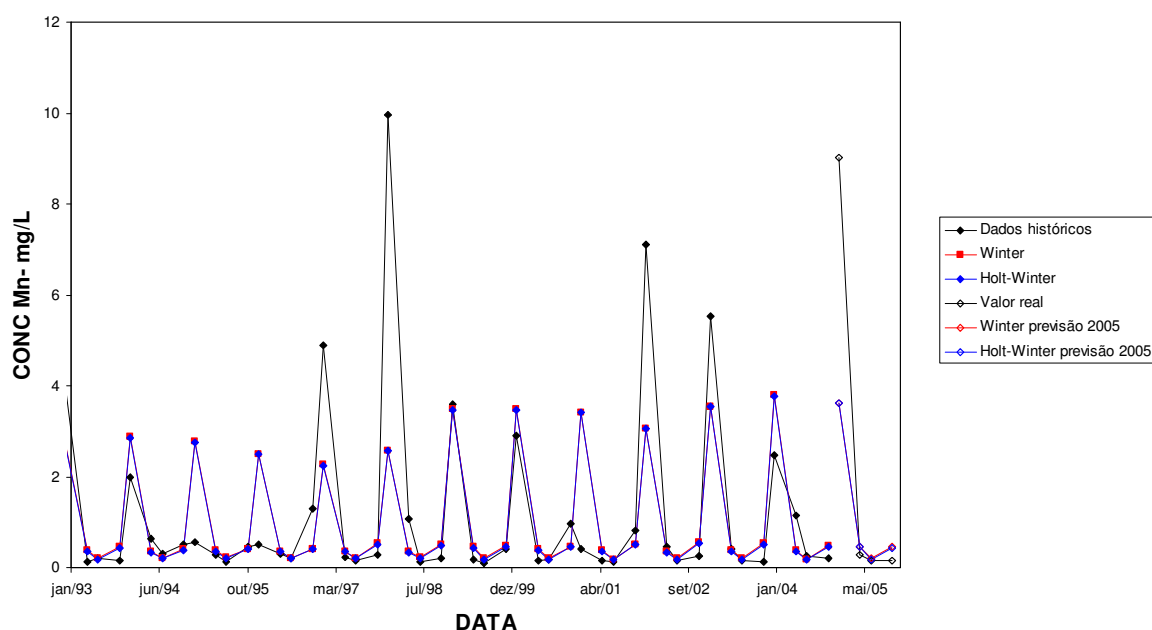


FIGURA 36 - Comparação dos dados históricos da concentração de manganês no rio Itabirito com os ajustes exponenciais dos modelos de Winter e Holt – Winter período de janeiro de 1993 a outubro de 2004 e as previsões em 2005

Os dados históricos apresentam comportamento sazonal com concentrações de manganês total maiores nos meses de janeiro e menores em julho. Não se observa tendência nos dados durante o período analisado. Os ajustes com cada modelo apresentam valores coincidentes durante o período de janeiro de 1993 a janeiro de 2004. As diferenças de ajustes embora pequenas são observadas a partir de julho de 2004. Os valores previstos para 2005 são também coincidentes com os dois modelos, mas o modelo de Winter foi escolhido por ter apresentado menor valor de BIC.

A TABELA 33 apresenta os valores previstos e limites mínimos e máximos de concentração de manganês total empregando o modelo de Winter, para o ano de 2005.

TABELA 33 – Valores reais e previstos para manganês no rio Itabirito, BV035, no período de 2005 a 2007 pelo modelo de Winter

DATA	PREVISÃO			VALOR REAL EM 2005
	Mínimo	Previsão	Máximo	
jan/05	0,451	3,63	6,81	9,01
abr/05	0	0,469	3,65	0,271
jul/05	0	0,196	3,37	0,147
nov/05	0	0,449	3,63	0,154
jan/06	0,451	3,63	6,81	
abr/06	0	0,469	3,65	
jul/06	0	0,194	3,37	
nov/06	0	0,449	3,63	
jan/07	0,451	3,63	6,81	
abr/07	0	0,469	3,65	
jul/07	0	0,194	3,37	
nov/07	0	0,449	3,63	

Com exceção do mês de janeiro, a concentração de manganês no rio Itabirito em 2005 foi menor que nos anos anteriores e ficou abaixo dos limites de classe 3. As concentrações medidas nos meses de abril, julho e novembro de 2005 são mais baixas que os limites médios de previsão. Embora os limites médios e máximos da concentração de manganês prevista para o mês de janeiro sejam bem maiores que nos outros meses a concentração medida nesse mês ainda excede esses valores.

As previsões para os anos de 2006 e 2007 estimam concentrações de manganês no rio Itabirito menores que o valor limite de classe 3 nos meses de abril, julho e novembro.

8.6 Avaliação dos teores de oxigênio dissolvido

O oxigênio dissolvido é essencial para a manutenção de processos de autodepuração em sistemas aquáticos naturais e estações de tratamento de esgotos. Durante a estabilização da matéria orgânica, as bactérias fazem uso do oxigênio nos seus processos respiratórios, podendo vir a causar uma redução de sua concentração no meio. Através da medição do teor de oxigênio dissolvido, os efeitos de resíduos oxidáveis sobre águas receptoras e a eficiência do tratamento dos esgotos, durante a oxidação bioquímica, podem ser avaliados. A concentração de oxigênio dissolvido em águas de classe 1 não deve ser inferior a 6 mg/L e em águas de classe 3 não deve ser inferior a 4 mg/L.¹⁻⁷

8.6.1 Estações ribeirão do Onça próximo de sua foz no rio das Velhas (BV154) e rio das Velhas logo a jusante do ribeirão do Onça (BV105)

A avaliação da concentração de oxigênio dissolvido nessas estações tem como finalidade avaliar o comportamento desse parâmetro indicativo de água de boa qualidade e compará-lo com os outros parâmetros analisados indicativos de poluentes como fosfatos totais e nitrogênio amoniacal

As concentrações de oxigênio dissolvido foram avaliadas com periodicidade trimestral, no período compreendido entre janeiro de 1994 a novembro de 2004 no ribeirão do Onça, na estação BV154, e janeiro de 1993 a novembro de 2004 no rio das Velhas a jusante do ribeirão do Onça, na estação BV105. Os resultados das campanhas em 2005 foram utilizados na comparação dos valores previstos por cada modelo.

A TABELA 34 apresenta a comparação dos valores das constantes de alisamento, coeficiente bayesiano, BIC, e erro quadrático médio de previsão, EQM, para os modelos avaliados para as séries de valores de concentração de oxigênio dissolvido nessas estações.

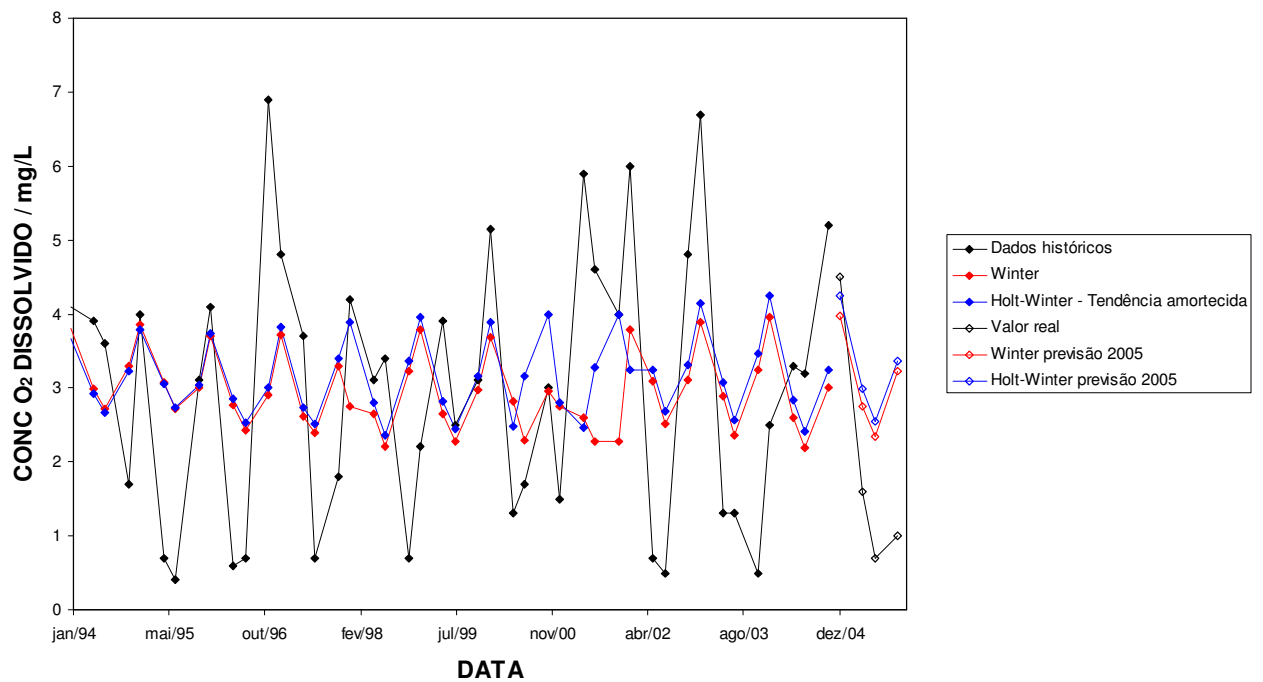
TABELA 34- Comparação de modelos de previsão para a concentração de oxigênio dissolvido em mg/L na estação ribeirão do Onça e rio das Velhas a jusante do ribeirão do Onça, BV154 e BV105

ESTAÇÃO	MODELO	NÍVEL	TENDÊNCIA	SAZONALIDADE	BIC	EQM
BV154	Winter	0,0276	-	0,0752	1,88	1,73
	Holt– Winter	0,0218	0,0169	0,0702	1,96	1,72
	Holt-Winter <i>Tendência amortecida</i>	0,0103	0,219 0,965	0,0597	2,04	1,71
BV105	Winter	0	-	0,0980	1,61	1,49
	Holt – Winter	0	0,0630	0,0964	1,68	1,49
	Holt-Winter <i>Tendência amortecida</i>	0	0,0413 0,991	0,0987	1,75	1,49

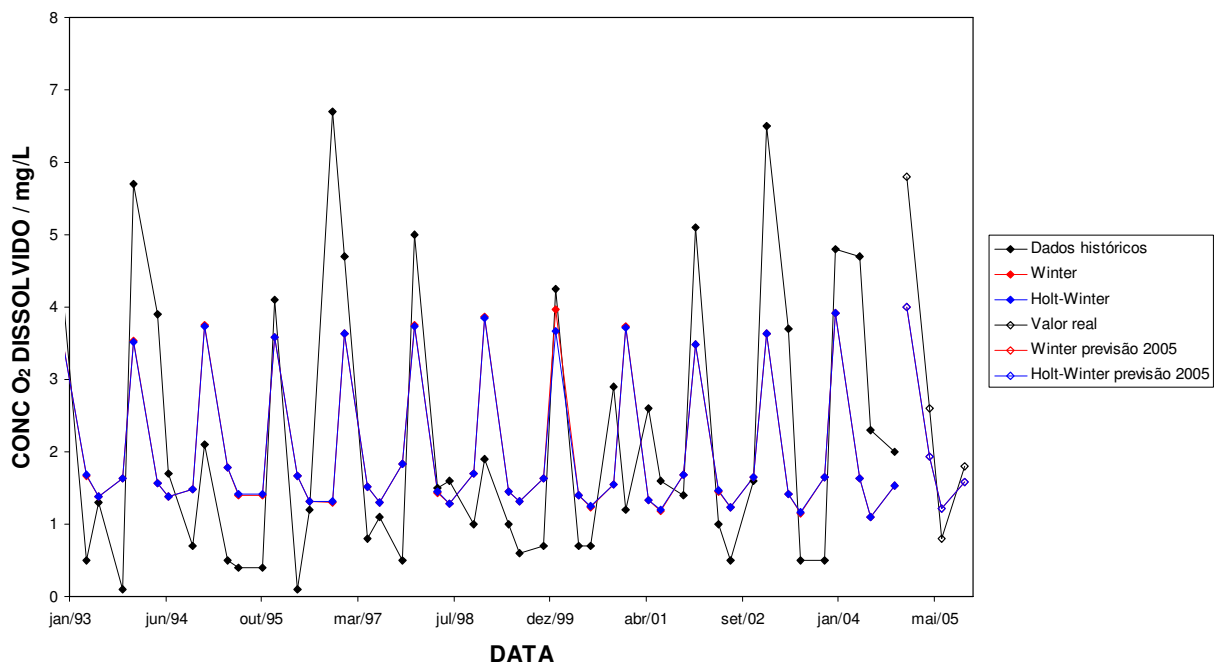
* As constantes de nível, tendência e sazonalidade são calculadas de forma a minimizar a soma do erro quadrático médio (EQM) de acordo com as equações 88 e 89. O EQM e o BIC são calculados pelas equações 83 e 96, respectivamente.

O modelo de Winter, aplicado aos dados de concentração de oxigênio dissolvido do ribeirão do Onça, apresenta menor valor do coeficiente BIC em relação aos demais e o modelo de Holt –Winter com tendência amortecida apresenta valor mais baixo do coeficiente EQM. No rio das Velhas a jusante do ribeirão do Onça o modelo de Winter apresenta menor coeficiente BIC em relação aos demais. Nos dados de cada estação os três modelos de ajuste apresentam o mesmo valor de erro quadrático médio.

A FIGURA 37 apresenta uma comparação dos ajustes dos modelos de Winter e Holt – Winter com tendência amortecida em relação aos valores reais e previsão para o ano de 2005 de concentração de oxigênio dissolvido do ribeirão do Onça.



PAINEL (A)



PAINEL (B)

FIGURA 37 - Comparação dos dados históricos da concentração de oxigênio dissolvido no ribeirão do Onça, PAINEL (A), com os ajustes exponenciais pelos modelos de Winter e Holt – Winter com tendência amortecida e Winter e Holt – Winter no rio das Velhas a jusante do ribeirão do Onça, PAINEL (B)

Os dados históricos mostram que o mesmo comportamento sazonal é observado no ribeirão do Onça, BV154 (PAINEL A) e no rio das Velhas a jusante do ribeirão do Onça, BV105 (PAINEL B) com maiores concentrações de oxigênio dissolvido nos meses de janeiro e menores nos meses de julho. Esse comportamento sazonal é contrário às variações nas concentrações de fosfatos e nitrogênio amoniacal que são maiores em julho e menores em janeiro. Esse comportamento antagônico é justificado pelo maior consumo de oxigênio dissolvido para deposição de material orgânico.

No ribeirão do Onça, os modelos de Winter e Holt – Winter com tendência amortecida não são muito eficientes no ajuste ao comportamento dos dados históricos e também na previsão dos dados em 2005 principalmente nas concentrações de oxigênio dissolvido mais baixas. O modelo de Holt – Winter com tendência amortecida apresentou melhor ajuste às concentrações mais altas de oxigênio dissolvido.

No rio das Velhas a jusante do ribeirão do Onça, os ajustes e previsão de concentrações de oxigênio dissolvido pelos modelos de Winter e Holt–Winter são coincidentes, apresentando uma pequena diferença em janeiro de 1999. Por ter apresentado menor BIC, o modelo de Winter foi escolhido para determinar as previsões de concentração de oxigênio dissolvido em 2005.

A TABELA 35 apresenta os valores previstos e limites mínimos e máximos pelos modelos de Holt-Winter com tendência amortecida para estação BV154 e Winter para estação BV105.

TABELA 35 – Valores reais e previstos para concentração de oxigênio dissolvido em mg/L nas estações ribeirão do Onça e no rio das Velhas a jusante do ribeirão do Onça no período de 2005 a 2007 pelos modelos de Holt – Winter com tendência amortecida e Winter respectivamente

DATA	PREVISÃO						VALOR REAL EM 2005	
	Mínimo		Previsão		Máximo		BV154	BV105
	BV154	BV105	BV154	BV105	BV154	BV105		
jan/05	0,0	0,0	3,98	4,01	8,88	10,2	4,50	5,80
abr/05	0,0	0,0	2,76	1,94	6,17	4,95	1,60	2,60
jul/05	0,0	0,0	2,34	1,21	5,24	3,09	0,700	0,800
nov/05	0,0	0,0	3,22	1,58	7,20	4,03	1,00	1,80
jan/06	0,0	0,0	3,98	4,01	8,88	10,2		
abr/06	0,0	0,0	2,76	1,94	6,17	4,95		
jul/06	0,0	0,0	2,34	1,21	5,25	3,09		
nov/06	0,0	0,0	3,22	1,58	7,21	4,03		
jan/07	0,0	0,0	3,98	4,01	8,89	10,2		
abr/07	0,0	0,0	2,76	1,94	6,18	4,95		
jul/07	0,0	0,0	2,34	1,21	5,25	3,09		
nov/07	0,0	0,0	3,22	1,58	7,21	4,03		

As concentrações medidas de oxigênio dissolvido nos meses de julho e abril foram menores que os valores médios estimados para as duas estações indicando maiores impactos de lançamentos de matérias no rio nesses meses. As concentrações de oxigênio dissolvido nos meses de janeiro e novembro, embora maiores que nos outros meses, são inferiores aos limites de classe 3, com exceção do mês de janeiro. Os valores estimados nas duas estações são mais próximos nos reais nos meses de janeiro e novembro e as previsões para os anos de 2006 e 2007 estimam concentrações de oxigênio dissolvido menores que o limite de classe 3.

Embora as concentrações de fosfato totais e nitrogênio amoniacal sejam menores no rio das Velhas, a jusante do ribeirão do Onça, que no próprio ribeirão do Onça, as concentrações de oxigênio dissolvido não apresentaram grandes diferenças nessas estações. Esse comprometimento da qualidade das águas do rio das Velhas deve ser contornado com ações mais rígidas no controle dos efluentes de ribeirões que passam por cidades.

8.6.2 Ribeirão Jequitibá próximo de sua foz no rio das Velhas (BV140)

A estação BV140 foi caracterizada pela concentração de nitratos como indica o gráfico *biplot* da análise de componentes principais apresentada no capítulo 6. A presença de uma significativa concentração de nitratos caracteriza a capacidade de autodepuração dos compostos nitrogenados presentes na matéria orgânica presente na água. A capacidade de autodepuração é maior quanto maior for a concentração de oxigênio dissolvido. Assim sendo, essa estação foi escolhida para análise temporal das concentrações de oxigênio dissolvido para avaliação da capacidade de autodepuração da matéria orgânica em seu curso.

Os teores de oxigênio dissolvido foram avaliados com periodicidade trimestral no período compreendido entre janeiro de 1993 a novembro de 2004, totalizando 48 medidas. Os resultados das campanhas em 2005 foram utilizados na comparação dos valores previstos por cada modelo.

A TABELA 36 apresenta a comparação dos valores das constantes de alisamento, coeficiente bayesiano, BIC e erro quadrático médio de previsão, EQM, para os modelos empregados para a série temporal teores de oxigênio dissolvido na estação BV140.

TABELA 36 - Comparação de modelos de previsão para teores de oxigênio dissolvido no ribeirão Jequitibá, estação BV140*

MODELO	NÍVEL	TENDÊNCIA	SAZONALIDADE	BIC	EQM
Winter	$2,30 \times 10^{-4}$	-	0,0866	0,977	0,901
Holt – Winter	0,0260	0,00358	0,0836	1,02	0,908
Holt-Winter <i>Tendência amortecida</i>	$8,30 \times 10^{-4}$	0,0255 0,916	0,0859	1,06	0,901

* As constantes de nível, tendência e sazonalidade são calculadas de forma a minimizar a soma do erro quadrático médio (EQM) de acordo com as equações 88 e 89. O EQM e o BIC são calculados pelas equações 83 e 96, respectivamente.

O modelo de Winter apresenta menores coeficientes BIC e EQM em relação aos demais.

A FIGURA 38 apresenta uma comparação dos ajustes dos modelos de Winter e Holt – Winter em relação aos valores reais e previsão para o ano de 2005.

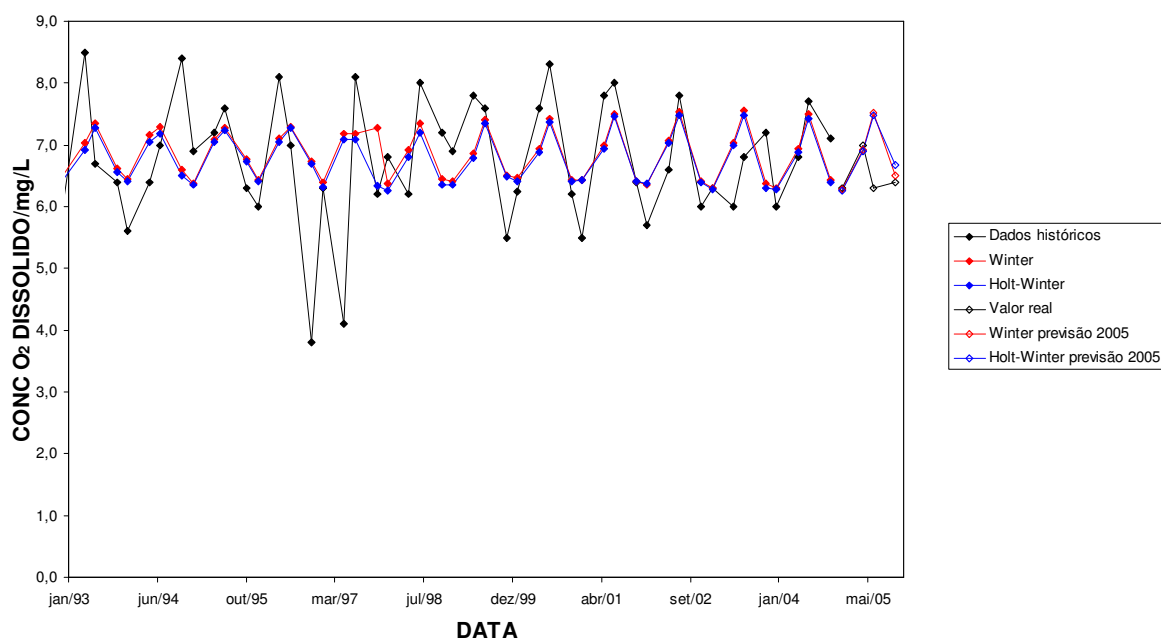


FIGURA 38 - Comparação dos dados históricos da concentração de oxigênio dissolvido no ribeirão Jequitibá com os ajustes exponenciais dos modelos de Winter e Holt – Winter período de janeiro de 1994 a outubro de 2004 e as previsões em 2005

Os dados históricos apresentaram comportamento sazonal com concentrações maiores nos meses de julho e menores em janeiro. Os meses mais críticos novembro de 1996 e abril de 1997 com concentrações de oxigênio dissolvido próximas a 4,0 mg/L. Nos outros meses as concentrações mínimas de oxigênio são próximas ao limite de águas de classe 2, que é de 5,0 mg/L. A diminuição da concentração de oxigênio dissolvido pode estar associada ao aumento de lançamento de efluentes com maior concentração de matéria orgânica. A estação BV140 apresentou concentração mais alta de oxigênio dissolvido dos que as estações BV154 e BV105, confirmando assim sua melhor qualidade também indicada pela presença de nitratos.

As variações nas concentrações de oxigênio dissolvido foram mais regulares e também apresentaram níveis mais baixos nos meses de janeiro e novembro. Os modelos de Winter e Holt – Winter apresentaram pequena diferença nos ajustes e previsões do parâmetro e o modelo de Winter foi escolhido para as previsões por ter apresentado menores coeficientes BIC e EQM e por não ser observada tendência importante nos dados.

A TABELA 37 apresenta os valores previstos e limites mínimos e máximos pelos modelos de Winter com tendência amortecida para estação BV140.

TABELA 37 - Valores reais e previstos para concentração de oxigênio dissolvido em mg/L na estação no ribeirão Jequitibá no período de 2005 a 2007 pelo modelo de Winter

DATA	PREVISÃO			VALOR REAL EM 2005
	Mínimo	Previsão	Máximo	
jan/05	4,50	6,28	8,05	6,30
abr/05	4,96	6,92	8,89	7,00
jul/05	5,38	7,51	9,64	6,30
nov/05	4,66	6,50	8,34	6,40
jan/06	4,50	6,28	8,05	
abr/06	4,96	6,92	8,89	
jul/06	5,38	7,51	9,64	
nov/06	4,66	6,50	8,34	
jan/07	4,50	6,28	8,05	
abr/07	4,96	6,92	8,89	
jul/07	5,38	7,51	9,64	
nov/07	4,66	6,50	8,34	

As concentrações menores de oxigênio dissolvido no ano de 2005 são superiores ao limite de águas de classe 1, caracterizando a boa qualidade das águas do ribeirão Jequitibá. Os limites médios previstos para as campanhas de 2005 são próximos aos valores reais, com exceção do valor previsto para o mês de julho que apresentou valor real inferior aos outros meses. A concentração de oxigênio dissolvido na campanha de julho de 2005 foi menor que nos outros anos.

As previsões para os de 2006 e 2007 apresentam o mesmo comportamento que 2005.

8.6.3 Rio Cipó a montante da foz do rio Paraúna (BV162)

A estação BV162 foi destacada com uma das estações com melhor índice de qualidade de águas entre as estações monitoradas ao longo do rio das Velhas, no estudo apresentado no capítulo 6. A avaliação temporal dos dados nessa estação se justifica e tem como finalidade a avaliação da qualidade da águas durante o período de janeiro de

2000 a novembro de 2004, totalizando 20 medidas feitas. Os resultados das campanhas em 2005 foram utilizados na comparação dos valores previstos por cada modelo.

A TABELA 38 apresenta a comparação dos valores das constantes de alisamento, coeficiente bayesiano, BIC, e erro quadrático médio de previsão, EQM, para os modelos avaliados para concentração de oxigênio dissolvido nessa estação.

TABELA 38 - Comparação de modelos de previsão para concentração de oxigênio dissolvido no rio Cipó, BV162*

MODELO	NÍVEL	TENDÊNCIA	SAZONALIDADE	BIC	EQM
Winter	0,221	-	0,228	0,558	0,481
Holt – Winter	0,220	$6,00 \times 10^{-5}$	0,225	0,602	0,481
Holt-Winter <i>Tendência amortecida</i>	0,219	0,938	0,224	0,649	0,481

* As constantes de nível, tendência e sazonalidade são calculadas de forma a minimizar a soma do erro quadrático médio (EQM) de acordo com as equações 88 e 89. O EQM e o BIC são calculados pelas equações 83 e 96, respectivamente.

O modelo de Winter apresenta menores coeficientes BIC em relação aos demais e mesmo EQM que os modelos de Holt – Winter e Holt – Winter com tendência amortecida.

A FIGURA 39 apresenta uma comparação dos ajustes dos modelos de Winter e Holt – Winter com tendência amortecida em relação aos valores reais e previsão para o ano de 2005.

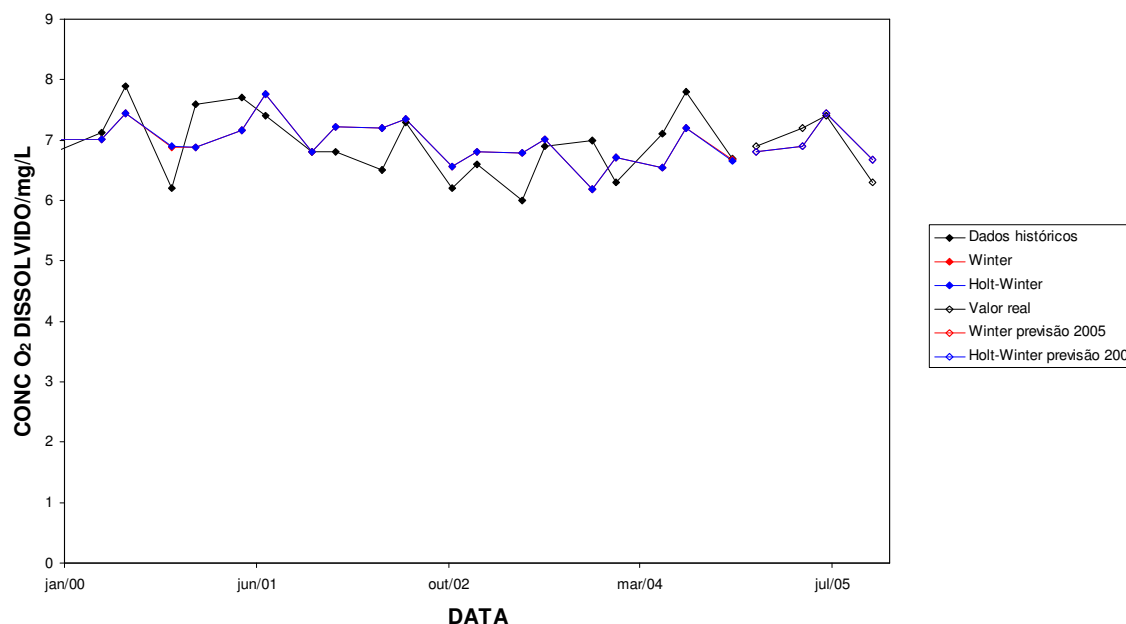


FIGURA 39 - Comparação dos dados históricos da concentração de oxigênio dissolvido no rio Cipó com os ajustes exponenciais dos modelos de Winter e Holt – Winter período de janeiro de 2000 a outubro de 2004 e as previsões em 2005

A concentração de oxigênio dissolvido esteve acima do limite de classe 1, que é de 6,0 mg/L, durante todo o período analisado, mesmo nos meses mais críticos como novembro de 2000 e 2002. As maiores concentrações foram observadas nos meses de julho de 2000 e julho de 2004. Não é observada nenhuma tendência nos dados.

Os modelos de Winter e Holt – Winter apresentam comportamento coincidente em relação aos ajustes e previsões. A dificuldade de ajuste para essa série é maior do que nas anteriores devido ao reduzido número de dados.

O modelo de Winter foi escolhido para previsões nos anos de 2005 a 2007 por não haver tendência nos dados e ainda por ter apresentado menor BIC e mesmo EQM que os demais.

A TABELA 39 apresenta os valores previstos e limites mínimos e máximos de concentração de oxigênio dissolvido pelo modelo de Winter para estação BV140.

TABELA 39 - Valores reais e previstos para concentração de oxigênio dissolvido em mg/L na estação BV162 de 2005 a 2007 pelo modelo de Winter

DATA	PREVISÃO			VALOR REAL EM 2005
	Mínimo	Previsão	Máximo	
jan/05	5,72	6,80	7,88	6,90
abr/05	5,78	6,90	8,02	7,20
jul/05	6,22	7,45	8,68	7,40
nov/05	5,54	6,68	7,82	6,30
jan/06	5,62	6,80	7,99	
abr/06	5,68	6,90	8,12	
jul/06	6,13	7,45	8,77	
nov/06	5,44	6,68	7,92	
jan/07	5,52	6,80	8,08	
abr/07	5,58	6,90	8,21	
jul/07	6,04	7,45	8,86	
nov/07	5,34	6,68	8,02	

Nas campanhas de 2005, também é observada maior concentração de oxigênio dissolvido no mês de julho e menor no mês de novembro. Apesar dessas variações a concentração de oxigênio dissolvido está acima do limite de classe 1.

Os valores médios previstos pelo modelo de Winter para o ano de 2005 são bem próximos aos reais. A concentração média prevista para os anos de 2006 a 2007 tem o mesmo comportamento de 2005, com valores acima do limite de classe 1, indicando assim, uma preservação da boa qualidade das águas do Rio Cipó.

8.7 Conclusões do capítulo

A análise temporal dos parâmetros fosfatos totais, nitrogênio amoniacal e oxigênio dissolvido nos ribeirões do Onça e Arrudas e em estações no rio das Velhas a jusante de cada ribeirão permitiu avaliar o comprometimento da qualidade das águas do rio das Velhas pelos efluentes dos ribeirões. Embora menores, as concentrações de fosfatos totais e nitrogênio nas estações a jusante dos ribeirões estão ainda acima dos limites de classe 3 e para oxigênio dissolvido as concentrações mínimas medidas no ribeirão do Onça (0,700 mg/L) e no rio das Velhas a jusante do Onça (0,800 mg/L) são bem próximas.

Esse comportamento é preocupante uma vez que ações de controle vêm sendo realizadas pela COPASA como as estações de tratamento de água, ETE, e a operações de identificação de esgotos. O aumento das concentrações de fosfatos e nitrogênio amoniacal nos ribeirões do Onça e Arrudas no ano de 2005 sugere lançamentos ainda não identificados de esgotos domésticos.

O observado aumento da concentração de manganês também deve resultar no aumento de outros metais mais tóxicos como níquel, cobre e chumbo geralmente correlacionados ao manganês. Na estação rio das Velhas a jusante do ribeirão Arrudas as concentrações mais altas nos meses mais chuvosos como janeiro e menores em meses de estiagem como julho e novembro indica o carreamento desse metal do solo para o rio em períodos chuvosos. No rio Itabirito também é observado o mesmo comportamento sazonal com concentrações máximas em período de chuva até 20 vezes superior ao limite de classe 3. As previsões para o ano de 2005 estão em boa concordância com as concentrações medidas, com exceção do mês de janeiro; o mesmo comportamento é esperado em relação às previsões para os anos de 2006 e 2007.

As águas do rio Itabirito e do rio das Velhas a jusante do ribeirão Arrudas, região de mineração, apresentaram violação dos limites de classe na concentração de manganês apenas no mês de janeiro, período de chuvas. Para uma melhor avaliação dessa estação a variação na concentração de manganês deve ser observada por um período maior.

A boa qualidade das águas do ribeirão Jequitibá e do rio Cipó foi mantida nas campanhas de 2005 e as previsões feitas para a concentração de oxigênio dissolvido estimam o mesmo comportamento em 2006 e 2007. Esses afluentes ajudam a aumentar a concentração de oxigênio dissolvido águas do rio das Velhas.

As análises de série temporal permitiram avaliar o comportamento dos parâmetros analisados mais importantes e para as estações como o rio das Velhas a jusante dos ribeirões do Onça e Arrudas, rio Itabirito, rio Jequitibá e rio Cipó os valores previstos para 2005 são bem próximos dos reais.

O pequeno número de medidas nas séries históricas analisadas não permitiu bons ajustes e previsões pelos modelos para estações do ribeirão do Onça e Arrudas. O inevitável aumento do número de medidas dos dados históricos, garantido pela manutenção do programa de monitoramento, produzirá melhores ajustes dos modelos e previsões no futuro.

Finalmente é importante ressaltar que para o correto ajuste dos modelos de séries temporais é essencial que as amostragens e as análises sejam realizadas nas mesmas épocas do ano, evitando que variações sazonais não previstas influenciem a caracterização da série histórica.

Capítulo IX – Conclusões finais

O objetivo desse trabalho foi a proposição de uma metodologia para tratamento estatístico da base de dados mantida pelo IGAM relativa à sub-bacia do rio das Velhas buscando o estabelecimento de correlações importantes entre a presença de poluentes, as regiões do rio e as fontes de contaminação foi plenamente realizado. Além disso, esse trabalho permitiu avaliar o desempenho dos métodos estatísticos no tratamento de dados ambientais, compará-los e desenvolver uma estratégia metodológica que poderá ser aplicada a outros estudos de dados ambientais.

A análise fatorial associada à análise de variância permitiu a caracterização de regiões com maiores impactos de contaminações associadas a efluentes domésticos, industriais e de mineração. Essas análises permitiram a elaboração de modelos preliminares de regressão não lineares entre os parâmetros mais correlacionados. A dificuldade na obtenção de um modelo prévio entre os parâmetros está associada à alta variabilidade de dados ambientais, sujeitos a variações sazonais, influências antrópicas, fontes distintas de contaminação como urbana e industrial e fontes difusas de poluição.

As análises de componentes principais e de agrupamento permitiram a melhor caracterização das estações de monitoramento pela identificação dos principais poluentes e ainda a seleção dos poluentes mais importantes na descrição da variabilidade dos dados. A análise de componentes principais permitiu também a proposição de uma redução da base de parâmetros avaliados nos programas de monitoramento, mantendo a mesma caracterização das estações da macro-rede de monitoramento e também da rede dirigida às estações do alto curso do rio das Velhas, com a identificação dos parâmetros mais importantes entre aqueles responsáveis pelos agrupamentos das estações de acordo com suas similaridades.

De acordo com a classificação das estações, foi possível identificar quatro grupos. No grupo de estações com águas de pior qualidade estão os ribeirões do Onça e Arrudas, receptores dos efluentes domésticos e industriais da região metropolitana de Belo Horizonte; com teores elevados de fosfatos totais, nitrogênio amoniacal, DBO, DQO, coliformes fecais e totais, turbidez, manganês, chumbo e níquel; chegando a exceder os limites de classe 3 em até 1000 vezes para coliformes fecais e 10 vezes para DBO e chumbo.

O segundo grupo formado pelas estações no rio das Velhas a jusante dos ribeirões Arrudas, do Onça, da Mata e na ponte Raul Soares em Lagoa Santa também apresenta poluentes característicos de contaminação por lançamento de esgotos, como coliformes totais e fecais e também metais como manganês associado a características do solo e efluentes de mineração.

O terceiro grupo destaca estações pertencentes ao Quadrilátero Ferrífero de Minas Gerais onde as atividades minerárias são importantes. Os altos teores de ferro nas águas caracterizaram a maior parte das estações desse grupo. Também nesse grupo, estão incluídas estações com águas de melhor qualidade como o rio Cipó, destacado por apresentar os maiores teores de oxigênio dissolvido na água.

O quarto grupo associa estações do médio e baixo curso do rio das Velhas e é caracterizado pelos teores de nitratos nas águas, pelas temperaturas da água e do ar e pela basicidade das águas. A presença de nitratos na água é um indicador da capacidade de autodepuração da carga orgânica presente. A boa qualidade das águas em estações do

médio curso do rio das Velhas, como a do ribeirão Jequitibá, destaca a capacidade de autodepuração da carga orgânica lançada nas cidades.

O tratamento da base dados do alto curso do rio das Velhas permitiu identificar como críticas as estações do ribeirão Água Suja próximo de sua foz no rio das Velhas e Córrego da Mina a montante do rio das Velhas, localizadas no município de Nova Lima, pertencentes à região metropolitana de Belo Horizonte. Essas estações são caracterizadas por poluentes presentes em rejeitos de mineração e esgoto. Entre esses poluentes destaca-se a presença de arsênio, elemento tóxico, cujo principal fonte de contaminação é o beneficiamento do ouro é geralmente presente em formas inorgânicas de alta toxicidade e orgânicas. A presença de arsênio no ribeirão Água Suja, que atingiu no passado teores alarmantes, chegando a alcançar o valor correspondente a 58 vezes o limite de classe 3 em julho de 1995, vem sendo reduzida a partir de 2002 com fiscalizações efetivas da FEAM. Esta redução teve grande contribuição das obras de reabilitação ambiental do antigo depósito do Morro do Galo, localizado as margens do ribeirão Água Suja, e que constituía uma fonte potencial de arsênio para o rio das Velhas. Não obstante, os teores de arsênio em alguns pontos no Rio das Velhas ainda são elevados e nas últimas campanhas em 2006 a concentração maior de arsênio excedeu em 2 vezes o seu limite de classe 3.

A presença de mercúrio, utilizado na extração do ouro, embora crítica nas campanhas realizadas entre julho de 1993 a julho de 1997, não ocorreu nas campanhas realizadas até novembro de 2006, apresentando concentrações abaixo de 0,0002 mg/L, correspondente ao limite de classe 3, em todas estações da macro-rede de monitoramento.

Na região do alto curso do rio das Velhas está também localizada a rede de estações dirigida ao abastecimento doméstico. Entre essas, as estações localizadas no rio das Velhas em Bela Fama na alça direita e na alça esquerda foram distinguidas pela presença de sólidos em suspensão e turbidez nas águas, característicos de atividades minerárias. Essas estações são importantes no abastecimento da cidade de Belo Horizonte e sua contaminação aumenta o custo no tratamento da água.

Foi possível estabelecer um novo mapa de monitoramento das estações do alto curso do rio das Velhas, reduzindo o número de parâmetros avaliados e sua frequência de monitoramento. Essa redução foi determinada através da análise de componentes principais, que atribui maiores cargas aos parâmetros mais relevantes nas principais componentes, e não apresenta variação na caracterização da base de dados com a eliminação de parâmetros pouco relevantes para descrição da variabilidade dos dados. Além disso, a identificação de estações correlacionadas foi também empregada na proposta uma nova rede de monitoramento a ser empregada nas campanhas em 2007. Esse resultado faz parte do “Relatório de Monitoramento do Alto Curso da Sub-bacia do Rio das Velhas” publicado pelo IGAM em 2005⁹⁹.

Não foi possível a proposição de uma redução no número de estações de amostragem na macro-rede de monitoramento do rio das Velhas devido ao ainda baixo número de estações com a razão de 0,98 estações/1000 km² nessa rede; próxima, mas inferior a meta estabelecida de 1 estação/ 1000 km², indicada na TABELA 1, do capítulo 1.

As análises de séries temporais dos parâmetros mais críticos permitiram a avaliação de tendências, sazonalidade e a proposição de modelos de previsão. Os parâmetros mais críticos destacados nas análises de componentes principais e na análise fatorial também apresentaram os maiores índices de violação dos limites de classe de enquadramento durante o período de estudo; dentre eles destacaram-se: fosfatos, manganês, nitrogênio amoniacal e oxigênio dissolvido.

Os resultados quantitativos, representados pelas equações de regressão e pelos modelos de previsão, requerem ainda um incremento no número de observações para definição de melhores determinações que não sejam influenciadas por variações temporárias que não façam parte da sazonalidade e tendência naturais das estações.

Dessa forma, é importante que esse programa de monitoramento seja mantido para que se tenha um banco de dados atualizado de modo a permitir a construção de uma base de dados sólida e consistente e, posteriormente a agregação de informações de uso e ocupação urbana e rural para implementação de modelos que avaliem tanto a carga de poluição pontual como a difusa.

Capítulo X – Recomendações

Com base no estudo feito e atendendo as sugestões apresentadas pela banca examinadora desta tese, apresentamos aqui algumas recomendações para a melhoria da qualidade das águas da Bacia do Rio das Velhas:

- As atividades minerárias ainda são fonte de contaminação principalmente no alto curso do rio das Velhas, identificando como críticas as estações do ribeirão Água Suja próximo de sua foz no rio das Velhas e do Córrego da Mina a montante do rio das Velhas, localizadas no município de Nova Lima, pertencentes à região metropolitana de Belo Horizonte. Nessas estações os teores de arsênio observados nas últimas campanhas de 2006 ainda são elevados, apesar da importante redução observada nos anos anteriores. Novas ações são necessárias para reduzir esses teores até valores abaixo do seu limite de classe 3.
- O monitoramento das estações de abastecimento doméstico identificou contaminações importantes que merecem atenção. Entre essas, as estações localizadas no rio das Velhas em Bela Fama na alça direita e na alça esquerda

apresentam elevada contaminação de sólidos em suspensão e turbidez nas águas, característicos de atividades minerárias. Essas estações são importantes no abastecimento da cidade de Belo Horizonte e o controle das atividades minerárias na região reduzirá o custo no tratamento da água.

- Na análise temporal das concentrações de fosfatos nos ribeirões do Onça e Arrudas, as previsões para os anos de 2005 a 2007 indicam o mesmo comportamento do período compreendido entre 1994 a 2004, com teores mais altos de fosfatos no mês de julho e mais baixos em janeiro. No entanto, observou-se nos ribeirões do Onça e Arrudas um aumento atípico da concentração de fosfatos nas campanhas nos meses de abril e novembro em 2005. O ribeirão Arrudas apresentou concentrações máximas de fosfatos no mês de julho superiores às do ribeirão do Onça. Essa inesperada discrepância deve ser conferida nos resultados dos monitoramentos de 2006 e 2007, buscando confirmar se é indicativa de uma flutuação ou de uma nova tendência de crescimento.
- Os teores de fosfatos no rio das Velhas a jusante dos ribeirões do Onça e Arrudas embora apresentem uma tendência de decréscimo ainda são elevados. As previsões para o ano de 2005 apresentam boa concordância com os valores reais medidos nesse ano para os dois ribeirões, não apresentando as discrepâncias observadas nas estações nos ribeirões. É previsto um mesmo comportamento, com tendência de decréscimo nas concentrações de fosfatos nas campanhas de 2006 e 2007. Essa tendência deve ser conferida nos resultados do monitoramento.
- Embora haja uma tendência de decréscimo nas concentrações nitrogênio amoniacal no rio das Velhas a jusante do ribeirão Arrudas, são observadas tendências de crescimento nas concentrações de nitrogênio amoniacal no ribeirão Arrudas, no ribeirão do Onça e no rio das Velhas a jusante do ribeirão do Onça. Para o ribeirão do Onça as previsões estimam valores altos e ainda críticos. No ribeirão Arrudas, as concentrações altas de nitrogênio amoniacal observadas em 2005 estão acima dos limites máximos estimados e esse comportamento atípico invalida as previsões feitas para os anos de 2006 e 2007.

Os demais valores previstos para as campanhas de 2005 aproximam dos valores medidos nesse ano e para as campanhas dos anos de 2006 e 2007 os valores previstos também estimam concentrações inferiores aos limites de classe 3 como em 2005. Essas previsões nos teores de nitrogênio amoniacal devem ser conferidas.

- Como evidenciado pelas análises multivariada e temporal as principais fontes de contaminação do rio das Velhas vêm dos efluentes dos ribeirões do Onça e Arrudas. Esse fato é preocupante uma vez que ações de controle vêm sendo realizadas pela COPASA como a instalação das ETEs nesses dois ribeirões. Uma avaliação dessas ações é necessária.
- Atualmente, 85% dos esgotos de Belo Horizonte são coletados e desse percentual 40% são tratados pela ETE Arrudas; enquanto na ETE Onça apenas 15 % são tratados dos 29% coletado. Na ETE Arrudas o esgoto recebe tratamento preliminar, primário e secundário e na ETE Onça recebe tratamento preliminar e secundário¹⁰¹. Além do aumento das capacidades de coleta e tratamento dos esgotos, a introdução do tratamento terciário nas ETEs permitiria a remoção de poluentes específicos como compostos de nitrogênio e fósforo, metais pesados, além da completa remoção da matéria orgânica.^{101, 102}. Não há consenso se a remoção de compostos de nitrogênio e fósforo é considerada um tratamento em nível secundário. Quando sua remoção ocorre na etapa biológica do tratamento dos esgotos, usualmente se diz que o tratamento é secundário, como remoção de compostos de nitrogênio e fósforo. Quando há necessidade específica de uma etapa posterior, configura-se mais claramente o nível terciário de tratamento. Os principais mecanismos de remoção de compostos de nitrogênio em lagoas de estabilização são por volatilização da amônia, assimilação por algas e sedimentação do nitrogênio orgânico particulado e nitrificação-desnitrificação. Os fosfatos podem ser removidos por precipitação em condições de elevado pH. A remoção desses nutrientes pode também envolver sistemas de lodos ativados e reatores aeróbicos com biofilmes e ainda por processos físico-químicos envolvendo a adição de coagulantes ou alcalinizantes e filtração ou flotação. Normalmente os processos físico-químicos

repercutem em custos operacionais mais elevados do que os associados à remoção biológica.¹⁰²

- Além de um tratamento de qualidade superior para os esgotos, um maior investimento em programas já existentes na COPASA como o “Programa de Monitoramento de Corpos Receptores”, “Programa Caça-Esgoto”, o “Programa de Recebimento e Controle de Efluentes para Usuários Não Domésticos - PRECEND” e a implantação de novas estações de tratamento de esgoto trariam melhor qualidade às águas da sub-bacia do rio das Velhas e também do rio Paraopeba.

O “Programa de Monitoramento de Corpos Receptores”, visa o monitoramento de estações a montante e jusante das estações de tratamento de esgoto em operação, em implantação, com projeto aprovado e ações planejadas, como forma de aferir a eficiência das mesmas e atendimento às condicionantes do licenciamento ambiental. No PRECEND, os estabelecimentos que optarem pelo lançamento de seus efluentes na rede pública coletora de esgotos, irão repassar a COPASA a responsabilidade pela destinação correta de seus efluentes. Esses efluentes serão tratados conforme as suas características e isso poderá prevenir a introdução de poluentes que passam pelas ETEs e continuam a poluir os cursos d’água. O “Programa Caça-Esgoto”, visa à identificação de ligações de esgoto não cadastradas e também um aumento da capacidade de atendimento do sistema coletor já existente.¹⁰¹

- O aumento da capacidade de tratamento das ETEs já existentes e a construção novas estações de tratamento em Nova Contagem (Contagem), Jardim Canadá (Nova Lima), na comunidade do Galo (Raposos), em Cachoeira e Petrovale (Betim) e em Juatuba, além, de redes coletoras, interceptoras, elevatórias e novas ligações prediais¹⁰¹ serão fundamentais para o sucesso da meta estabelecida pelo Projeto Manuelzão para o rio das Velhas, que tem como objetivo tornar o rio novamente adequado para práticas de natação e pesca em 2010.

Capítulo XI – Referências bibliográficas

1. FUNDAÇÃO ESTADUAL DO MEIO AMBIENTE. **Relatório de qualidade das águas superficiais do Estado de Minas Gerais em 1997**, Belo Horizonte, 1998, 1v.
2. FUNDAÇÃO ESTADUAL DO MEIO AMBIENTE. **Relatório de qualidade das águas superficiais do Estado de Minas Gerais em 1998**, Belo Horizonte, 1999, 1v.
3. FUNDAÇÃO ESTADUAL DO MEIO AMBIENTE. **Relatório de qualidade das águas superficiais do Estado de Minas Gerais em 1999**, Belo Horizonte, 2000, 1v.
4. FUNDAÇÃO ESTADUAL DO MEIO AMBIENTE. **Relatório de qualidade das águas superficiais do Estado de Minas Gerais em 2000**. Belo Horizonte, 2001, 1v. (Projeto: Águas de Minas - Monitoramento da Qualidade das Águas Superficiais do Estado de Minas Gerais)
5. INSTITUTO MINEIRO DE GESTÃO DAS ÁGUAS. **Relatório de monitoramento das águas superficiais na Bacia do Rio São Francisco em 2001 . Sub – Bacia do Rio das Velhas**. Belo Horizonte, 2002, 1v.

6. INSTITUTO MINEIRO DE GESTÃO DAS ÁGUAS. **Relatório de monitoramento das águas superficiais na Bacia do Rio São Francisco em 2002 – Sub – Bacia do Rio das Velhas.** Belo Horizonte, 2003, 1v.
7. INSTITUTO MINEIRO DE GESTÃO DAS ÁGUAS. **Relatório de monitoramento das águas superficiais na Bacia do Rio São Francisco em 2003 – Sub – Bacia do Rio das Velhas.** Belo Horizonte, 2004, 1v.
8. INSTITUTO MINEIRO DE GESTÃO DAS ÁGUAS. **Plano diretor de recursos hídricos da bacia hidrográfica do Rio das Velhas 2004 – 2010.** Belo Horizonte: IGAM, 2004, 43p.
9. BRASIL. Lei no 9.433, de 8 de janeiro de 1997. Institui a Política Nacional de Recursos Hídricos, cria o Sistema Nacional de Gerenciamento de Recursos Hídricos, **Lex**, ano 61, 1997, p.18-29
10. GUIMARÃES, A. P. Paisagem física do Rio das Velhas: Belo Horizonte: UFMG/IGC, 1953.123p. Tese de Doutorado. – Instituto de Geociências. Universidade Federal de Minas Gerais, Belo Horizonte, Minas Gerais, 1953.
11. SECRETARIA DO ESTADO DO MEIO AMBIENTE E DESENVOLVIMENTO SUSTENTÁVEL DE MINAS GERAIS. **Aperfeiçoamento do monitoramento da qualidade das águas da Bacia do Alto Curso do Rio das Velhas.** Belo Horizonte: 2003, 66p. (Relatório 1 - Marco Zero).
12. SECRETARIA DO ESTADO DO MEIO AMBIENTE E DESENVOLVIMENTO SUSTENTÁVEL DE MINAS GERAIS. **Aperfeiçoamento do monitoramento da qualidade das Águas da Bacia do Alto Curso do Rio das Velhas.** Belo Horizonte: 2004, 147p. (Relatório 2 – Avaliação dos Resultados do ano de 2003).
13. CONSELHO NACIONAL DO MEIO AMBIENTE. Resolução. **CONAMA 357** . Disponível em: <<http://www.mma.gov.br>.> Acesso em: 10 ago. 2006.
14. SECRETARIA DO ESTADO DO MEIO AMBIENTE E DESENVOLVIMENTO SUSTENTÁVEL DE MINAS GERAIS. **Sistema de cálculo da qualidade da água (SCQA): estabelecimento das equações do índice de qualidade das águas (IQA).** (Relatório 1-Divulgação das informações alcançadas sobre as condições de qualidade e quantidade das águas na bacia do Alto Rio das Velhas).2005, 16p.

-
15. SENA, M. M.; POPPI, R. J.; FRIGHETTO, R. T.S.; VALARINI, P. J.; Avaliação do uso de métodos quimiométricos em análise de solos. **Química Nova**, v.23, n.4, p.547-556, 2000.
 16. EINAX, J. W.; TRUCKENBRODT D., KAMPE K.; River pollution data interpreted by means of chemometrics methods. **Microchemical Journal**, v.58, p.315-324, 1998.
 17. ARUGA, R., NEGRO, G., OSTACOLI, G.; Multivariate data analysis applied to the investigation of river pollution. **Fresenius J. Anal. Chem.**, v.346, p.968-975, 1993
 18. ARUGA, R., GASTALDI D., NEGRO, G., OSTACOLI, G.; Pollution of a river basin and its evolution with time studied by multivariate statistical analysis. **Analytica Chimica Acta**, v.310, p.15-25, 1995.
 19. MARENGO, E.; GENNARO, M.C.; GIACOSA, D. et al. How chemometrics can helpfully assist in evaluating environmental data. Lagoon water. **Analytica Chimica Acta**. v.317, p.53-63, 1995.
 20. SMOLINSKI, A; WACZAK, B.; EINAX, J. W; Hierarchical clustering extend with visual complements of environmental data set. **Chemometrics and Intelligent Laboratory Systems**64, p.45-54, 2002.
 21. VONCINA, D. B.; DOBCNIK, D; NOVIC, M.; ZUPAN, J.; Chemometrics characterization of quality of river water. **Analytica Chimica Acta**. v. 462, p.87-100, 2002
 22. PANTSAR –KALLIO, M.; MUJUNEN, S., HATZLIMIHALIS, G.; et al; Multivariate data analysis of key pollutants in sewage samples: a case study. **Analytica Chimica Acta** .v. 393, p.181-191, 1999.
 23. VIGHI, M.; ALTENBURGER, R; et al. Water quality objectives for mixtures of toxic chemicals: problems and perspectives. **Ecotoxicology and Environmental Safety**. v.54, p.139-150, 2003.
 24. VÁZQUEZ, A.; COSTOYA, M.; PEÑA, R; et al. A rainwater quality monitoring network: a preliminary study of the composition of rainwater in Galicia (NW Spain). **Chemosphere**. v.51, p.375-386, 2003.
 25. REGHUNATH, R.; MURTHY, T.; RAGHAVAN, B. The utility of multivariate statistical techniques in hydro geochemical studies: an example from Karnataka, India. **Water Research**. v.36, p.2437 – 2442, 2002.

-
26. SIMEONOV, V.; STRATIS, J.; SAMARA, C.; et al. Assessment of the surface water quality in Northern Greece. **Water Research**. 37, p.4119 – 4124, 2003.
 27. SINGH, K.; MALIK, A; MOHAN, D.; SINHA, S. Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India) – a case study. **Water Research**. v.38, p.3980-3992, 2004.
 28. QIN, J.; HUH, Y.; EDMOND, J. ; DU, G.; RAN, J. Chemical and physical weathering in the Min Jiang, a headwater tributary of Yangtze River. **Chemical Geology**. v.227,p. 53-69, 2006.
 29. MINGOTI, S. A., **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: UFMG, 2004. 295p.
 30. ANDERSON, C.M.; BRP, R. Review: Practical aspects of PARAFAC modeling of fluorescence excitation-emission data. **Journal of Chemometrics**. v.17, p.200-215, 2003.
 31. PARADKAR, M.M.; SIVAKESAVA, S. IRUDAYARAJ, J. Discrimination and classification of adulterants in maple syrup with the use of infrared spectroscopic techniques. **Journal of Science of Food and Agriculture**.v 83, p. 714-721, 2003.
 32. DOBLE, P.; SANDERCOCK, M. PASQUIER, E.; et al. Classification of premium and regular gasoline by gas chromatography/mass spectrometry, principal component analysis and artificial neural networks. **Forensic Science International**. v.132, p. 26-39, 2003
 33. FERNANDES, A; SANTOS, M.; LEMOS, S.; FERREIRA, M.C; et al. Pattern recognition applied to mineral characterization of Brazilian coffees and sugarcane spirits. **Spectrochimica Acta Part B** , v.60, p.717-724, 2005
 34. OTTO MATTHIAS; **Chemometrics: statistics and computer application in analytical chemistry**. New York: Wiley , 1999.314p.
 35. Gabriel, K.R. *Biometrika*, 58, 3, 453-467, 1971
 36. GILBERT, R.O., **Statistical methods for environmental pollution monitoring**. New York: Van Nostrand Reinhold, 1987.
 37. CARVALHO, M.A, **Métodos estatísticos para análise de dados de monitoração ambiental**.2003. 134p. Tese doutorado, IPEN, São Paulo, São Paulo, 2003.

-
38. FORINA, M.; ARMANINO, C.; RAGGIO, V.; Clustering with dendrograms on interpretation variables. **Analytica Chimica Acta**, v. 454, p.13-19, 2002
 39. MARTENS, H.; NAES T.; **Multivariate Calibration**, John Wiley and Sons, Chichester, UK, 1989. 419 p.
 40. HUANG, W.,CAMPREDON; R., et al; Variation of heavy metals in recent sediments from Piratininga Lagoon (Brazil): interpretation of geochemical data with the aid of multivariate analysis. **Environmental Geology**, v.23, p. 241-247, 1994.
 41. MINGOTI, S. A.; AQUINO, V.; BARBOSA, F.A .R.; et al; Índices físico – químico e biológico de qualidade da água do rio Piracicaba e parte do médio rio Doce: uma primeira aproximação. **Revista de Matemática e Estatística da UNESP**, v. 16, n. 2, p.05-16, 1998.
 42. MOITA NETO, J. M.; MOITA, G. C.; Um introdução à análise exploratória de dados multivariados, **Química Nova**,v. 21 n.4, p.467-469, 1998
 43. ZHANG,C.; SELINUS, O.; SCHEDIN, J.; Statistical analysis for heavy metal contents in till and root samples in an area of southeastern Sweden, **The science of the total environment**, v. 212, p.217-232, 1998.
 44. EGREJA FILHO, F.B.; REIS, E. L.; JORDÃO, C. P.; Avaliação quimiométrica da distribuição de metais pesados em composto de lixo urbano domiciliar, **Química Nova**, v. 22, n.3, p. 324-328, 1998
 45. FERREIRA, M.M.C; FARIA, C. G.; PAES, E.T.; Oceanographic characterization of northern São Paulo Coast: a chemometric study. **Chemometrics and intelligent laboratory systems**, v. 47, p.289-297, 1999.
 46. QU, W.; KELDERMAN, P.; Heavy metal contents in the Delft canal sediments and suspend solids of River Rhine: multivariate analysis for source tracing, **Chemosphere**, v. 45, p.919-925, 2001.
 47. 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.
 48. GALLEGO, J.L.R.; ORDÓÑEZ, A.; LOREDO, J.; Investigation of trace element sources from an industrialized area (Avilés, northern Spain) using multivariate statistical methods, **Environmental International**, v.27, p.589-596, 2002.

-
49. SOLIDORO, C.; PASTRES, R., COSSARINI, G.; et al, Seasonal and spatial variability of water quality parameters in lagoon of Venice. **Journal of Marine Systems**, v. 51, p.7-18, 2004.
50. MARVIN, C.H.; PAINTER, S.; CHARLTON, M.N.; et al, Trends in spatial and temporal levels of persistent organic pollutants in Lake Erie sediments, **Chemosphere**, v.54, p.33-40, 2004.
51. BARTOLOMEO, A .; POLETTI, L.; SANCHINI, G.; et al, Relationship among parameters of lake polluted sediments evaluated by multivariate statistical analysis, **Chemosphere**, v. 55, p.1323-1329, 2004
52. CHEN, T.; ZHENG, Y.; Lei, M.; Assessment of heavy metal pollution in surface soils of urban parks in Beijing, China, **Chemosphere**, v.60, p.542-551, 2005.
53. PIZZEGHELLO, D.; ZANELLA, A .; CARLETTI, P. et al, Chemical and biological characterization of dissolved organic matter from silver fir and beech forest soils, **Chemosphere**, v.65, p.190-200, 2006.
54. . MathSoft Inc . **User's Guide - S-PLUS**. Seattle WA, 1995. 620 p.
55. MathSoft Inc. **S-PLUS 4 Guide to Statistics, Data Analysis**. Seattle, WA, 1997. 877p
56. JOHNSON, R.,A., **Applied Multivariate Statistical Analysis**.4. ed. Upper Saddle River: Prentice-Hall, 1998. 816 p.
57. ARUGA, R.; Treatment of responses below the detection limit: some current techniques compared by factor analysis on environmental data, **Analytica Chimica Acta**, v.354, p. 255-262, 1997.
58. LIU, C.; LIN, K.; KUO, Y.; Application of factor analysis in the assessment of groundwater quality in a blackfoot disease area in Taiwan. **The science of total environment**, v.313, p.77-89, 2003.
59. SABINO, C.V.S; NEVES; O . F. N.; FRANCO, M. B.; Estudo arqueométrico de cerâmicas do sítio Guará, Goiás, Brasil, **Química Nova**, v. 25, n.3, p.368-371, 2002.
60. SABINO, C.V.S; NEVES; KASTNER, G. F.; AMARAL, A .M.; Estudo da biodisponibilidade de metais nos sedimentos da lagoa da Pampulha. **Química Nova**, v.27, n.2, p.231-235, 2005.
61. MAZUCHELI, J.; ACHCAR, J.A.; Algumas considerações em regressão não linear. **Acta Scientiarum**, v. 24, n.6 , p.1761-1770, 2002.

-
62. BELSLEY, D. A ., et al; **Regression diagnostic: identifying influential data and sources of variations**. New York: John Wiley , 1980. 310p.
 63. COOK, R. D.; WEISBERG, S.; **Residuals and influence in regression**. London: Chapman and Hall, 1982. 230p.
 64. DRAPER, N. R.; SMITH, H.; **Applied regression analysis**. 2. ed, New York: John Wiley and Sons, 1981. 269p.
 65. SEBER, G. A .F.; **Linear regression analysis**. New York: John Wiley and Sons, 1977. 465p.
 66. SEBER, G. A .F; WILD, C. J. **Nonlinear regression analysis**. New York: John Wiley and Sons, 1989. 792p
 67. BATES, D. M.; WATTS, D. G.; **Nonlinear regression analysis and its applications**. New York: John Wiley and Sons, 1988. 384p.
 68. KHURI, A . I.; CORNELL, J. A . **Response surface: designs and analysis**. New York: Marcel Dekker, 1987. 472p.
 69. RATKOWSKY, D. A .; **Handbook of nonlinear regression models**. New York: Marcel Dekker, 1990.
 70. RALSTON, M.L.; JENNRICH, R.I. DUD, a derivates-free algorithm for nonlinear least squares. **Technometrics**, v. 20, p.7-14, 1978
 71. SEARLE, S. R. **Linear models**. New York: John Wiley , 1971. 190p.
 72. JENNRICH, R.I. Asymptotic properties of nonlinear least squares estimators. **Ann Math. Stat.**, v.40, p.633-643, 1969.
 73. RATKOWSKY, D. A.; **Nonlinear regression models**. New York: Marcel Dekker, Inc, 1983. 250p.
 74. GRAYBILL, F.A.; **Matrices with applications in statistic**, 2. ed. Belmont: Wadsworth Publishing, 1983. 480p
 75. MORETTIN, P. A.A, TOLOI, C. M., **Análise de séries temporais**. São Paulo. Edgard Blucher, 2004. 535p.
 76. TRAUTWEIN, B. e VIEIRA, G. E., Análise de modelos de ajuste exponencial para previsão de consumo de um sistema de distribuição de água. In: CONGRESSO BRASILEIRO DE ENGENHARIA SANITÁRIA E AMBIENTAL, 23, 2004. Campo Grande, 2004.
 77. SCHLINK, U.; HERBARTH, O.; A component time-series model for SO₂ data: forecasting, interpretation and modification. **Atmospheric Environment**, v. 31, n. 9, p.1285-1295, 1997.

-
78. DONGEN, G.V.; GEUENS, L.; Multivariate time series analysis for design and operation of biological wastewater treatment plant, **Water Research.**, v.32, n.3, p.691-700, 1998.
 79. GEIB, S.; EINAX, J.; Multivariate correlation analysis – a method for the analysis of multidimensional time series in environmental studies, **Chemometrics and intelligent laboratory systems**, v.32, p. 57-65, 1996.
 80. SALCEDO, R.L.R.; ALVIM FERRAZ, M.C.M.; ALVES; C. A .; MARTINS, F.G.; Time series analysis of air pollution data. **Atmospheric Environment**, v.33, p. 2361-2372, 1999.
 81. AHN, H.; Statistical modeling of total phosphorus concentrations measured in south Florida rainfall. **Ecological modeling**, v.116, p.33-44, 1999
 82. CIMINO, G.; DEL DUCE, G.; KADONAGA, L.K.; et al; Time series analysis of geological data. **Chemical geology**, v.161, p. 253-270, 1999.
 83. REICK, C.H.; PAGE, B.; Time series prediction by multivariate next neighbor methods with application to zooplankton forecasts. **Mathematics and Computers in Simulation**, v.52, p.289-310, 2000.
 84. TEMME, C.; EBINGHAUS, R.; EINAX, J.; Time series analysis of long-term data sets of atmospheric mercury concentrations. **Anal. Bioanal. Chem.** v.380, p.493-501, 2004.
 85. DUENAS, C.; FERNANDEZ, M.C.; CANETE, S; et al; Stochastic model to forecast ground-level ozone concentration at urban and rural areas, **Chemosphere**, 61, 1379-1389, 2005.
 86. WANG, Xie – Kang and LU, Wei-Zhen. Seasonal variation of air pollution index: Hong Kong case study. **Chemosphere**, v. 63, n. 8, p.1261-1272, 2006.
 87. BOX, G.E.P. and JENKINS, G.M. Time series analysis: Forecasting and control. San Francisco: Holden- Day, 1976
 88. SYSTAT Software SPSS Inc .SYSTAT 7.0 for Windows, 1997.
 89. FOREST PRO for Windows 3.5, extended edition. Business Forecast Systems, Inc. Belmont, Massachusetts, USA, 1997.
 90. AMERICAN PUBLIC HEALTH ASSOCIATION. **Standard Methods for the examination of water and wastewater**, 20. ed., Washington: APHA, 1998, 1v.
 91. FUNDAÇÃO ESTADUAL DO MEIO AMBIENTE. **Monitoração da qualidade das águas superficiais da sub-bacia do rio das Velhas**. Relatório técnico do período de julho/93 a maio/94 , Belo Horizonte, 1995,1v.

-
92. FUNDAÇÃO ESTADUAL DO MEIO AMBIENTE. **Monitoração da qualidade das águas superficiais da sub-bacia do rio das Velhas**. Relatório técnico do período de julho/94 a novembro/95 , Belo Horizonte, 1996,1v.
 93. FUNDAÇÃO ESTADUAL DO MEIO AMBIENTE. **Monitoração da qualidade das águas superficiais da sub-bacia do rio das Velhas**. Relatório técnico do período de julho/96 a julho/97 , Belo Horizonte, 2000,1v.
 94. EINAX, J. W.; KAMPE O.; TRUCKENBRODT, D.; Assessing the deposition and remobilization behavior of metals between river water and river **sediment using partial least squares regression**. **Fresenius journal of analytical chemistry**. v.93, n.361. p.149-154, 1998.
 95. VEGA, M.; PARDO, R, BARRADO, E.; et al; Assessment of seasonal and polluting effects on the quality of river water by exploratory data analysis. **Water Research**., v.32, n.2,p.3581-3592, 1998.
 96. ROBINSON, R. B; WOODA M. S; SMOOTA J. L.; MOORE S. E.. Parametric modeling of water quality and sampling strategy in a high-altitude Appalachian stream. **Journal of Hydrology**, v. 287, p.62–73, 2004.
 97. CERÓN, J. C.; JIMÉNEZ-ESPINOSA, R.; PULIDO-BOSCH, A .; Numerical analysis of hydrogeochemical data: a case study (Alto Guadalentín, southeast Spain. **Applied Geochemistry**, v.15, p.1053-1067, 2000.
 98. LOPES, J. F.; DIAS, J.M.; CARDOSO, A .C. ,et al; The water quality of the Rio de Aveiro lagoon, Portugal: From the observations to the implementation of a numerical model. **Marine Environmental Research**,v. 60, p.594-628, 2005.
 99. SECRETARIA DO ESTADO DO MEIO AMBIENTE E DESENVOLVIMENTO SUSTENTÁVEL DE MINAS GERAIS. **Aperfeiçoamento do monitoramento da qualidade das Águas da Bacia do Alto Curso do Rio das Velhas**. Belo Horizonte: 2005, 29p. (Relatório-Tratamento estatístico dos parâmetros da qualidade das águas da sub-bacia do Alto Curso do Rio das Velhas).
 100. INSTITUTO MINEIRO DE GESTÃO DAS ÁGUAS. Projeto de Gerenciamento Integrado das Atividades Desenvolvidas em Terra na Bacia do Rio São Francisco. GEF/PNUMA/OEA/SRH. Sub-projeto 1.2. **Relatório Final: Resultados consolidados e plano de ações ambientais. Volume I. Avaliação das Interferências Ambientais da Mineração sobre Recursos Hídricos na Bacia**

- do Alto Rio das Velhas.** Belo Horizonte: Golder Associates Brasil LTDA, 2001. 129p.
- 101.COMPANHIA DE SANEAMENTO DE MINAS GERAIS. **Esgotamento Sanitário.** Disponível em: <<http://www.copasa.com.br>.> Acesso em: 14 mar. 2007.
- 102.VON SPERLING, M., **Introdução à qualidade das águas e ao tratamento de esgotos.** 3. ed. Belo Horizonte: UFMG, 2005. v. 1.452p.