

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE CIÊNCIAS AGRONÔMICAS
CÂMPUS DE BOTUCATU**

**ASSOCIAÇÃO ENTRE VARIÁVEIS CLIMÁTICAS E QUALIDADE DA
ÁGUA PARA CONSUMO HUMANO POR MEIO DE TÉCNICAS
MULTIVARIADAS**

SERGIO AUGUSTO RODRIGUES

Tese apresentada à Faculdade de Ciências
Agronômicas da UNESP – Campus de
Botucatu, para obtenção do título de Doutor
em Agronomia (Energia na Agricultura).

**BOTUCATU - SP
Maio - 2015**

**UNIVERSIDADE ESTADUAL PAULISTA “JÚLIO DE MESQUITA FILHO”
FACULDADE DE CIÊNCIAS AGRONÔMICAS
CÂMPUS DE BOTUCATU**

**ASSOCIAÇÃO ENTRE VARIÁVEIS CLIMÁTICAS E QUALIDADE DA
ÁGUA PARA CONSUMO HUMANO POR MEIO DE TÉCNICAS
MULTIVARIADAS**

SERGIO AUGUSTO RODRIGUES

Orientador: Prof. Dr. Carlos Roberto Padovani

Tese apresentada à Faculdade de Ciências
Agronômicas da UNESP – Campus de
Botucatu, para obtenção do título de Doutor
em Agronomia (Energia na Agricultura).

**BOTUCATU - SP
Maio - 2015**

FICHA CATALOGRÁFICA ELABORADA PELA SEÇÃO TÉCNICA DE AQUISIÇÃO E TRATAMENTO
DA INFORMAÇÃO - DIRETORIA TÉCNICA DE BIBLIOTECA E DOCUMENTAÇÃO - UNESP - FCA
- LAGEADO - BOTUCATU (SP)

Rodrigues, Sergio Augusto, 1976-
R696a Associação entre variáveis climáticas e qualidade da água
para consumo humano por meio de técnicas multivariadas /
Sergio Augusto Rodrigues. - Botucatu : [s.n.], 2015
xi, 118 f. : fots. color., grafs. color., tabs.

Tese (Doutorado) - Universidade Estadual Paulista, Fa-
culdade de Ciências Agrônômicas, Botucatu, 2015
Orientador: Carlos Roberto Padovani
Inclui bibliografia

1. Água - Qualidade. 2. Água - Análise. 3. Mudanças cli-
máticas. 4. Análise multivariada. I. Padovani, Carlos Rober-
to. II. Universidade Estadual Paulista "Júlio de Mesquita
Filho" (Câmpus de Botucatu). Faculdade de Ciências Agronômi-
cas. III. Título.



UNIVERSIDADE ESTADUAL PAULISTA
CAMPUS DE BOTUCATU
FACULDADE DE CIÊNCIAS AGRONÔMICAS DE BOTUCATU

CERTIFICADO DE APROVAÇÃO

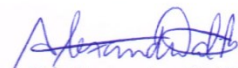
TÍTULO: ASSOCIAÇÃO ENTRE VARIÁVEIS CLIMÁTICAS E QUALIDADE DA ÁGUA PARA CONSUMO HUMANO POR MEIO DE TÉCNICAS MULTIVARIADAS

AUTOR: SÉRGIO AUGUSTO RODRIGUES

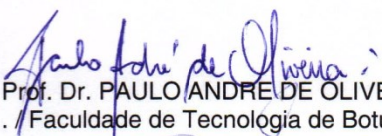
ORIENTADOR: Prof. Dr. CARLOS ROBERTO PADOVANI

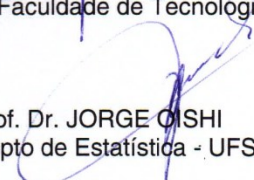
Aprovado como parte das exigências para obtenção do Título de DOUTOR EM AGRONOMIA (ENERGIA NA AGRICULTURA), pela Comissão Examinadora:


Prof. Dr. CARLOS ROBERTO PADOVANI
Departamento de Bioestatística / Instituto de Biociências de Botucatu


Prof. Dr. Alexandre Dal Pai
Depto de Bioprocessos e Biotecnologia - FCA


Profa. Dra. LILIAN CRISTINA TREVIZAN FELIPE
Sabesp/Botucatu


Prof. Dr. PAULO ANDRÉ DE OLIVEIRA
Faculdade de Tecnologia de Botucatu


Prof. Dr. JORGE OISHI
Depto de Estatística - UFSCAR (Aposentado)

Data da realização: 13 de maio de 2015.

DEDICATÓRIA

Aos meus pais, Angelo Rodrigues e Benilda Marchiori Rodrigues, minha família e a todos meus queridos amigos!!

AGRADECIMENTOS

Agradeço a Deus pela oportunidade de viver com saúde e ao lado de pessoas muito especiais.

Ao meu grande e querido orientador, professor Dr. Carlos Roberto Padovani, pela acolhida, dedicação, companheirismo e apoio nesta jornada. Um grande exemplo a ser seguido, tanto profissionalmente quanto pessoalmente.

Aos meus pais, pelo amparo e incentivo em todos os momentos de minha vida.

Aos meus queridos amigos e familiares, que me ajudaram nos momentos difíceis e que contribuíram intensamente para que esse momento fosse possível.

À minha namorada, Paola G. Fernandes, pela sua importância na etapa final deste trabalho.

A todos os meus professores pela dedicação e conhecimentos transmitidos durante minha trajetória na graduação, mestrado e doutorado, em especial ao Prof. Dr. Jorge Oishi, pela forma encantadora de ensinar.

Aos membros da minha banca de defesa: Prof. Dr. Jorge Oishi, Prof. Dr. Paulo André de Olivera, Prof. Dr. Alexandre Dal Pai, Dra. Lilian Cristina Trevisan Felipe e Prof. Dr. Carlos Padovani.

À Faculdade de Tecnologia de Botucatu e ao Programa de Pós-Graduação em Ciências Agrônomicas, Área de Concentração Energia na Agricultura da Universidade Estadual Paulista “Júlio de Mesquita Filho”, pela oportunidade.

Agradeço também à Companhia de Saneamento Básico do Estado de São Paulo e à Faculdade de Ciências Agrônomicas (UNESP), pelo fornecimento dos dados e informações indispensáveis.

Enfim, a todos que contribuíram de alguma forma para a realização deste trabalho.

SUMÁRIO

	Página
LISTA DE TABELAS	VII
LISTA DE FIGURAS	IX
LISTA DE QUADROS	XI
1 RESUMO	1
2 SUMMARY	2
3 INTRODUÇÃO	3
4 REVISÃO DE LITERATURA	6
4.1 Características climáticas e atividades humanas	6
4.2 Recursos hídricos e disponibilidade brasileira	9
4.3 Caracterização da água e seus diferentes usos	11
4.3.1 Características de qualidade da água para o consumo humano	13
4.3.2 Características de qualidade da água na rede de distribuição	15
4.4 Parâmetros de qualidade da água	16
4.4.1 Turbidez	16
4.4.2 Cor Aparente	17
4.4.3 Flúor	17
4.4.4 Cloro Residual Livre	18
4.4.5 pH	18
4.4.6 Temperatura	18
4.4.7 Bactérias Heterotróficas	19
4.4.8 Coliformes Totais	19
4.5 Variáveis climáticas	19
4.5.1 Temperatura Mínima, Máxima e Média	20
4.5.2 Precipitação Pluviométrica e Número de Dias Chuvosos	20
4.5.3 Umidade Relativa do Ar	20
4.5.4 Velocidade dos Ventos	21
4.5.5 Evaporação	21
4.5.6 Insolação	21
4.5.7 Radiação Solar Global	22

4.6 Técnicas de análise estatística na avaliação da qualidade da água	22
4.7 Associação entre as variáveis climáticas e as de qualidade da água	27
5 MATERIAIS E MÉTODOS.....	29
5.1 Materiais	29
5.1.1 Variáveis de qualidade da água.....	32
5.1.2 Variáveis climáticas	33
5.2 Métodos	37
5.2.1 Conjunto de dados.....	38
5.2.2 Teste T^2 de Hotelling para comparação de dois vetores de médias	40
5.2.3 Análise de Correlação Canônica	45
5.2.3.1 Cargas canônicas e variância total explicada	50
5.2.3.2 Inferência sobre os coeficientes de correlação canônica	51
5.2.4 Análise Fatorial	53
6 RESULTADOS E DISCUSSÕES.....	56
6.1 Teste T^2 de Hotteling na comparação da qualidade da água na rede de distribuição e no final de tratamento	62
6.2 Correlação Canônica.....	65
6.2.1 Associação entre o conjunto de variáveis de qualidade da água observado na rede de distribuição e o conjunto de variáveis climáticas.....	65
6.2.2 Associação entre o conjunto de variáveis de qualidade da água observado no final do tratamento e o conjunto de variáveis climáticas	77
6.3 Análise Fatorial.....	83
6.3.1 Análise do conjunto de dados formado por todas as variáveis (qualidade da água na rede e climáticas)	84
6.3.2 Análise do conjunto de dados formado somente pelas variáveis de qualidade da água na rede de distribuição	89
6.3.3 Análise do conjunto de dados formado somente pelas variáveis climáticas.....	92
6.3.4 Correlação linear simples entre os fatores climáticos e os de qualidade da água	96
7 CONCLUSÕES	98
8 REFERÊNCIAS	101
APÊNDICES	109

LISTA DE TABELAS

	Página
Tabela 1. Representação genérica dos dados das características físico-químicas e microbiológicas analisadas na rede de distribuição e final de tratamento.....	41
Tabela 2. Representação genérica dos dados das variáveis de qualidade da água e climáticas, fixando o i – <i>esimo</i> grupo.....	46
Tabela 3. Média e desvio padrão mensal das variáveis de qualidade da água e proporção de amostras com ausência de coliformes totais, considerando os dados da rede de distribuição.	57
Tabela 4. Média e desvio padrão mensal das variáveis de qualidade da água e proporção de amostras com ausência de coliformes totais, considerando o final de tratamento.	58
Tabela 5. Média e desvio padrão mensal das variáveis climáticas.....	59
Tabela 6. Medidas de posição das variáveis de qualidade da água avaliadas na rede de distribuição e no período de Janeiro de 2007 a dezembro 2011.....	61
Tabela 7. Medidas de posição das variáveis de qualidade da água observadas no final do tratamento	61
Tabela 8. Medidas de posição das variáveis climáticas no período de janeiro de 2007 a dezembro de 2011	62
Tabela 9. Vetores de médias das variáveis de qualidade da água.	62
Tabela 10. Limite inferior (LI) e superior (LS) dos intervalos de confiança simultâneos de Hotteling para a diferença entre as médias dos dois grupos.....	64
Tabela 11. Variâncias, covariâncias ($S(1,1)$) e correlações ($R(1,1)$) das variáveis de qualidade da água avaliadas na rede de distribuição*.....	66
Tabela 12. Variâncias, covariâncias ($S(2,2)$) e correlações ($R(2,2)$) das variáveis climáticas *	66
Tabela 13. Matriz de variâncias e covariâncias ($S(1,2)$) entre as variáveis do primeiro e segundo conjunto, considerando a rede de distribuição	67
Tabela 14. Matriz de correlação ($R(1,2)$) entre as variáveis do primeiro e segundo conjunto, considerando a rede de distribuição.....	67
Tabela 15. Autovalor, correlação canônica e resultado do teste estatístico – Rede de distribuição	68
Tabela 16. Variâncias, covariâncias ($S(1,1)$) e correlações ($R(1,1)$) das variáveis de qualidade da água no final de tratamento *	77
Tabela 17. Matriz de covariâncias ($S(1,2)$) entre as variáveis do primeiro e segundo conjunto, considerando o final de tratamento da água	78
Tabela 18. Matriz de correlação ($R(1,2)$) entre as variáveis do primeiro e segundo conjunto, considerando o final de tratamento da água	78

Tabela 19. Autovalor, correlação canônica e resultado do teste estatístico – Final de Tratamento da água.....	79
Tabela 20. Cargas fatoriais, autovalores, percentual acumulado de explicação da variância, comunalidade e unicidade das variáveis de qualidade da água e climáticas (Varimax)	86
Tabela 21. Coeficientes dos fatores – todas as variáveis.....	88
Tabela 22. Cargas fatoriais, autovalores, percentual acumulado de explicação variância, comunalidade e unicidade das variáveis de qualidade da água na distribuição (Varimax).....	90
Tabela 23. Coeficientes dos fatores – variáveis de qualidade da água.....	92
Tabela 24. Cargas fatoriais, autovalores, percentual acumulado de explicação variância, comunalidade e unicidade das variáveis climáticas (Varimax).....	94
Tabela 25. Coeficientes dos fatores – variáveis climáticas	95
Tabela 26. Coeficientes de correlações entre os fatores de qualidade da água e climáticos	96

LISTA DE FIGURAS

	Página
Figura 1. Localização de captação e estação de tratamento	30
Figura 2. Localização dos bairros com pontos de coleta de amostras de água.....	31
Figura 3. Termômetro de máxima e mínima	34
Figura 4. Veneziana abrigando os equipamentos	34
Figura 5. Pluviômetro <i>Ville de Paris</i>	34
Figura 6. Psicrômetro	35
Figura 7. Anemômetro.....	35
Figura 8. Tanque evaporímetro	36
Figura 9. Heliógrafo	36
Figura 10. Actinógrafo	37
Figura 11. Intervalos de confiança simultâneos de Hotteling para a diferença entre as médias dos dois grupos (* $p < 0,05$).....	64
Figura 12. Gráfico de dispersão do primeiro par de variáveis canônicas padronizadas, considerando a rede de distribuição.....	69
Figura 13. Gráfico das correlações de cada variável original com as variáveis canônicas (u_1 e v_1), considerando a rede de distribuição.....	70
Figura 14. Gráfico de dispersão do segundo par de variáveis canônicas padronizadas, considerando a rede de distribuição.....	73
Figura 15. Gráfico das correlações de cada variável original com as variáveis canônicas (u_2 e v_2), considerando a rede de distribuição.....	74
Figura 16. Gráfico de dispersão do primeiro par de variáveis canônicas padronizadas, considerando o final de tratamento.....	80
Figura 17. Gráfico das correlações de cada variável original com as variáveis canônicas (u_1 e v_1), considerando o final de tratamento.....	81
Figura 18. Gráfico das correlações entre as variáveis climáticas e de qualidade da água.....	84
Figura 19. Distribuição dos fatores para as variáveis avaliadas	87
Figura 20. Gráfico das correlações entre as variáveis de qualidade da água.....	90
Figura 21. Distribuição dos fatores para as variáveis de qualidade da água	91
Figura 22. Gráfico das correlações entre as variáveis climáticas	93
Figura 23. Distribuição dos fatores para as variáveis climáticas.....	94

Figura 24. Evolução mensal dos dados das variáveis de qualidade da água na rede de distribuição entre janeiro de 2007 a dezembro de 2011	110
Figura 25. Evolução mensal dos dados das variáveis de qualidade da água no final de tratamento entre janeiro de 2007 a dezembro de 2011	111
Figura 26. Evolução mensal dos dados das variáveis climáticas entre janeiro de 2007 a dezembro de 2011	112
Figura 27. Médias mensais das variáveis de qualidade da água considerando a rede de distribuição	113
Figura 28. Médias mensais das variáveis de qualidade da água considerando o final de tratamento	114
Figura 29. Médias mensais das variáveis climáticas	115
Figura 30. Gráfico de probabilidade <i>QQ-Plot</i> multivariado para os dados de qualidade da água na rede de distribuição (a) e no final de tratamento (b).....	116
Figura 31. Gráfico de probabilidade <i>QQ-Plot</i> multivariado para os dados de qualidade da água na rede de distribuição e climáticos.....	117
Figura 32. Gráfico de probabilidade <i>QQ-Plot</i> multivariado para os dados de qualidade da água no final de tratamento e climáticos.....	118

LISTA DE QUADROS

	Página
Quadro 1. Padrões para qualidade da água distribuída à população segundo Portaria 518/04.....	15
Quadro 2. Descrição das variáveis físico-químicas e microbiológicas das amostra de água.....	38
Quadro 3. Descrição das variáveis climáticas observadas para o município de Botucatu.....	39
Quadro 4. Resumo da associação entre o componente de qualidade da água na rede de distribuição (u_1) e a caracterização climática (v_1).....	75
Quadro 5. Resumo da associação entre o componente de qualidade da água na rede de distribuição (u_2) e a caracterização climática (v_2).....	76
Quadro 6. Resumo da associação entre o componente de qualidade da água em final de tratamento (u_1) e a caracterização climática (v_1).....	83

1 RESUMO

Vários fatores podem contribuir para alterar os parâmetros de qualidade da água, destacando-se, em especial, os climáticos. Sabendo que a água é um recurso natural essencial para as atividades humanas, estudos relacionados com seu uso racional e sustentável tornam-se fundamentais. Destinada ao consumo humano, a água deve apresentar características físicas, químicas e microbiológicas não prejudiciais à saúde, as quais são delimitadas por órgãos de saúde pública e caracterizadas como padrões de potabilidade. A presente pesquisa envolveu características físico-químicas e microbiológicas consideradas importantes para o monitoramento da qualidade da água, analisadas em amostras coletadas diariamente em pontos sorteados na rede de distribuição e final de tratamento do município de Botucatu-SP, no período de 2007 a 2011. Entre essas variáveis destacam-se o cloro residual, pH, turbidez, cor aparente, flúor e temperatura da água, enquanto que entre as microbiológicas, observou-se a ocorrência de coliformes totais e o número de unidades formadoras de colônias de bactérias heterotróficas. Neste contexto, o estudo atual tem como objetivo avaliar o relacionamento simultâneo das características de qualidade da água tratada e distribuída no município de Botucatu com um conjunto de variáveis climáticas. Para isso, utilizou-se das técnicas de análise multivariada de dados com o intuito de aprofundar e melhorar a qualidade das informações e, também, oferecer um procedimento metodológico não usual nas áreas envolvidas. Os resultados obtidos indicam a existência de associação entre determinadas situações climáticas e algumas características de qualidade da água.

RELATION BETWEEN CLIMATE VARIABLES AND WATER QUALITY FOR HUMAN CONSUMPTION THROUGH MULTIVARIATE TECHNIQUES. Botucatu, 2015. 120p, Tese (Doutorado em Agronomia – Energia na Agricultura) – Faculdade de Ciências Agrônômicas, Universidade Estadual Paulista.

Author: SERGIO AUGUSTO RODRIGUES

Adviser: CARLOS ROBERTO PADOVANI

2 SUMMARY

Several factors may contribute for changing water quality parameters, especially climate variables. Knowing that water is a natural paramount resource for human activities, researches observing its rational and sustainable use are crucial. Intended for human consumption, water must have physical, chemical and microbiological characteristics up to harmless levels to health which are determined by public health agencies as well as characterized by its potability standards. This research observed physicochemical and microbiological characteristics which are important for monitoring water quality analysed through daily collected samples at randomized sections within water distribution network and also at the final water treatment sector in Botucatu, SP, Brazil, from 2007 to 2011. Among physical and chemical variables it was observed residual chlorine, pH, turbidity, apparent color, fluorine and water temperature. While on microbiological characteristics it was observed total coliforms occurrence and the number of heterotrophic bacteria colony forming units. This paper aims at evaluating simultaneously treated water characteristics and distribution in Botucatu, SP, Brazil with a set of climate variables. Multivariate data analysis technique was used aiming to develop and improve the quality of information as well as an unusual methodological procedure in the areas involved. Results showed the association between specific climate situations and some water quality characteristics.

Key Words: Multivariate statistics, water quality, climate characteristics.

3 INTRODUÇÃO

O crescimento econômico, o desenvolvimento de novas tecnologias e, conseqüentemente, o aumento populacional são considerados os principais responsáveis pelo aumento na demanda mundial por recursos naturais, em especial a água.

A água é um recurso natural essencial para as atividades humanas. É imperativo seu uso racional e sustentável nas áreas urbanas e rurais, já que seu uso inadequado e a facilidade de apresentar sua qualidade comprometida pela poluição podem promover sua escassez, tanto em quantidade quanto em qualidade.

No setor agrícola, a produtividade assume uma posição de destaque, tanto entre culturas de característica alimentar quanto entre culturas de produção de energias renováveis. Nesse sentido, o perfil climático (temperaturas do ar, radiação solar global, precipitações pluviométricas, umidade relativa, velocidade dos ventos e evaporação) e a disponibilidade de água são determinantes para a produtividade agrícola.

Órgãos oficiais do estado de São Paulo (Instituto Nacional de Meteorologia - INMET, IPMet - Centro de Meteorologia de Bauru, Estação meteorológica da Unesp Botucatu, entre outros) monitoram e armazenam diariamente várias características climáticas, estando disponíveis séries históricas para diversas localidades, entre elas o município de Botucatu. Estudos dessas séries históricas podem contribuir para um melhor planejamento agrícola, uma utilização racional dos recursos hídricos na agricultura e um monitoramento da qualidade da água disponibilizada para a população.

A crescente preocupação com o meio ambiente, destacando a má utilização da água disponível para o consumo humano, faz aumentar o interesse por

estudos relacionados às características climáticas e outras relacionadas à degradação ambiental, tais como: desmatamento, ocupação desordenada dos leitos de rios e morros, uso inadequado do solo na agricultura, má utilização e qualidade da água disponível. No entanto, existem na literatura poucos estudos avaliando a associação entre qualidade da água distribuída para a população com as características climáticas.

No contexto de planejamento urbano, sabe-se que a precipitação pluviométrica é uma das características climáticas com maior influência na vida humana, tanto em relação à sua falta como a seu excesso. O conhecimento desse fato contribui com o planejamento de atividades agrícolas, com a prevenção dos processos erosivos e das inundações, com a geração de energia elétrica e com disponibilidade de água potável para o consumo humano, a qual normalmente é captada em rios, lagos ou lagoas (conhecidas como águas superficiais) e extremamente dependente da situação pluviométrica.

Para o abastecimento urbano, sabendo que a água é uma possível fonte de transmissão de doenças, as estações de tratamento de água dos municípios são essenciais para garantir a potabilidade da água e, conseqüentemente, evitar contaminações e disseminação de doenças. Antes de chegar ao consumo humano, a água pode conter impurezas que alteram suas características física, química e microbiológica, passando por processo de tratamento até alcançar níveis não prejudiciais aos humanos, os quais são delimitados por órgãos de saúde pública e caracterizados como padrões de potabilidade.

As análises físico-químicas e microbiológicas da água, considerando a aferição das variáveis residuais de cloro, turbidez, cor, flúor, pH, temperatura da água, coliformes totais e bactérias heterotróficas, são realizadas pela SABESP (Companhia de Saneamento Básico do Estado de São Paulo) com base na Portaria 2914/11, a qual estabelece que a água distribuída deve ser controlada, com uma quantidade mínima de amostras para análises laboratoriais e com parâmetros físico-químicos e microbiológicos em conformidade com os padrões estabelecidos.

A Portaria 2914/11 passou a ser utilizada integralmente pela SABESP de Botucatu somente a partir do ano de 2012, exigindo uma adequação nos procedimentos e equipamentos de aferição das variáveis. Para efeito de padronização dos dados, uma vez que, para a obtenção de um histórico de cinco anos haveria a necessidade de trabalhar com dados obtidos a partir de duas portarias (a atual e a Portaria 518/04), optou-se em utilizar para este estudo o período de janeiro de 2007 a dezembro de 2011, o qual foi regido somente pela Portaria 518/04 (BRASIL, 2004 e BRASIL, 2011).

Ponderando que as características físico-químicas e microbiológicas da água dependem de características climáticas e sabendo da preocupação atual com o aumento da produtividade, um aprofundamento sobre o comportamento de variáveis climáticas associadas às características físico-químicas e microbiológicas da água, faz-se necessário pela sua alta importância no uso racional da água e na utilização sustentável dos recursos hídricos. Estudos desta natureza podem contribuir com a melhoria das técnicas de monitoramento da qualidade da água.

No entanto, estudar as variáveis de forma independente não se configura o melhor procedimento de análise estatística de dados. Nos casos onde há uma estrutura de relacionamento simultâneo entre variáveis, utilizar técnicas multivariadas constitui-se na melhor alternativa para a análise exploratória e procedimento inferencial.

Neste contexto, as técnicas de Estatística Multivariada podem ser utilizadas para indicar a importância de cada característica climática no monitoramento dos parâmetros de qualidade da água nas estações de tratamento do município de Botucatu e na rede de distribuição, fornecendo informações relevantes para ajudar na tomada de decisões e na elaboração de políticas públicas nas diversas atividades econômicas, tais como uso racional de energia e recursos hídricos na agricultura e em áreas urbanas dos Municípios.

Pautado nisto, o intuito deste trabalho é contribuir com estudos que buscam associar a qualidade da água com as características climáticas, apresentando uma ferramenta metodológica estatística não usual na área, mas que pode melhorar em muito a qualidade das informações e, conseqüentemente, propiciar aos gestores das empresas responsáveis pela água potável dos municípios uma melhoria no entendimento das associações entre as variáveis envolvidas no processo de tratamento da água.

Assim sendo, objetiva-se neste estudo:

- 1) Comparar a qualidade da água na rede de distribuição com a saída dos reservatórios (final de tratamento, após a passagem pela ETA) considerando simultaneamente as variáveis físico-químicas e microbiológicas.
- 2) Analisar o relacionamento entre variáveis climáticas e características físico-químicas e microbiológicas da água distribuída no município de Botucatu, identificando as variáveis climáticas de maior impacto na qualidade da água, envolvendo a rede de distribuição e o final de tratamento. Conseqüentemente, destacar as características da água que devem ser monitoradas e/ou adequadas para garantir sua potabilidade em diversas situações climáticas.

4 REVISÃO DE LITERATURA

4.1 Características climáticas e atividades humanas

Tendo em vista o interesse mundial pelo aumento da produtividade agrícola e consequentemente, pela preocupação em minimizar os impactos ambientais gerados por essas atividades econômicas, observam-se muitas pesquisas que utilizam dados de variáveis climáticas. Para um planejamento agrícola dentro do contexto de produtividade e da preocupação com a preservação ambiental, entender o comportamento das características climáticas de uma região é fundamental.

A simulação de dados climáticos de uma região pode ajudar a minimizar possíveis perdas na produção agrícola por meio de um melhor dimensionamento dos sistemas de irrigação, adubação, entre outros fatores importantes para agricultura. Virgens Filho (2001) desenvolve uma ferramenta de simulação de dados climáticos sobre precipitação pluviométrica, radiação solar e temperatura, baseada na distribuição de probabilidade estimada a partir de séries históricas dessas variáveis. Além disso, estabelece a frequência de dias chuvosos por mês, utilizando a probabilidade dessa ocorrência estimada pelo processo de Markov.

No que diz respeito aos estudos sobre o perfil climático de uma região, Cardim (2001) utiliza-se de técnicas de estatística multivariada para entender o comportamento de variáveis climáticas importantes para o planejamento agrícola, tais como: altitude, precipitação pluviométrica e temperatura média mensal, identificando

regiões homogêneas do estado de São Paulo. Para isso, a técnica de análise de agrupamento é utilizada, identificando regiões do estado com o mesmo perfil climático. Já pelas técnicas de componentes principais e de geoestatística são construídas três classificações climáticas individualizadas, apresentadas por mapas com características térmicas, pluviométricas e pluviotérmicas distintas.

Oliveira (2003) destaca que pesquisas com o objetivo de conhecer melhor as características climáticas por meio de simulações vêm evoluindo bastante nos últimos anos, contribuindo com um melhor planejamento agrícola e com a avaliação dos impactos ambientais. Nesse cenário, o objetivo de seu trabalho é desenvolver um modelo para geração de séries sintéticas de precipitação, fornecendo informações mais detalhadas tais como: precipitação total diária, duração desse evento, além de valores de intensidade da precipitação e do tempo de sua ocorrência, possibilitando a obtenção de um perfil para cada precipitação.

Para Tucci (1997 citado por OLIVEIRA, 2003), a pluviometria ou precipitação auxilia na gestão de atividades como a irrigação, geração de energia e utilização da água.

O rendimento de culturas agrícolas sofre profunda influência das condições climáticas, portanto o conhecimento antecipado dessas condições, bem como o entendimento de sua variação ao longo do tempo e espaço tornam-se extremamente importantes. Procurando entender a variabilidade temporal e espacial das precipitações em Uberaba-MG, Silva, Guimarães e Tavares (2003) apresentam um estudo estatístico exploratório dos dados de precipitações mensais entre os anos de 1914 a 2000, utilizando geoestatística. Verifica-se que os dados de precipitação possuem assimetria acentuada para os meses de maio a setembro e tendência à simetria para os demais meses e também para os valores anuais. Observa-se ainda uma dependência temporal fraca ou não existente das precipitações mensais, indicando que os métodos de estimação, considerando independência entre as amostras, podem ser utilizados para previsão temporal das precipitações.

Para Ayoade (2004), a evaporação das águas do solo, das águas superficiais e a transpiração vegetal são alguns dos fatores que contribuem para o vapor de água presente na atmosfera. É esse vapor a origem de todas as formas de condensação e precipitação, sendo o conhecimento dessa característica ambiental bastante importante para diversas atividades humanas. Como exemplo desta importância, destaca-se o

desenvolvimento de organismos patógenos nas plantações, em seres humanos e em animais, o qual está condicionado ao excesso ou baixa no teor de umidade do ar.

Silva (2006) ressalta que as características ambientais, mais precisamente o comportamento pluviométrico, podem influenciar no planejamento das atividades econômicas de uma região e, portanto, um melhor conhecimento da estrutura das características ambientais colabora para tornar mais eficiente essas atividades. O autor procura, por meio de técnicas de estatística multivariada, mapear áreas homogêneas nas regiões Sul e Sudeste do Brasil em relação à precipitação pluviométrica. Utilizando-se de técnicas de análise de agrupamento, de componentes principais e de geoestatística, 13 áreas homogêneas em relação à precipitação são identificadas e referenciadas por meio de mapas. A partir desse mapeamento, um estudo comparativo das 13 áreas em relação aos resultados de precipitação é realizado, verificando, entre outros fatores, que as 13 áreas são suficientes para entender e representar as variabilidades espacial e sazonal nas regiões Sul e Sudeste.

Gherardi, Dourado Neto e Sentelhas (2008) observam que o entendimento das condições climáticas de uma região é necessário para um planejamento agrícola mais adequado e apresentam um modelo para a caracterização espacial e temporal da temperatura média do ar (diária e mensal) de diferentes localidades do estado do Mato Grosso. Nesse sentido, para localidades sem informações meteorológicas, é possível estimá-las utilizando a técnica de regressão linear múltipla com dados de temperatura, altitude, latitude e longitude, obtidos junto ao Instituto Nacional de Meteorologia. A modelagem proposta pelos autores apresenta bom desempenho para estimação de dados mensais, mas não é tão eficiente para dados diários.

Furlan (2009) estuda a variabilidade no tempo e no espaço da precipitação no estado de Rondônia com o objetivo de verificar se o desmatamento levou a alguma alteração no perfil hídrico e pluviométrico dessa região. Devido à falta de informações pluviométricas em alguns momentos e a algumas falhas na coleta dos dados, um gerador de dados climáticos é utilizado, para completar a série histórica. Em seguida, procura-se ajustar a distribuição de probabilidade Gama para identificar regiões homogêneas em relação à probabilidade de chuva e complementa a investigação com a análise de agrupamento a fim de identificar localidades e meses com o mesmo perfil de precipitação. Os resultados das análises desses dados mostram que, apesar do aumento do

desmatamento em Rondônia, não houve mudanças significativas no clima da região, mas sugere a utilização de séries mais longas.

Virgens Filho et al. (2009) destacam que as situações extremas de umidade relativa do ar podem interferir na saúde dos seres humanos, no crescimento de plantas, na incidência de pragas e doenças em animais e plantas, entre outras situações relacionadas com as atividades humanas.

Objetivando conhecer os perfis de clima e de produtividade e suas possíveis inter-relações nas microrregiões do estado do Mato Grosso, Oliveira e Padovani (2010) utilizam a técnica multivariada de análise de agrupamento para identificar regiões semelhantes com relação ao perfil climático (temperatura, umidade relativa do ar, evaporação e precipitação) e à produtividade de culturas tais como algodão, arroz, banana, cana de açúcar, feijão, mandioca, milho, soja e sorgo. A técnica utilizada permitiu construir um mapa, destacando as regiões com o mesmo perfil climático e de produtividade das culturas analisadas. Dessa forma, os resultados da análise dos agrupamentos indicou inter-relacionamento entre as regiões com evapotranspiração e temperaturas mais baixas com as de maiores produtividades de milho e sorgo, já nas regiões com temperaturas mais altas, menor produtividade de arroz.

Vários estudos apontam para importância do conhecimento das condições climáticas para um melhor gerenciamento e definição de estratégias adequadas para o manejo agrícola. Além disso, Gadgil e Iyengar (1980 citado por Cardim, 2001) destacam a importância do estudo das características climáticas no contexto da utilização racional dos recursos hídricos na agricultura.

Outros autores apresentam métodos de estimação para características climáticas a partir do histórico de dados climáticos de uma região, entre os quais se destacam os trabalhos apresentados por Baena et al. (2005), Zanetti et al. (2005), Rickli (2006) e Zanetti, Oliveira e Pruski (2006).

4.2 Recursos hídricos e disponibilidade brasileira

Do total de água existente no planeta, 97,5% representa água salgada e apenas 2,5% doce. Da água doce disponível, 68,8% encontra-se em áreas congeladas e em regiões polares, 30% em reservatórios subterrâneos e 0,9% estão

presentes na umidade do solo, na biomassa e no vapor da atmosfera. Apenas 0,3% aproximadamente do volume total de água doce da terra são de águas superficiais, presentes em rios e lagos, mais acessíveis ao uso humano (SETTI et al., 2001; UNITED NATIONS, 2014).

O Brasil possui um grande potencial de recursos hídricos a ser explorado. De acordo com Plano Nacional de Recursos Hídricos (BRASIL, 2006a), aproximadamente 13% da disponibilidade mundial de água doce encontra-se em território brasileiro.

No entanto, essa ideia de abundância gerou nos brasileiros uma cultura de uso abusivo da água de rios e lagos. Além disso, essa disponibilidade hídrica não é homogênea em seu território, tanto em termos de disponibilidade para o consumo, irrigação e outras atividades, quanto em termos de potencial para geração de energia elétrica. Enquanto na Região Amazônica, habitada por 5% da população brasileira, concentram-se 73% da água doce do país, nas demais regiões observa-se que 27% dos recursos hídricos estão disponíveis com maior facilidade de acesso para 95% da população. Já mais precisamente nas Sul e Sudeste, com aproximadamente 32% da população brasileira, observa-se apenas 6% da água doce superficial disponível (BRASIL, 2006b; BRASIL, 2013).

Além disso, a oferta de água tratada não é uniforme entre as regiões brasileiras, apresentando alguns tipos de contrastes, tais como o percentual de domicílios atendidos por rede de distribuição de água tratada, totalizando 87,5% dos domicílios da região Sudeste e 58,7% da região Nordeste (BRASIL, 2013).

Uma questão interessante é apresentada na Lei de Política Nacional de Recursos Hídricos, Lei n.9433/97 (BRASIL, 1997), descrevendo que “a água é um recurso natural limitado, dotado de valor econômico”, sendo que em situações de escassez, “o uso prioritário dos recursos hídricos é o consumo humano e a dessedentação de animais”. Observa-se ainda, no inciso IV, que na ausência de escassez a gestão dos recursos hídricos, compostos pelas águas superficiais (rios, lagos, lagoas e bacias e mar territorial) e subterrâneas (águas originadas do interior do solo), deve sempre proporcionar o uso múltiplo das águas, garantindo às gerações futuras a disponibilidade de água com padrões aceitáveis de qualidade.

Portanto, o gerenciamento dos recursos hídricos brasileiros deve ser pautado no uso harmonioso, sendo fundamental um equilíbrio entre os setores usuários

da água, evitando conflitos entre a utilização dos recursos hídricos para a agricultura irrigada, geração de energia elétrica, navegação, abastecimento urbano, uso industrial, saneamento básico, pesca e lazer.

4.3 Caracterização da água e seus diferentes usos

A água, essencial para vida humana, é uma fonte natural renovável, porém, finita e aleatória. Ela é renovável por sua capacidade de se recompor rapidamente em quantidade, principalmente pelas chuvas, no entanto é um recurso finito, pois pode ser comprometida facilmente pela poluição, tornando-se inadequada ao uso humano. É de ocorrência aleatória, pois sua distribuição é irregular no tempo e por região, sendo influenciada pelas condições climáticas e meteorológicas (CONEGO, 1993).

Dependendo da quantidade hídrica demandada e das perdas em quantidade e qualidade, os usos da água são classificados como de forma consuntiva e não consuntiva. No uso consuntivo, uma determinada quantidade de água é retirada do manancial e devolvida em uma quantidade menor ou com qualidade inferior após sua utilização. Já no uso não consuntivo, a água é utilizada no próprio manancial ou, quando captada, é devolvida em quantidade e qualidade idêntica (REBOUÇAS, BRAGA JUNIOR e TUNDISI, 2002; ESTEVES, 2011).

De acordo com Rebouças, Braga Junior e Tundisi (2002), é considerada de uso consuntiva a utilização da água para a agricultura, pecuária, irrigação de plantações agrícolas, processamento industrial e para o abastecimento urbano e rural. Já os usos não consuntivos destacam-se o transporte, lazer, piscicultura, geração de energia elétrica e entre outros.

Para garantir o uso múltiplo das águas brasileiras, há necessidade da satisfação simultânea de critérios mínimos de qualidade. Esses parâmetros são estipulados pela resolução CONAMA 357 /2005 (BRASIL, 2005), a qual estabelece os padrões de qualidade necessários para cada finalidade ou utilização das águas superficiais. Já os critérios mínimos para a água distribuída para o consumo humano foram regidos até 2011 pela Portaria 518/04 e atualmente pela Portaria 2914/11.

Segundo Von Sperling (2005), entre os setores usuários da água, o abastecimento urbano e industrial necessita de um tratamento prévio, pois há uma

exigência maior sobre os parâmetros de qualidade, sendo necessário atender a diversos critérios. Entretanto, existem situações em que essa exigência é menor, tal como a irrigação.

Dependendo da localidade e da situação de origem, a água pode apresentar padrões de qualidade variáveis. A disponibilidade de água doce com qualidade é fundamental para a qualidade de vida, desenvolvimento econômico e sustentabilidade do planeta (TUNDISI, 2003).

A possibilidade de escassez de água se dá por meio de seu uso indiscriminado, da poluição e também das características climáticas, tais como a restrição pluviométrica. O artigo 3º, inciso III da Lei da Política Nacional do Meio Ambiente apresenta o conceito de água poluída e, conseqüentemente imprópria para o consumo humano como sendo “qualquer alteração química, física ou biológica que possa importar em prejuízo à saúde, à segurança e ao bem estar das populações, causar dano à flora e fauna, ou comprometer o seu uso para finalidades sociais e econômicas”.

Os fenômenos naturais e as atividades humanas são fatores determinantes para a qualidade da água, ou seja, suas características são afetadas pelo escoamento e infiltrações no solo das águas de chuvas, incorporando diversas impurezas sólidas ou dissolvidas e pela interferência do homem ao gerar dejetos domésticos e industriais, bem como a utilização de defensivos agrícolas (VON SPERLING, 2005, p.15).

Dessa forma, além do ciclo da água na natureza (precipitação, escoamento superficial, infiltração, evaporação e transpiração), podem-se definir alguns caminhos em que as características da água são alteradas: água bruta (águas superficiais ou subterrâneas captadas e possuindo uma determinada qualidade), água tratada (água tratada para o abastecimento urbano e industrial), esgoto bruto (água utilizada sofrendo transformação em sua qualidade, constituindo-se em um despejo líquido), esgoto tratado (tratamento do despejo líquido visando remover os principais poluentes antes do lançamento nos corpos d'água), água pluvial (água que escoou pelo solo e foi captada com novos constituintes), corpo receptor (o esgoto tratado e a água pluvial atingindo os corpos d'água e pelo processo de diluição atingem nova característica de qualidade) e reúso (VON SPERLING, 2005, p.20).

Portanto, enquanto a demanda mundial por água aumenta, verifica-se uma deterioração na qualidade dos recursos hídricos em função de sua utilização pouco

sustentável, uma ocupação humana desordenada e condições ambientais (SILVA et al., 2009, TAVARES et al, 2009).

4.3.1 Características de qualidade da água para o consumo humano

Souza et al (2014) reforçam o conceito de que a qualidade da água utilizada pelo ser humano é uma característica que depende estritamente de seu uso, apresentando necessidades diferentes quando utilizada para o consumo humano, irrigação, transporte, lazer entre outros.

Neste sentido, para que a água disponível a uma população seja adequada para o consumo é necessário apresentar certas características. Alguns minerais devem estar presentes, pois são essenciais para saúde e para deixá-la com odor e gosto agradáveis. Não deve apresentar valores elevados de cor aparente, matéria suspensa, turbidez, agentes patogênicos, dureza, sabor, odor, produtos orgânicos e inorgânicos perigosos à saúde e não deve ser corrosiva para não prejudicar as tubulações (GRAY, 2008 citado por MORENO, 2009).

Entre as principais fontes de contaminação da água, Bettega et al (2006) destacam a falta de tratamento de esgotos lançados em rios e lagos; os aterros sanitários de alguns municípios, muitas vezes estabelecidos de forma irregular, podendo afetar os lençóis freáticos; os defensivos agrícolas escoados com a chuva para os rios e algumas indústrias que utilizam os rios para lançar seus resíduos tóxicos.

De acordo com o Ministério da Saúde, uma água apropriada para o consumo humano deve atender um padrão de potabilidade, considerando algumas características físicas, químicas, microbiológicas e radioativas. No sentido de garantir seu consumo sem riscos à saúde, as estações de tratamento são responsáveis pelos processos de desinfecção antes de sua distribuição para as residências (BRASIL, 2013).

Segundo Moreno (2009), a partir da Portaria 518/08 foi estabelecido que a água adequada para o consumo humano é aquela que, além de respeitar um padrão de potabilidade, deve ser tratada e distribuída sem apresentar riscos à saúde humana.

Em alguns municípios, a população tem acesso à água de bicas ou fontes naturais, as quais normalmente são consumidas sem passarem por um tratamento prévio. Para verificar a potabilidade da água de bicas dos municípios de Santos e São

Vicente (SP), considerando os critérios de qualidade para o consumo humano estabelecidos pela Portaria 518/04, Tavares et al. (2009) avaliam amostras de água de 31 bicas, coletando dados das variáveis: presença de coliformes totais e termotolerantes, turbidez, cor aparente, sólidos totais dissolvidos, sulfato, pH, odor, nitrato, nitrito, dureza, ferro e cloro residual livres (no caso de bicas com tratamento). A partir de um estudo descritivo para cada variável, quantificando as amostras aprovadas e não aprovadas, com base nos padrões estabelecidos pela portaria, os autores verificam que o parâmetro nitrato apresentou um maior número de amostras com resultados insatisfatórios, seguido pelos coliformes e pH. Observam também que apenas 19% das bicas foram aprovadas em todos os parâmetros.

Sabendo que a água pode ser um transmissor de patógenos, podendo contaminar os alimentos no momento de sua higienização, Siqueira et al. (2010) realizam uma análise microbiológica da água de consumo, empregada em unidades de alimentação, avaliando a presença de coliformes totais e termotolerantes em diferentes estabelecimentos da cidade de Recife-PE. Para isso, 40 estabelecimentos abastecidos com a rede pública de distribuição e com armazenamento de água por meio de cisternas e/ou caixas d'água são utilizados para coleta das amostras. Os resultados deste estudo apontam que 62,5% dos estabelecimentos não atendem os padrões de qualidade da legislação, pois se notou, nessas amostras, a presença de coliformes. Considerando que a água chega aos estabelecimentos dentro dos padrões de qualidade, o estudo indica que a contaminação observada está relacionada à conservação inadequada dos reservatórios. Sugere-se, portanto, um trabalho de conscientização a fim de uma melhor higienização desses reservatórios.

Caselani (2014) elabora um diagnóstico da percepção de risco sanitário dos responsáveis por propriedades rurais da bovinocultura de corte no Estado de São Paulo em relação à utilização de insumos alimentares, agrícolas e para a saúde animal. Assim, são avaliados por meio de técnicas da estatística multivariada, mais precisamente a análise de correspondência múltipla, dados relativos à utilização da água, alimentos, produtos agrícolas e veterinários no rebanho dessas propriedades rurais, além de fatores socioeconômicos. Entre os diversos resultados obtidos, verifica-se que os produtores com menos de 100 animais e escolaridade intermediária possuem uma tendência maior a acreditar que a qualidade da água utilizada não interfere nos riscos sanitários e ambientais inerentes ao processo de produção, ou seja, essa água não transmite doenças aos animais.

4.3.2 Características de qualidade da água na rede de distribuição

Entre os principais usos da água, a distribuição para o consumo humano deve merecer uma atenção especial em relação à sua qualidade, sendo necessária ser considerada potável, atendendo critérios definidos pela legislação para não apresentar riscos à saúde da população (BRASIL, 2004).

Por intermédio da água, várias enfermidades podem ser transmitidas, sendo que a ingestão de água que contenha agentes nocivos à saúde é um dos principais mecanismos de transmissão (BRASIL, 2006c).

Pela Portaria 518/04 (BRASIL, 2004), alguns dos padrões mínimo e máximo para qualidade da água distribuída à população são apresentados no Quadro 1 (considerando apenas as variáveis restritas a este estudo). Entre eles, destacam-se os padrões de potabilidade para substâncias químicas que podem proporcionar riscos à saúde humana, mais precisamente o cloro residual livre em altas concentrações, além de padrões físicos de aceitação da água, nos quais estão a cor aparente, turbidez e pH. Além disso, estabelecem-se também as técnicas adotadas para detectar os microrganismos que podem estar presentes na água (BRASIL, 2004).

Quadro 1. Padrões para qualidade da água distribuída à população segundo Portaria 518/04

Características Físico-químicas e microbiológicas	Limite permitido	Unidade
Turbidez	Máximo de 1,0 <i>uT</i> em 95% das amostras, caso seja encontrado algum valor superior a 1,0 <i>uT</i> , não é permitido ultrapassar 5 <i>uT</i> .	Unidade Nefelométrica de Turbidez (<i>uT</i>)
Cor aparente	Máximo 15 <i>UC</i>	Unidade de Cor (<i>UC</i>)
Flúor	Máximo 1,5 <i>mg/L</i>	Miligramas por litro (<i>mg/L</i>)
Cloro residual	Mínimo de 0,5 <i>mg/L</i> e máximo de 2,0 <i>mg/L</i>	Miligramas por litro (<i>mg/L</i>)
pH	De 6,0 a 9,5	-
Bactéria heterotrófica	Máximo de 500 UFC por <i>ml</i>	Unidade formadora de colônia (UFC)
Coliformes totais	Ausência em 95% das amostras analisadas no mês (amostras de 100 <i>ml</i>)	Número mais provável (<i>NMP</i>)

Estes padrões constituem uma referência no território nacional, exceto a quantidade de flúor, o qual pode variar de acordo com o clima de cada região e o consumo médio diário de água por indivíduo (MACEDO, 2004).

Em 2011 o Ministério da Saúde publica a quinta versão da norma brasileira de qualidade da água para consumo, estabelecendo por meio da Portaria 2914/11 em substituição à Portaria 518/04 os procedimentos atualizados de controle e vigilância da qualidade da água para consumo humano (RIBEIRO, 2012).

De acordo com a SABESP (QUALIDADE ...2012), a Portaria 2914/11 determina que a água distribuída para a população deve ser controlada, exigindo uma quantidade mínima de amostras de água para análises laboratoriais, sendo que as variáveis monitoradas em seu controle de qualidade e utilizadas no presente estudo são de natureza físico-químicas e microbiológicas. O monitoramento destas variáveis deve ser realizado desde o momento da captação até o da distribuição para a população. Em Botucatu a SABESP realiza esse monitoramento, tendo como base os padrões e limites estabelecidos por essa legislação para os parâmetros físico-químicos e microbiológicos da água.

4.4 Parâmetros de qualidade da água

Como já descrito anteriormente, a água para o consumo humano deve atender alguns parâmetros de qualidade, os quais dependem de sua utilização. Para a água distribuída nos domicílios, as características de caráter físico, químico e biológico devem ser limitadas até um nível não prejudicial ao consumo humano, sendo que a Portaria 518/04 estabelece esses limites no período de análise desta pesquisa (BRASIL, 2004).

Com base em Brasil (2004), Macêdo (2004) e Eaton et al. (2005), são caracterizados a seguir os parâmetros de qualidade da água, tratados neste estudo por variáveis.

4.4.1 Turbidez

A variável turbidez é uma medida da passagem de luz pela água, ou seja, indica a resistência da água à penetração da luz, ocasionada pela presença de

partículas em suspensão existentes, alterando sua transparência. Esta variável apresenta uma característica estética da água, sendo seu valor máximo permitido pela Portaria 518/04 de 1,0 unidade nefelométrica de turbidez (uT) em 95% das amostras e, caso esse valor seja ultrapassado, não é permitido ser superior 5,0 uT . Na nova portaria (2914/11), o valor máximo de turbidez foi reduzido para 0,5 uT em tratamentos de filtração rápida (mantendo-se em 1,0 uT na filtração lenta), não devendo ser superior a 5,0 uT em 5% das amostras. Quanto maior for o valor desta variável, pior a qualidade da água para o consumo humano, pois os sólidos em suspensão podem, em algumas situações, exercer o papel de meios de fixação para alguns microrganismos nocivos à saúde, impedindo a ação do cloro e deixando-a mais vulnerável.

4.4.2 Cor Aparente

A variável cor aparente, avaliada em unidades de cor (UC), mede o comprimento de ondas, objetivando indicar possíveis presenças de substâncias na água. Esta variável tem efeito estético, sendo que na Portaria 518/04 a cor aparente possui como parâmetro de qualidade o valor máximo de 15 UC . Quanto menor o valor dessa variável, melhor é a qualidade da água nesse aspecto.

4.4.3 Flúor

O flúor é um componente adicionado à água na estação de tratamento e um importante agente na proteção dos dentes humanos contra a cárie. Esta variável, medida em miligrama por litro (mg/L), indica a quantidade de flúor na água do ponto de coleta. A portaria do Ministério da Saúde estabelece, para o estado de São Paulo, um teor entre 0,6 a 0,8 mg/L , no entanto, essa variável é bastante influenciada pelo clima e pela temperatura de cada região. O excesso (acima de 1,5 mg/L) desse componente deteriora a qualidade da água no sentido de ocasionar manchas nos dentes dos consumidores. Já sua falta não afeta a qualidade, no entanto, pode aumentar os índices de cáries, principalmente, em populações mais carentes.

4.4.4 Cloro Residual Livre

A variável cloro residual livre, medida em miligrama por litro (mg/L), indica a quantidade de cloro ainda presente na água ao analisá-la na rede de distribuição e no final de tratamento. O cloro é um agente que contribui para redução de bactérias que podem estar presentes na água e é adicionado na estação de tratamento para eliminar os microrganismos, porém uma quantidade de cloro permanece para assegurar a desinfecção durante a distribuição na rede. A quantidade de cloro que deixou de ser utilizada na eliminação das bactérias é chamada de cloro residual livre.

De acordo com o Ministério da Saúde (Portaria 518/04) esta variável deverá ter uma concentração mínima de $0,5\ mg/L$ e máxima de $2,0\ mg/L$ para garantir a potabilidade. A partir de 2012 (Portaria 2914/11) a concentração mínima e máxima foi alterada para $0,2\ mg/L$ e $5,0\ mg/L$.

4.4.5 pH

O pH ou potencial hidrogeniônico da água é uma importante variável para o monitoramento da qualidade da água. É um valor que irá indicar se a água distribuída é ácida, alcalina ou neutra. Valores baixos de pH são considerados ácidos e podem deteriorar as tubulações por meio de corrosões, já valores altos (próximos ou acima de 9,0) podem provocar entupimentos. O ideal para o consumo humano e recomendado pelo Ministério da Saúde são valores na faixa entre 6,0 a 9,5, mesmo não apresentando riscos à saúde em qualquer valor.

4.4.6 Temperatura

A temperatura, dada em graus Celsius ($^{\circ}C$) é uma medida da intensidade de calor da água no momento da coleta. A temperatura da água pode alterar-se dependendo de condições naturais (energia solar) ou de despejos industriais.

4.4.7 Bactérias Heterotróficas

A variável bactérias heterotróficas, medida em *UFC/ml*, indica a quantidade de unidades formadoras de colônias de bactérias heterotróficas por mililitro de água. Essas bactérias são definidas como microrganismos que necessitam de carbono orgânico como nutrientes (DOMINGUES et al., 2007).

4.4.8 Coliformes Totais

A variável coliformes totais é mensurada em unidades formadoras de colônias por 100 *ml* (*UFC/100 ml*). No entanto, a SABESP considera esta variável como dicotômica, sendo que uma resposta zero indica a ausência ou a falta de sensibilidade para identificar a ocorrência de grupos de bactérias na água, e uma resposta igual a um indica a presença de coliformes na água independente da quantidade mensurada. Entre as espécies avaliadas como coliformes totais destacam-se como principais *Escherichia*, *Klebsiella*, *Enterobacter* e *Citrobacter* (BRASIL, 2004).

4.5 Variáveis climáticas

Cunha e Martins (2009) apresentam a caracterização climática dos municípios de Botucatu e São Manuel e destacam que os dois municípios apresentam a mesma classificação climática pelo método *Koppen* e uma pequena diferença pelo método de *Thornthwaite*, devido ao índice de umidade. De uma forma geral, Botucatu e São Manuel apresentam um clima úmido com uma pequena deficiência hídrica nos meses de abril, julho e agosto, enquanto que, nos meses de primavera-verão (final de setembro ao início de março), os índices hídricos são relativamente elevados. A temperatura média anual é de 20,3°C, apresentando temperatura média mais baixa no mês de julho (17,1°C) e mais alta em fevereiro (23,1°C), caracterizando-se como um clima temperado quente.

A seguir estão apresentadas, em detalhes, as características climáticas.

4.5.1 Temperatura Mínima, Máxima e Média

A temperatura do ar à superfície, também conhecida como temperatura do ar à sombra, fornece uma medida do aquecimento da atmosfera próxima à superfície terrestre. Sua unidade de medida é a escala Celsius ($^{\circ}\text{C}$) e, de acordo com Varejão-Silva (2006), sua mensuração é realizada sob uma condição padrão, para permitir comparações com outras localidades. Essa condição padrão estabelece que a temperatura seja registrada a uma altura entre 1,5 a 2,0 metros acima da superfície, em uma área plana, gramada e ao abrigo do sol, impedindo a incidência de raios solares nos equipamentos, e com venezianas que permitam a ventilação natural.

O termômetro é o instrumento de medida, possuindo características próprias para mensuração da temperatura mínima (termômetro de álcool) e máxima (termômetro de mercúrio). Já a temperatura média é calculada com base na metodologia do Instituto Nacional de Meteorologia (INMET), a qual se baseia em duas medidas realizadas em horários padronizados (normalmente às 9 e 21 horas) além das medidas de temperatura mínima e máxima (VAREJÃO-SILVA, 2006).

4.5.2 Precipitação Pluviométrica e Número de Dias Chuvosos

Segundo Varejão-Silva (2006), a quantidade de precipitação pluviométrica ou altura pluviométrica utilizada neste estudo, expressa em milímetros (mm), é mensurada com base no volume de água que se formaria em uma superfície horizontal impermeável com 1 m^2 de área, ou seja, é o volume de água precipitado por unidade de área horizontal da superfície terrestre. Portanto, uma precipitação de 1 mm representa uma queda de 1 litro de água por m^2 de superfície.

O número de dias chuvosos é estabelecido a partir da contagem de dias com uma quantidade mínima de precipitação, estabelecida por órgãos oficiais, tais como o Instituto Nacional de Meteorologia (INMET).

4.5.3 Umidade Relativa do Ar

A umidade do ar pode ser descrita como uma medida da quantidade de vapor de água existente da atmosfera. A umidade relativa do ar é dada pela relação

percentual entre a umidade relativa em um determinado momento e a umidade de saturação (capacidade máxima de vapor de água que a atmosfera consegue reter, dependente da temperatura e da pressão atmosférica). Desta forma, sua unidade de medida é em percentual (%) e, fisicamente, representa a proporção da umidade máxima possível que já se encontra preenchida. Existem vários instrumentos para sua mensuração, tais como o psicrômetro e sensores capacitivos (VAREJÃO-SILVA, 2006; BISCARO, 2007).

4.5.4 Velocidade dos Ventos

Os deslocamentos de ar no sentido horizontal à superfície terrestre decorrentes de diferentes valores de pressão atmosférica e de temperatura são chamados de ventos. Sua caracterização é dada pela identificação da direção e velocidade. O vento se desloca das áreas de maior pressão (regiões mais frias) para as de menor pressão (mais quentes), sendo sua velocidade influenciada pela diferença de pressão e de atrito com a superfície. A velocidade do vento é normalmente expressa em quilômetros por hora (*km/h*) ou metros por segundo (*m/s*) (VAREJÃO-SILVA, 2006; BISCARO, 2007).

4.5.5 Evaporação

Evaporação é o processo de passagem de um elemento em estado líquido para o estado gasoso. Em meteorologia a evaporação é usada para mensurar a passagem de água para a atmosfera em forma de vapor de água, observada no solo úmido sem vegetação, rios, lagos, oceanos ou qualquer outra superfície hídrica. Sua unidade de medida, assim como a precipitação, é dada em milímetros (*mm*), representando a transferência de vapor de água para atmosfera para cada metro quadrado de superfície horizontal hídrica (VAREJÃO-SILVA, 2006; BISCARO, 2007).

4.5.6 Insolação

O número de horas, durante um dia, em que o sol é visível por uma pessoa na superfície terrestre é chamado de insolação, ou seja, pode-se dizer que, de uma forma geral, a insolação é o intervalo de tempo entre o nascimento e o ocaso do sol,

descontando seu ocultamento por nuvens ou outros fenômenos da natureza (VAREJÃO-SILVA, 2006).

A partir da medida de insolação, é possível calcular duas novas variáveis, as quais não são objetos deste estudo. Uma destas variáveis é a chamada Razão de Insolação, obtida pela divisão da Insolação pelo fotoperíodo (total de horas do nascimento até o por do sol), indicando o percentual do dia em que o sol brilhou, sendo que o 0 significa dia inteiro nublado e uma resposta igual a 1 céu inteiro aberto e limpo. Outra variável é a Nebulosidade, dada pelo complementar da Razão de Insolação, indicando a proporção de Nebulosidade de um determinado dia.

4.5.7 Radiação Solar Global

A radiação solar global é a quantidade de radiação eletromagnética emitida pelo sol que atinge uma superfície horizontal ao nível do solo em um dado instante, promovendo calor e iluminação ao planeta. Representa a soma da radiação oriunda diretamente do sol (quando esse se mostra visível) com a radiação difundida na atmosfera (radiação que atinge o local considerado após ter sofrido um ou mais desvios na atmosfera). Sua unidade de medida é dada em quantidade de caloria por centímetro quadrado por dia ($cal.cm^{-2}.dia^{-1}$). A intensidade da radiação depende do horário, sendo a radiação máxima recebida por volta do meio dia (VAREJÃO-SILVA, 2006).

4.6 Técnicas de análise estatística na avaliação da qualidade da água

Existem várias pesquisas publicadas nos últimos anos relacionadas à análise de parâmetros físico-químicos e microbiológicos da água para o consumo humano. Entretanto, pesquisas que utilizam técnicas estatísticas multivariadas no entendimento da qualidade da água, principalmente na rede de distribuição, ainda são bastante escassas. Por isso, pretende-se, no presente estudo, enfatizar a importância das técnicas multivariadas no entendimento do problema prático.

Pinto e Maheschwari (2011) destacam que alguns países vêm buscando desenvolver metodologias mais eficientes para avaliar a qualidade da água para o consumo humano e espécies bióticas. Relatam que as técnicas de estatística multivariada

estão se destacando, pois facilitam a compreensão do fenômeno biológico, o qual, normalmente envolve conjuntos de dados com muitas variáveis correlacionadas entre si.

Entre as pesquisas relacionadas com a análise de características da água, Toledo e Nicolella (2002) avaliam a qualidade da água em microbacias de uso agrícola e urbano por meio de 10 variáveis, entre elas, fósforo total e dissolvido, pH, oxigênio dissolvido, amônia, nitrato, condutividade elétrica, turbidez, sólidos em suspensão e clorofila. Utilizando a técnica multivariada de análise fatorial, os autores construíram um indicador para qualidade da água, observando que as variáveis que mais influenciaram na composição do indicador de qualidade são oxigênio dissolvido, fósforo total, amônia e condutividade elétrica. Segundo os autores, vários índices podem ser utilizados para uma avaliação da qualidade das águas de uma bacia hidrográfica, sendo que as peculiaridades de cada região influenciam nesse processo, o que dificulta o estabelecimento de um indicador padrão para qualquer situação.

Campos, Farache Filho e Faria (2002) analisaram a influência da tubulação nos parâmetros de qualidade da água distribuída à população de Araraquara-SP, utilizando amostras de água coletadas em três diferentes locais: nas proximidades do reservatório de distribuição, em pontos intermediários e nas pontas da rede de distribuição. Utilizando um teste estatístico univariado (Teste de *Kruskal-Wallis*) verificaram que não foi possível encontrar diferença significativa entre os locais para todos os parâmetros avaliados.

Já em outro estudo realizado por Campos, Farache Filho e Faria (2003), os efeitos dos reservatórios domiciliares de água para o consumo no município de Araraquara-SP são avaliados em relação aos físico-químicos e microbiológicos. Para isso, realizam um estudo descritivo destas variáveis nas amostras coletadas dos reservatórios e posteriormente faz-se a comparação dos parâmetros analisados com a rede pública de distribuição, utilizando os testes de *Mann-Whitney* e de *Kruskal-Wallis* e, para comparação isolada das variáveis físico-químicas com o padrão de potabilidade, o teste Qui-quadrado. Os autores mostram que os reservatórios alteram os parâmetros de qualidade da água distribuída nos domicílios; porém eles não consideram o comportamento simultâneo das variáveis, utilizando-se de técnicas estatísticas envolvendo análises univariadas.

Bueno, Galbiatti e Borges (2005) comparam algumas variáveis de qualidade da água em duas condições de uso do solo: uma considerando pontos de coleta de dados ao longo do curso d'água em regiões de mata nativa e, outra, em regiões de

plantação de eucalipto. Utilizando testes para comparação de dois grupos em cada variável isolada, os autores observam que não existem diferenças significativas para as variáveis oxigênio dissolvido, temperatura, pH e matéria orgânica; no entanto, salientam que a cobertura vegetal é responsável por variações na temperatura da água, sugerindo que o uso do solo é importante para a qualidade da água de rios ou lagoas.

Para verificar se existem diferenças entre o período seco (meses de junho, julho, agosto e outubro) e chuvoso (meses de dezembro, janeiro, fevereiro e abril) quanto ao comportamento das variáveis físico-químicas e biológicas analisadas em amostras de água de uma represa que abastece tanques e viveiros para piscicultura na cidade de Pirassununga-SP, Bufon e Landim (2007) aplicam a metodologia multivariada de análise de agrupamento. As amostras são avaliadas em relação à temperatura, oxigênio dissolvido, condutividade, transparência, pH, alcalinidade, nitrito e nitrado, amônia, nitrogênio total, fósforo total, fosfato total dissolvido, material em suspensão e pigmento total. Por meio da técnica multivariada utilizada, verifica-se que os valores de fosfatos e fósforo total se relacionam diretamente com o material de suspensão na água apenas no período seco.

Sabino et al. (2008) realizam um estudo com o objetivo de reconhecer a qualidade da água por meio de análise multivariada, considerando a distribuição geográfica e temporal de características químicas e bacteriológicas na bacia do Rio Paraopeba, em Minas Gerais. Os dados são coletados em períodos denominados de chuva (janeiro, fevereiro e março) e de estiagem (julho, agosto e setembro), utilizando diversas variáveis físico-químicas e microbiológicas. Após uma análise da associação entre as variáveis pela correlação de Spearman, utiliza-se a análise de componentes principais para redução do número de variáveis, complementada com o teste de *Kruskal-Wallis* para verificar se existem diferenças significativas entre os pontos avaliados ao longo da bacia.

Felski et al. (2008) determinam a qualidade de amostras de águas tratada e não-tratada de áreas urbana e rural do município de Guarapuava-PR, utilizando análise descritiva de características físico-químicas e microbiológicas. São realizadas medidas das variáveis cloro residual, fluoreto, turbidez e coliformes em todas as amostras e comparadas aos padrões de potabilidade da Portaria 518/04. Os autores verificam que os sistemas de tratamento são eficientes ao atingir os padrões estabelecidos pela legislação em relação à desinfecção e ao controle da turbidez, mas chamam atenção para os níveis de cloro residual livre, pois seu excesso pode ser prejudicial à saúde humana.

Barbosa, Lage e Badaró (2009) realizam um estudo das características microbiológicas da água em bebedouros de um campus universitário de Ipatinga-MG. A quantidade de bactérias heterotróficas, a presença ou ausência de coliformes e a temperatura da água são avaliadas em 80 amostras coletadas entre os meses de agosto a dezembro de 2007. Estas observações são comparadas com os padrões de potabilidade da água estabelecidos pela Portaria 518/04 e o coeficiente de correlação de Pearson é utilizado para verificar a associação entre a temperatura da água e a quantidade de bactérias heterotróficas. Os resultados apontam que a qualidade da água pode ser classificada como potável, apesar de algumas amostras apresentarem uma quantidade de bactérias heterotróficas acima do estabelecido. Além disso, a correlação entre a temperatura e a quantidade de bactérias apresentou-se de forma negativa e fraca (abaixo de -0,2).

Para identificar possíveis associações entre as variáveis de qualidade da água para consumo humano e a observação de casos de diarreia no município de Vitória-ES, Queiroz, Heller e Silva (2009) utilizam o modelo de regressão linear simples e múltipla. Observam que a ocorrência de diarreia apresentou associações significativas com a turbidez, os coliformes totais e termotolerantes.

Freire (2012) apresenta estudo semelhante ao estudo de Campos, Farache Filho e Faria (2002), verificando os efeitos dos reservatórios domiciliares da cidade de Recife-PE nas variáveis físico-químicas e microbiológicas da água. Por meio de um estudo descritivo dessas variáveis e comparando com os padrões de potabilidade da legislação, verifica-se que o armazenamento domiciliar é um fator deteriorador da qualidade de água, pois foram observados altos índices de coliformes totais e *E. coli*, além do baixo nível de cloro residual. Para realizar esse estudo e obter esses resultados, são observados dados na saída da estação de tratamento, na rede de distribuição e na saída dos reservatórios domiciliares.

A fim de examinar a qualidade da água para o consumo humano no sistema de abastecimento público na cidade de Guarabira-PB, Oliveira et al. (2012a) realizam um estudo com dados de variáveis físico-químicas (cloro residual livre, cor aparente, pH, turbidez, alcalinidade e oxigênio consumido) e microbiológicas (coliformes totais e termotolerantes) obtidos a partir de coletas realizadas, de quatro em quatro dias, na torneira de uma residência (antes do armazenamento domiciliar), totalizando sete coletas. Em cada uma dessas coletas, sete amostras de água foram obtidas em horários distintos e

encaminhadas para as análises em laboratório. Os resultados são resumidos por meio da média, desvio padrão e coeficiente de variação para posteriormente serem comparados com os padrões de potabilidade da legislação vigente. Além disso, a técnica de análise de variância, complementada com o teste de comparações múltiplas de *Tukey* é utilizada para verificar se existem diferenças significativas entre as diferentes coletas. Os autores certificam-se de que todos os parâmetros analisados estão dentro dos padrões de qualidade estabelecidos pela Portaria 518/04, indicando que a água distribuída no município é apropriada ao consumo.

Com o mesmo objetivo e utilizando uma metodologia similar ao estudo anterior, Oliveira et al. (2012b) estudam a qualidade físico-química e microbiológica da água distribuída na cidade de Limoeiro do Norte-CE, avaliando se o sistema de abastecimento e distribuição do município está atendendo adequadamente os critérios estabelecidos pela legislação. Diferente do estudo anterior, quatro residências são analisadas, nas quais quatro coletas são realizadas a cada cinco dias. Da mesma forma que o estudo apresentado por Oliveira et al (2012a), os resultados apontam que a água distribuída à população de Limoeiro do Norte atende todos os requisitos estabelecidos pela legislação.

Buscando aprimorar a compreensão das informações dos estudos sobre a qualidade das águas, Almeida, Fernandes e Sieciechowicz (2013) aplicam técnicas de análise estatística multivariada para identificar os parâmetros de qualidade que mais se destacam. Entre os parâmetros analisados, os que mais se destacam no entendimento da qualidade da água na Bacia do Alto Iguaçu, localizada na região metropolitana de Curitiba-PR, são os referentes à degradação de material orgânico, aos nitrogênios e aos sólidos.

Freire e Castro (2014) aplicam o procedimento multivariado envolvendo a técnica de correlação canônica em um estudo de associação entre o uso e a ocupação do solo e a qualidade da água da bacia hidrográfica do rio Itapemirim no estado do Espírito Santo em um período de seca. Verificam que os locais com uma maior atividade humana se associam com um índice de degradação da qualidade de água mais elevado. Observam também que a atividade humana está relacionada à exposição do solo e às pastagens, já o índice de qualidade da água com a concentração de nitratos e oxigênio dissolvido.

4.7 Associação entre as variáveis climáticas e as de qualidade da água

Para avaliar a influência de variáveis climáticas no consumo anual de água na cidade de Phoenix, Arizona, EUA, Balling Jr. e Gober (2006) utilizam técnicas de estatística multivariada e verificam que o consumo anual de água está relacionado com a temperatura, precipitação e com os períodos de seca da região.

Silva et al. (2008a) descrevem a influência da precipitação pluviométrica nas características de qualidade da água do Rio Purus, Amazonas, observando alterações na turbidez, temperatura, condutividade, oxigênio dissolvido sólidos suspensos totais e pH.

Silva et al. (2008b) sugerem um modelo que busca relacionar variáveis socioeconômicas e climáticas com o consumo per capita de água, baseado em dados de um levantamento amostral de bairros do município de Cuiabá (MT). Em um estudo semelhante, Dias (2008) avalia a influência da renda familiar no consumo de água em domicílios amostrados em Belo Horizonte (MG).

Prathumratana et al. (2008) realizam um estudo com dados secundários (já existentes) de variáveis climáticas (precipitação, evaporação e temperatura do ar) e hidrológicas (o nível médio da água e o fluxo de descarga) do Rio Mekong, em quatro países da Ásia banhados pelo rio, avaliando as associações entre as variáveis de qualidade da água (total de sólidos suspensos-TSS, oxigênio dissolvido-DO, carbono orgânico dissolvido-COD, pH, condutividade, alcalinidade e metais, tais como Sódio-Na, Potássio-K, Ferro-Fe, Magnésio-Mg, Cálcio-Ca, Silício-Si, Cloro-Cl, Sulfato- SO_4^{-2} , Fosfato- PO_4^{-3} , ânion de Nitrato- NO_3^-) obtidas por amostras analisadas em laboratório. Para avaliar a relação entre essas variáveis, os autores utilizam o coeficiente de correlação linear de Pearson e verificam que a correlação entre as variáveis climáticas, hidrológicas e de qualidade da água possuem a mesma tendência nos diferentes pontos do rio avaliados. Observam-se correlações positivas significativas consideradas de razoável a forte entre as variáveis: precipitação, nível médio da água, fluxo de descarga e temperatura média do ar com as seguintes variáveis de qualidade da água: TSS, NO_3^- , PO_4^{-3} e COD. Já as variáveis DO, pH, condutividade, Ca, Mg, Na, K, alcalinidade, Cl, SO_4^{-2} e Si apresentaram correlações negativas significativas com todas as variáveis hidrológicas.

Com o objetivo de avaliar o relacionamento de variáveis climáticas com a qualidade de água em lagoas (pesque-pague) do município de Dracena (SP), Sandre

et al. (2009) utilizam dados climáticos diários de temperatura ambiente máxima e mínima, umidade relativa do ar, precipitação pluviométrica e velocidade do vento e dados da temperatura da água, concentração de oxigênio dissolvido, pH, transparência, alcalinidade e concentração de amônia total na superfície, obtidos de amostra de água coletadas a cada três dias dos viveiros. A técnica de análise de variância univariada é utilizada para comparar o comportamento de cada variável em dois períodos do ano (no verão e no inverno) e o coeficiente de correlação linear de Pearson para identificar o grau de relacionamento entre as variáveis levantadas. O estudo indica que, entre as variáveis climáticas avaliadas, somente a temperatura ambiente apresenta associação com os parâmetros de qualidade da água dos pesque-pague.

Souza et al (2014) apresentam um estudo da qualidade da água no rio Almada, sul da Bahia, avaliando as variáveis turbidez, pH, oxigênio dissolvido, temperatura, condutividade elétrica, fósforo total e coliformes fecais em 6 pontos distintos ao longo do rio. As amostras de água são coletadas a cada dois meses entre novembro de 2011 a outubro de 2012, abrangendo um período chuvoso e outro menos chuvoso. Para verificar possíveis diferenças espaciais e temporais, o teste de *Kruskal-Wallis* é utilizado. Já a análise de componentes principais é aplicada para explicar conjuntamente o comportamento das variáveis. Entre os resultados, verificam que a turbidez, os coliformes fecais, o pH e o oxigênio dissolvido apresentam valores mais elevados no período chuvoso, enquanto que temperatura e condutividade se mostram inferiores no período menos chuvoso. Com a análise de componentes principais, identificam que as unidades amostrais apresentaram uma tendência de se agregarem nos eixos de acordo com períodos mais e menos chuvosos. Observam também diferenças significativas entre alguns pontos de coletas, os quais estão relacionados com diferentes usos da água.

Em síntese, a revisão de literatura indica que são cada vez mais frequentes estudos que utilizam técnicas de estatística multivariada para caracterização climática ou da qualidade da água. Porém, poucos estudos nestas áreas procuram avaliar as associações entre conjuntos de variáveis considerando toda a estrutura de variabilidade dos dados, ou seja, a maioria releva as associações conjuntas entre as variáveis. Portanto, mesmo de caráter exploratório, as técnicas multivariadas podem expressar resultados mais próximos da realidade observada.

5 MATERIAIS E MÉTODOS

5.1 Materiais

O estudo exploratório do relacionamento entre as variáveis climáticas e as de qualidade da água foi realizado no município de Botucatu, interior do estado de São Paulo, a partir de dados do período de janeiro 2007 a dezembro de 2011, fornecidos pela estação meteorológica da FCA-UNESP Botucatu e SABESP, respectivamente.

Localizado na região centro-sul do Estado, com altitude média de 840 metros aproximadamente, o município de Botucatu apresenta um clima subtropical úmido. A água distribuída para a população é captada no Rio Pardo (Represa do Mandacaru, com altitude de 820 metros) e bombeada para a estação de tratamento de água (ETA), distante aproximadamente 5 km do rio e com altitude de 878 metros (Figura 1).

Os dados diários de variáveis climáticas para o município de Botucatu foram obtidos junto à Estação Meteorológica do Departamento de Recursos Naturais - Ciências Ambientais da Faculdade de Ciências Agronômicas de Botucatu, localizada nas coordenadas de 22°50'47,45 de latitude sul, 48°25'54,14 de longitude oeste em uma altitude de 786 metros.

Para a rotina de monitoramento da qualidade da água distribuída à população, algumas variáveis são analisadas periodicamente pelas empresas responsáveis pela captação e distribuição da água. Em Botucatu, a SABESP é responsável por este

trabalho, realizando o controle da qualidade da água em todo sistema de abastecimento do município, desde as estações de tratamentos e captação até o cavalete das residências.

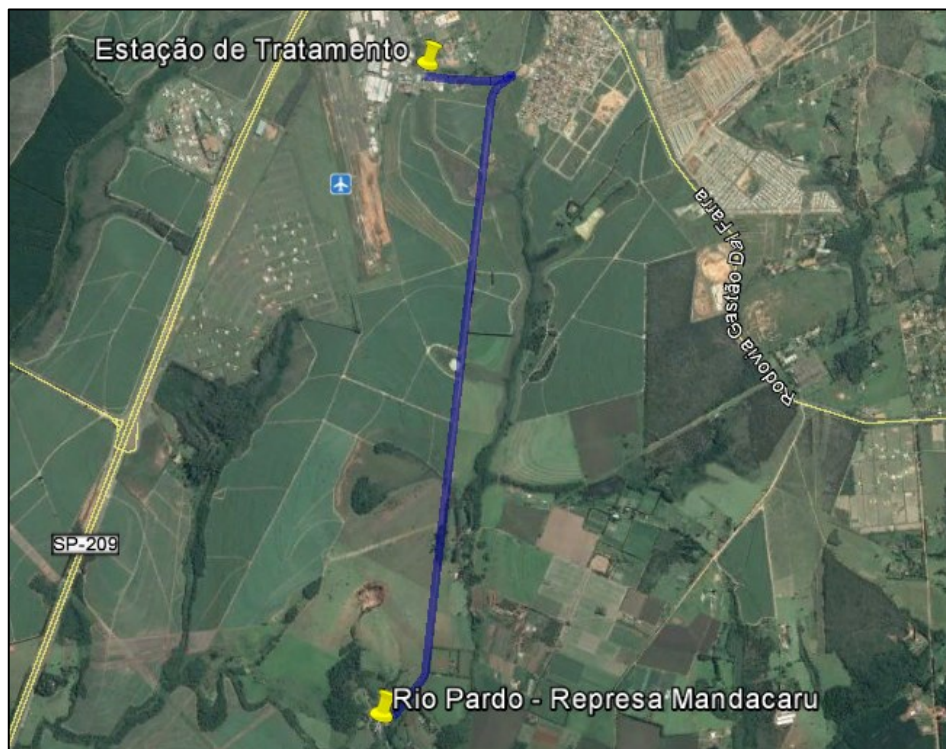


Figura 1. Localização de captação e estação de tratamento

Fonte: Google Earth (2014)

Este controle é realizado por meio de amostras diárias coletadas em aproximadamente 148 pontos de distribuição além da estação de tratamento (antes, durante e após o tratamento). Em algumas situações, amostras são realizadas por meio de solicitação dos usuários, mas esses casos não fizeram parte deste estudo.

A Figura 2, apresentada por Campos (2014), destaca no mapa do município de Botucatu os bairros com pontos de coleta de água na rede de distribuição (indicados pelos números). Alguns bairros apresentam mais de um ponto de coleta, no entanto, somente o bairro foi identificado no mapa.

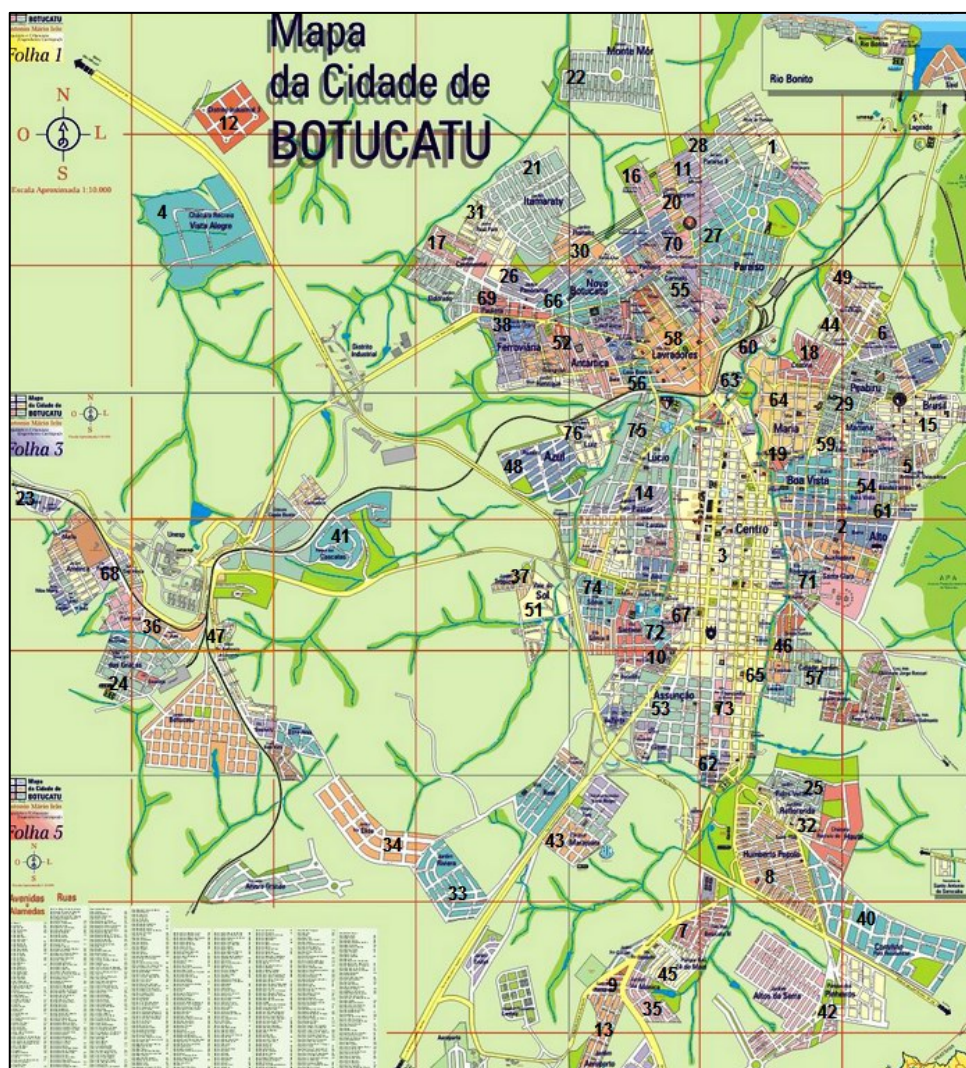


Figura 2. Localização dos bairros com pontos de coleta de amostras de água
Fonte: Campos (2014)

A quantificação dos valores numéricos das variáveis que trazem informações a respeito da qualidade da água para consumo humano foi realizada por análises físico-químicas e microbiológicas das amostras diárias dos pontos sorteados na rede de distribuição, complementadas por amostras coletadas na saída dos reservatórios, ou melhor, coletas realizadas após o tratamento da água realizado pela Estação de Tratamento da SABESP (ETA).

5.1.1 Variáveis de qualidade da água

As variáveis de qualidade da água, medidas pelo laboratório da SABESP a partir das amostras diárias coletadas na rede e no final de tratamento, são apresentadas a seguir, com seus respectivos equipamentos de medição.

- **Turbidez (TURB):** medida em unidades de turbidez (uT) pelo método nefelométrico (com um colorímetro/turbidímetro da marca Hach, modelo 2100 NA).
- **Cor aparente (CORAP):** avaliada em unidades de cor (UC) por meio de um espectrofotômetro e o equipamento colorímetro/turbidímetro Hach modelo 2100 NA.
- **Quantidade de flúor (F):** o equipamento utilizado para medir a concentração de flúor na água, em miligrama por litro (mg/L), foi o pHmetro da marca Orion modelo 720.
- **Cloro residual livre (CRL):** a mensuração do cloro residual livre, em miligrama por litro (mg/L), foi realizada por meio do método colorimétrico-DPD, utilizando um colorímetro AquaColor Policontrol.
- **Potencial hidrogeniônico (pH):** para sua mensuração do pH foi utilizado o pHmetro portátil, modelo pHtestr 30, da marca Oakton, considerando o método eletrométrico.
- **Temperatura da água (T):** para sua mensuração da temperatura da água no momento da coleta (em graus *Celsius*) foi utilizado um termômetro acoplado ao equipamento de medida do pH, ou seja, o pHmetro pHtestr 30, marca Oakton. Este equipamento é inserido em um béquer com a amostra de água coletada, fornecendo simultaneamente os valores de temperatura da água e pH.
- **Unidades formadoras de colônia de bactérias heterotróficas (HET):** esta variável é amplamente utilizada como indicador de qualidade bacteriológica da água potável e para sua quantificação, em unidades formadoras de colônias de bactérias por mililitro de água (UFC/ml), foi utilizado o método *Pour Plate*.
- **Ocorrência de coliformes totais (CT):** para mensurar a presença ou ausência de coliformes totais nas amostras de água foi utilizado o método da membrana filtrante.

No presente estudo, considera-se a qualidade da água definida por parâmetros de cor aparente, turbidez, flúor, cloro residual livre, pH, unidades formadoras de colônias de bactérias heterotróficas e proporção de ocorrências de coliformes totais estabelecidos dentro dos limites da legislação (Portaria 518/04), acrescida da mensuração da variável temperatura da água, a qual não pode ser controlada, no entanto, deve ser monitorada para possíveis ações de minimização do risco de presença de bactérias.

Apesar de atualmente estar em vigor a Portaria 2914/11, o presente estudo considera a Portaria 518/04, devido à necessidade de um volume maior de dados seguindo uma mesma legislação e oriundos dos mesmos equipamentos de análise dos parâmetros de qualidade. Esse fato limitou o presente trabalho no sentido de não ser possível utilizar dados dos últimos anos.

Os padrões de potabilidade indicados pela portaria brasileira para os parâmetros físico-químicos e microbiológicos, bem como seus métodos de mensurações, visam atender as especificações e metodologias padronizadas para a análise da água definidas em publicação da *American Public Health Association* (EATON, et al, 2005).

5.1.2 Variáveis climáticas

Em relação às características climáticas, foi utilizada a base de dados disponibilizada pela estação meteorológica mecânica da Faculdade de Ciências Agrônomicas da UNESP de Botucatu, considerando mensurações diárias relativas a um período de cinco anos. As variáveis deste estudo, bem como seus respectivos processos de mensuração são descritos a seguir.

- **Temperatura mínima, máxima e média (TMIN, TMAX, TMEDIA):** Para mensuração da temperatura mínima e máxima (em graus *Celsius*) foram utilizados e destacados na Figura 3, respectivamente, os termômetros de álcool (termômetro inferior) e de mercúrio (termômetro superior). Esses termômetros estavam abrigados em veneziana considerando as condições padrões para sua medição (Figura 4). A temperatura média foi calculada considerando a metodologia do Instituto Nacional de Meteorologia (INMET).



Figura 3. Termômetro de máxima e mínima



Figura 4. Veneziana abrigando os equipamentos

- **Precipitação pluviométrica média e Número de dias chuvosos (PP e DIASC):** O instrumento de medida utilizado para precipitação (em milímetros) foi o pluviômetro *Ville de Paris* (Figura 5), instalado a 1,5 m de altura. A partir da observação da presença de precipitação pluviométrica diária, definiu-se um indicador de presença (precipitação diferente de zero) ou ausência (precipitação igual à zero) de chuva no dia para a mensuração da variável número de dias chuvosos no mês.



Figura 5. Pluviômetro *Ville de Paris*

- **Umidade relativa do ar (UR):** No período considerado para este estudo, a mensuração da umidade relativa do ar (em %) foi realizada pelo psicrômetro, instalado dentro de abrigo termométrico (Figura 6). O psicrômetro possui um termômetro responsável por medir a temperatura de bulbo seco e um segundo para a temperatura do bulbo úmido, sendo que, com estes resultados, e por meio da equação psicrométrica, calcula-se a umidade do ar.



Figura 6. Psicrômetro

- **Velocidade média dos ventos (VV):** A velocidade dos ventos considerada neste estudo foi mensurada pela média da velocidade (em *km/h*) observada a uma altura de dois metros da superfície, utilizando um anemômetro (Figura 7).



Figura 7. Anemômetro

- **Evaporação (EV):** A mensuração da evaporação (em milímetros) foi realizada por um recipiente de água chamado tanque evaporímetro classe A (Figura 8), sendo observada diariamente a variação em seu nível no mesmo horário.



Figura 8. Tanque evaporímetro

- **Insolação (INSOL):** Para mensuração da insolação foi utilizado um instrumento chamado heliógrafo (Figura 9), o qual é constituído por uma esfera de cristal que projeta a radiação solar em uma tira de papel, produzindo sua queima e registrando, em horas, o período de insolação (VAREJÃO-SILVA, 2006).



Figura 9. Heliógrafo

- **Radiação solar global (RAD):** O instrumento para medir a radiação solar global (em quantidade de calor por centímetro quadrado por dia) foi o actinógrafo (Figura 10). Este instrumento é formado por duas placas de materiais metálicos distintos que, submetidos à mesma energia dilatam de forma diferenciada. Esta dilatação entre as placas é registrada no actinograma (cartão fixo em um eixo girante, calibrado para realizar o percurso de uma volta por dia).



Figura 10. Actinógrafo

5.2 Métodos

Para avaliar os dados mensais das variáveis climáticas e de características físico-químicas e microbiológicas da água para consumo humano, relativos aos cinco anos estudados, foi utilizado procedimento estatístico, de caráter exploratório, envolvendo técnicas de análise multivariada que consideram simultaneamente todas as variáveis e toda a estrutura de variação existente nos dados, ou seja, a inter e a intravariabilidade. Para isso, foram consideradas as seguintes técnicas multivariadas: Teste T^2 de Hotelling para comparação de vetores de médias, Análise de Correlação Canônica e Análise Fatorial. Todos os resultados analíticos foram discutidos no nível de 5% de significância (JOHNSON; WICHERN, 2007).

Os procedimentos analíticos no contexto da estatística inferencial devem ser planejados a partir da utilização de amostras aleatórias independentes. Porém, em algumas situações práticas o caráter do estudo torna-se observacional, utilizando amostras com ausência da aleatoriedade, situação encontrada neste estudo, no qual observações temporais das variáveis foram avaliadas. Nestes casos, o estudo exploratório permite que esses procedimentos possam ser utilizados, considerando que as conclusões ficam restritas ao conjunto de dados coletados.

Para os cálculos e execução das técnicas de análise de dados propostas foram utilizadas planilhas eletrônicas e o software livre R-Gui (THE COMPREHENSIVE...,2012).

5.2.1 Conjunto de dados

A partir dos dados diários das variáveis climáticas e de qualidade da água, fornecidos pela FCA-UNESP e pela SABESP de Botucatu, respectivamente, um novo banco de dados foi organizado de acordo com os objetivos da pesquisa e para que fosse compatível com os procedimentos de análise estatística.

Em relação às variáveis de qualidade da água, a unidade amostral para análise foi estabelecida pela média mensal dos pontos sorteados, considerando o período de cinco anos e referenciadas conforme descrição dada no Quadro 2. A única exceção é a variável ocorrência de coliformes totais (X_8), a qual representa a proporção mensal de amostras com ausência de bactérias.

Quadro 2. Descrição das variáveis físico-químicas e microbiológicas das amostras de água

Código	Variável	Unidade	Descrição
X_1	TURB	uT	Média mensal das mensurações de amostras diárias de turbidez.
X_2	CORAP	UC	Média mensal das mensurações de amostras diárias de cor aparente.
X_3	F	mg/l	Média mensal das mensurações de flúor realizadas nos pontos de coletas sorteados.
X_4	CRL	mg/l	Quantidade mensal média de cloro presente na água coletada.
X_5	pH		Valor médio mensal do potencial hidrogeniônico da água observada nas amostras diárias.
X_6	T	$^{\circ}C$	Temperatura média mensal da água no momento da coleta.
X_7	HET	UFC/ml	Indicação da quantidade média mensal de unidades formadoras de colônias de bactérias heterotróficas por ml de água.
X_8	CT		Variável indicando a proporção mensal de amostras com ausência de bactérias.

Estas variáveis são analisadas tanto na rede de distribuição de água para os domicílios, local que será chamado de grupo 1, quanto no final do tratamento da água, local chamado de grupo 2.

No mesmo sentido, a unidade amostral das variáveis climáticas é dada pela média mensal dos dados diários observados no período da pesquisa e referenciadas conforme descrição do Quadro 3. A exceção é a variável frequência de dias

chuvosos (X_9), a qual representa a frequência de dias com presença de chuvas no determinado mês.

Quadro 3. Descrição das variáveis climáticas observadas para o município de Botucatu

Código	Variável	Unidade	Descrição
X_9	DIASC	dias	Frequência de dias chuvosos no mês.
X_{10}	TMIN	°C	Média das temperaturas mínimas diárias observadas no mês, em graus Celsius.
X_{11}	TMAX	°C	Média das temperaturas máximas diárias observadas no mês, em graus Celsius.
X_{12}	TMED	°C	Média das temperaturas médias diárias do mês, em graus Celsius.
X_{13}	PP	mm	Precipitação pluviométrica média em cada mês, dada em milímetros.
X_{14}	UR	%	Umidade relativa do ar média no mês, em %.
X_{15}	VV	km/h	Velocidade média do vento, em quilômetros por hora.
X_{16}	EV	mm	Evaporação média, em milímetros.
X_{17}	ISOL	horas	A insolação indica o número médio mensal de horas de brilho solar entre o nascimento e o caso do Sol
X_{18}	RAD	$\frac{cal}{cm^2 \cdot dia}$	Radiação solar global média no mês. Esta variável é medida em <i>Caloria por cm² por dia</i> .

Denotando-se o número de variáveis físico-químicas e microbiológicas da água por p , o número de variáveis climáticas por q e i representando o local onde as p variáveis físico-químicas e microbiológicas da água foram avaliadas, com $i = 1$ para a observação na rede de distribuição (grupo 1) e $i = 2$ no final de tratamento (grupo 2), um vetor aleatório de dimensão $p + q$ é representado por:

$$\mathbf{X}'_i = [X_{i1} \quad \dots \quad X_{ip} \quad X_{i(p+1)} \quad \dots \quad X_{i(p+q)}]$$

sendo $X_1, \dots, X_p, X_{p+1}, \dots, X_{p+q}$ variáveis aleatórias avaliadas em $i = 1$ e 2 .

Considerando-se que as $p + q$ variáveis foram avaliadas em n_i meses em cada grupo, os dados dessas variáveis são representados genericamente por x_{ijk} , indicando a resposta da k – ésima variável (com $k = 1, \dots, p + q$) medida no j – ésimo mês (com $j = 1, \dots, n_i$) do i – ésimo grupo.

Dessa maneira, essas mensurações podem ser agrupadas em uma matriz genérica de dados $\mathbb{X}_{n \times (p+q)}$, definida por:

$$\mathbb{X} = \begin{bmatrix} x_{1\ 1\ 1} & \cdots & x_{1\ 1\ p} & x_{1\ 1\ (p+1)} & \cdots & x_{1\ 1\ (p+q)} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ x_{1\ n_1\ 1} & \cdots & x_{1\ n_1\ p} & x_{1\ n_1\ (p+1)} & \cdots & x_{1\ n_1\ (p+q)} \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ x_{2\ n\ 1} & \cdots & x_{2\ n\ p} & x_{2\ n\ (p+1)} & \cdots & x_{2\ n\ (p+q)} \end{bmatrix} = \begin{bmatrix} \mathbb{X}_1^{(1)} & \mathbb{X}_1^{(2)} \\ \mathbb{X}_2^{(1)} & \mathbb{X}_2^{(2)} \end{bmatrix} = [\mathbb{X}_\bullet^{(1)} \quad \mathbb{X}_\bullet^{(2)}]$$

sendo que as $n = n_1 + n_2$ linhas são observações multivariadas representadas genericamente pelo vetor linha de resposta $\mathbf{x}'_{ij}$, ou seja, a j – ésima observação (com $j = 1, \dots, n$) das $p + q$ variáveis no i – ésimo grupo.

As submatrizes $\mathbb{X}_i^{(1)}$ e $\mathbb{X}_i^{(2)}$ (com $i = 1, 2$), de dimensões $n_i \times p$ e $n_i \times q$, respectivamente, representam respectivamente as observações das p variáveis físico-químicas e microbiológicas da água e das q variáveis climáticas no i – ésimo grupo.

Genericamente, essas matrizes são definidas por $\mathbb{X}_i^{(l)}$, onde $l = 1$ para o conjunto de variáveis físico-químicas e microbiológicas da água e $l = 2$ para o conjunto de variáveis climáticas, com $\mathbb{X}_\bullet^{(l)}$ envolvendo os dados dos dois grupos conjuntamente.

Como os dados das variáveis climáticas são considerados para o município de Botucatu, não sendo observados em dois locais diferentes, as submatrizes $\mathbb{X}_1^{(2)}$ e $\mathbb{X}_2^{(2)}$ são formadas pelos mesmos dados (submatrizes idênticas).

5.2.2 Teste T^2 de Hotelling para comparação de dois vetores de médias

O teste T^2 de Hotelling consiste em uma generalização do procedimento univariado (teste t de Student) para testar hipóteses a respeito da média de uma variável com distribuição normal, avaliada em dois grupos independentes. Quando várias variáveis são avaliadas com o interesse de compará-las em duas populações multivariadas, deve-se considerar toda a estrutura de variação existente dentro e entre estas variáveis. Desta forma, torna-se necessário um teste para comparação de dois vetores de médias, situação que exige a estatística T^2 de Hotelling (JOHNSON; WICHERN, 2007).

Pretende-se nesse estudo comparar o comportamento das variáveis físico-químicas e microbiológicas da água na rede de distribuição (Grupo 1) e no final da estação de tratamento (Grupo 2).

Genericamente, os dados dessas variáveis são representados por x_{ijk} , indicando a mensuração da k – ésima variável no j – ésimo mês (com $j = 1, \dots, n_i$) do i – ésimo grupo. Para esse estudo somente as p variáveis físico-químicas e microbiológicas da água são consideradas nos dois grupos, ou seja, $k = 1, \dots, p$. Dessa maneira, a estrutura dos dados utilizados pode ser observada na Tabela 1.

Tabela 1. Representação genérica dos dados das características físico-químicas e microbiológicas analisadas na rede de distribuição e final de tratamento

Grupo	Mês	Variável				Vetor de resposta
		X_1	X_2	...	X_p	
Grupo 1 (Rede de Distribuição)	1	$x_{1\ 1\ 1}$	$x_{1\ 1\ 2}$	...	$x_{1\ 1\ p}$	$\mathbf{x}'_{1\ 1}$
	2	$x_{1\ 2\ 1}$	$x_{1\ 2\ 2}$	...	$x_{1\ 2\ p}$	$\mathbf{x}'_{1\ 2}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	n_1	$x_{1\ n_1\ 1}$	$x_{1\ n_1\ 2}$	...	$x_{1\ n_1\ p}$	$\mathbf{x}'_{1\ n_1}$
Grupo 2 (Final de Tratamento)	1	$x_{2\ 1\ 1}$	$x_{2\ 1\ 2}$	...	$x_{2\ 1\ p}$	$\mathbf{x}'_{2\ 1}$
	2	$x_{2\ 2\ 1}$	$x_{2\ 2\ 2}$	...	$x_{2\ 2\ p}$	$\mathbf{x}'_{2\ 2}$
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	n_2	$x_{2\ n_2\ 1}$	$x_{2\ n_2\ 2}$	...	$x_{2\ n_2\ p}$	$\mathbf{x}'_{2\ n_2}$

Considerando-se cada grupo com $n_i = 60$ unidades amostrais independentes e $n = n_1 + n_2$, então o vetor linha de resposta multivariada $\mathbf{x}'_{i\ j}$ representa a j – ésima observação das $p = 8$ variáveis no i – ésimo grupo.

A representação matricial dos dados apresentados na Tabela 1 é dada por uma matriz de dimensão $n \times p$, indicada por $\mathbb{X}_{\bullet}^{(1)}$, ou seja,

$$\mathbb{X}_{\bullet}^{(1)} = \begin{bmatrix} x_{1\ 1\ 1} & \dots & x_{1\ 1\ p} \\ \vdots & \ddots & \vdots \\ x_{1\ n_1\ 1} & \ddots & x_{1\ n_1\ p} \\ \vdots & \vdots & \vdots \\ x_{2\ n\ 1} & \dots & x_{2\ n\ p} \end{bmatrix}.$$

Nesse contexto, o vetor aleatório de dimensão p do i – éximo grupo é representado por:

$$\mathbf{X}_i^{(1)} = [X_{i1} \quad X_{i2} \quad \cdots \quad X_{ip}].$$

O entendimento do comportamento biológico, considerando-se simultaneamente as respostas das variáveis físico-químicas e microbiológica da água consumida em Botucatu em função do local onde essas variáveis foram avaliadas (rede de distribuição e final de tratamento), pode ser estabelecido por medidas que envolvam tanto a tendência central como a variabilidade dos dados.

Sendo $\boldsymbol{\mu}_1$ e $\boldsymbol{\mu}_2$ os vetores p -dimensionais de médias populacionais dos grupos 1 e 2 respectivamente, as hipóteses estatísticas sobre a igualdade de respostas dos grupos são dadas por:

$$H_0: \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 = \mathbf{0} \quad \text{e} \quad H_1: \boldsymbol{\mu}_1 - \boldsymbol{\mu}_2 \neq \mathbf{0},$$

sendo $\mathbf{0}$ um vetor nulo composto por p zeros.

Uma suposição necessária para a realização do teste T^2 de Hotelling é que o vetor aleatório $\mathbf{X}_i^{(1)}$ (com $i = 1, 2$) tenha aderência a uma distribuição normal multivariada com vetor de valores esperados $\boldsymbol{\mu}'_i = (\mu_{i1} \quad \cdots \quad \mu_{ip})$ e matriz de variância-covariância não singular dada por:

$$\Sigma_{i \, p \times p} = \begin{bmatrix} \sigma_{11}^{(i)} & \sigma_{12}^{(i)} & \cdots & \sigma_{1p}^{(i)} \\ \sigma_{21}^{(i)} & \sigma_{22}^{(i)} & \cdots & \sigma_{2p}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1}^{(i)} & \sigma_{p2}^{(i)} & \cdots & \sigma_{pp}^{(i)} \end{bmatrix} \text{ para } i = 1, 2.$$

Segundo Mardia (1974) este teste é robusto a não normalidade quando os dados têm distribuições aproximadamente simétricas, ou seja, a falta de normalidade não afeta a conclusão do teste.

Considerando amostras aleatórias independentes das p variáveis para os dois grupos, o vetor de médias do grupo i pode ser estimado por $\bar{\mathbf{X}}'_i = (\bar{x}_{i1} \quad \cdots \quad \bar{x}_{ip})$ sendo

$$\bar{x}_{ik} = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ijk}, \text{ para } k = 1, \dots, p \text{ e } i = 1, 2.$$

Já a matriz de variância-covariância do grupo i pode ser estimada

$$\text{por } S_{i \, p \times p} = \begin{bmatrix} s_{11}^{(i)} & s_{12}^{(i)} & \dots & s_{1p}^{(i)} \\ s_{21}^{(i)} & s_{22}^{(i)} & \dots & s_{2p}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ s_{p1}^{(i)} & \dots & \dots & s_{pp}^{(i)} \end{bmatrix} \text{ onde}$$

$$s_{kk}^{(i)} = \left(\sum_{j=1}^{n_i} x_{ijk}^2 - n_i \bar{x}_{ik}^2 \right) (n_i - 1)^{-1} \text{ e}$$

$$s_{kk'}^{(i)} = \left(\sum_{j=1}^{n_i} x_{ijk} x_{ijk'} - n_i \bar{x}_{ik} \bar{x}_{ik'} \right) (n_i - 1)^{-1}, \quad k \neq k' \text{ e } i = 1, 2.$$

Dois situações quanto à estrutura geral de variabilidade devem ser observadas para realização deste teste: quando as matrizes de variância-covariância dos dois grupos são consideradas iguais e quando são consideradas desiguais. Para isso, há necessidade de realizar um teste multivariado de homogeneidade de matrizes de variâncias-covariâncias, tal como o teste de Box, apresentado em Johnson e Wichern (2007).

Considerando a situação na qual as matrizes de variâncias-covariâncias são iguais, ou seja, $\Sigma_1 = \Sigma_2$, a estatística de teste dos vetores de médias é expressa por:

$$T^2 = [(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)]' \left[\left(\frac{1}{n_1} + \frac{1}{n_2} \right) \mathbb{S} \right]^{-1} [(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)]$$

$$\text{com } \mathbb{S} = \frac{(n_1-1)\mathbb{S}_1 + (n_2-1)\mathbb{S}_2}{n_1 + n_2 - 2}.$$

A regra de decisão é baseada no quantil de ordem $100(1 - \alpha)\%$ da distribuição F , ou seja, se o valor da estatística T^2 for maior que $\frac{(n_1+n_2-2)p}{n_1+n_2-p-1} F_{(\alpha, p, n_1+n_2-p-1)}$ a hipótese inicial $H_0 (\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2)$ deve ser rejeitada.

A situação na qual as matrizes de variâncias-covariâncias são desiguais, ou seja, $\Sigma_1 \neq \Sigma_2$, a estatística T^2 de Hotelling para os vetores de médias é expressa por:

$$T^2 = [(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)]' \left[\frac{1}{n_1} \mathbb{S}_1 + \frac{1}{n_2} \mathbb{S}_2 \right]^{-1} [(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)].$$

A hipótese inicial H_0 ($\boldsymbol{\mu}_1 = \boldsymbol{\mu}_2$) é rejeitada quando o valor calculado da estatística T^2 é superior ao quantil de ordem $100(1 - \alpha)\%$ da distribuição de probabilidade qui-quadrado dado por $\chi^2_{(\alpha, p)}$.

Caso a hipótese inicial de igualdade dos vetores de médias (H_0) seja rejeitada, os intervalos de confiança simultâneos (ICS) tornam-se interessantes.

Neste sentido, o ICS de Hotelling para as diferenças entre os componentes dos vetores de médias na situação em que $\Sigma_1 = \Sigma_2$, com limites de $100(1 - \alpha)\%$ de confiança são estabelecidos por:

$$ICS[\mathbf{a}'(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)]: \mathbf{a}'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \pm \sqrt{\frac{(n_1+n_2-2)p}{n_1+n_2-p-1} F_{(\alpha; p; n_1+n_2-p-1)} \frac{n_1+n_2}{n_1 n_2} \mathbf{a}' \mathbb{S} \mathbf{a}},$$

sendo que $\mathbf{a}' = [a_1, a_2, \dots, a_p]$ é um vetor indicador, o qual é composto por 0's e 1, sendo que o número 1 assume a posição que indica a diferença de médias de interesse.

Na situação em que $\Sigma_1 \neq \Sigma_2$, o ICS de Hotelling é expresso por:

$$ICS[\mathbf{a}'(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)]: \mathbf{a}'(\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \pm \sqrt{\chi^2_{(\alpha, p)} \mathbf{a}' \left(\frac{1}{n_1} \mathbb{S}_1 + \frac{1}{n_2} \mathbb{S}_2 \right) \mathbf{a}}.$$

Os resultados dos intervalos de confiança simultâneos de Hotelling para as diferenças entre os componentes dos vetores de médias indicam a existência de diferença entre os componentes dos dois grupos quando o valor zero não está dentro do intervalo, caso contrário não é possível dizer que há uma diferença significativa.

5.2.3 Análise de Correlação Canônica

O objetivo da análise de correlação canônica é avaliar o grau de associação entre dois conjuntos de variáveis aleatórias quantitativas, resumindo as informações de cada conjunto em combinações lineares. Para cada conjunto de variáveis, várias combinações lineares são determinadas, as quais são chamadas de variáveis canônicas. As correlações entre as variáveis canônicas do primeiro conjunto com as variáveis canônicas do segundo são chamadas de correlações canônicas. Os coeficientes dessas combinações lineares são determinados de forma a se obter a maior associação possível entre as variáveis canônicas de cada conjunto e combinações lineares das variáveis originais não correlacionadas (JOHNSON; WICHERN, 2007 e MINGOTI, 2005).

A técnica será aplicada para analisar a associação entre o conjunto de variáveis físico-químicas e microbiológicas da água avaliadas na rede de distribuição do município de Botucatu (grupo 1) e o conjunto de variáveis climáticas. A mesma estrutura pode ser utilizada para elaboração de um conjunto de dados com as variáveis da água observadas o final do tratamento (grupo 2).

Dessa forma, considerando p variáveis aleatórias, representando os parâmetros físico-químicos e microbiológicos da água, e q variáveis climáticas, para aplicação da técnica de correlação canônica considerando a avaliação no i – ésimo grupo, o vetor aleatório de dimensão $p + q$ dado por

$$\mathbf{X}'_i = [X_{i1} \quad \cdots \quad X_{ip} \quad X_{i(p+1)} \quad \cdots \quad X_{i(p+q)}]$$

é dividido em dois conjuntos de vetores aleatórios e expresso por:

$$\mathbf{X}'_i = [\mathbf{X}'^{(1)}_i \quad \mathbf{X}'^{(2)}_i]$$

sendo que $\mathbf{X}'^{(1)}_i = (X_{i1} \quad \cdots \quad X_{ip})$ é um vetor aleatório representando o conjunto de p variáveis físico-químicas e microbiológicas da água e $\mathbf{X}'^{(2)}_i = (X_{i(p+1)} \quad \cdots \quad X_{i(p+q)})$ um vetor representando o conjunto de q variáveis climáticas.

Considerando o estudo realizado para um determinado grupo, ou seja, fixando o i – ésimo grupo (variáveis da água observadas na rede de distribuição ou no final de tratamento), os dados dessas $p = 8$ variáveis físico-químicas e microbiológicas da água, e $q = 10$ variáveis climáticas, mensuradas no período de 60 meses ($n_i = 60$) e expressas pelas médias mensais, estão organizados genericamente na Tabela 2 com a resposta da k – ésima variável (com $k = 1, \dots, p + q$), medida no j – ésimo mês (com $j = 1, \dots, n_i$) representada por x_{jk} .

Tabela 2. Representação genérica dos dados das variáveis de qualidade da água e climáticas, fixando o i – ésimo grupo

Mês	Variáveis da água				Variáveis climáticas			Vetor de resposta
	X_1	X_2	...	X_p	X_{p+1}	...	X_{p+q}	
1	$x_{1\ 1}$	$x_{1\ 2}$	...	$x_{1\ p}$	$x_{1\ p+1}$	...	$x_{1\ p+q}$	$\mathbf{x}'_1$
2	$x_{2\ 1}$	$x_{2\ 2}$	...	$x_{2\ p}$	$x_{2\ p+1}$	...	$x_{2\ p+q}$	$\mathbf{x}'_2$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
n_i	$x_{n_i\ 1}$	$x_{n_i\ 2}$	...	$x_{n_i\ p}$	$x_{n_i\ p+1}$	...	$x_{n_i\ p+q}$	$\mathbf{x}'_{n_i}$

O vetor linha de resposta $\mathbf{x}'_j$ representa a j – ésima medida das $p + q$ variáveis observadas no i – ésimo grupo.

Esta estrutura de dados é representada pela matriz de dimensão $n_i \times (p + q)$ expressa por:

$$\mathbb{X}_i^{(l)} = \begin{pmatrix} x_{1\ 1} & \dots & x_{1\ 8} & x_{1\ 9} & \dots & x_{1\ 18} \\ \vdots & \ddots & & \ddots & & \vdots \\ x_{n_i\ 1} & \dots & x_{n_i\ 8} & x_{n_i\ 9} & \dots & x_{n_i\ 18} \end{pmatrix}$$

com i fixo definindo o local de observação das variáveis físico-químicas e microbiológicas da água (Grupo 1 ou 2) e l indicando os dois conjuntos de variáveis ($l = 1$ para o conjunto de variáveis da água e $l = 2$ as variáveis climáticas).

Para simplificação da notação, como o índice i será fixo para indicar se os dados são provenientes da rede de distribuição ou do final de tratamento, optou-se pela omissão do mesmo na representação genérica dos dados. Dessa forma, o vetor de médias amostral do vetor aleatório $\mathbf{X}'$ é dado por

$$\bar{\mathbf{x}}' = (\bar{\mathbf{x}}'^{(1)} \quad \bar{\mathbf{x}}'^{(2)}),$$

onde $\bar{\mathbf{x}}'^{(1)} = (\bar{x}_1 \quad \dots \quad \bar{x}_p)$, $\bar{\mathbf{x}}'^{(2)} = (\bar{x}_{p+1} \quad \dots \quad \bar{x}_{p+q})$ e $\bar{x}_k = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{jk}$, para $k = 1, \dots, p + q$.

Já a matriz de variância-covariância do vetor $\mathbf{X}'$ pode ser estimada por

$$\mathbb{S}_{(p+q) \times (p+q)} = \begin{pmatrix} \mathbb{S}^{(1,1)} & \mathbb{S}^{(1,2)} \\ \mathbb{S}^{(2,1)} & \mathbb{S}^{(2,2)} \end{pmatrix}.$$

sendo que $\mathbb{S}^{(1,1)}$ e $\mathbb{S}^{(2,2)}$ são matrizes de dimensões $p \times p$ (ou 8×8) e $q \times q$ (ou 10×10), estimando, respectivamente, as associações existentes entre as variáveis aleatórias que estão no primeiro e no segundo conjuntos, ou seja, dos vetores $\mathbf{X}'^{(1)}$ e $\mathbf{X}'^{(2)}$. A matriz $\mathbb{S}^{(1,2)}$, de dimensão $p \times q$ (ou 8×10), apresenta as covariâncias amostrais entre as variáveis do vetor $\mathbf{X}'^{(1)}$ e do $\mathbf{X}'^{(2)}$, enquanto que $\mathbb{S}^{(2,1)}$ denota a matriz transposta de $\mathbb{S}^{(1,2)}$.

Portanto, o elemento genérico da matriz $\mathbb{S}$ pode ser representado por $s_{kk'}$, de tal forma que:

- para $k = 1, \dots, p$ e $k' = 1, \dots, p$ representará os elementos genéricos $(s_{kk'}^{(1,1)})$ da matriz $\mathbb{S}^{(1,1)}$,
- para $k = 1, \dots, p$ e $k' = p + 1, \dots, p + q$ representará os elementos genéricos $(s_{kk'}^{(1,2)})$ da matriz $\mathbb{S}^{(1,2)}$,
- para $k = p + 1, \dots, p + q$ e $k' = 1, \dots, p$ representará os elementos genéricos $(s_{kk'}^{(2,1)})$ da matriz $\mathbb{S}^{(2,1)}$ e
- para $k = p + 1, \dots, p + q$ e $k' = p + 1, \dots, p + q$ representará os elementos genéricos $(s_{kk'}^{(2,2)})$ da matriz $\mathbb{S}^{(2,2)}$.

A análise de correlação canônica procura estudar as associações existentes entre os dois conjuntos de variáveis $(\mathbf{X}'^{(1)}$ e $\mathbf{X}'^{(2)})$ por meio de variáveis canônicas. Nesse contexto, as variáveis canônicas construídas a partir de combinações lineares das variáveis físico-químicas e microbiológicas da água, representadas por u_t , e a

partir das variáveis climáticas, representadas por v_t , com $t = 1, \dots, \min(p, q)$, são expressas por:

$$u_t = \mathbf{a}'_t \mathbf{X}'^{(1)} = a_{t1}X_1 + \dots + a_{tp}X_p$$

$$v_t = \mathbf{b}'_t \mathbf{X}'^{(2)} = b_{t1}X_{p+1} + \dots + b_{tq}X_{p+q}$$

Os vetores $\mathbf{a}'_t$ e $\mathbf{b}'_t$ com $t = 1, \dots, \min(p, q)$, de dimensões $p \times 1$ e $q \times 1$, respectivamente, são obtidos de forma a maximizar a correlação entre as variáveis canônicas u_t e v_t no conjunto das combinações lineares com variâncias iguais a um ($Var(u_t) = Var(v_t) = 1$) e apresentar todos os pares de combinações lineares, (u_t, v_t) e $(u_{t'}, v_{t'})$, não correlacionados entre si (para $t \neq t'$).

O coeficiente de correlação canônico é a correlação em valor absoluto entre as variáveis canônicas (JOHNSON; WICHERN, 2007 e MINGOTI, 2005) e é expresso por:

$$r_{u_t, v_t} = \frac{\mathbf{a}'_t \text{Cov}(\mathbf{X}'^{(1)}, \mathbf{X}'^{(2)}) \mathbf{b}_t}{\sqrt{(\mathbf{a}'_t \text{Var}(\mathbf{X}'^{(1)}) \mathbf{a}_t) (\mathbf{b}'_t \text{Var}(\mathbf{X}'^{(2)}) \mathbf{b}_t)}} = \frac{\mathbf{a}'_t \mathbb{S}^{(1,2)} \mathbf{b}_t}{\sqrt{(\mathbf{a}'_t \mathbb{S}^{(1,1)} \mathbf{a}_t) (\mathbf{b}'_t \mathbb{S}^{(2,2)} \mathbf{b}_t)}}$$

Para este estudo, como $p = 8$ e $q = 10$, é possível estabelecer oito pares de variáveis canônicas, ou seja, $t = 1, \dots, \min(p; q)$ e, consequentemente, estabelecer oito correlações canônicas entre os dois conjuntos de variáveis. Essas correlações são determinadas em ordem decrescente de magnitude, ou seja, o primeiro par de variáveis canônicas é o que apresenta maior correlação; o segundo, a segunda maior correlação e assim sucessivamente.

Posteriormente, as correlações canônicas serão avaliadas por meio de testes de hipóteses para verificar se os resultados são significativos e, consequentemente, verificar se as variáveis canônicas correspondentes são variáveis importantes para resumir as informações dos dois conjuntos de dados do estudo.

Segundo Mingoti (2005), a construção matemática para determinar os vetores de constantes $\mathbf{a}'_t$ e $\mathbf{b}'_t$, para $t = 1, \dots, \min(p; q)$, satisfazendo os critérios apresentados anteriormente para estimar as variáveis canônicas, não é comum. No entanto, é sabido que esses vetores são soluções não triviais do seguinte sistema de equações:

$$\begin{cases} \left(\mathbb{S}^{(1,2)} (\mathbb{S}^{(2,2)})^{-1} \mathbb{S}^{(2,1)} - \lambda_t \mathbb{S}^{(1,1)} \right) \mathbf{a}_t = \mathbf{0} \\ \left(\mathbb{S}^{(2,1)} (\mathbb{S}^{(1,1)})^{-1} \mathbb{S}^{(1,2)} - \lambda_t \mathbb{S}^{(2,2)} \right) \mathbf{b}_t = \mathbf{0} \end{cases}$$

sendo λ_t o t – *ésimo* autovalor por ordem de magnitude, obtidos a partir da matriz $(\mathbb{S}^{(1,1)})^{-1} \mathbb{S}^{(1,2)} (\mathbb{S}^{(2,2)})^{-1} \mathbb{S}^{(2,1)}$ ou, de forma equivalentemente, da matriz $(\mathbb{S}^{(2,2)})^{-1} \mathbb{S}^{(2,1)} (\mathbb{S}^{(1,1)})^{-1} \mathbb{S}^{(1,2)}$, com $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_t$, e os vetores $(\mathbf{a}'_t$ e $\mathbf{b}'_t)$ são obtidos a partir dos autovetores associados a cada autovalor.

Com esses resultados é possível determinar os t pares de variáveis canônicas (u_t, v_t) para $t = 1, \dots, \min(p; q)$, pois, como descrito anteriormente, os componentes $(\mathbf{a}'_t$ e $\mathbf{b}'_t)$ são dados a partir dos autovetores associados a cada autovalor λ_t . Além disso, as correlações canônicas entre os dois conjuntos de variáveis podem ser calculadas por:

$$r_{u_t, v_t} = \sqrt{\lambda_t}, \text{ com } t = 1 \dots \min(p, q)$$

Da mesma forma, os autovalores e seus respectivos autovetores e, conseqüentemente, as correlações e variáveis canônicas também podem ser determinados utilizando-se de variáveis padronizadas e, nesse caso, deve-se substituir as matrizes de variâncias e covariâncias pelas respectivas matrizes de correlações das variáveis originais, ou seja:

$$\begin{cases} \left(\mathbb{R}^{(1,2)} (\mathbb{R}^{(2,2)})^{-1} \mathbb{R}^{(2,1)} - \lambda_t \mathbb{R}^{(1,1)} \right) \mathbf{a}_t = \mathbf{0} \\ \left(\mathbb{R}^{(2,1)} (\mathbb{R}^{(1,1)})^{-1} \mathbb{R}^{(1,2)} - \lambda_t \mathbb{R}^{(2,2)} \right) \mathbf{b}_t = \mathbf{0} \end{cases}$$

sendo que $\mathbb{R}^{(1,1)}$ e $\mathbb{R}^{(2,2)}$ são as matrizes de correlações estimadas das variáveis dos vetores $\mathbf{X}'^{(1)}$ e $\mathbf{X}'^{(2)}$, respectivamente. Já $\mathbb{R}^{(1,2)}$ (ou sua transposta $\mathbb{R}^{(2,1)}$) é a matriz de correlação entre as variáveis do vetor $\mathbf{X}'^{(1)}$ e aquelas que estão no vetor $\mathbf{X}'^{(2)}$.

5.2.3.1 Cargas canônicas e variância total explicada

Para auxiliar na interpretação das variáveis canônicas e contribuir com a qualidade das análises, as correlações das variáveis canônicas com as variáveis originais e o percentual da variância total explicada, podem ser utilizados (MINGOTI, 2005).

As correlações entre as variáveis canônicas u_t e v_t , para $t = 1, \dots, \min(p; q)$, e as variáveis originais de seus respectivos conjuntos, conhecidas como cargas canônicas, são dadas por:

$$\begin{aligned} R_{u_t, X^{(1)}} &= \mathbb{R}^{(1,1)} \mathbf{a}_t \\ R_{v_t, X^{(2)}} &= \mathbb{R}^{(2,2)} \mathbf{b}_t \end{aligned}$$

Já as correlações entre a variável canônica u_t e as variáveis originais do outro conjunto (variáveis do vetor $\mathbf{X}^{(2)}$), e as correlações da variável canônica v_t com as variáveis canônicas do vetor $\mathbf{X}^{(1)}$ (correlações cruzadas), conhecidas como cargas canônicas cruzadas, são dadas por:

$$\begin{aligned} R_{u_t, X^{(2)}} &= \mathbb{R}^{(2,1)} \mathbf{a}_t \\ R_{v_t, X^{(1)}} &= \mathbb{R}^{(1,2)} \mathbf{b}_t \end{aligned}$$

Para determinar a proporção da variância total dos dados originais explicada por cada variável canônica u_t e v_t , para $t = 1, \dots, \min(p; q)$, representada por $R_{X^{(1)}|u_t}^2$ e $R_{X^{(2)}|v_t}^2$, respectivamente, faz-se:

$$\begin{aligned} R_{X^{(1)}|u_t}^2 &= \frac{\sum_{k=1}^p (r_{u_t X_k})^2}{p} \\ R_{X^{(2)}|v_t}^2 &= \frac{\sum_{k=p+1}^{p+q} (r_{v_t X_k})^2}{q} \end{aligned}$$

sendo que $r_{u_t X_k}$ são as cargas canônicas do primeiro conjunto de variáveis, ou seja, é a correlação entre a variável canônica u_t e a variável X_k , com $k = 1, \dots, p$. Já $r_{v_t X_k}$ são as cargas canônicas do segundo conjunto, ou seja, é a correlação entre a variável canônica v_t e a variável X_k com $k = p + 1, \dots, p + q$ (JOHNSON; WICHERN, 2007).

5.2.3.2 Inferência sobre os coeficientes de correlação canônica

Para verificar se as correlações canônicas encontradas são significativas, utiliza-se um procedimento estatístico, baseado no teste da razão de verossimilhanças com uma correção proposta por Bartlett (1939 citado por JOHNSON; WICHERN, 2007), o qual rejeita a hipótese nula, definida como sendo a ausência de correlação linear canônica entre os dois conjuntos de variáveis, se o resultado da estatística teste for maior ou igual ao correspondente valor crítico de sua distribuição de probabilidade.

Algumas suposições são necessárias para generalizações dos resultados da análise de correlação canônica. Uma destas suposições é que o vetor de resposta $\mathbf{x}'_j = (\mathbf{x}'^{(1)}_j \ \mathbf{x}'^{(2)}_j)$, para $j = 1, \dots, n_i$, seja uma amostra aleatória de uma população com distribuição normal multivariada, ou seja, $\mathbf{N}_{p+q}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, na qual $\boldsymbol{\mu}$ é o vetor

de médias populacionais e $\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_{p \times p}^{(1,1)} & \boldsymbol{\Sigma}_{p \times q}^{(1,2)} \\ \boldsymbol{\Sigma}_{q \times p}^{(2,1)} & \boldsymbol{\Sigma}_{q \times q}^{(2,2)} \end{pmatrix}$, a qual foi verificada pelo gráfico *QQ-Plot* Multivariado. Também foram avaliadas, de forma prática pela observação dos

diagramas de dispersão e correlações lineares simples, as suposições de linearidade, multicolinearidade, homocedasticidade de variâncias, bem como a ausência de autocorrelação (HAIR JUNIOR, 2005).

Como a ideia básica da análise de correlação canônica é simplificar a representação da associação de dois conjuntos de variáveis por meio de combinações lineares das mesmas, é interessante verificar quantos pares de variáveis canônicas são realmente necessários. Para isso, primeiramente deve ser testado se todas as correlações canônicas são iguais a zero, o que equivale testar a seguinte hipótese:

$$H_0: \boldsymbol{\Sigma}^{(1,2)} = \mathbf{0} \text{ versus } H_1: \boldsymbol{\Sigma}^{(1,2)} \neq \mathbf{0}$$

onde $\Sigma^{(1,2)}$ é a matriz covariâncias populacionais entre as variáveis do vetor aleatório $\mathbf{X}'^{(1)}$ e do $\mathbf{X}'^{(2)}$ e $\mathbf{0}$ uma matriz nula de dimensão $p \times q$.

A hipótese nula (H_0) é rejeitada, para um nível de significância α , quando:

$$-\left(n-1-\frac{p+q+1}{2}\right) \ln \prod_{i=1}^t (1-\lambda_i) > \chi_{pq}^2(\alpha)$$

onde $\chi_{pq}^2(\alpha)$ é o (100α) percentil superior da distribuição qui-quadrado com pq graus de liberdade.

Se a hipótese nula for rejeitada, é necessário verificar quais correlações canônicas são significativas. Considerando que essas correlações estejam ordenadas da maior para a menor, é natural assumir inicialmente a hipótese de que a primeira correlação é diferente de zero e as $t-1$ correlações restantes não são significativas. Se essa hipótese é rejeitada, verifica-se que as duas primeiras são diferentes de zero e as $t-2$ não são significativas e, assim, sucessivamente até que a hipótese nula não seja rejeitada. Dessa forma, as seguintes hipóteses são consideradas:

$$H_0: \rho_1 \neq 0, \rho_2 \neq 0, \dots, \rho_m \neq 0, \rho_{m+1} = \dots = \rho_t = 0$$

$$H_1: \rho_l \neq 0, \text{ para algum } l \geq m+1$$

na qual ρ_t é o t -ésima correlação canônica populacional. Dessa forma, estão sendo testadas as m primeiras correlações canônicas e, conseqüentemente, identificando as variáveis canônicas que deverão ser consideradas para caracterização da associação dos dois conjuntos de variáveis.

Segundo Bartlett (1939 citado por JOHNSON; WICHERN, 2007), a hipótese nula (H_0) da m -ésima hipótese é rejeitada, considerando um nível de significância α , quando:

$$-\left(n-1-\frac{p+q+1}{2}\right) \ln \prod_{l=m+1}^t (1-\lambda_l) > \chi_{(p-m)(q-m)}^2(\alpha)$$

na qual $\chi_{(p-m)(q-m)}^2(\alpha)$ é o (100α) percentil superior da distribuição qui-quadrado com $(p-m)(q-m)$ graus de liberdade.

5.2.4 Análise Fatorial

Busca-se, com a análise fatorial, a redução da dimensão do vetor de respostas para que a estrutura de variabilidade possa ser melhor compreendida. Assim, o objetivo principal desse procedimento de análise multivariada é explicar a variabilidade original de todas as variáveis por meio de um número menor de variáveis não observadas, denominadas variáveis latentes ou fatores comuns, construídas a partir da combinação linear das variáveis originais, maximizando o poder de explicação simultâneo de todas as variáveis. Contudo, deve ser considerado que uma parte da variabilidade total dos dados originais é explicada pelos fatores e, conseqüentemente, a parte restante pelas variáveis latentes que não são incluídas no modelo, ou melhor, incorporadas ao erro aleatório (HAIR JUNIOR, 2005 e MINGOTI, 2005).

Basicamente, a construção dos fatores considera que as variáveis podem ser agrupadas de acordo com suas correlações, combinando-se dentro de um mesmo fator as variáveis mais correlacionadas entre si, mas que possuem correlações muito baixas com as variáveis que compõem outros fatores. Dessa forma, é possível determinar os coeficientes que relacionam as variáveis originais com os fatores comuns, os quais são chamados de cargas fatoriais (ou *loadings*) e são utilizados para representar os fatores subjacentes. Após a identificação dos fatores, é possível encontrar seus valores numéricos para cada unidade amostral, os quais são chamados de escores fatoriais (JOHNSON; WICHERN, 2007).

Essa técnica permite identificar as variáveis mais representativas para formação dos fatores, indicando como algumas variáveis climáticas podem se somar ou contrastar com as variáveis de qualidade da água na rede de distribuição.

Para este estudo o vetor de resposta é considerado genericamente como apresentado na Tabela 2, ou seja, $\mathbf{X}'_i = [X_{i1} \quad \dots \quad X_{ip} \quad X_{i(p+1)} \quad \dots \quad X_{i(p+q)}]$, sendo a repetição dada por cada mês de observação das $p + q = 18$ variáveis analisadas em um período de 5 anos, com $p = 8$ variáveis de qualidade da água observadas no i – ésimo grupo e $q = 10$ variáveis climáticas. O índice i será fixo para indicar se os dados da água foram avaliados na rede de distribuição ou no final de tratamento e, por isso, optou-se por sua omissão na representação a seguir.

Desta forma, quando se observa um grande número de variáveis correlacionadas entre si, as mesmas podem ser sumarizadas por um número menor de

variáveis latentes não correlacionadas e que explicam de forma reduzida as variáveis originais.

De acordo com Johnson e Wichern (2007), o modelo de análise fatorial, construído a partir da matriz de correlação, é dado por:

$$\mathbf{Z} = \mathbb{L} \mathbf{F} + \mathbf{e},$$

onde $\mathbf{Z}_{(p+q) \times 1}$ é um vetor representando o conjunto das $p + q$ variáveis padronizadas Z_k com média nula e variância unitária ($k = 1, \dots, p + q$), $\mathbf{F}_{m \times 1}$ é um vetor aleatório com os m fatores comuns não observáveis (sendo $1 \leq m \leq p$), $\mathbb{L}_{(p+q) \times m}$ é uma matriz de parâmetros que necessitam serem estimados por algum mecanismo e $\mathbf{e}_{(p+q) \times 1}$ é um vetor de erros aleatórios, erros esses representando a variação de Z_k que não é explicada pelo fatores comuns incluídos no modelo.

A matriz $\mathbb{L}_{(p+q) \times m}$ é formada por coeficientes l_{kg} , normalmente chamados de cargas fatoriais ou *loadings*, indicando o coeficiente da k – ésima variável padronizada no g – ésimo fator ($g = 1, \dots, m$). As cargas fatoriais indicam o grau de relacionamento linear entre a variável padronizada original e os fatores.

A partir desse modelo, pode-se dizer que a variabilidade da variável padronizada Z_k é decomposta em duas partes: a primeira que representa a quantidade de variância explicada pelos m fatores incluídos no modelo, chamada de comunalidade, e a segunda representando a quantidade da variabilidade de Z_k associada ao erro aleatório, ou seja, variabilidade não explicada pelos fatores e, comumente chamada de unicidade ou variância específica.

Neste estudo, utilizou-se o método dos componentes principais na estimação da matriz $\mathbb{L}$ a partir da matriz de correlação, por ser um método comumente usado como análise exploratória de dados e não exigir suposições sobre a distribuição de probabilidade do vetor $\mathbf{Z}$ (MINGOTI, 2005).

Nesse método, as raízes características da matriz de correlação amostral dos dados originais, designadas também por autovalores (λ), são determinadas e organizadas em ordem decrescente para, posteriormente, determinarem-se os autovetores normalizados correspondentes a cada autovalor. Esses componentes são utilizados na estimação dos parâmetros do modelo.

Os autovalores também auxiliam na interpretação dos resultados, uma vez que representam a parcela de variância total dos dados explicada por cada fator. Dessa forma, o primeiro fator está associado ao maior autovalor e retém a maior proporção de explicação da variabilidade dos dados originais. Segundo Johnson e Wichern (2007), quando o método dos componentes principais é utilizado na estimação dos parâmetros do modelo, a proporção explicada por um dado fator, F_g , é determinada por $\frac{\lambda_g}{(p+q)} 100\%$.

A rotação varimax foi utilizada para facilitar a interpretação dos fatores, não alterando o percentual acumulado de explicação da variância, as comunalidades e as variâncias específicas.

Para avaliar a viabilidade do modelo de análise fatorial, ou seja, verificar se é adequado analisar os dados com a técnica de análise fatorial, utilizou-se o teste de Bartlett (HAIR JR et al., 2005) e o critério de Kaiser-Meyer-Olkin (KMO), proposto inicialmente por Kaiser (KAISER, 1970 citado por MINGOTI, 2005).

O teste de Bartlett para a matriz de correlação verifica se a matriz de correlação populacional é próxima ou não da matriz identidade. A análise fatorial pressupõe que as variáveis são correlacionadas entre si e, portanto, quando é rejeitada a hipótese de que a matriz de correlação é igual à matriz identidade, a análise fatorial pode ser realizada. O critério KMO consiste em uma medida variando de 0 a 1 de adequação amostral para aplicação da análise fatorial, podendo ser obtida tanto de uma forma global quanto individualmente para cada variável. Essa medida compara os coeficientes de correlações com as correlações parciais, sendo considerada aceitável a análise fatorial para valores superiores a 0,5 (HAIR JUNIOR, 2005 e MINGOTI, 2005).

Com o objetivo de determinar o número de fatores a serem utilizados no modelo de análise fatorial, toma-se como base os autovalores da matriz de correlação amostral. Entre os critérios utilizados para esse fim, foram utilizados o critério baseado na observação do gráfico *Scree-plot* (CARTTELL, 1966 citado por MINGOTI, 2005) e o critério do número de autovalores maiores que a unidade (KAISER, 1958 citado por MINGOTI, 2005).

6 RESULTADOS E DISCUSSÕES

Os dados das oito variáveis ($p = 8$) relacionadas com as características físico-químicas e microbiológicas da água (na rede de distribuição e no final de tratamento) e das dez variáveis climáticas ($q = 10$) analisadas em 60 meses do período de janeiro de 2007 a dezembro de 2011 são apresentados de forma resumida nas figuras do Apêndice 1.

Observando as Figuras 24 e 25 (Apêndice 1), nota-se uma pequena alteração no comportamento das variáveis Flúor, Cloro e pH a partir do ano de 2010. Em relação aos anos anteriores, as variáveis Flúor e pH apresentaram uma queda, enquanto que os níveis de Cloro mostraram-se mais elevados. De acordo com Felipe (2014), algumas alterações nos procedimentos da Estação de tratamento podem ter causado essas oscilações, destacam-se: a mudança de equipamento para determinação de fluoreto, podendo levar a pequenas alterações na dosagem do flúor e a substituição do alcalinizante e do coagulante, gerando alterações na medida do pH. Em relação ao cloro, não se destacou qualquer causa plausível, no entanto, uma suposição levantada foi que houve a necessidade de aumento no nível de cloro devido à expansão da rede de distribuição.

Inicialmente, os dados das variáveis dos três conjuntos (qualidade da água na rede de distribuição, no final de tratamento e o conjunto de variáveis climáticas) foram resumidos por meio da média e do desvio padrão das cinco observações mensuradas em cada mês (Tabela 3, 4 e 5 respectivamente). Os gráficos apresentados nas Figuras 27, 28 e 29 do Apêndice 2 foram construídos para uma melhor visualização do comportamento das respectivas médias.

Considerando os resumos somente do conjunto de variáveis de qualidade da água (Tabela 3 e 4), percebe-se que não há grandes alterações nas médias destas variáveis tanto na rede de distribuição quanto no final de tratamento, indicando um padrão médio dos parâmetros de qualidade da água de Botucatu equivalente durante todos os meses. Apenas a variável “Bactérias heterotróficas”, denominada HET, apresenta algumas alterações entre os meses, no entanto, essas alterações estão dentro do limite aceitável (máximo de 500 UFC por *ml*) para o consumo humano, considerando os parâmetros estipulados pela legislação brasileira (Figuras 27 e 28 do Apêndice 2).

Tabela 3. Média e desvio padrão mensal das variáveis de qualidade da água e proporção de amostras com ausência de coliformes totais, considerando os dados da rede de distribuição.

Mês	Variáveis de qualidade da água *							
	<i>TURB</i>	<i>CORAP</i>	<i>F</i>	<i>CRL</i>	<i>pH</i>	<i>T</i>	<i>HET</i>	<i>CT</i>
Janeiro	0,688 (0,19)	6,746 (1,26)	0,684 (0,04)	1,439 (0,38)	7,807 (0,31)	25,085 (1,18)	1,926 (1,86)	0,988
Fevereiro	0,976 (0,43)	9,362 (3,96)	0,695 (0,02)	1,536 (0,46)	7,822 (0,34)	25,817 (1,57)	1,414 (1,63)	0,996
Março	0,644 (0,36)	6,537 (2,11)	0,688 (0,04)	1,626 (0,41)	7,830 (0,42)	26,851 (0,83)	1,250 (1,10)	1,000
Abril	0,686 (0,27)	6,310 (1,37)	0,687 (0,03)	1,705 (0,41)	7,801 (0,40)	25,360 (0,70)	6,296 (6,43)	0,990
Maio	0,666 (0,29)	6,322 (1,38)	0,686 (0,05)	1,826 (0,42)	7,939 (0,53)	23,390 (1,18)	3,522 (4,92)	0,998
Junho	0,755 (0,17)	7,109 (1,10)	0,693 (0,04)	1,866 (0,34)	7,672 (0,43)	21,085 (0,65)	1,248 (1,60)	1,000
Julho	0,895 (0,40)	7,404 (1,64)	0,683 (0,05)	1,855 (0,38)	7,618 (0,34)	20,199 (0,50)	0,798 (0,86)	0,998
Agosto	0,693 (0,54)	7,177 (3,46)	0,685 (0,02)	1,828 (0,39)	7,746 (0,60)	21,181 (1,02)	1,290 (1,47)	1,000
Setembro	0,554 (0,23)	6,265 (0,89)	0,707 (0,06)	1,752 (0,35)	7,604 (0,43)	22,951 (0,81)	0,950 (0,74)	1,000
Outubro	0,787 (0,38)	7,296 (1,96)	0,716 (0,07)	1,683 (0,39)	7,506 (0,50)	23,939 (1,35)	0,940 (0,77)	1,000
Novembro	0,463 (0,22)	5,915 (1,50)	0,694 (0,04)	1,655 (0,43)	7,674 (0,37)	25,618 (2,24)	5,972 (12,10)	0,990
Dezembro	0,749 (0,32)	7,666 (2,11)	0,715 (0,06)	1,628 (0,40)	7,566 (0,20)	25,033 (1,52)	5,522 (11,27)	0,960

* Variáveis: *TURB*(X_1)=Turbidez, *CORAP*(X_2)=Cor Aparente, *F*(X_3)=Flúor, *CRL*(X_4)=Cloro, *pH*(X_5), *T*(X_6)=Temperatura, *HET*(X_7)=Bactérias Heterotróficas, *CT*(X_8)= Ocorrência de Coliformes Totais

Nota-se também, na Tabela 3, que a variável HET apresentou um desvio padrão bastante elevado nos meses de novembro e dezembro, comparado com os

demais meses. Avaliando as médias mensais dessa variável no período de cinco anos, percebe-se que nos meses de novembro e dezembro de 2009 (Figura 24–Apêndice 1) a mesma apresentou uma média acima dos valores dos demais meses (valores médios superiores a 25 *UFC/ml*). Estas médias elevadas foram obtidas, pois uma amostra do mês de novembro de 2009 apresentou 435 *UFC/ml* e uma, em dezembro, 146 *UFC/ml*.

Tabela 4. Média e desvio padrão mensal das variáveis de qualidade da água e proporção de amostras com ausência de coliformes totais, considerando o final de tratamento.

Mês	Variáveis de qualidade da água *							
	<i>TURB</i>	<i>CORAP</i>	<i>F</i>	<i>CRL</i>	<i>pH</i>	<i>T</i>	<i>HET</i>	<i>CT</i>
Janeiro	0,604 (0,22)	6,346 (1,23)	0,684 (0,03)	1,629 (0,38)	7,798 (0,30)	23,086 (1,69)	3,552 (6,47)	1,000
Fevereiro	0,535 (0,15)	6,173 (1,26)	0,703 (0,01)	1,674 (0,40)	7,782 (0,33)	24,193 (1,97)	7,138 (8,38)	1,000
Março	0,523 (0,20)	5,874 (1,62)	0,691 (0,05)	1,770 (0,39)	7,940 (0,49)	24,363 (1,15)	3,333 (4,00)	0,988
Abril	1,241 (1,17)	8,309 (4,19)	0,690 (0,03)	1,908 (0,44)	7,959 (0,43)	23,190 (0,29)	3,973 (3,84)	1,000
Maio	0,497 (0,08)	5,487 (0,52)	0,699 (0,06)	1,978 (0,36)	7,906 (0,49)	20,749 (0,98)	8,172 (11,20)	1,000
Junho	0,520 (0,12)	5,384 (0,32)	0,681 (0,03)	2,042 (0,32)	7,724 (0,37)	18,863 (0,64)	1,806 (3,14)	1,000
Julho	0,525 (0,20)	5,453 (0,55)	0,678 (0,04)	1,912 (0,36)	7,844 (0,30)	18,380 (0,79)	117,752 (255,1)	1,000
Agosto	0,480 (0,23)	5,703 (0,96)	0,678 (0,03)	1,927 (0,44)	7,665 (0,49)	19,668 (1,19)	1,292 (1,64)	1,000
Setembro	0,599 (0,25)	5,574 (0,46)	0,703 (0,04)	1,884 (0,30)	7,631 (0,58)	20,963 (1,23)	2,112 (2,02)	1,000
Outubro	0,543 (0,14)	5,781 (0,33)	0,701 (0,07)	1,770 (0,30)	7,449 (0,50)	24,627 (4,31)	4,825 (9,96)	1,000
Novembro	0,401 (0,15)	5,482 (0,76)	0,698 (0,03)	1,746 (0,38)	7,705 (0,27)	22,586 (1,95)	1,567 (2,25)	1,000
Dezembro	0,662 (0,30)	6,642 (1,55)	0,690 (0,04)	1,676 (0,28)	7,651 (0,17)	23,577 (1,48)	0,686 (0,69)	0,974

* Variáveis: *TURB*(X_1)=*Turbidez*, *CORAP*(X_2)=*Cor Aparente*, *F*(X_3)=*Flúor*, *CRL*(X_4)=*Cloro*, *pH*(X_5), *T*(X_6)=*Temperatura*, *HET*(X_7)=*Bactérias Heterotróficas*, *CT*(X_8)=*Ocorrência de Coliformes Totais*

Destacam-se também os resultados da média e desvio padrão dos valores encontrados em julho para a variável HET (Bactérias Heterotróficas), considerando o final de tratamento (Tabela 4). Percebeu-se, avaliando os dados diários obtidos em julho de 2008 (Figura 25 – Apêndice 2), que uma amostra apresentou resultado, para essa variável, superior a 5701 *UFC/ml*, alterando consideravelmente sua média. Siqueira et al

(2010) descrevem que não há uma limitação para a quantidade de coliformes totais presentes na água potável, porém a legislação (Portaria 518/04) indica a necessidade de se tomar providências imediatas de caráter preventivo e corretivo, caso seja observada a presença de coliformes totais. Destacam ainda que, do ponto de vista microbiológico, a ausência de coliformes totais e termotolerantes em amostras de 100 ml de água para consumo é condição necessária para que seja considerada potável.

Já na Tabela 5, observam-se as médias mensais das variáveis climáticas com os respectivos desvios padrões, enquanto que na Figura 29 do Apêndice 2 é possível visualizar a evolução anual das médias mensais das temperaturas máxima, mínima e média do ar, precipitação e número de dias chuvosos, umidade relativa do ar, velocidade dos ventos, evaporação, insolação e radiação solar do município de Botucatu, SP.

Tabela 5. Média e desvio padrão mensal das variáveis climáticas

Mês	Variáveis climáticas*									
	<i>DIASC</i>	<i>TMIN</i>	<i>TMAX</i>	<i>TMED</i>	<i>PP</i>	<i>UR</i>	<i>VV</i>	<i>EV</i>	<i>ISOL</i>	<i>RAD</i>
Janeiro	20,60 (1,82)	19,03 (0,60)	27,41 (1,31)	22,69 (1,21)	416,89 (171,70)	73,57 (7,35)	82,98 (23,89)	4,98 (0,72)	4,48 (0,85)	328,96 (26,71)
Fevereiro	11,00 (3,54)	19,44 (0,47)	28,99 (0,75)	23,60 (1,11)	142,91 (42,00)	63,58 (15,99)	76,62 (14,89)	5,96 (0,80)	6,23 (0,83)	367,43 (30,65)
Março	9,80 (2,59)	18,84 (0,48)	28,14 (1,64)	23,51 (0,98)	103,60 (48,10)	62,50 (9,45)	84,20 (22,25)	5,79 (0,61)	6,40 (1,25)	371,00 (50,07)
Abril	6,80 (4,21)	16,99 (0,82)	26,62 (0,65)	21,77 (0,63)	85,62 (32,50)	62,76 (18,32)	78,32 (12,27)	4,67 (0,48)	7,65 (1,24)	341,03 (38,14)
Maio	5,20 (2,68)	13,89 (0,76)	23,78 (0,85)	18,84 (0,76)	55,30 (37,50)	64,17 (8,65)	79,94 (15,23)	3,80 (0,29)	7,59 (0,62)	315,34 (34,25)
Junho	5,00 (3,24)	12,51 (1,48)	23,68 (1,99)	17,75 (1,48)	45,92 (33,60)	61,51 (9,92)	77,34 (14,35)	3,61 (0,43)	7,12 (1,03)	297,52 (21,54)
Julho	4,40 (4,28)	13,14 (0,70)	24,14 (1,55)	18,60 (0,97)	75,74 (78,90)	59,50 (15,79)	82,16 (23,35)	4,13 (0,54)	6,69 (1,85)	285,70 (48,38)
Agosto	3,40 (3,21)	14,10 (1,15)	25,04 (1,95)	19,57 (0,93)	43,59 (49,70)	57,53 (11,29)	102,86 (17,63)	5,07 (0,45)	7,88 (0,76)	358,02 (8,95)
Setembro	5,80 (4,46)	15,72 (1,48)	26,15 (2,17)	20,63 (1,47)	50,88 (61,00)	60,96 (11,67)	113,36 (31,03)	5,84 (1,01)	6,92 (1,48)	380,05 (88,05)
Outubro	9,80 (3,56)	16,72 (2,41)	27,21 (2,07)	20,94 (3,25)	157,80 (120,20)	63,60 (9,34)	109,00 (12,55)	6,17 (0,71)	6,77 (0,60)	396,46 (38,92)
Novembro	12,20 (2,49)	17,18 (1,58)	27,88 (1,14)	21,28 (3,04)	155,44 (84,80)	63,29 (9,99)	103,28 (25,79)	6,26 (0,25)	6,93 (0,98)	451,03 (72,68)
Dezembro	14,60 (3,71)	18,11 (0,73)	28,52 (1,31)	23,14 (0,78)	207,15 (81,40)	65,43 (6,38)	91,58 (12,97)	6,90 (1,15)	6,95 (1,82)	455,87 (56,50)

* Variáveis: *DIASC*(X_9)=Dias chuvosos, *TMIN*(X_{10})=Temperatura mínima, *TMAX*(X_{11})= Temperatura máxima, *TMED*(X_{12})= Temperatura média, *PP*(X_{13})=Precipitação, *UR*(X_{14})=Umidade relativa, *VV*(X_{15})=Velocidade dos ventos, *EV*(X_{16})=Evaporação, *ISOL*(X_{17})=Insolação, *RAD*(X_{18})=Radiação solar

Observou-se que o período com maior precipitação pluviométrica se concentra entre os meses de outubro a março, destacando-se dezembro e janeiro como os meses mais chuvosos, chegando em média a 416,89 *mm* em janeiro. O mesmo acontece com o número de dias chuvosos no mês. Em contrapartida, entre os meses de maio a setembro as chuvas são muito escassas, observando em média cinco dias com chuva por mês, e o mês de agosto apresentando uma média de 43,59 *mm* de precipitação (Tabela 5). Esses resultados estão muito próximos da classificação climática dos municípios de Botucatu e São Manuel apresentada por Cunha e Martins (2009), na qual indicam os meses de abril, julho e agosto com baixa precipitação, e o período compreendido entre final de setembro ao início de março com índices hídricos mais elevados.

Ainda baseado nos resultados da Tabela 5, Botucatu apresenta temperaturas médias amenas durante todos os meses do ano, com mínimo de 17,75°C no mês de junho e máxima no mês de fevereiro por volta de 23,60°C. Observa-se ainda que o período menos chuvoso corresponde aos meses mais frios. Resultados semelhantes foram apresentados por Cunha e Martins (2009), indicando uma temperatura média anual de 20,3°C, com temperatura mais baixa no mês de julho (17,1°C em média) e fevereiro mostrando-se como o mês mais quente (23,1°C).

Após a análise mensal dos dados, as medidas de posição das variáveis de cada conjunto de dados foram determinadas, considerando os 60 meses do período de estudo, obtendo-se, dessa forma, os vetores de médias, quartis e valores máximos e mínimos.

Nas Tabelas 6 e 7 é possível observar o resumo dos dados das variáveis físico-químicas e microbiológicas da água mensuradas na rede de distribuição e no final de tratamento respectivamente. Comparando os valores das médias e medianas, verifica-se que estão bastante próximos, exceto na variável HET, indicando que os dados são aproximadamente simétricos.

Em relação à variável HET, como destacado anteriormente, observou-se em julho de 2008 (Final de tratamento), novembro e dezembro de 2009 (Rede de distribuição) valores elevados em algumas amostras de água, fatos que refletiram em valores máximo e médio bastante elevados no final de tratamento, ou seja, 574 e 13,017 *UFC/ml*, respectivamente (Tabela 7). Este tipo de situação, quando ocorre pontualmente em amostras isoladas, a Portaria 518/04 estabelece que é necessário realizar a limpeza dos reservatórios, eliminando as possíveis causas desta ocorrência (BRASIL, 2004).

Tabela 6. Medidas de posição das variáveis de qualidade da água avaliadas na rede de distribuição e no período de Janeiro de 2007 a dezembro 2011

Variáveis	Média	Desvio Padrão	Q1	Mediana	Q3	Mínimo	Máximo	Limites da portaria
TURB (X_1)	0,713	0,327	0,450	0,681	0,928	0,234	1,623	Max 5
CORAP (X_2)	7,009	2,087	5,451	6,361	7,969	5,000	15,211	Max 15
F (X_3)	0,695	0,148	0,663	0,700	0,718	0,617	0,800	Max 1,5
CRL (X_4)	1,700	0,382	1,412	1,559	2,082	1,076	2,410	0,5 a 2,0
pH (X_5)	7,716	0,397	7,408	7,713	7,946	6,915	8,513	6,0 a 9,5
T (X_6)	23,86	2,341	21,87	23,81	25,74	19,56	28,24	-
HET (X_7)	2,594	5,309	0,220	0,920	2,280	0,000	27,59	Max 500
CT (X_8)	0,993	0,032	1,000	1,000	1,000	0,830	1,000	Min 0,95

Tabela 7. Medidas de posição das variáveis de qualidade da água observadas no final do tratamento

Variáveis	Média	Desvio Padrão	Q1	Mediana	Q3	Mínimo	Máximo	Limites da portaria
TURB (X_1)	0,594	0,405	0,419	0,499	0,643	0,229	3,150	Max 5
CORAP (X_2)	6,017	1,593	5,000	5,606	6,178	5,000	13,133	Max 15
F(X_3)	0,691	0,032	0,670	0,693	0,714	0,613	0,771	Max 1,5
CRL(X_4)	1,826	0,355	1,533	1,708	2,119	1,309	2,603	0,5 a 2,0
pH (X_5)	7,754	0,397	7,453	7,772	8,083	6,733	8,600	6,0 a 9,5
T (X_6)	22,020	2,666	19,909	21,971	24,004	17,619	32,107	-
HET (X_7)	13,017	73,870	0,000	1,250	4,655	0,000	574,000	Max 500
CT (X_8)	0,997	0,032	1,000	1,000	1,000	0,867	1,000	Min 0,95

Comparando os resultados apresentados nas Tabelas 6 e 7 com os padrões de qualidade estabelecidos pela portaria (apresentados na última coluna), percebe-se que, em média, todas as variáveis estão de acordo com a legislação, tanto nas amostras de água coletadas na rede de distribuição, quanto as coletadas no final de tratamento.

Já a Tabela 8 expressa os vetores de médias, desvio padrão, quartis, valores máximos e mínimos dos dados das variáveis climáticas para o município de Botucatu. Observa-se que no período deste estudo, Botucatu apresentou, em média, nove dias por mês com ocorrência de chuvas, resultando em uma média mensal pluviométrica de 128,403 mm. Os valores médios e medianos de praticamente todas as variáveis situam-se

muito próximos, com destaque para a variável precipitação (PP) com um mínimo de zero e máximo de 712,250 mm de precipitação por mês.

Tabela 8. Medidas de posição das variáveis climáticas no período de janeiro de 2007 a dezembro de 2011

Variáveis	Média	Desvio Padrão	Q1	Mediana	Q3	Mínimo	Máximo
DIASC (X_9)	9,050	5,771	5,000	9,000	13,000	0,000	23,000
TMIN (X_{10})	16,306	2,563	13,850	16,730	18,578	10,900	20,110
TMAX (X_{11})	26,464	2,292	24,835	26,695	28,153	20,850	30,540
TMED (X_{12})	21,027	2,422	19,175	21,565	22,850	15,460	24,800
PP (X_{13})	128,403	124,887	47,275	102,750	155,850	0,000	712,250
UR (X_{14})	63,201	11,226	55,938	65,145	72,063	34,740	86,030
VV (X_{15})	90,137	22,018	72,400	89,050	105,075	45,100	141,900
EV (X_{16})	5,263	1,185	4,263	5,235	6,268	3,090	8,110
ISOL (X_{17})	6,800	1,363	5,830	7,055	7,780	3,830	9,200
RAD (X_{18})	362,368	67,645	316,550	358,260	400,138	223,110	545,710

6.1 Teste T^2 de Hotelling na comparação da qualidade da água na rede de distribuição e no final de tratamento

Considerando somente os dados de qualidade da água, o teste T^2 de Hotelling é utilizado para comparação simultânea dessas variáveis, observadas na rede de distribuição (Grupo 1 – G1) e no final de tratamento (Grupo 2 – G2), a fim de verificar se existem diferenças significativas entre os vetores de médias desses dois grupos.

Os vetores de médias das variáveis de qualidade da água observadas na rede de distribuição (G1) e no final de tratamento (G2) são destacados na Tabela 9.

Tabela 9. Vetores de médias das variáveis de qualidade da água.

Grupo	Estatística	TURB (X_1)	CORAP (X_2)	F (X_3)	CRL (X_4)	pH (X_5)	T (X_6)	HET (X_7)	CT (X_8)
G1	$\bar{x}'_1$	0,713	7,009	0,695	1,700	7,715	23,876	2,594	0,993
G2	$\bar{x}'_2$	0,594	6,017	0,691	1,826	7,754	22,020	13,017	0,997

Já as matrizes de variâncias e covariâncias amostrais do Grupo 1 (representada por S_1) e do Grupo 2 (S_2) são dadas por:

$$S_1 = \begin{bmatrix} 0,107 & & & & & & & \\ 0,619 & 4,356 & & & & & & \\ 0,003 & 0,017 & 0,022 & & & & & \\ -0,037 & -0,204 & -0,011 & 0,146 & & & & \\ 0,057 & 0,260 & 0,007 & -0,099 & 0,158 & & & \\ 0,003 & 0,266 & 0,021 & -0,353 & 0,225 & 5,481 & & \\ 0,055 & 0,231 & 0,019 & -0,080 & 0,336 & 4,223 & 28,182 & \\ 0,001 & 0,002 & 0,001 & 0,001 & 0,000 & -0,005 & -0,009 & 0,001 \end{bmatrix} \quad e$$

$$S_2 = \begin{bmatrix} 0,164 & & & & & & & \\ 0,539 & 2,538 & & & & & & \\ 0,004 & 0,014 & 0,001 & & & & & \\ -0,037 & -0,137 & -0,009 & 0,126 & & & & \\ 0,067 & 0,130 & 0,007 & -0,075 & 0,158 & & & \\ 0,132 & 0,836 & 0,013 & -0,342 & 0,072 & 7,105 & & \\ 1,374 & -7,848 & -0,014 & -2,077 & 3,237 & -39,206 & 5456,79 & \\ 0,001 & 0,003 & 0,000 & 0,000 & 0,000 & -0,001 & 0,042 & 0,001 \end{bmatrix}.$$

A princípio, por inspeção visual do gráfico de probabilidade *QQ Plot* (*Chi-Square plot*) multivariado, apresentado na Figura 30 do Apêndice 3, verificou-se no contexto global a tendência das distâncias alinharem-se ao redor de uma linha reta, salvo quatro maiores distâncias que se mostraram com leve fuga da tendência das demais na rede de distribuição e seis no final de tratamento. Esta situação, na prática, sugere uma aderência confortável dos dados à distribuição normal multivariada de probabilidade. Além disso, como observado anteriormente, os dados das características da água na rede e no final de tratamento são aproximadamente simétricos, condição importante para a robustez do teste em relação à não normalidade, de acordo com Mardia (1974).

Para verificar a homogeneidade das matrizes de variâncias-covariâncias o teste multivariado proposto por Box foi realizado, indicando que a hipótese de igualdade das matrizes em questão deve ser rejeitada (estatística de teste igual a 269,78 e $p < 0,05$).

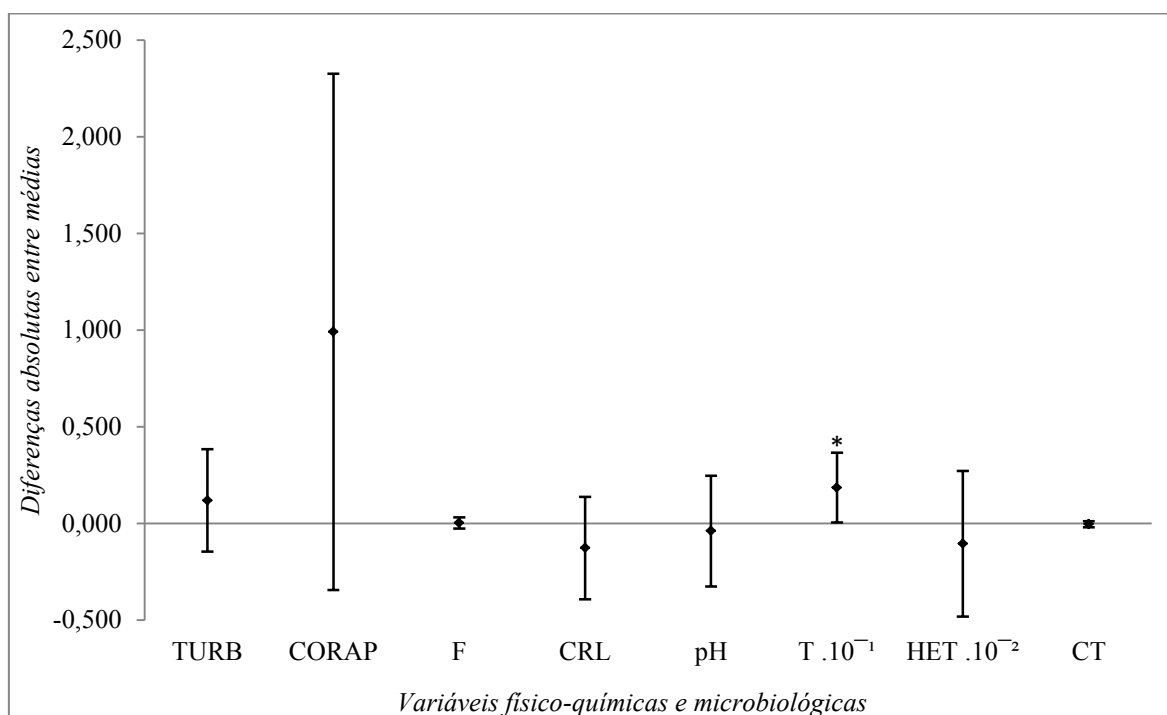
Dessa forma, foi utilizado o teste T^2 de Hotelling para matrizes de variâncias-covariâncias desiguais, verificando que existe diferença significativa entre o

vetor de médias das variáveis de qualidade da água na rede de distribuição e o vetor considerando o final de tratamento ($T^2 = 28,35$ e $p < 0,05$).

Na Tabela 10, observam-se os limites inferiores e superiores dos intervalos de confiança simultâneos de Hotteling para a diferença entre as médias encontradas a partir dos dados da rede de distribuição e as médias obtidas no final de tratamento. Já, na Figura 11, é possível visualizar graficamente esses intervalos.

Tabela 10. Limite inferior (LI) e superior (LS) dos intervalos de confiança simultâneos de Hotteling para a diferença entre as médias dos dois grupos

Estatística	TURB (X_1)	CORAP (X_2)	F (X_3)	CRL (X_4)	pH (X_5)	T (X_6)	HET (X_7)	CT (X_8)
$(\bar{x}_1 - \bar{x}_2)'$	0,119	0,992	0,003	-0,127	-0,039	1,855	-10,424	-0,004
LI	-0,146	-0,343	-0,026	-0,391	-0,325	0,052	-48,075	-0,019
LS	0,384	2,327	0,032	0,138	0,247	3,659	27,228	0,012



Conforme se observa na Tabela 10 e na Figura 11, para a variável temperatura da água no momento da coleta, os resultados mostraram que existe diferença significativa entre os dois grupos estudados ($p < 0,05$), pois o intervalo de confiança simultâneo não envolve o zero, e, como essa diferença é positiva, é possível dizer que a

temperatura da água na rede é superior à temperatura observada no final de tratamento. Já para as demais variáveis não foi possível identificar diferença significativa entre a análise realizada na rede de distribuição e no final de tratamento ($p > 0,05$).

Em um estudo realizado por Campos, Farache Filho e Faria (2002), não foram observadas diferenças significativas em diferentes pontos da rede de distribuição de água no município de Araraquara (proximidades do reservatório de distribuição, em pontos intermediários e nas pontas da rede de distribuição). Em outro estudo, Campos, Farache Filho e Faria (2003) verificaram diferenças significativas na qualidade da água observada nos reservatórios dos domicílios e na rede de distribuição, indicando condições desfavoráveis para água dos reservatórios domiciliares nas variáveis pH, Turbidez, concentração de cloro e ferro, bem como um percentual maior de amostras com presença de coliformes totais.

6.2 Correlação Canônica

São apresentados, nesta seção, os resultados do estudo da associação entre variáveis climáticas e características físico-químicas e microbiológicas da água potável distribuída no município de Botucatu, considerando amostras de água na rede de distribuição e no final de tratamento.

6.2.1 Associação entre o conjunto de variáveis de qualidade da água observado na rede de distribuição e o conjunto de variáveis climáticas

As medidas resumos dos dados das variáveis físico-químicas e microbiológicas da água e das características climáticas, com os respectivos vetores de médias, quartis e valores máximos e mínimos, foram apresentadas respectivamente nas Tabelas 6 e 8.

Para as variáveis do primeiro conjunto (características de qualidade da água), suas respectivas variâncias e covariâncias (matriz $\mathbb{S}^{(1,1)}$), bem como suas correlações (por meio do coeficiente de correlação linear de Pearson - matriz $\mathbb{R}^{(1,1)}$) podem ser visualizadas na Tabela 11. Já na Tabela 12 as mesmas informações são apresentadas para o segundo conjunto de variáveis (características climáticas).

Tabela 11. Variâncias, covariâncias ($S^{(1,1)}$) e correlações ($R^{(1,1)}$) das variáveis de qualidade da água avaliadas na rede de distribuição*

	TURB	CORAP	F	CRL	pH	T	HET	CT
TURB	0,1074	0,91	0,21	-0,30	0,44	0,00	0,03	0,07
CORAP	0,6195	4,3565	0,19	-0,26	0,31	0,05	0,02	0,05
F	0,0030	0,0167	0,0019	-0,67	0,38	0,21	0,08	0,11
CRL	-0,0373	-0,2040	-0,0111	0,1459	-0,65	-0,40	-0,04	0,06
pH	0,0575	0,2596	0,0066	-0,0988	0,1583	0,24	0,16	0,06
T	0,0027	0,2658	0,0207	-0,3532	0,2247	5,4806	0,34	-0,09
HET	0,0546	0,2307	0,0194	-0,0795	0,3356	4,2230	28,1880	-0,09
CT	0,0005	0,0022	0,0001	0,0005	0,0005	-0,0049	-0,0110	0,0005

* As variâncias estão na diagonal, as covariâncias e as correlações são apresentadas abaixo e acima da diagonal, respectivamente.

Considerando o conjunto de variáveis de qualidade da água (Tabela 11), destacam-se as três correlações significativas absolutas com grau mais elevado, ou seja, a correlação positiva entre as variáveis turbidez e cor aparente (0,91) e as correlações negativas entre a variável cloro e as variáveis flúor (-0,67) e pH (-0,65), todas com valor $p < 0,0001$.

Tabela 12. Variâncias, covariâncias ($S^{(2,2)}$) e correlações ($R^{(2,2)}$) das variáveis climáticas *

	DIASC	TMIN	TMAX	TMED	PP	UR	VV	EV	INSOL	RAD
DIASC	33,303	0,63	0,36	0,45	0,81	0,37	-0,19	0,21	-0,67	0,09
TMIN	9,326	6,567	0,72	0,87	0,56	0,25	0,02	0,60	-0,39	0,38
TMAX	4,797	4,210	5,253	0,77	0,29	-0,10	0,08	0,76	-0,06	0,51
TMED	6,252	5,376	4,279	5,868	0,41	0,18	0,09	0,62	-0,20	0,41
PP	586,100	179,327	83,864	123,387	15596,844	0,26	-0,21	0,20	-0,55	0,02
UR	24,121	7,146	-2,676	4,791	366,065	126,034	-0,11	-0,09	-0,30	0,05
VV	-23,580	0,925	3,810	5,012	-590,208	-26,330	484,786	0,41	0,14	0,33
EV	1,449	1,812	2,067	1,787	30,183	-1,256	10,662	1,404	0,14	0,72
INSOL	-5,250	-1,370	-0,172	-0,666	-93,973	-4,567	4,227	0,219	1,857	0,29
RAD	33,741	65,037	79,303	66,856	163,961	40,285	488,663	57,385	27,124	4575,860

* As variâncias estão na diagonal, as covariâncias e as correlações são apresentadas abaixo e acima da diagonal, respectivamente.

Com base nas correlações lineares entre as variáveis do conjunto climático (Tabela 12), percebe-se que o número de dias chuvosos possui correlações significativas positivas, em grau mais elevado, com a temperatura mínima (0,63) e precipitação (0,81), já correlação negativa com a insolação (-0,67). Nota-se também que temperatura máxima apresentou maior correlação positiva significativa com evaporação (0,76), e esta possui associação positiva com a radiação solar (0,72).

Respectivamente, as covariâncias e correlações amostrais entre as variáveis de qualidade da água (primeiro conjunto) e climáticas (segundo conjunto) são dadas na Tabela 13 e 14.

Tabela 13. Matriz de variâncias e covariâncias ($S^{(1,2)}$) entre as variáveis do primeiro e segundo conjunto, considerando a rede de distribuição

	DIASC	TMIN	TMAX	TMED	PP	UR	VV	EV	INSOL	RAD
TURB	0,212	0,014	-0,011	0,069	5,455	0,187	-1,771	-0,071	-0,081	-5,096
CORAP	1,950	0,459	0,242	0,749	30,720	1,754	-5,317	-0,190	-0,580	-18,293
F	0,008	0,023	0,011	0,028	-0,212	0,161	0,264	0,005	-0,011	0,557
CRL	-0,838	-0,405	-0,115	-0,318	-9,026	-2,537	-0,797	-0,007	0,232	-3,065
pH	0,224	0,143	-0,124	0,092	2,993	1,678	-1,533	-0,111	-0,090	-3,488
T	6,320	4,687	3,129	3,981	94,001	6,058	-4,021	1,083	-0,814	71,216
HET	4,524	2,319	1,220	1,911	144,899	-11,883	-33,296	-0,304	-0,657	62,804
CT	-0,052	-0,013	-0,008	-0,012	-0,699	0,001	0,012	-0,006	0,004	-0,111

Tabela 14. Matriz de correlação ($R^{(1,2)}$) entre as variáveis do primeiro e segundo conjunto, considerando a rede de distribuição

	DIASC	TMIN	TMAX	TMED	PP	UR	VV	EV	INSOL	RAD
TURB	0,11	0,02	-0,01	0,09	0,13	0,05	-0,25	-0,18	-0,18	-0,23
CORAP	0,16	0,09	0,05	0,15	0,12	0,07	-0,12	-0,08	-0,20	-0,13
F	0,03	0,21	0,11	0,27	-0,04	0,33	0,28	0,10	-0,19	0,19
CRL	-0,38	-0,41	-0,13	-0,34	-0,19	-0,59	-0,09	-0,01	0,45	-0,12
pH	0,10	0,14	-0,14	0,10	0,06	0,38	-0,18	-0,24	-0,17	-0,13
T	0,47	0,78	0,58	0,70	0,32	0,23	-0,08	0,39	-0,26	0,45
HET	0,15	0,17	0,10	0,15	0,22	-0,20	-0,28	-0,05	-0,09	0,17
CT	-0,39	-0,22	-0,14	-0,21	-0,24	0,00	0,02	-0,21	0,14	-0,07

A maior correlação (com $p < 0,0001$) entre qualquer variável do conjunto de qualidade da água e qualquer variável climática equivale a 0,78 (temperatura da água e temperatura ambiente mínima). Destacam-se também as correlações positivas entre a temperatura da água e a temperatura ambiente máxima (0,58) e média (0,70), além da correlação negativa (-0,59) entre a variável cloro e umidade relativa do ar, todas com $p < 0,0001$ (Tabela 14).

A partir da matriz de correlação do vetor $\mathbf{X}' = (\mathbf{X}'^{(1)} \quad \mathbf{X}'^{(2)})$, foram obtidas as variáveis canônicas: u_t definidas pela combinação linear das variáveis de qualidade da água e v_t pela combinação linear das variáveis climáticas. Os coeficientes de correlação canônica, com seus respectivos testes de significância, são apresentados na Tabela 15.

Tabela 15. Autovalor, correlação canônica e resultado do teste estatístico – Rede de distribuição

Pares de variáveis canônicas	Autovalor	Correlação Canônica	Estatística de teste	gl	Valor p
(u_1, v_1)	0,847	0,920	208,66	80	< 0,001
(u_2, v_2)	0,656	0,810	118,22	63	< 0,001
(u_3, v_3)	0,364	0,604	65,64	48	0,05
(u_4, v_4)	0,274	0,524	43,11	35	0,16
(u_5, v_5)	0,227	0,477	26,79	24	0,31
(u_6, v_6)	0,130	0,360	13,24	15	0,58
(u_7, v_7)	0,091	0,302	5,76	8	0,67
(u_8, v_8)	0,008	0,092	0,48	3	0,92

Das oito correlações canônicas encontradas, apenas as duas primeiras se mostraram significativas ($p < 0,001$), apresentando valores iguais a 0,920 e 0,810, respectivamente. Com estes resultados, é possível dizer que existem duas variáveis canônicas que expressam características distintas a respeito da qualidade da água (u_1 e u_2) e duas expressando determinadas situações climáticas (v_1 e v_2), bem como a existência de associação entre os pares de variáveis canônicas, (u_1, v_1) e (u_2, v_2) , ou seja, entre uma situação climática e uma característica de qualidade da água distribuída à população. Esta situação aponta para uma possível influência das condições climáticas nas características de qualidade da água distribuída aos domicílios de Botucatu.

A suposição de normalidade multivariada foi avaliada visualmente por meio do gráfico de probabilidade *QQ Plot (Chi-Square plot)* multivariado apresentado na Figura 31 do Apêndice 3. Com exceção das duas maiores distâncias ordenadas, percebe-se uma tendência das demais a se alinharem ao redor da bissetriz, sugerindo, na prática, aderência dos dados à distribuição normal multivariada de probabilidade. Observou-se também, por uma análise exploratória dos dados, que estão adequadas as suposições de linearidade, multicolinearidade, homocedasticidade de variâncias. O mesmo não se pode afirmar a respeito da ausência de autocorrelação nos resíduos, devido ao aspecto temporal dos dados originais, no entanto, considerando que este estudo apresenta mais um caráter exploratório do que confirmatório, a suposição em questão não foi considerada relevante.

A Figura 12 apresenta, para o primeiro par de variáveis canônicas (u_1 e v_1), os valores padronizados que estas variáveis assumem em cada mês. Nota-se uma

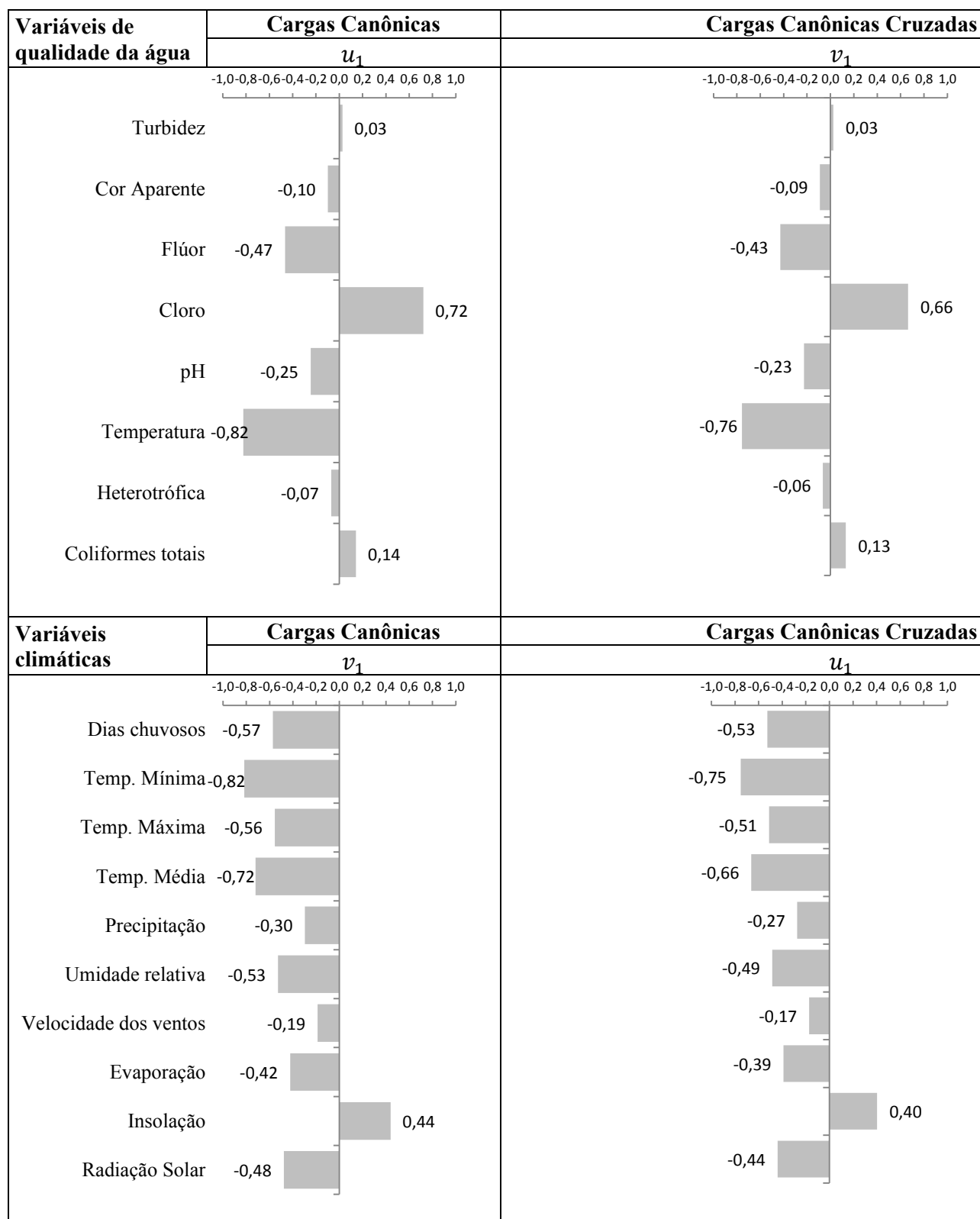


Figura 13. Gráfico das correlações de cada variável original com as variáveis canônicas (u_1 e v_1), considerando a rede de distribuição

Observando as cargas canônicas de u_1 , as variáveis originais de qualidade da água que apresentam cargas mais altas são: temperatura da água (-0,82), cloro (0,72) e flúor (-0,47). Desta forma, a variável canônica u_1 pode ser interpretada pelo antagonismo entre a concentração de cloro com a concentração de flúor e a temperatura da água. Esse padrão sugere que o componente u_1 pode ser representado por “substâncias químicas do tratamento da água”, assumindo valores mais elevados quando a água apresenta concentrações maiores de cloro, menores concentrações de flúor e temperaturas mais baixas.

Já ao analisar as cargas canônicas de v_1 , as variáveis climáticas originais que apresentam maiores correlações com o componente climático v_1 são temperatura mínima (-0,82), média (-0,72), máxima (-0,56), número de dias chuvosos (-0,57), umidade relativa do ar (-0,53), radiação solar (-0,48) e insolação (0,44). Esse padrão sugere v_1 como um componente relacionado com “características de umidade e temperatura”, assumindo valores maiores nos meses em que as temperaturas são mais baixas, com menos dias chuvosos, menores umidades relativas do ar, radiações solares e maiores insolações.

Portanto, com base nas análises das cargas canônicas e conforme resultado apresentado na Figura 12, observa-se que os meses de junho e julho apresentam valores altos para as duas variáveis canônicas, indicando um período mais seco e temperaturas mais baixas associado a um período em que a água distribuída a população apresenta temperaturas mais baixas, menos flúor e mais cloro. Contudo, os meses de novembro a março se mostraram como um período mais úmido, com mais dias chuvosos e temperaturas mais elevadas associados com meses onde a água apresentou-se com menos cloro, temperaturas mais elevadas e maior concentração de flúor.

Além disso, observando as cargas canônicas cruzadas entre as variáveis originais de qualidade da água com o componente climático, destacam-se com maiores associações: temperatura da água (-0,76), cloro (0,66) e flúor (-0,43), indicando que o componente climático alto associa-se com concentrações altas de cloro, menores temperaturas e baixas concentrações de flúor.

Observa-se também que as cargas canônicas cruzadas mais elevadas entre as variáveis climáticas e o componente canônico da água são: temperaturas mínima (-0,75) e média (-0,66), frequência de dias chuvosos (-0,53), temperatura máxima (-0,51), umidade relativa (-0,49) e radiação solar global (-0,44). Estes resultados indicam

que o componente canônico, representando uma característica de qualidade da água, assume valores altos associados com valores baixos da temperatura do ar (mínima, média e máxima), número de dias chuvosos, umidade relativa do ar e radiação solar.

Utilizando também a análise de correlação canônica, Freire e Castro (2014) verificam que existe associação entre um componente representado a qualidade da água e outro representando o uso e ocupação do solo em um estudo realizado em algumas sub-bacias da bacia hidrográfica do Rio Itapemirim-ES. Destacam também que a qualidade da água pode ser resumida em uma variável canônica relacionada com a alta concentração de nitratos e baixa concentração de oxigênio dissolvido. Para este estudo, as variáveis de qualidade da água utilizadas foram: coliformes termotolerantes (*UFC/100 ml*), oxigênio dissolvido (*mg/l*), fósforo (*mg/l*), nitratos (*mg/l*) e amônia (*mg/l*), enquanto que o percentual de áreas edificadas, de áreas agrícolas, de pastagens e de solo expostos formam o conjunto de variáveis de uso e ocupação do solo.

Considerando agora o segundo par de variáveis canônicas (u_2 e v_2), a Figura 14 destaca seus valores padronizados, enquanto que as correlações entre as variáveis originais com esse segundo par são apresentadas na Figura 15.

No segundo par canônico são destacados algumas particularidades de respostas em determinados meses e anos (Figura 14). Verifica-se que os meses de novembro (2009), dezembro (2009) e abril (2011) apresentam altos valores nas duas variáveis canônicas (u_2 e v_2), em contrapartida, valores baixos para os dois conjuntos de variáveis destacam-se nos meses entre julho e agosto de 2007. Nos meses de janeiro e fevereiro de 2011, verificaram-se valores baixos para o componente de qualidade da água associado com valores altos do componente climático.

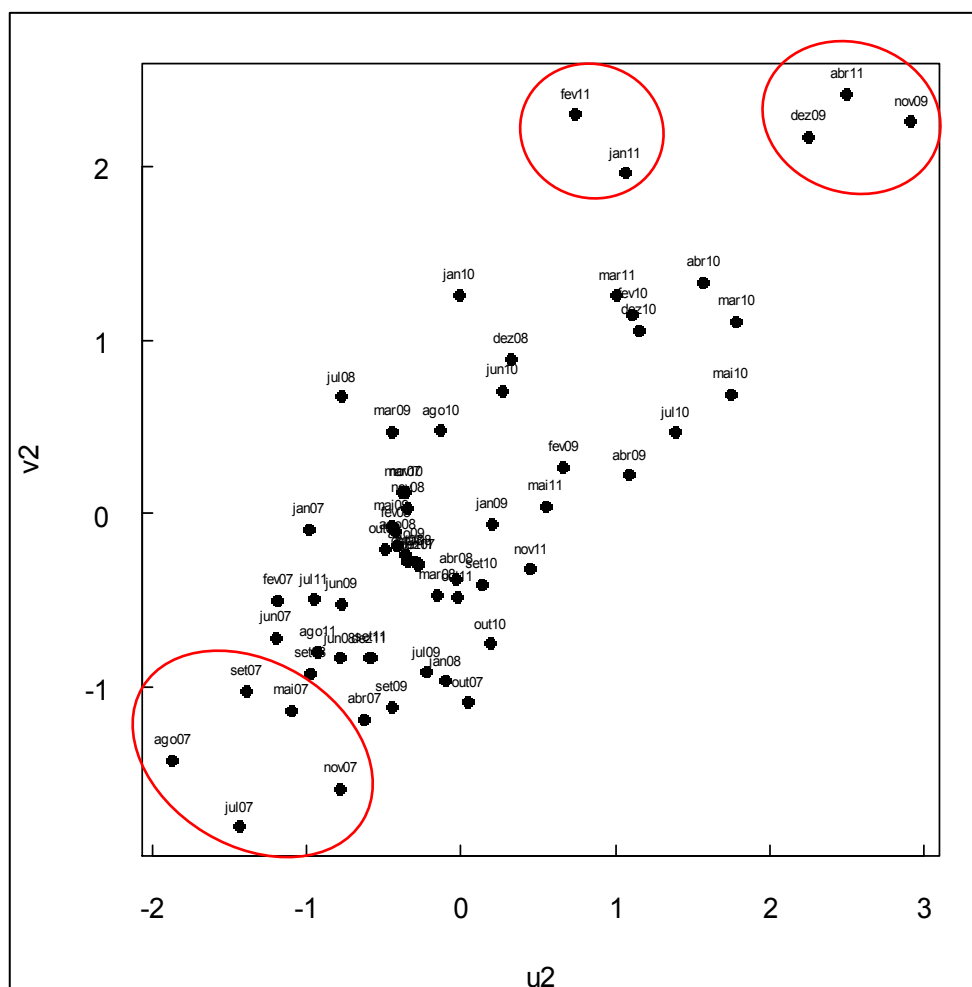


Figura 14. Gráfico de dispersão do segundo par de variáveis canônicas padronizadas, considerando a rede de distribuição

Analisando as cargas canônicas apresentadas na Figura 15, as variáveis originais de qualidade da água que manifestam maior associação com componente canônico u_2 (representando uma característica de qualidade da água) são: quantidade de bactérias heterotróficas (0,72), temperatura da água (0,48), cloro (0,27) e flúor (-0,34). Percebe-se, portanto, que este componente exibe valores altos em situações em que se observam concentrações menores de flúor, conjuntamente com temperatura da água elevada, maiores concentrações de cloro e de quantidade de bactérias heterotróficas.

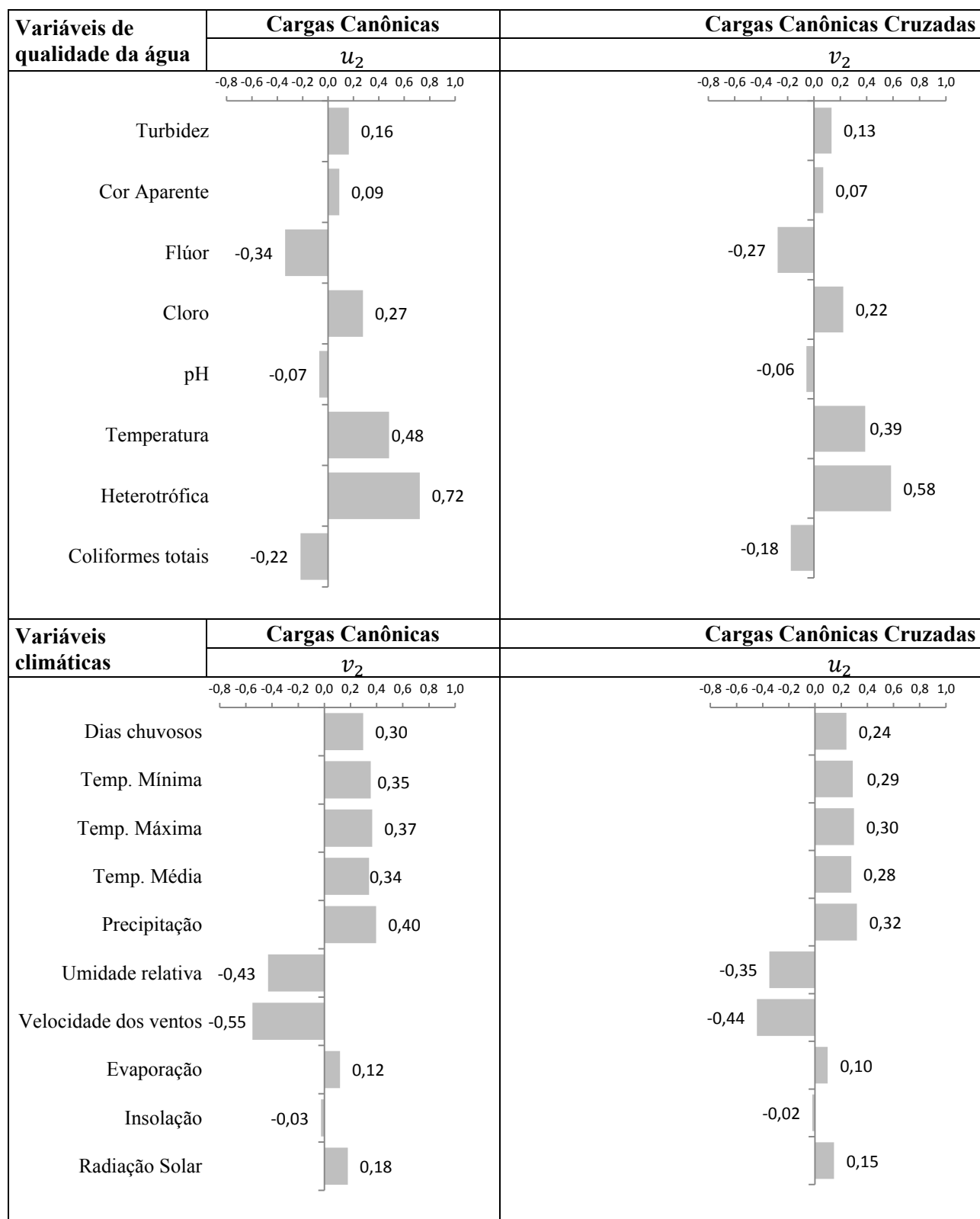


Figura 15. Gráfico das correlações de cada variável original com as variáveis canônicas (u_2 e v_2), considerando a rede de distribuição

Observando as cargas canônicas do conjunto de variáveis climáticas, as variáveis que apresentam maiores correlações com v_2 (componente

representando uma situação climática), são velocidade dos ventos (-0,55), umidade relativa (-0,43), precipitação (0,40), temperaturas (mínima =0,35, máxima=0,37, média=0,34) e quantidade de dias chuvosos (0,30). Assim, este componente assume valores mais elevados quando a velocidade dos ventos e a umidade relativa forem baixas, concomitantemente, a altos índices de precipitação e temperaturas mais elevadas.

Os mesmos destaques podem ser observados quando se cruzam os conjuntos de variáveis, isto é, resultados similares aos da análise das correlações canônicas (não cruzadas).

Portanto, de uma forma geral, os resultados discutidos anteriormente indicam que é possível resumir as informações das variáveis relacionadas com a qualidade da água na rede de distribuição do município e das variáveis climáticas em dois pares de componentes ou variáveis canônicas com correlações significativas, (u_1, v_1) e (u_2, v_2) , sendo uma expressando determinada característica de qualidade da água e outra uma característica climática.

Para uma melhor compreensão da correlação positiva entre u_1 e v_1 , o Quadro 4 apresenta o comportamento das variáveis climáticas, por meio de setas nas variáveis mais importantes para explicar v_1 , quando o componente de qualidade da água (u_1) apresenta valores mais altos ($\uparrow$) ou mais baixos ($\downarrow$).

Quadro 4. Resumo da associação entre o componente de qualidade da água na rede de distribuição (u_1) e a caracterização climática (v_1)

Variáveis climáticas	Componente de qualidade da água - u_1	
	Assumindo valores altos $\uparrow$ (Cloro $\uparrow$, Temperatura $\downarrow$ e Flúor $\downarrow$)	Assumindo valores baixos $\downarrow$ (Cloro $\downarrow$, Temperatura $\uparrow$ e Flúor $\uparrow$)
Dias chuvosos	$\downarrow$	$\uparrow$
Temp. mínima	$\downarrow$	$\uparrow$
Temp. máxima	$\downarrow$	$\uparrow$
Temp. média	$\downarrow$	$\uparrow$
Precipitação		
Umidade relativa	$\downarrow$	$\uparrow$
Veloc. dos ventos		
Evaporação		
Insolação	$\uparrow$	$\downarrow$
Radiação solar	$\downarrow$	$\uparrow$

Percebe-se que o componente de qualidade da água, expressa por u_1 , apresenta-se mais elevado (água com maior concentração de cloro, temperaturas mais baixas e maior quantidade de flúor) nos meses com temperaturas ambientes mais baixas, com baixa umidade relativa e uma quantidade menor de dias chuvosos. Situação inversa ocorre quando o componente de qualidade da água assume um valor mais baixo (Quadro 4).

Da mesma forma, a variável canônica u_2 , representando outra característica de qualidade da água, assume um valor mais elevado (água apresentando menos flúor, maior concentração de cloro, temperaturas mais elevadas e quantidade maior de bactérias heterotróficas) nos meses com uma maior precipitação, temperaturas (mínima, máxima e média) mais elevadas, baixa umidade relativa de ar e baixa velocidade dos ventos (Quadro 5).

Quadro 5. Resumo da associação entre o componente de qualidade da água na rede de distribuição (u_2) e a caracterização climática (v_2)

Variáveis climáticas	Componente de qualidade da água - u_2	
	Assumindo valores altos ↑ (Flúor ↓, Cloro ↑, Temperatura ↑ e Heterotróficas ↑)	Assumindo valores baixos ↓ (Flúor ↑, Cloro ↓, Temperatura ↓ e Heterotróficas ↓)
Dias chuvosos		
Temp. mínima	↑	↓
Temp. máxima	↑	↓
Temp. média	↑	↓
Precipitação	↑	↓
Umidade relativa	↓	↑
Veloc. dos ventos	↓	↑
Evaporação		
Insolação		
Radiação solar		

6.2.2 Associação entre o conjunto de variáveis de qualidade da água observado no final do tratamento e o conjunto de variáveis climáticas

O resumo dos dados das variáveis de qualidade da água observadas no final da estação de tratamento (saída dos reservatórios) e das características climáticas, com os respectivos vetores de médias, quartis e valores máximos e mínimos, pode ser visualizado nas Tabelas 7 e 8.

As variâncias e covariâncias, bem como as correlações lineares entre as variáveis climáticas foram apresentadas na seção anterior. Para as variáveis de qualidade da água avaliadas no final de tratamento, esses resultados são destacados na Tabela 16.

Assim como na rede de distribuição, observa-se, no final de tratamento da água, correlação positiva significativa ($p < 0,0001$) mais elevada entre turbidez e cor aparente (0,83), enquanto que correlações significativas negativas mais moderadas (todas com $p < 0,0001$) são observadas entre flúor e cloro (-0,68) e entre o cloro e pH (-0,54).

Tabela 16. Variâncias, covariâncias ($\mathbb{S}^{(1,1)}$) e correlações ($\mathbb{R}^{(1,1)}$) das variáveis de qualidade da água no final de tratamento *

	TURB	CORAP	F	CRL	pH	T	HET	CT
TURB	0,1641	0,83	0,26	-0,26	0,40	0,12	0,05	0,07
CORAP	0,5386	2,5381	0,24	-0,24	0,20	0,20	-0,07	0,11
F	0,0041	0,0144	0,0015	-0,68	0,44	0,13	0,00	0,20
CRL	-0,0374	-0,1372	-0,0092	0,1256	-0,54	-0,36	-0,08	-0,02
pH	0,0652	0,1297	0,0067	-0,0754	0,1580	0,07	0,11	-0,01
T	0,1319	0,8361	0,0127	-0,3417	0,0717	7,1049	-0,20	-0,03
HET	1,3793	-7,8434	-0,0092	-2,0795	3,2377	-39,209	0,0055	0,03
CT	0,0005	0,0033	0,0001	-0,0001	-0,00008	-0,0013	0,0405	0,0003

* As variâncias são apresentadas na diagonal, já as covariâncias e as correlações abaixo e acima da diagonal, respectivamente.

As covariâncias e correlações amostrais entre as variáveis do primeiro (qualidade da água no final de tratamento) e do segundo conjunto (climáticas) são dadas na Tabela 17 e 18.

Tabela 17. Matriz de covariâncias ($S^{(1,2)}$) entre as variáveis do primeiro e segundo conjunto, considerando o final de tratamento da água

	DIASC	TMIN	TMAX	TMED	PP	UR	VV	EV	ISOL	RAD
TURB	0,046	0,116	0,044	0,125	4,231	0,378	-0,568	-0,012	0,015	0,270
CORAP	1,498	0,834	0,439	0,792	26,637	1,945	1,651	0,066	-0,134	0,364
F	0,015	0,016	0,004	0,020	-0,283	0,146	0,161	-0,001	-0,011	0,270
CRL	-0,761	-0,379	-0,115	-0,288	-7,852	-2,404	-1,313	-0,047	0,191	-4,527
pH	0,131	0,126	-0,113	0,054	1,049	1,446	-1,850	-0,118	-0,060	-4,196
T	6,382	4,354	3,380	2,855	85,547	5,141	0,043	1,361	-0,967	76,357
HET	-92,628	-23,438	-17,136	-17,211	-1345,266	-217,052	130,997	-4,409	21,848	-191,770
CT	-0,021	-0,007	-0,007	-0,009	-0,236	0,005	0,023	-0,006	-0,001	-0,156

Tabela 18. Matriz de correlação ($R^{(1,2)}$) entre as variáveis do primeiro e segundo conjunto, considerando o final de tratamento da água

	DIASC	TMIN	TMAX	TMED	PP	UR	VV	EV	ISOL	RAD
TURB	0,020	0,110	0,050	0,130	0,080	0,080	-0,060	-0,020	0,030	0,010
CORAP	0,160	0,200	0,120	0,210	0,130	0,110	0,050	0,040	-0,060	0,000
F	0,070	0,170	0,040	0,210	-0,060	0,340	0,190	-0,010	-0,210	0,100
CRL	-0,370	-0,420	-0,140	-0,330	-0,180	-0,600	-0,170	-0,110	0,400	-0,190
pH	0,060	0,120	-0,120	0,060	0,020	0,320	-0,210	-0,250	-0,110	-0,160
T	0,410	0,640	0,550	0,440	0,260	0,170	0,000	0,430	-0,270	0,420
HET	-0,220	-0,120	-0,100	-0,100	-0,150	-0,260	0,080	-0,050	0,220	-0,040
CT	-0,200	-0,160	-0,180	-0,200	-0,100	0,020	0,060	-0,260	-0,040	-0,130

O gráfico de probabilidade *QQ Plot (Chi-Square plot)* multivariado pode ser visualizado na Figura 32 do Apêndice 3. Percebem-se também neste gráfico algumas distâncias não seguindo o padrão das demais no sentido de se alinharem ao redor de uma linha reta. No entanto, considerando o tamanho amostral e a simetria dos dados, como já observada anteriormente, pode-se considerar a aderência dos mesmos à distribuição normal. Em relação às suposições de linearidade, multicolinearidade, homocedasticidade e autocorrelação, destacam-se, nesta análise de caráter exploratório, as mesmas considerações observadas na rede de distribuição.

Assim como observado na rede de distribuição, das oito correlações canônicas encontradas com os dados de qualidade da água no final de tratamento, somente as duas primeiras foram estatisticamente significativas ($p < 0,05$), apresentando correlações iguais a 0,872 e 0,694 respectivamente (Tabela 19). No entanto, essas correlações apresentaram-se em valores inferiores ao observado na rede, mas nos dois casos, estão indicando que existe uma associação entre as características climáticas e as relacionadas à qualidade da água.

Tabela 19. Autovalor, correlação canônica e resultado do teste estatístico – Final de Tratamento da água

Pares de variáveis canônicas	Autovalor	Correlação Canônica	Estatística de teste	gl	Valor p
(u_1, v_1)	0,760	0,872	153,37	80	< 0,001
(u_2, v_2)	0,481	0,694	84,33	63	0,038
(u_3, v_3)	0,327	0,572	52,20	48	0,314
(u_4, v_4)	0,241	0,491	32,41	35	0,594
(u_5, v_5)	0,160	0,399	18,24	24	0,791
(u_6, v_6)	0,092	0,304	9,11	15	0,872
(u_7, v_7)	0,044	0,210	3,92	8	0,865
(u_8, v_8)	0,025	0,158	1,43	3	0,698

Dessa forma, os valores padronizados do par de variáveis canônicas com maior correlação, $(u_1$ e $v_1)$, podem ser visualizados na Figura 16. Observa-se que valores altos da variável canônica que representa as variáveis climáticas (v_1) estão associados com valores altos da variável canônica que representa as variáveis de qualidade da água (u_1). Além disso, nota-se um padrão semelhante da rede de distribuição em relação aos meses avaliados, ou seja, principalmente nos meses de junho (2010) e julho (2008 e 2010) verificam-se valores altos tanto para v_1 quanto para u_1 , enquanto que nos meses mais quentes do ano (de novembro a fevereiro) as duas variáveis canônicas apresentam valores mais baixos.

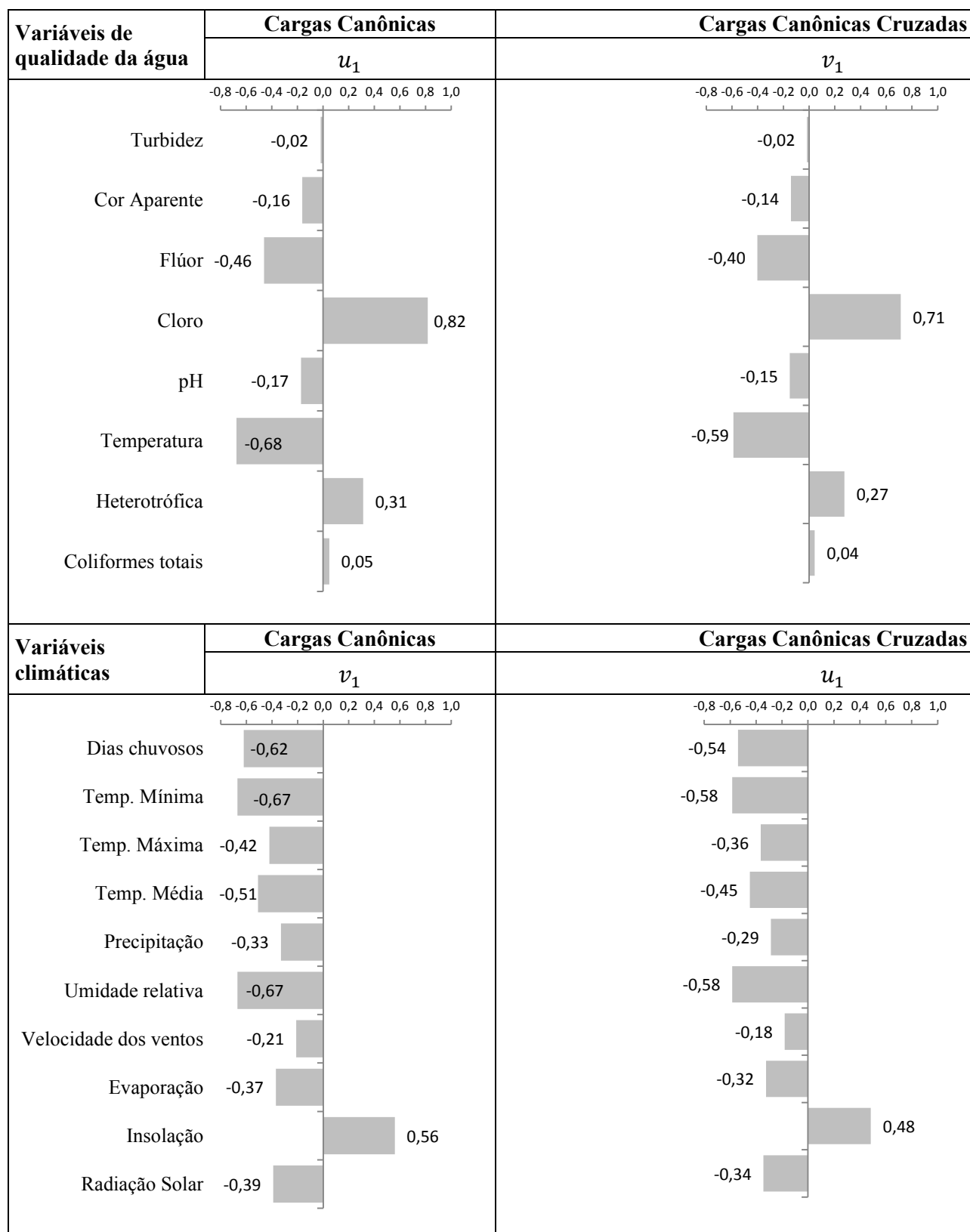


Figura 17. Gráfico das correlações de cada variável original com as variáveis canônicas (u_1 e v_1), considerando o final de tratamento

Todavia, observando as cargas canônicas entre as variáveis originais de qualidade da água com o componente u_1 , destacam-se o cloro (0,82), a temperatura da água (-0,68) e o flúor (-0,46) com cargas mais altas. Portanto, a variável canônica u_1 expressa um contraste entre a concentração de cloro com a temperatura da água e o flúor. Assim, este componente canônico se apresenta com valores mais elevados quando há concentrações maiores de cloro e valores menores de temperatura da água e concentrações de flúor.

Já as cargas canônicas entre as variáveis originais do conjunto climático com o componente v_1 , as que mais se destacam são temperatura mínima e umidade relativa do ar (-0,67), número de dias chuvosos (-0,62) e insolação (0,56). Essas correlações indicam que o componente climático v_1 expressa um contraste entre a insolação com a umidade relativa, número de dias chuvosos e temperatura mínima. Neste sentido, este componente aponta valores mais altos em situações nas quais a insolação é alta e são baixas a temperatura mínima, a umidade relativa e a frequência de dias chuvosos.

Em relação às cargas canônica cruzadas, as variáveis originais de qualidade da água que possuem maiores associações com o componente climático são cloro (0,71), temperatura da água (-0,59) e flúor (-0,40). Desta forma, observa-se no final de tratamento que o componente climático associa-se positivamente com concentrações de cloro e negativamente com flúor e com a temperatura da água, ou melhor, quando o componente climático apresenta valores altos (alta insolação e baixa temperatura mínima, umidade relativa e frequência de dias chuvosos) observa-se que a concentração de cloro na água é mais elevada e a concentração de flúor e a temperatura da água apresentam-se baixas.

As variáveis climáticas com maiores associações com o componente canônico da água são temperatura mínima e umidade relativa (-0,58), número de dias chuvosos (-0,54) e insolação (0,48). Os resultados indicam que o componente canônico da água com valores altos (alta concentração de cloro e baixa temperatura e concentração de flúor) estão associados com valores altos de insolação e baixos da temperatura mínima, da umidade relativa do ar e do número de dias chuvosos.

Um resumo da associação entre as variáveis canônicas u_1 e v_1 está apresentado no Quadro 6, destacando-se a situação climática quando o componente expressando uma característica de qualidade da água assume valores mais elevados ($\uparrow$) ou mais baixos ($\downarrow$). Com u_1 associado positivamente com v_1 , pode-se dizer que o

componente u_1 assume valores mais elevados (maiores concentrações de cloro, menor temperatura e baixas concentrações de flúor na água em final de tratamento) nos meses de maior insolação, menor umidade relativa, menos dias chuvosos e com temperatura mínima mais baixa. Situação inversa ocorre nos meses com o componente de qualidade da água assumindo valores mais baixos

Quadro 6. Resumo da associação entre o componente de qualidade da água em final de tratamento (u_1) e a caracterização climática (v_1)

Variáveis climáticas	Componente de qualidade da água - u_1	
	Assumindo valores altos ↑ (Cloro ↑, Temperatura ↓ e Flúor ↓)	Assumindo valores baixos ↓ (Cloro ↓, Temperatura ↑ e Flúor ↑)
Dias chuvosos	↓	↑
Temp. mínima	↓	↑
Temp. máxima		
Temp. média		
Precipitação		
Umidade relativa	↓	↑
Veloc. dos ventos		
Evaporação		
Insolação	↑	↓
Radiação solar		

6.3 Análise Fatorial

Primeiramente, a técnica de análise fatorial foi utilizada considerando as oito variáveis relacionadas com os parâmetros de qualidade da água na rede de distribuição e as 10 variáveis climáticas, buscando determinar novas variáveis (não observáveis) formadas por subconjuntos das variáveis originais, especificadas de acordo com o maior interesse para informação prática contida nas novas variáveis.

Posteriormente, a análise fatorial é apresentada separadamente para o conjunto de variáveis climáticas e de qualidade da água, criando indicadores que expressam as informações práticas das mesmas em fatores, os quais permitem um estudo simplificado de associação.

6.3.1 Análise do conjunto de dados formado por todas as variáveis (qualidade da água na rede e climáticas)

O intuito de se utilizar a análise fatorial, considerando simultaneamente as variáveis de qualidade da água e climáticas, foi de identificar possíveis somas ou contrastes entre algumas variáveis climáticas e características da água na rede de distribuição.

Para uma melhor visualização das associações entre as variáveis analisadas, a Figura 18 apresenta a matriz de correlação amostral, bem como quadrados abaixo da diagonal indicando por cores a intensidade do grau de associação entre as mesmas. Os quadrados coloridos em tons de azul indicam correlações positivas e os quadrados vermelhos correlações negativas. Quanto mais escuro a tonalidade da cor, maior o grau de associação, tanto positiva quanto negativa.

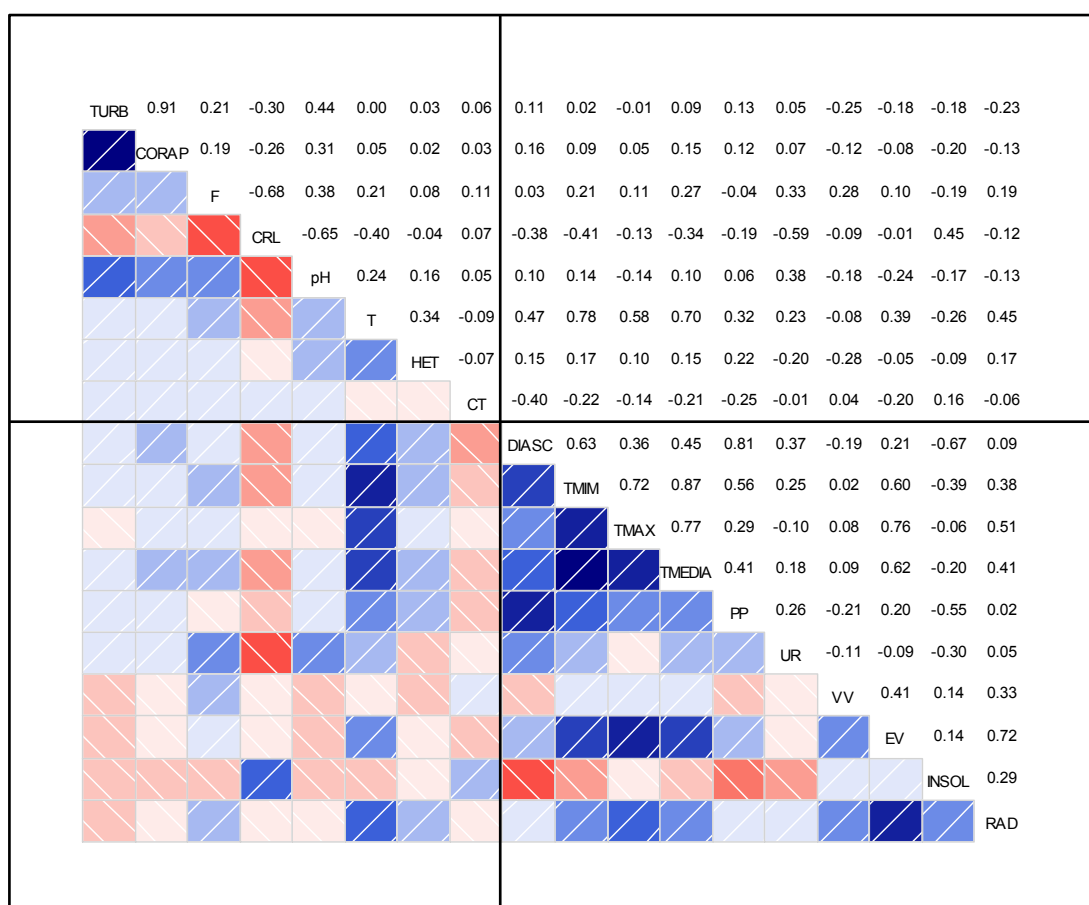


Figura 18. Gráfico das correlações entre as variáveis climáticas e de qualidade da água

Para verificar se a matriz de correlação pode ser considerada igual à matriz identidade, indicando que não há correlações entre as variáveis, foi utilizado o teste de *Bartlett*, o qual rejeitou a hipótese inicial de que as variáveis não estão correlacionadas entre si ($p < 0,01$). A medida global de adequação amostral para análise fatorial (KMO) apresentou um resultado igual 0,64, indicando uma adequação aceitável ($KMO > 0,5$). Já as medidas de adequação amostral para cada variável apresentou resultados insatisfatórios para as variáveis TURB (0,47), CORAP (0,47), HET (0,31), CT (0,42), UR (0,48) e VV (0,36) e satisfatórios para as variáveis F(0,71), CRL (0,62), pH (0,66), T (0,78), DIASC (0,73), TMIN (0,78), TMAX (0,71), TMED (0,79), PP (0,65), EV (0,65), INSOL (0,74) e RAD (0,61).

Por apresentarem medidas de adequação amostral muito baixas, decidiu-se retirar as variáveis bactérias heterotróficas (HET) e velocidade dos ventos (VV) na aplicação dos procedimentos da análise fatorial. Com isso, o valor do KMO global passou a ser igual a 0,79 e nenhuma variável apresentou valor inferior a 0,5.

Foram obtidos quatro fatores associados à autovalores maiores que a unidade, os quais explicam 74,7% da variância total das variáveis originais. Os resultados das estimativas das cargas fatoriais (*loadings*), utilizando a rotação varimax, das comunidades e das variâncias específicas (unicidades) são apresentados na Tabela 20.

Percebe-se que apenas a variável CT não apresentou uma comunalidade elevada em relação à variância específica (comunalidade < variância específica), indicando que apenas 27,4% da variabilidade dessa variável é explicada pelos fatores, e o restante (72,6%) é a variabilidade associada ao erro aleatório (variância específica). No entanto, de uma forma geral, os quatro fatores informam adequadamente o comportamento biológico apresentado pelas variáveis analisadas.

Tabela 20. Cargas fatoriais, autovalores, percentual acumulado de explicação da variância, comunalidade e unicidade das variáveis de qualidade da água e climáticas (Varimax)

Variável	Carga fatorial dos fatores*				Comunalidade	Variância específica
	F1	F2	F3	F4		
TURB	-0,077	-0,172	-0,039	-0,955	0,950	0,050
CORAP	0,017	-0,116	-0,063	-0,939	0,900	0,100
F	0,217	-0,744	0,182	-0,124	0,649	0,351
CRL	-0,162	0,883	0,191	0,158	0,867	0,133
pH	-0,098	-0,719	0,004	-0,340	0,641	0,359
T	0,675	-0,335	-0,297	0,001	0,656	0,344
CT	-0,124	-0,147	0,480	-0,079	0,274	0,726
DIASC	0,246	-0,173	-0,888	-0,058	0,882	0,118
TMIM	0,747	-0,240	-0,514	-0,037	0,880	0,120
TMAX	0,884	0,104	-0,182	-0,085	0,832	0,168
TMEDIA	0,815	-0,183	-0,317	-0,131	0,815	0,185
PP	0,192	-0,034	-0,839	-0,088	0,750	0,250
UR	-0,053	-0,726	-0,283	0,144	0,630	0,370
EV	0,884	0,159	-0,019	0,104	0,817	0,183
INSOL	0,121	0,317	0,764	0,133	0,716	0,284
RAD	0,770	-0,071	0,218	0,232	0,700	0,300
Autovalor	5,24	3,18	1,96	1,58		
% Acumulada	32,7	52,6	64,9	74,7		

* Interpretações dos fatores: F1 – Componente Climático, F2 – Físico-Químico, F3 – Microbiológico e F4 – Estético

A partir das cargas fatoriais ou *loadings* (Tabela 20) é possível identificar as variáveis com maiores associações com os respectivos fatores. Essas associações podem ser visualizadas por meio de representações bidimensionais (Figura 19), representando o cruzamento das cargas fatoriais de dois fatores quaisquer.

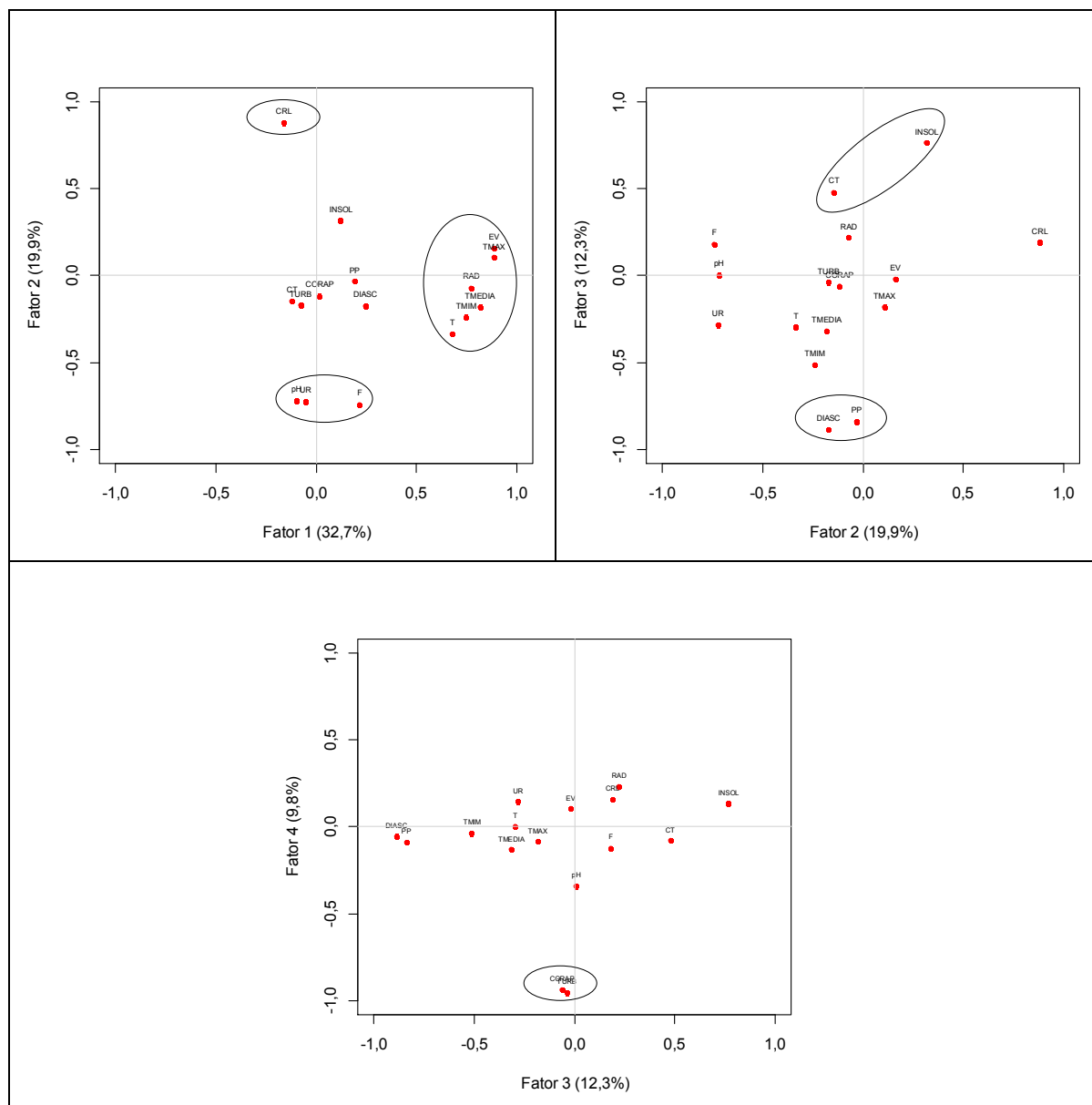


Figura 19. Distribuição dos fatores para as variáveis avaliadas

Com essas representações gráficas, juntamente com a observação das cargas fatoriais da Tabela 20, é possível caracterizar as informações biológicas expressas por cada um dos quatro fatores.

Esses resultados permitem caracterizar o primeiro fator (F1), representando 32,7% da variabilidade total dos dados, como um componente climático no qual se somam as variáveis temperatura do ar (temperatura máxima, mínima e média), evaporação, radiação solar e temperatura da água distribuída. O segundo fator (F2), representando 19,9%, caracteriza-se como indicativo de tratamento da água (componente físico-químico), no que se observa um contraste entre o cloro residual livre com a soma do

flúor, pH e umidade relativa do ar. Já o terceiro fator (F3), com 12,3%, caracteriza-se como um componente microbiológico com precipitação, representando um contraste entre a soma de ausência de coliformes totais com insolação versus a soma entre número de dias chuvosos e precipitação. O quarto fator (F4), com 9,8%, representa o somatório de turbidez e cor da água distribuída, caracterizando-se como um fator estético.

Tabela 21. Coeficientes dos fatores – todas as variáveis

Variável	F1	F2	F3	F4
TURB	-0,014	0,059	0,058	-0,696
CORAP	0,042	0,094	0,020	-0,359
F	0,041	-0,216	0,120	0,023
CRL	-0,003	0,596	-0,092	-0,086
pH	-0,020	-0,183	0,051	0,004
T	0,083	-0,065	0,006	0,012
CT	0,008	-0,037	0,058	-0,002
DIASC	-0,097	0,074	-0,552	0,028
TMIM	0,229	-0,048	-0,142	0,009
TMAX	0,277	0,128	0,040	-0,060
TMEDIA	0,204	-0,017	0,005	-0,028
PP	-0,049	0,087	-0,265	-0,003
UR	-0,042	-0,185	-0,020	0,060
EV	0,276	0,103	0,102	-0,023
INSOL	0,103	0,032	0,214	-0,017
RAD	0,162	-0,046	0,139	0,018

A Tabela 21 apresenta os coeficientes da combinação linear para a determinação dos escores fatoriais, em que cada valor representa o peso de ponderação para cada variável original padronizada. Os escores fatoriais mensais obtidos para o fator são determinados pelo produto dos pesos pelos respectivos valores padronizados das variáveis originais.

Para os escores do fator 1 (F1) destacam-se como mais contributivas as variáveis temperatura máxima com peso de 0,277, seguida da evaporação com 0,276; temperatura mínima, com 0,229; temperatura média, com 0,204; radiação solar global com 0,162 e temperatura da água com 0,083.

Na determinação do escore mensal do fator 2 (F2), a variável com maior peso é o cloro (0,596), seguida das variáveis flúor, umidade relativa do ar e pH, as quais contribuem negativamente com -0,216, -0,185 e -0,183 de seus respectivos valores. Isso indica que uma variação de uma unidade da variável cloro é compensada com uma variação acumulada na mesma direção nas variáveis flúor, umidade relativa e pH.

Nos escores mensais do fator 3 (F3), as variáveis com maior peso são frequência de dias chuvosos (-0,552) e precipitação (-0,265), contrastando com os pesos da insolação (0,214) e proporção de amostras com ausência de coliformes totais (0,058).

Já nos escores mensais do fator 4, (F4) a variável com maior peso é turbidez, contribuindo com -0,696, seguida da variável cor aparente com -0,359. Observa-se que a contribuição da turbidez é quase duas vezes a contribuição da cor aparente.

6.3.2 Análise do conjunto de dados formado somente pelas variáveis de qualidade da água na rede de distribuição

Considerando apenas as variáveis de qualidade da água na rede de distribuição, a matriz de correlação amostral com o indicativo do grau de correlação por meio de quadrados coloridos é apresentada na Figura 20.

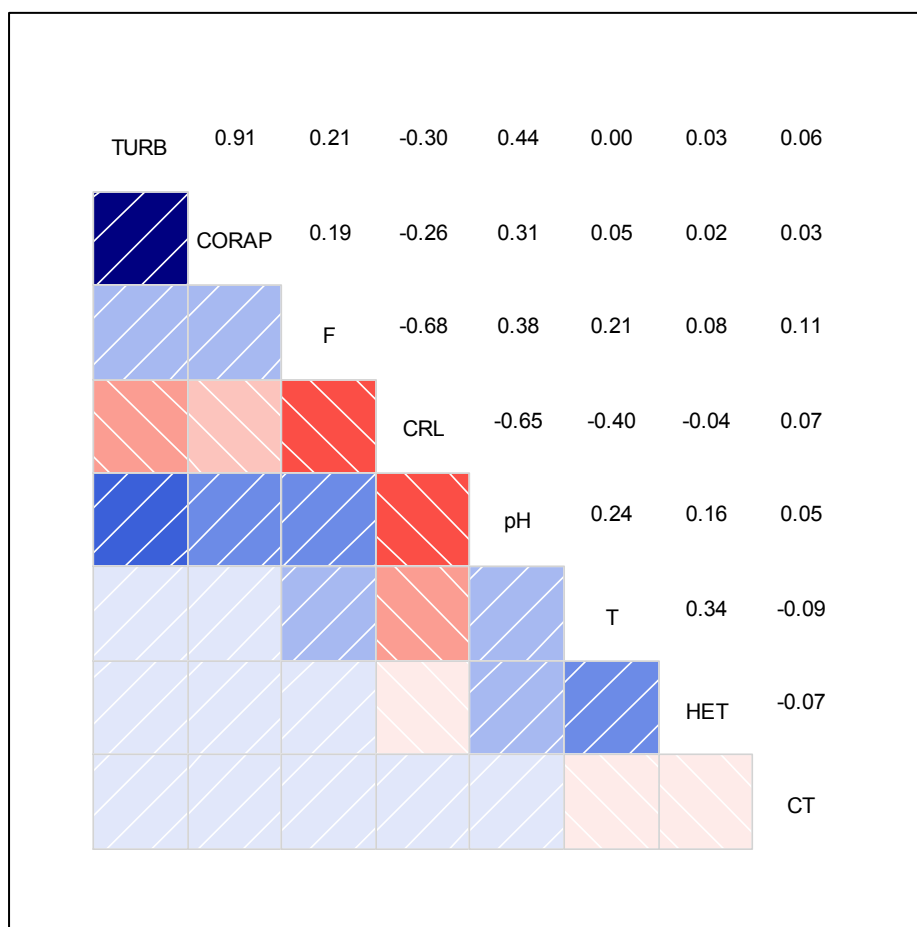


Figura 20. Gráfico das correlações entre as variáveis de qualidade da água

Tabela 22. Cargas fatoriais, autovalores, percentual acumulado de explicação variância, comunalidade e unicidade das variáveis de qualidade da água na distribuição (Varimax)

Variável	Carga fatorial			Comunalidade	Variância específica
	F1A	F2A	F3A		
TURB	0,187	-0,960	0,026	0,957	0,043
CORAP	0,124	-0,952	-0,009	0,921	0,079
F	0,834	-0,040	0,107	0,709	0,291
CRL	-0,898	0,151	0,087	0,837	0,163
pH	0,694	-0,354	-0,108	0,618	0,382
T	0,450	0,087	-0,636	0,615	0,385
HET	0,085	-0,028	-0,755	0,577	0,423
CT	0,169	0,011	0,561	0,343	0,657
Autovalor	2,86	1,57	1,15		
% Acumulada	28,4	53,2	69,7		

* Interpretações dos fatores: F1A – Componente Físico-Químico, F2A – Estético e F3A – Microbiológico

Os três fatores obtidos, associados aos autovalores maiores que um, explicam 69,7% da variância total das variáveis originais. A rotação varimax foi utilizada para facilitar a interpretação dos fatores e os resultados das estimativas das cargas fatoriais (*loadings*), das comunidades e das variâncias específicas (unicidades) são apresentados na Tabela 22.

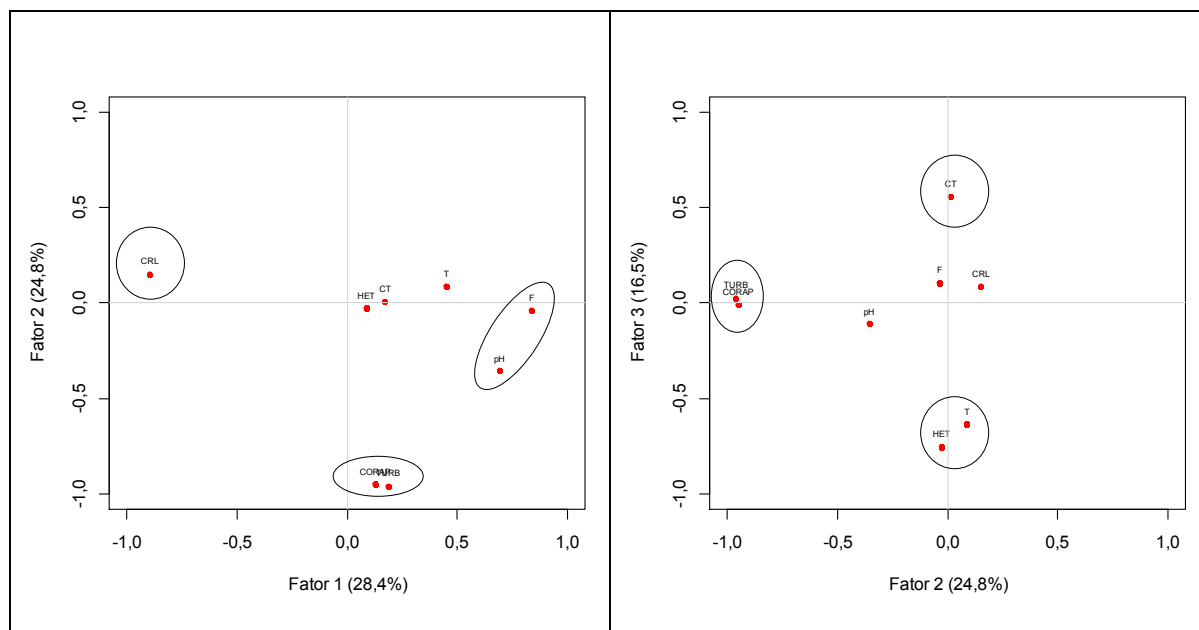


Figura 21. Distribuição dos fatores para as variáveis de qualidade da água

Observando a Figura 21, conjuntamente com as cargas fatoriais apresentadas na Tabela 22, é possível caracterizar três fatores, simbolizando os parâmetros de qualidade da água.

O primeiro fator (F1A), correspondendo a 28,4% da variabilidade total dos dados, é caracterizado como um componente físico-químico, definindo-se como um contraste entre a soma das contribuições do flúor e do pH com o cloro. Já o segundo fator (F2A), correspondendo a 24,8% da variabilidade, pode ser compreendido como um componente estético da água, mostrando a soma das características de turbidez e cor aparente da água. O terceiro fator (F3A), com 16,5% da variabilidade total dos dados, é um componente microbiológico, denotando um contraste entre a proporção de ausência de coliformes totais e a soma das contribuições da temperatura da água e quantidade de unidades formadoras de colônias de bactérias heterotróficas.

Tabela 23. Coeficientes dos fatores – variáveis de qualidade da água

Variável	F1A	F2A	F3A
TURB	-0,038	-0,672	0,101
CORAP	-0,135	-0,395	-0,135
F	0,360	0,076	0,254
CRL	-0,619	-0,105	-0,047
pH	0,183	0,011	-0,031
T	0,074	0,017	-0,522
HET	-0,057	-0,022	-0,615
CT	0,070	0,019	0,310

Os pesos de ponderação de cada variável de qualidade da água padronizada para determinação dos escores fatoriais são apresentados na Tabela 23.

Destacam-se na composição do fator 1 (F1A) as variáveis cloro com peso negativo igual a -0,619 e as variáveis flúor e pH com pesos positivos iguais a 0,360 e 0,183, respectivamente. Mais uma vez, pode-se dizer que uma variação no valor da variável cloro é compensada com uma variação acumulada na mesma direção nas variáveis flúor e pH.

Para a determinação do escore mensal do fator 2 (F2A), a variável com maior peso é turbidez, contribuindo com -0,672, seguida da variável cor aparente com -0,395.

Nos escores mensais do fator 3 (F3A), tem-se as variáveis com maior peso, a quantidade de bactérias heterotróficas (-0,615) e a temperatura da água (-0,522), contrastando com o peso da proporção de amostras com ausência de coliformes totais (0,310).

6.3.3 Análise do conjunto de dados formado somente pelas variáveis climáticas

Considerando apenas as variáveis climáticas, a matriz de correlação amostral com o indicativo do grau de correlação por meio de quadrados coloridos é destacada na Figura 22.

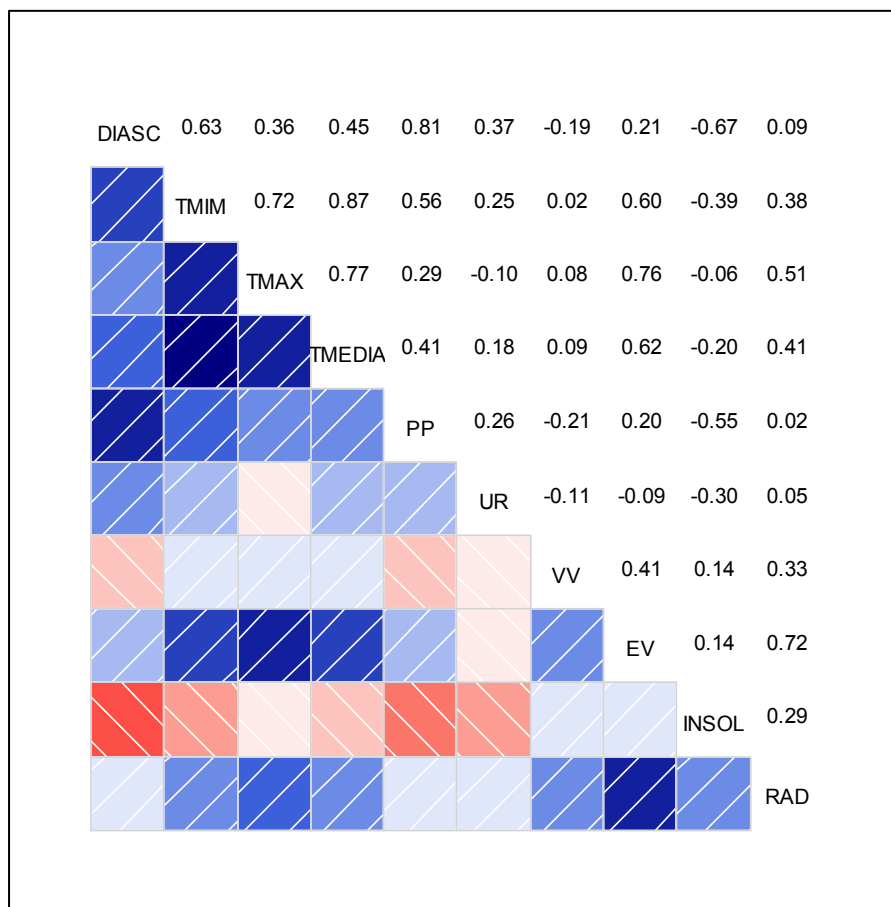


Figura 22. Gráfico das correlações entre as variáveis climáticas

Os dois fatores obtidos, associados aos autovalores maiores que um, explicam 67,5% da variância total das variáveis originais. A rotação varimax foi utilizada para facilitar a interpretação dos fatores, e os resultados das estimativas das cargas fatoriais (*loadings*), das comunidades e das variâncias específicas (unicidades) são apresentados na Tabela 24.

Tabela 24. Cargas fatoriais, autovalores, percentual acumulado de explicação variância, comunalidade e unicidade das variáveis climáticas (Varimax)

Variável	Carga fatorial		Comunalidade	Variância específica
	F1C	F2C		
DIASC	-0,281	0,870	0,836	0,164
TMIM	-0,737	0,568	0,865	0,135
TMAX	-0,863	0,170	0,773	0,227
TMEDIA	-0,796	0,379	0,776	0,224
PP	-0,236	0,817	0,723	0,277
UR	0,043	0,504	0,256	0,744
VV	-0,384	-0,415	0,320	0,680
EV	-0,931	-0,083	0,874	0,126
INSOL	-0,095	-0,820	0,682	0,318
RAD	-0,773	-0,229	0,650	0,350
Autovalor	4,22	2,53		
% Acumulada	42,2%	67,5%		

* Interpretações dos fatores: F1C – Componente de Temperatura e F2C – Precipitação

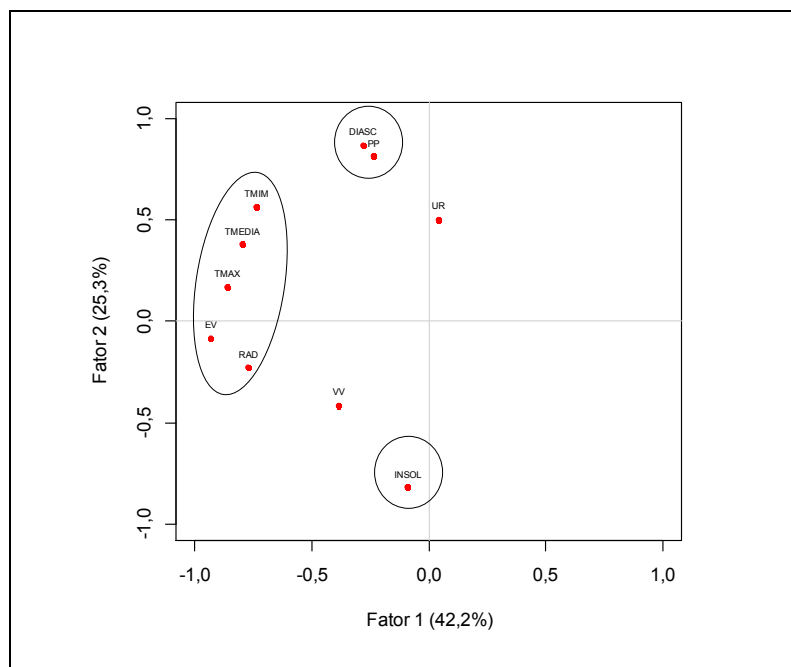


Figura 23. Distribuição dos fatores para as variáveis climáticas

Percebe-se que apenas dois fatores são necessários para reter 67,5% da variabilidade total dos dados das características climáticas, e a caracterização dos

mesmos pode ser realizada, observando as cargas fatoriais apresentadas na Tabela 24 e destacadas pelo gráfico da Figura 23.

O primeiro fator (F1C) caracteriza-se como um componente de temperatura, representando o somatório das condições de temperatura do ar (mínima, máxima e média), evaporação e radiação solar. Já o segundo fator (F2C) representa um contraste entre a soma de insolação e velocidade dos ventos e a soma da precipitação, número de dias chuvosos e umidade relativa do ar.

Tabela 25. Coeficientes dos fatores – variáveis climáticas

Variável	F1C	F2C
DIASC	0,033	0,419
TMIM	-0,212	0,230
TMAX	-0,202	-0,030
TMEDIA	-0,164	0,059
PP	0,024	0,236
UR	0,020	0,061
VV	-0,048	-0,067
EV	-0,445	-0,243
INSOL	-0,082	-0,232
RAD	-0,145	-0,113

Os pesos de ponderação de cada variável climática padronizada para determinação dos escores fatoriais são mostrados na Tabela 25.

Para determinação dos escores do fator 1 (F1C), destacam-se com maiores pesos as variáveis evaporação (-0,445), temperatura mínima, máxima e média (-0,212, -0,202 e -0,164) e radiação solar global com -0,145.

Já para os escores mensais do fator 2 (F2C), as variáveis que mais contribuem são: frequência de dias chuvosos (0,419), precipitação (0,236), evaporação (-0,243) e insolação (-0,232), indicando que a soma entre frequência de dias chuvosos e precipitação se antagoniza com a soma de evaporação e insolação.

6.3.4 Correlação linear simples entre os fatores climáticos e os de qualidade da água

Considerando apenas os três fatores obtidos a partir da análise fatorial com as variáveis de qualidade da água (F1A, F2A e F3A) e os dois fatores obtidos a partir das variáveis climáticas (F1C e F2C), os quais foram definidos nas seções 4.4.2 e 4.4.3, os coeficientes de correlação linear amostral de Person entre estes fatores, com seus respectivos *valores p* do teste de significância, são apresentados na Tabela 26, sendo as correlações significativas indicadas por um *.

Tabela 26. Coeficientes de correlações entre os fatores de qualidade da água e climáticos

Correlações	Coeficientes	Valor p
F1A x F1C	-0,15	0,242
F1A x F2C	0,29 *	0,023
F2A x F1C	-0,13	0,315
F2A x F2C	-0,18	0,171
F3A x F1C	0,34 *	0,008
F3A x F2C	-0,37 *	0,003

Observando os resultados da Tabela 26, verifica-se que o F3A apresenta correlações significativas de forma negativa com F2C (-0,37) e positiva com F1C (0,34), sendo as maiores correlações apresentadas. Outra correlação significativa, mas de menor magnitude de associação, foi entre F1A e F2C (0,29).

Estes resultados indicam que valores altos do fator F3A estão associados a valores baixos do fator F2C. Desta forma, valores altos para F3A são observados quando a proporção de amostras com ausência de coliformes totais é alta e a temperatura da água conjuntamente com a quantidade de bactérias heterotróficas são baixas. Já, valores baixos para F2C são observados quando as medidas de insolação e da velocidade dos ventos são elevadas, enquanto que a frequência de dias chuvosos, precipitação e umidade relativa são baixas.

Em relação à segunda maior medida de correlação entre os fatores, verifica-se que valores altos do fator F3A são observados quando os valores de F1C também são altos. Isto significa que quando a proporção de ausência de coliformes é alta e

a temperatura da água e a quantidade de bactérias heterotróficas assumem valores baixos, também são observados valores baixos para temperatura ambiente (máxima, mínima e média), evaporação e radiação solar.

Outra correlação significativa foi entre F1A e F2C. Este resultado indica que situações com baixa quantidade de cloro na água e alta concentração de flúor e pH estão associadas com baixos valores de insolação e velocidade dos ventos e altos valores da frequência de dias chuvosos, precipitação e umidade relativa.

7 CONCLUSÕES

Estudos avaliando características físico-químicas e microbiológicas da água são bastante comuns, porém, a utilização de uma metodologia que considera toda estrutura de variabilidade dos dados (estatística multivariada) com o intuito de avaliar a associação destas características com variáveis climáticas se apresentou muito pouco utilizada.

A análise de dados envolvendo estruturas complexas e multidimensionais por meio de técnicas estatísticas multivariadas, metodologia empregada nesta pesquisa, utiliza-se de modelos estatísticos mais próximos da realidade biológica dos dados, obtendo-se resultados que expressam melhor a situação de relacionamento entre as variáveis climáticas e de qualidade da água.

Na comparação entre os vetores médios com características de qualidade da água avaliados na rede de distribuição e no final de tratamento, verificou-se diferença significativa, sendo a variável determinante para esta diferenciação a temperatura da água. Constatou-se que a temperatura média da água foi superior na rede de distribuição. No entanto, a temperatura da água é um aspecto independente do tratamento que não pode ser controlado, pois depende do meio ambiente. Certificou-se também que os demais parâmetros de qualidade da água não evidenciaram diferenças significativas entre os dois locais de análise e se apresentaram, em média, dentro dos padrões de qualidade estabelecidos pela legislação.

Assim, pode-se afirmar que, no município de Botucatu, o padrão de qualidade da água mantém-se entre o final de tratamento até a rede de distribuição

(cavaletes das residências), ou seja, o padrão de qualidade permanece estável em todo o sistema. Desta forma, o procedimento multivariado de T^2 de Hotteling pode ser utilizado como uma técnica de avaliação do padrão de qualidade da água nos diversos municípios e empresas de saneamento.

Embora, nos intervalos de confiança simultâneos de Hotteling, não fosse possível verificar significância entre os dois locais de análise para cor aparente e cloro residual livre, sobre o aspecto prático para um estudo mais aprofundado, o valor médio do parâmetro cor aparente se mostrou superior na rede de distribuição, enquanto que a concentração média de cloro residual foi maior no final de tratamento. Estes resultados exploratórios sugerem a hipótese de que essa característica da água está relacionada com o tipo de tubulação (pois o ferro e o manganês utilizado atualmente nas tubulações podem influenciar na cor) e com a temperatura ambiente mais elevada (ocasionando uma maior evaporação do cloro na rede). Se realmente essa hipótese for comprovada, estes resultados sugerem atenção em relação ao material utilizado nas tubulações e apontam para a necessidade de aumento da concentração de cloro no tratamento em dias mais quentes.

No estudo da associação conjunta entre as variáveis de qualidade da água na rede de distribuição e as climáticas foram identificadas duas correlações positivas significativas. No contexto das cargas canônicas tem-se para o componente de qualidade da água destaque para flúor, cloro e temperatura e para o componente climático a frequência de dias chuvosos, temperaturas, umidade relativa, radiação solar e insolação. Os meses com alta umidade, temperaturas mais elevadas, maiores níveis de radiação solar e menores insolações estão associados com meses em que a água na rede se apresenta com temperaturas mais elevadas, maiores concentrações de flúor e menores concentrações de cloro.

Já no estudo da associação conjunta entre as variáveis de qualidade da água no final de tratamento e as climáticas também foram identificadas duas correlações positivas significativas. No contexto das cargas canônicas tem-se para o componente de qualidade da água os mesmos destaques apresentados na rede de distribuição, sobressaindo o cloro no final de tratamento, enquanto que, na rede, a temperatura.

Na análise fatorial envolvendo as variáveis da qualidade da água e climáticas conjuntamente foram estabelecidos quatro fatores que explicam acumuladamente 74,7% da variabilidade total dos dados. O primeiro fator envolve as variáveis temperatura da água, temperatura ambiental (mínima, máxima e média),

evaporação e radiação solar global; o segundo, flúor, cloro residual livre, pH e umidade relativa do ar; o terceiro, proporção de amostras com ausência de coliformes totais, frequência de dias chuvosos, precipitação e insolação; e o último, turbidez e cor aparente.

Para os conjuntos isolados de variáveis de qualidade de água e climáticas, tem-se, para o primeiro o emprego de três fatores com retenção acumulada de 69,7% da variância total e, para o segundo dois fatores, com 67,5%. Na qualidade da água, destacam-se no primeiro fator flúor, cloro residual livre e pH; no segundo, turbidez e cor aparente; e no terceiro temperatura da água, quantidade de colônias de bactérias heterotróficas e proporção de amostras com ausência de coliformes totais. No conjunto climático, destacam-se no primeiro fator, temperatura ambiente, evaporação e radiação solar global e no segundo, frequência de dias chuvosos, precipitação e insolação. Neste último conjunto, duas variáveis não participaram com destaque na composição dos fatores: umidade relativa e velocidade dos ventos.

Portanto, nos meses mais quentes e com maiores precipitações há necessidade de se aumentar a quantidade de cloro, pois o mesmo se perde por evaporação em temperaturas mais elevadas, ficando mais propícios à contaminação por bactérias heterotróficas. Estas conclusões confirmam o procedimento adotado atualmente de se aumentar o cloro no verão e ressaltam a necessidade de alterações nas concentrações de cloro de acordo a temperatura ambiente e os níveis de precipitação.

8 REFERÊNCIAS

- ALMEIDA, C. C.; FERNANDES, C. V. S.; SIECIECHOWICZ, M. S. F. Análise estatística multivariada em gestão de recursos hídricos: estudo de caso da bacia do Alto Iguaçu. In: SIMPÓSIO BRASILEIRO DE RECURSOS HÍDRICOS, 20., 2013, Bento Gonçalves, RS. **Anais do XX Simpósio Brasileiro de Recursos Hídricos**: Associação Brasileira de Recursos Hídricos, 2013. p. 1-8.
- AYOADE, J. O. **Introdução à climatologia para os trópicos**. 10. ed. Rio de Janeiro: Bertrand Brasil, 2004. 332 p.
- BAENA, L. G. N. et al. Programa computacional para geração de séries sintéticas de dados climáticos. **Engenharia na Agricultura**, Viçosa, MG, v.13, n.3, p.210-220, jul./set. 2005.
- BALLING JR., R. C.; GOBER, P. Climate Variability and residential Water use in the city of Phoenix, Arizona. **Journal of applied meteorology and climatology**, Boston, MA, v.46, p.1130-1137. 2006. Disponível em: <<http://journals.ametsoc.org/doi/pdf/10.1175/JAM2518.1>> Acesso em: 06 out. 2011.
- BARBOSA, D. A.; LAGE, M. M.; BADARÓ, A. C. L. Qualidade Microbiológica da Água dos Bebedouros de um Campus Universitário de Ipatinga, Minas Gerais. **Nutrir Gerais**, Ipatinga, MG, v.3, n 5, p.505-517, ago. 2009.
- BETTEGA, J. M. P. R. et al. Métodos analíticos no controle microbiológico da água para consumo humano. **Ciência agrotecnologia**, Lavras, MG, v.30, n.5, p. 950-954, set./out. 2006.
- BISCARO, G. A. **Meteorologia Agrícola Básica**. v.1, Cassilândia, MS: UNIGRAF, 2007. 87p. Disponível em <<http://www.do.ufgd.edu.br/guilhermebiscaro/arquivos/meteorologia.pdf>>. Acesso em: 08 jan. 2014.

BRASIL. Lei Federal n. 9433/97. **Política Nacional de Recursos Hídricos**. Brasília, DF, 1997. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/19433.htm>. Acesso em: 20 abr. 2012.

BRASIL. Ministério da Saúde. Portaria nº 518 de 25 de março de 2004. Estabelece os procedimentos e responsabilidades relativas ao controle e vigilância da qualidade da água para o consumo humano e seu padrão de potabilidade e dão outras providências. **Diário Oficial da República Federativa do Brasil**. Brasília, DF, p. 266-9, 26 de mar. 2004. Seção 1.

BRASIL. Resolução Conselho Nacional do Meio Ambiente - CONAMA nº 357, de 17 de março de 2005. Dispõe sobre a classificação dos corpos de água e diretrizes ambientais para o seu enquadramento, bem como estabelece as condições e padrões de lançamento de efluentes, e da outras providências. **Diário Oficial da União**, Brasília, DF, n. 53, p. 58-63, 18 mar. 2005. Seção 1.

BRASIL. Ministério do Meio Ambiente. Secretaria de Recursos Hídricos. **Programas nacionais e metas**. Brasília, DF. 2006a. (Plano Nacional de Recursos Hídricos, v. 4). Disponível em: <http://www.mma.gov.br/estruturas/161/_publicacao/161_publicacao03032011025031.pdf>. Acesso em: 02 maio 2013.

_____. **Panorama e estado dos recursos hídricos do Brasil**. Brasília, DF, 2006b. (Plano Nacional de Recursos Hídricos, v. 1). Disponível em: <http://www.mma.gov.br/estruturas/161/_publicacao/161_publicacao03032011025312.pdf>. Acesso em: 02 maio 2013.

_____. **Vigilância e controle da qualidade da água para consumo humano**. Brasília, DF, 2006c. 221p. (Série B: Textos básicos de Saúde). Disponível em: <http://bvsms.saude.gov.br/bvs/publicacoes/vigilancia_controle_qualidade_agua.pdf>. Acesso em: 20 dez 2014.

BRASIL. Ministério da Saúde. Portaria nº 2914 de 14 de dezembro de 2011. Dispõe sobre os procedimentos de controle e de vigilância da qualidade da água para o consumo humano e seu padrão de potabilidade. **Diário Oficial da República Federativa do Brasil**. Brasília, DF, p. 39, 14 de dez. 2011. n. 239. Seção 1.

BRASIL. Ministério do Meio Ambiente. Secretaria de Recursos Hídricos. **Água: um recurso cada vez mais ameaçado**. Brasília. DF. Disponível em: <http://www.mma.gov.br/estruturas/sedr_proecotur/_publicacao/140_publicacao09062009025910.dpf>. Acesso em: 10 abr. 2013.

BUENO, L. F.; GALBIATTI, J. A.; BORGES, M. J. Monitoramento de variáveis de qualidade da água do horto Ouro Verde - Conchal - SP. **Engenharia Agrícola**, Jaboticabal, SP, v.25, n.3, p.742-748, set/dez. 2005. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-69162005000300020&lng=pt&nrm=iso&tlng=pt>. Acesso em: 15 abr. 2013.

BUFON, A. G. M.; LANDIM, P. M. B. Análise da qualidade da água por metodologia estatística multivariada na Represa Velha (CEPTA/IBAMA/PIRASSUNUNGA/SP). **Holos**

Environment, Rio Claro, SP, v.7, n 1, p.42-59, 2007. Disponível em: <<http://www.periodicos.rc.biblioteca.unesp.br/index.php/holos/article/view/972>>. Acesso em: 24 ago. 2013.

CAMPOS, L. P. **Perfil geográfico multivariado da água consumida no município de Botucatu**. 2014. 71 f. Dissertação (Mestrado em Biometria), Instituto de Biociência, Botucatu, Universidade Estadual Paulista, Botucatu, SP. 2014. Disponível em <<http://base.repositorio.unesp.br/bitstream/handle/11449/110367/000788256.pdf?sequence=1&isAllowed=y>>. Acesso em 27 nov. 2014.

CAMPOS, J. A. D. B., FARACHE FILHO, A., FARIA, J. B. Qualidade sanitária da água distribuída para consumo humano pelo sistema de abastecimento público da cidade de Araraquara-SP, **Alimentos e Nutrição**. Araraquara, SP, vol. 13, p.117-129, 2002.

_____. Qualidade da água armazenada em reservatórios domiciliares: Parâmetros Físico-Químicos e Microbiológicos. **Alimentos e Nutrição**. Araraquara, SP, v.14, n.1, p. 63-67. 2003.

CARDIM, M. **Mapeamento do comportamento multivariado das principais variáveis climáticas de interesse agrícola do Estado de São Paulo**. 2001. 124 f. Tese (Doutorado em Agronomia - Energia da Agricultura) - Faculdade de Ciências Agrônômicas, Universidade Estadual Paulista, Botucatu, SP. 2001.

CASELANI, K. **Percepção de risco sanitário na bovinocultura de corte no estado de São Paulo**. 2014. 212 f. Tese (Doutorado - Programa de Pós-graduação em Ciências Agrárias e Veterinárias), Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal, SP, 2014.

CONEGO, J. G. L., A outorga de usos da água como instrumento de gerenciamento dos recursos hídricos. **Revista de Administração Pública**, Rio de Janeiro, RJ, v.27, n.2, p.28-62, abr./jun. 1993. Disponível em<<http://www.spell.org.br/documentos/ver/14000/a-outorga-de-usos-da-agua-como-instrumento-de-gerenciamento-dos-recursos-hidricos>>. Acesso em: 15 abr. 2012.

CUNHA, A. R., MARTINS, D. Classificação climática para os municípios de Botucatu e São Manuel, SP. **Irriga**, Botucatu, SP. v.14, n.1, p. 1-11, jan/mar. 2009. Disponível em: <<http://irriga.fca.unesp.br/index.php/irriga/article/view/1909>>. Acesso em: 13 out. 2014.

DIAS, D. M. **Avaliação do impacto da renda sobre o consumo hidrometrado de água em domicílios residências urbanos: um estudo de caso para regiões de Belo Horizonte-MG**. 2008. 152 f. Dissertação (Mestrado em Saneamento, Meio Ambiente e Recursos Hídricos) - Universidade Federal de Minas Gerais, Belo Horizonte, MG. 2008. Disponível em < <http://dspace.lcc.ufmg.br/dspace/bitstream/1843/REPA-7PTJ35/1/david336.pdf>> Acesso em 03 out. 2011.

DOMINGUES, V. O. et al. Contagem de bactérias heterotróficas na água para consumo humano: comparação entre duas metodologias. **Revista Saúde (Santa Maria)**, Santa Maria,

RS, v. 33, n 1, p 15-19, 2007. Disponível em: <<http://cascavel.ufsm.br/revistas/ojs-2.2.2/index.php/revistasauade/article/view/6458>>. Acesso em: 20 out. 2014.

EATON, A. D., CLESCERI, L. S., RICE, E. W., GREENBERG, A. E. (Coord.). **Standard methods for examination of water and wastewater**. 21 ed., Washington-DC: American Public Health Association (APHA); American Water Works Association (AWWA); Water Environment Federation (WEF), 2005.

ESTEVEZ, F. A., **Fundamentos de limnologia**, 3 ed. Rio de Janeiro: Interciência, 2011.

FELIPE, L. C. T. **Resultados ETA**. [mensagem pessoal] Mensagem recebida por: <lfelipe@sabesp.com.br> em: 18 nov. 2014.

FELSKI, G.; ANAISSI, F. J.; QUINÁIA, S. P. Avaliação da qualidade da água consumida pela população do município de Guarapuava, Paraná. **Revista Eletrônica Lato Sensu**. Paraná. Ano 3, v.1, p.1-25, mar. 2008.

FREIRE, R. C. Qualidade da água nos reservatórios domiciliares na região metropolitana da cidade do Recife-PE. **Journal of Management and Primary Health Care**, Recife, PE, v.3, n.2, p.102-105, 2012. Disponível em: <<http://www.jmphc.com/ojs/index.php/01/article/viewArticle/59>>. Acesso em: 20 nov. 2014.

FREIRE, A. P., CASTRO, E. C. Análise da Correlação do uso e Ocupação do Solo e da Qualidade da Água. **RBRH-Revista Brasileira de Recursos Hídricos**. Porto Alegre, RS, v.19, n.1, p. 41-49, jan./mar. 2014. Disponível em: <http://www.abrh.org.br/SGCv3/UserFiles/Sumarios/c720cf8418e55a3f8a9c4280dafd717b_90f8621cd34013b4293a537120beb593.pdf>. Acesso em: 20 nov. 2014.

FURLAN, D. N. **Variabilidade temporal e espacial das chuvas e do balanço hídrico no estado de Rondônia: caracterização e análise de tendência**. 2009. 129 f. Dissertação (Mestrado em Agronomia – Área de concentração: Física do ambiente agrícola), Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo, Piracicaba, SP, 2009.

GHERARDI, B., DOURADO NETO, D., SENTELHAS, P. C. Caracterização espacial e temporal da temperatura média do ar, em escala diária e mensal, no Estado de Mato Grosso. **Revista Brasileira de Agrometeorologia**, Santa Maria, RS, v.16, n.2, p.111-120, 2008. Disponível em: <<http://www.sbagro.org.br/rbagro/ojs/index.php/rbagro/article/view/92>>. Acesso em: 20 nov. 2014.

HAIR JUNIOR, J. F. et al. **Análise Multivariada de dados**. 6ª ed. Porto Alegre: Bookman, 2009, 688p.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**, 6.ed. New Jersey: Prentice-Hall, 2007, 773 p.

MACEDO, J. **Águas & Águas**. Belo Horizonte: CRQ-MG, 2004. 977 p.

MARDIA, K.V. Assessment of multinormality and the robustness of Hotelling T^2 test. **Applied Statistics**, v.24, p.163-171, 1974.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada**: uma abordagem aplicada. Belo Horizonte, MG: Editora UFMG, 2005. 297 p.

MORENO, J. **Avaliação e gestão de riscos no controle da qualidade da água em redes de distribuição: Estudo de caso**. 2009. 579 f. Tese (Doutorado- Programa de Pós-graduação e Área de Concentração em Engenharia Hidráulica e Saneamento), Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, SP, 2009.

OLIVEIRA, V. P. S. **Modelo para a geração de séries sintéticas de precipitação**, 2003. 156 f. Tese (*Doctor Scientiae* em Engenharia Agrícola) - Universidade Federal de Viçosa, Viçosa, MG. 2003. Disponível em: <ftp://ftp.ufv.br/dea/GPRH/teses/ds_oliveira/Tese_Doutorado_Vicente_de_Paulo_Santos_de_Oliveira.pdf> Acesso em: 03 out. 2011.

OLIVEIRA, P.R.T., PADOVANI, C. R. Utilização de análise de agrupamento na caracterização do perfil agrícola e climático da região sudeste do estado de Mato Grosso. **Revista Energia na Agricultura**, Botucatu, SP. v.25, n 4: p 97-109, 2010. Disponível em: <<http://energia.fca.unesp.br/index.php/energia/article/view/93>>. Acesso em: 23 dez. 2013.

OLIVEIRA, A. S. et al. Qualidade da água para consumo humano distribuída pelo sistema de abastecimento público em Guarabira - PB. **Revista Verde de Agroecologia e Desenvolvimento Sustentável**, Mossoró, RN, v. 7, n. 2, p.199-205, 2012 (a). Disponível em <http://www.gvaa.com.br/revista/index.php/RVADS/article/viewFile/1230/pdf_478>

_____. Qualidade da água para consumo humano ofertada na cidade de Limoeiro do Norte, Ceará. **Tecnologia & Ciência Agropecuária**. João Pessoa, PB, v.6, n.2, p.1-5, 2012 (b). Disponível em: <http://www.emepa.org.br/revista/volumes/tca_v6_n2_jun/tca6201.pdf>. Acesso em: 23 dez. 2013.

PINTO, U., MAHESWARI, B. L. River health assesment in peri-urban landscapes: na application of multivariate analysis to identify the key variables. **Water Research**, New York, US, v.45, p.3915-3924, 2011.

PRATHUMRATANA, L., STHIANNOPKAO, S., KIN, K. W. The relationship of climatic and hydrological parameters to surface water quality in the lower Mekong River. **Environment International**, New York, US, v. 34, p.860-866, 2008.

QUALIDADE da Água. Disponível em: <<http://site.sabesp.com.br/site/interna/Default.aspx?secaoId=493>>. Acesso em: 15 ago. 2012.

QUEIROZ, J. T. M., HELLER, L., SILVA, S. R. Análise da correlação de ocorrência de doença diarreia aguda com a qualidade da água para consumo humano no município de Vitória-ES. **Saúde Sociedade**. São Paulo, v.18, n.3, p.479-489, jul./set. 2009. Disponível

em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-12902009000300012>. Acesso em: 22 mar. 2013.

REBOUÇAS, A. C., BRAGA JUNIOR, B. J., TUNDISI, J. G. **Águas doces no Brasil: capital ecológico, uso e conservação**. 3ed. São Paulo: Escrituras Editora, 2002. 768 p.

RIBEIRO, M. C. M. Nova portaria de potabilidade de água: busca de consenso para viabilizar a melhoria da qualidade de água potável distribuída no Brasil. **Revista DAE**, São Paulo, n 189: p 8-14, maio/ago. 2012.

RICKLI, L. I. **Eficiência de estimadores, geradores e algoritmos na simulação de dados diários de precipitação pluviométrica utilizando a distribuição Gama**. 2006. 84 f. Tese (Doutorado em Agronomia - Energia da Agricultura) - Faculdade de Ciências Agrônomicas, Universidade Estadual Paulista, Botucatu, SP. 2006.

SABINO, C. V. S.; ABREU, J. F.; LOBATO, W.; SABINO, G. S.; KNUP, E. A. N. Análise de alguns aspectos da qualidade da água da Bacia do Rio Paraopeba utilizando estatística Multivariada. **Revista de Biologia e Ciências da Terra**. João Pessoa, PB, v.8, n.2, p.6-18, 2008. Disponível em: <<http://eduep.uepb.edu.br/rbct/sumarios/pdf/estatisticamultivariada.pdf>> Acesso em: 03 out. 2011.

SANDRE, L. C. G. et al. Influência dos fatores climáticos na qualidade da água em pesques pagues. **Veterinária e Zootecnia**, Botucatu, SP, v.16, n.3, p.509-518, 2009. Disponível em: <[http://www.fmvz.unesp.br/revista/volumes/vol16_n3/VZ16_3\(2009\)_509-518.pdf](http://www.fmvz.unesp.br/revista/volumes/vol16_n3/VZ16_3(2009)_509-518.pdf)> Acesso em: 06 out. 2011.

SETTI, A. A. et al. **Introdução ao gerenciamento de recursos hídricos**. 2 ed. Brasília: Agência Nacional de Energia Elétrica, Superintendência de Estudos e Informações Hidrológicas, 2001. 207 p. Disponível em: <http://www.comitepcj.sp.gov.br/download/livro_Introd-Gerenc-Rec-Hidr.pdf>. Acesso em: 20 set. 2012.

SILVA, J. W.; GUIMARÃES, E. C.; TAVARES, M. Variabilidade temporal da precipitação mensal e anual na estação climatológica de Uberaba-MG. **Ciência e Agrotecnologia**. Lavras, MG, v.27, n. 3, p.1-17, maio/jun. 2003. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-70542003000300023&lng=pt&nrm=iso&tlng=pt>. Acesso em: 26 abr. 2012.

SILVA, E. S. **Variabilidade da precipitação pluviométrica nas regiões sudeste e sul do Brasil**. 2006. 188 f. Tese (Doutorado em Agronomia-Energia da Agricultura) - Faculdade de Ciências Agrônomicas, Universidade Estadual Paulista, Botucatu, SP. 2006.

SILVA, A. E. P.; ANCELIS, C. F.; MACHADO, L. A. T.; WAICHAMAN, A. V. Influência da precipitação na qualidade da água do rio Purus. **Acta Amazônica**, Manaus, AM, v.38, n.4, p.733-742, 2008 (a). Disponível em: <<http://www.scielo.br/pdf/aa/v38n4/v38n4a17.pdf>> Acesso em: 03 out. 2011.

SILVA, W. T. P.; SILVA, L. M.; CHICHORRO, J. F. Gestão de recursos hídricos: perspectivas do consumo per capita de água em Cuiabá. **Engenharia Sanitária e Ambiental**, Rio de Janeiro, v.13, n.1, 2008 (b). Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-41522008000100002> Acesso em: 03 out. 2011.

SILVA, A.P.S., DIAS, H. C. T., BASTOS, R. K. X., SILVA, E. Qualidade da água do reservatório da usina hidrelétrica (UHE) de Peti, Minas Gerais. **Revista Árvore**. Viçosa, MG, v.33, n.6, p. 1063-1069, 2009. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622009000600009&lng=en&nrm=iso>. Acesso em: 03 out. 2013.

SIQUEIRA, L. P. et al. Avaliação microbiológica da água de consumo empregada em unidades de alimentação. **Ciência & Saúde Coletiva**, Rio de Janeiro, RJ, v.15, n. 1, jan. 2010. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232010000100011>. Acesso em: 18 jun. 2013.

SOUZA, J. R. et al, A importância da qualidade da água e os seus múltiplos usos: caso do Rio Almada, sul da Bahia, Brasil. **Revista eletrônica do Prodepa**, Fortaleza, CE, v.8, n.1, p.26-45, abr. 2014. Disponível em: <<http://www.revistarede.ufc.br/revista/index.php/rede/article/viewArticle/217>>. Acesso em: 20 jan. 2015.

TAVARES, D. S., et al., Qualidade da água de bicas localizadas nos municípios de Santos e São Vicente, Estado de São Paulo, Brasil. **Revista do Instituto Adolfo Lutz**, São Paulo, v.68, n.2, p.237-244, 2009. Disponível em: <http://periodicos.ses.sp.bvs.br/scielo.php?script=sci_arttext&pid=S0073-98552009000200010&lng=pt&nrm=iso>. Acesso em: 20 dez. 2014.

THE COMPREHENSIVE R ArchiveNetwork. Disponível em: <<http://cran.r-project.org/>>. Acesso em: 10 mar. 2012.

TOLEDO, L. G.; NICOLELLA, G. Índice de qualidade de água em Microbacia sob uso agrícola e urbano. **Scientia Agrícola**, Piracicaba, SP, v.59, n.1, p.181-186, 2002. Disponível em: <<http://www.scielo.br/pdf/sa/v59n1/8092.pdf>>. Acesso em: 04 out. 2011.

TUNDISI, J. G. **Água no século XXI: enfrentando a escassez**. São Paulo: Editora Rima, 2003, 247 p.

UNITED NATIONS. **The volume of freshwater resources on Earth is around 35 million km³**. Disponível em <<http://www.unwater.org/statistics/statistics-detail/en/c/211801/>>. Acesso em: 10 dez. 2014.

VAREJÃO-SILVA, M. A., **Meteorologia e Climatologia**. Versão Digital 2. Recife, PE, 2006, 449 p. Disponível em: <www.agritempo.gov.br/publish/publicacoes/livros/METEREOLOGIA_E_CLIMATOLOGIA_VD2_Mar_2006.pdf>. Acesso em: 08 jan. 2014.

VIRGENS FILHO, J. S. **Ferramenta computacional para simulação de séries climáticas diárias, baseada na parametrização dinâmica das distribuições de probabilidade.** 2001. 93 f. Tese (Doutorado em Agronomia - Energia na Agricultura) - Faculdade de Ciências Agrônômicas, Universidade Estadual Paulista, Botucatu, SP. 2001.

VIRGENS FILHO, J. S. et al. Modelo computacional estocástico para simulação de séries climáticas diárias de umidade relativa do ar, baseada na parametrização dinâmica das distribuições de probabilidade decorrente da retroalimentação de dados. **Revista Brasileira de Climatologia**, Presidente Prudente, SP, v.5, n.1. p.133-151. 2009.

VON SPERLING, M. **Princípios do tratamento biológico de águas residuárias:** Introdução a qualidade das águas e ao tratamento de esgotos. 3 ed., Belo Horizonte: Departamento de Engenharia Sanitária e Ambiental - Universidade Federal de Minas Gerais, 2005. 542 p.

ZANETTI, S. S. et al. Programa computacional para geração de séries sintéticas de precipitação. **Engenharia Agrícola**. Jaboticabal, SP, v. 25, n. 1, p.96-104, jan./abr. 2005. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-69162005000100011>. Acesso em: 20 fev. 2012.

ZANETTI, S. S.; OLIVEIRA, V. P. S.; PRUSKI, F. F. Validação do modelo ClimaBR em relação ao número de dias chuvosos e a precipitação total diária. **Engenharia Agrícola**, Jaboticabal, SP, v. 26, n. 1, p.96-102, jan./abr. 2006. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-69162006000100011>. Acesso em: 20 fev. 2012.

APÊNDICES

APÊNDICE 1. Evolução mensal dos dados

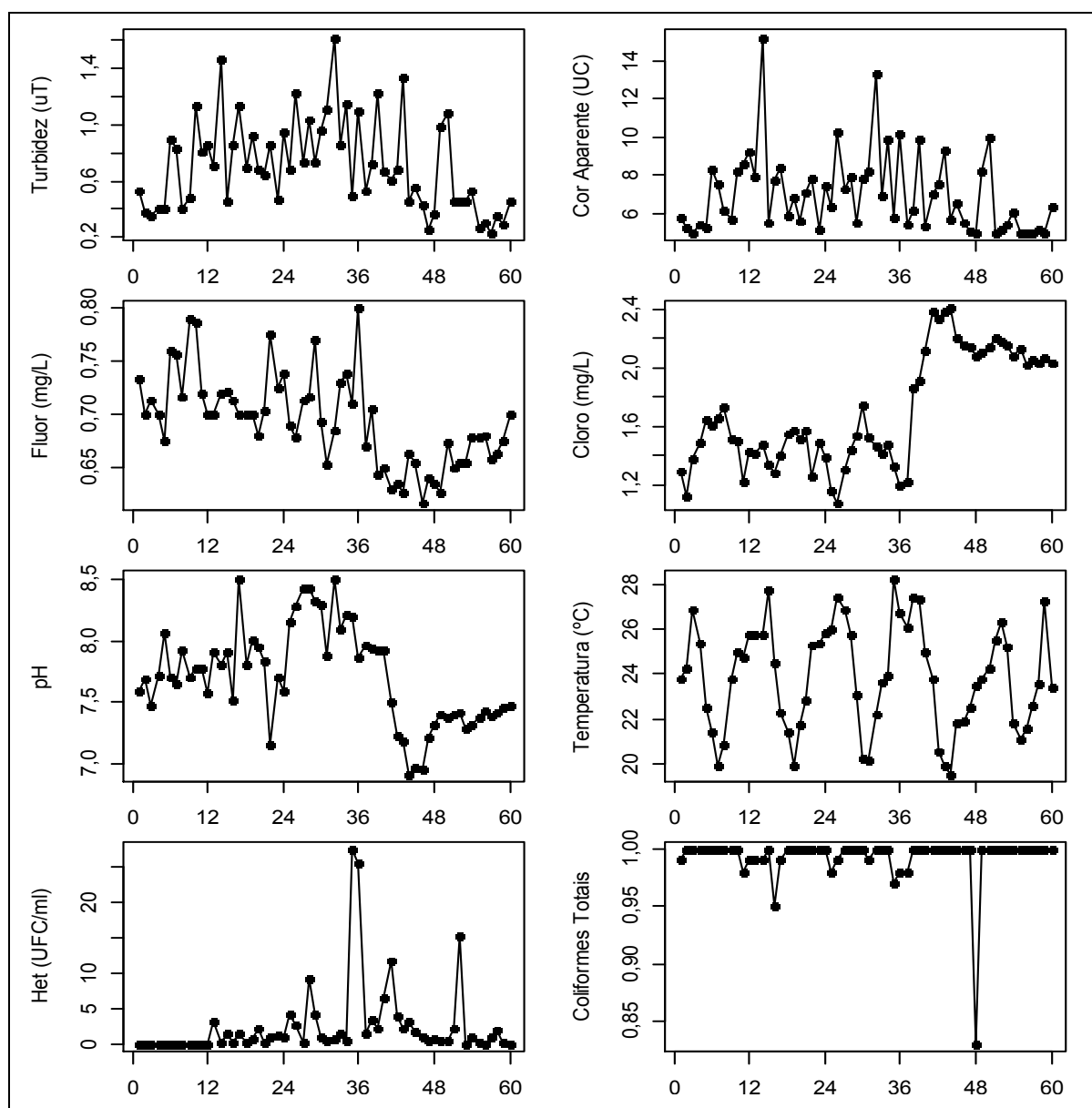


Figura 24. Evolução mensal dos dados das variáveis de qualidade da água na rede de distribuição entre janeiro de 2007 a dezembro de 2011

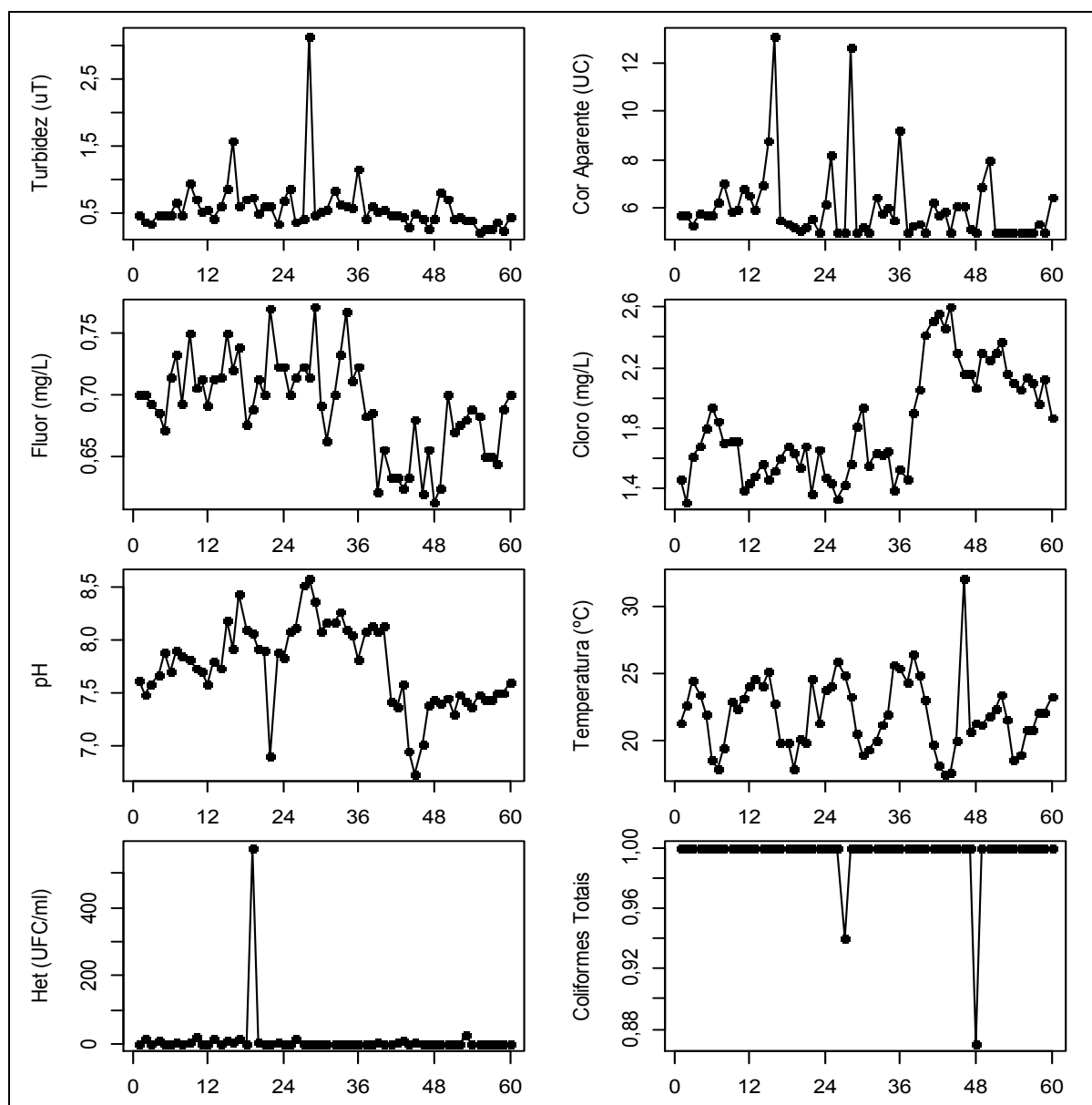


Figura 25. Evolução mensal dos dados das variáveis de qualidade da água no final de tratamento entre janeiro de 2007 a dezembro de 2011

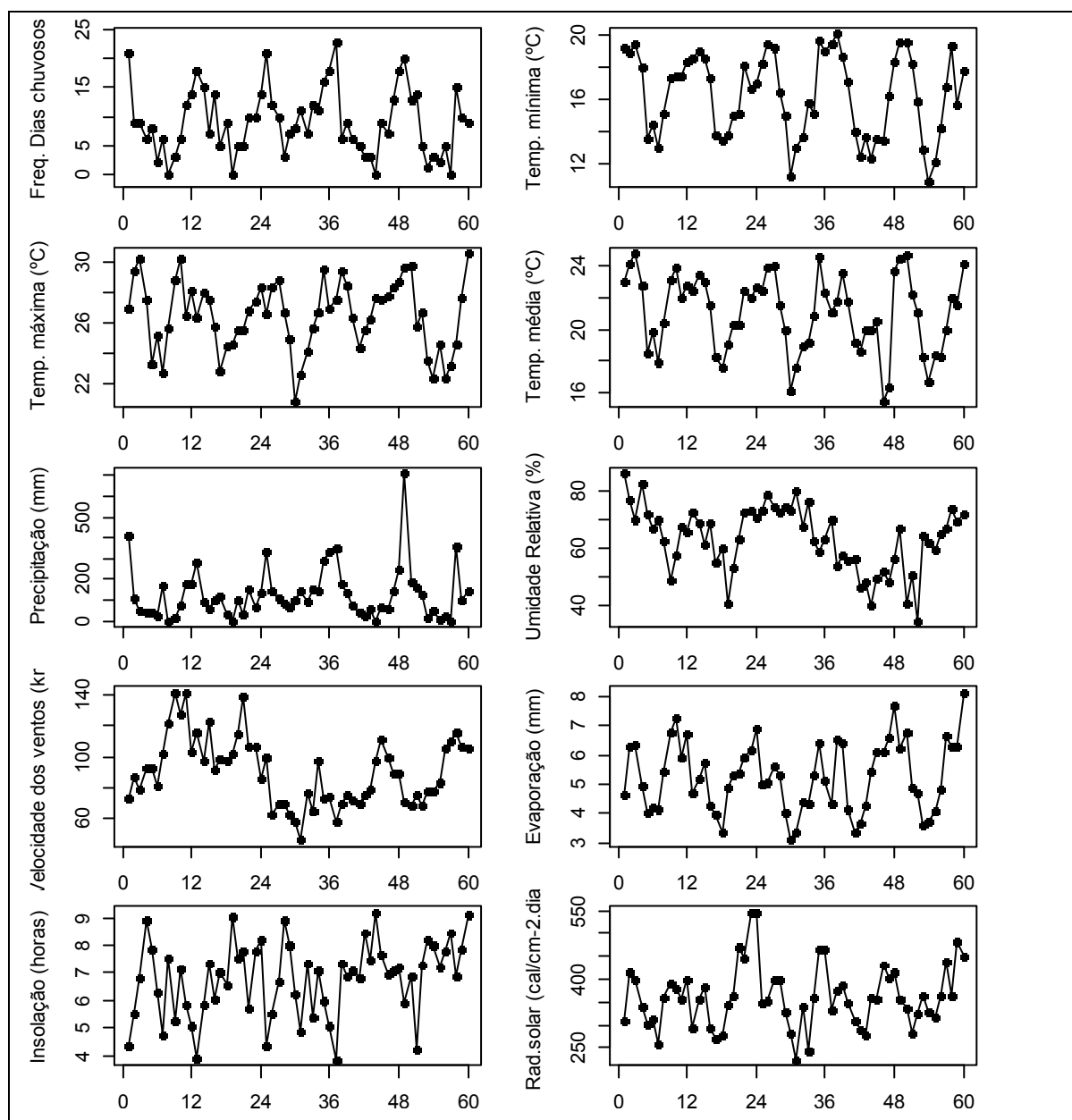


Figura 26. Evolução mensal dos dados das variáveis climáticas entre janeiro de 2007 a dezembro de 2011

APÊNDICE 2. Médias mensais

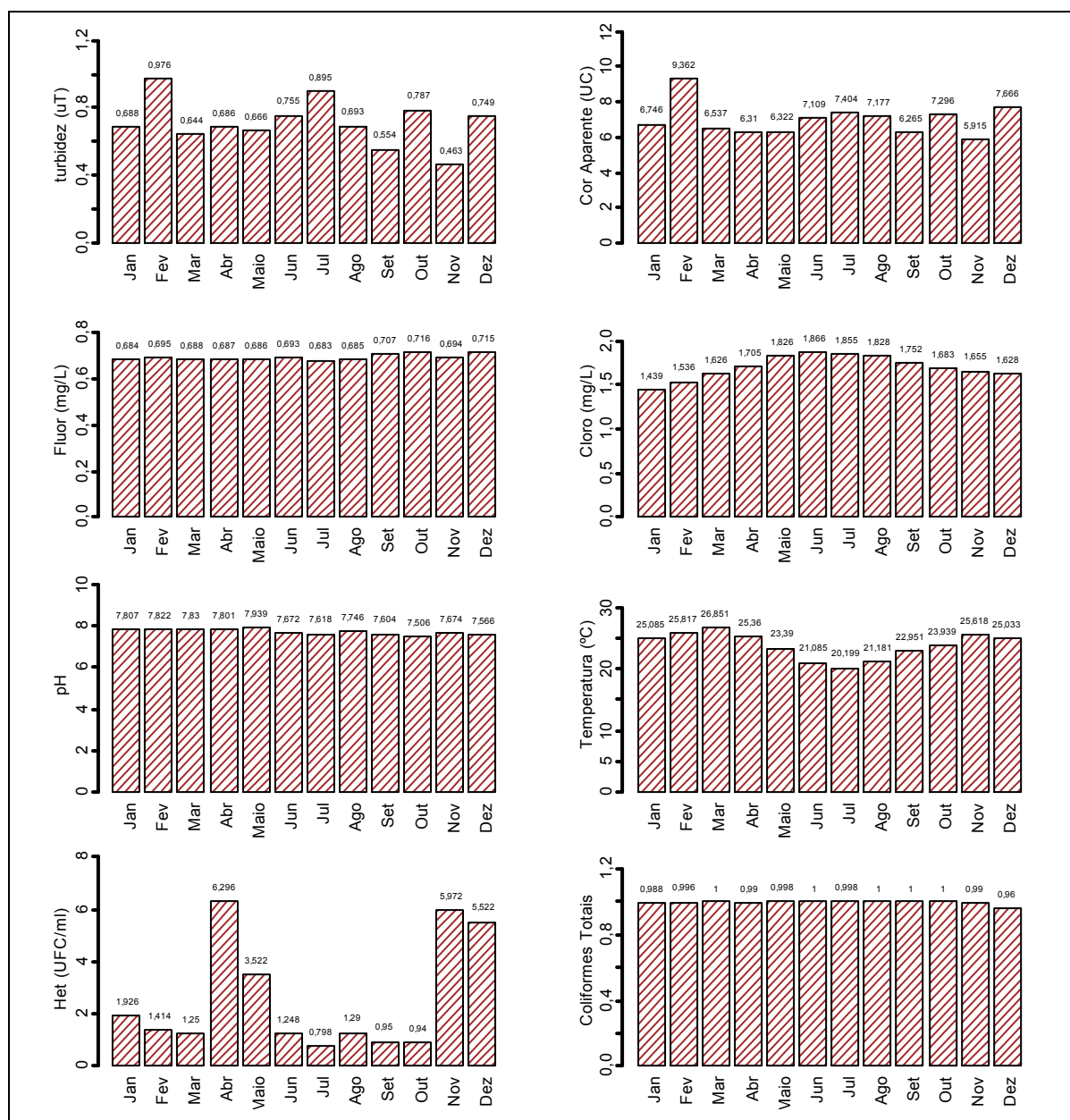


Figura 27. Médias mensais das variáveis de qualidade da água considerando a rede de distribuição

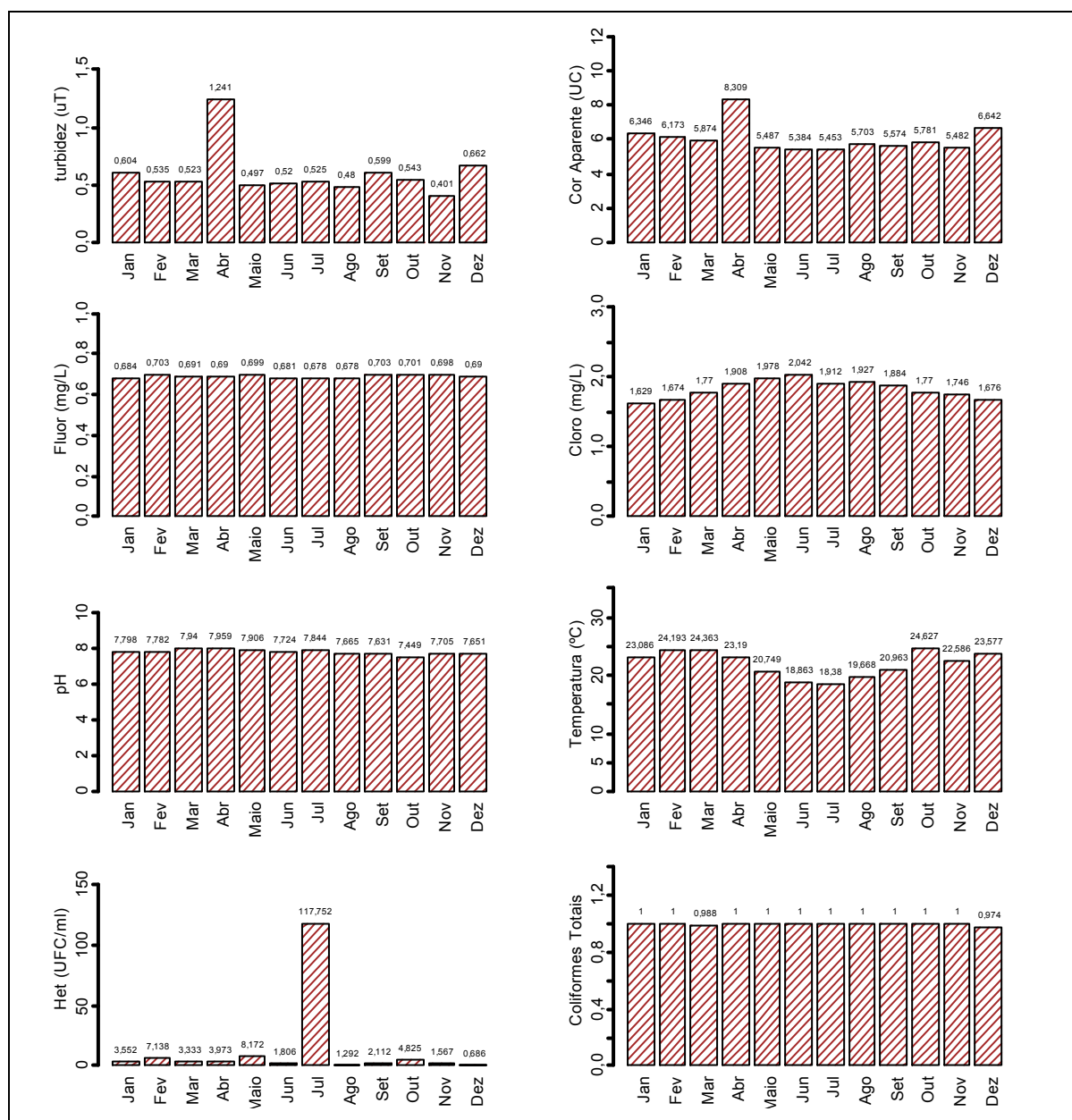


Figura 28. Médias mensais das variáveis de qualidade da água considerando o final de tratamento

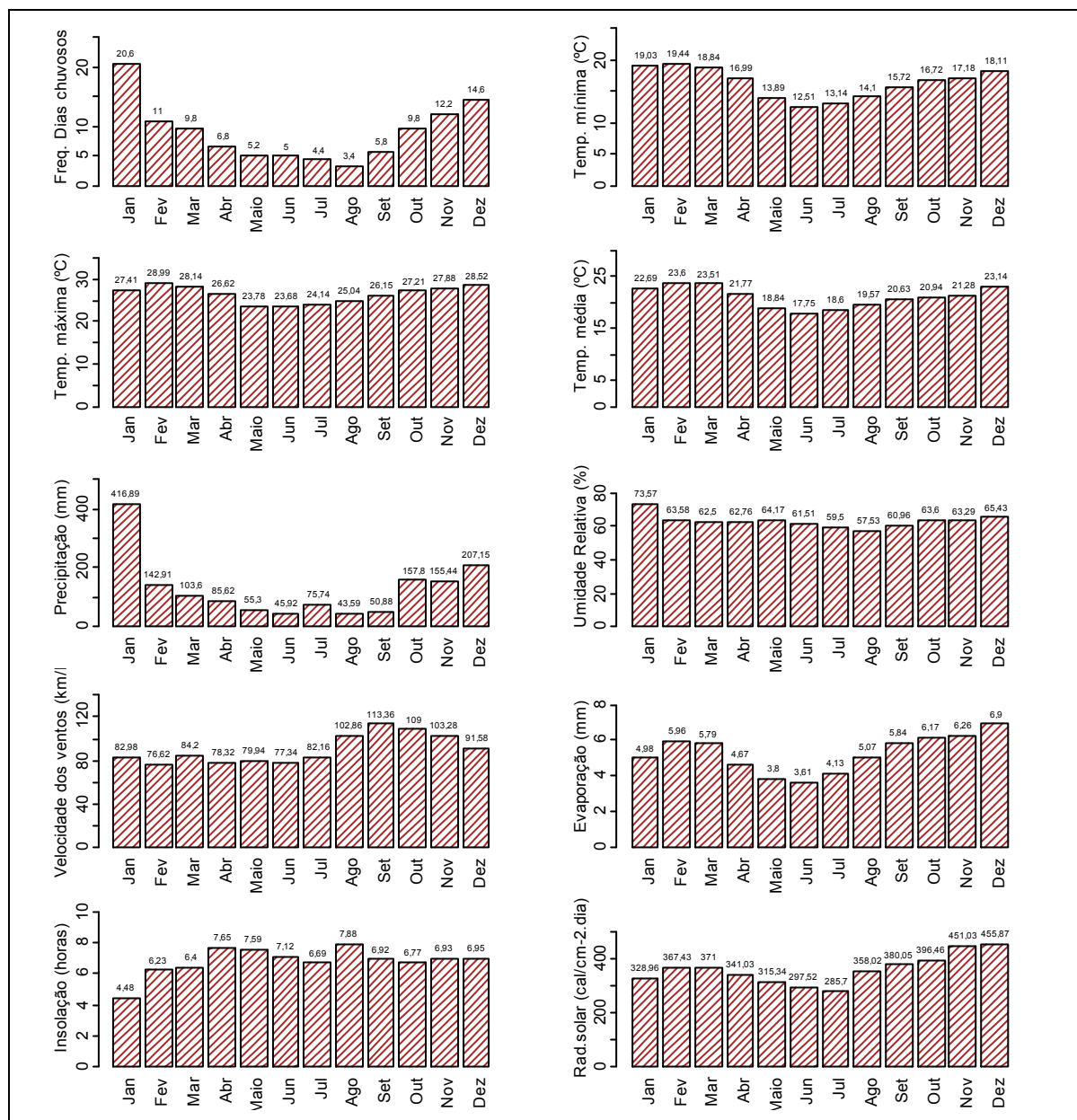


Figura 29. Médias mensais das variáveis climáticas

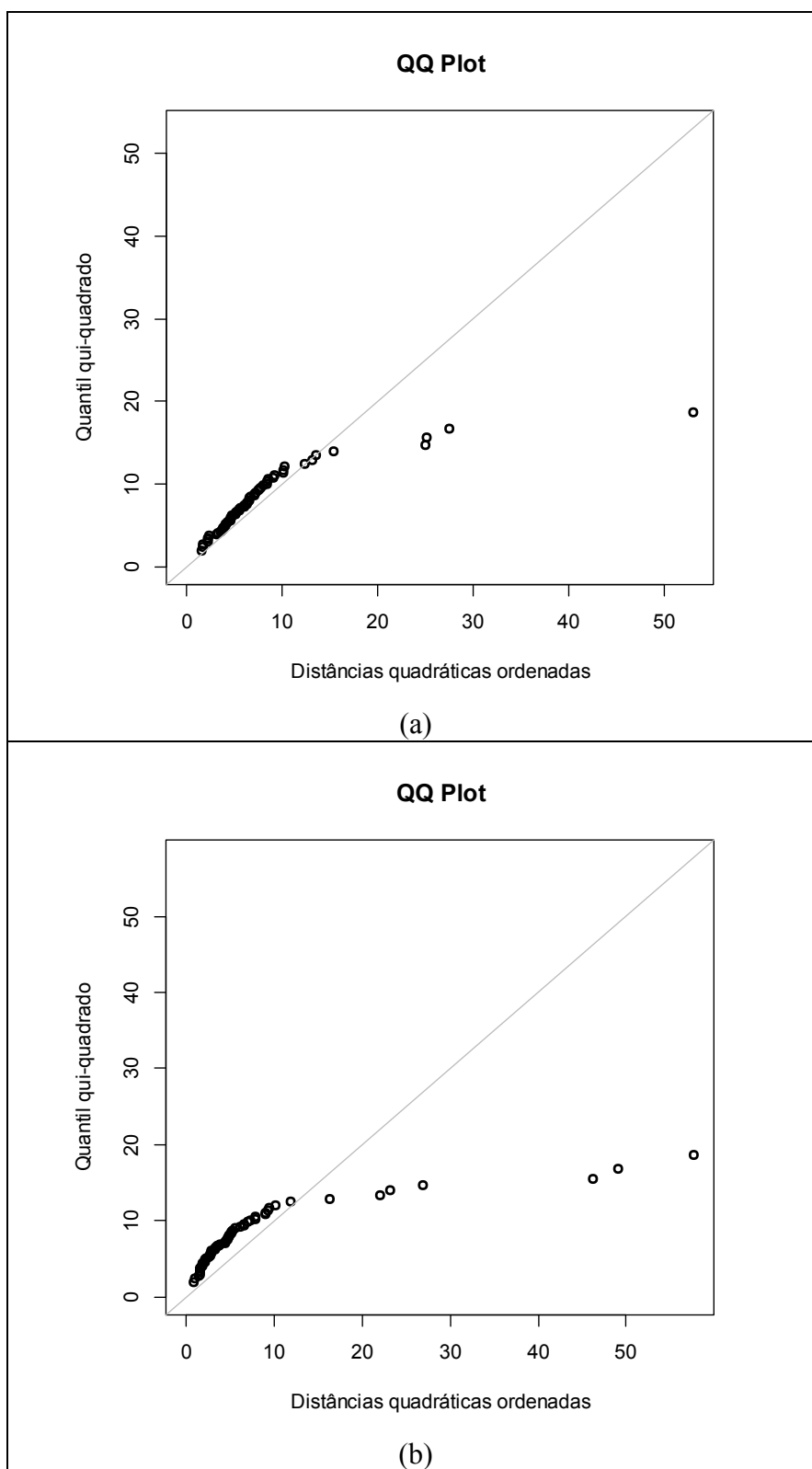
APÊNDICE 3. Gráficos *QQ Plot* Multivariados

Figura 30. Gráfico de probabilidade *QQ-Plot* multivariado para os dados de qualidade da água na rede de distribuição (a) e no final de tratamento (b)

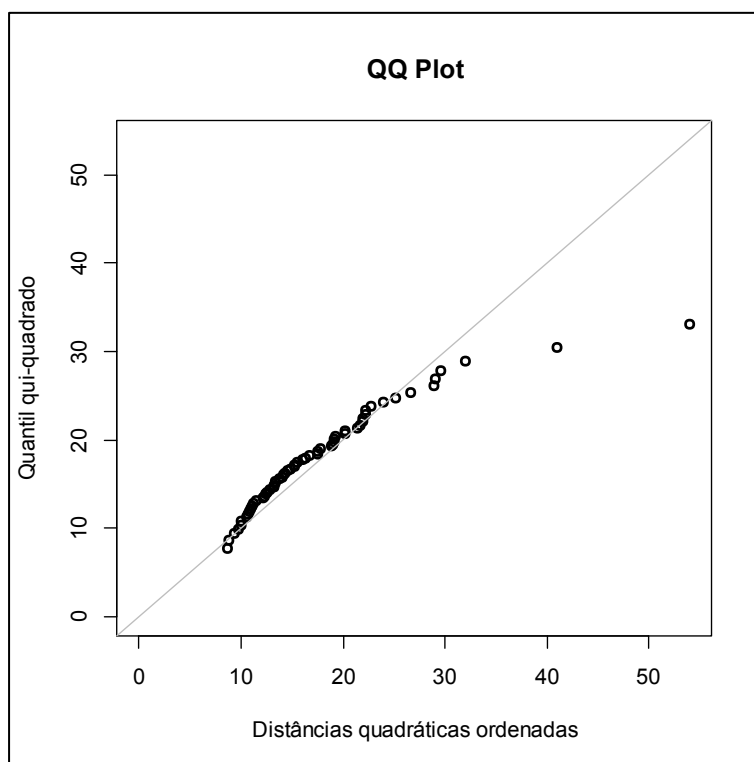


Figura 31. Gráfico de probabilidade *QQ-Plot* multivariado para os dados de qualidade da água na rede de distribuição e climáticos

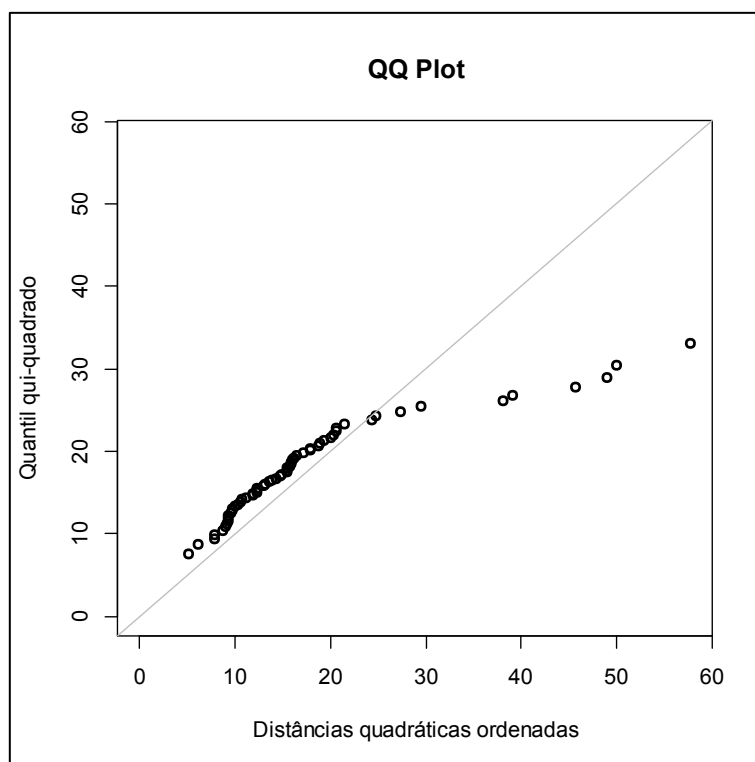


Figura 32. Gráfico de probabilidade *QQ-Plot* multivariado para os dados de qualidade da água no final de tratamento e climáticos