



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO
Diretoria de Engenharia e Qualidade Ambiental

RELATÓRIO TÉCNICO FINAL
“ESTATÍSTICA APLICADA À GESTÃO DA QUALIDADE
DAS ÁGUAS SUPERFICIAIS E SUBTERRÂNEAS”

CONTRATO FEHIDRO 139/2012 – CORHI 124
TOMADOR: CETESB – COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO

OUTUBRO / 2016



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO
Diretoria de Engenharia e Qualidade Ambiental

Diretor Presidente
Carlos Roberto dos Santos

Diretor de Gestão Corporativa
Carlos Roberto dos Santos

Diretor de Controle de Poluição Ambiental
Geraldo do Amaral Filho

Diretora de Avaliação de Impacto Ambiental
Ana Cristina Pasini da Costa

Diretor de Engenharia e Qualidade Ambiental
Eduardo Luis Serpa

Gerente do Departamento de Qualidade Ambiental
Maria Helena R. B. Martins



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO
Diretoria de Engenharia e Qualidade Ambiental

UNIDADES RESPONSÁVEIS PELA EXECUÇÃO DO PROJETO

Divisão de Programas e Projetos de Qualidade Ambiental - Coordenação
Lilian Barrella Peres

Setor de Programas e Ações Institucionais
Roberto Xavier De Oliveira

Divisão de Qualidade das Águas e do Solo
Nelson Menegon Júnior

Setor de Águas Subterrâneas e do Solo
Rosângela Pacini Modesto

Setor de Águas Interiores
Fabio Netto Moreno

Setor de Hidrologia
Luis Altivo Carvalho Alvim

Setor de Águas Litorâneas
Claudia Condé Lamparelli



INTRODUÇÃO

O presente relatório tem a finalidade de apresentar as ações decorrentes da execução total do contrato de financiamento FEHIDRO 139/2012, cujo objetivo principal foi implementar rotinas de tratamento e interpretação dos dados gerados nas redes de monitoramento da CETESB, bem como capacitar suas equipes em estatística aplicada, com vistas à condução de estudos, voltados para a gestão da qualidade ambiental e de recursos hídricos, relacionados ao comportamento das variáveis de qualidade ambiental em águas superficiais e subterrâneas.

Para tanto, neste projeto a CETESB adquiriu licença de uso de software de estatística e contratou duas empresas prestadoras de serviços especializados. Uma para o desenvolvimento de um estudo estatístico com base em dados das redes de monitoramento da CETESB e outra empresa para a realização de uma pesquisa de opinião sobre a divulgação do monitoramento de praias litorâneas.

1. ATIVIDADES EXECUTADAS

As atividades executadas e os produtos obtidos no empreendimento em pauta foram os seguintes:

- 1) Levantamento de Necessidades Específicas;
- 2) Aquisição de Licença de Software;
- 3) Pesquisa de Opinião sobre o Monitoramento de Praias;
- 4) Desenvolvimento de Análise Estatística sobre dados das Redes de Monitoramento da Qualidade das Águas da CETESB;
- 5) Capacitação de técnicos da CETESB em estatística aplicada.

A descrição das atividades e produtos encontra-se no Anexo1 deste relatório.

As atividades foram iniciadas em meados de 2012 com um levantamento de necessidades específicas junto às áreas técnicas e terminou em dezembro de 2015 com a entrega final dos estudos estatísticos com base nos dados das redes de monitoramento da qualidade das águas da CETESB.

3. CONSIDERAÇÃO FINAL

No Anexo 1 deste relatório são apresentadas as atividades executadas, mas em resumo considera-se que o desenvolvimento do projeto atendeu seus objetivos, tendo em vista que surgiram subsídios técnicos para tomada de decisão, por meio do conhecimento do comportamento das variáveis de qualidade ambiental em águas superficiais e subterrâneas, a partir dos dados da rede de monitoramento da CETESB, bem como foram capacitadas as equipes e implantada tecnologia aplicada à estatística, com vistas à elaboração de informações para a gestão da qualidade ambiental e de recursos hídricos.



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO
Diretoria de Engenharia e Qualidade Ambiental

4. CONCLUSÃO

Diante do exposto consideram-se concluídas todas as atividades constantes do cronograma físico-financeiro aprovado por esse agente técnico.

São Paulo, 20 de setembro de 2016.

Lilian Barrella Peres
Gerente da Divisão de Programas e Projetos
de Qualidade Ambiental



ANEXO – ATIVIDADES EXECUTADAS

LEVANTAMENTO DE NECESSIDADES ESPECÍFICAS

A partir da assinatura do Contrato Fehidro 139, em junho de 2012, foi realizado um levantamento de necessidades específicas junto às áreas responsáveis pelas redes de monitoramento da Cetesb. Em fevereiro de 2013, com base neste levantamento foi encaminhado ao Agente Técnico/DAEE um plano de trabalho sobre as atividades previstas no empreendimento em pauta.

Neste processo de levantamento verificou-se a necessidade de um estudo de tendência e de correlação de variáveis, bem como a atualização de uma licença de um software de estatística, a fim de aprimorar o tratamento dos dados gerados nos monitoramentos realizados pela Cetesb.

Também foi levantada a necessidade de um estudo específico para o monitoramento de praias. Dessa forma, em maio de 2013, foi solicitada a autorização do Agente Técnico para a realização deste estudo sobre pesquisa de opinião, mediante a readequação do cronograma físico-financeiro do projeto.

A partir da autorização do Agente Técnico foram providenciados os processos licitatórios para a execução dos dois estudos.

AQUISIÇÃO DE LICENÇA DE SOFTWARE

Em dezembro de 2012, foram adquiridas 3 (três) licenças de Upgrade do software IBM SPSS STATISTICS FOR WINDOWS V.10.05 PARA 20.0 com os módulos Statistics Base, Custom Tables, Advanced Statistics, Forecasting e Regression.

As licenças instaladas em computadores da CETESB estão sendo utilizadas por técnicos para o tratamento dos dados de monitoramento de qualidade das águas no estado de São Paulo.

Os módulos adquiridos permitem as seguintes funcionalidades.

IBM SPSS Statistics Base

O IBM SPSS Statistics Base é o software parte integrante dos processos analíticos, fornecendo funcionalidades para acesso, gerenciamento, preparação, análise e relatório de dados. Permite trabalhar de forma confiável com os módulos opcionais e outros produtos da família IBM SPSS Statistics, que fornecem capacidade para planejamento, coleta de dados e implementação e também adição de novas funcionalidades para áreas do IBM SPSS Statistics Base.

Na versão 20 com:

- Mapeamento – Funcionalidade que melhora sua capacidade de direcionar, prever e planejar por região geográfica, assim como expandir suas funcionalidades de relatórios usando modelos de mapas pré-construídos.
- Tabelas mais rápidas – Gera tabelas de saídas completamente interativas e editáveis cerca de cinco vezes mais rápido.
- Suporte de idioma melhorado – A interface disponível em Português (Brasil).



IBM SPSS Custom Tables

Resume e transmite até mesmo os resultados mais complicados num formato tabular usando uma interface altamente intuitiva baseada em arrastar e soltar. É possível girar colunas e tabelas cruzadas.

IBM SPSS Advanced Statistics

Este módulo opcional permite analisar relacionamentos complexos usando poderosas técnicas multivariadas. Na versão 20.0 com:

- Modelos mistos lineares generalizados (GLMM) para dados hierárquicos;
- Modelos lineares gerais (GLM) e procedimentos de modelos mistos;
- Modelos lineares generalizados (GENLIN), incluindo os modelos estatísticos mais utilizados como regressão linear para respostas normalmente distribuídas, modelos logísticos para dados binários e modelos log-lineares para dados de frequência;
- O GENLIN que oferece os modelos estatísticos mais utilizados através da sua formulação geral de modelos;
- Os procedimentos de Equações de Estimação Generalizadas(GEE) que ampliam os modelos lineares generalizados para ajustar dados longitudinais correlacionados e dados agrupados.

IBM SPSS Forecasting

O IBM SPSS Forecasting tem técnicas mais avançadas de estatística para trabalhar com dados de séries temporais, utilizado por especialistas em estatística.

IBM SPSS Regression

O IBM SPSS Regression permite aplicar modelos mais sofisticados aos dados, usando uma grande variedade de modelos de regressão não-lineares.



PESQUISA DE OPINIÃO SOBRE O MONITORAMENTO DE PRAIAS

1. INTRODUÇÃO

Águas recreacionais são águas doces, salobras e salinas destinadas à recreação de contato primário, sendo este definido como um contato direto e prolongado com a água (natação, mergulho, esqui-aquático etc.), no qual, a possibilidade do banhista ingerir quantidades apreciáveis de água é elevada. O contato secundário refere-se àquele associado a atividades em que o contato com a água é esporádico ou acidental e a possibilidade de ingerir quantidades apreciáveis de água é pequena, como na pesca e na navegação.

A qualidade da água para fins de recreação de contato primário constitui a balneabilidade, sendo necessário para sua avaliação o estabelecimento de critérios baseados em indicadores, que se possam identificar as condições de balneabilidade em um determinado local, inclusive com classes de balneabilidade para uma melhor orientação dos usuários.

O monitoramento e a avaliação da qualidade das águas é uma atribuição da CETESB. Para este fim a empresa realiza de forma sistemática a coleta e análise da água em 156 pontos fixos distribuídos em 144 praias de 14 municípios do litoral do Estado e as classifica em própria ou imprópria para banho.

Um dos principais objetivos do programa de balneabilidade é informar a população sobre a condição das praias, disponíveis à população no site da CETESB e na sinalização por meio de bandeiras instaladas nas praias monitoradas.

É importante destacar que o objetivo da divulgação da informação de balneabilidade da praia visa à proteção da saúde, tendo em vista que nas praias impróprias para o banho, o risco de contaminação aumenta.

Sendo assim, em dezembro de 2013, foi assinado o contrato com a empresa FOCO OPINIÃO E MERCADO LTDA – EPP para a realização de uma pesquisa quantitativa de satisfação sobre o sistema de informação e divulgação do monitoramento de qualidade das praias do estado de São Paulo e teve como finalidade auxiliar na identificação de possíveis falhas na estratégia de comunicação.

Os principais aspectos da pesquisa estão descritos a seguir, mas podem ser consultados na íntegra por meio do relatório apresentado pela empresa FOCO OPINIÃO E MERCADO LTDA – EPP. (Arquivo em CD anexo)

Também fez parte desse trabalho a apresentação dos resultados dessa pesquisa ao corpo funcional da CETESB, cujo evento aconteceu em 12 de novembro de 2014, na sede da cia em Pinheiros/São Paulo.

2. OBJETIVOS

O trabalho foi iniciado em janeiro de 2014 e teve os seguintes objetivos:



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO

Diretoria de Engenharia e Qualidade Ambiental

Objetivo geral

Medir a satisfação e a compreensão dos usuários quanto às informações geradas pelo serviço do monitoramento da qualidade das praias do litoral do Estado de São Paulo e suas formas de divulgação.

Objetivos específicos

- Identificar o perfil dos banhistas que se utilizam das informações geradas pelo serviço do monitoramento das praias da CETESB, inclusive o socioeconômico (Critério Brasil – versão 2014);
- Identificar os motivos e hábitos de uso das informações produzidas pelo serviço do monitoramento das praias efetuado pela CETESB;
- Medir a satisfação da população quanto aos diversos tipos de sinalização e informes de qualidade das praias utilizados para divulgação;
- Verificar associações de algumas características do perfil dos usuários na satisfação geral;
- Fazer comparações de resultados de satisfação entre as populações dos diversos municípios e regiões do litoral (Litoral Norte, Baixada Santista e Litoral Sul);
- Fazer comparações de resultados de satisfação entre a população fixa e a flutuante;
- Avaliar a compreensão da abrangência dos serviços prestados pela CETESB.

3. METODOLOGIA

A. Universo da pesquisa e público-alvo

O universo da pesquisa foi formado pelos banhistas das praias monitoradas pela CETESB. O público-alvo foi formado pelos moradores e banhistas, residentes na cidade pesquisada, denominados população fixa. Além disso, compõe o público alvo os visitantes às praias pesquisadas, denominados população flutuante. Porém, apenas fizeram parte da população de estudo os indivíduos com 18 anos ou mais.

B. Tipo da pesquisa e instrumento de coleta

A pesquisa teve caráter quantitativo, realizada pela técnica de survey, por levantamento amostral, sendo a coleta executada através de entrevistas pessoais, face a face, nas cidades selecionadas para compor a amostra da pesquisa. Para o caso da população fixa as entrevistas foram domiciliares e para a população flutuante, realizada em pontos de fluxo nas praias de cada município.

A coleta de dados foi realizada com equipamento eletrônico do tipo tabletes. No município de Itanhaém, particularmente, durante a execução da coleta de dados o equipamento da pesquisadora foi roubado, de modo que apenas 20,7% (17) das entrevistas programadas (82 no total) para aquele local foram realizadas neste formato de coleta. As demais foram realizadas em questionários impressos, mediante autorização da CETESB.

O instrumento de pesquisa foi o questionário estruturado, composto por 44 perguntas abertas e fechadas, cujo conteúdo pode ser consultado no relatório apresentado pela empresa contratada.

C. Amostragem e tamanho da amostra

O tamanho da amostra foi de 1250 entrevistas em 14 (quatorze) cidades do litoral de São Paulo, sendo 625 para cada tipo de população. A amostra foi construída por cotas representativas da população fixa e flutuante, proporcionais inicialmente à densidade populacional de cada município, conforme dados apurados no Censo 2012, do IBGE (para a população fixa) e em dados disponibilizados pela Fundação SEADE, 2012 (para a população flutuante). A tabela 1 a seguir apresenta os dados da população e amostra estratificados por município e tipo de população.

Tabela 1: Amostra da pesquisa

Município	Censo 2012 - Fixa		Seade 2012 - Flutuante		População fixa + flutuante	Amostra populacional	
	Qtidade	%	Qtidade	%		fixa	flutuante
Ubatuba	80604	4,24%	106810	7,81%	187414	27	49
Caraguatatuba	104150	5,48%	109296	7,99%	213446	34	50
Ilhabela	29308	1,54%	14049	1,03%	43357	10	6
São Sebastião	76344	4,02%	63282	4,63%	139626	25	29
Bertioga	50304	2,65%	62518	4,57%	112822	17	29
Guarujá	294669	15,50%	194597	14,22%	489266	95	89
Santos	419614	22,07%	82937	6,06%	502551	138	38
São Vicente	336809	17,72%	51178	3,74%	387987	111	23
Praia Grande	272390	14,33%	379697	27,76%	652087	90	173
Mongaguá	47984	2,52%	90736	6,63%	138720	16	41
Itanhaém	89332	4,70%	115314	8,43%	204646	29	53
Peruíbe	61030	3,21%	65409	4,78%	126439	20	30
Iguape	29055	1,53%	12999	0,95%	42054	10	6
Ilha Comprida	9376	0,49%	19196	1,40%	28572	3	9
Total	1900969	58,15%	1368018	41,85%	3268987	625	625

Fonte: CETESB/Foco Opinião e Mercado

Considerou-se um coeficiente de confiança de 95% com margem de erro de +4% para cada população e de +2,8% quando considerado o total da amostra (1250 entrevistas).

D. Seleção da amostra populacional fixa

A seleção dos domicílios ocorreu, inicialmente por sorteio do setor censitário. Identificado o setor, o ponto de início da varredura geográfica foi o indicado no mapa do IBGE como ponto inicial. Em cada setor foram realizadas entre 3 e 6 entrevistas, de acordo com a população do município, conforme segue:



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO
Diretoria de Engenharia e Qualidade Ambiental

Tabela 2: Total de entrevistas por setor censitário

População do município	Entrevistas/setor censitário
até 50.000 hab.	3
entre 50.000 e 100.000 hab	4
acima de 100.000 hab	6

A pesquisa foi domiciliar e uma vez selecionado o domicílio de estudo, apenas um dos moradores foi entrevistado. A amostra não abrangeu residentes em território brasileiro que não falam a língua portuguesa, deficientes mentais e outros indivíduos incapacitados de responder ao questionário.

No caso da impossibilidade de entrevistar o morador do domicílio selecionado, este foi substituído por outro no mesmo setor censitário. Domicílios não elegíveis, como domicílios vagos, domicílios coletivos (hotéis, alojamentos, asilos, etc.), estabelecimento agropecuário, estabelecimento de ensino, estabelecimento de saúde, estabelecimento de outras finalidades e edificações em construção foram substituídos por outro no mesmo setor censitário.

E. Seleção da amostra populacional flutuante

A população flutuante foi pesquisada nas praias do litoral em cada município, em pontos de fluxo, com um intervalo mínimo de 10 minutos de uma abordagem bem sucedida para outra.

F. Período da coleta de dados

A etapa de coleta de dados aconteceu simultaneamente nas três regiões litorâneas, sendo que as entrevistas com a população flutuante, obrigatoriamente, ocorreram em dias consecutivos no feriado do carnaval (sábado, domingo, segunda e terça-feira) ou, quando da não completude da amostra nesta data (por conta de questões climáticas e outros fatores externos alheios ao controle logístico da pesquisa), no sábado e domingo subsequente. A população fixa foi visitada em qualquer tempo nos dias programados para a realização da pesquisa.

Desta forma, a coleta de dados ocorreu entre os dias 3 e 26 de março de 2014.

G. Equipe de Campo

A equipe de trabalho de campo foi formada por profissionais com escolaridade mínima de ensino Médio, além de habilidades exigidas para o bom desempenho da função. A pesquisa foi realizada no intervalo de horário das 8h00 às 20h30 em dias úteis e não úteis, visando contemplar os 3 (três) turnos do dia.

Os profissionais que realizaram as entrevistas estavam devidamente identificados com crachá, contendo nome e logomarca da empresa prestadora do serviço, além de portarem documento oficial de identificação com foto.

4. PROCESSAMENTO DOS DADOS

A. Checagem, validação e consistência dos dados

A checagem foi realizada por telefone, obedecendo ao seguinte procedimento: a) Confirmação da existência do entrevistado; b) Verificação da aplicação integral do questionário. A checagem

telefônica foi realizada em 30% da amostra, aleatoriamente por entrevistador, totalizando 375 checagens.

B. Banco de dados

Após checagem foi realizado teste de consistência em 100% dos dados, garantindo a correta e completa digitação.

5. RESULTADOS

A seguir são apresentados os resultados da pesquisa com relação aos seguintes aspectos:

- 5.1 Perfil dos pesquisados.
- 5.2 Hábitos relacionados ao uso das praias;
- 5.3 Informações sobre a balneabilidade das praias;
- 5.4 Acesso ao serviço de divulgação;
- 5.5 Nível de conhecimento sobre o que é balneabilidade e monitoramento das praias
- 5.6 Imagem da CETESB;

Dentre estes itens foram selecionados os principais, porém todos os resultados da pesquisa podem ser consultados no relatório apresentado pela empresa contratada. (Arquivo em CD anexo)

5.1 PERFIL DOS PESQUISADOS

Em resumo, o Quadro 1 mostra o perfil da população entrevistada na pesquisa.

Quadro 1 – Perfil dos pesquisados



População Flutuante:

- **Sexo:** variável controlada
- **Idade:** Predominam jovens adultos - 75% até 44 anos
- **Escolaridade elevada:**
 - 42,4% com ensino superior (incompleto, completo ou pós-graduação)
 - 38,7% com médio completo
 - Média de anos de estudo: 11,76 anos
- **Renda média familiar:** R\$ 4.248,81
- **Classe:** A E B1 (50%)
- **Trabalham:** 77%



População Fixa:

- **Sexo:** variável controlada
- **Idade:** Predominam adultos maduros - 67% acima de 35 anos
- **Escolaridade mediana:**
 - 35,8% com médio completo
 - 36,8% até fundamental completo
 - Média de anos de estudo: 9,67 anos
- **Renda média familiar:** R\$ 3.012,76
- **Classe:** B2 e C1 (54%)
- **Trabalham:** 61%

5.2 HÁBITOS RELACIONADOS AO USO DAS PRAIAS

5.2.1 Frequência de ida a praia

De modo geral, 41,6% dos entrevistados costumam frequentar a praia ocasionalmente e outros 19,8% apenas em feriados prolongados ou no período de férias. Cerca de um terço da população é frequentador assíduo e costuma ir à praia pelo menos três vezes ao mês.

Gráfico 1: Frequência de ida a praia



Fonte: Pesquisa Foco Opinião e Mercado

Do total dos pesquisados da população local (fixa), 50,1% costuma frequentar as praias ocasionalmente. Já entre a população flutuante, 36,96% costuma frequentar apenas nos feriados prolongados no verão. Comparando os públicos, a população local é frequentadora mais assídua e 41,16% costuma ir a praia até três vezes ao mês.

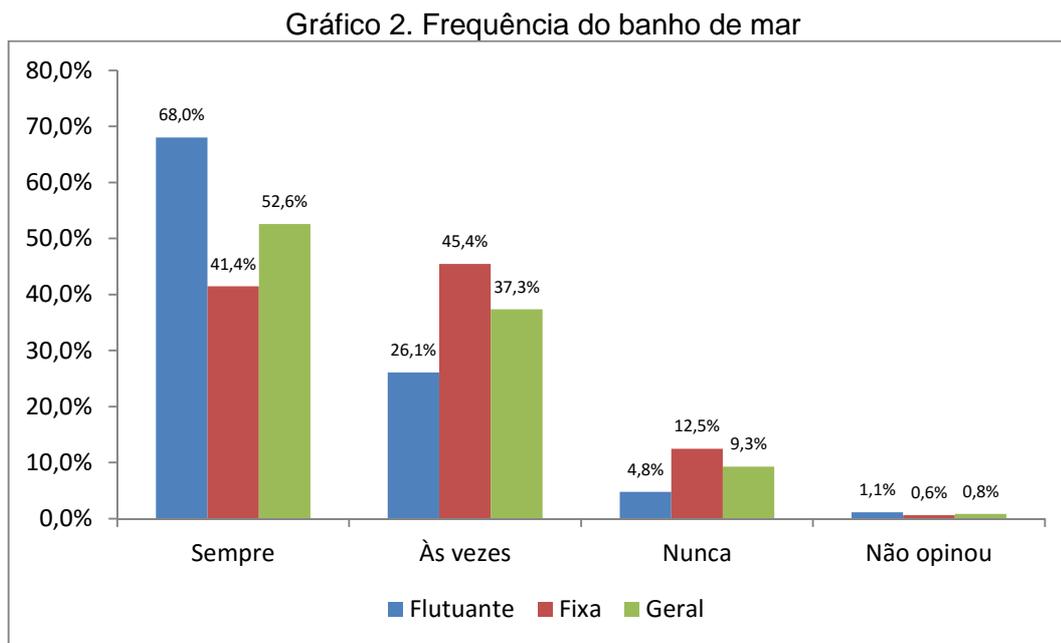
Tabela 3: Dados da frequência do banhista

Opções	Flutuante	Fixa	Geral
Praticamente todo o final de semana	13,76%	26,08%	20,92%
Até três vezes por mês	10,72%	14,08%	12,67%
Apenas nos feriados prolongados no verão	36,96%	7,52%	19,84%
Somente nas férias de final de ano	8,80%	2,24%	4,99%
Ocasionalmente	29,76%	50,08%	41,58%

Fonte: Pesquisa Foco Opinião e Mercado

5.2.2 Frequência com que costuma tomar banho de mar

Quando questionados sobre a frequência do banho de mar, quando vai à praia, a maioria respondeu que sempre toma banho (52,56%). Comparando-se as populações, os visitantes a frequência de banho é maior entre os visitantes em relação aos moradores, conforme mostrado no gráfico 2.



Com relação aos que nunca tomam banho quando vão à praia (9,3%), o principal motivo apontado foi a condição da água (suja, imprópria, tem dúvidas em relação à qualidade da água). Desta forma, pode-se afirmar que:

Tabela 4: Motivos que não tomam banho de mar

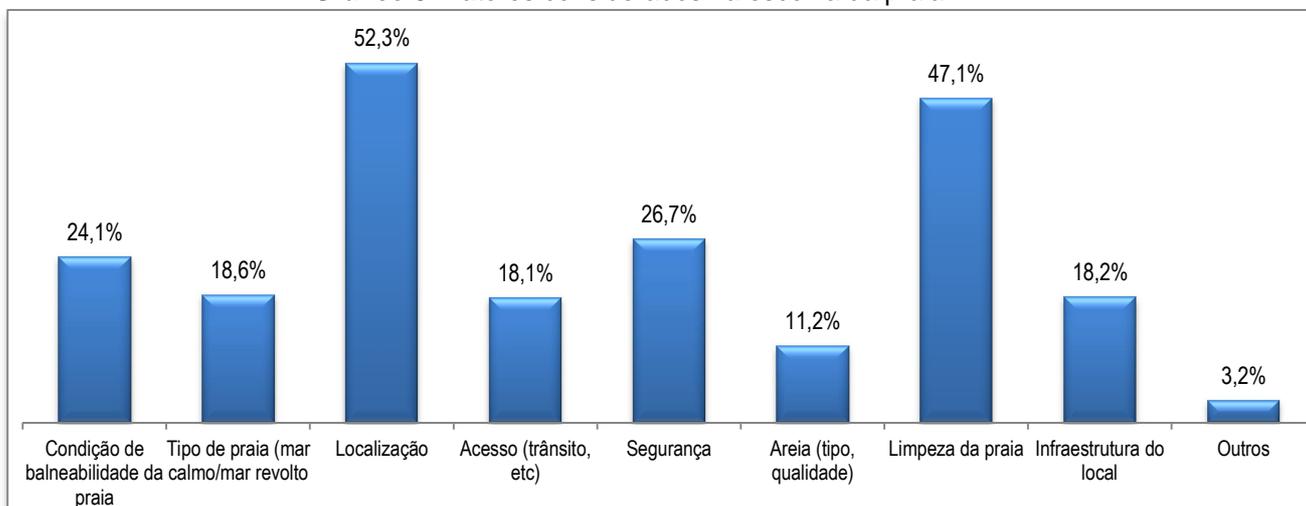
Opções	Flutuante	Fixa	Geral
Condição da água (suja, imprópria, tem dúvidas em relação à qualidade da água).	53,33%	58,97%	56,61%
Por outra razão (não gosta da água, do mar, tem alergias, tem medo de água etc.)	46,67%	41,03%	43,39%

Fonte: Pesquisa Foco Opinião e Mercado

5.2.3 Fatores considerados na escolha da praia

Os pesquisados foram questionados a respeito dos motivos que consideram importantes para a escolha da praia que frequentam. A localização e a limpeza são as condições mais prezadas na escolha, seguidos da segurança do local. As condições de balneabilidade aparecem na 4ª posição, citadas por 24,1% dos entrevistados.

Gráfico 3: Fatores considerados na escolha da praia



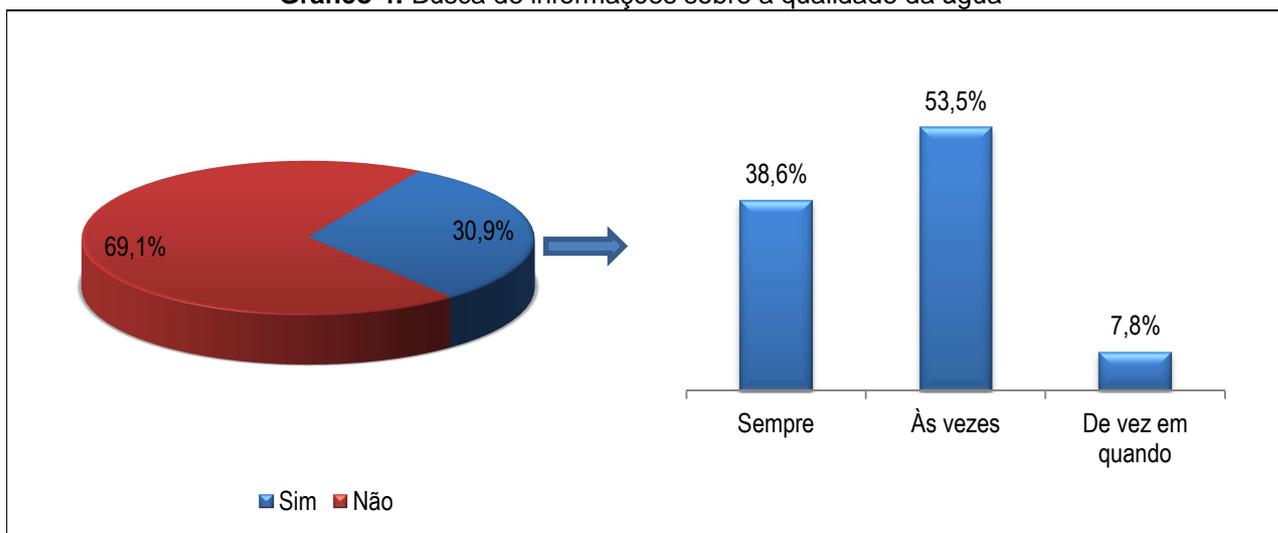
Fonte: Pesquisa Foco Opinião e Mercado

5.3 INFORMAÇÕES SOBRE BALNEABILIDADE DAS PRAIAS

5.3.1 Busca de informações sobre a balneabilidade das praias

A busca de informação sobre a qualidade da água, ou seja, a balneabilidade das praias que frequentam, não é comum entre os entrevistados. Apenas 30,9% do total de entrevistados costuma acessar informações desta natureza.

Gráfico 4: Busca de informações sobre a qualidade da água



Fonte: Pesquisa Foco Opinião e Mercado

Do total de entrevistados que procuram informações sobre a balneabilidade, a população visitante é a mais interessada (33,9%), quando comparada com a de moradores locais (28,8%).

5.3.2 Informação sobre a balneabilidade da praia onde estava

Dos visitantes entrevistados nas praias do litoral (população flutuante) apenas 17,4% respondeu que buscou informações sobre a condição de balneabilidade daquela praia onde estavam.

Esta população buscou tal informação, em sua grande maioria, pela internet em geral, porém não associa diretamente à CETESB. A maioria (61,5%) citou sites em geral (google, prefeituras, cidade, surf). As bandeiras indicativas nas praias também foram usadas como referência por 12,8% dos visitantes. A CETESB, seja através de seu site ou painel, foi citada por apenas dois entrevistados, o que representa 1,8% dos que buscam informação e 0,3% do total de turistas entrevistados.

Tabela 5: Meios de informação consultados sobre a balneabilidade das praias

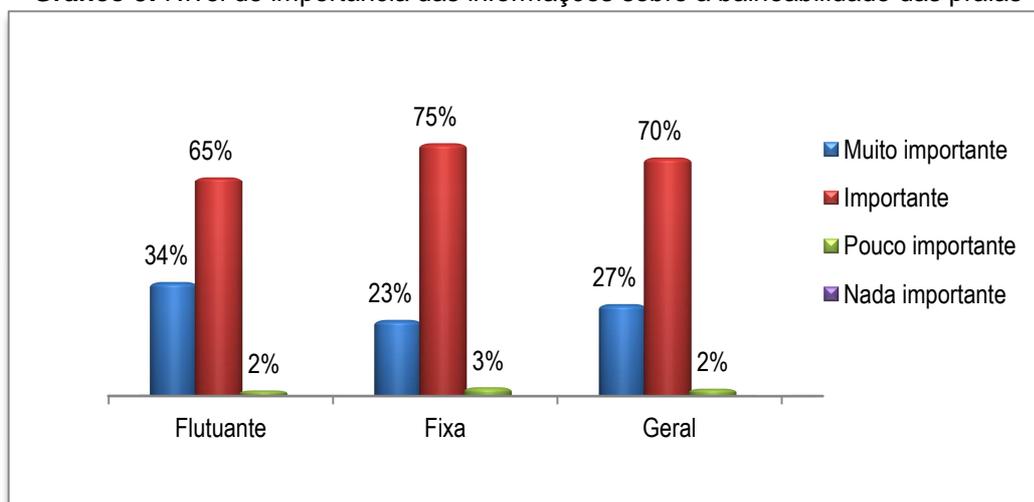
Opções	Percentual
Internet em geral (google, site de prefeituras, da cidade, de surf)	61,5%
Amigos, conhecidos, familiares	15,6%
Bandeiras nas praias	12,8%
TV	4,6%
População local	4,6%
Hotéis onde se hospeda	2,8%
Placas indicativas nas praias	1,8%
Aplicativos no celular	1,8%
Site da CETESB	0,9%
Painel da CETESB (nas praias)	0,9%
Jornais impressos	0,9%
Frequentadores	0,9%

Fonte: Pesquisa Foco Opinião e Mercado

5.3.3 Importância da informação sobre a balneabilidade das praias

Apesar de não se informarem a respeito da balneabilidade, praticamente a totalidade dos entrevistados (97%) considera importante ou muito importante conhecer a informação.

Gráfico 5: Nível de importância das informações sobre a balneabilidade das praias



Fonte: Pesquisa Foco Opinião e Mercado

Além disso, quanto maior a busca pela informação, mais importante ela é considerada. Entre os que costumam informar-se, 52,1% consideram muito importante.

Tabela 6: Importância da informação sobre balneabilidade de acordo com quem costuma ou não se informar

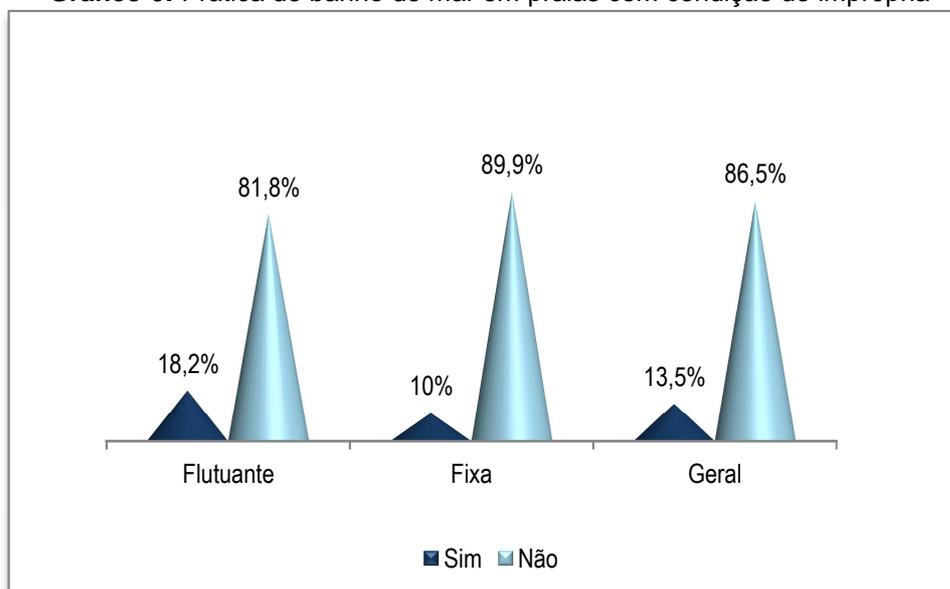
Importância	Costuma Informar-se	Não costuma se informar
Muito importante	52,1%	17%
Importante	47,1%	80%
Pouco importante	0,8%	2,7%
Nada importante	--	0,3%
Respondentes	392	858

Fonte: Pesquisa Foco Opinião e Mercado

5.3.4 Comportamento do banhista em caso de água do mar imprópria

Em 86,5% dos casos os entrevistados afirmaram que não entram na água quando está indicando que as condições são impróprias ao banho.

Gráfico 6: Prática de banho de mar em praias com condição de imprópria



Fonte: Pesquisa Foco Opinião e Mercado

5.3.5 Percepção de agravos à saúde causada pela água imprópria para banho

Para 84,3% dos entrevistados, a água imprópria para banho seguramente causa prejuízos à saúde do banhista.

Tabela 7: Percepção de agravos à saúde causados pela água imprópria para banho

Opções	Flutuante	Fixa	Geral
Sim, tenho certeza	83%	85,3%	84,3%
Acho que sim, mas não tenho certeza	14,1%	11,4%	12,5%
Seguramente não causa	1,6%	0,3%	0,9%
Não sabe opinar a respeito	1,3%	3%	2,3%
Total	100%	100%	100%

Fonte: Pesquisa Foco Opinião e Mercado

5.3.6 Informações sobre a balneabilidade das praias por estratificações

Gráfico 7: Busca de informações sobre a balneabilidade das praias por região do litoral

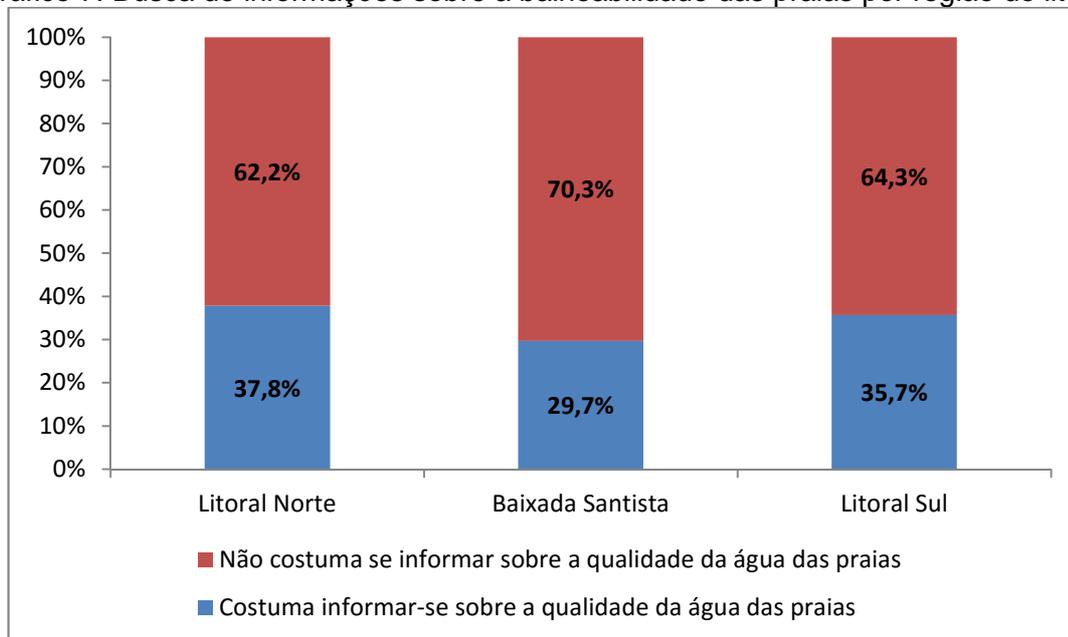
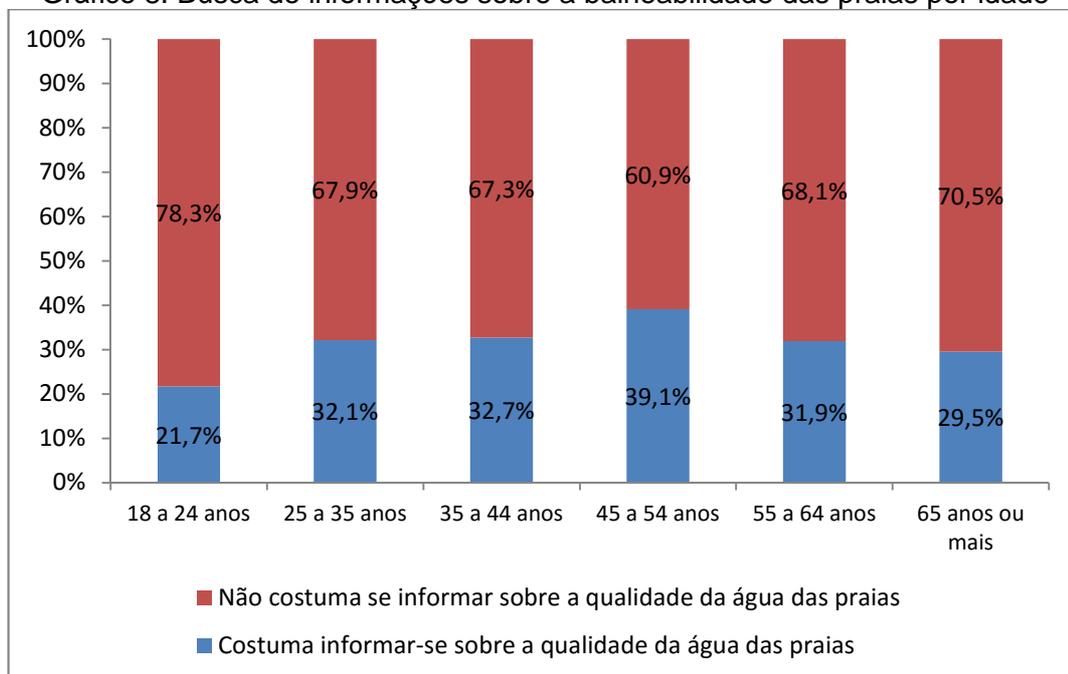


Gráfico 8: Busca de informações sobre a balneabilidade das praias por idade



5.3.7 Estratificação por importância da informação sobre a balneabilidade das praias

Gráfico 9: Importância da informação sobre a balneabilidade das praias por região do litoral

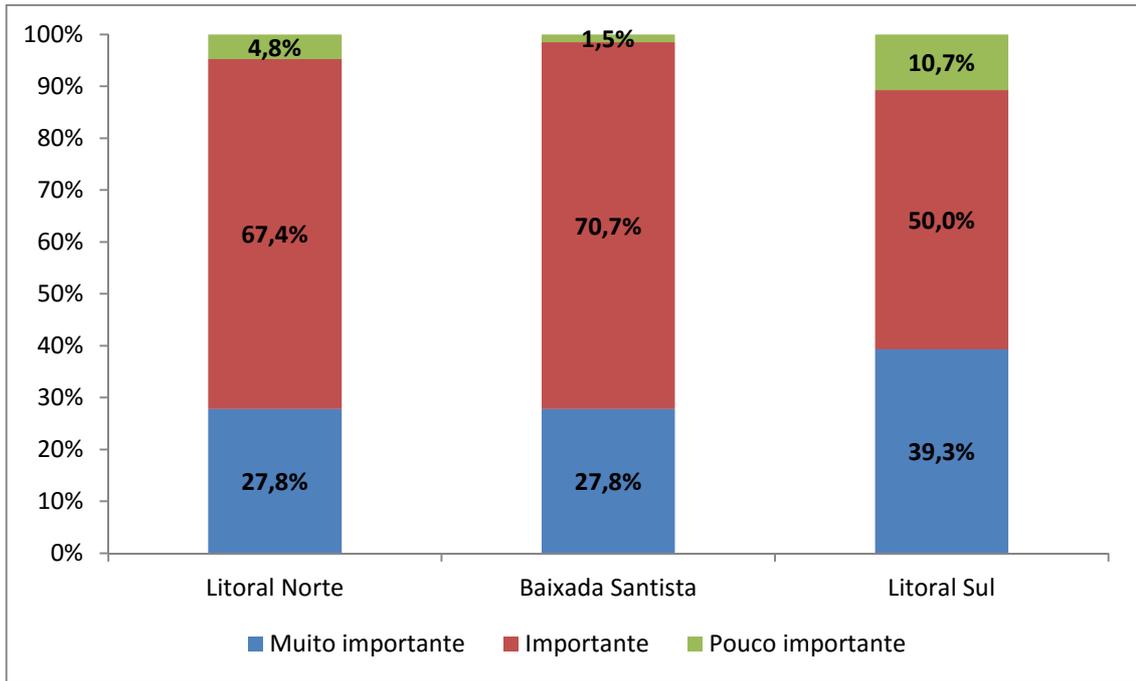
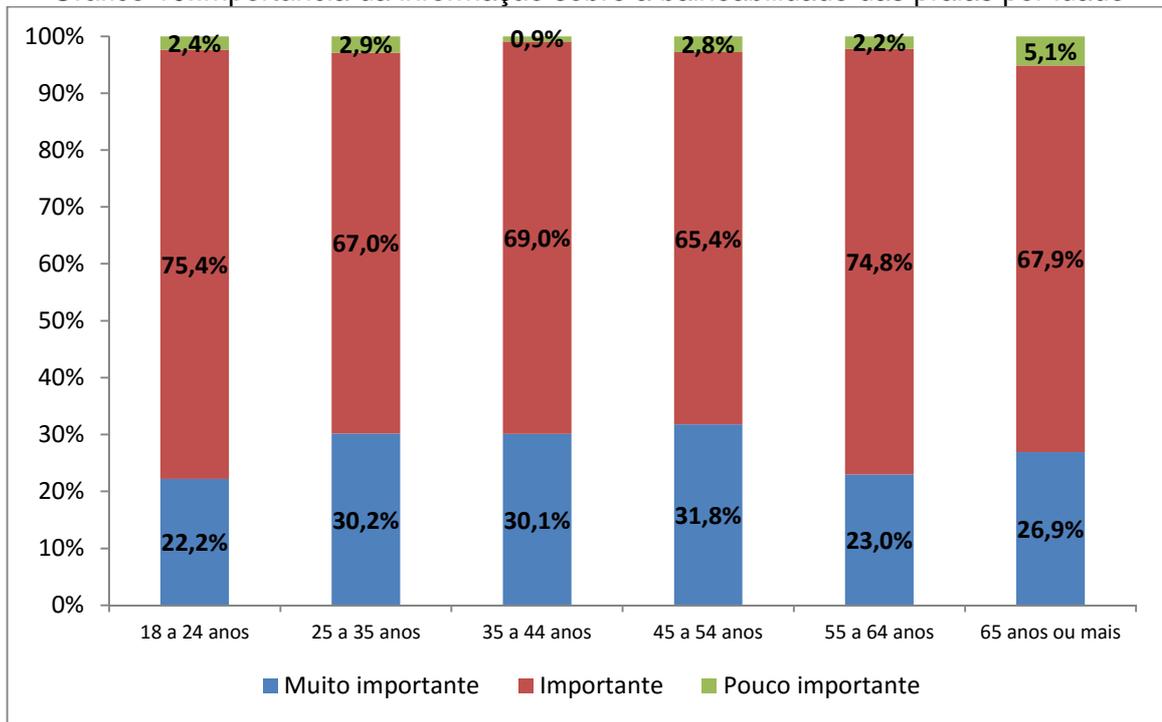


Gráfico 10: Importância da informação sobre a balneabilidade das praias por idade



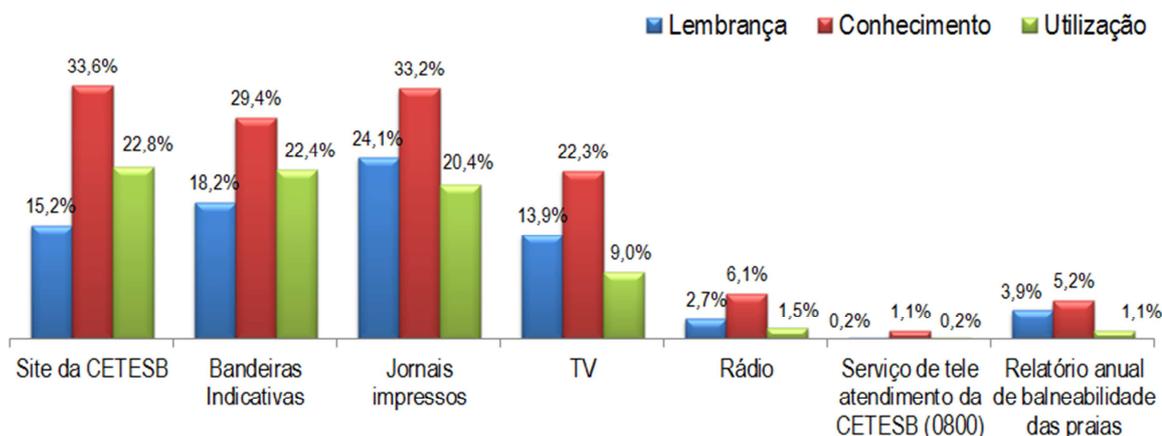
5.4 ACESSO AO SERVIÇO DE DIVULGAÇÃO

5.4.1 Nível de conhecimento dos canais de divulgação

O nível de conhecimento dos canais de divulgação das informações sobre balneabilidade foi pesquisado por meio de três questões diferentes e o resultado foi o seguinte.

- Quando questionados a respeito dos canais de divulgação: espontaneamente mencionaram os jornais impressos (24,1%) e as bandeiras indicativas (18,2%)
- Quando estimulados, ou seja, questionados se conhecem aquele canal como fonte de informação sobre a balneabilidade das praias, o site da CETESB ocupa a primeira posição, conhecido por 33,6% dos entrevistados.
- Quando questionados sobre a utilização dos canais, o site da CETESB é também o mais utilizado, por 22,8% da população.

Gráfico 11: Nível de conhecimento dos canais de divulgação

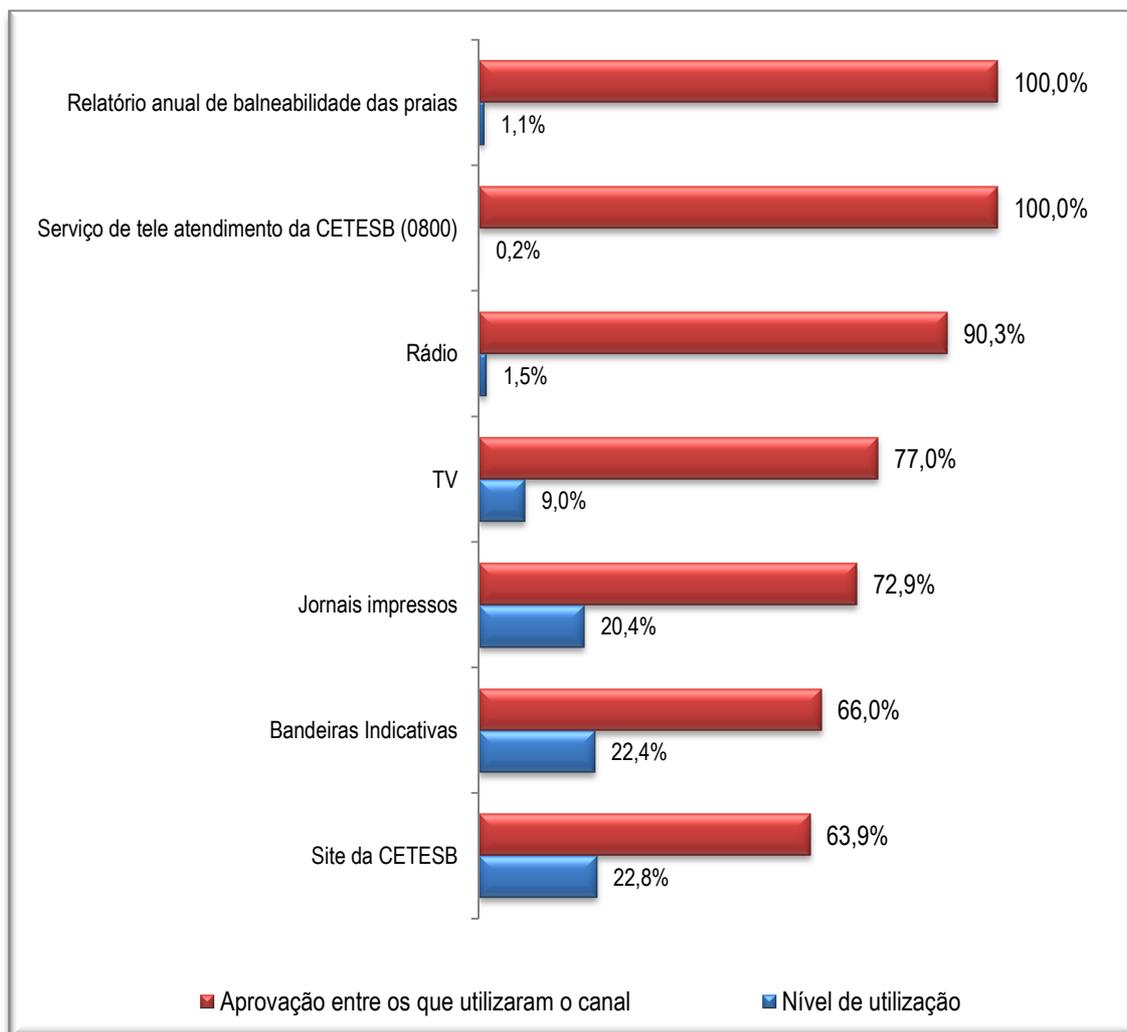


Fonte: Pesquisa Foco Opinião e Mercado

5.4.2 Nível de utilização e avaliação dos canais de divulgação

De acordo com a pesquisa, verificou-se que, entre os entrevistados, os canais de divulgação sobre balneabilidade das praias mais utilizados são o site da CETESB (22,8%), as bandeiras indicativas (22,4%) e jornais (20,4%). No entanto são também estes os canais com menor avaliação entre os observados, conforme mostra o gráfico x.

Gráfico 12: Nível de utilização e avaliação dos canais de divulgação

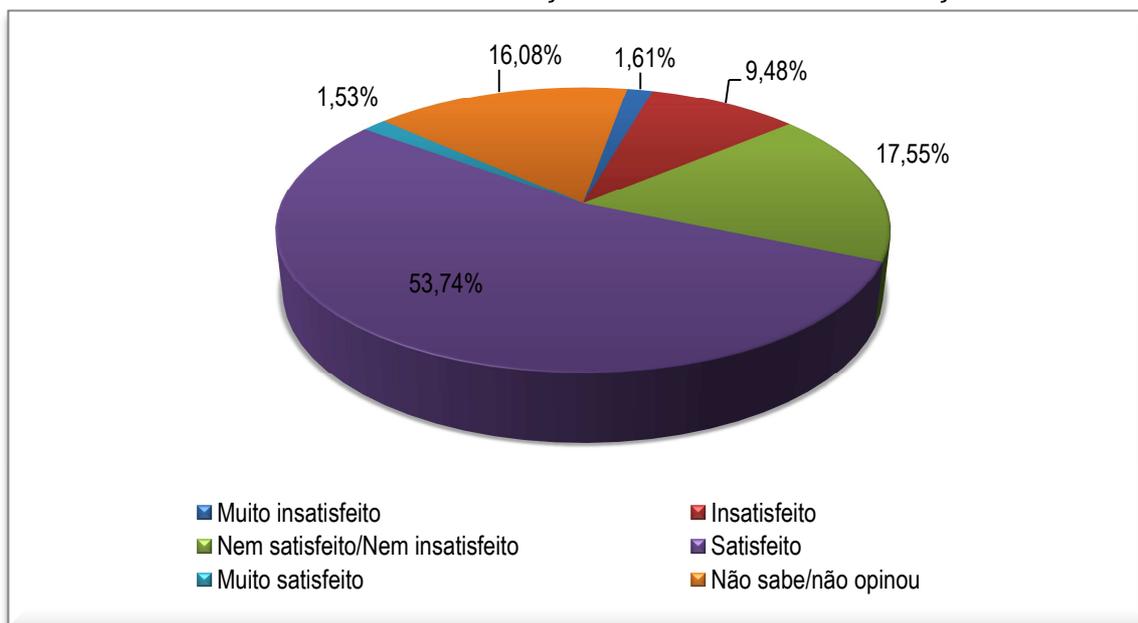


Fonte: Pesquisa Foco Opinião e Mercado

5.4.3 Nível de satisfação com os canais de informação

Independentemente do conhecimento, os entrevistados declararam seu nível de satisfação com os canais de informação existentes sobre a balneabilidade das praias. Do total, 55,27% dos entrevistados estão satisfeitos ou muito satisfeitos com os canais que têm à disposição.

Gráfico 13: Nível de satisfação com os canais de informação



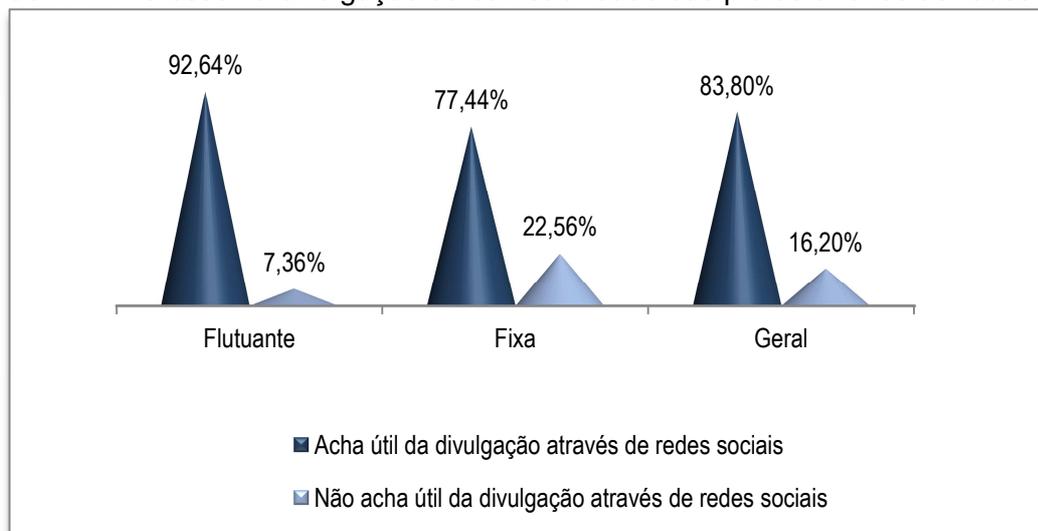
Fonte: Pesquisa Foco Opinião e Mercado

Observou-se também que o nível de satisfação é similar na população fixa(54,8%) e flutuante(52,1%).

5.4.4 Divulgação da balneabilidade das praias através das redes sociais

Os entrevistados consideram importante a divulgação da balneabilidade das praias nas redes sociais (83,80%). Entre a população de visitantes o interesse neste canal é ainda maior, de 92,64%.

Gráfico 14: Interesse na divulgação da balneabilidade das praias através de redes sociais



Fonte: Pesquisa Foco Opinião e Mercado



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO

Diretoria de Engenharia e Qualidade Ambiental

Com relação às ferramentas de mídias sociais a serem usadas, citam o Facebook (90,19%) como principal canal de divulgação, seguido do Instagram (13,27%).

5.4.5 Sugestões para melhoria na divulgação da qualidade da água nas praias

Os entrevistados foram questionados sobre sugestões para melhorar a divulgação da balneabilidade das praias. De modo geral, concentraram suas opiniões na intensificação da divulgação, especialmente utilizando a internet, principalmente, as redes sociais.

Em resumo, as sugestões abordaram os seguintes aspectos:

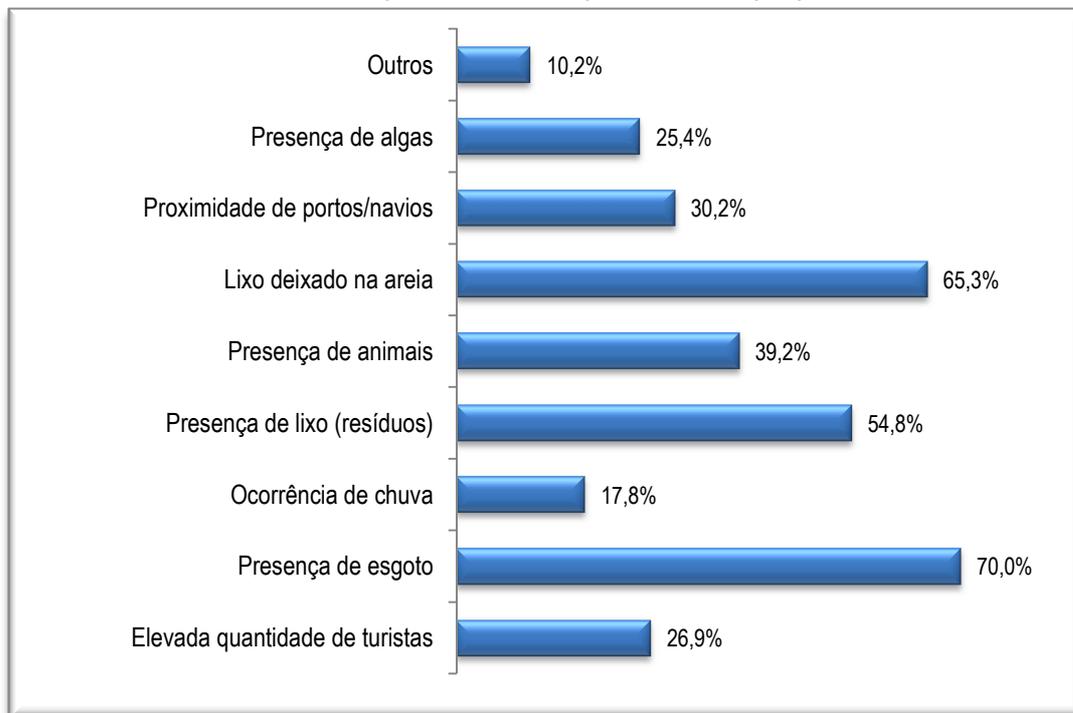
- Anúncios na televisão, Internet, redes sociais, rádios (53,7%);
- Mais informações, cartazes, panfletos, banners, placas, bandeiras, painel eletrônico, quiosques eletrônicos (29,5%);
- Informar mais a comunidade: nas ruas, na praia, na cidade (8,5%)
- Conscientizar, realizar campanhas em geral, divulgar nas escolas e estimular as pessoas para o cuidado com a praia (5,1%);
- Ponto de informação e mais pessoas na praia (2,3%)
- Fiscalização (2,0%)

5.5 NÍVEL DE CONHECIMENTO SOBRE BALNEABILIDADE E MONITORAMENTO DAS PRAIAS

5.5.1 Fatores que levam uma praia a ficar imprópria

A presença de esgoto, de lixo na praia e na areia são os principais motivos que levam a praia a ficar imprópria (70,0%, 54,8% e 65,3%, respectivamente) na opinião dos entrevistados.

Gráfico 15: Motivos que levam uma praia ficar imprópria



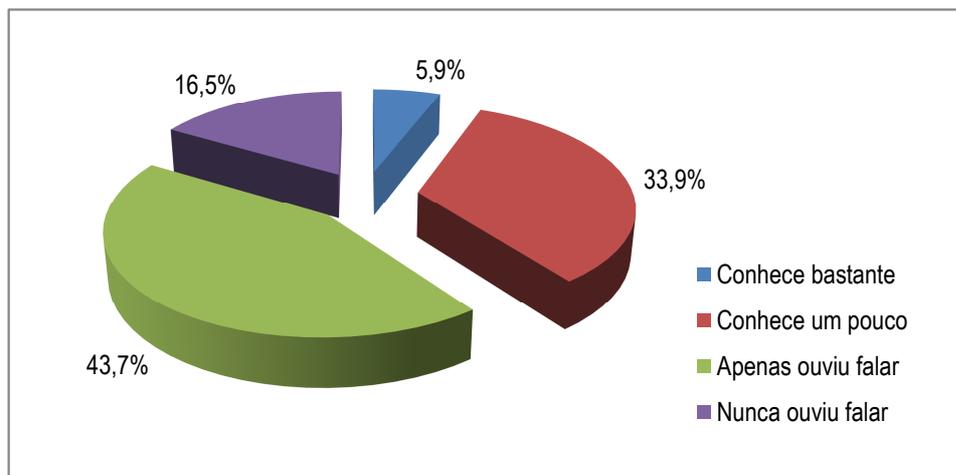
Fonte: Pesquisa Foco Opinião e Mercado

5.6 IMAGEM DA CETESB

5.6.1 Nível de conhecimento da CETESB

O nível de conhecimento sobre a CETESB é baixo, tanto população fixa quanto na flutuante. Para 16,5% dos entrevistados a instituição inexistente e para 43,7% o conhecimento é superficial, visto que apenas ouviram falar da empresa.

Gráfico 16: Nível de conhecimento da CETESB

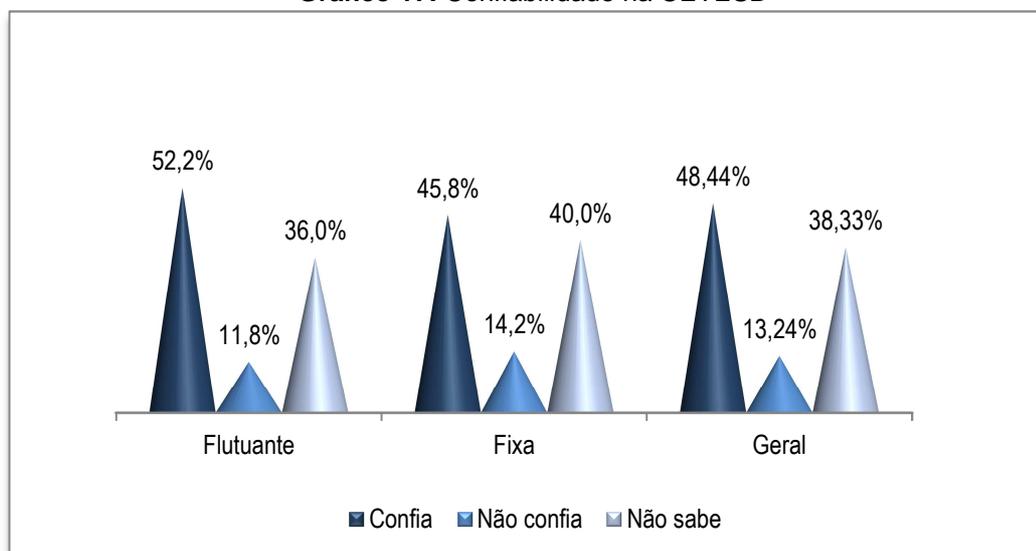


Fonte: Pesquisa Foco Opinião e Mercado

5.6.2 Confiabilidade na CETESB

De modo geral, os consultados consideram a CETESB uma companhia confiável (48,4%), principalmente para a população flutuante (52,2%).

Gráfico 17: Confiabilidade na CETESB



Fonte: Pesquisa Foco Opinião e Mercado

A confiabilidade na CETESB cresce com o conhecimento que possuem da instituição. Quanto mais sabem a respeito maior é a confiança na empresa.

Tabela 8: Nível de confiança sobre a CETESB

	Conhece muito ou conhece pouco	Apenas ouviu falar Nunca ouviu falar
É confiável	61%	41%
Não é confiável	19%	9%
Não sabe, não opinou	20%	50%

5.6.3 Avaliação sobre a CETESB

Independentemente do conhecimento, a avaliação em relação à CETESB foi positiva, sendo que a maioria (46,9%) dos que conhecem e responderam classificou como ótima ou boa.

Tabela 9: Avaliação sobre a CETESB

Avaliação	Flutuante	Fixa	Geral
Ótimo	2,5%	2,4%	2,4%
Bom	47,8%	42,2%	44,5%
Regular	10,8%	10,0%	10,3%
Ruim	1,1%	3,2%	2,3%
Péssimo	0,3%	0,4%	0,4%
Não sabe, não respondeu	37,2%	41,6%	39,8%

Fonte: Pesquisa Foco Opinião e Mercado

5.6.4 Pontos fortes percebidos na CETESB

Os entrevistados que avaliaram a atuação da CETESB como ótima ou boa (46,9%) foram questionados sobre a motivação dessa avaliação. A maioria (21,3%) não soube indicar um motivo que justifique tal impressão, porém destacam-se os que apontaram as seguintes razões pela avaliação positiva:

- os serviços de análise da água e informações prestadas - 16,2%;
- credibilidade da empresa - 11,1%;
- consideram o trabalho da CETESB ótimo ou bom por perceberem melhorias na qualidade da água das praias e atribuírem tal mudança ao seu trabalho.

5.6.5 Pontos que deixam a desejar na atuação da CETESB

Os entrevistados que avaliaram a atuação da CETESB como regular, ruim ou péssima (13%) foram questionados sobre a motivação dessa avaliação. A maior parcela justifica a necessidade de maior divulgação, da companhia em si e do resultado de seu trabalho. Um segmento atribui a



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO

Diretoria de Engenharia e Qualidade Ambiental

condição atual das praias à responsabilidade da CETESB, motivo pelo qual não aprova sua atuação. A falta de conhecimento sobre o trabalho da CETESB, também levou os entrevistados responder as seguintes razões:

- falta de divulgação (59,3%);
- precisa melhorar a atuação em geral (18,5%);
- precisa melhorar a qualidade da água (12,3%).

6. CONSIDERAÇÕES FINAIS

Com base nos resultados da pesquisa, pode-se inferir que apesar de considerarem importante conhecer a condição de balneabilidade das praias que frequentam (97%) e de acreditarem que a água imprópria causa dano à saúde (97%), apenas:

- 24,1% consideram as condições de balneabilidade ao escolher a praia;
- 30,9% costumam informar-se sobre a balneabilidade da praia que frequentam;
- 17,4% dos visitantes efetivamente se informaram sobre a balneabilidade da praia onde estavam;
- 13,5% entram na água mesmo sabendo que esta é imprópria para banho;
- 89,9% tomam banho de mar quando vão a praia.

Apesar de considerarem importante conhecer a condição de balneabilidade das praias que frequentam (97%) e de acreditarem que a água imprópria causa dano à saúde (97%):

- 56 % NÃO LEMBRAM DE NENHUM CANAL de divulgação de informações sobre balneabilidade;
- 41% NÃO CONHECEM nenhum canal de divulgação de informações sobre balneabilidade;
- 53 % NÃO UTILIZARAM NENHUM CANAL de divulgação de informações sobre balneabilidade.

Em resumo, a pesquisa apontou os seguintes aspectos que devem ser explorados:

- Imagem da instituição CETESB: predomina a impressão positiva, mas não é consolidada pelo baixo nível de conhecimento;
- Há dúvidas relacionadas à atividade fim da instituição;
- O nível de conhecimento a respeito do assunto balneabilidade também é baixo.
- A utilização das redes sociais, na atual conjuntura, configura um excelente canal para divulgação entre os visitantes (jovens elevada escolaridade, renda superior à dos moradores)
- As campanhas educativas por veículos mais presenciais são mais indicadas para a população local.
- Existe interesse no assunto, porém faz-se necessário:
 - Divulgar o que é balneabilidade
 - Porque é importante



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO

Diretoria de Engenharia e Qualidade Ambiental

- Onde buscar a informação.

7. CONCLUSÃO

A pesquisa mostrou que a CETESB tem credibilidade junto à população, mas falta divulgação sobre o conteúdo de suas atividades, principalmente entre o público jovem. Dessa forma, a partir desse trabalho, as áreas técnicas e de comunicação da empresa trabalharam no sentido de buscar formas de melhorar essa percepção.

Sendo assim, com base nos resultados da pesquisa e nas considerações levantadas, a CETESB realizou, até o momento, as seguintes ações para ampliar a divulgação da atividade de monitoramento das praias:

- Apresentação da pesquisa aos funcionários da CETESB em 12/11/2014 na sede da empresa, a fim de conscientizar o corpo técnico sobre o grau de conhecimento da empresa pelo público em geral e a importância da divulgação de suas atividades;
- Em dezembro de 2014, foi criada uma página no Facebook com as informações de balneabilidade das praias litorâneas monitoradas pela CETESB, que pode ser acessada por meio do link: <https://www.facebook.com/qualidadedaspraias/?fref=ts>
- Em meados de 2015, um vídeo institucional sobre Balneabilidade das Praias foi inserido na citada página do Facebook e no aplicativo SAP, criado pela Secretaria do Meio Ambiente para download gratuito em telefones celulares e tablets com informações sobre o sistema ambiental paulista. (Consulta pelo link: <http://www.ambiente.sp.gov.br/aplicativos/sap/>)



SERVIÇOS ESPECIALIZADOS EM ESTATÍSTICA APLICADA À GESTÃO DA QUALIDADE DAS ÁGUAS SUPERFICIAIS E SUBTERRÂNEAS

1. INTRODUÇÃO

A Companhia Ambiental do Estado de São Paulo - CETESB iniciou em 1974 a operação da Rede de Monitoramento de Qualidade das Águas Superficiais do Estado de São Paulo. As informações obtidas por meio do monitoramento têm possibilitado o conhecimento das condições ambientais nos principais rios e reservatórios situados nas 22 Unidades de Gerenciamento de Recursos Hídricos (UGRHs), em que se divide o Estado de São Paulo de acordo com a Lei Estadual n.º 9.034 de 27 de dezembro de 1994.

A rede de monitoramento de águas superficiais da CETESB em operação há mais de 35 anos, iniciou com a implantação de 47 pontos de amostragem e em 2011 contou com 354 estações manuais e 12 estações automáticas de monitoramento das águas, que geram dados em tempo real e 30 estações de avaliação de balneabilidade, totalizando 396 pontos de monitoramento, bem como fez uso de 50 variáveis de qualidade de água. Para o mesmo ano, a Rede de Monitoramento de Praias Litorâneas contou com 156 pontos de amostragem, em 137 praias do litoral do Estado de São Paulo. Toda a rede de monitoramento gerou um volume de dados anual superior a 60.000 análises químicas, físicas e biológicas, realizadas por seus laboratórios sediados em São Paulo e nas agências ambientais distribuídas no Estado. (CETESB, 2012)

O monitoramento de águas subterrâneas executado pela CETESB tem seu histórico iniciado em 1977, em atendimento ao regulamento da Lei 997/76, aprovado pelo Decreto Estadual 8468/76, quando foi publicado um relatório sobre a poluição das águas subterrâneas no Estado de São Paulo, elaborado com base em dados disponíveis nos cadastros da CETESB e do Departamento de Águas e Energia Elétrica - DAEE.

Em 1990, a rede estadual de monitoramento ambiental começou a ser efetivamente implementada, também, em atendimento à Lei 6.134 de 02.07.1988, que dispõe sobre a preservação dos depósitos naturais de águas subterrâneas, regulamentada pelo Decreto Estadual 32.955 de 07.02.1991.

A CETESB realiza o monitoramento regional, que além do critério de vulnerabilidade dos aquíferos, também considera a utilização da água subterrânea para poço de abastecimento público. Assim, os primeiros pontos de amostragem foram selecionados nas áreas de afloramento do Aquífero Guarani e no Aquífero Bauru, no centro e oeste do Estado, sendo eleitos poços de abastecimento devidamente cadastrados no DAEE e SABESP.

Em 2007/2009, foram monitorados semestralmente 184 pontos de amostragem selecionados de forma a abranger os principais aquíferos do Estado, incluindo, além dos poços de abastecimento público, poços e nascentes de águas minerais na Região Metropolitana de São Paulo - RMSP, bem como poços particulares de uso industrial. (CETESB, 2010)

O acervo produzido por meio desses monitoramentos apresenta uma massa de dados históricos, que permite análises estatísticas apuradas, com vistas ao aprimoramento do diagnóstico ambiental no que diz respeito ao conhecimento do comportamento e da tendência das variáveis de qualidade das águas monitoradas pela CETESB. Os dados dos



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO

Diretoria de Engenharia e Qualidade Ambiental

monitoramentos da CETESB encontram-se armazenados em bancos de dados, porém ainda faltava um refinamento tecnológico e o conhecimento mais específico em análise estatística destes dados de forma a permitir a geração de informações que possam subsidiar os gestores na tomada de decisão.

Dessa forma, em dezembro de 2013, foi contratada a empresa STATSOFT SOUTH AMERICA COMÉRCIO DE SOFTWARE LTDA para realizar serviços especializados em estatística.

Os principais aspectos dos serviços executados estão descritos a seguir, mas podem ser consultados na íntegra por meio dos relatórios apresentados pela empresa contratada. (Arquivos em CD anexo)

2. OBJETIVOS

A empresa foi contratada visando à execução de serviços especializados para o desenvolvimento dos seguintes produtos:

- A. Realização de análise estatística, visando à implementação de rotinas de tratamento e interpretação de dados dos monitoramentos de qualidade ambiental da CETESB, que são apresentados nos relatórios de qualidade das águas superficiais, anualmente, e das águas subterrâneas, trienalmente
- B. Relatórios dos seguintes estudos:
 - Estudo da tendência da qualidade das águas, referente às redes de monitoramento superficial e subterrânea da CETESB, no tocante às variáveis: sanitárias, metais, ecotoxicológicas e hidrobiológicas.
 - Avaliação da correlação de variáveis de qualidade ambiental, com base nos dados da rede de monitoramento CETESB, visando respostas para fenômenos como eutrofização em reservatórios, impropriedade de praias em função de lançamentos de esgoto, nitrificação do lençol freático, e outras situações que sejam relevantes, bem como estabelecer a similaridade entre variáveis para obtenção de mesma avaliação (Demanda Bioquímica de Oxigênio e Carbono Orgânico Total), visando a otimização da rede de monitoramento.
 - Estudo da série histórica da rede automática de monitoramento de águas superficiais da CETESB, objetivando avaliar o comportamento das variáveis: Oxigênio Dissolvido, pH, Condutividade e Turbidez, identificando situações de anormalidade.
- C. Capacitação dos técnicos da área de qualidade ambiental em desenvolvimento de estudos estatísticos.

3. METODOLOGIA

Com base em dados selecionados a partir dos resultados das Redes de Monitoramento de Qualidade das Águas da CETESB foram realizados estudos de tendência e correlação de variáveis, conforme a seguir.



3.1 ESTUDOS DE TENDÊNCIA

- 3.1.1 ESTUDO TENDÊNCIA – REDE BALNEABILIDADE
- 3.1.2 ESTUDO TENDÊNCIA – REDE ÁGUAS INTERIORES
- 3.1.3 ESTUDO TENDÊNCIA – REDE AUTOMÁTICA
- 3.1.4 ESTUDO TENDÊNCIA – REDE ÁGUAS SUBTERRÂNEAS

3.2 ESTUDOS DE CORRELAÇÃO DE VARIÁVEIS

- 3.2.1 ESTUDO CORRELAÇÃO – REDE ÁGUAS INTERIORES
- 3.2.2 ESTUDO CORRELAÇÃO – REDE ÁGUAS INTERIORES

3.1 ESTUDOS DE TENDÊNCIA

Os estudos de tendência foram realizados com base no monitoramento das seguintes redes da CETESB:

- Balneabilidade das Praias Litorâneas;
- Águas Interiores (Rede Básica);
- Rede de Monitoramento Automático de Águas Superficiais;
- Águas Subterrâneas (Poços de Abastecimento Público).

Nos itens a seguir são apresentados os principais aspectos com relação aos resultados obtidos nos respectivos estudos, porém todo conteúdo pode ser consultado, na íntegra, nos relatórios apresentados pela empresa contratada (Arquivos em CD anexo)

3.1.1 ESTUDO TENDÊNCIA – REDE BALNEABILIDADE

O objetivo deste estudo foi evidenciar estatisticamente a tendência do parâmetro *Enterococos*, mensurado semanalmente em 91 praias da costa do Estado de São Paulo no período de 2004 a 2013.

Para modelar todas as praias foi utilizada a estrutura SARIMA (p,d,q)(P,D,Q)L focando exclusivamente na avaliação da presença ou não da estacionariedade, cujos aspectos associados na previsão não serão explorados nas análises.

Resumo dos procedimentos estatísticos

Parte 1: Preparação da base de dados.

- Os dados originais devem ser substituídos pelos dados transformados através da transformação logarítmica de base 10.

Parte 2: Análise Descritiva.

- Gerar o gráfico da série temporal (parâmetro).
- Gerar dois gráficos Box-Plot por ano, um excluindo o mês de janeiro e o outro incluindo somente o mês de janeiro.
- Produzir uma tabela com as estatísticas mediana, 3º Quartil, percentil de 90%, percentil de 95% e valor máximo. Gerar um gráfico de linhas, por ano, com essas estatísticas.



Parte 3: Metodologia Box & Jenkins.

- Gerar os gráficos de Função de Autocorrelação e Função de Autocorrelação Parcial. Analisar a estacionariedade da série temporal.
- Caso a série temporal seja estacionária, conclui que a mesma não possui tendência. Um modelo AR(1) é adequado a série.
- Caso a série temporal não seja estacionária, proceder com diferenças na série até torná-la estacionária.
- Gerar os gráficos de Função de Autocorrelação e Função de Autocorrelação Parcial. Modelar a série.
- Construir uma tabela com os coeficientes do modelo de Box & Jenkins.

Parte 4: Análise residual.

- Gerar os gráficos de Função de Autocorrelação dos resíduos do modelo. Verificar as correlações entre os resíduos.
- Caso os resíduos sejam correlacionados, buscar outro modelo para a série temporal. Caso os resíduos sejam não-correlacionados, conclui sobre o comportamento da série diferenciada ser ruído branco.

De modo geral, esse procedimento foi adotado nos estudos das demais redes de monitoramento.

DESENVOLVIMENTO ANALÍTICO POR PRAIA

O estudo de tendência foi realizado por praia de cada município do litoral.

A título de exemplo segue a análise da praia de Boracéia em Bertioga.

O conjunto de dados analisado corresponde aos resultados das concentrações de *Enterococos* (UFC/100mL) nas amostras de água colhidas semanalmente nessas praias no período de 10 anos (2004-2013).

Praia de BORACÉIA

Foram observados os resultados temporais associados à praia de Centro. Durante a análise, transformou-se as informações semanais a partir do operador matemático \log_{10} com o intuito de estabilizar a variabilidade. No gráfico 1, a série transformada:

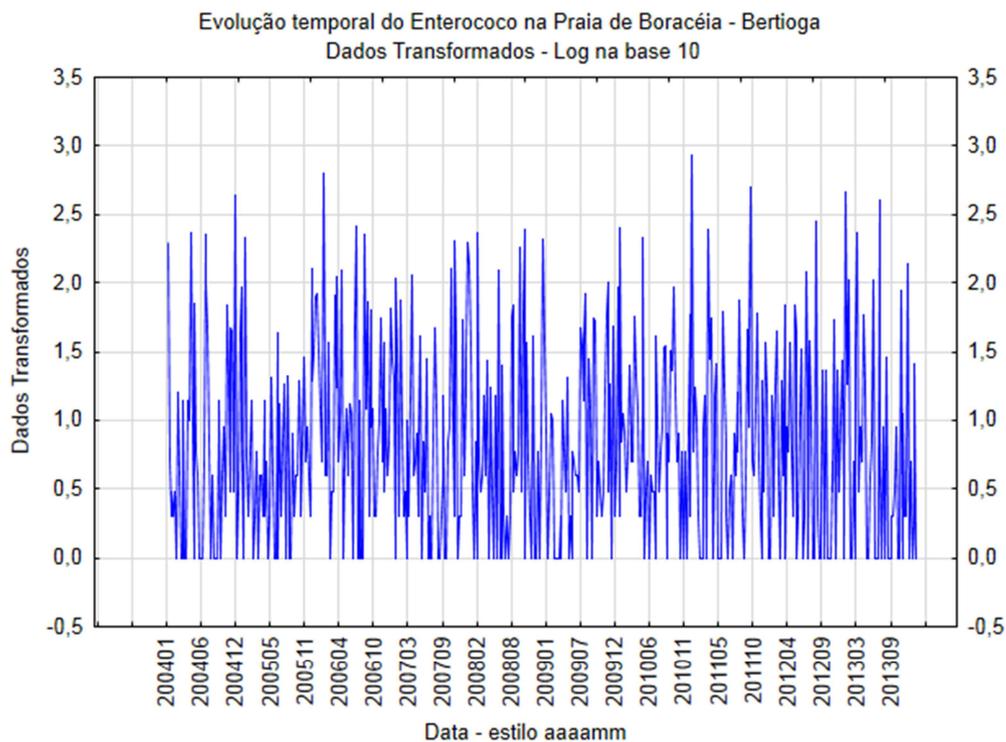


Gráfico 1: Evolução temporal da concentração de Enterococos - Praia de Boracéia - Bertioga.

A série pode ser exibida e decomposta em duas vertentes: anual (exceto os meses de janeiro) e apenas o mês de janeiro para todos os anos:

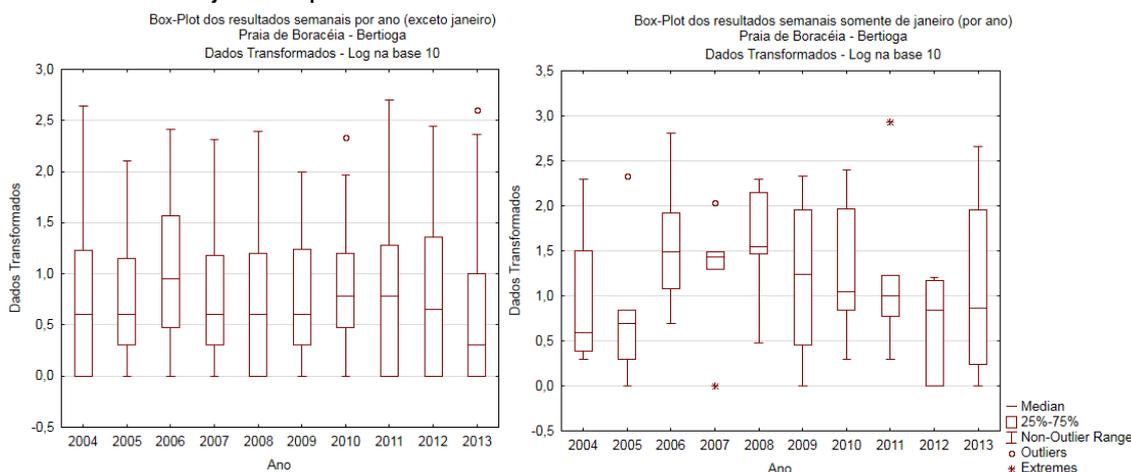


Gráfico 2: Box-Plot dos resultados semanais de enterococos por ano (exceto janeiro) e Box-Plot dos resultados semanais por ano apenas para os meses de janeiro - Praia de Boracéia – Bertioga.

A distribuição do parâmetro apresenta queda no último ano. Além disso, quando observado isoladamente os meses de janeiro dos anos, nota-se grande heterogeneidade do parâmetro. O gráfico 3 é apresentado com indicadores de tendência central e de valores extremos:

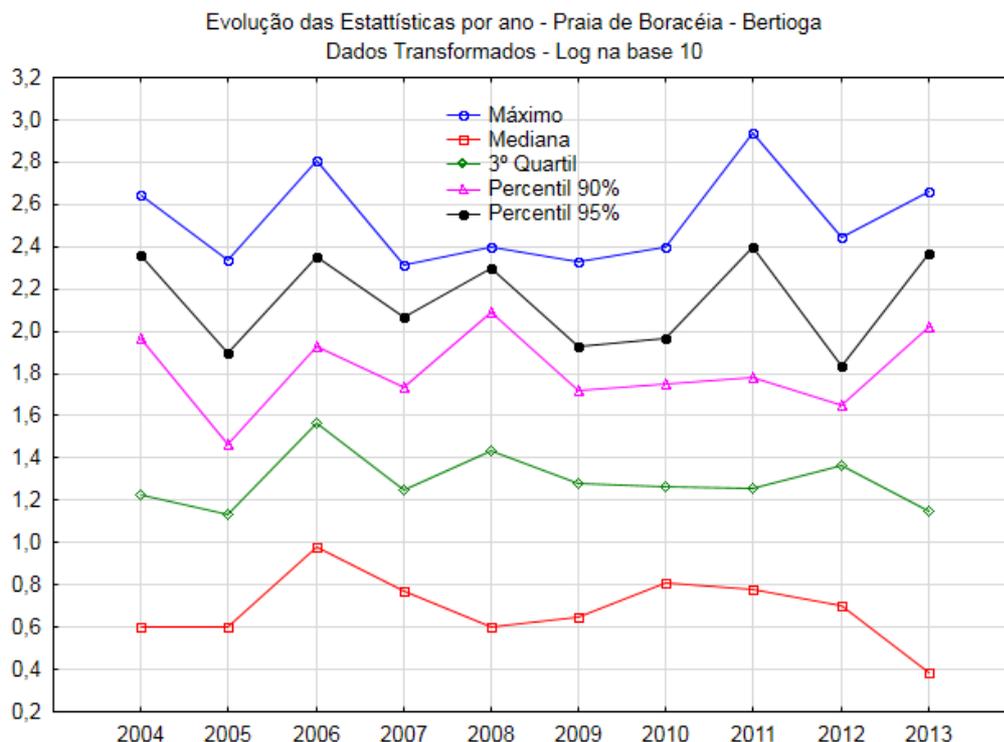


Gráfico 3: Distribuição dos resultados semanais de enterococos por ano – Praia de Boracéia - Bertioga.

O gráfico 3 é uma representação holística das principais estatísticas robustas associadas à distribuição do parâmetro e também, uma alternativa para “lapidar” a distribuição dos dados. O comportamento temporal da mediana e terceiro quartil não apresentaram sinais de tendência na redução do parâmetro.

Metodologia de Box & Jenkins

Identificação

Para a identificação dos modelos apropriados, deve-se analisar o gráfico do tempo da série em estudo, gráfico 4. A análise descritiva prévia do gráfico 4 revelou a presença de estacionariedade por todo o período do estudo.

O próximo passo foi analisar as funções de autocorrelações (FAC) e de autocorrelações parciais (FACP) da série. O comportamento dessas funções auxilia na verificação da estacionariedade e na proposição do modelo.

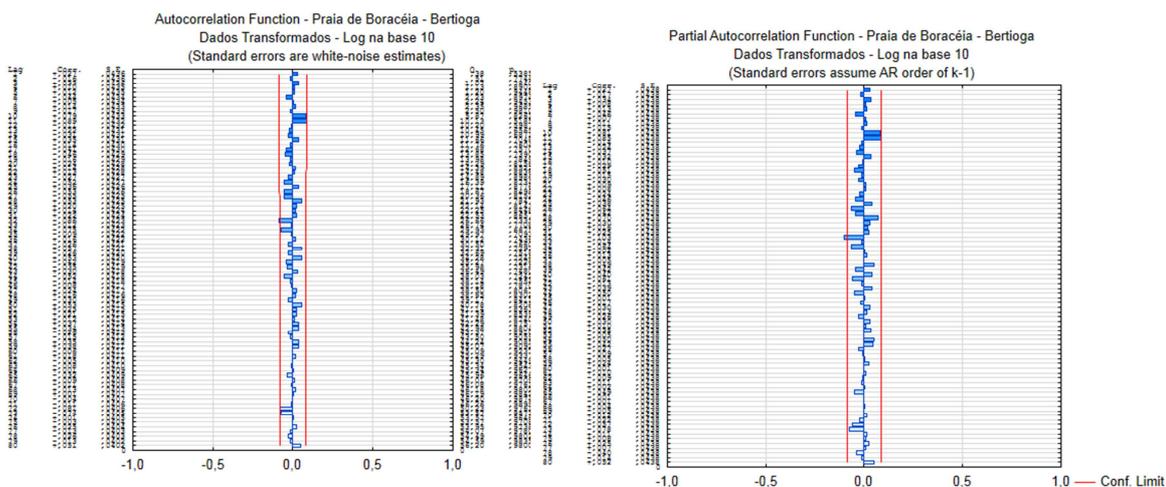


Gráfico 4: FAC e FACP da série transformada - Praia de Boracéia – Bertioga.

A função FAC mostra ausência da autocorrelação, apresenta todas as defasagens (lags não significativas). Assim há uma indicação de que a ordem do modelo auto-regressivo é $p = 1$, neste AR(1).

Normalmente, a ordem “d” da série é obtida através da FAC que apresentar a menor flutuação em suas defasagens após aplicadas as diferenciações. Mas, a série sem ser diferenciada já apresenta a FAC com a menor flutuação, portanto $d = 0$.

Estimação

Uma vez indicado os valores de p, d e q , passa-se para a estimativa dos parâmetros do modelo proposto. As estimativas são apresentadas na tabela 1:

Tabela 1: Estimativas dos parâmetros do modelo AR – Praia de Boracéia - Bertioga.

Paramet.	BERTIOGA_BORACÉIA Model: AR(1,0,0) MS Residual= ,49349					
	Param.	Asympt. Std.Err.	Asympt. t(520)	P	Lower 95% Conf	Upper 95% Conf
Constant	0,815750	0,031626	25,79332	0,000000	0,753619	0,877881
p(1)	0,026915	0,043940	0,61254	0,540450	-0,059406	0,113236

A série é então melhor descrita por um comportamento constante e igual a 0,81575 na escala logarítmica.

Verificação

Essa etapa consistiu em verificar se o modelo identificado foi adequado. Em caso negativo, será necessário identificar outro modelo e repetir as etapas de estimação e verificação. As formas de verificação utilizadas foram:

- **Análise de Resíduos:** os resíduos devem apresentar comportamento de “ruído branco” se o modelo estiver adequadamente especificado, isto é, suas correlações devem ser não significativas e seguirem distribuição normal. Para isto utilizou-se o teste de Ljung-Box para reforçar essa afirmativa.

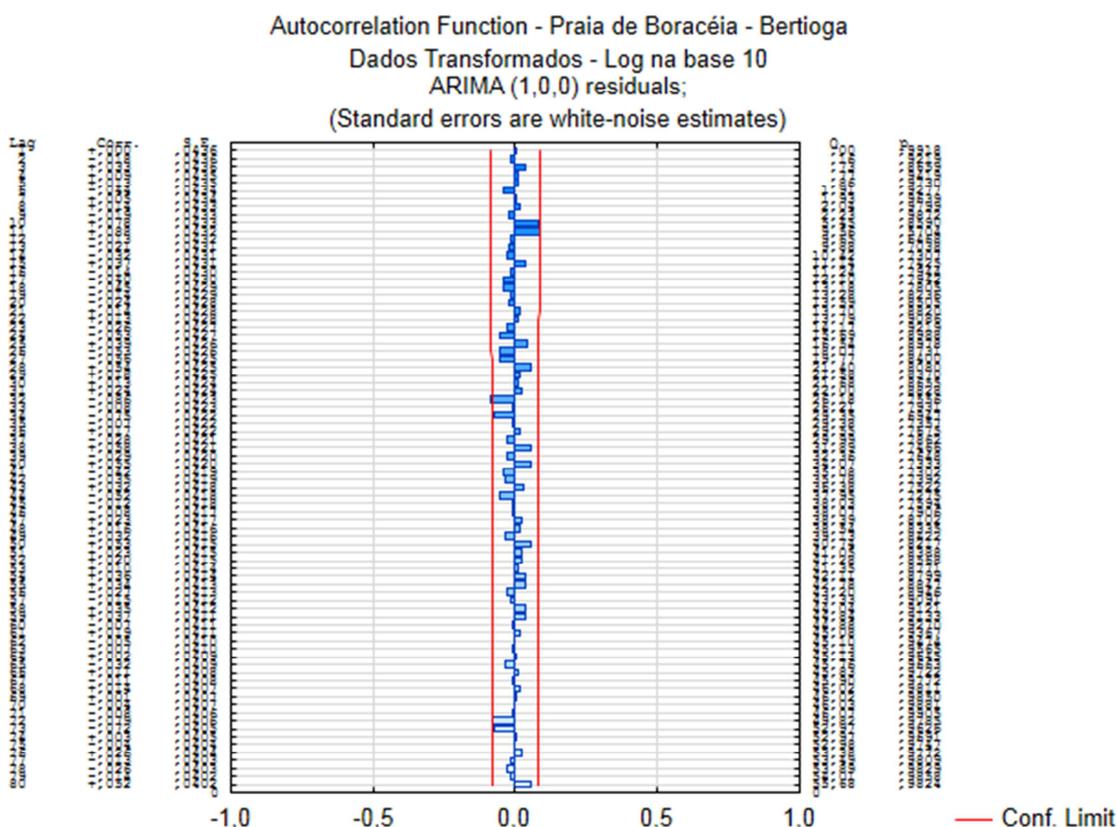


Gráfico 5: FAC dos resíduos - Praia de Boracéia – Bertiooga.

O modelo AR(1,0,0) proposto pela análise da FAC e FACP mostrou-se adequado, devido ao comportamento dos resíduos, que se comportam como ruído branco.

Com base nesta análise, **a conclusão geral** é que a série temporal da praia de Boracéia em Bertiooga **não apresenta tendência e sazonalidade estatisticamente significativas**.

Esta metodologia foi utilizada para análise das demais praias objeto do estudo e podem ser consultadas no relatório **“Desenvolvimento de Análise Estatística sobre dados de Pesquisas da Qualidade das Águas de São Paulo Estudo sobre a Qualidade das praias litorâneas”**.

3.1.2 ESTUDO TENDÊNCIA – REDE ÁGUAS INTERIORES

O objetivo deste estudo foi evidenciar estatisticamente a tendência de 10 parâmetros e um índice de qualidade ambiental em 16 pontos de águas superficiais da Rede da CETESB no Estado de São Paulo.

Resumo dos pontos por período.

Ponto	Parâmetros									
	Alumínio Dissolvido	Condutividade	DBO	<i>E. coli</i>	Ferro Dissolvido	Fósforo Total	IQA	Nitrato	Nitrogênio Amoniacal	Oxigênio Dissolvido
ATIB02065	2006-2012	1986-2012	1983-2012	1983-2012	2006-2012	1983-2012	1983-2012	1983-2012	1983-2012	1983-2012
ATIB02800	2007-2012	2002-2012	2000-2012	2000-2012	2007-2012	2003-2012	2007-2012	2003-2012	2003-2012	2000-2012
CARO02800	2006-2012	2001-2012	2001-2012	2001-2012	2006-2012	2001-2012	2001-2012	2001-2012	2001-2012	2001-2012
JAGR02800	2006-2012	1986-2012	1983-2012	1983-2012	2006-2012	1983-2012	1983-2012	1983-2012	1983-2012	1983-2012
JUNA04200	2011-2013	2002-2013	2000-2013	2002-2011	2011-2013	2003-2013	2009-2013	2003-2013	2003-2013	2000-2013
JUNA04900	2006-2012	1995-2012	1993-2012	1993-2012	2006-2012	1993-2012	1993-2012	1993-2012	1993-2012	1993-2012
MOGU02200	2006-2012	1986-2012	1983-2012	1983-2012	2006-2012	1983-2012	1983-2012	1983-2012	1983-2012	1983-2012
MOGU02300	2006-2012	1986-2012	1983-2012	1983-2012	2006-2012	1983-2012	1983-2012	1983-2012	1983-2012	1983-2012
PARB02300	2006-2012	1986-2012	1984-2012	1984-2012	2006-2012	1984-2012	1984-2012	1984-2012	1984-2012	1984-2012
PARB02900							1983-2012			
PCAB02800	2006-2012	1986-2012	1983-2012	1983-2012	2006-2012	1983-2012	1983-2012	1983-2012	1983-2012	1983-2012
PRET02800	2006-2012	1986-2012	1983-2012	1983-2012	2006-2012	1983-2012	1983-2012	1983-2012	1983-2012	1983-2012
RIBE02500	2006-2012	1986-2012	1983-2012	1983-2012	2006-2012	1983-2012	1983-2012	1983-2012	1983-2012	1983-2012
TIET04150							1983-2012			
TIPI04900	2006-2012	1986-2012	1983-2012	1983-2012	2006-2012	1983-2012	1983-2012	1983-2012	1983-2012	1983-2012
TURV02800	2006-2012	1986-2012	1983-2012	1983-2012	2006-2012	1983-2012	1983-2012	1983-2012	1983-2012	1983-2012

O estudo de tendência foi realizado por ponto de acordo com o período de dados disponíveis.

A título de exemplo segue a análise do ponto ATIB02065, no Rio Atibaia em Campinas, para o parâmetro *E. coli*.

**Análise de tendência do ponto ATIB02065 (Rio Atibaia – Captação de Campinas)
Parâmetro: *E.coli***

Para avaliação dos resultados os dados foram transformados usando logaritmo na base 10 devido à assimetria da distribuição dos dados.

Inicialmente avaliou-se o gráfico da série utilizando os valores observados em cada bimestre ao longo do período de medição. Ainda, transformou-se as informações do parâmetro a partir do operador matemático \log_{10} com o intuito de estabilizar a sua variabilidade.

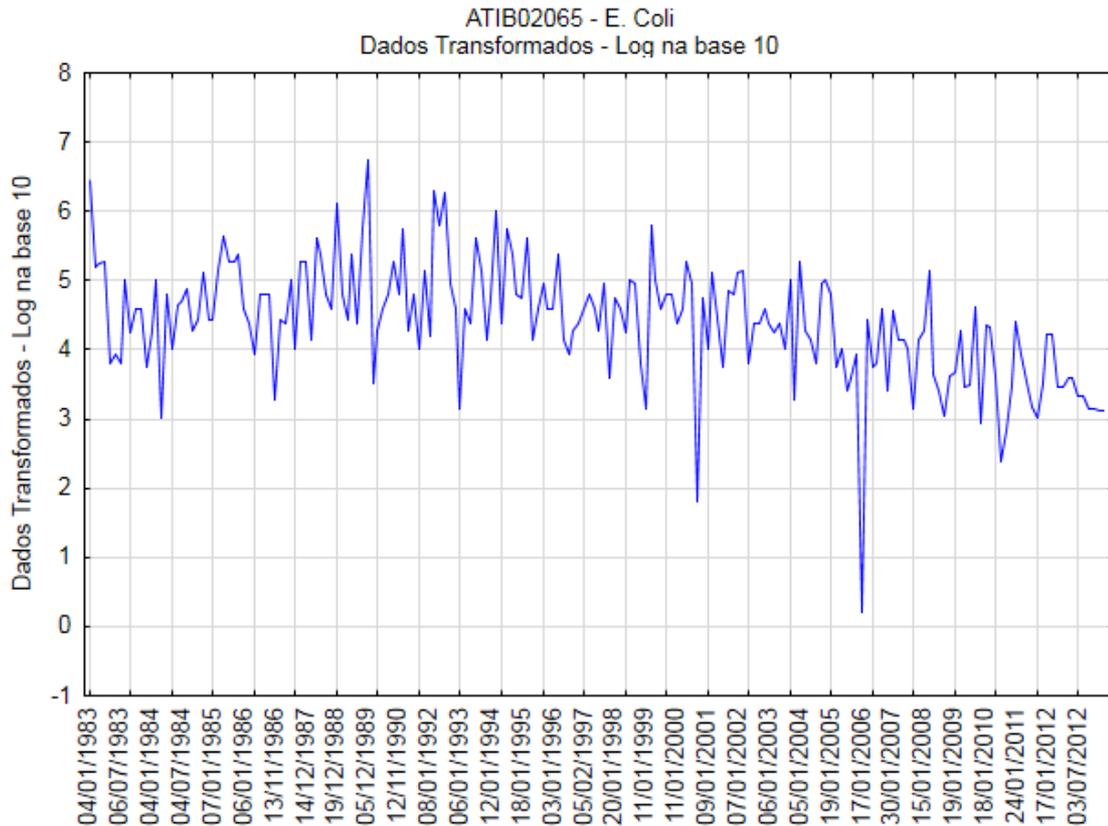


Gráfico 6 – Evolução temporal de E. Coli – ATIB02065.

A partir do Gráfico 6, verificou-se a ausência de estacionariedade da série, isto é, a série não converge para sua média e possui uma tendência polinomial. Há possível queda nos registros.

A série também pode ser exibida e decomposta em duas vertentes: anual e por períodos de chuvas e estiagens. Para este tipo de análise considerou-se todos os meses onde ocorreram medições.

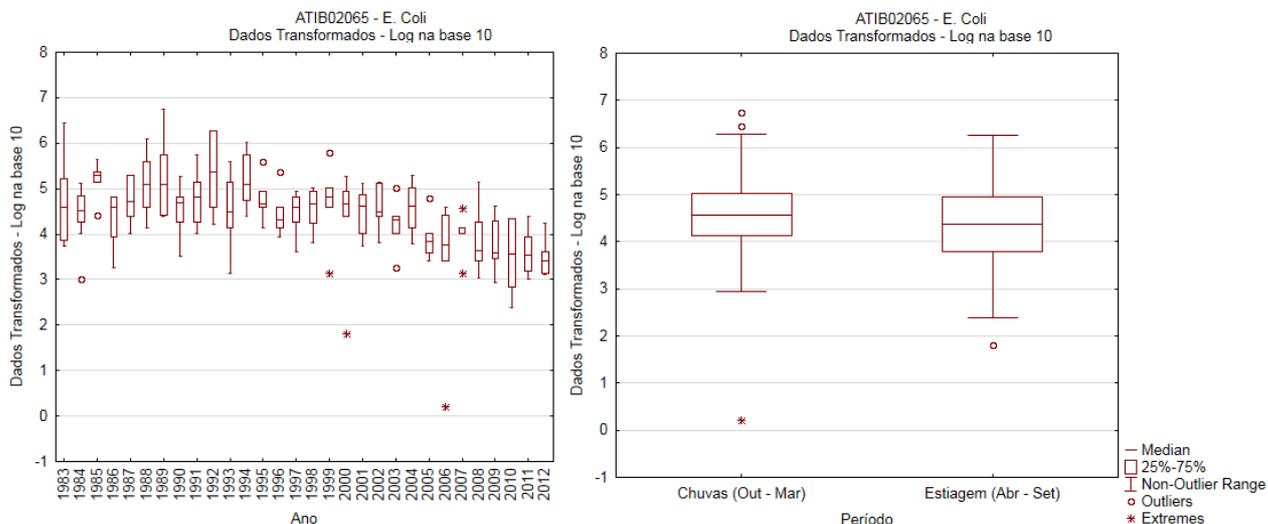


Gráfico 7 – BoxPlot por ano e por período – E. Coli – ATIB02065.

Por meio do Gráfico 7 verificou-se que as variabilidades dos registros permanecem semelhantes após o início dos anos 90. Também, não há evidências de que haja diferença significativa entre os registros do parâmetro em ambos os períodos do ano.

Para ter uma noção de evolução das estatísticas ao longo do tempo, elaborou-se o Gráfico 8, a partir do qual foi possível verificar queda nos valores do parâmetro.

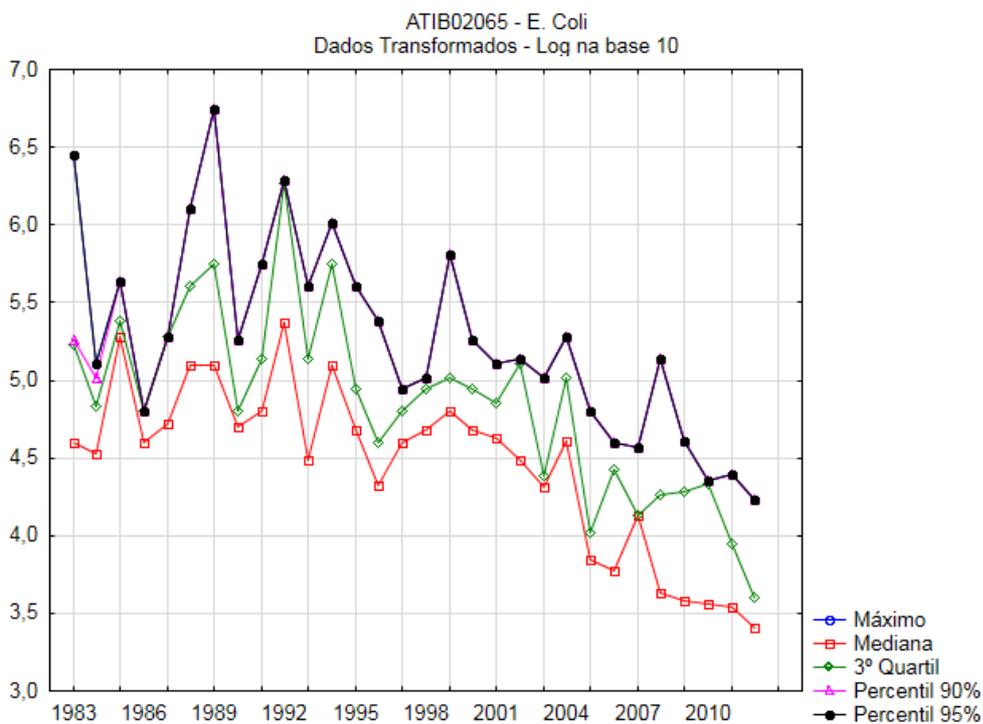


Gráfico 8 – Evolução das estatísticas descritivas, por ano, de E. Coli – ATIB02065.

Metodologia de Box & Jenkins

Identificação

Para modelar a série no tempo se fez necessário recorrer às médias de cada bimestre devido à presença de variabilidade nas medições de cada bimestre dependendo do ano. Dessa forma, construiu-se o gráfico a seguir.

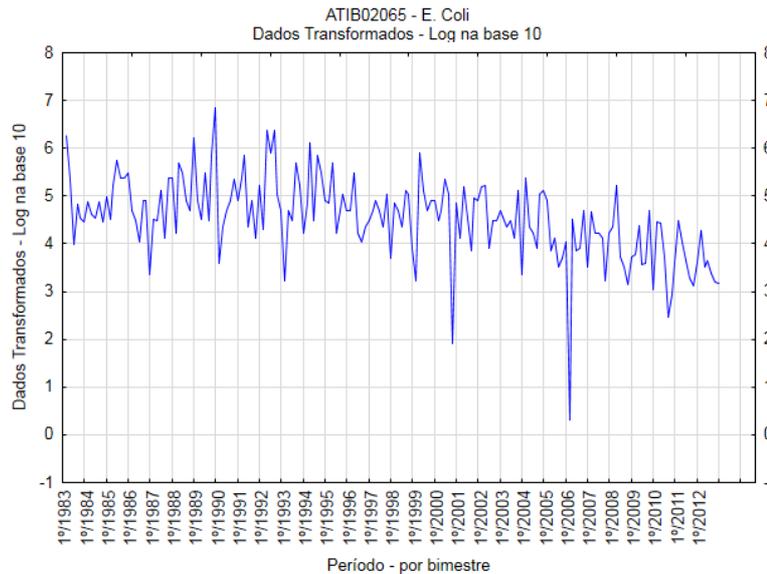


Gráfico 9 – Série das médias de cada bimestre dos dados de E. Coli – ATIB02065.

A seguir, as FAC e FACP dos valores médios de cada bimestre para os dados:

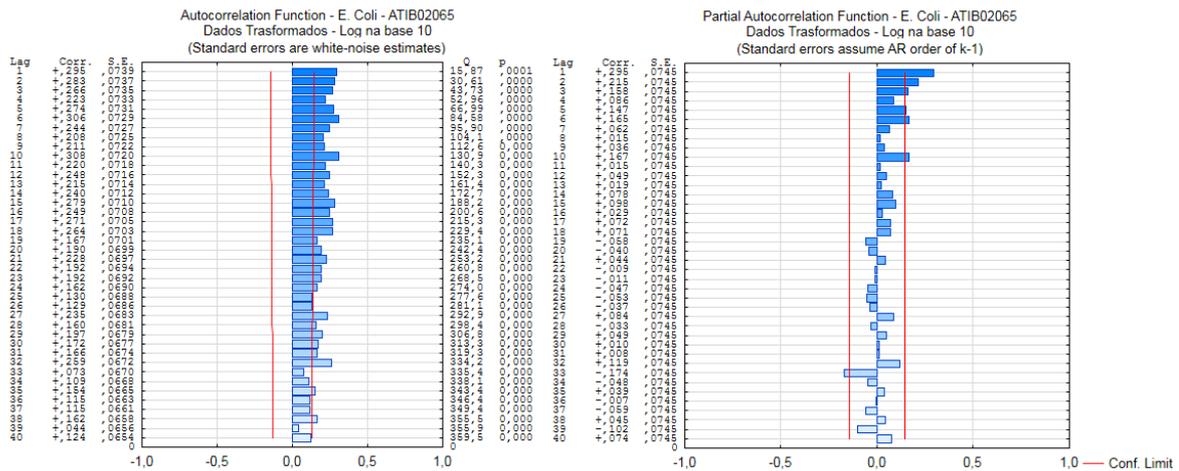


Gráfico 10 – FAC e FACP da série temporal das médias de cada bimestre dos dados de E. Coli – ATIB02065.

Através da análise descritiva e visual da série apresentada notou-se a ausência de estacionariedade, então foi iniciada a modelagem tirando a primeira diferença da série, isto é, define-se uma nova variável que representa a diferença entre duas observações no tempo

definidas por $Y_t = X_t - X_{t-1}$. A série diferenciada é apresentada no Gráfico 2.6, onde nota-se a presença de estacionariedade.

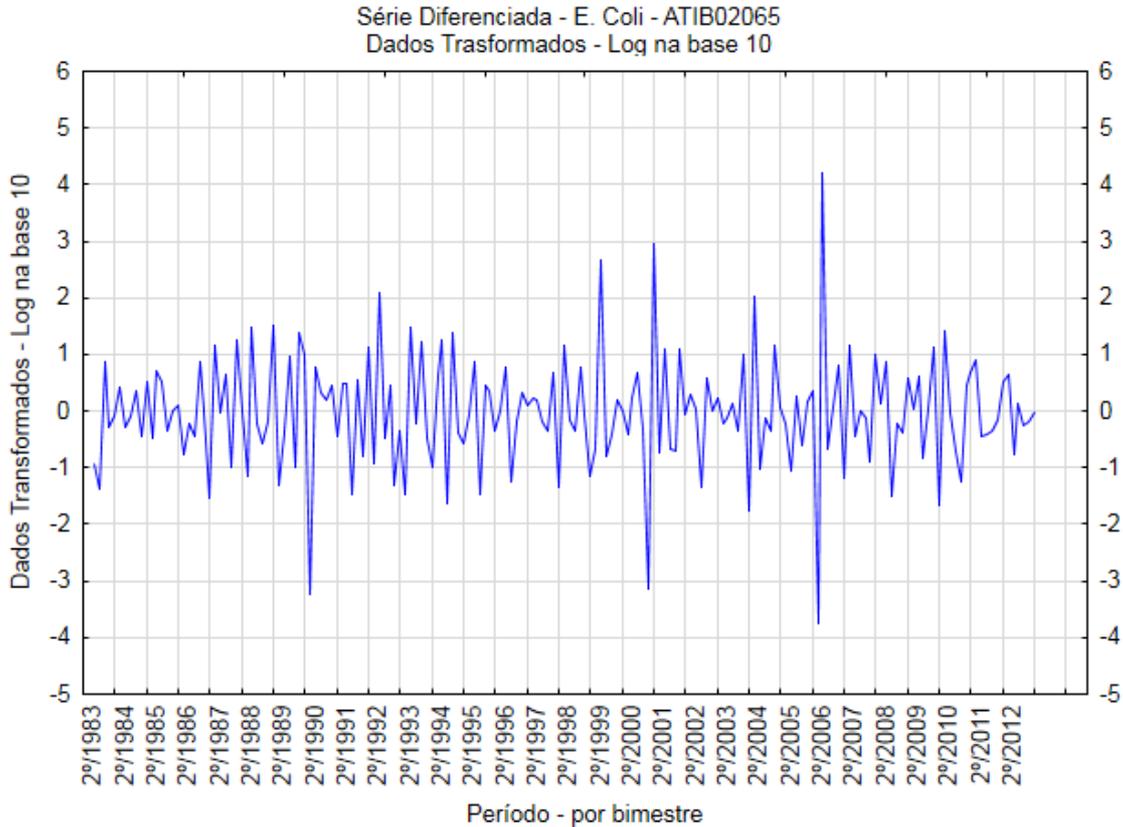


Gráfico 2.6 – Série temporal diferenciada de E. Coli – ATIB02065.

O próximo passo foi analisar as funções de autocorrelações (FAC) e de autocorrelações parciais (FACP) da série diferenciada. O comportamento dessas funções auxiliou na verificação da estacionariedade e na identificação da ordem do modelo.

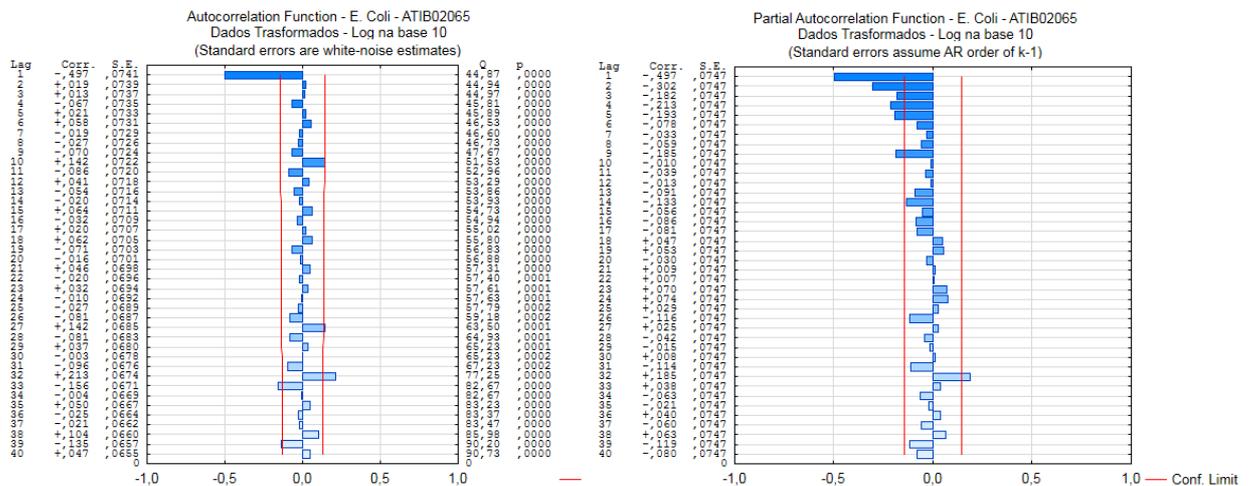


Gráfico 11 – FAC e FACP da série temporal diferenciada de E. Coli – ATIB02065.



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO

Diretoria de Engenharia e Qualidade Ambiental

Por meio do Gráfico 11, iniciou-se o processo de identificação da ordem do modelo. Para identificação do modelo observou-se que a FAC apresentou valores significativos para o lag 32 (possível sazonalidade) e no lag 1, indicando parâmetro de Média Móvel igual a 1. E a FACP foi significativa para o lag 32 (indicando possível sazonalidade) e um decaimento leve até o lag 5.

Dessa forma, foi testado inicialmente um modelo SARIMA(0,1,1)(0,0,1)₃₂ e que se mostrou adequado, pois tanto a FAC quanto a FACP para os resíduos apresentaram valores insignificantes para os lags.

Assim, sugeriu-se o modelo SARIMA(0,1,1)(0,0,1)₃₂.

Verificação

Para validação do modelo escolhido SARIMA(0,1,1)(0,0,1)₃₂ foi necessário analisar seus resíduos e verificar se eles poderiam ser considerados um processo ruído branco, ou seja, que as correlações entre tempos distintos fossem não significativas e próximas de zero.

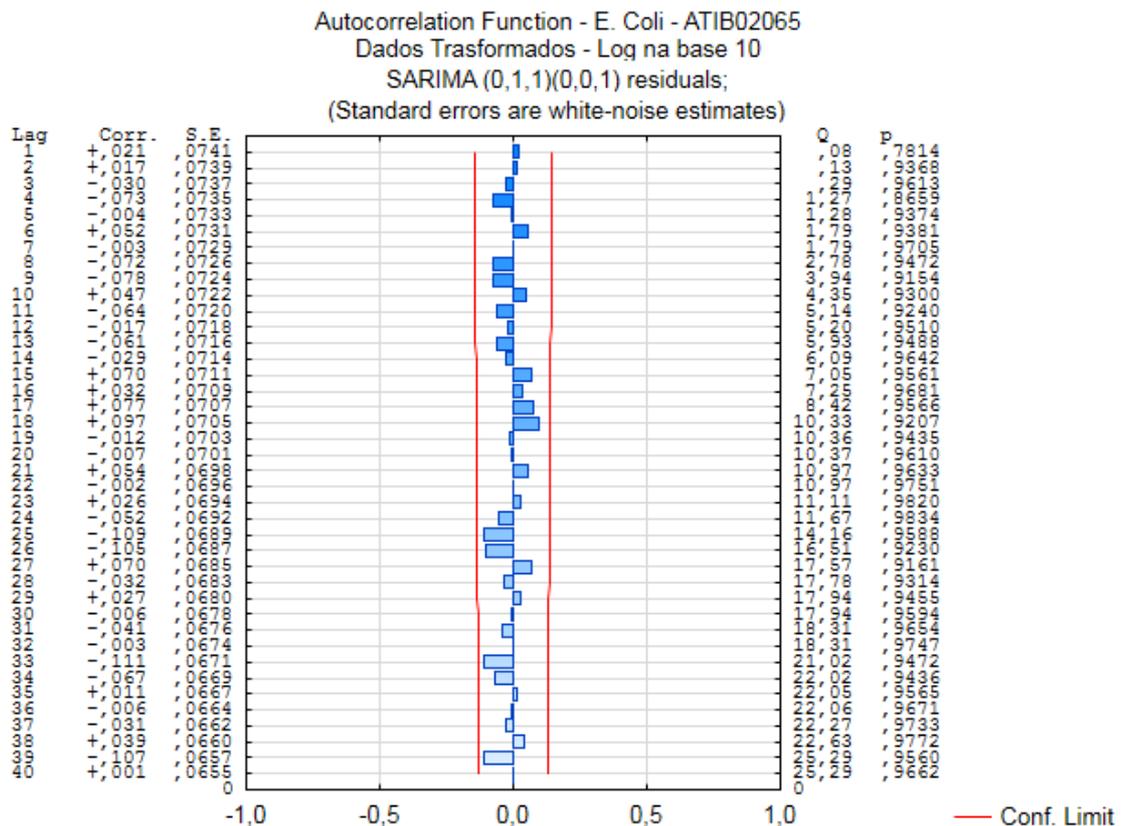


Gráfico 12 – FAC dos resíduos – E. Coli – ATIB02065.

O Gráfico 12 mostra a FAC dos resíduos do modelo proposto. Notou-se que nenhum lag é significativamente diferente de 0 e também pelo teste de Ljung-Box (p-valor apresentado na coluna p no lado direito do gráfico), onde testou-se se as correlações conjuntas até o lag k foram iguais a zero. Dessa forma, não houve evidências de que as correlações conjuntas até o lag 40 foram significativamente diferentes de zero, o que indicou que o modelo foi corretamente especificado.

Estimação

Uma vez identificada a ordem do modelo, estimou-se os seus respectivos parâmetros. As estimativas são apresentadas na tabela 2

Tabela 2 – Estimativa dos parâmetros do modelo – E. Coli – ATIB02065.

Paramet.	Dados Transformados - Log na base 10: ATIB02065_EColi Model: SARIMA(0,1,1)(0,0,1) Seasonal lag: 32 MS Residual= ,57643					
	Param.	Asympt. Std.Err.	Asympt. t(177)	p	Lower 95% Conf	Upper 95% Conf
q(1)	0,838880	0,043713	19,19043	0,000000	0,752613	0,925147
Qs(1)	-0,160430	0,078648	-2,03986	0,042851	-0,315637	-0,005222

Verifica-se que todos os parâmetros do modelo são significativos, ou seja, o modelo está bem especificado.

Conclusão

O modelo SARIMA(0,1,1)(0,0,1)₃₂ proposto pela análise da FAC e FACP se mostrou adequado, devido ao comportamento dos resíduos, que se comportam como ruído branco.

Com base nessa análise, a **conclusão** é que a **série apresenta tendência de queda** estatisticamente significativa com sazonalidade.

Esta metodologia foi utilizada para análise dos demais pontos objeto do estudo, podendo ser consultados no relatório “**Desenvolvimento de Análise Estatística sobre dados de Pesquisas da Qualidade das Águas de São Paulo Estudo sobre a Qualidade das Águas Superficiais**”. (Arquivos em CD anexo)

3.1.3 ESTUDO TENDÊNCIA – REDE AUTOMÁTICA

O objetivo deste estudo foi evidenciar estatisticamente a tendência dos parâmetros qualitativos (oxigênio dissolvido, pH, condutividade, turbidez e temperatura) e quantitativos (vazão ou volume), referente aos resultados obtidos no monitoramento das Estações Laranjal Paulista, Guarapiranga e Taquacetuba da Rede Automática da CETESB.

O estudo de tendência foi realizado por estação automática de acordo com o período de dados disponíveis.

A título de exemplo segue a análise da estação Laranjal Paulista para o parâmetro Condutividade Elétrica.

Análise de Tendência na Estação Laranjal Paulista (Rio Tietê - TIET02450)

Parâmetro: Condutividade Elétrica

O gráfico 13 apresenta a série temporal dos valores médios diários antes e após a recomposição e aplicação dos critérios de inclusão de dados. O total de casos válidos para

cada uma das séries temporais é apresentado na Tabela 3. Observa-se que, de acordo com a metodologia utilizada, os anos de 2008 e 2012 foram excluídos da análise.

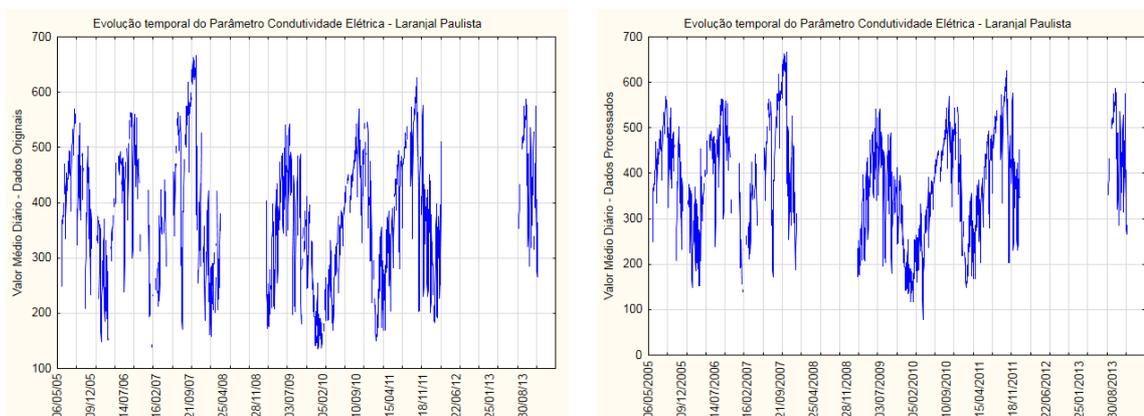


Gráfico 13: Valores médios diários para o parâmetro condutividade elétrica antes da recomposição (esquerda) e após a recomposição e aplicação dos critérios de inclusão de dados (direita).

Tabela 3: Total de casos válidos, por ano, antes e após recomposição e aplicação dos critérios de inclusão de dados.

Ano	Antes		Depois	
	Casos Válidos	Porcentagem de Casos Válidos	Casos Válidos	Porcentagem de Casos Válidos
2005	167	45,75	185	50,68
2006	227	62,19	266	72,88
2007	247	67,67	271	74,25
2008	72	19,67	0	0,00
2009	288	78,90	308	84,38
2010	291	79,73	329	90,14
2011	327	89,59	341	93,42
2012	68	18,58	0	0,00
2013	108	29,59	108	29,59

A Figura 1 apresenta o Box-Plot dos valores médios diários por ano, sem distinção de meses, e por mês, sem distinção de anos. Em ambos os casos, apresenta-se os gráficos para os dados antes da recomposição e após a recomposição e aplicação dos critérios de inclusão de dados.

Observa-se uma variação nos valores médios diários ao longo dos anos. Além disso, um indicativo de que a partir de 2009 os valores médios do parâmetro estão crescendo. Em relação ao comportamento dos valores médios diários, ao longo dos meses, observa-se uma possível sazonalidade. A maior concentração de *outliers* (observações atípicas) encontra-se nos meses de janeiro, julho, agosto e setembro.

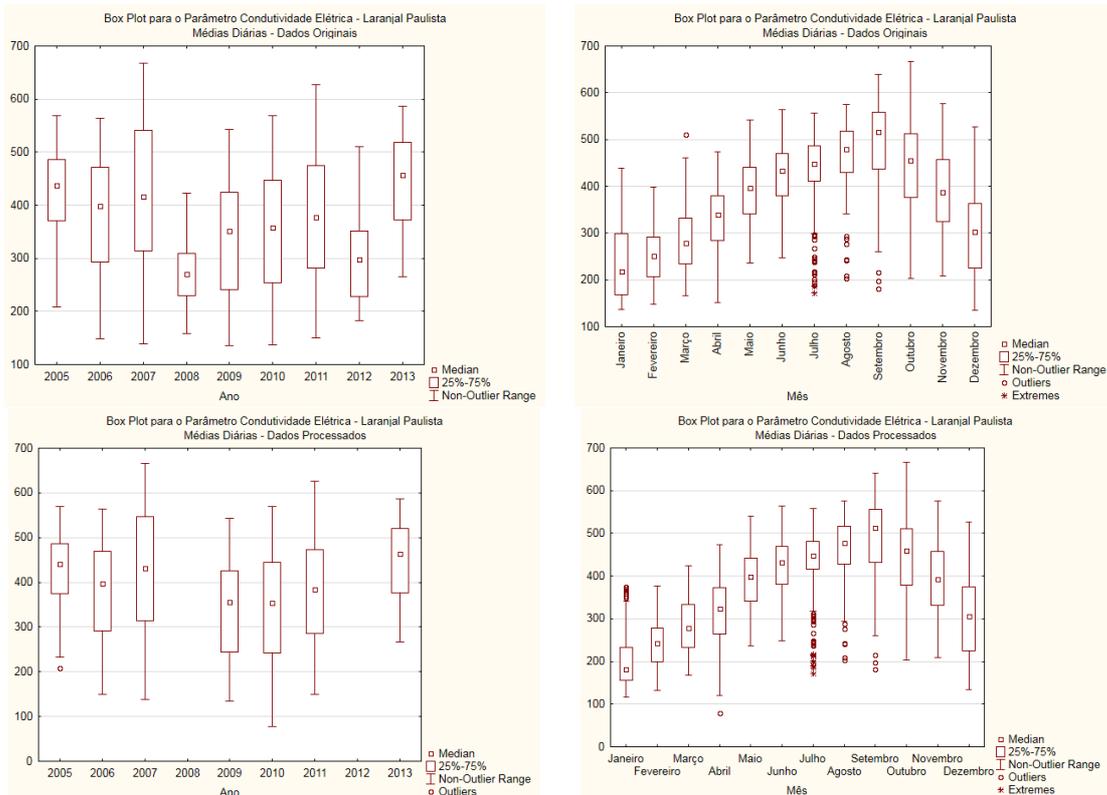


Figura 1: Box-Plot dos valores médios diários por ano (esquerda) e por mês (direita), para os dados antes da recomposição (primeira linha, de cima para baixo) e após a recomposição e aplicação dos critérios de inclusão de dados (segunda linha, de cima para baixo).

O gráfico 14 apresenta a série temporal das principais estatísticas descritivas associadas à distribuição do parâmetro, para os dados antes da recomposição e após a recomposição e aplicação dos critérios de inclusão de dados. Observa-se tanto crescimento quanto decréscimo dos valores ao longo dos anos de forma que não é possível afirmar qual o padrão apresentado pelas estatísticas.

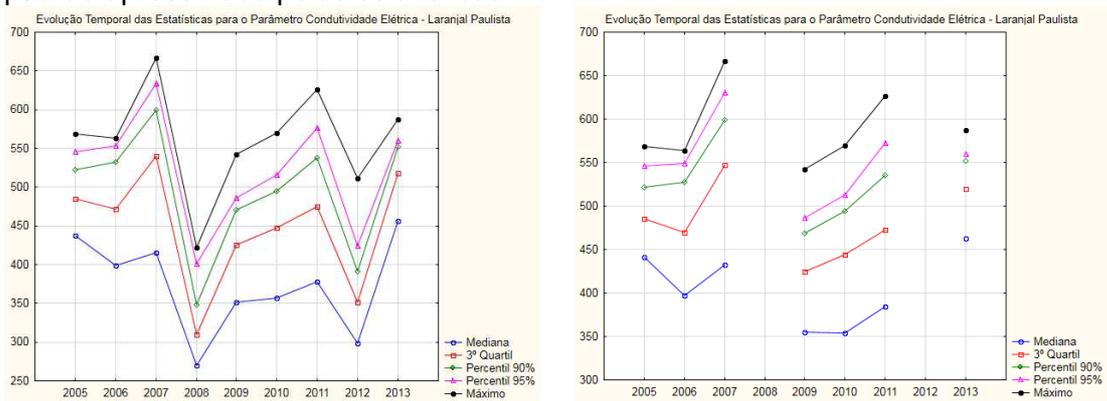


Figura 1.3: Estatísticas descritivas (mediana, Q75, P90, P95 e máximo) para o parâmetro condutividade elétrica antes da recomposição (esquerda) e após a recomposição e aplicação dos critérios de inclusão de dados (direita).



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO
Diretoria de Engenharia e Qualidade Ambiental

Seguindo os critérios adotados na metodologia, aplica-se a regressão robusta para avaliação de tendência do parâmetro. Os resultados da estimação são apresentados na Tabela 1.2 e indicam a existência de tendência de decrescimento ($\beta_1 = -0,0188$, p-valor = 0,0000).

Tabela 4: Regressão linear robusta para o parâmetro Condutividade Elétrica.

```
Conductividade Elétrica
O algoritmo convergiu após 40 iterações. Erro quadrático = 7.538393e-07

Call:
robust::lmRob(formula = yy ~ t)

Residuals:
    Min       1Q   Median       3Q      Max
-346.3263  -65.2476  -0.4889   -0.1206  225.6876

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 4.606e+02  1.534e+00  300.23  <2e-16 ***
t           -1.881e-02  7.648e-04  -24.59  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 29.73 on 3103 degrees of freedom
Multiple R-Squared: 0.06061

Test for Bias:
      statistic p-value
M-estimate    5.459 0.06525
LS-estimate  123.912 0.00000
```

Para a validação do modelo de regressão robusta foram realizados testes de hipóteses para verificar a normalidade (teste Shapiro-Wilk) e a existência de correlação entre os resíduos (testes Box-Pierce e Box-Ljung). Os resultados dos testes são reportados a seguir. Verifica-se que, para o modelo da Tabela 4 tais suposições não foram satisfeitas.

```
Shapiro-Wilk normality test
data: m$residuals
W = 0.8868, p-value < 2.2e-16

Box-Pierce test
data: m$residuals
X-squared = 57131, df = 20, p-value < 2.2e-16

Box-Ljung test
data: m$residuals
X-squared = 57360, df = 20, p-value < 2.2e-16
```

Conclusão

Devido à presença de correlação serial entre os resíduos do modelo e a não normalidade dos mesmos, não é possível realizar inferência sobre a significância dos coeficientes e, portanto, o resultado da análise foi inconclusivo.

Esta metodologia foi utilizada para análise dos demais parâmetros das estações automáticas objeto do estudo, podendo ser consultados no relatório “**Desenvolvimento de Análise Estatística sobre dados de Pesquisas da Qualidade das Águas de São Paulo Estudo sobre a Qualidade das Águas de Rede Automática**”. (Arquivos em CD anexo)



3.1.4 ESTUDO TENDÊNCIA – REDE ÁGUAS SUBTERRÂNEAS

O objetivo deste estudo foi evidenciar estatisticamente a tendência de 8 parâmetros de qualidade ambiental em 10 poços de águas subterrâneas monitorados pela Rede da CETESB no Estado de São Paulo.

Os parâmetros analisados foram: Nitrato Total, Condutividade, pH, Cloreto Total, Sódio Total, Potássio Total, Cálcio Total e Magnésio Total.

Aspectos Descritivos

- Série Temporal para a totalidade das medições e por período (seco e chuvoso);
- Observar o Box-Plot Temporal apenas para os períodos (seco e chuvoso);
- Avaliar a tendência através da metodologia de Regressão Linear Robusta e do algoritmo EM (Expectation-Maximization). Neste trabalho, o interesse foi apenas estimar e testar hipótese sobre o parâmetro de inclinação ou coeficiente angular.

O estudo de tendência foi realizado por ponto de acordo com o período de dados disponíveis.

A título de exemplo segue a análise do ponto São José do Rio Preto para o parâmetro Nitrato.

Análise de tendência do ponto SÃO JOSÉ DO RIO PRETO (127)

Parâmetro: NITROGÊNIO NITRATO

A série pode ser exibida e decomposta em duas vertentes: períodos chuvoso e seco. Para este tipo de análise considerou-se todos os anos onde ocorreram medições.

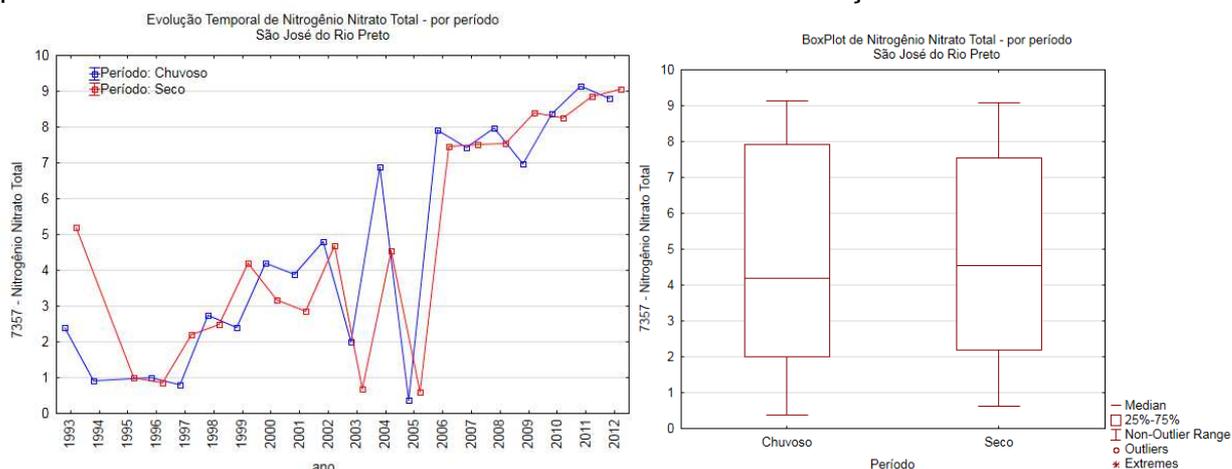


Gráfico 15: Evolução Temporal e BoxPlot dos resultados por período – Nitrogênio Nitrato – São José do Rio Preto.

O Gráfico 15 mostra tendência de aumento nos valores do parâmetro, independente do período, com redução atípica nos anos de 2003 e 2005.

Processo de Inferência

Para contornar o problema da escala temporal na variável de entrada x , padronizou-se calculando a quantidade de dias a partir de 01/01/1993, ano da primeira medição, até a data de cada medição.

Para estimar os coeficientes do modelo para Nitrogênio-Nitrato, foi utilizada a metodologia EM em conjunto com Regressão Linear Robusta, devido à presença de missings durante período de coleta dos dados.

Os coeficientes da reta de regressão foram estimados e seus valores seguem na tabela 5:

Tabela 5: Coeficientes estimados para a reta de regressão – Nitrogênio Nitrato – São José do Rio Preto.

Parâmetros	Valor	Erro Padrão	t de Student	p-valor
b (intercepto)	0.0314	0.3533	0.0888	0.9297
a (coef. angular)	0.0013	0.0001	15.6334	0.0000

O gráfico 16 de dispersão conjugado com reta estimada:

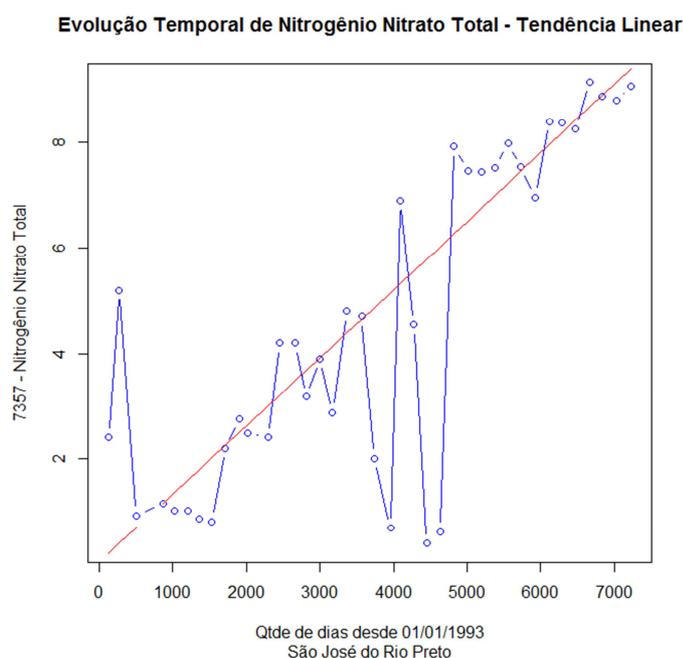


Gráfico 16 - Evolução Temporal desde 01/01/1993: Nitrogênio Nitrato – Tendência Linear – São José do Rio Preto.

Os resultados da estimação dos coeficientes da regressão indicam a existência de tendência de crescimento para o parâmetro ($a = 0.0013$, p -valor = 0.0000). A validação deste resultado é dada após a análise residual.

Análise dos Resíduos – Validação do Modelo

A descrição da análise dos resíduos para verificar a adequabilidade do modelo de reta de regressão estimado é apresentada a seguir.



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO
Diretoria de Engenharia e Qualidade Ambiental

A normalidade dos resíduos foi testada utilizando-se o teste de Shapiro-Wilk. Os resultados do teste são descritos na Tabela 6. Observa-se que, para $\alpha=0,01$, o teste de Shapiro-Wilk não rejeita a hipótese de que os resíduos atendem a suposição de normalidade exigida para o modelo.

Tabela 6: Teste de Shapiro-Wilk para normalidade dos resíduos – Nitrogênio Nitrato – São José do Rio Preto.

Estatística de Shapiro-Wilk	Valor	p-valor
W	0.9855	0.8859

Para testar a correlação entre os resíduos consideraram-se os testes de Ljung-Box e Box-Pierce. Ambos não rejeitaram a hipótese nula, portanto a conclusão é que não existe correlação entre os resíduos.

Box-Ljung test

Box-Pierce test

data: teste\$residuals
X-squared = 26.0158, df = 20, p-value = 0.1653

data: teste\$residuals
X-squared = 18.7389, df = 20, p-value = 0.5389

Conclusão

A série de resultados do monitoramento do poço São José do Rio Preto **apresenta tendência de crescimento estatisticamente significativa.**

Esta metodologia foi utilizada para análise dos demais poços objeto do estudo e podem ser consultados no relatório “**Desenvolvimento de Análise Estatística sobre dados de Pesquisas da Qualidade das Águas de São Paulo Estudo sobre a Qualidade das Águas Subterrâneas**”. (Arquivos em CD anexo)

3.2 ESTUDOS DE CORRELAÇÃO

Os estudos de correlação foram realizados com base no monitoramento das seguintes redes da CETESB:

- Águas Interiores (Rede Básica);
- Rede de Monitoramento Automático de Águas Superficiais.

Nos itens a seguir são apresentados os principais aspectos com relação aos resultados obtidos nos respectivos estudos, porém todo conteúdo pode ser consultado, na íntegra, nos relatórios apresentados pela empresa contratada. (Arquivo em CD anexo)

3.2.1 ESTUDO CORRELAÇÃO – REDE ÁGUAS INTERIORES

O objetivo deste estudo foi evidenciar estatisticamente as correlações entre pares de parâmetros de interesse envolvidos em processos de eutrofização de reservatórios, estudos de similaridades entre variáveis e influência de arraste do solo em águas interiores do Estado de São Paulo.

A título de exemplo segue a análise de correlação de parâmetros em processo de eutrofização de reservatórios em ponto na Represa Billings.



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO
Diretoria de Engenharia e Qualidade Ambiental

Eutrofização de Reservatórios

Ponto BILL02030 (Represa Billings próximo à barragem de Pedreira)

Na tabela abaixo apresenta-se algumas estatísticas descritivas para os parâmetros analisados no ponto BILL02030. Observa-se que o desvio padrão do parâmetro "Cobre Dissolvido" é igual a zero. Uma inspeção no banco de dados revelou que tal parâmetro apresenta valores constantes (0,009). Sendo assim, tal parâmetro foi removido da análise.

Tabela 7: Estatísticas Descritivas do ponto BILL02030.

Parâmetro	N	Média	Mínimo	Máximo	Std.Dev.
Clorofila-a	36	154,810	10,160	794,770	201,477
Cobre Dissolvido	37	0,009	0,009	0,009	0,000
Fósforo Total	45	0,212	0,007	0,800	0,160
Número de Células de Cianobactérias	37	250785,649	1600,000	1326150,000	331058,568
Temperatura da Água	42	22,639	17,200	28,900	3,020

Para as análises que seguem todos os parâmetros FT, CL e T foram mantidos em sua escala original enquanto que a série do parâmetro NCC passou pela transformação log10.

A Tabela 8 mostra os valores dos coeficientes de correlação para os cinco parâmetros estudados, enquanto que a figura que segue apresenta os gráficos de dispersão para os pares de parâmetros utilizados nos cálculos das correlações. As correlações para os pares de interesse (isto é, os pares descritos na Tabela 7) aparecem destacadas em negrito. A conclusão dessa análise é dada a seguir.

Tabela 8: Coeficiente de correlação de para os parâmetros.

Parâmetro	Clorofila-a	Fósforo Total	Log10 de Número de Células de Cianobactérias	Temperatura da Água
Clorofila-a	1,000000	0,798993	0,661565	0,228935
Fósforo Total	0,798993	1,000000	0,567597	0,151996
Log10 de Número de Células de Cianobactérias	0,661565	0,567597	1,000000	0,224605
Temperatura da Água	0,228935	0,151996	0,224605	1,000000

Nota: Para os valores fora da diagonal principal, os valores são estatisticamente diferentes de zero (p -valor < 0,05).

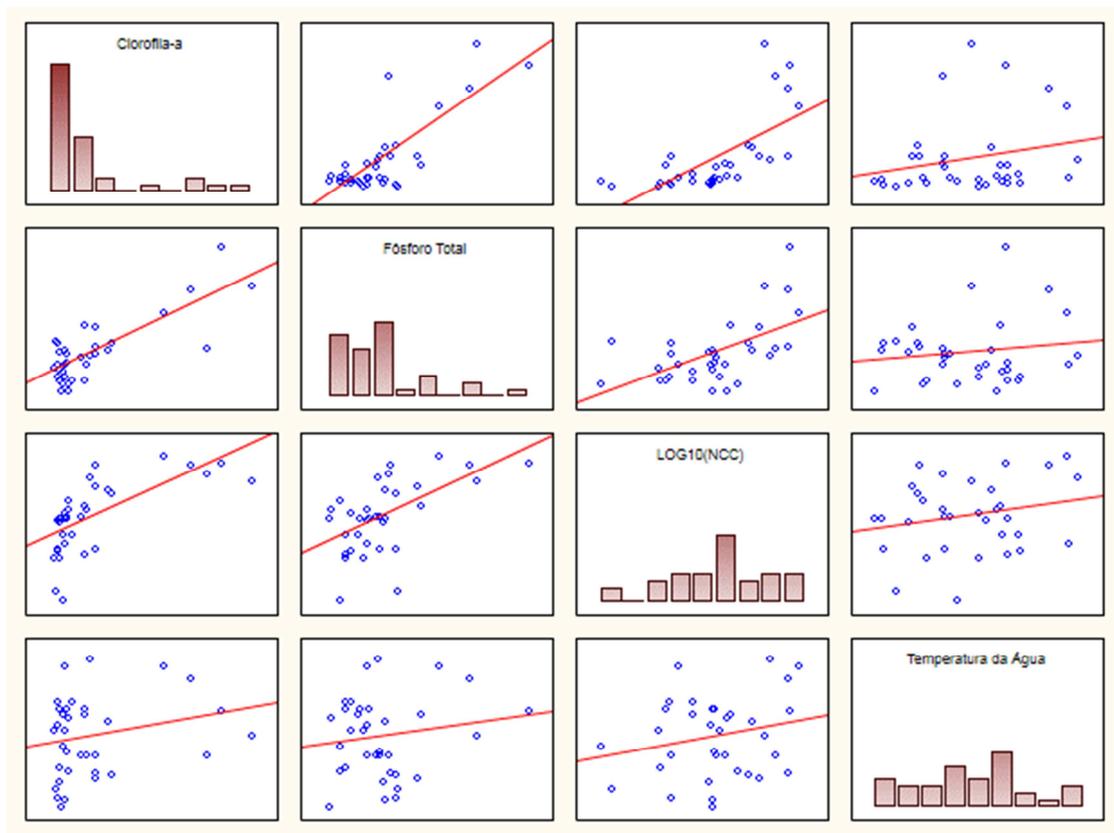


Figura 2: Histogramas (diagonal) e gráficos de dispersão para os parâmetros analisados.
Variável dependente: LOG10(NCC).
Variáveis independentes: Clorofila-a, Cobre Dissolvido, Fósforo Total, Temperatura da Água.

Os resultados da análise de regressão são apresentados nas tabelas que seguem. Os resultados reportados, tanto para a regressão simples quanto para a regressão múltipla, correspondem aos modelos selecionados conforme os critérios definidos na descrição da metodologia. As conclusões são reportadas após as tabelas.



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO
Diretoria de Engenharia e Qualidade Ambiental

Tabela 9: Regressão linear simples com “Fósforo Total” como variável independente.

```
Residuals:
      Min       1Q   Median       3Q      Max
-1.68200 -0.39912 -0.06179  0.24734  0.78765

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.6826     0.1356  34.527 < 2e-16 ***
x$FT        2.0357     0.4971   4.095 0.000237 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5146 on 35 degrees of freedom
Multiple R-Squared: 0.2737

Test for Bias:
              statistic p-value
M-estimate    0.5831  0.7471
LS-estimate    2.5403  0.2808
8 observations deleted due to missingness

      shapiro-wilk normality test

data:  simples$residuals
W = 0.9345, p-value = 0.03094
```

Tabela 10: Regressão linear simples com “Clorofila-a” como variável independente.

```
Residuals:
      Min       1Q   Median       3Q      Max
-1.33717 -0.38269  0.07738  0.18114  0.82189

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.8562674  0.0929314  52.256 < 2e-16 ***
x$CL         0.0017594  0.0003585   4.907 2.27e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3958 on 34 degrees of freedom
Multiple R-Squared: 0.3925

Test for Bias:
              statistic p-value
M-estimate    1.808  0.4050
LS-estimate    2.328  0.3122
9 observations deleted due to missingness

      shapiro-wilk normality test

data:  simples$residuals
W = 0.9487, p-value = 0.0951
```



COMPANHIA AMBIENTAL DO ESTADO DE SÃO PAULO
Diretoria de Engenharia e Qualidade Ambiental

Tabela 11: Regressão linear simples com “Temperatura” como variável independente.

```
Residuals:
  Min       1Q   Median       3Q      Max
-1.44684 -0.42877  0.04987  0.40875  1.03281

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.08471    1.16978   3.492  0.00142 **
x$T          0.04343    0.05099   0.852  0.40071
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6511 on 32 degrees of freedom
Multiple R-Squared:  0.0483

Test for Bias:
              statistic p-value
M-estimate    1.1602  0.5598
LS-estimate   -0.6008  1.0000
11 observations deleted due to missingness

      shapiro-wilk normality test

data:  simples$residuals
W = 0.9735, p-value = 0.5645
```

Tabela 12: Regressão linear múltipla com seleção de variáveis pelo método *stepwise*.

```
Residuals:
  Min       1Q   Median       3Q      Max
-1.33717 -0.38269  0.07738  0.18114  0.82190

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.8562642  0.0929387  52.252 < 2e-16 ***
x$CL         0.0017594  0.0003586   4.907 2.27e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3957 on 34 degrees of freedom
Multiple R-Squared:  0.3925

Test for Bias:
              statistic p-value
M-estimate    1.809  0.4047
LS-estimate    2.330  0.3120
9 observations deleted due to missingness

      shapiro-wilk normality test

data:  last$residuals
W = 0.9487, p-value = 0.0951
```

Conclusões - Correlação:

- log₁₀(NCC) vs. FT: correlação positiva, estatisticamente significativa e igual a 0,568.
- log₁₀(NCC) vs. CL: correlação positiva, estatisticamente significativa e igual a 0,662.
- log₁₀(NCC) vs. T: correlação estatisticamente igual a zero (0,225).
- FT vs. CL: correlação positiva, estatisticamente significativa e igual a 0,799.



Conclusões - Regressões Simples:

- A hipótese de normalidade para os resíduos, verificada através do teste de Shapiro-Wilk, foi satisfeita para todos os modelos ($\alpha = 0,01$).
- O coeficiente do parâmetro T foi considerado não significativo na respectiva regressão simples.
- Dentre os modelos de regressão simples, o que apresentou o melhor desempenho foi aquele que considera CL como variável independente.
- Para todos os modelos ajustados, a estatística R^2 foi sempre menor do que 0,5 indicando que menos do que 50 % da variabilidade do parâmetro NCC pode ser explicada pelos demais parâmetros.

Conclusões - Regressão Múltipla:

- O modelo de regressão múltipla resultante apresenta como variáveis independentes os parâmetros CL sendo que foi considerado estatisticamente significativo.

Esta metodologia foi utilizada para análise de correlações entre pares de parâmetros de interesse nos demais estudos e podem ser consultados no relatório "**Desenvolvimento de Análise Estatística sobre dados de Pesquisas da Qualidade das Águas de São Paulo Estudo sobre a Qualidade das Águas Superficiais**". (Arquivo em CD anexo)

3.2.2 ESTUDO CORRELAÇÃO – REDE AUTOMÁTICA

O objetivo deste estudo foi evidenciar estatisticamente a correlação de variáveis de qualidade da água referente a resultados obtidos nas Estações Laranjal Paulista, Guarapiranga e Taquacetuba da Rede Automática da CETESB.

Para cada estação de monitoramento, foi avaliada a correlação entre os seguintes parâmetros:

- oxigênio dissolvido x pH;
- oxigênio dissolvido x condutividade;
- oxigênio dissolvido x turbidez;
- oxigênio dissolvido x temperatura da água;
- oxigênio dissolvido x vazão ou volume;
- condutividade x turbidez;
- condutividade x vazão ou volume;
- turbidez x vazão ou volume.

A título de exemplo segue a análise de correlação entre parâmetros em ponto na Represa Billings.

Estação Taquacetuba (Represa Billings)

A Tabela 13 apresenta o período com o maior número de observações, para cada um dos pares de parâmetros para os quais deseja-se analisar a correlação. Observa-se que, de acordo com as regras definidas na metodologia, não é possível calcular a correlação em três casos: Oxigênio Dissolvido x Turbidez, Condutividade Elétrica x Turbidez e Turbidez x Volume.

Tabela 13: Período com o maior número de observações

Correlações entre Parâmetros	Período	Total de Observações
Oxigênio Dissolvido x pH	12/07/2011 a 30/12/2013	903
Oxigênio Dissolvido x Condutividade Elétrica	12/07/2011 a 20/12/2012	528
Oxigênio Dissolvido x Turbidez	21/02/2013 a 30/12/2013	313
Oxigênio Dissolvido x Temperatura d'Água	12/07/2011 a 30/12/2013	903
Oxigênio Dissolvido x Volume	12/07/2011 a 30/12/2013	903
Condutividade Elétrica x Turbidez	21/02/2013 a 30/11/2013	283
Condutividade Elétrica x Volume	05/05/2011 a 20/12/2012	596
Turbidez x Volume	21/02/2013 a 30/12/2013	313

As Figuras 3 a 7 apresentam as séries temporais das médias diárias e gráficos de dispersão para cada um dos pares de parâmetros para os quais deseja-se obter as correlações. Os períodos correspondentes a cada gráfico são dados na Tabela 13.

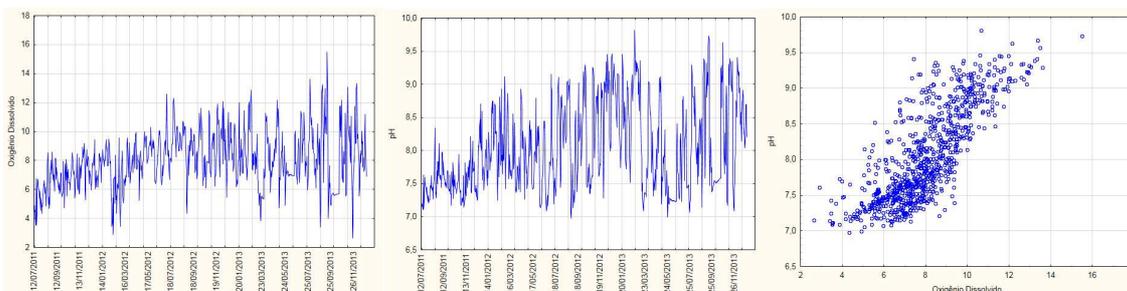


Figura 3. Série temporal das médias diárias para os Oxigênio Dissolvido (esquerda) e pH (centro), no período de 12/07/2011 a 30/12/2013, e o gráfico de dispersão para este par de parâmetros (direita).



Figura 4. Série temporal das médias diárias para os Oxigênio Dissolvido (esquerda) e Condutividade Elétrica (centro), no período de 12/07/2011 a 20/12/2012, e o gráfico de dispersão para este par de parâmetros (direita)



Figura 5. Série temporal das médias diárias para os Oxigênio Dissolvido (esquerda) e Temperatura d'Água (centro), no período de 21/02/2011 a 30/12/2013, e o gráfico de dispersão para este par de parâmetros (direita)



Figura 6. Série temporal das médias diárias para os Oxigênio Dissolvido (esquerda) e Volume (centro), no período de 12/07/2011 a 30/12/2013, e o gráfico de dispersão para este par de parâmetros (direita).



Figura 7. Série temporal das médias diárias para os Condutividade Elétrica (esquerda) e Volume (centro), no período de 05/05/2011 a 20/12/2012, e o gráfico de dispersão para este par de parâmetros (direita).

A Figura 8 apresenta os gráficos do coeficiente ρ_{DCCA} para diferentes valores de escala $\nu \in \{10, 12, 14, \dots, 208\}$, para cada um dos pares de interesse. A linha vermelha pontilhada representa o valor médio de ρ_{DCCA} calculado sob todas as escalas. Tal valor é reportado na Tabela 14.

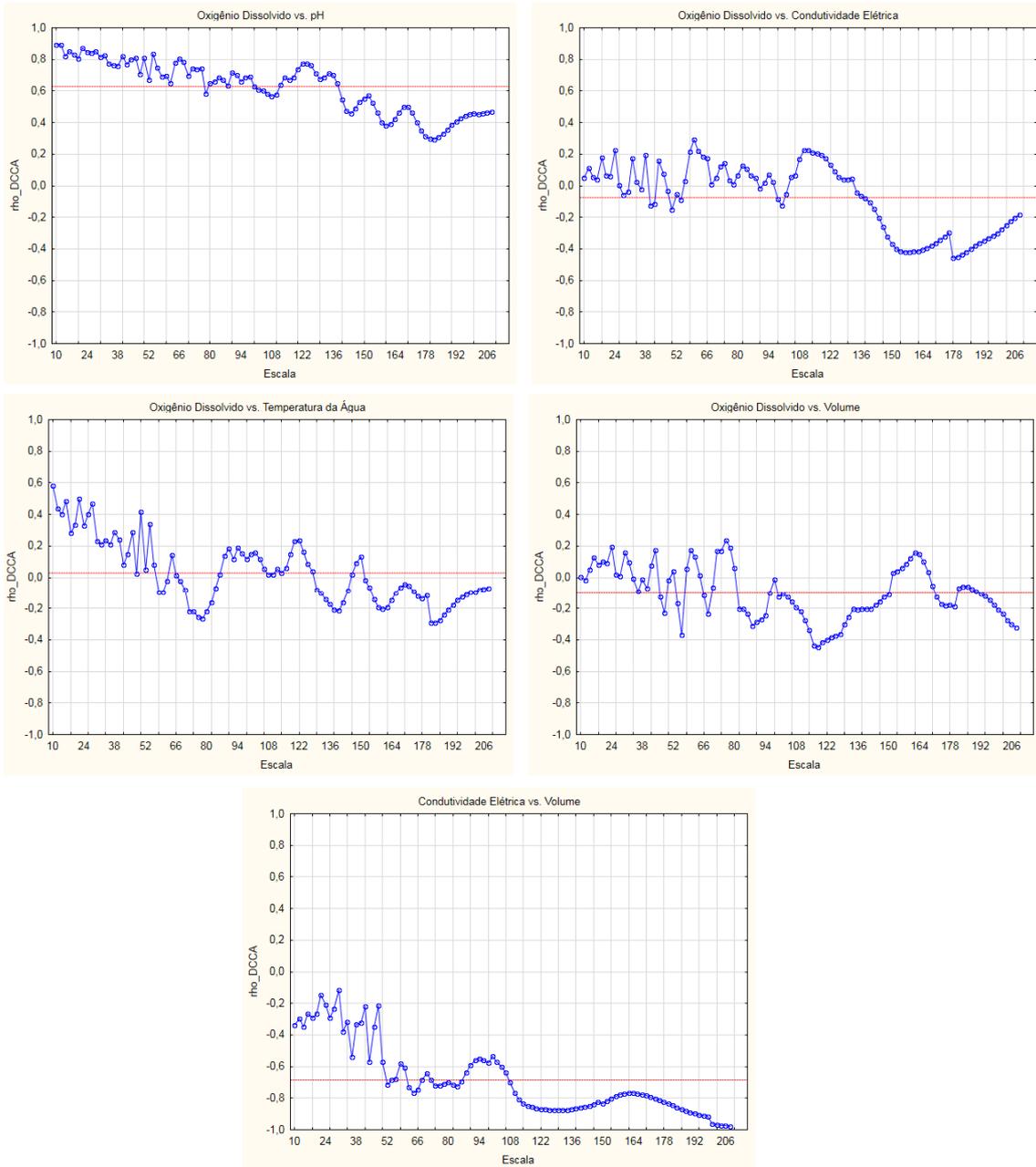


Figura 8. Valores de ρ_{DCCA} para diferentes escalas ν , para os pares: Oxigênio Dissolvido e pH, Oxigênio Dissolvido e Condutividade Elétrica, Oxigênio Dissolvido e Temperatura d'Água, Oxigênio Dissolvido e Volume, Condutividade Elétrica e Volume. A linha pontilhada representa a média tomada sob todas as escalas.

Tabela 14. Valor médio do coeficiente de correlação ρ_{DCCA} , calculado para diferentes valores de escala $v \in \{10, 12, 14, \dots, 208\}$, para os pares de parâmetros de interesse.

Parâmetro	Parâmetro	ρ_{DCCA} (Média)
Oxigênio dissolvido	pH	0,6264
Oxigênio dissolvido	Condutividade elétrica	-0,0761
Oxigênio dissolvido	Temperatura da água	0,0307
Oxigênio dissolvido	Volume	-0,0990
Condutividade elétrica z	Volume	-0,6833

Observa-se que para todos os pares de parâmetros, o valor da correlação ρ_{DCCA} é influenciado pela escala escolhida. Esse resultado é esperado devido a natureza não estacionária dos dados envolvidos no problema. Mais especificamente, à medida que a escala aumenta, a remoção da tendência fica comprometida. Dessa forma, aconselha-se a olhar os resultados para escalas pequenas (≤ 60).

Em termos de força e sinal (positiva ou negativa) da correlação, podemos resumir os resultados de acordo com a Tabela 15. Lembramos que as escalas grandes são utilizadas para ilustrar a influência da não estacionariedade no cálculo da correlação, enquanto as escalas pequenas são aquelas nas quais foi focada a atenção.

Tabela 15. Conclusões sobre a correlação entre os pares de parâmetros para diferentes escalas, seguindo a convenção: F1 = Fraca; F2 = Fraca/Intermediária; F3 = Intermediária; F4 = Intermediária/Forte; F5 = Forte; N = Negativa; P = Positiva.

Par de Parâmetros	$v \leq 60$		$60 < v < 120$		$v \geq 120$	
	Força	Sinal	Força	Sinal	Força	Sinal
Oxigênio Dissolvido vs. pH	F4	P	F4	P	F3	P
Oxigênio Dissolvido vs. Condutividade Elétrica	F2 e F3 ⁽¹⁾	P	F1 e F2 ⁽¹⁾	P	F2	N
Oxigênio Dissolvido vs. Temperatura da Água	F2	P	F2	P	F2	P
Oxigênio Dissolvido vs. Volume	F1	P	F2	P	F2	N
Condutividade Elétrica vs. Volume	F2	N	F4 ¹⁾	N	F4	N

Notas: (1) Os dois casos ocorrem em proporções muito próximas.

Conclusão

Devido ao grande número de dados faltantes não foi possível estabelecer uma base amostral representativa para cálculo da correlação para todos os pares que possuem Turbidez como um dos parâmetros.

O par Oxigênio Dissolvido e pH é o que apresenta maior correlação, classificada como Intermediária/Forte e positiva, para escalas pequenas (≤ 60). Nos demais casos, obteve-se sempre correlação Fraca ou Intermediária/Fraca, não podendo descartar a hipótese de que essas correlações sejam iguais a zero.

O par Condutividade Elétrica e Volume foi o que apresentou maior discrepância nos valores da correlação, para escalas pequenas e grandes. Tal fato era esperado pois este é o caso em que a não estacionariedade das séries temporais é mais pronunciada (para referência, veja as Figuras 3 a 7).

Esta metodologia foi utilizada para análise de correlação nas demais estações automáticas objeto do estudo e podem ser consultadas no relatório “**Desenvolvimento de Análise Estatística sobre dados de Pesquisas da Qualidade das Águas de São Paulo Estudo sobre a Qualidade das Águas de Rede Automática**”. (Arquivos em CD anexo)

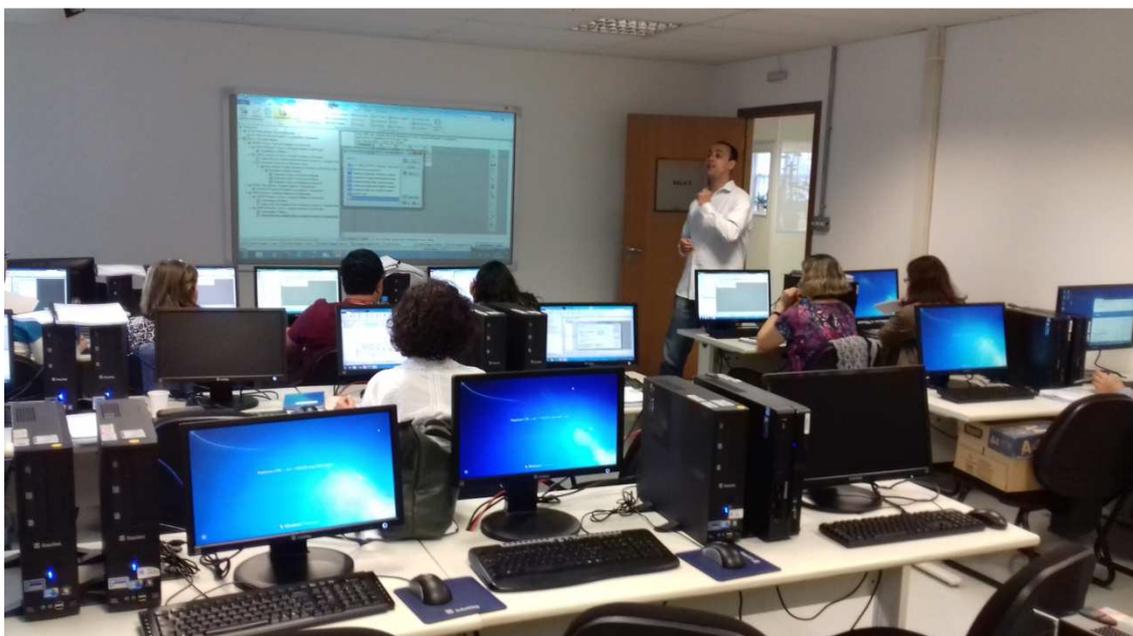
4. CAPACITAÇÃO EM ESTATÍSTICA

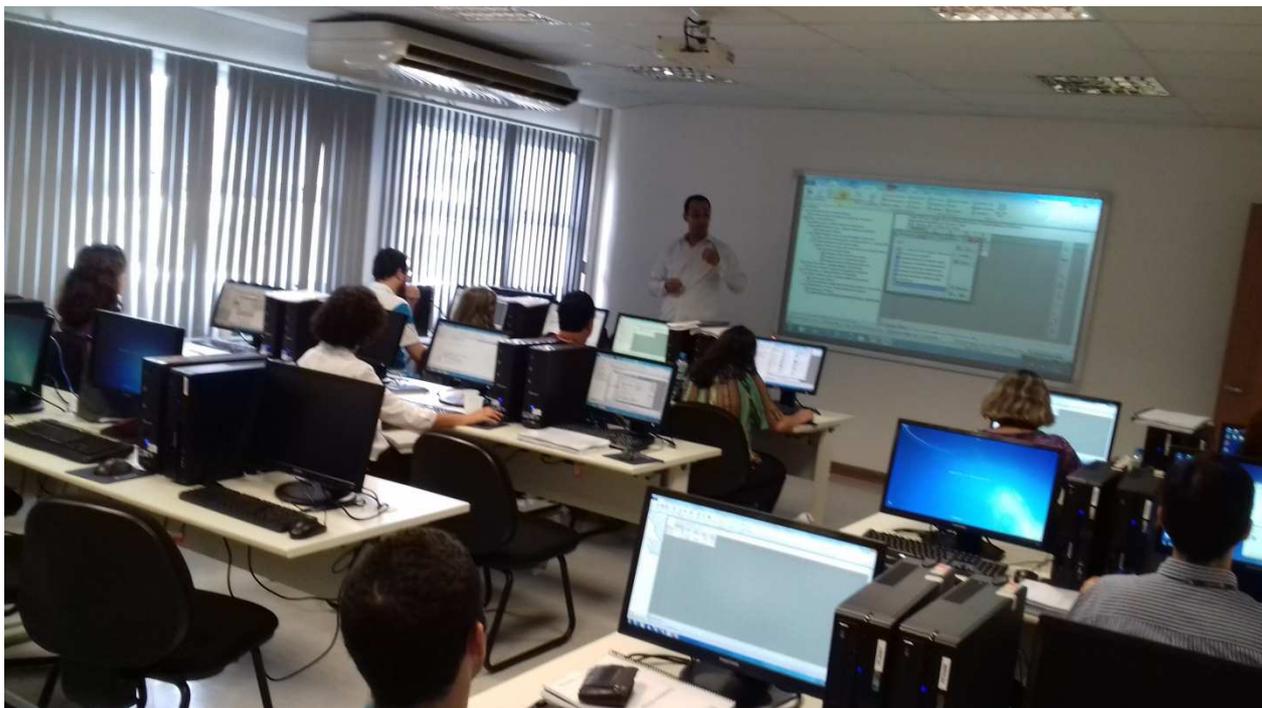
Em atendimento ao contrato, a empresa Statsoft realizou treinamento on-site, nos dias 26 a 29 de outubro de 2015, das 8 às 17 horas (duração de 32 horas), para um grupo de 15 profissionais da CETESB atuantes nas áreas objeto dos estudos.

O treinamento teve o objetivo de capacitar os técnicos para avaliar:

- O comportamento das variáveis monitoradas, identificando situações de anormalidade, por meio do levantamento de tendência e sazonalidade, através do ajuste de modelos de Séries Temporais; e
- A estrutura de dependência entre as variáveis de qualidade ambiental, visando respostas para fenômenos como eutrofização em reservatórios, identificação de pontos com comportamento similar ao longo do tempo, visando a otimização da rede de monitoramento e outras situações que sejam relevantes.

O programa, as ementas e a lista de presença encontram-se arquivados no CD anexo e a seguir são apresentadas fotos do treinamento.





6. CONCLUSÃO

Diante do exposto e tendo em vista os produtos obtidos com a contratação da STATSOFT SOUTH AMERICA COMÉRCIO DE SOFTWARE LTDA para realizar serviços especializados em estatística, considera-se atendidos os objetivos do empreendimento Fehidro Contrato 124/2012, no que se refere à criação de subsídios técnicos para tomada de decisão, por meio do conhecimento do comportamento das variáveis de qualidade ambiental em águas superficiais e subterrâneas, a partir dos dados da rede de monitoramento da CETESB, bem como capacitar suas equipes e implantar tecnologia aplicada à estatística, com vistas à elaboração de informações para a gestão da qualidade ambiental e de recursos hídricos.