



**UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

TESE DE DOUTORADO

Avaliação do funcionamento de uma estação de tratamento de esgoto doméstico e desenvolvimento de um novo método para determinação de DQO usando espectrometria NIR e quimiometria

Antônio Cícero de Sousa

João Pessoa-PB-Brasil

Junho/2007



**UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

TESE DE DOUTORADO

Avaliação do funcionamento de uma estação de tratamento de esgoto doméstico e desenvolvimento de um novo método para determinação de DQO usando espectrometria NIR e quimiometria

Antônio Cícero de Sousa

Tese apresentada como requisito para obtenção do título de Doutor em Química Analítica pela Universidade Federal da Paraíba

Orientador: Prof. Dr. Mário César Ugulino de Araújo

2º Orientador: Dr. Wallace Duarte Fragoso

João Pessoa-PB-Brasil

Junho/2007

S725a Sousa, Antônio Cícero de.

Avaliação do funcionamento de uma estação de tratamento de esgoto doméstico e desenvolvimento de um novo método para determinação de DQO usando espectrometria NIR e quimiometria/Antônio Cícero de Sousa.-João Pessoa, 2007.

87p.:il.

Orientadores:Mário César Ugulino de Araújo e Wallace Duarte Fragoso.

Tese (Doutorado)-UFPB/CCEN

1.Esgoto doméstico-tratamento. 2.DQO.
3.Séston. 4.NIR. 5.Quimiometria.

UFPB/BC

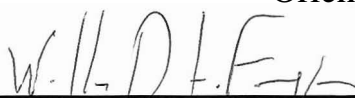
CDU: 628.3.033(043)

Tese de Doutorado submetida ao Corpo Docente do Programa de Pós-Graduação em Química do Departamento de Química do Centro de Ciências Exatas e da Natureza da Universidade Federal da Paraíba como parte dos requisitos para a obtenção do grau de Doutor em Ciências.

Aprovada pela banca examinadora:



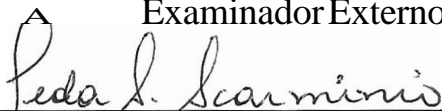
Prof. Dr. Mário César Ugulino de Araújo
Departamento de Química do CCEN/UFPB
Orientador



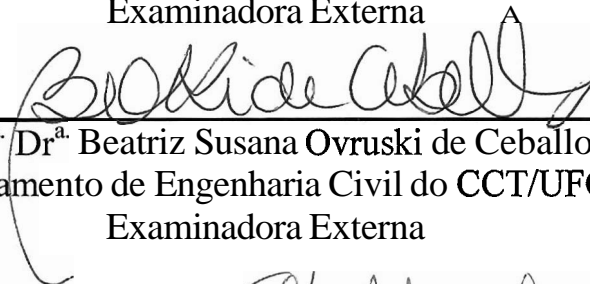
Prof. Dr. Wallace Duarte Fragoso
Departamento de Química do CCEN/UFPB
2º. Orientador



Prof. Dr. Wilson de Figueiredo Jardim
Instituto de Química da UNICAMP
Examinador Externo



Prof.ª Dr.ª Iêda Spacino Scarmínio
Departamento de Química da UEL/PR
Examinadora Externa



Prof.ª Dr.ª Beatriz Susana Ovruski de Ceballos
Departamento de Engenharia Civil do CCT/UFCG
Examinadora Externa



Prof.ª Dr.ª Teresa Cristina Bezerra Saldanha
Departamento de Química do CCEN/UFPB
Examinadora Interna

Dedicatória

Dedico esta tese de doutorado a meus pais em especial a minha mãe Joana Bastos de Sousa (*in memoriam*), pela minha criação e formação profissional.

Ao meu amor Verinha pela sua companhia e presença permanente nos momentos de alegria e também de dificuldades da minha vida.

Ao meu querido filho Arthur, uma obra Divina presente em minha vida.

Agradecimentos

- A Deus pela minha existência e graças concedidas.
- Ao prof. Dr. Mário César Ugulino pela honrosa dedicação na orientação deste trabalho.
- Ao Dr. Wallace Duarte Fragoso pela valiosa contribuição e dedicação na co-orientação deste trabalho.
- Ao professor Dr. Roberto Kawakami do ITA pelas discussões e contribuição no andamento desse trabalho.
- À professora Tereza pelas sugestões ao longo da realização desse trabalho.
- Ao prof. Dr. Célio Pasquini da UNICAMP por conceder acesso ao laboratório para execução das medidas NIR.
- Ao Centro Federal de Educação Tecnológica da Paraíba (CEFET) pelo apoio da coordenação de recursos naturais e colaboração de todos os alunos bolsistas que contribuíram direta ou indiretamente para realização deste trabalho. Além disso, por conceder acesso as instalações físicas dos laboratórios e transportes para as coletas.
- À CAGEPA pela disponibilidade do sistema estudado (ETE-Mangabeira) e do laboratório de microbiologia.
- À SUDEMA pela concessão de uma bolsa para estagiário do CEFET.
- A todos os professores da Pós-graduação em Química.
- À Glauciene, Alessandra, Edilene, Mônica e Ovídio, colaboradores deste trabalho.
- A todos os alunos de IC e de mestrado do LAQA/Química Ambiental que colaboraram na parte experimental deste trabalho.
- A todos que fazem o LAQA pela convivência e amizade cultivada.

- Ao colega Heronides pela ajuda nas medidas NIR na UNICAMP.
- Ao colega Carlos pela colaboração na correção dos textos em inglês.
- A todos os meus amigos pela amizade e confiança.
- Aos meus pais, pelo ensino dos valores fundamentais que carregarei por toda minha vida e a toda minha família pelo apoio e incentivo.
- Ao CNPq pelo financiamento do projeto referente ao processo (305729/2003-9) intitulado “Uso da quimiometria e espectrometria UV-VIS e NIR para o desenvolvimento de um método de avaliação de qualidade de águas do rio Mumbaba e da lagoa de estabilização de Mangabeira da cidade de João Pessoa-PB” que culminou com a realização deste trabalho.

Este trabalho relata o uso de métodos quimiométricos em dois estudos sobre química aquática. O primeiro estudo tem como foco principal o funcionamento de uma Estação de Tratamento de Esgoto Doméstico (ETE), situada no bairro de Mangabeira da cidade de João Pessoa - Paraíba - Brasil. Esta ETE, denominada ETE-Mangabeira, foi inicialmente projetada para funcionar como uma ETE composta de uma lagoa aerada seguida de lagoa de maturação. Porém, por um problema mecânico nos aeradores, foi adaptada para operar como uma ETE do tipo sistema australiano, ou seja, lagoa anaeróbia seguida de lagoa facultativa. Neste estudo, de fevereiro/2005 a fevereiro/2006, foram coletadas amostras em três diferentes pontos da ETE-Mangabeira, das quais foram determinados parâmetros físicos, químicos e bacteriológicos. Para compreender a dinâmica aquática, a sazonalidade, os processos de decomposição anaeróbia de matéria orgânica e o processo fotossintético que ocorrem neste sistema, foram usados os modelos de equilíbrio químico e as seguintes ferramentas quimiométricas: gráficos de *Box-Whisker*, análise de componentes principais (PCA), rotação varimax e análise discriminante (DA). Através da PCA e DA foi possível distinguir as etapas do tratamento responsáveis pela melhoria na qualidade do esgoto e pela influência da sazonalidade. Através da rotação varimax foi possível identificar as variáveis mais importantes nos processos de fotossíntese e de decomposição anaeróbia de matéria orgânica. O segundo estudo teve como finalidade o desenvolvimento de um novo método para determinação de DQO na ETE-Mangabeira, usando espectros de reflectância difusa no infravermelho próximo do séston e a técnica quimiométrica mínimos quadrados parciais (PLS). O erro estimado para o método proposto foi igual a $19 \text{ mg O}_2 \text{ L}^{-1}$ e é comparável ao erro encontrado com o método de referência que é de $21 \text{ mg O}_2 \text{ L}^{-1}$. Os resultados obtidos para ambos os métodos apresentaram uma correlação de 0,97, que pode ser considerada muito boa.

The use of chemometric methods in two studies involving aquatic chemistry is described in this work. The aim of the first study was to investigate the operation of a domestic wastewater treatment plant (DWTP), located in Mangabeira, district of João Pessoa city (Paraíba, Brazil). Named DWTP-Mangabeira, it was initially projected in order to operate as a DWTP that was composed of an aerobic pond followed by maturation pond. However, due to mechanic problem in aerators, it was adapted to work as an Australian system (an anaerobic pond followed by facultative pond). In this study, samples were collected at three points of DWTP-Mangabeira within the February/2005 - February/2006 period, and chemical, physical and bacteriological parameters were determined. In order to understand the processes of aquatic chemical dynamics, seasonality, anaerobic degradation of organic matter and photosynthesis that occur in this system, models of chemical equilibrium and the following chemometrics tools: Box and Whisker plot, Principal Component Analysis (PCA), Varimax Rotation and Discriminant Analysis (DA) were used. With PCA and DA, it was possible to distinguish the stages of the treatment responsible for the improvement in the wastewater quality and the influence of the seasonality. By using the varimax rotation, it was possible to identify the most important variables for the two main processes that occur in ponds (organic matter degradation and photosynthesis). The purpose of the second study was to develop a novel method to determine COD in DWTP-Mangabeira by using near-infrared reflectance spectra of seston and PLS (Partial Least Square) chemometric method. The estimated error of the proposed method was $19 \text{ mg O}_2 \text{ L}^{-1}$ and it is comparable to the error of the reference method, which was $21 \text{ mg O}_2 \text{ L}^{-1}$. These results present a good agreement with a correlation estimated in 0.97.

LISTA DE FIGURAS.....	v
LISTA DE TABELAS.....	viii
LISTAS DE ABREVIATURAS E SIGLAS.....	ix
CAPÍTULO 1 - Introdução.....	1
1.1 Química aquática.....	2
1.2 A problemática das águas.....	3
1.3 Parâmetros de qualidade de Águas.....	4
1.4 Revisão bibliográfica do uso da quimiometria em águas.....	9
1.5 Objetivos do trabalho.....	14
CAPÍTULO 2 - Investigação da eficiência do funcionamento da ETE- Mangabeira.....	15
2.1 Tratamento de esgoto doméstico.....	16
2.2 Tratamento de esgoto doméstico baseado em Lagoa de estabilização.....	17
2.3 Técnicas quimiométricas.....	20
2.3.1 Gráficos de Box-Whisker.....	20
2.3.2 Análise fatorial.....	21
2.3.3 Análise discriminante.....	25
2.4 OBJETIVOS.....	26
2.5 EXPERIMENTAL.....	26
2.5.1. Sistema hídrico investigado.....	26
2.5.2. Dados de precipitação pluviométrica.....	30
2.5.3 Coleta das amostras e períodos de amostragem.....	31
2.5.4 Parâmetros monitorados.....	31
2.5.5 Tratamento dos dados.....	35
2.6 RESULTADOS E DISCUSSÃO.....	35
2.6.1 – Análise univariada dos dados da ETE-Mangabeira.....	36

	Páginas
2.6.2 – Análise multivariada dos dados da ETE-Mangabeira.....	41
CAPÍTULO 3 - Desenvolvimento de um novo método para determinação de DQO em esgoto doméstico.....	47
3.1 Determinação de DQO.....	48
3.2 Espectrometria NIR.....	49
3.3 Espectrometria NIR de reflectância difusa (NIRR).....	49
3.4 A transformada wavelet para minimização de ruído.....	51
3.5 Calibração multivariada.....	53
3.5.1 Calibração PLS.....	54
3.6 Revisão da literatura.....	55
3.7 OBJETIVO.....	56
3.8 EXPERIMENTAL.....	56
3.8.1 Coleta e períodos de amostragem.....	56
3.8.2 Processamento das amostras.....	57
3.8.3 Método de referência para determinação de DQO.....	57
3.8.4 Medidas dos espectros NIRR.....	57
3.8.5 Pré-processamentos e programas computacionais.....	58
3.9. RESULTADOS E DISCUSSÃO.....	59
3.9.1 Pré-tratamento e pré-processamento dos espectros NIRR.....	59
3.9.2 Escolha do volume de filtração das amostras.....	60
3.9.3 Seleção das amostras do conjunto de calibração e de previsão.....	63
3.9.4 Construção dos modelos <i>PLS</i>	64
CAPÍTULO 4 - Conclusões e propostas	70
4.1 CONCLUSÕES.....	71
4.2 PROPOSTAS	73
Referências bibliográficas.....	74
Anexos.....	81

LISTA DE FIGURAS

Figura 2.1 Ilustração gráfica de Box-Whisker.....	21
Figura 2.2 Ilustração de uma análise de rotação fatorial rtogonal, daptado de Hair...24	24
Figura 2.3 Mapa de localização da ETE-Mangabeira.....	27
Figura 2.4 Diagrama esquemático da ETE-Mangabeira. Região de tratamento primário (a) e de tratamento secundário (b) e (c).....	28
Figura 2.5 Fotografia das duas lagoas anaeróbias da ETE-Mangabeira.....	29
Figura 2.6 Precipitação pluviométrica mensal de fevereiro de 2005 a fevereiro de 2006.....	30
Figura 2.7 Esquema da seqüência de processos na digestão anaeróbia que ocorre em sistemas de tratamento de esgoto. ^[42]	36
Figura 2.8 Gráficos de <i>Box-Whisker</i> para a DQO do afluente anaeróbio (AA), do efluente anaeróbio (EA) e do efluente facultativo (EF) da ETE-Mangabeira.....	37
Figura 2.9 Gráficos de <i>Box-Whisker</i> para pH (a) e sulfato (b) do afluente anaeróbio (AA), do efluente anaeróbio (EA) e do efluente facultativo (EF) da ETE-Mangabeira.....	37
Figura 2.10 Diagrama de distribuição da fração (α) em função do pH ^[1] de algumas espécies em equilíbrio responsáveis pela alcalinidade em sistemas aquáticos.....	38
Figura 2.11 Gráficos de <i>Box-Whisker</i> para a alcalinidade total do afluente anaeróbio (AA), do efluente anaeróbio (EA) e do efluente facultativo (EF) da ETE-Mangabeira.....	38
Figura 2.12 Gráficos de <i>Box-Whisker</i> para os sólidos totais (a) e turbidez (b) do afluente anaeróbio (AA), do efluente anaeróbio (EA) e do efluente facultativo (EF) da ETE-Mangabeira.....	39
Figura 2.13 Gráfico de <i>Box-Whisker</i> para OD do afluente anaeróbio (AA), do efluente anaeróbio (EA) e do efluente facultativo (EF) da ETE-Mangabeira.....	40
Figura 2.14 Gráfico de <i>Box-Whisker</i> para amônia e fosfato do afluente anaeróbio (AA), do efluente anaeróbio (EA) e do efluente facultativo (EF) da ETE-Mangabeira.....	40
Figura 2.15 Gráfico de <i>Box-Whisker</i> para C. Term do afluente anaeróbio (AA), do efluente anaeróbio (EA) e do efluente facultativo (EF) da ETE-Mangabeira.....	41

Figura 2.16 Gráfico dos escores e dos pesos de PCA da ETE-Mangabeira, em três níveis de índices pluviométricos: alto (■), médio (▣) e baixo (□), onde, AA–afluente anaeróbico (■), EA-efluente anaeróbico(■) e EF- efluente facultativo (■)...41	41
Figura 2.17 Gráfico dos escores da rotação varimax fator1 (F1) versus fator3 (F3) para as amostras da lagoa de estabilização, AA (■) , EA (■) e EF (■)44	44
Figura 2.18 Escores Z das amostras da lagoa de estabilização, AA (■) , EA (■) e EF (■)45	45
Figura 3.1 Refletância difusa resultante da incidência de um feixe de radiação eletromagnética em uma amostra constituída de partículas de forma e tamanho irregulares. Luz incidente = a ; radiação difusa refletida = b ; suporte da cela = c ; janela da cela = d ; cela da amostra = e ^[52]50	50
Figura 3.2 Implementação de banco de filtro de transformada wavelet com N níveis de decomposição. As seqüências c^k e d^k são denominadas, respectivamente, coeficientes de aproximação e de detalhe do k-ésimo nível de decomposição.....52	52
Figura 3.3 Representação ilustrativa das etapas de calibração multivariada, onde, as linhas horizontais de 1 a n e de 1 a m e as colunas de 1 a p e de 1 a q representam, respectivamente, as amostras, as variáveis espectrais e os parâmetros das matrizes X e Y das etapas de calibração e de previsão.....54	54
Figura 3.4 Espectrômetro NIR. A é a fonte, B é o detector, C acessório de refletância difusa e (D) recipiente de material refratário onde era colocada a membrana contendo o séston das amostras de esgoto analisadas.....58	58
Figura 3.5 Espectro NIRRR na faixa espectral de 10.000 a 4.000 cm^{-1} do séston das amostras coletadas nos pontos AA (—), EA (—) e EF (—).....59	59
Figura 3.6 Espectros registrados para as amostras do AA (a), EA (b) e EF (c) para volumes de (—10), (—25), (—50), (—100), (—150) e (—200) ml.....60	60
Figura 3.7 o espectro NIRRR do material sestônico de uma amostra coletada no ponto EA (a); o espectro derivativo resultante da aplicação da derivação Savitzky-Golay (b); e os espectros derivativos resultantes da minimização de ruídos com a transformada de <i>wavelet</i> (c) com 4 e 5 níveis de decomposição.....62	62
Figura 3.8. O gráfico <i>scree</i> dos modelos <i>PLS</i> gerados usando os coeficientes da transformada wavelet (modelos PLS_{TW}). A seta indica o ponto associado ao número de variáveis latentes adotado no modelo.....64	64

Figura 3.9. O gráfico *scree* dos modelos *PLS* gerados usando os espectros brutos (modelos PLS_B). A seta indica o ponto associado ao número de variáveis latentes adotado no modelo.....65

Figura 3.10. O gráfico *scree* dos modelos *PLS* gerados usando os os espectros derivados com Savitzky-Golay (modelos PLS_{SG}). A seta indica o ponto associado ao número de variáveis latentes adotado no modelo.....65

Figura 3.11 Valores de DQO previstos pelo modelo PLS_{tw} *versus* valores referência. A barra de erro indica o desvio padrão obtido na previsão de cada uma das 10 amostras do conjunto de previsão.....67

Figura 3.12 Valores previstos de DQO *versus* referência para as 31 amostras da validação cruzada para cada amostra. A barra de erro indica o desvio padrão para cada amostra de previsão.....68

LISTA DE TABELAS

Tabela 2.1 Condições técnicas de operação das lagoas anaeróbia e facultativa, fornecidos pela CAGEPA.....	29
Tabela 2.2 Eficiência de DBO ₅ , DQO e C. Term para ETE-Mangabeira, fornecidos pela CAGEPA.....	30
Tabela 2.3 Pesos das variáveis para os três primeiros fatores rotacionados	43
Tabela 2.4 Parâmetros de avaliação da significância estatística da função discriminante.....	46
Tabela 3.1 Valores de correlação e RMSEP dos PLS_{TW} , PLS_B e PLS_{SG} escolhidos, com, respectivamente, seis, sete e seis variáveis latentes.....	66

LISTA DE ABREVIATURAS E SIGLAS

PNRH:	Programa Nacional de Recursos Hídricos
ETE:	Estação de Tratamento de Esgoto
CONAMA:	Conselho Nacional do Meio Ambiente
IBGE:	Instituto Brasileiro de Geografia e Estatística
DBO:	Demanda Bioquímica de Oxigênio
DQO:	Demanda Química de Oxigênio
CF:	Coliformes Fecais
AA:	Afluente Anaeróbio
EA:	Efluente Anaeróbio
EF:	Efluente Facultativo
HCA:	Análise de Agrupamento Hierárquico
PCA:	Análise de Componentes Principais
PC:	Componente Principal
AF :	Análise Fatorial
DA:	Análise Discriminante
PARAFAC:	Análise de Fatores Paralelos
OD:	Oxigênio Dissolvido
C.Term:	Coliformes Termotolerantes
TOC:	Carbono Orgânico Total
ST:	Sólidos Totais
CETESB:	Companhia de Tecnologia de Saneamento Ambiental
CAGEPA:	Companhia de Água e Esgoto da Paraíba
SUDEMA:	Superintendência de Desenvolvimento do Meio Ambiente
NIRR:	Espectrometria de Refletância Difusa no Infravermelho Próximo
TW:	Transformada Wavelet
H:	Filtração Passa Baixa
G:	Filtração Passa Alta
SIMCA:	<i>Soft Independent Modelling of Class Analogies</i>
PLS:	Regressão por Mínimos Quadrados Parciais
RMSECV:	Raiz quadrada do Erro Médio Quadrático da Validação Cruzada
RMSEP:	Raiz Quadrada do Erro Médio Quadrático de Previsão
KS:	Kennard-Stone
S/N:	Relação Sinal Ruído
SÉSTON:	Conjunto das partículas, orgânicas ou não, que se encontram dispersas na coluna de água.

CAPÍTULO 1

Introdução

1.1 Química Aquática

A Química aquática é uma subárea da química ambiental que estuda os fenômenos associados à distribuição e à circulação de espécies químicas em águas naturais, águas subterrâneas, rios, lagos, mares e estuários. Estes estudos buscam a compreensão da composição química de sistemas aquáticos através da investigação de fontes e transportes de substâncias bem como da caracterização de reações químicas considerando toda espécie de interferência do ambiente geofísico e social.^[1]

A interpretação de fenômenos químicos em sistemas aquáticos é muito importante no estudo de qualidade de água. Esta deve ter bases técnico-científicas que associem as propriedades específicas da água a processos simultâneos que são descritos por diferentes modelos de reações como: ácido-base, solubilidade, oxirredução e complexação, sob os pontos de vista termodinâmico e cinético. Porém a complexidade do tratamento simultâneo dos processos termodinâmicos e cinéticos tem levado à descrição da maioria dos fenômenos químicos em corpos aquáticos através de modelos simples freqüentemente baseados no conceito de equilíbrio químico.^[1] Embora estes modelos sejam limitados com relação à exatidão e não sejam absolutos, podem resultar em generalizações úteis para visualização das condições que determinam os processos químicos em sistemas aquáticos naturais.

A dinâmica de sistemas aquáticos é multifatorial e de complexa compreensão, devido principalmente à interação entre os constituintes do meio aquático e entre os compartimentos ambientais: biosfera, solo, ar e água, dificultando o estudo e a interpretação de sistemas aquáticos.^[2] Assim, os estudos ambientais se caracterizam pela sua natureza multivariada, em decorrência da geração de bancos de dados multidimensionais com visualização e interpretação complexas. Estes estudos envolvem sempre interação e efeitos antagônicos ou sinérgicos sobre estes tipos de conjunto de dados que não são facilmente tratados pelas ferramentas estatísticas convencionais, geralmente aplicadas a dados normais e homoscedásticos.^[2] No contexto da química ambiental, a quimiometria tem se mostrado uma ferramenta

poderosa para a compreensão das questões associadas à variabilidade inerente aos dados ambientais devido à influência de fatores naturais e antrópicos.^[2]

1.2 A problemática das Águas

A preocupação com problemas ambientais que levam à escassez de água e afetam a qualidade da mesma, torna a implementação de programas de monitoramento de recursos hídricos uma temática importante.^[2] Assim, são definidas políticas de gestão e administrativas, fundamentadas em estudos sobre riscos químicos à saúde pública e ambiental, buscando o aperfeiçoamento da legislação e a prática do monitoramento de qualidade de águas. No Brasil, destaca-se neste contexto a Política Nacional de Recursos Hídricos (PNRH), instituída pela Lei 9.433 de 8 de janeiro de 1997 e a Resolução CONAMA número 357 de março de 2005.^[3]

O impacto do lançamento de efluentes originados de estações de tratamento de esgotos em corpos d'água é motivo de grande preocupação para a maioria dos países. Uma série de leis ambientais, critérios, políticas e revisões influem tanto na seleção dos locais de descarga quanto no nível de tratamento exigido para garantir que os impactos ambientais provocados pela disposição destes efluentes sejam aceitáveis.^[4] Vários autores^[5-8] publicaram estudos sobre a verificação do atendimento aos padrões de lançamento e sobre o desempenho de processos de tratamento, considerando a qualidade dos efluentes. Todos esses trabalhos dão subsídios às agências ambientais para uma definição adequada dos padrões de lançamento nos corpos d'água. No Brasil, apesar de estudos e avaliações em escala piloto ou em Estações de Tratamentos de Esgotos (ETEs) individuais, o conhecimento sobre o desempenho das tecnologias de tratamento de esgotos em operação no país é relativamente esparso, havendo poucas consolidações estruturadas em termos de uma avaliação global.^[4,9,10] Segundo o Instituto Brasileiro de Geografia e Estatística^[11], no Brasil, 47,8% dos municípios não possuem coleta de esgoto sanitário, 32% só coletam e 20,2% coletam e tratam. Nos domicílios apenas 33,5% tem serviços de coleta de

esgoto. O resultado é que as regiões metropolitanas e as grandes cidades concentram grandes volumes de esgoto coletado que são despejados nos corpos receptores sem nenhum tipo de tratamento.

De maneira geral os esgotos sanitários possuem mais de 98% de sua composição constituída por água, porem, há contaminantes entre os quais destacam-se: sólidos suspensos, compostos orgânicos, nutrientes, metais , sólidos dissolvidos inorgânicos, sólidos inertes, sólidos grosseiros, compostos não biodegradáveis, organismos patogênicos, e ocasionalmente, contaminantes tóxicos decorrentes de atividades industriais ou acidentais.^[12]

As características dos esgotos domésticos urbanos variam de acordo com os costumes, hábitos e situação sócio-econômico da população de cada região.^[13] Por isso são necessários métodos específicos para o tratamento, além do acompanhamento através de medidas de parâmetros físicos, químicos e biológicos, utilizados no monitoramento para determinar a eficiência do tratamento nas estações de tratamentos de esgotos.^[14]

1.3 Parâmetros de Qualidade de Águas

No projeto de uma estação de tratamento de esgoto, normalmente não há interesse em se determinar os diversos compostos presentes na água de esgoto, devido não somente às dificuldades em se executar vários destes testes em laboratório, mas também pelo fato dos resultados em si não serem diretamente utilizáveis como elemento de projeto e operação.^[14] Assim é preferível a utilização de parâmetros indiretos que traduzam o caráter ou o potencial poluidor dos esgotos. Os parâmetros associados à qualidade dos esgotos domésticos são:

i) Coliformes termotolerantes

A quase totalidade dos microorganismos patogênicos é incapaz de viver em sua forma adulta ou reproduzir-se fora do organismo que lhe serve de hospedeiro e, portanto, têm vida limitada quando se encontram na água.

São vários os agentes de destruição natural de organismos patogênicos nas águas armazenadas. Além da temperatura, destacam-se os efeitos da luz, a presença ou não de oxigênio dissolvido, parasitas ou predadores de bactérias, substâncias tóxicas ou antibióticos produzidos por outros microorganismos como algas e fungos, etc. Entre os principais tipos de organismos patogênicos que podem ser encontrados na água estão as bactérias, os vírus, os protozoários e os helmintos.

Devido à grande dificuldade para identificação dos vários organismos patogênicos encontrados na água, dá-se preferência, para isso, a métodos que permitam a identificação de bactérias do grupo coliformes mais comumente coliformes termotolerantes (C Term) que, por serem habitantes normais do intestino humano, existem, obrigatoriamente, em águas contaminadas por matéria fecal. Este grupo é portanto utilizado como indicador de contaminação fecal e da provável presença de microorganismos patogênicos.^[15]

ii) pH

A importância dos valores de pH está associada à influência que este parâmetro exerce nos processos que ocorrem no meio aquático e na toxicidade de alguns compostos. Segundo Esteves^[16], observa-se uma estreita interdependência do pH com as comunidades vegetais e animais do meio aquático. Este fenômeno ocorre na medida em que as comunidades aquáticas interferem no pH, assim como o pH interfere de diferentes maneiras no metabolismo destas comunidades, como por exemplo, através da assimilação de CO₂, durante o processo fotossintético, as macrófitas aquáticas e as algas podem elevar o pH do meio.

iii) Alcalinidade

A alcalinidade de uma água indica sua capacidade de neutralizar ácidos, sendo, em geral, função das concentrações de carbonatos e bicarbonatos do meio aquático. Influenciam também a alcalinidade, embora em menor porcentagem e com menor frequência, boratos, silicatos, fosfatos e ácidos orgânicos fracos.^[1] A

alcalinidade da água é importante no tratamento de água, na química e biologia de água natural.

iv) Oxigênio Dissolvido

O oxigênio nas águas é proveniente de duas fontes: a fotossíntese e a aeração. A fotossíntese é o processo no qual o dióxido de carbono, que pode ser derivado de HCO_3^{-1} , é convertido a carbono e oxigênio pelas plantas clorofiladas em presença de luz solar difusa ou direta. A aeração consiste na difusão do ar atmosférico na água, com conseqüente aumento do teor de oxigênio da mesma. Por outro lado, as perdas de oxigênio são originadas pela decomposição por matéria orgânica, por evasão para atmosfera, pela respiração de organismos aquáticos e pela oxidação de íons metálicos, como por exemplo ferro e manganês.^[16]

v) Amônia, nitrito e nitrato

O nitrogênio é um dos elementos limitantes do crescimento celular, embora é abundante na natureza e bastante importante em sistemas de tratamentos de efluentes. Na biodegradação de matéria orgânica nitrogenada, nos processos biológicos de tratamentos de esgotos ocorre a conversão em compostos mais simples como amônia (NH_4^+), nitrito (NO_2^-), nitrato (NO_3^-) e nitrogênio molecular (N_2).^[17] Dentre as diferentes formas, o nitrato e o íon amônio assumem grande importância nos ecossistemas aquáticos uma vez que representam as principais fontes de nitrogênio para os produtores primários.^[16] Além das formas inorgânicas de nitrogênio, os produtores primários podem assimilar formas orgânicas como: uréia, aminoácidos, peptídeos, etc. Estes compostos, juntamente com outros (muitos deles quimicamente ainda não identificados), formam um grupo de compostos genericamente denominado de nitrogênio orgânico dissolvido.^[16]

vi) Fósforo

O fósforo na água apresenta-se principalmente nas formas: ortofosfatos, polifosfatos e fósforo orgânico. Os ortofosfatos estão diretamente disponíveis para o metabolismo sem necessidade de conversões a formas mais simples. As formas que os ortofosfatos se apresentam na água são: PO_4^{-3} , HPO_4^{-2} , H_2PO_4^- , H_3PO_4 . A importância do fósforo associa-se principalmente à sua atuação como um nutriente essencial e fator limitante para o crescimento e reprodução das comunidades do meio aquático e responsável pelos processos de eutrofização e alteração do seu equilíbrio dinâmico.^[16]

No esgoto bruto há quantidades substanciais de fósforo, nas formas: orgânica e inorgânica complexa (polifosfatos), como aquelas provenientes de detergentes e pesticidas; e ortofosfato inorgânico solúvel, que é um produto final no ciclo do fósforo é a forma mais prontamente disponível para uso biológico. Durante o processo de tratamento biológico, os compostos orgânicos são degradados, podendo disponibilizar ortofosfatos solúveis e polifosfatos que, quando hidrolisados, são convertidos em ortofosfatos. Em um efluente orgânico bem estabilizado, ou seja, submetido a tratamento secundário, o ortofosfato é a forma predominante do fósforo.^[17]

vii) Sulfato

Compostos de enxofre são comumente encontrados nas águas. Dentre as suas várias formas na água, o SO_4^{-2} e o gás sulfídrico (H_2S) são os mais freqüentes, sendo que o sulfato assume maior importância na produtividade do ecossistema, visto que constitui a principal fonte de enxofre para os produtores primários.^[16] Os compostos de enxofre orgânico são muito comumente encontrados em sistemas aquáticos naturais e na decomposição de compostos provenientes de importantes processos microbiológicos. Na presença de oxigênio, um grupo de bactérias pode converter a forma de enxofre reduzida (sulfeto, S^{-2}) a íon sulfato.^[12] Por outro lado, em processos anaeróbios, bactérias redutoras de sulfato também chamadas de

dessulfurantes ou de respiração de sulfato podem converter o íon sulfato (receptor de elétrons) a gás sulfídrico. Como doadores de hidrogênio são utilizados ácidos orgânicos, hidrogênio molecular e alcalóides.^[16]

viii) Temperatura

A temperatura influencia processos biológicos, reações químicas e bioquímicas que ocorre na água e também outros processos como a solubilidade dos gases dissolvidos. O aumento da temperatura aumenta a velocidade das reações químicas, de tal forma que em lagos tropicais, a decomposição da matéria orgânica é muito rápida, implicando no alto consumo de oxigênio dissolvido. Por outro lado, na superfície, estimula a fotossíntese e portanto a produção de O₂

ix) Turbidez

A turbidez é atribuída principalmente às partículas sólidas em suspensão e representa o grau de resistência à passagem de luz através da água, conferindo uma aparência turva à mesma. Pode ser provocada por materiais biológicos (bactérias e algas, por exemplo) e detritos orgânicos e inorgânicos.

x) Sólidos

Todas as impurezas da água, com exceção dos gases dissolvidos, contribuem para a carga dos sólidos. Os sólidos podem ser classificados de acordo com o seu tamanho (sólidos dissolvidos ou em suspensão), sua decantabilidade (sedimentáveis ou não) e suas características químicas (voláteis e fixos). Devido a suas características químicas quando submetidos a uma temperatura elevada (550°C), os sólidos têm sua fração orgânica volatilizada (sólidos voláteis), permanecendo apenas a fração inerte não oxidada (sólidos fixos). Considera-se os sólidos sedimentáveis aqueles que sejam capazes de sedimentar no cone Imhoff no período de 1 hora, a fração que não se sedimenta representa os sólidos. Os sólidos totais incluem

os sólidos totais suspensos (fração retida no filtro) e os sólidos totais dissolvidos (filtrado).^[13]

xi) Matéria Orgânica Carbonácea - Demanda Química de Oxigênio (DQO)

A matéria orgânica presente nos esgotos é essencial em sua caracterização, pois ela é a causadora do principal problema de poluição das águas. As substâncias orgânicas presentes nos esgotos são constituídas principalmente por proteínas (40% a 60%), carboidratos (25% a 50%), óleos e graxas (10%), uréia, surfactantes e outros, em menores quantidades.^[12] Na prática, não é feita a caracterização da matéria orgânica em termos de proteínas, gorduras, carboidratos, etc, devido à dificuldade dessas determinações e por não ser necessário, na Engenharia Sanitária se calcula a matéria orgânica biodegradável (DBO_5) e a matéria orgânica mais recalcitrante (DQO).^[12]

1.4 Revisão bibliográfica do uso da quimiometria em águas

Andrade *et al.*,^[18] aplicaram as técnicas PCA e análise fatorial em programa de monitoramento de águas profundas numa região da Espanha. Foram investigadas 29 variáveis físicas e químicas, das quais foram selecionadas 6 variáveis como sendo as mais representativas das variações sazonais e espaciais. Assim eles conseguiram uma redução significativa de custos e horas de trabalho.

Marengo *et al.*,^[19] investigaram a importância relativa entre mais de um fator ambiental sobre as alterações na qualidade de água da lagoa de Veneza na Itália. Foram investigados 18 amostras e 13 variáveis físicas e químicas provenientes de 3 estações de coleta em 6 diferentes períodos de amostragem. Dos resultados de análise de agrupamento hierárquico (HCA) e PCA eles concluíram que a posição geográfica dos pontos de coleta foi o fator mais significativo para as alterações na qualidade das águas do sistema aquático investigado. Eles destacaram mudanças nos diferentes pontos de coleta ocasionados pelas variações sazonais e que podem ser melhor

retratadas pelas medidas de pH, temperatura, OD, fosfato e nitrato que são variáveis associadas à atividades de fitoplânctons.

Vega *et al.*,^[20] usaram as técnicas HCA e PCA para investigar a influência de poluição antrópica e sazonalidade na qualidade de água do rio Pisuerga na Espanha. Com a análise exploratória associada ao conhecimento de hidroquímica foi possível reduzir o número de componentes principais (PCs) para 4. A realização de uma rotação varimax destas PCs conduziu a um número reduzido de fatores, sendo cada um deles relacionado a um pequeno grupo de variáveis experimentais. Eles concluíram que o conteúdo mineral do rio está relacionado ao fator 1, a poluição antropogênica ao fator 2 e a temperatura da água ao fator 3.

Einax *et al.*,^[21] usaram métodos quimiométricos para investigar contaminação por fontes poluidoras como descargas de efluentes, bem como o comportamento da deposição e remobilização de metais na água e no sedimento do rio Saale na Alemanha. Neste trabalho foram amostrados 29 locais de setembro de 1993 a agosto de 1994. Os resultados da rotação varimax na análise dos dados de sedimento mostraram que a alta concentração de Mn, Co e Ni, proveniente de influência geogênica da montanha Thuringian está relacionada ao fator 1, as altas concentrações de Cu e Hg provenientes do efluente da indústria Middle German ao fator 2 e a carga de Zn e Pb originada do efluente da indústria de fibra têxtil Thuringian ao fator 3. A modelagem de regressão por mínimos quadrados parciais (PLS) se mostrou uma ferramenta importante na análise quantitativa do comportamento de metais pesados entre o sedimento e a água do rio.

Abollino *et al.*,^[22] aplicaram as técnicas quimiométricas (HCA, PCA e DA) para investigar a concentração de 14 elementos coletados de água do mar Ross (Antártica) no período de 1994 a 1996. Foram encontrados níveis de elementos mais importantes e menos importantes e de elementos traços dissolvidos. Observou-se também uma diminuição após desagregação da banquisa de gelo devido à diluição da água e à bioacumulação sendo a última particularmente relevante para as

concentrações de metais pesados. Os resultados também mostraram a formação de 2 grupos de amostras respectivamente antes e depois da desagregação da banquisa de gelo. Concluíram também que a distribuição dos pesos nas 2 PCs examinadas parece ser indicativa de que o comportamento dos elementos dissolvidos não só está relacionado com o equilíbrio químico mas também com fatores físicos e biológicos.

Ana Vazquez *et al.*,^[23] aplicaram as técnicas de análise exploratória para monitorar a qualidade água da chuva em termos de espécies iônicas em 17 estações de amostragem de uma rede de monitoramento de água pluvial. Os pontos de amostragem foram escolhidos, estrategicamente numa área geográfica da Galícia (Nordeste da Espanha) de janeiro a dezembro de 1999. Concluíram que existem diferenças na composição da água da chuva entre os pontos de amostragem, baseado na sua localização geográfica. Concluíram também que foram detectados três grupos referentes aos pontos de amostragem com base na maior contribuição das variáveis: acidez parcialmente neutralizado (*partially neutralized acidity component - PNA*), influência marinha (*sea-salt base component - SSB*), atividade antropogênica (*anthropogenic activity component - AAC*) e acidez total (*total acidity component - TAC*).

Brengaine e Marhaba,^[24] usaram PCA no monitoramento de variáveis físicas, químicas e biológicas para avaliar a qualidade de água em 12 locais ao longo do rio Passaic em Nova Jersey nos EUA, durante o ano de 1998. Os resultados mostraram impacto de testes padrões orgânicos, biológicos, e químicos. Foram observados através de PCA que estes padrões eram de origens natural, urbana, industrial e agrícola. Durante o período de seca os parâmetros orgânicos eram os mais fortemente correlacionados.

Mendiguchía C. *et al.*,^[25] aplicaram as técnicas quimiométricas HCA, PCA e análise fatorial (AF) para estudos de efeitos causados por fatores antropogênicos na água do rio Guadalquivir localizado no Sul da Espanha. Foram medidos vários parâmetros físicos e químicos em 26 pontos de amostragem localizado ao longo do rio

no período de 2001 a 2002. Os resultados identificaram quatro zonas diferentes do rio com qualidade de águas diferentes. Concluíram também que as variáveis Mn^{2+} , NO_2^- e NH_4^+ estão relacionados ao fator 1 (poluição urbana), PO_4^{3-} e Sólidos Suspensos (SS) ao fator 2 (atividade agrícola) e finalmente variáveis originadas de fontes diferentes como NO_3^- ou Cu^{2+} ao fator 3.

Parinet *et al.*,^[26] usaram PCA para estudar a caracterização do estado trófico de lagos. Foram medidos 18 variáveis de qualidade de água em 21 pontos de amostragem em 10 lagos localizados na cidade de Yamoussoukro situada no centro da Costa de Marfim. Eles concluíram que o potencial trófico dos lagos investigados estariam muito bem definidos ao se considerar apenas as variáveis pH, condutividade, absorvância a 254 nm e índice de permanganato (em meio ácido), relacionadas aos efeitos de eutrofização.

Miettinen, T. *et al.*,^[27] usaram PCA e análise paralela de fatores (PARAFAC) para caracterizar os diferentes estágios do sistema de tratamento de lagoa biológica em uma estação de tratamento de esgoto em Melbourne na Austrália. Foram estudadas 19 variáveis físicas e químicas no período de 7 anos. Concluíram que PCA e PARAFAC são ferramentas eficientes na visualização de perfis multivariados de sistemas complexos de tratamento de esgoto biológico. Foram identificados também os sucessivos estágios do processo de purificação.

Singh, K. P. *et al.*,^[28] utilizaram HCA, PCA, AF e DA para avaliar a qualidade de água do rio Gomti no Nordeste da Índia. Foram monitorados 24 parâmetros em 8 diferentes pontos de amostragem, durante um período de 5 anos, selecionados de acordo com o nível de poluição. Os resultados da HCA mostraram que os 8 pontos de amostragem agrupam-se em três classes com propriedades semelhantes que se refletem nas características da qualidade de água e de fontes de poluição (natural e antropogênica). Este resultado sugere uma redução no número de pontos de amostragem. Os seis fatores obtidos da AF indicam que os parâmetros

responsáveis pela variação na qualidade de água estão relacionados principalmente aos sais solúveis (natural) e a carga de poluição orgânica (de fonte antropogênica).

Os resultados de DA mostram que 5 parâmetros foram capazes de discriminar as variações sazonais e apenas 9 parâmetros para discriminar as variações espaciais nas três diferentes regiões do rio.

Tomasz *et al.*,^[29] usaram HCA, PCA, AF e DA, para classificar a qualidade de água do rio Brda localizado no norte da Polônia no período de 1994 a 2002. Os resultados de HCA mostraram que os pontos de amostragem encontram-se arranjados em quatro grupos de amostras que podem estar associados a poluição proveniente de descarga de esgoto urbano e esgoto proveniente de cultura de peixes. Os resultados de DA confirmaram a formação destes grupos. A AF mostra que o fator 1 está relacionado a poluição química e o fator 2 está associada a composição biológica ou eutrofização.

Shrestha S. *et al.*,^[30] aplicaram HCA, PCA, AF e DA para avaliar a variação sazonal e espacial no monitoramento de 12 parâmetros de qualidade de água do rio Fuji no Japão. As amostras foram coletadas de 13 diferentes pontos no período de 1995 a 2002. Os resultados de HCA mostraram que os 13 pontos de coleta foram arranjados dentro de 3 grupos de amostras associados a características semelhantes na qualidade de água. Concluíram também que os fatores da análise fatorial indicam que as variações dos parâmetros responsáveis pela qualidade de água, estão principalmente relacionados a descarga pluvial e temperatura; poluição orgânica de esgoto doméstico em áreas relativamente menos poluídas; nutrientes de fonte difusa, agricultura e plantações de pomares, em áreas de poluição média; e poluição orgânica e de nutrientes proveniente de fontes pontual, estações de tratamento de esgoto doméstico e industrial em áreas altamente poluídas.

1.5 Objetivos do trabalho

Dentro do tema tratamento de esgotos, os objetivos deste trabalho foram:

- i) investigar o funcionamento de uma estação de tratamento de esgoto doméstico, que foi inicialmente projetada para funcionar como uma ETE composta de uma lagoa aerada seguida de lagoa de maturação. Porém, por um problema mecânico nos aeradores, foi adaptada para operar como uma ETE do tipo lagoa anaeróbia seguida de lagoa facultativa, situada no bairro de Mangabeira da cidade de João Pessoa - Paraíba - Brasil;
- ii) desenvolver um novo método para determinação de DQO em lagoas de estabilização, usando espectros de reflectância difusa do séston na região do infravermelho próximo e calibração multivariada.

CAPÍTULO 2

Investigação do funcionamento da ETE-Mangabeira

2.1 Tratamento de esgoto doméstico

A remoção dos contaminantes no tratamento de esgoto, de forma a adequar o posterior lançamento a um sistema hídrico a uma qualidade desejada ou a um padrão de qualidade vigente, está associada aos níveis e a eficiência do tratamento.^[12]

O tratamento de esgoto é usualmente dividido nas seguintes etapas:

i) Tratamento primário

Consiste na remoção dos sólidos grosseiros por meio de grades e a sedimentação (caixa retentora de areia e decantadores) ou flotação de materiais constituídos principalmente de partículas em suspensão sedimentáveis e sólidos flutuantes. Neste nível de tratamento também faz parte do sistema um canal para realização de medição de vazão, que usualmente é formado por uma calha de dimensões padronizadas (calha de Parshall) ou vertedores retangulares ou triangulares.^[12]

ii) Tratamento secundário

Este tratamento destina-se principalmente a degradação biológica de compostos carbonáceos. Quando é feita a degradação, naturalmente ocorre a decomposição de carboidratos, óleos, graxas e proteínas à compostos mais simples, como por exemplo: CO_2 , H_2O , NH_3 , H_2S e outros, dependendo do tipo de processo. As bactérias que efetuam o tratamento, por outro lado, se reproduzem e tem a sua massa total aumentada em função da quantidade de matéria degradada.^[12] Existe uma grande variedade de métodos de tratamento de esgoto de nível secundário, os mais comuns são: lagoas de estabilização, lodos ativados, fossas sépticas e filtros biológicos.^[12]

iii) Tratamento terciário

Esta fase do tratamento tem como objetivo principal a remoção de contaminantes específicos (usualmente nutrientes N e P), geralmente fundamentada em processos biológicos realizados em fases subseqüentes

denominadas de nitrificação e desnitrificação. Fazem parte também desta fase do tratamento os processos químicos, como desinfecção e precipitação do fósforo.^[12]

2.2 Tratamento de esgoto doméstico baseado em lagoa de estabilização

As lagoas de estabilização são sistemas que quando bem projetados e construídos permitem o tratamento de esgotos domésticos e despejos industriais com altos teores de matéria orgânica.^[12]

No Brasil a maioria dos sistemas de tratamentos de esgotos sanitários envolve o uso de lagoas de estabilização devido ao seu baixo custo de construção e simplicidade de operação.^[31]

Von Sperling^[32] menciona que as lagoas de estabilização são bastante indicadas para regiões de clima quente e para países em desenvolvimento, devido à disponibilidade de área superficial em um grande número de localidades, clima favorável, com temperatura e insolação elevadas, simplicidade operacional e utilização de pouco ou nenhum equipamento.

Os sistemas de tratamento mais usados são: lagoas anaeróbias, lagoas facultativas, lagoas aeradas, lagoas de maturação, lagoas anaeróbias seguidas de lagoa facultativa (conhecido como sistema australiano) e lagoas de estabilização como pós-tratamento de reatores anaeróbios.^[32] Dentre estes tipos de lagoas discutiremos em mais detalhes os seguintes tipos:

i) Lagoas anaeróbias

As lagoas anaeróbias constituem um sistema de tratamento, onde a existência de condições estritamente anaeróbias é essencial. Nesta configuração típica da lagoa anaeróbia ocorre a ausência de oxigênio livre na quase totalidade do seu volume, não há agitação por via externa e o fluxo no interior do reator é mais no sentido horizontal, o que favorece a sedimentação de sólidos suspensos dos esgotos ao longo do percurso do líquido entre a entrada e a saída. Portanto, uma parte significativa do volume útil da lagoa funciona como um sedimentador, sendo, o fundo da lagoa anaeróbia a região mais ativa, na qual ocorrem as reações. Neste sentido se desenvolve uma biomassa essencialmente para a digestão da matéria

orgânica particulada sedimentada, limitando a eficiência na remoção de DBO total, daí a necessidade de tratamento complementar para a fração mais solúvel. Uma modificação introduzida em anos recentes tem sido o uso de uma rede de canalização para melhor distribuição do afluente em vários pontos no fundo da lagoa, com a finalidade de melhorar o contato biomassa-esgoto, buscando aproximar este tratamento ao de um reator anaeróbio de manta de lodo (UASB).^[12]

As lagoas anaeróbias apresentam as seguintes características:^[12]

- i) tempo de detenção hidráulica na faixa 1 a 6;
- ii) carga orgânica volumétrica entre 0,1 e 0,3 Kg DBO₅/m³.dia;
- iii) profundidade entre 2 e 5 metros;
- iv) Eficiência de remoção da DBO₅ entre 50% e 60%, dependendo da temperatura média dos esgotos.

Por estes valores típicos, verifica-se que uma lagoa anaeróbia é um reator de grandes dimensões (tempo de detenção hidráulica na ordem de dias), além disso, a carga orgânica aplicada pode ser considerada baixa se comparada a outros reatores anaeróbios, que pode acomodar cargas bem superiores, uma vez que o tempo de detenção hidráulica nestes casos são da ordem de horas.^[12]

ii) Lagoas facultativas

As lagoas facultativas são caracterizadas pela combinação de grupos de bactérias anaeróbias, aeróbias e facultativas. Elas operam basicamente com tempo de detenção por um período de tempo longo o suficiente para que os processos naturais de estabilização da matéria orgânica se desenvolvam. As vantagens desse tipo de lagoa são a grande simplicidade e a confiabilidade da operação, uma vez que os processos naturais são via de regra confiáveis (sem necessidade de equipamentos). Por outro lado, os processos são lentos necessitando de longos tempos de detenção (de 15 a 45 dias) para que as reações se completem, o que implica em grandes requisitos de área para sua implementação. A eficiência do sistema é usualmente satisfatória, podendo chegar a níveis comparáveis aos da maior parte dos tratamentos secundários. Com base nos critérios de área e volume,

a profundidade é um compromisso entre essas duas variáveis. O conhecimento disponível é ainda limitado para se otimizar a profundidade da lagoa, de forma a obter maiores benefícios . A tendência atual adota lagoas com profundidade variando entre 1,5 a 3,0 m.^[32]

iii) Lagoas aeradas

As lagoas aeradas, diferentemente das lagoas de estabilização facultativas, não são processos estritamente naturais, e assim dependem da introdução artificial do oxigênio através de aeradores, que servem não somente para garantir a oxigenação do meio, mas também para manter os sólidos em suspensão dispersos no meio líquido, favorecendo o desenvolvimento dos organismos decompositores aeróbios da matéria orgânica solúvel e finamente particulada.

Segundo Jordão e Pessoa^[33], devido ao fornecimento contínuo de oxigênio e à capacidade de mistura dos equipamentos de aeração, nas lagoas aeradas permite-se adotar maiores profundidades e menores tempos de detenção que nas lagoas de estabilização, sendo, portanto, menor o requisito de área em comparação com as dessas lagoas clássicas. Ainda assim, existe um tempo de detenção mínimo, abaixo do qual a lagoa aerada tem o seu desempenho seriamente comprometido. Trabalhos de pesquisa em laboratório e medições em escala real indicam que após um certo tempo de detenção é negligenciável a obtenção de maior remoção da matéria orgânica solúvel. Assim, tempos de detenção acima de 4 dias não produzem melhorias significativas nos efluentes das lagoas aeradas, sendo razoável admitir tempos de detenção compreendidos entre 2 e 4 dias.^[14] Usualmente, adotam-se para esse tipo de lagoa, profundidades na faixa de 2,5 a 4,5 m.

iv) Lagoas de maturação

As lagoas de maturação possibilitam um “polimento” no efluente de qualquer sistema de tratamento de esgotos. O principal objetivo das lagoas de maturação é remover os microorganismos patogênicos e não uma remoção

adicional de matéria orgânica em termos de DBO. Esse tipo de lagoa constitui uma alternativa bastante econômica à desinfecção do efluente por métodos convencionais, como cloração. As lagoas de maturação são projetadas com baixas profundidades (0,8 a 1,5 m) de forma a maximizar a fotossíntese e os efeitos bactericidas da luz solar.^[32]

v) Sistema australiano

O sistema australiano de tratamento associa as características das lagoas de estabilização anaeróbias e facultativas. Desta forma constituem um sistema de tratamento de esgoto em série, que ocupa uma área menor que a de uma lagoa facultativa única e permite que 50% da DBO seja estabilizada na lagoa anaeróbia (mais profunda e com menor volume), enquanto parte da DBO remanescente é removida na lagoa facultativa.^[32]

2.3 Técnicas quimiométricas

As técnicas quimiométricas empregadas neste capítulo foram:

- a) Gráficos de Box-Whisker;
- b) Análise Fatorial;
- c) Análise Discriminante.

2.3.1 Gráficos de Box-Whisker

Os gráficos de Box-Whisker não são técnicas de representação multivariada de dados, mas são muito úteis para comparar diferentes grupos de objetos em relação a uma característica. Eles representam a distribuição dos parâmetros (dispersão, variância, etc) das variáveis. Esta ferramenta pode ser usada para dar uma idéia visual dos testes t e F . Todo grupo é apresentado como uma caixa em um diagrama^[34], como mostrado na **Figura 2.1**.

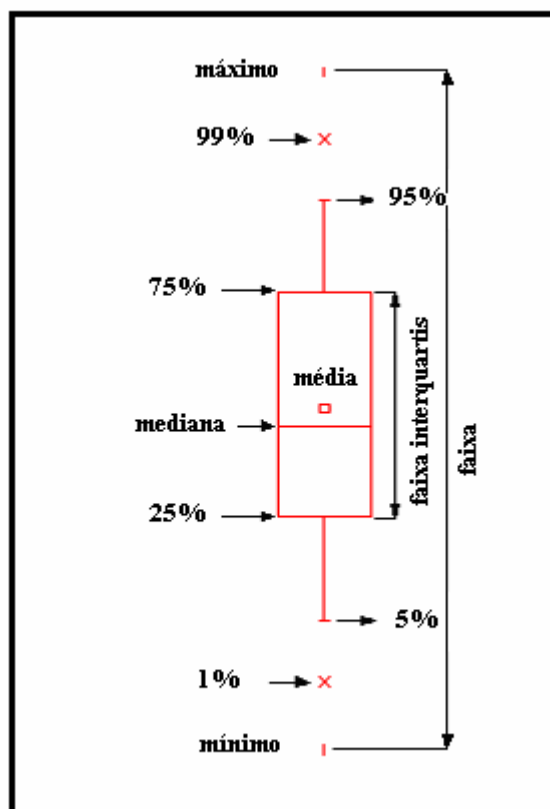


Figura 2.1 Ilustração gráfica de Box-Whisker

2.3.2 Análise fatorial

A análise fatorial^[35] é uma técnica de análise multivariada cuja finalidade é reduzir a dimensionalidade das colunas da matriz original (variáveis em estudo) por meio da identificação de fatores latentes, sintetizando a informação comum a grupo de variáveis. O modelo pode ser descrito pela **Equação 2.1**, em que as variáveis são expressas como uma função linear de um número reduzido de fatores latentes.^[30]

$$X_{ji} = p_{t1}t'_{1i} + p_{t2}t'_{2i} + p_{t3}t'_{3i} + \dots + p_{tm}t'_{mi} + e_{ii} \quad (2.1)$$

onde, X é o valor medido de uma variável; p = peso do fator, t' = fator; e = termo residual que responde por erros ou outra fonte de variação; i = número de amostras; j = número de variáveis e m = número total de fatores.

Dois grupos de técnicas quimiométricas de análise fatorial foram utilizados neste trabalho:

i) Análise por componentes principais

A análise por componentes principais consiste em projetar os objetos (pontos), na direção do espaço dimensional das variáveis que contenham a maior quantidade de informação (variância) possível. Como resultado, obtém-se um novo sistema de eixos ortogonais denominados componentes principais. Como as primeiras PC contêm a maior variância dos dados, é possível representar as informações mais relevantes do conjunto de dados usando um número pequeno de componentes.^[36-37] Por outro lado, as informações irrelevantes, erros aleatórios, ruídos nos dados são mais significativas nas últimas componentes principais.

O princípio geral da PCA pode ser resumido através da **Equação 2.2**, na qual uma matriz de dados originais (**X**) é decomposta em uma matriz de escores (**T**) e em uma matriz de pesos (**P**), que expressam a variação principal dos dados; e uma matriz de resíduo (**E**), que representa a contribuição não sistemática. O sobrescrito **t** representa uma matriz transposta.

$$X = TP^t + E \quad (2.2)$$

Os dados multidimensionais em **X** estão dispostos em um conjunto de componentes ortogonais, que correspondem a uma combinação linear aproximada dos dados originais.^[36]

ii) Rotação varimax

O método de rotação varimax tem sido muito bem sucedido como uma abordagem analítica para a obtenção de uma rotação ortogonal de fatores.^[35] Para uma melhor interpretação dos fatores, faz-se necessário realizar uma rotação dos eixos relativos aos fatores comuns na matriz de pesos fatoriais.

A rotação consiste em girar os eixos em um ângulo (θ), oferecendo uma nova estrutura para os pesos, de tal forma que cada variável tenha peso alto em um único fator e pesos mais baixo ou médio nos demais fatores.^[38] A rotação não produz uma estrutura visível quando o número de fatores é maior que dois, neste caso programas computacionais são usados para executar a rotação.

A rotação pode ser realizada no sentido horário (**equação 2.3**) e anti-horário (**equação 2.4**) e produz as matrizes \mathbf{R}^+ e \mathbf{R}^- .

$$\mathbf{R}^+ = \begin{bmatrix} \cos(\theta) & \text{sen}(\theta) \\ -\text{sen}(\theta) & \cos(\theta) \end{bmatrix} \quad (2.3)$$

$$\mathbf{R}^- = \begin{bmatrix} \cos(\theta) & -\text{sen}(\theta) \\ \text{sen}(\theta) & \cos(\theta) \end{bmatrix} \quad (2.4)$$

Os novos pesos para uma rotação de dois fatores^[38] são calculados conforme a **Equação 2.5**.

$$\hat{\Lambda} = \mathbf{P} \mathbf{R} \quad (2.5)$$

onde, θ é o ângulo de rotação, \mathbf{P} é a matriz dos pesos estimados originais, $\hat{\Lambda}$ é a nova matriz dos pesos estimados e \mathbf{R} é a matriz de rotação. Desse modo, os fatores serão colocados em uma única posição tal que cada fator possa ser interpretado pelos maiores pesos possíveis relacionados ao menor número de variáveis.^[38]

A rotação varimax pode ser melhor visualizada através de uma representação gráfica em que, os fatores são rotacionados de um ângulo θ , no sentido horário, onde, cinco variáveis imaginárias (V_1, V_2, V_3, V_4 e V_5) são projetadas em um diagrama fatorial ortogonal bidimensional apresentado na **Figura 2.2**.^[35]

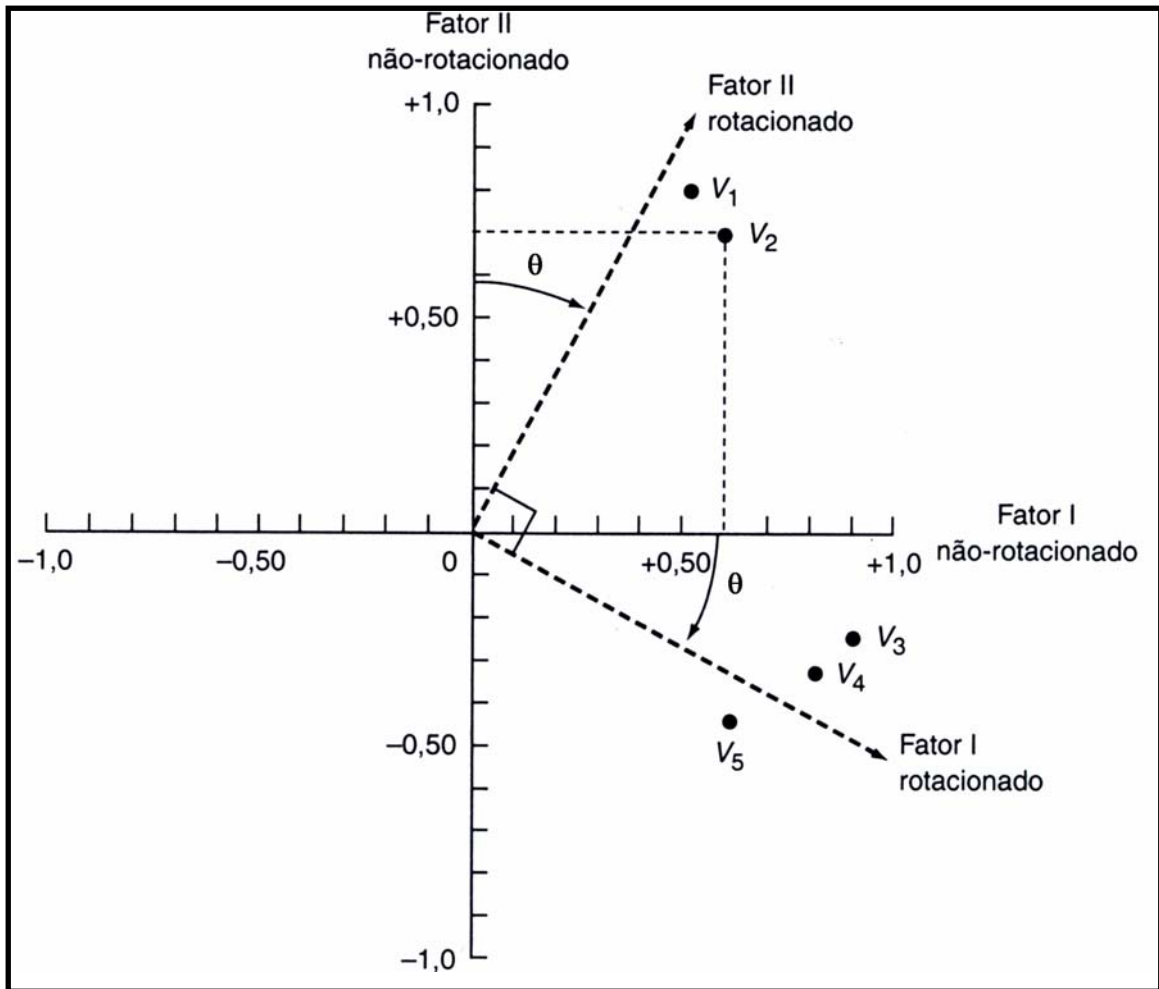


Figura 2.2 Ilustração de uma análise de rotação fatorial ortogonal, adaptado de Hair^[35].

Na **Figura 2.2** os eixos com linhas contínuas representam a análise fatorial não rotacionada e os eixos com as linhas tracejadas a análise fatorial rotacionada. Os eixos são rotulados com zero na origem e prolongados para +1,0 ou -1,0, onde os números nos eixos representam os pesos fatoriais.

O peso fatorial para a variável V_2 no fator II não-rotacionado é determinado desenhando-se uma linha tracejada horizontalmente a partir do ponto projetado até o eixo vertical do fator II. De modo similar, para determinar o peso da variável V_2 no fator I não rotacionado uma linha vertical é tracejada a partir da variável V_2 até o eixo horizontal do fator I. Um procedimento similar é adotado para calcular os pesos fatoriais das outras variáveis para as análises não rotacionadas.

Podemos observar na **Figura 2.2** que:

- i) todas as variáveis possuem pesos fatoriais elevados no primeiro fator não rotacionado;
- ii) os pesos das variáveis V_1 e V_2 são muito altos na direção positiva no segundo fator não rotacionado;
- iii) a variável V_5 tem peso moderadamente alto na direção negativa do segundo fator não rotacionado;
- iv) V_3 e V_4 têm pesos relativamente baixos na direção negativa do segundo fator não rotacionado;

Por outro lado, a partir de uma inspeção visual da **Figura 2.2**, podemos observar que após a rotação, fica óbvio que há dois agrupamentos de variáveis:

- i) um grupo formado pelas variáveis V_1 e V_2 , com pesos altos no fator II;
- ii) um outro grupo formado pelas variáveis V_3 , V_4 e V_5 , com pesos altos no fator I.

Contudo, tal padrão de variáveis não ficaria tão óbvio a partir da análise fatorial não rotacionada como observado acima.^[35]

2.3.3 Análise Discriminante

A análise discriminante consiste na separação a priori de determinadas classes de objetos. A separação é obtida estabelecendo os pesos na variável estatística (uma combinação linear de duas ou mais variáveis independentes) para cada variável com a finalidade de maximizar a variância-covariância entre as classes relativa a variância-covariância dentro das classes.^[35] Novos objetos são classificados por meio de uma função discriminante chamada de escores Z discriminante mostrada na **Equação 2.6**.

$$Z_{jk} = a + W_1 X_{1k} + W_2 X_{2k} + W_n X_{nk} \quad 2.6$$

onde, Z_{jk} = escore Z discriminante da função discriminante j para o objeto k

a = intercepto

W_i = peso discriminante para a variável independente i

X_{ik} = variável independente i para o objeto k

As medidas de λ de *Wilks* são testes estatísticos usados para avaliar o nível de significância estatística do poder discriminatório da função discriminante.

O parâmetro tolerância calculado como $1-R_i^2$ (onde, R_i^2 é a quantidade de variância da variável independente i explicada por todas as outras variáveis independentes), representa a proporção da variação nas variáveis independentes não explicada pelas variáveis que já estão no modelo (função). Uma tolerância de zero significa que a variável independente sob consideração é uma combinação linear perfeita das variáveis independentes já incluídas no modelo enquanto que uma tolerância de um significa que uma variável independente é totalmente independente das outras variáveis que já estão no modelo.^[35]

2.4 OBJETIVOS

- i) Observar os fenômenos que envolvem os processos de decomposição anaeróbia da matéria orgânica em termos da ciclagem do carbono e o processo fotossintético, que acontece durante o tratamento de esgoto da ETE-Mangabeira;
- ii) Investigar as influências sazonais e eventuais durante o processo de tratamento de esgoto da ETE-Mangabeira;
- iii) Investigar o funcionamento da ETE-Mangabeira diante da adaptação do projeto de construção e operação realizado na mesma.

2.5 EXPERIMENTAL

2.5.1. Sistema Hídrico Investigado

O sistema hídrico investigado é a estação de tratamento de esgoto (ETE), inserida na bacia do rio Piratibe. Esta bacia está situada no bairro de Mangabeira, localizado à -7,19722 graus de latitude (S) e -34,81306 graus de longitude (W) da cidade de João Pessoa - Paraíba - Brasil (**Figura 2.3**).

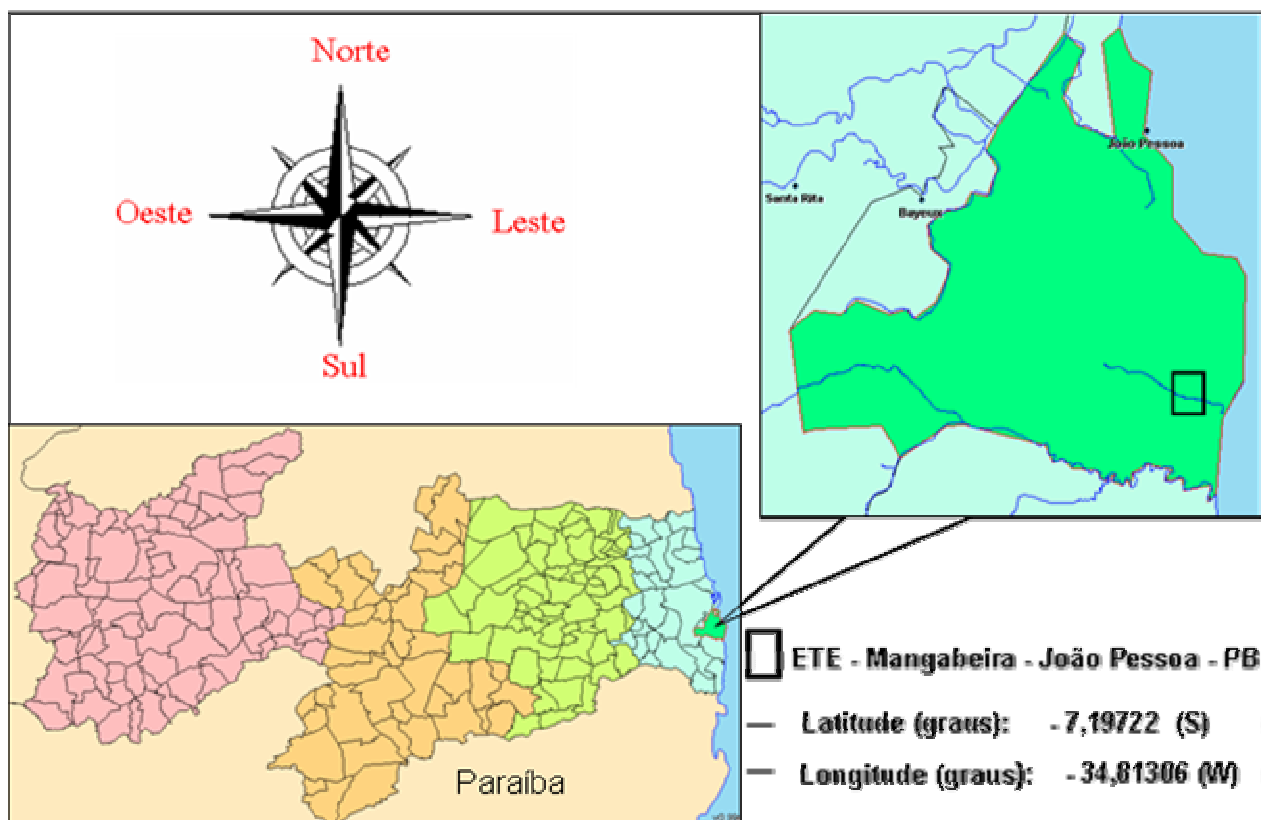


Figura 2.3 Mapa de localização da ETE-Mangabeira

A ETE-Mangabeira foi originalmente construída de modo a operar com duas unidades de tratamento de esgoto em paralelo. Cada unidade era constituída de duas lagoas aeradas em paralelo seguida de uma lagoa de maturação. A ETE-Mangabeira foi projetada para atender uma população estimada de 33.000 habitantes, que gera, estimadamente, uma vazão afluyente para cada unidade de tratamento de $8.275 \text{ m}^3/\text{dia}$.^[39]

Um problema mecânico nos aeradores das duas lagoas de ambas unidades de tratamento, acrescido das dificuldades financeiras do órgão administrador (a Companhia de Água e Esgoto da Paraíba - CAGEPA), fez com que o sistema originalmente projetado fosse adaptado, gerando um novo sistema de tratamento. Esta nova ETE-Mangabeira é agora constituída de duas lagoas anaeróbias em série, seguida de uma lagoa facultativa. Assim, o sistema passou a ser uma ETE do tipo sistema australiano, ou seja, lagoa anaeróbia seguida de lagoa facultativa (**Figura 2.4**).

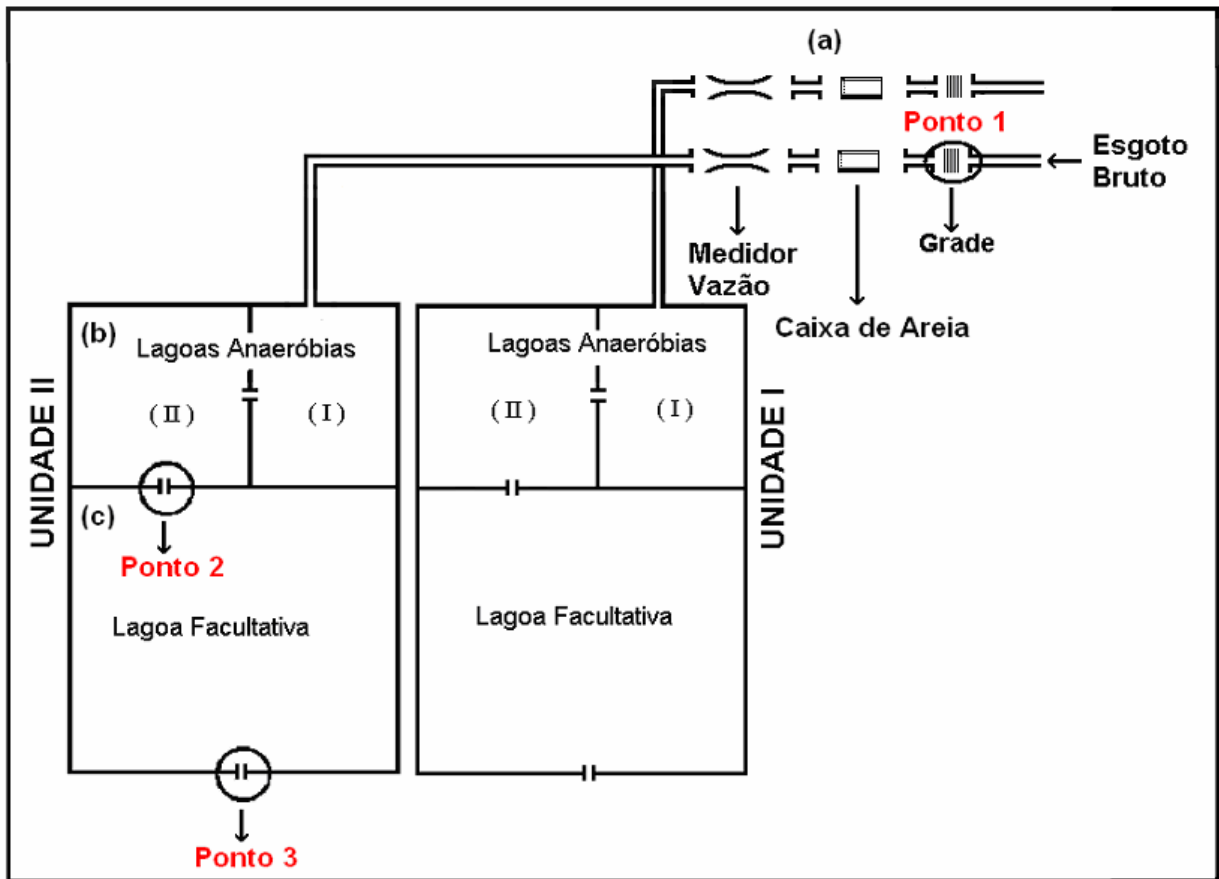


Figura 2.4 Diagrama esquemático da ETE-Mangabeira. Região de tratamento primário (a) e de tratamento secundário (b) e (c).

Três pontos de amostragem foram monitorados, na unidade II da ETE-Mangabeira como indicado na **Figura 2.4**. Estes pontos são caracterizados como:

- i) ponto 1 - afluente anaeróbio (AA);
- ii) ponto 2 - efluente anaeróbio (EA);
- iii) ponto 3 - efluente facultativo (EF).

A **Figura 2.5** mostra uma fotografia das duas lagoas anaeróbias da unidade II do sistema ETE-Mangabeira.

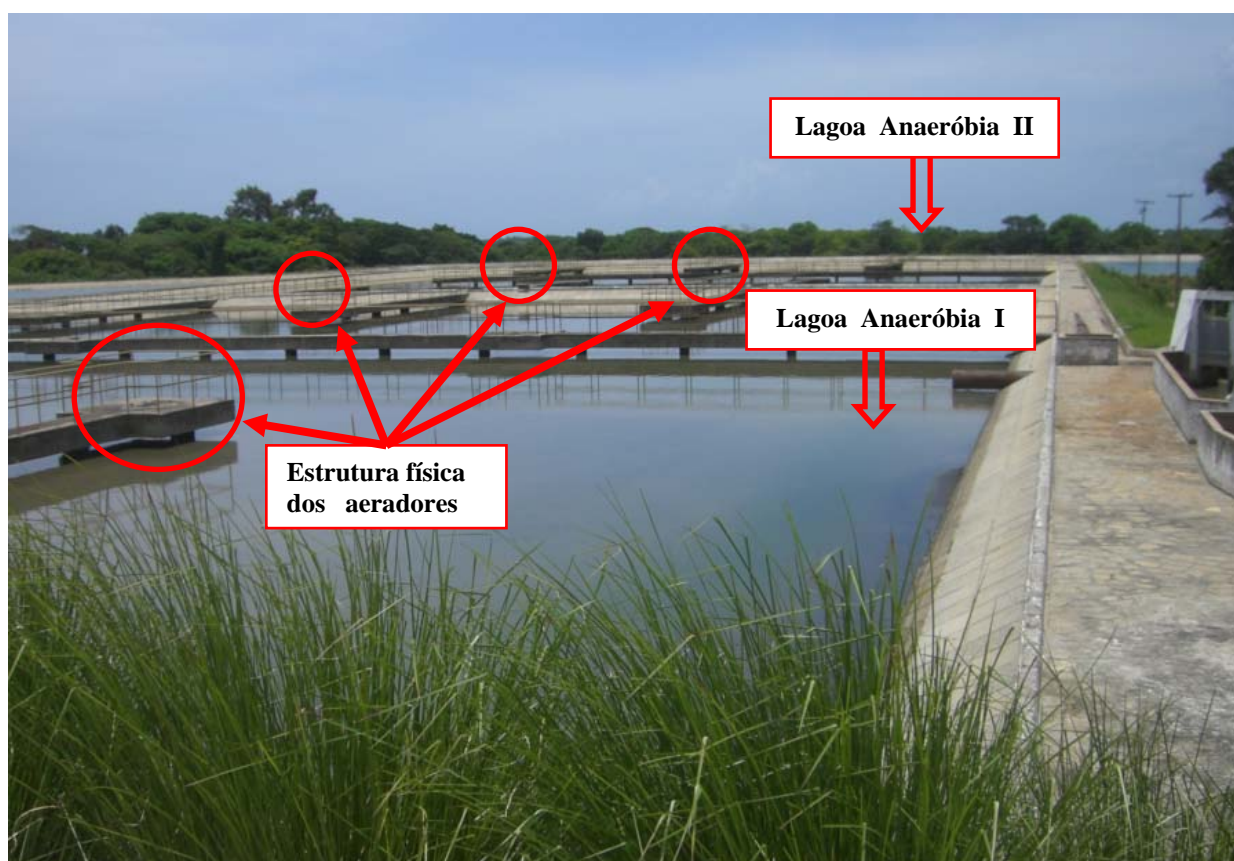


Figura 2.5 Fotografia das duas lagoas anaeróbias da ETE-Mangabeira.

Como pode ser observado na **Figura 2.5**, as estruturas físicas que sustentavam os aeradores das duas lagoas aeróbias, ainda permanecem nas lagoas anaeróbias da ETE-Mangabeira.

As condições técnicas de operação das lagoas anaeróbias e facultativa e os valores da eficiência da ETE-Mangabeira em porcentagem de DBO₅, DQO e C. Term são apresentadas nas **Tabelas 2.1 e 2.2**.

Tabela 2.1 – Condições técnicas de operação das lagoas anaeróbia e facultativa, fornecidos pela CAGEPA

Lagoas	Área (ha)	Profundidade útil (m)	Volume útil (m ³)	Tempo de detenção (dias)
*Anaeróbia (I e II)	0,40	3,0	12.150	1,45
Facultativa	3,2	1,4	44.350	5,4

* valores unitários para cada lagoa

Tabela 2.2 – Eficiência de DBO₅, DQO e C. Term para ETE-Mangabeira, fornecidos pela CAGEPA

Mes/Ano	*Eficiência (%)		
	DBO	DQO	C. Term
Fevereiro/2005	90,52	79,14	99,35
Março/2005	92,18	71,99	96,09
Abril/2005	89,49	74,55	99,75
Mai/2005	89,40	71,59	99,75
Junho/2005	92,45	57,37	99,64
Julho/2005	89,52	54,72	99,93
Agosto/2005	91,59	66,63	99,39
Setembro/2005	91,59	66,63	99,39
Novembro/2005	92,43	68,32	-
Dezembro/2005	93,90	71,35	-
Janerio/2006	93,29	76,07	-
Fevereiro/2006	88,70	58,81	-

2.5.2. Dados de precipitação pluviométrica

A **Figura 2.6** mostra os valores da precipitação pluviométrica mensal acumulada na região investigada no período de fevereiro de 2005 à fevereiro de 2006, fornecidos pela SUDEMA.

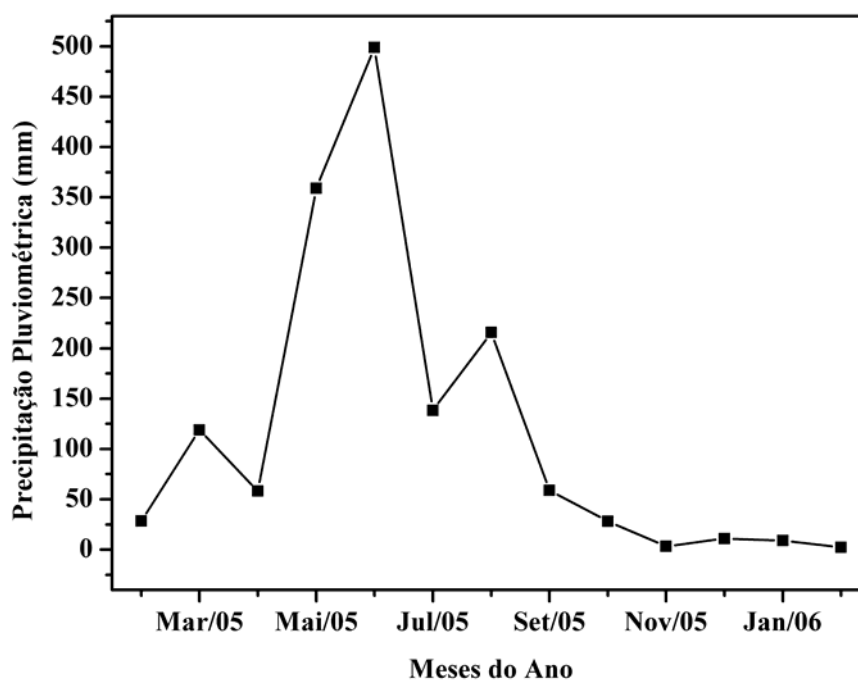


Figura 2.6 Precipitação pluviométrica mensal de fevereiro de 2005 a fevereiro de 2006

Para facilitar a interpretação dos resultados, foram definidas neste trabalho três categorias de amostras de acordo com a ocorrência de precipitação pluviométrica nos respectivos meses de coleta:

- i) índice pluviométrico baixo (abaixo de 20 mm),
- ii) índice pluviométrico médio (entre 20 e 120 mm)
- iii) índice pluviométrico alto (acima de 120 mm).

2.5.3 Coleta das amostras e períodos de amostragem

As amostras de esgoto doméstico e dos efluentes foram coletadas no turno da manhã (entre 7:00 e 9:00 horas) semanalmente em duplicatas de fevereiro a junho de 2005 e mensalmente em triplicatas de agosto de 2005 a fevereiro de 2006 nos pontos de amostragem mostrados na **Figura 2.4**. Por razões operacionais não foram coletadas amostras nos meses de julho e outubro de 2005.

2.5.4 Parâmetros monitorados

A escolha dos parâmetros físicos, químicos e bacteriológicos monitorados na lagoa de estabilização foi baseada:

- i) no projeto de operação da ETE-Mangabeira;
- ii) na norma L1-009 de dezembro de 1986^[40] contida no manual técnico de operação e manutenção de lagoas anaeróbias e facultativas da Companhia de Tecnologia de Saneamento Ambiental de São Paulo (CETESB);
- iii) na disponibilidade do orçamento do projeto;
- iv) na infra-estrutura dos laboratórios;
- v) na alocação de recursos humanos.

Os parâmetros escolhidos foram determinados em um conjunto de 105 amostras, de acordo com os métodos de referência da **APHA - American Public Health Association**.^[41] Os métodos e materiais utilizados são apresentados a seguir:

i) pH

O pH foi determinado pelo método potenciométrico, utilizando um aparelho da marca Quimis modelo Q 400 2.01, no laboratório de Físico-Química/Bacteriologia do Centro Federal de Educação Tecnológica da Paraíba (LFQB-CEFET-PB).

ii) Oxigênio dissolvido (OD)

As medidas de OD foram realizadas no campo. Foi adotado método eletrométrico, usando uma membrana seletiva de oxigênio DurOx 325, acoplada a um Oxímetro de campo WTW Wissenschaftlich, modelo oxi 315i com medidor de temperatura. Após ligar o aparelho fez-se a checagem das baterias e esperou-se 15 min para estabilização da temperatura. O eletrodo foi mantido em um ambiente de ar saturado com vapor d'água e em seguida fez-se a calibração do instrumento seguindo as indicações do fabricante. Após calibração do aparelho procedeu-se a leitura de OD das amostras.

iii) Turbidez

Para determinação da turbidez foi utilizado o método turbidimétrico, empregando-se um turbidímetro modelo AP 2000-IR da Policontrol e suspensões padrão de 10, 100 e 1000 unidades nefelométricas de turbidez (UNT). As amostras foram diluídas com água deionizada na razão de 1:9 (v/v) e em seguida foram analisadas. As medidas foram realizadas no LFQB-CEFET-PB.

iv) Temperatura

A temperatura das amostras foi determinada diretamente na massa do esgoto com um termômetro de filamento de mercúrio de marca INCOTERM com precisão de leitura de 0,1°C.

v) Sulfato

Foi empregado o método turbidimétrico. Inicialmente construiu-se uma curva analítica com padrões de sulfato na faixa de concentração de 0 - 40 mg L⁻¹.

Em seguida, as amostras foram filtradas numa membrana Whatman GF/C (de 5 cm de diâmetro e porosidade em torno de 1,0 μm), num sistema de filtração a vácuo e analisadas em um Espectrofotômetro UV-VIS de varredura da HP (Hewlett Packard), modelo 8453, empregando uma cubeta de quartzo com 1,0 cm de caminho óptico. As determinações de sultato foram realizadas no laboratório de Automação em Química Analítica/Quimiometria (LAQA) e de Química Ambiental (LQAA) da Universidade Federal da Paraíba (UFPB).

vi) Fosfato solúvel

Foi determinado pelo método espectrofotométrico do azul de molibdênio com ácido ascórbico, em amostras filtradas numa membrana Whatman GF/C (de 5 cm de diâmetro e porosidade em torno de 1,0 μm), num sistema de filtração a vácuo. Neste procedimento as formas de fósforo dissolvidas são separadas das suspensas. As leituras das soluções de calibração e das amostras foram feitas no mesmo espectrofotômetro UV-VIS da seção anterior, usando a mesma cubeta e o comprimento de onda de 880 nm. As determinações de fosfato foram realizadas no LAQA-UFPB e LQAA-UFPB.

vii) Alcalinidade Total

Foi empregado o método titulométrico com ácido sulfúrico 0,04 mol L⁻¹ determinando-se os pontos de inflexão da fenoftaleína (pH 8,3) e alaranjado de metila (pH 4,3). As amostras foram diluídas na proporção de 1:9 v/v e em seguida analisadas. As medidas foram realizadas no LFQB-CEFET-PB.

viii) Sólidos totais (ST)

Empregou-se o método gravimétrico.^[41] Inicialmente cápsulas de porcelana foram colocadas na mufla a temperatura de 550°C por 30 min e resfriadas em dessecador por 40 min. Em seguida 125 ml das amostras foram colocadas nas cápsulas, evaporadas em banho-maria até à secura e levadas à estufa a temperatura de 105°C por 12 horas. Finalmente as amostras foram resfriadas em dessecador e pesadas. As medidas foram feitas no LFQB-CEFET-PB.

ix) Amônia

Foi empregado o método do indofenol.^[41] Filtraram-se as amostras numa membrana Whatman GF/C (de 5 cm de diâmetro e porosidade em torno de 1,0 µm), num sistema de filtração a vácuo e analisou-se em seguida usando um espectrofotômetro UV-VIS, já mencionado, em comprimento de onda de 640 nm. As medidas foram feitas no LAQA-UFPB e LQAA-UFPB.

x) Demanda química de oxigênio (DQO).

Foi empregado o método titulométrico com refluxação fechada^[41] no qual baseia na oxidação da matéria orgânica por dicromato de potássio em meio ácido, precedido de digestão em um termoreator Spectroquant da Merck modelo TR 420. Na etapa de digestão, adicionou-se ao tubo digestor 2,5 ml da amostra, 3,5 ml de uma solução de $\text{Ag}_2\text{SO}_4 / \text{H}_2\text{SO}_4$ e 1,5 ml de solução digestora preparada a partir de dicromato de potássio, ácido sulfúrico concentrado e sulfato de mercúrio. Os tubos foram colocados no bloco digestor já pré-aquecido por um período de 2 horas a temperatura de 150°C, resfriados a temperatura ambiente e tituladas com solução padrão de sulfato ferroso amoniacal na presença de indicador ferroin. A determinação de DQO foi realizada no LFQB-CEFET-PB.

xi) Coliformes termotolerantes (C. Term)

As amostras foram processadas pela técnica de membrana filtrante, usando uma membrana Millipore tipo HAWG 047-SO de 47 mm de diâmetro e poros de 0,45 µm. As amostras foram diluídas em solução tampão fosfato ou tampão fosfato-cloreto de magnésio em pH 7,2, preparando-se diluições seriadas decimais até 10^{-4} . O meio de cultura utilizado foi ágar m-FC (DIFCO). A incubação efetuava-se a 44,5°C durante 24 horas. A determinação de coliformes termotolerantes foi realizada no laboratório de Microbiologia da Companhia de Água e Esgoto da Paraíba (CAGEPA-PB).

2.5.5 Tratamento dos dados

Para o tratamento e análise estatístico dos dados foram usados os pacotes computacionais: *STATISTICA FOR WINDOWS* versão 6.0, Matlab versão 6.5 e *OriginLab* versão 7.0. As análises de PCA, rotação varimax e DA foram feitas usando os dados autoescalados.

2.6 RESULTADOS E DISCUSSÃO

A digestão anaeróbia é um processo bioquímico complexo, composto por várias reações seqüenciais, cada uma com sua população bacteriana específica.^[12] A **Figura 2.7** mostra uma representação esquemática dos vários processos que ocorre na digestão anaeróbia. Para digestão anaeróbia de material orgânico complexo, como proteínas carboidratos e lipídios (a maior parte da composição do material orgânico em águas residuárias é formada por esses grupos), podem-se distinguir quatro etapas diferentes no processo global da conversão:^[12]

- i) hidrólise, conversão do material orgânico particulado em compostos dissolvidos de menor massa molecular;
- ii) acidogênese, compostos dissolvidos, gerados no processo de hidrólise ou liquefação, são absorvidos nas células das bactérias fermentativas e, após esta etapa, escretadas como substâncias orgânicas simples como ácidos graxos voláteis de cadeia curta, álcoois, ácido lático, e compostos minerais entre outros.
- iii) acetogênese, conversão dos produtos da acidogênese em compostos que formam os substratos para produção de metano: acetato, hidrogênio e CO₂;
- iv) metanogênese, o metano é produzido pelas bactérias acetotróficas a partir da redução de ácido acético, ou pelas bactérias hidrogenotróficas, a partir da redução de dióxido de carbono.

Uma outra etapa que pode ocorrer, quando da presença de sulfato é a sufetogênese.

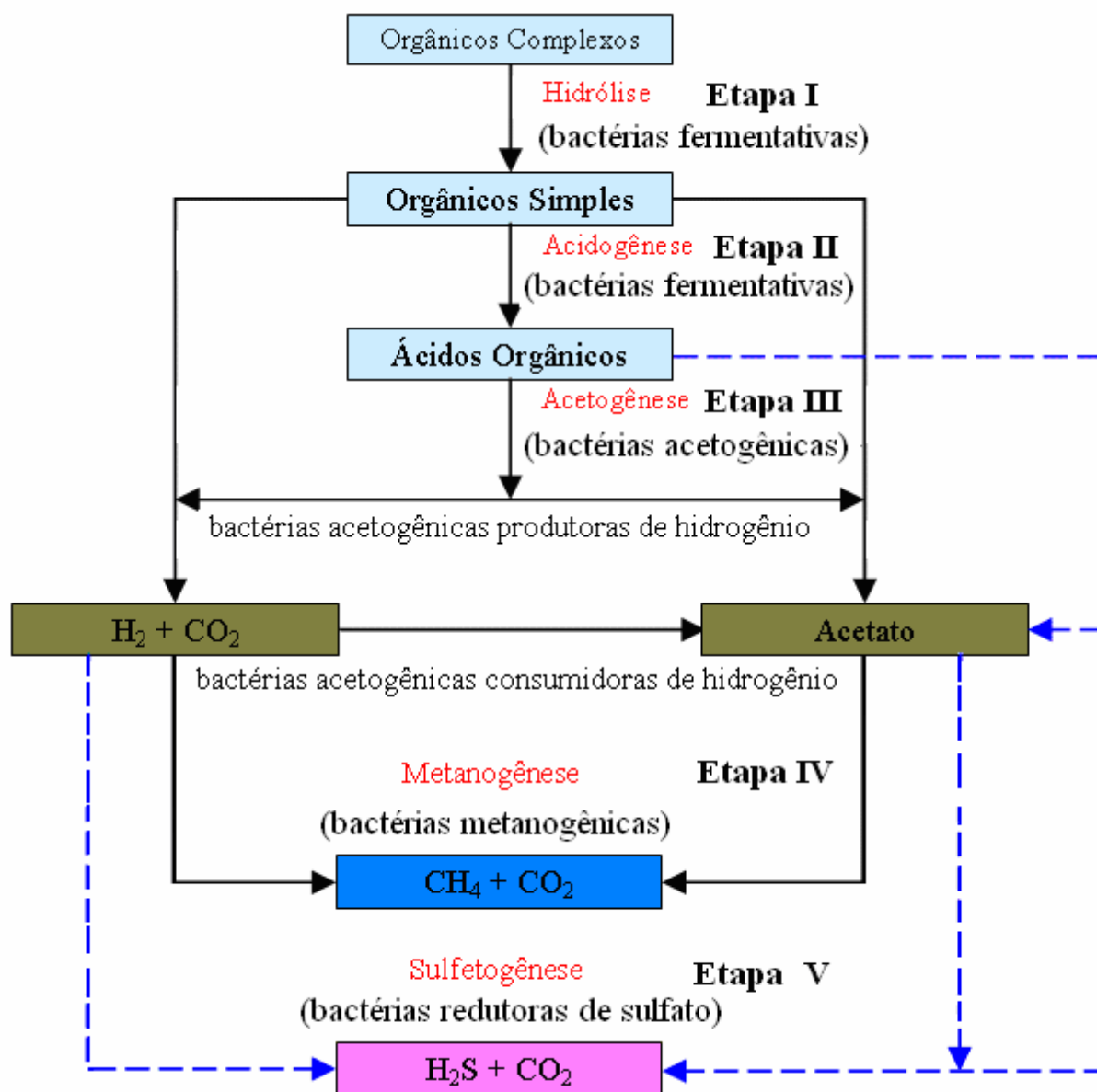


Figura 2.7 Esquema da seqüência de processos na digestão anaeróbia que ocorre em sistemas de tratamento de esgoto.^[42]

2.6.1 – Análise univariada dos dados da ETE-Mangabeira

i) Processos anaeróbios de decomposição de matéria orgânica

Das etapas de I a IV da **Figura 2.7** podemos observar que a matéria orgânica complexa pode ser convertida para uma forma oxidada (CO_2) e uma reduzida (CH_4). A remoção efetiva de matéria orgânica do sistema ocorre quando uma parte significativa desta é convertida para CH_4 , pois este é liberado para a atmosfera. Este processo foi observado na ETE-Mangabeira devido à significativa

redução de DQO (**Figura 2.8**) do ponto AA (entrada de esgoto bruto) para ponto EA (saída da lagoa anaeróbia).

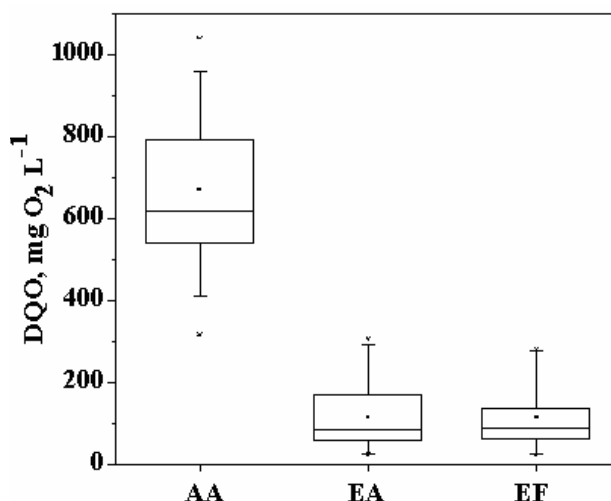


Figura 2.8 Gráficos de *Box-Whisker* para o DQO do afluente anaeróbio (AA), do efluente anaeróbio (EA) e do efluente facultativo (EF) da ETE-Mangabeira.

A formação de ácidos orgânicos, representada na etapa II, e de H₂S, na etapa V em lagoas anaeróbias (**Figura 2.7**) são produzidas pela ação de bactérias denominadas, respectivamente de fermentativas e redutoras de sulfato. Estes dois processos estão ocorrendo na lagoa anaeróbia da ETE-Mangabeira, uma vez que pode ser observada uma redução do pH e do teor de sulfato do ponto AA para o ponto EA, como mostrado nas **Figuras 2.9(a) e 2.9(b)**.

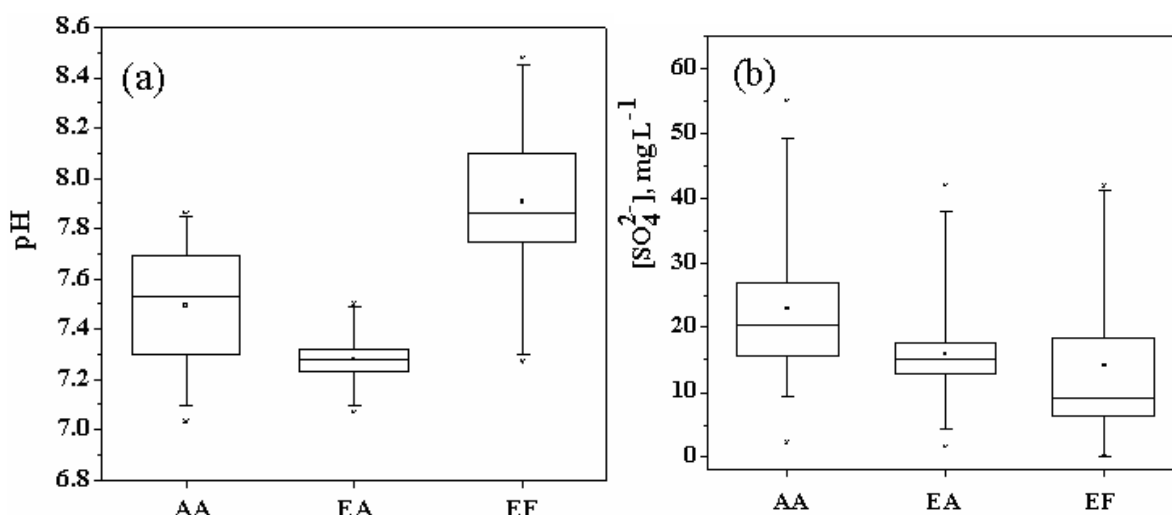


Figura 2.9 Gráficos de *Box-Whisker* para o pH (a) e o sulfato (b) do afluente anaeróbio (AA), do efluente anaeróbio (EA) e do efluente facultativo (EF) da ETE-Mangabeira.

O aumento da concentração de íons H⁺, produzido pela liberação de ácidos orgânicos e expresso pelo decréscimo do pH, desloca o equilíbrio para a

esquerda na **Equação 2.7**, reduzindo o teor de íons bicarbonato (HCO_3^-), como pode ser verificado na **Figura 2.10**.

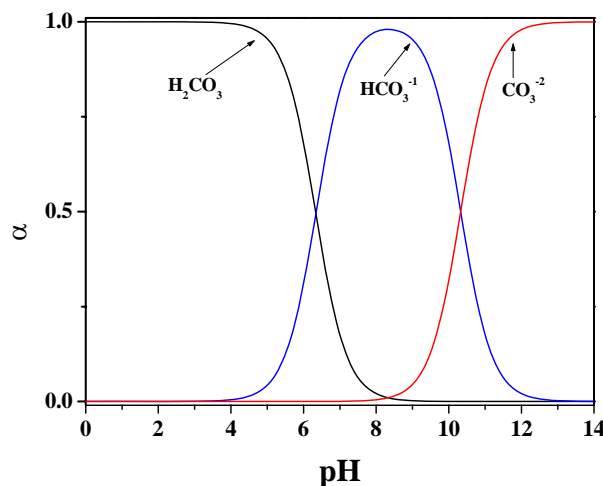


Figura 2.10 Diagrama de distribuição da fração (α) em função do pH ^[1] de algumas espécies em equilíbrio responsáveis pela alcalinidade em sistemas aquáticos.

Como pode ser observado na **Figura 2.11**, ocorreu também uma redução do teor de bicarbonato na lagoa anaeróbia da ETE-Mangabeira, uma vez que a alcalinidade total diminuiu do ponto AA para o ponto EA.

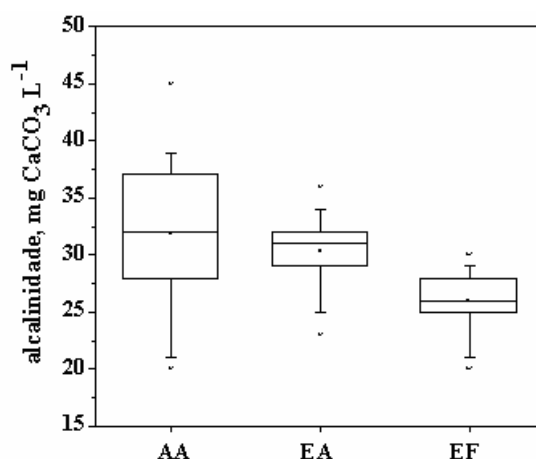


Figura 2.11 Gráficos de *Box-Whisker* para a alcalinidade total do aflente anaeróbio (AA), do efluente anaeróbio (EA) e do efluente facultativo (EF) da ETE-Mangabeira.

A redução de matéria orgânica do ponto AA para o ponto EA da ETE-Mangabeira, levou também a uma redução dos parâmetros sólidos totais e turbidez, como mostrado na **Figura 2.12**.

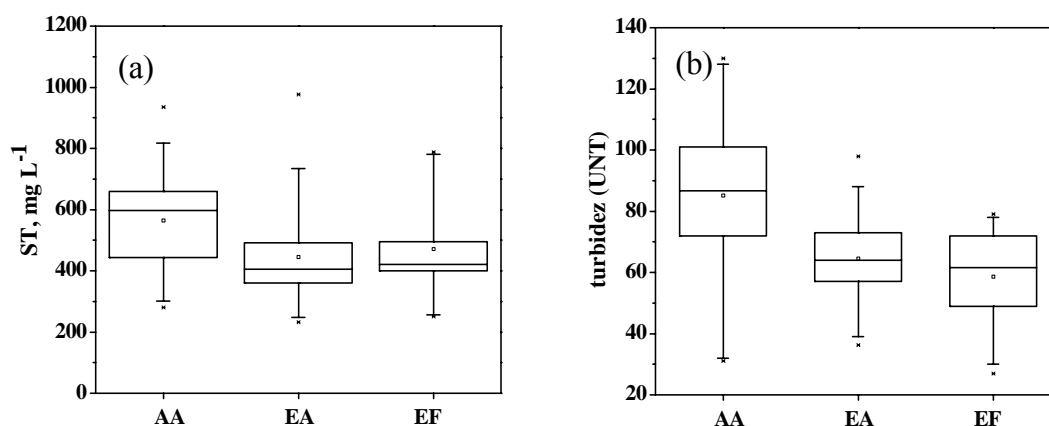
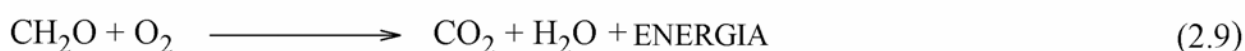
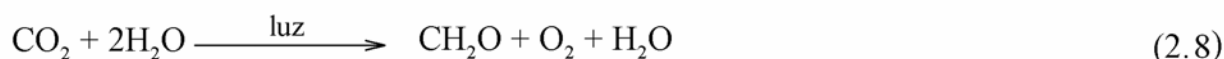


Figura 2.12 Gráficos de *Box-Whisker* para sólidos totais (a) e turbidez (b) do afluente anaeróbio (AA), do efluente anaeróbio (EA) e do efluente facultativo (EF) da ETE-Mangabeira.

ii) Processos fotossintéticos

Nas lagoas facultativas ao longo de vários dias, uma série de mecanismos contribui para purificação dos esgotos. Os processos mais freqüentes que ocorrem neste tipo de lagoas facultativas são: a fotossíntese, favorecida pela penetração da luz solar e pela baixa turbidez e a oxidação de matéria orgânica por meio de respiração aeróbia. O dióxido de carbono produzido nesta oxidação orgânica serve como uma fonte de carbono para as algas.^[17] As **Equações 2.8** e **2.9** apresentam simplificadaamente esses processos.^[32]



A fotossíntese, por depender da energia solar, é mais eficiente próximo à superfície da lagoa. Já em maiores profundidades, o processo de respiração é predominante. No cômputo geral, os organismos autotróficos (fotossintéticos) realizam muito mais fotossíntese do que oxidação. Isto gera um saldo maior de compostos orgânicos que constituem a reserva de energia para os seres heterótrofos, além de um *superavit* de oxigênio que permite a respiração dos outros organismos.^[32]

A produção de algas e o consumo de CO₂ por atividades fotossintéticas, representadas pela **Equação 2.8**, em geral, ocorre em lagoas facultativas^[31] em condições de pH elevado, reduzindo o teor de íons bicarbonato. Esta redução foi também observada na ETE-Mangabeira, uma vez que a alcalinidade total diminuiu

do ponto EA para o ponto EF, como mostrado na **Figura 2.11**. Além disso, a atividade fotossintética leva a produção de oxigênio gasoso. Este processo também foi observado na ETE-Mangabeira, uma vez que se percebe uma presença significativa de OD na lagoa facultativa como mostra a **Figura 2.13**.

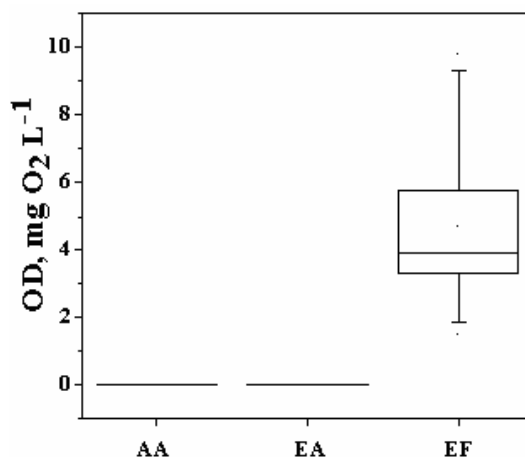


Figura 2.13 Gráfico de *Box-Whisker* para OD do afluente anaeróbio (AA), do efluente anaeróbio (EA) e do efluente facultativo (EF) da ETE-Mangabeira.

Durante o metabolismo fotossintético, as algas consomem nutrientes (principalmente compostos de nitrogênio e fósforo) para o seu desenvolvimento.^[43] Este processo foi também observado na ETE-Mangabeira, uma vez que, ocorreu uma redução gradativa do teor de amônia e de fosfato do ponto EA para o ponto EF, como mostra as **Figuras 2.14 (a) e 2.14 (b)**.

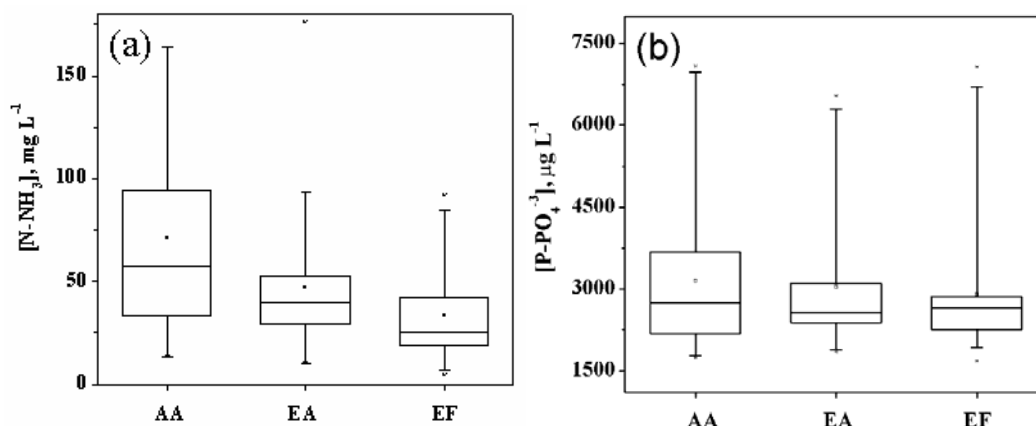


Figura 2.14 Gráfico de *Box-Whisker* para amônia (a) e fosfato (b) do afluente anaeróbio (AA), do efluente anaeróbio (EA) e do efluente facultativo (EF) da ETE-Mangabeira.

Em sistemas do tipo australiano, como a ETE-Mangabeira, espera-se sempre uma redução significativa de coliformes termotolerantes (C. Term) nas lagoas facultativas. Este comportamento foi também observado, do ponto EA para o ponto EF na ETE-Mangabeira, como mostra a **Figura 2.15**.

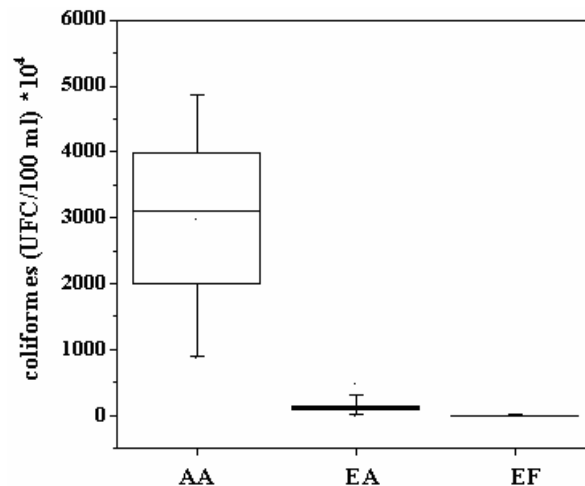


Figura 2.15 Gráfico de *Box-Whisker* para C. Term do afluente anaeróbico (AA), do efluente anaeróbico (EA) e do efluente facultativo (EF) da ETE-Mangabeira.

2.6.2 – Análise multivariada dos dados da ETE-Mangabeira

A **Figura 2.16** mostra o gráfico dos escores e dos pesos, de PC1×PC2 de uma PCA realizada em amostras rotuladas segundo os pontos de coleta (AA, EA e EF) e o índice de precipitação pluviométrica.

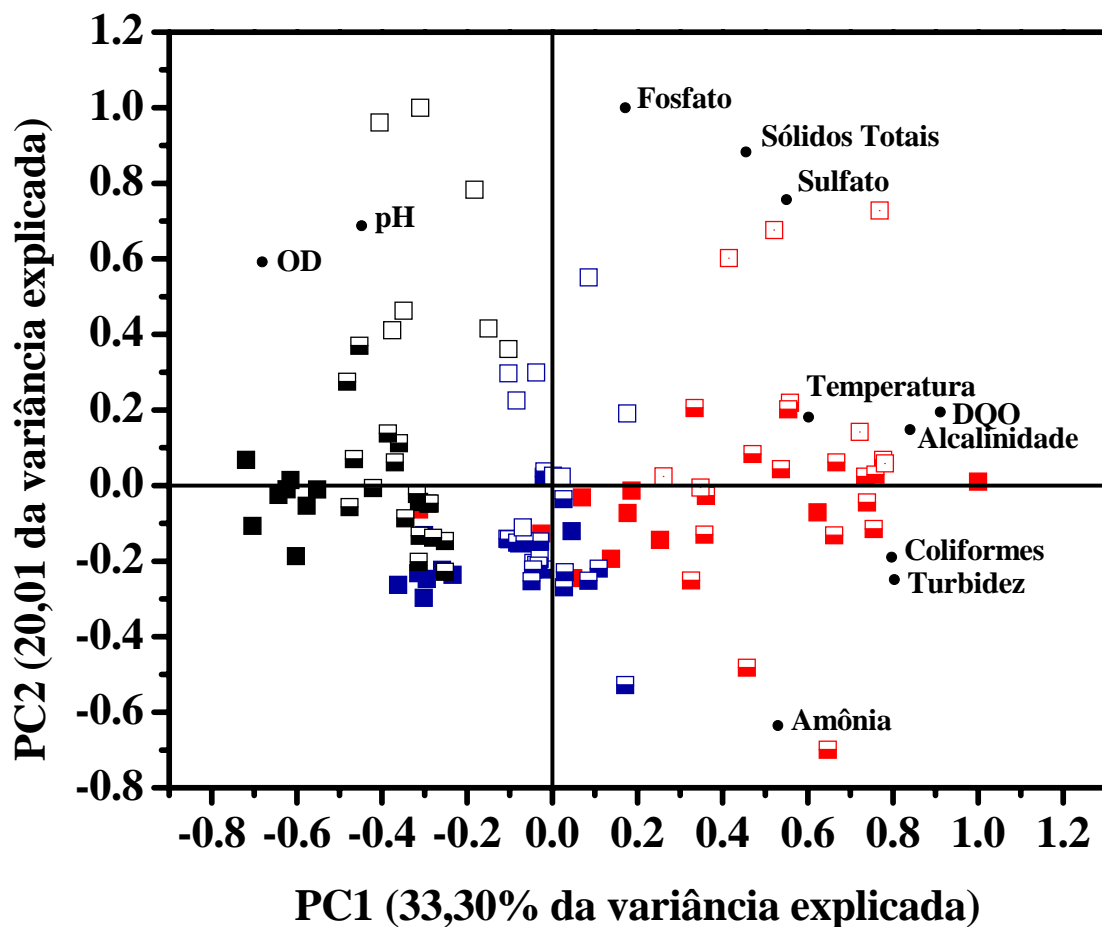


Figura 2.16 Gráfico dos escores e dos pesos de PCA da ETE-Mangabeira, em 3 níveis de índices pluviométricos: alto (■), médio (▒) e baixo (□), onde, AA -afluente anaeróbico (■), EA- efluente anaeróbico(■) e EF- efluente facultativo (■).

A PC1, que explica 33,3% da variância dos dados, está relacionada com a qualidade do esgoto, uma vez que os parâmetros como DQO, C. Term, alcalinidade e turbidez têm pesos bastante positivos nesta componente, enquanto OD e pH têm pesos bastante negativos.

Os escores de PC1 distribuem-se segundo os pontos de coleta, AA, EA e EF, da direita para esquerda, concordando também como nos casos dos pesos com a direção da qualidade do esgoto durante as etapas do tratamento.

Ambos os comportamentos são esperados para uma ETE do tipo australiano, evidenciando o bom funcionamento da ETE-Mangabeira, do sistema adaptado.

A PC2, que explica 20,0% da variância, está associada aos fenômenos da sazonalidade que ocorrem em cada lagoa de estabilização. À medida que se desloca na direção de menor índice pluviométrico, ou seja, na direção de valores positivos de escores e pesos em PC2, percebe-se um aumento da concentração do OD, e do pH. Isto se deve provavelmente a uma maior produção de OD, resultante de atividades fotossintéticas, as quais são mais favorecidas em períodos secos (de baixo índice pluviométrico), devido a uma maior incidência da luz solar. Vale salientar que a fonte de oxigênio proveniente de atividades fotossintéticas em períodos secos predomina sobre outras fontes, como por exemplo, o O₂ atmosférico e da reaeração produzida pelo revolvimento do esgoto no período de chuva. Da mesma forma, também foi observado um aumento de ST neste período. Este comportamento deve estar associado ao aumento de sólidos em suspensão, devido ao crescimento mais intenso das algas.

As concentrações de fosfato e sulfato também aumentam em períodos secos, como se observa pelos altos valores positivos em PC2. Este comportamento que ocorre na ETE-Mangabeira pode estar associado respectivamente a decomposição de compostos orgânicos de fósforo e oxidação do H₂S, mais intenso neste período.

Por outro lado, ocorre uma redução da concentração de amônia em períodos secos, como se observa pelos altos valores negativos em PC2. Isto se deve

provavelmente a uma maior conversão da mesma em nitrato por processos de nitrificação favorecida pelo aumento da atividade fotossintética em período seco.

A rotação varimax pode levar a interpretação mais clara dos fatores em termos das variáveis. ^[35] Os resultados dos três primeiros fatores obtidos a partir da rotação varimax estão apresentados na **Tabela 2.3**.

Tabela 2.3 Pesos das variáveis para os três primeiros fatores rotacionados

Parâmetros	Fator 1	Fator 2	Fator 3
Turbidez (UNT)	0,68	0,03	0,32
Alcalinidade (mg L ⁻¹)	0,49	0,42	0,42
Fosfato (mg L ⁻¹)	-0,16	0,85	-0,03
Temperatura (°C)	0,40	0,32	0,21
pH	0,02	0,07	-0,94
Sólidos Totais (mg L ⁻¹)	0,28	0,72	-0,16
OD (mgO ₂ L ⁻¹)	-0,25	6,6E-4	-0,86
Amônia (mg L ⁻¹)	0,59	-0,38	0,25
C. Term (UFC/100ml)	0,85	-0,05	0,06
Sulfato (mg L ⁻¹)	0,26	0,72	0,05
DQO (mg O ₂ L ⁻¹)	0,85	0,28	0,04

O **fator 1** tem principalmente a contribuição dos parâmetros DQO, turbidez e C. Term, com valores elevados no ponto AA (**Figuras 2.8, 2.12b e 2.16**) que diminuem significativamente no ponto EA em sua maior parte devido ao processo de decomposição de matéria orgânica carbonácea que ocorre na lagoa anaeróbia. O **fator 2** tem uma contribuição maior dos parâmetros fosfato, sulfato e sólidos totais. Finalmente, o **fator 3** tem uma participação maior dos parâmetros pH e OD que aumentam bastante devido aos processos fotossintéticos que ocorrem na lagoa de estabilização do tipo facultativa. Assim, os parâmetros distribuem-se em grupos coerentes com os principais processos esperados para o sistema de tratamento de esgoto do tipo australiano:

- i) os processos de decomposição anaeróbia que estão relacionados aos parâmetros associados ao **fator 1**;

- ii) as atividades fotossintéticas que estão relacionados aos parâmetros associados ao **fator 3**.

A **Figura 2.17** mostra o gráfico dos escores rotacionados do **fator 1** versus **fator 3** para todas as amostras da ETE-Mangabeira.

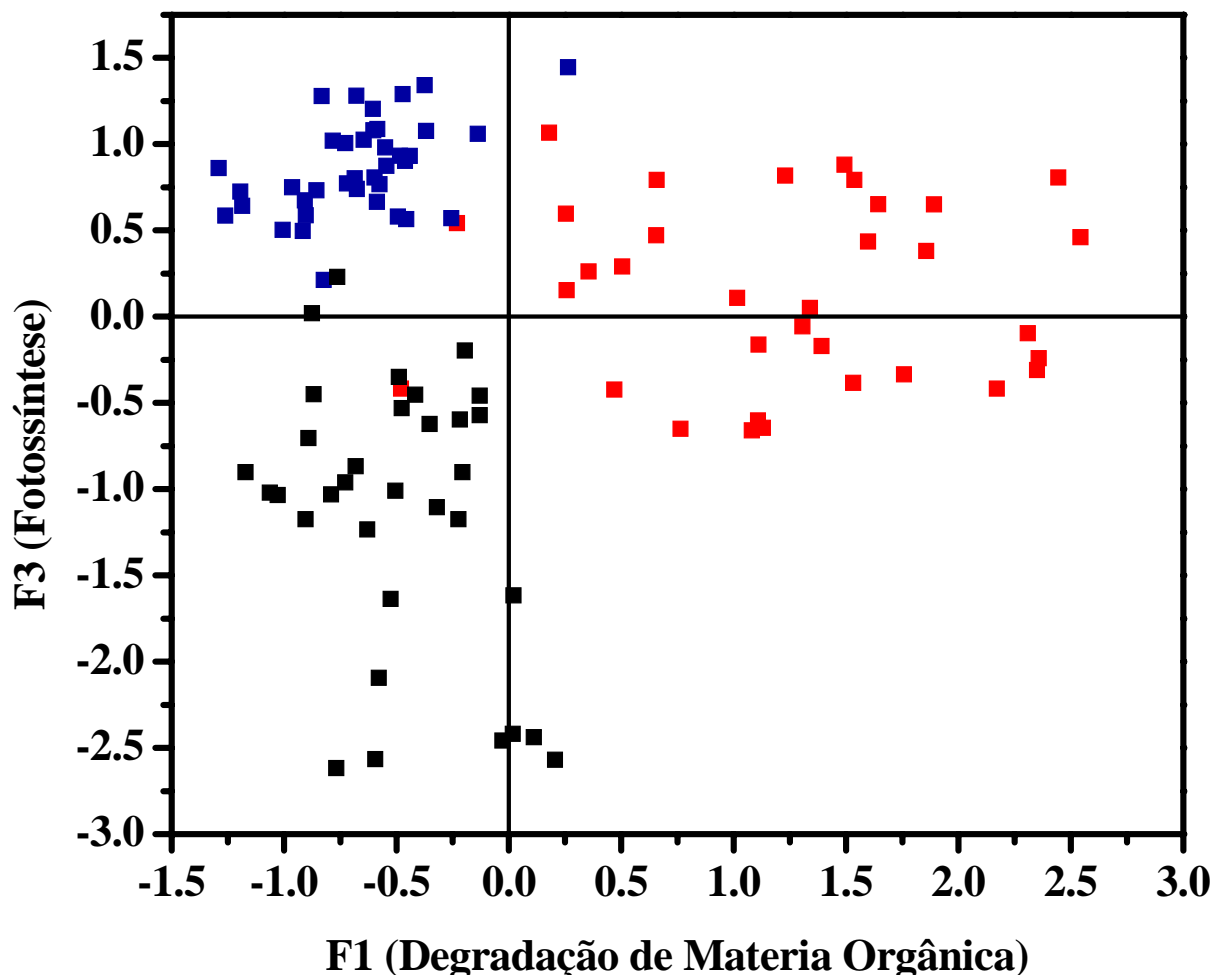


Figura 2.17 Gráfico dos escores da rotação varimax **fator1** (F1) versus **fator3** (F3) para as amostras da lagoa de estabilização, AA (■), EA (■) e EF (■).

O gráfico da **Figura 2.17** mostra que quando nos deslocamos na direção dos escores negativos do **fator1** ocorre uma queda da matéria orgânica relacionado à passagem do esgoto bruto para a lagoa anaeróbia. Também foi observado no **fator3** que quando nos deslocamos no sentido de menores valores de escores percebe-se um aumento da concentração de oxigênio dissolvido (processo fotossintético) relacionado a passagem da lagoa anaeróbia para facultativa.

O gráfico dos escores Z da análise discriminante apresentado na **Figura 2.18** mostra a formação de três agrupamentos de amostras bem definidos e

coerentes com o comportamento observado na ETE-Mangabeira. Estes resultados vêm de encontro com os fatos observados na aplicação da rotação varimax mostrada na **Figura 2.17**. Além disso, a identificação com os três pontos de coleta relacionados aos processos distintos, que ocorrem nas 3 etapas do tratamento de esgotos indicam que a ETE-Mangabeira se comporta de acordo com o esperado para um sistema australiano de tratamento de esgotos.

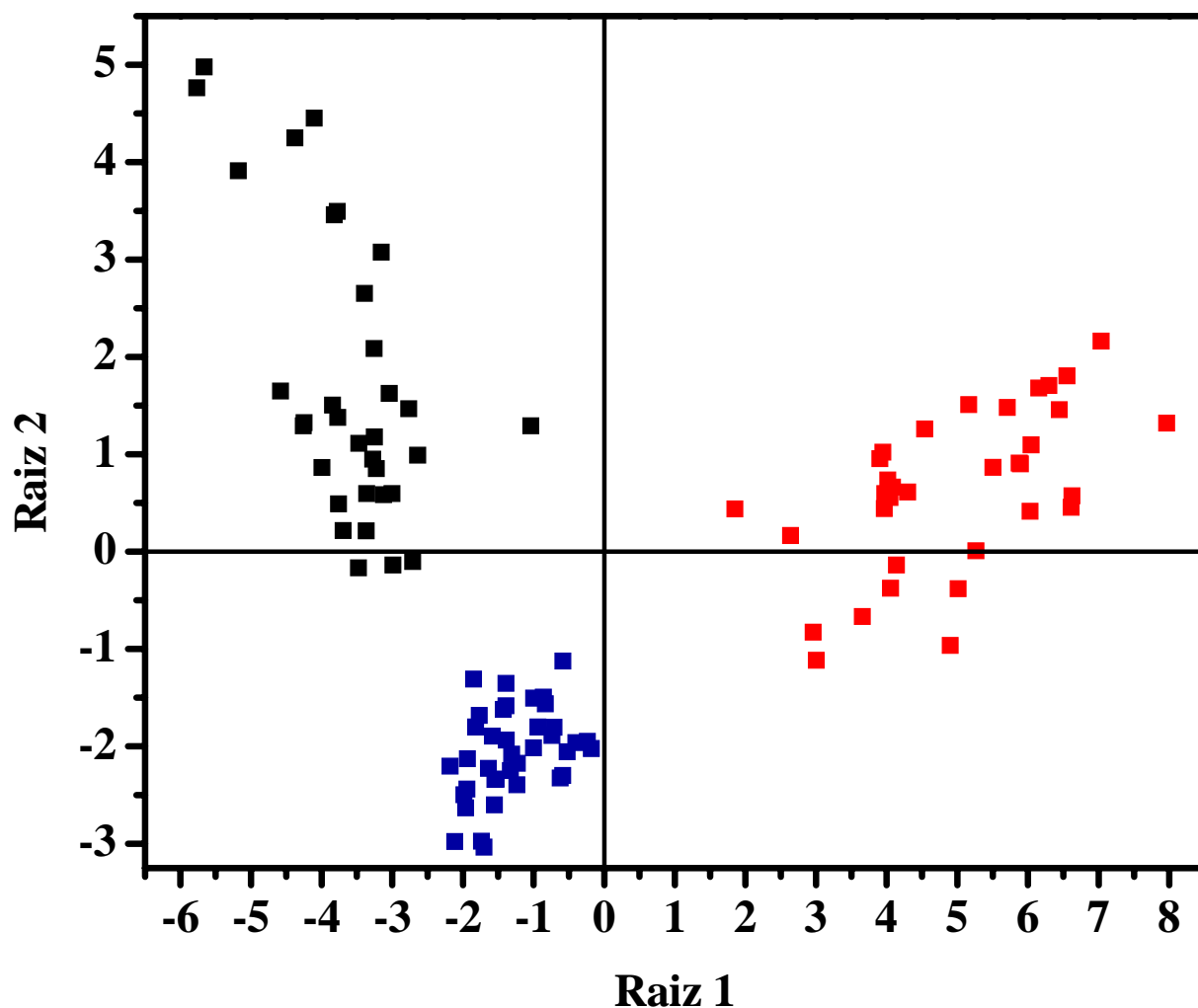


Figura 2.18 Escores Z das amostras da lagoa de estabilização, AA (■), EA (■) e EF (■)

Para avaliar a significância estatística do poder discriminatório da função discriminante foi usada a medida de *Wilks*, cujos parâmetros, λ parcial e tolerância de *Wilks* estão apresentados na **Tabela 2.4**.

Observa-se na **Tabela 2.4** que os parâmetros C. Term, DQO e OD apresentam menores valores de λ parcial de *Wilks*. Isto se deve a uma contribuição mais significativa dos processos de decomposição de matéria orgânica que ocorre

na lagoa anaeróbia (relacionados às variáveis DQO e C. Term) e das atividades fotossintéticas (relacionada principalmente a OD) que ocorrem na lagoa facultativa. Os altos valores de tolerâncias sugerem que não há redundância dos parâmetros monitorados que foram usadas na construção do modelo DA.

Tabela 2.4-Parâmetros de avaliação da significância estatística da função discriminante

Parâmetros monitorados	Parâmetros de <i>Wilks</i>	
	λ Partial	Tolerância
Turbidez (UNT)	0,948697	0,609065
Alcalinidade (mg L ⁻¹)	0,840985	0,600069
Fosfato (µg L ⁻¹)	0,949422	0,498383
Temperatura (°C)	0,963687	0,642143
pH	0,923206	0,595309
Sólidos Totais (mg L ⁻¹)	0,992790	0,664292
OD (mgO ₂ L ⁻¹)	0,647244	0,622399
Amônia (mg L ⁻¹)	0,987324	0,773374
C. Term (UFC/100ml)	0,646634	0,673404
Sulfato (mg L ⁻¹)	0,959604	0,500563
DQO (mg O ₂ L ⁻¹)	0,696708	0,638009

Os resultados da DA, assim como os resultados da análise fatorial (rotacionada e não rotacionada) mostram que as amostras coletadas nos três pontos AA, EA e EF são realmente distintas, o que reforça a idéia de que diferentes processos estão ocorrendo como discutido acima. Além disso, todos os processos evidenciados e as lagoas onde eles ocorrem, concordam com a dinâmica esperada para uma ETE do tipo sistema australiano. Assim, os resultados multivariados adicionados aos resultados univariados observados nos gráficos de *Box - Whisker* demonstram que, apesar das adaptações não inicialmente projetadas, a ETE-Mangabeira está funcionando satisfatoriamente.

CAPÍTULO 3

Desenvolvimento de um novo método para determinação de DQO
em esgoto doméstico

3.1 Determinação de DQO

Nos últimos anos os problemas ambientais relacionados com atividades humanas (esgotos domésticos e industriais) têm aumentado consideravelmente nos centros urbanos, em consequência do aumento populacional. Podemos citar, por exemplo, a poluição de rios e mares. Estes problemas têm exigido das autoridades ambientais a criação de programas de monitoramento que empreguem métodos modernos, práticos, rápidos e de baixo custo usados na determinação de parâmetros de qualidade de águas, afim de que possibilitem tomadas de decisão em tempo hábil, de modo a minimizar o lançamento destes poluentes nos corpos d'águas.^[2,44]

Dentre os vários parâmetros utilizados para caracterizar a qualidade dos esgotos domésticos, pode-se mencionar a demanda química de oxigênio (DQO) como parâmetro que reflete grande parte da matéria orgânica presente nos esgotos e portanto aparece em grande parte da legislação mundial.^[45-47] Este parâmetro é medido através da oxidação de compostos orgânicos produzidos por agentes químicos. O método de referência para a determinação de DQO descrito na literatura^[41,44] apresenta sérios inconvenientes:

- (i) o procedimento padrão fundamenta-se na oxidação da matéria orgânica pelo dicromato de potássio em meio sulfúrico e elevadas temperaturas, aumentando os riscos de acidentes;
- (ii) faz uso de catalisadores como sulfato de prata que elevam o custo da análise;
- (iii) gera resíduos como cromo e mercúrio, prejudiciais ao meio ambiente;
- (iv) envolve uma etapa de digestão que dura de 2 a 3 horas implicando um tempo de análise elevado;
- (v) possui baixa reprodutibilidade levando a desvios de $\pm 5\%$.

Neste contexto é de grande importância o desenvolvimento de métodos analíticos práticos, rápidos e baratos para a determinação de parâmetros de qualidade de águas e de caracterização dos esgotos.

3.2 Espectrometria NIR

A espectrometria NIR é a espectroscopia vibracional que abrange a faixa do espectro eletromagnético aproximadamente entre 12.800 e 4.000 cm^{-1} (780 a 2.500 nm), contendo bandas de absorção de sobretons de vibrações fundamentais e de suas combinações de C-H, N-H, S-H, e/ou O-H, devido a grande anarmonicidade dessas vibrações envolvendo energia de átomos de hidrogênio. Entretanto, alguns outros efeitos associados com termos de alta ordem podem ser observados e contribuem para a complexidade dos espectros NIR, como por exemplo acoplamento ou ressonância entre diferentes vibrações de um mesmo grupo funcional.^[48-51]

Dentre as vantagens da espectrometria NIR podemos mencionar: uma técnica rápida, de natureza não destrutiva, não invasiva e de simplicidade na preparação da amostra.^[51] A maior desvantagem é provavelmente a baixa sensibilidade para maioria dos constituintes.^[51]

3.3 Espectrometria NIR de reflectância difusa (NIRR)

Uma das técnicas mais utilizadas em medidas quantitativas na espectrometria NIR, especialmente em aplicações com amostras sólidas, baseia-se em medidas da refletância difusa. Na espectrometria de reflectância difusa (NIRR), os espectros de refletância embora não sejam idênticos aos espectros de absorção correspondentes, são semelhantes na aparência e carregam a mesma informação, podendo assim ser utilizados tanto para análises qualitativas como quantitativas.^[50] A refletância difusa, é obtida a partir de uma superfície rugosa que pode apresentar-se na forma contínua ou de pó. Estritamente, não se trata de um fenômeno de superfície, pois, ele requer a ocorrência de absorção pela interação entre a amostra e a radiação infravermelha incidente. Em uma medida instrumental, o feixe de radiação eletromagnética penetra a superfície, interage com amostra e é refletido de forma difusa,^[52] como mostra a **Figura 3.1**.

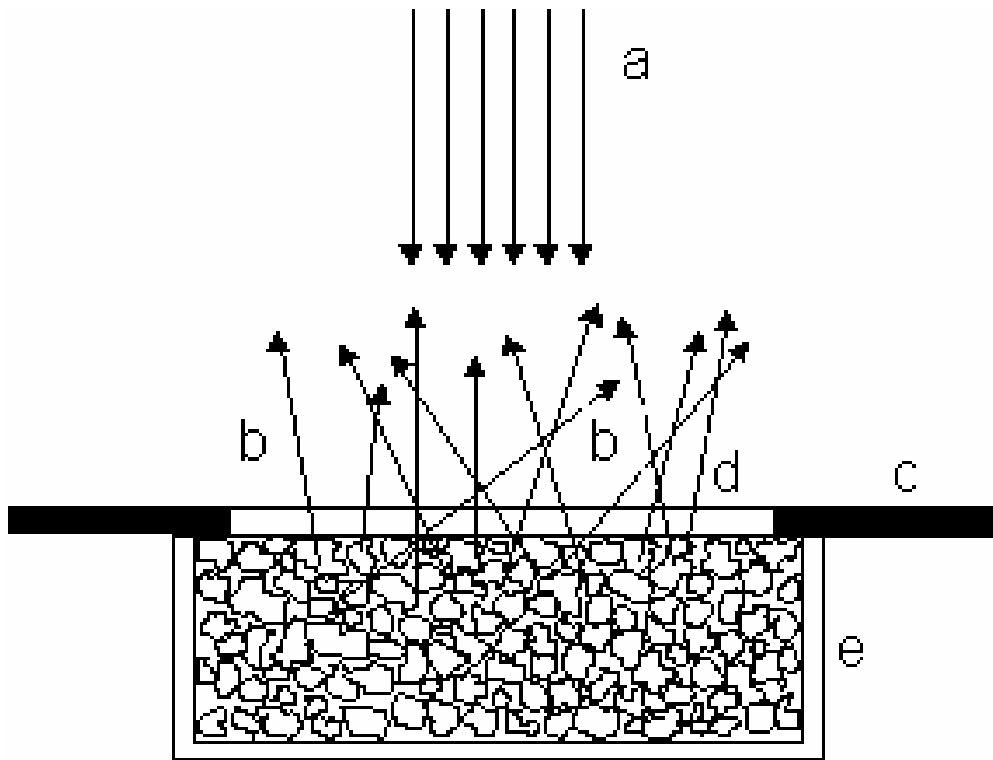


Figura 3.1 Refletância difusa resultante da incidência de um feixe de radiação eletromagnética em uma amostra constituída de partículas de forma e tamanho irregulares. Luz incidente = **a**; radiação difusa refletida = **b**; suporte da cela = **c**; janela da cela = **d**; cela da amostra = **e**.^[52]

Em 1931, Kubelka e Munk,^[52] desenvolveram uma teoria considerando que uma camada é composta de partículas absorventes uniformemente espalhadas e aleatoriamente distribuída, de dimensões muito menores que a espessura da camada. Eles consideraram somente a direção da radiação incidente e refletida perpendicular à superfície da camada, assumindo que a radiação é monocromática e que a espessura é infinita (o aumento da espessura não resulta em diferença na reflexão).

A resolução das equações diferenciais obtidas pela concepção desta teoria resulta na **Equação 3.1**, conhecida como equação de Kubelka e Munk.

$$F(R_{\infty}) = \frac{(1 - R_{\infty})^2}{2R_{\infty}} = \frac{k}{s} \quad (3.1)$$

onde, R_{∞} corresponde a reflexão da amostra sob condições em que a espessura da amostra é infinita, k é o coeficiente de absorção da amostra, que é proporcional à

concentração, e s é o coeficiente de espalhamento da amostra. O coeficiente de espalhamento determina a extensão da interação da luz incidente com a amostra antes da radiação retornar a superfície. Para uma certa extensão, s controla a profundidade na qual a luz penetra na amostra, ou seja, quanto menor o valor de s , maior é o valor $F(\mathbf{R}_\infty)$. Entretanto, este comportamento é válido somente para partículas que efetivamente espalham a radiação incidente.^[52] Pode-se notar que o termo $1-R_\infty$ inverte o espectro, de forma que o espectro resultante se apresenta com o mesmo aspecto do espectro de absorbância convencional.^[52]

As medidas de refletância difusa não requerem grande preparo de amostra, contudo algumas precauções devem ser levadas em consideração:^[50]

- i) Para tratamento de amostras sólidas na forma de pós, para melhores resultados é necessário que o tamanho das partículas seja uniforme;
- ii) O grau de compactação da amostra na cela não deve sofrer variações significativas e a superfície da amostra deve ser plana.

3.4 A transformada wavelet para minimização de ruído

A transformada wavelet (TW) é uma ferramenta de processamento de sinal^[53] que tem várias aplicações, sendo uma delas a minimização de ruído de sinais instrumentais.^[54-58] A TW de um espectro é dado por:

$$\mathbf{x} = [x(\nu_1) \ x(\nu_2) \ \dots \ x(\nu_J)] \quad (3.2)$$

onde ν_k é o k -ésimo número de ondas e, pode ser obtida usando uma estrutura de banco de filtro digital^[59] descrito na **Figura 3.2**.

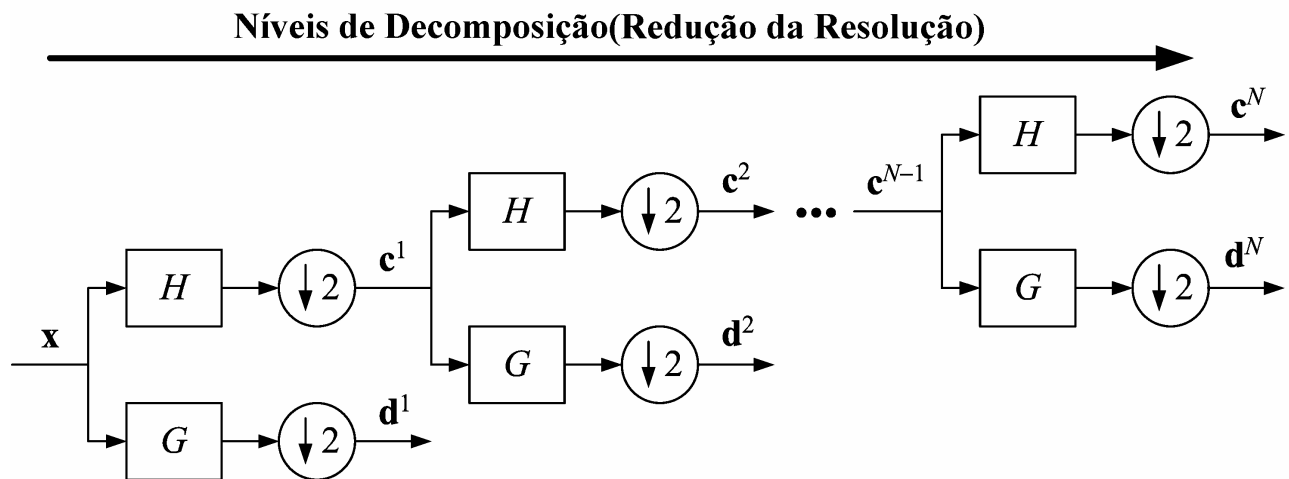


Figura 3.2 Implementação de banco de filtro de transformada wavelet com N níveis de decomposição. As seqüências c^k e d^k são denominadas, respectivamente, coeficientes de aproximação e de detalhe do k -ésimo nível de decomposição.

A estrutura básica do banco de filtro consiste em um par de filtros passa-baixa (H) e passa-alto (G), seguidos por uma operação de sub-amostragem ($\downarrow 2$) que é realizada em cada filtro. Nesta operação, um entre cada dois pontos do sinal resultante filtrado é descartado. O sinal resultante sub-amostrado pelo filtro passa baixa é uma aproximação suavizada do sinal original em uma menor resolução. A operação de sub-amostragem pode ser re-aplicada aos coeficientes de aproximação gerada no primeiro nível e este processo é realizado até um número de níveis de resolução especificado pelo analista.

Espectros NIRR normalmente apresentam características inerentes tanto na linha base (perfil de linha base), como no sinal espectral (ruídos) e estas características não estão relacionadas com as variações dos parâmetros de interesse (por exemplo, a concentração de um analito, DQO, etc). Um procedimento padrão para remover perfil de linha base consiste em usar um processo derivativo.^[59] Este processo tende, em geral, a aumentar o ruído espectral, assim um algoritmo suavizador, como o método polinomial Savitzky-Golay^[60] normalmente é aplicado, antes da derivação. Porém, o ajuste do comprimento de janela para este método pode não ser direto, porque janelas pequenas retêm muito ruído, e janelas grandes podem

introduzir distorções e perda de informação. Por isto, no presente trabalho, um procedimento Savitzky-Golay derivativo^[60] é usado com uma janela de comprimento pequeno, seguida por minimização de ruído usando a transformada de wavelet.

Neste trabalho, foi adotado o filtro wavelet symlet 8 da família symlet.^[61] Foi empregado o número de máximo de níveis de resolução para os quais os coeficientes de aproximação não exibiram uma distorção notável com respeito ao sinal original. Os coeficientes de aproximação obtidos desta maneira foram usados para modelagem PLS e previsão subsequente de valores de DQO em novas amostras. Tal procedimento é semelhante ao processo descrito por Trygg, J. S. Wold,^[62] para calibração multivariada empregando medidas de espectros NIR.

3.5 Calibração multivariada

O processo geral de calibração^[60,63] é obtido em duas etapas:

- i) modelagem, em que estabelece uma relação matemática entre a matriz X (variáveis independentes) e a matriz Y (resposta) no conjunto de calibração e da validação, que otimiza a relação no sentido de uma melhor descrição do analito(s) de interesse;
- ii) previsão, que utiliza a modelagem estabelecida na etapa anterior para estimar propriedades desconhecidas de amostras de interesse a partir do seu sinal analítico.

A **Figura 3.3** ilustra em uma forma simplificada o procedimento da calibração multivariada.

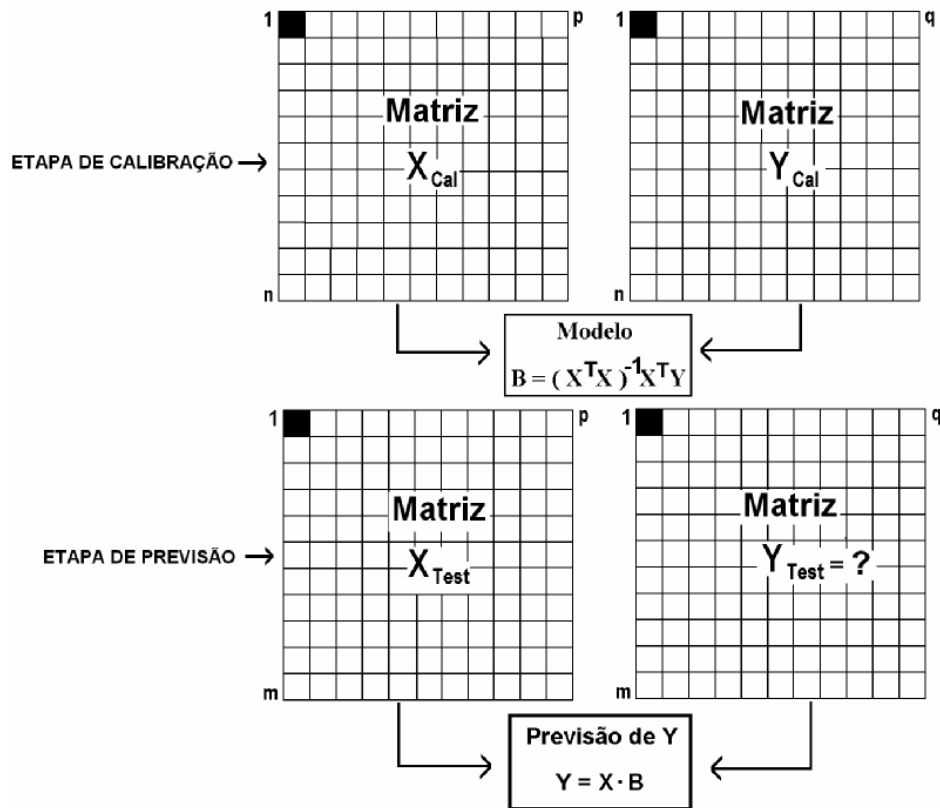


Figura 3.3 Representação ilustrativa das etapas de calibração multivariada, onde, as linhas horizontais de 1 a n e de 1 a m e as colunas de 1 a p e de 1 a q representam, respectivamente, as amostras, as variáveis espectrais e os parâmetros das matrizes X e Y das etapas de calibração e de previsão.

3.5.1 Calibração PLS

Dentre os diversos métodos de calibração multivariada a regressão por mínimos quadrados parciais é a mais usada.^[60]

A regressão PLS consiste em modelar tanto a matriz X como a matriz Y em termos de componentes principais, formando dois grupos de matrizes de escores e pesos (T, U e P, Q) e dois blocos de **Equações 3.3 e 3.4**:

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad 3.3$$

$$\mathbf{Y} = \mathbf{UQ}^T + \mathbf{F} \quad 3.4$$

onde, T e P são os escores e pesos de X, U e Q os escores e pesos de Y, e E e F representam as matrizes dos resíduos. O modelo final consiste em relacionar

linearmente os escores do bloco X (matriz **T**) com escores do bloco Y (matriz **U**) através da **Equação 3.5**.

$$\mathbf{U} = \mathbf{BT} + \mathbf{G} \quad \mathbf{3.5}$$

onde, B é a matriz dos coeficientes da regressão linear e G é a matriz dos resíduos. A previsão das amostras futuras é obtida através da **Equação 3.6**:

$$\mathbf{Y} = \mathbf{BTQ}^T + \mathbf{H} \quad \mathbf{3.6}$$

onde H é representa a matriz dos resíduos.

O uso dessa técnica com espectrometria NIR tem sido bastante utilizadas nas mais diversas aplicações.^[60,64] Entretanto, estudos vêm demonstrando a viabilidade de se usar a espectrometria NIR e métodos de análise multivariada no monitoramento de parâmetros físicos, químicos e bacteriológicos associados à qualidade de águas.^[65-69]

3.6 Revisão da Literatura

Nilsson *et al.*^[65] mostraram que espectros *NIR* de sedimento de lagos da Suécia e PLS podem ser utilizados na previsão de:fósforo total, pH e carbono orgânico total das águas. Eles demonstraram que informações sobre pH, fósforo total e carbono orgânico total (TOC) da água do lago podem ser deduzidos do sedimento através do uso de funções de transferências obtidas de modelos construídos usando espectros NIR e variáveis químicas da água em conjunto de dados de calibração multivariada.

Malley *et al.*^[66] aplicaram a técnica PLS, regressão linear múltipla (MLR) e espectrometria NIR para determinação de carbono, nitrogênio e fósforo em material particulado de 10 lagos no nordeste de Ontario no Canadá. Os resultados mostraram que excelentes modelos de calibração podem ser desenvolvidos dos espectros NIR e dos dados de referência de C, N e P para material em suspensão.

Dabakk *et. al.*^[67] estudaram a relação entre espectros NIRR do seston e os parâmetros de qualidade de água, carbono orgânico total, fósforo total, absorvância em 420 nm e pH de 271 lagos do nordeste da Suécia, empregando PLS. Eles demonstraram que os espectros NIRR do seston desses lagos podem ser usados para prever os parâmetros de qualidade de água mencionados.

Chisty *et. al.*^[68] conseguiram discriminar grupos de amostras de lagos da Noruega, de acordo com sua composição orgânica natural, usando espectroscopia NIR, PCA e PLS. Os modelos PLS demonstraram que 4 biopolímeros (carboidratos, açúcares com grupos *N*-acetil amino, proteínas e poliídros aromáticos) de amostras proveniente de matéria orgânica natural podem ser determinados com precisão aceitável. A PCA mostrou que as amostras dos lagos foram agrupadas em três classes diferentes e que estes grupos concordam com aqueles obtidos por técnicas caras e laboriosas.

Medeiros *et al.*^[69] propuseram um método de análises para screening do seston de águas superficiais do rio Mumbaba, localizado João Pessoa - Paraíba - Brasil. Eles usaram a espectrometria NIRR, uma estratégia de seleção de variáveis baseada em Transformada Wavelet e SIMCA com intuito de perceber alterações de propriedades espectrais causadas por descarga de efluente industrial.

3.7 OBJETIVO

Desenvolvimento de um novo método para determinação de DQO na ETE-Mangabeira empregando espectrometria NIRR do seston e PLS.

3.8 EXPERIMENTAL

3.8.1 Coleta e períodos de amostragem

Neste trabalho, foram utilizadas as amostras coletadas no turno da manhã (entre 7:00 e 9:00 horas) nos pontos de amostragem da ETE-Mangabeira, indicados

na **Figura 2.4**. As amostras foram coletadas em duplicatas de março a junho de 2005 e em triplicatas de agosto de 2005 a fevereiro de 2006.

3.8.2 Processamento das amostras

O seston foi coletado por filtrações de 10, 25, 50, 100, 150 e 200 ml de cada amostra numa membrana Whatman GF/C (de 5 cm de diâmetro e porosidade em torno de 1,0 μm) num sistema de filtração a vácuo. As membranas foram colocadas em placas de Petri, secas em estufa com temperatura controlada a 50°C durante 6 horas e esfriada em dessecador por 40 min até atingir a temperatura ambiente.^[67] Um círculo de 10 mm de diâmetros foi cortado aleatoriamente em cada membrana para aquisição dos espectros.

3.8.3 Método de referência para determinação de DQO

Num conjunto de 90 amostras do esgoto bruto, sendo 29 amostras do ponto AA, 31 do ponto EA, e 30 do ponto EF, foram determinados DQO segundo o método de referência da *APHA*.^[41] O método e os materiais foram apresentado em detalhes no item x da **seção 2.5.4** deste trabalho.

3.8.4 Medidas dos espectros NIR

Os espectros foram registrados usando um espectrofotômetro NIR com Transformada de Fourier da marca Bomem, modelo MB 160 D, fonte a base de quartzo-halogênio e detector a base de arseneto de gálio e índio (InGaAs) (**Figura 3.4**).

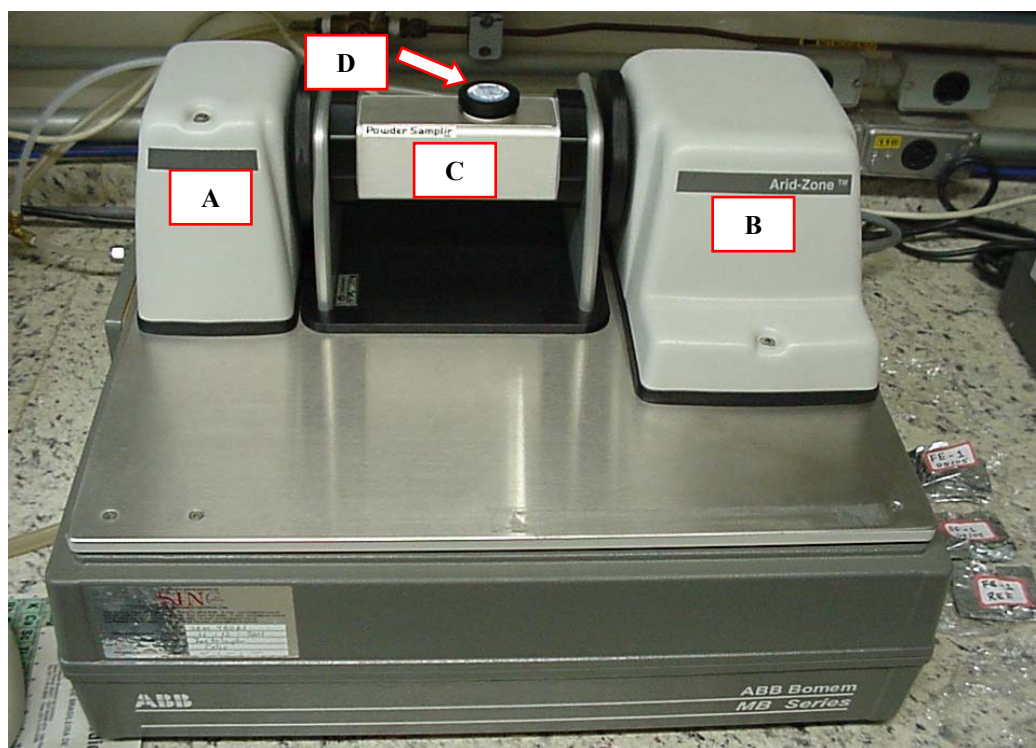


Figura 3.4 Espectrômetro NIR. **A** (fonte), **B** (detector), **C** (acessório de refletância difusa) e **(D)** recipiente de material refratário colocado sobre a membrana contendo o séston das amostras de esgoto analisadas.

Os espectros foram registrados na faixa espectral de 10.000 a 4.000 cm^{-1} . Cada espectro foi obtido como uma média de 32 varreduras, com uma resolução de 4 cm^{-1} , numa temperatura de 24°C e com cerca de 49,0% de umidade relativa do ar. O branco foi obtido registrando o espectro de uma membrana whatman GF/C, após filtração com água destilada e aplicando o mesmo procedimento acima. O instrumento apresentado na **Figura 3.4** pertence ao laboratório do Instituto de Química da Unicamp-Campinas-SP.

3.8.5 Pré-processamentos e programas computacionais

Para melhorar a capacidade preditiva dos modelos afetada pela variação sistemática de linha de base e a baixa relação sinal/ruído principalmente na região de segundo sobreton foram empregados a primeira derivada com polinômio de grau 2 após suavização por Savitzky-Golay e uma janela 25 pontos e Transformada Wavelet

usando a Symlet 8 com 4 e 5 níveis de decomposição nos espectros NIRR das lagoas anaeróbia e facultativa.

A derivação dos espectros após suavização Savitzky–Golay e a modelagem PLS foram realizadas usando o pacote quimiométrico Unscrambler, versão 9.5, da CAMO-AS. O procedimento de Wavelet para minimização de ruído foi implementado com o Toolbox Wavelet do Mathworks Matlab.

3.9. RESULTADOS E DISCUSSÃO

3.9.1 Pré-tratamento e pré-processamento dos espectros NIRR

Para a seleção da faixa de trabalho foram analisados os espectros NIRR brutos das amostras dos pontos AA, EA e EF da ETE-Mangabeira. Os espectros registrados na região de 10.000 a 4.000 cm^{-1} são mostrados na **Figura 3.5**.

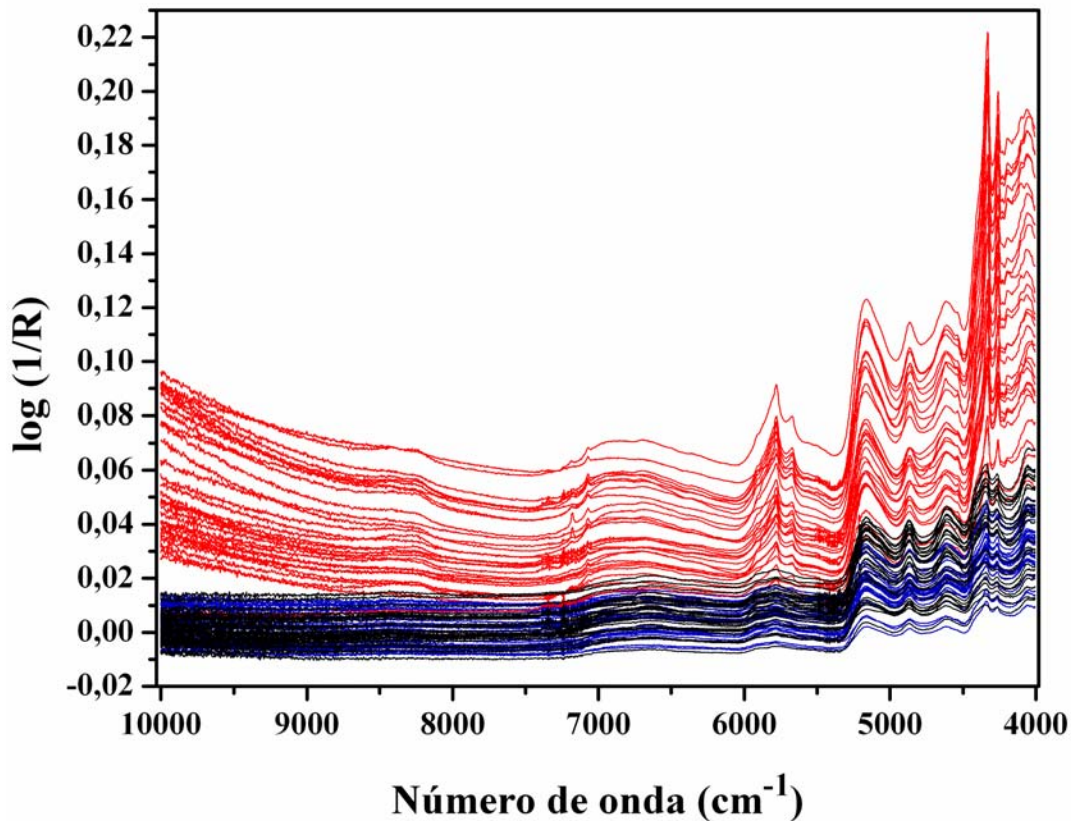


Figura 3.5 Espectro NIRR na faixa espectral de 10.000 a 4.000 cm^{-1} do séston das amostras coletadas nos pontos AA (—), EA (—) e EF (—).

Como podemos observar na **Figura 3.5**, a faixa espectral de 10.000 a 8.000 cm^{-1} é bastante ruidosa e não se percebe a existência de bandas espectrais que possam ser relacionadas à informação química. Assim, para esse trabalho foi selecionada a faixa espectral de 8.000 a 4.000 cm^{-1} por ser uma região de maior relação S/R e com bandas espectrais que podem ser relacionadas à informação química.

3.9.2 Escolha do volume de filtração das amostras

Para escolher o volume adequado de amostra a ser usado na filtração das amostras de esgotos coletados, seis diferentes volumes (10, 25, 50, 100, 150 e 200 ml) foram testados e os espectros do séston são mostrados na **Figura 3.6**.

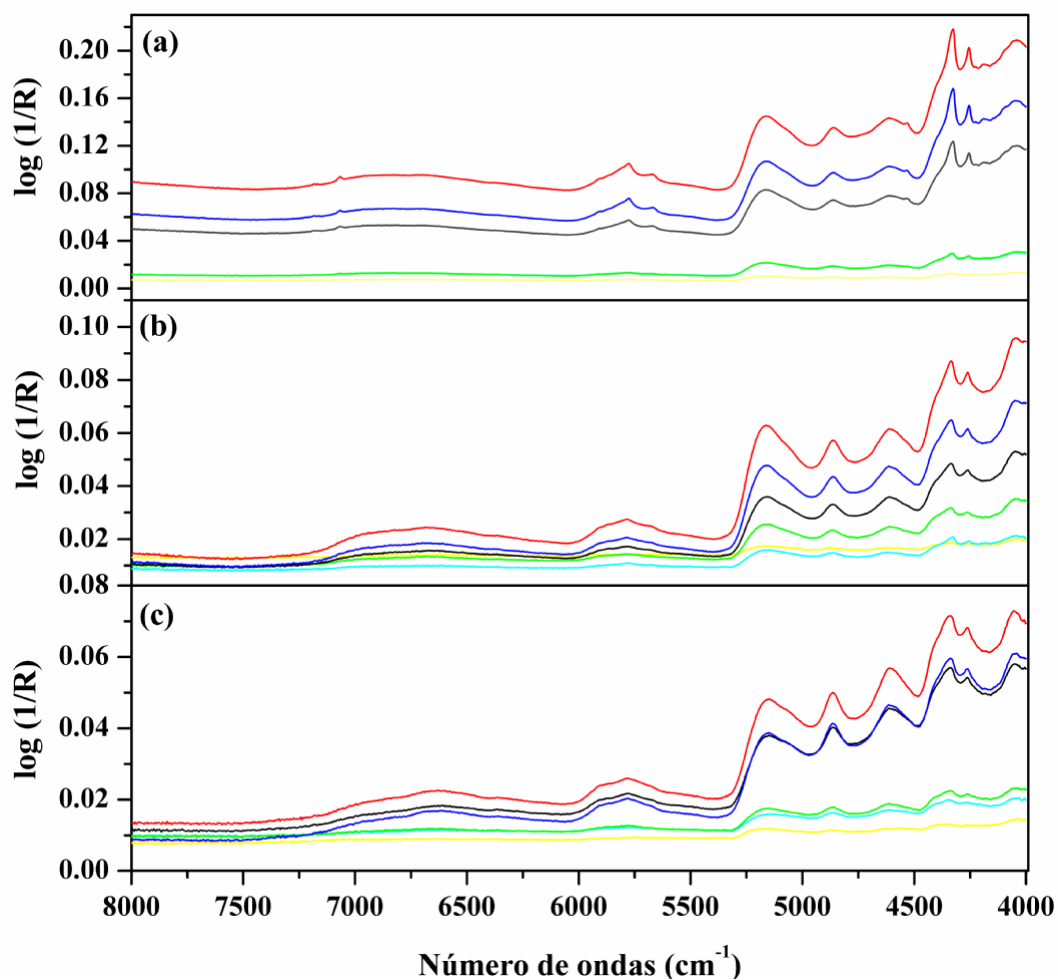


Figura 3.6 Espectros registrados para as amostras do AA (a), EA (b) e EF (c) para volumes de (—10), (—25), (—50), (—100), (—150) e (—200) ml.

Como pode ser visto na **Figura 3.6**, os espectros NIRRR do seston coletados usando volumes de amostras menores que 100 ml apresentam baixos valores de absorbâncias. A filtração de volumes de amostra maiores que 100 ml demoram muito tempo.

Como a filtração de 100 ml de amostra demora menos de 10 minutos e os espectros NIRRR do seston apresentam sinais analíticos comparáveis com aqueles obtidos quando foram usados volumes de 150 e 200 ml que demoram mais de uma hora de filtração, adotou-se, então, na filtração de todas as 90 amostras do conjunto de dados, um volume de 100 ml como sendo o melhor compromisso entre o tempo de filtração e perda de sinal analítico.

Os gráficos *Box-Whisker* do parâmetro DQO (**Figura 2.8**) mostraram uma redução expressiva dos valores deste parâmetro do ponto AA para os pontos EA e EF. Adicionalmente, de acordo com um teste-*t* emparelhado, em nível de 95% de confiança, não há diferença significativa entre os valores de DQO dos pontos EA e EF. Assim, uma vez que a maior parte da degradação de matéria orgânica afluyente ou seja, a maior redução de DQO ocorre na lagoa anaeróbia, utilizou-se no desenvolvimento de um novo método para determinação DQO, apenas as amostras coletadas no ponto EA.

Na figura abaixo são apresentados: o espectro NIRRR do material sestônico de uma amostra coletada no ponto EA na faixa espectral de 8.000 a 4.000 cm^{-1} (**Figura 3.7a**); o espectro derivativo resultante da aplicação a este espectro NIRRR de um processo de derivação Savitzky-Golay, empregando um polinômio de segunda ordem e uma janela de 25 pontos (**Figura 3.7b**); e os espectros derivativos resultantes da minimização de ruídos empregando a transformada de *wavelet* (**Figura 3.7c**) com 4 e 5 níveis de decomposição.

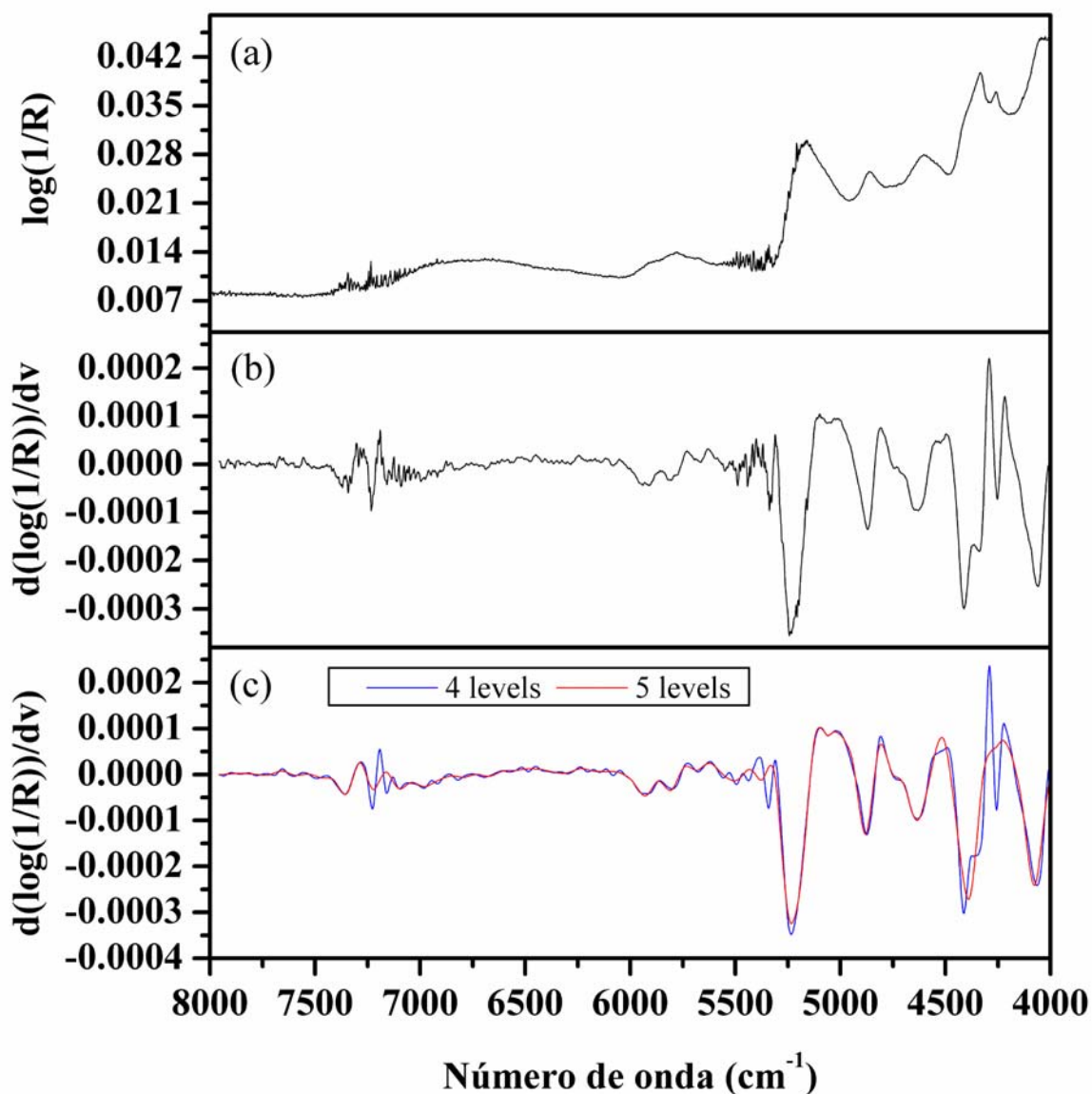


Figura 3.7 o espectro NIRR do séston de uma amostra coletada no ponto EA (a); o espectro derivativo resultante da aplicação da derivação Savitzky-Golay (b); e os espectros derivativos resultantes da minimização de ruídos com a transformada de *wavelet* (c) com 4 e 5 níveis de decomposição.

O espectro mostrado na **Figura 3.7a** exibe um perfil de linha de base característica de espectros NIRR do séston, que foi removido pela derivação Savitzky-Golay como apresentado na **Figura 3.7b**. O ruído ainda presente no espectro derivativo foi minimizado pelo emprego da transformada de *wavelet* com 4 e 5 níveis de decomposição como mostrado na **Figura 3.7c**.

Verifica-se na **Figura 3.7c** que o uso de 5 níveis de decomposição na transformada de *wavelet* introduz uma distorção no sinal espectral, especificamente nos dois picos em torno de 4250 cm^{-1} . Por esse motivo, optou-se por empregar, neste trabalho, uma minimização de ruído aplicando a transformada de *wavelet* com 4 níveis de decomposição. Este processo, resulta em 142 coeficientes de aproximação que serão usados na construção dos modelos PLS.

3.9.3 Seleção das amostras do conjunto de calibração e de previsão

O conjunto de dados é formado por 31 amostras e 142 variáveis (coeficientes de aproximação *wavelet*). Das 31 amostras usadas na análise, 21 foram escolhidas na construção do modelo PLS (conjunto de calibração) e 10 foram usadas como um conjunto independente (conjunto de previsão) que foi empregado na avaliação da capacidade de previsão do modelo PLS construído.

As amostras do conjunto de calibração e de previsão foram escolhidas de acordo com o seguinte procedimento proposto aqui neste trabalho:

- i) Inicialmente, 12 amostras com diferentes valores de DQO foram selecionadas aplicando o algoritmo de amostragem uniforme de Kennard-Stone (KS)^[70,71] ao vetor \mathbf{y} contendo os valores de DQO das 31 amostras usadas na análise;
- ii) As duas primeiras amostras deste conjunto são aquelas que contem o maior e o menor valor de DQO e, portanto, elas foram incluídas no conjunto de calibração para evitar problemas de extrapolação do modelo PLS a ser construído;
- iii) As 10 amostras remanescentes da seleção pelo algoritmo KS foram tomadas como sendo o conjunto de previsão. Uma vez que, o algoritmo KS varre uniformemente toda a faixa de valores de DQO das 31 amostras analisadas, este conjunto de previsão selecionado é representativo e assegura que a capacidade preditiva do modelo seja avaliada de uma maneira justa ao longo da faixa inteira de calibração;

iv) As 19 amostras não selecionadas pelo algoritmo KS, adicionadas das 2 amostras com valores extremos de DQO, compõem as 21 amostras do conjunto de calibração.

3.9.4 Construção dos modelos *PLS*

Os modelos PLS foram construídos usando o processo de validação cruzada^[56] aplicada às 21 amostras do conjunto de calibração. Na construção desses modelos foram utilizados a matriz dos espectros NIRR dessas amostras sem pré-processamento (espectros brutos) e com pré-processamento, empregando derivação Savitzky-Golay e a derivação Savitzky-Golay seguida da minimização de ruído pela transformada de *wavelet* com 4 níveis de decomposição, que gera uma matriz de 142 coeficientes *wavelets* de aproximação. O gráfico *scree* dos modelos PLS gerados usando os coeficientes da transformada wavelet (modelos PLS_{TW}) e os espectros brutos (modelos PLS_B) e derivados com Savitzky-Golay (modelos PLS_{SG}) são apresentados nas **Figuras 3.8, 3.9 e 3.10**.

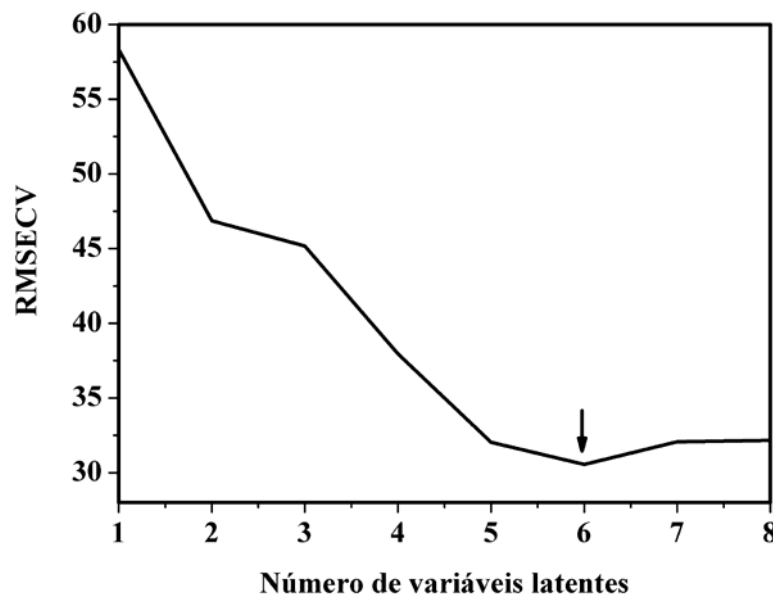


Figura 3.8. O gráfico *scree* dos modelos PLS gerados usando os coeficientes da transformada wavelet (modelos PLS_{TW}). A seta indica o ponto associado ao número de variáveis latentes adotado no modelo.

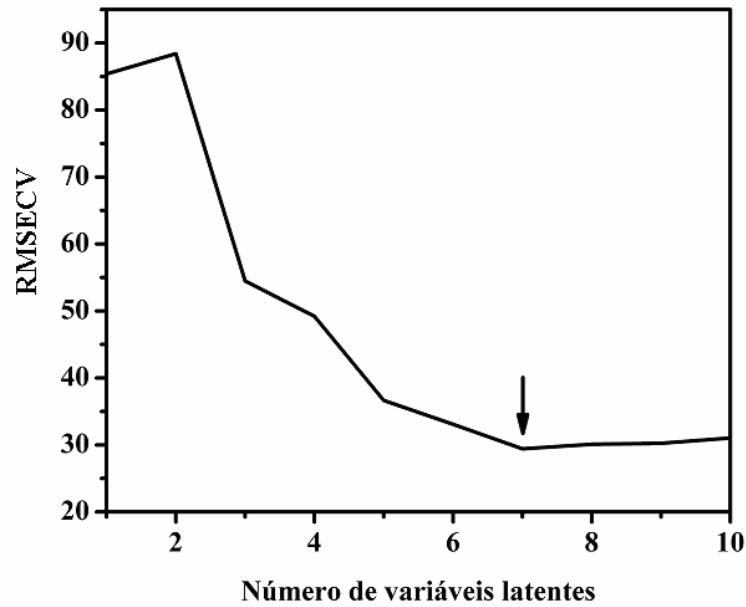


Figura 3.9. O gráfico *scree* dos modelos PLS gerados usando os espectros brutos (modelos PLS_B). A seta indica o ponto associado ao número de variáveis latentes adotado no modelo.

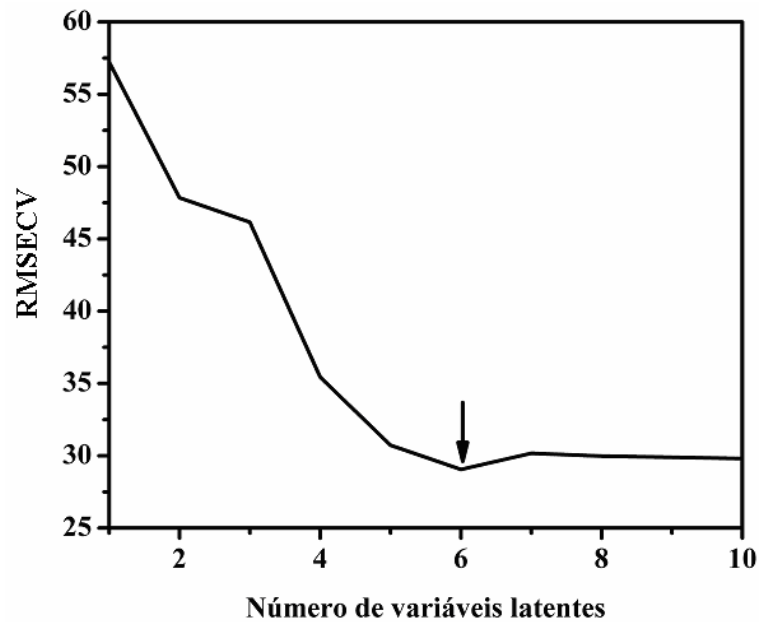


Figura 3.10. O gráfico *scree* dos modelos PLS gerados usando os espectros derivados com Savitzky-Golay (modelos PLS_{SG}). A seta indica o ponto associado ao número de variáveis latentes adotado no modelo.

Como pode ser observado nas **Figuras 3.8, 3.9 e 3.10**, modelos PLS_{TW}, PLS_B e PLS_{SG} com, respectivamente, seis, sete e seis variáveis latentes foram escolhidos para estimar os valores de DQO das amostras do conjunto de previsão,

uma vez que eles fornecem menores valores da raiz quadrada do erro médio quadrático da validação cruzada (RMSECV), definido como:

$$RMSECV = \sqrt{\frac{1}{M_c} \sum_{k=1}^{M_c} (y_{cv}^k - \hat{y}_{cv}^k)^2} \quad (3.7)$$

onde, y_{cv}^k e \hat{y}_{cv}^k são os valores de DQO obtidos pelo método de referência e pela previsão dos modelos PLS construídos na k -ésima amostra do conjunto de calibração e $M_c = 21$ é o número de amostras do conjunto de calibração.

Os resultados expressos em termos da correlação entre os valores de DQO das 10 amostras do conjunto de previsão obtidos pelo o método de referência e pelos modelos PLS_{TW}, PLS_B e PLS_{SG} escolhidos são apresentados na **Tabela 3.1**.

Tabela 3.1 Valores de correlação e RMSEP dos PLS_{TW}, PLS_B e PLS_{SG} escolhidos, com, respectivamente, seis, sete e seis variáveis latentes.

Modelos	Correlação	RMSEP (mg O ₂ L ⁻¹)
PLS _B	0,91	30
PLS _{SG}	0,97	23
PLS _{TW}	0,97	19

Também são mostrados na **Tabela 3.1** os valores da raiz quadrada do erro médio quadrático de previsão (RMSEP) das 10 amostras do conjunto de previsão, definido como:

$$RMSEP = \sqrt{\frac{1}{M_p} \sum_{k=1}^{M_p} (y_p^k - \hat{y}_p^k)^2} \quad (3.8)$$

onde, y_p^k e \hat{y}_p^k são os valores de DQO obtidos pelo método de referência e pela previsão dos modelos PLS nas amostras do conjunto de previsão e $M_p = 10$ é o número de amostras do conjunto de previsão.

Observa-se na **Tabela 3.1** que o uso do modelo de calibração PLS_{SG} na previsão das 10 amostras do conjunto de previsão apresenta um maior valor de correlação ($= 0,97$) do que o modelo PLS_B ($= 0,91$), bem como uma redução nos valores do RMSEP (de 30 para $23 \text{ mg O}_2 \text{ L}^{-1}$). Esta melhor capacidade preditiva é esperada, pois na construção do modelo PLS_{SG} foram removidas características de linha base e o ruído espectral foi minimizado, ou seja, foram eliminando fontes de variabilidade dos dados espectrais não diretamente relacionados com a propriedade DQO prevista.

A aplicação do modelo de calibração PLS_{TW} para previsão das 10 amostras do conjunto de previsão quando comparado com o modelo PLS_{SG} reduziu ainda mais o RMSEP (de 23 a $19 \text{ mg O}_2 \text{ L}^{-1}$). Assim, para este trabalho adotou-se o modelo PLS_{TW} como sendo o melhor modelo de calibração usada na previsão das 10 amostras do conjunto de previsão escolhidas pelo procedimento proposto aqui neste trabalho.

Na **Figura 3.11** é apresentada a curva dos valores previstos de DQO pelo modelo PLS_{TW} *versus* valores obtido pelo método de referência da APHA, bem como a barra de erro indicando o desvio padrão obtido na previsão de cada uma das 10 amostras do conjunto de previsão.

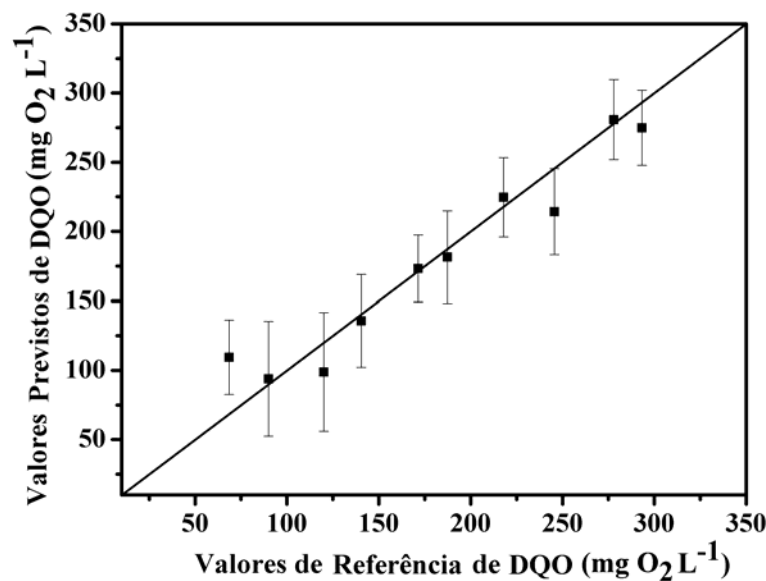


Figura 3.11 Valores de DQO previstos pelo modelo PLS_{TW} *versus* valores referência. A barra de erro indica o desvio padrão obtido na previsão de cada uma das 10 amostras do conjunto de previsão.

Como pode ser visto na **Figura 3.11**, os pontos estão distribuídos aleatoriamente em torno da linha da bissetriz, indicando ausências de erros sistemáticos. O valor do RMSEP e a correlação obtida (**Tabela 3.1**) para estas amostras foram respectivamente $19 \text{ mg O}_2 \text{ L}^{-1}$ e 0,97. Tais resultados podem ser aceitos para esta aplicação quando comparado com a estimativa conjunta do erro padrão do método de referência calculado para DQO que apresentou valor de $21 \text{ mg O}_2 \text{ L}^{-1}$.

A fim de fornecer uma evidência adicional para validação do método proposto, um modelo PLS_{TW} foi construído usando técnica de validação cruzada com todas as 31 amostras. Na **Figura 3.12** é apresentada a curva dos valores previstos de DQO por esse modelo PLS_{TW} *versus* valores obtido pelo método de referência da APHA, bem como a barra de erro indicando o desvio padrão obtido na previsão de cada uma das 31 amostras de todo o conjunto.

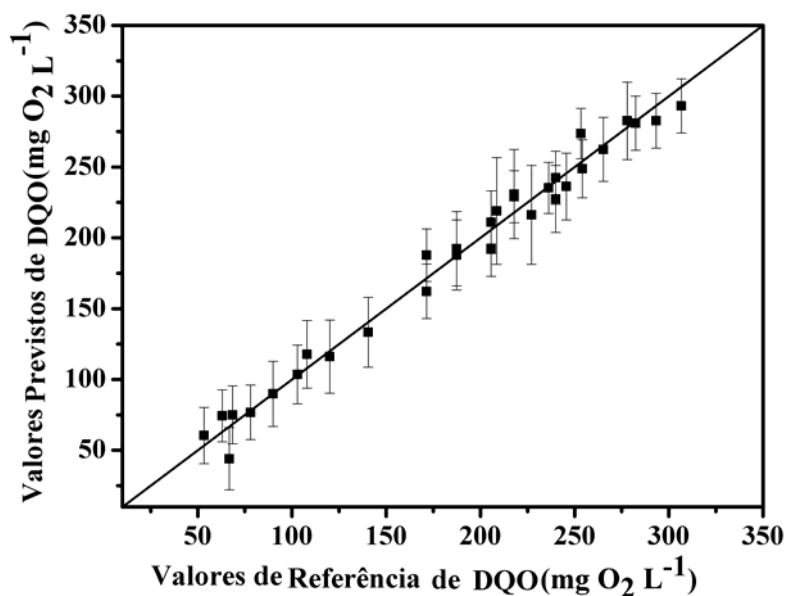


Figura 3.12 Valores previstos de DQO *versus* referência para as 31 amostras da validação cruzada para cada amostra. A barra de erro indica o desvio padrão para cada amostra de previsão.

Verifica-se na **Figura 3.12** que os pontos estão distribuídos aleatoriamente em torno da linha da bissetriz, indicando ausências de erros sistemáticos. Neste caso, o valor da correlação obtida e o RMSECV foram, respectivamente, iguais a

0,96 e 22 mg O₂ L⁻¹. Novamente, estes resultados são aceitáveis quando comparado com a estimativa conjunta do erro padrão do método de referência calculado para DQO que apresentou valor de 21 mg O₂ L⁻¹.

CAPÍTULO 4

Conclusões e propostas

4.1 CONCLUSÕES

A partir da investigação aqui realizada com o intuito de avaliar a ETE-Mangabeira, que inicialmente foi projetada para operar como sendo um sistema de tratamento de esgoto doméstico do tipo lagoa aerada seguida de lagoa de maturação, e que foi adaptada para um sistema de lagoa anaeróbia seguida de lagoa facultativa (sistema do tipo australiano) pode-se concluir que:

- i) Os três pontos de coletas escolhidos na investigação são de fato de qualidade diferentes, reforçando a idéia de que diferentes processos estão ocorrendo, ou seja, a matriz muda na direção de interesse de um bom tratamento;
- ii) Os resultados apontam para maior concentração de OD e valores mais altos de pH no período de seca. Isto se deve as reações fotossintéticas (com produção de oxigênio) que são mais favorecidas neste período. Da mesma forma, também foi observado um aumento de ST neste período. Este comportamento estar associado ao aumento de sólidos em suspensão, devido ao crescimento mais intenso das algas. Percebe-se também um aumento nas concentrações de fosfato e sulfato neste período, quando comparado com o período de chuva. Este comportamento que ocorre na ETE-Mangabeira pode estar relacionado respectivamente por decomposição de compostos orgânicos de fósforo à fosfato e oxidação do H_2S , mais intenso neste período.
- iii) Por outro lado, observa-se uma redução na concentração de amônia que deve estar associada uma maior conversão da mesma em nitrato por processos de nitrificação favorecida pelo aumento da atividade fotossintética em período seco.
- iv) Todos os processos (decomposição anaeróbia de matéria orgânica e fotossínteses) estudados e as lagoas onde eles ocorrem, concordam com a dinâmica esperada para uma ETE do tipo sistema australiano;

- v) Finalmente, pode-se inferir que a ETE estudada está funcionando adequadamente, apesar de operar em sistema distinto daquele inicialmente projetado.

Com relação ao novo método desenvolvido para determinação de DQO em esgoto doméstico pode-se concluir que:

- i) O séston retido na membrana filtrante guarda informação química relativa à matéria orgânica carbonácea do esgoto doméstico, representada em termos de DQO, capaz de ser capturado pelo espectro NIRR;
- ii) O novo método pode ser considerado uma alternativa viável para o monitoramento de DQO em esgoto doméstico, uma vez que se trata de um método relativamente rápido, que não requer manipulação analítica, ou seja, precisa de apenas filtração;
- iii) Diferente do método de referência da *APHA*, que é muito utilizado no monitoramento rotineiro de águas, o método aqui proposto não necessita do uso de reagentes, catalisadores ou solventes, tais como H_2SO_4 , $K_2Cr_2O_7$, $HgSO_4$, entre outros, que são perigosos ao ser humano e ao meio ambiente;
- iv) Além do mais, foi proposto, durante o desenvolvimento deste método um novo procedimento para a escolha do conjunto de calibração e de previsão baseado no algoritmo de amostragem uniforme de Kennard-Stone (KS)^[70-71] que demonstrou ser bastante eficiente para a construção e validação de modelos PLS. Vale salientar, que a tarefa de escolher um conjunto de calibração e de previsão de um *pool* de amostras é de grande valor para a construção de modelos de calibração multivariada, principalmente quando as aplicações analíticas envolvem matrizes complexas cuja variabilidade da composição das amostras reais não pode ser facilmente reproduzida através do uso de técnicas de planejamento otimizado, tais como planejamento fatorial, planejamento composto central, planejamento de Brereton, entre outros.

4.2 PROPOSTAS

Vislumbram-se, para continuidade dos temas que foram investigados neste trabalho, as seguintes propostas:

- i) Implementação de um sistema de monitoramento da variável DQO no sistema estudado usando o novo método para determinação de DQO aqui desenvolvido;
- ii) Fazer, periodicamente, uma avaliação da capacidade preditiva do modelo PLS_{TW} construído analisando novas amostras, de modo a verificar sua robustez com o tempo e, se necessário, fazer ajustes no modelo proposto por introduzir novas amostras ao modelo de calibração;
- iii) Investigar a implementação de modelos PLS_{TW} de modo que ele possa também prever outros parâmetros físicos, químicos e biológicos e possa investigar outros processos que ocorrem em outras lagoas da ETE-Mangabeira ou em outras ETEs, bem como no estudo de sistemas aquáticos naturais (eutrofizados);
- iv) Realizar um estudo de análise exploratória comparativa com outro sistema de tratamento de esgoto doméstico similar adaptado ou não adaptado;

Referências Bibliográficas

1. Manahan, S. E. **Environmental chemistry**. 2. ed. New York: Lewis Publishers, 1994.
2. Medeiros, V. M. **O uso de quimiometria em três estudos de dados ambientais multivariado envolvendo a temática qualidade de água**, João Pessoa, Programa de Pós-Graduação em Química, UFPB, 2004, Tese de Doutorado, 87p.
3. Ministério do meio ambiente, Resolução **CONAMA**, n. 357, 2005.
4. Oliveira, S. M. A. C., Sperling, M. V.; **Avaliação de 166 ETEs em operação no País, compreendendo diversas tecnologias. Parte I - Análise de desempenho**. *Engenharia Sanitária e Ambiental*, **10: 347, 2005**.
5. Dean, R. B.; Forsythe, S. L., **Estimating the reliability of advanced waste treatment. Part 1 and 2**. *Water & Sewage Works*, **1976a and 1976b**.
6. Niku, S.; Schroeder, E. D.; Samaniego F. J., **Performance of activated sludge process: reliability, stability, and variability**. *Environmental Protection Agency, EPA. Grant Number R805097-01, p.1, 1981*.
7. Berthouex, P. M.; Hunter, W. G., **How to construct reference distributions to evaluate treatment plant effluent quality**. *JWPCF*, **55: 1417, 1983**.
8. Smith, E. P.; Keying, Ye.; Hughes, C.; Shabman, L., **Statistical assessment of violations of water quality standards under Section 303(d) of the Clean Water Act**. *Environ. Sci. Technol.*, **35: 606, 2001**.
9. Andrade Neto, C. O., **Sistemas simples para tratamento de esgotos sanitários. Experiência Brasileira**. *ABES*, p. 302, 1997.
10. Sperling, M. V.; Chernicharo, C. A. I., **Urban wastewater treatment technologies and the implementation of discharge standards in developing countries**, *Urban Water*, **4: 105, 2002**.
11. Pesquisa Nacional de Saneamento Básico. Instituto Brasileiro de Geografia e Estatística - IBGE, 2000. Disponível em:

<<http://www.ibge.gov.br/home/estatistica/populacao/condicaodevida/pnsb/pnsb.pdf>>. Acessado em: 25 de maio de 2007.

12. Campos, J. R. (coordenador). **Tratamento de esgotos sanitários por processo anaeróbio e disposição controlada no solo**. PROSAB. Rio de Janeiro: ABES, 1999.
13. Sperling, M. V. **Introdução a qualidade de águas e ao tratamento de esgoto**. 2. ed. Belo Horizonte, 1996. V. 1.
14. Sperling, M. V. **Princípios básicos do tratamento de esgotos**. 2. ed. Belo Horizonte, 1996. V. 2.
15. Ceballos, B. S. O. **Utilização de indicadores microbiológicos na tipologia de ecossistemas aquáticos do trópico semi-árido**, São Paulo, Instituto de Ciências Biomédicas da Universidade de São Paulo, 1995, Tese de Doutorado, 192p.
16. Esteves, F. A. **Fundamentos de limnologia**, Rio de Janeiro: Interciência, 1988.
17. George, T.; Franklin L. B., (Revised), Metcalf and Eddy, INC. **Wastewater engineering: treatment, disposal and reuse**. 3. ed. New York: McGraw-Hill, third edition, 1991.
18. Andrade, J. M.; Prado, D.; Alonso, E.; López, P., Muniategui, S.; Fuente, P.; Quijano, M. A., **Selection of analytical variables to optimize laboratory efforts in future groundwater studies**. *Analytical Chimica Acta*, **292: 253, 1994**.
19. Marengo, E.; Gennaro, M. C.; Giacosa, D.; Abrigo, C.; Saini, G.; Avignone, M. T., **How chemometrics can helpfully assist in evaluating environmental data lagoon water**. *Analytical Chimica Acta*, **317: 53, 1995**.
20. Vega, M.; Pardo, R.; Barrado, E.; Debán, L., **Assessment of seasonal and pollution effects on the quality of river water by exploratory data analysis**. *Water Res.*, **32: 3581, 1998**.

21. Einax, J. W.; Truckenbrodt, D.; Kampe, O., **River pollution data interpreted by means of chemometrics methods.** *Microchemical journal*, **58: 315, 1998.**
22. Abollino, O.; Aceto, M.; Gioia, C. L.; Sarzanini, C.; Mentasti, E., **Spatial and seasonal variations of major, minor and trace elements in Antarctic seawater. Chemometrics investigation of variable and site correlations.** *Advances in environmental research*, **6: 29, 2001.**
23. Ana, V. A.; Costoya, M.; Pena, R. M.; Garcia, S.; Herrero, C., **A rainwater quality monitoring network: a preliminary study of the composition of rainwater in Galicia (NW Spain).** *Chemosphere*, **51: 375, 2003.**
24. Brengaine, K.; Marhaba, T. F., **Using principal component analysis to monitor spatial and temporal changes in water quality.** *Journal of Hazardous Materials*, **100: 179, 2003.**
25. Mendiguchía, C.; Moreno, C.; Galindo-Riano, M. D.; Garcia-Vargas, M., **Using chemometrics tools to assess anthropogenic effects in river water. A case study: Guadalquivir River (Spain).** *Analytica Chimica Acta*, **515: 143, 2004.**
26. Parinet, B.; Lhote, A.; Legube, B., **Principal component analysis: an appropriate tool for water quality evaluation and management – application to a tropical lake system.** *Ecological Modelling*, **178: 295, 2004.**
27. Miettinen, T.; Hurse, T. J.; Connor, M. A.; Reinikainen, S. P.; Minkkinen, P., **Multivariate monitoring of a biological wastewater treatment process: a case study at Melbourne Water's Western Treatment Plant.** *Chemometr. Intell. Lab. Systems*, **73: 131, 2004.**
28. Singh, K. P.; Malik, A.; Mohan, D.; Sinha, S., **Multivariate statistical techniques for the evaluation of spatial and temporal variations in water quality of Gomti River (India)-a case study.** *Water Res.*, **38: 3980, 2004.**

29. Tomasz, K.; Zbytniewski, R.; Szpejna, J.; Buszewski, B., **Application of chemometrics in river water classification**, *Water Res.*, **40**: 744, 2006.
30. Shrestha S.; Kazama, F., **Assessment of surface water quality using multivariate statistical techniques: a case study of the Fuji river basin, Japan**. *Environmental Modeling e Software*, **22**: 464, 2007.
31. Silva, S. R.; Mendonça, A. S. F., **Correlação entre DBO e DQO em esgotos domésticos para a região da Grande Vitória – ES**. *Engenharia Sanitária e Ambiental*, **4**: 213, 2003.
32. Sperling, M. V. **Lagoa de estabilização**, 2. ed. Belo Horizonte, 1996 V. 3.
33. Jordão, E. P.; Pessoa, C. A. **Tratamento de esgotos domésticos**. 4. ed. Rio de Janeiro: ABES, 1995.
34. Einax, J. W.; Zwanziger, H. W. **Chemometrics in environmental analysis**. Weinheim (Germany): VCH A Wiley company, 1997
35. Hair, J. F.; Tathan, R. L.; Anderson, R. E. **Multivariate data analysis**. 5. ed. Porto Alegre: Bookman, 2005.
36. Massart, D. L.; Vandeginste, B. G. M.; Buyens, L. M. C.; Jong, S. De Lewi, P. J.; Smeyers-Verbeke, J. **Handbook of Chemometrics and Qualimetrics: Part A**. Amsterdam: ELSEVIER, 1997.
37. Wishart, D.; Clustan 1C., **User manual**. Edinburg: edinburg University Program Library Unit., 1978.
38. Rencher, A. C. **Methods of multivariate analysis**. 2. ed. Canadá: John Wiley interscience, 2002.
39. Projeto de construção e operação, Estação de Tratamento de Esgoto da Companhia de Água e Esgoto da Paraíba -CAGEPA-PB, GE-169, João Pessoa, 1981.
40. Normas técnicas da CETESB, **Manual de operação e manutenção de lagoas anaeróbias e facultativas**. L-1-009, São Paulo, 1986.
41. APHA. **Standard Methods for Examination of Water and Wastewater**. 20. ed. Washington: American Public Health Association, 1998.

42. Guimarães, J. R., Nour, E. A. A. **Tratando nossos esgotos: processos que imitam a natureza.** *Química nova na escola*, edição especial, p. 21, maio 2001.
43. Arceivala, S. J. **Wastewater treatment and disposal.** New York: Marcel Dekker, 1981.
44. Zamora, P. P.; Cordeiro, A. G.; Nagata, N., **Utilização de regressão multivariada para avaliação espectrofotométrica da demanda química de oxigênio em amostras de relevância ambiental.** *Química Nova*, **28: 838, 2005.**
45. Dulekgurgen, E.; Dogruel, S.; Karahan, O.; Orhon, D., **Size distribution of wastewater COD fractions as an index for biodegradability.** *Water Res.* **40: 273, 2006.**
46. Aiuk, S.; Amoako, J.; Raskinb, L.; Haandelc, A. V.; Verstraetea. W., **Removal of carbon and nutrients from domestic wastewater using a low investment, integrated treatment concept.** *Water Res.* **38: 3031, 2004.**
47. Hu, Z.; Chandran, K.; Smets, B. F.; Domenico, G., **Evaluation of a rapid physical chemical method for the determination of extant soluble COD.** *Water Res.* **36: 617, 2002.**
48. Bokobza, L., **Near infrared spectroscopy.** *Journal. Near Infrared Spectroscopy*, **6: 3, 1998.**
49. Pasquini, C., **Near infrared spectroscopy: fundamentals, practical aspects and analytical applications.** *J. Braz. Chem. Soc.*, **14: 198, 2003.**
50. Williams, P.; Norris, K. **Near-infrared tecnologia in the agricultura and food industriais.** 2. ed. Saint Paul (USA): American Association of Cereal Chemists, INC., 2001.
51. Workman, J. J. Jr., **Interpretive spectroscopy for near infrared appl.** *Spect. Rev.*, **31: 251, 1996.**
52. Braga, J. W. B. Avaliação de figuras de mérito em calibração multivariada, aplicada na determinação de carbamazepina por espectroscopia no

- infravermelho próximo e médio, Campinas, Programa de pós-graduação em química, UNICAMP, 2004, dissertação de mestrado.
53. Walczak, B. (Ed.). **Wavelets in chemistry**. New York: Elsevier Science, 2000.
54. Galvão, R. K. H.; Dantas Filho, H. A.; Martins, M. N.; Araújo, M. C. U.; Pasquini, C., **Sub-optimal wavelet denoising of coaveraged spectra employing statistics from individual scans**. *Analytical Chimica Acta*, **581: 159, 2007**.
55. Barclay, V. J.; Bonner, R. F.; Hamilton, I. P., **Application of wavelet transforms to experimental spectra: smoothing, denoising, and data set compression**. *Anal. Chem.*, **69: 78, 1997**.
56. Cai, C.; Harrington, P. B., **Different discrete wavelet transforms applied to denoising analytical data**. *J. Chem. Inf. Comput. Sci.*, **38: 1161, 1998**.
57. Pasti, L.; Walczak, B.; Massart, D. L.; Reschiglian, P., **Optimization of signal denoising in discrete wavelet transform**. *Chemometr. Intell. Lab. Systems*, **48: 21, 1999**.
58. Depczynski, U.; Jetter, K.; Molt, K.; Niemoller, A., **The fast wavelet transform on compact intervals as a tool in chemometrics: II. Boundary effects, denoising and compression**. *Chemometr. Intell. Lab. Systems*, **49: 151, 1999**.
59. Vetterli, M.; Kovacevic, J. **Wavelets and subband coding**, New Jersey: prentice-Hall, 1995.
60. Beebe, K. R.; Pell, R. J. Seasholtz, B. **Chemometrics: A Practical Guide**. New York: Wiley, 1998. V. 348.
61. Alsberg, B. K.; Woodward, A. M.; Winson, M. K.; Rowland, J. J.; Kell, D. B., **Variable selection in wavelet regression models**. *Analytical Chimica Acta*, **368: 29, 1998**.
62. Trygg, J.; Wold, S., **PLS regression on wavelets compressed NIR spectra**. *Chemometr. Intell. Lab. Systems*, **42: 209, 1998**.

- 63.Ferreira, M. M. C.; Antunes, A. M; Melgo, M. S; e Volpe, P. L. O., **Quimiometria I: calibração multivariada, um tutorial**, *Química Nova*, **22: 724, 1999.**
- 64.Bruns, D. A.; Ciurczak, E. W. **Handbook of near-infrared analysis NIR.** 2. ed. New York: Marcel Dekker, 2001. V. 27.
- 65.Nilsson M. B.; Dabakk, E.; Korsman T.; Renberg, I., **Quantifying relationships between near-infrared reflectance spectra of lake sediments and water chemistry.** *Environ. Sci. Technol.*, **30: 2586, 1996.**
- 66.Malley D. F.; Williams P. C.; Stainton, M. P., **Rapid measurement of suspended C, N, and P from precambrian shield lakes using near-infrared reflectance spectroscopy.** *Water Res.*, **30: 1325, 1996.**
- 67.Dabakk, E.; Nilson, M.; Geladi, P.; Wold, S.; Renberg, I., **Inferring lake water chemistry from filtered seston using NIR spectrometry.** *Water Res.*, **34: 1666, 1999.**
- 68.Chisty, A. A.; Egeberg, P. K., **Characterisation of natural organic matter from the Nordic typing project water samples by chemometric analysis of their near infrared spectral profiles.** *Chemometr. Intell. Lab. Systems*, **50: 225, 2000.**
- 69.Medeiros, V. M.; Araújo, M. C. U.; Galvão, R. K. H.; Silva, E. C.; Saldanha, T. C. B.; Toscano, I. A. S.; Oliveira, M. S. R.; Freitas, S. K. B.; Mariano Neto, M., **Screening analysis of river seston downstream of an effluent discharge point using near infrared-reflectance spectrometry and wavelet-based spectral region selection.** *Water Res.*, **39: 3089, 2005.**
- 70.Kennard, R. W.; Stone, L. A., **Technometrics.** **11: 137, 1969.**
- 71.Despaigne, F.; Massart, D. L., **Development of Robust calibration medel for nonlinear in-line process data.** *Anal. Chem.*, **72: 1657, 2000.**

Anexos



A method for determination of COD in a domestic wastewater treatment plant by using near-infrared reflectance spectrometry of seston

Antonio C. Sousa^{a,b}, Maria Mônica L.M. Lucio^a, Ovídeo F. Bezerra Neto^b,
Glauciene P.S. Marcone^a, Alessandra F.C. Pereira^a, Edilene O. Dantas^a,
Wallace D. Fragoso^a, Mario Cesar U. Araujo^{a,*}, Roberto K.H. Galvão^c

^a Departamento de Química, Universidade Federal da Paraíba, João Pessoa, PB, Brazil

^b Centro Federal de Educação Tecnológica da Paraíba, João Pessoa, PB, Brazil

^c Divisão de Engenharia Eletrônica, Instituto Tecnológico de Aeronáutica, São José dos Campos, SP, Brazil

Received 23 October 2006; received in revised form 7 February 2007; accepted 13 February 2007

Available online 20 February 2007

Abstract

This paper proposes a method for determination of chemical oxygen demand (COD) in domestic wastewater. The proposed method is based on near-infrared reflectance (NIRR) measurements of seston collected from wastewater samples by filtration. The analysis does not require any special reagent, catalyst or solvent. Inherent baseline and noise features present in NIRR spectra are removed by a Savitzky–Golay derivative procedure followed by wavelet denoising. The resulting wavelet approximation coefficients are used for partial-least-squares modelling and subsequent prediction of COD values in new samples. The model is calibrated by using COD values obtained according to the American Public Health Association (APHA) reference method. The proposed method is applied to effluent samples from the anaerobic ponds of the Mangabeira municipal wastewater treatment plant in the city of João Pessoa (Paraíba, Brazil). By comparing the NIRR prediction results with the APHA reference values, a root-mean-square error of prediction (RMSEP) of 19 mg O₂ L⁻¹ and a correlation of 0.97 were obtained. Such results are deemed adequate in view of the joint estimate of the standard error of the reference method, which was calculated as 21 mg O₂ L⁻¹.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Chemical oxygen demand; Domestic wastewater analysis; Seston; Near-infrared reflectance spectrometry; Partial-least-squares; Wavelet transform

1. Introduction

In the recent past, environmental problems related to human activities (both domestic and industrial) have increased considerably in urban centres as a consequence of population growth. In particular, pollution of water resources has become a major concern for environmental monitoring agencies, which need to take appropriate actions in a timely manner in order to mitigate the impact of effluent discharges in reservoirs [1,2]. In this context, the development of practical, fast and low-cost analytical methods for the determination of water quality parameters is of great value. For this purpose, near-infrared

reflectance (NIRR) spectrometry [3–5] in conjunction with multivariate analysis methods [6] has been shown to be a useful tool for monitoring physical, chemical, and bacteriological parameters associated to water quality. Such a technique can be used without the need for any special reagents or solvents.

Koraman et al. [7] employed NIRR spectral measurements to infer lake-water pH from sediment cores of lakes in southern Sweden by using Partial Least Squares (PLS) regression. Nilsson et al. [8] showed that NIRR spectra of superficial sediments from lakes in Sweden can be used to determine total phosphorus, pH and total organic carbon in the water by PLS modelling. Malley et al. [9] employed NIRR spectrometry and Multiple Linear Regression (MLR) for the determination of C, N and P in particulate matter of lake water. Dabakk et al. [10] demonstrated that NIRR spectra of sestonic matter from 271 lakes in north-eastern Sweden contain information regarding water quality. By using spectral pre-treatment procedures and PLS models, the authors determined total organic carbon, total

* Corresponding author at: Universidade Federal da Paraíba – Departamento de Química, Laboratório de Automação e Instrumentação em Química Analítica/Quimiometria (LAQA), Caixa Postal 5093, CEP 58051-970 - João Pessoa, PB, Brazil. Tel.: +55 83 3216 7438; fax: +55 83 3216 7437.

E-mail address: laqa@quimica.ufpb.br (M.C.U. Araujo).

phosphorus, absorption at 420 nm, and pH in good agreement with the American Public Health Association (APHA) methods. Chisty and Egeberg [11] employed NIRR spectrometry and PLS regression to determine the biopolymer input (carbohydrates, amino sugars, proteins and polyaromatic hydrocarbons) of nine different water sources in southern Norway. Moreover, the NIRR spectra were used to group the water samples in three classes, which were found to be in agreement with the results of earlier classification studies based on other expensive and time-consuming analytical techniques. Medeiros et al. [1] proposed a method for screening analysis of river seston in order to monitor changes associated to industrial effluent discharges. For this purpose, a variable selection process in the wavelet domain and Soft Independent Modelling of Class Analogies (SIMCA) were employed. The methodology was applied to samples from river Mumbaba in northeast Brazil.

Among the parameters that are currently in use for assessing the quality of domestic wastewaters, an important index is the chemical oxygen demand (COD) [12–14], which reflects a large part of the organic matter in the water. However, according to [2], the APHA reference method for determination of COD [15] has some inconveniences: (i) strong experimental conditions involved in the oxidation of organic matter by potassium dichromate in sulphuric acid at high temperatures; (ii) use of catalysts such as silver sulphate, which increases the cost of analysis; (iii) generation of residues such as chromium, silver and mercury, which are harmful to the environment; and (iv) low analytical speed, owing to the use of a digestion stage that lasts from 2 to 3 h.

The present work proposes a method for COD determination in domestic wastewater samples on the basis of NIRR spectra of sestonic matter. In order to reduce the noise level of the reflectance spectra, a wavelet denoising procedure is employed. The wavelet approximation coefficients of the spectra are then used for PLS modelling. The proposed method is applied to COD determination of samples from the Mangabeira municipal wastewater treatment plant in the city of João Pessoa (Paraíba, Brazil). The COD values predicted by the NIRR method are compared with those obtained according to the APHA standard method.

2. Experimental

The wastewater treatment plant considered in this work comprises two parallel units. Each unit consists of an anaerobic pond followed by a facultative pond. The anaerobic pond is further divided in two modules, which will be henceforth termed anaerobic pond #1 and anaerobic pond #2. Three sampling sites of the second unit were considered for investigation: affluent of anaerobic pond #1 (site 1), effluent of anaerobic pond # 2 (site 2) and effluent of the facultative ponds (site 3).

2.1. Sample collection

The samples were collected from March 2005 to February 2006 on a monthly basis in the morning period (7:00–9:00 am)

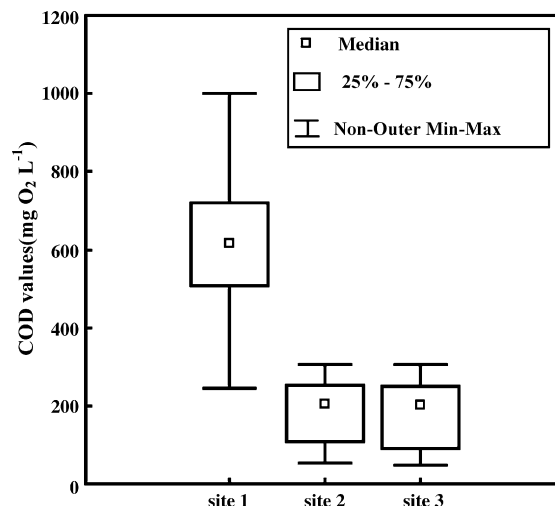


Fig. 1. Box-and-whisker plots for COD values ($\text{mg O}_2 \text{L}^{-1}$) at the three sites of the wastewater treatment plant.

and subsequently stored in 5 L plastic flasks. A total of 90 samples were obtained from sites 1 (29 samples), 2 (31 samples), and 3 (30 samples).

2.2. COD determination

The reference COD value for each sample was determined according to the method prescribed by APHA [15]. Fig. 1 presents box-and-whisker plots of the resulting COD values for the three sites. An expressive COD reduction is observed from site 1 to site 2. On the other hand, there is no significant difference between COD levels from site 2 to site 3, according to a paired *t*-test at 95% confidence level. It can be argued that the main depuration processes responsible for the COD reduction occur in the anaerobic ponds between sites 1 and 2. Therefore, the focus of the investigation was placed on the determination of COD for the 31 samples from site 2. These samples were divided in calibration and prediction sets as described in Section 3.3. Samples from site 3 could be used to increase the number of calibration objects, since the COD values in sites 2 and 3 are similar. However, samples from site 3 correspond to a distinct chemical matrix because of the biological processes (especially photosynthesis) that take place in the facultative pond between both sites. Therefore, only samples from site 2 were included in the calibration data set.

2.3. Seston extraction

Seston was extracted from 100 ml of each sample on Whatman GF/C filters (diameter 47 mm, pore size 1.0 μm) by vacuum filtration. The filters were placed in Petri dishes and heated at 50 °C during 6 h, then dried in desiccators at room temperature (24 °C) during 40 min. Such a drying procedure was based on the study carried out by Dabakk et al. [10]. A 10 mm-diameter circle was randomly cut from each filter and then subjected to the NIRR measurement process. The circle diameter was chosen to be compatible with the size of the diffuse reflectance accessory specified below.

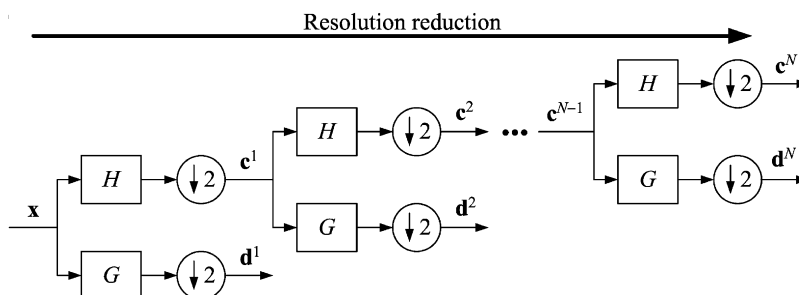


Fig. 2. Filter bank implementation of the wavelet transform with N decomposition levels. Sequences c^k and d^k are termed approximation and detail coefficients at the k th resolution level, respectively. In this work, the PLS model was based on coefficients c^N . The choice of N is clarified in Section 3.

2.4. NIR spectra measurements

The spectra were recorded with a FT-NIR spectrometer Bomem MB 160 D with spectral resolution of 4 cm^{-1} and 32 scans in the near-infrared range $4000\text{--}8000\text{ cm}^{-1}$. The blank was taken as the spectrum of the Whatman GF/C filter after filtration of distilled water and application of the procedures described above. The measurements were carried out at 24°C and 49% relative humidity by using a Bomem diffuse reflectance accessory.

2.5. Wavelet denoising procedure

NIR spectra usually present inherent baseline features that are not related to variations of the parameter of interest. A standard procedure to remove such features consists of using the first derivative of the signal [6]. Since the derivative process tends to magnify instrumental noise, a smoothing algorithm is normally applied. In this context, the Savitzky–Golay polynomial method [6] is widely adopted. However, the adjustment of the window length for this method may not be straightforward, because small windows retain too much noise, whereas large windows may introduce distortions and loss of information. For this reason, in the present work, a Savitzky–Golay derivative procedure is used with a window small enough (25 points) to avoid distorting the spectra, followed by wavelet denoising.

The wavelet transform (WT) is a signal processing tool [16] that has found several applications for the denoising of instrumental signals [17–21]. The WT of a spectrum $x = [x(\nu_1) x(\nu_2) \dots x(\nu_J)]$, where ν_k is the k th wavenumber, can be obtained by using a digital filter bank structure [22] of the form depicted in Fig. 2.

The basic structure of the filter bank consists of a pair of low-pass (H) and high-pass (G) filters, followed by a downsampling operation, which discards one in every two points of the filtering outcome. The downsampled output of the low-pass filter is a smoothed approximation of the signal at a coarser resolution. This operation can be reapplied to the approximation coefficients up to the number of resolution levels specified by the analyst.

In this work, the Symlet 8 wavelet filters from the Symlet family [23] were adopted. The maximum number of resolution levels for which the approximation coefficients did not display a noticeable distortion with respect to the original signal was employed. The approximation coefficients obtained in this

manner were used for PLS modelling and subsequent prediction of COD values in new samples. Such a procedure is similar to the process described in [24] for multivariate calibration of NIR data.

2.6. Software

Spectrum differentiation, Savitzky–Golay smoothing, and PLS modelling were carried out in the CAMO-AS Unscrambler 9.5 software. The wavelet denoising procedure was implemented by using the Mathworks Matlab Wavelet Toolbox.

3. Results and discussion

3.1. Choice of filtration volume

In order to choose an adequate filtration volume for seston collection, six different volumes (10, 25, 50, 100, 150 and 200 ml) of a sample from site 2 were tested. The NIR spectra obtained in this manner are shown in Fig. 3.

As can be seen in Fig. 3, the NIR signal is very small for filtration volumes smaller than 100 ml. On the other hand, the filtration of volumes larger than 100 ml required more than 1 h, whereas the filtration of 100 ml was completed in less than 10 min. Therefore, the use of 100 ml was adopted as the best

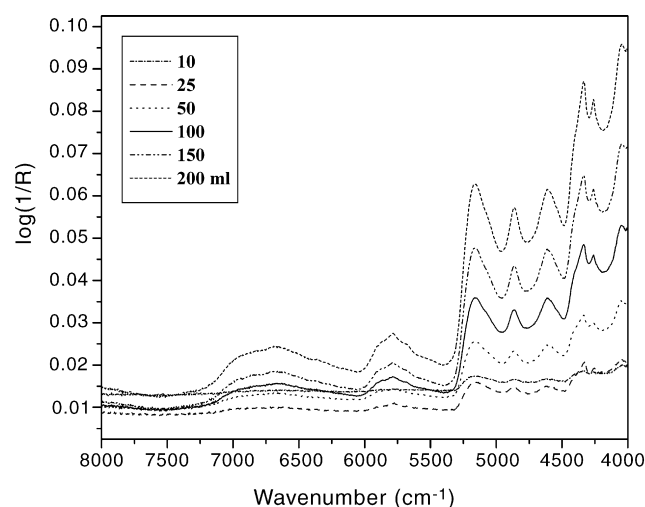


Fig. 3. NIR spectra of seston for different filtration volumes for site 2.

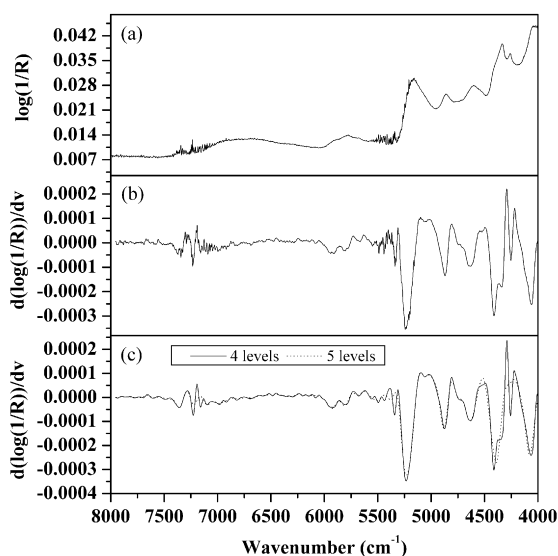


Fig. 4. (a) Raw, (b) derivative and (c) denoised NIRR spectrum of an illustrative sample from site 2.

compromise between analytical speed and NIRR signal sensitivity.

3.2. Pre-treatment of the NIRR spectra

The NIRR spectrum of an illustrative sample is shown in Fig. 4a. The spectrum displays baseline features, which were removed by a Savitzky–Golay derivative procedure with a 2nd-order polynomial and a 25-point window.

As shown in Fig. 4b, the resulting derivative spectrum presents noisy features, which were subsequently removed by the wavelet denoising procedure. Fig. 4c compares the denoising results (that is, the spectra reconstructed from the approximation coefficients by using the inverse wavelet transform) obtained by using four and five wavelet decomposition levels. As can be seen, the use of five levels introduces a significant distortion in the signal, especially in the two peaks around 4250 cm^{-1} . Therefore, four decomposition levels were adopted, resulting in 142 wavelet approximation coefficients for the PLS modelling and prediction procedures.

3.3. PLS modelling

From the overall set of 31 samples, 21 were chosen for PLS calibration and 10 were used as an independent set to evaluate the prediction ability of the resulting model. For this purpose, 12 COD values were selected by applying the Kennard–Stone (KS) uniform sampling algorithm [25,26] to the y vector of 31 elements. The two initial COD values selected in this manner always correspond to the lowest and highest extremes of the interval. Therefore, the two corresponding samples were included in the calibration set in order to avoid extrapolation problems. The samples corresponding to the 10 remaining COD values in the KS selection were taken as the prediction set. Since the KS algorithm spreads the selected y -values uniformly across the COD range, the resulting prediction set is representative and

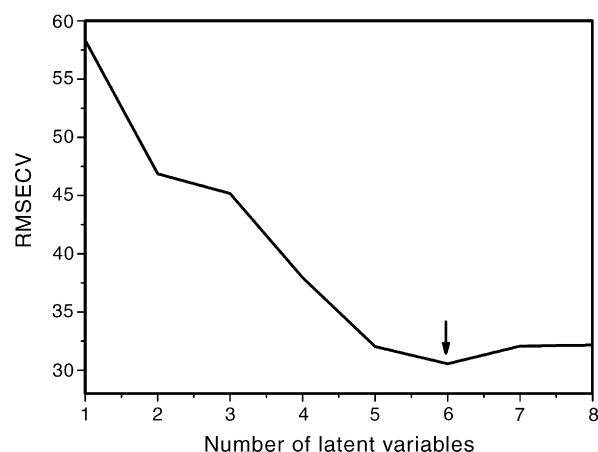


Fig. 5. Cross-validation scree plot for the PLS model constructed with 21 calibration samples and 142 wavelet approximation coefficients. The arrow indicates the point associated to the number of latent variables adopted in the model.

ensures that the predictive ability of the model is evaluated in a fair manner along the entire calibration range. The 19 samples not selected by the KS algorithm, in addition to the two samples with extreme COD values, were taken as the calibration set.

By using the default settings of the Unscrambler software with full cross-validation, six PLS latent variables were selected. Such a finding is corroborated by the minimum point of the cross-validation scree plot presented in Fig. 5. The root-mean-square error of cross-validation (RMSECV) index presented in this graph is defined as

$$\text{RMSECV} = \sqrt{\frac{1}{M_c} \sum_{k=1}^{M_c} (y_{cv}^k - \hat{y}_{cv}^k)^2} \quad (1)$$

where y_{cv}^k and \hat{y}_{cv}^k are the reference and predicted COD values in the k th calibration sample used in the cross-validation procedure, and $M_c = 21$ is the number of calibration samples.

In order to investigate the utility of the wavelet denoising procedure, two additional models were constructed by using the raw spectra (without any pre-treatment) and the derivative spectra obtained with the Savitzky–Golay technique. The prediction results obtained in this manner are presented in Table 1. The results are expressed in terms of correlation between predicted and expected (reference) COD values, as well as root-mean-square error of prediction (RMSEP). As can be seen, both correlation and RMSEP values improve as the derivative is applied to the raw spectra. The number of latent variables also decreases, as expected, since the removal of baseline features eliminates a source of data variability. The application of the

Table 1
Prediction results of the PLS models obtained with different pre-treatment procedures

Pre-treatment	Correlation	RMSEP ($\text{mg O}_2\text{ L}^{-1}$)	LVs ^a
None (raw spectra)	0.91	30	7
Derivative	0.97	23	6
Wavelet	0.97	19	6

^a Number of latent variables.

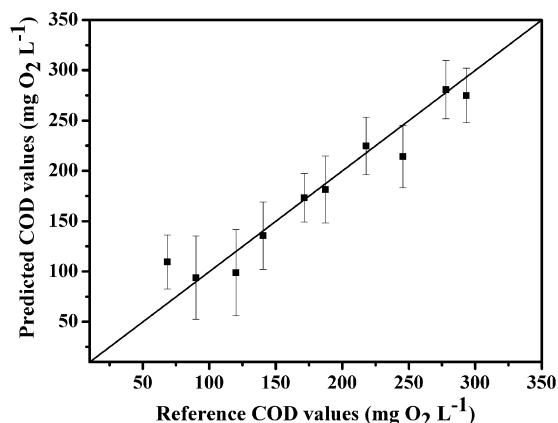


Fig. 6. Reference vs. predicted COD values for the 10 independent samples employed to assess the prediction ability of the PLS model. Error bars indicating the standard deviation of the predictions are displayed.

wavelet treatment further improves the results by decreasing the RMSEP from 23 to 19 mg O₂ L⁻¹ as compared to the model generated with derivative spectra. Henceforth, the discussions will be restricted to the model obtained with the wavelet coefficients.

Fig. 6 is a graph of reference vs. predicted COD values for the 10 independent samples that were not used in the PLS calibration procedures. As can be seen, the points are randomly distributed around the bisectrix line, which indicates that systematic errors are not present. The RMSEP and correlation values obtained for these samples are 19 mg O₂ L⁻¹ and 0.97, respectively. Such results can be deemed acceptable for this application, because the joint estimate of the standard error of the reference method, calculated from the measured COD values of the overall set, is 21 mg O₂ L⁻¹.

In order to provide additional evidence supporting the validity of the proposed method, a full cross-validation (leave-one-out) study was also carried out. In this case, each of the 31 samples was used once for validation. The resulting correlation and RMSECV values were 0.96 and 22 mg O₂ L⁻¹, respectively, which are again acceptable in view of the standard error estimated for the reference method (21 mg O₂ L⁻¹). Fig. 7 is the graph of reference vs. predicted COD values for the cross-validation study. Again, systematic errors are not present.

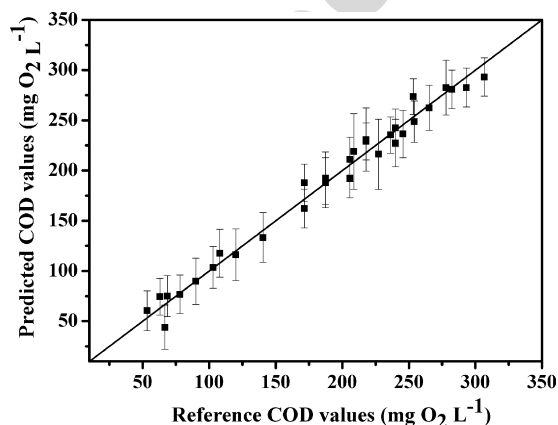


Fig. 7. Reference vs. predicted COD values for the cross-validation study. Error bars indicating the standard deviation of the predictions are displayed.

4. Conclusions

This paper proposed a method for COD determination in domestic wastewater based on NIR measurements of sestonic matter collected by filtration. The proposed method does not require any special reagents, catalysts or solvents. The baseline and noise features of the NIR spectra are removed by a Savitzky–Golay derivative procedure followed by wavelet denoising. The resulting wavelet coefficients are employed for PLS modelling and prediction.

The proposed method was applied to COD determination of samples collected from the Mangabeira municipal wastewater treatment plant in the city of João Pessoa (Paraíba, Brazil). A statistical analysis revealed that the main depuration processes responsible for COD reduction occur in the anaerobic ponds. Therefore, the investigation was concerned with the effluent of these ponds.

The results of the NIR proposed method were found to be in good agreement with the COD values obtained according to the APHA standard method (RMSEP and correlation values of 19 mg O₂ L⁻¹ and 0.97, respectively). Future extensions of this work could include the determination of other physical, chemical, and bacteriological parameters associated to water quality.

Finally, it is worth noting that the composition of the incoming wastewater in stabilization ponds can vary according to the habits of the population, as well as the weather and other external factors. Therefore, the application of empirical spectroscopic models to different treatment plants and time periods should be handled with care to avoid possible extrapolations of the calibration limits. In the present work, seasonal variations were taken into account by using calibration data from an entire year including periods of low and high rain index. If the model were to be applied to a different treatment plant, the calculated uncertainty limits for the predicted values could be used as a criterion for validating the prediction and checking for possible extrapolation problems, as suggested elsewhere [6]. If necessary, additional samples from other treatment plants could be inserted in the calibration set in order to improve the generality of the resulting model. However, such a procedure tends to produce more complex models, in terms of the number of latent variables required to describe the additional sources of variation in the data. In addition, a global model is expected to be less accurate than a model specifically developed for a particular treatment plant [27].

Acknowledgments

This work was supported by CNPq (grants Universal 475204/2004-2 and 476406/2003-0, Pronex 015/98, scholarships and research fellowships), CAPES (PROCAD 0081/05-1 and scholarships), FAPESQ-PB/CNPq (research fellowship), SUDEMA (scholarship). The authors are also indebted to Dr. Celio Pasquini (Instituto de Química Universidade Estadual de Campinas) for granting access to the FT-NIR Bomem spectrometer as well as for his valuable comments and suggestions. The authors also wish to thank Companhia de água e

esgoto do estado da Paraíba (CAGEPA) and Superintendência de Administração do Meio Ambiente do estado da Paraíba (SUDEMA).

References

- [1] V.M. Medeiros, M.C.U. Araujo, R.K.H. Galvão, E.C. Silva, T.C.B. Saldanha, I.A.S. Toscano, M.S.R. Oliveira, S.K.B. Freitas, M.M. Neto, *Water Res.* 39 (2005) 3089.
- [2] P.P. Zamora, A.G. Cordeiro, N. Nagata, *Quim. Nova* 28 (2005) 838.
- [3] C. Pasquini, *J. Braz. Chem. Soc.* 14 (2003) 198.
- [4] L. Bokobza, *J. Near Infrared Spectrosc.* 6 (1998) 3.
- [5] J.J. Workman Jr., *Appl. Spect. Rev.* 31 (1996) 251.
- [6] K.R. Beebe, R.J. Pell, M.B. Seasholtz, *Chemometrics: A Practical Guide*, Wiley, New York, 1998.
- [7] T. Koraman, M. Nilson, J. Ohman, I. Renberg, *Environ. Sci. Technol.* 26 (1992) 2122.
- [8] M.B. Nilsson, E. Dabakk, T. Korsman, I. Renberg, *Environ. Sci. Technol.* 30 (1996) 2586.
- [9] D.F. Malley, P.C. Williams, M.P. Stainton, *Water Res.* 30 (1996) 1325.
- [10] E. Dabakk, M. Nilson, P. Geladi, S. Wold, I. Renberg, *Water Res.* 34 (1999) 1666.
- [11] A.A. Chisty, P.K. Egeberg, *Chemom. Intell. Lab. Syst.* 50 (2000) 225.
- [12] E. Dulekgurgen, S. Dogruel, O. Karahan, D. Orhon, *Water Res.* 40 (2006) 273.
- [13] S. Aiyuk, J. Amoako, L. Raskinb, A.V. Haandelc, W. Verstraetea, *Water Res.* 38 (2004) 3031.
- [14] Z. Hu, K. Chandran, B.F. Smets, G. Domenico, *Water Res.* 36 (2002) 617.
- [15] APHA, *Standard Methods for Examination of Water and Wastewater*, 20th ed., American Public Health Association, Washington, DC, USA, 1998.
- [16] B. Walczak (Ed.), *Wavelets in Chemistry*, Elsevier Science, New York, 2000.
- [17] R.K.H. Galvão, H.A.D. Filho, M.N. Martins, M.C.U. Araújo, C. Pasquini, *Anal. Chim. Acta* 581 (2007) 159.
- [18] V.J. Barclay, R.F. Bonner, I.P. Hamilton, *Anal. Chem.* 69 (1997) 78.
- [19] C. Cai, P.B. Harrington, *J. Chem. Inf. Comput. Sci.* 38 (1998) 1161.
- [20] L. Pasti, B. Walczak, D.L. Massart, P. Reschiglian, *Chemom. Intell. Lab. Syst.* 48 (1999) 21.
- [21] U. Depczynski, K. Jetter, K. Molt, A. Niemoller, *Chemom. Intell. Lab. Syst.* 49 (1999) 151.
- [22] M. Vetterli, J. Kovacevic, *Wavelets and Subband Coding*, Prentice-Hall, New Jersey, 1995.
- [23] B.K. Alsberg, A.M. Woodward, M.K. Winson, J.J. Rowland, D.B. Kell, *Anal. Chim. Acta* 368 (1998) 29.
- [24] J. Trygg, S. Wold, *Chemom. Intell. Lab. Syst.* 42 (1998) 209.
- [25] R.W. Kennard, L.A. Stone, *Technometrics* 11 (1969) 137.
- [26] R.K.H. Galvão, M.C.U. Araujo, G.E. José, M.J.C. Pontes, E.C. Silva, T.C.B. Saldanha, *Talanta* 67 (2005) 736.
- [27] F. Despaigne, D.L. Massart, *Anal. Chem.* 72 (2000) 1657.