



UEPB

**UNIVERSIDADE ESTADUAL DA PARAÍBA
CAMPUS I – CAMPINA GRANDE
CENTRO DE CIÊNCIAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA E TECNOLOGIA
AMBIENTAL**

EMANOEL DA SILVA CARVALHO

**MODELOS QUIMIOMÉTRICOS PARA A IDENTIFICAÇÃO E
QUANTIFICAÇÃO DE CAFEÍNA EM BAIXAS CONCENTRAÇÕES EM MEIO
AQUOSO**

**CAMPINA GRANDE
2017**

EMANOEL DA SILVA CARVAVALHO

**MODELOS QUIMIOMÉTRICOS PARA A IDENTIFICAÇÃO E
QUANTIFICAÇÃO DE CAFEÍNA EM BAIXAS CONCENTRAÇÕES EM MEIO
AQUOSO**

Trabalho de Dissertação apresentada ao Programa de Pós-Graduação em Ciência e Tecnologia Ambiental da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de Mestre.

Área de concentração: tecnologias para o tratamento de água e efluentes, para a obtenção do título de Mestre.

Orientador: Prof^a Dr^a. Simone da Silva Simões

CAMPINA GRANDE

2017

É expressamente proibida a comercialização deste documento, tanto na forma impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que na reprodução figure a identificação do autor, título, instituição e ano da dissertação.

C331m Carvalho, Emanuel da Silva.
Modelos quimiométricos para a identificação e quantificação de cafeína em baixas concentrações em meio aquoso [manuscrito] / Emanuel da Silva Carvalho. - 2017.
61 p.
Digitado.
Dissertação (Mestrado em Ciência e Tecnologia Ambiental) - Universidade Estadual da Paraíba, Centro de Ciências e Tecnologia, 2017.
"Orientação: Profa. Dra. Simone da Silva Simões, Departamento de Engenharia Sanitária e Ambiental".

1. Tratamento de água. 2. Xenobióticos. 3. Aquafotomica.
4. Cafeína. 5. Quimiometria. I. Título.

21. ed. CDD 628.162

EMANOEL DA SILVA CARVAVALHO

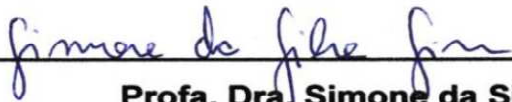
**MODELOS QUIMIOMÉTRICOS PARA IDENTIFICAÇÃO E QUANTIFICAÇÃO
DE CAFÉINA EM BAIXAS CONCENTRAÇÕES EM MEIO AQUOSO**

Trabalho de Dissertação apresentada ao Programa de Pós-Graduação em Ciência e Tecnologia Ambiental da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de Mestre.

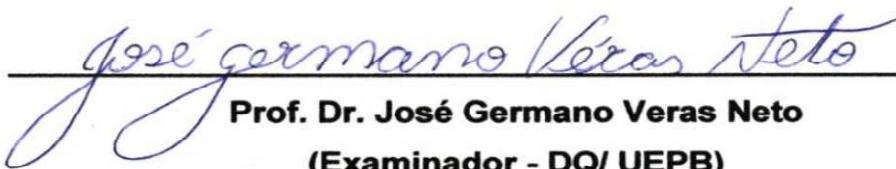
Área de concentração: tecnologias para o tratamento de água e efluentes, para a obtenção do título de Mestre.

APROVADO EM: 28 / 04 / 2017

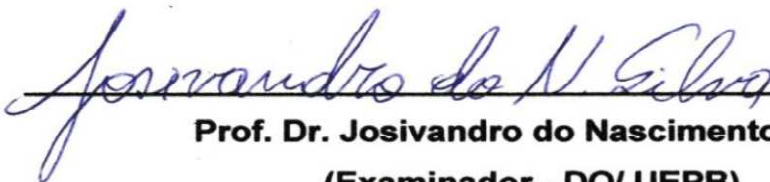
BANCA EXAMINADORA



Prof. Dra. Simone da Silva Simões
(Orientadora – DQ/ UEPB)



Prof. Dr. José Germano Veras Neto
(Examinador - DQ/ UEPB)



Prof. Dr. Josivandro do Nascimento Silva
(Examinador - DQ/ UEPB)

RESUMO

O aumento da demanda da população por bens de consumo industrializados e conseqüentemente o aumento da geração de resíduos, tem relação direta com estado de saúde da população, devido à gravidade das implicações relativas a bioacumulação de compostos como medicamentos, produtos de higiene pessoal, substâncias orgânicas tóxicas, também conhecidos como xenobióticos. Além disso, a inadequação dos sistemas de tratamento e legislação relacionados a água é alarmante pois a qualidade da água de abastecimento fornecida na maioria das capitais brasileiras, de acordo com pesquisa realizada pelo Instituto Nacional de Ciências e Tecnologias Analíticas Avançadas (INCTAA), precisa ser melhorada, apesar de atender os padrões de potabilidade. Nas últimas décadas, muitos avanços foram feitos em relação aos métodos instrumentais destinados a identificação e quantificação contaminantes em água, e atualmente o maior desafio desta área está relacionado ao desenvolvimento de métodos destinados ao tratamento das amostras. Diante do exposto, desenvolveu-se uma metodologia de baixo custo, baseada nos princípios da aquafotomica capaz de identificar e quantificar contaminantes em concentrações traço em meio aquoso. Para isso foi realizado um extenso levantamento bibliográfico sobre esta nova área da ciência, a aquafotomica. A parte experimental foi baseada na espectroscopia UV-vis-NIR aliada a técnicas de reconhecimento de padrões, como a PCA, e de calibração multivariada, como PLS e MLR. A construção dos modelos de identificação e quantificação foi baseada em 62 espectros de padrões sintéticos contendo cafeína em baixas concentrações ($10e^{-5}$ a $10e^{-9}$) em meio aquoso. Os modelos PCA para identificação foi capaz de distinguir as amostras contendo cafeína daquelas de água pura. Já os modelos de quantificação mostraram uma capacidade preditiva satisfatória para amostras com concentrações superiores a $10e^{-6}g.L^{-1}$ em cafeína.

Palavras-Chave: Tratamento de água. Xenobióticos. Aquafotomica. Cafeína. Quimiometria

ABSTRACT

The increase in population for manufactured consumer it is a consequence from the waste generated, has a direct consequence without health status of the population, due to the serious consequences with the bioaccumulation of composed, such as medicines, personal hygiene products, toxic organic substances, also known as xenobiotics. Furthermore, an alarming inadequacy from the treatment systems and legislation related to the water, because the quality of the water supply in most Brazilian capitals, according to research provided by the Instituto Nacional de Ciências e Tecnologias Analíticas Avançadas (INCTAA), need to be revised, even though it is according to the standards of potability. In last decades, some advances have been made in relation to instrumental methods destined to identification and quantification contaminants in water, and lately the main challenges are related to the development intended for treatment of the samples. On what was exposed, developed a methodology with a low cost, based on aquafotomics standards capable of identifying and quantifying contaminants in trace concentrations in aqueous environment. Therefore, an extensive bibliographical survey was carried out on this new area of science, an aquafotomics. The experimental part was based on spectroscopy for UV-vis-NIR combined with pattern recognition techniques spectroscopy, such as a PCA, and multivariate calibration, such as PLS and MLR. The construction of the identification and quantification models was made in 62 models of synthetic standards containing caffeine in low concentrations ($10e^{-5}$ to $10e^{-9}$) in aqueous environment. The PCA models for identification were able to distinguish as samples containing caffeine from those of pure water. The quantification models gives a satisfactory predictive capacity for samples with concentrations higher than $10e^{-6}$ g.L⁻¹ in caffeine.

Keywords: Water treatment. Xenobiotics. Aquafotomic. Caffeine. Chemometrics

LISTA DE FIGURAS

Figura 1: 100 milhões de Registros no CAS em 50 anos de existência (2016) Fonte: https://www.cas.org/news/media-releases/100-millionth-substance	9
Figura 2: Modelo de estrutura molecular da cafeína com numeração dos átomos. Fonte: MEJRI et al., 2009	16
Figura 3 Estrutura cristalina da cafeína monohidratada.	17
Figura 4: Simulação de nuvem de moléculas de água em interação com cafeína. Tavagnacco et al. (2011).....	19
Figura 5: Etapas para a construção de aquagramas relacionados a transformação da estrutura da água durante a formação da fibrina.....	22
Figura 6: Os três modos de vibração da molécula de água, sendo o módulo v_2 relacionado principalmente a deformação angular e os módulos v_1 e v_3 de deformação axial.	26
Figura 7: Espectro na região UV-vis-NIR da água pura	37
Figura 8: Espectro na região UV-vis-NIR das soluções contendo cafeína nas concentrações de 1×10^{-5} a 4×10^{-9}	37
Figura 9: gráfico de escores gerado pela PCA utilizando todas as amostras..	38
Figura 10: gráfico de escores gerado pela PCA utilizando apenas as amostras de água pura e contendo cafeína em baixa concentração.....	39
Figura 11: A) Gráfico de escores gerado pela PCA utilizando apenas as amostras contendo cafeína e B) Gráfico de pesos para PC1.....	40
Figura 12: Coeficientes selecionados pelo algoritmo Jack Knife para a construção do modelo PLS.	42
Figura 13: A) Gráfico dos valores medidos versus os valores previstos pelo modelo e B) Detalhe do gráfico para amostras com concentração inferior a 1×10^{-6} g.L ⁻¹ em cafeína.	44
Figura 14: Modelo MLR-SPA com dados brutos (a) variáveis selecionadas vs RMSEC e (b) valores medidos vs valores preditos pelo modelo.....	46
Figura 15: Variáveis selecionadas pelo SPA para os dados brutos.	47
Figura 16: Modelo MLR-SPA com dados preprocessados (A) variáveis selecionadas vs RMSE e B) valores medidos vs valores preditos pelo modelo.	48

LISTA DE TABELAS

Tabela 1: intensidades relativas de bandas de estiramento C-H para região infravermelho e sobretons de Infravermelho Próximo. Fonte: Water Journal...	20
Tabela 2: Parâmetros para avaliação dos modelos PLS	43
Tabela 3: Parâmetros para avaliação do modelo MLR-SPA	45

LISTA DE SIGLAS E ABREVIATURAS

AAS - Ácido Acetilsalicílico

CLS - Método dos Mínimos Quadrados Clássicos

ETA - Estações de Tratamento de Água

ETE - Estações de Tratamento de Esgoto

ERD- Espectroscopia de Relaxação Dielétrica

HCA - Análise de Agrupamentos Hierárquicos

IR - Infravermelho

INCTAA - Instituto Nacional de Ciências e Tecnologias Analíticas

INSA - Instituto Nacional do Semiárido

LOD - Limite de Detecção

JK - Jack Knife

UV - Ultravioleta

Vis - Visível

NIR - Infravermelho Próximo

PCA - Análise de Componentes Principais

PCR - Regressão por Componentes Principais

PLS - Partial Least Squares

MLR - Regressão Linear Múltipla

MSC - Multiplicative Scatter Correction

NIPALS - Non-linear Interactive Partial Least Squares

CAS – Chemical Abstract Services

FTIR – Infravermelho com Transformada de Fourier

RMN - Ressonância Magnética Nuclear

SPA - Algoritmo das Projeções Sucessivas

VL - Variáveis Latentes

WASP - Water Spectral Pattern

WAMACs - Water Matrix Coordinates

SUMÁRIO

1. INTRODUÇÃO	9
2. OBJETIVO	12
2.1 Objetivo Geral:	12
2.2 Objetivos específicos:	12
3. JUSTIFICATIVA	13
4. FUNDAMENTAÇÃO TEÓRICA	15
4.1 Xenobióticos	15
4.3 Aquafotomica.....	19
4.3.2 O uso da aquafotomica	23
4.4 Espectroscopia na região do infravermelho	26
4.5 Quimiometria.....	29
4.6 Métodos de Calibração Multivariada	30
4.6.1 Regressão por mínimos quadrados parciais (PLS).....	32
4.6.3 Algoritmo das Projeções Sucessivas.....	33
5. MATERIAIS E MÉTODOS	34
5.1 Tipo de pesquisa	34
5.2 Área de estudo.....	34
5.3 Amostras	34
5.4 Software.....	34
5.5 Metodologia.....	34
6. RESULTADOS E DISCUSSÕES	36
6.1 Caracterização físico-química da água pura e das soluções contendo cafeína	36
6.2 Modelo multivariado para identificação da cafeína em meio aquoso	36
6.3 Metodologia multivariadas para quantificação da cafeína em meio aquoso. 41	
7. CONCLUSÕES	49
REFERENCIAS	51

1. INTRODUÇÃO

Após a Segunda Guerra Mundial, a produção e o uso indiscriminado de pesticidas marcaram o início de um período de desenvolvimento acelerado das atividades industriais. Foi observado até mesmo um retorno dos investimentos massivos em tecnologia militar, principalmente por meio da aplicação de muitas tecnologias intermediárias na melhoria dos bens de consumo (DE MOURA, 2000).

Na medida que a sociedade e a indústria se modernizam os números de registros da CAS (Chemical Abstracts Service) que se trata de um serviço de registros de substâncias química, aumentam exponencialmente. Nos últimos cinquenta anos a CAS registrou mais de 100 milhões de substâncias químicas, como pode ser observado na figura 1, na última década a taxa de registro superou todos os anos anteriores.

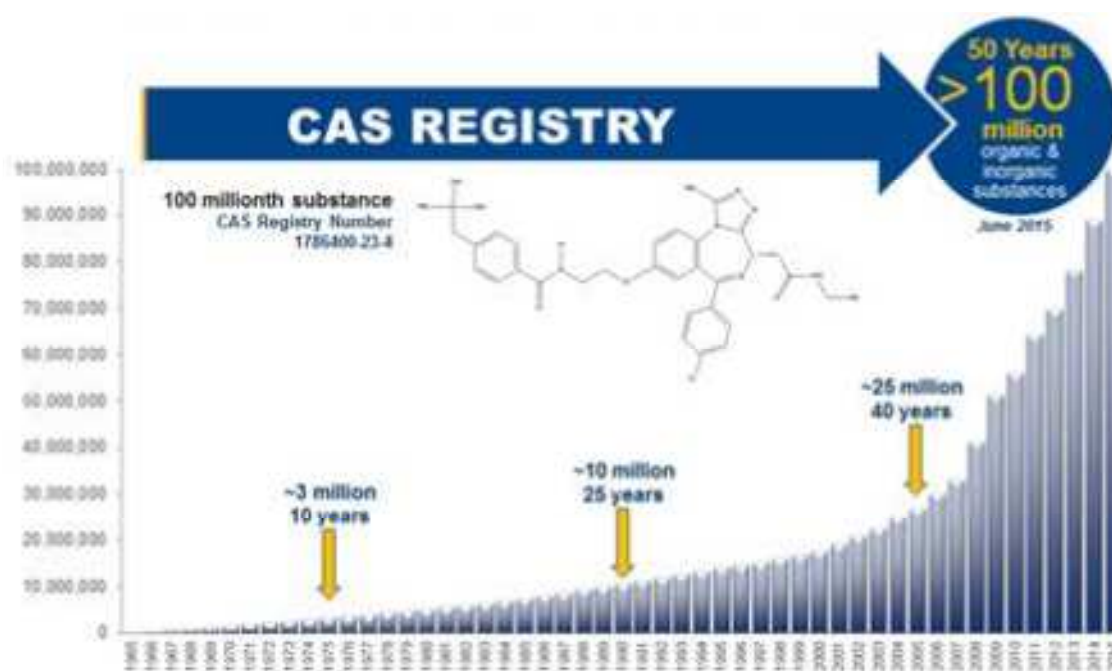


Figura 1: 100 milhões de Registros no CAS em 50 anos de existência (2016)
 Fonte: <https://www.cas.org/news/media-releases/100-millionth-substance>

O aumento da demanda da população por bens de consumo industrializados, aliado ao crescimento gradual da geração de resíduos, tem preocupado a comunidade como um todo, devido à gravidade das

consequências relativas a bioacumulação de compostos xenobióticos, como medicamentos, produtos de higiene pessoal, etc. Essa preocupação é ampliada na medida que se observa a inadequação dos sistemas de tratamento e legislação relacionados a água. A qualidade da água de abastecimento fornecida na maioria das capitais brasileiras, de acordo com pesquisa realizada pelo Instituto Nacional de Ciências e Tecnologias Analíticas Avançadas (INCTAA), precisa ser melhorada, apesar de atender os padrões de potabilidade para águas de abastecimento definidos pelo Ministério da Saúde (PORTARIA Nº 2.914, DE 12 DE DEZEMBRO DE 2011). De acordo com o Prof. Dr. Wilson Jardim “A portaria 2914, do Ministério da Saúde, que normatiza a qualidade da água potável é muito estática, e a nossa sociedade é dinâmica”. (JORNAL DA UNICAMP, 2013).

Pode-se denominar xenobióticos (do grego xenos=estranho) os compostos químicos sintetizados, bem como os subprodutos da indústria química, diferentes modos de ação e toxicidade (SILVA, 2005). Dentre eles pode-se citar plásticos, solventes, lubrificantes, detergentes, agrotóxicos assim como uma variedade de subprodutos da química. Muitos deles são poluentes orgânicos que não participam dos ciclos globais do carbono, nitrogênio ou enxofre e exibem efeitos tóxicos não desejados aos organismos vivos. (NEILSON, A. H. et al 1985). Neste trabalho, a cafeína foi considerada um xenobiótico devido ao alto consumo de bebidas, alimentos e medicamentos contendo este composto químico.

Os xenobióticos são encontrados no meio ambiente geralmente em concentrações traço e geralmente são quantificados e identificados por cromatografia (DRUMMER, 1999). Que, apesar de ser uma técnica consolidada, apresenta um custo elevado devido a quantidade de reagentes utilizados que também podem ser prejudiciais ao meio ambiente. Além disso, necessita de diversas técnicas e etapas de preparação de amostra, o que pode não ser interessante quando se trabalha com substâncias em concentração traço.

Aliado a este fato, as técnicas convencionais para remoção ou minimização destes compostos como as membranas de ultrafiltração (SCHNEIDER e TSUTYA, 2001), osmose inversa (CORNELL COOPERATIVE EXTENSIÓN, NEW CORK STATE COLLEGE OF HUMAN ECOLOGY, 1995),

associação dos processos de coagulação/floculação/sedimentação (C/F/S) e nanofiltração (NF) (TATEOKA, M. S., 2014) ou utilização de materiais como carvão ativado pulverizado (MUCCIACITO, J. C., 2006), não tem se mostrado eficientes.

Diante do exposto, a procura por metodologias capazes de identificar e quantificar compostos prejudiciais ao meio ambiente, em concentração traço, de forma rápida, pouco dispendiosa, com precisão e exatidão, é uma realidade.

A aquafotomica, (do inglês: aquaphotomic), é uma metodologia, ainda pouco divulgada e utilizada, que define a água como um sistema multielementar e utiliza suas bandas de absorção e os seus padrões de absorção para fornecer um conhecimento de estruturas de interações entre a água e outros componentes do sistema aquoso, adquirindo informações da substância em estudo indiretamente a partir do espectro da água. Neste contexto, as espectroscopias na região do ultravioleta, visível, infravermelho próximo e médio, Raman e Terahertz permitem medidas não invasivas e multivariadas, servindo com uma fonte de informação sistematizada para um banco de dados chamado aquafotome que serve de base para análise de amostras contendo compostos em concentração traço. (TSENKOVA, 2009)

2. OBJETIVO

2.1 Objetivo Geral:

Realizar um levantamento sobre as metodologias baseadas no princípio da aquafotomica para a identificação de compostos em baixa concentração em meio aquoso, bem como desenvolver uma metodologia, baseada nestes princípios para identificar e quantificar cafeína em meio aquoso.

2.2 Objetivos específicos:

- Conhecer e divulgar o uso da aquafotomica para a identificação e quantificação de produtos em baixa concentração em meio aquoso.
- Desenvolver um modelo quimiométrico de identificação de cafeína em água.
- Avaliar o desempenho de metodologias de calibração multivariada como PLS e MLR na quantificação de cafeína em concentrações traço em água.

3. JUSTIFICATIVA

No período da segunda guerra mundial (1939-1945) o investimento no desenvolvimento de novos compostos, como produtos farmacêuticos e alimentícios, foi impulsionado e sua distribuição ocorreu sem nenhuma preocupação relativa aos possíveis contaminantes que poderiam ser lançados no ambiente (DRUCKER, 1967; DE MOURA, 2000) e por muitos anos o controle e efeitos destes contaminantes na saúde foi negligenciado. No entanto, a partir da década de 70, despertou-se o interesse por pesquisas nesta área e conseqüentemente promoveu-se aumento no monitoramento destes compostos xenobióticos no meio ambiente.

Os compostos xenobióticos podem ser definidos como: substâncias sintetizadas artificialmente que não existem na natureza, mas que podem surgir no ambiente por meio de ação antrópica. Uma boa parte dos fármacos por exemplo, após ingeridos podem ser parcialmente metabolizados pelo organismo e outra parte excretada, chegando às estações de tratamento de esgoto, onde podem persistir e chegar aos leitos aquáticos, dentre outros meios, contaminando sistemas biológicos e/ou diversos compartimentos ambientais (GHISELLI; JARDIM; SODRÉ et al., 2007).

Um estudo coordenado por um grupo de pesquisa da Universidade Estadual de Campinas pesquisadores detectaram que 20 contaminantes emergentes nas águas de abastecimento das capitais brasileiras e alertou quanto a necessidade de evoluir os sistemas de tratamento e a próprias legislação para acompanhar o ritmo da sociedade (JORNAL DA UNICAMP, 2013)

Nas últimas décadas, muitos avanços foram feitos em relação aos métodos instrumentais (equipamentos) destinados a identificação e quantificação de contaminantes em água. Deste modo, atualmente o maior desafio desta área é o desenvolvimento de metodologias simples, com baixo limite de detecção e quantificação, de baixo custo, precisa; além de técnicas que possam isolar o analito de sua matriz, mesmo que estes se encontrem em concentração traço.

Este trabalho se propõe a desenvolver uma metodologia de baixo custo, baseada nos princípios da aquafotomica (TSENKOVA, 2009), para identificar e quantificar contaminantes em concentrações traço em águas. Conhecimento que poderá contribuir para as análises de água e diagnósticos ambientais.

4. FUNDAMENTAÇÃO TEÓRICA

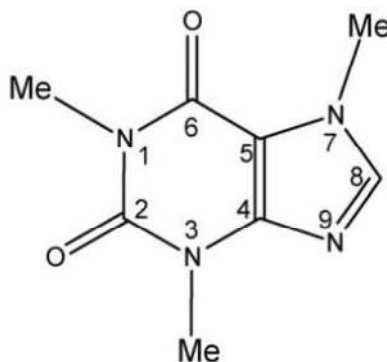
4.1 Xenobióticos

A inexistência de planejamentos baseados em critérios ambientais, deu origem a um cenário no qual o lançamento de esgoto vem contribuindo para o aporte de muitas espécies exógenas, incluindo-se microrganismos patógenos e substâncias orgânicas xenobióticas em corpos aquáticos receptores. Dependendo das características do corpo receptor, da substância química e igualmente das estratégias de saneamento, muitos compostos podem ser transferidos para estações de tratamento de água (ETA) e para a água potável. Dentre estes compostos podemos citar alguns que sofrem pouca ou nenhuma alteração quando da passagem pelo metabolismo humano como o caso da cafeína e paracetamol; e outros que formam metabolitos secundários, como o caso dos hormônios. (GHISELLI; JARDIM; SODRÉ et al., 2007)

Os xenobióticos encontrados podem ser de diferentes classes como: analgésicos, anti-inflamatórios, drogas psiquiátricas, antilipêmicos, antibióticos (de uso humano e veterinário), hormônios e esteroides, além de componentes presentes em protetores solares, produtos de higiene pessoal como fragrâncias contendo grupos nitro e ftalatos, inseticidas repelentes e antissépticos, bem como os surfactantes bastantes utilizados no consumo doméstico. Estes compostos também podem ser chamados de contaminantes emergentes. (HEUETT & NUBIA, 2015). No Brasil os principais fármacos encontrados em águas de abastecimento, com relação a sua abundância, são o ibuprofeno, o diclofenaco, o paracetamol e o ácido acetilsalicílico (AAS) (GHISELLI; JARDIM, 2007).

Além dos xenobióticos citados, a cafeína apresenta um papel essencial neste estudo. Apesar de existir naturalmente em mais de 60 espécies de plantas além do próprio café, tais como, ervas medicinais e guaraná é muito explorada pelo homem na fabricação medicamentos para resfriados, alergias e em analgésicos, em concentrações na faixa de 15 a 64 mg, por unidade, moderadores de apetite apresentam de 50 a 200 mg por unidade, essa exploração contribui com o aumento da sua concentração nos leitos aquáticos

(SRISUPHAN; BRACKEN, 1986). Mesmo não havendo comprovações sobre sua ação no sistema endócrino e também não se conhece seu efeito à biota, em alguns casos foi associada com elevadas concentrações de nitratos no meio aquático (CHEN et al., 2002) ou à presença de coliformes totais, sendo comumente usada como traçador de contaminação fecal, pois pelos mesmos meios que esses elementos conseguem passar pelas estações de tratamento, a cafeína também passa. (PIOCOS & DE LA CRUZ, 2000).



(1,3,7-trimetilxantina) $C_8H_{10}N_4O_2$

Figura 2: Modelo de estrutura molecular da cafeína com numeração dos átomos. Fonte: MEJRI et al., 2009

A cafeína (figura 2) é um composto largamente consumido no Brasil e no mundo, da classe das metil xantinas que são substâncias que possuem a propriedade de estimular o sistema nervoso central, sendo provavelmente a substância psicoativa mais consumida no mundo (GARDINALI; ZHAO, 2002; DIOGO et al., 2013). A cafeína tem alta solubilidade em água e é recalcitrante aos processos normalmente usados nas ETA, posicionando-o como um composto indicador de contaminação antrópica em variados estudos realizados ao redor do mundo. Nos últimos anos, muitas outras evidências têm mostrado que a presença de cafeína em águas brasileiras está diretamente relacionada ao aporte de esgoto, muitas vezes bruto, em ambientes aquáticos naturais (SODRÉ et al., 2007; FROEHNER ET AL., 2010; 2011B; MONTAGNER; JARDIM, 2011; GOMES et al., 2013). No Brasil, o uso da cafeína como indicador da qualidade de águas foi advogado em 2005 por pesquisadores da Fundação Osvaldo Cruz com base em estudos voltados à detecção de cafeína em águas superficiais do estado do Rio de Janeiro (FERREIRA, 2005; FERREIRA; CUNHA, 2005).

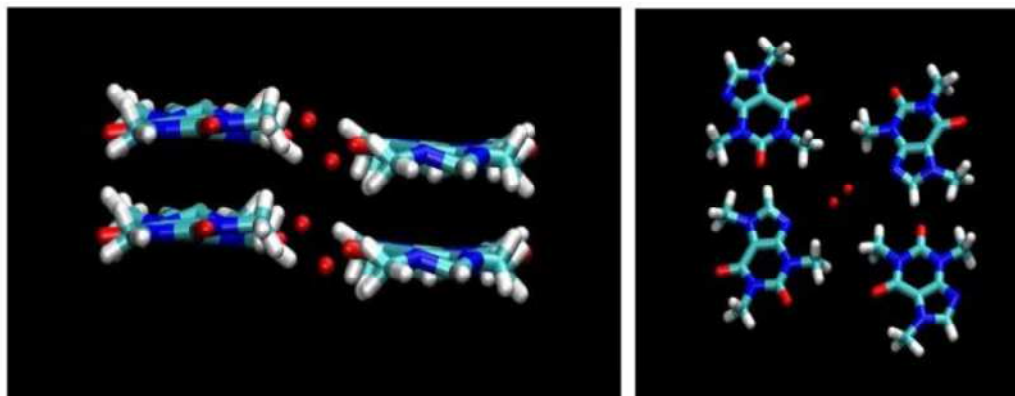


Figura 3 Estrutura cristalina da cafeína monohidratada.
Fonte: Tavagnacco et al. (2011)

As moléculas de cafeína não possuem grupos doadores de hidrogênio e, portanto, são incapazes de se auto associar por ligação de hidrogênio diretamente. A figura 3 apresenta dois pontos de vista da estrutura cristalina da cafeína monohidratada, que ilustra a posição da molécula de água (como esferas vermelhas sobre o átomo de oxigênio), ligados ao átomo de nitrogênio na posição 9 (N9) da molécula de cafeína fazendo uma segunda ligação com outra molécula de água. (MEJRI et al., 2009; TAVAGNACCO et al., 2011)

Depois de administrada, muitas drogas são transformadas em um ou mais metabólitos e excretadas, através da urina e fezes, em uma mistura contendo o fármaco que não foi metabolizado pelo organismo e seus metabólitos. No caso do ibuprofeno, por exemplo, apenas 15% do que foi ingerido, é eliminado na forma original, enquanto que 26% são excretados nas formas hidroxibuprofeno e 43% como carboxibuprofeno, essas porcentagens podem variar de acordo com o fármaco (WEIGEL ET AL., 2004).

Nas ETE's há três destinos possíveis para qualquer fármaco individual: (1) pode ser biodegradado, ou seja, mineralizado a gás carbônico e água, (2) pode sofrer algum processo metabólico ou ser parcialmente degradado; (3) ou pode ser persistente. Pouco se conhece sobre as rotas dos fármacos no ambiente, mas sabe-se que os esgotos domésticos e hospitalares são as principais fontes de contaminação, sobretudo de matrizes aquáticas. Mesmo em concentrações-traço, alguns compostos exógenos, sintéticos ou naturais, têm sido detectados em amostras de águas superficiais em todos os

continentes do planeta, principalmente em função da atividade antrópica. (RAIMUNDO, 2007)

4.2 Água

A água é um dos mais importantes e bem distribuídos componentes do meio biológico e é um constituinte básico para a maioria das estruturas. Várias técnicas têm sido utilizadas para estudar as propriedades físicas e químicas da água, dentre elas podemos citar: Espectrometria de Raio-X (HUANG, C. et al., 2009; HUSE, N. et al., 2009; IKEHATA, A. et al., 2010), Espectroscopia de infravermelho (ROBERTSON, W. H. et al., 2003; WEBER, J. M. et al., 2000), Espectroscopia de Espalhamento Raman (WALRAFEN, G., 1972), Ressonância Magnética Nuclear (RMN) Espectroscopia (RUAN, R. R. & CHEN, P. L., 1998), Espectroscopia de Relaxação Dielétrica (ERD) (KABIR, S. R. et al., 2003; SHINYASHIKI, N. et al., 2009), Espectroscopia Terahertz (EBBINGHAUS, S. et al., 2007), Difração de Nêutrons (WINKEL, K. et al., 2009). Cada uma destas técnicas é capaz de fornecer valiosas informações em diferentes aspectos para o sistema água. Para entender a água do ponto de vista de um sistema, algumas ferramentas são necessárias para monitorar em tempo real e de forma não invasiva o sistema biológico e aquoso.

Do ponto de vista espectroscópico, pode-se descrever a água como um solvente e também uma matriz, ou seja, um cruzamento molecular de conformações da molécula da água com diferentes populações de ligações de hidrogênio e diferentes distribuições para acomodar e facilitar várias funções, inclusive as biológicas (ROY, R. et al., 2005). A força da ligação das moléculas de hidrogênio e a sua localização ao redor de outra molécula diferente da água definem sua funcionalidade, isto pode ser descrito como um corpo multifuncional tendo simultaneamente várias faces, ou seja, várias estruturas.

Muitas estruturas moleculares são estudadas utilizando a espectroscopia vibracional no infravermelho (IR) e as regiões vis-NIR (Visível-Infravermelho Próximo). A água é um forte absorvente da energia infravermelha e, como resultado, este método só é adequado para análise de amostras com pequenas porções de água. Assim, a espectroscopia no IR torna-se inútil quando a tarefa é a de avaliar um complexo sistema biológico e

aquoso. Deste modo, para análises de amostras com alto teor de água o ideal é usar a região do vis-NIR (TSENKOVA, 2010).

O espectro da água como um meio depende do modo de vibração das moléculas do soluto, quando existe um. Ou seja, as vibrações da molécula da água (conformações da molécula da água) são influenciadas pela interação das moléculas da água com as outras moléculas presentes na solução. Deste modo, uma boa compreensão das interações entre a energia NIR e a água, em suas diversas conformações moleculares, é fundamental para a compreensão de um sistema aquoso. (ROY, R. et al., 2005; TSENKOVA, 2010)

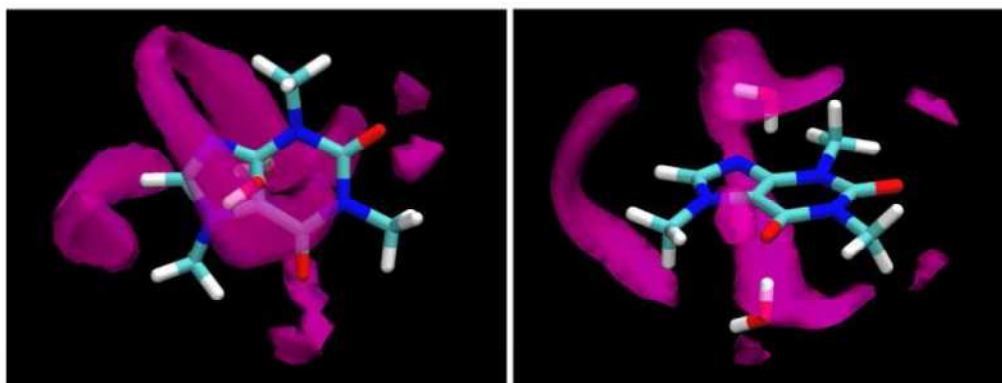


Figura 4: Simulação de nuvem de moléculas de água em interação com cafeína. Tavagnacco et al. (2011)

A simulação computacional realizada pelo grupo de pesquisa de Tavagnacco et al. (2011) mostra a agregação da cafeína em solução aquosa, Figura 4. A partir desta pode-se observar a organização da nuvem de moléculas de água interagindo com as regiões onde se encontram os grupos polares, como o átomo de nitrogênio na posição 9 (N9) ou os dois grupos carbonila, mostrando como as moléculas de água se comportam na presença de um soluto com diversos grupos polares.

4.3 Aquafotomica

Em estudo desenvolvido por Tsenkova (2009), soluções de proteína, soluções de água biomoleculares, suspensões de pequenas partículas em água e vários sistemas biológicos, tais como células, fluidos corporais de animais e tecidos do corpo foram analisados com espectroscopia NIR. Nestes estudos, as bandas absorvância da água foram observadas em várias

situações (isto é, na presença de partículas manométricas de vários tamanhos, de moléculas que não absorvem a radiação na faixa do infravermelho e com variações temperatura, e intensidade luminosa etc). Espectros foram adquiridos e analisados a partir de métodos multivariados e como resultado, foram identificadas inúmeras bandas de absorbância da água. Muitas dessas bandas estavam em concordância com os relatos e cálculos de bandas de água já publicados sobre IR região. (GOWEN et al., 2009)

Estudos realizados (HERZBERG, G., 1945; LUCK, W. A. P., 1998; BUJIS, K. & CHOPPIN, G. R., 1963; CHAPLIN, M. F., 1999; FORNÉS, V. & CHAUSSIDON, J., 1978; DOSTER, W., 1986) relataram diversas bandas de absorbância da água. Uma análise mais aprofundada na região vis-NIR mostrou que esses picos ocorreram sistematicamente ao analisar amostras de origem diferente. A única coisa que todas estas amostras tinham em comum era a constituinte água. Estas bandas de absorbância NIR foram atribuídas a um certo número de conformações moleculares de água.

Tabela 1: intensidades relativas de bandas de estiramento C-H para região infravermelho e sobretons de Infravermelho Próximo. Fonte: Water Journal

Band	Wavelength region	Relative Intensity	Pathlength	Optical Density
Fundamental band (n)	3380 - 3510 nm	100	0.01 mm	2
1st overtone (2n)	1690 - 1755 nm	1	1.0 mm	2
2nd over tone (3n)	1127 - 1170 nm	0.1	1.0 cm	2
3rd overtone (4n)	845 - 878 nm	0.01	10.0 cm	2
4th overtone (5n)	690 - 780 nm	0.005	10.0 cm	1

Em 2003, estudos (ROBERTSON, W. H. et al., 2003; WEBER, J. M. et al., 2000) na região do IR (Infravermelho) em várias moléculas individuais ou pequeno número de moléculas de água, revelaram diversas bandas, o que levou a localização de seus sobretons na região do NIR (tabela 1). Além disso foram estudadas as mudanças espectrais na região do NIR para a água sob

diversas perturbações, tais como temperatura e concentração de sacarose (MAEDA, H. et al., 1995; SASIC, S. et al., 2002; CZARNIK-MATUSEWICZ, B. et al., 1999; GIANGIACOMO, R. et al., 2009). Estes resultados para vários sistemas biológicos, soluções aquosas e faixas de frequência, juntamente com os obtidos em estudos mais atuais como o de Kovacs (2015) que aborda o monitoramento da qualidade da água utilizando seus padrões espectrais como marcadores, esses resultados devem ser organizados de forma sistematizada para a construção de um banco de dados de bandas de absorvância de água, chamado aquafotome (do inglês: *aquaphotome*).

As bandas de absorção da água podem ser identificadas de duas formas: experimentalmente e por cálculos das frequências dos sobretons das bandas de água já atribuídas. A identificação experimental de bandas de água na faixa de comprimento de onda NIR depende de análise multivariada de dados. Estas podem ser relacionadas com as bandas de sobretons e calculados a partir das bandas fundamentais observadas. (TSENKOVA, 2009; GOWEN, 2011).

Numa série de experiências, espectros NIR foram obtidos, e os modelos de regressão das respectivas perturbações foram analisadas para elucidar a existência de bandas de água comuns em vários sistemas biológicos, tais como células cultivadas sob estresse oxidativo ou na presença de hormônios ambientais e fluidos corporais de vacas saudáveis ou acometidas de uma enfermidade. (TSENKOVA, 2009) A partir destes resultados, descobriu-se que as coordenadas da matriz aquática (WAMACs, do inglês "Water Matrix Coordinates") combinadas com seus respectivos valores dos padrões espectrais da água (WASP, do inglês: "Water Spectral Pattern"), levantam de forma consistente dados de sistemas biológicos específicos em modelos de regressão, prevendo as perturbações investigadas. (TSENKOVA, 2009)

4.3.1 Aquagrama

Aquagramas (do inglês: "Aquagram") são um tipo de gráfico (ou "*star chart*") que foi concebido para visualizar os padrões espectrais da água (WASP) dividindo-os em diferentes estados vibracionais. Os aquagramas exibem os valores de absorção de água normalizados para várias bandas nos

eixos provenientes do centro do gráfico. Os valores de absorvância das coordenadas da matriz aquática (WAMACs) são colocados nos respectivos eixos radiais, indicando a organização das moléculas de água.

A figura 5 demonstra a construção de um aquagrama realizada por Chatani em 2014, para identificar o comprimento de onda relacionado a formação da fibrina. Quatorze bandas de absorção características de água foram usadas como eixos, estas bandas englobam a maioria das variadas estruturas da água (WAMACS). Os valores do experimentais mostrados no aquagrama estão de acordo com os calculados pela seguinte equação. (CHATANI, 2014).

$$A'_\lambda = (A_\lambda - \mu_\lambda) / \sigma_\lambda$$

Onde A'_λ , μ_λ e σ_λ são, respectivamente, as absorvâncias depois do prétratamento por correção multiplicativa do sinal (MSC); o centro do espectro e desvio padrão de todo o espectro, no comprimento de onda λ . (CHATANI, 2014).

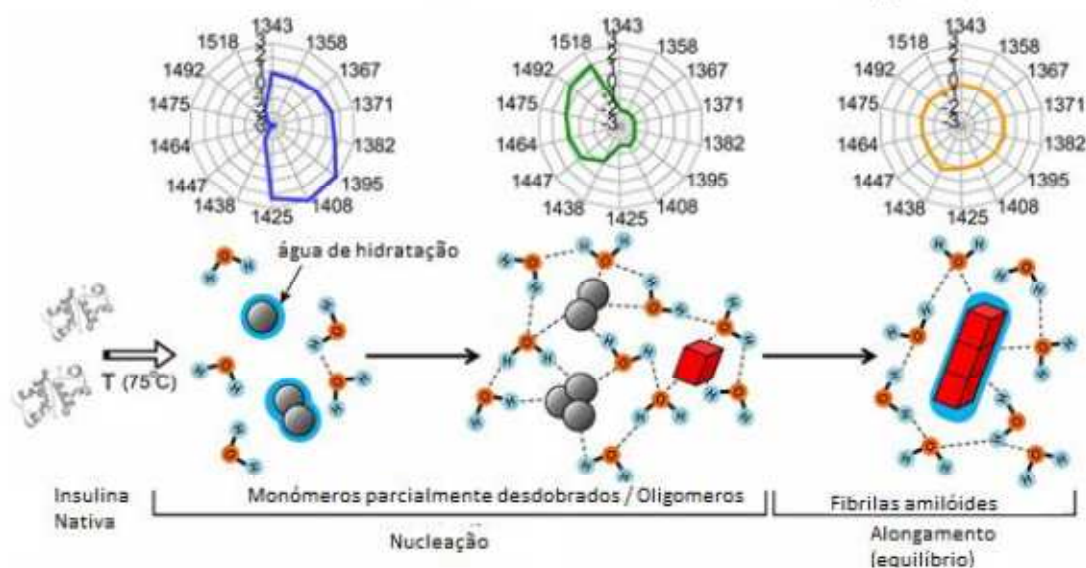


Figura 5: Etapas para a construção de aquagramas relacionados a transformação da estrutura da água durante a formação da fibrina.
Fonte: Journal One doi:10.1371/journal.pone.0101997.g006

4.3.2 O uso da aquafotomica

Dos 39 artigos encontrados nos bancos de dados utilizados neste trabalho (periódicos CAPES, Science Direct, Publimed, Scopus), cinco foram selecionados por tratarem principalmente de análises de contaminantes em meio aquoso e monitoramento de qualidade de águas utilizando a aquafotomica e como potencialidade para aplicação na área ambiental. Dentre estes está o trabalho desenvolvido por Yingtian Hu, Xiaoping Wang, em 2017, onde foi estudado o espectro de absorção de oito contaminantes típicos que podem ser agrupados em três categorias de acordo com os grupos funcionais e foi proposto quatro parâmetros integrais em diferentes faixas espectrais relativas aos diferentes grupos funcionais e a turbidez. Os parâmetros alternativos propostos podem ser utilizados para analisar e prever rapidamente as categorias de contaminantes com base nas suas alterações sem qualquer modelagem. Utilizou-se um algoritmo de particionamento dinâmico baseado no espectro derivativo de quarta ordem para adaptar o movimento das bandas de absorção. Utilizando este método, realizou-se monitoramento a longo prazo das águas residuais de uma fábrica de tingimento e impressão. A rápida análise e o potencial de previsão das diferentes categorias de contaminantes foi demonstrada. Os resultados da previsão podem fornecer evidências para análise detalhada e monitoramento da qualidade da água.

Outro trabalho referente ao tema foi publicado por Kovacs et al, no ano de 2016, tendo como objetivo introduzir uma abordagem complementar para a triagem de água, observando as mudanças do sistema molecular da água a partir do uso da aquafotomica e o método de Gráfico de Controle de Qualidade (QCC). A aplicabilidade do conceito proposto foi avaliada através do monitoramento das mudanças estruturais da molécula de água em soluções aquosas de ácido, açúcar e sal em um nível de concentração milimolar, além de águas subterrâneas. Os resultados mostraram o potencial da abordagem proposta para usar o monitoramento do padrão espectral de água como biomarcador de qualidade da água.

Em 2015, Gowen et al realizaram um trabalho do tipo investigativo que avaliou a viabilidade do NIRS para detectar contaminantes em água utilizando apenas soluções de sal como sistemas-modelo. Eles verificaram que a primeira

região de sobretom do espectro de infravermelho próximo (1300-1600nm) foi mais eficaz para a predição da concentração de sal em soluções aquosas. Os modelos de calibração construídos utilizando esta faixa de comprimentos de onda e empregando o pré-tratamento espectral de correção multiplicativa estendida (MSE) resultaram em erro quadrático médio de valores de predição variando de 0,004 a 0,01 mol.L⁻¹. O limite de detecção (LOD) foi estimado em cerca de 0,1% (massa/massa) ou 1000 ppm. No âmbito da aquafotomica, foi possível examinar o efeito de diferentes sais sobre os espectros NIR da água na região do primeiro sobretom.

Em 2014, Munćan et. Al utilizaram a análise multivariada para a classificação de águas minerais comerciais com base em seus espectros de infravermelho próximo (NIR). Além disso, a aquafotomica foi aplicada como uma nova abordagem para a interpretação de espectros do infravermelho próximo da água, elucidando sobre a organização das moléculas de água em cada uma dessas amostras (águas minerais). Abordando as diferentes conformações que as moléculas de água assumem, como clusters por exemplo.

Em 2013, Gowen, Amigo e Tsenkova, investigaram a aplicação da técnica de Multivariate Curve Resolution – Alternate Least Squares (MCR-ALS) para caracterizar os efeitos de várias perturbações no espectro NIR da água em termos de ligação de hidrogénio. Para o efeito, as perturbações criadas pela alteração da temperatura e a adição de quatro soluções de diferentes forças iônicas e a acidez de Lewis (NaCl, KCl, MgCl₂ e AlCl₃, com concentrações que variam de 0,2 a 1 mol.L⁻¹, com intervalo de 0,2 mol.L⁻¹) foram estudados. Os espectros de transmissão de todas as soluções de sal e água pura foram obtidos a temperaturas variando 28-45°C. As soluções de sal apresentaram tendências semelhantes em relação à perturbação da temperatura, enquanto que a localização dos picos dos componentes puros gerados pelo MCR-ALS variou de acordo com a força iônica do sal utilizado.

A utilização da abordagem adotada pela aquafotomica na análise e monitoramento de água, tem alcançado resultados relevantes na área ambiental, quando testada frente a diversas situações em que a água foi perturbada por efeitos químicos e físicos. Conseguindo resultados cada vez mais precisos e detalhados quando associados a métodos de análise

multivariada dos dados, conseguindo extrair informações referentes a conformações das moléculas de água (cluster, dímeros, etc), que é um tema pouco explorado nas publicações do banco de periódicos da Capes.

Um dos objetivos da aquafotomica e dos pesquisadores que trabalham com essa nova metodologia é os registros dos espectros da água a partir de técnicas como a espectroscopia na região infravermelho próximo, ultravioleta e visível, além de trabalhos com espectroscopia Raman e Terahertz que permitem medidas não invasivas e multivariadas de amostras de água em diferentes temperaturas, pH e concentrações de substâncias como cloreto de sódio e contaminantes orgânicos, por exemplo. Tornando o banco de dados aquaphotome mais completo ao fim de cada pesquisa.

A aplicação dos conceitos da aquafotomica são recentes e tiveram suas primeiras publicações a partir do ano de 2008, contando com um número crescente de publicações que duplicou nos últimos três anos. Mostrando o potencial dessa metodologia para reduzir custos com reagentes, perda de amostra em etapa de preparações e tornar a análise de água mais prática, com potencial para monitoramento em tempo real de leitos aquáticos.

4.4 Espectroscopia na região do infravermelho

A espectroscopia estuda a interação da radiação eletromagnética com a matéria, sendo um dos seus principais objetivos a determinação dos níveis de energia de átomos ou moléculas. Quase todos os compostos que tenham ligações covalentes, sejam orgânicos ou inorgânicos, absorvem frequências de radiação eletromagnética na região do infravermelho.

A absorção de radiação IR (região espectral do infravermelho) é limitada principalmente a espécies moleculares que possuem grandes diferenças de energia entre diversos estados vibracionais e rotacionais. Para absorver a radiação IR, uma molécula deve sofrer uma variação no momento dipolo durante seu movimento rotacional ou vibracional. Apenas sob estas circunstâncias o campo elétrico alternado da radiação pode interagir com a molécula e causar variações na amplitude de um de seus movimentos. As vibrações moleculares podem ser classificadas em deformações axiais e deformações angulares (figura 6). Uma vibração de deformação axial é um movimento rítmico ao longo do eixo da ligação que faz com que a distância interatômica aumente e diminua alternadamente. As vibrações de deformação angular correspondem às variações ritmadas de ligações que têm um átomo em comum ou o movimento de um grupo de átomos em relação ao resto da molécula sem que as posições relativas dos átomos do grupo se alterem (SILVERSTEIN, 2007; PAIVA, 2012).

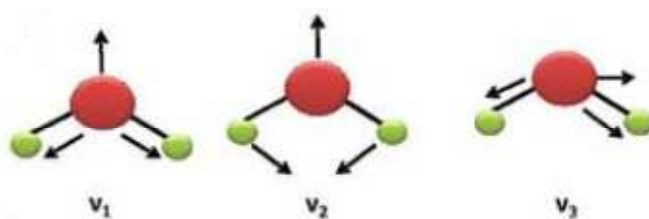


Figura 6: Os três modos de vibração da molécula de água, sendo o módulo v₂ relacionado principalmente a deformação angular e os módulos v₁ e v₃ de deformação axial.

Fonte: Sergio Kogikoski, et al 2012

No processo de absorção são absorvidas as frequências de radiação no infravermelho que equivalem às frequências vibracionais naturais da

molécula em questão, e a energia absorvida serve para aumentar a amplitude dos movimentos vibracionais das ligações na molécula. Nem todas as ligações em uma molécula são capazes de absorver energia no infravermelho (BARBOSA, 2007; SALA, 2008; PAVIA et al, 2010).

Considerando o modelo clássico, uma molécula diatômica pode ser comparada com um sistema formado por esferas de massa m_1 e m_2 representando os átomos, unidas por uma ligação com propriedades parecidas com as de uma mola, cuja constante de força seja f . Analisado inicialmente o caso mais simples de uma molécula diatômica A-B, vemos que esta representa três graus de liberdade translacionais e dois rotacionais (nenhuma energia é detectada no processo envolvendo a rotação em torno do eixo da ligação A-B). Os átomos podem também se mover um em relação ao outro, ou seja, os comprimentos das ligações A-B podem aumentar ou diminuir. Nesse caso ocorre compressão ou estiramento da ligação, processos denominados genericamente vibração ou estiramento. Essa vibração corresponde a um grau de liberdade vibracional. A energia necessária para promover o estiramento depende das massas dos átomos e da constante de força (f) da ligação. Isso pode ser entendido observando a lei de Hooke, que estabelece que, se uma mola é esticada de determinada quantidade x , existirá uma força oposta de restauração (F), que é proporcional ao deslocamento x , conforme equação (I) (BARBOSA, 2007; SILVERSTEIN, 2007).

$$F = -fx \quad (I)$$

Na realidade, o processo de absorção de radiação de um nível de energia para outro é quantizado, ocorrendo apenas em determinados comprimentos de onda, de modo que a energia da transição é dada pela equação $\Delta E = hv$ (h = constante de Planck; v = frequência da radiação) (BARBOSA, 2007; SILVERSTEIN, 2007; HOLLER, 2009)

A região denominada de infravermelho (IR) pode ser dividida em 3 regiões, infravermelho próximo, médio e distante. O infravermelho próximo compreende a faixa de números de ondas entre 14290 cm^{-1} a 200 cm^{-1} . A região que apresenta número de ondas entre 4000 cm^{-1} a 400 cm^{-1} , é denominado infravermelho médio e a que compreende a faixa de 200 a 10 cm^{-1} é chamada de Infravermelho distante.

O comprimento de onda que vai 100000 a 25000cm^{-1} caracteriza a região do ultravioleta (UV) que possui energia capaz de causar transições eletrônicas em moléculas orgânicas. A ordem de grandeza da energia das radiações UV é comparável à de energia de ligações covalentes. Assim, muitos compostos orgânicos, quando irradiados com radiação UV, decompõem-se. A radiação na região do infravermelho (IR), no entanto, como possui maior comprimento de onda, apresenta menor energia. Dessa forma, esse tipo de radiação, quando interage com moléculas orgânicas, é suficiente apenas para causar alterações em modos vibracionais e rotacionais das moléculas. (BARBOSA, 2007).

Para exemplificar o uso dos espectros obtidos no infravermelho, um estudo realizado por Gallignani et al., (2008) amostras de café foram analisadas e caracterizadas pelo espectro de absorção na região do infravermelho (IR), em pastilhas de KBr, na região de 400 a 4000cm^{-1} (SILVERSTEIN & WEBSTER, 2000). O espectro de IR da cafeína pura nas situações A e B (figura 7) foram caracterizados pelas três principais bandas proeminentes e as regiões espectrais de 1500 - 1800 e de 2800 - 3000cm^{-1} .

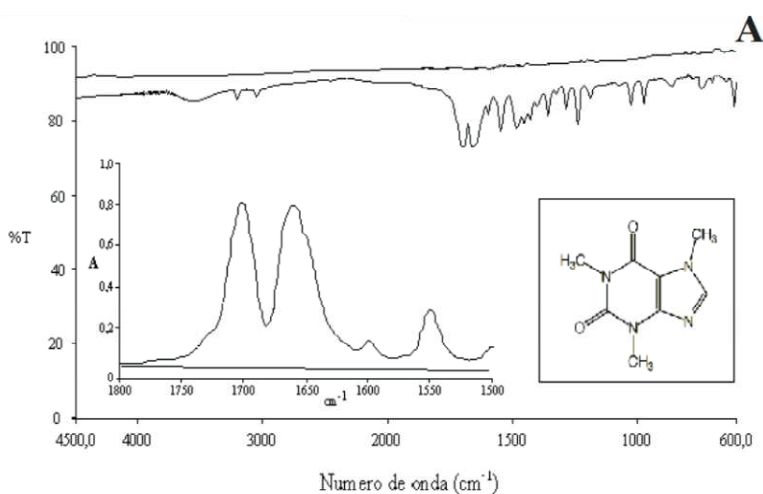


Figura 7: Espectro FTIR da transmitância da cafeína sólida em pastilhas de KBr. A figura mostra a amplitude na absorbância, na região entre 1300cm^{-1} e 1800cm^{-1} e uma representação da estrutura da cafeína.

Fonte: Gallignani et al., (2008)

Foi relatado o uso de NIRS para a previsão da concentração de metais em soluções aquosas usando espectroscopia de transmissão NIR, com limites relatados de detecção variando 10 - 40ppm (SAKUDO, et al.; 2006). Embora os

metais não absorvam radiação no NIR, a sua presença é detectável devido à interação de íons metálicos com ligações OH em água.

A aquafotomica visa explorar tais interações entre água e a radiação do UV-visNIR para extrair informações sobre o estado de sistemas aquosos. Justamente porque a água é o principal componente em vários sistemas biológicos (e muitos não biológicos). No entanto, a sua estrutura é perturbada pela presença de vários componentes tais como sais, açúcares, proteínas e outras moléculas biológicas. Com isto em mente, a aquafotomica tem como objetivo caracterizar o efeito de várias perturbações na estrutura da água usando os espectros do vis-NIR entre outras fontes de radiação como a ultravioleta. (TSENKOVA, 2009; GOWEN, 2011)

4.5 Quimiometria

A Quimiometria é uma área da química que utiliza conceitos estatísticos e matemáticos para o planejamento e otimização das condições experimentais e para extração de informações químicas provindas de dados químicos multivariados.

Os métodos espectroscópicos aliados a quimiometria tornam-se uma alternativa viável e econômica, onde modelos matemáticos multivariados correlacionam a concentração do analito com toda a faixa espectral, possibilitando a criação de modelos de predições estatisticamente confiáveis, mesmo na presença de interferentes. A utilização destas técnicas permite uma redução significativa no custo, tempo e resíduos produzidos quando comparadas às técnicas cromatográficas, principalmente por necessitarem de pouco ou nenhum preparo de amostra (S. Wold & S. D. Brown; 1994).

As áreas mais importantes da quimiometria são o planejamento de experimentos, a análise exploratória de dados, a construção de modelos quantitativos de calibração e a construção de modelos qualitativos de classificação.

Na análise exploratória, os métodos mais simples e utilizados são a análises de componentes principais (PCA) e a análise de agrupamentos hierárquicos (HCA). Já em relação às técnicas de calibração multivariada (modelos quantitativos, os métodos mais utilizados são de regressão por componentes principais (PCR) e regressão por mínimos quadrados parciais

(PLS) (ROBINSONS, 2001 e VOET et al., 2000). No tocante aos modelos de classificação pode-se citar o SIMCA (Modelagem independente e flexível por analogia de classes) (SJÖSTRÖM, KOWALSKI, 1979; WOLD et al., 1983) e o LDA (Análise Independente Linear) (BELHUMEUR et al., 1997).

4.6 Métodos de Calibração Multivariada

A Calibração multivariada é possivelmente a campo da Quimiometria que tem atraído um maior interesse em aplicações da espectrofotometria na região do infravermelho (ESBENSEN, 2002). O propósito da calibração é estabelecer uma relação matemática quantitativa entre os dados de saída de um instrumento e o parâmetro físico ou químico de interesse anteriormente determinado por uma técnica independente (FERREIRA et al, 1999). O processo geral de calibração consiste basicamente em três etapas: calibração, validação e previsão. A etapa de calibração estabelece um modelo matemático, obtido relacionando-se a matriz de dados das variáveis medidas (matriz X), por exemplo, espectros no infravermelho, com a matriz de dados das propriedades de interesse determinados por um método de referência (matriz Y). Na etapa de validação, o modelo é avaliado utilizando-se ferramentas estatísticas de diagnóstico. Por fim, o modelo validado é aplicado para prever propriedades de amostras desconhecidas, que possuam a mesma natureza das amostras do conjunto de calibração, a partir do seu sinal espectral (PERALTA-ZAMORA et al, 1997; FERREIRA et al, 1999).

Os métodos de calibração multivariada mais tradicionais são o método dos mínimos quadrados clássicos (CLS) e a regressão linear múltipla (MLR).

O CLS sugere a representação do problema multivariado como um conjunto de problemas univariados, para cada um dos analitos analisados. No entanto, o procedimento de CLS tem como principal limitação a necessidade de definir a priori as concentrações de todas as espécies espectroscopicamente ativas no conjunto de calibração, o que nem sempre é possível nos problemas práticos.

O MLR é uma técnica mais simples que não propõe a decomposição do problema original em conjuntos menores de problemas de calibração univariada.

Na construção de um modelo MLR cada variável dependente, isto é cada vetor \mathbf{y}_{cal} da matriz \mathbf{Y}_{cal} , é expressa como uma combinação linear das variáveis independentes da matriz \mathbf{X}_{cal} e um vetor, \mathbf{b}_{MLR} , que contém os coeficientes de regressão e é dada pela seguinte equação:

$$\mathbf{y}_{cal} = \mathbf{X}_{cal} \mathbf{b}_{MLR} + \mathbf{e} \quad (1)$$

O vetor \mathbf{b}_{MLR} dos coeficientes lineares é estimado, na etapa de calibração, utilizando o critério dos “Mínimos Quadrados” que minimiza o vetor dos resíduos ($\mathbf{e} = \mathbf{y}_{cal} - \mathbf{X}_{cal} \mathbf{b}_{MLR}$) e é calculado por:

$$\mathbf{b}_{MLR} = \left(\mathbf{X}_{cal}^T \mathbf{X}_{cal} \right)^{-1} \mathbf{X}_{cal}^T \mathbf{y}_{cal} \quad (2)$$

Onde os índices sobrescritos -1 e T representam a inversão e transposição de uma matriz ou vetor, respectivamente.

Uma estimativa para a previsão do valor da concentração ou do parâmetro de interesse em uma amostra desconhecida pode ser obtida por:

$$\hat{\mathbf{y}}_{desc} = \mathbf{X}_{desc} \mathbf{b}_{MLR} + \mathbf{e} \quad (3)$$

A regressão linear múltipla apresenta alguns problemas que limitam sua aplicação. Um destes problemas é que o número de amostras deve ser igual ou superior ao número de variáveis, uma vez que o modelo consiste na resolução de um sistema de equações lineares simultâneas, essa condição necessita ser satisfeita caso contrário o sistema torna-se indeterminado.

O outro problema do MLR é que a matriz ($\mathbf{X}_{cal}^T \mathbf{X}_{cal}$) pode não apresentar inversa devido à alta correlação entre as variáveis (VANDEGINSTE, B. G. M., et al., 1998). Neste caso, pode ser necessário escolher um número de variáveis de calibração menor que o número de analitos, o que pode compor um procedimento demorado e tedioso. Este problema pode ser solucionado usando um algoritmo de seleção de variáveis.

Neste cenário foram propostas técnicas de calibração multivariada capazes de realizar a calibração de forma eficientemente, mesmo na presença de interferentes, e que não necessitassem do conhecimento do número e da natureza dos mesmos, como o caso da regressão em mínimos quadrados parciais (PLS). (FERREIRA et al, 1999).

4.6.1 Regressão por mínimos quadrados parciais (PLS)

O método de calibração multivariada PLS foi desenvolvido por Herman Wold em meados de 1964 (DRAPER, N.R & SMITH, H., 1998). Para a construção dos modelos, o método utiliza tanto às informações da matriz de dados espectroscópicos (matriz \mathbf{X}), como as informações da matriz de concentrações (matriz \mathbf{Y}).

A regressão por mínimos quadrados parciais estende o conceito do modelo inverso (propriedade como função da resposta instrumental) trocando as variáveis originais por um sub-conjunto truncado das variáveis latentes dos dados originais (WOLD, S., 2001; BEEB, K. R. & KOWAALAKI, B. R., 1987). Considerando um caso geral para a determinação de mais de uma espécie de interesse, as matrizes \mathbf{X}_{cal} e \mathbf{Y}_{cal} são decompostas em suas matrizes de pesos e escores (BLANCO, M. & ALCALA, M., 2006) respectivamente, como mostrado nas **Equações 4 e 5**.

$$\mathbf{X} = \mathbf{TP}^T + \mathbf{E} \quad (4)$$

$$\mathbf{Y} = \mathbf{UQ}^T + \mathbf{F} \quad (5)$$

Onde: \mathbf{T} e \mathbf{U} são as matrizes dos escores e \mathbf{P} e \mathbf{Q} são as matrizes dos pesos das matrizes \mathbf{X} e \mathbf{Y} , respectivamente. \mathbf{E} é a matriz de resíduos espectrais e \mathbf{F} é a matriz dos resíduos de concentração.

A decomposição em escores e pesos pode ser realizada empregando diferentes tipos de algoritmos. Todavia, eles produzem no final resultados similares. Um exemplo desses algoritmos é o NIPALS (do inglês, Non-linear Iterative Partial Least Squares). (GERALDI, P. & KOWALSKI, B. R., 1986)

O modelo final em PLS consiste em relacionar linearmente os escores da matriz \mathbf{X} com os escores da matriz \mathbf{Y} , de acordo com as seguintes equações:

$$\mathbf{U} = \mathbf{BT} + \mathbf{G} \quad (6)$$

$$\mathbf{Y} = \mathbf{BTQ}^T + \mathbf{H} \quad (7)$$

Onde \mathbf{B} é a matriz dos coeficientes de regressão; \mathbf{G} é a matriz de resíduos dos escores e \mathbf{H} a matriz de resíduos de concentração.

4.6.3 Algoritmo das Projeções Sucessivas

O SPA é uma técnica de seleção que parte de uma variável X_k e vai incorporando em cada interação uma nova variável com menor multicolineariedade possível, em relação às variáveis já selecionadas. Para escolher qual variável a ser incorporada em cada interação, cálculos de projeções são realizados. (ARAÚJO, et al 2001)

A variável de partida x_k , que guarda maior relação com a resposta a ser modelada e o número ideal de variáveis para construir o modelo MLR não é conhecido inicialmente. Assim os subconjuntos de variáveis gerados são avaliados com base no RMSE obtido pela comparação entre os valores previstos e os valores de referência.

Os resultados são armazenados na matriz RMSE possuem dimensões (MxK). O RMSE(M, K) está associado a cadeia com a variável de partida x_k e um total de m variáveis. A melhor cadeia de variáveis é definida pelo menor valor de RMSE.

Para realizar a seleção de amostras, o SPA é aplicado à matriz X_{cal}^t em vez de X_{cal} . Desta maneira, o SPA seleciona um subconjunto de amostras que são minimamente redundantes (pouca multicolineariedade entre as colunas selecionadas de X_{cal}^t) e ainda fornecem informações representativas para a calibração de modelo MLR. (GALVÃO, et al 2004)

5. MATERIAIS E MÉTODOS

5.1 Tipo de pesquisa

Foi utilizado como pressuposto teórico metodológico, as pesquisas bibliográficas e experimental. (GIL, 2008).

5.2 Área de estudo

A citada pesquisa na sua parte experimental, onde todas as variáveis são controladas, poderá ser aplicada para o controle de xenobióticos em águas de diversas regiões.

5.3 Amostras

Amostras são compostas por águas ultrapuras, coletadas no Instituto Nacional do Semiárido (INSA) e soluções da mesma contendo concentrações predeterminadas de cafeína.

5.4 Software

A análise por componentes principais e a construção dos modelos foi realizada utilizando-se o programa Unscrambler® versão 9.8 e MATLAB 2015.

5.5 Metodologia

A etapa experimental do trabalho visou desenvolver metodologias, baseadas no princípio da aquafotomica para a identificação e quantificação de cafeína em meio aquoso.

Tendo em vista que o espectro da água pode ser afetado por diversos fatores como temperatura, entre outros (TSENKOVA, 2009), as amostras de água pura e contendo cafeína foram caracterizadas em termos de algumas de suas propriedades físico-químicas como pH, temperatura e condutividade. Para isso, 10 amostras da água ultra-pura e 10 amostras das soluções contendo a cafeína foram analisadas, em termos de pH, temperatura e

condutividade, , utilizando um pHmetro T-1000 da fabricante TEKNA e um condutivímetro de bancada mCA-150 da MS TECNOPON.

As amostras da água contendo cafeína em baixas concentrações foram preparadas a partir de diluições em série apropriadas de uma solução estoque contendo cafeína na concentração de cafeína $1,9 \times 10^{-2}$ g/L. Para preparação desta solução, 0,0194 g de cafeína pura cedida pela farmácia manipulação Roval foi pesada e diluída em 1L de água ultrapura. Cada diluição foi realizada em triplicata.

Os espectros UV-vi-NIR de todas as amostras foram registrados, faixa de 190 a 1100nm, utilizando intervalo de 0,1 nm, em temperatura de aproximadamente 25oC, no Laboratório de Pesquisa em Ciências Ambientais (LAPECA) utilizando um espectrômetro da marca Shimadzu, modelo UV-2550.

A construção dos modelos de identificação (PCA) e quantificação da cafeína foi realizada de acordo com as seguintes etapas, que são comuns para as duas abordagens: a) Avaliação e pré-tratamento dos espectros; b) a partir das características espectrais, foram selecionados os pré-tratamentos a serem aplicados. Estas etapas foram realizadas utilizando o software The Unscrambler. Os modelos PCA e PLS também foram construídos utilizando o software The Unscrambler. Já a seleção dos conjuntos de calibração e predição, bem como a seleção de variáveis espectrais foram feitas utilizando programas escritos em linguagem Matlab.

6. RESULTADOS E DISCUSSÕES

6.1 Caracterização físico-química da água pura e das soluções contendo cafeína

De acordo com R. Tsenkova (2009) os espectros na região vis-NIR (680 a 2500 nm) da água são fortemente afetados pela temperatura e pH, deste modo torna-se necessário caracterizar as amostras em termos de pH, temperatura e condutividade. O pH médio das 10 amostras de água ultrapura analisadas foi $6,76 \pm 0,42$ enquanto a condutividade apresentada foi $1,64 \pm 0,28 \mu\text{S}/\text{cm}$ e a temperatura média das mesmas foi $21 \pm 0,49$ °C. As soluções contendo cafeína também foram caracterizadas e apresentaram um comportamento físico-químico, em relação aos parâmetros analisados, semelhantes ao da água pura. O que indica que a adição da cafeína não altera significativamente as propriedades físico-químicas estudadas (pH e condutividade) da água pura.

6.2 Modelo multivariado para identificação da cafeína em meio aquoso

Os espectros de todas as amostras foram registrados na faixa de 190 a 1100 nm (52.631 a 9090 cm^{-1}), utilizando a cubeta vazia como branco. A figura 7 mostra o espectro da água pura, nela pode-se notar uma banda proeminente e característica na região de 700 a 1100 nm (SAKUDO; BABA; IKUTA, 2012). Já a figura 8 apresenta o espectro das soluções de cafeína nas concentrações de 1×10^{-5} a 7×10^{-9} , nela pode-se notar a banda característica da cafeína na região de 245,8 a 299,8nm (DE MARIA; MOREIRA, 2007). Nestas soluções também pode-se observar a banda característica da água.

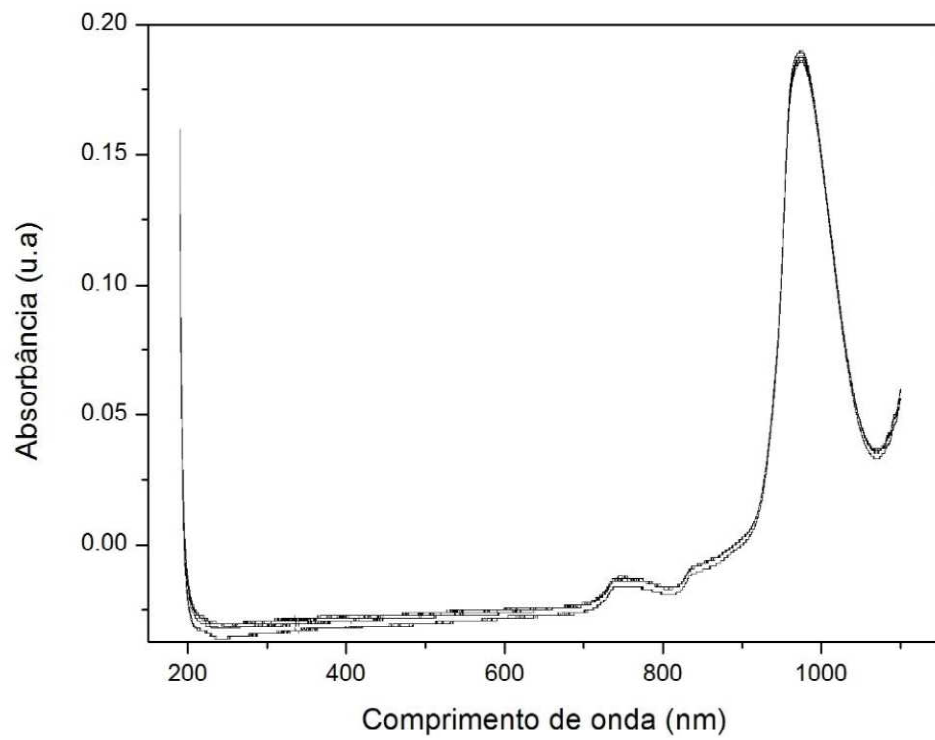


Figura 7: Espectro na região UV-vis-NIR da água pura

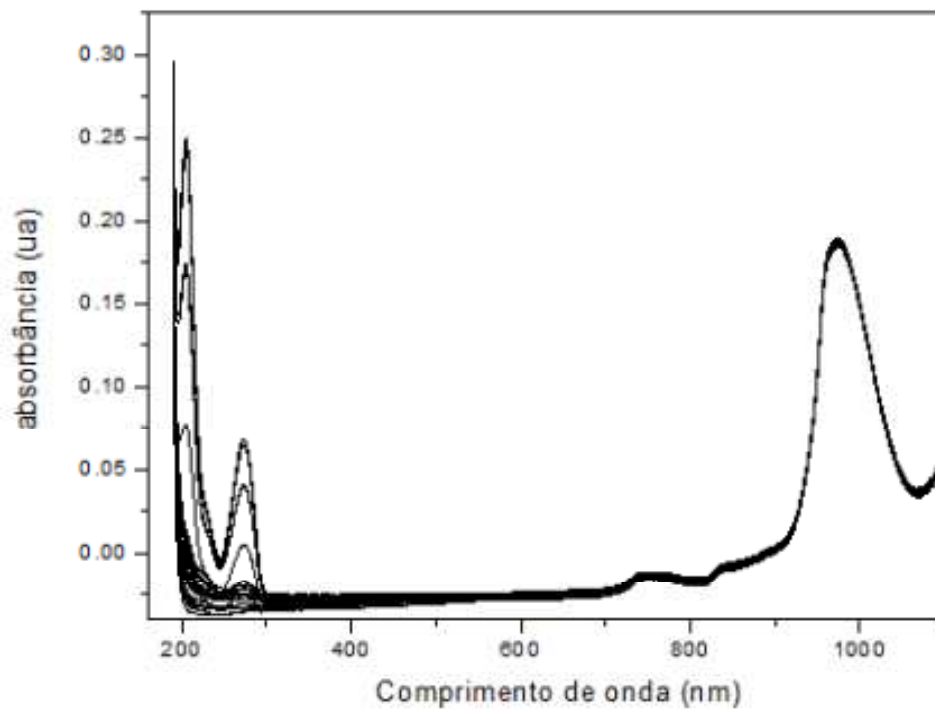


Figura 8: Espectro na região UV-vis-NIR das soluções contendo cafeína nas concentrações de 1×10^{-5} a 4×10^{-9}

Para avaliar se os espectros vis-NIR são capazes de distinguir entre amostras de água pura e amostras contendo cafeína (mesmo que em baixas concentrações) realizou-se uma análise em componentes principais (PCA). A figura 9 apresenta o gráfico de escores, obtido através da PCA. Nele pode-se observar a formação de agrupamentos naturais. Nos escores positivos de PC1 estão agrupadas as soluções com amostras mais concentradas de cafeína (1×10^{-5} a 7×10^{-6} g/L) e nos escores negativos de PC1 estão amostras com concentrações mais baixas de cafeína 1×10^{-7} a 4×10^{-9} g/L. Nos escores negativos de PC1 também estão agrupadas as amostras de água pura (destacadas na figura 9). As amostras de água pura e aquelas com baixa concentração de cafeína estão ambas no quadrante negativo de PC1.

O pico localizado na região próxima a 670 cm^{-1} , foi considerado como ruído característico de apenas uma das amostras. O espectro desta amostra particular foi excluído da matriz de dados.

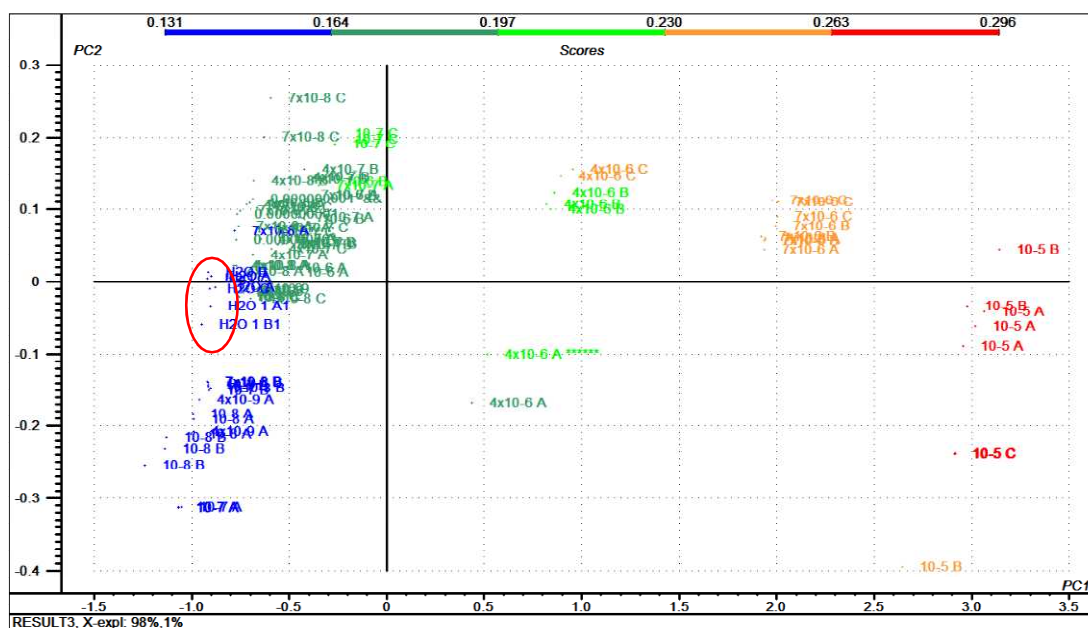


Figura 9: gráfico de escores gerado pela PCA utilizando todas as amostras.

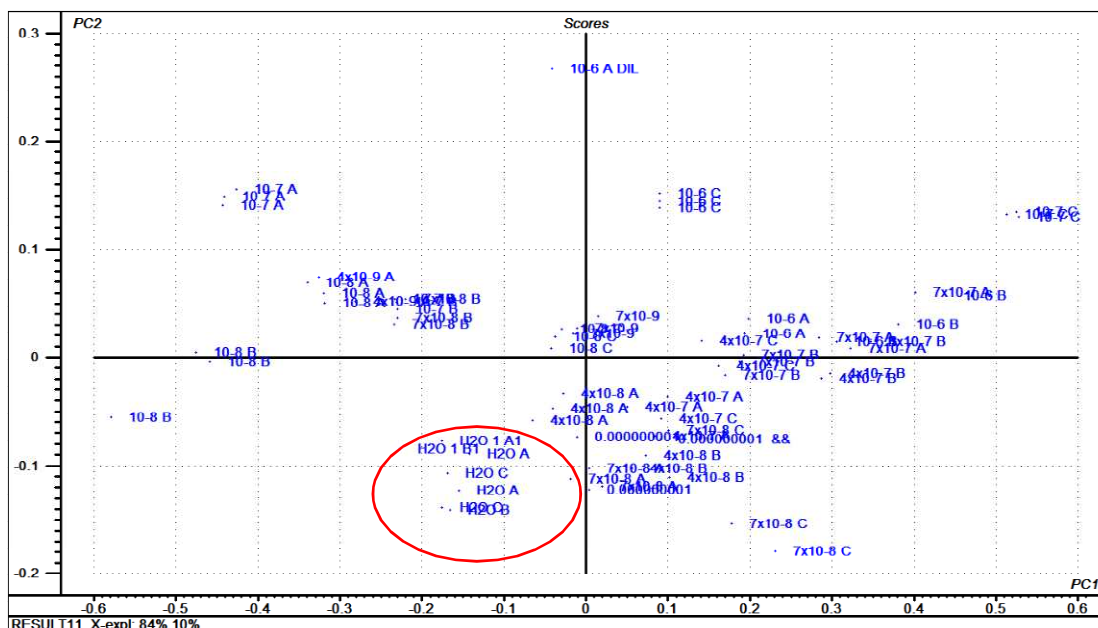


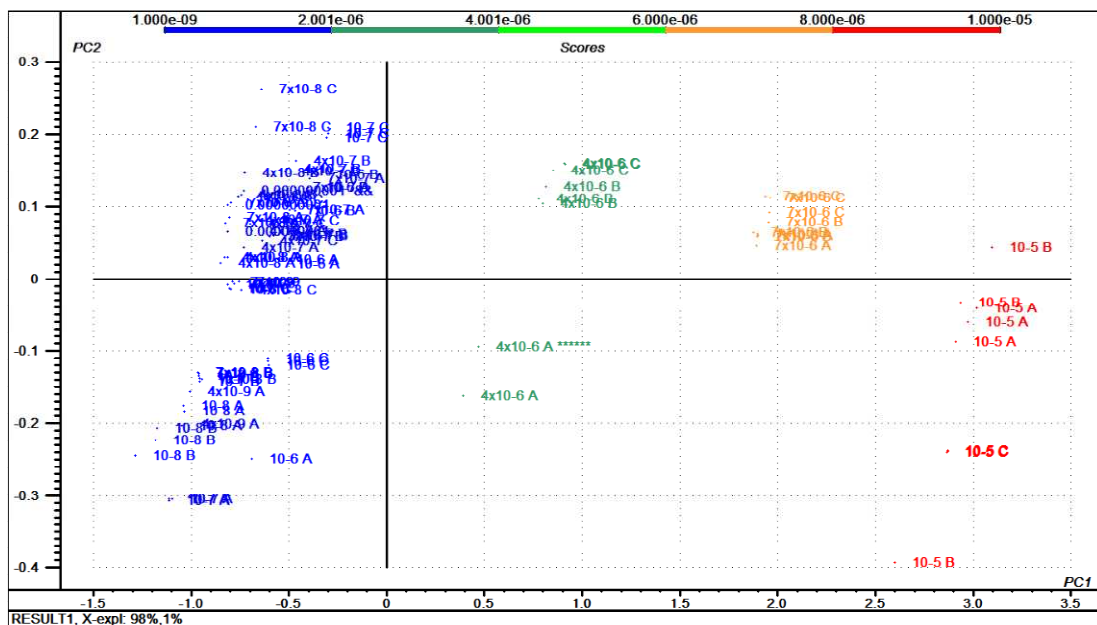
Figura 10: gráfico de escores gerado pela PCA utilizando apenas as amostras de água pura e contendo cafeína em baixa concentração.

Ao se realizar uma PCA apenas com as amostras menos diluídas e a água percebe-se uma melhor distinção entre os agrupamentos (Figura 10), de modo que a amostra de água pura forma um agrupamento bem definido no quadrante negativo esquerdo do gráfico de escores da figura 10.

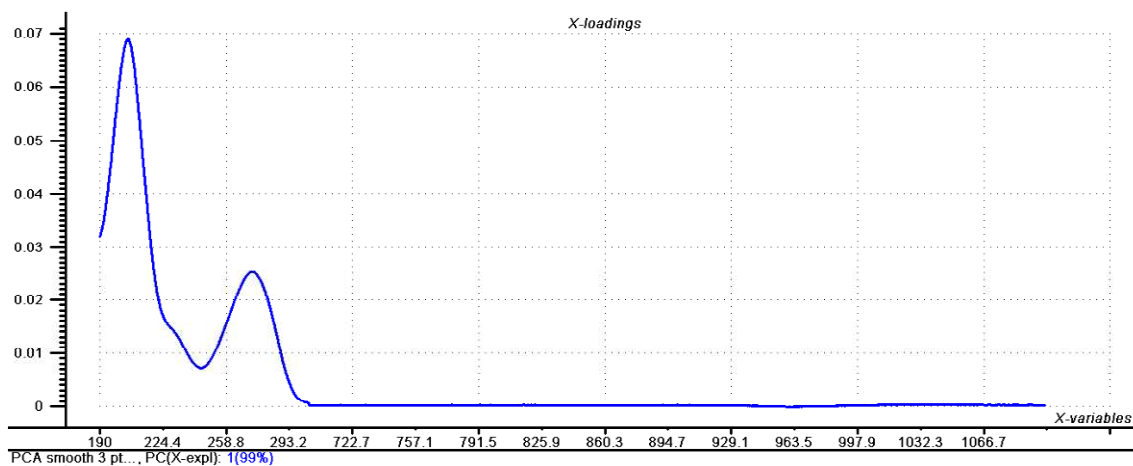
Deste modo, de acordo com os dados obtidos na PCA, decidiu-se trabalhar com dois modelos: uma para amostras com concentração cafeína de 1×10^{-5} a 7×10^{-6} g/L e outro nas concentrações de 1×10^{-7} a 4×10^{-9} g/L.

A PCA realizada apenas com amostras contendo cafeína (figura 11 A) mostra claramente os dois grupos separados por PC1, o primeiro grupo localizado nos escores negativos de PC1 (referente a amostras com baixo teor de cafeína) e o segundo grupo, localizado nos escores positivos de PC1 (referente as amostras com maior concentração de cafeína). O que também indica que a melhor estratégia é a construção de modelos independentes.

A figura 11 B representa o gráfico de pesos para PC1, como pode-se observar, este gráfico se assemelha bastante ao espectro de padrões contendo concentrações elevadas de cafeína. O que mostra que o modelo matemático foi capaz de extrair informação relativa a cafeína.



(A)



(B)

Figura 11: A) Gráfico de escores gerado pela PCA utilizando apenas as amostras contendo cafeína e B) Gráfico de pesos para PC1.

6.3 Metodologia multivariadas para quantificação da cafeína em meio aquoso.

Os modelos de calibração multivariada destinados a quantificação de cafeína em águas foram construídos utilizando o método dos mínimos quadrados parciais (PLS) e a regressão linear múltipla (MLR). Para a construção dos modelos multivariados foi retirada a região central dos espectros (300 a 699nm) e a faixa espectral de trabalho compreendeu as regiões entre 190 a 300 nm e 699 a 1100 nm, onde se encontram as informações químicas de interesse, eliminando o centro que não apresentou informações relevantes no estudo.

Os conjuntos de calibração e predição foram selecionados utilizando o algoritmo SPXy. Deste modo, o conjunto de calibração foi composto por 61 amostras e o de predição por 31 amostras, totalizando 92 amostras. O mesmo conjunto de calibração selecionado foram utilizados para a construção dos modelos PLS e MLR e sua capacidade preditiva.

Para a minimização de características espectrais não relacionadas a informação de interesse, como os ruídos espectrais, espalhamento de sinal, deslocamento de linha de base, entre outros, foram testados pré-tratamentos espectrais como suavização, correção multiplicativa do sinal e correção de linha de base. Foram construídos modelos PLS utilizando toda faixa espectral de trabalho (PLS-full) e também utilizado dos coeficientes de regressão mais informativos para o modelo calculados pelo algoritmo Jack Knife (PLS-JK). Já o modelo MLR construído foram utilizadas as variáveis mais informativas selecionadas pelo o algoritmo SPA.

O modelo PLS construído com os coeficientes de regressão selecionados pelo Jack Knife não apresentou melhora nos parâmetros de avaliação (RMSECV, correlação, r^2 , etc) em relação ao modelo construído com todas as variáveis da faixa espectral de trabalho, já que foram selecionadas a maioria das variáveis espectrais como mostra a figura 12.

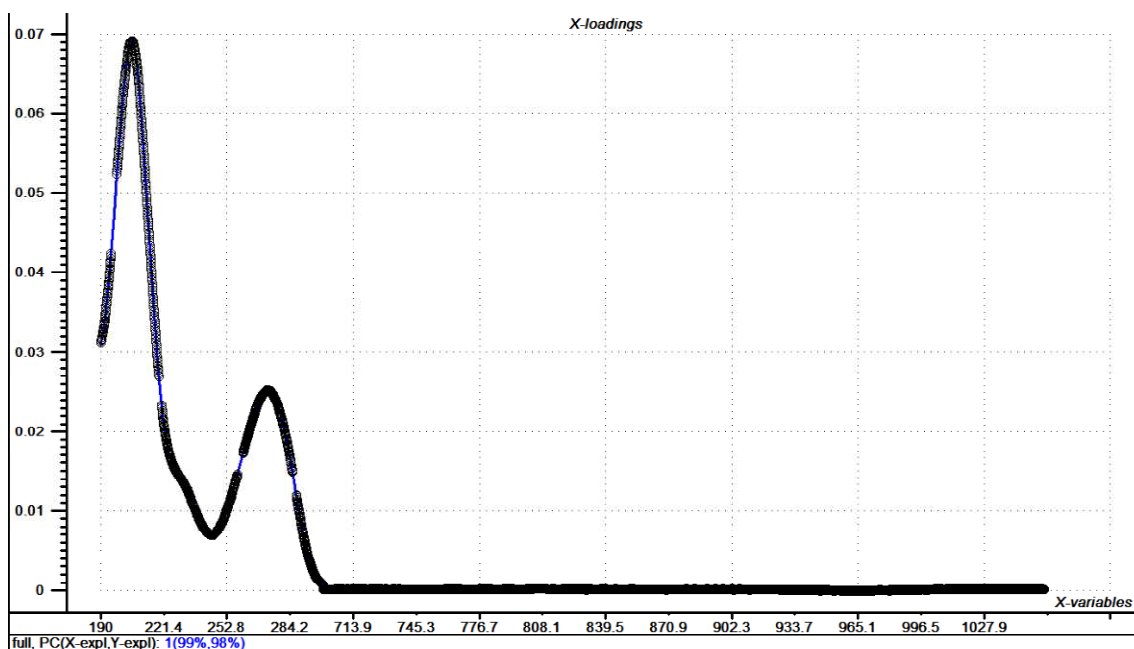


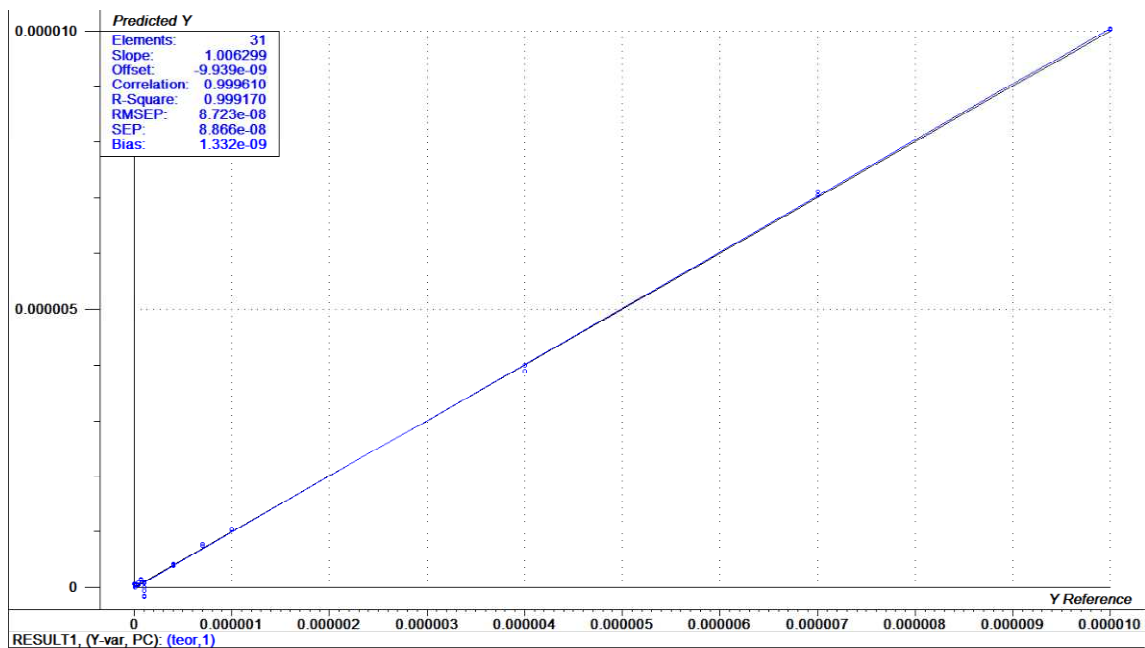
Figura 12: Coeficientes selecionados pelo algoritmo Jack Knife para a construção do modelo PLS.

Os parâmetros de avaliação do modelo PLS, utilizando vários pré-processamentos espectrais e utilizando toda faixa espectral de trabalho ou os coeficientes Jack Knife são apresentados na tabela 2. Apenas os melhores resultados são apresentados na tabela.

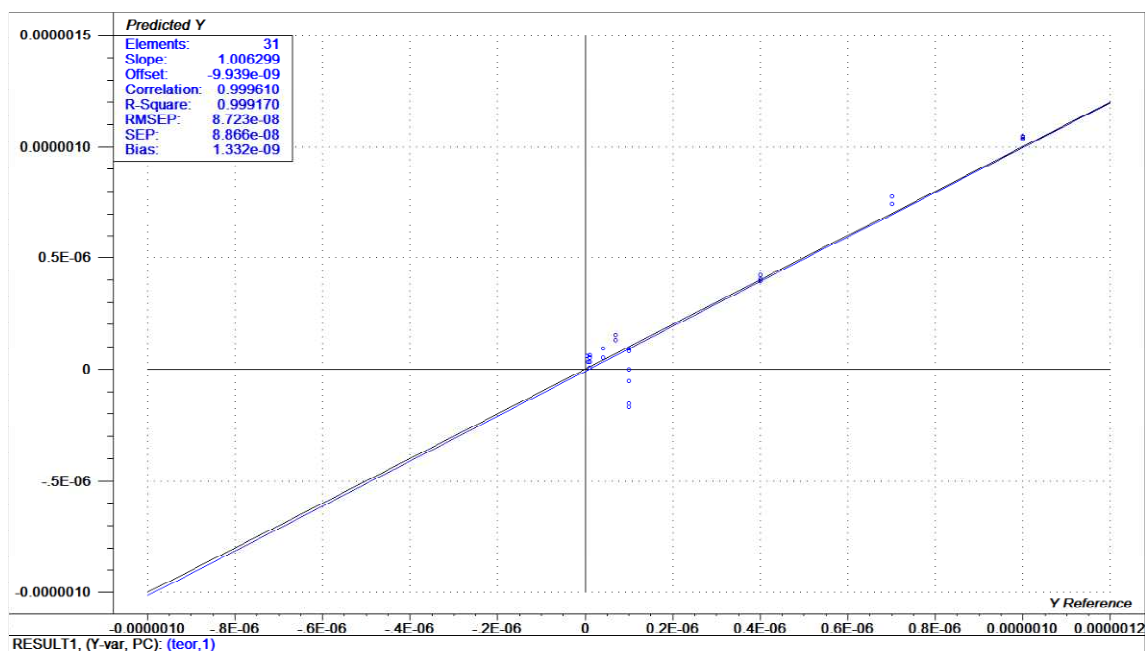
A figura 13 apresenta o gráfico dos valores medidos versus os valores previstos para o melhor modelo (usando toda faixa espectral de trabalho e o pré-processamento derivada com polinômio de 2 ordem e janela de 11 pontos). Pode-se observar uma boa concordância entre os valores medido e previstos pelo modelo. No entanto o modelo não apresenta uma boa performance preditiva para amostras com concentração inferior a $2 \times 10^{-6} \text{ g.L}^{-1}$ em cafeína.

Tabela 2: Parâmetros para avaliação dos modelos PLS

	Brutos	Suavização com janela de 3 pontos	Derivada com janela de 3 pontos		Derivada com janela de 11 pontos	
			Full	Jack Knife	Full	Jack Knife
n_{cal}	61	61	61	61	61	61
n_{pred}	31	31	31	31	31	31
VL	2	2	2	2	1	1
Slope	0.9999	0.9999	0.9984	0.9991	0.9992	0.9867
Offset	$2.11e^{-9}$	$2.11e^{-9}$	$3.84e^{-9}$	$2.09e^{-9}$	$1.91e^{-9}$	$2.29e^{-8}$
Correlação	0.9999	0.9999	0.9992	0.9995	0.9996	0.9995
r^2	0.9999	0.9999	0.9984	0.9993	0.9992	0.9990
RMSEC	$1.02e^{-7}$	$1.02e^{-7}$	$1.38e^{-7}$	$1.02e^{-7}$	$9.7e^{-8}$	$1.13e^{-7}$
SEC	$1.03e^{-7}$	$1.03e^{-7}$	$1.29e^{-7}$	$1.03e^{-7}$	$9.67e^{-8}$	$1.14e^{-7}$
RMSECV	$1.17e^{-7}$	$1.17e^{-7}$	$1.99e^{-7}$	$1.14e^{-7}$	$1.01e^{-7}$	$1.01e^{-7}$
RMSEP	$1.31e^{-7}$	$1.32e^{-7}$	$1.66e^{-7}$	$9.93e^{-8}$	$8.72e^{-8}$	$8.76e^{-8}$
Bias	$-1.702e^{-14}$	$-1.702e^{-14}$	$1.32e^{-7}$	$1.32e^{-7}$	$-4.31e^{-14}$	$-9.21e^{-14}$



(A)



(B)

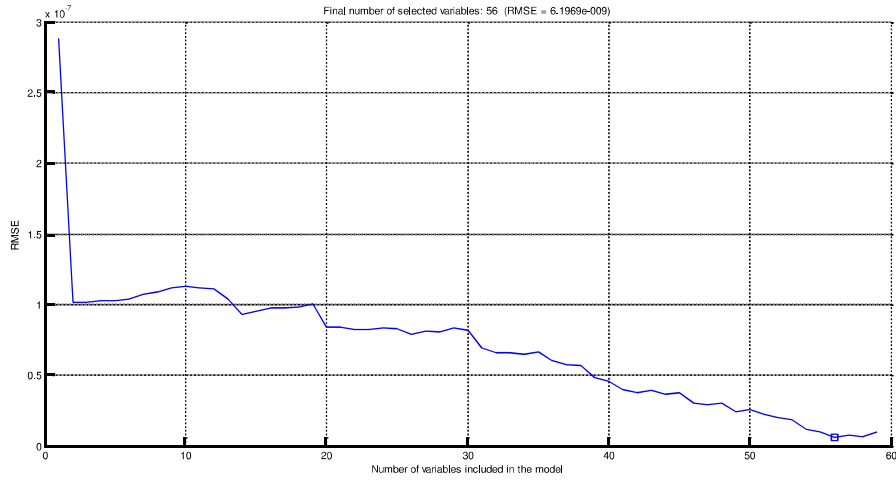
Figura 13: A) Gráfico dos valores medidos versus os valores previstos pelo modelo e B) Detalhe do gráfico para amostras com concentração inferior a $1e^{-6}$ g.L⁻¹ em cafeína.

Também foi construído um modelo MLR utilizando as variáveis espectrais selecionadas pelo algoritmo das projeções sucessivas (MLR-SPA). Utilizando tantos os dados brutos (figura 14), quanto os preprocessados com derivada de

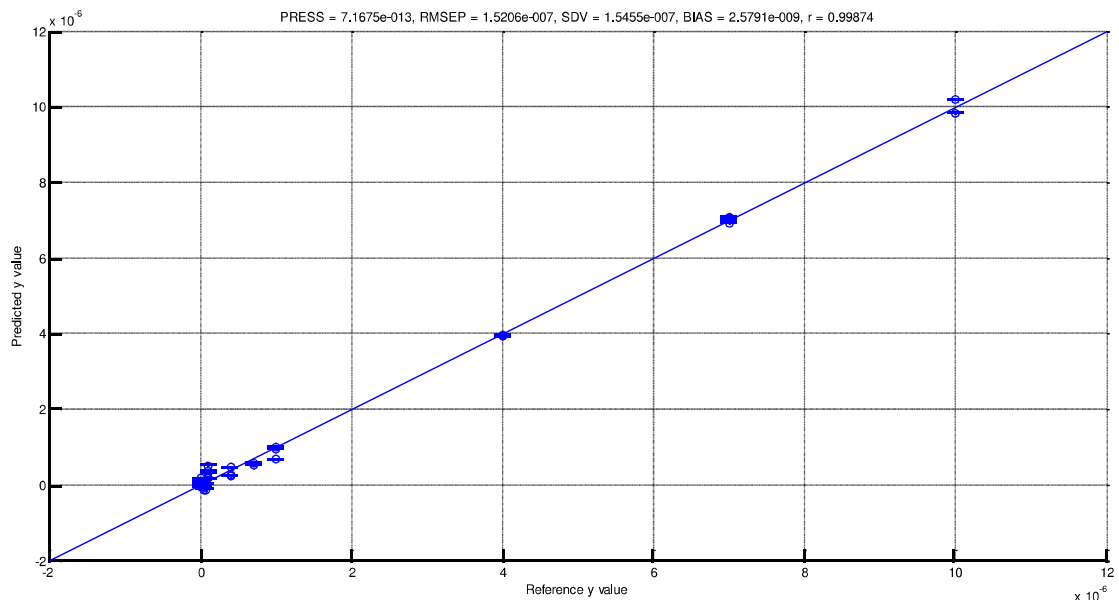
11 pontos e polinômio de segunda ordem (figura 16). Ambos os modelos construídos mostraram uma boa concordância entre os valores medidos e os previstos pelo modelo (figuras 14B e 16B). O modelo construído com os dados brutos utilizou um maior número de variáveis espectrais que os modelos construídos com os dados preprocessados, 56 e 11 respectivamente. Em relação aos valores de RMSEC e RMSEP (tabela3), o modelo construído com os dados brutos mostrou uma melhor concordância entre estes dois valores, mostrando ser um modelo mais robusto. Ambos os modelos não apresentam uma boa capacidade preditiva para concentrações inferiores $1e^{-6}$. Em relação as variáveis selecionadas pelo SPA para os dados bruto (Figura 15), em sua maioria se referem a banda da água, no entanto, também foram selecionados comprimentos de onda na região aos relativos a cafeína (~275 nm). A bandas selecionadas pelo SPA para os dados preprocessados foram 190 a 190.4; 192.1; 192.5; 192.7; 193.2; 193.6; 194.7; 195.2 e 4.698 nm, que estão localizadas nos extremos da faixa espectral de trabalho, não apresentando significado químico. A tabela 3 resume os parâmetros para avaliação dos modelos.

Tabela 3: Parâmetros para avaliação do modelo MLR-SPA

	Brutos	Preprocessados
N de variáveis selecionadas	56	13
RMSEC	$6.19e^{-9}$	$4.23e^{-7}$
RMSEP	$1.52 e^{-7}$	$4.14 e^{-7}$
PRESS	$7.17 e^{-13}$	$5.33 e^{-12}$
SDV	$1.54 e^{-7}$	$4.22 e^{-7}$
Bias	$2.5 e^{-9}$	$- 4.47e^{-9}$
R	0.9987	0.9905



(a)



(b)

Figura 14: Modelo MLR-SPA com dados brutos (a) variáveis selecionadas vs RMSEC e (b) valores medidos vs valores preditos pelo modelo.

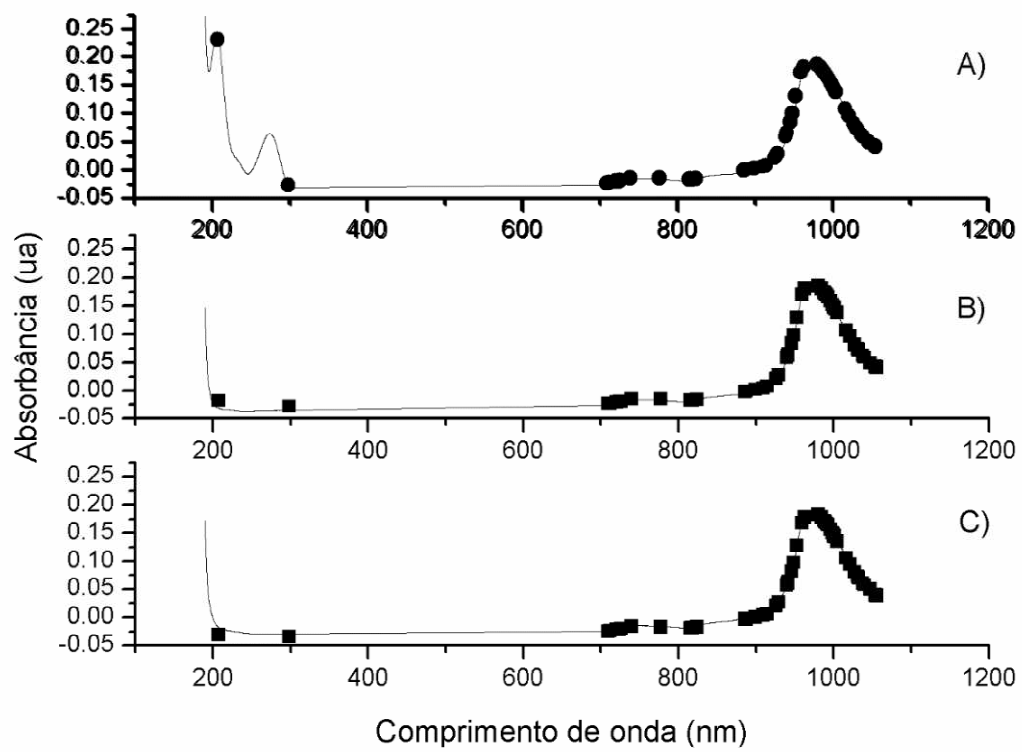
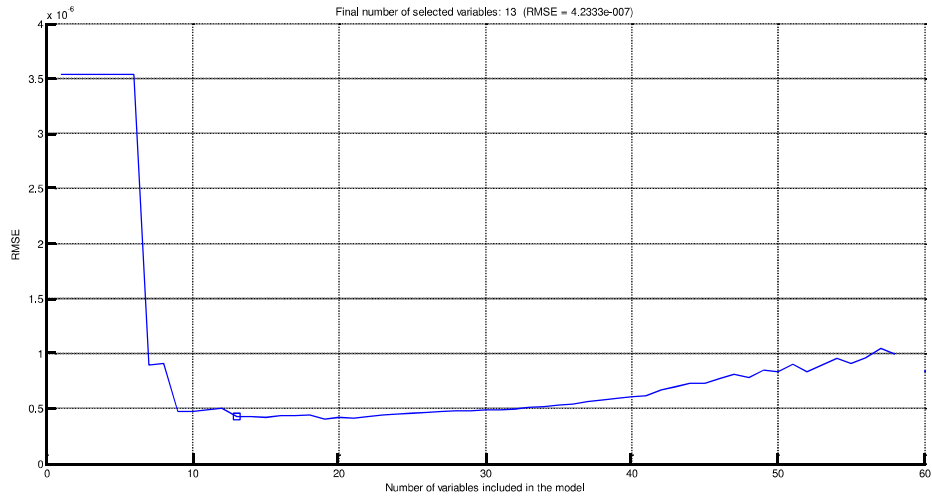
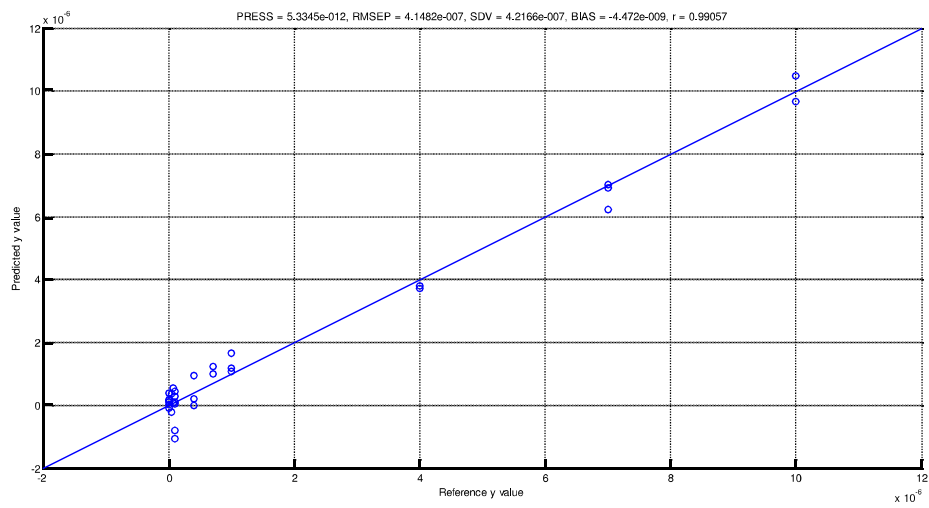


Figura 15: Variáveis selecionadas pelo SPA para os dados brutos.



(A)



(B)

Figura 16: Modelo MLR-SPA com dados preprocessados (A) variáveis selecionadas vs RMSE e B) valores medidos vs valores preditos pelo modelo.

7. CONCLUSÕES

Os artigos encontrados demonstraram que o emprego da Aquafotomica reduz custos, tempo e quantidade de reagentes a serem utilizados, análise em tempo real sem destruição de amostra e análise in-vivo que são seus pontos fortes quando comparados as técnicas já em uso mencionadas no trabalho, além da gama de aplicações na área ambiental e outras áreas que utilizam a água como fonte de dados. Apesar do seu potencial, a aquafotomica ainda é uma técnica pouco conhecida e pouco aplicada, como pode ser comprovado de acordo com a pesquisa bibliográfica realizada neste trabalho. O que é justificado pelo fato da pouca divulgação da técnica, como também aos limites de detecção que não são tão pequenos em relação às práticas já estabelecidas para a análise de água. Em contrapartida, esta técnica está apta a trabalhar abaixo dos limites estabelecidos pelos órgãos nacionais e internacionais de água, que estimam concentrações de riscos para os contaminantes encontrados em água de abastecimento e reservatórios naturais.

Neste trabalho foi utilizada a espectroscopia UV-vis-NIR aliada a métodos reconhecimento de padrões e de calibração multivariada como, PCA e o PLS e MLR, respectivamente, para construção de modelos para a identificação e quantificação de cafeína em água. A princípio, a mesma metodologia poderia ser utilizada para outros xenobióticos.

A metodologia de identificação, baseada em PCA se mostrou promissora e capaz de distinguir amostras de água daquelas contendo cafeína em concentrações inferiores a $10e^{-7}$. O que poderia ser utilizado para realizar análises de triagem quanto ao nível de cafeína na água e indiretamente seu grau de contaminação.

Em relação aos modelos de quantificação, foram utilizadas estratégias de calibração multivariada baseadas em PLS e MLR, utilizando toda faixa espectral de trabalho ou apenas a variáveis mais informativas selecionadas pelo algoritmo SPA ou Jack Knife. Também foram testados alguns métodos de preprocessamento espectral. De forma geral, as duas estratégias de calibração apresentaram performances preditivas comparáveis obtendo RMSEP's da ordem de $10e^{-8}$.

Diante do exposto pode-se concluir que as metodologias baseadas na espectroscopia UV-vis-NIR aliada a métodos de calibração multivariadas se tornam viáveis e para a identificação e quantificação de cafeína em baixas concentrações em meio aquoso.

REFERENCIAS

ARAÚJO, M. C. U., SALDANHA, T. C. B., GALVÃO, R. K. H., YONEYAMA, T., CHAME, H. C. VISANI, V. **“The Successive Projections Algorithm for variable selection in spectroscopic multicomponent”**, Chemometrics and Intelligent Laboratory Systems, 57 (2) 65-73, 2001.

BARBOSA, L.C.A. **Espectroscopia no infravermelho na caracterização de compostos orgânicos**. Viçosa: UFV; 2007.

BARCELO, D., **Analytical and bioanalytical chemistry** [1618-2642], 2004 vol.378 iss.3 pg.547-548.

BEEB, K.R., KOWASLSKI, B.R., **An introduction to multivariation calibration analysis**, Analytical Chemistry, 59:A1007, 1987.

BELHUMEUR, P. N., HESPANHA, J. P., AND KRIEGMAN, D. J. **Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. IEEE Transactions on Pattern Analysis and Machine Intelligence** ;(1997). 19(7):711-720.

BLANCO, M.; ALCALA, M.; **Content uniformity and tablet hardness testing of intact pharmaceutical tablets by near infrared spectroscopy - A contribution to process analytical technologies**, Analytica Chimica Acta, 557:353, 2006.

BOERIS, V.; ARANCIBIA, J. A.; OLIVIERI, A. C **Determination of five pesticides in juice, fruit and vegetable samples by means of liquid chromatography combined with multivariate curve resolution**. Anal. Chim. Acta 2014, 814, 23.

BRASIL, Ministério da Saúde. Portaria n.º 2.914, de 12 de Dezembro de 2011. **Dispõe sobre normas de potabilidade de água para o consumo humano.** Brasília: SVS, 2011;

BUSER H-R., POIGER T., MÜLLER M. D.; **Occurrence and environmental behavior of the chiral pharmaceutical drug ibuprofen in surface waters and in wastewater;** Environ. Sci. Technol.; 33; 2529-2535, 1999.

BUJIS, K. & CHOPPIN, G. R., **Near infrared studies of the structure of water.** I. Pure water, J. Chem. Phys. 39, 2035–2041 (1963). doi: 10.1063/1.1734579.

CHAPLIN, M. F., **A proposal for the structuring of water,** Biophys. Chem. 83(3), 211 (1999). doi: 10.1016/S0301-4622(99)00142-8.

CHATANI E, Tsuchisaka Y, Masuda Y, Tsenkova R **Water Molecular System Dynamics Associated with Amyloidogenic Nucleation as Revealed by Real Time Near Infrared Spectroscopy and Aquaphotomics.** Plus One 9(7) (2014): e101997. doi:10.1371/journal.pone.0101997.

CORNELL COOPERATIVE EXTENSIÓN, New Cork State College of Human Ecology, 1995. **Tratamiento de agua por ósmosis inversa,** Power Argentina pg.1 Dezembro de 1995.

CZARNIK-MATUSEWICZ, B., Murayama, K., Tsenkova, R. & Ozaki, Y., **Analysis of near-infrared spectra of complicated biological fluids by two-dimensional correlation spectroscopy: protein and fat concentration-dependent spectral changes of milk,** Appl. Spectrosc. 53(12), 1582 (1999).

DE JUAN, A.; NAVEA, S.; DIEWOK, J.; TAULER, R. Local rank exploratory analysis of evolving rank-deficient systems. Chemom. Intell. Lab. Syst. 2004, 70,11.

DE JUAN, A.; TAULER, R. **Multivariate curve resolution (MCR) from 2000: progress in concepts and applications** Crit. Rev. Anal. Chem. **2006**, 36, 163.

DE MOURA A. F.; **A inovação tecnológica e o avanço científico: A química em perspectiva**; Quim. Nova; 23; 851-853; 2000.

DOSTER, W., Bachleitner, A., Dunau, R., Hiebl, M. & Luscher, E.; **Thermal properties of water in myoglobin crystals and solutions at subzero temperatures**, Biophys. J. 50, 213 (1986). doi: 10.1016/S0006 3495(86)83455-5.

DRAPER, N. R.; Smith, H. **Applied Regression Analysis, 3rd ed.**; Wiley, New York, 1998.

DRUCKER, P.F. **In Technology in Western Civilization, volume 2**; Kransberg, M.; Pursell Jr., C.W., Ed.; Oxford University Press; New York; 1967, Chapter 3.

DRUMMER, O. H.; J. **Chromatographic screening techniques in systematic toxicological analysis**. Chromatogr. B. 1999, 733, 27

EBBINGHAUS, S., Seung, J.K., Heyden, M., Yu, X., Heugen, U., Gruebele, M., Leitne, D.M. & Havenit, M., **An extended dynamical hydration shell around proteins**, Proc. Natl. Acad. Sci. USA 104, 20749–20752 (2007). doi: 10.1073/pnas.0709207104.

ESBENSEN, K. H. **Multivariate Data Analysis – in practice: an introduction to multivariate analysis and experimental design**. 5 ed. Aalborg University, Esbjerg: CAMO, 2002.

FERREIRA, M. M. C.; ANTUNES, A. M.; MELGO, M. S., VOLPE, P. L. O. **Quimiometria I: calibração multivariada, um tutorial**. Química Nova. 22: 724, 1999.

FORNÉS, V. & CHAUSSIDON, J., **“An interpretation of the evolution with temperature of the $v_2 + v_3$ combination band in water”**, J. Chem. Phys. 68(10), 4667 (1978). doi: 10.1063/1.435576.

GALVÃO, R. K. H., JOSÉ, G. E., DANTAS FILHO, H. A., ARAÚJO, M. C. U., SILVA, E. C., PAIVA, H. M., SALDANHA, T. C. B., **Optimal Wavelet Filter Construction using X and Y data**, Chemometrics and Intelligent Laboratory Systems, 70 (2004) 1 – 10.

GELADI, P., KOWALSKI, B.R., **Partial least square regression: a tutorial**, *Analytica Chimica Acta*, 185:1, 1986.

GIANGIACOMO, R., Pani, P. & Barzaghi, S., **Sugars as a perturbation of the water matrix**, J. Near Infrared Spectrosc. 17(6), 329–335 (2009). doi: 10.1255/jnirs.861.

GIL, Antonio Carlos. **Como elaborar projetos de pesquisa**. 4. ed. São Paulo: Atlas, 2008.

GOLUB G.H.; VAN LOAN C.F. **Matrix computations**. 3rd ed. Baltimore (MD): The Johns Hopkins University Press; 1996.

GOWEN, A. A., R. Tsenkova, C. Esquerre, G. Downey and C. O'Donnell, **“Water matrix coefficients and absorbance patterns in mushrooms (*Agaricus bisporus*) subjected to mechanical damage using hyperspectral imaging”**, J. Near Infrared Spectrosc. 17(6), 363–371 (2009).

GOWEN, A. A., Yutaro Tsuchisaka, Colm O'Donnell, Roumiana Tsenkova. **Investigation of the Potential of Near Infrared Spectroscopy for the Detection and Quantification of Pesticides in Aqueous Solution**. American Journal of Analytical Chemistry, 2011, 2, 53-62.

GOWEN, A.A., Amigo, J.M., Tsenkova, R. **Characterisation of hydrogen bond perturbations in aqueous systems using aquaphotomics and multivariate curve resolution-alternating least squares**. *Analytica Chimica Acta*, 2013.

GOWEN, A. A., Marini, F., Tsuchisaka, Y., De Luca, S., Bevilacqua, M. O'Donnell, C., Downey, G., Tsenkova R. **On the feasibility of near infrared spectroscopy to detect contaminants in water using single salt solutions as model systems**. *Talanta*, 2015.

HAYKIN S. **Neural networks: a comprehensive foundation**. New York: Prentice Hall; 1999.

HERZBERG, G., **Molecular Spectra and Molecular Structure. II. Infrared and Raman Spectra of Polyatomic Molecules**. Van Nostrand Company, New York. USA (1945).

HEUETT, Nubia, **Journal of hazardous materials [0304-3894]**, 2015 vol.282 pg.41 -50.

HOLLER, F. JAMES; SKOOG, DOUGLAS A; CROUCH, STANLEY R. **Princípios de análise instrumental**. Tradução para o português: CÉLIO PASQUINI et al [Coord.]. 6ª ed. Porto alegre. 2009.

HUANG, C., Wikfeldt, K.T., Tokushima, T., Nordlund, D., Harada, Y., Bergmann, U., Niebuhr, M., Weiss, T. M., Horikawa, Y., Leetmaa, M., Ljungberg, M.P., Takahashi, O., Lenz, A., Ojamae, L., Lyubartsev, A.P., Shin, S., Pettersson, L.G.M. & Nilsson, A., **The inhomogeneous structure of water at ambient conditions**, *Proc. Nat. Acad. Sci.* 106, 15214 (2009). doi: 10.1073/pnas.0904743106.

HUSE, N., Wen, H., Nordlund, D., Szilagyi, E., Daranciang, D., Miller, T.A., Nilsson, A., Schoenlein, R.W. & Lindenberg, A., **Probing the hydrogen-bond**

network of water via time-resolved soft x-ray spectroscopy”, Phys. Chem. Chem. Phys. 11, 3951 (2009).

IKEHATA, A, Hydrogen bonding of water in aqueous alcohol solution: spectral analyses by an extended concept of molar absorption coefficient, Bunseki-kagaku 59, 13–21 (2010). doi: 10.2116/bunsekikagaku.59.13.

JORNAL DA UNICAMP, Campinas, 23 de setembro de 2013 a 29 de setembro de 2013 – ANO 2013 – Nº 576.

KABIR, S. R., Yokoyama, K., Mihashi, K., Kodama, T. & Suzuki, M., Hypermobile water is induced around actin filaments, Biophys. J. 85, 3154–3161 (2003). doi: 10.1016/S0006-3495(03)74733-X.

KOVACS, Zoltan, György Bázár, Mitsue Oshima, Shogo Shigeoka, Mariko Tanaka, Akane Furukawa, Airi Nagai, Manami Osawa, Yukari Itakura, Roumiana Tsenkova. Water spectral pattern as holistic marker for water quality monitoring. Talanta, vol 147, 598-608 (2015), Elsevier.

LUCK, W. A. P., “The importance of cooperativity for the properties of liquid water”, J. Mol. Struct. 448, 131–142 (1998). doi: 10.1016/S0022-2860(98)00343-3.

MAEDA, H., OZAKI, Y., TANAKA, M., HAYASHI, N. e KOJIMA, T. Near infrared spectroscopy and chemometrics studies of temperature-dependent spectral variations of water: relationship between spectral changes and hydrogen bonds. J. Near Infrared Spectrosc. 3:191–201.

MEJRI, A. M. et al. Hydration and self-association of caffeine molecules in aqueous solution: Comparative effects of sucrose and -cyclodextrin. Spectrochimica Acta. Zaghouan, Tunísia, 8 jan. 2009. p. 6-10.

MALINOWSKI, E. R.; Factor Analysis in Chemistry, 3rd ed., Wiley: New York, 2002.

DE MARIA, CARLOS A. B. e MOREIRA, Ricardo F. A. **Cafeína: revisão sobre métodos de análise.** Química Nova, Rio de Janeiro, v. 30, p.99-105, fev. 2007.

MARÇO, P. H.; POPPI, R. J.; SCARMINIO, I. S.; TAULER, R. **Investigation of the pH effect and UV radiation on kinetic degradation of anthocyanin mixtures extracted from Hibiscus acetosella.** Food Chem. 2011, 125, 1020

MASHAL N, FAUST M. HENDLER T. **The role of the right hemisphere in processing nonsalient metaphorical meanings: application of principal components analysis to fMRI data.** Neuropsychologia. 2005.

MENDIETA, J.; DÍAZ-CRUZ, M. S.; ESTEBAN, M.; TAULER, R. **Multivariate Curve Resolution: A possible Tool in the Detection of Intermediate Structures in Protein Folding.** Biophys. J. 1998, 74, 2876.

MUCCIACITO, J. C. **conceitos e aplicações do carvão ativado,** Revista e Portal Meio Filtrante, ano V, edição nº22 setembro/outubro de 2006.

MUNČAN, Jelena S. Lidija Matija, Simić-Krstić, Jovana B., Nijemčević, Srećko S., Koruga, Djuro Lj. **Discrimination of mineral waters using near-infrared spectroscopy and aquaphotomics.** Hemijska Industrija (Industria Química), 2014.

NEILSON, A H; Allard, A S; Remberger, M; **biodegradatiosn and transformations of recalcitrante compounds.** On Hutzinger, O. (ed.) **The handbook of enviromental chemistry,** Berlin: springer, 1985 v.2 part C. p29-86.

PARASTAR, H.; TAULER, R. **Multivariate curve resolution of hyphenated and multidimensional chromatographic measurements: a new insight to address current chromatographic challenges.** Anal. Chem. 2014, 86, 286.

PAVIA, DONALD. L., GARY M. LAMPMAN, GEORGE S. KRIZ E JAMES R. VYVYAN F. **Introdução à espectroscopia**. Trad. Pedro Barros. – 4ª Ed. São Paulo: Cengage Learning, 2010.

PERALTA-ZAMORA, P.; CORNEJO-PONCE, L., NAGATA, N., POPPI, R. J. **Alternativas quimiométricas para a resolução de problemas analíticos clássicos. Determinação espectrofotométrica de misturas de zircônio e háfnio**. Química Nova. 20: 469, 1997.

PIOCOS E. A., DE LA CRUZ A. A.; **SPE-HPLC-DAD of chemical indicators of human fecal contamination in water.**; JLC&RTech; 23; 1281-1291; 2000.

RAIMUNDO, C. C. M.; **Ocorrência de interferentes endócrinos e produtos farmacêuticos nas águas superficiais da bacia do rio Atibaia**. Dissertação de mestrado. Instituto de Química da Universidade Estadual de Campinas, 2005.

RAVINA M., CAMPANELLA L., KIWI J.; **Accelerated mineralization of the drug Diclofenac via Fenton reactions in a concentric photo-reactor.**; Water Res.; 36; 3553-3560, 2002.

ROBERTSON, W. H., DIKEN, E.G., PRICE, E.A., SHIN, J.W. & M.A. JOHNSON, **Spectroscopic determination of the OH solvation shell in the OH⁻ (H₂O)_n clusters**, Science 299, 1367 (2003). Doi :10.1126/science.1080695.

ROBINSON, C. **Genetic modification technology and food**. Brussels:ILSI Europe, 2001.

ROY, R., Tiller, W.A., Bell, I. & Hoover, M.R., **The structure of liquid water; novel insights from materials research; potential relevance to homeopathy**, Mat. Res. Innovat. 9(4), 93–124 (2005).

RUAN, R. R. & CHEN, P. L., **Water in Foods and Biological Materials: A Nuclear Magnetic Resonance Approach**. CRC Press, Boca Raton, Florida, USA, p. 279 (1998).

SAKUDO, Akikazu; BABA, Koichi; IKUTA, Kazuyoshi. **Analysis of Vis–NIR spectra changes to measure the inflammatory response in the nasal mucosal region of influenza A and B virus-infected patients**. Journal Of Clinical Virology. Osaka, p. 334-338. dez. 2012.

SALA, O. **Fundamento da Espectroscopia Raman e no infravermelho**. 2ed São Paulo: Eitora UNESP, 2008.

SASIC, S., Segtnan, V.H. e Ozaki, Y., **Self-modeling curve resolution study of temperature-dependent nearinfrared spectra of water and the investigation of water structure**, J. Phys. Chem. A 106, 760 (2002). doi: 10.1021/jp013436p.

SCHENEIDER, R. P.; TSUTIYA, M. T. **Membranas filtrantes para o tratamento de água, esgoto e água de reuso**. 1º edição. São Paulo: ABES – Associação Brasileira de Engenharia Sanitária Ambiental. 2001. 324 p.

SHINYASHIKI, N., Yamamoto, W., Yokoyama, A., Yoshinari, T., Yagihara, S., Kita, R., Ngai, K.L. & Capaccioli, S., **Glass transitions in aqueous solutions of protein (bovine serum albumin)**”, J. Phys. Chem. B 113(43), 14448–14456 (2009). doi: 10.1021/jp905511w.

SILVA, C. M. M. S.; FAY, E. F. Agrotóxicos e meio ambiente in: SILVA, C. M. M. S.; FAY, E. F., ed. **Biotransformação de agrotóxicos e biorremediação**. São Paulo: Jaguariuna, 2005. P. 145-192.

SILVERSTEIN, R. M.; WEBSTER, F. X. **Identificação Espectrométrica de Compostos Orgânicos**. 6ª ED. LTC editora, Rio De Janeiro, p. 77-88, 2000.

SILVERSTIEN, R. M. **Identificação espectrométrica de compostos orgânicos**. Tradução: Ricardo Bicca de Alencastro. Rio de Janeiro: LTC, 2007.

SJÖSTRÖM, M., KOWALSKI, B.R. **A comparison of five pattern recognition methods based on the classification results from six real data bases**. Anal. Chim. Acta, Amsterdam, v.112, p.11-30, 1979.

SMITH, L.I. **A tutorial on principal components analysis** [Internet]. 2002 [citado 22 maio e 2011]. Disponível em: http://www.sccg.sk/~haladova/principal_components.pdf.

TATEOKA, M. S., **Tratamento de água de abastecimento por nanofiltração e osmose inversa para remoção dos carbamatos: carbaril, carbofurano e metomil**. Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Ambiental, Florianópolis, 2014.

TAULER, R.; SMILDE, A.; KOWALSKI, B.; J. CHEMOM. 1995, 9, 31.

TERNES T. A.; **Occurrence of drugs in German sewage treatment plants and rivers.**; Water Res.; 32; 3245-3260; 1998.

TSENKOVA, R., **Aquaphotomics: water in the biological and aqueous world scrutinised with invisible light**. Spectroscopy Europe 22: 6–10, 2010.

TSENKOVA, R., **“Introduction Aquaphotomics: Dynamic Spectroscopy of Aqueous and Biological Systems Describes Peculiarities of Water,”** Journal of Near Infrared Spectroscopy, Vol. 17, No. 6, 2009, pp. 303-313.

TSENKOVA, R. et al. **Near infrared spectra of cows' milk for milk quality evaluation: disease diagnosis and pathogen identification**. J. Near Infrared Spec. 14(6), 3632370 (2006).

TAVAGNACCO, L.; Udo Schnupf, Philip E. Mason, Marie-Louise Saboungi, Attilio Cesàro, and John W. Brady. **Molecular Dynamics Simulation Studies**

of Caffeine Aggregation in Aqueous Solution. J Phys Chem B. 2011 September 22; 115(37): 10957–10966.

VANDEGINSTE, B.G.M., MASSART, D.L., BUYDENS, L.M.C., JING, S., LEWI, P.J., SMEYERS-VERBEKE, J., **Handbook of Chemometrics and Qualimetrics: Part B.** Amsterdam: Elsevier, 1998.

VOET, D.; VOET, J.; PRATT, C.W. **Fundamentos de bioquímica.** Porto Alegre: Artmed, 2000.

WALRAFEN, G., **Raman and infrared spectral investigations of water structure, in Water: A Comprehensive Treatise, Vol. 1, Ed by F. Frank.** Plenum Publishing, London, UK (1972).

WEBER, J. M., Kelley, J.A., Nielsen, S.B., Ayotte, P. & M.A. Johnson, **Isolating the spectroscopic signature of a hydration shell with the use of clusters: superoxide tetrahydrate,** Science 287, 2461 (2000). doi: 10.1126/science.287.5462.2461.

WEIGEL S., BERGER U., JENSEN E., KALLENBORN R., THORESEN H., HUHNERFUSS H.; **Determination of selected pharmaceuticals and caffeine in sewage and seawater from Tromsø/Norway with emphasis on ibuprofen and its metabolites.**; Chemosphere; 56; 583-592; 2004.

WINKEL, K., Bowron, D.T., Loerting, T., Mayer, E., & Finney, J. L., **Relaxation effects in low density amorphous ice: Two distinct structural states observed by neutron diffraction,** J. Chem. Phys. 130, 204502 (2009). doi: 10.1063/1.3139007.

WOLD, S., ALBANO, C., DUNN, W.J., ESBENSEN, K., HELLBERG, S., JOHANSSON, E., SJÖSTRÖM, M. **Pattern recognition: Finding and using regularities in multivariate data.** In: MARTENS, H., RUSSWURM, H. Food Research and Data Analysis, eds. Essex: Applied Science, 1983. p.147-188.

WOLD, S.; **Personal memories of the early PLS development; Chemometrics and Intelligent Laboratory Systems**, 58:83, 2001.

XIA, A-L.; WU, H. L.; ZHANG, Y; ZHU, S. H.; HAN, Q. J.; YU, R. Q. **A novel efficient way to estimate the chemical rank of high-way data arrays.** Anal. Chim. Acta 2007, 598, 1.

XU, C.J.; GOURVÉNEC, S.; LIANG, Y. Z.; MASSART, D.L. **Local factor analysis of rank-deficient reaction systems.** Anal. Chim. Acta 2006, 575, 1.

YE J.; JANARDAN R.; LI Q., GPCA: an efficient dimension reduction scheme for image compression and retrieval [Internet]. In: **Conference on Knowledge Discovery in Data Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. Seattle (WA); 2004.**