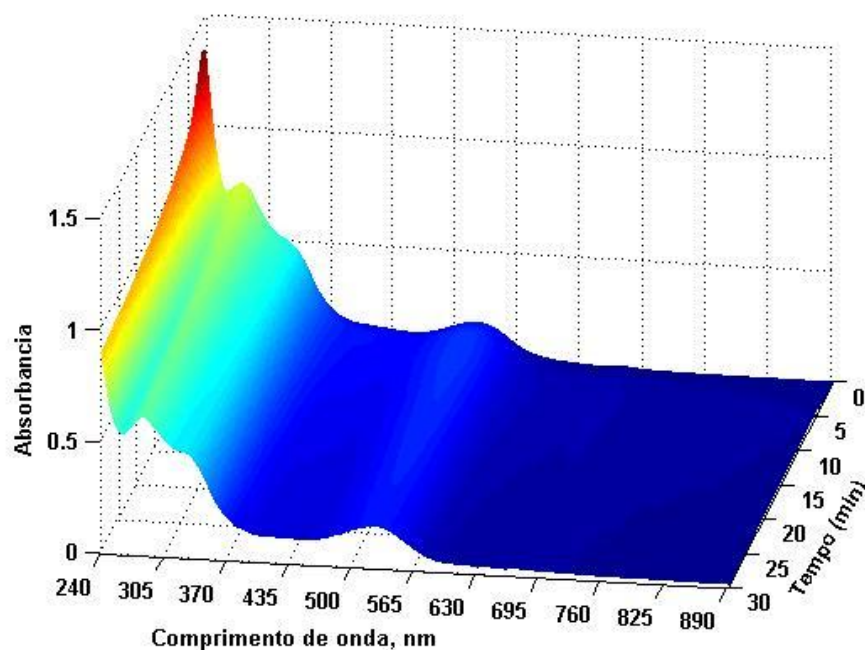




Controlo de Qualidade em Tintas e Vernizes por Espectroscopia de Infravermelho Próximo

Carla Alexandra de Rocha Ferreira

Mestrado em Química

Departamento de Química e Bioquímica

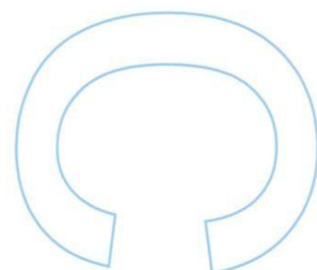
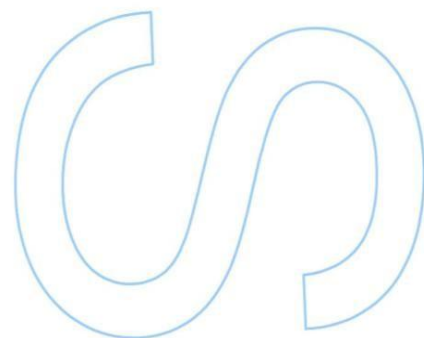
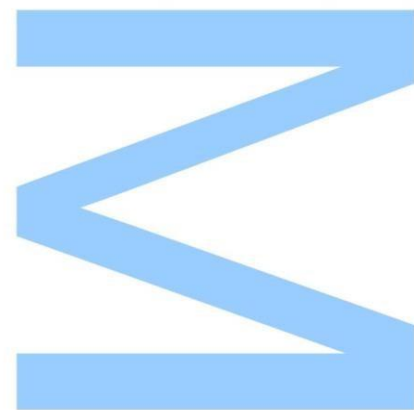
2014

Orientador

Prof. Cosme Moura, Faculdade Ciências Universidade do Porto

Coorientador

Dr.^a Etelvina Veludo, CIN-Corporação Industrial do Norte, S.A.

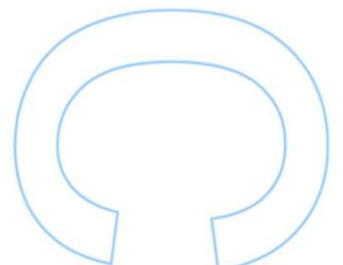
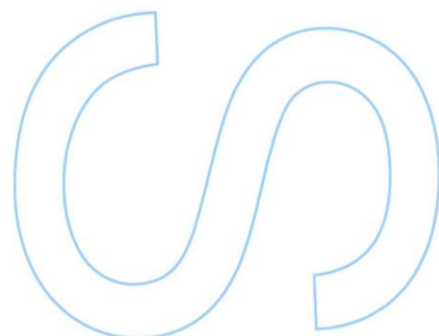
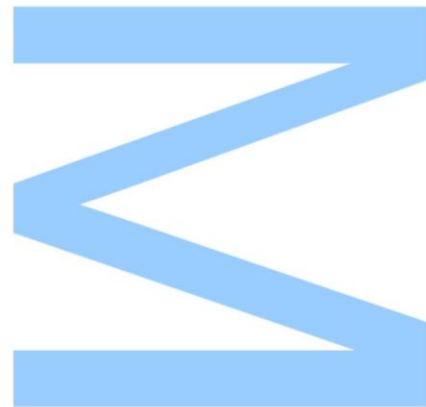




Todas as correções determinadas pelo júri, e só essas, foram efetuadas.

O Presidente do Júri,

Porto, ____/____/____



Agradecimentos

Um especial agradecimento à CIN, na pessoa do Eng.º João Machado, por ter aceite a realização deste projeto e, em particular, à Dr.^a Etelvina Veludo, por ter aceitado coordená-lo e, acima de tudo, por toda a ajuda prestada durante a sua execução.

A todo o laboratório de I&D, no qual passei momentos de boa disposição e no qual me receberam muito bem. Em especial, à Conceição Dias, que foi a pessoa mais próxima com a qual trabalhei no L.I.R.M.P, por toda a ajuda que me prestou e pelos conhecimentos que me transmitiu.

Ao Nuno, por todo o apoio incondicional que me deu, pelas horas passadas ao meu lado até à conclusão deste trabalho e pelas palavras de incentivo que me dirigiu em todo o meu percurso.

Aos meus pais, por me terem permitido viver esta experiência das mais variadas formas e por terem estado sempre do meu lado nos bons e nos maus momentos.

Por último, mas não menos importantes, aos meus amigos, um grande obrigado por terem estado sempre presentes.

Resumo

O presente trabalho insere-se no âmbito do estágio curricular integrado no Mestrado em Química da FCUP e foi desenvolvido na empresa CIN – Corporação Industrial do Norte. O estágio consistiu na execução de um projeto na área da Química Analítica, tendo como principal objetivo a aplicação da técnica de espectroscopia de infravermelho próximo (NIR) no controlo de qualidade de vernizes, endurecedores, soluções intermédias e tintas.

A aplicação da técnica NIR à análise química exige pouca ou nenhuma preparação da amostra e um curto intervalo de tempo necessário para a obtenção de resultados, em comparação com os métodos clássicos de análise, aumentando a simplicidade dos ensaios e a consequente redução dos tempos de espera, além da possibilidade de realização de medições do processo de produção.

Com base nas vantagens apresentadas da espectroscopia de infravermelho próximo, o projeto propunha a criação e implementação de uma aplicação que permitisse realizar a correta identificação, por um método de “pass-fail”, dos diversos vernizes existentes no universo CIN, tentando para isso discriminar as diferenças químicas existentes entre amostras. O estudo de uma segunda aplicação foi iniciada neste projeto e consistia na possibilidade de determinação do teor de água em tintas de base aquosa, no entanto, esta aplicação não foi aprofundada, ficando-se pela obtenção de mais do que uma calibração, passíveis de, futuramente, serem melhoradas e otimizadas.

O passo inicial deste trabalho foi a aquisição dos espectros de infravermelho próximo dos diversos produtos em estudo, para que posteriormente fossem obtidos os modelos para as aplicações desejadas.

As ferramentas utilizadas no tratamento dos dados obtidos, com base nos dos espectros de infravermelho próximo, foram técnicas quimiométricas, associadas aos métodos de análise multivariável PCA (*Principal Component Analysis*) e PCA (*Partial Least Squares*).

Para o modelo final de identificação dos vernizes foram realizadas algumas validações, das milhares de variáveis possíveis, como também a utilização da aplicação de controlo de qualidade em paralelo com as técnicas de referência normalmente utilizadas.

Palavras-Chave: Análise multivariada; NIR; Controle de qualidade; PCA; PLS; Quimiometria

Abstract

This work falls within the scope of integrated MSc in Chemistry FCUP and was developed at the company CIN - Corporação Industrial do Norte, SA. The internship consisted of develop a project in the field of Analytical Chemistry and the main goal was using near-infrared spectroscopy (NIR) in the quality control of lacquers, hardeners, intermediate solutions and paints.

The application of NIR technique for chemical analysis requires little or no sample preparation and short time required to obtain results, compared with the classical methods of analysis, increasing the simplicity of the tests and with the consequent reduction of quality control times, plus the ability to perform measurements of the production process.

Based on the presented advantages of near infrared spectroscopy, the project provides for the creation and implementation of an application that would allow carry proper identification, by a method of "pass-fail", of many of the existing CIN varnishes, by trying to discriminate the chemical differences between samples. The study of a second application was started in this project and consisted in the possibility of determining the water content in water-based paints; however, this application has not been optimized.

The initial step of the study was the acquisition of near-infrared spectra of the various products that are intended to study, so one can later obtain models for the desired applications.

Chemometric techniques were used as tools for the treatment of data obtained (near-infrared spectra), namely methods of multivariate analysis PCA (*Principal Component Analysis*) and PCA (*Partial Least Squares*).

Some validations were performed in the final model for varnishes identification, out of thousands of possible variables, as well as applying this quality control tool in parallel with the commonly used techniques.

Keywords: Multivariate Analysis; NIR; Quality control; PCA; PLS; Chemometrics

Índice

Agradecimentos.....	IV
Resumo.....	V
Abstract	VI
Índice.....	VIII
Lista de Figuras	XI
Lista de Abreviaturas.....	XIII
1. Introdução.....	1
1.1. Objetivos e Programa de Trabalho Realizado	1
1.2. Entidade de Acolhimento.....	2
2. Tintas e Vernizes	4
2.1. Definições.....	4
2.1.1. Características Principais dos Constituintes.....	4
2.1.2. Classificação de Tintas e Vernizes.....	6
3. Espectroscopia de Infravermelho Próximo	7
3.1. Referências Históricas	7
3.2. Princípios Teóricos	8
3.3. Instrumentação – Espectrofotómetro de Infravermelho Próximo com Transformada de Fourier	12
3.4. Fatores que Afetam a Aquisição dos Espectros NIR	16
3.5. Vantagens e Desvantagens da Técnica NIR.....	17
4. Quimiometria.....	19
4.1. Análise dos Componentes Principais	20
4.2. Regressão dos Mínimos Quadrados Parciais.....	22
4.3. Tratamento Prévio de Dados.....	24
4.3.1. Correção Multiplicativa de Sinal (MSC).....	25
4.3.2. Primeira Derivada	26
4.3.3. Normalização.....	27
4.3.4. Centralização da Média.....	28
4.4. Estratégia para a Construção das Calibrações	28
4.5. Validação da Calibração	29
5. Equipamentos e Materiais.....	32
6. Resultados e Discussão.....	34
6.1. Identificação de Vernizes.....	34
6.1.1. Validação do Modelo NIR Qualitativo	40
6.2. Quantificação do Teor de Água em Tintas.....	41

7.	Conclusão e Perspetivas de Trabalho Futuro	46
8.	Bibliografia.....	48

Lista de Tabelas

Tabela 1: Etapas do processo de construção de calibrações de um modelo NIR.

Tabela 2: Condições operacionais de funcionamento do equipamento NIR utilizado no estudo.

Tabela 3: Parâmetros utilizados nas calibrações desenvolvidas para o modelo NIR qualitativo para vernizes transparentes de base solvente.

Tabela 3: Parâmetros utilizados nas calibrações desenvolvidas para o modelo NIR qualitativo para vernizes transparentes de base aquosa.

Tabela 4: Resultado de alguns parâmetros determinados para o modelo NIR quantitativo de determinação do teor de água em tintas de base aquosa, utilizando diferentes pré-tratamentos.

Tabela 5: Parâmetros utilizados para avaliar as calibrações desenvolvidas no estudo.

Lista de Figuras

Figura 1: Espetrómetro rudimentar.

Figura 2: Modos vibracionais de alongamento ou estiramento (stretching).

Figura 3: Modos vibracionais de deformação angular (bending).

Figura 4: Principais bandas de absorção na região do infravermelho próximo e a sua localização no espectro eletromagnético.

Figura 5: Localização da radiação infravermelha no espectro eletromagnético.

Figura 6: Representação esquemática de um espectrofotómetro FT-NIR com interferómetro de Michelson.

Figura 7: Representação gráfica do somatório dos produtos dos vetores *Scores* e *Loadings*.

Figura 8: Representação de uma componente principal (CP) no caso de duas variáveis: (A) representa o vetor *Loadings* e (B) representa o vetor *Scores*. Notar que os dados estão centrados na média.

Figura 9: Representação esquemática das matrizes X e Y para o método PLS. A matriz X contém os dados espectrais, M representa o número de amostras analisadas e N os comprimentos de onda percorridos para a obtenção do espectro. A matriz Y contém os valores das concentrações, sendo as colunas referentes aos valores de referência dos vários componentes.

Figura 10: Espectros de absorvância NIR (a) antes da aplicação do pré-tratamento MSC e (b) após aplicação do pré-tratamento MSC.

Figura 11: Aumento da resolução através do uso de derivadas: sinal original (linha sólida) e aplicação da segunda derivada ao sinal original (linha pontilhada).

Figura 12: Espectrofotómetro NIRVIS da BUHLER utilizado neste estudo.

Figura 13: Espectros de absorvância originais, da região NIR dos vernizes A.

Figura 14: Espectros de absorvância, pré-tratados, da região NIR dos vernizes A.

Figura 15: Gráfico *X-PRESS* com base no qual se poderá seleccionar o número de componentes principais.

Figura 16: Gráfico *Factors* com base no qual se poderá seleccionar o número de componentes principais.

Figura 17: Gráfico que apresenta os *Clusters* formados, para cada produto, após uma análise PCA.

Figura 18: Espectros de absorvância na região NIR de todas as amostras em estudo.

Figura 19: Gráfico % água predita/ % água teórica.

Figura 20: Gráfico *Factors* com base no qual se poderá seleccionar o número de variáveis latentes.

Figura 21: Gráfico *Consistency* com base no qual se poderá verificar a qualidade da calibração.

Lista de Abreviaturas

NIR – Near-Infrared Spectroscopy – Espectroscopia de Infravermelho Próximo

FT-NIR – Fourier Transform Near-Infrared – Espetrofotómetro de Infravermelho Próximo com Transformada de Fourier

DRIFTS – Diffuse Reflectance Infrared Fourier Transform Spectrometry – Espectroscopia por Reflexão Difusa no Infravermelho com Transformada de Fourier

MLR – Multiple Linear Regression – Regressão Linear Múltipla

PCA – Principal Component Analysis – Análise dos Componentes Principais

PLS – Partial Least Squares – Regressão dos Mínimos Quadrados Parciais

CP's – Componentes Principais

SVD – Singular Value Decomposition – Decomposição do Valor Singular

R – Regression Coefficient – Coeficiente de Correlação

MSC – Multiplicative Scatter Correction – Correção Multiplicativa de Sinal

SEC – Standard Error of Calibration – Erro Quadrático de Calibração

SEP – Standard Error of Prediction – Erro Quadrático de Previsão

RMSECV – Root Mean Square Error of Cross Validation – Raiz Quadrada do Erro Médio Quadrático de Validação Cruzada

RMSEP – Root Mean Square Error of Prediction – Raiz Quadrada do Erro Médio Quadrático de Previsão

IV – Infravermelho

1. Introdução

1.1. Objetivos e Programa de Trabalho Realizado

O presente trabalho de estágio curricular consistiu na execução de um projeto na área do controlo da qualidade de tintas e vernizes produzidos na unidade fabril da CIN na Maia (sede do grupo). O principal objetivo deste projeto teve por base a criação de uma aplicação computacional quimiométrica, que permitisse a correta identificação de tintas e vernizes transparentes de base aquosa e de base solvente e o consequente reconhecimento entre produtos aceites e rejeitados por espectroscopia de infravermelho próximo (NIR) com recurso a técnicas quimiométricas de análise multivariada. A implementação desta aplicação quimiométrica irá provocar a diminuição dos atuais tempos de controlo das tintas e vernizes e, consequentemente, uma diminuição dos custos associados para a empresa.

Um segundo objetivo consistiu na obtenção de uma aplicação quimiométrica que permitisse a análise quantitativa de modo a determinar o teor de água em tintas de base aquosa.

Numa fase inicial procedeu-se a uma revisão bibliográfica do material existente sobre a técnica em estudo e à recolha dos produtos necessários ao desenvolvimento e cumprimento do plano de trabalho proposto.

O trabalho prosseguiu com a aquisição dos espectros de infravermelho próximo das tintas e vernizes, para seguidamente iniciar-se o tratamento daqueles espectros por técnicas quimiométricas de análise multivariada, nomeadamente a análise dos componentes principais (PCA) e a regressão dos mínimos quadrados parciais (PLS), a fim de se construir as calibrações multivariadas que permitiriam estabelecer o modelo de identificação e reconhecimento da qualidade das tintas e vernizes, bem como o modelo de determinação do teor de água em tintas de base aquosa.

Apesar de se ter conseguido obter mais do que uma calibração para o modelo de determinação do teor de água em tintas de base aquosa, não foi possível ser otimizado de modo a que pudesse ser implementado no decurso do estágio, pelo que é referido neste relatório de modo muito sucinto e apenas como um trabalho ainda em estudo.

1.2. Entidade de Acolhimento

A CIN é uma empresa cuja atividade industrial se centra na comercialização e produção de tintas, vernizes e produtos afins para as mais diversas áreas do mercado, que pode ir desde acessórios, repintura automóvel, indústria, proteção anticorrosiva e decorativos. Atualmente, o Grupo CIN é líder ibérico do mercado de tintas e em 2014 encontrava-se no 37.º lugar do Coatings World, ranking mundial de empresas produtoras de tintas.

A história da empresa tem início no ano de 1917 quando é fundada a *Companhia Industrial do Norte Lda.*, cuja atividade consistia no fabrico e introdução de vários produtos comerciais no mercado, tais como sabões, tintas, óleos e vernizes. Em 1985 é constituída a *CIN – Corporação Industrial do Norte S.A.*, nome que vigora até aos dias atuais.

Desde o início da sua fundação a CIN nunca parou de evoluir, tendo conseguido aumentar a sua cota de mercado e a liderança em acabamentos industriais nos anos 1950. No entanto, foi nos anos 1970 que se verificou um grande crescimento na área das tintas decorativas.

A forte aposta numa estratégia de crescimento e internacionalização do Grupo CIN, levou à criação em 1970 e em 1973 das empresas de Tintas Angola e Tintas Moçambique, SARL, respetivamente. Contudo, a política de expansão seguida pela empresa não se cingiu apenas a estes dois países, mas também teve em atenção as ilhas dos Açores e da Madeira, e ainda países como Espanha e França, com a aquisição de outras empresas do ramo das tintas e com a constituição de parcerias com outras empresas inseridas no mercado das tintas. Assim, com vista à internacionalização, a CIN adquiriu a *Barnices Valentine* em Espanha e a *Celliose* em França. A nível nacional, no mesmo seguimento de aquisições, foram adicionadas ao grupo a *Fábrica de Tintas Lacose Lda.* e a *Sotinco – Sociedade de Tintas de Construção, S.A.*, no fim dos anos 1980 e inícios dos anos 1990, respetivamente.

Em 1991, a CIN vem revolucionar o mercado das tintas decorativas com o lançamento do sistema COLORMIX.

Em pleno século XXI, o Grupo CIN continua a implementar a sua política de crescimento com o reforço da participação no capital de diversas empresas e com o estabelecimento de novas parcerias a nível nacional e internacional, de modo a complementar a sua oferta.

Atualmente, o Grupo CIN é constituído em Portugal pela empresa mãe, *CIN*, pela *Sotinco* e pela *NITIN*, em Espanha pela *Barnices Valentine*, pela *Industria de La Pintura* e pela *CIN Canárias*, em França pela *Artilin* e pela *Celliose*, em Angola por *Tintas CIN de Angola*, SARL e, por fim, em Moçambique por *Tintas CIN Moçambique*, SARL [1].

1.3. Inserção ao Tema

Uma das atividades inerentes à política de gestão da qualidade é a procura e a consequente implementação de novos e expeditos métodos de análise de baixo custo, para a concretização do controlo de qualidade de produtos, processos e serviços de uma unidade de produção. Estes requisitos são preenchidos na sua totalidade pela técnica de espectroscopia de infravermelho próximo, o que lhe confere uma relevância cada vez maior em controlo de qualidade e processos e produtos industriais. Esta é uma técnica rápida, simples, não destrutiva e que, habitualmente, dispensa qualquer tipo de pré-tratamento prévio da amostra, evitando desse modo o consumo de reagentes e, consequentemente, a formação de resíduos resultantes da análise [2]. A estas características acresce o facto dos espectros NIR conterem informação diversificada sobre a amostra ensaiada, permitindo uma análise simultânea dos vários componentes presentes na amostra, com níveis de exatidão e precisão comparáveis aos de outras técnicas de referência.

O NIR vem sendo aplicado há mais de 25 anos, num ritmo sucessivamente crescente, em diferentes áreas, entre as quais se destaca a indústria farmacêutica. Esta começou por ser aplicada apenas na análise de matérias-primas, mas mais recentemente tem sido igualmente aplicada na análise de formulações sólidas e líquidas, para o controlo de qualidade do produto final, bem como para a monitorização de operações intermédias de produção. A capacidade de reconhecimento atribuída aos espectros NIR provém do facto da radiação na região do infravermelho ser suficientemente energética para causar múltiplas e diversificadas transições vibracionais de energia, que basicamente resultam numa impressão digital química e física da amostra. Assim, basta a aquisição do espectro NIR da amostra a ensaiar para que a identidade e, eventualmente, a qualidade dessa amostra seja confirmada a partir de algoritmos de padrões conhecidos e que são construídos com recurso a técnicas quimiométricas. Basicamente, estas técnicas permitem extrair de forma organizada a complexa informação contida nos espectros, possibilitando identificar as amostras ou agrupamento de amostras, bem como deteção de amostras não conformes, para além de conseguir também estimar valores de parâmetros analíticos através do estabelecimento prévio de calibrações apropriadas [3].

2. Tintas e Vernizes

2.1. Definições

Uma tinta pode ser definida como um produto de pintura homogéneo pigmentado, que pode encontrar-se na forma líquida, em pó ou em pasta, que quando aplicada sob um substrato forma uma película opaca, dotada de propriedades protetoras, decorativas ou propriedades específicas ao seu uso final. Os seus principais constituintes são o veículo fixo ou ligante (i), o veículo volátil ou solvente (ii), os pigmentos (iii), os aditivos (iv) e as cargas (v) [4].

Um verniz pode ser definido por um composto homogéneo não pigmentado, que se apresenta sob a forma de líquido, sólido ou pastoso, que quando aplicado sob uma superfície apropriada numa camada fina, no estado em que é fornecido ou após diluição, é convertível numa película sólida, ao fim de um dado espaço de tempo, contínua, transparente ou translúcida, que pode ser mais ou menos dura. Os constituintes fundamentais de um verniz são o veículo fixo ou ligante (i), o veículo volátil ou solvente (ii) e os aditivos (iv) [4]. Uma das principais diferenças entre uma tinta e um verniz está na inexistência de pigmento na formulação deste último, de modo a ressaltar a sua textura e cor natural.

2.1.1. Características Principais dos constituintes

- (i) O veículo fixo, designado também por ligante, é constituído por resinas utilizadas no fabrico de tintas, vernizes e produtos similares. Este componente é responsável pelas principais características do revestimento da película, sendo a resina responsável pela formação do filme seco depois do produto final aplicado e seco. O veículo seco serve como estrutura do pigmento e atua como uma barreira de corrosão, como também confere propriedades de resistência mecânica, elasticidade, durabilidade e brilho [4]. Assim, as resinas utilizadas em tintas e vernizes podem ser acrílicas, vinílicas, epóxi, endurecedoras de epóxi, acrílicas em emulsão, poliuretano, etc.
- (ii) O veículo volátil é constituído por solventes utilizados no fabrico de tintas, vernizes e produtos similares, que modo geral é eliminado por evaporação no processo de secagem. Estes têm um papel fundamental no ajuste da viscosidade dos produtos em que entram na formulação, apesar de não manifestarem qualquer reatividade química, e no controlo de algumas propriedades, como por exemplo reológicas, de aplicação, tal com afeta a

estabilidade do produto enquanto se encontra no estado líquido [4]. Os solventes mais utilizados são hidrocarbonatos e solventes oxigenados (álcoois, cetonas, éteres e ésteres). No caso de produtos aquoso o principal solvente é a água, enquanto que em produtos de base solvente pode existir uma combinação de diferentes solventes, mistura designada de diluente, que pode conter hidrocarbonetos alifáticos, aromáticos, álcoois, cetonas, etc.

- (iii) Os pigmentos são substâncias sólidas, em geral, finamente divididas, praticamente insolúveis no meio, usadas no fabrico de tintas, com a finalidade de lhe conferir cor e opacidade (elevada opacidade). A sua incorporação em tintas permite igualmente auxiliar no controlo de brilho do revestimento, contribuir para o aumento do poder de cobertura, resistência mecânica e às intempéries, de aderência e propriedades anticorrosivas [4]. Estes podem ser classificados com pigmentos orgânicos, inorgânicos e metálicos, sendo exemplo, o dióxido de titânio (pigmento inorgânico); ftalocianina de cobre (pigmento orgânico) e alumínio, ligas de cobre, zinco e bronze (pigmentos metálicos).
- (iv) Os aditivos são substâncias incorporadas em pequenas quantidades em tintas, vernizes e produtos similares com o objetivo de lhes alterar acentuadamente determinadas características, tais como, contribuir para facilitar a aplicação e resolver defeitos que possam surgir na película. Estes são agentes dispersantes/molhantes, secantes, biocidas, anti espuma, inibidores de corrosão, entre outros, podendo encontrar-se no estado sólido ou líquido e são classificados apenas pela sua função, distinguindo-se em quatro classes: cinéticos, reológicos, dispersivos e para preservação dos produtos em que são incorporados [4,5].
- (v) As cargas são substâncias minerais, sob a forma de partículas de granulometria variável, com fraco poder de cobertura (baixa opacidade), tal como os pigmentos são insolúveis no meio, que são empregues em formulações de tintas. Este componente tem como finalidade modificar algumas propriedades da película, tais como a resistência química, o brilho, a velocidade de sedimentação, a resistência à abrasão, o comportamento anticorrosivo, a viscosidade, a espessura e demãos, a permeabilidade a gases, a humidade e o teor de sólidos. São exemplos de cargas o carbonato de cálcio, a barita, talcos, caulinos, entre outros [4].

2.1.2. Classificação de Tintas e Vernizes

As tintas e vernizes podem ser classificados seguindo diferentes critérios, tais como, a natureza do solvente, o processo de cura, a natureza do substrato a que se destinam, o meio e o método de aplicação [4]. Assim temos,

- (i) Pela natureza do solvente pode-se subdividir estes produtos em dois grandes grupos. Produtos de base solvente, em que o ligante geralmente se encontra disperso na fase aquosa e produtos de base solvente, em que geralmente o ligante está disperso num solvente ou diluente adequado.
- (ii) O processo de cura é um critério de classificação que permite distribuir os produtos por quatro classes: produtos de secagem física, em que a película é originada pela evaporação dos solventes; produtos de secagem por reação química; produtos de secagem em que a formação do filme surge à medida que os componentes regem com o oxigénio presente no meio; produtos de secagem por reação com a radiação ultravioleta e, por último, produtos de secagem a altas temperaturas.
- (iii) Com base na natureza do substrato pode-se encontra-se produtos para superfícies minerais, metálicas, plásticas e de madeira.
- (iv) O meio de aplicação dos produtos classificadas os produtos para a indústria automóvel, construção civil, indústria mobiliária e indústria naval.
- (v) O método de aplicação divide os produtos em dois grupos, sendo eles produtos de aplicação manual e produtos de aplicação automática.

3. Espectroscopia de Infravermelho Próximo

3.1. Referências Históricas

A descoberta da radiação infravermelha teve um primeiro impulso em 1666 por Sir Isaac Newton, quando este demonstrou, com base num prisma, que a luz solar se decompunha nas cores do arco-íris, vermelho, laranja, amarelo, verde, azul, anil e violeta. Com base neste princípio, Sir Frederick William Herchel, em 1800, realizou uma experiência com o auxílio de um termómetro e de um prisma de vidro, transparente à radiação, em que pretendia descobrir qual a contribuição que cada uma das cores do espectro solar tinha no aumento de temperatura das substâncias a ela expostas. No decorrer da experiência, verificou que a temperatura aumentava gradualmente em direção à luz vermelha. Com a continuidade desta experiência para a região logo acima à luz vermelha (região invisível da radiação), Sir Herchel verificou a existência de um aumento significativo de temperatura, levando-o a concluir que se tratava de um novo tipo de radiação não-visível. Após a realização de mais algumas experiências, Frederick Herschel determinou que o novo tipo de radiação recém-descoberta sofria efeitos tais como reflexão, refração, absorção e transmissão, tal como a radiação visível. Esta zona do espectro foi denominada de radiação infravermelha [6].

Em 1905, Coblentz apresentou os resultados de uma experiência, na qual utilizou um espectrómetro (Figura 1) construído a partir de sal de rocha (sal-gema) e com uma pilha acoplada a um galvanómetro de espelho, com o qual conseguiu obter os primeiros espectros de compostos puros e verificar a importância da sua utilização para a identificação de grupos funcionais de compostos orgânicos. Desta forma, Coblentz compreendeu que cada composto está associado a um espectro característico, sendo este específico e único da sua constituição química [6].

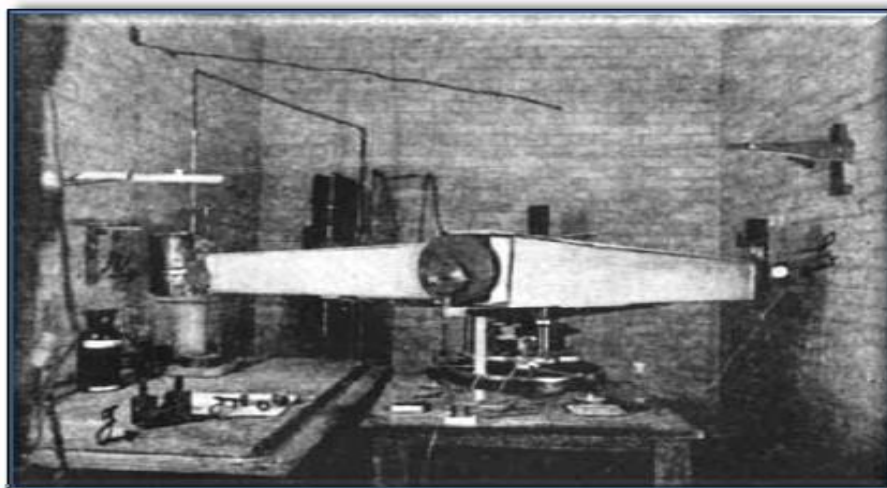


Fig. 1 - Espectrómetro rudimentar. (adaptado da referência 7)

A técnica de NIR teve um dos seus grandes desenvolvimentos durante a Segunda Guerra Mundial, onde foi utilizada como técnica de análise. Este tipo de utilização permitiu impulsionar o desenvolvimento da tecnologia usada na época, nomeadamente os espectrofotómetros [8].

Até à década de 60 do século passado, a técnica de NIR não registou um desenvolvimento apreciável, mas nas décadas seguintes foi objeto de estudo mais aprofundado devido ao trabalho desenvolvido por Karl Norris, no departamento Americano da Agricultura. O sucesso obtido por esses estudos proporcionou o interesse de outras indústrias no uso desta técnica, tais como a farmacêutica, a alimentar, a petrolífera e a bioquímica, por proporcionar uma alternativa credível aos métodos de análise usados na época para monitorizar os produtos acabados, que naquelas indústrias, estão sujeitos a rigorosa regulamentação.

Nos últimos anos, a espectroscopia de infravermelho próximo evoluiu ao ponto de se ter tornado uma técnica indispensável para a investigação científica, bem como para uma vasta gama de aplicações no controlo de qualidade de produtos industriais.

3.2. Princípios Teóricos

Quando a radiação infravermelha interatua com a matéria, uma parte dessa radiação é absorvida pelas ligações químicas das unidades estruturais de que a matéria é composta, fazendo com que estas vibrem de formas diferentes. Tal como acontece com as cordas de um instrumento, as unidades estruturais, como por exemplo as moléculas, possuem frequências características de vibração que são atribuídas pelos chamados modos vibracionais. Os modos mais comuns de

vibração das moléculas são de dois tipos: alongamento ou estiramento e a deformação angular. Os modos vibracionais do primeiro tipo são provocados por oscilações radiais das distâncias entre os núcleos, podendo exibir uma configuração simétrica ou assimétrica (Figura 2). Enquanto os modos vibracionais do segundo tipo envolvem a alteração dos ângulos formados entre o plano que contém a ligação e um plano de referência, podendo apresentar-se sob a forma de tesoura, baloiço, meneio e torção (Figura 3).

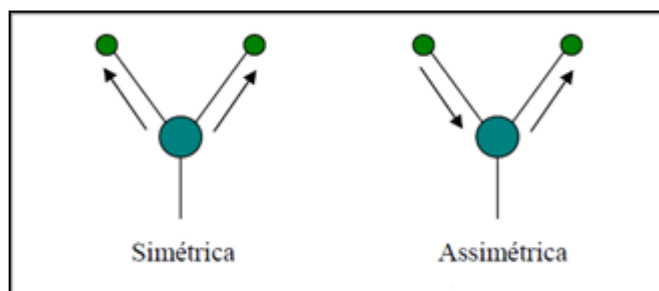


Fig. 2 - Modos vibracionais de alongamento ou estiramento (*stretching*). (adaptado da referência 9)

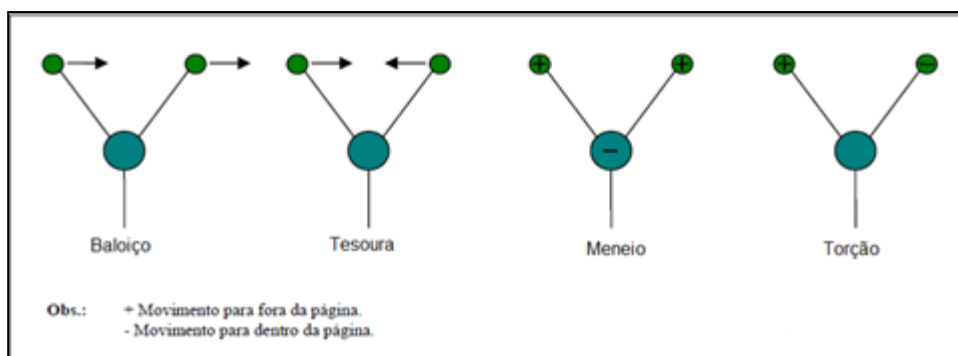


Fig. 3 - Modos vibracionais de deformação angular (*bending*). (adaptado da referência 9)

Nem sempre os modos vibracionais dão origem a bandas observadas no espectro de IV, uma vez que nem todos os tipos de vibração podem ser excitados pela radiação IV. As condições necessárias para que se verifique absorção de radiação IV pela matéria são descritas por duas regras de seleção. A primeira está relacionada com o modo como o vetor elétrico característico da radiação incidente interage com a molécula. Quando a molécula vibra com a mesma frequência que o vetor elétrico, ela entra em ressonância e absorve a energia da radiação. Mas para que a ligação seja observável em espectroscopia de IV é necessário que, durante a vibração, ocorra variação (não nula) do momento dipolar da molécula em relação à distância internuclear de equilíbrio.

A segunda condição fica-se a dever ao facto das moléculas poderem ocupar apenas alguns níveis de energia vibracional, perfeitamente definidos e quantizados. Os níveis de energia vibracional que uma determinada ligação pode ocupar são representados por números quânticos vibracionais (n) e, para uma dada ligação transitar, por exemplo, do nível $n=0$ para o nível $n=1$ (transição fundamental), é necessário que absorva radiação com uma energia igual à diferença entre as energias dos dois níveis.

Se os modos normais de vibração atuarem como um oscilador harmónico ideal, a regra de seleção permite a ocorrência de transições que envolvam apenas a variação unitária do número quântico vibracional. Todavia, os modos vibracionais de uma molécula afastam-se, em maior ou menor grau, do comportamento de um oscilador harmónico ideal, pelo que é frequente observar-se transições entre estados vibracionais com $\Delta n = +2, +3, +4, \dots$. Estas transições não-fundamentais denominam-se por *overtones*, que são caracterizados por possuírem valores de frequência que correspondem a múltiplos inteiros dos valores das frequências das vibrações fundamentais, que ocorrem na região do infravermelho médio. Tomando como exemplo, um valor de ν de frequência de um dado modo vibracional fundamental, os overtones dão origem a valores aproximados de 2ν , 3ν e assim sucessivamente. Como surgem em consequência da quebra da regra de seleção, os *overtones* são normalmente bandas muito mais fracas do que as bandas de transição fundamental. A banda para o primeiro overtone é 10 a 100 vezes mais fraca do que a transição fundamental.

Embora também sejam formalmente proibidas, é também frequente observar-se mudanças energéticas simultâneas em dois ou mais modos vibracionais, o que leva ao aparecimento de bandas de absorção correspondentes a combinações de vibrações fundamentais. Estas bandas são conhecidas por bandas de combinação e resultam da combinação linear de frequências fundamentais ou de múltiplos inteiros destas. Por exemplo, seja ν_a e ν_b valores de frequências vibracionais, podem ocorrer as seguintes bandas de combinações $(\nu_a + \nu_b)$, $(\nu_a - \nu_b)$ e $(\nu_a + 2\nu_b)$.

As frequências dos *overtones* e das bandas de combinação são necessariamente superiores às das bandas de vibrações fundamentais, indo ocupar a zona do infravermelho próximo. Desta forma, os espectros de absorção na zona do infravermelho são de interpretação complexa devido ao facto das absorções serem menos intensas e de se darem sobreposições de ligações contendo CH-, -OH e -NH dos vários grupos funcionais. Para compreender a informação fornecida é necessário recorrer-se a métodos estatísticos e computacionais (Quimiometria) para a realização de uma calibração multivariada, a partir de vários comprimentos de onda. A Figura 4 mostra as principais bandas de absorção na região do infravermelho próximo e a sua localização no espectro eletromagnético

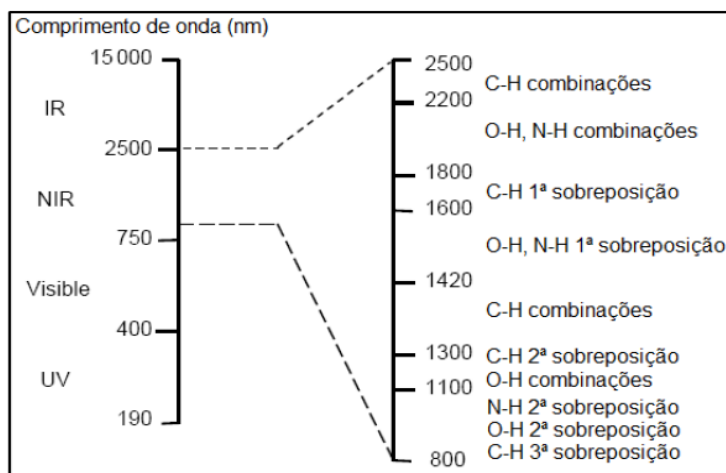


Fig. 4 - Principais bandas de absorção na região do infravermelho próximo e a sua localização no espectro eletromagnético. (adaptado da referência 10)

A região espectral do infravermelho enquadra-se entre as regiões do visível e do micro-ondas, sendo subdivida em infravermelho próximo, infravermelho médio e infravermelho longínquo. A espectroscopia de infravermelho próximo corresponde à zona de comprimento de onda compreendido entre 750 e 2500 nm (13300 a 4000 cm^{-1}) e tem uma energia de fóton ($h\nu$) que varia desde $2,65 \times 10^{-19}$ e $7,96 \times 10^{-20}$ J [9].

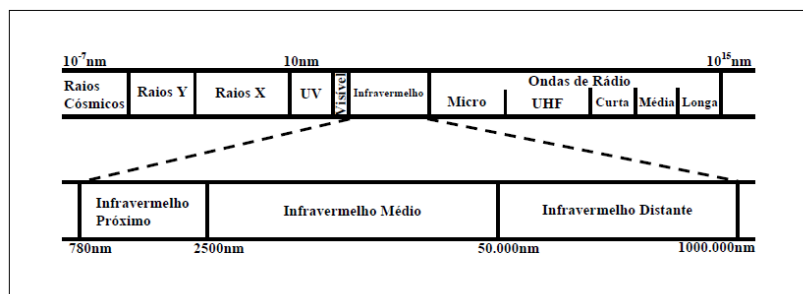


Fig. 5 – Localização da radiação infravermelha no espectro eletromagnético (adaptado da referência 9).

Como a intensidade de uma banda de absorção é proporcional à concentração da espécie química que causa essa banda, a quantidade de um dado componente existente numa amostra pode ser determinada através de uma curva de calibração construída com base em amostras de concentração conhecida do componente em questão, através de uma análise multivariada devido ao elevado número de variáveis obtidas num espectro. No entanto, um espectro NIR também permite obter curvas de calibração para a identificação de amostras desconhecidas, tendo por base a análise dos componentes principais.

A intensidade das bandas NIR está dependente da mudança do momento dipolar da molécula e não da harmonicidade da ligação. Assim, as principais bandas tipicamente observadas na região NIR correspondem a ligações que contenham hidrogénio ou outros átomos leves (C-H, N-H, O-H, S-H), pois são os que apresentam um maior número de vibrações e um maior desvio ao comportamento do oscilador harmónico ideal.

Contudo, existem interações entre diferentes átomos numa molécula, tais como pontes de hidrogénio ou interações dipolo-dipolo, que podem alterar os estados de energia vibracional, havendo uma deslocação das bandas de absorção existentes e dando origem a novas bandas, a partir das diferenças nas estruturas cristalinas. Deste modo, e com base na técnica NIR, é possível que diferentes estruturas cristalinas de uma mesma substância sejam distinguidas, possibilitando a determinação de várias propriedades físicas, como por exemplo a viscosidade, a densidade, o tamanho de partícula e a percentagem de água.

Em suma, a informação contida num espectro NIR passa da composição química para a determinação de componentes presentes na substância, como também contém informação física que pode ser útil na determinação das propriedades físicas das amostras.

3.3. Instrumentação – Espectrofotómetro de Infravermelho Próximo com Transformada de Fourier

Os espectrofotómetros baseados na radiação infravermelha próxima evoluíram muito rapidamente ao longo dos últimos anos, devido às inúmeras vantagens que a técnica NIR possui e pelas diversas aplicações a nível industrial.

De modo geral, um equipamento NIR é composto por uma fonte luminosa, um monocromador, um dispositivo que possa armazenar a amostra e/ou uma interface que contacte diretamente com a amostra e um detetor de sinal que permita realizar medições de transmitância e/ou refletância.

A fonte luminosa utilizada neste tipo de equipamentos pode ser uma lâmpada de halogéneo, quartzo ou LEDs [11]. No caso do detetor de sinal, um equipamento NIR pode possuir um detetor de silício para baixos comprimentos de onda ou um detetor de sulfureto de chumbo/índio-gálio-arsénio para altos comprimentos de onda. O detetor de sulfureto de chumbo é o que apresenta um maior tempo de vida útil, sendo este mais barato e altamente sensível, no entanto, apresenta limitações na linearidade, saturação e velocidade de resposta. Este tipo de limitações podem ser atenuados com o uso de detetores de InGaAs [12].

A aquisição dos espectros NIR pode ser realizada a partir de diferentes tipos de espectrofotómetros, tendo por base diferentes princípios teóricos. A escolha do equipamento mais adequado deve ter por base os seguintes critérios, como por exemplo a aplicação a que se destina, a sensibilidade pretendida das determinações e a robustez dos resultados.

No mercado existem diversos equipamentos NIR adaptados a diferentes ambientes de utilização, tais como ambientes fabris em que os equipamentos estão sujeitos a variações de temperatura e a vibrações, ambientes controlados em que requerem a presença de ar condicionado de modo a que seja mantida uma temperatura controlada, ou ainda, ambientes isentos de vibrações, sendo este último equipamento o mais adequado ao uso em ambiente laboratorial.

Os espectrofotómetros mais utilizados atualmente funcionam com Transformada de Fourier (FT-NIR). Este tipo de aparelhos tem por base um interferómetro de Michelson, em que o funcionamento consiste na divisão, por um *beamsplitter*, da radiação emitida pela fonte luminosa em dois feixes iguais. Os dois feixes resultantes dirigem-se para dois espelhos, um numa posição fixa e outro numa posição móvel, que irão refletir os feixes, sendo estes novamente divididos. Como consequência da nova divisão, parte da radiação será emitida na direção do detetor de sinal. Este sistema de funcionamento permite abranger todos os comprimentos de onda em estudo de forma rápida e sem necessitar de uma grande quantidade de amostra.

À medida que o espelho de posição móvel percorre uma certa distância, é originado um interferograma a partir da recombinação dos feixes absorvidos pelo detetor de sinal, isto é, o interferograma será constituído pelo somatório de todas as ondas de diferentes amplitudes e frequências que chegam ao detetor.

O interferograma criado possui toda a informação da amostra sendo contudo de complexa interpretação. Para ultrapassar esta limitação é aplicado ao interferograma a operação matemática da transformada de Fourier (Equação 1), que, quer por ação de um microprocessador acoplado ao espectrofotómetro, quer por um computador associado externamente ao aparelho, distribui a energia absorvida pela amostra por cada comprimento de onda, apresentando desse modo os valores de cada frequência, resultando numa leitura e compreensão dos resultados mais simplificada [12].

$$F(x) = \int_{-\infty}^{\infty} f(x) e^{-2\pi i k x} dx$$

Equação 1

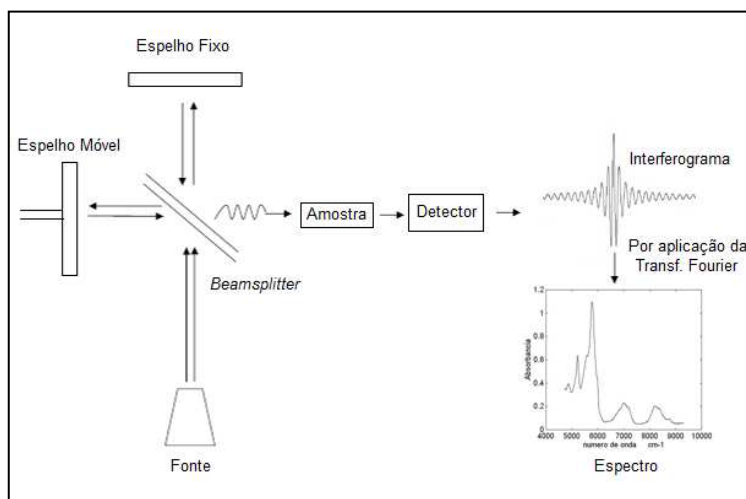


Fig. 6 - Representação esquemática de um espectrofotómetro FT-NIR com interferómetro de Michelson. (adaptado da referência 13)

Tal como já foi referido, estes aparelhos podem funcionar em modo de transmitância ou de refletância difusa, ou ainda num modo combinado dos dois, designado de transfletância. Este último foi o modo de trabalho empregado no estudo realizado.

A unidade de transmitância baseia-se na absorção da radiação de parte da radiação incidente, isto é, mede o decréscimo da intensidade da radiação em função do número de onda, quando esta atravessa a amostra e é, novamente, absorvida pelo detetor. Neste tipo de medição, é necessário ter em consideração a concentração e a homogeneidade, tal como o tamanho das partículas. Este é um método indicado para medições em soluções líquidas, dado que nestas o efeito da reflexão é desprezável [15]. A transmitância pode ser descrita pela razão entre a intensidade transmitida após a absorção e a intensidade da luz incidente, tal como demonstrado pela equação 2.

$$T = \frac{I}{I_0} = 10^{-A} \quad \text{Equação 2}$$

(I – intensidade transmitida após absorção; I_0 – intensidade da luz incidente; A - absorvância)

A unidade de refletância difusa com transformada de Fourier é conhecida como DRIFTS (*Diffuse Reflectance Infrared Fourier Transform Spectrometry*). Este modo de funcionamento ocorre quando a radiação incide numa matriz descontínua, em que parte da radiação é absorvida e parte é refletida em todas as direções, contendo as informações espectrais. No momento em que a radiação contacta com a amostra parte é absorvida pelas ligações vibracionais e a fração não absorvida poderá ser refletida de volta. A última fração de radiação será detetada de forma a permitir o cálculo da intensidade da radiação absorvida.

A refletância difusa é uma metodologia que pode ser utilizada em amostras em pó ou com partículas muito finas, em que estas dispersam bem a radiação incidente e absorvem em menor quantidade [16]. Assim a refletância é traduzida pela Equação 3:

$$R = \frac{I_R}{I_0} \quad \text{Equação 3}$$

(I_R – intensidade da luz refletida; I_0 – intensidade da luz incidente)

Os dois modos de medição, refletância e transmitância, apresentam uma importante diferença, que reside no caminho ótico percorrido pela luz. Isto é, na transmitância o caminho ótico percorrido pela luz é constante para todos os números de onda, enquanto que na refletância o caminho ótico pode ser variável.

Outro facto importante é que radiação que ocorre numa amostra obedece ao princípio da Lei de Lambert-Beer, ou seja, a absorção da radiação por uma amostra é relacionada com a concentração de um determinado analito. Contudo, deve-se ter em consideração que a absorvância de cada analito presente na amostra é aditiva. A lei de Lambert-Beer é definida pela seguinte equação:

$$A = \varepsilon \times b \times c \quad \text{Equação 4}$$

(ε – absortividade molar do analito [$\text{mol}^{-1} \text{ L cm}^{-1}$]; b – percurso ótico percorrido pela luz [cm]; c – concentração do analito [mol L^{-1}])

Outras teorias foram desenvolvidas de forma a explicar o fenómeno de refletância difusa, entre elas encontra-se a teoria de Kubelka-Munk, em que se considera que a amostra é constituída por partículas aleatoriamente distribuídas, que absorvem e dispersam a luz uniformemente, sendo a dimensão das partículas muito menor que a espessura da camada [17].

A teoria de Kubelka-Munk é traduzida por uma razão de linearidade entre a intensidade de reflexão e a concentração da amostra (Equação 5). Assim, assume-se uma diluição infinita numa matriz não absorvente, o coeficiente de dispersão constante e uma camada de amostra “infinitamente espessa”, isto é, o aumento da espessura não afeta a reflexão [18].

$$F(R) = \frac{(1-R)^2}{2R} = \frac{kc}{S^*} \quad \text{Equação 5}$$

(R – valores de refletância; K – coeficiente de absorção molar; c – concentração; S – coeficiente de dispersão da amostra)

Através da equação de Kubelka-Munk é possível relacionar os espectros NIR de refletância difusa com a concentração da amostra, transformando os espectros de reflexão num formato semelhante ao de um espectro de absorvância.

Para qualquer uma das metodologias, refletância difusa e/ou transmitância, os espectros resultantes na espectroscopia NIR são apresentados em termos de absorvância, e consoante os dados obtidos forem baseados em transmitância ou refletância, esta é definida por $A = \log(1/T)$ ou $A = \log(1/R)$ [19].

3.4. Fatores que Afetam a Aquisição dos Espectros NIR

De acordo com a literatura, existem diversos fatores que podem afetar tanto a aquisição dos espectros como, posteriormente, a comparação com espectros de amostras desconhecidas. Seguidamente são apresentados alguns desses fatores [20]:

- **Temperatura da Amostra:** no caso de amostras líquidas ou soluções aquosas, ou ainda sólidos que contenham solventes como água, as variações de temperaturas podem dar origem a alterações na estrutura dos espectros obtidos.
- **Solventes e Humidade:** as ligações de hidrogénio, provenientes dos solventes, que se ligam ao princípio ativo ou aos excipientes podem alterar o espetro.
- **Espessura da amostra:** é um dos fatores mais importantes e do qual se deve ter um controlo mais rigoroso, pois quer no modo de transmitância usando um comprimento ótico fixo, quer no modo de refletância em que o comprimento ótico é variável, deve-se verificar a profundidade de penetração da luz, para que não altere o espetro se a espessura da amostra aumentar.
- **Propriedades Óticas da Amostra:** quando as substâncias são sólidas, a superfície de espalhamento da amostra deve ser tida em consideração quando se procede à calibração, pois a morfologia e o índice de refração afetam as propriedades de dispersão dos materiais sólidos. No caso de amostras em pó, o tamanho das partículas e a sua densidade influenciam a dispersão e, consequentemente o espetro obtido.
- **Polimorfismo:** as variações da estrutura cristalina de amostras com a mesma composição química irão influenciar o espetro final. No caso de diferentes polimorfismos ou formas sólidas amorfas podem ser distinguidos pelos espectros NIR obtidos, bem como diferentes estados de hidratação cristalina ou estados de solvatação do mesmo material podem resultar em espectros distintos dos utilizados na calibração.
- **Idade da Amostra:** em amostras antigas podem manifestar-se alterações químicas, óticas e físicas, pelo que se deve ter em consideração as amostras e padrões utilizados

na construção das calibrações, pois estes podem não estar nas condições mais adequadas a uma correta análise espectral.

3.5. Vantagens e Desvantagens da Técnica NIR

A espectroscopia de infravermelho próximo apresenta vantagens e desvantagens, tal como qualquer outra técnica de espectroscopia usada para fins analíticos. De seguida são descritas algumas dessas vantagens [20] e desvantagens [21].

Vantagens:

- Os espectros NIR podem ser recolhidos sem necessidade de preparação das amostras, não sendo um método destrutivo nem invasivo quer para amostras sólidas ou líquidas.
- A utilização de espectrofotómetros de infravermelho próximo não implica o consumo de reagentes, nem a produção de resíduos.
- Este método é rápido na obtenção de espectros, e após construídas as calibrações, é possível obter a apresentação de cálculos e resultados de um modo quase imediato, o que permite a sua utilização em procedimentos de controlo em tempo real.
- Método versátil e fiável, que permite análises qualitativas (identificação) e quantitativas.
- A aquisição de um único espectro NIR permite a quantificação de vários componentes químicos e a determinação de várias propriedades físicas, mediante as calibrações previamente realizadas.
- A informação contida num espectro NIR é sensível a alterações de formulações ou impurezas existentes no produto, eventualmente não detetáveis por técnicas analíticas convencionais.
- A robustez dos equipamentos utilizados por esta técnica permite a análise em ambientes hostis, como por exemplo, em linhas de produção. Este tipo de aparelhos tem uma manutenção reduzida, sendo apenas necessário a substituição da lâmpada e do laser periodicamente.

- A sua utilização pode ser efetuada ao longo de todo o processo produtivo, isto é, desde a receção das matérias-primas até ao controlo do produto final, já que também permite análises remotas, importantes em casos de materiais tóxicos ou perigosos.
- Um mesmo equipamento pode monitorizar diversos produtos de diferentes linhas de produção, diminuindo dessa forma os custos com o controlo de qualidade.

Desvantagens

- Os espectros obtidos com a técnica de espectroscopia de infravermelho próximo são pouco seletivos e de difícil interpretação, pelo que têm que ser aplicadas técnicas quimiométricas adequadas aos dados do modelo pretendido para que possam ser extraídas informações relevantes.
- A construção das calibrações robustas e precisas implica a utilização de um número elevado de amostras, para que se possa abranger todas as possíveis variações das propriedades químicas e físicas.
- A necessidade de incorporar o maior número de variabilidades químicas e físicas implica a construção de calibrações intrincadas, utilizando técnicas quimiométricas complexas, consoante o tipo de amostra que se pretende controlar, além da necessidade da existência de mais do que uma calibração por propriedade ou até para a mesma propriedade.
- Para a construção das calibrações, principalmente as que são direcionadas à determinação de teor de algum componente (determinação quantitativa), é necessário a determinação das propriedades por métodos analíticos de referência.
- Reduzida aceitação da técnica NIR para efeitos de controlo de qualidade, pelos organismos reguladores, pois trata-se de uma técnica recente.

4. Quimiometria

A quimiometria é uma área de estudo que aplica conjuntamente métodos matemáticos, estatísticos e computacionais associadas às ciências químicas, com o intuito de selecionar e otimizar procedimentos experimentais.

As análises quantitativas realizadas em ambiente laboratorial são, na sua maioria, relativamente demoradas e com resultados pouco precisos, pelo que tendem a ser substituídas por técnicas instrumentais mais precisas, entre elas a espectroscopia de infravermelho próximo. Com a permanente evolução das técnicas instrumentais, impulsionada com o crescente desenvolvimento de microprocessadores e de microcomputadores em ambiente laboratorial, tornou-se primordial encontrar ferramentas que facilitassem o tratamento dos dados mais complexos do ponto de vista matemático e estatístico.

No caso da espectroscopia NIR, tal como referido anteriormente, os espetros apresentam bandas correspondentes a combinações de vibrações fundamentais e de overtones, que geram um elevado número de sinais. Assim estes espetros são de difícil compreensão direta, sendo necessário recorrer às técnicas quimiométricas de análise multivariada, tanto para a identificação de amostras como, também, para a determinação de concentrações de componentes pretendidos.

De acordo com o sistema de dados em estudo, a obtenção de uma calibração multivariada tem que ter em consideração as especificações e características pretendidas, para que desse modo a escolha do método de análise multivariada e dos pré-tratamentos seja a mais adequada. Os métodos de análise multivariada mais utilizados pelas técnicas quimiométricas são:

- PCA – Análise dos Componentes Principais
- PLS- Regressão dos Mínimos Quadrados Parciais
- MLR – Regressão Linear Múltipla

Nos pontos 4.1 e 4.2 são apresentados os métodos PCA, como método qualitativo, e PLS, como método quantitativo, sendo os que foram utilizados no decorrer deste trabalho.

No entanto, antes de se abordar os métodos acima indicados é necessário entender como são preparados os dados que serão, posteriormente, processados. Assim, para cada amostra é selecionada a média de vários espetros, os quais serão organizados numa matriz M_{ij} , que corresponderá ao valor de absorvância da amostra i no comprimento de onda j . Por exemplo, para o dado espectral X temos:

$$M = [X_{11} \ X_{12} \ \dots \ X_{1j}]$$

Equação 6

Para duas ou mais amostras teremos:

$$M = \begin{bmatrix} X_{11} & X_{12} & \dots & X_{1j} \\ X_{21} & X_{22} & \dots & X_{2j} \\ \vdots & \vdots & \ddots & \vdots \\ X_{nm} & X_{nm} & \dots & X_{nm} \end{bmatrix} \quad \text{Equação 7}$$

De modo análogo, organiza-se o conjunto de variáveis desejáveis ou analito sob a forma de um conjunto de dados N , que contém os valores da propriedade j da amostra i , que será utilizado na derivação das equações de correlação com a matriz M .

4.1. Análise dos Componentes Principais

A análise dos componentes principais ou *Principal Component Analysis* (PCA) é uma ferramenta quimiométrica utilizada para a redução de dados e visualização de amostras, reorganizando a informação obtida pelos dados de amostragem. Este método é, normalmente, utilizado na identificação de grupos distintos, ou seja, agrupando amostras com características semelhantes e permitindo a deteção de amostras *outliers*, isto é, amostras que por um qualquer motivo não pertencem ao sistema em estudo [22].

O PCA consiste num método estatístico que visa calcular, com base num conjunto de dados recolhidos, um subconjunto que abranja o máximo da variabilidade dos dados originais. No caso de dados espectrais, este método tenta encontrar um subconjunto de espectros através da combinação linear das absorvâncias determinadas inicialmente, de todo o espectro, por *Loadings* adequados, de forma a descrever a maioria das variações espectrais observadas dos dados iniciais. A combinação linear das absorvâncias obtida é designada por *combinação linear* ou *factor*.

Os componentes principais são obtidos por ordem decrescente de quantidade de informação estatística, ou seja, o primeiro componente principal é o vetor próprio mais elevado, o segundo componente principal é o segundo vetor com o valor mais elevado e assim por diante. O segundo e terceiros vetores são ortogonais aos vetores anteriores, descrevendo a maior fração da variabilidade restante [16]. Após a aplicação do método de análise multivariável aos componentes principais, as novas variáveis serão capazes de se correlacionar entre si e serão, igualmente, combinações lineares das variáveis originais.

A utilização do método PCA consiste na decomposição da matriz M que contém os dados espectrais X (onde as m linha são as amostras e as n colunas são as variáveis), numa soma de h novas matrizes de valor igual a 1, como apresentada na Equação 8:

$$X = M_1 + M_2 + M_3 + \dots + M_h \quad \text{Equação 8}$$

As novas matrizes criadas são o produto e vetores chamados *Scores*, t_h , e *Loadings*, p_h , respetivamente. Estas duas matrizes podem ser determinadas por um ajuste dos mínimos quadrados, um cálculo semelhante à decomposição dos valores singulares (SVD, *Singular Value Decomposition*). A equação pode ser representada na forma vetorial,

$$X = t_1 p'_1 + t_2 p'_2 + \dots + t_h p'_h \quad \text{Equação 9}$$

ou na forma matricial,

$$X = TP' \quad \text{Equação 10}$$

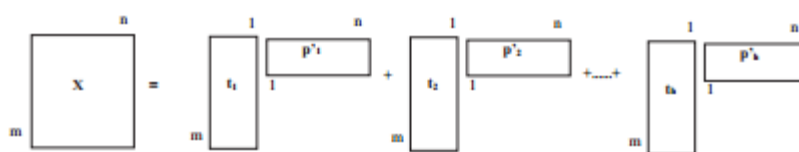


Fig. 7 – Representação gráfica do somatório dos produtos dos vetores Scores e Loadings (adaptado da referência 22).

A Figura 8 pretende exemplificar os vetores *Scores* e *Loadings* num plano bidimensional de duas variáveis x_1 e x_2 . Nesta figura é apresentada uma componente principal, em que a reta aponta na direção de maior variabilidade das amostras (pontos 1 e 6), da Figura 8B. Os *Scores* são as projeções das amostras na direção da componente principal e os *Loadings* são os cossenos dos ângulos formados entre a componente principal e cada variável (Figura 8A).

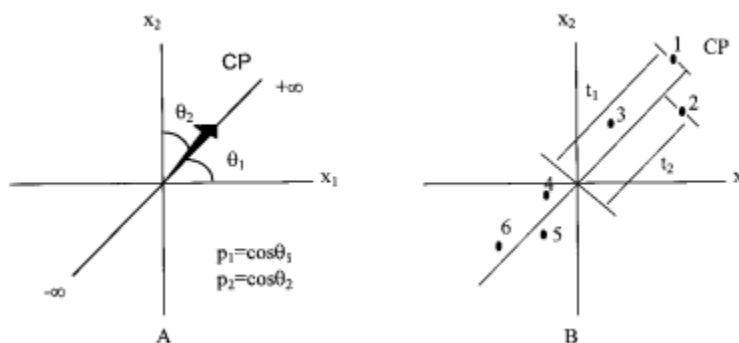


Fig. 8 – Representação de uma componente principal (CP) no caso de duas variáveis: (A) representa o vetor *Loadings* e (B) representa o vetor *Scores*. Notar que os dados estão centrados na média (adaptado da referência 22).

As novas variáveis originadas, os componentes principais, são ortogonais entre si, não sendo por isso correlacionáveis. Normalmente são os primeiros componentes principais que explicam a maior variabilidade contida nos dados espectrais.

Em resumo, a análise dos componentes principais é um método que possui como intuito principal a redução de dados a partir de combinações lineares das variáveis originais, sendo um dos métodos mais utilizados pela espectroscopia NIR, uma vez que permite visualizar milhares de variáveis a duas dimensões. A utilização deste método permite também a extração apenas da informação mais relevante, excluindo as variações associadas ao ruído [16].

4.2. Regressão dos Mínimos Quadrados Parciais

A regressão dos mínimos quadrados parciais ou *Partial Least Squares* (PLS), é uma técnica de análise de dados multivariados utilizada com o intuito de relacionar uma ou mais variáveis (matriz Y) com diversas variáveis independentes (matriz X), correlacionando-as de forma a obter uma relação linear, baseada no uso de componentes principais. Este método permite realizar determinações quantitativas, com base na construção de uma reta de calibração multivariada [22].

Dependendo do objetivo de estudo e das quantificações que se pretendem determinar, o número de amostras necessárias ao desenvolvimento da calibração varia entre as 20 e 200 amostras. Isto é, quanto mais complexa for a amostra, maior será o número de espectros a serem analisados para gerar a calibração.

Um outro fator importante a ter em consideração num estudo deste tipo é a região espectral usada, pois parte da informação contida no espectro NIR pode ser redundante para o estudo em questão, uma vez que parte dos sinais medidos em certos comprimentos de onda podem não ser lineares, constituir ruído ou conter informação inútil para a previsão do valor analítico. A escolha da região espectral deve ser realizada com base num maior valor do coeficiente de correlação (R) ou menor erro de previsão.

O processamento matemático consiste na determinação de duas matrizes que irão traduzir as variações espectrais e os valores de referência, seguidamente serão gerados os vetores próprios. Estes vetores serão designados de componentes principais, mas também designadas de variáveis latentes em alguma literatura, e poderão ser utilizados na previsão da concentração do componente pretendido, uma vez que os espectros NIR originais contêm todas as informações relevantes para o sistema em estudo. Esta decomposição em componentes principais apresenta como vantagem o facto da informação mais relevante ficar comprimida e poder ser utilizada na calibração [3].

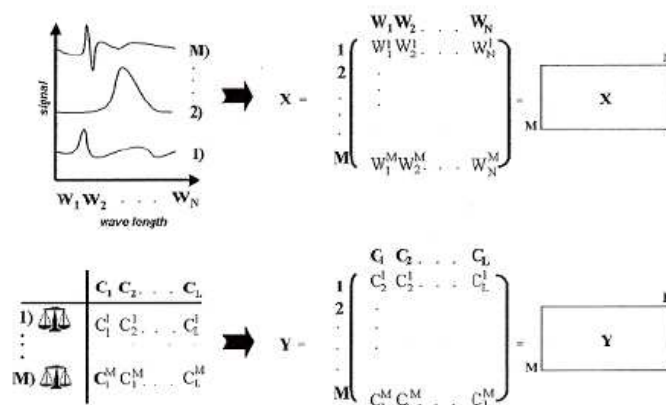


Fig. 9 - Representação esquemática das matrizes X e Y para o método PLS. A matriz X contém os dados espectrais, M representa o número de amostras analisadas e N os comprimentos de onda percorridos para a obtenção do espectro. A matriz Y contém os valores das concentrações, sendo as colunas referentes aos valores de referência dos vários componentes. (adaptado da referência 3)

O primeiro fator caracteriza as principais alterações observadas nos espectros NIR, no entanto quando este fator não é suficiente para explicar todas as alterações observadas, é gerado um segundo fator que poderá ser perpendicular ou ortogonal ao primeiro. Assim, com o aumento do número de fatores gerados por este método, são caracterizadas todas as alterações nas estruturas cada vez mais pequenas. No fim do processo de modelagem normalmente obtém-se uma matriz de resíduos que deverá apenas conter as alterações não explicadas associadas ao ruído. Um baixo número de fatores indica que apenas as alterações na estrutura espectral foram caracterizadas, contudo, um elevado número de fatores representa, principalmente, que foram caracterizadas as distribuições associadas ao ruído.

O ruído dos dados espectrais é independente do ruído dos dados de concentração, não sendo possível definir uma relação linear, necessitando-se de empregar o conceito dos mínimos quadrados.

As matrizes originais X e Y serão decompostas, cada uma, nos vetores *Scores* (T e U) e *Loading* (P e Q) e nas matrizes residuais E e F, respetivamente, correspondendo a Equação 11 para os dados espectrais e a Equação 12 para os valores de concentrações.

$$X = TP' + E = \sum t_h p'_h + E \quad \text{Equação 11}$$

$$Y = UQ' + F = \sum u_h q'_h + F \quad \text{Equação 12}$$

Quando o objetivo passa pela determinação de apenas uma variável dependente, têm-se a calibração PLS1 e y consiste só no vetor, no caso de se desejar prever mais do que uma variável dependente em simultâneo, na mesma calibração, Y passa a consistir numa matriz e o método denomina-se de PLS2.

A correlação linear entre as matrizes X e Y consiste numa relação linear obtida pelo coeficiente de regressão linear (Equação 13).

$$U_h = b_h t_h \quad \text{Equação 13}$$

Os valores de b_h são agrupados na matriz diagonal B , matriz essa que contém os coeficientes de regressão entre a matriz de *Scores* U de Y e a matriz de *Scores* T de X . A matriz Y pode ser determinada a partir de u_h :

$$Y = TBQ' + F \quad \text{Equação 14}$$

Sendo a concentração de novas amostras prevista com base nos novos *Scores*, T^* , substituídos na seguinte equação,

$$Y = T^*BQ' \quad \text{Equação 15}$$

A importância do método PLS em análises químicas surge do facto da simultaneidade e mútua dependência da fatorização dos dados presentes nas matrizes originais, conseguindo corresponder uma alteração dos dados espectrais a um valor de concentração de um dado componente.

4.3. Tratamento Prévio de Dados

Tal como já foi referido, os espectros NIR são de elevada complexidade e de difícil interpretação direta, sendo em alguns casos necessário recorrer ao uso de pré-tratamentos, gerando melhorias significativas no desenvolvimento de calibrações.

Assim, à medida que a complexidade do sistema em estudo vai aumentando, os espectros NIR podem ser influenciados por efeitos que não são facilmente quantificados, sendo alguns desses efeitos sistemáticos e não correlacionáveis com as variações a modelar. Tais efeitos podem ser:

- Interação complexa entre os componentes da amostra;
- Dispersão da luz em amostras sólidas e amostras líquidas turvas;
- Dificil reprodutibilidade das medições;
- Distorções espectrais causadas pelo *hardware* do equipamento, tais como: deslocamento da linha de base; efeitos de não-linearidade do detetor, ruído no detetor, saturação do detetor, etc;

- Falhas no completo varrimento do intervalo comprimento de onda escolhido;
- Variações no espectro devido à temperatura da amostra.

Por vezes existe a ocorrência de ruído em virtude do equipamento em utilização, no entanto, este pode ser atenuado com o aumento do número de varrimentos de comprimento de onda, embora seja comprometida a velocidade de análise, mas também pode ser atenuada com o uso de métodos de pré-tratamentos.

De seguida são apresentados alguns dos pré-tratamentos passíveis de serem utilizados, que têm como objetivo eliminar ou remover variações sistemáticas, corrigir a linha de base, minimizar o ruído, melhorar a definição dos picos que se encontram sobrepostos na mesma região, evidenciando os parâmetros de interesse e consequentemente aumentando a seletividade, remover a informação irrelevante e diminuir os efeitos causados por diferenças no tamanho de partículas. Apesar de todas estas possíveis melhorias dos espectros NIR, é necessário assegurar que o uso de pré-tratamentos não elimina a informação química e física de interesse para o estudo.

4.3.1. Correção Multiplicativa de Sinal (MSC)

A técnica de correção multiplicativa de sinal ou *Multiplicative Scatter Correction* (MSC), tem como objetivo a eliminação de efeitos de ordem física que possam surgir na aquisição dos espectros, tal como, a dispersão da luz, com a finalidade de linearizar os dados espectrais e diminuir a variação de ruído, permitindo dessa forma a aplicação de técnicas de regressão linear [23].

Este é um dos métodos utilizados quando ocorrem variações multiplicativas nos espectros NIR obtidos a partir das amostras analisadas. Assim, esta técnica é muito utilizada em leituras de refletância difusa, pois visa corrigir o efeito da dispersão da luz causada pela falta de homogeneidade das amostras [16].

Neste sentido, o espectro de cada amostra é modelado como uma função linear do espectro médio do conjunto total de espectros estudados, sendo o efeito final ilustrado pela Figura 10.

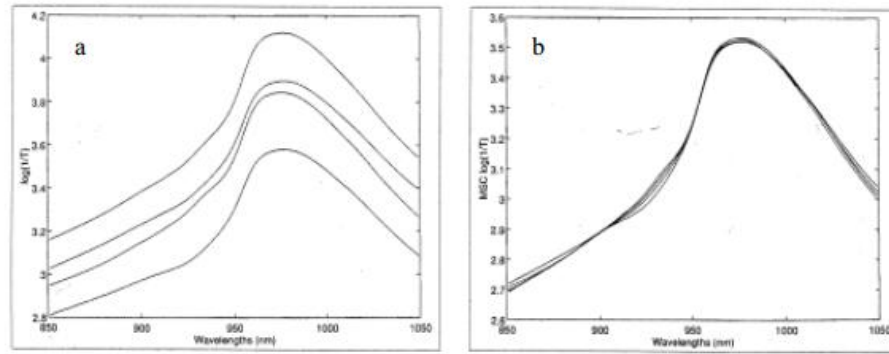


Fig. 10 – Espectros de absorvância NIR (a) antes da aplicação do pré-tratamento MSC e (b) após aplicação do pré-tratamento MSC. (adaptado da referência 24)

O MSC é um tratamento matemático que consiste no cálculo de um espectro médio determinado a partir de todos os espectros presentes no conjunto de calibração. Assim, cada espectro da amostra i (X_i) é transformado de acordo com a Equação 16 [14].

$$X_i' = u + v X_i$$

Equação 16

(X_i' – espectro transformado para a amostra i ; u e v – são escolhidos de acordo a que a diferença entre o espectro transformado e o espectro médio seja a mínima possível)

Este método tem como principais vantagens a redução do número de componentes principais necessários, simplificando a calibração e o aumento da linearidade [19].

4.3.2. Primeira Derivada

Um dos pré-tratamentos mais utilizados são as derivadas, que poderão ser de primeira ou segunda ordem dos espectros NIR originais, e têm como objetivo dar maior ênfase aos picos, permitindo um aumento da resolução e a eliminação dos efeitos aditivos. Os métodos derivados podem também eliminar interferências causadas pelo tamanho de partículas [14].

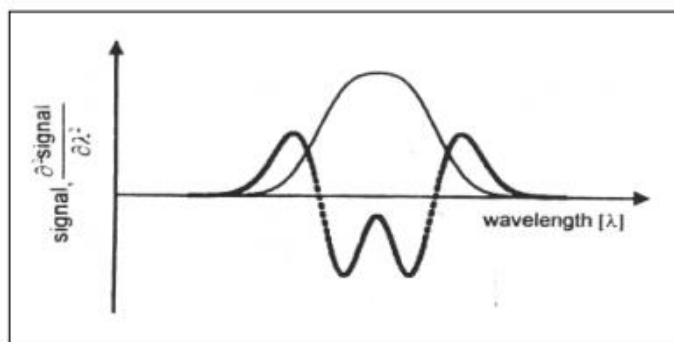


Fig. 11 - Aumento da resolução através do uso de derivadas: sinal original (linha sólida) e aplicação da segunda derivada ao sinal original (linha pontilhada). (adaptada a referência 25)

Quando aplicados métodos derivativos aos espectros originais, o ruído existente neste é amplificado, a menos que o método seja sensível a variações de sinal, podendo ser necessário antes da aplicação de derivadas o uso de filtros que suavizem o ruído.

A aplicação da primeira derivada tem como objetivo a remoção de efeitos aditivos da linha de base do espectro e evidência, maioritariamente, características mais pequenas, face às características com maior intensidade. No caso da segunda derivada, os mesmos sinais podem ser estudados, apesar do ruído espectral ser mais forte, sendo por isso aplicada em zonas do espectro muito restritas.

4.3.3. Normalização

A normalização é um tipo de pré-tratamento que tem como objetivo reduzir a influência de variações indesejadas, tais como, variação do percurso ótico, efeitos causados pela fonte e pelo detetor, etc., presentes no conjunto de dados, garantindo que cada observação é representada de forma adequada e consistente.

Este processo matemático consiste na divisão de cada uma das variáveis de uma dada amostra i por uma constante ou fator de normalização. De um modo geral, é através do emprego de um valor máximo, médio ou mínimo da intensidade apresentada por um comprimento de onda do espectro, que os dados de cada espectro serão escalados. Como resultado obtém-se todas as amostras numa mesma escala.

$$x_{ij} = \frac{x_{ij}}{\|x_i\|}, j = 1, 2, \dots, J$$

Equação 17

4.3.4. Centralização da Média

A centralização da média ou *Mean Centering* baseia-se na subtração da absorvância em cada comprimento de onda pela absorvância média a esse mesmo comprimento de onda para o total conjunto de amostras. Isto é, o valor de absorvância média resultante será igual a zero em cada comprimento de onda [26].

Do ponto de vista estatístico, a centralização tem como objetivo prevenir que os pontos mais distantes do centro dos dados tenham maior influência que os pontos mais próximos, dando desse modo importância não à intensidade verificada, mas antes à sua distância do valor médio. Com a utilização deste pré-tratamento, geralmente, diminui-se o número de componentes principais a utilizar na construção da calibração.

4.4. Estratégia para a Construção das Calibrações

Uma dos principais objetivos dos métodos de análise multivariável é a construção de calibrações com capacidade de prever, com a máxima exatidão possível, as propriedades que se pretendem analisar de amostras desconhecidas. O processo que é seguido para a criação e desenvolvimento de uma calibração envolve a passagem por várias etapas. A tabela 1 pretende descrever de um modo sucinto cada uma das etapas deste processo.

Tabela 1: Etapas do processo de construção de calibrações de um modelo NIR ^[20, 24].

Etapa	Objetivo
1 ^a	Seleção de um conjunto de amostras representativas da propriedade a calibrar e da variabilidade existente.
2 ^a	Determinação por métodos de referência dos valores das propriedades a calibrar, de um modo preciso e exato, o rigor dos valores obtidos irá refletir-se na calibração.
3 ^a	Aquisição dos espectros NIR, obtendo a informação físico-química de uma forma reprodutível.
4 ^a	Aplicação de técnicas quimiométricas, de modo a que seja possível reduzir as contribuições indesejadas nos espectros.
5 ^a	Construção da calibração, com o estabelecimento de uma relação espectro-propriedade, utilizando para esse efeito métodos multivariados.
6 ^a	Validação da calibração, de modo a que seja assegurada a exatidão com que prevê a(s) propriedade(s) de interesse em amostras diferentes das utilizadas na calibração.
7 ^a	Disponibilização da aplicação para o controlo de qualidade, para que seja possível prever, com rapidez, a(s) propriedade(s) de interesse em amostras a controlo.

4.5. Validação da Calibração

Após a aquisição do conjunto de amostras que irão integrar a calibração é efetuada uma divisão, aleatória, em dois novos conjuntos, em que no primeiro conjunto será utilizado para estabelecer o modelo de calibração (3/4 das amostras), e o segundo conjunto de amostras será utilizado para fazer-se a validação da calibração (1/4 das amostras). A distribuição uniforme das amostras pelos dois conjuntos é um fator importante, pois não poderão ser atribuídos diferentes pesos às regiões com maior ou menor peso de amostras, o que conduziria a uma debilitação da capacidade de previsão da calibração construída.

A seleção do número ótimo de componentes principais ou dos fatores é de grande relevo para a determinação da qualidade do modelo PLS obtido. Quando é feita uma escolha de um número de fatores menor do que os necessários, pode corresponder a uma calibração que origine resultados pouco precisos, pois possui menos informação do que a necessária na realidade (*Underfitting*), pelo contrário, com um elevado número de fatores a análise é deteriorada, uma vez que é incorporado mais ruído e variações sistemáticas (*Overfitting*). Assim sendo, no processo de construção da calibração é da máxima importância encontrar o número ideal de fatores que irão fornecer uma melhor previsão relativamente às novas amostras, com menor um erro, sendo os

métodos com um menor número de fatores os mais desejados, uma vez que são os mais simples e, supostamente, os mais robustos.

Um dos métodos mais utilizados para se proceder à escolha do número de fatores é a validação cruzada. Neste procedimento, a cada n amostras uma é deixada de fora, de seguida $n-1$ amostras são usadas na calibração, a validação passa pela previsão da concentração da amostra que não foi usada na calibração. Este procedimento é empregado até que todas as amostras tenham sido excluídas pelo menos uma vez. Posteriormente, os valores previstos serão comparados com os de referência e determina-se o erro, sendo a calibração com o menor valor de erro a mais adequada.

Durante a realização da calibração são determinados dois tipos de erros, um que está associado ao conjunto de amostras inseridas na calibração e se denomina erro quadrático de calibração (SEC) e o erro do coeficiente de correlação (R), de acordo com as Equações 18 e 19, respetivamente. Para o grupo de amostras utilizadas na validação é calculado o erro quadrático de previsão (SEP) pela Equação 20.

$$SEC = \sum_{j=1}^n \sqrt{\frac{(y_{j,NIR} - y_{j,EXP})^2}{n-m-1}} \quad \text{Equação 18}$$

$$R = \frac{\sum_{j=1}^n (y_{j,NIR} - \bar{y}_{j,EXP})}{\sqrt{\sum_{j=1}^n (y_{j,NIR} - \bar{y}_{j,EXP})^2}} \quad \text{Equação 17}$$

$$SEP = \sum_{j=1}^n \sqrt{\frac{(y_{j,NIR} - y_{j,EXP})^2}{n-1}} \quad \text{Equação 20}$$

(n – número de amostras utilizadas na calibração; m – número de variáveis independentes; y_j – valores previstos pelo espectrofotómetro de infravermelho próximo (NIR) e medidos experimentalmente (EXP); $\bar{y}_{EXP}$ – média de todo os valores medidos experimentalmente)

Os erros SEC e SEP podem ser utilizados como critérios de diagnóstico de modelagem da calibração pretendida, no entanto, os seus valores não devem exceder em demasia o desvio-padrão da técnica analítica utilizada para determinar os valores de referência da propriedade.

Um dos métodos mais usados na escolha do número adequado de componentes principais é designada por validação cruzada ou *Cross Validation*, no entanto não foi a metodologia utilizada sendo apenas referidas algumas noções básicas de como se precederia caso fosse utilizado.

Assim, neste tipo de procedimento uma amostra é deixada de fora de um conjunto de n amostras, de seguida $n-1$ amostras são utilizadas na calibração, realizando-se uma previsão da concentração da amostra que não foi usada na construção da calibração. Este procedimento é realizado até que todas as amostras tenham sido excluídas uma vez. Seguidamente compara-se os

valores previstos com os valores de referência do parâmetro desejado, permitindo a determinação da raiz quadrada do erro médio quadrático de validação cruzada (RMSECV), o qual é representado pela Equação 21 [14].

$$RMSECV = \sqrt{\frac{(\sum_{i=1}^n y_{i \text{ referência}} - y_{i \text{ previsto}})^2}{n-1}}$$

Equação 21

(n – número de amostras inseridas na calibração; $y_{i \text{ referência}}$ – valor de referência da amostra i; $y_{i \text{ previsto}}$ – valor previsto para a amostra i com base no modelo elaborado)

Com base neste valor é possível proceder-se à escolha da melhor calibração de previsão, ou seja, a calibração que apresentar um menor valor de RMSECV será a mais adequado ao modelo de previsão de parâmetro desejado [20].

Para finalizar o modelo PLS elaborado é necessário proceder ao seu teste e para isso são usadas amostras externas à calibração de concentração conhecida que irão pertencer apenas ao conjunto de validação. Assim, com base neste procedimento é calculado um novo erro, designado de raiz quadrada do erro médio quadrático de previsão (RMSEP) [24]. Este erro traduz a diferença entre os valores previstos da amostra i ($y_{i \text{ previsto}}$) e os valores de referência da amostra i ($y_{i \text{ referência}}$) das n amostras do conjunto de validação.

$$RMSEP = \sqrt{\frac{(\sum_{i=1}^n y_{i \text{ referência}} - y_{i \text{ previsto}})^2}{n}}$$

Equação 22

A validação cruzada deverá ser utilizada como método de validação nas situações em que não existe número de amostras suficientes para que seja construído um grupo de amostras para a calibração e um grupo de amostras para a validação.

5. Equipamentos e Materiais

Na execução deste projeto de estágio foi utilizado um espectrofotómetro de infravermelho próximo da marca BÜHLER, modelo NIRVIS, equipado com uma sonda de transmitância e um acessório refletância difusa, permitindo o seu funcionamento no modo combinado de transletância.

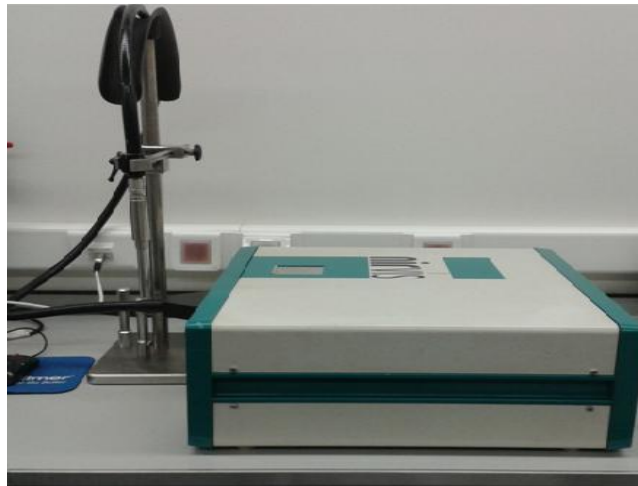


Fig. 12 – Espectrofotómetro NIRVIS da BÜHLER utilizado neste trabalho.

A aquisição dos espectros NIR foi realizada à temperatura ambiente, com intervalo de número de onda de 4008 a 9996 cm^{-1} , uma resolução de 12 cm^{-1} e um número mínimo de varrimentos de 6. Os espectros das substâncias ensaiadas foram obtidos em triplicado.

Tabela 2: Condições operacionais de funcionamento do equipamento NIR utilizado no estudo.

Resolução do laser/cm^{-1}	12
Detetor	InGaAs
Número de varrimentos por espectro	6
Percurso ótico/mm	1,5
Célula	Quartzo

O procedimento experimental seguido na execução do trabalho consistiu na aquisição de espectros de 78 vernizes transparentes de base solvente, 17 vernizes transparentes de base aquosa e de 68 tintas de base aquosa.

Para a aplicação de identificação de vernizes de base solvente e base aquosa foram construídos dois modelos NIR distintos, isto é, um para a identificação de vernizes de base solvente e um para a identificação de vernizes de base aquosa.

Seguidamente procedeu-se à divisão dos dois grandes grupos de vernizes em conjuntos mais pequenos, de modo a que o modelo NIR fosse o mais preciso possível. Dessa forma, o critério utilizado na divisão foi o tipo de resina(s) utilizada(s) na formulação. Assim, grupo de vernizes transparentes de base solvente foi dividido em dez subgrupos, que deram origem a dez calibrações, sendo estas agrupadas num único modelo NIR. Tendo em conta o mesmo critério de separação dos produtos seguido anteriormente, os vernizes transparentes de base aquosa foram divididos em três subgrupos, que originaram três calibrações, que posteriormente foram agrupadas num único modelo NIR.

No caso da aplicação de quantificação do teor de águas em tintas de base aquosa foram utilizadas todas as amostras ensaiadas numa mesma calibração, utilizando-se como valor de referência do teor de água, na construção do modelo NIR, o calculado teoricamente na formulação de cada formulação ensaiada.

Os materiais usados neste trabalho foram tintas de base aquosa e vernizes transparentes de base aquosa e de base solvente. As amostras ensaiadas provieram da unidade arquivo e da unidade produção, com um tempo de condicionamento de alguns dias até ao máximo de 1 ano, e foram fornecidas em frascos de vidro de volume variável (50 a 250 mL) e em embalagens de folha-de-flandres e foram utilizados sem a aplicação de nenhum pré-tratamento.

O tratamento dos espectros NIR obtidos, para a obtenção dos modelos PLS e PCA, foi realizado com o auxílio do *software* NIRCal 4.21, igualmente fornecido para marca BÜHLER.

Os espectros NIR obtidos no decurso deste estudo são representativos de cada produto ensaiado, uma vez evidenciam diferenças entre as formulações estudadas. Isto é, poderemos encontrar espectros em que a diferença está na intensidade das bandas NIR (diferentes percentagens dos componentes presentes nas formulações) ou espectros em que surgem novas bandas características (novos componentes presentes nas formulações).

6. Resultados e Discussão

Neste capítulo será apresentada uma das calibrações desenvolvidas para a criação do modelo NIR para a identificação de vernizes, e seguidamente, será apresentado parte do estudo iniciado para a obtenção do modelo de previsão do teor de água em tintas de base aquosa, pois este não foi concluído de modo a que pudesse ser implementado no decurso do estágio.

6.1. Identificação de Vernizes

Cada produto analisado neste estudo possui um espectro próprio, que o irá diferenciar dos restantes. Neste tipo de produtos poderá encontrar-se vernizes com a mesma formulação, no entanto, os constituintes poderão estar presentes em diferentes percentagens, ou ainda, poderá encontrar-se vernizes com formulações completamente distintas. Este tipo de diferenças permite distinguir os vários produtos na espectroscopia de infravermelho próximo. Para a construção deste modelo foi utilizado o método multivariável da análise dos componentes principais (PCA).

A Figura 13 mostra os espectros originais de vários vernizes A, a partir dos quais se pode verificar os vários produtos inseridos nesta calibração (cada produto é diferenciado por um cor).

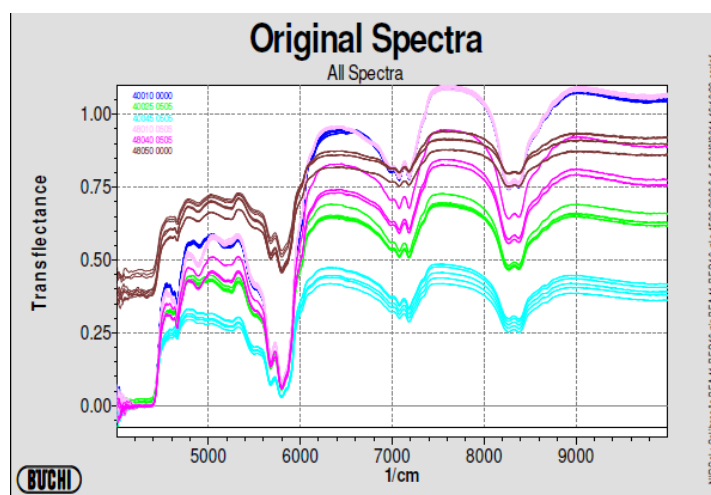


Fig. 13 – Espectros de absorvância originais, da região NIR dos vernizes A.

Os espectros dos diferentes produtos mostram as características químicas de cada substância presente na formulação. No entanto, nem todas estas diferenças estão associadas às pesquisadas no estudo. As “falsas” variações podem surgir devido à má homogeneização da amostra, ou ainda, à variação do tamanho de partícula. Assim, este tipo de condicionantes e a sua

influência podem ser reduzidas ou, parcialmente, eliminadas com a ajuda da aplicação de pré-tratamentos apropriados citados na secção 3.3, durante o desenvolvimento da calibração.

Assim, o primeiro passo consistiu pela aplicação de vários pré-tratamentos, de modo a poder escolher-se qual o mais vantajoso, ou seja, qual o que apresenta um melhor conjunto de parâmetros que permitissem melhorar a calibração. No entanto, como já foi referido anteriormente (secção 3.4), deverá ter em atenção o número de componentes principais ou fatores, como é designado no *software*, uma vez que poderia estar a utilizar-se um número excessivo componentes principais ou considerar um elevado número de *outliers*.

No caso em estudo foi utilizado como pré-tratamento uma normalização, usualmente foi este o pré-tratamento mais utilizado nos espectros NIR estudados neste trabalho, antecedendo qualquer outro tipo de pré-tratamento a ser aplicado (Figura 14).

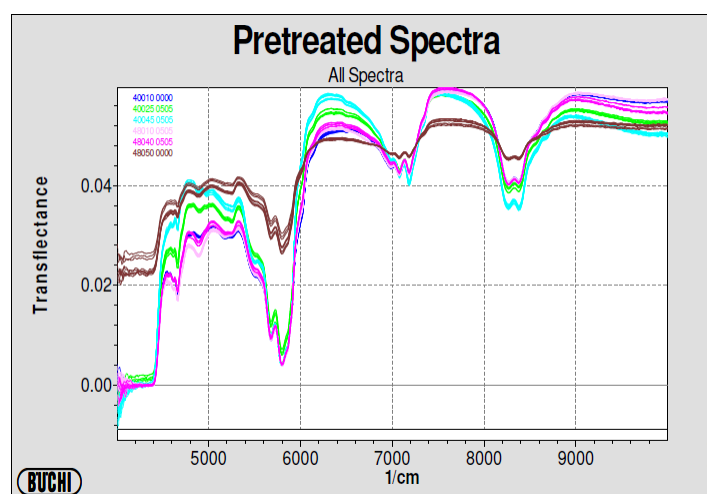


Fig. 14 – Espectros de absorvância, pré-tratados, da região NIR dos vernizes A.

Após a seleção do pré-tratamento a aplicar ao conjunto de espectros NIR, o número de componentes principais PCA a ser utilizado na calibração foi o passo seguinte na obtenção da melhor calibração. Para esta escolha o *software* fornece vários parâmetros que permitem fazer a seleção mais adequada do número de componentes principais, de modo a que não se verifique *overfitting* ou *underfitting*. Estes parâmetros serão apresentados seguidamente e consistem no gráfico *X-PRESS* e no gráfico *Factors*, na forma das Figuras 15 e 16, respetivamente.

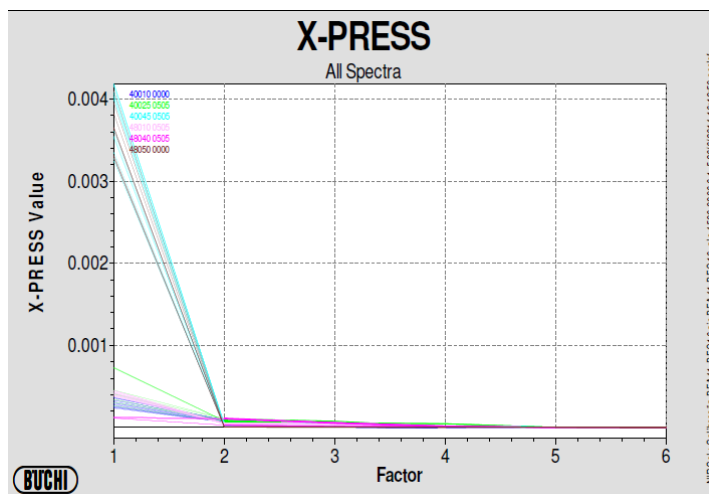


Fig. 15 – Gráfico *X-PRESS* com base no qual se poderá seleccionar o número de componentes principais.

O gráfico *X-PRESS* (*Predicted Residual Sum Squared*) consiste na soma dos quadrados dos erros de previsão residuais e permite observar o efeito do número necessário de componentes principais na construção do espetros. Assim, a escolha do número de componentes principais deverá recair sob aquele que a partir do qual não se observa variações significativas no gráfico, ou seja, a partir do qual o valor do *X-PRESS* é constante. Este valor é calculado pelo *software* para os vários componentes principais com base na Equação 23.

$$\text{PRESS} = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad \text{Equação 23}$$

(n – número de amostras do conjunto de normalização; y_i – valor de referência; $\hat{y}_i$ – valor previsto pelo modelo para a i -ésima amostra)

Com base no que foi descrito, no caso desta calibração o valor de 6 componentes principais, é aquele a partir do qual estão satisfeitos estes critérios.

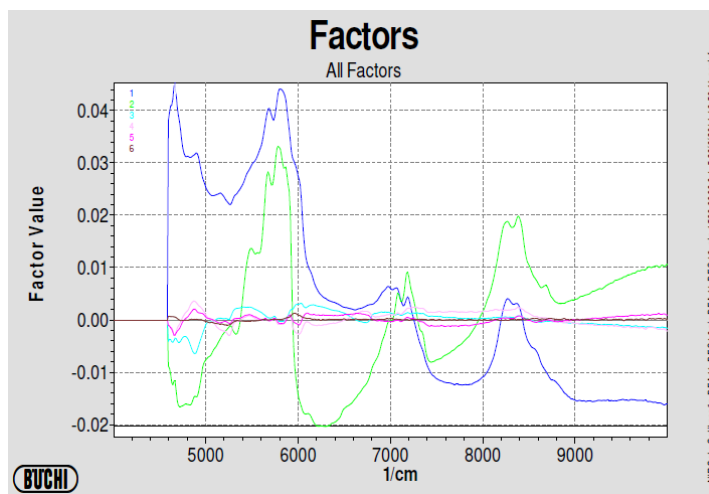


Fig. 16 – Gráfico Factors com base no qual se poderá selecionar o número de componentes principais.

O gráfico *FACTORS* consiste na diferença entre os espectros estudados, em que apenas os principais componentes com características espectrais devem ser utilizados.

Com base neste último gráfico também se poderia proceder à seleção do número de componentes principais. Neste ponto poderá avaliar-se a importância dos componentes principais, pois valores elevados significam mais informação, em sentido oposto, valores pequenos significam menos informação. Assim, também o valor de 6 componentes principais pareceu ser o mais adequado a aplicar na construção da calibração em estudo.

Após a aplicação dos pré-tratamentos mais apropriados e de selecionado o número de componentes principais mais importantes é criada uma nova ilustração da informação espectral, em que o objetivo é apresentar as principais diferenças entre os dados analisados. A esta nova representação (Figura 17) designa-se de *clusters* e é obtida a partir do método de análise PCA. Cada *cluster* formado representa uma propriedade, ou seja, cada propriedade representa um produto inserido na calibração, e pretende-se que estes se encontrem bem separados nas regiões de tolerância.

Tabela 3: Parâmetros utilizados nas calibrações desenvolvidas para o modelo NIR qualitativo dos vernizes transparentes de base solvente.

Calibração	λ/cm^{-1}	Fatores Primários	Pontos na calibração	Pontos na validação	Pré-tratamento	Q-Value
A	4596-9996	8	27	9	Normalização	0,972297
B	4596-9996	7	43	14	MSC + Normalização	0,907008
C	4596-9996	7	39	12	Normalização	0,912189
D	4596-9996	6	27	9	Normalização + 1ª Derivada	0,920919
E	4596-9996	6	43	14	Normalização	0,917254
F	4596-9996	10	75	24	Normalização + 1ª Derivada	0,913545
G	4596-9996	9	63	21	Normalização + 1ª Derivada	0,813123
H	4596-9996	6	48	36	Normalização + 1ª Derivada	0,905080
I	4596-9996	8	124	96	Normalização + 1ª Derivada	0,892155
J	4596-9996	9	50	16	Normalização + 1ª Derivada	0,918555

Tabela 4: Parâmetros utilizados nas calibrações desenvolvidas para o modelo NIR qualitativo dos vernizes transparentes de base aquosa.

Calibração	λ/cm^{-1}	Fatores Primários	Pontos na calibração	Pontos na validação	Pré-tratamento	Q-Value
A	4596-9996	6	32	16	Sem pré-tratamento	0,905275
B	4596-9996	6	28	14	Normalização	0,93941
C	4596-9996	5	20	10	Normalização	0,97973

6.1.1. Validação do Modelo NIR Qualitativo

O *software* utilizado neste trabalho, recorre à construção de dois grupos distintos de espectros, em que num primeiro grupo se encontram os espectros que serão usados na calibração (3/4) e num segundo grupo encontram-se os espectros que serviram para validar a calibração (1/4), como referido na seção 3.4.

No entanto, para realizar uma validação mais robusta optou-se pela validação externa do modelo NIR. Isto é, recorreu-se a novas amostras não utilizadas durante o processo, realizando alterações à formulação das mesmas. Deste modo, objetivo seria apurar-se quais os limites de aprovação/reprovação do modelo, com base em diferentes procedimentos de validação.

Os procedimentos seguidos para a validação do modelo qualitativo de identificação foram:

1. Troca da identificação do produto controlado, por uma em que a formulação fosse semelhante;
2. Alteração da percentagem de resina(s) presente(s) na formulação em 5, 7.5 e 10%;
3. Troca de uma resina da formulação, por uma outra semelhante na mesma percentagem;
4. A realização em paralelo do controlo de qualidade pelas técnicas de referência e o modelo NIR desenvolvido.

Os resultados destas validações foram, na grande maioria os espectados, os seja, o modelo aprovou os produtos nos casos em que deveria aprovar e reprovando quando isso deveria acontecer, como nos procedimentos 1 e 3.

Contudo, no procedimento 2, de alteração da percentagem de resinas presentes na formulação, quando se aplicada uma alteração de 5% na formulação o modelo aprovou as amostras testadas, uma vez que esta é a percentagem limite de erro na pesagem aceite na conceção dos produtos em fábrica. Na percentagem de 7.5% os resultados variaram entre a aprovação e a reprovação dos produtos analisados, uma vez que as amostras utilizadas provieram de lotes da produção, já poderia existir algum erro de pesagem que se refletisse na percentagem em estudo. Por fim, quando aplicada uma alteração de 10% na formulação, todas as amostras examinadas foram reprovadas pelo modelo desenvolvido.

No procedimento 4, foram controlados diversos produtos diferentes, em que os resultados obtidos foram na sua grande maioria a de aprovação pelo modelo (foram controlados 72 produtos, previamente controlados pelas técnicas de referência, dois quais 5 foram rejeitados). Após verificados os motivos da reprovação de certos produtos, prosseguiu-se com a otimização das calibrações nas quais estavam inseridos os produtos reprovados. Para a realização desta melhoria, procedeu-se à elaboração em laboratório das formulações em causa e com a percentagem de cada componente rigorosamente conhecida, posteriormente realizou-se a aquisição dos espetros destas amostras para serem inseridas nas calibrações.

As amostras rejeitadas foram, novamente testadas após a otimização do modelo, onde estas foram aprovadas.

6.2. Quantificação do Teor de Água em Tintas

O teor de água em tintas aquosas é um fator muito importante, pois determina várias características físicas, químicas e mecânicas no desempenho que estas terão como produto final.

Para a construção de um modelo de calibração desta propriedade, utilizou-se um conjunto de amostras de arquivo, em que o valor do teor de água usado, na elaboração da calibração, foi o teórico calculado na formulação, e que estava contido num intervalo de 24,15 a 94,08%.

Assim, para o desenvolvimento do modelo NIR quantitativo para a determinação do teor de água em tintas de base aquosa foi usado o método multivariável da regressão dos mínimos quadrados parciais, sendo os procedimentos adotados muito semelhantes aos seguidos na secção 6.1.

A partir do *software* foi aplicado o método PLS, iniciando pelo emprego de vários pré-tratamentos de modo a escolher-se o mais vantajoso, assim como, a escolha do número mais adequado de variáveis latentes, também designadas de fatores, ou seja, o número que irá conter a informação espectral mais importante. Ainda com base no *software* é possível determinar os

valores do erro quadrático de previsão, o erro quadrático de calibração e o coeficiente de correlação, que permitiriam determinar a qualidade da calibração elaborada, como também, permitem determinar o intervalo de comprimento de onda mais adequado à propriedade em estudo.

Assim, para a construção do modelo de regressão multivariável foram, inicialmente, traçados os espectros de absorvância na região NIR das tintas de base aquosa (figura 19). Seguidamente foram aplicados diversos pré-tratamentos de forma a ser escolhido aquele que apresenta-se o menor valor do erro quadrático de previsão.

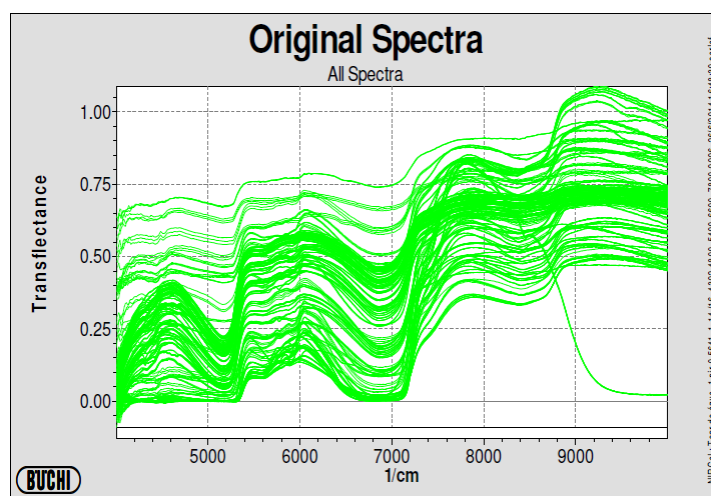


Fig. 18 – Espectros de absorvância na região NIR de todas as amostras em estudo.

A escolha do pré-tratamento mais adequado tem por base o erro SEP, a região espectral consoante o parâmetro de previsão, o número ótimo de variáveis latentes e, ainda, o menor número de amostras *outliers* que este considera.

Na elaboração deste estudo foram testados vários tipos de pré-tratamentos, sendo que três deles apresentaram os melhores resultados, a aplicação de uma suavização, sem aplicação de pré-tratamentos e a aplicação de uma suavização seguida de uma normalização.

Na tabela 5, são apresentados os resultados para os diversos parâmetros calculados na construção das calibrações, tais como o intervalo de comprimento de onda, o número ótimo de variáveis latentes, etc., que tendem a permitir a escolha da calibração que melhor se adequa ao modelo de previsão do parâmetro pretendido, neste caso, a quantificação do teor de água em tintas de base aquosa.

Tabela 5: Resultado de alguns parâmetros determinados para o modelo NIR quantitativo de determinação do teor de água em tintas de base aquosa, utilizando diferentes pré-tratamentos.

Parâmetro	Suavização	Sem pré-tratamento	Suavização+ Normalização
λ/cm^{-1}	4392-4800; 5400-6600; 7800-9996	4392-4800; 5400-6600; 7800-9996	4392-4800; 5400-6600; 7800-9996
Fatores Primários	14	14	14
R calibração	0,944196	0,943353	0,944589
R validação	0,944669	0,944029	0,944122
SEC/%	5,74844	5,79047	5,72874
SEP/%	5,77672	5,80480	5,81128
Consistência	99,5104	99,7531	98,5797
Q-Value	0,564116	0,563955	0,550514

As calibrações desenvolvidas foram avaliadas usando o erro quadrático de previsão e o erro do coeficiente de correlação da calibração (tabela 6).

Tabela 6: Parâmetros utilizados para avaliar as calibrações desenvolvidas no estudo.

Parâmetro	Suavização	Sem pré-tratamento	Suavização+ Normalização
R calibração	0,944196	0,943353	0,944589
SEP/%	5,77672	5,80480	5,81128

A Figura 19 mostra a calibração que utiliza a suavização como pré-tratamento e que poderia ser aquela que se usaria no modelo NIR quantitativo.

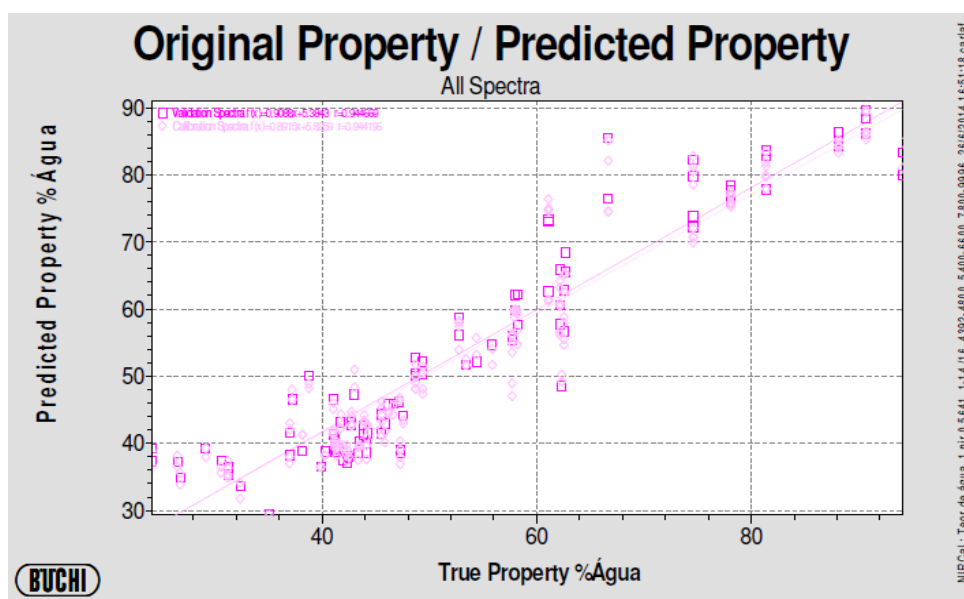


Fig. 19 – Gráfico %água obtida/ %água teórica.

No gráfico apresentado pela Figura 19 é possível observar-se duas retas representativas dos valores correspondentes às amostras inseridas na calibração e a os valores correspondentes às amostras inseridas na validação, das quais são apresentadas em seguida as respetivas equações:

$$\square \text{ Validation Spectra } f(x) = 0.9088x + 5.3843 \quad r = 0.9444669 \quad \text{Equação 25}$$

$$\diamond \text{ Calibration Spectra } f(x) = 0.8915x + 5.8259 \quad r = 0.944196 \quad \text{Equação 26}$$

Assim, com base nos resultados obtidos, foi possível observar que as três calibrações expostas tiveram valores semelhantes para os parâmetros avaliados, sendo os mais importantes de referir o SEC e SEP, assim como, o número de variáveis latentes (componentes principais) foi igual nos três caso (14). A correlação linear, R, determinada entre os valores de referência e os espectros NIR apresentam valores muito semelhantes, de aproximadamente 0,89 tanto para a calibração como para a validação.

Um outro parâmetro que permite avaliar a melhor calibração é a consistência, definida como a percentagem de SEC dividida pelo SEP e multiplicada por 100. Este valor deve estar o mais próximo possível de 100, no entanto, aceitam-se valores entre 80 e 100, e a qual suporta a qualidade da calibração desenvolvida. No estudo apresentado este parâmetro variou entre 0,995104 e 0,985797.

Durante a elaboração de cada uma das calibrações o *software* fornece diversos gráficos e diversos outros parâmetros que permitem fazer uma correta avaliação de cada uma das calibrações desenvolvidas. As figuras que se seguem são os gráficos resultantes de alguns parâmetros já referidos.

A Figura 20 permite fazer uma avaliação do número ótimo de variáveis latentes para os intervalos de comprimento de onda escolhidos no estudo, neste caso os intervalos foram de 4392 – 4800; 5400 – 6600; 7800 – 9996 cm^{-1} .

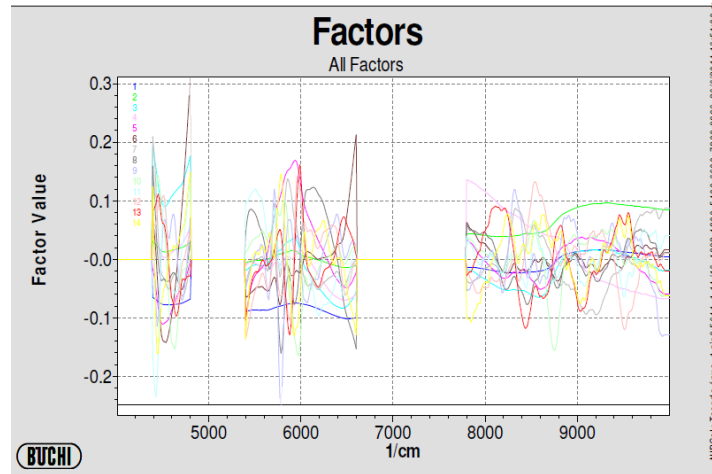


Fig. 20 – Gráfico *Factors* com base no qual se poderá seleccionar o número de variáveis latentes.

Na Figura 21, é possível observar o comportamento da consistência para as várias variáveis latentes, permitindo verificar quando esta fica perto do 100%.

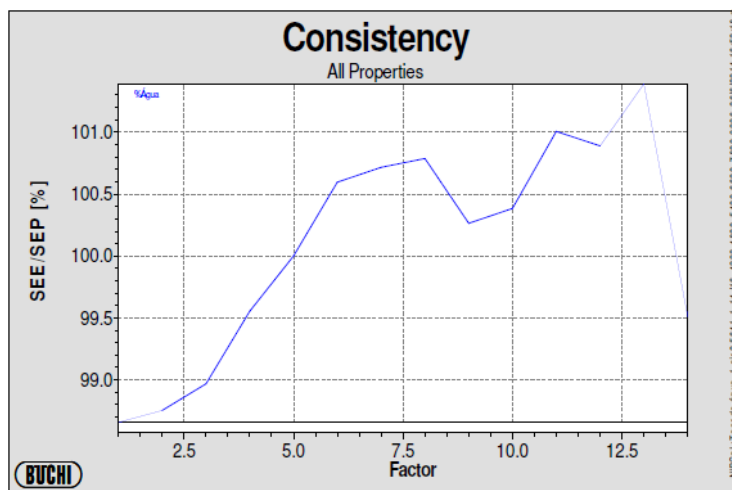


Fig. 21 – Gráfico *Consistency* com base no qual se poderá verificar a qualidade da calibração.

7. Conclusão e Perspetivas de Trabalho Futuro

Um dos principais objetivos definidos no início da execução deste trabalho consistia na elaboração de uma aplicação que permitisse realizar em tempo real o controlo de qualidade de vernizes de base aquosa e vernizes de base solvente. Partindo deste princípio foram elaborados dois modelos NIR que permitissem a correta identificação por “pass-fail” destes produtos.

Assim, foram construídas no total treze calibrações, sendo divididas em dez calibrações para o conjunto de vernizes transparentes de base solvente e em três calibrações para o conjunto de vernizes transparentes de base aquosa, com base em técnicas quimiométricas e utilizando o método multivariável de análise dos componentes principais para esse efeito.

O tipo de resina ou resinas utilizadas na formulação dos vernizes serviu como critério de divisão dos grupos maiores de vernizes, de modo que a análise fosse baseada em produtos mais semelhantes entre si e por isso com uma maior previsão. No entanto, esta separação gera um problema no número de amostras ensaiadas para cada produto, uma vez que em alguns casos não existia mais do que duas ou três amostras para certos produtos. Este facto torna as calibrações menos robustas do que o que seria de esperar.

Assim, será de esperar que futuramente seja realizado o aperfeiçoamento das calibrações, especialmente das referentes a produtos com uma menor número de amostras, permitindo desse modo tornar a calibração mais robusta.

No que respeita ao número de componentes principais necessários à construção de cada calibração devera-se realizar um estudo que permita verificar se em algumas calibrações não se procedeu ao *overfitting* e se com isso não se comprometeu o modelo de identificação.

Após a construção das calibrações e de as agrupar nos modelos NIR respetivos, procedeu-se à sua validação, neste caso, foi apenas iniciada a validação do modelo NIR de identificação e vernizes transparentes de base solvente.

Tal como anteriormente referido, a validação consistiu, principalmente, na adição de diferentes percentagens de resinas pertencentes à formulação e no controlo direto por NIR em comparação com os métodos de controlo de referência. No primeiro caso, quando a adição de resina foi de 10% obteve-se uma taxa de reprovação de 100%, no entanto, quando a percentagem de resina foi apenas de 5% a taxa de reprovação desceu significativamente. Neste caso, deverá futuramente estudar-se de um modo mais aprofundado qual o intervalo de erro do modelo NIR para que este aprove ou reprove um produto.

No segundo método de validação utilizado, a taxa de aprovação foi de cerca de 90%. Neste caso, os produtos reprovados pelo modelo NIR foram alguns em que as suas formulações seriam, relativamente, semelhantes à de outros produtos inseridos no modelo. Inicialmente, de forma a contornar este problema foi necessário recorrer a amostras produzidas em laboratório e de concentrações conhecidas, as quais foram usadas para enriquecer as respetivas calibrações. Assim, caso futuramente surja o mesmo problema, será necessário recorrer a amostras produzidas em laboratório de modo a minimizar os erros que possam acontecer no processo de controlo de qualidade.

Um segundo objetivo proposto para execução neste trabalho seria a elaboração de uma aplicação que permitisse a determinação do teor de água em tintas de base aquosa, a partir a construção de um modelo NIR, com base em técnicas quimiométricas utilizando o método multivariável da regressão dos mínimos quadrados parciais.

Assim, foram analisadas por NIR diversas amostras de tintas, com um teor de água variável, o que permitiu construir três possíveis calibrações. No entanto, estas não foram otimizadas, sendo necessário proceder ao seu aperfeiçoamento e enriquecimento com novas amostras que possam alargar e melhorar as gamas operacionais. Posteriormente será necessário proceder à validação, de modo a que o modelo seja capaz de quantificar com exatidão o teor de água em tintas.

De realçar que os modelos desenvolvidos, para a identificação de vernizes, requerem atualização nos casos em que haja alteração das formulações para que possam continuar a responder eficazmente. Após a adição de novas amostras aos conjuntos de calibração e validação, é necessário que se execute novamente os procedimentos de validação.

8. Bibliografia

- [1] CIN, 2009, *Manual de acolhimento*.
- [2] Canha, N., 2008, *Determinação da estabilidade oxidativa do biodiesel e da sua relação com as condições operatórias do produto*. Dissertação de Mestrado, Instituto Superior Técnico Universidade de Lisboa.
- [3] Almeida, F., 2009, *Espectroscopia de Infravermelho Próximo com Transformada de Fourier (FT-NIR) na Caracterização de Farinhas para Alimentação Pueril*. Tese de Doutoramento, Instituto Superior Técnico Universidade de Lisboa.
- [4] CIN, 2009, *Manual de Construção e Decoração*.
- [5] Goldschmidt, Artur and Streitberger, Hans-Joachim, 2003, *BASF handbook on basics of coating technology*, William Andrew.
- [6] Pasquini, C., 2003. *Near Infrared Spectroscopy: Fundamentals, Practical Aspects and Analytical Applications*. J. Braz. Chem. Soc., vol. 14, no. 2, pp. 198-219.
- [7] Costa, P., Alves, T. e Menezes, J., 2001, *Monitorização de Processos Biológicos: Monitorização at-line de Bioprocessos Industriais com FT-NIR*. Boletim Biotechnol. (Soc. Port. Bioquímica) no.69, pp. 11-16.
- [8] Ciurczak, E., 2006, *Near-infrared Spectroscopy*. Modern Instrumental Analysis. vol. 47, pp. 157 - 170.
- [9] Araújo, A., 2007, *Monitoramento de processos em tempo real via espectrofotometria no infravermelho próximo*. Tese de Doutoramento, Universidade Tiradentes – UNIT.
- [10] Osborne, B. G., 2000, *Near-infrared Spectroscopy in Food Analysis*. *Encyclopedia of Analytical Chemistry*. John Wiley & Sons Ltd.
- [11] Ilharco, L., 2009, *Métodos Avançados de Análise e Técnicas Hifenadas*. Faculdade de Ciências Universidade de Lisboa.
- [12] Bakeev, K., 2005, *Analytical Technology*. Blackwell Publishig Ltd.
- [13] Conzen, J. P., 2006, *Multivariate Calibration. A Practical Guide for the Method Development in the Analytical Chemistry*. Bruker Optick GmbH. 2nd edition.
- [14] Perkins, W. D., 1986, *Fourier Transform – Infrared Spectroscopy – Instrumentation*. Journal Chemometrics, pp. A5 – A10.
- [15] MPA – Multiple Purpose Analyzers, 2007, *User Manual*. 2nd updated, Cruker Optik Gmhbb.

- [16] Sabin, J. G., Ferrão, M. F., Furtado, J. C., 2004, *Análise multivariada aplicada na identificação de fármacos antidepressivos. Parte II: Análise por componentes principais (PCA) e o método de classificação SIMCA*. Brazilian Journal of Pharmaceutical Sciences, vol. 40, pp. 387-396.
- [17] Birth, G. S., Hecht, H. G., 1987, *The physics of near-infrared reflectance in near-infrared technology in the agricultural and food industries*. St. Paul: American Association of Cereal Chemists.
- [18] Ollinger, J. M., Griffiths, P. R., 1993, *Effects of sample dilution and particle size / morphology on diffuse reflection spectra of carbohydrate systems in the near and mid-infrared. Part I: single analytes*. Applied Spectroscopy, pp. 687-694.
- [19] Naes, T., 2002, *A User-Friendly Guide to Multivariate Calibration and Classification*. NIR Publications.
- [20] Leitão, T., 2012, *Aplicações da espectroscopia de infravermelho próximo em Ciências Farmacêuticas*. Dissertação de Mestrado, Faculdade de Ciências Universidade Fernando Pessoa
- [21] Blanco, M., Villarroya, I., 2002, *NIR spectroscopy: a rapid-response analytical tool*. Trends in analytical chemistry, vol. 21, no. 4.
- [22] Poppi, Ronei J., 2000, *Avaliação do uso de métodos quimiométricos em análise de solos*. Química Nova, vol. 23 (4), pp. 547 – 556.
- [23] O'Haver, T., *Introduction to Signal Processing in Chemical Analysis*. Department of Chemistry and Biochemistry, University of Maryland.
- [24] Naes, T., Isaksson, T., 1994, *Some aspects of scatter-correction in calibration*. Part I, NIR News. vol. 5 (2), pp. 4-5, 15.
- [25] Siesler, H., 2002, *Near-infrared spectroscopy: principles, instruments, applications*. Weinheim, Wiley-VCH.
- [26] Thomas, E. V., 1994, *A primer on multivariate calibration*. Analytical Chemistry. pp A795-A804.