



M 2020

DETERMINAÇÃO DO TEOR DE RESINA E TEOR DE VOLÁTEIS EM PAPEL IMPREGNADO USANDO ESPECTROSCOPIA DE INFRAVERMELHO PRÓXIMO

ANA SOFIA DOS REIS GONÇALVES

DISSERTAÇÃO DE MESTRADO APRESENTADA

À FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO EM
ENGENHARIA QUÍMICA

Mestrado Integrado em Engenharia Química

Determinação do Teor de Resina e Teor de Voláteis em papel impregnado usando espectroscopia de infravermelho próximo

Dissertação de Mestrado

de

Ana Sofia dos Reis Gonçalves

Desenvolvida no âmbito da unidade curricular de Dissertação

realizado em

EuroResinas - Indústrias Químicas S.A.



Orientador na FEUP: Prof. Luísa Carvalho

Prof. Fernão Magalhães

Coordenador na EuroResinas: Doutora Nádia Paiva



julho de 2020

Agradecimentos

Primeiramente, gostaria de agradecer à Faculdade de Engenharia da Universidade do Porto e EuroResinas - Indústrias Químicas S.A. por permitirem a realização desta tese em ambiente empresarial.

De seguida, gostaria de agradecer à coordenadora deste projeto, a Doutora Nádia Paiva, e aos orientadores na FEUP, a Professora Luísa Carvalho e o Professor Fernão Magalhães, por todo o apoio e disponibilidade no esclarecimento de dúvidas ao longo deste projeto. Gostaria também de agradecer ao Professor Jorge Martins e ao Engenheiro Roberto Magalhães pela partilha de conhecimento.

Não poderia deixar de agradecer a toda a equipa da EuroResinas pela simpatia e às minhas colegas de estágio com quem tive o prazer de partilhar esta experiência.

Todo este projeto não teria sido possível sem o apoio incondicional da minha família e dos meus amigos. Assim, gostaria de deixar um agradecimento especial aos meus pais, Amélia e José, ao meu irmão, Diogo, e ao meu namorado, Daniel.

Apesar dos tempos incertos vividos no decorrer desta dissertação, as medidas tomadas pelos orientadores deste projeto e pela equipa da EuroResinas, permitiram minimizar o impacto da pandemia nesta dissertação. Considero que esta experiência foi bastante positiva e enriquecedora, pelo que, certamente me preparou para entrada no mercado de trabalho.

A Prof.^a Luísa H. Carvalho e o Prof. Fernão Magalhães, orientadores desta dissertação, são membros integrados do LEPABE - Laboratório de Engenharia de Processos, Ambiente Biotecnologia e Energia, financiado por: Financiamento Base - UIDB/00511/2020 da Unidade de Investigação - Laboratório de Engenharia de Processos, Ambiente, Biotecnologia e Energia - LEPABE - financiada por fundos nacionais através da FCT/MCTES (PIDDAC).

Projeto Innosurf - Superfícies Inovadoras, projeto financiado pelo programa Portugal 2020 com o n.º POCI-01-0247-FEDER-33768, cofinanciado pelo Fundo Europeu de Desenvolvimento Regional (FEDER) através do COMPETE2020-Programa Operacional Competitividade e Internacionalização (POCI).

Resumo

No processo de produção do papel impregnado, os dois parâmetros importantes a controlar são o teor de resina (RC) e o teor de voláteis (VC). Atualmente, na fábrica da EuroResinas, este processo é controlado de modo *off-line* e requer entre cinco e dez minutos para obtenção de resultados. A implementação da espectroscopia de infravermelho próximo (NIR) na linha de produção permite acelerar e automatizar esta determinação. O objetivo deste estudo consistiu na criação de modelos para a quantificação destas propriedades recorrendo a métodos quimiométricos como a Regressão por mínimos quadrados parciais (PLS).

Previamente à construção de modelos, estudou-se o efeito dos parâmetros escolhidos na aquisição dos espectros de modo a reduzir, sem prejudicar os resultados obtidos, o tempo de análise. Conclui-se que, das combinações estudadas, a que permitiu chegar a este objetivo foi: 16 cm^{-1} de resolução, 128 varrimentos e 20 kHz de frequência de varrimento, correspondendo a um tempo de aquisição de 22 segundos.

Procedeu-se à construção de modelos para a quantificação do teor de resina e voláteis. Idealmente, um modelo adequado deve apresentar, nomeadamente, alto R^2 e baixa raiz quadrada do erro médio quadrático da validação cruzada, $RMSE_{CV}$. Esta variável é indicativa da diferença entre os valores previstos pelo modelo e os obtidos pelo método de referência (pesagem diferencial). Deste modo, este valor deveria ser o mais próximo possível de zero.

O primeiro papel impregnado estudado foi o B1. Para este utilizou-se o procedimento das médias para a aquisição de espectros. O modelo construído para a quantificação do RC após o primeiro banho, RC_1 , obteve um valor de $RMSE_{CV}$ de 0,929 % (m/m) com um R^2 de 0,90 para um número de fatores de 5. No caso do modelo construído para o RC após o segundo banho, RC_2 , o $RMSE_{CV}$ obtido foi de 1,01 % (m/m) com um R^2 de 0,82 para um número de fatores de 6. Optou-se por criar um modelo geral para a quantificação do teor de resina, RC . Para este, o $RMSE_{CV}$ obtido foi de 1,05 % (m/m), com um R^2 de 0,98 para um número de fatores de 6. Foi também criado um modelo para a quantificação do VC após o segundo banho, VC_2 . Assim, o $RMSE_{CV}$ obtido foi de 0,386 % (m/m), com um R^2 de 0,91 com um número de fatores de 5. Criou-se novamente um modelo geral, porém, para a quantificação do teor de voláteis, VC . Para este, o $RMSE_{CV}$ obtido foi de 0,4 % (m/m), com um R^2 de 0,91 para um número de fatores de 6.

Estudou-se também um segundo papel impregnado com apenas um banho, designado por L. Para este foram estudados dois procedimentos para a aquisição de espectros: o das médias e o do pórtico. O modelo construído para a quantificação do RC após o primeiro banho, RC_1 , utilizando o procedimento das médias, obteve um $RMSE_{CV}$ de 1,47 % (m/m), com um R^2 de

0,91 para um número de fatores de 7. Por sua vez, para o procedimento do p $\acute{o}$ rtico, o $RMSE_{CV}$ obtido foi de 1,06 % (m/m), com um R^2 de 0,96 para um número de fatores de 4. No caso do VC, o modelo obtido utilizando o procedimento da m $\acute{e}$ dias obteve um $RMSE_{CV}$ de 0,804 % (m/m), com um R^2 de 0,86 para um número de fatores de 4. Utilizando o procedimento do p $\acute{o}$ rtico, o $RMSE_{CV}$ obtido foi de 0,624 % (m/m), com um R^2 de 0,93 com um número de fatores de 9.

Conclui-se que a capacidade de previs $\tilde{a}$ o dos modelos foi mais baixa que o desejado. No entanto, os resultados obtidos permitiram aprofundar o conhecimento na $\acute{a}$ rea da espectroscopia de NIR. Espera-se que este estudo permita apoiar trabalhos futuros do mesmo tipo, de modo a proceder $\grave{a}$ implementa $\tilde{c}$ ao futura de m $\acute{e}$ todos deste tipo em ambiente fabril.

Palavras Chave: Espectroscopia de infravermelho pr $\acute{o}$ ximo, teor de resina, teor de vol $\acute{a}$ teis, m $\acute{e}$ todos quimiom $\acute{e}$ tricos.

Abstract

In the process of production of impregnated paper, the two main controlled parameters are the resin content (RC) and the volatile content (VC). Currently, in the industrial plant of EuroResinas, this process is controlled off-line and requires up to five to ten minutes to obtain these results. The implementation of near infrared spectroscopy in the production line allows the automatization of those measurements, as well as, the reduction of the amount of time needed to perform it. The main objective of this study was to create models for the quantification of these properties using chemometric methods such as, partial least squares regression (PLS).

Before the construction of these models, the effect of the chosen spectra acquisition parameters was studied in order to reduce the total acquisition time, without jeopardizing the quality of the obtained results. Amongst the studied combinations, the one who allowed this objective to be met was: 16 cm^{-1} of resolution, 128 scans and 20 kHz of scanning frequency, resulting in a total acquisition time of 22 seconds.

Models were then constructed to quantify the content of resin and volatiles. Ideally, a robust model should present, namely, a high R^2 and a low root mean square error of cross-validation, $RMSE_{CV}$. This variable is related to the difference between the predicted values by the model and the obtained values using the reference method (differential weighing). Thus, this value should be as close to zero as possible.

The first impregnated paper studied was B1. For this type of paper, the averaging procedure was used for the acquisition of spectra. The model for the quantification of the RC after the first bath, RC_1 , obtained a $RMSE_{CV}$ of 0,929 % (m/m) with a R^2 of 0,90 for a *rank* of 5. In the case of the model constructed for the quantification of RC after the second bath, RC_2 , the obtained $RMSE_{CV}$ was of 1,01 % (m/m) with a R^2 of 0,82 and a *rank* of 6. The hypotheses of a generic model for the quantification of the resin content, RC , was studied. The obtained $RMSE_{CV}$ for this model was of 1,05 % (m/m), with a R^2 of 0,98 for a *rank* of 6. A model for the quantification of the volatile content after the second bath, VC_2 , was also created. The obtained $RMSE_{CV}$ for this model was of 0,386 % (m/m), with a R^2 of 0,91 for a *rank* of 5. Lastly, a generic model for the quantification of the volatile content, VC , was studied. For this model, the obtained $RMSE_{CV}$ was of 0,4 % (m/m), with a R^2 of 0,91 for a *rank* of 6.

A second type of impregnated paper submitted to only one bath, L, was studied. For this type of paper two different procedures for the acquisition of spectra were studied: the averaging and the gantry procedure. The model constructed for the quantification of RC after the first bath, RC_1 , using the averaging procedure, resulted in a $RMSE_{CV}$ of 1,47 % (m/m), with a R^2 of 0,91 for a *rank* of 7. With the use of the gantry procedure the obtained $RMSE_{CV}$ was of

1,06 % (m/m), with a R^2 of 0,96 for a *rank* of 4. In the case of the volatile content, the model obtained using the averaging procedure resulted in a $RMSE_{CV}$ of 0,804 % (m/m), with a R^2 of 0,86 for a *rank* of 4. With the use of the gantry procedure the obtained $RMSE_{CV}$ was of 0,624 % (m/m), with a R^2 of 0,93 for a *rank* of 9.

It was concluded that the predictive capacity of the created models was lower than desired. However, the obtained results contributed to further the knowledge in the area of near infrared spectroscopy. It is expected this study will facilitate future studies in this area allowing the implementation of similar methods in the production line.

Keywords: Near infrared spectroscopy, Resin content, volatile content, chemometric methods.

Declaração

Declara, sob compromisso de honra, que este trabalho é original e que todas as contribuições não originais foram devidamente referenciadas com identificação da fonte.

Ana Sofia dos Reis Gonçalves

Porto, 25 de junho de 2020

Índice

1	Introdução.....	1
1.1	Enquadramento e Apresentação do Projeto	1
1.2	Apresentação da Empresa	1
1.3	Contributos do(a) autor(a) para o Trabalho.....	2
1.4	Organização da Tese	2
2	Contexto e Estado da Arte	3
2.1	Resinas amino.....	3
2.1.1	Resinas melamina-formaldeído (MF)	3
2.2	Papel impregnado.....	4
2.2.1	Produção de papel impregnado	5
2.3	Espectroscopia de infravermelho próximo.....	6
2.3.1	Funcionamento do equipamento.....	8
2.4	Métodos quimiométricos.....	10
2.4.1	Pré-processamentos	10
2.4.2	Análise dos componentes principais (PCA)	12
2.4.3	Regressão por mínimos quadrados parciais (PLS).....	14
2.4.4	Validação dos modelos	15
3	Materiais e Métodos	17
3.1	Impregnação laboratorial de papel	17
3.2	Papel impregnado.....	17
3.3	Teor de resina (RC) e Teor de voláteis (VC)	18
3.4	Espectroscopia de infravermelho próximo (NIR).....	19
3.4.1	Descrição do equipamento	20
3.4.2	Estudo das condições de aquisição dos espectros	20
3.4.3	Aquisição dos espectros.....	20
3.4.4	Construção de modelos utilizando o <i>software</i> OPUS 6.5	21
3.4.5	Remoção de amostras anómalas	24

4	Resultados e Discussão.....	25
4.1	Papel B1	25
4.1.1	Criação de um modelo PLS para determinar o RC_1	25
4.1.2	Criação de um modelo PLS para determinar o RC_2	26
4.1.3	Criação de um modelo PLS para determinar o RC	28
4.1.4	Criação de um modelo PLS para determinar o VC_2	30
4.1.5	Criação de um modelo PLS para determinar o VC	31
4.1.6	Validação externa com diferentes papéis impregnados	33
4.2	Papel L	33
4.2.1	Criação de modelos PLS para determinar o RC_1	34
4.2.2	Criação de modelos PLS para determinar o VC_1	36
4.3	Causas de erros na previsão do VC e RC	39
5	Conclusões	43
6	Avaliação do trabalho realizado.....	45
6.1	Objetivos Realizados	45
6.2	Outros Trabalhos Realizados.....	45
6.3	Apreciação Final	46
7	Referências	47
	Anexo A - Algoritmo NIPALS	51
	Apêndice A - Escolha dos parâmetros.....	53
A.1	Resolução.....	53
A.2	Número de varrimentos.....	54
A.3	Frequência de varrimento	55
A.4	Otimização dos parâmetros	56
	Apêndice B - Espectros e Loadings	59
B.1	Papel B1	59
B.2	Papel L	61
	Apêndice C - Resultados das validações	63
C.1	Papel B1	63

C.2	Papel L	68
------------	----------------------	-----------

Lista de Figuras

Figura 2.1 - Reações de formação das resinas melamina-formaldeído.	4
Figura 2.2 - Estrutura de um HPL.....	5
Figura 2.3 - Exemplo de um processo industrial de impregnação de papel decorativo.	6
Figura 2.4 - Bandas de absorção mais pronunciadas na região do NIR.	7
Figura 2.5 - Medições de transmitância, refletância e transrefletância.....	8
Figura 2.6 - Interferômetro de Michelson.	9
Figura 3.1 - Espectros adquiridos em papel impregnado industrialmente com o espaçamento de 0,2 mm.	21
Figura 3.2 - Exemplo do resultado da ferramenta <i>Optimize</i>	22
Figura 3.3 - Exemplo de uma representação gráfica de $RMSE_{CV}$ em função do número de fatores.	23
Figura 4.1 - Validação cruzada para a metodologia de quantificação do RC_1	26
Figura 4.2 - Validação cruzada para a metodologia de quantificação do RC_2	27
Figura 4.3 - Validação cruzada para a metodologia de quantificação do RC	29
Figura 4.4 - Validação cruzada para a metodologia de quantificação do VC_2	30
Figura 4.5 - Validação cruzada para a metodologia de quantificação do VC	32
Figura 4.6 - Validação cruzada para a metodologia de quantificação do RC_1 , utilizando o procedimento das médias.	35
Figura 4.7 - Validação cruzada para a metodologia de quantificação do VC_1 , utilizando o procedimento do pórtico.	36
Figura 4.8 - Validação cruzada para a metodologia de quantificação do VC_1 , utilizando o procedimento das médias.	37
Figura 4.9 - Validação cruzada para a metodologia de quantificação do VC_1 , utilizando o procedimento do pórtico.	38
Figura 4.10 - Defeitos associados ao papel impregnado laboratorialmente.....	40
Figura 4.11 - Colinearidade e não colinearidade entre amostras.....	42
Figura 4.12 - Avaliação da colinearidade entre o VC e o RC.	42

Figura A.1 - Demonstração do efeito da resolução nos espectros obtidos.	53
Figura A.2 - Variância média ponto a ponto para os espectros obtidos com as resoluções de 8 e 16 cm ⁻¹	54
Figura A.3 - Variância média ponto a ponto para os espectros obtidos com o número de varrimentos de 32, 64, 128 e 256.	55
Figura A.4 - Variância média ponto a ponto para os espectros obtidos com uma frequência de varrimento de 5, 10 e 20 kHz.	56
Figura A.5 - Variância média ponto a ponto para os espectros obtidos com as combinações de parâmetros ótimos.....	57
Figura B.1 - Espectros utilizados no modelo de quantificação do RC_1	59
Figura B.2 - Primeiros três <i>loadings</i> do PCA do conjunto de espectros utilizados no modelo de quantificação do RC_1 , pré-processados pelo SNV.	59
Figura B.3 - Espectros utilizados no modelo de quantificação do RC_2	60
Figura B.4 - Primeiros três <i>loadings</i> do PCA do conjunto de espectros utilizados no modelo de quantificação do RC_2 , pré-processados pelo SNV.	60
Figura B.5 - Espectros utilizados no modelo de quantificação do RC_1 , obtidos recorrendo ao procedimento das médias.	61
Figura B.6 - Primeiros três <i>loadings</i> do PCA do conjunto de espectros adquirido pelos procedimento das médias e utilizados no modelo de quantificação do RC_1 , pré-processados pelo SNV.....	61
Figura B.7 - Espectros utilizados no modelo de quantificação do RC_1 , obtidos recorrendo ao procedimento do pórtico.	62
Figura B.8 - Primeiros três <i>loadings</i> do PCA do conjunto de espectros adquirido pelos procedimento do pórtico e utilizados no modelo de quantificação do RC_1 , pré-processados pelo SNV.....	62

Lista de Tabelas

Tabela 3.1 - Quantidade e características das amostras analisadas.	18
Tabela 3.2 - Valores ótimos de RC e VC para os papeis impregnados utilizados na construção de modelos.	19
Tabela 3.3 - Procedimento para aproximar os valores dos ótimos.	19
Tabela 3.4 - Número de amostras anômalas para cada modelo.	24
Tabela 4.1 - Comparação entre o desempenho dos modelos para o RC.	29
Tabela 4.2 - Comparação entre o desempenho dos modelos para o VC.	32
Tabela 4.3 - Resultados da validação externa com diferentes papeis impregnados.	33
Tabela 4.4 - Resultados da validação externa para os modelos baseados no procedimento das médias e do pórtico.	39
Tabela 4.5 - Comparação entre modelos contruídos apenas com amostras laboratoriais.....	40
Tabela A.1 - Combinações a testar para otimizar os parâmetros de leitura de espectros.....	56
Tabela C.1 - Resultados da validação para o modelo para determinar o RC_1	63
Tabela C.2 - Resultados da validação para o modelo para determinar o RC_2	64
Tabela C.3 - Resultados da validação para o modelo para determinar o RC	65
Tabela C.4 - Resultados da validação para o modelo para determinar o VC_2	66
Tabela C.5 - Resultados da validação para o modelo para determinar o VC	67
Tabela C.6 - Resultados da validação para o modelo para determinar o RC_1 , utilizando o procedimento das médias.	68
Tabela C.7 - Resultados da validação para o modelo para determinar o RC_1 , utilizando o procedimento do pórtico.	69
Tabela C.8 - Resultados da validação para o modelo para determinar o VC_1 , utilizando o procedimento das médias.	70
Tabela C.9 - Resultados da validação para o modelo para determinar o VC_1 , utilizando o procedimento do pórtico.	71

Notação e Glossário

A	Absorvância
a_0	Valor médio do espectro a ser corrigido
a_1	Desvio padrão do espectro
B	Matriz dos coeficientes da regressão
b	Vetor dos coeficientes de regressão
b_0	Parâmetro escalar
$b_{ref,1}$	Parâmetro escalar de referência
C	Matriz dos pesos da matriz dos valores de referência
c	Vetor dos pesos da matriz dos valores de referência
E	Matriz residual da matriz dos dados
e	Parte não representada
$E_{eletrônica}$	Energia associada a elétrons em diversas orbitais exteriores da molécula
$E_{rotacional}$	Energia associada à rotação das moléculas relativamente ao seu centro de gravidade
$E_{vibracional}$	Energia da molécula total devido a vibrações interatômicas
F	Matriz residual da matriz dos valores de referência
G	Termo das diferenças entre a parte modulada e observada
I_0	Intensidade da luz incidente
I_R	Intensidade da luz refletida
m	Número de variáveis
m_1	Massa do papel impregnado após o primeiro banho
m_2	Massa do papel impregnado após o segundo banho
m_i	Massa do papel seco
m_{R1}	Massa do papel impregnado antes da secagem
m_{R2}	Massa do papel impregnado após secagem
n	Número de amostras
n_c	Número de espectros do conjunto de calibração
n_p	Número de amostras do conjunto de previsão
np	Número de componentes principais
P	Matriz dos <i>loadings</i> da matriz dos dados
p_k	Vetor dos <i>loadings</i>
Q	Matriz dos <i>loadings</i> da matriz dos valores de referência
R	Espectro em refletância
r	Número total de fatores ou <i>rank</i>

R^2	Coeficiente de determinação
RC	Teor de resina geral
RC_1	Teor de resina após o primeiro banho
RC_2	Teor de resina após o segundo banho
$RMSE_{CV}$	Raiz quadrada do erro médio quadrático da validação cruzada
$RMSE_P$	Raiz quadrada do erro médio quadrático da previsão
T	Matriz dos scores da matriz dos dados
t_k	Vetor dos <i>scores</i>
U	Matriz dos <i>scores</i> da matriz dos valores de referência
u	Vetor dos <i>scores</i> da matriz dos valores de referência
VC	Teor de voláteis geral
VC_1	Teor de voláteis após o primeiro banho
VC_2	Teor de voláteis após o segundo banho
W	Matriz dos pesos da matriz dos dados
w	Vetor dos pesos da matriz dos dados
X	Matriz de dados
x_1	Abcissa do primeiro ponto escolhido
x_2	Abcissa do segundo ponto escolhido
x_{corr}	Espectro corrigido
x_{org}	Espectro da amostra original
x_{ref}	Espectro de referência
Y	Matriz dos valores de referência
y	Vetor das variáveis dependentes
$\bar{y}$	Média dos valores de referência
y_1	Ordenada do primeiro ponto escolhido
y_2	Ordenada do segundo ponto escolhido
y_i	Valor medido pelo método de referência
$\hat{y}_i$	Valor estimado pelo modelo

Letras gregas

ε	Permitividade relativa
ε'	Parte real da permitividade relativa
ε''	Parte imaginária da permitividade relativa
λ_k	Valor próprio

Índices

C	Calibração
corr	Corrigido
CV	Validação cruzada
<i>i</i>	Contador
<i>k</i>	Contador de fatores
org	Original
P	Previsão
ref	Referência

Lista de Siglas

ASTM	<i>American Society of Testing and Materials</i>
COE	Eliminação de um <i>offset</i> constante
FiPLS	<i>Forward</i> mínimos quadrados parciais intervalar
FT-NIR	Infravermelho Próximo com transformada de Fourier
HPL	Termolaminado decorativo de Alta Pressão
MF	Melamina-formaldeído
MFC	Placas revestidas de melamina
MSC	Correção Multiplicativa de sinais
MUF	Melamina-ureia-formaldeído
NIPALS	Algoritmo não-linear iterativo dos mínimos quadrados parciais
NIR	Infravermelho Próximo
PC	Componentes principais
PCA	Análise dos Componentes Principais
PF	Fenol-Formaldeído
PLS	Regressão por mínimos quadrados parciais
RC	Teor de Resina
RC _{ótimo}	Teor de Resina pretendido
SG	Savitzky e Golay
SNV	Padronização Normal da Variável
UF	Ureia-formaldeído
VC	Teor de Voláteis
VC _{ótimo}	Teor de Voláteis pretendido

1 Introdução

1.1 Enquadramento e Apresentação do Projeto

Atualmente, na fábrica da EuroResinas, no processo de produção de papel impregnado, são realizadas duas análises de controlo de qualidade: determinação do teor de resina (RC) e determinação do teor de voláteis (VC). A realização destas análises é efetuada de modo *off-line* e requer entre cinco e dez minutos para obtenção de resultados. A implementação da espectroscopia de infravermelho próximo (NIR) na linha de produção permite acelerar e automatizar esta determinação. Deste modo, permite aumentar a produtividade e reduzir o desperdício gerado no tempo de espera das análises. Contudo, devido à complexidade dos espectros obtidos, esta metodologia requer o desenvolvimento de modelos de análise multivariável e o conhecimento de métodos quimiométricos.

Este projeto tem como objetivo o desenvolvimento de métodos de análise baseados em espectroscopia de NIR para a determinação do teor de resina e de voláteis de papel impregnado branco. Pretende-se também otimizar os métodos previamente desenvolvidos, diminuindo o tempo de aquisição de espectros.

1.2 Apresentação da Empresa

A Sonae Arauco foi criada através da parceria de duas empresas líderes no setor de painéis derivados de madeira: a Sonae Indústria e a Arauco. Esta parceria foi criada com o objetivo de criar soluções sustentáveis inspiradas na madeira, e engloba todas as unidades de produção de painéis derivados de madeira, químicos e papel impregnado na Europa e África do Sul. Os produtos da Sonae Arauco chegam a mais de 75 países, pelo que é uma das principais empresas mundiais de painéis derivados de madeira [1].

A Sonae Indústria é uma empresa multinacional de painéis derivados de madeira fundada em 1959. A empresa desenvolve produtos para as indústrias de mobiliário, construção e decoração, tendo sempre em conta a sustentabilidade. Por sua vez, a Arauco foi criada em 1970 e, atualmente, é um dos maiores produtores de recursos florestais em todo o mundo [1].

A EuroResinas - Indústrias Químicas S.A. é uma unidade industrial da Sonae Arauco que se dedica ao fabrico e comercialização de formaldeído, resinas sintéticas e papel impregnado [2].

Foi constituída em novembro de 1994 e instalou-se na área industrial de Sines em 2001. São produzidas uma gama variada de resinas sintéticas sendo estas do tipo: fenol-formaldeído (PF), melamina-formaldeído (MF), ureia-formaldeído (UF) e melamina-ureia-formaldeído (MUF). A sua principal utilização é como agente aglutinador na produção de diversos produtos derivados de madeira e papel [2].

O papel impregnado produzido é de dois tipos diferentes: papel decorativo (impregnado com resina Melamina-Formaldeído) e papel *kraft* (impregnado com resina Fenol-Formaldeído), utilizados para a produção de laminados de alta pressão (HPL - *High Pressure Laminates*) [2].

1.3 Contributos do(a) autor(a) para o Trabalho

Após este trabalho experimental foi possível reduzir o tempo total de aquisição dos espectros, sem comprometer a qualidade dos resultados. O tempo otimizado foi de 22 segundos, sendo que, o previamente estipulado era de 1 minuto e 40 segundos.

Após esta otimização, foram desenvolvidos modelos de quantificação do RC e VC para os papeis impregnados B1 e L. Para o papel B1, já foram criados modelos para este efeito em estudos prévios. Num primeiro estudo, verificou-se a heterogeneidade do papel impregnado. Assim, os modelos com a maior capacidade de previsão recorreram a um procedimento das médias em três pontos distintos da amostra para a aquisição de espectros [3]. Num estudo mais recente, utilizou-se em paralelo com este procedimento um pórtico de mesa para a aquisição de espectros ao longo do papel impregnado, melhorando a capacidade de previsão dos modelos. No entanto, a gama de propriedades utilizada neste estudo é bastante restrita, e considera-se que a capacidade de previsão dos modelos já desenvolvidos ainda pode ser melhorada [4].

O papel denominado de L não foi analisado em outros estudos. Para este papel, foi possível testar de modo mais aprofundado duas formas distintas de aquisição de espectros. Apesar da capacidade de previsão dos modelos previstos ser inferior à desejada, espera-se que os resultados obtidos neste estudo permitam apoiar trabalhos futuros do mesmo tipo, de modo a proceder à implementação futura de métodos deste tipo em ambiente fabril.

1.4 Organização da Tese

Esta tese organiza-se encontra-se dividida em seis partes: Introdução, Contexto e Estado da Arte, Materiais e Métodos, Resultados e Discussão, Conclusões e por fim uma Avaliação do trabalho realizado.

Na Introdução é feita uma apresentação da empresa e do objetivo do projeto. Por sua vez, no Contexto e Estado da Arte são apresentadas informações relevantes sobre as resinas amino, papel impregnado, a espectroscopia de infravermelho próximo (NIR) e os métodos quimiométricos utilizados para interpretação dos espectros obtidos. No terceiro capítulo são apresentados os materiais e métodos utilizados neste projeto. No capítulo seguinte apresentam-se os resultados obtidos e a sua respetiva discussão. No capítulo cinco apresentam-se as principais conclusões deste projeto. E por fim, no capítulo 6 é feita uma reflexão sobre o trabalho realizado e o grau de concretização dos objetivos inicialmente previstos.

2 Contexto e Estado da Arte

2.1 Resinas amino

As resinas amino são polímeros termoendurecíveis resultantes da reação de um aldeído com um composto com grupos amino ($-NH_2$), como a ureia e a melamina. O aldeído mais comumente utilizado é o formaldeído [5] [6].

As resinas ureia-formaldeído (UF) constituem 80 % da produção de resinas amino sendo o restante constituído por resinas melamina-formaldeído (MF). Outros aldeídos e compostos amino são produzidos com uma frequência muito mais reduzida [7].

Os principais benefícios associados às resinas amino são: a sua solubilidade inicial em água, transparência e alta resistência a solventes após a cura, dureza e resistência à abrasão e ao calor. A sua solubilidade inicial em água permite uma fácil aplicação em vários materiais. Por sua vez, a transparência destas resinas após a cura facilita a sua conjugação com pigmentos [5] [7].

Contudo, estas resinas apresentam também as suas limitações como a libertação de formaldeído durante a cura e, em alguns casos, após esta. Por sua vez, as resinas ureia-formaldeído, quando expostas a condições cíclicas de humidificação e secura, podem causar fissuras na superfície do aglomerado em que são utilizadas [7].

As resinas amino são utilizadas numa grande variedade de produtos, nomeadamente como adesivos utilizados na produção de painéis derivados de madeira. Para esta aplicação, as resinas ureia-formaldeído são as mais utilizadas [6] [8].

2.1.1 Resinas melamina-formaldeído (MF)

As resinas melamina-formaldeído são formadas pela reação entre a melamina ($C_2H_6N_6$) e formaldeído (CH_2O). Estas resinas são utilizadas nomeadamente em adesivos para madeira e em impregnação de papel decorativo [9].

Contrariamente às resinas UF, as resinas melamina-formaldeído apresentam alta resistência à água. No entanto, as resinas de MF têm um preço aproximadamente 2,5 vezes mais elevado que o preço das UF. Assim, uma quantidade variável de ureia pode ser adicionada à formulação de forma a encontrar um compromisso entre custo e performance, originando as resinas melamina-ureia-formaldeído (MUF) [10] [11] [12].

O processo de formação das resinas MF segue duas etapas: metilolação e condensação. Estas encontram-se esquematizadas na figura 2.1 [9] [13].

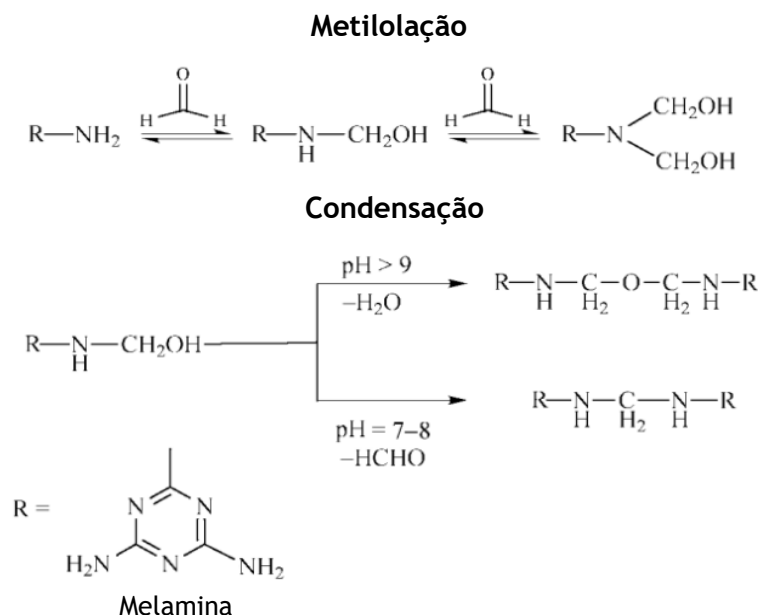


Figura 2.1 - Reações de formação das resinas melamina-formaldeído (adaptado de [14]).

Na etapa de metilolação, ocorre a formação de compostos metilolados pela reação do formaldeído com a melamina. Na segunda etapa, a condensação, podem ser formadas dois tipos de pontes por condensação: metileno-éter ou metileno. Nesta reação podem ser libertadas moléculas de água ou formaldeído dependendo do pH do meio. Os dois passos podem ser catalisados de forma ácida ou básica e, são reações de equilíbrio [14].

Os parâmetros mais importantes para a síntese das resinas MF são: o tipo de monómeros, a sua razão molar, o procedimento utilizado e a pureza das matérias primas. O procedimento do processo de síntese pode variar em diversos parâmetros, tais como: pH, temperatura, tipo e quantidade de catalisadores ácidos e básicos, a sequência de adição das matérias primas e a duração das diversas etapas. De acordo com o tipo de aplicação, os parâmetros referidos podem ser ajustados [9] [14].

As resinas MF apresentam habitualmente um teor de sólidos de 54 a 60 % e uma viscosidade baixa de 15 a 27 s (medida com um copo do tipo DIN 4) [11]. Quando utilizadas para a impregnação de papel, a sua viscosidade mais baixa e alto teor de sólidos são fundamentais. Desta forma, estas resinas conseguem penetrar o papel adequadamente, secar sem perder a sua capacidade de adesão e, mais tarde, formar uma ligação forte com o substrato [12].

2.2 Papel impregnado

O uso de papel impregnado tem vindo a aumentar recentemente nas indústrias de mobiliário e pavimento laminado na Europa [11]. Quando impregnado com resina melamínica, é utilizado frequentemente para a produção de placas revestidas de melamina (MFC - *Melamine Faced Boards*) e termolaminados de alta pressão (HPL) [15] [16].

Segundo a norma EN 438-1:2016 um laminado de alta pressão consiste num conjunto de camadas de material fibroso celulósico (normalmente papel) impregnado com resinas termoendurecíveis e unidas por um processo de alta pressão. Sendo que, as condições de operação típicas para estes são de uma temperatura ≥ 120 °C e de pressão ≥ 5 MPa [16] [17].

Os HPL, representados na figura 2.2, são constituídos por três tipos de papel impregnado: *kraft*, decorativo e *overlay* [18].

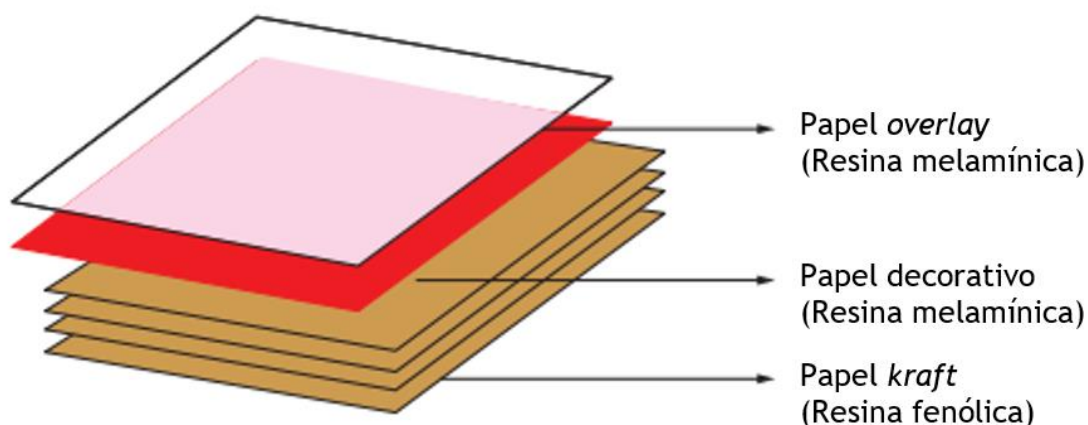


Figura 2.2 - Estrutura de um HPL (adaptado de [18]).

O papel *kraft*, impregnado com resina fenólica, é utilizado nas camadas inferiores do HPL. Este tipo de papel é responsável por conferir diversas propriedades ao HPL, como: resistência mecânica (dependente da orientação das fibras), térmica e à humidade [18].

O papel decorativo, impregnado com resina melamínica, é constituído por α -celulose. Este tipo de papel deve ter boas propriedades mecânicas, rápida absorção de resina combinada com resistência mecânica suficiente quando húmido, alta opacidade, e distribuição homogênea das fibras. O papel decorativo pode apresentar diferentes padrões e cores. A utilização de gramagens mais elevadas pode ser necessária em papeis brancos, de modo a esconder a cor castanha do papel kraft. Outra solução poderá ser recorrer a um papel barreira [18] [19].

O papel *overlay*, tal como o papel decorativo, é composto maioritariamente por α -celulose e é geralmente impregnado com resina melamínica. Após prensado no topo do papel decorativo, torna-se transparente permitindo ver a superfície do papel decorativo. A sua utilização contribui para um aumento da abrasão e da resistência ao tingimento, o que permite o uso em pavimentos [18] [19].

2.2.1 Produção de papel impregnado

No processo de impregnação de papel decorativo, ilustrado na figura 2.3, o papel é impregnado com resina MF. Contudo, pode ser utilizada resina UF num primeiro banho. Neste caso, o papel será seguidamente submetido a um segundo banho, *coating*, com resina MF [9] [20].

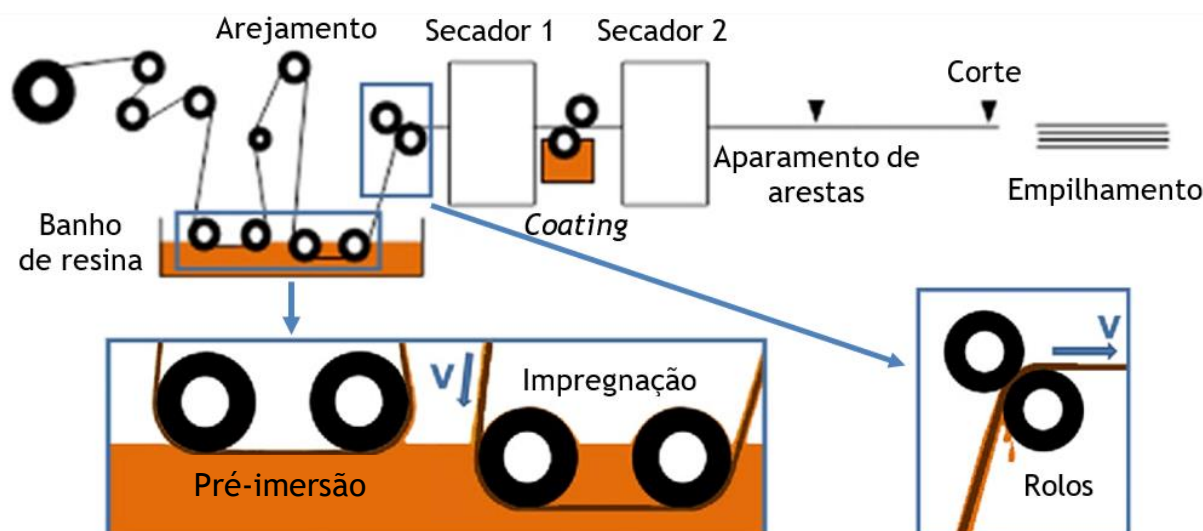


Figura 2.3 - Exemplo de um processo industrial de impregnação de papel decorativo (adaptado de [21]).

Quando utilizadas resinas UF e MF, o papel é movido até um primeiro banho onde é impregnado apenas de um lado (pré-imersão) com resina UF. De seguida, o papel é movido até a uma altura adequada que permita a penetração da resina e a remoção de ar. Após ser imerso na resina, o papel passa entre dois rolos que ajustam o seu teor de resina. O papel passa então para um primeiro secador até apresentar um teor de voláteis de aproximadamente 10 % (m/m). Após secagem, o papel passa para um segundo banho de resina MF (*coating*). Este é seguido por uma segunda secagem até atingir um teor de voláteis próximo de 5 % (m/m). Os secadores operam a uma temperatura de 110 a 140 °C. Por fim, o papel é aparado, cortado e empilhado [20] [21] [22].

O teor de resina (RC) e o teor de voláteis (VC) do papel impregnado são dois parâmetros importantes a controlar ao longo do processo de produção. O teor de resina pode influenciar o escoamento da resina e, conseqüentemente, o aspeto do HPL. No caso do teor de voláteis, se este for demasiado alto, as folhas enroladas ou empilhadas poderão aderir umas às outras antes de serem enviadas para a prensa para o processo de cura. Por outro lado, um teor baixo de VC pode indicar que a secagem foi excessiva e que, a resina reagiu demasiado na secagem. Assim, o papel não irá apresentar adesão suficiente na cura [9].

Estas análises podem ser realizadas *off-line* por pesagem diferencial do papel após impregnação. Contudo, também podem ser agilizadas e automatizadas pela implementação da espectroscopia de infravermelho próximo (NIR) na linha de produção [9].

2.3 Espectroscopia de infravermelho próximo

A espectroscopia de infravermelho próximo tem vindo a substituir técnicas analíticas demoradas na monitorização de processos e no controlo de qualidade industrial [23].

De acordo com a ASTM (*American Society of Testing and Materials*), o infravermelho próximo corresponde à região do espectro eletromagnético imediatamente a seguir à região do visível, com um comprimento de onda entre 780 e 2500 nm [23] [24].

A espectroscopia de NIR é baseada na absorção de radiação eletromagnética. Quando esta incide numa partícula elementar (átomo, molécula ou ião) pode ocorrer a absorção de energia. Contudo, a absorção só ocorre se a energia incidente for igual à diferença de energia entre o estado fundamental e o estado excitado. As moléculas, quando excitadas por radiação, podem ter três tipos de transições: eletrónica, vibracional e rotacional. A energia total E associada a uma molécula é dada pela equação 2.1.

$$E = E_{\text{eletrónica}} + E_{\text{vibracional}} + E_{\text{rotacional}} \quad (2.1)$$

Onde, a $E_{\text{eletrónica}}$ corresponde à energia associada a eletrões em diversas orbitais exteriores da molécula, $E_{\text{vibracional}}$ é a energia da molécula total devido a vibrações interatómicas e a $E_{\text{rotacional}}$ corresponde à energia associada à rotação das moléculas relativamente ao seu centro de gravidade [24].

As bandas de absorção mais pronunciadas na região do NIR são as correspondentes às sobreposições e combinações das vibrações fundamentais dos grupos funcionais -CH, -NH, -OH e -SH. Estas encontram-se representadas na figura 2.4 [24] [25].

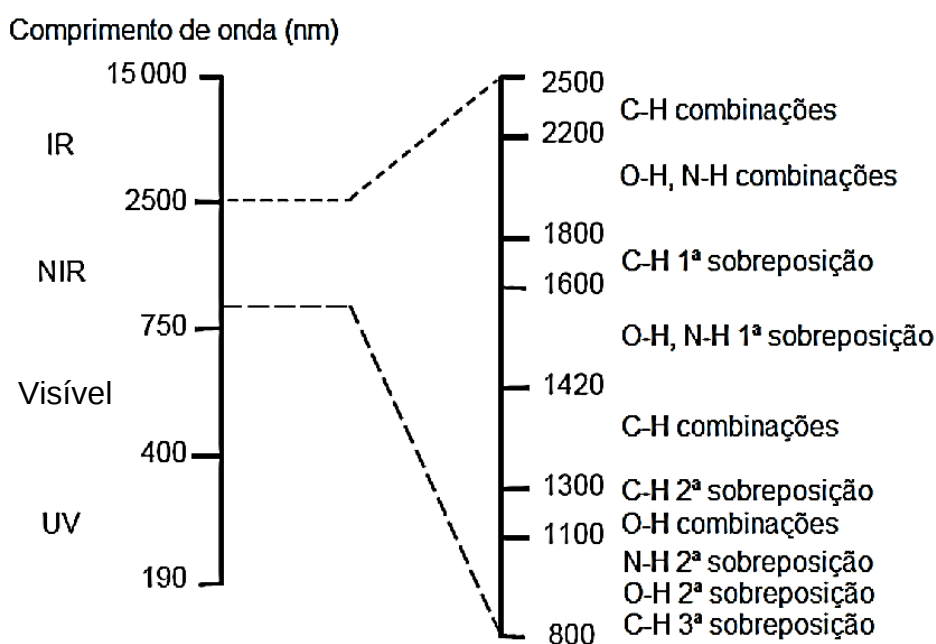


Figura 2.4 - Bandas de absorção mais pronunciadas na região do NIR [26].

As bandas de absorção na região NIR são largas, sobrepostas e com uma amplitude menor do que as características da região do infravermelho médio. Contudo, a menor absorção de radiação nesta região é também uma vantagem, pelo que permite uma maior penetração da radiação na amostra. Assim, é possível analisar amostras que dispersam ou absorvem bastante radiação sem pré-tratamentos [9] [24].

A espectroscopia de NIR apresenta outras vantagens, tais como: poder ser realizada *on-line* ou *off-line*, as amostras requerem pouca preparação, podem também ser sólidas ou líquidas, a análise é rápida, simples e não é destrutiva [9] [24].

Contudo, esta tecnologia também apresenta as suas desvantagens. Os espectros obtidos são complexos, pelo que, precisam de ser processados. Assim, o uso da espectroscopia de NIR tem de ser efetuado juntamente com métodos quimiométricos. Estes permitem a correlação entre os dados obtidos e os parâmetros a analisar [9] [24].

2.3.1 Funcionamento do equipamento

Os principais componentes de um espectrofotómetro de NIR são geralmente: uma fonte luminosa, um monocromador, um suporte para a amostra, e um detetor [24] [27].

A fonte luminosa é habitualmente uma lâmpada de tungsténio-halógeno. No caso dos detetores, estes são do tipo multicanal pelo que contêm vários elementos de deteção dispostos em filas (raios díodo) ou planos (dispositivo de carga acoplada - CCD). Deste modo, são capazes de ler vários comprimentos de onda simultaneamente, aumentando a velocidade de aquisição dos espectros. Os tipos de detetores mais usados são: silicone, sulfureto de chumbo (PbS) e índio-gálio-arsénio (InGaAs) [9] [24] [25].

Existem diferentes técnicas de medição que podem ser classificadas como: transmitância, refletância e transrefletância. Estas encontram-se representadas na figura 2.5.

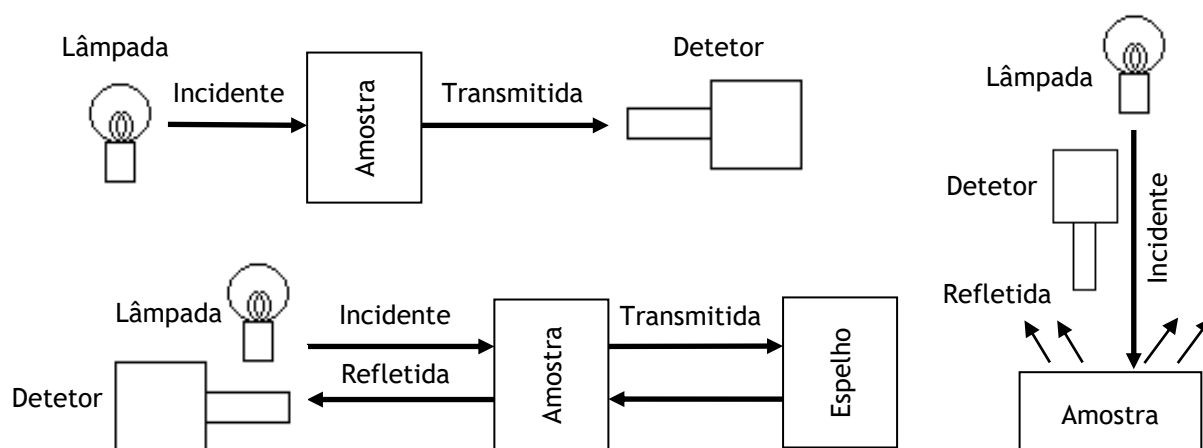


Figura 2.5 - Medições de transmitância, refletância e transrefletância (adaptado de [24]).

Nas medições de transmitância, parte da radiação emitida atravessa a amostra, esta fração transmitida é enviada para o detetor. Em medições de refletância, a radiação emitida é parcialmente refletida na interface sonda/amostra e enviada para o detetor. No caso das medições de transrefletância, a radiação refletida na superfície atinge o detetor juntamente com a radiação que atravessa a amostra. Esta radiação transmitida, sendo refletida por um espelho, volta a atravessar a amostra antes de atingir o detetor [9] [24].

A escolha da técnica de medição deve ser tomada tendo em conta as propriedades da amostra. Assim, medições de refletância são mais utilizadas para sólidos, as de transmitância para líquidos e as de transrefletância para emulsões ou líquido turvos [9] [24].

No caso de a medição ser efetuada em refletância pura, como se irá verificar neste estudo, a conversão para absorvância será dada pela equação 2.2.

$$A = \log \left(\frac{1}{R} \right) \quad (2.2)$$

Em que A corresponde à absorvância e R corresponde ao espectro em refletância [25] [28].

A refletância é dada pela equação 2.3.

$$R = \frac{I_R}{I_0} \quad (2.3)$$

Em que R corresponde ao espectro em refletância, I_R à intensidade da luz refletida e I_0 e à intensidade da luz incidente [29].

- **Espectrofotômetros de NIR com transformada de Fourier (FT-NIR)**

Os espectrofotômetros com transformada de Fourier não apresentam um elemento dispersivo. Em vez de um monocromador, um interferómetro é usado para produzir padrões de interferência que contêm a informação espectral. Os interferómetros mais simples são do tipo Michelson, esquematizados na figura 2.6 [23] [28].

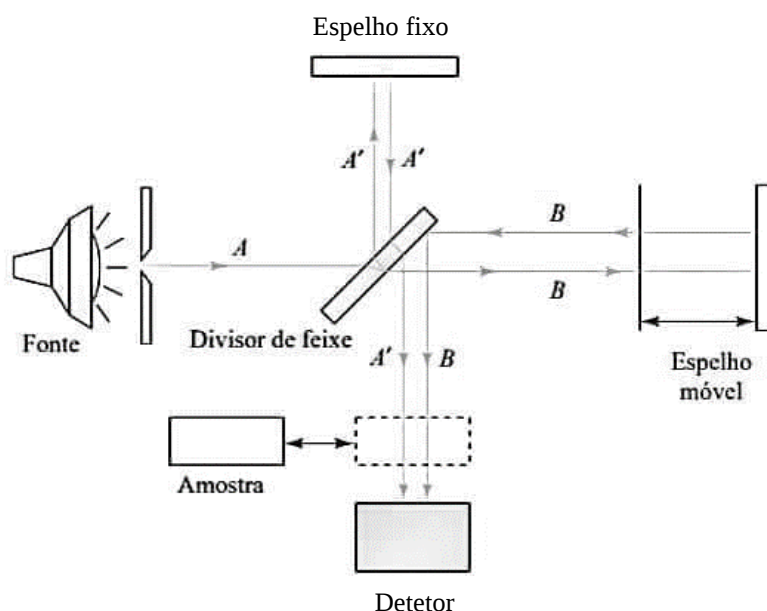


Figura 2.6 - Interferómetro de Michelson (adaptado de [28]).

Um interferómetro de Michelson consiste numa fonte de luz, um espelho fixo, um espelho móvel perpendicular a este, um divisor de feixe e um detetor [23] [28].

Neste equipamento, o divisor do feixe reflete parcialmente o feixe de luz vindo da fonte para o espelho fixo e o restante para o móvel. Este move-se a uma velocidade constante ao longo

de um eixo. Os feixes refletidos pelos espelhos são recombinados pelo divisor do feixe e reencaminhados para a amostra e, de seguida, para o detetor. A interferência entre estes dois feixes com um diferente caminho ótico ou atraso cria o interferograma, que contém a informação do espectro da amostra. Cada comprimento de onda tem uma frequência de modulação distinta (em kHz), que pode ser transformada por um computador na sua respetiva frequência eletromagnética. Este processo é realizado pelos computadores atuais recorrendo à transformada de Fourier [23] [28].

Os espectrofotómetros com transformada de Fourier são de feixe único. Assim, é necessário realizar primeiramente um espectro de fundo (*background*). De seguida, pode efetuar-se o espectro da amostra. Por fim, a razão entre o espectro da amostra e o de fundo é calculada, e a absorvância ou transmitância em função do comprimento de onda é registada [28].

As principais vantagens dos espectrofotómetros FT-NIR relativamente aos dispersivos são: maior velocidade e precisão, resolução constante, uma calibração mais exata do comprimento de onda e espectros sem descontinuidades [23] [28].

2.4 Métodos quimiométricos

2.4.1 Pré-processamentos

O pré-processamento dos dados de espectroscopia de NIR é uma etapa fundamental aquando a utilização de métodos quimiométricos. O seu objetivo consiste em remover a influência de fenómenos físicos nos espectros de modo a melhorar a regressão multivariável que lhe sucede [30].

Dois fatores afetam negativamente a análise de amostras sólidas por espectroscopia de NIR: a dispersão da luz e o tamanho da partícula. No caso da dispersão da luz, esta ocorre na superfície do material e depende da sua natureza física e do seu tamanho de partícula. A interação entre o meio e a luz ocorre no interior do material. Logo, o tamanho da partícula define o caminho espectral e a variação do tamanho das partículas origina num desvio na linha base do espectro [31].

A interferência destes fatores pode resultar em desvios na linha base e alterações de declives. Deste modo, são adicionadas informações físicas ao espectro que podem não ser relevantes para o estudo, pelo que, é útil recorrer ao pré-processamento dos espectros obtidos [31].

As técnicas de processamento mais utilizadas podem ser divididas em duas categorias: métodos de correção da dispersão da luz e de derivada de espectros [30].

- **Correção multiplicativa de sinais (MSC)**

O pré-processamento de correção multiplicativa de sinais (MSC) é um dos métodos mais utilizados seguido da padronização normal da variável (SNV) e dos derivativos [30]. O MSC permite remover variações resultantes de fatores físicos como o tamanho de partícula que, pode resultar num desvio da linha de base. Este método é capaz de minimizar estes desvios tendo como vantagem que os espectros obtidos são semelhantes aos espectros originais facilitando a sua interpretação [32].

O método de MSC compreende dois passos descritos pelas equações 2.4 e 2.5 [30].

1. Estimativa dos coeficientes de correção (contribuições aditivas ou multiplicativas):

$$x_{org} = b_0 + b_{ref,1} \times x_{ref} + e \quad (2.4)$$

2. Correção do espectro obtido:

$$x_{corr} = \frac{x_{org} - b_0}{b_{ref,1}} \quad (2.5)$$

Em que, x_{org} corresponde ao espectro da amostra original medida pelo espectrofotómetro NIR, x_{ref} é um espectro de referência utilizado para todo o pré-processamento, e é a parte não representada pelo modelo de x_{org} , x_{corr} é o espectro corrigido e b_0 e $b_{ref,1}$ são parâmetros escalares que diferem para cada amostra. Normalmente, a média dos espectros utilizados para calibração é utilizada como espectro de referência.

- **Padronização normal da variável (SNV)**

O método de padronização normal da variável (SNV) é um pré-processamento utilizado para corrigir diferenças nas intensidades globais dos sinais, bem como reduzir efeitos da dispersão da luz e do tamanho da partícula. Neste pré-processamento cada espectro é centrado e corrigido pela divisão pelo seu respetivo desvio padrão [33] [34].

A correção efetuada em cada espectro pelo método de SNV é dada pela equação 2.6 [30].

$$x_{corr} = \frac{x_{org} - a_0}{a_1} \quad (2.6)$$

Em que, x_{org} corresponde ao espectro da amostra original medida e x_{corr} ao espectro corrigido. Por sua vez, a_0 corresponde ao valor médio do espectro a ser corrigido e a_1 ao desvio padrão do espectro.

- **Normalização Mín-máx**

Neste pré-processamento os valores mínimos e máximos de cada espectro são definidos para valores fixos. No *software* OPUS, que irá ser utilizado para realizar os modelos, com este pré-processamento, é subtraído a cada espectro um *offset* linear de modo a definir o valor mínimo como zero. De seguida, o valor máximo é definido como 2 pela multiplicação por uma constante. A sua utilização é semelhante à do SNV [34] [35].

- **Primeira e segunda derivada**

A aplicação da primeira ou segunda derivada permite remover os desvios na linha de base. Contudo, o efeito da segunda derivada remove também qualquer tendência linear no espectro. Os métodos derivativos têm também uma ação de suavização de modo a não reduzir demasiado a razão sinal-ruído [30].

O método derivativo mais utilizado é o de Savitzky e Golay (SG). De modo a determinar a derivada, este método recorre a um ajuste polinomial numa janela simétrica dos dados à esquerda e à direita do ponto que se pretende suavizar. O número de pontos a utilizar, assim como a ordem do polinómio de ajuste têm de ser escolhido previamente. De seguida, a derivada do ponto central é determinada. Este procedimento é então aplicado aos restantes pontos do espectro [30].

- **Eliminação de um *offset* constante (COE)**

O método da eliminação de um *offset* contante é utilizado para corrigir desvios na linha de base dos espectros. Um ou mais pontos, sem desvios significativos, são escolhidos para calcular a correção do *offset*. No caso mais simples, um valor constante é subtraído a todo o espectro. Este valor pode ser a intensidade de um determinado ponto ou a intensidade média de uma série de pontos adjacentes. Quando o ponto selecionado é o mínimo do sinal, a correção efetuada corresponde a subtrair o mínimo a todo o espectro. Contudo, uma linha reta para subtrair a todo o espectro pode ser determinada utilizando dois pontos. A equação desta reta é determinada pela equação 2.7 [34].

$$y(i) = \left[\frac{(y_2 - y_1)}{(x_2 - x_1)} \right] \cdot x(i) + \left[y_1 - \left(\frac{(y_2 - y_1)}{(x_2 - x_1)} \right) \cdot x_1 \right] \quad (2.7)$$

Em que, (x_1, y_1) e (x_2, y_2) são as coordenadas dos dois pontos escolhidos.

- **Subtração de uma linha reta**

O método da subtração de uma linha reta, à semelhança do COE, é utilizado para corrigir desvios na linha de base dos espectros. Neste pré-processamento uma linha reta é ajustada à base do espectro. Esta reta é subtraída ao espectro de modo a retirar a sua inclinação [35].

2.4.2 Análise dos componentes principais (PCA)

A análise dos componentes principais (PCA) tem como principal objetivo representar um conjunto de dados num espaço dimensional menor, com a menor perda de informação possível. Assim, o PCA é útil quando a quantidade de dados é elevada, sendo o método mais utilizado deste tipo em várias áreas científicas [36] [37].

Esta técnica é utilizada na espectroscopia para: selecionar espectros para calibração ou validação de modelos, detetar espectros semelhantes e detetar valores anómalos [9] [35].

O PCA organiza conjuntos de dados em componentes ortogonais (independentes) cujas combinações lineares se aproximam aos dados originais com um grau de exatidão escolhido. As novas variáveis, os componentes principais (PC) ou fatores, correspondem aos maiores valores próprios da matriz das covariâncias [23] [25].

O PCA é aplicado a uma matriz de dados, X , com m linhas e n colunas, sendo que cada variável corresponde a uma coluna e cada amostra a uma linha. O produto dos primeiros vetores $t_1 p_1^T$ é então determinado de forma a contabilizar o máximo de variância, sendo t_1 e p_1 os primeiros *scores* e *loadings*, respetivamente. Os restantes *scores* e *loadings*, ortogonais aos primeiros vetores, são determinados até aos vetores número k . A equação 2.8 representa equação matemática do PCA [38] [39].

$$X = \sum_{k=1}^{np} t_k \cdot p_k^T + E = T \cdot P^T + E \quad (2.8)$$

Onde np corresponde ao número de componentes principais, T em sobrescrito à transposta da matriz, T à matriz dos *scores*, P à matriz dos *loadings* e E corresponde ao termo do erro.

Os vetores *scores*, t_k , contêm a informação acerca das relações entre as amostras, os vetores *loadings*, p_k , descrevem como as variáveis do processo se relacionam entre si. Por sua vez, E , o termo do erro, contêm toda a informação não descrita pela quantidade np de componentes principais [39].

Em várias situações verifica-se que um número reduzido de componentes principais é suficiente para descrever a maioria da informação da matriz X sendo a restante informação correspondente a ruído e erros nas medições. Na espectroscopia, os dados são comumente espectros, e o número de PC dever ser menor ou igual ao número de variáveis ou ao número de espectros, dependendo do que for menor [23] [39].

No PCA, t_k e p_k podem ser determinados utilizando o método de decomposição em valores singulares (SVD). Os vetores p_k , *loadings*, são determinados pelo cálculo dos vetores próprios da matriz de covariância de X , tal como descrito pela equação 2.9 [39].

$$\text{cov}(X) \cdot p_k = \lambda_k \cdot p_k \quad (2.9)$$

Em que, λ_k corresponde ao valor próprio associado ao vetor próprio p_k .

Relativamente aos *scores*, t_k , estes são determinados pela equação 2.10.

$$t_k = X \cdot p_k \quad (2.10)$$

Deste modo, t_k é por vezes referido como uma projeção de X em p_k . Os escalares, λ_k , são utilizados como medida da variância descrita pelo par $t_k p_k$. Com o cálculo sucessivo de pares, cada um descreve o máximo possível de variância residual dos dados. Assim, o primeiro PC contabiliza o máximo de variância em todas as combinações lineares e os seguintes cada vez menos [23] [25] [39].

2.4.3 Regressão por mínimos quadrados parciais (PLS)

Apesar de o PCA ser um bom método para avaliar as diferenças entre amostras, este método não permite quantificar um dado analito numa amostra. Para tal, é necessário construir um modelo de calibração que relacione a variável de interesse, dependente, com a independente por um coeficiente. No caso da espectroscopia, as variáveis independentes são frequentemente os valores de refletância ou absorvância a diferentes comprimentos de onda e, a variável dependente poderá ser a concentração do analito. Um exemplo da relação entre estas variáveis é representada pelas equações 2.11 e 2.12 [35].

$$y = X \times b \quad (2.11)$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_n \end{bmatrix} = \begin{bmatrix} \text{Espectro 1} \\ \text{Espectro 2} \\ \dots \\ \text{Espectro n} \end{bmatrix} \times b \quad (2.12)$$

Onde y é o vetor das variáveis dependentes, b corresponde ao vetor dos coeficientes de regressão e X é a matriz dos dados das absorvâncias.

Para se realizar esta calibração, um dos métodos utilizados é a regressão por mínimos quadrados parciais (PLS). A relação obtida é linear e baseada no uso de componentes principais [29] [40].

À semelhança do PCA, são determinados os vetores próprios (PC) das matrizes X e Y . A decomposição em fatores da matriz X e Y são representadas nas equações 2.13 e 2.14, respetivamente [29].

$$X = \sum_{k=1}^r t_k \cdot p_k^T + E = T \cdot P^T + E \quad (2.13)$$

$$Y = \sum_{k=1}^r u_k \cdot q_k^T + F = U \cdot Q^T + F \quad (2.14)$$

Em que, T em sobrescrito corresponde à transposta da matriz. Assim, as matrizes de dados X e Y são decompostas em produtos das matrizes dos *scores*, T e U , e dos *loadings*, P e Q , e nas suas matrizes residuais, E e F , respetivamente. O número total de fatores, r , é também denominado de *rank* [29] [35].

Os *scores* da matriz X podem ser utilizados para prever Y de acordo com a equação 2.15.

$$Y = \sum_{k=1}^r c_k \cdot t_k^T + G = T \cdot C^T + G \quad (2.15)$$

Em que, C corresponde à matriz dos pesos de Y e G contabiliza as diferenças entre a parte modulada e observada [41].

O primeiro PC representa o máximo de variância em todas as combinações lineares e os seguintes componentes correspondem sucessivamente a menores valores de variância [29].

A decomposição pode ser feita utilizando o algoritmo não-linear iterativo dos mínimos quadrados parciais (NIPALS) [41]. Este algoritmo encontra-se descrito no anexo A.

Para recorrer a este método e obter correlações adequadas, é necessário um número adequado de amostras. Dependendo do tipo de amostra e do analito a analisar, o número de amostras necessárias pode variar entre 20 e 200 [29]. Por exemplo, uma calibração utilizada para uma amostra com apenas um componente poderá ser constituída por 20 amostras. Contudo, em situações em que diversos componentes estão presentes será necessário um maior número de amostras [35].

Na realização de modelos utilizando PLS, é importante ter em conta a região espectral a utilizar. Uma vez que, algumas regiões podem conter sinais não-lineares, ruído ou informação não relevante para a análise em causa [29]. A escolha das regiões a utilizar pode ser realizada por diversos métodos, sendo que, o utilizado neste estudo baseou-se no *Forward* mínimos quadrados parciais intervalar (FiPLS).

- **Forward mínimos quadrados parciais intervalar (FiPLS)**

O FiPLS divide os dados espectrais num número de intervalos escolhido. De seguida, modelos PLS são determinados para cada intervalo. O intervalo que gerar o menor erro na validação cruzada é selecionado. Por fim, a combinação do intervalo escolhido com cada um dos restantes intervalos é testada. O procedimento acaba quando todos os intervalos forem testados, sendo a combinação escolhida a que gerar menor erro na validação cruzada [42].

2.4.4 Validação dos modelos

Após o desenvolvimento dos modelos é necessário proceder à sua validação. Devido a erros nas medições os parâmetros utilizados nos modelos de calibração podem não ser suficientemente exatos. Assim, a robustez dos modelos é testada através do uso de diferentes parâmetros estatísticos, tais como: o coeficiente de determinação (R^2), a raiz quadrada do erro médio quadrático da validação cruzada ($RMSE_{CV}$), o enviesamento (*bias*) e a raiz quadrada do erro médio quadrático da previsão ($RMSE_P$) [29].

- **Coeficiente de determinação (R^2)**

O coeficiente de determinação é uma medida da qualidade do ajuste da reta de regressão e o seu valor pode variar entre 0 e 1, sendo a reta de regressão um modelo perfeito quando toma o valor 1. Este parâmetro é dado pela equação 2.16 [35] [43].

$$R^2 = 1 - \frac{\sum(\hat{y}_i - y_i)^2}{\sum(y_i - \bar{y})^2} \quad (2.16)$$

Em que, $\hat{y}_i$ corresponde ao valor estimado pelo modelo, y_i ao valor medido pelo método de referência e $\bar{y}$ à média dos valores de referência.

- **Raiz quadrada do erro médio quadrático da validação cruzada ($RMSE_{CV}$)**

No método da validação cruzada, um espectro no conjunto de calibração é rejeitado e, procede-se à realização de uma calibração por PLS com os restantes espectros. A amostra rejeitada é utilizada para validar o modelo e, o procedimento é repetido com todas as amostras restantes [33].

O erro de previsão do modelo é definido como a raiz quadrada do erro médio quadrático da validação cruzada ($RMSE_{CV}$). Este parâmetro é dado pela equação 2.17 e é expresso nas unidades da quantificação efetuada [33].

$$RMSE_{CV} = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n_C - 1}} \quad (2.17)$$

Em que, n_C corresponde ao número de espectros do conjunto de calibração, $\hat{y}_i$ ao valor estimado pelo modelo e y_i ao valor medido pelo método de referência.

- **Enviesamento (*bias*)**

O parâmetro enviesamento é uma medida do desvio sistemático entre os valores lidos e os estimados pelo modelo. Este parâmetro é dado pela equação 2.18 [35].

$$bias = \frac{\sum_{i=1}^n (\hat{y}_i - y_i)}{n_C} \quad (2.18)$$

Em que, n_C corresponde ao número de espectros do conjunto de calibração, $\hat{y}_i$ ao valor estimado pelo modelo e y_i ao valor medido pelo método de referência. Este parâmetro pode também ser utilizado em conjuntos de previsão substituindo n_C por n_P , que corresponde ao número de amostras do conjunto de previsão.

- **Raiz quadrada do erro médio quadrático da previsão ($RMSE_P$)**

O erro de previsão do modelo é definido como a raiz quadrada do erro médio quadrático da previsão ($RMSE_P$). Para obter este parâmetro, um conjunto de amostras preparadas separadamente e excluídas da calibração é utilizado para validar o modelo desenvolvido. O número do conjunto de amostras deve ser elevado de modo a obter valores indicativos do desempenho do método. O $RMSE_P$ é dado pela equação 2.19 e é expresso nas unidades da quantificação efetuada [33].

$$RMSE_P = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n_P}} \quad (2.19)$$

Em que, n_P corresponde ao número de amostras do conjunto de previsão, $\hat{y}_i$ ao valor estimado pelo modelo e y_i ao valor medido pelo método de referência.

Conclui-se que, idealmente, um modelo adequado deve apresentar: baixo $RMSE_{CV}$, $RMSE_P$, enviesamento próximo de zero, bem como, alto R^2 [29] [43].

3 Materiais e Métodos

3.1 Impregnação laboratorial de papel

Para além da utilização de papel decorativo impregnado industrialmente utilizou-se papel impregnado em laboratório. O papel industrial tem de obedecer a um intervalo de aceitação assim, os seus respetivos teores de resina e voláteis encontram-se numa banda restrita. Contudo, na criação de modelos para a quantificação de parâmetros é necessária uma gama de amostras superior à que se pretende analisar. Uma gama alargada torna o modelo mais robusto o que é especialmente útil quando se espera obter valores anómalos [35].

A formulação das resinas utilizadas na impregnação de papel decorativo é constituída por resina MF, água e aditivos, nomeadamente endurecedor, desmoldante, humectante, anti-pó e *anti-blocking*. A adição de água permite diminuir a viscosidade e facilitar a impregnação. Por sua vez, o endurecedor facilita o processo de cura. O desmoldante impede que a resina adira e danifique o prato da prensa do HPL. O humectante permite que a resina adira mais facilmente ao papel. O aditivo anti-pó evita a adesão de pó e sujidade presente no ar à superfície do papel impregnado. Por fim, o *anti-blocking* evita a adesão das folhas de papel impregnado no armazenamento [3] [44]. A formulação do segundo banho é a que contém um maior número de aditivos.

Na primeira etapa de impregnação laboratorial de papel decorativo, folhas A4 são impregnadas de ambos os lados com a formulação do primeiro banho. De seguida, a folha é colocada num secador a 140 °C durante o tempo necessário para se obter o VC e RC pretendidos. Na terceira etapa, a folha de papel é impregnada de ambos os lados com a respetiva formulação utilizada no segundo banho. Por fim, a folha é colocada num secador a 140 °C durante o tempo necessário para se obter os VC e RC definidos.

3.2 Papel impregnado

No total foram estudados cinco tipos de papel impregnado, sendo que, para as amostras B1 e L foram construídos modelos para a quantificação do VC e RC. Os conjuntos das amostras estudadas para estes dois tipos de papel impregnado foram divididos de modo a 70 % ser utilizado para calibração e 30 % para validação dos modelos. Os restantes conjuntos de amostras de papel impregnado foram utilizados para validação do modelo do papel B1. Assim, foi possível verificar se o mesmo modelo seria capaz de prever amostras com gramagens e cores diferentes. A caracterização dos papeis impregnados estudados encontra-se na tabela 3.1.

Tabela 3.1 - Quantidade e características das amostras analisadas.

Nome da amostra	Cor	Gramagem (g/m ²)	Número de banhos	Número de amostras	
				Industriais	Laboratoriais
B1	Branco	80	1	31	69
			2	92	84
B2	Branco	105	2	28	-
B3	Branco	120	2	20	-
A	Azul	80	2	35	-
L	Cinzeno	80	1	-	130

A gama analisada das amostras B1 para o RC do primeiro banho foi de 34,38 a 52,16 % (m/m), para o RC do segundo banho de 46,38 a 64,97 % (m/m) e para os VC foi de 2,64 a 9,98 % (m/m). Relativamente ao papel L, a gama analisada das amostras para o RC do primeiro banho foi de 25,81 a 53,75 % (m/m), e para os VC foi de 1,34 a 13,25 % (m/m).

3.3 Teor de resina (RC) e Teor de voláteis (VC)

Como método de referência para a construção de modelos de quantificação de RC e VC, recorreu-se ao cálculo de diferenças de massas. No caso do RC, este corresponde à percentagem em massa de resina que o papel reteve após a impregnação no primeiro ou segundo banho. Relativamente ao VC, este corresponde ao teor de voláteis presentes no papel após o primeiro ou segundo banho e respetivas secagens.

Primeiramente, as amostras de papel são cortadas de modo a obter-se uma rodela com 100 cm² de área. De seguida, nas amostras submetidas a apenas um banho, o teor de resina (RC_1) é determinado pela equação 3.1.

$$RC_1 = \frac{m_1 - m_i}{m_1} \times 100 \quad (3.1)$$

Sendo m_1 a massa do papel impregnado após o primeiro banho e m_i a massa do papel seco. Para uma gramagem de 80 g/m², a massa de papel seco considerada foi de 0,8 g. Para os restantes papéis com diferentes gramagens, o raciocínio aplicado foi o mesmo.

No caso das amostras submetidas a dois banhos, o teor de resina (RC_2) é determinado pela equação 3.2.

$$RC_2 = \frac{m_2 - m_i}{m_2} \times 100 \quad (3.2)$$

Em que, m_2 corresponde à massa do papel impregnado após o segundo banho e m_i à massa do papel seco.

Por fim, para a determinação do teor de voláteis, a rodela da amostra é submetida a uma secagem a 160 °C durante 5 minutos. O teor de voláteis é então determinado pela equação 3.3.

$$VC = \frac{m_{R1} - m_{R2}}{m_{R1}} \times 100 \quad (3.3)$$

Em que, m_{R1} corresponde à massa do papel impregnado antes da secagem e m_{R2} à massa do papel após secagem.

Para cada papel impregnado existem valores ótimos de RC e VC. No caso dos papéis impregnados utilizados na construção de modelos, estes valores encontram-se na tabela 3.2.

Tabela 3.2 - Valores ótimos de RC e VC para os papéis impregnados utilizados na construção de modelos.

Amostra	RC (% (m/m))		VC (% (m/m))	
	1º banho	2º banho	1º banho	2º banho
B1	[38;40]	[54;56]	[4,5;5]	[6,0;7,0]
L	[38;40]	-	[4,5;5]	-

O papel B1 pode ser utilizado como papel barreira pelo que, nesse caso, é impregnado apenas com um banho. Assim, o VC é controlado após o primeiro banho. Contudo, se for utilizado como papel decorativo, o VC é apenas controlado após o segundo banho. Relativamente ao papel L, este apenas é impregnado com um banho.

Se os valores obtidos de RC e VC não forem adequados, será necessário tomar medidas para os aproximar dos ótimos. Os procedimentos a tomar encontram-se representados na tabela 3.3 [3].

Tabela 3.3 - Procedimento para aproximar os valores dos ótimos [3].

Parâmetro	Procedimento
RC > RC _{ótimo}	Aplicar menor quantidade de resina no próximo papel a impregnar
RC < RC _{ótimo}	Aplicar maior quantidade de resina no próximo papel a impregnar
VC > VC _{ótimo}	Aumentar o tempo de secagem após o 2º banho
VC < VC _{ótimo}	Diminuir o tempo de secagem após o 2º banho

3.4 Espectroscopia de infravermelho próximo (NIR)

Atualmente, na fábrica da EuroResinas, as análises de VC e RC ainda são realizadas *off-line* e requerem alguns minutos para obter resultados. A implementação da espectroscopia de NIR na linha de produção permite acelerar e automatizar esta determinação. Atualmente, já foram aplicados modelos com este propósito, nomeadamente para a análise do teor de voláteis com um coeficiente de determinação de 0,93 [45].

Existem, de momento, modelos previamente desenvolvidos para o papel impregnado B1 com um tempo de análise de aproximadamente 1 minuto e 40 segundos. Considera-se que estes modelos precisam de ser otimizados, diminuindo o tempo de aquisição de espectros, bem como, aumentando o valor de R^2 e diminuindo os valores de $RMSE_{CV}$ e $RMSE_P$. Pretende-se também aumentar a gama de propriedades dos modelos desenvolvidos [3] [4].

3.4.1 Descrição do equipamento

O equipamento utilizado para a aquisição dos espectros de papel impregnado foi espectrofotómetro FT-NIR do modelo Matrix-F da Bruker. O detetor utilizado foi do tipo TE-InGaAs e a sonda foi do modelo QR-400-7-VIS-BX 400 μm 2m Vis/NIR da Ocean Optics. O *software* utilizado para aquisição dos espectros foi o OPUS 6.5 da Bruker. Previamente à leitura dos espectros das amostras, foi realizada uma leitura do espectro de fundo (*background*) num circuito ótico fechado, no qual a fonte de luz se encontra diretamente ligada ao detetor. Todos os espectros foram convertidos de refletância para absorvância pelo *software*. Os espectros foram adquiridos num intervalo de número de onda de 4000 a 12000 cm^{-1} . Para a análise quimiométrica recorreu-se ao pacote QUANT 2 do *software* OPUS 6.5 da Bruker.

3.4.2 Estudo das condições de aquisição dos espectros

Previamente à aquisição de espectros para a construção de um modelo, estudou-se o efeito de três parâmetros de modo reduzir o tempo de análise, sem comprometer a qualidade dos resultados. Este estudo foi realizado em papel B1 impregnado industrialmente e, os parâmetros e os seus respetivos valores testados foram os seguintes:

- Resolução (cm^{-1}): 8, 16, 32, 64;
- Número de varrimentos: 32, 64, 128, 256;
- Frequência de varrimento (kHz): 5, 10 e 20.

3.4.3 Aquisição dos espectros

Como já foi confirmado em outros estudos, o papel impregnado tem uma distribuição heterogénea de resina [3] [4]. Na figura 3.1 encontra-se uma representação gráfica de espectros lidos ao longo da superfície de um papel impregnado industrialmente com o espaçamento de 0,2 mm, utilizando 8 varrimentos.

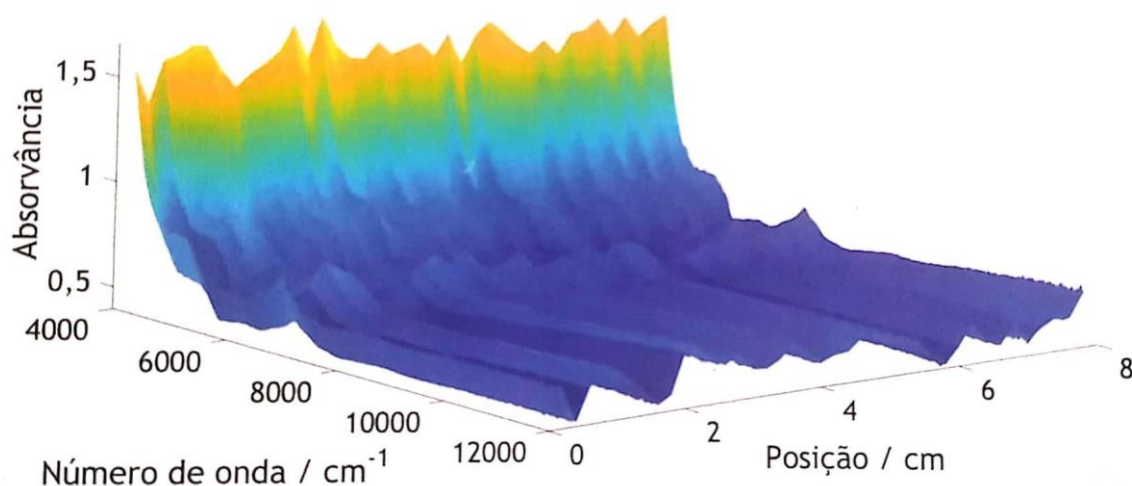


Figura 3.1 - Espectros adquiridos em papel impregnado industrialmente com o espaçamento de 0,2 mm [3].

De modo a tentar fazer uma aproximação à superfície heterogénea do papel impregnado, foram estudados dois procedimentos distintos para a aquisição de espectros NIR. O primeiro procedimento consistiu em, manualmente, registar três espectros da amostra em três posições diferentes. De seguida, a média destas três medições foi utilizada para a construção de modelos para a quantificação de RC e VC.

Para o segundo procedimento recorreu-se a um pórtico de mesa de alta precisão comandado por computador. Este permitiu realizar a recolha de varrimentos ao longo da superfície do papel impregnado movimentando a sonda NIR. Deste modo, espera-se que os espectros adquiridos sejam mais representativos da superfície heterogénea do papel impregnado. Assim, a sonda foi acoplada ao pórtico e movimentada a uma velocidade previamente definida com elevada precisão. Este procedimento permite simular a leitura *on-line* em ambiente fabril onde seriam aplicados os métodos desenvolvidos. Deste modo, é possível simular este ambiente em modo *off-line* em amostras à escala laboratorial sem ter de proceder à paragem ou alteração de processos industriais [4].

Para o papel B1 foi seguido o procedimento das médias. No caso do papel L, os dois procedimentos foram estudados em paralelo. Deste modo, foi possível concluir a melhor forma de contabilizar a superfície heterógena do papel impregnado.

3.4.4 Construção de modelos utilizando o *software* OPUS 6.5

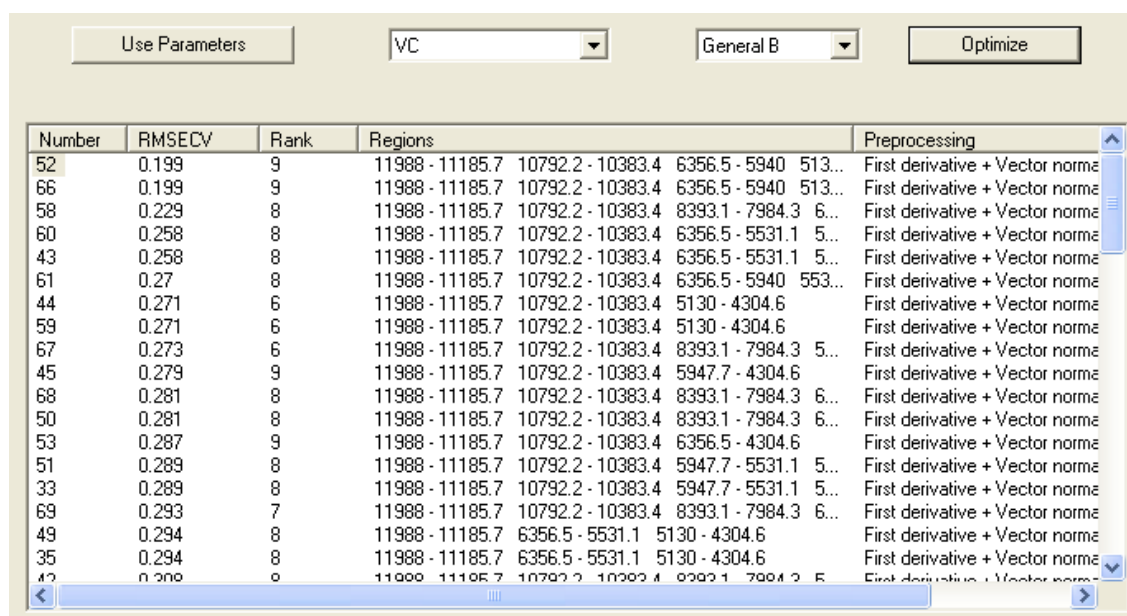
Para a construção dos modelos PLS seguiu-se o procedimento explicado nesta secção. Primeiramente, os espectros NIR obtidos foram introduzidos no *software* juntamente com os seus respetivos valores de RC e VC determinados pelo método de referência (pesagem diferencial).

De seguida, recorreu-se à ferramenta *Optimize*. Quando selecionada a opção *General B* no tipo de otimização, o *software* recorre a um procedimento semelhante ao fiPLS. A diferença entre estes procedimentos é que, no *General B*, quando são adicionados novos intervalos ao modelo, é testada a hipótese de uma das regiões previamente selecionadas estar a deteriorar a previsão. Assim, as regiões previamente selecionadas são removidas uma a uma e o valor de $RMSE_{CV}$ é analisado. O *software* permite definir no máximo 10 intervalos para esta análise [35]. Optou-se por construir os modelos dividindo a região espectral em análise em 5, 10 e 20 intervalos.

Assim, a análise utilizando 20 intervalos foi efetuada manualmente. Inicialmente, analisou-se os valores de $RMSE_{CV}$ e $RMSE_P$ dos modelos construídos utilizando 5 e 10 intervalos e, selecionou-se o pré-processamento mais adequado. De seguida, seguiu-se o procedimento fiPLS, combinando os melhores intervalos, até que os valores de $RMSE_{CV}$ não melhorassem [44]. Por fim, verificou-se se uma das regiões selecionadas estaria a degradar o modelo. As regiões selecionadas foram então removidas uma a uma e o valor de $RMSE_{CV}$ foi analisado.

Quando utilizada a ferramenta *Optimize*, para cada combinação de intervalos obtida, o *software* testa os diferentes pré-processamentos disponíveis e escolhe o número de fatores (*rank*) correspondentes a um menor $RMSE_{CV}$ [29] [35].

Os resultados da utilização da ferramenta *Optimize* são apresentados em forma de lista, ordenados de forma crescente pelo seu valor de $RMSE_{CV}$ ou $RMSE_P$. Um exemplo da janela obtida encontra-se representada na figura 3.2.



Number	RMSECV	Rank	Regions	Preprocessing
52	0.199	9	11988 - 11185.7 10792.2 - 10383.4 6356.5 - 5940 513...	First derivative + Vector norma
66	0.199	9	11988 - 11185.7 10792.2 - 10383.4 6356.5 - 5940 513...	First derivative + Vector norma
58	0.229	8	11988 - 11185.7 10792.2 - 10383.4 8393.1 - 7984.3 6...	First derivative + Vector norma
60	0.258	8	11988 - 11185.7 10792.2 - 10383.4 6356.5 - 5531.1 5...	First derivative + Vector norma
43	0.258	8	11988 - 11185.7 10792.2 - 10383.4 6356.5 - 5531.1 5...	First derivative + Vector norma
61	0.27	8	11988 - 11185.7 10792.2 - 10383.4 6356.5 - 5940 553...	First derivative + Vector norma
44	0.271	6	11988 - 11185.7 10792.2 - 10383.4 5130 - 4304.6	First derivative + Vector norma
59	0.271	6	11988 - 11185.7 10792.2 - 10383.4 5130 - 4304.6	First derivative + Vector norma
67	0.273	6	11988 - 11185.7 10792.2 - 10383.4 8393.1 - 7984.3 5...	First derivative + Vector norma
45	0.279	9	11988 - 11185.7 10792.2 - 10383.4 5947.7 - 4304.6	First derivative + Vector norma
68	0.281	8	11988 - 11185.7 10792.2 - 10383.4 8393.1 - 7984.3 6...	First derivative + Vector norma
50	0.281	8	11988 - 11185.7 10792.2 - 10383.4 8393.1 - 7984.3 6...	First derivative + Vector norma
53	0.287	9	11988 - 11185.7 10792.2 - 10383.4 6356.5 - 4304.6	First derivative + Vector norma
51	0.289	8	11988 - 11185.7 10792.2 - 10383.4 5947.7 - 5531.1 5...	First derivative + Vector norma
33	0.289	8	11988 - 11185.7 10792.2 - 10383.4 5947.7 - 5531.1 5...	First derivative + Vector norma
69	0.293	7	11988 - 11185.7 10792.2 - 10383.4 8393.1 - 7984.3 6...	First derivative + Vector norma
49	0.294	8	11988 - 11185.7 6356.5 - 5531.1 5130 - 4304.6	First derivative + Vector norma
35	0.294	8	11988 - 11185.7 6356.5 - 5531.1 5130 - 4304.6	First derivative + Vector norma
42	0.299	9	11988 - 11185.7 10792.2 - 10383.4 8393.1 - 7984.3 5...	First derivative + Vector norma

Figura 3.2 - Exemplo do resultado da ferramenta *Optimize*.

Contudo, é de notar que nem sempre o método que apresenta um menor erro é o mais vantajoso. É necessário ter em atenção se não está a ocorrer o uso de fatores em excesso ou está a ser considerado um número de valores anómalos demasiado elevado. A escolha de um

número de fatores demasiado elevado poderá resultar em *overfitting*, adicionando à informação ruído e degradando o modelo. Por sua vez, a escolha de um número de fatores demasiado baixo poderá resultar em *underfitting* descartando informação importante para o modelo [29] [35].

De modo a verificar o número de fatores escolhido, o *software* representa graficamente os valores de $RMSE_{CV}$ em função do número de fatores. Um exemplo deste gráfico encontra-se representado na figura 3.3.

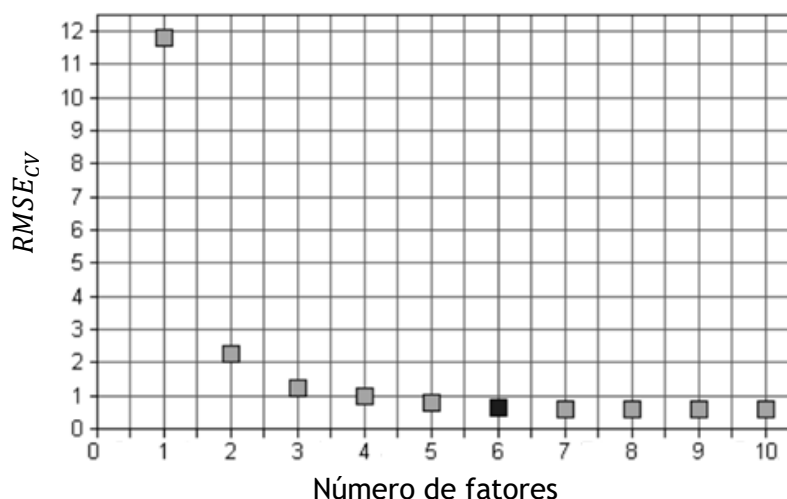


Figura 3.3 - Exemplo de uma representação gráfica de $RMSE_{CV}$ em função do número de fatores (adaptado de [35]).

Verifica-se neste exemplo que valores de número de fatores maiores que 6 não melhoram significativamente o modelo. Conclui-se que valores maiores apenas degradam o modelo, adicionando-lhe informação que apenas constitui ruído. Para todos os modelos, o *software* testa no máximo um valor de 10 fatores por ser considerado suficiente para determinar o valor ótimo. Esta restrição permite reduzir o tempo de processamento, especialmente quando utilizado um grande número de amostras [35].

Concluindo, a construção de modelos por PLS é um processo iterativo composto pelos seguintes passos:

1. Analisar os resultados da validação e selecionar o número de fatores mais adequado;
2. Para este número de fatores remover possíveis valores anómalos;
3. Determinar novamente o número de fatores ótimo;
4. Testar outras opções de pré-processamentos e, se necessário, alterar regiões espectrais utilizadas [35].

3.4.5 Remoção de amostras anômalas

A remoção de amostras anômalas foi efetuada de duas formas diferentes. O primeiro caso correspondeu a descartar amostras com valores de RC , VC ou espectros anômalos, previamente à construção dos modelos.

No segundo caso, durante a construção de modelos, a distância Mahalanobiana em função do valor residual dos espectros de cada amostra é representada. A distância Mahalanobiana refere-se à distância entre cada espectro e um espectro médio determinado pelo *software*. Através deste gráfico, é possível determinar os espectros que apresentaram uma probabilidade superior a 99 % de serem valores anômalos [29] [35].

A seleção e remoção de amostras anômalas permite diminuir o $RMSE_{CV}$ e o número de fatores utilizado na construção dos modelos [29]. Considerou-se que a remoção superior a 10 % do número total de amostras do mesmo tipo era excessiva, pelo que, este valor não foi ultrapassado [46].

O número total de amostras anômalas consideradas para cada modelo encontra-se na tabela 3.4.

Tabela 3.4 - Número de amostras anômalas para cada modelo.

Papel	Procedimento	Modelo					
		RC_1	RC_2	RC	VC_1	VC_2	VC
B1	Médias	7	11	18	3	8	11
L	Médias	8	-	-	10	-	-
	Pórtico	8	-	-	11	-	-

4 Resultados e Discussão

Previamente à construção de modelos para a quantificação dos teores de resina e voláteis no papel impregnado, estudou-se o efeito dos parâmetros na aquisição dos espectros. Pretende-se reduzir, sem prejudicar os resultados obtidos, o tempo de análise e, se possível, torná-lo inferior a um minuto. A escolha dos parâmetros de aquisição encontra-se descrita em detalhe no apêndice A.

Conclui-se que, das combinações estudadas a que permitiu chegar a este objetivo foi: 16 cm^{-1} de resolução, 128 varrimentos e 20 kHz de frequência de varrimento correspondendo a um tempo total de análise de 22 segundos.

4.1 Papel B1

4.1.1 Criação de um modelo PLS para determinar o RC_1

Primeiramente, procedeu-se à criação de um modelo para a quantificação do RC_1 . Os espectros utilizados foram adquiridos utilizando o procedimento das médias e encontram-se representados na figura B.1 do apêndice B.

Para identificar as regiões espectrais com maior variância, aplicou-se o PCA ao conjunto de espectros obtidos de modo a determinar os *loadings*. Previamente a esta análise, os espectros foram tratados utilizando o pré-processamento SNV. Assim, foi possível remover variações nos espectros resultantes de fatores físicos, tais como, os desvios na linha de base. Os primeiros três *loadings* obtidos encontram-se representados graficamente na figura B.2 do apêndice B.

Através da análise da figura B.2, conclui-se que a região espectral que compreende a maior variância entre espectros situa-se entre 7000 e 4000 cm^{-1} . Deste modo, é esperado que a ferramenta *General B* selecione esta região para construção dos modelos do RC_1 . Verifica-se também que, a região entre 12000 e 7000 cm^{-1} é a que continha mais ruído. Contudo, para outros pré-processamentos, este poderá ser atenuado.

Após esta análise, procedeu-se à construção de um modelo PLS. De modo a escolher o número de intervalos e pré-processamento a utilizar, analisou-se os valores dos parâmetros $RMSE_{CV}$, $RMSE_p$ e R^2 . Assim, foram feitas duas validações: a validação cruzada, que permite avaliar a metodologia utilizada, e uma validação externa para avaliar o modelo. A validação externa foi efetuada com amostras que não foram utilizadas na calibração, correspondendo a 30 % do número total das amostras com um banho. Estas amostras foram selecionadas pelo PCA, de modo a cobrir toda a gama estudada. Os resultados obtidos encontram-se representados na tabela C.1 do apêndice C. As regiões referidas na tabela C.1 correspondem aos intervalos selecionados que, foram numerados por ordem decrescente de número de onda.

Analisando os resultados obtidos, conclui-se que o modelo mais adequado não utilizou pré-processamentos e recorreu à ferramenta *General B* com 10 intervalos para selecionar as regiões $[11988;10383,4] \text{ cm}^{-1}$ e $[5600,6;3996] \text{ cm}^{-1}$. Para a validação cruzada, representada na figura 4.1, o $RMSE_{CV}$ obtido foi de 0,929 % (m/m), com um R^2 de 0,90 com um número de fatores de 5 e um enviesamento de 0,00443. Por sua vez, para a validação externa, o $RMSE_p$ obtido foi de 0,923 % (m/m), com um R^2 de 0,86 com um número de fatores de 2 e um enviesamento de 0,0411. A região delimitada a azul representa o intervalo de aceitação da propriedade em estudo.

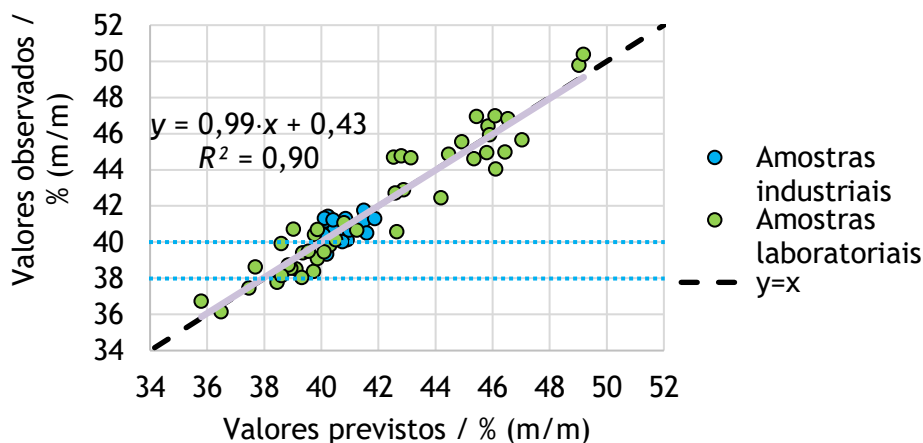


Figura 4.1 - Validação cruzada para a metodologia de quantificação do RC_1 .

Verifica-se que, tal como se tinha previsto pelos *loadings* do PCA, parte da região entre 7000 e 4000 cm^{-1} foi selecionada pelo que está associada a picos das ligações C-H, N-H e O-H presentes na resina. Relativamente ao intervalo $[11988;10383,4] \text{ cm}^{-1}$, neste está compreendido um pico associado à presença de água [3] [47]. É de notar que, o teor de água também influencia o RC final. Os valores determinados de $RMSE_{CV}$ e $RMSE_p$ foram elevados, pelo que, foram significativamente superiores a zero. Assim, o erro máximo, para a validação cruzada, correspondente à diferença entre o valor previsto e o verdadeiro, foi de -2,17 % (m/m). Apesar do valor de R^2 ser elevado, manteve-se inferior a 0,95. Relativamente, ao número de fatores, este foi considerado adequado, pelo que, manteve-se inferior a oito [46].

Deste modo, conclui-se que a capacidade de previsão do teor de resina pelo modelo foi baixa. Admite-se que, embora o procedimento das médias seja mais eficaz do que o uso de apenas uma medição, não foi suficiente para contabilizar a superfície heterogênea do papel impregnado.

4.1.2 Criação de um modelo PLS para determinar o RC_2

De seguida, procedeu-se à criação de um modelo para a quantificação do RC_2 . Os espectros utilizados foram adquiridos utilizando o procedimento das médias e encontram-se representados na figura B.3 do apêndice B.

De modo a proceder à análise realizada anteriormente e, identificar as regiões espectrais com maior variância, aplicou-se novamente o PCA ao conjunto de espectros obtidos. Os primeiros três *loadings* obtidos encontram-se representados graficamente na figura B.4 do apêndice B.

Através da análise da figura B.4, conclui-se que a região espectral que compreende a maior variância entre espectros situa-se entre 7500 e 4000 cm^{-1} . Deste modo, é esperado que a ferramenta *General B* selecione esta região para construção dos modelos do RC_2 . Por sua vez, a região entre 12000 e 8000 cm^{-1} é a que continha mais ruído. Contudo, para outros pré-processamentos, este poderá ser atenuado.

O modelo PLS foi construído recorrendo aos critérios utilizados previamente. A validação externa foi efetuada com amostras que não foram utilizadas na calibração, correspondendo a 30 % do número total das amostras com dois banhos. Estas amostras foram selecionadas pelo PCA. Os resultados obtidos para as validações efetuadas encontram-se representados na tabela C.2 do apêndice C.

Analizando os resultados obtidos, conclui-se que o modelo mais adequado utilizou como pré-processamento o SNV e recorreu à ferramenta *General B* com 10 intervalos para selecionar as regiões [11988;11185,7], [10391,1;7992] e [6395,1;4790,6] cm^{-1} . Para a validação cruzada, representada na figura 4.2, o $RMSE_{CV}$ obtido foi de 1,01 % (m/m), com um R^2 de 0,82 com um número de fatores de 6 e um enviesamento de 0,0442. Por sua vez, para a validação externa, o $RMSE_p$ obtido foi de 1,3 % (m/m), com um R^2 de 0,74 com um número de fatores de 4 e um enviesamento de 0,0303. A região delimitada a rosa representa o intervalo de aceitação da propriedade em estudo.

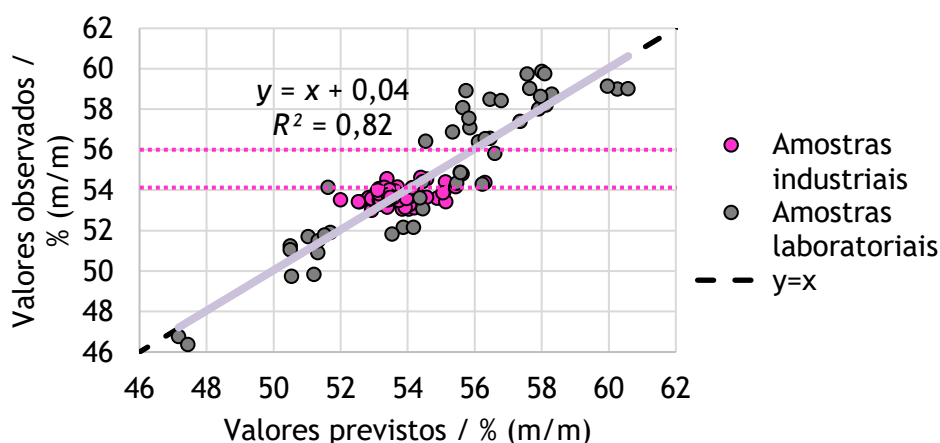


Figura 4.2 - Validação cruzada para a metodologia de quantificação do RC_2 .

Verifica-se que, parte da região prevista pelo PCA foi selecionada. O intervalo [6395,1;4790,6] cm^{-1} compreende picos das ligações C-H, N-H e O-H presentes na resina. Relativamente aos intervalos [11988;11185,7] e [10391,1;7992] cm^{-1} , estes contêm picos associados à presença de água [3] [47]. Novamente, os valores determinados de $RMSE_{CV}$ e $RMSE_p$ foram elevados, pelo que, foram significativamente superiores a zero. Assim, o erro

máximo da validação cruzada foi de -3,19 % (m/m). O R^2 obtido é considerado baixo pois, foi inferior a 0,95. Relativamente ao número de fatores, este foi considerado adequado, pelo que, manteve-se inferior a oito [46].

Conclui-se que, o modelo para a quantificação do RC após o segundo banho teve um desempenho semelhante ao criado para as amostras com um banho. Deste modo, a capacidade de previsão do teor de resina pelo modelo foi baixa.

4.1.3 Criação de um modelo PLS para determinar o RC

Os modelos desenvolvidos para a quantificação do RC_1 e RC_2 quantificam a mesma resina melamínica, embora com formulações diferentes. Assim, idealmente, seria possível juntar os modelos construídos e criar um modelo geral para quantificar o RC.

Optou-se por testar esta hipótese. Para a criação deste novo modelo, considerou-se os dois conjuntos de calibração utilizados anteriormente, bem como, por coerência, as mesmas amostras anómalas.

As regiões espectrais obtidas pela análise dos *loadings* para o modelo do RC_1 e RC_2 foram semelhantes o que indica que poderá ser possível juntar estes modelos. Visto que, as mesmas regiões foram selecionadas não se procedeu a essa análise para este modelo.

O novo modelo PLS geral foi construído recorrendo aos critérios utilizados previamente. A validação externa foi efetuada com um conjunto de validação obtido pela junção dos que foram utilizados para os modelos do RC_1 e RC_2 . Os resultados obtidos para as validações efetuadas encontram-se representados na tabela C.3 do apêndice C.

Analisando os resultados obtidos, conclui-se que o modelo mais adequado utilizou como pré-processamento o SNV e recorreu à ferramenta *General B* com 10 intervalos para selecionar as regiões [11193,4;8786,5] e [7197,4;3996] cm^{-1} . Para a validação cruzada, representada na figura 4.3, o $RMSE_{CV}$ obtido foi de 1,05 % (m/m), com um R^2 de 0,98 com um número de fatores de 6 e um enviesamento de 0,0144. Por sua vez, para a validação externa, o $RMSE_p$ obtido foi de 1,31 % (m/m), com um R^2 de 0,96 com um número de fatores de 5 e um enviesamento de 0,256. As regiões delimitadas a azul e rosa representam os intervalos de aceitação para o RC_1 e RC_2 , respetivamente.

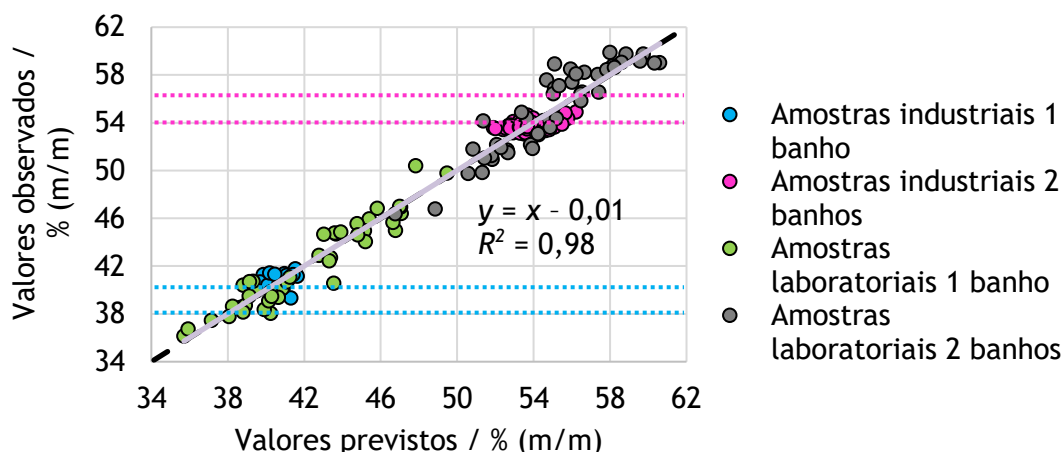


Figura 4.3 - Validação cruzada para a metodologia de quantificação do RC.

A região $[7197,4;3996]$ cm^{-1} foi selecionada tal como tinha sido previsto pelas análises dos *loadings* realizadas para os modelos específicos. Os valores determinados de $RMSE_{CV}$ e $RMSE_P$ mantiveram-se elevados, pelo que, foram significativamente superiores a zero. Assim, para a validação cruzada o erro máximo foi de -3,82 % (m/m). No entanto, o R^2 foi superior a 0,95, pelo que, se considera o valor obtido satisfatório. Relativamente ao número de fatores, este foi considerado adequado, pelo que, manteve-se inferior a oito [46].

Contudo, devido aos diferentes espaços amostrais utilizados na validação e calibração destes modelos, não é possível recorrer aos valores calculados de R^2 , $RMSE_{CV}$ e $RMSE_P$ como critério de comparação. Deste modo, foram realizadas mais duas validações externas para o modelo criado para o RC. Na primeira validação externa, utilizaram-se as amostras de validação com um banho e, para a segunda, as com dois banhos. A comparação entre o desempenho dos modelos encontra-se representada na tabela 4.1.

Tabela 4.1 - Comparação entre o desempenho dos modelos para o RC.

Modelo	Conjunto com 1 banho		Conjunto com 2 banhos	
	RC_1	RC	RC_2	RC
R^2	0,86	0,76	0,74	0,73
$RMSE_P$ (% (m/m))	0,923	1,21	1,2	1,34
Número de fatores	2	4	4	5
Erro máximo (% (m/m))	3,21	2,45	3,09	-4,15

Verifica-se que o modelo geral do RC não conseguiu prever com a mesma precisão as amostras com um banho. Contudo, o desempenho do modelo geral foi semelhante ao modelo específico para as amostras de dois banhos sendo que, não houve diferenças significativas nos valores de R^2 e $RMSE_P$. Conclui-se que o desempenho do modelo geral foi mais adequado para amostras com dois banhos. No entanto, a junção dos modelos introduziu um erro significativo na previsão do teor de resina das amostras com um banho. Assim, considera-se que a capacidade de previsão se manteve baixa para a quantificação do RC_1 e RC_2 .

4.1.4 Criação de um modelo PLS para determinar o VC_2

Procedeu-se à construção de um modelo para a quantificação teor de voláteis após o segundo banho, VC_2 . Utilizou-se os espectros do conjunto de calibração do modelo do RC_2 , à exceção dos valores anómalos que não foram identificados em ambas as amostragens. A seleção de valores anómalos diferentes no processo de amostragem pode ser devida a erros laboratoriais. Entre estes estão: a pesagem incorreta das amostras, erros na secagem do papel para a determinação do VC ou o corte incorreto das rodelas afetando maioritariamente o RC, visto que, este usa como referência uma massa fixa.

Não foi efetuada novamente a análise dos *loadings* do PCA, uma vez que, esta seria semelhante à da figura B.2 do apêndice B. Deste modo, sabe-se que a região espectral que compreende a maior variância entre espectros situa-se entre 7500 e 4000 cm^{-1} . Admitindo que o teor de voláteis é maioritariamente influenciado pela humidade do papel impregnado, espera-se que sejam selecionados picos associados à presença de água. Estes estão situados nos seguintes números de onda: 11,800, 8,600, 6,900, 5,620, 5,200, 5,150, e 4,115 cm^{-1} [47].

Durante a construção do modelo PLS recorreu-se aos critérios utilizados anteriormente. A validação externa foi efetuada com amostras que não foram utilizadas na calibração, correspondendo a 30 % do número total das amostras com dois banhos. Estas amostras foram selecionadas pelo PCA. Os resultados obtidos para as validações efetuadas encontram-se representados na tabela C.4 do apêndice C.

Analizando os resultados obtidos, conclui-se que o modelo mais adequado utilizou como pré-processamento a Normalização Mín-máx e recorreu à ferramenta *General B* com 10 intervalos para selecionar as regiões [11988;11185,7], [8794,3;6387,4] e [5600,6;4790,6] cm^{-1} . Para a validação cruzada, representada na figura 4.4, o $RMSE_{CV}$ obtido foi de 0,386 % (m/m), com um R^2 de 0,91 com um número de fatores de 5 e um enviesamento de -0,0003. Para a validação externa, o $RMSE_p$ obtido foi de 0,389 % (m/m), com um R^2 de 0,90 com um número de fatores de 4 e um enviesamento de -0,0548. A região a rosa delimita o intervalo de aceitação do VC_2 .

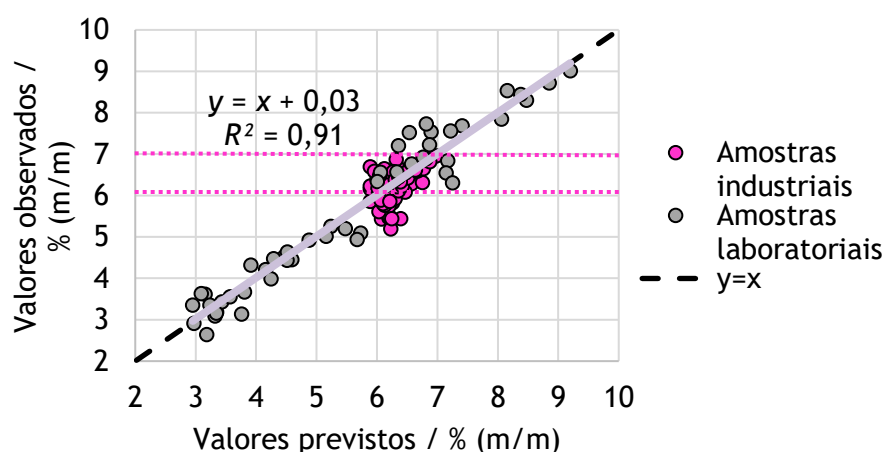


Figura 4.4 - Validação cruzada para a metodologia de quantificação do VC_2 .

Verifica-se que, as regiões selecionadas contêm picos associados à presença de água, bem como, parte da região prevista pelos *loadings* do PCA. Os valores determinados de $RMSE_{CV}$ e $RMSE_P$ foram significativamente inferiores aos valores obtidos para a quantificação do teor de resina, pelo que, o teor de voláteis foi previsto com maior precisão. Assim, para a validação cruzada, o erro máximo foi de 1,05 % (m/m). Considera-se que, apesar dos valores de $RMSE_{CV}$ e $RMSE_P$ serem ainda significativamente diferentes de zero, este modelo permitiu prever o VC das amostras. O R^2 obtido é considerado baixo pois, foi inferior a 0,95. Relativamente ao número de fatores, este foi considerado adequado, pelo que, manteve-se inferior a oito [46].

4.1.5 Criação de um modelo PLS para determinar o VC

No caso do papel B1 barreira, é determinado o teor de voláteis após a impregnação com apenas um banho. Assim, optou-se por criar um modelo geral para a quantificação do teor de voláteis, tal como se tinha efetuado para o teor de resina.

Utilizou-se as amostras de calibração com dois banhos utilizadas no modelo específico. Adicionaram-se ao conjunto de calibração 30 amostras industriais com um banho. Apenas foram consideradas 16 amostras laboratoriais pois, não foi alterado o tempo de secagem no primeiro banho. Assim, as amostras laboratoriais apresentaram um teor de voláteis semelhante, portanto, não se considerou relevante adicionar mais amostras deste tipo. Excluíram-se, por coerência, as mesmas amostras anómalas com dois banhos detetadas anteriormente. No caso das amostras com um banho, excluíram-se 3 amostras anómalas.

Optou-se por não efetuar a análise dos *loadings* do PCA, visto que, as figuras B.2 e B.4 do apêndice B permitiram selecionar as mesmas regiões.

Durante a construção do modelo PLS geral recorreu-se aos critérios utilizados anteriormente. A validação externa foi efetuada com o conjunto de validação utilizado no modelo do VC_2 . Os resultados obtidos para as validações efetuadas encontram-se representados na tabela C.5 do apêndice C.

Analisando os resultados obtidos, conclui-se que o modelo mais adequado utilizou como pré-processamento a Normalização Mín-máx e recorreu à ferramenta *General B* com 10 intervalos para selecionar as regiões [11193,4;10383,4] e [7197,4;4790,6] cm^{-1} . Para a validação cruzada, representada na figura 4.5, o $RMSE_{CV}$ obtido foi de 0,4 % (m/m), com um R^2 de 0,91 com um número de fatores de 6 e um enviesamento de -0,0024. Por sua vez, para a validação externa, o $RMSE_P$ obtido foi de 0,391 % (m/m), com um R^2 de 0,90 com um número de fatores de 4 e um enviesamento de -0,0117. As regiões delimitadas a azul e rosa representam os intervalos de aceitação para o VC_1 e VC_2 , respetivamente.

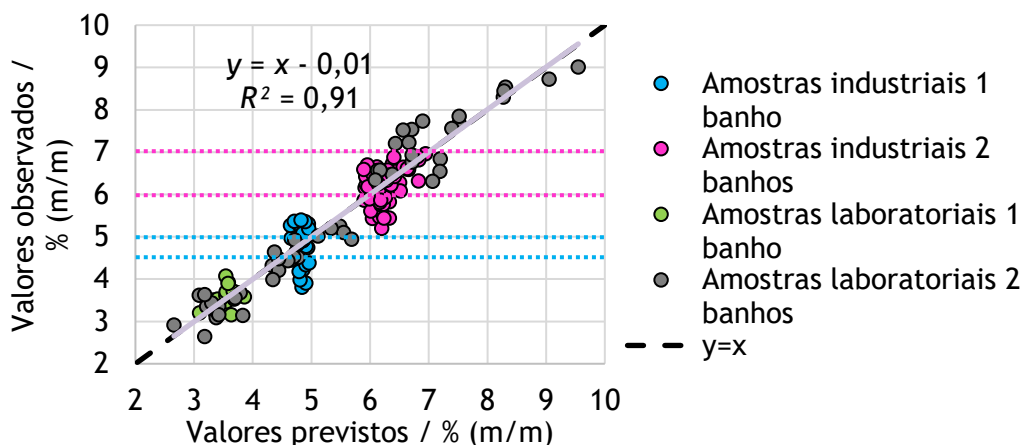


Figura 4.5 - Validação cruzada para a metodologia de quantificação do VC.

O intervalo $[7197,4;4790,6] \text{ cm}^{-1}$ foi selecionado, tal como previsto através análise dos *loadings*. Os valores determinados de $RMSE_{CV}$ e $RMSE_P$ mantiveram-se significativamente inferiores aos valores obtidos para a quantificação do teor de resina. Assim, para a validação cruzada, o erro máximo foi de 1,02 % (m/m). O R^2 obtido é considerado baixo pois, foi inferior a 0,95. O número de fatores, escolhido foi considerado adequado, pelo que, manteve-se inferior a oito [46].

Através da análise da figura 4.5, verifica-se que a previsão do teor de voláteis das amostras industriais com um banho foi anómala. Apesar de os valores observados do VC destas amostras variarem entre aproximadamente 4 e 5,5 % (m/m), os valores previstos mantiveram-se sem sofrerem alterações significativas.

De modo a verificar se tinha ocorrido deterioramento da capacidade de previsão do modelo, e compará-lo com o modelo específico, comparou-se os resultados da validação externa com as amostras de validação com dois banhos. A comparação entre o desempenho dos modelos encontra-se representada na tabela 4.2.

Tabela 4.2 - Comparação entre o desempenho dos modelos para o VC.

Modelo	Conjunto com 2 banhos	
	VC_2	VC
R^2	0,90	0,90
$RMSE_P$ (% (m/m))	0,389	0,391
Número de fatores	4	4
Erro máximo (% (m/m))	0,88	-0,93

Verifica-se que a adição do conjunto de amostras com um banho não resultou na deterioração significativa da capacidade de previsão do VC das amostras com dois banhos. Contudo, para tirar melhores conclusões, seria necessário proceder a uma análise mais detalhada como no caso do RC. Assim, teria de se impregnar amostras com um banho com uma gama alargada de VC para adicionar ao modelo e, criar um conjunto de validação para estas.

4.1.6 Validação externa com diferentes papeis impregnados

Após a criação dos modelos de quantificação RC e VC gerais e específicos, optou-se por fazer uma validação externa adicional. Esta análise tem como propósito verificar se os modelos desenvolvidos para o papel B1 podem ser aplicados em outros papeis impregnados com a mesma resina. Estudaram-se papeis impregnados industrialmente com dois banhos e diferentes gramagens, B2 e B3, e cores, A. É de notar que, visualmente, o papel B2 apresentou uma tonalidade de branco diferente dos papeis B1 e B3. Os resultados obtidos encontram-se na tabela 4.3.

Tabela 4.3 - Resultados da validação externa com diferentes papeis impregnados.

		<i>RMSE_p</i> (% (m/m))		
Papel		B2	B3	A
Modelo	<i>RC</i> ₂	0,996	4,67	1,77
	<i>RC</i>	2,3	2,37	1,46
	<i>VC</i> ₂	0,314	0,671	0,311
	<i>VC</i>	0,583	0,957	0,27

Através da análise da tabela 4.3, conclui-se que o parâmetro medido com maior precisão foi o teor de voláteis (VC), como se tinha verificado anteriormente. No geral, o papel cujas características foram previstas com maior precisão foi o papel A. Por sua vez, os modelos desenvolvidos apresentaram a pior capacidade de previsão para o papel de maior gramagem, B3. Verificou-se que, na maioria dos modelos, os valores previstos foram consistentemente superiores aos observados para os papeis de maior gramagem. Deste modo, poderão estar a ser contabilizadas as ligações do papel do tipo O-H, presentes na água, e C-H, sendo que ambas as ligações estão presentes na resina. Assim, é gerando um erro superior na previsão do teor de resina do que no teor de voláteis.

Conclui-se que, idealmente, deveriam ser criados modelos para cada um destes papeis impregnados. Contudo, a capacidade de previsão dos modelos construídos do teor de voláteis manteve-se para as amostras dos papeis B2 e A estudadas. Relativamente ao teor de resina, o melhor desempenho foi o do modelo do *RC*₂ para a amostra B2. Contudo, a capacidade de previsão para esta característica manteve-se baixa. Há que referir que, os papeis utilizados são todos industriais, pelo que, a sua gama de propriedades é bastante restrita. Assim, para obter valores de *RMSE_p* mais representativos e, retirar melhores conclusões, seria necessário alargar a gama em estudo.

4.2 Papel L

Procedeu-se à análise de um novo papel, L, de cor cinza e gramagem 80 g/m². Verificou-se para o papel B1 que uma das possíveis causas dos elevados erros de previsão poderia ser o

procedimento de aquisição dos espectros. Assim, optou-se por seguir também um novo procedimento, o do pósito. Os modelos baseados nestes dois procedimentos foram analisados em paralelo. Deste modo, foi possível verificar se a utilização do pósito permitiria a aquisição de espectros mais representativos da superfície heterogénea do papel impregnado.

4.2.1 Criação de modelos PLS para determinar o RC_1

- Procedimento das médias

Os espectros do papel L adquiridos utilizando o procedimento das médias encontram-se representados na figura B.5 do apêndice B.

De modo a identificar as regiões espectrais com maior variância aplicou-se o PCA ao conjunto de espectros obtidos de modo a determinar os *loadings*. Previamente a esta análise, os espectros foram tratados utilizando o pré-processamento SNV. Os primeiros três *loadings* obtidos encontram-se representados graficamente na figura B.6 do apêndice B.

Através da análise da figura B.6, conclui-se que a região espectral que compreende a maior variância entre espectros situa-se entre 7300 e 4000 cm^{-1} . Verifica-se que esta região é semelhante às obtidas para o papel B1, pelo que, conclui-se que não houve influência da cor do papel nesta seleção. Deste modo, é esperado que a ferramenta *General B* selecione esta região para construção dos modelos do RC_1 . Verifica-se também que, a região entre 12000 e 10500 cm^{-1} é a que continha mais ruído. Contudo, para outros pré-processamentos, este poderá ser atenuado.

Após esta análise, procedeu-se à construção de um modelo PLS. A validação externa foi efetuada com amostras que não foram utilizadas na calibração, correspondendo a 30 % do número total das amostras com um banho. Estas amostras foram selecionadas pelo PCA, de modo a cobrir toda a gama estudada. Os resultados obtidos encontram-se representados na tabela C.6 do apêndice C.

Analisando os resultados obtidos, conclui-se que o modelo mais adequado utilizou como pré-processamento o COE e recorreu à ferramenta *General B*, efetuada manualmente, com 20 intervalos para selecionar as regiões [11594,5;11185,7], [10391,1;9982,2], [6395,1;5986,3] e [5191,7;3996] cm^{-1} . Para a validação cruzada, representada na figura 4.6, o $RMSE_{CV}$ obtido foi de 1,47 % (m/m), com um R^2 de 0,91 com um número de fatores de 7 e um enviesamento de 0,0245. Por sua vez, para a validação externa, o $RMSE_p$ obtido foi de 1,54 % (m/m), com um R^2 de 0,88 com um número de fatores de 5 e um enviesamento de 0,106. A região delimitada a verde na figura 4.6 representa o intervalo de aceitação da propriedade em estudo.

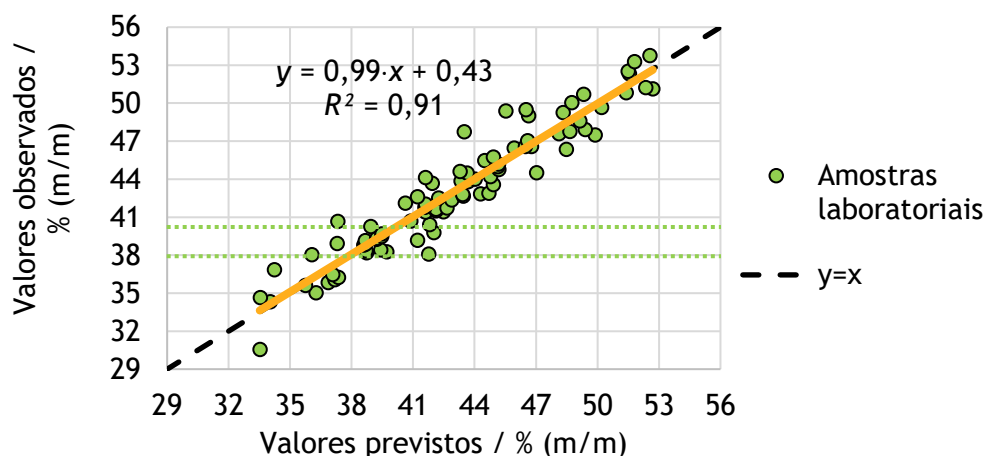


Figura 4.6 - Validação cruzada para a metodologia de quantificação do RC_1 , utilizando o procedimento das médias.

Verifica-se que, tal como se tinha previsto pelos *loadings* do PCA, parte da região entre 7300 e 4000 cm^{-1} foi selecionada pelo que está associada a picos das ligações C-H, N-H e O-H presentes na resina. Os valores determinados de $RMSE_{CV}$ e $RMSE_P$ foram elevados, pelo que, foram significativamente superiores a zero. Assim, o erro máximo, para a validação cruzada, correspondente à diferença entre o valor previsto e o verdadeiro, foi de -4,24 % (m/m). Apesar do valor de R^2 ser elevado, manteve-se inferior a 0,95. Por sua vez, o número de fatores, foi considerado adequado, pelo que, manteve-se inferior a oito [46].

Deste modo, conclui-se que a capacidade de previsão do teor de resina pelo modelo foi baixa. Admite-se que, tal como para o papel impregnado B1, o procedimento das médias não foi suficiente para contabilizar a superfície heterogênea do papel impregnado.

- **Procedimento do pórtico**

Os espectros utilizados foram adquiridos utilizando o procedimento do pórtico e encontram-se representados na figura B.7 do apêndice B.

De modo a proceder à análise realizada anteriormente e, verificar se tinha ocorrido alteração das regiões espectrais com maior variância, aplicou-se novamente o PCA ao conjunto de espectros obtidos. Os primeiros três *loadings* obtidos encontram-se representados graficamente na figura B.8 do apêndice B.

Através da análise da figura B.8, conclui-se não houve variações significativas nos *loadings* determinados pelo PCA com a alteração do procedimento. Assim, a região espectral correspondente a uma maior variância entre espectros mantém-se entre 7300 e 4000 cm^{-1} . Contudo, verifica-se um aumento do ruído na região entre 12000 e 10000 cm^{-1} .

O modelo PLS foi construído recorrendo aos critérios utilizados previamente. Para futuros efeitos comparativos, a validação externa foi efetuada com as amostras selecionadas

previamente para o modelo do procedimento das médias. Os resultados obtidos para as validações efetuadas encontram-se representados na tabela C.7 do apêndice C.

Conclui-se que o modelo mais adequado utilizou como pré-processamento a primeira derivada com o MSC e recorreu à ferramenta *General B* com 5 intervalos para selecionar a região $[8794,3;3996]$ cm^{-1} . Para a validação cruzada, representada na figura 4.7, o $RMSE_{CV}$ obtido foi de 1,06 % (m/m), com um R^2 de 0,96 com um número de fatores de 4 e um enviesamento de 0,00823. Por sua vez, para a validação externa, o $RMSE_P$ obtido foi de 0,765 % (m/m), com um R^2 de 0,97 com um número de fatores de 4 e um enviesamento de -0,232. A região delimitada a verde representa o intervalo de aceitação da propriedade em estudo.

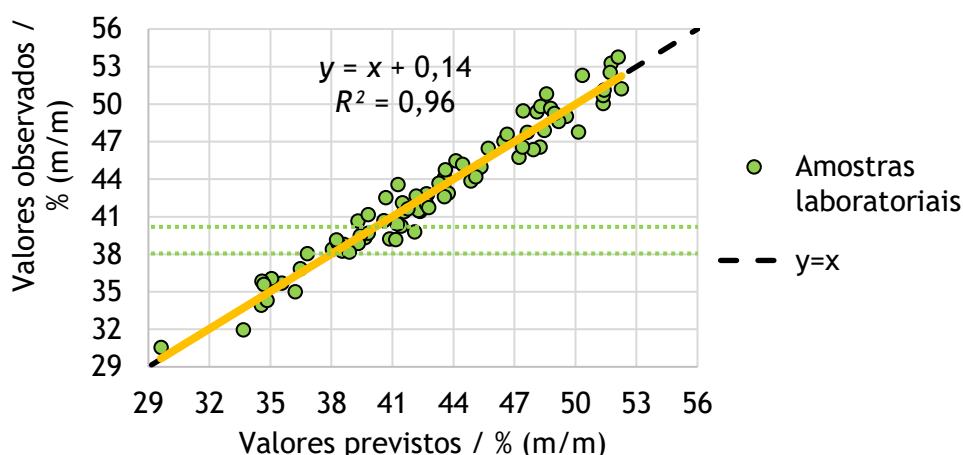


Figura 4.7 - Validação cruzada para a metodologia de quantificação do RC_1 , utilizando o procedimento do pórtico.

Verifica-se que, parte da região prevista pelo PCA foi selecionada, pelo que, compreende picos das ligações C-H, N-H e O-H associados à resina, bem como, picos associados à presença de água. Novamente, os valores determinados de $RMSE_{CV}$ e $RMSE_P$, embora tenham diminuído, mantiveram-se significativamente superiores a zero. Assim, o erro máximo da validação cruzada foi de 2,37 % (m/m). Verificou-se também uma melhoria no valor do R^2 obtido. Este foi considerado satisfatório pois, foi superior a 0,95. Por sua vez, o número de fatores foi considerado adequado, pelo que, manteve-se inferior a oito [46].

Conclui-se que, houve uma melhoria significativa dos valores de $RMSE_{CV}$ e $RMSE_P$. Assim, admite-se o procedimento do pórtico permitiu contabilizar com maior precisão a superfície heterogénea do papel impregnado. Contudo, a capacidade de previsão do teor de resina pelo modelo foi ainda inferior ao esperado, pelo que, os erros obtidos se mantiveram altos.

4.2.2 Criação de modelos PLS para determinar o VC_1

- Procedimento das médias

Procedeu-se à construção de um modelo para a quantificação teor de voláteis após o primeiro banho, VC_1 , utilizando os espectros obtidos pelo procedimento das médias. Estes

corresponderam aos utilizados no conjunto de calibração do modelo do RC_1 , obtidos com o mesmo procedimento, à exceção dos valores anómalos que não foram identificados em ambas as amostragens.

Para a criação deste modelo, não foi efetuada novamente a análise dos *loadings* do PCA, uma vez que, esta seria semelhante à da figura B.6. Assim, sabe-se que a região espectral que compreende a maior variância entre espectros situa-se entre 7300 e 4000 cm^{-1} .

Durante a construção do modelo PLS recorreu-se aos critérios utilizados anteriormente. A validação externa foi efetuada com amostras que não foram utilizadas na calibração, correspondendo a 30 % do número total das amostras com dois banhos. Estas amostras foram selecionadas pelo PCA. Os resultados obtidos para as validações efetuadas encontram-se representados na tabela C.8 do apêndice C.

O modelo considerado mais adequado utilizou não utilizou nenhum pré-processamento e recorreu à ferramenta *General B*, efetuada manualmente, com 20 intervalos para selecionar as regiões [11988;11586,8], [8794,3;8385,5], [7992;7583,1] e [5592,8;5184] cm^{-1} . Para a validação cruzada, representada na figura 4.8, o $RMSE_{CV}$ obtido foi de 0,804 % (m/m), com um R^2 de 0,86 com um número de fatores de 4 e um enviesamento de 0,00311. Por sua vez, para a validação externa, o $RMSE_P$ obtido foi de 0,819 % (m/m), com um R^2 de 0,89 com um número de fatores de 4 e um enviesamento de 0,0714. A região delimitada a verde representa o intervalo de aceitação para o VC_1 .

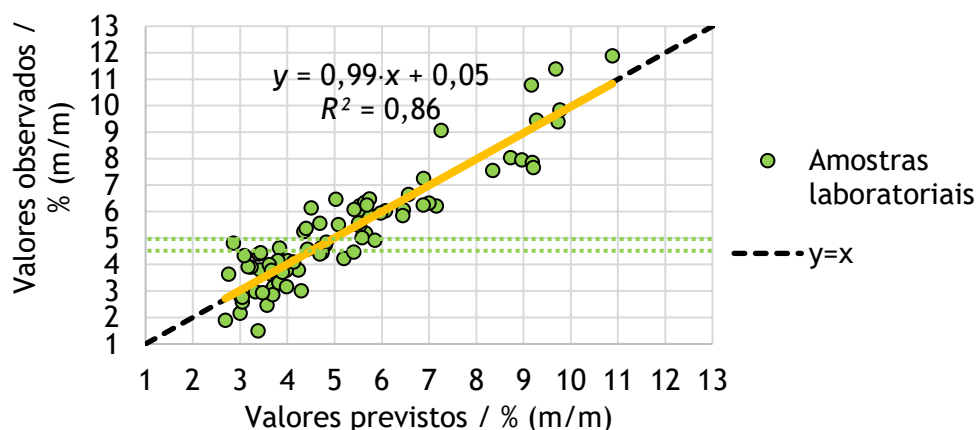


Figura 4.8 - Validação cruzada para a metodologia de quantificação do VC_1 , utilizando o procedimento das médias.

Verifica-se que, as regiões selecionadas contêm picos associados à presença de água, bem como, parte da região prevista pelos *loadings* do PCA. Os valores determinados de $RMSE_{CV}$ e $RMSE_P$ foram novamente significativamente inferiores aos valores obtidos para a quantificação do teor de resina, pelo que, o teor de voláteis foi previsto com maior precisão. Assim, para a validação cruzada, o erro máximo foi de -1,94 % (m/m). O R^2 obtido, embora elevado, foi inferior a 0,95. Relativamente ao número de fatores, este foi considerado adequado, pelo que,

manteve-se inferior a oito [46]. Considera-se que, contrariamente ao modelo desenvolvido para o papel B1, este modelo teve uma capacidade de previsão baixa.

- **Procedimento do pórtico**

Procedeu-se à construção de um modelo para a quantificação teor de voláteis após o primeiro banho, VC_1 , utilizando os espectros obtidos pelo procedimento do pórtico. Estes corresponderam aos utilizados no conjunto de calibração do modelo do RC_1 , obtidos com o mesmo procedimento, à exceção dos valores anómalos que não foram identificados em ambas as amostragens.

Para a criação deste modelo, não foi efetuada novamente a análise dos *loadings* do PCA, uma vez que, esta seria semelhante à da figura B.8. Assim, sabe-se que a região espectral que compreende a maior variância entre espectros situa-se entre 7300 e 4000 cm^{-1} .

O modelo PLS foi construído recorrendo aos critérios utilizados previamente. Para futuros efeitos comparativos, a validação externa foi efetuada com as amostras selecionadas previamente para o modelo do procedimento das médias. Os resultados obtidos para as validações efetuadas encontram-se representados na tabela C.9 do apêndice C.

O modelo considerado mais adequado e utilizou como pré-processamento o MSC e recorreu à ferramenta *General B*, efetuada manualmente, com 20 intervalos para selecionar as regiões [7992;5986,3] e [5592,8;4782,8] cm^{-1} . Para a validação cruzada, representada na figura 4.9, o $RMSE_{CV}$ obtido foi de 0,624 % (m/m), com um R^2 de 0,93 com um número de fatores de 9 e um enviesamento de -0,00155. Por sua vez, para a validação externa, o $RMSE_p$ obtido foi de 0,679 % (m/m), com um R^2 de 0,92 com um número de fatores de 4 e um enviesamento de -0,025. A região delimitada a verde representa o intervalo de aceitação para o VC_1 .

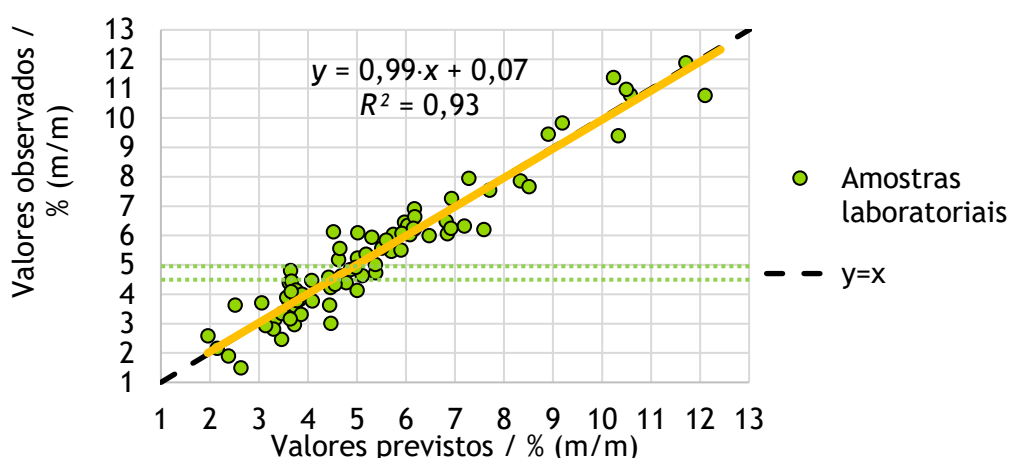


Figura 4.9 - Validação cruzada para a metodologia de quantificação do VC_1 , utilizando o procedimento do pórtico.

Verifica-se que, as regiões selecionadas contêm picos associados à presença de água, bem como, parte da região prevista pelos *loadings* do PCA. Os valores determinados de $RMSE_{CV}$ e

$RMSE_p$ foram novamente significativamente inferiores aos valores obtidos para a quantificação do teor de resina, pelo que, o teor de voláteis foi previsto com maior precisão. Assim, para a validação cruzada, o erro máximo foi de -1,61 % (m/m). O R^2 obtido, embora elevado, foi inferior a 0,95. No caso da validação cruzada, o valor do número de fatores foi superior a 8, porém, considerou-se que não ocorreu *overfitting*. Considera-se que, o procedimento do pórtico permitiu melhorar significativamente a capacidade de previsão do modelo. Contudo, os erros obtidos mantêm-se elevados comparativamente aos obtidos para o papel B1.

- **Comparação entre procedimentos**

De modo a avaliar a capacidade de previsão dos modelos, comparou-se os resultados das validações externas. A comparação entre o desempenho dos modelos encontra-se representada na tabela 4.4.

Tabela 4.4 - Resultados da validação externa para os modelos baseados no procedimento das médias e do pórtico.

Procedimento	RC_1		VC_1	
	Médias	Pórtico	Médias	Pórtico
R^2	0,88	0,97	0,89	0,92
$RMSE_p$ (% (m/m))	1,54	0,765	0,819	0,679
Número de fatores	5	4	4	4
Erro máximo (% (m/m))	3,13	2,1	-1,7	-1,75

Como já foi verificado noutros estudos, o papel impregnado tem uma superfície heterogénea, dificultando a aquisição de espectros representativos da amostra [3] [4]. Conclui-se que o procedimento do pórtico permitiu adquirir espectros mais representativos, melhorando a capacidade de previsão dos modelos criados e, diminuindo os valores de $RMSE_p$ e dos erros máximos. Esta melhoria foi mais notória no caso do RC, o que indica que a sua distribuição no papel é mais heterogénea do que o VC, embora estes estejam relacionados. A distribuição heterogénea da resina à superfície do papel pode resultar em aglomerados resultando na rugosidade do papel impregnado.

4.3 Causas de erros na previsão do VC e RC

- **Utilização exclusiva de papel impregnado laboratorialmente**

Analisando os valores dos erros máximos, $RMSE_{CV}$ e $RMSE_p$ dos modelos baseados no procedimento das médias para o papel L, verifica-se que foram significativamente superiores aos obtidos para os modelos do papel B1. Admite-se que esta discrepância de valores pode estar relacionada com o facto de os modelos para o papel L não recorrerem a amostras industriais.

Na figura 4.10 encontram-se representadas duas rodela de papel L impregnadas laboratorialmente, bem como, uma folha de papel L por impregnar.

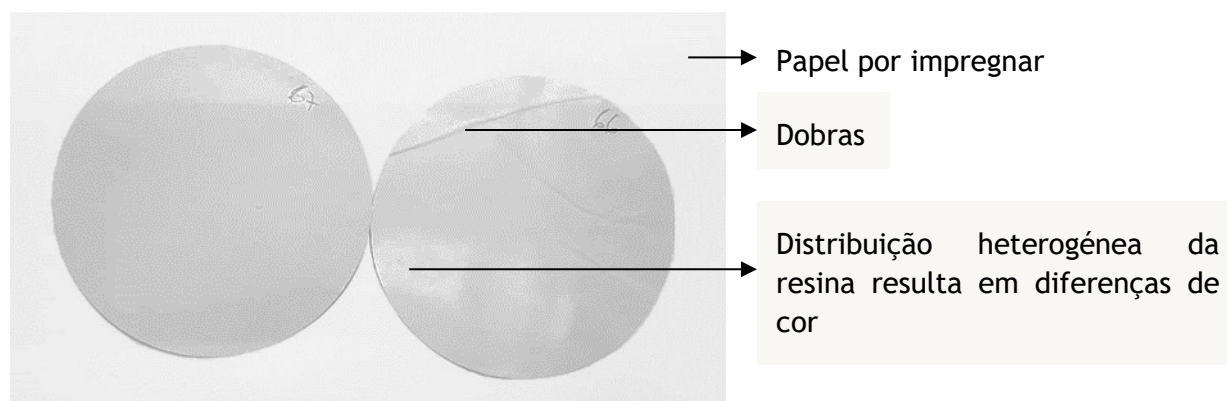


Figura 4.10 - Defeitos associados ao papel impregnado laboratorialmente.

Verifica-se que os papeis impregnados em laboratório apresentam diversos defeitos nomeadamente dobras e concavidades. Após a impregnação, o papel L escurece. Por sua vez, a impregnação manual resulta numa distribuição mais heterogênea da resina. Assim, no caso do papel L, uma impregnação manual resulta não só numa distribuição mais heterogênea de resina como em diferenças de cor na amostra. Esta diferença de cor poderá adicionar erros na previsão do teor de voláteis e resina.

Avaliou-se o efeito da ausência de amostras industriais na construção de modelos. Assim, retiraram-se aos modelos de RC_2 e VC_2 do papel B1 as amostras industriais e, realizou-se novamente a validação cruzada. Optou-se por não comparar com o modelo do RC_1 pois, não houve alteração no tempo de secagem das amostras utilizadas neste, aquando a impregnação. Os resultados obtidos, representados na tabela 4.5, foram comparados mantendo-se em conta que foram obtidos com metodologias e amostras diferentes.

Tabela 4.5 - Comparação entre modelos contruídos apenas com amostras laboratoriais.

Papel	RC		VC	
	L (médias)	B1	L (médias)	B1
R^2	0,91	0,82	0,86	0,95
$RMSE_{CV}$ (% (m/m))	1,47	1,52	0,804	0,417
Número de fatores	7	5	4	5
Erro máximo (% (m/m))	-4,24	3,18	-1,94	1,02

Através da análise da tabela 4.5, conclui-se a utilização exclusiva de amostras laboratoriais justificou na sua maioria os elevados erros obtidos na quantificação do teor de resina. Contudo, este fator não permitiu justificar os erros na quantificação do teor de voláteis. Admite-se que estes podem ser devidos à dificuldade de manter a distância entre a sonda e a amostra para o suporte utilizado na aquisição de espectros da amostra L. Este problema foi causado pelas dobras e concavidades inerentes às amostras laboratoriais. Assim, sugere-se que, em futuros

estudos, seja estudada a influência da distância entre a sonda e a amostra nos espectros obtidos, bem como, uma solução para manter esta distância constante.

Contudo, tendo em conta os erros obtidos para o papel B1 que, embora mais baixos que os do papel L, mantiveram-se superiores ao desejado assume-se que poderá haver outros fatores que estariam a influenciar negativamente a previsão do RC e VC.

- **Contabilização incompleta do teor de voláteis**

Após a impregnação do papel, este pode conter resina reagida e por reagir. Assim, após a secagem na estufa para a determinação do teor de voláteis, poderá estar a ocorrer a condensação de resina por reagir. Este processo resulta na libertação de água, cujo teor influencia na sua maioria o VC. Deste modo, o espectro adquirido não irá contabilizar a água formada durante este processo, resultado em erros de previsão no teor de voláteis. Esta hipótese já foi reconhecida num estudo prévio, mas ainda não foi estudada [3].

- **Contabilização de ligações do papel decorativo**

Outra possível causa de erros de previsão poderá ser a contabilização de ligações do papel. Como já foi referido anteriormente, o papel decorativo é constituído na sua maioria por α -celulose composta por ligações O-H, presentes na água, e C-H, sendo que ambas as ligações estão presentes na resina melamínica. Verificou-se que, os modelos do papel B1 quando aplicados em papéis de maior gramagem, B2 e B3, apresentaram uma capacidade de previsão baixa do VC e RC. E de notar também que, na maioria dos modelos, os valores previstos foram consistentemente superiores aos observados. Assim, admite-se que poderá haver a contabilização de ligações do papel decorativo. Contudo, para ser induzido erro nas previsões, teria de se verificar que as propriedades químicas do papel seriam também heterogêneas. Deste modo, sugere-se que, em futuros estudos, seja estudada a distribuição da composição do papel decorativo por impregnar pela aquisição de espectros ao longo deste, à semelhança da figura 3.1. Espera-se que a variação da composição não seja muito significativa pois, a diferença entre os valores previstos e observados para o papel B3 apresentou um *offset* aproximadamente constante relativamente à reta $y = x$.

- **Incerteza do método de referência**

Não é possível que os modelos PLS criados sejam mais precisos que o método de referência utilizado. Pelo que, admite-se que parte dos erros obtidos poderão ser devidos à incerteza deste. Verificou-se que as massas das amostras após impregnação, bem como, após secagem na estufa eram instáveis. Por sua vez, o RC e VC das amostras industriais encontraram-se frequentemente fora das especificações, o que poderá ter sido devido a erros nas pesagens. Aconselha-se que, para futuros estudos, o tempo entre a pesagem da amostra e aquisição dos espetros seja baixo e que a incerteza do método de referência seja testada e minimizada.

- **Colinearidade entre o teor de resina e voláteis**

Por fim, a última hipótese corresponde à colinearidade entre o teor de resina e voláteis. Na figura 4.11 encontram-se esquematizados exemplos de amostras colineares e não colineares.

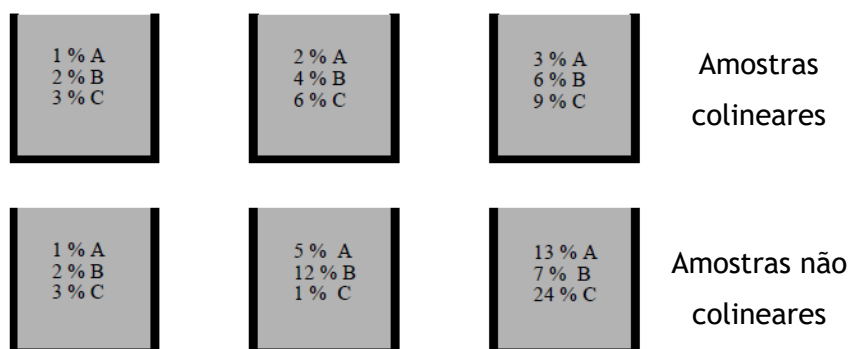


Figura 4.11 - Colinearidade e não colinearidade entre amostras [35].

De modo a avaliar a colinearidade das amostras de papel B1 e L utilizadas, recorreu-se a uma representação gráfica do VC em função do RC. Os resultados obtidos para as amostras do papel B1 e L encontram-se representados na figura 4.12.

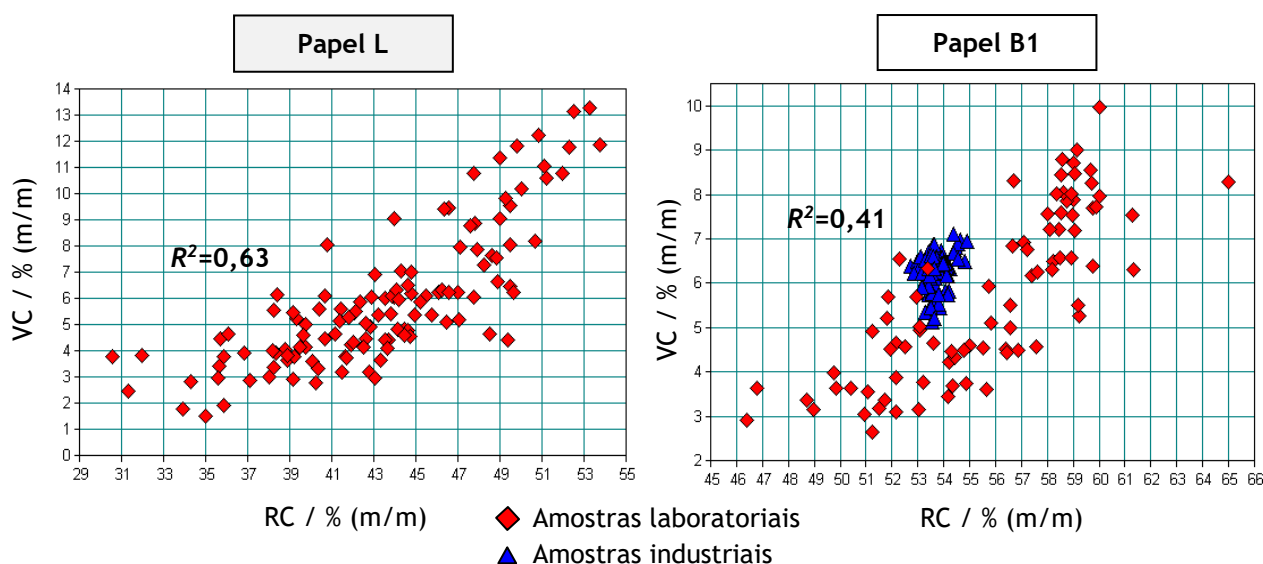


Figura 4.12 - Avaliação da colinearidade entre o VC e o RC.

Analisando os resultados obtidos, verifica-se que existe colinearidade entre o teor de resina e o teor de voláteis. Contudo, estes parâmetros resultam do mesmo processo de impregnação, por isso, haverá sempre alguma forma de colinearidade entre estes. O software OPUS considera que a colinearidade entre variáveis é elevada quando o R^2 é superior a 0,7 [35]. No caso do papel B1, o valor obtido é significativamente mais baixo, pelo que, se considera que a colinearidade foi baixa. No entanto, verifica-se que, no caso do papel L, o valor de R^2 é próximo do estipulado. Assim, conclui-se que a colinearidade entre amostras do papel L foi mais acentuada o que poderá ter contribuído para os erros de previsão mais elevados neste modelo. Para futuros estudos, segure-se que o tempo de secagem seja definido de uma forma aleatória de modo a evitar uma colinearidade alta entre as propriedades das amostras.

5 Conclusões

Este estudo visa a implementação da espectroscopia de infravermelho próximo na fábrica da EuroResinas. No processo de produção do papel impregnado, os dois parâmetros importantes a controlar são o teor de resina (RC) e o teor de voláteis (VC). De momento, este controlo é efetuado de modo *off-line* e requer entre cinco e dez minutos para obtenção de resultados. A implementação da espectroscopia de infravermelho próximo (NIR) na linha de produção permite acelerar e automatizar esta determinação, reduzindo o desperdício gerado no tempo de espera.

Existem, de momento, modelos PLS previamente desenvolvidos para o papel impregnado B1 com um tempo de análise de aproximadamente 1 minuto e 40 segundos. Considera-se que estes modelos precisam de ser otimizados, diminuindo o tempo de aquisição de espectros, bem como, aumentando o valor de R^2 e diminuindo os valores de $RMSE_{CV}$ e $RMSE_P$. Pretende-se também aumentar a gama de propriedades dos modelos desenvolvidos.

Previamente à construção de modelos, estudou-se o efeito dos parâmetros escolhidos na aquisição dos espectros de modo a reduzir, sem prejudicar os resultados obtidos, o tempo de análise. Considerou-se que, idealmente, este tempo deveria ser inferior a um minuto. Concluiu-se que, das combinações estudadas, a que permitiu chegar a este objetivo foi: 16 cm^{-1} de resolução, 128 varrimentos e 20 kHz de frequência de varrimento, correspondendo a um tempo de aquisição de 22 segundos.

Procedeu-se à construção de modelos para a quantificação do RC e VC para o papel B1. O modelo construído para a quantificação do RC após o primeiro banho, RC_1 , obteve um valor de $RMSE_{CV}$ de 0,929 % (m/m) com um R^2 de 0,90 para um número de fatores de 5. No caso do modelo construído para o RC após o segundo banho, RC_2 , o $RMSE_{CV}$ obtido foi de 1,01 % (m/m) com um R^2 de 0,82 para um número de fatores de 6. Optou-se por criar um modelo geral para a quantificação do teor de resina, RC . Para este, o $RMSE_{CV}$ obtido foi de 1,05 % (m/m), com um R^2 de 0,98 para um número de fatores de 6. Foi também criado um modelo para a quantificação do VC após o segundo banho, VC_2 . Assim, o $RMSE_{CV}$ obtido foi de 0,386 % (m/m), com um R^2 de 0,91 com um número de fatores de 5. Criou-se novamente um modelo geral, porém, para a quantificação do teor de voláteis, VC . Para este, o $RMSE_{CV}$ obtido foi de 0,4 % (m/m), com um R^2 de 0,91 para um número de fatores de 6.

Verificou-se que o desempenho do modelo geral do RC foi mais adequado para amostras com dois banhos. No entanto, a junção dos modelos específicos introduziu um erro significativo na previsão do RC das amostras com um banho. Relativamente ao modelo geral para a previsão do VC, verificou-se que a capacidade de previsão das amostras com dois banhos não foi significativamente alterada. Contudo, as amostras industriais com um banho foram previstas de forma anómala.

Conclui-se que, apesar da redução do tempo de análise, houve uma melhoria significativa do R^2 obtido para os modelos específicos de quantificação do RC, relativamente aos modelos previamente desenvolvidos. Foi também possível aumentar a gama de análise dos modelos. Relativamente aos valores de $RMSE_{CV}$, verificou-se uma melhoria significativa para o modelo do RC após um banho. Contudo, considerou-se que o único modelo com capacidade de previsão foi o do VC_2 , embora seja ainda inferior ao desejado pela empresa. Admite-se que, para serem obtidos valores mais satisfatórios, teria de ser utilizado o procedimento do pórtico para aquisição dos espectros, em vez do das médias. Deste modo, os espectros obtidos seriam mais representativos da superfície heterogénea do papel impregnado. Contudo, não foi possível no decorrer deste estudo utilizar o pórtico para a análise deste papel impregnado.

Num estudo adicional aplicaram-se os modelos desenvolvidos a papéis de diferentes gramagens, B2 e B3, e cores, L. Concluiu-se que o parâmetro medido com maior precisão foi o teor de voláteis (VC), como se tinha verificado anteriormente. Contudo, verificou-se que, na maioria dos modelos, os valores previstos foram consistentemente superiores aos observados para os papéis de maior gramagem. Deste modo, admite-se que, neste caso, podem ter sido contabilizadas ligações do papel do tipo O-H, presentes na água, e C-H, sendo que ambas as ligações estão presentes na resina.

Estudou-se de seguida um segundo papel impregnado com apenas um banho, L. Para este foram estudados os dois procedimentos de aquisição de espectros: o das médias e o do pórtico. O modelo construído para a quantificação do RC após o primeiro banho, RC_1 , utilizando o procedimento das médias, obteve um $RMSE_{CV}$ de 1,47 % (m/m), com um R^2 de 0,91 para um número de fatores de 7. Por sua vez, para o procedimento do pórtico, o $RMSE_{CV}$ obtido foi de 1,06 % (m/m), com um R^2 de 0,96 para um número de fatores de 4. No caso do VC, o modelo obtido utilizando o procedimento da médias obteve um $RMSE_{CV}$ de 0,804 % (m/m), com um R^2 de 0,86 para um número de fatores de 4. Utilizando o procedimento do pórtico, o $RMSE_{CV}$ obtido foi de 0,624 % (m/m), com um R^2 de 0,93 com um *número de fatores* de 9.

Conclui-se que, tal como sugerido por estudos prévios, o procedimento do pórtico permitiu adquirir espectros mais representativos, melhorando a capacidade de previsão dos modelos criados e, diminuindo significativamente os valores de $RMSE_{CV}$. Esta melhoria foi mais notória no caso do RC, o que indica que a sua distribuição no papel é mais heterogénea do que o VC, embora estes estejam relacionados. Contudo, a capacidade de previsão dos modelos manteve-se inferior ao desejado.

Apesar de não ter sido possível implementar os modelos desenvolvidos, considera-se que os resultados obtidos permitiram aprofundar o conhecimento na área da espectroscopia de NIR. Espera-se que este estudo permita apoiar trabalhos futuros do mesmo tipo, de modo a proceder à implementação futura de métodos deste tipo em ambiente fabril.

6 Avaliação do trabalho realizado

6.1 Objetivos Realizados

O primeiro objetivo deste estudo foi a redução do tempo total de aquisição dos espectros através da otimização dos parâmetros para este efeito, sem comprometer a qualidade dos resultados. O tempo otimizado foi de 22 segundos, sendo que, o previamente estipulado era de 1 minuto e 40 segundos. Assim, considera-se que o objetivo foi realizado. Contudo, aconselha-se que estas condições sejam novamente testadas utilizando o pórtico.

Após a otimização dos parâmetros de aquisição dos espectros, era proposto proceder à otimização de modelos previamente desenvolvidos para o papel B1. Conclui-se que foi possível aumentar significativamente a gama de análise destes modelos, bem como, os valores de R^2 . Devido ao facto de não ter sido possível utilizar o pórtico para realizar a aquisição dos espectros ao longo do papel, não houve uma melhoria significativa nos valores de $RMSE_{CV}$ e $RMSE_P$ obtidos, à exceção do modelo para o RC_1 . Considera-se que a capacidade de previsão destes modelos se manteve inferior ao desejado. Assim, não foi possível proceder à sua implementação em ambiente fabril.

O último objetivo consistia no desenvolvimento de novos modelos para a quantificação do RC e VC do papel L. Este papel não tinha sido estudado previamente. Foi possível desenvolver métodos para este papel, utilizando o procedimento das médias e do pórtico. Apesar da capacidade de previsão dos modelos se ter mantido baixa, considera-se que foram tiradas conclusões úteis à cerca da utilização destes procedimentos.

6.2 Outros Trabalhos Realizados

No decorrer deste estudo foram construídos modelos gerais para quantificação do teor de resina e voláteis. Isto é, modelos para a quantificação destas propriedades independentemente do número de banhos as amostras de papel B1 foram submetidas. Contudo, tal como para os modelos específicos, os valores de $RMSE_{CV}$ e $RMSE_P$ foram elevados. Espera-se que a capacidade de previsão destes modelos venha a ser melhorada em futuros estudos.

Foi também avaliada a capacidade de previsão dos modelos do papel B1 em papeis de diferentes cores e gramagens. Os modelos apresentaram uma baixa capacidade de previsão para os papeis de maior gramagem. No entanto, para o papel de cor azul estudado os resultados obtidos para a previsão do VC foram promissores.

6.3 Apreciação Final

Apesar das limitações encontradas no decorrer deste estudo, conclui-se que os resultados obtidos permitiram aprofundar o conhecimento à cerca da quantificação do teor de resina e voláteis em papel impregnado por espectroscopia de NIR. Espera-se que os resultados obtidos neste estudo permitam apoiar trabalhos futuros do mesmo tipo, de modo a proceder à implementação futura de métodos deste tipo em ambiente fabril.

7 Referências

- [1] Sonae Arauco, “Quem Somos”, 2020. Disponível em: https://www.sonaearauco.com/pt/empresa/quem-somos_2455.html. [Acedido: 2 de março de 2020]
- [2] R. Alvim, “EuroResinas - O ambiente na ZILS”, 2016. Disponível em: https://www.globalparques.pt/folder/galeria/ficheiro/58_Apresenta%C3%A7%C3%A3o%20Euroresinas%20-%20O%20Ambiente%20na%20ZILS%20-%202016.11.08%20VF_hyutr6b4sm.pdf. [Acedido: 2 de março de 2020]
- [3] M. Ribeiro, “Determinação do Teor de Resina e Teor de Voláteis em papel impregnado usando espectroscopia de infravermelho próximo”, Dissertação de Mestrado Integrado em Engenharia Química, Faculdade de Engenharia da Universidade do Porto, 2019.
- [4] M. Gonçalves, M. Ribeiro, N. Paiva, J. Ferra, J. Martins, F. Magalhães, L. Carvalho, “Resin and volatile content of melamine-impregnated paper assessed by near infrared spectroscopy, a simulation of the industrial process using a laboratory-scale gantry”.
- [5] S. Goodman, “Handbook of Thermoset Plastics”, Noyes Publications, 2.^a ed., New York, 2000.
- [6] N. Paiva, “Development of resins with very low formaldehyde emissions”, Programa Doutoral em Engenharia Química e Biológica, Faculdade de Engenharia da Universidade do Porto, 2015.
- [7] L. Williams, “Amino Resins”, Encyclopedia of Polymer Science and Technology, vol. 1, pp. 340-371, John Wiley and Sons, New Jersey, 2002.
- [8] J. Ferra, “Optimization of Ureia-Formaldehyde resins for the manufacture of wood-based panels”, Programa Doutoral em Engenharia Química e Biológica, Faculdade de Engenharia da Universidade do Porto, 2010.
- [9] A. Henriques, “Development of new resins and characterization techniques for laminates”, Programa Doutoral em Engenharia Química e Biológica, Faculdade de Engenharia da Universidade do Porto, 2018.
- [10] M. Ansell, “Wood composites”, Woodhead Publishing Series, in *Composites Science and Engineering*, n. 54, Elsevier, Amsterdam, 2015.
- [11] N. Ayrilmis, “Enhancement of dimensional stability and mechanical properties of light MDF by adding melamine resin impregnated paper waste”, *International Journal of Adhesion and Adhesives*, vol. 33, pp. 45-49, 2012.
- [12] A. Pizzi, “Melamine-Formaldehyde Adhesives”, in *Handbook of Adhesive Technology*, 2.^a ed., CRC Press, Boca Raton, Florida, 2003.
- [13] D. Merline, S. Vukusic, A. Abdala, “Melamine formaldehyde: curing studies and reaction mechanism”, *Polymer Journal*, vol. 45, pp. 413-419, 2013.

- [14] W. Binder, “Melamine-Formaldehyde Resins”, in *Encyclopedia of Polymer Science and Technology*, vol.10, pp. 369-385, John Wiley & Sons, New Jersey, 2004.
- [15] EN 14322: 2017, “Wood-based panels – Melamine faced board for interior uses – Definition, requirements and classification”, CEN.
- [16] EN 438-1: 2016, “High-pressure decorative laminates (HPL) – Sheets based on thermosetting resins (usually called laminates) – Part 1: Introduction and general information”, CEN.
- [17] P. Allen, M. Kemp, “Decorative Laminates”, in *Plastics Surface and Finish*, 2^a ed., pp. 113-136, Royal Society of Chemistry, Cambridge, 1995.
- [18] A. Henriques, C. Coelho, J. Ferra, J. Martins, F. Magalhães, L. Carvalho, “Chapter 1 - Introduction of advanced functionalities in laminates for wood-based panels: surface quality evaluation”, in J. Paulo Davim, Alfredo Aguilera (Eds.), *Wood Composites: Materials Manufacturing and Engineering*, Walter de Gruyter, Berlin, 2017 (ISBN 978-3-11-041607-7).
- [19] M. Ma, “Method of impregnating decorative paper with melamine resin”, European Patent office, EP 0 819 794 A2, 1998.
- [20] A. Hatam, M. Firouzabadi, H. Resalati, "Surface chemistry of gravure printed décor paper and adhesion of melamine formaldehyde resin coatings," *BioResources*, vol. 12, pp. 4239-4258, 2017.
- [21] M. Thébault, A. Kandelbauer, U. Müller, E. Zikulnig-Rusch, H. Lammer, “Factors influencing the processing and technological properties of laminates based on phenolic resin impregnated papers”, *Eur. J. Wood Prod*, vol. 75, pp. 785-806, 2017.
- [22] S. Bardak, B. Sarı, G. Nemli, H. Kırıcı, M. Baharoglu, “The effect of decor paper properties and adhesive type on some properties of particleboard”, *International Journal of Adhesion and Adhesives*, vol 31, pp. 412-415, 2011.
- [23] D. Burns, E. Ciurczak, “Handbook of Near-Infrared Analysis”, Practical Spectroscopy Series, CRC Press, vol. 35, 2007.
- [24] L. Carvalho, J. Ferra, M. Martins, A. Henriques, “Near Infrared Spectroscopy in the wood-bases products industry”, vol. 2, pp.1-20.
- [25] G. Reich, “Near-infrared spectroscopy and imaging: Basic principles and pharmaceutical applications”, *Advanced Drug Delivery Reviews*, vol. 57, pp. 1109-1143, 2005.
- [26] I. Marques, “Desenvolvimento de uma metodologia on-line para o controlo de síntese de resinas ureia formaldeído usando FT-NIR”, Tese de Mestrado Integrado em Engenharia Química, Faculdade de Engenharia da Universidade do Porto, 2014.
- [27] H. Cen, Y. He, “Theory and application of near infrared reflectance spectroscopy in determination of food quality”, *Trends in Food Science and Technology*, vol. 18, pp. 72-83, 2007.

- [28] D. Skoog, F. Crouch, J. Holler, D. West, “Fundamentos de Química Analítica”, Thomson Learning, Stamford, Connecticut, 2006.
- [29] F. Almeida, “Espectroscopia de Infravermelho Próximo com Transformada de Fourier (FT-NIR) na Caracterização de Farinhas para Alimentação Pueril”, Dissertação para obtenção do grau mestre em Engenharia Biológica, Universidade Técnica de Lisboa, 2009.
- [30] A. Rinnan, F. Berg, S. Engelsen, “Review of the most common pre-processing techniques for near-infrared spectra”, *Trends in Analytical Chemistry*, vol. 28, n. 10, pp. 1201-1222, 2009.
- [31] A. Candolfi, R. Maesschalck, D. Jouan-Rimbaud, P. Hailey, D. Massart, “The influence of data pre-processing in the pattern recognition of excipients near-infrared spectra”, *Journal of Pharmaceutical and Biomedical Analysis*, vol. 21, pp. 115-132, 1999.
- [32] M. Maleki, A. Mouazen, H. Ramon, J. Baerdemaeker, “Multiplicative Scatter Correction during On-line Measurement with Near Infrared Spectroscopy”, *Biosystems Engineering*, pp. 427-433, 2007.
- [33] A. Henriques, M. Gonçalves, N. Paiva, J. Ferra, J. Martins, F. Magalhães, L. Carvalho, “Determination of resin and moisture content in melamine formaldehyde paper using near infrared spectroscopy”, *Journal of Near Infrared Spectroscopy*, vol. 25, pp. 311-323, 2017.
- [34] M. Zeater, D. Rutledge, “Preprocessing Methods”, in *Comprehensive Chemometrics*, Elsevier, pp. 121-231, 2009.
- [35] “Manual III”, OPUS Spectroscopy Software, 2006.
- [36] F. Kherif, A. Latypova, “Principal component analysis”, em *Machine Learning*, pp. 209-225, 2020.
- [37] H. Thoemen, M. Irle, M. Sernek, “Wood-Based Panels – An Introduction for Specialists”, COST Office, pp. 236-244, 2010.
- [38] M. Mansouri, M. Harkat, H. Nounou e M. Nounou, “PCA and PLS-based generalized likelihood ratio for fault detection”, in *Data-Driven and Model-Based Methods for Fault Detection and Diagnosis*, pp. 12-45, Elsevier, Amsterdam, 2020.
- [39] Q. Yuan, B. Lennox, “Improved Model Predictive Control Using PCA”, in *Fault Detection, Supervision and Safety of Technical Processes*, pp. 102-107, Elsevier Science, Amsterdam, 2006.
- [40] S. Mudge, “Statistical analysis of oil spill chemical composition data”, in *Standard Handbook Oil Spill Environmental Forensics - Fingerprinting and Source Identification*, 2.^a ed., pp. 849-867, Academic Press, Cambridge, Massachusetts, 2016.
- [41] S. Wold, M. Sjostrom, L. Eriksson, “PLS-regression: a basic tool of chemometrics” *Chemometrics and Intelligent Laboratory Systems*, vol. 58, pp. 109-130, 2001.

- [42] Z. Xiaobo, Z. Jiewen, L. Yanxiao, "Selection of the efficient wavelength regions in FT-NIR spectroscopy for determination of SSC of 'Fuji' apple based on BiPLS and FiPLS models", *Vibrational Spectroscopy*, vol. 44, pp. 220-227, 2007.
- [43] P. Santos, "Aplicação da espectroscopia de infravermelho próximo na avaliação de alguns atributos de qualidade em pêra 'rocha' e maçã 'gala'", Dissertação para obtenção do Grau de Mestre em Engenharia Agronómica - Hortofruticultura e Viticultura, Instituto Superior de Agronomia, 2009.
- [44] M. Gonçalves, "Desenvolvimento de uma resina de impregnação que permita a produção de um termolaminado anti-fingerprint", Tese de Mestrado Integrado em Engenharia Química, Faculdade de Engenharia da Universidade do Porto, 2016.
- [45] P. Nakos, E. Minopoulou, "Novel Process Control for the Resin and Panel Industries Based on FT-NIR", em *5th European Wood-Based Panel Symposium*, p. 1-6, 2008.
- [46] P. Williams, P. Dardenne, Peter Flinn, "Tutorial: Items to be included in a report on a near infrared spectroscopy project", *Journal of Near Infrared Spectroscopy*, vol.25, n. 2, pp 85-90, 2017.
- [47] A. Henriques, P. Cruz, J. Ferra, J. Martins, F. Magalhães, L. Carvalho, "Determination of melamine content in amino resins by near-infrared spectroscopy Article", *Wood Science and Technology*, vol. 47, no. 5, 2013.
- [48] "Manual I", OPUS Spectroscopy Software, 2006.
- [49] G. Corder, D. Foreman, "Nonparametric Statistics for Non-Statisticians", *John Wiley and Sons*, New Jersey, 2009.
- [50] R. Pallas-Areny, J. Webster, "Sensors and Signal Conditioning", *John Wiley and Sons*, 2.^a ed., New Jersey, 2000.

Anexo A - Algoritmo NIPALS

Primeiramente, um vetor u é definido como sendo uma das colunas da matriz Y . Se houver apenas uma variável na matriz Y , então $u = Y$. De seguida, o vetor dos pesos, w , é determinado pela equação A.1 [41].

$$w = \frac{X^T \cdot u}{u^T \cdot u} \quad (\text{A.1})$$

Neste passo, w pode ser normalizado de modo a $\|w\| = 1,0$.

De seguida, são determinados os *scores* de X , t , pela equação A.2.

$$t = X \cdot w \quad (\text{A.2})$$

No passo seguinte, o vetor dos pesos de Y , c , é determinado recorrendo à equação A.3.

$$c = \frac{Y^T \cdot t}{t^T \cdot t} \quad (\text{A.3})$$

Por fim, os *scores* da matriz Y , u , são atualizados pela equação A.4.

$$u = \frac{Y \cdot c}{c^T \cdot c} \quad (\text{A.4})$$

Se houver mais do que uma variável na matriz Y , a convergência é testada em t . Se a condição da equação A.5 for verificada o processo segue, caso contrário, voltará à equação A.1.

$$\frac{\|t_{anterior} - t_{novo}\|}{\|t_{novo}\|} < 10^{-6} \quad (\text{A.5})$$

De seguida, os componentes determinados são retirados às matrizes X e Y e, as novas matrizes são utilizadas para o componente seguinte. Este processo é dado pelas equações A.6, A.7 e A.8.

$$p = \frac{X^T \cdot t}{t^T \cdot t} \quad (\text{A.6})$$

$$X = X - t \cdot p^T \quad (\text{A.7})$$

$$Y = Y - t \cdot c^T \quad (\text{A.8})$$

Todos os passos são repetidos para os restantes componentes até a validação cruzada indicar que não há mais informação relevante em X sobre Y .

Os coeficientes da regressão, B , são determinados pela equação A.9 [41].

$$B = W \cdot C^T \quad (\text{A.9})$$

Em que W e C são as matrizes dos pesos de X e Y do método PLS respetivamente.

Apêndice A - Escolha dos parâmetros

Para proceder à avaliação do efeito de cada parâmetro mantiveram-se os restantes parâmetros inalterados. Consideraram-se os parâmetros base iguais aos utilizados noutros estudos para a construção de modelos do mesmo tipo. Deste modo, as condições base utilizadas foram: 16 cm^{-1} de resolução, 256 varrimentos e 10 kHz de frequência de varrimento, correspondendo a um tempo de análise de aproximadamente 1 minuto e 40 segundos [3].

A.1 Resolução

A razão sinal-ruído é proporcional à raiz quadrada do tempo de aquisição. Assim, melhorar a resolução num fator de dois resulta num aumento de quatro vezes no tempo de aquisição para uma razão contante de sinal-ruído. Logo, a escolha de uma resolução mais alta do que o necessário deve ser evitada de modo a minimizar o tempo de aquisição [48].

Para proceder à avaliação do efeito da resolução nos espectros testaram-se os seguintes valores de resolução: 8, 16, 32, 64 cm^{-1} . Os espectros obtidos encontram-se representados na figura A.1.

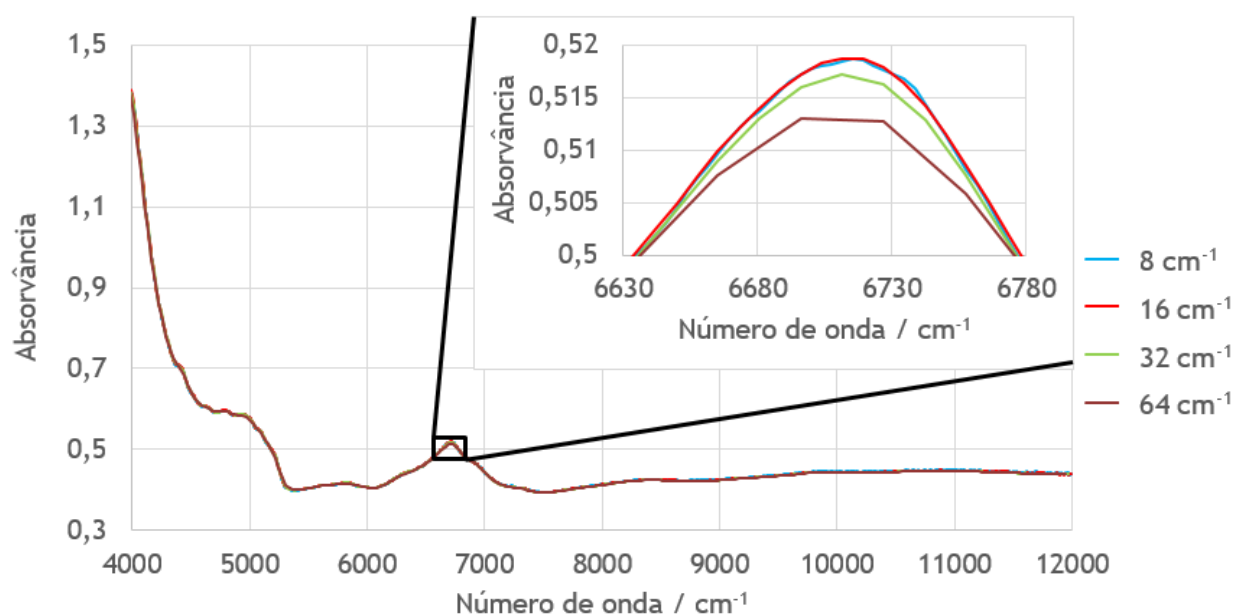


Figura A.1 - Demonstração do efeito da resolução nos espectros obtidos.

Visualmente, verificou-se que as resoluções de 64 e 32 cm^{-1} foram mais desadequadas para descrever as diversas sobreposições de bandas na região NIR. Se fossem realizadas análises de PCA com estes parâmetros, iria-se verificar uma maior dificuldade em diferenciar espectros, devido à perda de informação. Do mesmo modo, a criação de modelos por PLS também seria

afetada. Relativamente às resoluções de 8 e 16 cm^{-1} , estas conferiram maior reprodutibilidade dos espectros.

De modo a encontrar um critério de avaliação dos espectros semelhantes, optou-se por realizar uma análise da variância média ponto a ponto dos espectros. Mediram-se dez espectros na mesma posição da mesma amostra de papel impregnado industrialmente e analisou-se a variância ponto a ponto entre espectros. Os resultados obtidos encontram-se representados na figura A.2.

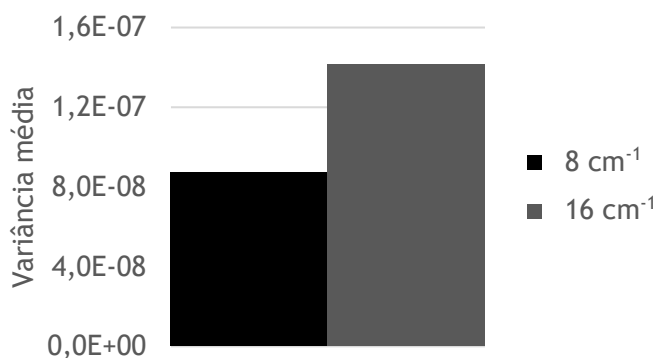


Figura A.2 - Variância média ponto a ponto para os espectros obtidos com as resoluções de 8 e 16 cm^{-1} .

Através da análise da figura A.2, conclui-se que, para a resolução mais baixa (16 cm^{-1}), ocorreu um aumento da variância média entre espectros. Conclui-se que as medições com a resolução de 16 cm^{-1} apresentam uma razão de sinal-ruído menor.

Contudo, considera-se que o tempo de análise utilizando a resolução de 8 cm^{-1} é excessivo. Assim, conclui-se que a resolução de 16 cm^{-1} , utilizada noutros estudos, é a opção mais adequada.

A.2 Número de varrimentos

O número de varrimentos consiste no número de leituras efetuadas para construir o espectro final da amostra. Um número adequado deve ser escolhido de modo a aumentar a razão sinal-ruído no espectro final. Contudo, o número de varrimentos escolhido deve ter também em conta o tempo total da análise. O número mínimo de varrimentos foi de 32 por ser próximo do mínimo considerado necessário para descrever uma distribuição normal de resultados [49].

Visualmente, verificou-se uma grande semelhança entre os espectros obtidos para cada número de varrimentos escolhido. Deste modo, optou-se por realizar a análise da variância média ponto a ponto dos espectros, referida anteriormente. Os resultados obtidos encontram-se representados na figura A.3.

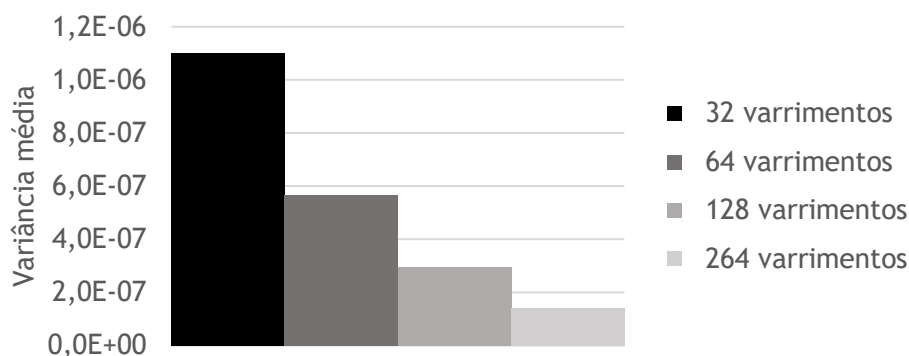


Figura A.3 - Variância média ponto a ponto para os espectros obtidos com o número de varrimentos de 32, 64, 128 e 256.

Através da análise da figura A.3, conclui-se que, com o aumento do número de varrimentos, ocorre uma diminuição acentuada da variância média entre espectros. Este comportamento é devido à razão sinal-ruído ser proporcional à raiz quadrada do tempo de aquisição, que aumenta com o número de varrimentos. Contudo, há que referir que, uma análise com 32 varrimentos tem uma duração de 13 segundos, enquanto que uma de 256 varrimentos dura aproximadamente 1 min e 40 segundos. Assim, considera-se que 128 varrimentos, correspondendo a 50 segundos, corresponde ao melhor compromisso entre a variância média e o tempo de análise.

A.3 Frequência de varrimento

O parâmetro seguinte a ser testado é denominado pelo *software* OPUS 6.5 por velocidade do *scanner*, expressa em kHz. Contudo, será referido por frequência de varrimento. Este parâmetro está relacionado com o tempo de integração que corresponde ao intervalo de tempo no qual os sensores recolhem carga. No final deste tempo, esta é enviada para o terminal de saída onde é convertida para voltagem ou corrente [50]. Assim, o aumento da frequência de varrimento resulta na diminuição do tempo de integração, reduzindo o tempo de análise.

O *software* não permite introduzir um número para este parâmetro, pelo que, foram testadas três opções existentes. Para os modelos já desenvolvidos utilizou-se uma frequência de varrimento de 10 kHz, recomendada pelo manual do *software* para o tipo de detetor usado. Assim, para testar o efeito deste parâmetro testaram-se frequências de varrimento de 5, 10 e 20 kHz.

Tal como no parâmetro anterior, verificou-se visualmente uma grande semelhança entre os espectros obtidos para cada número de varrimentos escolhido. Deste modo, optou-se por realizar a mesma análise da variância média ponto a ponto dos espectros. Os resultados obtidos encontram-se representados na figura A.4.

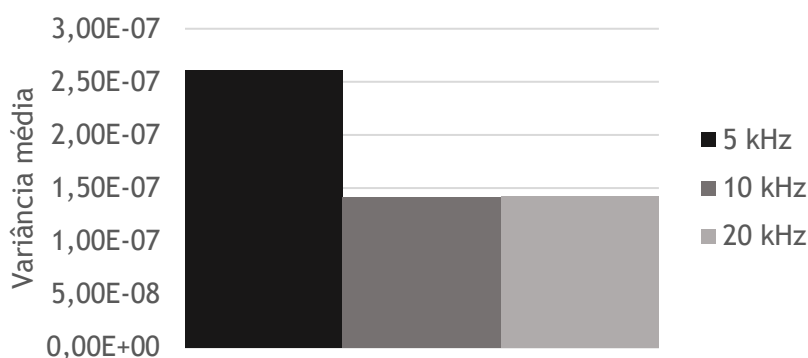


Figura A.4 - Variância média ponto a ponto para os espectros obtidos com uma frequência de varrimento de 5, 10 e 20 kHz.

Através da análise da figura A.4, conclui-se que, a frequência de varrimento de 5 kHz, causou uma maior variância entre espectros. Este comportamento pode ter sido causado pelo tempo excessivo da análise (3 minutos e 20 segundos) ter resultado na saturação do detetor. Outra hipótese poderá ser que esta frequência de varrimento é mais vulnerável a fatores externos, como a luz do próprio laboratório.

Verificou-se também que as frequências de varrimento de 10 e 20 kHz tiveram comportamentos semelhantes. Contudo, é de notar que, o tempo de análise utilizando estes parâmetros foi de 1 minuto e 40 segundos e 45 segundos, respetivamente. Assim, considera-se que uma frequência de varrimento de 20 kHz poderá ser a opção mais adequada.

A.4 Otimização dos parâmetros

Após analisados os parâmetros escolhidos, procedeu-se à sua otimização. Escolheram-se os com o melhor desempenho relativamente à razão sinal-ruído e ao tempo de aquisição dos espectros. As combinações a testar encontram-se descritas na tabela A.1.

Tabela A.1 - Combinações a testar para otimizar os parâmetros de leitura de espectros.

Resolução (cm^{-1})	Número de varrimentos	Frequência de varrimento (kHz)	Tempo (s)
16	256	10	90
		20	45
	128		22

Realizou-se novamente uma análise da variância média ponto a ponto, seguindo o procedimento previamente referido. Os resultados obtidos encontram-se representados na figura A.5.

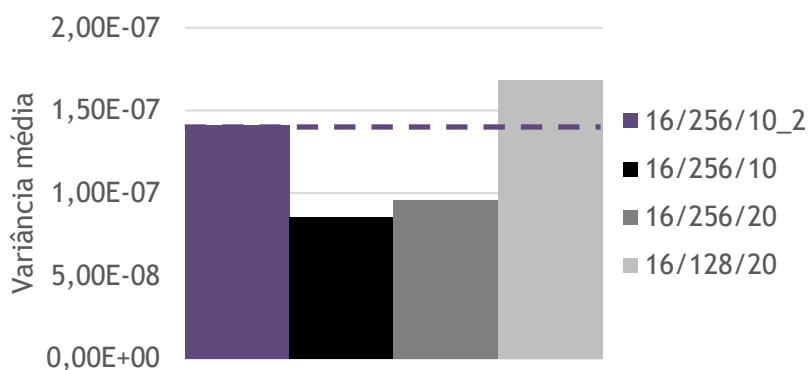


Figura A.5 - Variância média ponto a ponto para os espectros obtidos com as combinações de parâmetros ótimos.

Para facilitar a interpretação da figura A.5, representou-se a mesma análise efetuada em dois dias diferentes para os parâmetros de referência (16 cm^{-1} , 256 varrimentos e 10 kHz). A diferença de variâncias entre estes dois casos foi utilizada como critério para a análise dos novos parâmetros. Os resultados correspondentes às barras a tons de cinzento foram realizados no mesmo dia e, os da barra a roxo num dia diferente. Verifica-se que a diferença da variância obtida com os parâmetros testados e os de referência é desprezável pois, é semelhante à diferença observada em dias diferentes.

Conclui-se que o conjunto de parâmetros mais adequado correspondeu a 16 cm^{-1} de resolução, 128 varrimentos e 20 kHz de frequência de varrimento, com um tempo de análise de aproximadamente 22 segundos. Este conjunto permitiu otimizar o tempo de aquisição, mantendo-o inferior a um minuto, sem comprometer a qualidade dos resultados.

No decorrer deste estudo não foi possível estudar estas combinações ao longo do papel impregnado, utilizando o pórtico. Contudo, em futuros estudos, aconselha-se que as combinações referidas sejam testadas novamente juntamente com a combinação de 16 cm^{-1} de resolução, 128 varrimentos e 10 kHz de frequência de varrimento.

Apêndice B - Espectros e *Loadings*

B.1 Papel B1

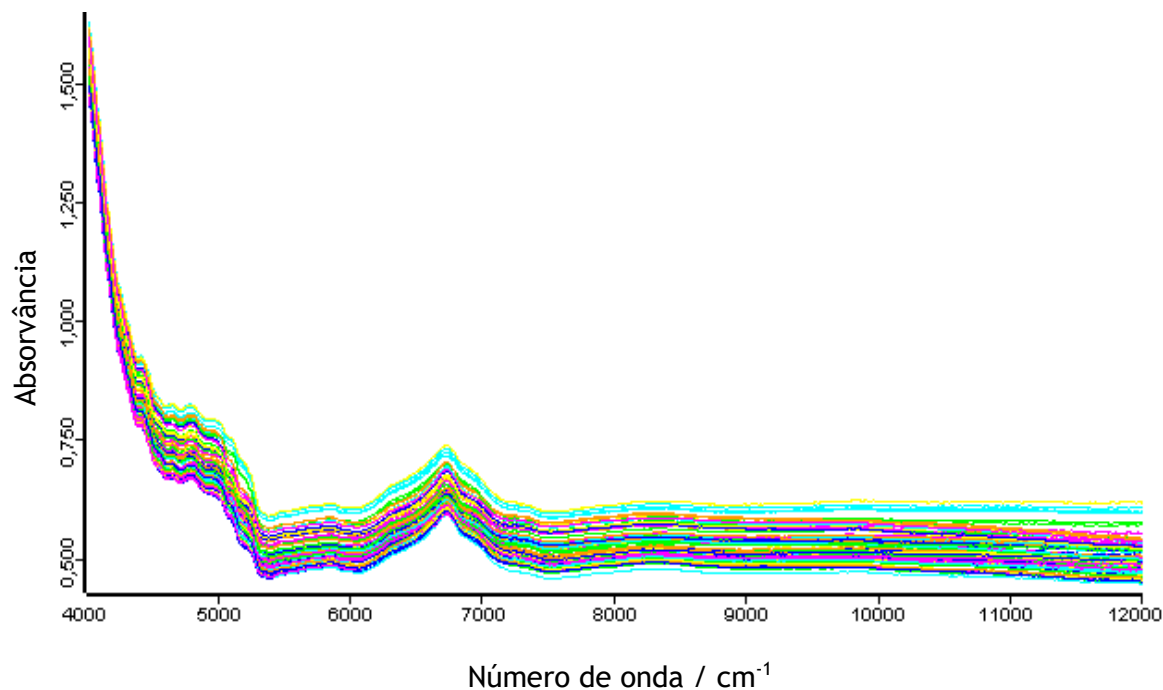


Figura B.1 - Espectros utilizados no modelo de quantificação do RC_1 .

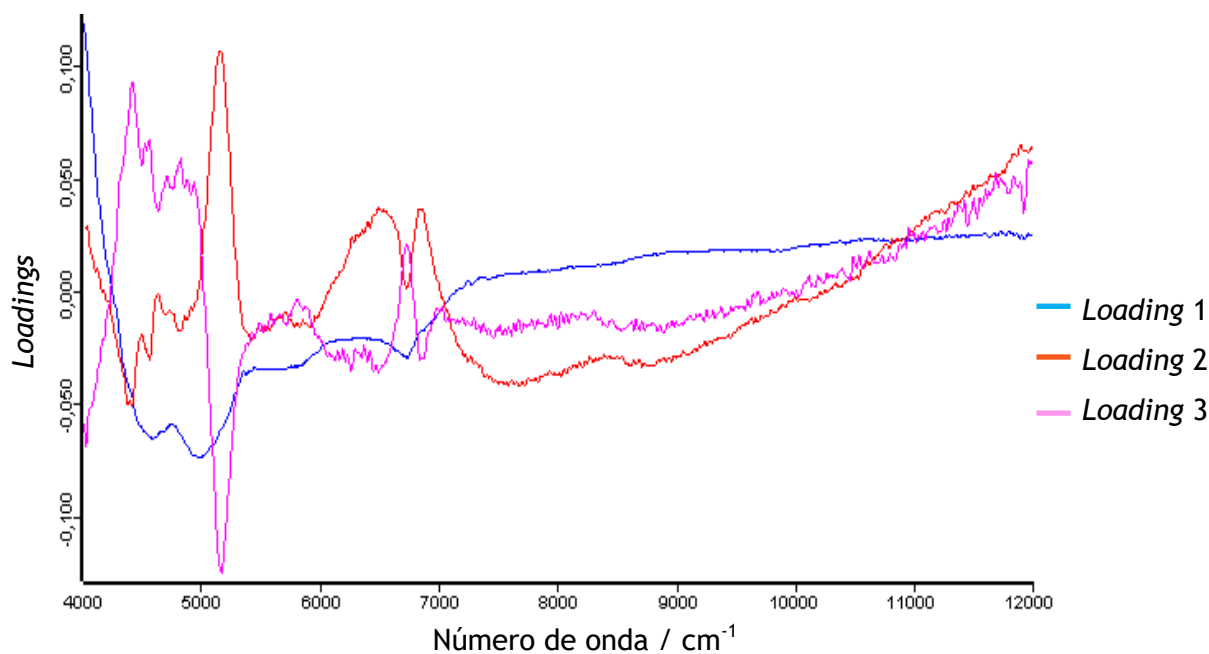


Figura B.2 - Primeiros três *loadings* do PCA do conjunto de espectros utilizados no modelo de quantificação do RC_1 , pré-processados pelo SNV.

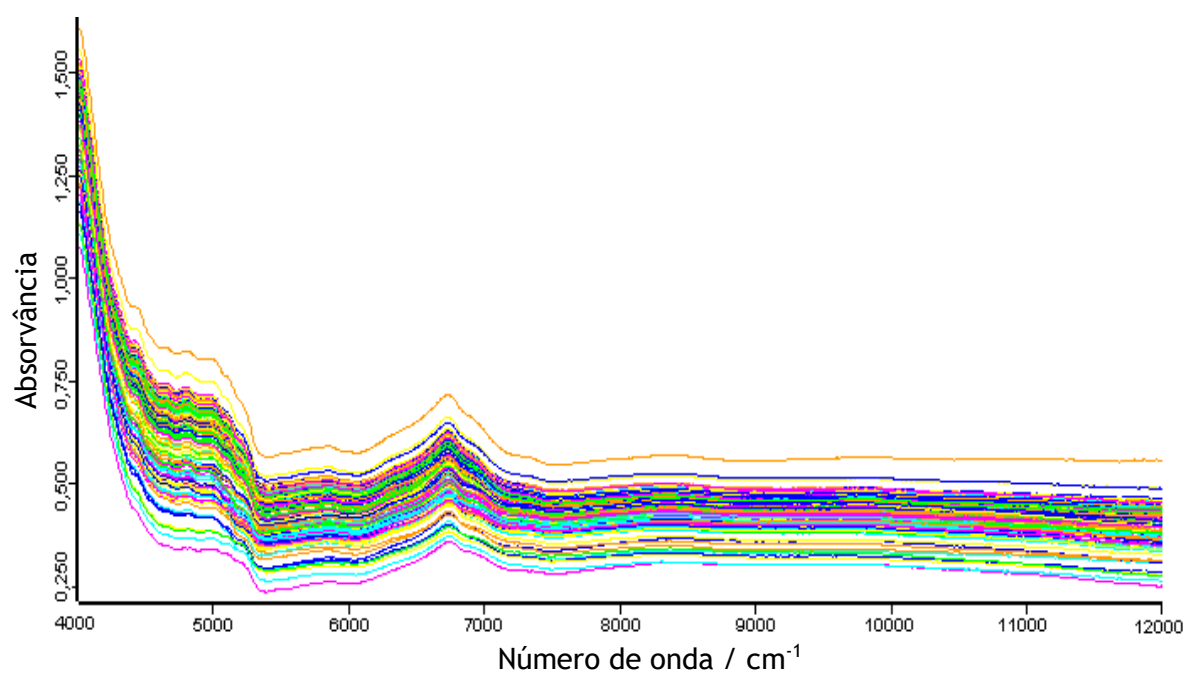


Figura B.3 - Espectros utilizados no modelo de quantificação do RC_2 .

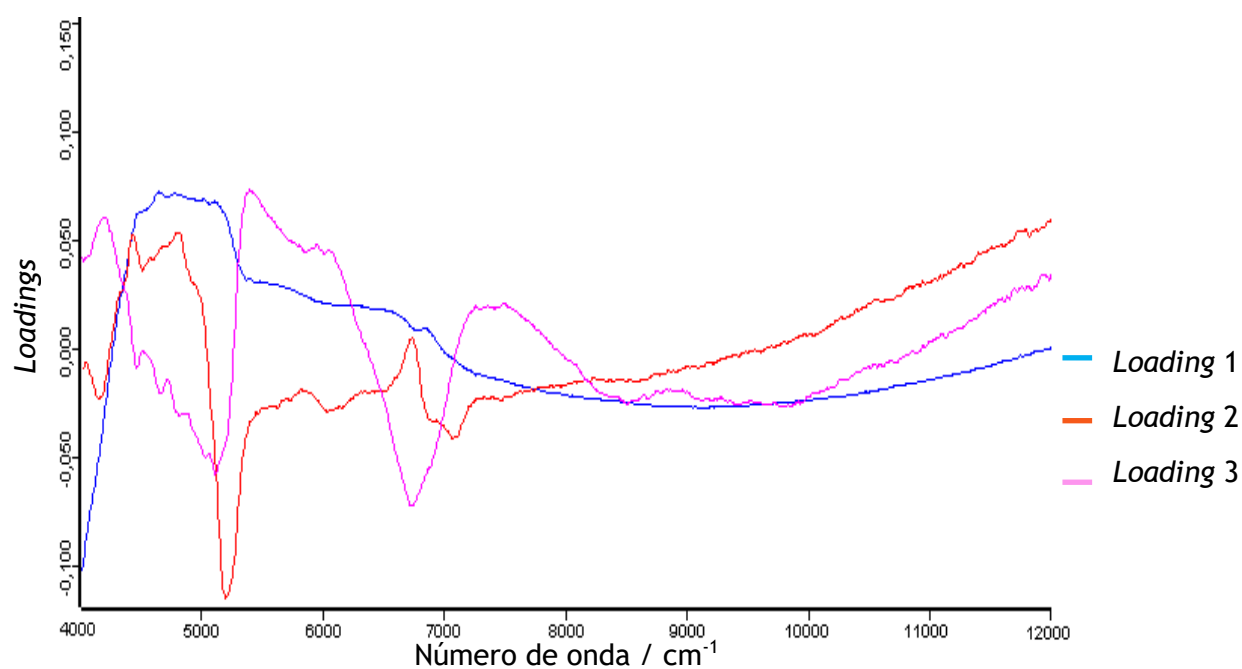


Figura B.4 - Primeiros três *loadings* do PCA do conjunto de espectros utilizados no modelo de quantificação do RC_2 , pré-processados pelo SNV.

B.2 Papel L

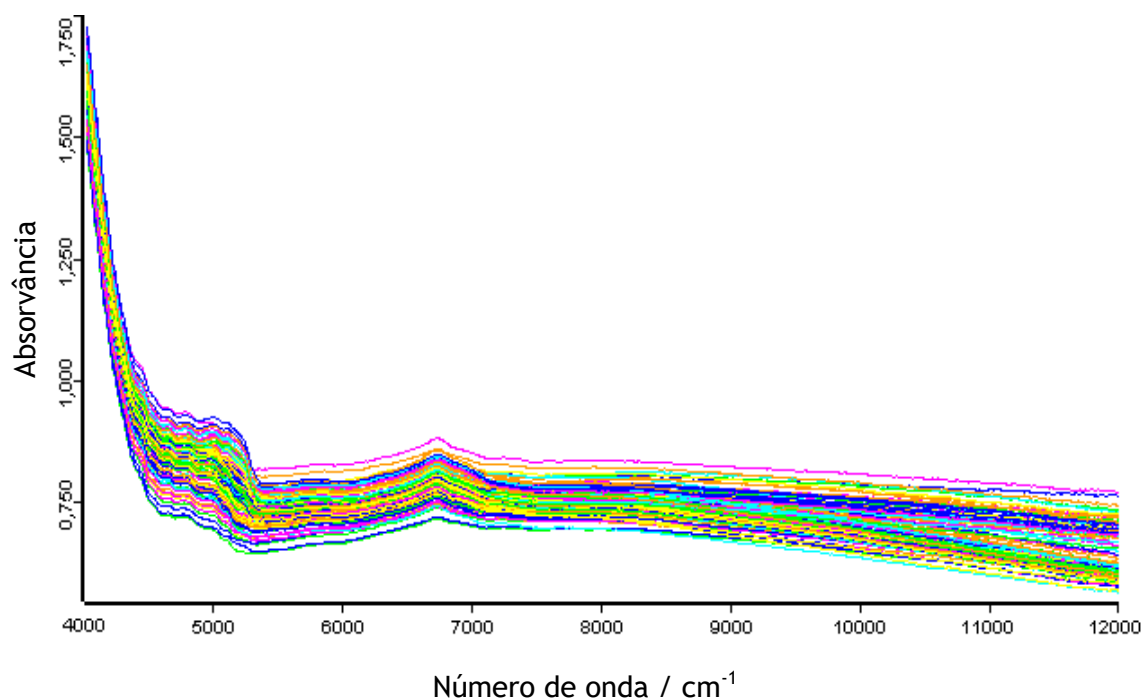


Figura B.5 - Espectros utilizados no modelo de quantificação do RC_1 , obtidos recorrendo ao procedimento das médias.

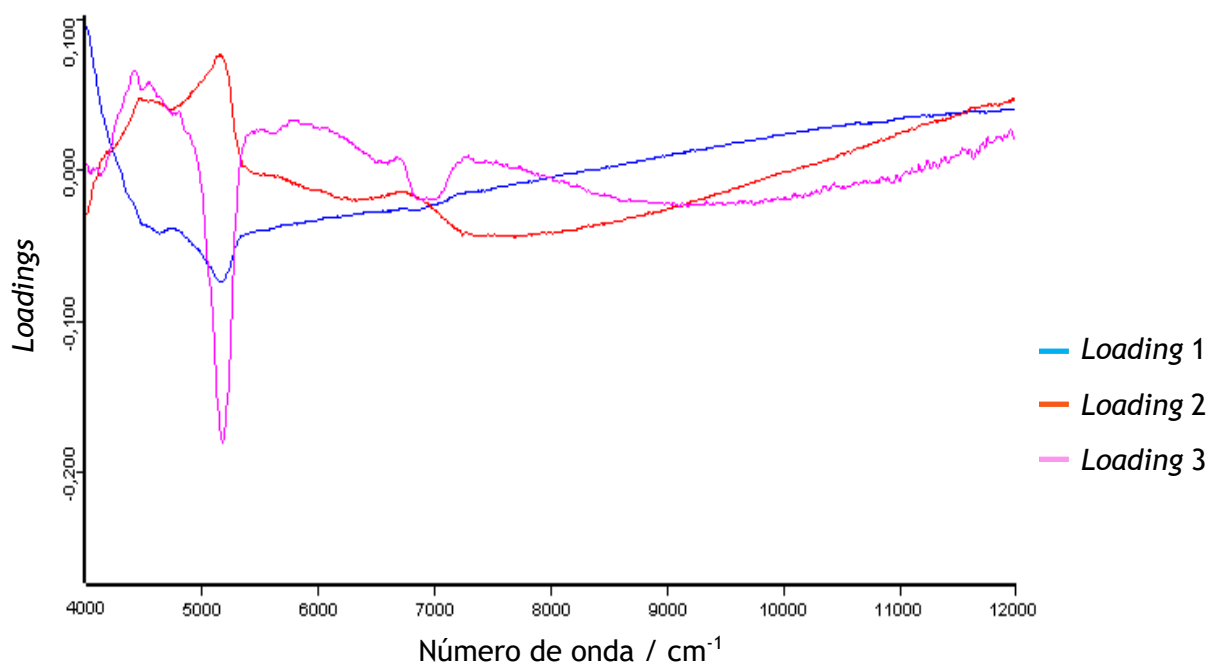


Figura B.6 - Primeiros três *loadings* do PCA do conjunto de espectros adquirido pelos procedimento das médias e utilizados no modelo de quantificação do RC_1 , pré-processados pelo SNV.

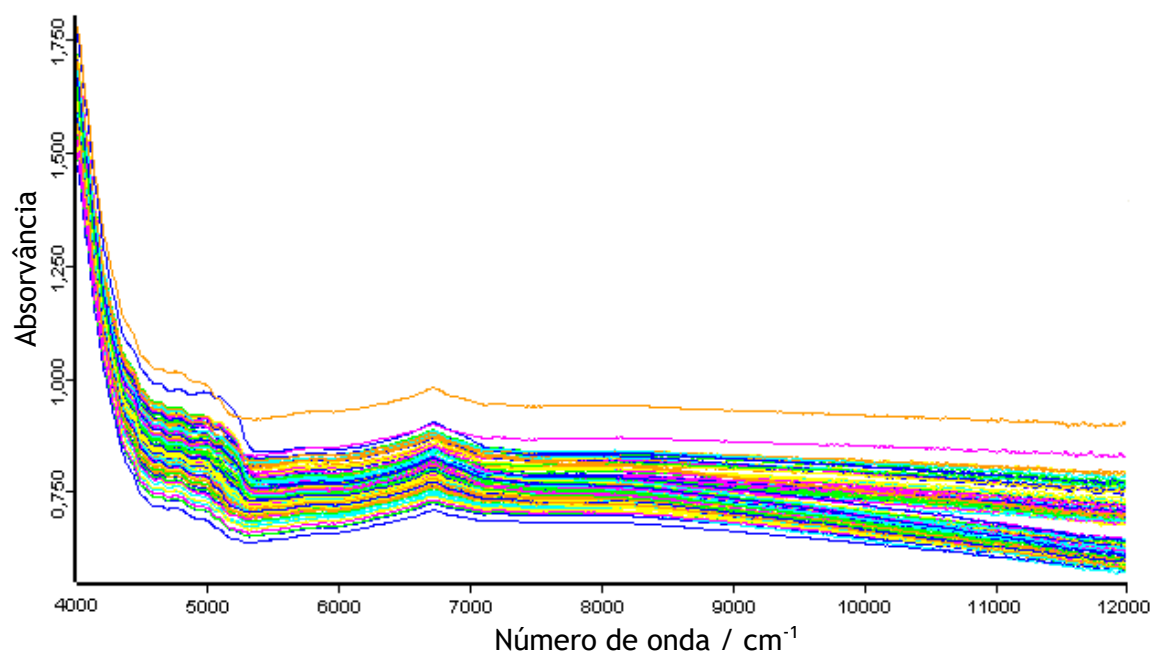


Figura B.7 - Espectros utilizados no modelo de quantificação do RC_1 , obtidos recorrendo ao procedimento do pórtico.

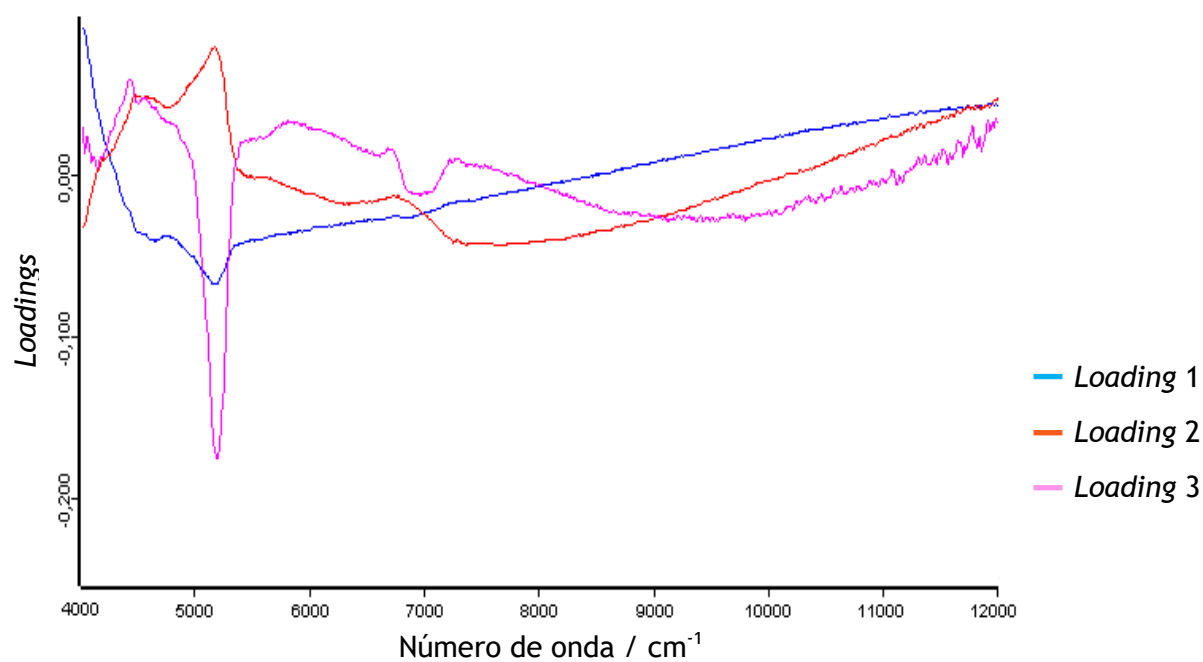


Figura B.8 - Primeiros três loadings do PCA do conjunto de espectros adquirido pelos procedimento do pórtico e utilizados no modelo de quantificação do RC_1 , pré-processados pelo SNV.

Apêndice C - Resultados das validações

C.1 Papel B1

Tabela C.1 - Resultados da validação para o modelo para determinar o RC_1 .

Pré-processamento	Intervalos	Região	Validação cruzada			Validação externa		
			$RMSE_{cv}$ (% (m/m))	R^2	Rank	$RMSE_p$ (% (m/m))	R^2	Rank
MSC	5	1,4,5	0,98	0,89	5	1,2	0,76	2
	10	1,4,5,7-10	0,94	0,90	5	1,14	0,78	3
SNV	5	1,2,5	1,02	0,88	4	0,923	0,85	2
	10	2,6,7,10	0,997	0,89	4	0,976	0,84	2
Normalização Mín-máx	5	1,4	1,02	0,88	3	1,09	0,80	2
	10	4,8	0,981	0,89	2	1,24	0,74	2
1ª derivada	5	1-5	1,12	0,86	4	1,13	0,79	3
	10	1,2,4,6-9	1,09	0,87	5	0,971	0,84	4
2ª derivada	5	1-5	1,12	0,86	4	1,13	0,79	3
	10	1,2,5-7,10	1,24	0,83	4	1,08	0,80	2
COE	5	1,3,5	0,931	0,90	5	1,03	0,82	3
	10	1,3,7,8,10	0,911	0,91	4	1,09	0,80	1
Subtração de uma linha reta	5	1,3-5	1,04	0,88	6	1,19	0,76	3
	10	1,2,5-10	1,04	0,88	6	1,19	0,76	3
1ª derivada + Subtração de uma linha reta	5	1-5	1,25	0,83	4	1,06	0,81	5
	10	1,7,10	1,26	0,82	2	1,04	0,81	2
1ª derivada + SNV	5	1,3-5	1,27	0,82	3	1,17	0,77	1
	10	4,5,7-9	1,23	0,83	4	1,08	0,80	2
1ª derivada + MSC	5	1-5	1,27	0,82	3	1,04	0,82	4
	10	5-10	1,28	0,82	3	1,08	0,81	8
Nenhum	5	1-3,5	0,952	0,90	5	0,918	0,86	3
	10	1,2,9,10	0,929	0,90	4	0,923	0,86	2
	20	2,16	0,98	0,89	3	0,958	0,85	3

Tabela C.2 - Resultados da validação para o modelo para determinar o RC_2 .

Pré-processamento	Intervalos	Região	Validação cruzada			Validação externa		
			$RMSE_{CV}$ (% (m/m))	R^2	Rank	$RMSE_P$ (% (m/m))	R^2	Rank
MSC	5	1-5	1,11	0,78	7	1,51	0,65	4
	10	1-9	1,12	0,78	7	1,3	0,74	8
SNV	5	1,2,4,5	1,12	0,78	7	1,49	0,66	6
	10	1,3-5,8,9	1,01	0,82	6	1,3	0,74	4
	20	14-16, 18-19	1,13	0,77	4	1,51	0,65	2
Normalização Mín-máx	5	1,4,5	1,11	0,78	7	1,52	0,65	6
	10	1,2,5,7-10	1,09	0,79	7	1,54	0,64	6
1ª derivada	5	1-4,7-10	1,28	0,71	4	1,53	0,64	2
	10	3,4,7-9	1,25	0,72	4	1,63	0,59	2
2ª derivada	5	2-5	1,39	0,65	4	1,6	0,61	2
	10	3-10	1,39	0,65	4	1,6	0,61	2
COE	5	1,3-5	1,13	0,77	7	1,66	0,58	4
	10	1,5,7-10	1,08	0,79	7	1,66	0,59	6
Subtração de uma linha reta	5	3-5	1,21	0,74	5	1,55	0,66	2
	10	2,3,5-10	1,21	0,74	5	1,5	0,66	2
1ª derivada + Subtração de uma linha reta	5	2,4,5	1,28	0,71	4	1,57	0,62	2
	10	1,3,7-10	1,27	0,71	4	1,57	0,62	2
1ª derivada + SNV	5	4-5	1,26	0,72	4	1,69	0,56	1
	10	7-10	1,26	0,72	4	1,69	0,56	1
1ª derivada + MSC	5	2,4,5	1,26	0,72	4	1,69	0,56	1
	10	3,7-10	1,26	0,72	4	1,69	0,56	1
Nenhum	5	1,2,4,5	1,07	0,79	7	1,69	0,56	3
	10	1,2,4,6,7, 9,10	1,14	0,77	7	1,67	0,57	3

Tabela C.3 - Resultados da validação para o modelo para determinar o *RC*.

Pré-processamento	Intervalos	Região	Validação cruzada			Validação externa		
			$RMSE_{CV}$ (% (m/m))	R^2	Rank	$RMSE_P$ (% (m/m))	R^2	Rank
MSC	5	1-5	1,08	0,97	7	1,35	0,95	6
	10	1-3,7-10	1,14	0,97	5	1,54	0,94	5
SNV	5	1,3-5	1,07	0,97	7	1,31	0,95	6
	10	2-4,7-10	1,05	0,98	6	1,31	0,96	5
	20	3,5,9,14-16,18,20	1,06	0,97	6	1,37	0,95	5
Normalização Mín-máx	5	1,2,4,5	1,07	0,97	8	1,5	0,94	7
	10	1,5,7-10	1,07	0,97	8	1,47	0,94	7
1ª derivada	5	1,2,4,5	1,16	0,97	6	1,6	0,93	4
	10	3-5,7-10	1,15	0,97	7	1,58	0,93	6
2ª derivada	5	1,4,5	1,38	0,96	4	1,77	0,92	4
	10	1,4,7-10	1,38	0,96	4	1,77	0,92	4
COE	5	2,4,5	1,13	0,91	7	1,47	0,94	7
	10	1-5,7-10	1,09	0,97	8	1,56	0,94	8
Subtração de uma linha reta	5	2-5	1,24	0,96	6	1,58	0,93	6
	10	4-10	1,23	0,97	6	1,55	0,94	6
1ª derivada + Subtração de uma linha reta	5	1-5	1,31	0,96	5	1,68	0,93	4
	10	1,4,7-10	1,31	0,96	5	1,69	0,93	4
1ª derivada + SNV	5	2-5	1,23	0,97	6	1,61	0,93	4
	10	3,5,7-10	1,23	0,97	6	1,62	0,93	4
1ª derivada + MSC	5	2-5	1,24	0,97	6	1,62	0,93	4
	10	3,5-10	1,23	0,97	6	1,62	0,93	4
Nenhum	5	3,4,5	1,15	0,97	6	1,47	0,94	7
	10	4,7-10	1,12	0,97	8	1,61	0,93	7

Tabela C.4 - Resultados da validação para o modelo para determinar o VC_2 .

Pré-processamento	Intervalos	Região	Validação cruzada			Validação externa		
			$RMSE_{CV}$ (% (m/m))	R^2	Rank	$RMSE_P$ (% (m/m))	R^2	Rank
MSC	5	1,3-5	0,411	0,90	4	0,438	0,87	2
	10	1-6,9	0,408	0,90	2	0,383	0,90	2
SNV	5	3-5	0,408	0,90	4	0,436	0,87	2
	10	1,4-6,8,9	0,4	0,91	4	0,385	0,90	3
Normalização Mín-máx	5	2,3,5	0,428	0,89	4	0,442	0,87	2
	10	1,5-7,9	0,386	0,92	5	0,389	0,90	4
	20	4,9,13, 17,18	0,38	0,92	4	0,41	0,89	3
1ª derivada	5	1-5	0,447	0,88	2	0,38	0,87	2
	10	1-5,8-10	0,446	0,88	2	0,443	0,87	2
2ª derivada	5	1,5	0,462	0,87	2	0,463	0,86	2
	10	1,3,5,8-10	0,46	0,87	2	0,462	0,86	2
COE	5	3-5	0,44	0,89	4	0,44	0,87	4
	10	4-10	0,439	0,89	4	0,451	0,86	4
Subtração de uma linha reta	5	1-5	0,436	0,89	3	0,427	0,89	3
	10	1-5,8-10	0,437	0,89	2	0,443	0,87	2
1ª derivada + Subtração de uma linha reta	5	1-3,5	0,451	0,88	2	0,44	0,87	2
	10	1-6,8-10	0,45	0,88	2	0,441	0,87	2
1ª derivada + SNV	5	1,2,5	0,434	0,89	2	0,426	0,88	2
	10	1,2,4,5,8,9	0,416	0,90	2	0,381	0,90	2
1ª derivada + MSC	5	1-5	0,435	0,89	2	0,42	0,88	2
	10	1-6,9	0,413	0,90	2	0,372	0,91	2
Nenhum	5	1,2,4,5	0,44	0,89	5	0,453	0,86	4
	10	1,2,4,5,7, 9,10	0,44	0,89	5	0,427	0,88	5

Tabela C.5 - Resultados da validação para o modelo para determinar o VC.

Pré-processamento	Intervalos	Região	Validação cruzada			Validação externa		
			$RMSE_{CV}$ (% (m/m))	R^2	Rank	$RMSE_P$ (% (m/m))	R^2	Rank
MSC	5	1,3-5	0,437	0,90	4	0,421	0,88	3
	10	3-6,8,9	0,41	0,91	4	0,404	0,89	2
SNV	5	4-5	0,422	0,90	5	0,383	0,90	4
	10	3,5-7,9	0,415	0,91	5	0,388	0,90	3
Normalização Mín-máx	5	2,4,5	0,432	0,90	5	0,42	0,88	4
	10	2,7-9	0,4	0,91	6	0,391	0,90	4
	20	2,17,19	0,432	0,90	5	0,403	0,89	4
1ª derivada	5	1-5	0,455	0,89	4	0,437	0,87	3
	10	1-6,8-10	0,454	0,89	4	0,44	0,87	3
2ª derivada	5	2,3,5	0,5	0,87	3	0,466	0,86	3
	10	3-6,8-10	0,498	0,87	3	0,463	0,86	3
COE	5	1,4,5	0,447	0,89	6	0,422	0,88	5
	10	5,7-10	0,44	0,90	5	0,426	0,88	4
Subtração de uma linha reta	5	2-5	0,438	0,90	4	0,42	0,88	3
	10	1-4,6-10	0,439	0,90	4	0,418	0,88	3
1ª derivada + Subtração de uma linha reta	5	1-5	0,458	0,89	4	0,438	0,87	3
	10	2,5,7-9	0,457	0,89	3	0,437	0,87	2
1ª derivada + SNV	5	1,3-5	0,441	0,90	4	0,423	0,88	3
	10	1,2,4- 6,8,9	0,399	0,91	4	0,401	0,89	2
1ª derivada + MSC	5	1-5	0,44	0,90	4	0,421	0,88	3
	10	2-9	0,423	0,90	4	0,388	0,90	2
Nenhum	5	3-5	0,452	0,89	5	0,45	0,86	4
	10	3,4,8-10	0,444	0,89	5	0,42	0,88	4

C.2 Papel L

Tabela C.6 - Resultados da validação para o modelo para determinar o RC_1 , utilizando o procedimento das médias.

Pré-processamento	Intervalos	Região	Validação cruzada			Validação externa		
			$RMSE_{CV}$ (% (m/m))	R^2	Rank	$RMSE_P$ (% (m/m))	R^2	Rank
MSC	5	1-3,5	1,7	0,88	4	1,71	0,85	4
	10	2,6-8,10	1,69	0,89	6	1,83	0,83	3
SNV	5	3,5	1,72	0,88	5	1,63	0,86	5
	10	7,9,10	1,72	0,88	4	1,66	0,86	4
Normalização Mín-máx	5	1,5	1,66	0,89	6	1,59	0,87	5
	10	1,3,9,10	1,66	0,89	6	1,63	0,87	5
1ª derivada	5	1,5	1,74	0,88	3	1,64	0,86	3
	10	1,2,9,10	1,74	0,88	3	1,64	0,86	3
2ª derivada	5	2,4,5	1,91	0,85	2	1,82	0,83	3
	10	1,3,6,7,10	1,87	0,86	2	1,93	0,81	2
COE	5	2-5	1,66	0,89	5	1,51	0,88	5
	10	2,3,7,10	1,61	0,90	5	1,6	0,87	4
	20	2,5,15,18-20	1,47	0,91	7	1,54	0,88	5
Subtração de uma linha reta	5	1-5	1,64	0,89	5	1,62	0,87	5
	10	2-5,7-10	1,63	0,89	4	1,58	0,87	4
1ª derivada + Subtração de uma linha reta	5	1,3,5	1,76	0,88	3	1,72	0,85	3
	10	1-3,6,7,9,10	1,75	0,88	3	1,71	0,85	3
1ª derivada + SNV	5	1,3-5	1,72	0,88	3	1,57	0,88	3
	10	6-8,10	1,74	0,88	1	1,67	0,86	2
1ª derivada + MSC	5	1-5	1,73	0,88	3	1,6	0,87	3
	10	2-10	1,74	0,88	3	1,6	0,87	3
Nenhum	5	1,2,5	1,58	0,90	5	1,88	0,82	5
	10	1-3,7,10	1,6	0,90	6	1,63	0,87	5

Tabela C.7 - Resultados da validação para o modelo para determinar o RC_1 , utilizando o procedimento do pórtico.

Pré-processamento	Intervalos	Região	Validação cruzada			Validação externa		
			$RMSE_{CV}$ (% (m/m))	R^2	Rank	$RMSE_P$ (% (m/m))	R^2	Rank
MSC	5	4,5	1,13	0,95	5	0,802	0,97	5
	10	4,5,9,10	1,11	0,95	4	0,809	0,97	4
SNV	5	3,5	1,12	0,85	5	0,809	0,97	6
	10	5-9	1,06	0,96	10	1,24	0,92	4
Normalização Mín-máx	5	4,5	1,09	0,95	6	0,834	0,96	6
	10	7-10	1,09	0,95	6	0,836	0,96	6
1ª derivada	5	2-5	1,08	0,96	5	0,775	0,97	5
	10	2,4-10	1,07	0,96	5	0,83	0,96	5
2ª derivada	5	2,4,5	1,17	0,95	4	0,938	0,96	4
	10	4,7-10	1,14	0,95	4	0,999	0,95	4
COE	5	2,4,5	1,14	0,95	7	0,973	0,95	7
	10	5,7,8,10	1,11	0,95	6	0,842	0,96	6
Subtração de uma linha reta	5	4,5	1,12	0,95	5	0,805	0,97	5
	10	2,6-8,10	1,08	0,96	4	0,914	0,96	4
1ª derivada + Subtração de uma linha reta	5	3-5	1,09	0,95	4	0,772	0,97	4
	10	2,4,6-8,10	1,07	0,96	4	0,973	0,95	2
1ª derivada + SNV	5	3-5	1,07	0,96	4	0,765	0,97	4
	10	2,7-10	1,06	0,96	4	0,912	0,96	3
1ª derivada + MSC	5	3-5	1,06	0,96	4	0,765	0,97	4
	10	4-10	1,06	0,96	4	0,888	0,96	3
	20	12, 15-17,19	1,05	0,96	5	1,08	0,94	4
Nenhum	5	3,5	1,15	0,94	7	0,902	0,96	6
	10	4-10	1,11	0,95	6	0,931	0,96	5

Tabela C.8 - Resultados da validação para o modelo para determinar o VC_1 , utilizando o procedimento das médias.

Pré-processamento	Intervalos	Região	Validação cruzada			Validação externa		
			$RMSE_{CV}$ (% (m/m))	R^2	Rank	$RMSE_P$ (% (m/m))	R^2	Rank
MSC	5	1-5	0,802	0,86	3	0,841	0,88	3
	10	3-5,9,10	0,852	0,84	2	0,822	0,88	2
SNV	5	1-5	0,81	0,86	3	0,844	0,88	3
	10	1,6-10	0,808	0,86	4	0,859	0,87	3
Normalização Mín-máx	5	1,3-5	0,822	0,85	3	0,832	0,88	3
	10	1,4,5,8-10	0,809	0,86	3	0,831	0,88	3
1ª derivada	5	1,2,5	0,832	0,85	1	0,834	0,88	1
	10	1-3,8-10	0,832	0,85	1	0,837	0,88	1
2ª derivada	5	2,5	0,813	0,85	1	0,87	0,87	1
	10	4,6,9,10	0,812	0,85	1	0,868	0,87	1
COE	5	1,3-5	0,835	0,85	3	0,817	0,89	3
	10	1,3,9	0,828	0,85	3	0,806	0,89	3
Subtração de uma linha reta	5	1-3,5	0,804	0,86	2	0,836	0,88	1
	10	3,4,6-9	0,836	0,85	2	0,903	0,86	1
1ª derivada + Subtração de uma linha reta	5	2,4,5	0,836	0,85	1	0,842	0,88	1
	10	3-6,8-10	0,834	0,85	1	0,846	0,8	1
1ª derivada + SNV	5	3,5	0,837	0,85	1	0,856	0,87	1
	10	5,6,8-10	0,836	0,85	1	0,855	0,88	1
1ª derivada + MSC	5	1,3-5	0,837	0,85	1	0,848	0,88	1
	10	1,4,6,8-10	0,835	0,85	1	0,854	0,88	1
Nenhum	5	1-5	0,823	0,85	4	0,816	0,89	4
	10	1,5,7-9	0,808	0,86	4	0,829	0,88	4
	20	1,9,11,17	0,804	0,86	4	0,819	0,89	4

Tabela C.9 - Resultados da validação para o modelo para determinar o VC_1 , utilizando o procedimento do pórtico.

Pré-processamento	Intervalos	Região	Validação cruzada			Validação externa		
			$RMSE_{CV}$ (% (m/m))	R^2	Rank	$RMSE_P$ (% (m/m))	R^2	Rank
MSC	5	1,2,4,5	0,76	0,90	5	0,779	0,90	2
	10	6,7,9,10	0,668	0,92	8	0,686	0,92	4
	20	11-15, 17,18	0,624	0,93	9	0,679	0,92	4
SNV	5	1-3,5	0,851	0,87	6	0,766	0,90	3
	10	6-9	0,693	0,91	10	0,824	0,88	2
Normalização Mín-máx	5	4,5	0,87	0,86	2	0,788	0,89	2
	10	5-7,9,10	0,703	0,9	10	0,732	0,91	4
1ª derivada	5	1-3,5	0,879	0,86	1	0,769	0,90	1
	10	6-10	0,779	0,89	8	0,814	0,87	1
2ª derivada	5	1,3-5	0,788	0,89	4	0,834	0,88	1
	10	5-7,9	0,712	0,91	4	0,853	0,88	1
COE	5	1,2,4,5	0,87	0,87	4	0,654	0,93	3
	10	6,7,9,10	0,8	0,89	6	0,791	0,89	4
Subtração de uma linha reta	5	1,4,5	0,792	0,89	4	0,802	0,89	1
	10	6,7,9	0,745	0,90	9	0,716	0,91	2
1ª derivada + Subtração de uma linha reta	5	2,4,5	0,816	0,88	3	0,788	0,89	1
	10	6,9	0,762	0,90	6	0,836	0,88	1
1ª derivada + SNV	5	4,5	0,875	0,86	1	0,785	0,89	1
	10	6-9	0,711	0,91	7	0,664	0,92	2
1ª derivada + MSC	5	1,4,5	0,867	0,87	2	0,793	0,89	2
	10	6-9	0,658	0,92	8	0,746	0,91	1
Nenhum	5	1,3,5	0,836	0,88	8	0,92	0,92	5
	10	8-10	0,857	0,87	3	0,739	0,91	3