



UnB

Instituto de Química
Programa de Pós-Graduação em Química

DISSERTAÇÃO DE MESTRADO

IDENTIFICAÇÃO DE MADEIRAS DO GÊNERO *DALBERGIA* COM
ESPECTRÔMETRO NIR PORTÁTIL POR PLS-DA E SIMCA.

FILIPE AUGUSTO SNEL DE OLIVEIRA BARROS

ORIENTADOR: JEZ WILLIAN BATISTA BRAGA

COORIENTADORA: TEREZA CRISTINA MONTEIRO PASTORE

Brasília, DF

2018

INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

IDENTIFICAÇÃO DE MADEIRAS DO GÊNERO *DALBERGIA* COM
ESPECTRÔMETRO NIR PORTÁTIL POR PLS-DA E SIMCA.

FILIFE AUGUSTO SNEL DE OLIVEIRA BARROS

ORIENTADOR: JEZ WILLIAN BATISTA BRAGA

COORIENTADORA: TEREZA CRISTINA MONTEIRO PASTORE

DISSERTAÇÃO DE MESTRADO EM QUÍMICA

Brasília, DF

2018

FOLHA DE APROVAÇÃO

Comunicamos a aprovação da Defesa de Dissertação do (a) aluno (a) **Fillipe Augusto Snel de Oliveira Barros**, matrícula nº **17/0097498**, intitulada ***“Identificação de madeiras do gênero Dalbergia com espectrômetro NIR portátil por PLS-DA e SIMCA”***, apresentada no (a) Auditório Azul do Instituto de Química (IQ) da Universidade de Brasília (UnB) em 14 de dezembro de 2018.

Prof. Dr. Jez Willian Batista Braga
Presidente de Banca

Prof. Dr. Alexandre Fonseca
Membro Titular

Dra. Fernanda Piccolo Pieruzzi
Membro Titular Serviço Florestal Brasileiro
(SFB)

Prof. Dr. Paulo José Prudente de Fontes
Membro Suplente

Em 14 de dezembro de 2018.

*Dedico este trabalho à minha vó,
à minha mãe, ao meu irmão e à minha tia..*

AGRADECIMENTOS

A Deus, Pai eterno e de infinita bondade.

À minha amada vó Dita Paula Snel de Oliveira, mulher guerreira e dedicada a família, que acompanhou a reta final desse trabalho de uma posição privilegiada, por sempre me dar amor e conhecimento.

À minha mãe Marília Viviane Snel de Oliveira, tradução do que é amor e conhecimento, por todo o apoio, científico e psicológico, e acompanhamento durante as longas madrugadas de trabalho.

À minha família que é a minha base por sempre me dar forças para seguir firme e compreender a minha ausência em algumas ocasiões.

Ao meu orientador, Prof. Dr. Jez Willian Batista Braga, por todo o ensinamento, críticas construtivas e discussões que foram fundamentais ao longo de todo o percurso.

À minha coorientadora, Dra. Tereza Cristina Monteiro Pastore, pelo apoio incondicional e incentivo que foram fundamentais à minha evolução na área científica.

Ao Dr. Alessandro Moreira, que em 2014 me deu a oportunidade de estagiar e conhecer o LPF, pelo empréstimo de livros e pelos diversos diálogos construtivos.

À Dra. Vera Teresa Rauber Coradin, pelas contribuições e identificação das amostras.

À equipe de trabalho: Diego Carvalho, Priscila Veras, Liz Soares, Dhébora Juliana e Rafaela Dornelas, por proporcionar um ambiente de trabalho extremamente prazeroso.

Aos meus amigos por todo o suporte e incentivo durante esses dois anos.

Ao programa de pós graduação do Instituto de Química da Universidade de Brasília pela oportunidade de realizar o mestrado.

Ao Laboratório de Produtos Florestais (LPF) por proporcionar as amostras de madeira.

Aos laboratórios visitados pela disponibilidade em me receber e permitir a aquisição dos espectros utilizados nessa dissertação.

À CAPES pela bolsa de mestrado.

Às agências fomentadoras CITES - ITTO, CNPq (308748/2015-8), INCTBio (465389/2014-7), FAPDF (193.001.643/2017) pelo apoio financeiro.

SUMÁRIO

1. INTRODUÇÃO E OBJETIVOS	13
2. REVISÃO BIBLIOGRÁFICA	17
2.1 <i>Dalbergia nigra</i>	17
2.2 Métodos de discriminação de <i>Dalbergia</i>	21
2.3 Comércio madeireiro.	22
2.4 Métodos de identificação de espécie florestal.	25
2.4.1 Anatomia da madeira.....	25
2.4.2 Espectroscopia de massa.	26
2.4.3 Isótopos estáveis	27
2.4.4 Radiocarbono	28
2.4.5 Código de barra de DNA.....	29
2.4.6 Tecnologia NIRS	29
3. FUNDAMENTAÇÃO TEÓRICA	31
3.1 Componentes químicos da madeira	31
3.2 Espectroscopia no infravermelho próximo.....	34
3.3 Modelos Quimiométricos	36
3.3.1 Seleção de amostras e exclusão de <i>outliers</i>	36
3.3.2 Análise Discriminante por Mínimos Quadrados (PLS-DA).....	37
3.3.3 Modelagem Independente Flexível por Analogia de Classes (SIMCA).	38
3.3.4 Figuras de Mérito.....	40
4. MATERIAIS E MÉTODOS.....	42
4.1 Preparação das amostras	42
4.2 Espectrômetro NIR portátil.	42
4.3 Pré-processamento dos espectros no infravermelho próximo.	44
4.4 Organizações das classes discriminadas	44
5. RESULTADOS E DISCUSSÃO	47

5.1 Discriminação da <i>Dalbergia nigra</i> das demais 14 espécies de <i>Dalbergia</i>	48
5.1.1 PLS-DA.....	48
5.1.2 SIMCA	57
5.2 Discriminação da <i>Dalbergia spruceana</i> das demais 14 espécies de <i>Dalbergia</i>	63
5.2.1 PLS-DA.....	63
5.2.2 SIMCA	70
5.3 Discriminação individual das nove espécies de <i>Dalbergia</i>	75
5.3.1 PLS-DA.....	75
5.3.2 SIMCA	77
5.3.3 Discriminações de espécies que não tiveram sua classe modelada na fase de treinamento.....	80
6. CONCLUSÕES.....	82
7. REFERÊNCIAS	84

LISTA DE TABELAS

Tabela 1. Características físicas da <i>D. nigra</i> e <i>D. spruceana</i>	18
Tabela 2. Informações das amostras estudadas, número, espectros e origem.	43
Tabela 3. Quantidade de variáveis latentes e espectros anômalos excluído nos dois modelos de discriminação de <i>D. nigra</i>	51
Tabela 4. Figuras de mérito obtidas com base no conjunto de validação dos dois modelos discriminatórios de <i>D. nigra</i> com a ferramenta PLS-DA.	56
Tabela 5. Figuras de mérito obtidas com base no conjunto de validação da discriminação de <i>D. nigra</i> com ferramenta SIMCA.....	61
Tabela 6. Tabela confusão do modelo SIMCA com seleção variável do tipo 1.	62
Tabela 7. Tabela confusão do modelo SIMCA com seleção de variável do tipo 2.	62
Tabela 8. Quantidade de variáveis latentes e espectros anômalos nas fases de treinamento e validação dos dois tipos de modelos de discriminação de <i>D. spruceana</i>	64
Tabela 9. Figuras de mérito obtidas com base no conjunto de validação da discriminação de <i>D. spruceana</i> utilizando ferramenta PLS-DA utilizando os dois tipos de seleção de variável.....	69
Tabela 10. Figuras de méritos obtidas com base no conjunto de validação dos modelos de discriminação da <i>D. spruceana</i> utilizando a ferramenta SIMCA.....	73
Tabela 11. Tabela confusão do modelo SIMCA com seleção variável do tipo 1.	73
Tabela 12. Tabela confusão do modelo SIMCA com seleção variável do tipo 2.	74
Tabela 13. Quantidade de variáveis latentes e espectros anômalos nas fases de treinamento e validação dos modelos individuais.	76
Tabela 14. Figuras de méritos obtidas com base no conjunto de validação para os modelos individuais com a ferramenta PLS-DA.	77
Tabela 15. Figuras de mérito do modelo discriminatório das nove espécies de <i>Dalbergia</i> utilizando a ferramenta SIMCA.....	78
Tabela 16. Tabela confusão do modelo discriminatório das nove espécies do gênero <i>Dalbergia</i>	79

Tabela 17. Quantidade e porcentual de amostras do grupo “diversas” que o modelo PLS-DA discriminou de forma errônea.	80
--	----

LISTA DE FIGURAS

Figura 1. Mesa dos Tratados feita em jacarandá, exposta no palácio do Itamaraty em Brasília. ⁵	13
Figura 2. Violão feito de <i>D. nigra</i> . Foto de Gunesch. ¹⁹	17
Figura 3. Amostra da madeira da espécie <i>D. nigra</i> à esquerda e da madeira da espécie <i>D. spruceana</i> à direita (Acervo pessoal).....	18
Figura 4. Corte tangencial longitudinal com aumento de 100x: (A) <i>D. cearensis</i> ; (B) <i>D. miscolobium</i> ; (C) <i>D. nigra</i> ; (D) <i>D. spruceana</i> ; (E) <i>D. tucurensis</i> ; (F) <i>D. granadillo</i> ; (G) <i>D. stevensonii</i> ; (H) <i>D. nigra</i> . Fonte: Gasson <i>et al.</i> ²⁵	20
Figura 5. Fotos das madeiras do gênero <i>Dalbergia</i> estudadas nesse trabalho. ^{27,28,29,30}	21
Figura 6. Quantidade de celulose brasileira exportada em toneladas.	23
Figura 7. Esquema do ciclo vicioso entre raridade de espécie produtora de madeira, valor e comércio ilegal (Gan <i>et al.</i> ³⁸).	24
Figura 8. Representação da ligação glicosídica ($\beta 1 \rightarrow 4$) presente na estrutura da celulose. ...	31
Figura 9. Principais açúcares presentes na hemicelulose.	32
Figura 10. Estrutura dos três componentes básicos das ligninas: (a) álcool p-cumarílico; (b) álcool coniferilílico e (c) álcool sinapílico.	32
Figura 11. Esquema estrutural da lignina da angiosperma <i>Fagus sylvatica</i> proposto por Nimz. ⁷²	33
Figura 12. Tronco da madeira <i>Dalbergia Nigra</i> com perceptível variação visual entre alborno e cerne.	34
Figura 13. Exemplo gráfico de um modelo SIMCA genérico. ⁶⁰	39
Figura 14. Espectrômetro portátil microPhazir™ RX Analyzer (Fonte: Thermo Fisher Scientific) ⁹⁶	44
Figura 15. Esquema de seleção de espectros dos modelos discriminatório de <i>D. nigra</i> com seleção de variáveis do tipo 1 e 2.	46
Figura 16. Esquema de seleção de espectros do modelo discriminatório de <i>D. spruceana</i> com seleção de variáveis do tipo 2.	46

Figura 17. Espectros médios de cada espécie sem pré-processamento.	47
Figura 18 Espectros médios de cada espécie após pré-processamento com alisamento, 1ª derivada e SNV.	48
Figura 19. Valor de erro de validação cruzada (RMSECV) de acordo com o número de variáveis latentes do modelo de discriminação de <i>D. nigra</i> com seleção de variável do tipo 1.	49
Figura 20. Gráfico dos valores de T^2 de Hotelling e resíduos Q das amostras de treinamento do modelo com seleção de variáveis do tipo 1 (A) e com seleção de variáveis do tipo 2 (B). Legenda: <i>D. nigra</i> (▲) e Demais Dalbergias (*).	50
Figura 21. Gráfico dos valores de T^2 de Hotelling e resíduos Q dos espectros da fase de treinamento e de validação do modelo com seleção de variáveis 1 (A) e com seleção de variável 2 (B). Legenda: <i>D. nigra</i> (▲) e Demais Dalbergias (*).	52
Figura 22. Valor de classe predito dos espectros do modelo PLS-DA de discriminação de <i>D. nigra</i> das demais espécies com seleção de variável do tipo 1. Símbolos: vazios - fase de treinamento; cheios fase de validação.	54
Figura 23. Valor de classe predito dos espectros do modelo PLS-DA de discriminação de <i>D. nigra</i> das demais espécies com seleção de variável do tipo 2. Símbolos: vazios - fase de treinamento; cheios fase de validação.	55
Figura 24. Probabilidade do espectro pertencer à espécie <i>D. nigra</i> segundo o modelo SIMCA com seleção de variável do tipo 1.	59
Figura 25. Probabilidade do espectro pertencer à espécie <i>D. nigra</i> segundo modelo SIMCA com seleção de variável do tipo 2.	60
Figura 26. Valor de erro de validação cruzada (RMSECV) de acordo com número de variáveis Latentes do modelo de discriminação de <i>D. spruceana</i> com seleção de variável do tipo 2.	64
Figura 27. Gráfico dos valores de T^2 de Hotelling e resíduos Q dos espectros de treinamento do modelo com seleção de variável tipo 1 (a) e com seleção de variável tipo 2 (b). Legenda: <i>D. spruceana</i> (■) e Demais Dalbergias (*).	65
Figura 28. Gráfico dos valores de T^2 de Hotelling e resíduos Q dos espectros da fase de treinamento e de validação do modelo com seleção de variável tipo 1 (a) e com seleção de variável tipo 2 (b). Legenda: <i>D. spruceana</i> (■) e Demais Dalbergias (*).	66

Figura 29. Valor de classe predito dos espectros do modelo PLS-DA com seleção de variável do tipo 1.	67
Figura 30. Valor de classe predito dos espectros do modelo PLS-DA com seleção de variável do tipo 2.	68
Figura 31. Probabilidade do espectro pertencer a espécie <i>D. spruceana</i> utilizando o modelo com seleção de variável do tipo 1.	71
Figura 32. Probabilidade do espectro pertencer à espécie <i>D. spruceana</i> utilizando o modelo com seleção de variável do tipo 2.	72

LISTA DE ABREVIATURAS E ACRÔNIMOS

CITES	Convenção sobre o Comércio Internacional de Espécies da Flora e Fauna Selvagens em Perigo de Extinção (do inglês, <i>Convention on International Trade in Endangered Species of Wild Flora and Fauna</i>)
CP	Componente Principal
FN	Falso Negativo
FP	Falso Positivo
IBAMA	Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis
IQ	Instituto de Química
LPF	Laboratório de Produtos Florestais
MMA	Ministério do Meio Ambiente
NIR	Infravermelho Próximo (do inglês, <i>Near Infrared</i>)
NIRS	Espectroscopia no infravermelho próximo (do inglês <i>Near Infrared Spectroscopy</i>)
PCA	Análise de Componentes Principais (do inglês, <i>Principal Component Analysis</i>)
PLS	Mínimos Quadrados Parciais (do inglês, <i>Partial Least Squares</i>)
PLS-DA	Análise Discriminante por Mínimos Quadrados (do inglês, <i>Discriminant Analysis Partial Least Squares</i>)
SIMCA	Modelagem Independente Flexível por Analogia de Classes (do inglês, <i>Soft Independent Modeling of Class Analogies</i>)
UnB	Universidade de Brasília
UNODOC	Nações Unidas sobre Drogas e Crimes (UNODOC, do inglês <i>United Nations Office on Drugs and Crime</i>)
VL	Variável Latente

RESUMO

A madeira da espécie *Dalbergia nigra*, popularmente conhecida como jacarandá-da-Bahia e jacarandá-preto, foi altamente explorada por ser matéria prima de qualidade para móveis de luxo e instrumentos musicais. Como consequência dessa exploração, a espécie está em risco de extinção e a comercialização nacional e internacional sofre restrições e fiscalização. No entanto, a identificação da espécie, tanto a campo como em laboratório, é difícil devido às semelhanças visuais e anatômicas com outras espécies do mesmo gênero. Tendo em vista essa dificuldade de identificação de espécies, essa dissertação teve como objetivo o desenvolvimento de um método prático e eficiente para identificação da *D. nigra* através do uso da espectroscopia na região do infravermelho próximo (NIR) acoplada à análise multivariada, utilizando um espectrômetro portátil. Dois tipos de análises quimiométricas (SIMCA e PLS DA) e duas faixas espectrais foram testadas para a análise de dados espectroscópicos da madeira para distinguir 15 espécies do gênero *Dalbergia*: *D. bariensis*, *D. cearenses*, *D. congestiflora*, *D. decipularis*, *D. glaucescens*, *D. latifolia*, *D. nigra*, *D. melanoxylon*, *D. miscolobium*, *D. retusa*, *D. sisso*, *D. spruceana*, *D. stevensonii*, *D. tucurensis* e *D. villosa*. Foram coletados espectros de 359 amostras com espectrômetro portátil, e avaliaram-se três tipos de discriminações: a discriminação da *D. nigra* das demais 14 espécies; a discriminação da *D. spruceana*, espécie de alto interesse para o mercado brasileiro, das demais espécies incluindo a *D. nigra* e discriminações individuais de nove espécies de *Dalbergia*. A ferramenta PLS-DA apresentou maior eficiência do que a SIMCA em todas as análises, a discriminação da *D. nigra* apresentou eficiência superior a 80 %. A utilização de toda faixa espectral do equipamento teve maior capacidade discriminatória do que o uso de bandas mais relacionadas com os principais constituintes da madeira. Os resultados mostram que é possível identificar a *D. nigra*, assim como outras espécies de *Dalbergia*, utilizando um espectrômetro portátil associado ao modelo PLS-DA. O método desenvolvido pode ser considerado simples, rápido, que não exige treinamento extensivo de fiscais, podendo ser implementado nos pontos de fiscalização promovendo a identificação da *D. nigra* ainda a campo.

Palavras Chaves: *in loco*, Quimiometria, fiscalização, infravermelho

ABSTRACT

The wood of the *Dalbergia nigra* species, popularly known as Brazilian jacarandá and black jacaranda, was highly exploited as a quality raw material for luxury furniture and musical instruments. As a consequence of this exploitation, the species is at risk of extinction and national and international commercialization is subject to restrictions and inspection. However, the identification of the species, both in the field and in the laboratory, is difficult due to visual and anatomical similarities with other species of the same genus. Regarding the difficulty of identifying species, the aim of this dissertation was to develop a practical and efficient method to identify *D. nigra* through the use of infrared spectroscopy (NIR) coupled to multivariate analysis using a spectrophotometer portable. Two types of chemometric analyzes (SIMCA and PLS DA) and two spectral bands were tested to analyze the spectroscopic data of wood to distinguish 15 species of the genus *Dalbergia*: *D. bariensis*, *D. cearenses*, *D. congestiflora*, *D. decipularis*, *D. glaucescens*, *D. latifolia*, *D. nigra*, *D. melanoxylon*, *D. miscolobium*, *D. retusa*, *D. sisso*, *D. spruceana*, *D. stevensonii*, *D. tucurensis* e *D. villosa*. Spectra of 359 samples were collected with portable spectrometer, and three types of discrimination were evaluated: the discrimination of *D. nigra* from the other 14 species; the discrimination of *D. spruceana*, a species of high interest for the Brazilian market, of other species including *D. nigra* and individual discriminations of nine species of *Dalbergia*. PLS-DA showed higher efficiency than SIMCA in all analyzes, the discrimination of *D. nigra* showed efficiency higher than 80 %. The use of all spectral bands of the equipment had a greater discriminatory capacity than the use of bands more related to the main constituents of wood. The results show that it is possible to identify *D. nigra*, as well as other *Dalbergia* species, using a portable spectrometer associated to the PLS-DA model. The developed method can be considered simple, fast, does not require extensive fiscal agents tax training, and can be implemented at the inspection points, promoting the identification of *D. nigra* in the field.

Keywords: *in loco*, chemometric, fiscalization, infrared

1. INTRODUÇÃO E OBJETIVOS

O gênero *Dalbergia*, extensivamente explorado para fornecer madeira a diversos produtos nobres, está distribuído ao redor do mundo, sendo que 39 espécies são nativas do Brasil.¹ A *Dalbergia nigra*, comumente denominada de jacarandá, jacarandá-da-Bahia, caviúna e jacarandá-preto, é uma das espécies de *Dalbergia* naturais do Brasil e é de alto interesse para produção de móveis luxuosos e instrumentos musicais, como violões e até mesmo caixas para piano devido às propriedades acústicas.^{2,3,4}

Dentre diversos produtos oriundos de madeiras *Dalbergia*, um em específico chama atenção pela sua importância histórica para o Brasil. A Figura 1 é uma fotografia da Mesa dos Tratados na qual a princesa Isabel assinou a Lei Áurea, em 1888, essa mesa está exposta no palácio do Itamaraty, em Brasília, DF.



Figura 1. Mesa dos Tratados feita em jacarandá, exposta no palácio do Itamaraty em Brasília.⁵

A valiosa *D. nigra*, que ocorre na mata Atlântica, é facilmente confundida com a *D. spruceana*, que é nativa da Amazônia, quando as suas características anatômicas são comparadas macroscopicamente e microscopicamente. Ou seja, sabendo a origem da madeira pode-se definir qual é a espécie.⁶ Porém, a origem da madeira nem sempre é conhecida, dificultando a discriminação entre essas duas espécies. Adicionalmente, é comum transportadores burlarem as fiscalizações. Logo, para o controle efetivo da madeira de *D. nigra* é necessário um método eficiente que discrimine espécies.

Devido à exploração exacerbada e possível extinção do gênero *Dalbergia*, foram estabelecidos limites no comércio internacional pela Convenção sobre o Comércio Internacional de Espécies da Flora e Fauna Selvagens em Perigo de Extinção (CITES, do inglês, *Convention on International Trade in Endangered Species of Wild Flora and Fauna*) para os 183 países

seguidores da convenção. O objetivo da CITES é assegurar que o comércio de animais e plantas não ponha em risco a sua sobrevivência no estado selvagem. A CITES atribui diferentes graus de proteção a cerca de 30.000 espécies de plantas, sendo 60 espécies produtoras de madeira,⁶ inscritas em três Anexos (I, II e III) de acordo com o grau de proteção.

Cerca de 30.000 espécies de plantas, sendo 60 espécies produtoras de madeira, estão inscritas em três Anexos (I, II e III) da CITES. As espécies classificadas no Anexo I são espécies em perigo de extinção, portanto o comércio internacional é proibido, a movimentação internacional é apenas permitida em condições excepcionais como, por exemplo, para fins de pesquisa científica. As espécies do Anexo II não estão em perigo de extinção, porém estarão em risco caso não haja um controle rígido, compatível com a sua sobrevivência. As espécies do Anexo III são protegidas em pelo menos um país, o que requisitou assistência internacional para controlar o comércio.¹⁰ Para assegurar que o comércio das *Dalbergias* não ponha em risco a sua sobrevivência no estado nativo atribui o grau de proteção máximo à *Dalbergia nigra*, inserindo-a no Anexo I em 1992.³¹ Em 2017, para aumentar o controle, as demais espécies do gênero *Dalbergia* foram listadas no Anexo II e a *D. nigra* permaneceu no Anexo I.¹⁰

Segundo informações do Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis – IBAMA, de 17 de outubro de 2018, o metro cúbico de *Dalbergia nigra* está sendo exportado por 143 mil dólares americanos. Apesar do comércio internacional da *D. nigra* ser proibido no Brasil, desde 1996, o material ainda pode ser exportado quando a data de corte da árvore é comprovadamente anterior à entrada do Brasil na CITES. A madeira é obtida de móveis ou construções que foram desmanchadas.⁸ A madeira da espécie *D. nigra* pode ser comumente vendida no comércio interno. No site de vendas brasileiro Mercado Livre tábuas são vendidas por 300 reais a peça, sendo o valor do metro cúbico de aproximadamente R\$ 196.800,00.⁹

O alto valor econômico associado a essa espécie acarreta a exploração desordenada, descontrolada e ilegal. A *D. nigra* foi a primeira espécie produtora de madeira listada no Anexo I pela CITES, no ano de 1992. Enquanto as outras espécies de *Dalbergia* eram listadas nos Anexos II e III. Tendo em vista a dificuldade de controlar o comércio das espécies do gênero *Dalbergia*, desde o início de 2017, as demais espécies do gênero *Dalbergia* foram listadas no Anexo II da CITES e a *D. nigra* continuou sendo listada no Anexo I, para que dessa maneira aumente a proteção dessas espécies.¹⁰

O controle do gênero *Dalbergia* é bastante dificultado devido à semelhança visual entre as espécies desse mesmo gênero como *D. nigra* e *D. spruceana*, ambas nativas do Brasil. E também à semelhança com espécies de outros gêneros, por exemplo, a espécie *Machaerium scleroxylon* que é visualmente similar à *D. nigra*.¹¹ A identificação das *Dalbergias* tanto em laboratórios como nos postos de fiscalização é uma das dificuldades encontradas pelos governos para controlar o tráfico ilegal desse gênero.¹² O método convencional utilizado para identificar espécies florestais

é a análise anatômica da madeira, tanto a nível macroscópico que pode ser feita em campo como a nível microscópico que é feita em laboratório. Algumas vezes, recorre-se ao auxílio de chaves eletrônicas e lentes de aumento para facilitar o processo de identificação. Porém, essa metodologia apresenta algumas desvantagens cruciais, pois demanda muito treinamento, prática constante de pessoal e alto investimento econômico. Idealmente, todo posto de fiscalização e toda aduana necessitaria da presença de um especialista em anatomia de madeira.

Tendo em vista a grande diversidade de espécies florestais no Brasil, essas dificuldades são ampliadas e de forma geral é feito somente uma triagem nos postos de fiscalização. Portanto, diversas metodologias estão sendo propostas e estudadas para facilitar ou auxiliar a identificação de espécies, como por exemplo: código de barra do DNA, isótopos estáveis, espectroscopia de massas e a espectroscopia no infravermelho próximo.

A Espectroscopia no Infravermelho Próximo (NIRS, do inglês *Near Infrared Spectroscopy*) vem sendo desenvolvida e adaptada para se adequar à identificação da madeira de espécies florestais em análises em campo. A *tecnologia NIRS* (associação da espectroscopia no infravermelho próximo e quimiometria) faz uso da radiação não destrutiva no infravermelho próximo e análises estatísticas multivariadas possibilitando a identificação correta de espécies produtoras de madeira semelhante. Braga *et al.*,¹³ Pastore *et al.*,¹⁴ e Soares *et al.*,¹⁵ utilizaram a espectroscopia no infravermelho próximo e o modelo de discriminação por Análise Discriminante por Mínimos Quadrados (PLS-DA, do inglês *Discriminant Analysis Partial Least Squares*) para identificar as espécies visualmente similares à *Swietenia macrophylla* King, sendo elas: *Carapa guianensis* Aubl.(andiroba), *Cedrela odorata* L. (cedro), *Erismia uncinatum* Warm. (cedrinho), *Micropholis melinoniana* Pierre (curupixá), *Hymenea coubaril* L. (jatobá). Adedipe *et al.*,¹⁶ utilizaram NIRS e a Modelagem Independente Flexível por Analogia de Classes (SIMCA, do inglês, *Soft Independent Modeling of Class Analogies*) para classificar as espécies de carvalho branco e vermelho.

A tecnologia NIRS, além de estar se mostrando uma ferramenta em potencial para identificação de espécies, também se mostra capaz de identificar a origem da madeira proveniente de diferentes países da América Latina utilizando ferramentas quimiométricas.^{17,18} Em 2016 Bergo *et al.*¹⁷ usaram equipamento NIRS de bancada e discriminaram 155 amostras de *Swietenia macrophylla* provenientes de 27 países diferentes das espécies *Carapa guianensis* (andiroba), *Cedrela odorata* (cedro) e *Micropholis melinoniana* (curupixá), concluindo que a variação espectral entre as espécies é maior do que a variação relativa ao país de origem. Dois anos depois o grupo de pesquisa avançou o estudo de identificação das origens das espécies florestais produtoras de madeira com aparelhos NIRS portáteis. Em 2018, Silva *et al.*¹⁸ avaliaram dois espectrômetros com faixa espectral de 1595 até 2.396 nm e de 950 até 1.650 nm, ambos portáteis, e alcançaram eficiência entre 90 e 100 % ao discriminar as origens de madeira dos países Brasil,

Bolívia, Guatemala, México e Peru com a ferramenta quimiométrica PLS-DA. Nesse mesmo trabalho a ferramenta SIMCA também foi avaliada e alcançou a eficiência entre 67 % e 100 %.

Tendo em vista a capacidade de identificação de espécies através de análises multivariadas em dados espectroscópicos da região do infravermelho, essa dissertação objetiva avaliar a hipótese de que as informações presentes no espectro NIR e o uso de ferramentas quimiométricas, tais como PLS-DA e SIMCA, são capazes de distinguir espectros da espécie *D. nigra* dos espectros de outras 14 espécies do gênero *Dalbergia*: *D. bariensis* Pierre, *D. cearensis* Ducke, *D. congestiflora* Pittier, *D. decipularis* Rizzini e A. Mattos, *D. glaucescens* Mart. Ex Benth., *D. latifolia* Roxb., *D. melanoxylon* Guill. E Perr., *D. miscolobium* Benth., *D. retusa* Hemsl., *D. sissoo* DC., *D. spruceana* Benth., *D. stevensonii* Standl., *D. tucurensis* Donn Sm. e *D. villosa* Benth. Para essa discriminação ainda pretende-se avaliar o desenvolvimento de modelos PLS-DA utilizando todos os comprimentos de onda disponíveis em um espectrômetro portátil ou pela seleção de comprimentos de onda com base nas principais bandas relacionadas aos principais constituintes da madeira.

Como a *D. spruceana* é uma espécie muito confundida com a *D. nigra* e ambas são de origem brasileira, se faz necessário desenvolver uma ferramenta que identifique ou auxilie na identificação dessa espécie. A segunda identificação objetiva avaliar a hipótese de que SIMCA e PLS-DA são capazes de discriminar a *D. spruceana* das demais espécies do gênero *Dalbergia*. Seguindo o mesmo procedimento utilizado para *D. Nigra*, foram construídos modelos PLS-DA e SIMCA avaliando duas regiões espectrais distintas.

Além da discriminação dessas duas espécies nativas do Brasil, essa dissertação objetiva avaliar a capacidade das duas ferramentas quimiométricas em distinguir outras espécies do mesmo gênero: *D. cearensis*, *D. decipularis*, *D. latifolia*, *D. nigra*, *D. retusa*, *D. sissoo*, *D. spruceana*, *D. stevensonii* e *D. tucurensis*. Foram feitos modelos individuais para essas madeiras por possuírem quantidade de amostras suficiente. As demais espécies estudadas: *D. bariensis*, *D. congestiflora*, *D. glaucescens*, *D. melanoxylon*, *D. miscolobium* e *D. villosa* foram usadas com o objetivo de avaliar a capacidade das ferramentas quimiométricas em discriminar espécies que não foram utilizadas na criação do modelo e, portanto, não foram obtidos modelos individuais para elas.

2. REVISÃO BIBLIOGRÁFICA

2.1 *Dalbergia nigra*

O gênero *Dalbergia*, pertencente à família *Leguminosae* que está inserida no grupo das angiospermas, possui 39 espécies só no Brasil, incluindo árvores, arbustos e lianas. *Dalbergia nigra*, a espécie mais ameaçada de extinção atualmente, é muito procurada no Brasil e no mundo devido às características de sua madeira que propiciam a produção de instrumentos musicais (Figura 2) e móveis de alta qualidade.^{1,2}



Figura 2. Violão feito de *D. nigra*. Foto de Gunesch.¹⁹

A madeira da espécie *D. nigra*, popularmente conhecida como jacarandá, jacarandá-da-Bahia, caviúna e jaracandá-preto é principalmente encontrada na mata Atlântica, possui o cerne e alburno distintos por diferença de coloração, anéis de crescimento perceptíveis, possui brilho chamativo nas superfícies longitudinais e odor característico.^{1,20}

Atualmente, o que restou dessa espécie é encontrado principalmente na região que se estende da região sul da Bahia até o norte do estado de São Paulo.¹ Por possuírem a madeira visualmente parecida com a de outras espécies do gênero *Dalbergia*, como a *D. spruceana*, as espécies são confundidas durante a identificação. As espécies *D. nigra* e *D. spruceana* não são facilmente identificadas nem mesmo através da anatomia devido à similaridade, ambas possuem grandes vasos, raios estriados e os parênquimas axiais são aliformes e em bandas apotraqueais.^{6,11}

D. nigra e *D. spruceana*, ambas brasileiras, além de características anatômicas semelhantes, são visualmente similares, levando a identificações errôneas em postos de fiscalização (Figura 3). A *D. spruceana* tem coloração alaranjada ou marrom avermelhada com cerne e alburno diferenciado, já a cor da *D. nigra* varia de um marrom escuro até marrom avermelhado tendo também cerne e alburno diferenciado.^{4,21}



Figura 3. Amostra da madeira da espécie *D. nigra* à esquerda e da madeira da espécie *D. spruceana* à direita (Acervo pessoal).

As duas espécies possuem dureza bastante semelhante, no teste Janka a *D. nigra* apresentou 420 N a mais do que a *D. spruceana*. O teste Janka consiste em medir a força necessária para afundar metade de uma bola de aço na madeira. O módulo de ruptura, que indica a pressão máxima que a madeira suporta antes da ruptura ao receber a força na direção tangencial, é maior na *D. nigra* (135,0 MPa) do que na *D. spruceana* (116,9 MPa). A *D. nigra* também suporta mais pressão na direção paralela (67,2 MPa) do que a *D. spruceana* (58 MPa). Todos esses valores descritos são para as madeiras com 12 % de umidade.^{4,21} Apesar das características serem semelhantes, observa-se a superioridade da *D. nigra* em suportar pesos.

Tabela 1. Características físicas da *D. nigra* e *D. spruceana*.

	<i>Dalbergia nigra</i>	<i>Dalbergia spruceana</i>
Densidade (kg/m³)	835	1.085
Dureza* (N)	12.410	11.990
Módulo de ruptura**(MPa)	135,0	116,9
Força de compressão*** (MPa)	67,2	58,0

*O valor foi obtido através do teste Janka; **Pressão máxima que a madeira suporta antes da ruptura quando a força é aplicada perpendicularmente ao corpo; ***Pressão máxima que a madeira suporta antes da ruptura quando a força é aplicada paralelamente ao corpo da madeira.^{22,23,24}

Como toda espécie incluída no grupo das angiospermas, a *D. nigra* possui raiz, caule, folha, flor, semente e fruto próprio de sua espécie. Todas essas estruturas podem ser utilizadas para identificar a espécie, porém elas não acompanham a madeira durante o transporte, o que dificulta

o procedimento de identificação, pois apenas a madeira está disponível para ser analisada e para essa identificação é necessário um expert em anatomia de madeira.

A madeira da *D. nigra* possui vasos visíveis a olho nu com diâmetros que variam de 100 a mais de 200 μm e não possuem padrão de disposição definido, o parênquima axial e os raios são observados somente com o uso de lente 10x.²⁰ Estudos mostram que os raios são pouco frequentes, entre 5 e 10 por mm^2 segundo Coradin *et al.*²⁰ e em média possuem 2,8 vasos por mm^2 segundo Gasson *et al.*²⁵

Apesar do vasto conhecimento das características anatômicas da *D. nigra* a anatomia da madeira não é suficiente para separá-la de outras espécies do mesmo gênero. No trabalho de Gasson *et al.*²⁵ foram identificadas outras 5 espécies de *Dalbergia* que possuem frequência de vasos muito próximas: *D. miscolobium* (4,5), *D. spruceana* (3,5), *D. tucurensis* (2,8), *D. granadillo* (3,8), *D. stevensonii* (3,1). O diâmetro dos vasos também não é uma característica que é facilmente utilizada para identificar a *D. nigra*, pois essas mesmas 5 espécies citadas anteriormente possuem valores próximos aos valores da *D. nigra*.²⁵ Adicionalmente, os raios dessas espécies também não variam significativamente para que se possa diferenciar a *D. nigra* das demais (Figura 4).^{6,25} Somente a *D. cearensis* pode ser diferenciada, pois possui pequenos vasos e com grande frequência (Figura 4).²⁵

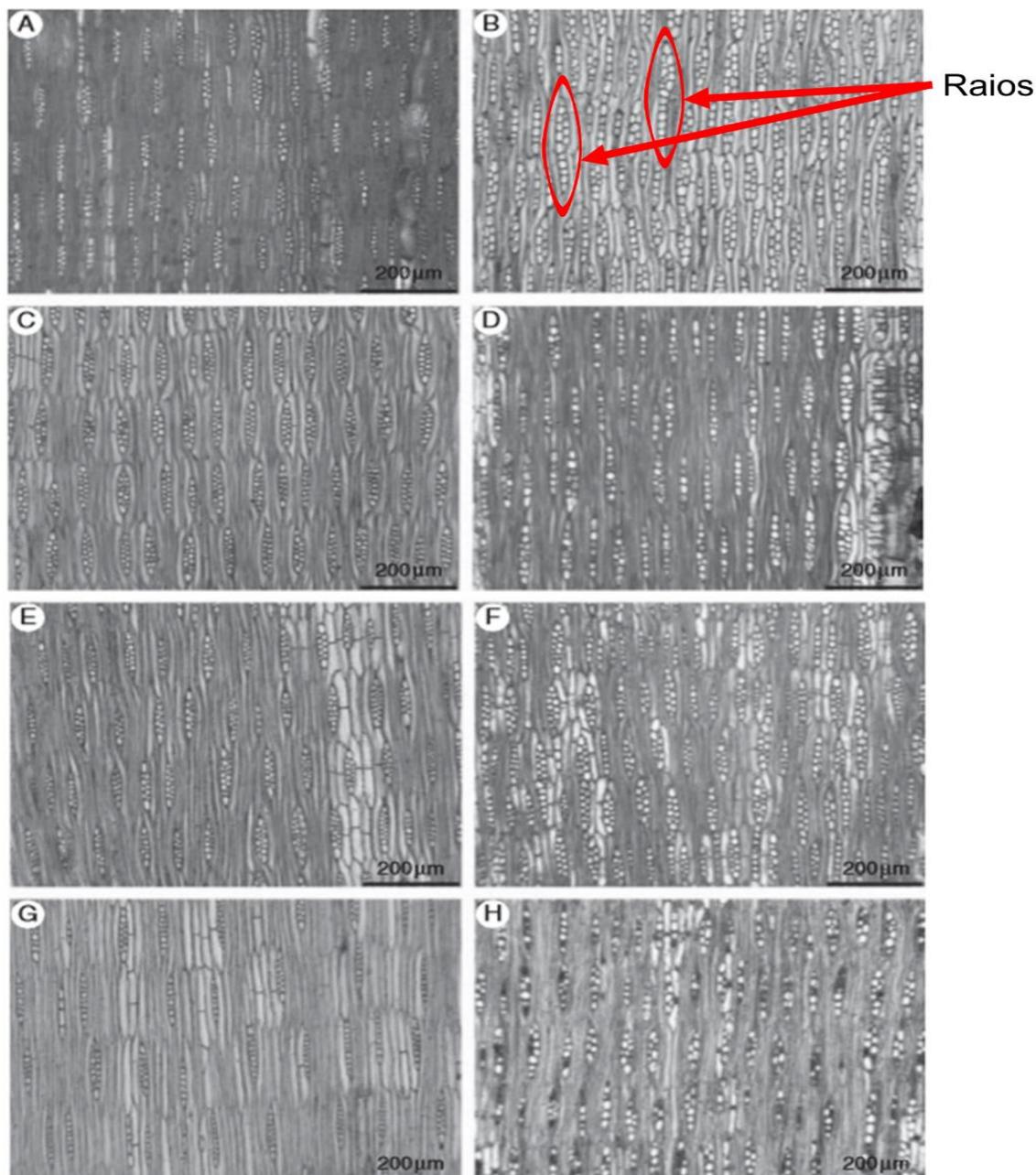


Figura 4. Corte tangencial longitudinal com aumento de 100x: (A) *D. cearensis*; (B) *D. miscolobium*; (C) *D. nigra*; (D) *D. spruceana*; (E) *D. tucurensis*; (F) *D. granadillo*; (G) *D. stevensonii*; (H) *D. nigra*. Fonte: Gasson *et al.*²⁵

A espécie *D. decipularis* possui uma característica importante que a distingue de *D. nigra*, apresenta poros em seus anéis enquanto a *D. nigra* não os possui.⁶

Em 2011 Gasson *et al.*²⁶ fizeram uma extensa análise de características anatômicas qualitativas de diversas espécies, incluindo três *Dalbergias* (*D. nigra*, *D. retusa* e *D. stevensonii*). Os autores concluíram que entre as três espécies de *Dalbergia*, as características semelhantes totalizavam 33.

Perante a esta quantidade de semelhanças visuais e anatômicas a dificuldade em identificar espécies de *Dalbergia*, tanto em postos de fiscalização como em laboratório, torna-se bastante evidente. A Figura 5 ilustra as espécies estudadas nesse trabalho.



Figura 5. Fotos das madeiras do gênero *Dalbergia* estudadas nesse trabalho.^{27,28,29,30}

Os trabalhos de anatomia da madeira mostram que é necessário investir em outras tecnologias para poder distinguir as espécies de *Dalbergia* entre si. Por esta razão, propõe-se a análise por meio da tecnologia NIRS para verificar se é possível separar as espécies deste gênero, com anatomia da madeira tão similar entre si.

Não apenas a comunidade internacional, como a CITES, reconhece o risco de extinção da *D. nigra*, através da portaria nº 37-Nde 1992, o IBAMA reconheceu *D. nigra* como uma espécie ameaçada de extinção e através da portaria nº 83 de 1996, o Brasil assumiu o compromisso com a CITES de proibir a exportação de *D. nigra* enquanto ela estiver no Anexo I.^{32,33}

A *D. spruceana* é encontrada na floresta Amazônia enquanto a *D. nigra*, na floresta Atlântica. Sendo assim, conhecendo a origem é possível discriminar essas duas espécies.^{6,34}

2.2 Métodos de discriminação de *Dalbergia*

Pigozzo em 2011³⁵ indicou que NIRS e quimiometria são ferramentas em potencial para discriminar *D. nigra* da *D. spruceana*. Os espectros das 28 amostras foram obtidos por reflexão difusa por sonda de fibra óptica e os dados espectroscópicos foram avaliados através da Análise de Componente Principal (PCA, do inglês, *Principal Component Analysis*) e Modelagem Independente Flexível por Analogia de Classes (SIMCA).

Outras metodologias sugeridas pelos especialistas para separar as duas espécies de *Dalbergias* brasileiras utilizam a fluorescência¹⁴, densidade básica¹⁴ e análises de compostos específicos através de Espectrometria de Massa por Análise Direta em Tempo Real – Tempo de Voo (DART-TOF-MS, do inglês, *Direct Analysis Real Time – Time of Flight – Mass Spectrometry*).^{11,36}

A análise de fluorescência possui uma desvantagem crítica: a presença de compostos que emitem fluorescência podem ocorrer ou não em uma mesma espécie determinada, portanto a fluorescência não pode ser considerada uma característica ampla da madeira e que pode ser utilizada para a identificação de muitas espécies. Além disso, exige extração de material e é necessário que a amostra seja levada ao laboratório. Outras duas espécies de *Dalbergia* (*D. granadillo* e *D. stevensonii*) apresentam características fluorescentes semelhantes à *D. nigra*³⁴, o que pode tornar a identificação dessa espécie mais complexa.

A análise de densidade básica é o método mais simples para diferenciar *D. nigra* e *D. spruceana*. A *D. nigra* possui densidade média de 0,85 g/cm³ e a *D. spruceana* de 1,02 g/cm³.⁶ Apesar da facilidade de diferenciação, esse método só é aplicável caso em que a amostra analisada seja ou *D. nigra* ou *D. spruceana*. No entanto, de acordo com a Norma NBr 7190,³⁷ para espécies não cultivadas existe incerteza de 10 % associada à medida de densidade básica. Como as densidades básicas das duas espécies são muito próximas a análise de discriminação pode levar a resultados não verdadeiros.

A Espectrometria de massa por Análise Direta em Tempo Real – Tempo de Voo (DART-TOF-MS, do inglês, Direct Analysis Real Time – Time of Flight – Mass Spectrometry), é um método analítico que demonstrou bons resultados na identificação de *D. nigra* e *D. spruceana*^{11,36}, porém também necessita do envio de material para o laboratório e o equipamento é de alto custo.

2.3 Comércio madeireiro.

No Brasil, o produto madeireiro mais exportado é a celulose. Só em 2016 mais de 13 milhões de toneladas foram exportados e de janeiro até setembro de 2017 foram exportados 10,5 milhões de toneladas. A celulose apresenta um histórico crescente de exportação como pode ser observado na Figura 6.³⁸

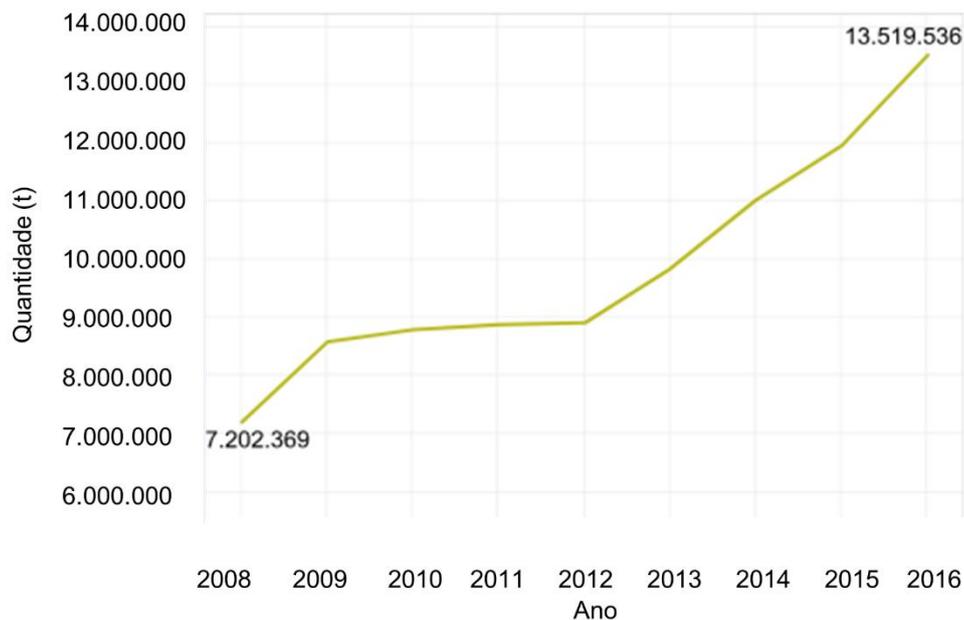


Figura 6. Quantidade de celulose brasileira exportada em toneladas.

Em 2016 os maiores exportadores, em valores totais, de produtos madeireiros brasileiros foram os Estados Unidos, China, Países Baixos, Itália e Argentina. Sendo que o estado do Paraná foi o maior exportador, totalizando mais de um bilhão de dólares em vendas em 2016.³⁸

Em uma perspectiva global, os Estados Unidos, Japão e União Europeia vinham sendo os maiores importadores de produtos de madeira tropical. Nos últimos anos, a China e Índia se tornaram os dois principais compradores de toras de madeira tropical. O aumento de importação da China e Índia é atribuído a três fatores:

- Rápido crescimento econômico de ambos países aumentaram a demanda por madeira, em especial madeira tropical devido aos valores sociais associados a ela;³⁹

A China ser uma economia que exporta produtos madeireiros secundários e é um país com limitações de recursos madeireiros, o país depende da importação de madeira para exportar;⁴⁰

Comerciantes preferem negociar com esses países, pois as legislações de importação são menos restritivas e praticam menores valores em relação aos Estados Unidos e à União Europeia.⁴¹

O termo “legal” tem diferentes significados e implicações de acordo com o sistema jurídico do país. No Brasil, para que a madeira seja considerada legal ela tem que ter tido autorização de corte pelo órgão responsável. Importante ressaltar que não necessariamente a madeira deve ter certificação para ser considerada legal. A certificação é um documento não obrigatório que algumas empresas buscam para provar que seu produto e o processo de produção seguem padrões de qualidade e sustentabilidade. Em 2002 criou-se o CERFLOR (Programa Brasileiro de Certificação Florestal) que é coordenado pelo INMETRO e em 2017 8,1 % das certificações mundiais eram brasileiras.³⁸

O desmatamento e comércio ilegal de madeira são problemas sérios que acontecem em todo globo terrestre, envolvendo muitos países principalmente os detentores das florestas tropicais, que são os maiores produtores de madeira nativa.¹² As espécies mais atingidas são, em sua maioria, espécies raras e em risco de extinção. A raridade da espécie eleva o valor da madeira incentivando os comerciantes ilegais a desflorestarem essas espécies, e quanto mais se corta dessa espécie mais rara ela se torna criando um ciclo vicioso ilustrado na Figura 7.⁴²

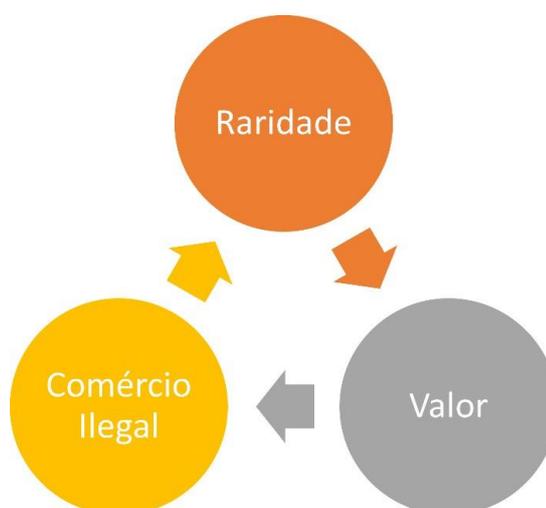


Figura 7. Esquema do ciclo vicioso entre raridade de espécie produtora de madeira, valor e comércio ilegal (Gan *et al.*⁴²).

O desmatamento irresponsável leva à extinção de espécies florestais não afetando apenas uma espécie específica ou gênero, mas sim todo o ecossistema. Expansão de áreas desflorestadas e destruição de comunidades locais são exemplos do efeito do comércio ilegal de madeira.¹²

Igualmente, o desmatamento e comércio ilegal afeta a economia mundial. O corte de madeira sem autorização da instituição responsável, sem atender as normas de segurança dos trabalhadores, não pagando todas as taxas e omitir ou até mesmo informar de forma incorreta a origem da madeira torna o seu custo significativamente menor. Essa redução de custo torna o mercado ilegal muito atrativo. O crime ambiental e econômico diminui a arrecadação governamental afetando toda a economia global. Na Indonésia, uma concessionária florestal gasta cerca de 85 US\$/m³ para entregar madeira de forma legal enquanto no mercado ilegal é gasto apenas 32 US\$/m³.⁴³ Na China, o mercado de instrumentos e móveis é alimentado por madeiras Dalbergia e como as reservas da China e de países ao redor se extinguiram, os chineses as importam da região Sava em Madagascar. O produto final feito de Dalbergia é vendido por milhares de dólares no comércio chinês enquanto os trabalhadores de Sava recebem apenas cerca de 5 US\$/dia.⁴⁴

No documento publicado pelo Escritório das Nações Unidas em Drogas e Crimes, (UNDOC, do inglês, United Nations Office on Drugs and Crime) 2016, afirma-se que a principal barreira no controle de espécies florestais é a dificuldade em identificar as espécies ainda nos postos de fiscalização. Portanto, é de extrema importância que se estude metodologias capazes de realizar ou auxiliar essa identificação ainda em condições de campo ou triagem.¹²

2.4 Métodos de identificação de espécie florestal.

Devido a demanda internacional de controle de tráfico de madeira, diversos métodos para determinação de espécie florestais estão sendo desenvolvidos para auxiliar o método convencional de identificação anatômica, como por exemplo: espectrometria de massas, isótopos estáveis, radiocarbono, DNA e espectroscopia no infravermelho próximo.

2.4.1 Anatomia da madeira.

A análise anatômica da madeira é feita tanto em nível microscópico como macroscópico de forma que vários elementos são investigados e comparados às descrições para que haja a identificação. Por exemplo: parênquima axial (difuso ou difuso em agregado) e visibilidade dos

poros a olho nu ou não, entre outros. Apesar de ser o método convencional, a anatomia da madeira nem sempre possibilita a identificação de algumas amostras a nível de espécie.⁴⁵ Como grande parte das madeiras examinadas em postos de fiscalização não possuem mais algumas estruturas que auxiliam na identificação como flores e sementes, o processo de identificação é ainda mais dificultado.⁴⁶ É necessária mão de obra qualificada e altamente treinada de especialistas em anatomia da madeira para realizar esse tipo de identificação. O estudo sobre as 16 características anatômicas analisadas através da Análise de Componentes Principais (PCA, do inglês, *Principal Component Analysis*) de sete espécies de *Dalbergia* feito por Gasson *et al.*²⁵ indicou que é apenas possível distinguir essas sete espécies em dois grupos, sendo o primeiro composto por *D. nigra*, *D. spruceana*, *D. tucurensis* Donn. Sm. e *D. stevensonii* Standl. E o segundo composto por *D. miscolobium* Benth, *D. cearenses* Ducke e *D. granadillo* Pittier. Ao analisar a densidade e sabendo o país de origem é possível fazer a identificação exata entre as espécies do grupo 1. As madeiras das espécies *D. nigra* e *D. spruceana* de origem brasileira, em alguns casos, podem ser diferenciadas de acordo com a densidade, assim como *D. tucurensis* e *D. stevensonii* que são da América Latina.^{20,46,47} Gasson *et al.*²⁵ afirmam que a *D. cearensis* é facilmente identificável devido a sua frequência de vasos, pois é a única que possui mais de dez vasos por milímetros quadrados.^{20,48}

2.4.2 Espectroscopia de massa.

A espectrometria de massa identifica a madeira principalmente de acordo com os extrativos presentes em cada espécie. Esses extrativos são ionizados e os íons gerados compõem o perfil químico de cada espécie. A dalnigrina, por exemplo, é um composto flavonóide que está presente apenas na espécie *D. nigra* e foi identificado pela primeira vez através da técnica de espectroscopia de massa por Kite *et al.*,³⁶. A espectrometria de massa convencional demanda longo tempo de análise devido ao preparo de amostra.⁴⁹ Devido a uma nova técnica, chamada Espectrometria de Massa por Análise Direta em Tempo Real – Tempo de Voo (DART-TOF-MS, do inglês, Direct Analysis Real Time – Time of Flight – Mass Spectrometry), que permite uma análise mais rápida devido a sua facilidade em ionizar a amostra, a utilização do método tem aumentado e facilitado as análises de madeira.^{49,50}

Lancaster e Espinoza¹¹ utilizaram a espectrometria de massa para discriminar 13 espécies de *Dalbergia*, mas o flavonóide dalnigrina não foi utilizado para a diferenciação entre *D. nigra* e *D. spruceana*, pois não foi possível distinguir dalnigrina e seu isômero kuhlmannina, ambos com razão massa carga de 299,0914. Mesmo não se baseando no composto dalnigrina os autores conseguiram discriminar essas duas espécies. Os íons predominantes da espécie *D.*

spruceana são Dalbergione I com razão massa carga de 255,10 e um composto não identificado com razão massa carga 209,12, enquanto os íons predominantes da espécie *D. nigra* são: caviúna com razão massa e carga de 375,11 e o íon dalbergina (m/z 269,08). A espécie *Machaerium scleroxylon*, visualmente semelhante a *D. nigra*⁵¹, conhecida como jacarandá e de origem na América do Sul foi diferenciada da madeira da espécie *D. nigra* devido à especificidade do DART-TOF que identificou o íon Melanettina ($m/z = 285,0758$) específico da *D. nigra* e o íon dimethoxidalbergione ($m/z = 285,1126$) específico da *Machaerium scleroxylon*. Nesse mesmo estudo Lancaster e Espinoza¹¹ almejavam discriminar 7 espécies de *Dalbergia*: *D. retusa*, *D. cearenses*, *D. melanoxylon*, *D. decipularis*, *D. stevensonii*, *D. madagascariensis* e *D. latifolia*. Utilizando validação cruzada da Análise Discriminante Linear (LDA, do inglês, Linear Discrimination Analysis) obtiveram 91,72 % de classificação correta indicando a capacidade do DART-TOF em discriminar essas espécies. Musah *et al.*⁵² ao utilizarem a análise discriminante de Kernel com informações do espectro de massas de 107,037 -527,155 razão massa carga notaram que *D. granadillo* e *D. retusa* possuem perfil químico similar.

2.4.3 Isótopos estáveis

Vários elementos estão presentes na constituição celular de uma árvore, como por exemplo o carbono, hidrogênio, oxigênio e nitrogênio, que são fundamentais para o crescimento, distribuição e ciclos biogeoquímicos das plantas.⁵³ Esses elementos, que podem possuir isótopos, são fornecidos pelo ambiente onde a árvore está localizada. Então, uma relação entre os isótopos estáveis e o ambiente de crescimento da planta é criado, e partir dessa relação é possível prever a origem do objeto de estudo. O teor de ¹⁸O na água que hidrata a planta possui informações geográficas de valor, como temperatura da formação da gota da chuva e localização da condensação da nuvem.⁵⁴ A relação de ¹⁸O e ¹⁶O nos tecidos das plantas, por exemplo, é utilizado para a identificação da origem da planta.⁵⁵ O somatório das informações providas pelo teor de ¹⁸O e de outros isótopos forma a impressão digital de uma área que pode ser relacionada à uma madeira de origem desconhecida, e então, identificar a origem dessa madeira. Essa técnica permite a utilização de qualquer parte da madeira da árvore para a análise, porém além de ser uma técnica que exige o encaminhamento da amostra ao laboratório, é necessário um extenso estudo de diversos ambientes para que se possa comparar com o padrão isotópico das amostras de madeira.

2.4.4 Radiocarbono

Assim como o método de isótopos, o método de radiocarbono utiliza as condições ambientais, porém, na técnica de radiocarbono a relação estudada é somente dos isótopos de carbono. Entretanto, não é possível obter uma relação entre a amostra e sua origem, pois a relação entre os isótopos de carbono é aproximadamente constante em todo o globo.⁵⁶ O chamado “pico da bomba” é um importante evento na área de datação por radiocarbono, pois alterou a constância da relação entre os isótopos de carbono: a relação $^{14}\text{C}/^{12}\text{C}$ na atmosfera foi elevada devido a testes e bombas nucleares.⁵⁷ A elevação repentina do nível de ^{14}C na atmosfera oportuniza a datação de materiais mais recentes.¹² O isótopo radioativo, o ^{14}C , é produzido principalmente na camada superior da atmosfera. A meia-vida desse isótopo não é bem conhecida: 5.568 ± 30 anos segundo Anderson e Libby,⁵⁸ 5.730 ± 40 anos de acordo com Godwin,⁵⁹ 5.660 ± 30 anos segundo estudos de Bella *et al.*,⁶⁰ estudos mais recentes de Chiu *et al.*,⁶¹ indicam que a meia vida do ^{14}C pode ser próxima de 6.000 anos.

A equação 1 representa a reação química que origina o ^{14}C .



Onde, n é nêutron que reage com nitrogênio resultando em ^{14}C e a liberação de um próton, representado como p .⁵⁶ O carbono ao ser oxidado a CO_2 entra no ciclo do carbono atingindo oceanos, atmosfera e biosfera. O CO_2 , que é incorporado à planta através da fotossíntese, está na forma de C^{14} e C^{12} . Apesar do decaimento do C^{14} para N^{14} (equação 2) a relação $\text{C}^{14}/\text{C}^{12}$ permanece em equilíbrio enquanto o organismo está vivo, pois constantemente há incorporação de mais C^{14} no organismo devido à fotossíntese. A partir do momento que não ocorre mais a fotossíntese, organismo morto, não há renovação do C^{14} . Devido ao decaimento da relação $\text{C}^{14}/\text{C}^{12}$ com o passar do tempo, é possível estimar o momento de corte das plantas.⁶² Ao analisar o teor de C^{14} e C^{12} de uma amostra orgânica e comparar à uma curva padrão estima-se a idade do material orgânico. A técnica de radiocarbono é importante ao determinar se a amostra de árvore em questão foi cortada antes ou depois da implementação de alguma lei.¹² Apesar de ser uma técnica efetiva ao determinar a data aproximada de corte, é uma técnica que exige partes específicas da madeira. A técnica tem como preferência a utilização do incremento do core de forma que cubra de 5 a 20 anéis ou a utilização de dois blocos: um da superfície mais externa e outro dos anéis entre 5 e 20.¹²

2.4.5 Código de barra de DNA

A técnica de código de barra de DNA consiste em identificar um indivíduo de acordo com o código de barras do seu DNA. Esse código de barra é um sequenciamento de regiões específicas do DNA. A região selecionada para ser feito o sequenciamento depende do tipo de indivíduo se do reino animal ou vegetal.⁶³ Em 2009, foi recomendado para plantas, pelo Consórcio pelo Código de Barra da Vida - Grupo de Trabalho em Plantas (CBOL, do inglês, Consortium for the Barcode of Life Plant Working Group), a utilização de determinadas regiões, do genoma, para o sequenciamento.⁶⁴ Entretanto, nem todas as espécies obtêm a maior eficiência de identificação baseando-se nas regiões indicadas pelo CBOL. A região ITS, por exemplo, foi utilizada na identificação de espécies da família Mahogany.⁶⁵ Porém, com intuito de identificar *Aquilaria spp.* utilizou-se outras regiões.⁶⁶ Recentemente, Yu *et al.*,⁶⁷ determinaram as melhores regiões para fazer a análise de amostras de *Dalbergia*. Essa técnica tem como facilidade a possibilidade de utilizar qualquer pedaço da madeira para a análise, porém ainda não há concordância com quais regiões do DNA devem ser analisadas, é necessário que se chegue a um consenso sobre a melhor região para cada tipo de identificação, além disso é necessário o envio da amostra para laboratório

2.4.6 Tecnologia NIRS

O Laboratório de Produtos Florestais do Serviço Florestal Brasileiro (LPF/SFB) em colaboração com o Instituto de Química da UnB foi pioneiro em estudos que combinam NIRS e quimiometria visando a identificação de espécies florestais tropicais nativas, para evitar a exploração ilegal de madeiras. Braga *et al.*¹³ e Pastore *et al.*¹⁴ descreveram um método utilizando PLS-DA que discrimina *Swietenia macrophylla* das espécies: *Carapa guianensis* Aubl., *Cedrela odorata* L. e *Micropholis melinoniana* Pierre. Em 2017, Soares *et al.*¹⁵ adicionaram as espécies *Erismia uncinatum* (cedrinho) e *Hymenaea courbaril* (jatobá) ao modelo previamente descrito. Todas essas espécies estudadas são facilmente confundidas entre si visualmente de acordo com Coradin *et al.*²⁰

Além da discriminação das espécies, Bergo *et al.*,¹⁷ e Silva *et al.*¹⁸ demonstraram que a utilização de NIRS e quimiometria também tem o potencial de identificar a origem da madeira de países diferentes da América Latina.

Pigozzo utilizou NIRS e quimiometria para distinguir as espécies *D. nigra* e *D. spruceana*. O estudo foi realizado com apenas 28 amostras, sendo 16 de *D. nigra* e 12 de *D. spruceana*. Através da análise de PCA observou-se que a primeira PC, que representa 71 % dos dados, foi o bastante para distinguir um grupo de 10 amostras de *D. nigra* de um grupo de 10

amostras de *D. spruceana*. Na análise SIMCA das 6 amostras da validação de *D. nigra* apenas uma amostra não foi corretamente identificada.³⁵

Mesmo com várias pesquisas utilizando o NIRS e a análise multivariada, não há estudo que utilize um número representativo de amostras e, portanto, modelos mais robustos e confiáveis, com essa tecnologia para identificar espécies do gênero *Dalbergia*.

3. FUNDAMENTAÇÃO TEÓRICA

3.1 Componentes químicos da madeira

Os componentes majoritários da madeira são: celulose, hemicelulose e lignina. Essas macromoléculas seguem uma proporção aproximada de, respectivamente, 50:25:25. Além dos polímeros citados, a madeira contém extrativos que são compostos orgânicos com massa molecular bem menor e com grande variedade de grupos funcionais.⁶⁸

A celulose, polissacarídeo mais abundante no mundo, é uma substância resistente e insolúvel em água que está presente nas paredes das células vegetais. É o polissacarídeo responsável por estruturar células vegetais e possui entre 10 e 15 mil unidades do seu monossacarídeo constituinte, D-glicose, unidas por uma ligação glicosídica entre os carbonos 1 e 4. A união de monômeros de glicose pode ocorrer através de ligações: tipo α , comum no amido e glicogênio, ou tipo β como é o caso da celulose (Figura 8).

Na ligação β o átomo de oxigênio que liga os monômeros, o grupo CH_2OH ligado ao carbono 5 da cadeia do monômero 1 e o grupo hidroxila do carbono 2 do segundo monômero se encontram todos na posição equatorial. Sendo assim, ocorre uma ligação de hidrogênio entre as hidroxilas (representada por traços na Figura 8) estabilizando a forma linear da estrutura. A estrutura também é estabilizada pela ligação de hidrogênio entre a hidroxila do C3 do primeiro monômero e o oxigênio presente no anel do monômero à esquerda. A forma linear permite que os segmentos dos polímeros se posicionem paralelamente e sejam interligados através das ligações de hidrogênio gerando um composto compacto e rígido capaz de estruturar a célula vegetal.⁶⁹

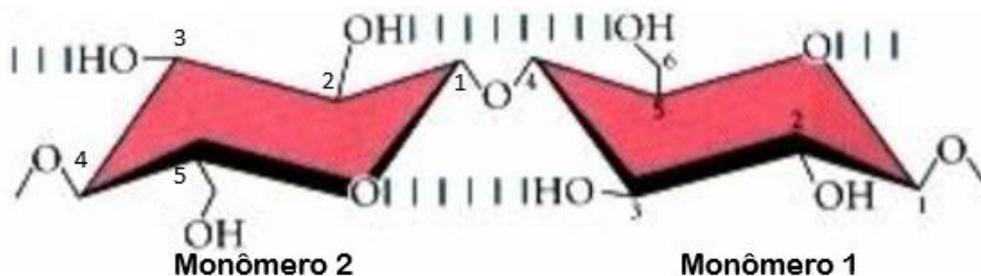


Figura 8. Representação da ligação glicosídica ($\beta 1 \rightarrow 4$) presente na estrutura da celulose.

A hemicelulose é uma classe de polímeros constituídos por monômeros de diversos açúcares presentes na madeira como ilustrado na Figura 9. Também se difere da celulose por apresentar cadeias significativamente menores e presença de ramificações. Dentro dessa classe

há duas subdivisões: homopolímero na qual a cadeia principal só apresenta um tipo de unidade e heteropolímero que a cadeia principal possui diversos tipos de monômeros.

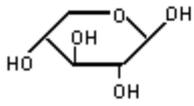
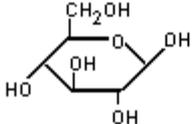
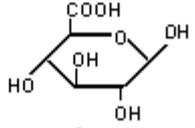
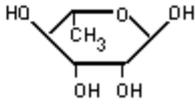
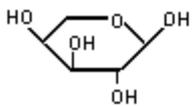
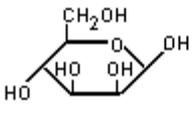
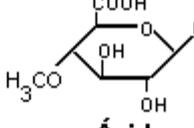
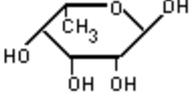
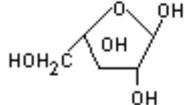
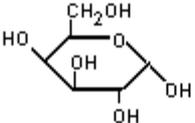
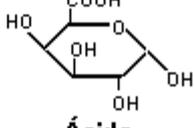
Pentoses	Hexoses	Ácidos Hexurônicos	Deoxi-hexoses
 <p>β-D-Xilose</p>	 <p>β-D-Glicose</p>	 <p>Ácido β-D-Glicurônico</p>	 <p>α-L-Ramnose</p>
 <p>α-L-Arabinopiranoose</p>	 <p>β-D-Manose</p>	 <p>Ácido α.D.O Metilglicurônico</p>	 <p>α-L-Fucose</p>
 <p>α.L.Arabinofuranose</p>	 <p>α.D.Galactose</p>	 <p>Ácido α.D.Galacturônico</p>	

Figura 9. Principais açúcares presentes na hemicelulose.

Assim como a celulose e hemicelulose, as ligninas são polímeros. Suas estruturas tridimensionais aromáticas são compostas por três substâncias: álcool p-cumarílico, álcool coniferilílico e o álcool sinapílico.^{68,70} Esses compostos estão ilustrados na Figura 10.

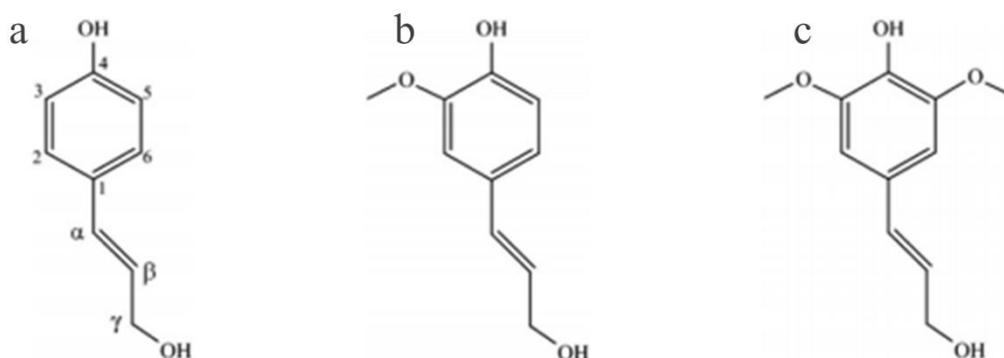


Figura 10. Estrutura dos três componentes básicos das ligninas: (a) álcool p-cumarílico; (b) álcool coniferilílico e (c) álcool sinapílico.

As porcentagens desses componentes básicos variam de acordo com a espécie e foram estimadas por Azadi *et al.*⁷⁰ de acordo com a classificação da árvore que produziu a madeira:

gimnosperma ou angiosperma. As madeiras das gimnosperma ou coníferas (*softwood*) possuem entre 90 e 95 % de álcool coniferilífico e entre 5 e 10 % de álcool sinapílico; as madeiras das angiospermas ou folhosas (*hardwood*) tem metade de álcool coniferilífico na sua composição, sendo a outra metade de álcool sinapílico. A Figura 11 representa a estrutura da lignina da angiosperma *Fagus sylvatica* que foi proposta por Nimz.⁷¹

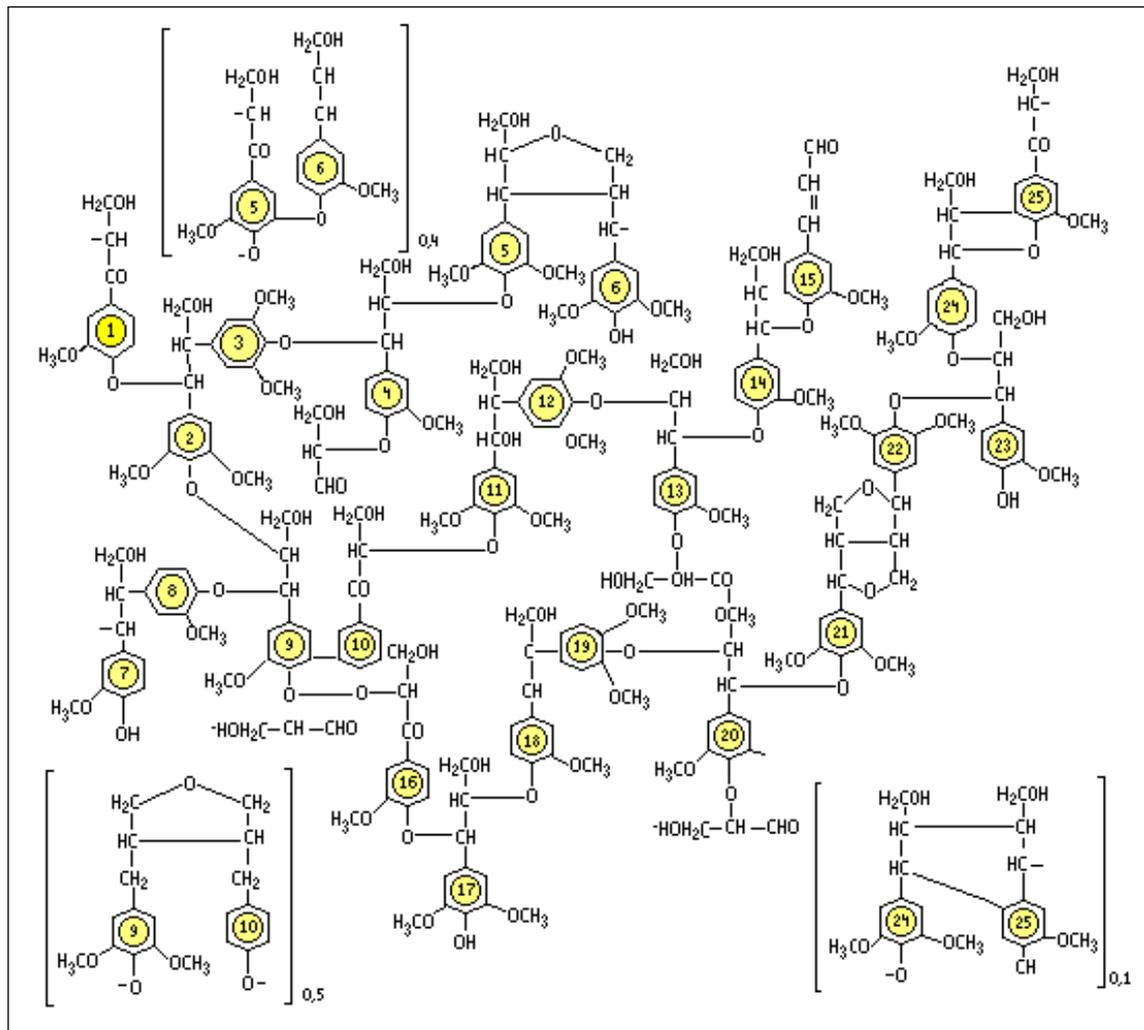


Figura 11. Esquema estrutural da lignina da angiosperma *Fagus sylvatica* proposto por Nimz.⁷¹

A caracterização da lignina é um processo complexo, pois há uma grande diversidade desses compostos na madeira e ocorrem variações estruturais na lignina de acordo com o método de isolamento empregado.⁷⁰

Os compostos denominados extrativos são componentes orgânicos minoritários da madeira que são extraídos com solventes neutros - etanol, acetona e diclorometano – e/ou mistura de solventes, podendo fazer sucessivas extrações. Os extrativos são obtidos de acordo

com sua interação química com solvente. Por exemplo: óleos, ceras, resinas são extraídos utilizando éteres; os componentes que colorem a madeira são extraídos com álcool. Esses componentes, apesar de contribuírem pouco no peso final, são muito influentes nas propriedades da madeira.⁷² A distinção visual entre cerne, região onde geralmente se encontra os extrativos, e alburno da madeira é influenciado por componentes que colorem a madeira, como os flavonoides (Figura 12).



Figura 12. Tronco da madeira *Dalbergia Nigra* com perceptível variação visual entre alburno e cerne.

3.2 Espectroscopia no infravermelho próximo

A espectroscopia no infravermelho próximo (NIRS, do inglês, Near Infrared Spectroscopy) é uma técnica espectroscópica que é baseada na interação entre as ondas eletromagnéticas, na região do infravermelho próximo, e as ligações químicas dos compostos constituintes do material de interesse. O espectrômetro é constituído, de maneira simplificada por uma fonte de radiação, sistema para seleção do comprimento de onda, detector e sistema de análise de dados. As fontes de radiação mais comuns são as lâmpadas de tungstênio ou halogênio, por conta da sua alta e constante intensidade de radiação na faixa do infravermelho próximo e serem lâmpadas de baixo custo. Os detectores mais comuns são baseados em arseneto de gálio e índio (InGaAs) ou seleneto de chumbo (PbS). Por outro lado, os seletores de comprimento de onda podem variar consideravelmente dependendo da configuração do equipamento.^{73,74}

Na análise de amostras sólidas, a técnica se baseia na absorção parcial da radiação emitida pela lâmpada pelos compostos presentes na superfície da amostra seguida da reflexão espalhada em todas as direções da radiação não absorvida, fenômeno conhecido como reflectância difusa. Esses compostos absorvem comprimentos de onda específicos de acordo com suas estruturas e os movimentos dos seus átomos. Então, ao receber a radiação do instrumento o uma molécula que possui vibração molecular que possua o mesmo valor energético que a radiação emitida pela lâmpada absorve essa radiação. Sendo assim, a radiação refletida de volta ao instrumento possuirá menor intensidade nos comprimentos de onda que foram absorvidos devido aos movimentos das moléculas, gerando assim um espectro característico da estrutura química da amostra.^{75,76}

Atualmente há diversos equipamentos portáteis disponíveis para compra. Mais detalhes sobre esses equipamentos podem ser encontrados em Tibola *et al.*⁷⁷

A região do infravermelho próximo (780 nm – 2.500 nm) contém muitas informações de sobretons e bandas de combinação. Sobretons são bandas correspondentes a múltiplos das vibrações fundamentais observadas na região do infravermelho médio e as bandas de combinação são combinações de bandas fundamentais e de múltiplos dessas combinações.⁷⁸ Apenas no infravermelho próximo o movimento de estiramento da ligação C-H apresenta quatro bandas de sobretons e quatro bandas de combinações. A região de 1.000 a 1.486 nm corresponde a diversos sobretons e combinações como por exemplo o primeiro e segundo sobretom da ligação O-H, primeiro sobretom das bandas de combinação de ligações C-H, segundo e terceiro sobretom do estiramento da ligação C-H e N-H. A região de 1.486 a 1.667 nm é preenchida principalmente pelo primeiro sobretom do estiramento da ligação O-H. Já a região de 1.667 a 2.000 nm representa o primeiro sobretom do estiramento alifático e aromático da ligação C-H e combinação da ligação OH. A região próxima a 2.100 nm representa a combinação da extensão das ligações O-H e C-H.^{79,80}

Como visto, o infravermelho próximo possui muita informação valiosa, porém devido às diversas sobreposições das bandas de diversas ligações ou grupos químicos diferentes, analisar apenas o espectro e tentar relacionar os picos aos compostos é uma tarefa bastante difícil em amostras contendo muitos constituintes, que quase sempre não é possível. Devido a isso, por muito tempo a região do infravermelho próximo foi deixada de lado pelos químicos analistas até o advento dos computadores.⁷⁴ Além das valiosas informações contidas na região do infravermelho próximo a técnica NIRS é uma técnica muito visada por ser muito prática e com vários benefícios. A técnica NIRS demanda nenhuma ou pouca preparação de amostra, possui rápida resposta, além disso, é uma técnica não destrutiva e não invasiva.^{74,81,82} De acordo

com Pasquini,⁷⁴ como NIRS é capaz de analisar amostras de forma bastante rápida e simples, logo, pode-se coletar uma imensidão de dados rapidamente. Devido à imensidão de amostras e quão rico de informações o espectro na região do infravermelho próximo é, utiliza-se a quimiometria para interpretar essas informações. A quimiometria é aplicação de métodos estatísticos e matemáticos em dados de fonte química. A combinação de NIRS e quimiometria, denominada tecnologia NIRS, é uma ferramenta que é utilizada em diversas áreas do conhecimento, como farmacêutica, alimentícia, agrônômica, etc.⁸³⁻⁸⁵

3.3 Modelos Quimiométricos

3.3.1 Seleção de amostras e exclusão de *outliers*

Para o desenvolvimento dos modelos de discriminação ou classificação as amostras de todas as espécies foram divididas em dois grupos: treinamento e validação. Como o nome indica, as amostras de treinamento, geralmente constituídas por dois terços do número total de amostras, foram utilizadas para treinar o modelo. Já as amostras de validação, geralmente compostas por um terço do número total de amostras, foram utilizadas para testar o modelo de discriminação e calcular as figuras de mérito da análise. A seleção das amostras de treinamento e de validação pode ser realizada por diferentes métodos, sendo um dos algoritmos descritos na literatura o Duplex.⁸⁶

O algoritmo Duplex é baseado na medida da distância euclidiana entre as amostras e uma sequência de etapas relativamente simples. Inicialmente, o duplex calcula as distâncias euclidianas entre todas as amostras disponíveis. A seguir, as duas amostras com maior distância são selecionadas para serem incluídas na fase de treinamento. Das amostras que restaram, as duas com maior distância entre si agora são escolhidas para a fase de validação. A seleção das amostras segue esse padrão até todas as amostras serem selecionadas.⁸⁶

Outro ponto importante na construção de modelos multivariados é a identificação e exclusão de amostras anômalas (“*outliers*”) na fase de treinamento. A utilização do espectrômetro portátil geralmente é simples, porém não é possível garantir que não haja erros na aquisição do espectro ou variações instrumentais. Devido a isso é preciso excluir essas amostras anômalas, que estão distantes do padrão das demais amostras, para que estas não prejudiquem o modelo. Os *outliers* são amostras que são muito distintas da maioria das amostras do banco de dados e a presença delas na fase de treinamento leva a criação de um modelo que contém valores que não representam a população descrita.⁸⁶ A detecção dessas amostras podem ocorrer de diversas

formas, nesse trabalho elas foram detectadas baseando-se nos valores de Q residual, Hotelling T² e os valores de y estimados para cada classe.^{15,87-91}

Q residual é o valor que indica a distância da posição da amostra original da sua projeção no modelo. Ou seja, quanto maior o valor de Q residual menos bem descrita está essa amostra no modelo. Hotelling T² é a distância que a amostra em questão está do centro de dados, esse conceito indica o quanto aquela amostra influencia no modelo. Um dos critérios mais aplicados para a identificação de outliers segundo esses parâmetros é considerar anomalias quando uma amostra apresentara valores de Hotelling T² e resíduos Q acima de um determinado limite estabelecido com certo grau de confiança, por exemplo 95 ou 99,9 %. Outro critério para a identificação de outliers é baseado nos valores estimados de classe (y), conforme descrito por Soares *et al.*¹⁵

3.3.2 Análise Discriminante por Mínimos Quadrados (PLS-DA).

Na análise PLS-DA os espectros são organizados em uma matriz **X** (I,M), onde I são as amostras e M representa os comprimentos de onda analisados. O desenvolvimento do modelo de discriminação ocorre a partir da correlação da matriz **X**, dos dados espectroscópicos, com o vetor **y** (I,1), onde I são as amostras. O vetor **y** contém a identificação das espécies/classes pelo número de classe, que é ou 0 ou 1. Sendo 0 para indicar que a amostra não pertence à classe a ser discriminada e 1 para indicar que a amostra pertence à classe discriminada. Ou seja, PLS-DA é uma análise binária, que indica se a amostra pertence ou não à determinada classe. A decomposição dos dados espectroscópicos e do vetor classe em variáveis latentes pode ser representada de forma resumida pelas seguintes equações:^{15,92,86}

$$\mathbf{X} = \sum_{a=1}^A \mathbf{t}_a \mathbf{p}_a^T + \mathbf{E} \quad (1)$$

$$\mathbf{y} = \sum_{a=1}^A \mathbf{t}_a \mathbf{q}_a^T + \mathbf{f} \quad (2)$$

Onde, \mathbf{t}_a é o vetor score, \mathbf{p}_a e \mathbf{q}_a são os loadings da primeira variável latente **a**, **E** e **f** são as matrizes que possuem informações não modeladas de **X** e **y**.

O limite de discriminação é o valor de classe (y) no qual a probabilidade da amostra pertencer à classe discriminada é igual a probabilidade dessa mesma amostra pertencer às outras classes. O limite de discriminação se baseia nos valores de y calculados para as amostras de treinamento, de forma a diminuir os números de falsos positivos e falsos negativos. Após estimar os valores de y para as amostras de treinamento calcula-se a probabilidades a priori e as funções de probabilidade. A probabilidade a priori da amostra pertencer a classe A (P(A)) e da amostra pertencer a classe B (P(B)) são definidas como:

$$P(A) = \frac{I_A}{I_A + I_B} \quad (3)$$

$$P(B) = \frac{I_B}{I_A + I_B} \quad (4)$$

As funções de densidade de probabilidade $p(\hat{y}_i|A)$ e $p(\hat{y}_i|B)$ são definidas como:

$$p(\hat{y}_i|A) = \frac{1}{S_A\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\hat{y}_i - \bar{y}_A}{S_A}\right)^2} \quad (5)$$

$$p(\hat{y}_i|B) = \frac{1}{S_B\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\hat{y}_i - \bar{y}_B}{S_B}\right)^2} \quad (6)$$

Baseando-se nessas duas estimativas as probabilidades de uma amostra i pertencer a classe A e a classe B são, respectivamente:

$$P(A|\hat{y}_i) = \frac{p(\hat{y}_i|A) \times P(A)}{p(\hat{y}_i|A) \times P(A) + p(\hat{y}_i|B) \times P(B)} \quad (7)$$

$$P(B|\hat{y}_i) = \frac{p(\hat{y}_i|B) \times P(B)}{p(\hat{y}_i|A) \times P(A) + p(\hat{y}_i|B) \times P(B)} \quad (8)$$

Como o limite de discriminação é o valor de y no qual a probabilidade da amostra pertencer à classe discriminada (A) é igual a probabilidade da amostra pertencer a classe não discriminada (B), iguala-se as equações 7 e 8. Como o denominador das duas equações são iguais, pode-se simplificar a igualdade da seguinte forma:

$$p(\hat{y}_i|A) \times P(A) = p(\hat{y}_i|B) \times P(B) \quad (9)$$

O resultado dessa igualdade é uma equação de segundo grau, sendo y a variável, resolvendo essa equação encontra-se o limite de discriminação. Maiores detalhes da resolução dessa igualdade é encontrado em Soares *et al.*¹⁵

3.3.3 Modelagem Independente Flexível por Analogia de Classes (SIMCA).

Essa técnica de classificação é baseada na decomposição dos dados em componentes principais (PC). A análise de componentes principais (PCA, do inglês, *Principal Component Analysis*) é uma ferramenta utilizada para extrair as informações que de melhor forma descrevem as tendências dos dados.⁹³ Para cada classe de espécie gera-se um modelo de componentes principais que representam essa classe, definindo uma região de confiança que caracteriza essa classe no espaço das PC. Deste modo, a classificação de uma amostra em uma

classe é obtida se essa amostra se encontra dentro da região de confiança para essa classe. A Figura 13 ilustra de forma simplificada, considerando um espaço de três dimensões, a projeção de três classes genéricas de acordo com SIMCA.

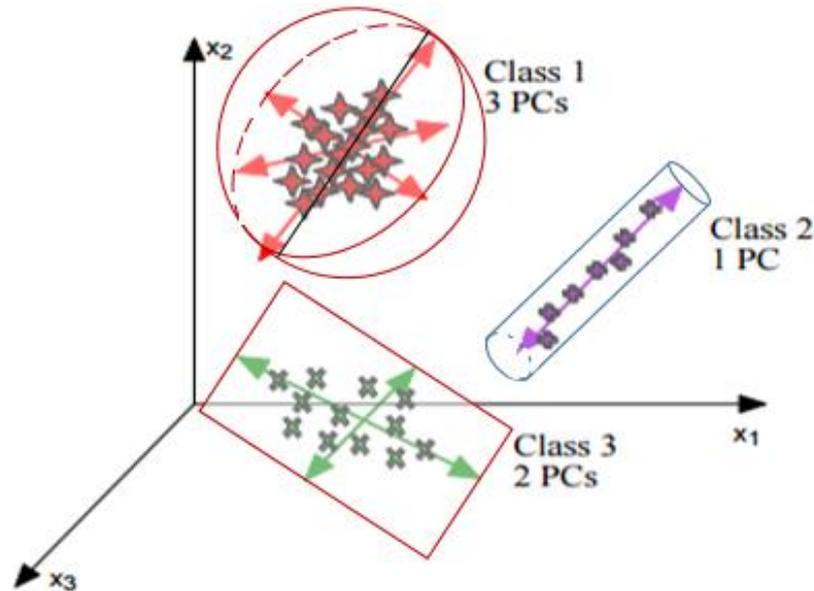


Figura 13. Exemplo gráfico de um modelo SIMCA genérico.⁹³

A decomposição PCA de uma classe w é expressa por:

$$\mathbf{X}_w = \sum_{a=1}^A \mathbf{t}_{a,w} \mathbf{p}_{a,w}^T + \mathbf{E}_w \quad (3)$$

Onde, \mathbf{X}_w é a matriz com os dados espectrais da classe w , $\mathbf{t}_{a,w}$ é a matriz de escores; $\mathbf{p}_{a,w}$ é a matriz de loading, \mathbf{E}_w é a matriz que contém os dados não incluídos no modelo e a é o número de componentes principais utilizados.⁹⁴ O número de componentes principais utilizadas para cada classe é determinado a partir da validação cruzada de forma que diminua os erros de classificação nas amostras de treinamento e aumente a eficiência nas predições das amostras de validação.

Na validação a amostra é projetada no modelo e as distâncias ($d_{i,w}$) da amostra i até a classe w do modelo é determinada, a classe que contiver a amostra dentro de sua região de confiança (critério stricto) ou apresentar a menor distância até a amostra (critério de classe mais provável) é a classe à qual a amostra pertencerá segundo o modelo. A distância d é calculada segundo a equação 4.

$$d_{i,w} = \sqrt{\left(\frac{T_{i,w}^2}{T_{crítico,w}^2}\right)^2 + \left(\frac{Q_{i,w}}{Q_{crítico,w}}\right)^2} \quad (4)$$

Onde, $d_{i,w}$ é a distância da amostra i até a classe w , $T_{i,w}^2$ é a distribuição T^2 de Hotelling, $Q_{i,w}$ é o somatório do quadrado do resíduo das amostras da classe w , $T_{crítico,w}^2$ e $Q_{crítico,w}$ são os limites de T^2 e Q com 99,9 % de nível de confiança para a classe w .⁹⁵ Segundo o critério stricto, caso a amostra apresente uma distância d maior que o valor de d crítico para todas as classes, a amostra é considerada *outlier* ou não pertencente à nenhuma classe.

3.3.4 Figuras de Mérito

As definições e equações apresentadas nessa seção são usadas para a determinação de figuras de mérito de métodos qualitativos de discriminação e classificação, podendo ser aplicadas tanto para a análise PLS-DA quanto SIMCA. O êxito do modelo é determinado de acordo com os resultados das amostras do grupo de validação. Considerando a classe w , o modelo irá determinar se a amostra ou pertence ou não à classe w . Importante destacar que as amostras de validação têm suas classes inicialmente determinadas por método convencional de referência, anatomia da madeira, assim como as de treinamento. As amostras verdadeiras positivas (VP) são as amostras que tanto o modelo quanto a técnica convencional identificaram como pertencente à classe w . As amostras verdadeiras negativas (VN) são as amostras que não são pertencentes à classe w de acordo com o modelo e a técnica convencional. As amostras falsas negativas (FN) são amostras que de acordo com a técnica convencional são pertencentes à classe w , porém o modelo as identificou como não pertencente à classe w . As amostras falsas positivas (FP) são amostras que o modelo identificou como pertencente à classe w , porém a técnica convencional não. A taxa de falso negativo (TFN) é a porcentagem das amostras da classe w que foram consideradas como não pertencente à classe w . A taxa de falso positivo (TFP) é a porcentagem das amostras não pertencentes à classe w identificadas como pertencente à classe w .⁹⁶

$$TFP = \frac{FP}{FP + VN} \times 100 \%$$

$$TFN = \frac{FN}{FN + VP} \times 100 \%$$

$$EFR = 100 \% - (TFP + TFN)$$

A eficiência do modelo é 100 % menos os valores de TFP e TFN.^{15,18} Como para todas as amostras os espectros foram coletados em triplicatas, analisou-se a predição dos espectros de cada amostra para decidir como a amostra como um todo foi predita pelo modelo. A amostra é considerada de determinada classe caso dois ou mais espectros dessa amostra forem preditos como pertencentes a essa classe.

4. MATERIAIS E MÉTODOS

4.1 Preparação das amostras

Foram utilizadas amostras de 15 espécies de *Dalbergia*: *D. bariensis*, *D. cearenses*, *D. congestiflora*, *D. decipularis*, *D. glaucescens*, *D. latifolia*, *D. melanoxyton*, *D. miscolobium*, *D. nigra*, *D. retusa*, *D. sisso*, *D. spruceana*, *D. stevensonii*, *D. tucurensis* e *D. villosa* pertencentes a coleção do Laboratório de Produtos Florestais do Serviço Florestal Brasileiro e de xilotecas credenciadas pela IAWA: Laboratório de Produtos Florestais do Serviço Florestal brasileiro (FPBw), Laboratório de Produtos Florestais do Departamento de Agronomia dos Estados Unidos de Madison - coleção original de Madison (MADw) e coleção Samuel J. Record (SJRw); Xiloteca do Instituto Florestal de São Paulo (SPSFw); Xiloteca do Jardim Botânico do Rio de Janeiro (RBw) e Xiloteca Dr. Calvino Mainieri do Instituto de Pesquisas Tecnológicas de São Paulo (BCTw). A quantidade de amostras, de espectros e a distribuição das amostras conforme xiloteca de origem está descrita na Tabela 2.

Os espectros foram medidos em três posições distintas, aleatoriamente de qualquer superfície do cerne da amostra. Antes da coleta dos espectros, as amostras foram lixadas com lixas grã 80 e o pó foi retirado com pincel. Objetiva-se com esse procedimento que o espectro coletado contenha apenas informações químicas de uma superfície de madeira que não esteja oxidada e fazer com que todas as amostras tenham passado pelo mesmo procedimento de preparo de superfície antes da construção do modelo, conforme descrito por Soares *et al.*¹⁵

4.2 Espectrômetro NIR portátil.

O instrumento utilizado para realizar as medidas de reflectância foi o microPhazir™ RX Analyzer que trabalha no intervalo de 1.595 – 2.396 nm com resolução de 9 nm. A Figura 14 mostra o instrumento, que possui formato anatômico que se assemelha a uma pistola e apresenta uma massa de 1,2 kg, sendo classificado como um espectrômetro portátil que permite medidas *in loco*.

A fonte de radiação consiste de uma lâmpada de tungstênio e a área de contato com a amostra corresponde a uma superfície de 1,2 x 1,2 cm, o que requer apenas pequena superfície da amostra de madeira. Para facilitar as análises de dados as informações espectrais foram exportadas para o computador e a análise foi feita com os programas Matlab R2011 e PLS Toolbox 7.03.

Tabela 2. Informações das amostras estudadas, número, espectros e origem.

Espécie	Nº amostra/ Nº espectro	Xiloteca de origem					
		FPBw ¹	MADw ² /SJRw ³	SPSFw ⁴	RBw ⁵	LPF ⁶	BCTw ⁷
<i>D. bariensis</i> Pierre	2 / 6	2					
<i>D. cearensis</i> Ducke	27 / 81	11	6	4	6		
<i>D. congestiflora</i> Pittier	2 / 6	2					
<i>D. decipularis</i> Rizzini & A. Mattos	28 / 84	8	16	2	2		
<i>D. glaucescens</i> Mart ex. Benth	1 / 3	1					
<i>D. latifolia</i> Roxb.	22 / 66	2	19	1			
<i>D. melanoxyton</i> Guill. & Perr.	2 / 6	2					
<i>D. miscolobium</i> Benth.	4 / 12	2		2			
<i>D. nigra</i> (Vell.) Benth.	169 / 507	14	28	6	5	116	
<i>D. retusa</i> Hemsl.	26 / 78	4	21	1			
<i>D. sissoo</i> DC.	21 / 63		19	1	1		
<i>D. spruceana</i> Benth	26 / 78	6	4	5	1		10
<i>D. stevensonii</i> Standl.	18 / 54	3	14		1		
<i>D. tucurensis</i> Donn.Sm.	10 / 30	1	9				
<i>D. villosa</i> Benth.	1 / 3	1					
Total	359 / 1077	59	136	22	16	116	10

¹FPBw - Laboratório de Produtos Florestais do Serviço Florestal brasileiro; ²MADw - Laboratório de Produtos Florestais do Departamento de Agronomia dos Estados Unidos de Madison - coleção original de Madison; ³SJRw - Laboratório de Produtos Florestais do Departamento de Agronomia dos Estados Unidos de Madison - coleção Samuel J. Record; ⁴SPSFw - Xiloteca do Instituto Florestal de São Paulo; ⁵RBw - Xiloteca do Jardim Botânico do Rio de Janeiro; ⁶LPF - Laboratório de Produtos Florestais do Serviço Florestal Brasileiro; ⁷BCTw - Xiloteca Dr. Calvino Mainieri do Instituto de Pesquisas Tecnológicas de São Paulo.



Figura 14. Espectrômetro portátil microPhazir™ RX Analyzer (Fonte: Thermo Fisher Scientific)⁹⁷

4.3 Pré-processamento dos espectros no infravermelho próximo.

O pré-processamento dos espectros tem como objetivo eliminar ou minimizar variações espectrais que não estão relacionadas à variação da propriedade que se deseja determinar (diferença entre as espécies), como por exemplo a variação de linha de base devido a espalhamentos aditivos e multiplicativos da radiação. O pré-processamento escolhido para essa análise foi a redução do ruído (alisamento) dos espectros, seguido da primeira derivada utilizando o algoritmo Savitzky-Golay com janela espectral de 11 pontos e polinômio de segunda ordem. Após o cálculo da derivada os espectros foram normalizados com o pré-processamento SNV (SNV, do inglês, Standard Normal Variate) e uso de dados centrados na média.

4.4 Organizações das classes discriminadas

Os modelos de discriminação ou classificação construídos nessa dissertação foram divididas em três grandes grupos:

- 1 – Discriminar/identificar a *D. nigra* das demais espécies;
- 2 – Discriminar/identificar a *D. spruceana* das demais espécies;
- 3 – Discriminar/identificar individualmente 9 espécies uma das outras.

Para as duas primeiras discriminações avaliou-se a eficiência das duas ferramentas quimiométricas de acordo com a região espectral utilizada. As duas faixas espectrais avaliadas foram:

1 – 1.622 – 1.852 nm + 1.960 – 2.307 nm: A primeira seleção de variável possui um número menor de comprimentos de onda, que foram selecionadas de acordo com os principais constituintes da madeira, de acordo com Schwanninger *et al.*⁷⁹:

- Banda mais relacionada a Extrativos centrada em 1.660 nm;
- Banda mais relacionada a Hemiceluloses centrada em 1.710 nm;
- Banda mais relacionada a Celulose centrada em 2.100 nm.

2 – 1.622 – 2.367 nm: O segundo tipo de seleção é um intervalo que representa todos os comprimentos de onda mensurados pelo espectrômetro portátil microPhazir™ RX Analyzer.

As duas primeiras discriminações desenvolvidas indicaram que ao utilizar toda a região espectral os modelos apresentaram mais eficiência, logo, para a discriminação/ identificação de cada espécie individualmente utilizou-se toda a região espectral. As 9 espécies escolhidas para essa fase foram selecionadas com base no critério de possuírem no mínimo 10 amostras.

Tendo em vista que se utilizou o algoritmo DUPLEX para seleção dos espectros da fase de treinamento e validação, para cada modelo o algoritmo fez uma seleção própria e diferente das demais. As Figuras 15 e 16 são esquemas que demonstram como os espectros de todas as espécies foram organizados nos modelos. As *Dalbergias* com menor número de amostras foram agrupadas em um grupo intitulado “Diversas”.

A separação dos espectros garantiu que os três espectros da mesma amostra estivessem ou na fase de treinamento ou na fase de validação, nunca uma mesma amostra teve espectros nas duas fases.

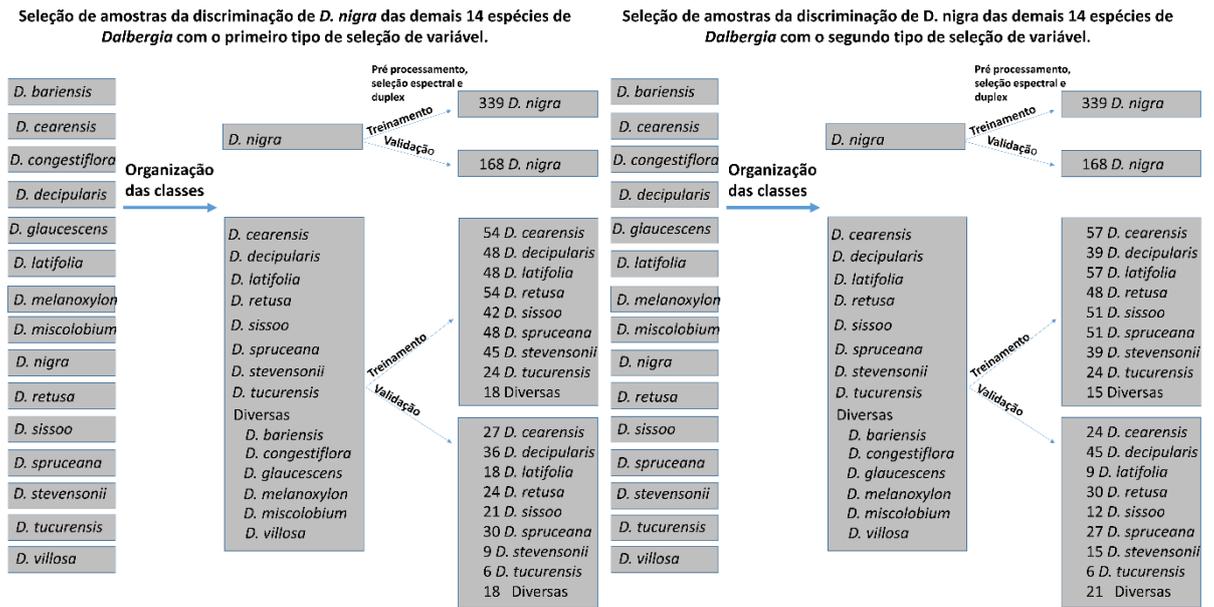


Figura 15. Esquema de seleção de espectros dos modelos discriminatório de *D. nigra* com seleção de variáveis do tipo 1 e 2.

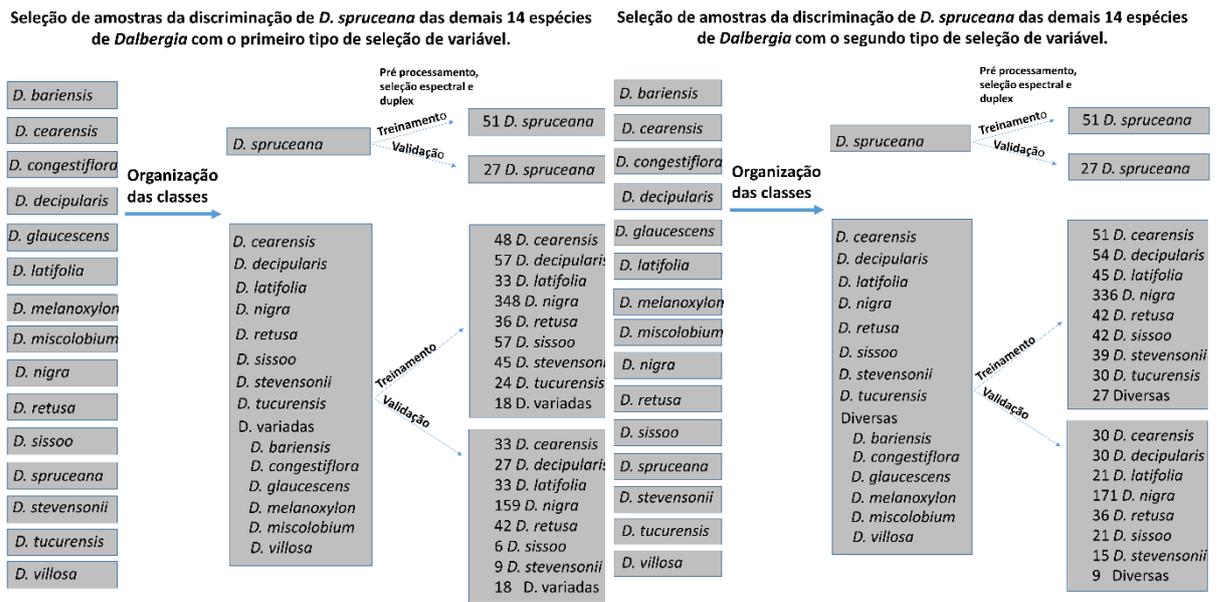


Figura 16. Esquema de seleção de espectros do modelo discriminatório de *D. spruceana* com seleção de variáveis do tipo 1 e 2.

5. RESULTADOS E DISCUSSÃO

A Figura 17 apresenta os espectros médios adquiridos de cada espécie estudada antes de qualquer pré-processamento. Observa-se a variação da linha de base entre uma espécie e outra e uma grande semelhança do formato dos espectros entre as espécies. A Figura 18 apresenta os mesmos dados da Figura 17, porém pré-processados com alisamento, 1ª derivada e SNV, nota-se que a variação da linha de base é significativamente minimizada. Além de minimizar a variação da linha de base os espectros pré-processados indicam que ocorre maior variação na intensidade de absorção entre as espécies em regiões específicas do espectro. Esses comprimentos de onda são relacionados com extrativos da madeira (~1660 nm), hemicelulose (~1710 nm) e celulose (2100 nm), essa variação de intensidade pode indicar variação da concentração desses compostos em cada espécie.⁷⁹ Na tentativa de se obter um método mais robusto e com menores erros de discriminação, nessa dissertação foram desenvolvidos modelos no qual as variáveis foram selecionadas de acordo com essas regiões e com o espectro completo.

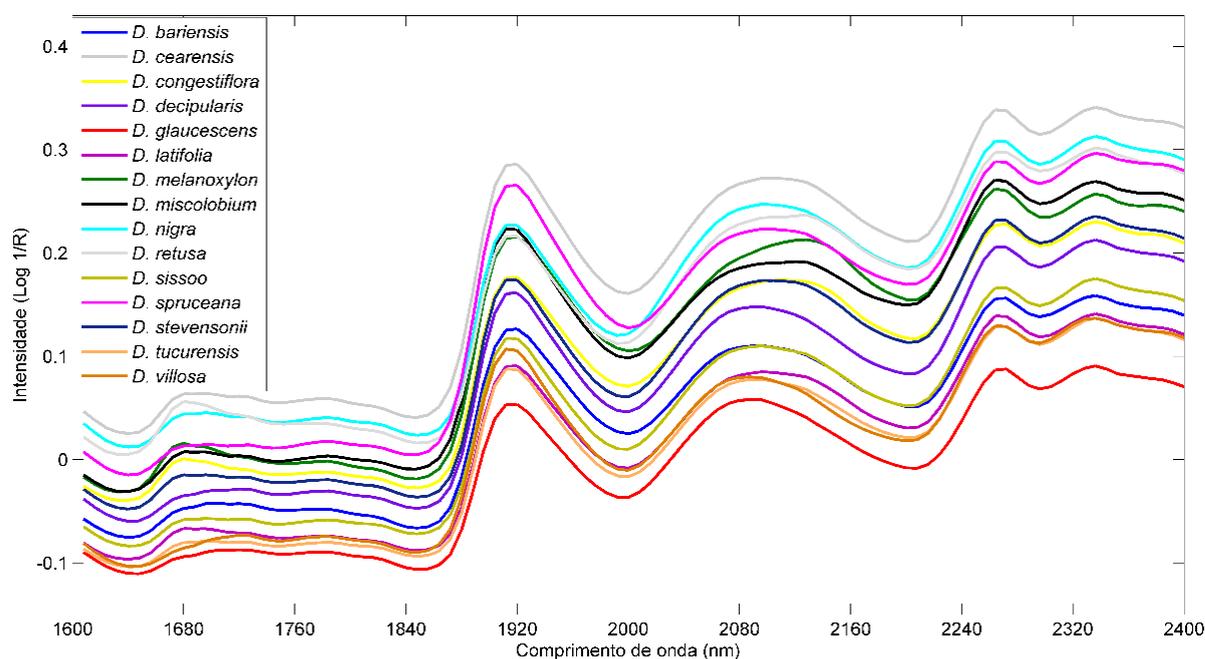


Figura 17. Espectros médios de cada espécie sem pré-processamento.

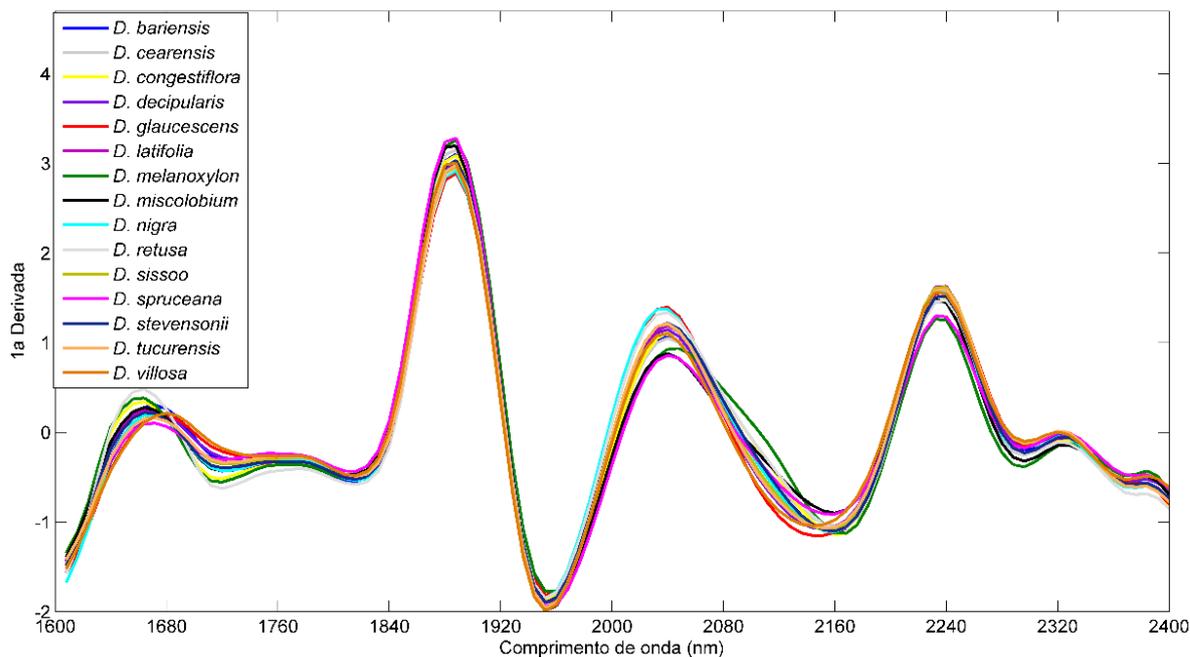


Figura 18 Espectros médios de cada espécie após pré-processamento com alisamento, 1ª derivada e SNV.

5.1 Discriminação da *Dalbergia nigra* das demais 14 espécies de *Dalbergia*

5.1.1 PLS-DA

O modelo PLS-DA, como descrito anteriormente em mais detalhes, faz uso de variáveis latentes para alcançar o coeficiente de regressão ideal para prever as qualidades das amostras de validação. A escolha do número de variáveis latentes é importante para criação do modelo e deve-se fazê-la com atenção para evitar o sobreajuste. Um modelo sobreajustado tende a representar muito bem as amostras utilizadas na fase de treinamento, porém como é um modelo extremamente específico para as amostras de treinamento, pequenas variações das amostras de validação não serão reconhecidas, decaindo a capacidade de predição do modelo e fazendo com que as amostras de validação sejam facilmente identificadas como outliers.

A seleção de variáveis latentes foi feita analisando a relação entre a raiz quadrada do erro de médio quadrático de validação cruzada (RMSECV) e a quantidade de variáveis latentes. A medida que aumenta-se o número de variáveis latentes o RMSECV geralmente diminui, pois mais componentes são usados para representar as amostras da fase de treinamento. Porém para evitar o sobreajuste escolheu-se a variável latente na qual o RMSECV se torna aproximadamente constante ou deixa de apresentar reduções significativas. Para ilustrar a relação do número de variáveis latentes com o erro de validação cruzada apresenta-se na Figura

19 a essa relação do modelo de separação de *D. nigra* das demais espécies ao utilizar as variáveis de acordo com os principais componentes da madeira.

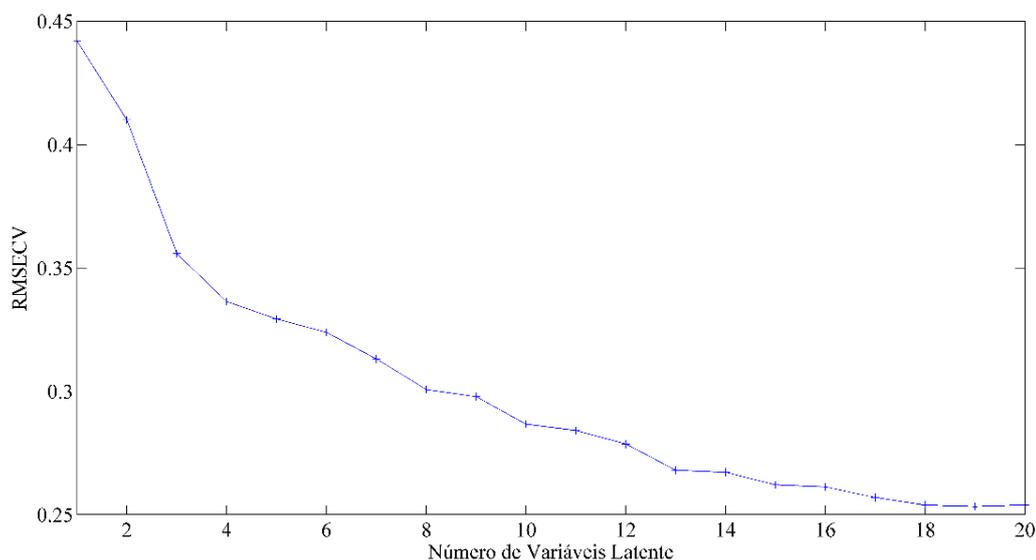


Figura 19. Valor de erro de validação cruzada (RMSECV) de acordo com o número de variáveis latentes do modelo de discriminação de *D. nigra* com seleção de variável do tipo 1.

Para o modelo citado escolheu-se treze variáveis latentes tendo em vista que ao utilizar mais variáveis latentes que treze o erro não diminui consideravelmente, sendo assim, as próximas variáveis latentes não são necessárias para representar bem o conjunto de amostras de treinamento (Figura 19). As variáveis latentes escolhidas para todos os modelos PLS-DA dessa dissertação foram determinadas seguindo esse mesmo critério.

Após a determinação da quantidade de variáveis latentes foi criado um primeiro modelo que deve ser otimizado com a exclusão de amostras anômalas de acordo com valor de T^2 , Q e y . As Figuras 20A e 20B, respectivamente, indicam os valores de T^2 de Hotelling e Resíduos Q dos espectros da fase de treinamento do modelo com seleção de variável do tipo 1 e do tipo 2.

Importante lembrar que em relação aos valores de Resíduos Q e T^2 de Hotelling, a amostra só foi considerada anômala caso ambos valores tenham sido superiores ao limite. Ou seja, se o espectro apresentar apenas valor de Resíduo Q acima do limite e não de T^2 de Hotelling o espectro não é considerado anômalo. Os espectros que apresentam altos valores de ambos parâmetros estão localizados no quadrante superior direito do gráfico de Resíduo Q em relação ao T^2 de Hotelling.

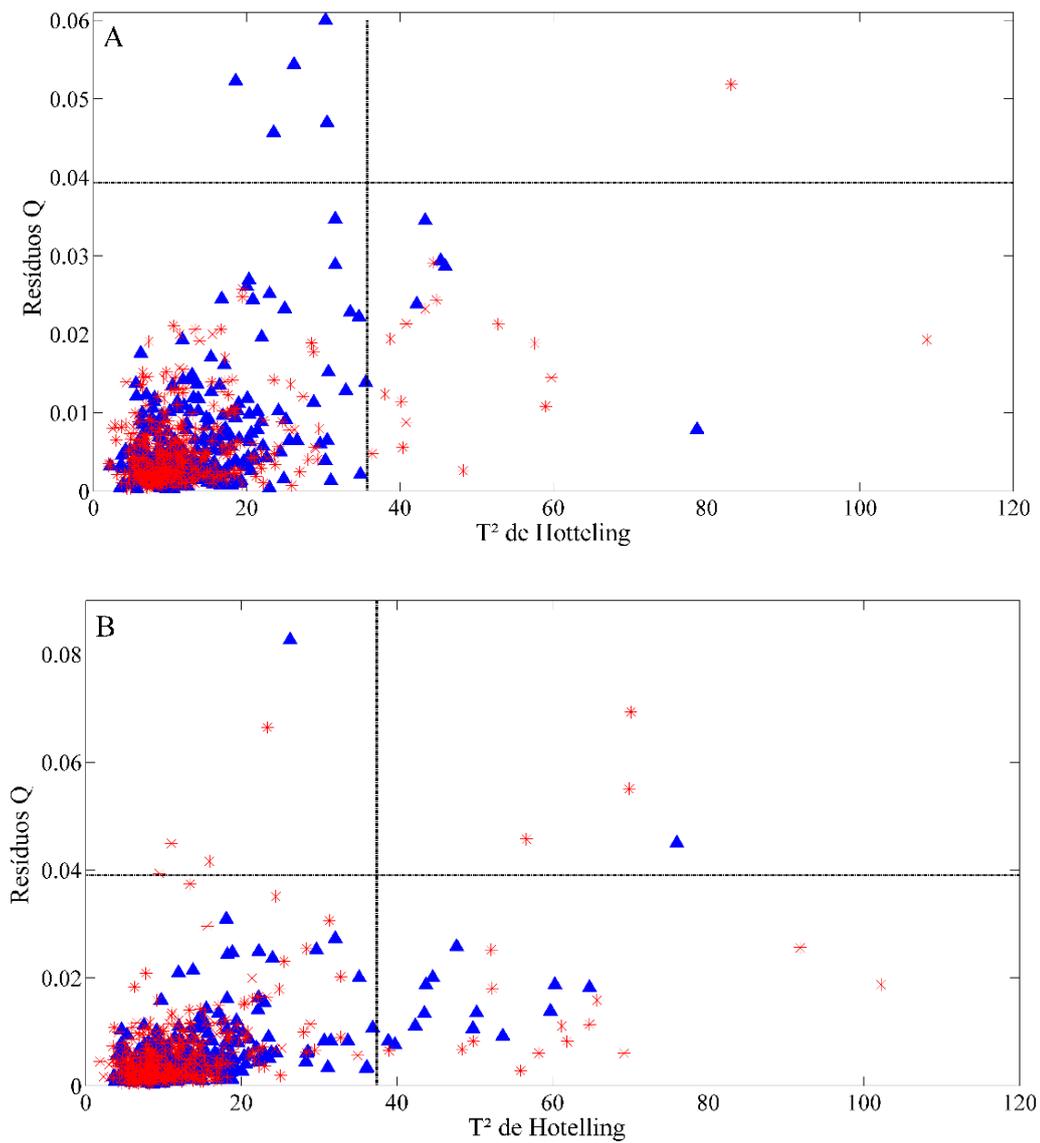


Figura 20. Gráfico dos valores de T^2 de Hotelling e resíduos Q das amostras de treinamento do modelo com seleção de variáveis do tipo 1 (A) e com seleção de variáveis do tipo 2 (B).
 Legenda: *D. nigra* (▲) e Demais Dalbergias (*).

A Tabela 3 indica a quantidade total de amostras anômalas identificadas e removidas da fase de treinamento, a tabela indica mais amostras do que as representadas na Figura 20, pois na tabela é apresentado as amostras anômalas de acordo com os três critérios: T^2 de Hotelling, Resíduos Q e y , a figura só apresenta os valores de T^2 de Hotelling e Resíduos Q.

Tabela 3. Quantidade de variáveis latentes e espectros anômalos excluído nos dois modelos de discriminação de *D. nigra*.

	1- $\lambda = 1.622 - 1.852 \text{ nm} +$ $1.960 - 2.307 \text{ nm}$	2- $\lambda = 1.622 - 2.367 \text{ nm}$
Número de variáveis latentes	13	14
Espectros anômalos na fase de treinamento	17	11
Espectros anômalos na fase de validação	4	14

Após a exclusão das amostras de treinamento o segundo modelo é criado e o limite de discriminação das classes é calculado. O modelo é então utilizado para prever os valores de y das amostras de validação. As amostras de validação também passaram por exclusão de acordo com os valores de T^2 de Hotelling, Q residual e valor de y adaptado: amostras da classe 1 foram excluídas apenas de acordo com o limite superior e as amostras da classe 0 de acordo com limite inferior. O motivo dessa adaptação está explicado na seção 3.3.1. A Figura 22 indica os valores de T^2 de Hotelling e Q residual de todas as amostras, incluindo treinamento e validação segundo o modelo de predição. Importante ressaltar que as amostras anômalas da validação foram desconsideradas para o cálculo das figuras de mérito. As Figuras 21A e 21B demonstram, respectivamente, os valores de Q residual e T^2 de Hotelling dos espectros nos modelos com seleção de variável do tipo 1 e do tipo 2.

A seleção adequada de espectros anômalos é importante para evitar a construção de um modelo que possa ter sua predição comprometida ou distorcida pela presença de espectros anômalos. Apesar da necessidade da exclusão de espectros anômalos, para que um modelo seja tenha maior capacidade preditiva é importante que não se exclua uma grande porção dos espectros. Pois, caso contrário, os espectros utilizados não estão sendo representativos das classes estudadas. A exclusão de um número elevado de outliers, por exemplo maior que 10 %, pode indicar erros de operação ou funcionamento do equipamento.⁸⁷

Observa-se que na fase de treinamento do primeiro tipo de modelo há 17 espectros anômalos que representam 2,4 % da quantidade total (Tabela 3). Na fase de validação do modelo, utilizando o segundo tipo de seleção de variável, 14 espectros foram considerados anômalos, o que representa apenas 3,9 % de espectros da fase de validação. Esses baixos percentuais de amostras anômalas identificadas e excluídas podem ser considerados aceitáveis

tendo em vista a variabilidade das amostras provenientes de diferentes regiões e diferentes espécies.

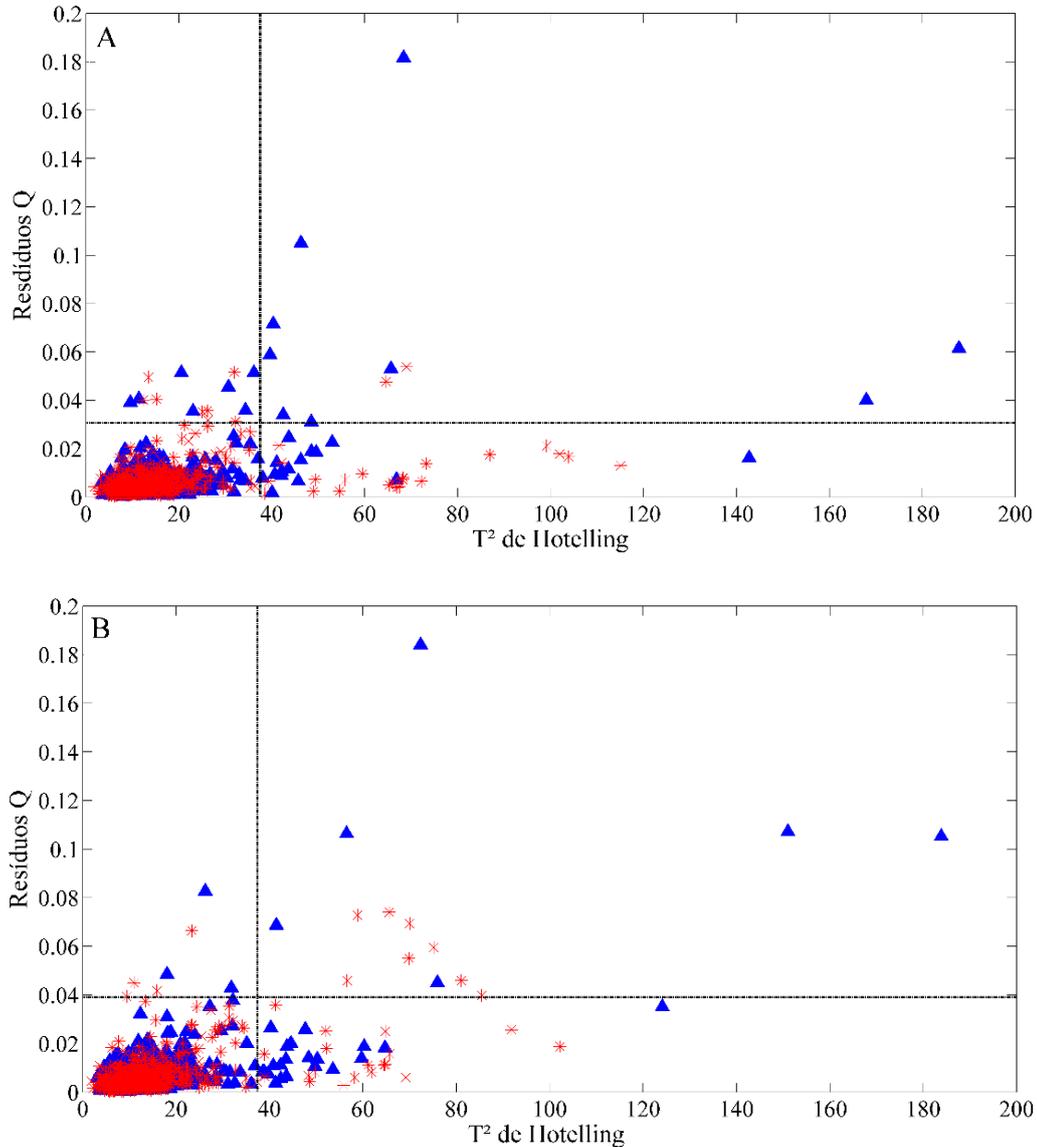


Figura 21. Gráfico dos valores de T^2 de Hotelling e resíduos Q dos espectros da fase de treinamento e de validação do modelo com seleção de variáveis 1 (A) e com seleção de variável 2 (B). Legenda: *D. nigra* (▲) e Demais Dalbergias (*).

Os valores estimados de classe (y predito) para os espectros com seleção de variáveis do tipo 1 e 2 estão, respectivamente, representados nas Figuras 22 e 23. O eixo das ordenadas indica o valor de classe estimado pelo modelo PLS-DA para o respectivo espectro, a indicação do número do espectro está no eixo das abscissas. Importante lembrar que para uma amostra de *D. nigra* ser considerada uma amostra verdadeira positiva ao menos dois espectros da amostra devem ter valor de classe predito superior ao limite de discriminação. O limite de discriminação

é representado por uma linha horizontal que corta o eixo das ordenadas no valor de 0,4663 no modelo com seleção de variável do tipo 1 (Figura 23) e 0,4645 no modelo com seleção de variável do tipo 2 (Figura 24). Com o mesmo raciocínio se identifica uma amostra verdadeira negativa (VN): dois espectros de uma amostra que não é *D. nigra* devem ter valor predito de classe menor que o limite de discriminação. Cada espécie é representada com símbolo/cor distintos de acordo com o indicado na legenda. Os símbolos vazios representam os espectros da fase de treinamento e os cheios os espectros da fase de validação. Apesar das espécies estarem individualmente representadas nas Figuras 22 e 23 é importante lembrar que todas as espécies, exceto a *D. nigra*, foram consideradas de uma única classe.

O modelo com seleção de variável do tipo 1 apresenta apenas duas amostras que foram erroneamente preditas pelo modelo como sendo *D. nigra*, são elas: uma amostra de *D. tucurensis* e uma que pertence ao grupo das Diversas. Observa-se na Figura 23 que um espectro da *D. cearensis* possui valor de classe predito (y) superior ao limite de discriminação. Contudo, tendo em vista que para se definir a qual classe a amostra pertence é preciso avaliar a maioria dos seus espectros, conclui-se que nenhuma amostra de *D. cearensis* foi erroneamente classificada no modelo com seleção de variável do tipo 2. Esse modelo apresentou 5 falsos positivos na fase de validação: duas amostras de *D. sissoo*, duas amostras de *D. spruceana* e uma de *D. tucurensis*.

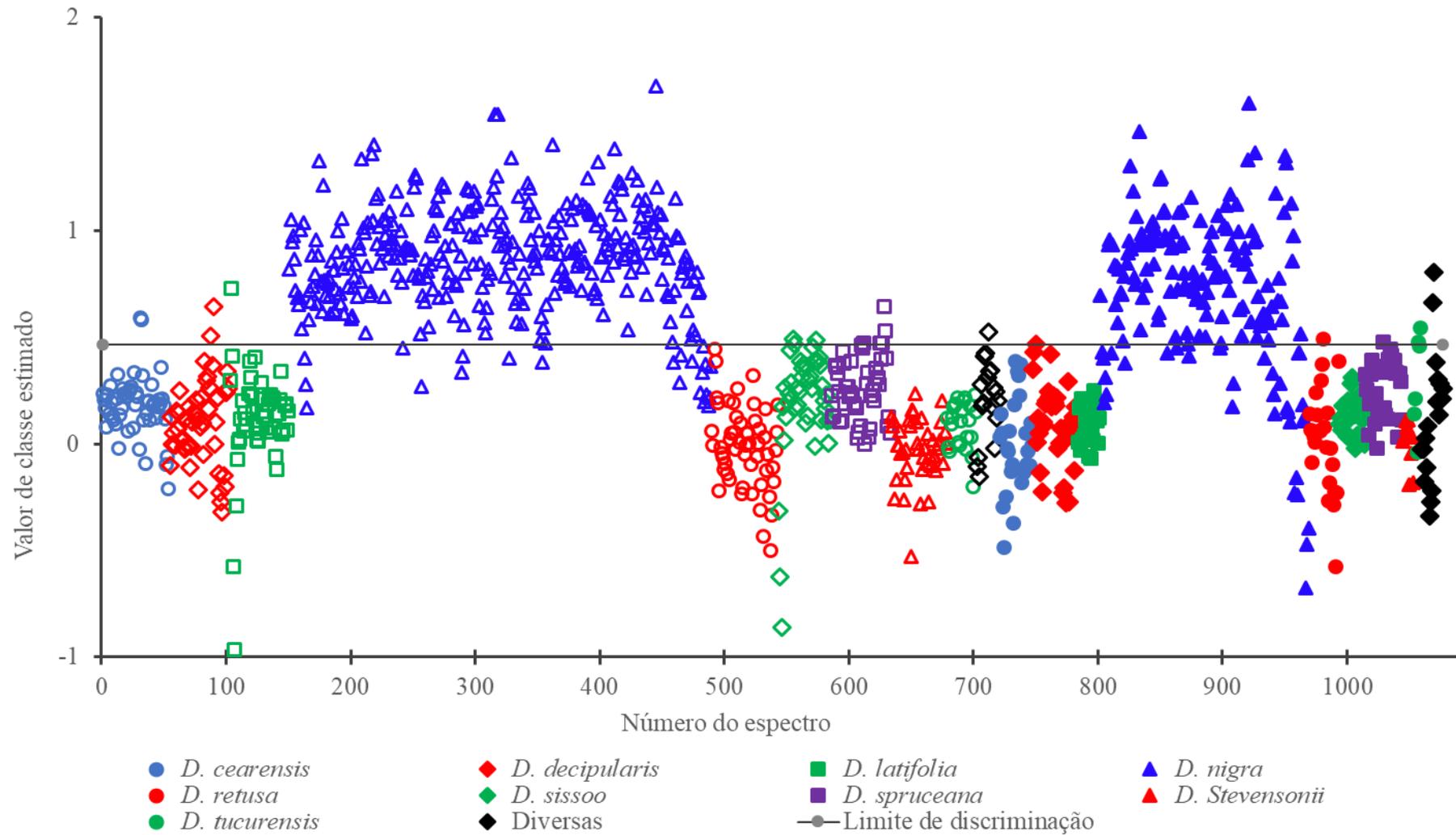


Figura 22. Valor de classe predito dos espectros do modelo PLS-DA de discriminação de *D. nigra* das demais espécies com seleção de variável do tipo 1. Símbolos: vazios - fase de treinamento; cheios fase de validação.

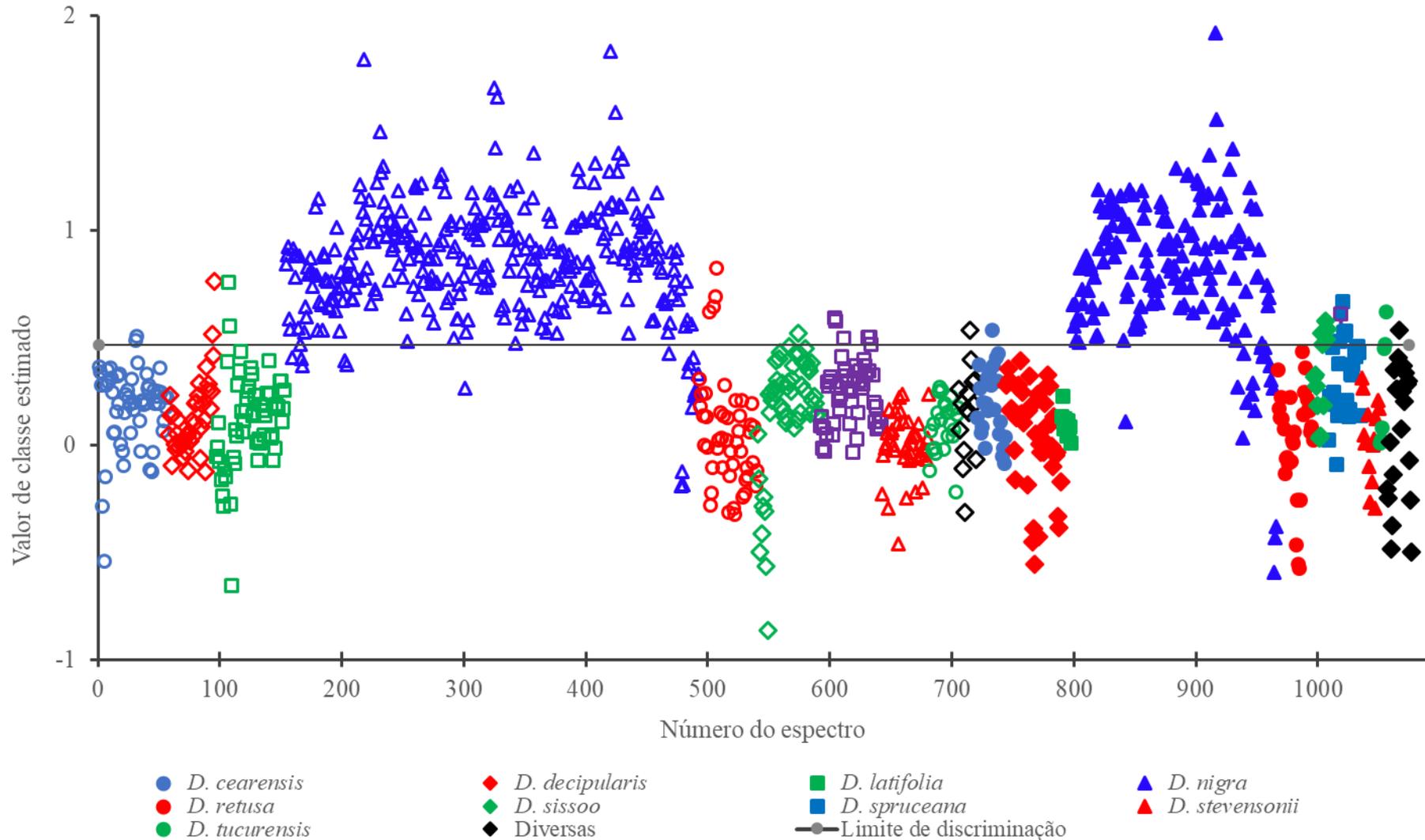


Figura 23. Valor de classe predito dos espectros do modelo PLS-DA de discriminação de *D. nigra* das demais espécies com seleção de variável do tipo 2. Símbolos: vazios - fase de treinamento; cheios fase de validação.

A Tabela 4 indica as figuras de mérito dos dois tipos de seleções de variáveis. Apesar de haver 56 amostras de *D. nigra* na fase de validação desse modelo, a tabela indica apenas 53 amostras (45 verdadeiras positivas e 8 falsas negativas) para o modelo com seleção de variável tipo 1 e 54 amostras (49 verdadeiras positivas e 5 falsas negativas) para o modelo com segundo tipo de variáveis, isso se deve ao fato de que três amostras de *D. nigra* foram excluídas como outliers no modelo com seleção de variável do tipo 1 e duas foram excluídas no modelo com seleção de variável do tipo 2.

Tabela 4. Figuras de mérito obtidas com base no conjunto de validação dos dois modelos discriminatórios de *D. nigra* com a ferramenta PLS-DA.

Figuras de mérito			
1- $\lambda = 1.622 - 1.852 \text{ nm} + 1.960 - 2.307 \text{ nm}$		2- $\lambda = 1.622 - 2.367 \text{ nm}$	
FP	2	FP	5
FN	8	FN	5
VP	45	VP	49
VN	58	VN	56
TFP (%)	3,3	TFP (%)	8,2
TFN (%)	15,1	TFN (%)	9,3
EF (%)	81,6	EF (%)	82,5

FP – Falso Positivo, **FN** – Falso Negativo, **VP** – Verdadeiro Positivo, **VN** – Verdadeiro Negativo, **TFP** – Taxa de Falso Positivo, **TFN** – Taxa de Falso Negativo e **EF** – Eficiência.

Observa-se que as eficiências entre os dois tipos de seleção de variável com a ferramenta PLS-DA diferem em apenas 0,9 % a favor do modelo com toda a faixa espectral do instrumento portátil. Para fins forenses é importante que haja baixa taxa de falsos positivos para que não ocorra apreensão de produtos legais, levando isso em consideração avaliamos que a seleção de variável do tipo 1 apresenta melhores resultados, tendo em vista que apresentou taxa de falso positivo de 3,3 % enquanto o segundo modelo apresentou taxa de 8,2 %.

Independente do modelo, nota-se que ambos tiveram altos valores de eficiência, indicando que o uso do espectrômetro portátil em conjunto com ferramenta PLS-DA é um método que pode auxiliar a discriminação da *D. nigra* das demais espécies de *Dalbergia* estudadas nesse trabalho. O mesmo tipo de distinção foi acusado por Espinoza *et al.*⁴⁶ ao utilizar a técnica DART-TOFMS. Estes autores discriminaram a *D. nigra* de outras espécies

visualmente semelhantes: *D. spruceana*, *D. stevensonii* e *D. tucurensis*. É importante apontar que a tecnologia NIRS permite a identificação em campo, diferente da técnica DAR-TOFMS que necessita de instrumentação laboratorial.

Apesar de haver várias espécies de *Dalbergia* que não foram estudadas nesse trabalho, observa-se que há distinção de informação espectral entre a *D. nigra* e outras *Dalbergias*, e espera-se que essa distinção ocorra também com as espécies ainda não estudadas, porém só com aquisição de mais espectros de outras espécies será possível chegar a essa conclusão.

5.1.2 SIMCA

A ferramenta SIMCA cria uma região com dimensões variadas, de acordo com o número de componentes principais inseridas no modelo, para definir uma região que caracteriza uma classe de interesse. Nessa análise modelou-se uma região para a classe *D. nigra* e na fase de validação as amostras foram projetadas no modelo e de acordo com a distância delas à região delimitada para a espécie *D. nigra* foi avaliado se as amostras pertencem a espécie *D. nigra* ou não pertencem.

Durante a fase de criação de modelos é necessário determinar o número de componentes principais que deve ser utilizado para criar a região da *D. nigra*. Para a decisão do número de componentes principais analisou-se a eficiência do modelo para prever a classe das amostras do conjunto de treinamento ao variar o número de componentes principais. Escolheu-se o número de componentes que a partir dela não havia aumento representativo da eficiência. Os valores de componentes principais utilizadas na construção dos modelos com seleção de variável do tipo 1 e do tipo 2 são, respectivamente, 7 e 8.

Os valores de probabilidade, em porcentagem, dos espectros pertencerem à classe *D. nigra*, utilizando a região espectral de acordo com os principais componentes da madeira, estão apresentados nas Figura 24. O modelo que utiliza toda a região espectral do instrumento está representado na Figura 25. Assim como nos modelos criados com PLS-DA, as figuras de mérito foram calculadas de acordo com as amostras e não os espectros de forma individual. Os pontos, que representam os espectros, são organizados da mesma forma que o PLS-DA, sendo os símbolos vazios os espectros da fase de treinamento e os símbolos cheios da fase de validação. Cada espécie é representada com símbolo/cor distintos de acordo com o indicado na legenda. Os símbolos vazios representam os espectros da fase de treinamento e os cheios os espectros da fase de validação.

Observa-se na Figura 24 que há muitos espectros de outras espécies com mais de 50 % de probabilidade de pertencer à classe *D. nigra*, ou seja, o modelo apresenta muitas amostras falsas positivas (12). Todas as figuras que representam os valores de probabilidades dos espectros pertencerem a alguma classe, figuras de análises com a ferramenta SIMCA, devem ser observadas com atenção. Por exemplo, ao analisar os espectros de *D. nigra* da fase de validação (Figura 24), se não observar os dados com atenção pode-se concluir que há maior densidade de pontos de *D. nigra* abaixo do limite do que acima, pois um olhar rápido leva em consideração apenas os pontos acima de 0 % e abaixo de 100 %, que são mais facilmente visualizados. Porém, deve-se notar que existem muitos espectros que tem 100 % de chance de pertencer à classe *D. nigra* e estão aglomerados na linha de 100 %. Tendo isso em vista, é válido ressaltar que há mais amostras verdadeiras positivas (36) do que falsas negativas (20) (Tabela 5). Em relação às demais espécies, 51 espectros foram corretamente classificados como não pertencentes à classe *D. nigra* e 12 foram erroneamente classificados como *D. nigra*.

Deve-se observar a Figura 25, que indica a probabilidade de um espectro pertencer à classe *D. nigra* com seleção de variável do tipo 2, com a mesma atenção. Esse tipo de seleção apresentou ainda mais amostras verdadeiras positivas (38) em relação ao modelo anterior e menos amostras das demais espécies foram erroneamente classificadas como *D. nigra* (FP) apenas oito.

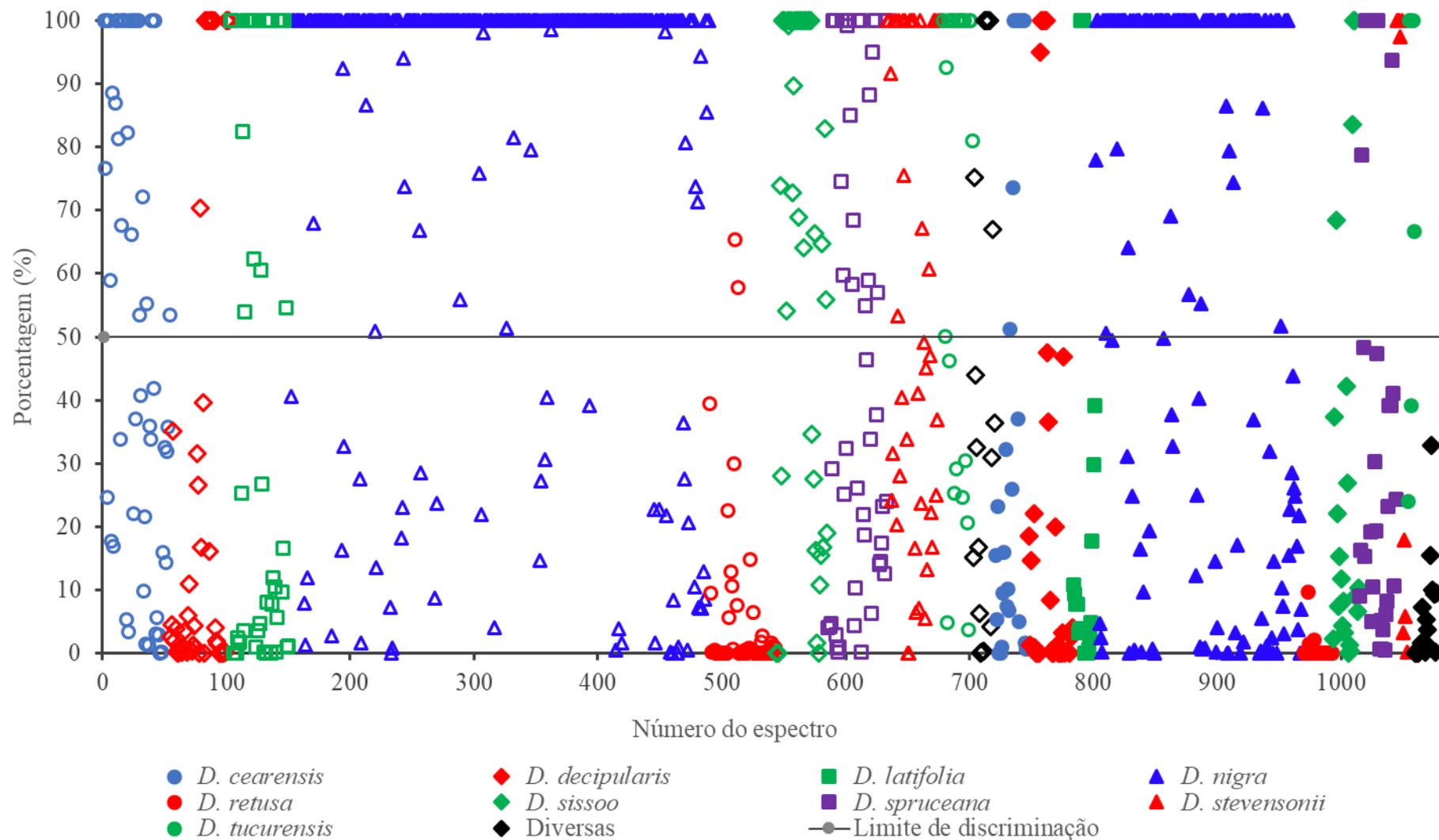


Figura 24. Probabilidade do espectro pertencer à espécie *D. nigra* segundo o modelo SIMCA com seleção de variável do tipo 1.

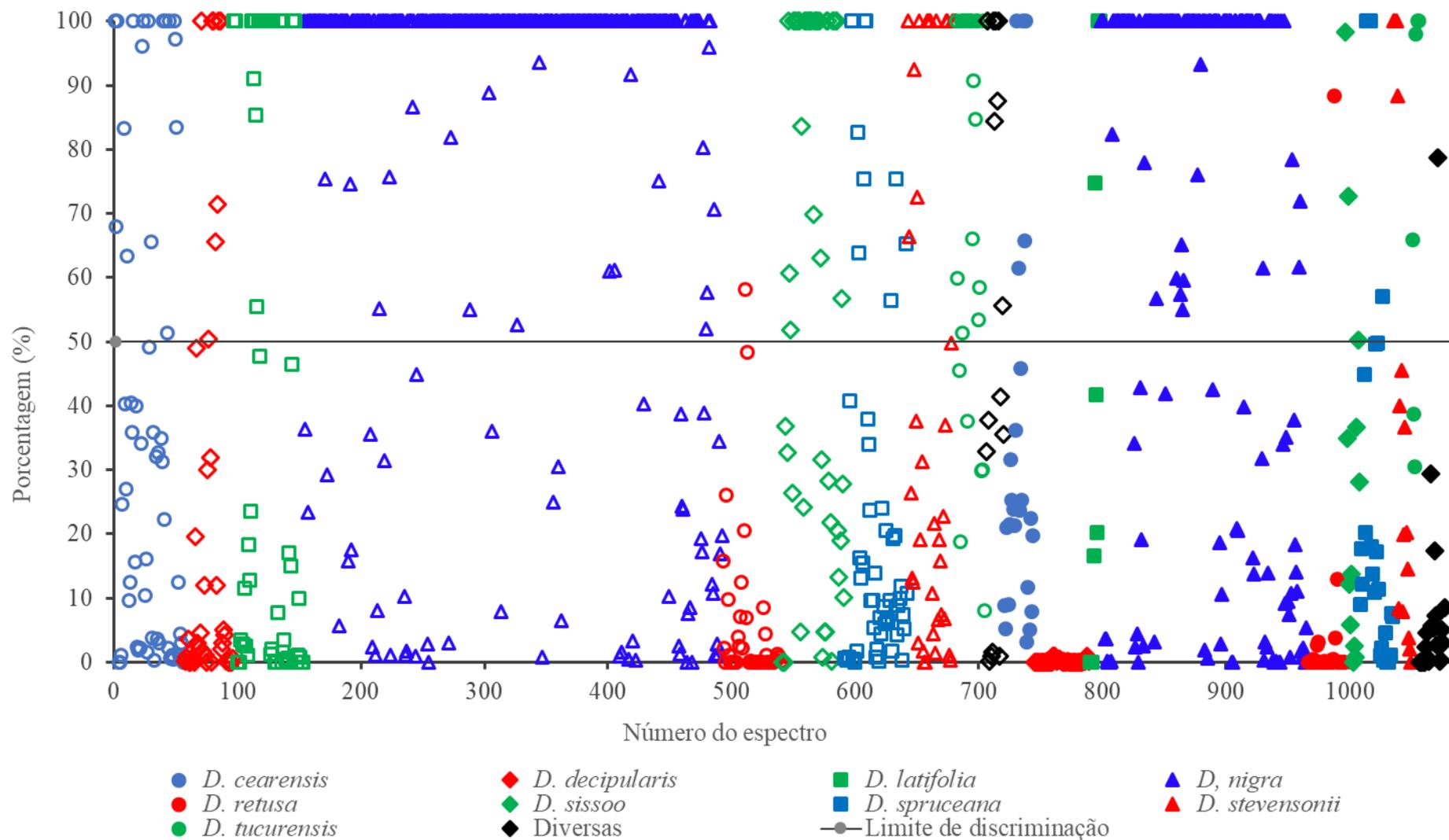


Figura 25. Probabilidade do espectro pertencer à espécie *D. nigra* segundo modelo SIMCA com seleção de variável do tipo 2.

As figuras de mérito de ambas análises realizadas com SIMCA estão na Tabela 5 na qual observa-se que ao usar maior região espectral a eficiência da discriminação é 10 % maior do que com região espectral reduzida. Ao usar todo o espectro o modelo discriminou corretamente duas amostras de *D. nigra* a mais e quatro amostras a mais das demais *Dalbergias*.

Tanto as análises PLS-DA quanto SIMCA apresentaram eficiências superiores ao utilizar toda a região espectral do equipamento, sugerindo que as informações espectrais de regiões além das que representam os principais constituintes da madeira são relevantes para discriminar *D. nigra* das demais amostras avaliadas.

Tabela 5. Figuras de mérito obtidas com base no conjunto de validação da discriminação de *D. nigra* com ferramenta SIMCA.

Figuras de mérito			
1- $\lambda = 1.622 - 1.852 \text{ nm} + 1.960 - 2.307 \text{ nm}$		2- $\lambda = 1.622 - 2.367 \text{ nm}$	
CP	7	CP	8
FP	12	FP	8
FN	20	FN	18
VP	36	VP	38
VN	51	VN	55
TFP (%)	19,0	TFP (%)	12,7
TFN (%)	35,7	TFN (%)	32,1
EF (%)	45,2	EF (%)	55,2

CP – Componentes Principais, **FP** – Falso Positivo, **FN** – Falso Negativo, **VP** – Verdadeiro Positivo, **VN** – Verdadeiro Negativo, **TFP** – Taxa de Falso Positivo, **TFN** – Taxa de Falso Negativo e **EF** – Eficiência.

As Tabelas 6 e 7 são as tabelas confusões das duas análises SIMCA. Essas tabelas indicam com mais detalhes quais amostras foram corretamente e incorretamente classificadas. O somatório de uma linha indica a quantidade de amostras daquela espécie na fase de validação do modelo. Já as quantidades em cada coluna indicam como o modelo estimou a classificação dessas espécies. No modelo de classificação com seleção de variável do tipo 1 observa-se que haviam 9 amostras de *D. cearensis* na fase de validação (somatório da primeira linha) e o modelo classificou corretamente 6 dessas amostras (Não predita como *D. nigra*) enquanto as outras três amostras foram classificadas como *D. nigra* segundo esse modelo (Tabela 6).

Tabela 6. Tabela confusão do modelo SIMCA com seleção variável do tipo 1.

Classe	Predita como <i>D. nigra</i>	Não predita como <i>D. nigra</i>
<i>D. cearensis</i>	3	6
<i>D. decipularis</i>	2	10
<i>D. latifolia</i>	1	5
<i>D. nigra</i>	36	20
<i>D. retusa</i>	0	8
<i>D. sissoo</i>	1	6
<i>D. spruceana</i>	1	9
<i>D. stevensonii</i>	2	1
<i>D. tucurensis</i>	2	0
Diversas	0	6

Tabela 7. Tabela confusão do modelo SIMCA com seleção de variável do tipo 2.

Classe	Predita como <i>D. nigra</i>	Não predita como <i>D. nigra</i>
<i>D. cearensis</i>	2	6
<i>D. decipularis</i>	1	14
<i>D. latifolia</i>	1	2
<i>D. nigra</i>	38	18
<i>D. retusa</i>	0	10
<i>D. sissoo</i>	1	3
<i>D. spruceana</i>	1	8
<i>D. stevensonii</i>	1	4
<i>D. tucurensis</i>	1	1
Diversas	0	7

Oito amostras das demais espécies de *Dalbergias* foram erroneamente classificadas como *D. nigra* (FP) no modelo com seleção de variável do tipo 2, a tabela de confusão desse modelo (Tabela 7) indica a quais espécies essas amostras pertencem: duas *D. cearensis*; uma *D. decipularis*, uma *D. latifolia*; uma *D. sissoo*, uma *D. spruceana*, uma *D. stevensonii* e uma *D. tucurensis*.

É válido chamar a atenção que nenhuma amostra de *D. retusa* em nenhum dos modelos apresentados até o momento, incluindo os modelos PLS-DA, foi erroneamente classificada

como *D. nigra*. Indicando que a classe de *D. retusa* criada com as informações espectrais das superfícies dessas amostras é razoavelmente diferente das informações da *D. nigra* apesar das diversas características anatômicas em comum entre essas duas espécies.²⁶

Ao avaliar as eficiências dos 4 modelos criados para discriminar a *D. nigra* conclui-se que a ferramenta PLS-DA (81,6 % e 82,5 %) apresentou eficiência discriminatória significativamente maior do que a ferramenta SIMCA (45,2 % e 55,2 %). Essa diferença discriminatória pode ocorrer devido ao fato de que a decomposição das matrizes feita pelo algoritmo PLS-DA é feita de forma a maximizar a correlação entre a matriz espectral e o vetor de classe enquanto o SIMCA busca os principais componentes que representam a matriz espectral, sem relacioná-las às classes. A ferramenta PLS-DA apresentou valores altos de eficiência para discriminar a *D. nigra*, espécie essa com grande importância nacional e que está em risco de extinção, sendo que uma das grandes barreiras para evitar a extinção é a discriminação correta da *D. nigra* de outras *Dalbergias* visualmente, e também anatomicamente, semelhantes.

5.2 Discriminação da *Dalbergia spruceana* das demais 14 espécies de *Dalbergia*

5.2.1 PLS-DA

Para a construção do modelo PLS-DA da discriminação de *D. spruceana* a escolha do número de variáveis latentes foi feita de acordo com o valor de RMSECV, como descrito anteriormente. A Figura 26 indica a variação de RMSECV de acordo com número de variáveis latentes usadas para construção do modelo no qual se utilizou todo o espectro disponível no equipamento.

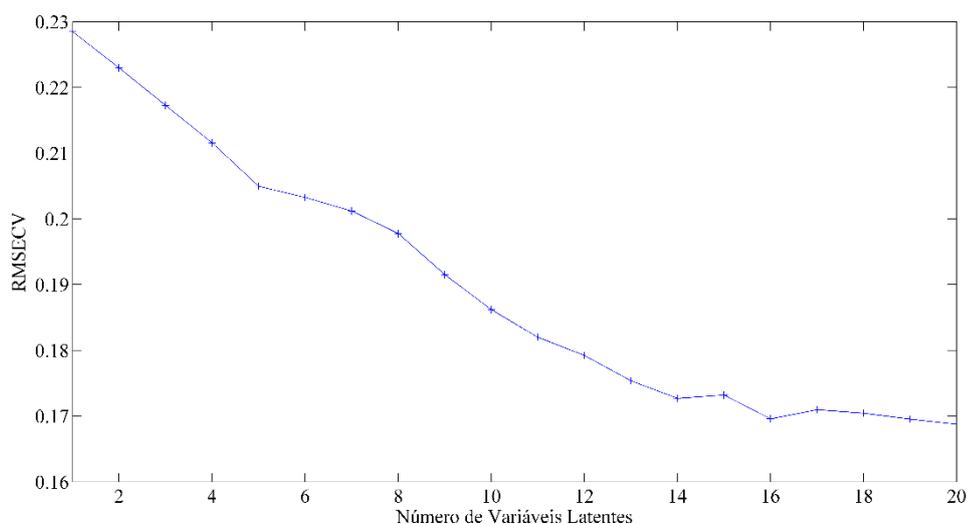


Figura 26. Valor de erro de validação cruzada (RMSECV) de acordo com número de variáveis Latentes do modelo de discriminação de *D. spruceana* com seleção de variável do tipo 2.

Para o modelo com seleção de variável do tipo 2 escolheu-se 14 variáveis latentes para a construção do modelo. Já para o modelo com seleção de variável do tipo 1 foram escolhidas 13 variáveis latentes (Tabela 8). A exclusão de amostras anômalas seguiu o mesmo procedimento como na discriminação da *D. nigra*. As Figuras 27A e 27B indicam os valores de Resíduos Q e T² de Hottelling dos espectros da fase de treinamento, respectivamente, dos modelos com seleção de variáveis do tipo 1 e do tipo 2. A Tabela 8 indica a quantidade total de espectros anômalos que foram excluídos na fase de treinamento, as quantidades da tabela são maiores do que a quantidade de pontos no quadrante superior direito, pois os valores da Tabela 8 incluem amostras excluídas de acordo com o parâmetro y.

Tabela 8. Quantidade de variáveis latentes e espectros anômalos nas fases de treinamento e validação dos dois tipos de modelos de discriminação de *D. spruceana*.

	1- $\lambda = 1.622 - 1.852 \text{ nm} + 1.960 - 2.307 \text{ nm}$	2- $\lambda = 1.622 - 2.367 \text{ nm}$
Número de variáveis latentes	13	14
Espectros anômalos na fase de treinamento	8	10
Espectros anômalos na fase de validação	5	15

Observa-se na Tabela 8 que o modelo com maior faixa espectral possui maior quantidade de espectros anômalos: 10 e 15. Esses valores representam apenas 1,4 % e 4,2 % das amostras nas fases de treinamento e validação, respectivamente. Esse baixo percentual de espectros anômalos indica que o modelo construído com os espectros de treinamento é válido para a análise das amostras de validação.

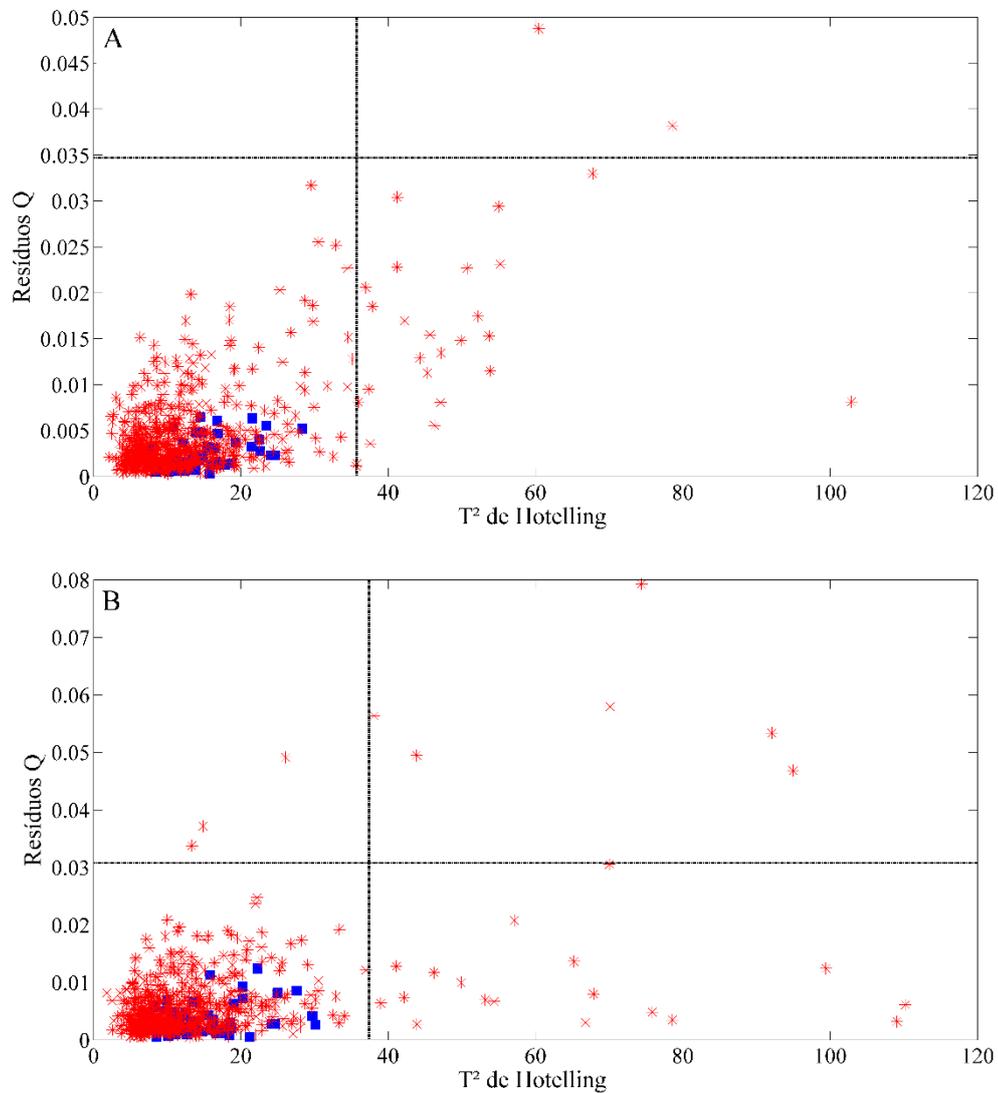


Figura 27. Gráfico dos valores de T^2 de Hotelling e resíduos Q dos espectros de treinamento do modelo com seleção de variável tipo 1 (a) e com seleção de variável tipo 2 (b). Legenda: *D. spruceana* (■) e Demais Dalbergias (*).

A figura 28A e 28B indicam os valores de Resíduos Q e T^2 de Hotelling dos espectros da fase de treinamento e validação, respectivamente, dos modelos com seleção de variáveis do tipo 1 e do tipo 2.

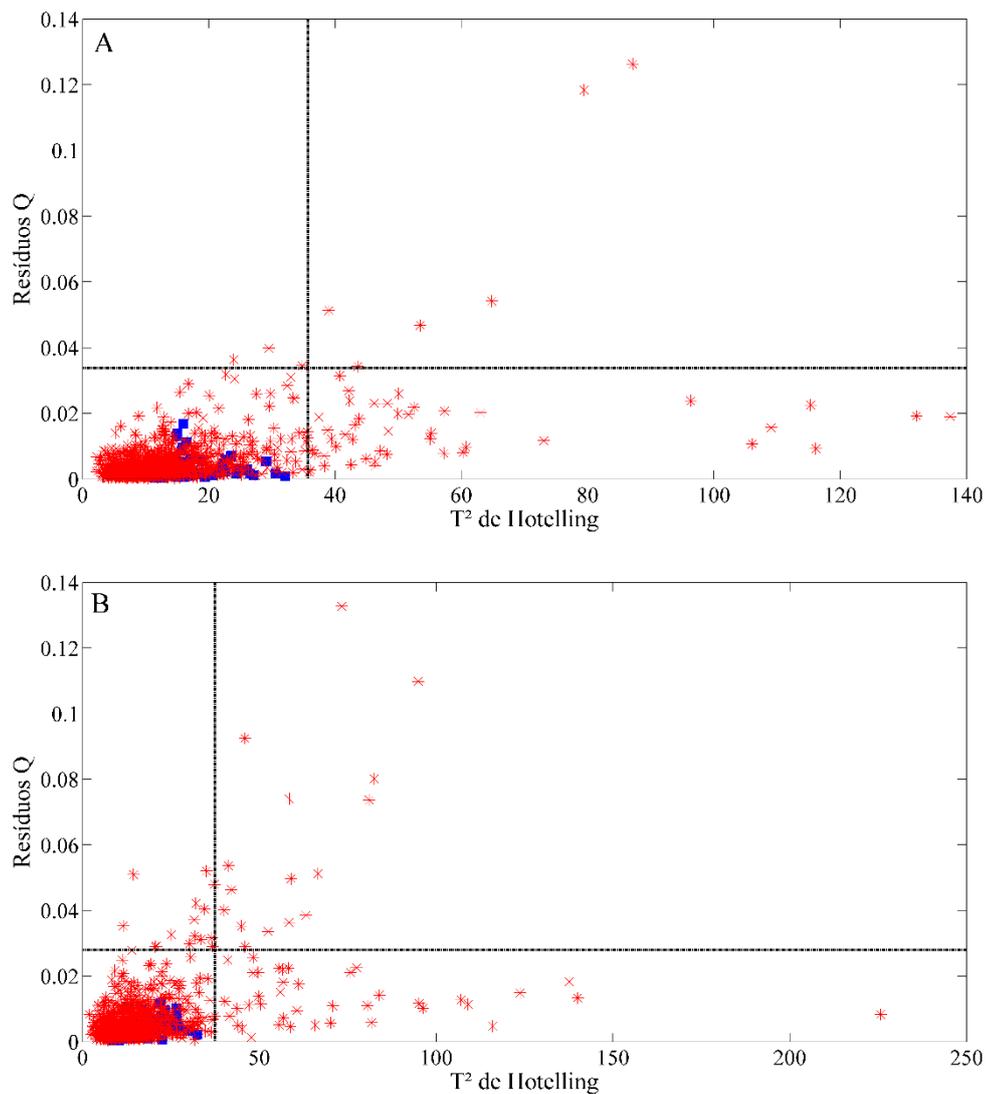


Figura 28. Gráfico dos valores de T^2 de Hotelling e resíduos Q dos espectros da fase de treinamento e de validação do modelo com seleção de variável tipo 1 (a) e com seleção de variável tipo 2 (b). Legenda: *D. spruceana* (■) e Demais *Dalbergias* (*).

As Figuras 29 e 30 indicam os valores de y e preditos para os espectros das amostras de acordo, respectivamente, com o modelo com seleção de variáveis do tipo 1 e 2.

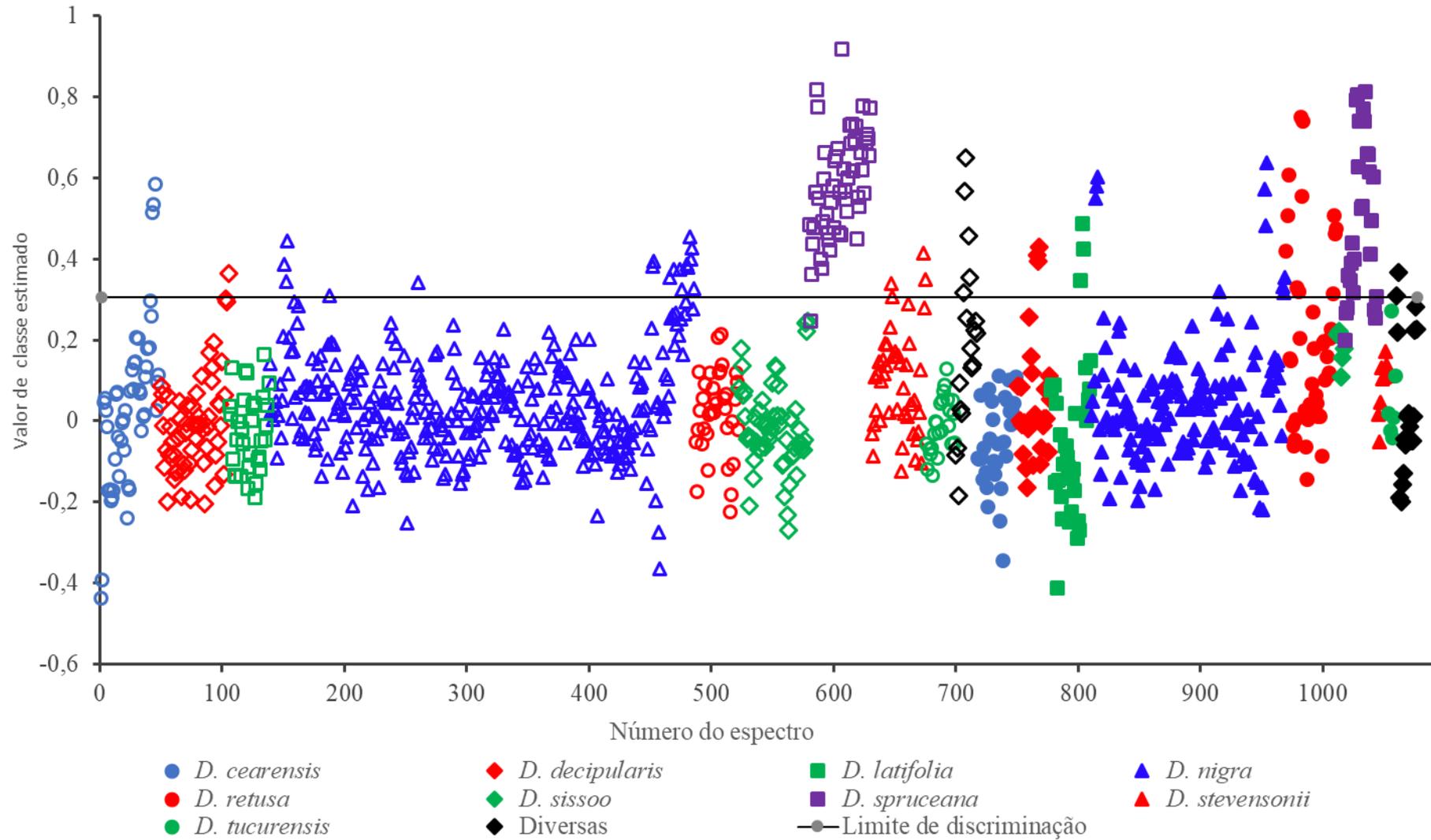


Figura 29. Valor de classe predito dos espectros do modelo PLS-DA com seleção de variável do tipo 1.

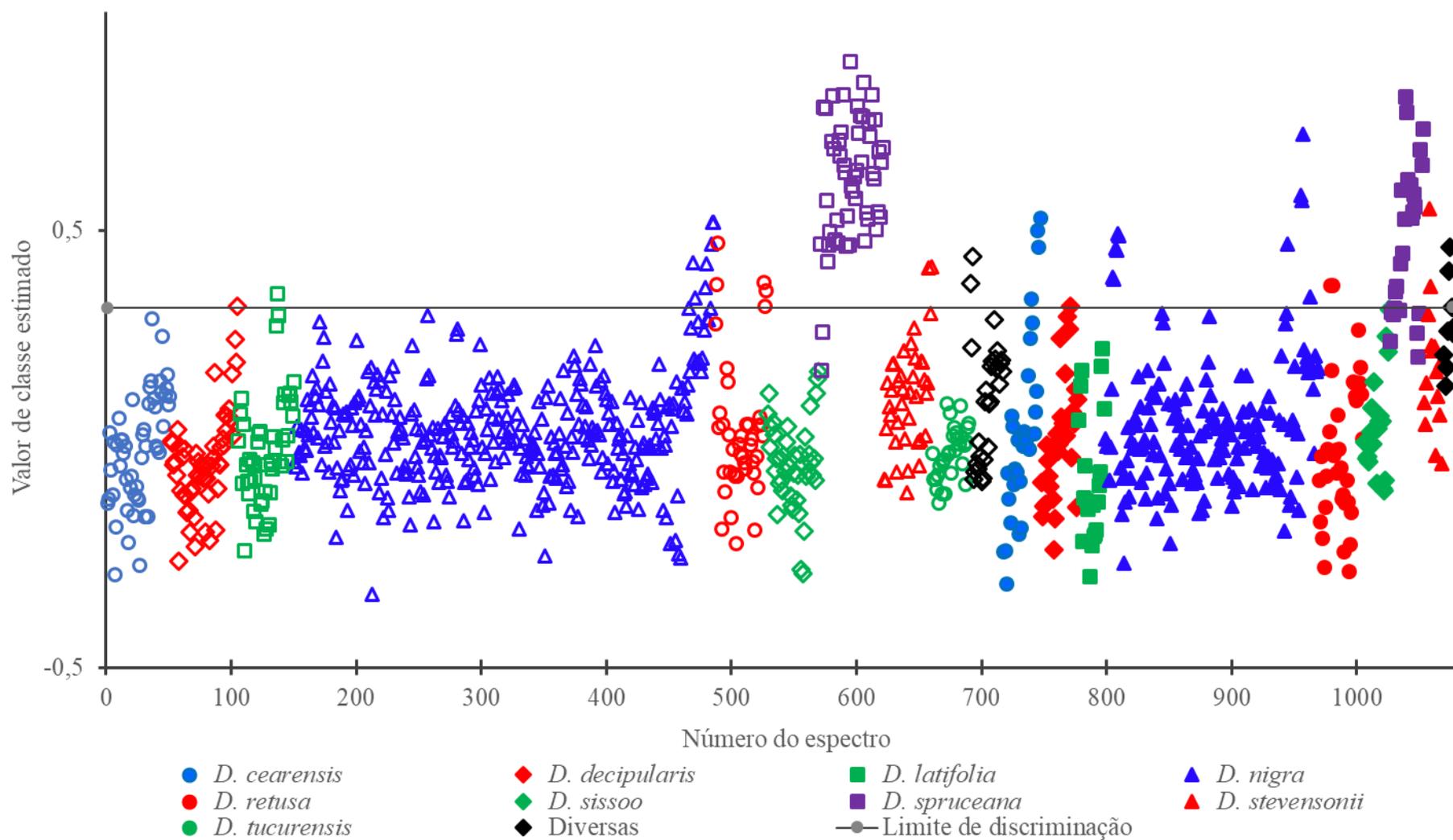


Figura 30. Valor de classe predito dos espectros do modelo PLS-DA com seleção de variável do tipo 2.

Observa-se na Tabela 9 que o modelo com seleção de variável de acordo com os principais componentes da madeira apresenta uma eficiência superior em 7,3 % do que o modelo que utiliza todo o espectro do instrumento portátil. Apesar de que o modelo com seleção do tipo 2 apresenta menor taxa de falso positivo (5,3 %) do que o primeiro (9,1 %), o modelo com menor região espectral foi capaz de predizer uma amostra de *D. spruceana* de forma correta (VP) a mais do que o segundo modelo, tendo assim uma taxa de falso negativo 11,1 % menor do que o segundo. Como há 9 amostras de *D. spruceana* na fase de validação apenas uma amostra de *D. spruceana* não classificada corretamente diminui a eficiência em 11,1 %.

Tabela 9. Figuras de mérito obtidas com base no conjunto de validação da discriminação de *D. spruceana* utilizando ferramenta PLS-DA utilizando os dois tipos de seleção de variável.

Figuras de mérito			
1- $\lambda = 1.622 - 1.852 \text{ nm} + 1.960 - 2.307 \text{ nm}$		2- $\lambda = 1.622 - 2.367 \text{ nm}$	
FP	10	FP	6
FN	2	FN	3
VP	7	VP	6
VN	100	VN	107
TFP (%)	9,1	TFP (%)	5,3
TFN (%)	22,2	TFN (%)	33,3
EF (%)	68,7	EF (%)	61,4

FP – Falso Positivo, **FN** – Falso Negativo, **VP** – Verdadeiro Positivo, **VN** – Verdadeiro Negativo, **TFP** – Taxa de Falso Positivo, **TFN** – Taxa de Falso Negativo e **EF** – Eficiência.

É importante notar o decaimento da eficiência de discriminação da *D. spruceana* com modelo PLS-DA em relação à discriminação de *D. nigra* com a mesma ferramenta (81,6 % e 82,5 %). Tendo em vista a grande diferença de quantidade de amostras entre essas duas *Dalbergias*, já se esperava que a predição da *D. spruceana* fosse menos eficiente que a predição da *D. nigra*. Ao modelar uma classe com muitas amostras cria-se um modelo com bastante informação a respeito da determinada classe, sendo assim o modelo é capaz de predizer amostras de validação com maior eficiência. A relativa pequena quantidade de amostras de *D. spruceana* é um fator que influenciou os valores de eficiência discriminatória.

5.2.2 SIMCA

Para a discriminação da *D. spruceana* utilizando a ferramenta SIMCA criou-se uma região que caracterizava essa classe com o uso de componentes principais (CP). Em ambos modelos utilizou-se 4 componentes principais. A metodologia para escolher a quantidade de CP é a mesma que se utilizou na discriminação de *D. nigra*.

As probabilidades dos espectros pertencerem à espécie *D. spruceana* ao utilizar a região espectral do tipo 1 estão indicadas na Figura 31 e as probabilidades utilizando a região espectral do tipo 2 estão indicadas na Figura 32.

Verifica-se que ao utilizar maior região espectral no modelo SIMCA a eficiência aumenta em 11,1 %, sugerindo novamente que informações de componentes minoritários da madeira são influentes na discriminação de espécies. Esse aumento de eficiência se deve ao fato de que o modelo com seleção espectral do tipo 2 foi capaz de predizer corretamente uma amostra a mais de *D. spruceana*. A capacidade de discriminar as amostras das outras espécies foi concordante entre os dois modelos, ambos predisseram corretamente 110 de 111 amostras (Tabela 10).

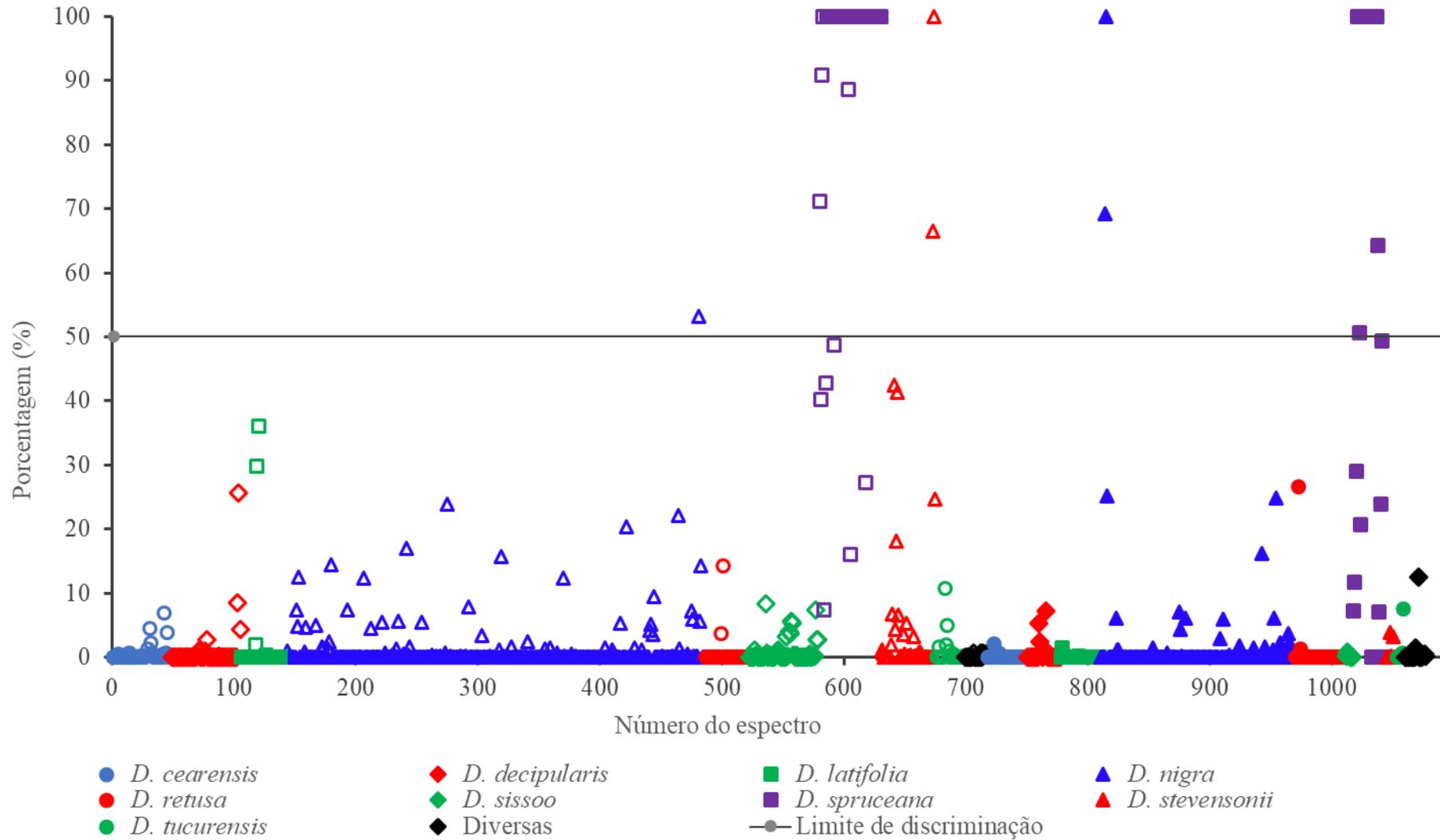


Figura 31. Probabilidade do espectro pertencer a espécie *D. spruceana* utilizando o modelo com seleção de variável do tipo 1.

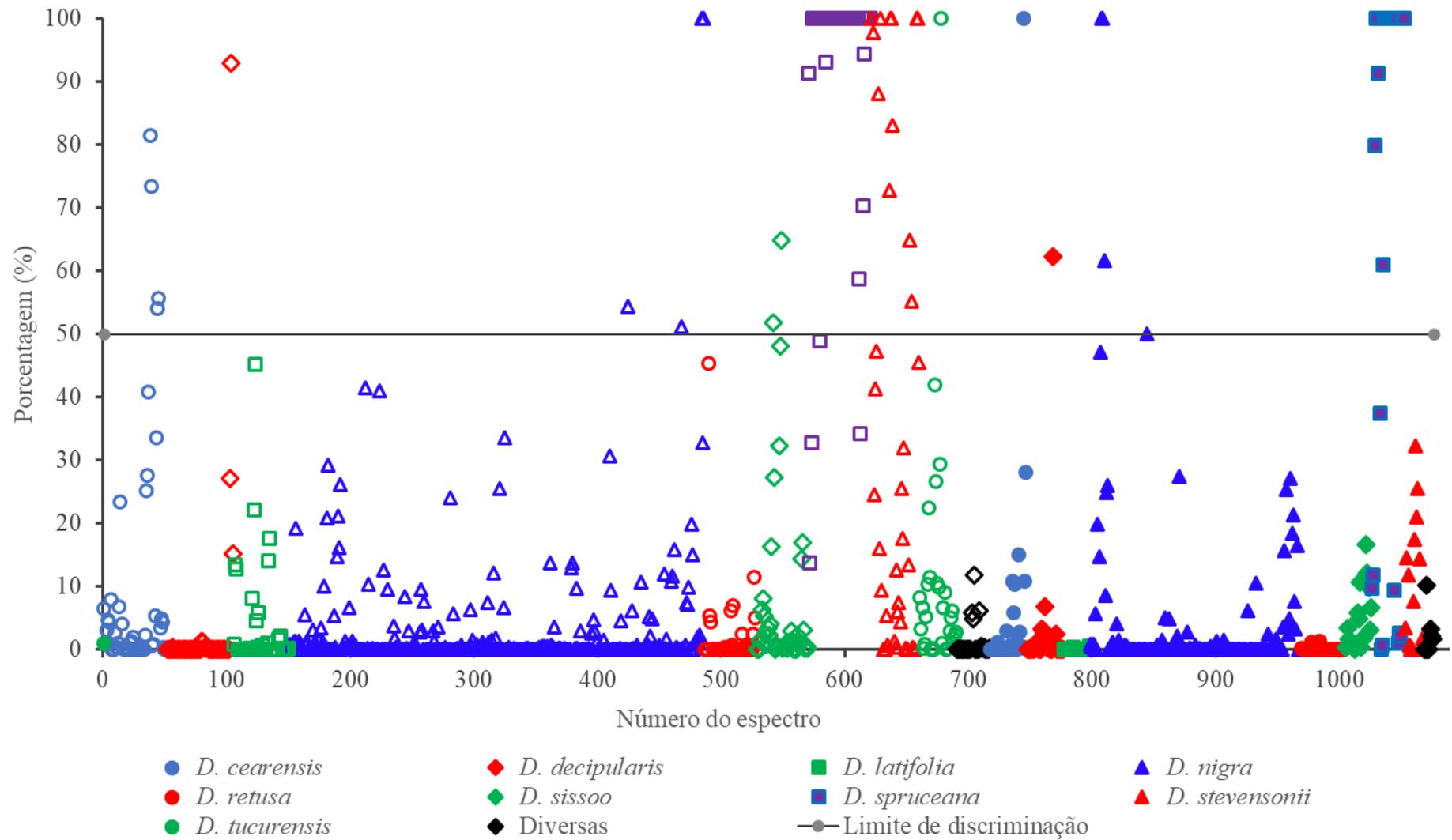


Figura 32. Probabilidade do espectro pertencer à espécie *D. spruceana* utilizando o modelo com seleção de variável do tipo 2.

Tabela 10. Figuras de méritos obtidas com base no conjunto de validação dos modelos de discriminação da *D. spruceana* utilizando a ferramenta SIMCA.

Figuras de mérito			
1- $\lambda = 1.622 - 1.852 \text{ nm} + 1.960 - 2.307 \text{ nm}$		2- $\lambda = 1.622 - 2.367 \text{ nm}$	
CP	4	CP	4
FP	1	FP	1
FN	4	FN	3
VP	5	VP	6
VN	110	VN	110
TFP (%)	0,9	TFP (%)	0,9
TFN (%)	44,4	TFN (%)	33,3
EF (%)	54,7	EF (%)	65,8

CP – Componentes Principais, **FP** – Falso Positivo, **FN** – Falso Negativo, **VP** – Verdadeiro Positivo, **VN** – Verdadeiro Negativo, **TFP** – Taxa de Falso Positivo, **TFN** – Taxa de Falso Negativo e **EF** – Eficiência.

As Tabelas 11 e 12 indicam como as amostras foram preditas, respectivamente, nos modelos com seleção de variáveis do tipo 1 e do tipo 2. Como só há uma amostra falsa positiva em ambos os modelos a TFP é baixa (0,9 %), porém os modelos erraram 44,4 % e 33,3 % das amostras de *D. spruceana*, tendo como resultado eficiências de 54,7 % e 65,8 %.

Tabela 11. Tabela confusão do modelo SIMCA com seleção variável do tipo 1.

Classe	Predita como <i>D. spruceana</i>	Não predita como <i>D. spruceana</i>
<i>D. cearensis</i>	0	11
<i>D. decipularis</i>	0	9
<i>D. latifolia</i>	0	11
<i>D. nigra</i>	1	52
<i>D. retusa</i>	0	14
<i>D. sissoo</i>	0	2
<i>D. spruceana</i>	5	4
<i>D. stevensonii</i>	0	3
<i>D. tucurensis</i>	0	2
Diversas	0	6

Tabela 12. Tabela confusão do modelo SIMCA com seleção variável do tipo 2.

Classe	Predita como <i>D. spruceana</i>	Não predita como <i>D. spruceana</i>
<i>D. cearensis</i>	0	10
<i>D. decipularis</i>	0	10
<i>D. latifolia</i>	0	7
<i>D. nigra</i>	1	56
<i>D. retusa</i>	0	12
<i>D. sissoo</i>	0	7
<i>D. spruceana</i>	6	3
<i>D. stevensonii</i>	0	5
Diversas	0	3

Observa-se que os modelos SIMCA para discriminação de *D. spruceana* apresentaram maiores eficiências em comparação aos mesmos tipos de modelos na discriminação de *D. nigra* (45,2 % e 55,2 %). É interessante ressaltar que nos modelos para discriminação de *D. spruceana* apenas uma amostra foi considerada falsa positiva, enquanto nos modelos de *D. nigra* houveram 8 e 12 amostras falsas positivas.

Futuros trabalhos devem ser direcionados ao aumento de dados espectrais dessa espécie, o que não é uma tarefa fácil, tendo em vista que ao visitar 6 coleções, incluindo a coleção Samuel J. Record e a coleção original de Madson que são internacionalmente conhecidas devido a sua história e a enorme quantidade de amostras, apenas 26 amostras de *D. spruceana* foram encontradas e tiveram os espectros coletados.

5.3 Discriminação individual das nove espécies de *Dalbergia*

Nesta etapa criou-se uma classe individual para as espécies que contém ao menos 10 amostras: *D. cearenses*, *D. descipularis*, *D. latifolia*, *D. retusa*, *D. sissoo*, *D. stevensonii* e *D. tucurensis*, com o intuito de analisar como os modelos PLS-DA e SIMCA discriminam essas espécies. As espécies com poucas amostras não tiveram classes específicas criadas, pois não haveria amostras suficientes para construir um modelo razoável e mais amostras para a fase de validação. Logo, para as espécies *D. bariensis*, *D. congestiflora*, *D. glaucescens*, *D. melanoxylon*, *D. miscolobium* e *D. villosa* não se criou uma classe específica, mas essas amostras foram usadas na fase de validação. A utilização dessas espécies na fase de validação visa analisar como os modelos predizem amostras de classes que não fizeram parte da fase de treinamento.

Nessa análise individual das espécies decidiu-se utilizar apenas a seleção de variável do tipo 2, pois ela se mostrou mais eficiente nos modelos anteriores.

5.3.1 PLS-DA

As escolhas das variáveis latentes para os modelos individuais seguiram o mesmo procedimento adotado nas análises anteriores: a relação de RMSECV e o número de variáveis latentes. A exclusão de espectros anômalos, assim como nas outras análises, foi baseada nos valores de resíduos Q, T² de Hotelling e valor estimado de classe (predito de y). A Tabela 13 indica a quantidade de variáveis latentes utilizadas em cada modelo e o número de espectros anômalos tanto na fase de treinamento quanto na fase de validação dos modelos individuais.

Observa-se na Tabela 13 que o número de espectros anômalos, tanto na fase de treinamento quanto na de validação, foi relativamente baixo. Na fase de treinamento o modelo que apresenta maior número de espectros anômalos é o da espécie *D. retusa* (14). Essa quantidade representa apenas 2,0 % dos espectros na fase de treinamento. Na fase de validação o modelo que apresenta maior número de espectros anômalos é o da *D. sissoo* (22) e representa apenas 5,9 % das amostras dessa fase. Tendo em vista que as porcentagens de amostras anômalas são baixas conclui-se que os modelos são válidos para essas discriminações.

A quantidade alta de variáveis latentes utilizadas se justifica devido ao fato que a madeira é um material complexo e apesar de que há diferenças na constituição da madeira de acordo com a espécie, a estrutura em geral da madeira é parecida, fazendo necessário a utilização de mais variáveis latentes para discriminar uma espécie da outra. Entretanto, chama-se a atenção para o modelo de *D. retusa* que utiliza apenas 7 variáveis latentes indicando que

nas amostras de *D. retusa* estudadas nessa dissertação há pouca variação na informação espectral sendo necessário poucas variáveis latentes para descrever essa classe.

Tabela 13. Quantidade de variáveis latentes e espectros anômalos nas fases de treinamento e validação dos modelos individuais.

	Espectros anômalos (treinamento)	Espectros anômalos (validação)	Número de variáveis latentes
<i>D. cearensis</i>	9	14	14
<i>D. decipularis</i>	11	14	9
<i>D. latifolia</i>	6	14	14
<i>D. nigra</i>	10	16	14
<i>D. retusa</i>	14	12	7
<i>D. sissoo</i>	3	22	17
<i>D. spruceana</i>	8	6	14
<i>D. stevensonii</i>	12	18	10
<i>D. tucurensis</i>	3	20	19

Observa-se na Tabela 14 que de forma geral o modelo PLS-DA alcançou boa eficiência na discriminação das espécies. Das 9 espécies, 7 alcançaram eficiência superior a 75 %. A *D. decipularis* (57,9 %) e a *D. spruceana* (44,8 %) foram as que obtiveram a menor eficiência discriminatória. Suas baixas eficiências foram muito influenciadas pela TFN, *D. decipularis* (33,3 %) e *D. spruceana* (50,0 %) indicando a dificuldade desses modelos de identificar como positivas as amostras dessas espécies, o que pode representar uma necessidade de um maior número de amostras de treinamento para essas espécies.

Chama-se a atenção para a espécie *D. sissoo* e *D. tucurensis* que obtiveram, respectivamente, 100 % e 95 % de eficiência. A alta eficiência discriminatória da espécie *D. sissoo* vem em concordância com resultado obtido por Snel *et al*⁹⁸ no qual discriminou-se *D. sissoo* com 97,6 % de eficiência. Os altos valores de eficiência indicam que os espectros coletados dessas espécies apresentam características espectrais diferentes das demais espécies estudadas, porém deve-se observar que a grande quantidade de espectros anômalos de validação sugere um pequeno sobreajuste no modelo.

Tabela 14. Figuras de méritos obtidas com base no conjunto de validação para os modelos individuais com a ferramenta PLS-DA.

	VP	VN	FP	FN	TFP	TFN	EF
<i>D. cearensis</i>	8	103	11	1	9,6	11,1	79,2
<i>D. decipularis</i>	6	104	10	3	8,8	33,3	57,9
<i>D. latifolia</i>	6	104	13	0	11,1	0,0	88,9
<i>D. nigra</i>	51	64	3	5	4,5	8,9	86,6
<i>D. retusa</i>	6	110	5	1	4,4	14,3	81,4
<i>D. sissoo</i>	7	112	0	0	0,0	0,0	100,0
<i>D. spruceana</i>	4	110	6	4	5,2	50,0	44,8
<i>D. stevensonii</i>	5	107	9	1	7,8	16,7	75,6
<i>D. tucurensis</i>	3	113	6	0	5,0	0,0	95,0

FP – Falso Positivo, **FN** – Falso Negativo, **VP** – Verdadeiro Positivo, **VN** – Verdadeiro Negativo, **TFP** – Taxa de Falso Positivo, **TFN** – Taxa de Falso Negativo e **EF** – Eficiência.

É interessante observar que, apesar da *D. cearensis* ser uma espécie que é facilmente distinta das demais através de sua anatomia, a discriminação dessa espécie com a técnica do infravermelho acoplada às análises multivariadas não foi a espécie que apresentou maior eficiência.²⁵

5.3.2 SIMCA

Para a criação do modelo SIMCA a escolha de componentes principais se deu da mesma forma que as análises anteriores, para todas as classes se utilizou 8 componentes principais. Diferentemente do modelo PLS-DA que discrimina de forma binária as amostras da fase de validação, o modelo SIMCA cria 9 regiões, uma para cada espécie, e na fase de validação analisa as chances das amostras pertencerem a essas classes. Sendo assim, as informações espectrais para cada classe são consideravelmente menores em relação a classe zero do modelo PLS-DA, tendo em vista que apenas uma espécie possui menos informação espectral do que as demais 8 espécies juntas.

Os valores das figuras de mérito para os 9 modelos estão indicados na Tabela 15. Observa-se que as eficiências dos modelos individuais utilizando SIMCA são muito inferiores em relação ao modelo PLS-DA. Apenas o modelo de discriminação da *D. nigra* apresenta eficiência acima de 60 %. Tendo em vista que essa é a espécie com maior número de amostras (169) é possível concluir que como há mais amostras de *D. nigra* foi possível modelar melhor

a classe dessa espécie levando a uma eficiência maior em relação às demais. É importante apontar que as taxas de falso positivo (TFP) são baixas, sendo que na maioria dos modelos a taxa é zero. Isso indica que apesar das eficiências com SIMCA serem baixas, poucas espécies são confundidas entre si. Contudo, apesar de haver muitas amostras falsas negativas, a grande maioria dessas amostras não foram preditas em outras classes, mas sim como não pertencente a nenhuma classe. Novamente, esse efeito pode refletir uma necessidade de maior número de amostras.

Tabela 15. Figuras de mérito do modelo discriminatório das nove espécies de *Dalbergia* utilizando a ferramenta SIMCA.

	VP	VN	FP	FN	TFP	TFN	EF
<i>D. cearensis</i>	3	115	1	6	0,86	66,67	32,47
<i>D. decipularis</i>	5	116	0	4	0,00	44,44	55,56
<i>D. latifolia</i>	0	118	0	7	0,00	100,00	0,00
<i>D. nigra</i>	40	64	5	16	7,25	28,57	64,18
<i>D. retusa</i>	2	117	0	6	0,00	75,00	25,00
<i>D. sissoo</i>	0	118	0	7	0,00	100,00	0,00
<i>D. spruceana</i>	2	117	0	6	0,00	75,00	25,00
<i>D. stevensonii</i>	0	119	0	6	0,00	100,00	0,00
<i>D. tucurensis</i>	0	122	0	3	0,00	100,00	0,00

A Tabela 16 é a tabela confusão do modelo SIMCA, na qual indica-se quais espécies foram confundidas entre si. Como destacado anteriormente, poucas amostras foram classificadas como pertencentes a outras classes de forma errônea. Essas amostras são: uma *D. cearensis* que foi predita como pertencente à espécie *D. nigra*, uma *D. stevensonii* predita como *D. cearensis*, duas *D. stevensonii* preditas como *D. nigra*, uma *D. miscolobium* predita como *D. nigra* e uma amostra de *D. congestiflora* também predita como *D. nigra*. O que causa a baixa eficiência é a dificuldade dos modelos de classificar corretamente a espécie discriminada. Observa-se que em alguns modelos nenhuma amostra foi predita a pertencer a sua classe de forma correta, como é o caso da *D. latifolia* que todas as suas 7 amostras foram preditas como não pertencentes a nenhuma classe.

Tabela 16. Tabela confusão do modelo discriminatório das nove espécies do gênero *Dalbergia*.

Classe predita/classe	<i>D. cearensis</i>	<i>D. decipularis</i>	<i>D. latifolia</i>	<i>D. nigra</i>	<i>D. retusa</i>	<i>D. sissoo</i>	<i>D. spruceana</i>	<i>D. stevensonii</i>	<i>D. tucurensis</i>	<i>Diversas</i>
<i>D. cearensis</i>	3							1		
<i>D. decipularis</i>		5								
<i>D. latifolia</i>			0							
<i>D. nigra</i>	1			40				2		2
<i>D. retusa</i>					2					
<i>D. sissoo</i>						0				
<i>D. spruceana</i>							2			
<i>D. stevensonii</i>								0		
<i>D. tucurensis</i>									0	
Não pertencente a nenhuma classe	5	4	7	16	6	7	6	3	3	10

5.3.3 Discriminações de espécies que não tiveram sua classe modelada na fase de treinamento.

Essa seção objetiva avaliar como as ferramentas SIMCA e PLS-DA discriminam/identificam amostras que não pertencem às classes que foram utilizadas para criar o modelo. As 12 amostras estudadas pertencem às espécies: *D. bariensis*, *D. congestiflora*, *D. glauscences*, *D. melanoxylon*, *D. miscolobiu* e *D. villosa*. Observa-se na Tabela 16 que através do modelo SIMCA apenas duas amostras foram erroneamente classificadas como *D. nigra*, alcançando um percentual de erro de classificação de 16,7 %.

Na análise PLS-DA criou-se um modelo para cada uma das 9 espécies que foram discriminadas. A Tabela 14 apresenta a quantidade de amostras que foram classificadas de forma errada em cada um dos modelos e o percentual.

Tabela 17. Quantidade e percentual de amostras do grupo “diversas” que o modelo PLS-DA discriminou de forma errônea.

Modelo	Amostras diversas preditas de forma errada.	Porcentual de amostras diversas preditas de forma errada
<i>D. cearensis</i>	5	41,7 %
<i>D. decipularis</i>	4	33,3 %
<i>D. latifolia</i>	4	33,3%
<i>D. nigra</i>	1	8,3 %
<i>D. retusa</i>	1	8,3 %
<i>D. sissoo</i>	0	0,0 %
<i>D. spruceana</i>	3	25,0 %
<i>D. stevensonii</i>	5	41,7 %
<i>D. tucurensis</i>	0	0,0%

Observa-se que os modelos de discriminação da espécie *D. cearensis*, *D. stevensonii*, *D. latifolia* e *D. decipularis* foram os modelos que mais erraram a predição das amostras da classe nominada “diversas”. Algo que esses modelos têm em comum é o fato de terem sido desenvolvidos com um conjunto reduzido de amostras de treinamento para suas respectivas classes.

A média percentual de erro dos 9 modelos do PLS-DA é 21,3 %, esse valor é 4,6 % superior ao valor alcançado pelo modelo SIMCA. Para o conjunto de amostras avaliados nessa

dissertação observou-se que há pouca diferença na eficiência da discriminação de amostras que não tiveram sua classe modelada na fase de treinamento nos modelos PLS-DA e SIMCA.

Tendo em vista que as duas ferramentas apresentaram eficiências próximas ao discriminar as espécies que não tiveram sua classe modelada na fase de treinamento e que PLS-DA apresentou eficiência significativamente maior que a ferramenta SIMCA em todas as análises infere-se que entre as ferramentas PLS-DA e SIMCA, a ferramenta PLS-DA é a mais indicada para discriminação dessas amostras de *Dalbergia*.

6. CONCLUSÕES

A primeira etapa desse trabalho confirmou a hipótese levantada de que a ferramenta espectroscopia NIR associada com PLS-DA é capaz de discriminar a espécie *D. nigra* das demais 14 espécies estudadas. Ambos os tipos de seleção de variável apresentaram altos valores de eficiência, mas levando em consideração as taxas de falso positivo concluiu-se que o melhor resultado foi ao utilizar a seleção de variável que foca nas bandas dos principais componentes da madeira (1.622 – 1.852 nm + 1.960 – 2.307 nm), pois apresentou taxa de falso positivo menor e eficiência discriminatória de 81,6 %. Baseados nesses resultados conclui-se que a espectroscopia NIRS associada à ferramenta PLS-DA é um método com grande potencial para identificação da espécie *D. nigra*.

A utilização da ferramenta SIMCA na discriminação das espécies de *D. nigra* não apresentou resultados satisfatórios, alcançando eficiências de apenas 45,2 % e 55,2 %. Ao menos com os resultados obtidos nessa dissertação, a hipótese levantada de que SIMCA seria uma ferramenta capaz de predizer corretamente se a amostra pertence ou não à classe *D. nigra* não pode ser considerada aceita.

Além de apresentar alta eficiência discriminatória, é importante ressaltar que essa ferramenta apresenta outras vantagens como: análise rápida, pouco preparo de amostra, equipamento de fácil manipulação e portátil, ou seja, pode ser utilizada em campo por um fiscal treinado. O treinamento necessário é de como manipular o equipamento e de preparar a amostra (lixar), algo que pode ser realizado em no máximo 4h. A eficiência, acompanhada da praticidade apresentada por esse método, pode ser um facilitador no controle do tráfico de madeira da espécie *D. nigra*. Trabalhos futuros tem como objetivo testar o uso desse método em postos de fiscalização para avaliar o comportamento da ferramenta em situações reais.

A ferramenta PLS-DA discriminou a espécie *D. spruceana*, também nativa do Brasil e semelhante à espécie *D. nigra*, com eficiências de 68,7 % e 61,4 % ao utilizar, respectivamente, a seleção de variável do tipo 1 e do tipo 2. A discriminação dessa espécie apresentou eficiência menor do que a discriminação da *D. nigra*. Nota-se que há diferença de informações espectrais da superfície das amostras de *D. spruceana* estudadas em relação às demais, porém os espectros de *D. spruceana* não foram capazes de descrever por completo a classe. Para a ferramenta SIMCA obteve-se 65,8 % ao utilizar a seleção de variável 2. Dessa forma, os resultados das duas ferramentas quimiométricas podem ser consideradas apenas regulares. Certamente a quantidade de amostras influenciou nos valores de eficiência alcançados por esse modelo, a

utilização de mais espectros para construção do modelo favorecerá a discriminação da madeira da espécie *D. spruceana*.

Em relação a faixa espectral nota-se que na maioria dos casos ao utilizar toda a região espectral há um considerável aumento de eficiência, o que sugere que compostos minoritários são importantes para a discriminação entre as espécies de *Dalbergia*.

Ao comparar os modelos PLS-DA e o modelo SIMCA das análises individuais de nove espécies de *Dalbergia* foi clara a vantagem do modelo PLS-DA em relação ao SIMCA. Levando em consideração a limitação do número de amostras, conclui-se que a ferramenta PLS-DA é mais indicada para discriminar as nove espécies estudadas nessa etapa.

A discriminação da espécie *D. sissoo* alcançou 100 % de eficiência, sugerindo que a composição química da superfície das amostras de *D. sissoo* difere das demais espécies. O alto valor de eficiência corrobora com os resultados do trabalho de Snel et al.⁹⁸ no qual utilizando também a ferramenta PLS-DA obtiveram eficiência discriminatória de 97,6 %.

Quanto à capacidade das ferramentas quimiométricas em discriminar espécies que não tiveram suas classes modeladas na fase de treinamento, pode-se considerar, ao menos para as espécies *D. nigra*, *D. retusa*, *D. sissoo* e *D. tucurensis*, que obteve-se bons resultados. Esse teste destaca a importância de se avaliar o modelo com amostras de classes diferentes das inseridas na fase de treinamento, pois em situações reais é possível que amostras duvidosas sejam de espécies diferentes das utilizadas na fase de treinamento.

7. REFERÊNCIAS

1. Carvalho, A. M.; *Brittonia* **1997**, 49, 87.
2. Lorenzi, H.; *Árvores brasileiras: manual de identificação e cultivo de plantas arbóreas nativas do Brasil*, 2nd ed.; Plantarum Ltda, Nova Udessa, São Paulo, Brasil, 1992.
3. Flynn, J. H.; Holder, C. D. *A guide to useful woods of the world*, 2^a ed.; Forest Products Society, Madison, EUA, 2001.
4. <https://www.wood-database.com/brazilian-rosewood/#pics>, acessado em 23 de outubro de 2018.
5. <https://casavogue.globo.com/Arquitetura/Edificios/noticia/2017/01/um-tour-arquitetonico-pelo-palacio-do-itamaraty.html>, acessado em 02 de junho de 2018.
6. http://www.fpl.fs.fed.us/documnts/fplrp/fpl_rp632.pdf, acessado em 12 de Dezembro de 2017.
7. <https://www.fws.gov/international/plants/current-cites-listings-of-tree-species.html>, acessado em 20 de Abril de 2018.
8. Boletim IBAMA 2018.
9. https://produto.mercadolivre.com.br/MLB-813201460-escala-de-jacaranda-da-amazniadalbergia-spruceana-_JM, acessado em 19 de outubro de 2018.
10. <https://cites.org/eng/app/appendices.php>, acessado em 19 de Abril de 2018.
11. Lancaster, C.; Espinoza, E.; *Rapid Commun. Mass Spectrom.* **2012**, 26, 1147.
12. UNODC - United Nations Office on Drugs and Crime. *Best practice guide for forensic timber identification*. United Nations, Nova Iorque, EUA, 2016.
13. Braga, J. W. B.; Pastore, T. C. M.; Coradin, V. T. R.; Camargos, J. A. A.; da Silva, A. R.; *IWA J.* **2011**, 32, 285.

14. Pastore, T. C. M.; Braga, J. W. B.; Coradin, V. T. R.; Magalhães, W. L. E.; Okino, E. Y. A.; Camargos, J. A. A.; De Muñiz, G. I. B.; Bressan, O. A.; Davrieux, F.; *Holzforschung* **2011**, *65*, 73.
15. Soares, L. F.; Silva, D. C. da; Bergo, M. C. J.; Coradin, V. T. R.; Braga, J. W. B.; Pastore, T. C. M.; *Quim. Nova* **2017**, *40*, 418.
16. Adedipe, O. E.; Dawson-Andoh A, B.; Slahor, J.; Osborn A, L.; *J. Near Infrared Spectrosc.* **2008**, *16*, 49.
17. Bergo, M. C. J.; Pastore, T. C. M.; Coradin, V. T. R.; Wiedenhoeft, A. C.; Braga, J. W. B.; *IWA J.* **2016**, *37*, 420.
18. Silva, D. C.; Pastore, T. C. M.; Soares, L. F.; Barros, F. A. S.; Bergo, M. C. J.; Coradin, V. T. H.; Gontijo, A. B.; Sosa, M. H.; Chacón, C. B.; Braga, J. W. B. *Holzforschung*, DOI: 10.1515/hf-2017-0160.
19. <https://www.wood-database.com/wood-articles/rosewoods-bubinga-really-banned-cites/>, acessado em 9 de outubro de 2018
20. Coradin, V.T.R., J.A.A. Camargos, L.F. Marques & E.R. da Silva Jr. 2009. Madeiras similares ao mogno (*Swietenia macrophylla King*): chave ilustrada para identificação anatômica em campo. Serviço Florestal Brasileiro, Brasília.
21. <https://www.wood-database.com/brazilian-rosewood/#pics>, acessado em 23 de outubro de 2018.
22. <https://www.wood-database.com/wood-articles/janka-hardness/>, acessado em outubro de 2018.
23. <https://www.wood-database.com/wood-articles/modulus-of-rupture/>, acessado em outubro de 2018.
24. <https://www.wood-database.com/wood-articles/crushing-strength/> acessado em outubro de 2018.

25. Gasson, P.; Miller, R.; Stekel, D. J.; Whinder, F.; Ziemińska, K.; *Ann. Bot.* **2010**, *105*, 45.
26. Gasson, P.; Baas, P.; Wheeler, E.; *IAWA J.* **2011**, *32*, 155.
27. <https://www.wood-database.com/rosewood-world-poster/>, acessado em outubro de 2018.
28. Lorenzi, H. *Árvores brasileiras: manual de identificação e cultivo de plantas arbóreas do Brasil*, vol. 1. 5ª. ed.; Plantarum: Nova Odessa, Brasil, 2008.
29. <https://www.delta-intkey.com/citeswood/en/www/fabfabri.htm>, acessado em outubro de 2018.
30. <http://rarewoodsandveneers.com>, acessado em outubro de 2018.
31. CITES (1992) Check list of CITES species. CITES. <http://checklistcites.org/#/en/search/outputlayout=alphabetical&leveloflisting=0&showsynonyms=1&showauthor1&showenglish=1&showspanish=1&showfrench=1&scientificname=Dalbergia+ni+gra&page=1&perpage=20>, acessado em 10 de março de 2017.
32. Portaria IBAMA nº 37-N, de 03 de Abril de 1992.
33. Portaria IBAMA nº 83, de 15 de outubro de 1996.
34. Camargos, J. A. A.; Coradin, V. T. R.; Czarneski, C. M.; de Oliveira, D.; Meguerditchian, I. *Catálogo de Árvores do Brasil*. 2ª ed.; IBAMA, Brasília, Brasil, 2001.
35. Pigozzo, R. J. B. Espectroscopia de infravermelho-próximo em madeiras neotropicais: aplicação na identificação e predição de propriedades físicas; Universidade de São Paulo: Dissertação de Mestrado, 2011.
36. Kite, G. C.; Green, P. W. C.; Veitch, N. C.; Groves, M. C.; Gasson, P. E.; Simmonds, M. S. J.; *Phytochemistry* **2010**, *71*, 1122.
37. Associação Brasileira De Normas Técnicas. Nbr 7190: Projeto De Estrutura De Madeira.
38. <http://www.florestal.gov.br/documentos/publicacoes/3230-boletim-snif-2017-ed1-final/file>, acessado em 21 de dezembro de 2018.
39. https://www.forest-trends.org/wp-content/uploads/imported/tropical-hardwood-flows-in-china-v12_12_3_2013-pdf.pdf, acessado em 20 de outubro de 2018.

40. Zhang, J.; Gan, J. Who Will Meet China's Demand for Imported Forest Products? *World Development* 2007, 35, 2150.
41. Giurca, A., Jonsson, R., Rinaldi, F.; Priyadi, H.; Ambiguity in Timber Trade Regarding Efforts to Combat Illegal Logging: Potential Impacts on Trade between Southeast Asia and Europe. *Forests* **2013**, 4, 730.
42. Gan, J.; Cerutti, P. O.; Masieiro, M.; Pettenella, D.; Andrighetto, N.; Dawson, T. *Quantifying Illegal Logging and Related Timber Trade*; Kleinschmit, D.; Mansourian, S.; Wildburger, C.; Purret, A., eds.; International Union of Forest Research Organizations: Vienna, Austria, 2016, ch 3.
43. http://www.cifor.org/publications/pdf_files/Books/BTaconi0401.pdf, acessado em 12 de Fevereiro de 2018.
44. <https://www.illegal-logging.info/sites/files/chlogging/uploads/madagascarreportrevisedfinalen.pdf>, acessado em 27 de Junho de 2017.
45. Gasson, P. *IAWA J.* **2011**, 32, 137.
46. Espinoza, E. O.; Wiemann, M. C.; Barajas-Morales, J.; Chavarria, G. D.; McClure, P. J.; *IAWA J.* **2015**, 36, 311.
47. <http://www.treesearch.fs.fed.us/pubs/40493>, acessado em 12 de Janeiro em 2018.
48. Kribs, D.A.; *Commercial foreign woods on the American market*; Edwards Brothers, Michigan, EUA, 1959.
49. Cody, R. B.; Dane, A. J.; Dawson-Andoh, B.; Adedipe, E.A.; Nkansah, K.; *J. Anal. Appl. Pyrolysis* **2012**, 95, 134.
50. Cody, R. B.; Laramée, J. A.; Durst, H. D.; *Anal Chem.* **2005**, 77, 2297.
51. Eyton, W. B.; Ollis, W. D.; Fineberg, M.; Gottlier, O. R.; Guimarães, S. S.; Magalhães, M. T. *Tetrahedron* 1965, 21, 2697.

52. Musah, R. A.; Espinoza, E. O.; Cody R. B.; Lesiak, A. D.; Christensen, E. D.; Moore, H. E.; Maleknia, S.; Drijfhout, F. P.; *Sci. Rep.* **2015**, *5*, 11520.
53. Dawson, T. E.; Mambelli, S.; Plamboeck, A. H.; Templer, P. H; Tu, K. P.; *Annu. Rev. Ecol. Syst.* **2002**, *33*, 507.
54. Craig, H.; *Science* **1961**, *133*, 1702.
55. Barbour, M. M.; *Funct. Plant Biol.* **2007**, *34*, 83.
56. Ramsey, C. B.; *Archaeometry* **2008**, *50*, 249.
57. Quan, H.; Barbetti, M.; Rakowski, A. Z.; *Radiocarbon* **2013**, *55*, 2059.
58. Anderson, E. C.; Libby, W. F.; *Phys. Rev.* **1951**, *81*, 64.
59. Godwin, H.; *Nature* **1962**, *195*, 984.
60. Bella, F.; Alessio, M.; Fratelli, P.; *Nuovo Cim. B*, **1968**, *58*, 232.
61. Chiu, T. C.; Fairbanks, R. G.; Cao, L.; Mortlock, R. A.; *Quat. Sci. Rev.* **2007**, *26*, 18.
62. Farias, R. F.; *QNEsc* **2002**, *16*, 6.
63. Cywinska, A.; Ball, S.L.; deWaard, J.R.; *Proc Biol Sci.* **2003**, *270*, 313.
64. CBOL Plant Working Group. *Proc. Natl. Acad. Sci U.S.A.* **2009**, *106*, 12794.
65. Muellner, A. N.; Schaefer, H.; Lahaye, R.; *Mol Ecol Resour.* **2011**, *11*, 450.
66. Jiao, L.; Yin, Y.; Cheng, Y. M.; Jiang, X.; *Holzforchung* **2014**, *68*, 487.
67. Yu, M.; Jiao, L.; Guo, J.; Wiedenhoeft, A. C.; He, T.; Jiang, X.; Yin, Y.; *Planta* **2017**, *246*, 1165.
68. Pastore, T. C. M. Estudos do efeito da radiação ultravioleta em madeiras por espectroscopia RAMAN (FT-RAMAN), de reflectância difusa no infravermelho (DRIFT) e no visível (CIE-L*a*b*). 2004. 131 f. Tese de doutorado - Universidade de Brasília: Brasília, 2004.
69. Lehninger, A. L.; Nelson, D. L.; Cox, M. M.; Simões, A. A. Princípios de Bioquímica, 4ª ed.; Sarvier: São Paulo, Brasil, 2006.

70. Azadi, P.; Inderwildi, O. R.; Farnood, R.; King, D. A. *Renew. Sust. Energ. Rev.* **2013**, *21*, 506.
71. Lewin, M.; Goldstein, I.S. *Wood structure and composition*. Marcel Dekker: New York, USA, 1991.
72. Fengel, D.; Wegener, G. *Wood: chemistry, ultrastructure, reactions*. Walter de Gruyter Berlin: New York, USA, 1989.
73. Prieto, N.; Pawluczyk, O.; Dugan, M. E. R.; Aalhus, J. L. *App. Spec.* **2017**, *71*, 1403.
74. Pasquini, C.; *J. Braz. Chem. Soc.* **2003**, *14*, 198.
75. Pavia, D. L.; Lampman, G. M.; Kriz, G. S.; Vyvyan, J. R.; *Introdução à Espectroscopia*, 3^a ed.; Cengage Learning, Washington, EUA, 2010.
76. Skoog, D. A.; Holler, F. J.; Crouch, S. R. *Principles of Instrumental Analysis*, 6^a ed.; Thomson: Califórnia, EUA, 2006.
77. Tibola, C. S.; Medeiros, E. P.; Simeone, M. L. F.; Oliveira, M. A. *Espectroscopia no Infravermelho Próximo para Avaliar Indicadores de Qualidade Tecnológica e Contaminantes em Grãos*; Embrapa: Brasília, Brasil, 2018.
78. Bokobza, L. *J. Near Infrared Spectrosc.* **1998**, *6*, 3.
79. Schwanninger, M.; Rodrigues, J. C.; Fackler, K. *J. Near Infrared Spectrosc.* **2011**, *19*, 287.
80. Li, X.; Sun, C.; Zhou, B.; He, Y. *Sci. Rep.* **2015**, *5*, 17210.
81. Burns, D. A.; Ciurczak, E.W.; *Handbook of Near-Infrared Analysis*, 3^a ed.; CRC Press, Florida, EUA, 2007.
82. Blanco, M.; Villarroya, I.; *Trends Anal. Chem.* **2002**, *21*, 240.
83. Hetrick, E. M.; Shi, Z.; Barnes, L. E.; Garrett, A. W.; Rupard, R. G.; Kramer, T. T.; Cooper, T. M.; Myers, D. P.; Castle, B. C.; *Anal. Chem.* **2017**, *89*, 9175.
84. Sánchez, M. T.; Pérez-Marín, D.; Torres, I.; Gil, B.; Garrido-Varo, A.; De la Haba, M. J.; *Postharvest Biol. Technol.* **2017**, *125*, 122.

85. Torres, I.; Pérez-Marín, D.; Haba, M. J.; Sánchez, M. T.; *Biosyst. Eng.* **2017**, *153*, 140.
86. Snee, R. D. *Technometrics* **1977**, *19*, 415.
87. Valderrama, P.; Braga, J. W. B; Poppi, R. J. *J. Agric. Food Chem.* **2007**, *55*, 8331.
88. Borin, A.; Poppi, R. J.; *J. Braz. Chem. Soc.* **2004**, *15*, 570.
89. da Silva, V. A. G.; Talhavini, M.; Zacca, J. J.; Maldaner, A. O.; Peixoto, I. C. F.; Braga, J. W. B.; *Microchem. J.* **2014**, *116*, 235.
90. da Silva, V. A. G.; Talhavini, M.; Zacca, J. J.; Trindade, B. R.; Braga, J. W. B.; *J. Braz. Chem. Soc.* **2014**, *25*, 1552.
91. Annual Book of ASTM Standards; Standards Practices for Infrared Multivariate Quantitative Analysis, E1655-05. ASTM International: West Conshohocken, 2012
92. Brereton, R.G. *Analyst* **2000**, *125*, 2125.
93. Wise, B. M.; Shaver, J. M.; Gallagher, N. B; Windig, W.; Bro, R.; Koch, R. S.; *PLS Toolbox Version 4.0 for use with MATLAB™*. Eigenvector Research Inc: Wenatchee, USA, 2006.
94. De Maesschalck, R.; Candolfi, A.; Massarta, D. L.; Heuerdingb, S.; *Chemom. Intel. Lab. Syst.* **1999**, *47*, 65.
95. Marini, F. *Curr. Analytical Chem.* **2010**, *6*, 72.
96. Botelho, B. G.; Reis, N.; Oliveira, L. S.; Sena, M. M.; *Food Chem.* 2015, *181*, 31.
97. <https://www.thermofisher.com/order/catalog/product/MICROPHAZIRRX>, acessado em 05 de maio de 2018.
98. Snel, F. A.; Braga, J. W. B.; Silva, D.; Wiedenhoeft, A. C.; Costa, A.; Soares, R.; Coradin, V. T. R.; Pastore, Tereza, C. M. *Wood Sci. Technol.* **2018**, *52*, 1411.