



**UNIVERSIDADE DE BRASÍLIA  
INSTITUTO DE CIÊNCIAS BIOLÓGICAS  
DEPARTAMENTO DE BOTÂNICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM BOTÂNICA**

**Espectroscopia no infravermelho próximo para predição de lignina siringil/guaiacil em  
*Eucalyptus* usando equipamento portátil e de bancada.**

**Carolina Pinto Diniz  
Orientador: Lúcio Flávio de Alencar Figueiredo**

**Brasília – DF  
Março de 2018**

**Espectroscopia no infravermelho próximo para predição de lignina siringil/guaiacil em *Eucalyptus* usando equipamento portátil e de bancada.**

Dissertação submetida ao Programa de Pós-Graduação em Botânica da Universidade de Brasília, como um dos requisitos para obtenção do título de Mestre em Botânica.

**Carolina Pinto Diniz (Bolsista CAPES)**  
**Orientador: Lúcio Flávio de Alencar Figueiredo**

## **Agradecimentos**

Agradeço a Universidade de Brasília (UnB), em especial ao Programa de Pós-Graduação em Botânica (PPG BOT), pelo curso e estrutura de laboratórios, equipamentos e professores oferecidos para a realização deste trabalho.

A Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pelo apoio financeiro para a concessão de bolsa ofertada durante o período do curso de mestrado.

A Empresa Brasileira de Pesquisa Agropecuária (Embrapa) pelo material (amostras de madeira de *Eucalyptus spp.*) pesquisado, a partir do qual o trabalho foi realizado.

A Fundação de Apoio a Pesquisa do Distrito Federal pelo apoio financeiro oferecido para a participação na 18ª Conferência Internacional de Espectroscopia do Infravermelho Próximo em Copenhague, Dinamarca de 11 a 15 de junho de 2017.

Agradeço ao meu orientador, Prof. Dr. Lúcio Flávio de Alencar Figueiredo, do PPG BOT/UnB primeiramente pela oportunidade para realizar em seu laboratório o estágio que possibilitou a realização do mestrado, mas sobretudo por todo o apoio, ensinamentos e orientação.

Agradeço ao Dr. Dario Grattapaglia, da Embrapa Recursos Genéticos e Biotecnologia, pelas amostras e equipamento de infravermelho próximo (FOSS 5500), pela participação nas bancas de projeto, qualificação e defesa final, assim como a disponibilidade para esclarecer diversas dúvidas ao longo da realização do trabalho.

Aos demais profissionais que participaram das bancas de projeto, qualificação e defesa final ao longo do curso e que contribuíram para a conclusão deste trabalho: Dra. Tereza Pastore (Laboratório de Produtos Florestais – Ministério do Meio Ambiente), Dr. Robélio Marchão (Embrapa Cerrados), Dra. Julia Sonsin (UnB – Departamento de Botânica) e Dr. Bruno de Lima (Fibria).

A todos os professores do Departamento de Botânica que tive contato ao longo do curso e que sempre se mostraram disponíveis para ajudar de alguma forma no desenvolvimento do trabalho.

Aos amigos e colegas de departamento, que estavam presentes em diversos momentos e que estavam sempre por perto para incentivar quando eu precisava: Larissa, Hérica, Débora, Karen, Isadora, Laísa, Rennan, Fernanda, Thiago, Janderson, Guilherme e Marcelo.

Aos amigos da vida, por sempre me apoiarem, compartilharem comigo os momentos felizes e por celebrarem minhas vitórias como se suas fossem: Ana Luiza, Isabela, Arthur, Camila, Natalia, Alessandra, Maria Raquel, Dani e Bia, Gui e Lud, Adriana, Aline, Lorena, Lu, Mari, Andrea, Taya, Taline, Tiago, Joacy, Gustavo e Aline Oliveira.

E agradeço a minha família, minha irmã Isabela e meu sobrinho e afilhado Bernardo, por todo apoio e todos os momentos felizes, e a meus pais Cláudio e Deo, aos quais eu devo tudo o que sou, é graças a vocês e por vocês.

## Sumário

Resumo.....	1
Abstract.....	2
Capítulo 1 – Referencial teórico .....	3
1. Introdução.....	3
2. Revisão bibliográfica.....	4
2.1. O gênero <i>Eucalyptus</i> L’Her.....	4
2.1.1. Biologia da planta .....	4
2.1.2. Morfologia do eucalipto.....	4
2.2. Cultivo do eucalipto no Brasil e no mundo.....	5
2.2.1. O cultivo no mundo .....	5
2.2.2. O cultivo no Brasil .....	5
2.2.3. Cultivo da planta.....	6
2.2.4. Destinação.....	6
2.3. <i>Eucalyptus benthamii</i> .....	7
2.4. <i>Eucalyptus pellita</i> .....	8
2.5. Química da madeira .....	9
2.6. Infravermelho .....	13
2.6.1. Infravermelho médio e infravermelho próximo.....	13
2.6.2. História do infravermelho .....	14
2.6.3. Uso da espectroscopia do infravermelho próximo em eucalipto .....	17
2.7. Quimiometria .....	18
3. Objetivos.....	19
3.1. Objetivo geral.....	19
3.2. Objetivos específicos.....	19
4. Referências bibliográficas .....	20
Capítulo 2 – Desenvolvimento de modelo para discriminação de espécies utilizando espectroscopia do infravermelho próximo com equipamento portátil e de bancada.....	27
Artigo: Comparative performance of bench and portable near infrared spectrometers for measuring wood samples of two <i>Eucalyptus</i> species ( <i>E. pellita</i> and <i>E. benthamii</i> ).....	28
Capítulo 3 – Desenvolvimento de modelos de predição da proporção siringil/guaiacil utilizando espectroscopia do infravermelho próximo.....	37
Artigo: Near infrared based models for lignin syringyl/guaiacyl ratio of <i>Eucalyptus benthamii</i> and <i>E. pellita</i> using a streamlined thioacidolysis procedure as the reference method.....	38

## Resumo

A madeira do eucalipto (*Eucalyptus* spp.) é uma das principais fontes de polpa de celulose para a produção de papel. No processo de extração da polpa de celulose, a proporção de lignina siringil e guaiacil (S/G) é a principal característica química. Quanto maior a proporção S/G, menores são os custos na extração da polpa de celulose. Todavia, as análises químicas são onerosas e demoradas, tornando-se necessária a utilização de métodos indiretos. Nesse contexto, a espectroscopia do infravermelho próximo (NIR) tem sido amplamente utilizada para o desenvolvimento de modelos de predição das características químicas em diversas amostras. Isso reduz os custos com novas análises químicas e fornece predições rápidas a partir da aquisição dos espectros. O objetivo deste trabalho foi utilizar a NIR para o desenvolvimento de modelos de predição a partir da serragem de 1.220 árvores de duas populações de eucaliptos (*Eucalyptus benthamii* e *E. pellita*), com cerca de quatro anos de idade, crescidas em dois biomas diferentes no Brasil. Foram utilizados dois espectrômetros, um de bancada (FOSS 5000) e um portátil (Viavi MicroNIR 1700). Os modelos de classificação e predição foram desenvolvidos a partir de análises multivariadas utilizando somente os espectros (n=6.100) e dos espectros correlacionados com a proporção S/G a partir de trinta árvores de cada população. A classificação das espécies foi muito boa no equipamento de bancada [acurácia superior a 95% na análise discriminante linear (LDA) e coeficiente de determinação ( $R^2$ ) superior a 0,90 na análise discriminante por regressão de mínimos quadrados parciais (PLS-DA)] e levemente menor do equipamento portátil (LDA=93% e  $R^2=0,80$ ). Nos modelos de predição quantitativa, os resultados da PLS feitas a partir dos espectros no equipamento de bancada foram superiores aos resultados do equipamento portátil baseado em quatro parâmetros de análise dos modelos desenvolvidos: RSD - desvio padrão relativo (coeficiente de variação entre o valor de referência e o predito),  $r_s$  - ranqueamento pela correlação de Spearman, RPD - razão da performance pelo desvio, e  $R^2$ . Os melhores modelos, com validação externa obtiveram os seguintes índices: 4,06; 0,95; 2,08 e 0,82 para o equipamento de bancada e 6,39; 0,0; 1,06 e 0,46 para o equipamento portátil, respectivamente. Dessa forma, os modelos no equipamento de bancada podem ser utilizados além de uma seleção inicial, o que seria sugerido baseando-se somente na RPD. Esses modelos podem ser utilizados em programas de melhoramento para as espécies desse estudo e como referência para a extração da polpa de celulose na indústria.

Palavras-chave: NIR, *Eucalyptus*, proporção lignina siringil/guaiacil, *E. benthamii*, *E. pellita*.

## Abstract

Eucalyptus (*Eucalyptus* spp.) wood is one of the main sources in the pulping industry for paper production. In the cellulose extraction process from pulp, the principal chemical characteristic of relevance is the lignin syringyl to guaiacyl (S/G) ratio. The higher the S/G ratio, the lower the cost for cellulose extraction. However, as chemical analyses are expensive and time consuming, the development and utilization of indirect methods for measurement are therefore necessary. In this context, near infrared spectroscopy (NIR) has been widely used to develop prediction models for chemical characteristics in various samples. This reduces cost of chemical analyses and provides rapid predictions from spectral acquisitions. The objective of this study was to use NIR for the development of prediction models from sawdust samples of 1,220 trees from two eucalyptus species populations (*Eucalyptus benthamii* and *E. pellita*), all approximately four years old, and cultivated in two different Brazilian biomes. Two spectrometers were employed in analyses, a bench spectrometer (FOSS 5000) and a portable spectrometer (Viavi MicroNIR 1700). Classification and prediction models were developed using multivariate analysis of spectral data (n=6,100), and spectral data correlated to S/G ratio from 30 trees from each population. Species classification was highly accurate when employing the bench equipment [accuracy higher than 95% in linear discriminant analysis (LDA) and coefficient of determination ( $R^2$ ) higher than 0.90 in partial least square regression for discriminant analysis (PLS-DA)] and slightly less accurate with the portable equipment (LDA=93% and  $R^2=0.80$ ). Quantitative prediction models for PLS were superior with the bench equipment, when compared to the portable device, considering four parameters: RSD - relative standard deviation (coefficient of variation between reference and prediction values),  $r_s$  - Spearman's rank correlation, RPD - ratio performance to deviation, and  $R^2$ . The best prediction models, with external validation, were as follows: 4.06, 0.95, 2.08 and 0.82 for the bench spectrometer and 6.39, 0.0, 1.06 and 0.46 for the portable spectrometer. Thus, the models obtained with the bench spectrometer are useful for application beyond initial screening, as suggested based only on the RPD. These models can be used in breeding programs for the species in this study and as references for the cellulose industry, with regard to cellulose extraction from pulp material.

Key-words: NIR, *Eucalyptus*, lignin syringyl/guaiacyl ratio, *E. benthamii*, *E. pellita*.

## Capítulo 1 – Referencial teórico

### 1. Introdução

O gênero *Eucalyptus* L'Hér, da família Myrtaceae, possui espécies florestais de grande importância que são cultivadas em extensas áreas no Brasil, com diferentes aplicações industriais. A maioria das espécies desse gênero é nativa da Austrália e essas espécies têm sido muito estudadas em programas de melhoramento visando o aumento de produção e maior resistência ou tolerância a pragas e adaptabilidade ao clima.

Os programas de melhoramento genético em *Eucalyptus* geram milhares de descendentes que precisam ser analisados de acordo com as características de interesse. No caso de *Eucalyptus*, a principal característica química de interesse é a proporção de lignina siringil/guaiacil visando a extração da celulose, sendo as análises químicas por via úmida longas e onerosas.

Uma alternativa às análises químicas tradicionais é a utilização da espectroscopia do infravermelho próximo (NIR). A NIR é uma técnica analítica que utiliza energia de fótons na região de comprimento de onda de 750 a 2.500 nm, na qual a energia que incide nas moléculas gera vibrações moleculares que resultam em transições harmônicas. Os comprimentos de onda nos quais estas vibrações ocorrem para um composto são funções de sua estrutura e composição, o que possibilita que NIR seja utilizada para análise qualitativa e quantitativa de compostos em diversos tipos diferentes de materiais (Pasquini, 2003). Os dados de absorvância ou refletância gerados pelo espectrofotômetro são analisados por meio de análises multivariadas, que permitem a interpretação do espectro.

O objetivo deste trabalho é desenvolver modelos de predição da proporção de lignina siringil/guaiacil a partir da serragem de árvores das populações de *Eucalyptus benthamii* e de *Eucalyptus pellita*, utilizando espectroscopia do infravermelho próximo (NIR). O total de árvores analisadas foi de 1.220, das quais 484 são de *E. benthamii* e 736 de *E. pellita*. Os modelos de classificação das espécies foram estabelecidos utilizando-se análise discriminante linear (LDA) e análise discriminante por regressão de mínimos quadrados parciais (PLS-DA). Os modelos de predição da proporção S/G foram desenvolvidos a partir dos dados de referência de 30 árvores de cada população, e seus respectivos espectros utilizando a regressão de mínimos quadrados parciais (PLS). Os modelos foram avaliados baseados no coeficiente de determinação ( $R^2$ ), razão da performance pelo desvio (RPD), desvio padrão relativo (RSD) e ranqueamento pela correlação de Spearman ( $r_s$ ).

## **2. Revisão bibliográfica**

### **2.1. O gênero *Eucalyptus* L'Her**

O eucalipto (*Eucalyptus spp* L'Hér), é uma planta nativa da Austrália, com algumas espécies de origem da Papua Nova Guiné, Indonésia e Ilhas da Oceania (Andrade, 1961), e apresenta 822 espécies descritas (The Plant List), divididas entre 10 subgêneros. Os subgêneros de maior destaque são *Symphyomyrtus*, *Eucalyptus* e *Eudesmia* (Bayly, 2016). O eucalipto foi descrito pela primeira vez em 1788 pelo botânico francês Charles Louis L'Héritier de Brutelle. Ele pertence à ordem Myrtales, família Myrtaceae e subfamília Leptospermoideae (Wilson et al., 2001). A ordem Myrtales pertence ao clado Rosids, e está inserida no clado das Eudicots, no grupo das angiospermas (Bremer et al., 2009). O gênero apresenta ampla distribuição mundial e se desenvolve em diferentes condições climáticas, porém apenas cerca de 1% das espécies são utilizadas industrialmente (Grattapaglia et al., 2012; Santos et al., 2001).

#### **2.1.1. Biologia da planta**

O eucalipto é a árvore modelo para as espécies lenhosas tropicais, devido a sua diversidade, fácil propagação, baixa ploidia (diploide), grande número de informações genômicas, sistema de propagação e transformação *in vitro* definidos, genoma sequenciado e pela sua importância econômica (Chaffey, 2002; Myburg et al., 2014).

O eucalipto é uma planta lenhosa com estruturas físicas diversas, podendo apresentar espécies de caules múltiplos e arbustos (Vuong et al., 2015). O gênero apresenta de árvores pequenas (3-5 m) até árvores de grande porte (40 - 80m) e longevidade média de 100 anos. Algumas espécies podem chegar a viver 450 anos (Benson & Mcdougall, 1998).

As árvores apresentam folhas perenes, que são mantidas durante todo o ano, mesmo na estação seca. É uma planta de metabolismo C3 (Lima et al., 2006) e fotoperíodo neutro. Devido a sua diversidade, cresce em climas variados, tendo um crescimento rápido em clima tropical (Rezende et al., 2014).

#### **2.1.2. Morfologia do eucalipto**

O principal órgão de interesse econômico do eucalipto é o caule, geralmente formado por um tronco reto e de grandes dimensões (Moura & Guimarães, 2003). Em todas as espécies, a camada mais externa da casca morre e, em algumas espécies, essa casca se solta. A casca pode ser persistente ou decídua e lisa ou áspera. As espécies de casca lisa costumam apresentar cores e texturas uniformes. Nos eucaliptos que apresentam casca áspera, as camadas de fibra da casca se acumulam



ao longo dos anos, gerando a textura áspera e irregular. Em algumas espécies ocorre a infusão da casca com goma, que a torna dura. Nas espécies de caules múltiplos, forma-se o lignotúber sob ou próximo ao solo (Wilson, 2011).

Suas folhas apresentam glândulas oleíferas na forma de pequenos pontos translúcidos, presentes também nas flores e frutos. Essas glândulas oleíferas são típicas das espécies da família Myrtaceae (Wilson, 2011). As folhas são simples, de bordo inteiro e peninérveas. Existem algumas particularidades nas folhas de algumas espécies de *Eucalyptus*, que apresentam dimorfismo foliar, podendo ser opostas, ovais ou arredondadas e, ocasionalmente, sem pecíolo na fase jovem, e após alguns anos de crescimento, passam a apresentar folhas alternadas, isobilaterais ou em alguns casos dorsiventrais (Wilson, 2011).

As flores são actinomorfas, diclamídeas, dialipétalas, raramente com pétalas de tamanho reduzido, polistêmone, as anteras são globosas, rimosas e bitecas. O ovário é ínfero, com número variável de lóculos e de óvulos. Os frutos são secos e produzem um grande número de sementes, cerca de 1 a 108 por fruto (Gill et al., 1992), e muitas vezes correspondendo ao número de óvulos (Wilson et al., 2001). As sementes possuem formatos variados e os diâmetros podem variar entre 4,2 – 38mm (Gill et al., 1992).

## **2.2. Cultivo do eucalipto no Brasil e no mundo**

### **2.2.1. O cultivo no mundo**

Os primeiros plantios datam do início do século XVIII e ocorreram na Europa, na Ásia e na África. Posteriormente, já no século XIX, começou a ser plantado na Espanha, Índia, Brasil, Argentina e Portugal (CIB, 2008).

As plantações de eucalipto atuais são dominadas pelas “nove grandes” espécies (*E. camaldulensis*, *E. grandis*, *E. tereticornis*, *E. globulus*, *E. nitens*, *E. urophylla*, *E. saligna*, *E. dunnii* e *E. pellita*) e seus híbridos, que somam mais de 90% de todo gênero *Eucalyptus* plantado (Stanturf et al., 2013).

### **2.2.2. O cultivo no Brasil**

Os primeiros registros da introdução do eucalipto no Brasil datam de 1868, no entanto, até o início do século XIX a espécie era utilizada como árvore decorativa e quebra-vento, pouco plantado com fins industriais e florestais. Somente ao final de 1904, Edmundo Navarro de Andrade começa a desenvolver trabalhos com eucalipto visando fornecer combustível para locomotivas, madeira para postes, moirões e dormentes (Andrade, 1961). Entre 1904 e 1909, foram realizados vários

experimentos comparativos entre eucaliptos e algumas espécies nativas como peroba, cabreúva, cedro, jequitibá, pinheiro do Paraná e exóticas de valor econômico (Andrade, 1961).

Os primeiros plantios com interesse comercial surgiram na segunda metade do século XIX, porém o incremento na produtividade da cultura só ocorreu no início do século XX, devido ao melhoramento genético, iniciado por Navarro de Andrade em 1904 (Castro et al., 2016), e ao desenvolvimento de práticas silviculturais voltadas às áreas tropicais. Apesar de ter origem em outro continente, o eucalipto se adaptou muito bem ao Brasil por ter encontrado condições ideais ao seu desenvolvimento como insolação intensa e abundância de recursos hídricos (Myburg et al., 2014).

Atualmente, o eucalipto é extensamente plantado, conhecido e utilizado no Brasil. Em 2017, plantações de *Eucalyptus* representavam mais de 78% das plantações de *Eucalyptus* e *Pinus*, os dois gêneros mais utilizados para silvicultura no país. As áreas das plantações de eucalipto correspondem a um total de 5,7 milhões de hectares, o que representa um aumento de 2,4% ao ano, nos últimos 5 anos. O destaque na produção de eucalipto no país é da região Sudeste (53%) e os estados que mais produzem são Minas Gerais, São Paulo e Mato Grosso (IBÁ 2017). Dentre as espécies mais cultivadas no país, podemos destacar *E. grandis*, *E. saligna*, *E. urophylla*, e híbridos de *E. grandis* x *E. urophylla*, os quais apresentam grande área plantada principalmente devido a utilização de material clonal (Higa et al., 2000). Outras espécies, plantadas no sul do Brasil, são *E. dunnii*, *E. benthamii*, *E. camaldulensis*, *E. viminalis*, *E. cloeziana* e *E. badjensis* (Serpe, 2015).

### **2.2.3. Cultivo da planta**

O eucalipto pode ser cultivado por meio de sementes ou estaquia, no caso de clones ou reprodução de híbridos. Outros métodos que também podem ser utilizados são a enxertia e a micropropagação (Yang et al., 1995).

A utilização de técnicas de silvicultura adequadas, o preparo e a fertilização do solo, o combate a pragas e doenças, e melhoramento genético através de plantios de variedades puras ou híbridas, assim como a clonagem de materiais com características favoráveis e aumento da homogeneidade dos povoamentos possibilitaram ganhos significativos na produção do gênero *Eucalyptus* (Higashi et al., 2000).

### **2.2.4. Destinação**

As características que contribuíram para o eucalipto ganhar destaque são o rápido crescimento, alta produtividade, diversidade de espécies, grande capacidade de adaptação e produção de madeira de importância comercial para diversos usos industriais. Destes fins, pode-se citar a produção de

papel, celulose, lenha, biomassa, carvão vegetal, painéis de madeira reconstituída, serraria, dormentes, postes, óleos para indústria farmacêutica, cosméticos, mel, ornamentação e quebra-vento (Santos et al., 2001). Além disso, o eucalipto é estudado para o uso na produção de biocombustíveis derivados de lignocelulose (Çetinkol et al., 2012; Mizrachi et al., 2012; Vega-Sánchez & Ronald, 2010).

Em áreas tropicais e subtropicais, o eucalipto é uma das árvores de maior produtividade. No Brasil, em algumas regiões, a produtividade excedeu  $80 \text{ m}^3\text{ha}^{-1}$  ao ano (Rezende et al., 2014). Em 2016, 206 milhões de metros cúbicos de madeira foram utilizados para fins industriais. Deste total, 45% corresponde à madeira de *Eucalyptus*, sendo 46% destinado ao segmento de papel e celulose, 30% lenha industrial, 14% para produção de carvão e os 14% restantes divididos em outros segmentos (IBÁ, 2017). No Brasil, 100% da matéria prima para produção de celulose e papel têm origem em florestas plantadas de eucalipto e pinheiro (BRACELPA, 2014).

Do ponto de vista econômico, o setor brasileiro de árvores plantadas é muito importante, representando cerca de 6,2% do produto interno bruto (PIB) industrial, gerando receita de R\$ 71,1 bilhões. Além disso, promove a geração de empregos e renda, com uma estimativa de criação de 3,7 milhões de empregos, sendo 510 mil postos diretos e o restante indiretos e resultantes do efeito renda (IBÁ, 2017).

### **2.3. *Eucalyptus benthamii***

O *Eucalyptus benthamii* Maiden et Cabbage, conhecido na Austrália como Camden White Gum, é uma espécie do subgênero *Symphyomyrtus*, seção *Maidenaria* (Hill & Johnson, 1995) e foi classificado na série Viminales, subsérie Viminalinae (Pryor & Johnson, 1971). As espécies mais próximas são a *Eucalyptus dorriigoensis* (descrita anteriormente como uma variedade de *E. benthamii*) e a *Eucalyptus kartzoffiana*, ambas apresentam distribuição relativamente restrita, onde *E. dorriigoensis* é encontrada próxima de Dorriigo, no litoral norte e *E. kartzoffiana* no litoral sul de New South Wales, Austrália (Benson, 1985).

*E. benthamii* é encontrada em áreas restritas e pouco extensas, se limitando a duas populações ao sudoeste de Sidney: uma pequena população e alguns indivíduos espalhados ao longo do Rio Nepean entre Wallacia e Camden; e outra população maior em Kedumba Creek ( $33^{\circ}49'S$ ,  $150^{\circ}22'$ ) a uns 5 km da confluência com o Rio Coxs (Benson, 1985). É considerada uma espécie ameaçada de extinção na Austrália, por se encontrar apenas duas populações.

*E. benthamii* alcança uma altura média de 40 m, com longevidade de mais 200 anos (Benson, 1985); ocorre preferencialmente em margens de rios, em regiões de altitude de até 100m; é capaz de

se desenvolver em solos bem drenados e rasos a moderadamente profundos (<1m), com pH entre 5,5 e 6,5 (Benson & Mcdougall, 1998). Apresenta caule com casca lisa, inflorescências brancas (março – maio), sua cápsula apresenta 4 a 5 mm de comprimento com dispersão das sementes entre junho e dezembro (Benson & Mcdougall, 1998). Nas regiões onde ocorre, a média de precipitação anual é de 730 a 1100 mm (regime uniforme de chuva no verão e tempo de seca de 5 meses). Nessas regiões, a média das temperaturas máximas fica entre 26 a 30°C, das mínimas de -1 a 3°C e a média anual está na amplitude de 13 a 17°C (Jovanovic & Booth, 2002).

As primeiras populações de *E. benthamii* no Brasil foram plantadas em Colombo, Paraná em 1988 pela EMBRAPA Florestas. O povoamento foi feito a partir de famílias compostas de dez matrizes provenientes de Wentworth Falls, NSW, Austrália (Graça et al., 1999). As populações no sul do Brasil estão sendo utilizadas para testar a resistência à geadas, as quais têm demonstrado uma alta resistência, com rápido crescimento, boa forma do fuste e alta homogeneidade do talhão (Graça et al., 1999). Um fator que pode ser limitante para programas de melhoramento genético e produção de mudas para plantios comerciais é que as estacas de *E. benthamii* apresentam baixo enraizamento (Graça et al., 1999).

No sul da África e em partes da América do Sul, destaca-se o interesse por *E. benthamii* para implantação de povoamentos florestais para a produção de celulose e papel (FAO, 2004). Os principais usos potenciais da espécie são como lenha, carvão e celulose (Higa & Pereira, 2003). Por outro lado, sua madeira pode apresentar elevada instabilidade dimensional e grande propensão à formação de rachaduras e empenos, não sendo recomendado para usos estruturais (Higa & Pereira, 2003; Müller et al., 2014).

O interesse com *E. benthamii* é grande, quer seja para uso como espécie pura ou para estabelecimento de híbridos com *E. grandis* em zonas mais frias, pois é uma espécie que apresenta resistência à geadas (Higa, 1999).

#### **2.4. *Eucalyptus pellita***

O *Eucalyptus pellita* F. Muell, popularmente conhecido na Austrália como “Red Mahogany” (mogno vermelho), é uma espécie do subgênero *Symphyomyrtus* e seção *Transversaria*, a qual faz parte juntamente com *E. resinifera* e *E. propinqua* (Araújo et al., 2002). *E. pellita* e *E. resinifera* são classificadas de forma agrupada e possuem várias características morfológicas similares, e são fortemente relacionadas a *E. urophylla* (Pryor et al., 1995). Este é um grupo de cerca de 20 espécies que incluem algumas das maiores e mais valiosas espécies de eucalipto (Bristow, 2008).

A distribuição geográfica de *E. pellita* se limita a uma área na Austrália: na costa norte de Queensland (latitude 12 a 18°S), com clima tropical, chuvas periódicas e períodos de 5 a 7 meses de seca (Ferreira & Silva, 2004). Pode-se também encontrar uma distribuição natural em Papua Nova Guiné e Irian Jaya (província da Indonésia) (Harwood et al., 1997), sendo encontrada em regiões climáticas úmidas e quentes, com altitude de até 600 m. A temperatura máxima média varia de 24 a 33° C e a temperatura mínima média varia de 12 a 16° C (Ferreira, 1979).

*E. pellita* é encontrado em zonas de clima quente e úmido, onde geadas são raras ou ausentes e a média de precipitação anual fica entre 900 – 2200 mm. Em florestas naturais, é uma árvore média a alta, chegando a 40m de altura e 1m de circunferência. Nas melhores condições, apresenta uma excelente forma de fuste, reto até metade da altura, e uma grande coroa fortemente ramificada (Harwood et al., 1997). Cresce preferencialmente em locais úmidos nas encostas mais baixas de grandes cordilheiras e ao longo de córregos, na parte mais seca e quente. O solo varia de areia rasa em cumes de arenito a areia rasa de podsols (Queensland) e solo profundo e limoso de florestas (Irian Jaya e Nova Guiné). É a espécie arbórea mais plantada nas áreas quentes e de alta precipitação ao nordeste de Queensland (Bristow, 2008).

Em comparação com outras espécies de eucalipto, é uma espécie que é propagada facilmente (Baker & Walker, 2006), apresenta rápido crescimento inicial em plantações (Bristow et al., 2006; Sun et al., 1996) e é pouco afetada por pestes e doenças. É uma espécie que consegue se desenvolver bem tanto em culturas mistas quanto em monoculturas com interesse industrial (madeira serrada e celulose integrada) (Bristow, 2008).

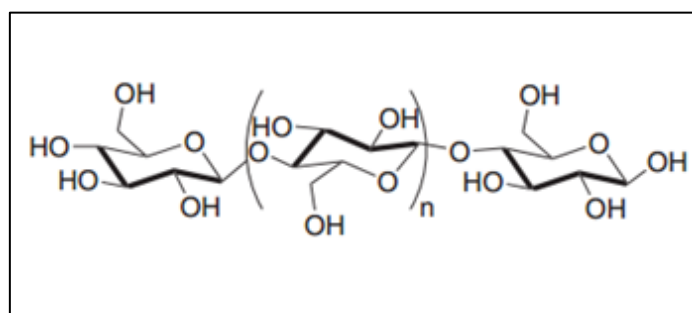
A espécie foi introduzida no Brasil possivelmente em São Paulo. As plantações pilotos foram feitas no Horto Guarani, da Cia Paulista de Estradas de Ferro, em solo muito pobre, e não apresentando resultados muito animadores (Ferreira, 1979). Entretanto, existem outras características que são determinantes para o interesse em *E. pellita* como a resistência a doenças (Harwood et al., 1997) e possuir uma madeira fácil de trabalhar, com um bom acabamento que pode ser usada para pisos, revestimento, painéis, soleiras, construções e estruturas (Zanata et al., 2010). *E. pellita* é uma madeira mais pesada e mais rica em lignina e extrativos do que *E. benthamii* (Pereira et al., 2000).

## **2.5. Química da madeira**

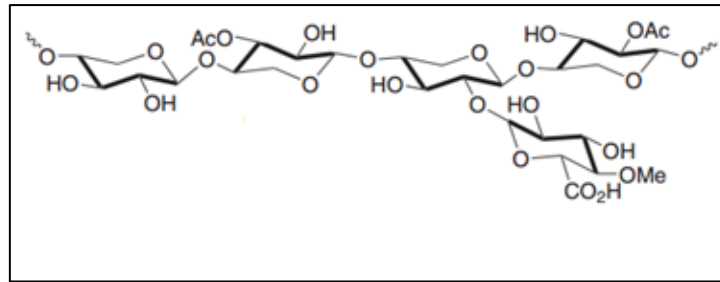
A composição química da madeira não pode ser definida como um valor fixo para todas as espécies de árvores e nem para todas as árvores de uma mesma espécie, pois esta varia de acordo com a estrutura analisada (raiz, caule ou ramo), tipo de madeira (normal, sobre tensão ou compressão), localização geográfica, clima e condições do solo, planta na fase juvenil ou adulta. Os

dois maiores componentes da madeira são carboidratos (65-75%) e lignina (18-35%) (Pettersen, 1984). Em *Eucalyptus*, esses valores correspondem em média a 48,8 - 55,3% para carboidratos e 24,9 - 31,5 para lignina (Çetinkol et al., 2012). Materiais como extrativos orgânicos e minerais inorgânicos são encontrados em valores menores (4-10%) para madeiras de clima temperado, podendo chegar aos 15% em madeiras de clima tropical (Da Silva et al., 2013). No geral, a composição elementar da madeira é de 50% de C, 44% O, 6% H e traços de vários íons metálicos (Pettersen, 1984).

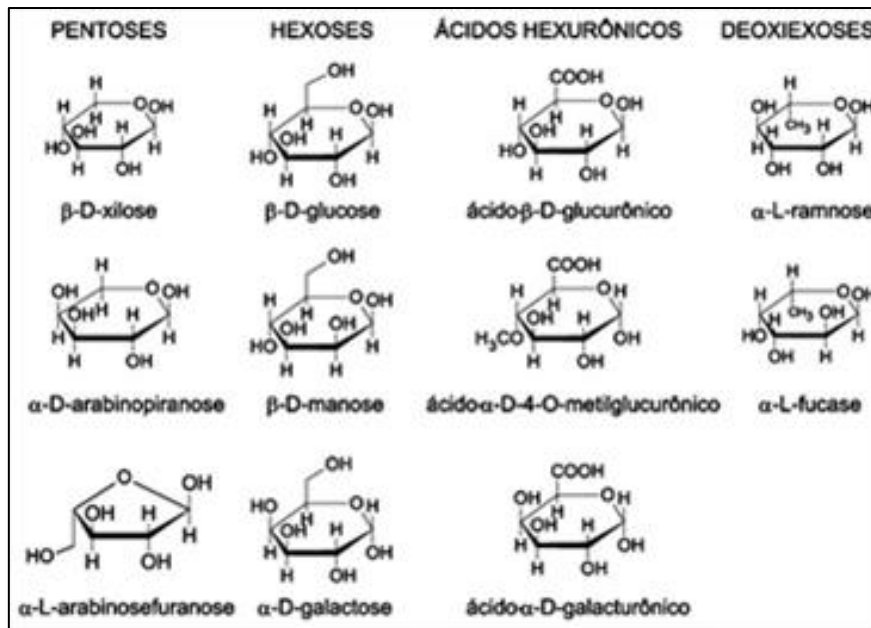
Os carboidratos são divididos em celulose e hemicelulose, que variam de 40-50% de celulose e 25-35% de hemicelulose (Pettersen, 1984). Em *Eucalyptus*, a celulose varia entre 33 - 47,8% (Çetinkol et al., 2012; Willför et al., 2005) e a hemicelulose corresponde a cerca de 20% (Carvalho, 2015). A celulose (Figura 1) é composta por uma cadeia linear e não ramificada de polímeros de glicose, com as unidades ligadas por uma ligação  $\beta$ -1,4, sendo constituída por 10 a 15 mil unidades de glicose. É uma substância fibrosa, resistente e insolúvel encontrada na parede celular vegetal, particularmente em caules, tronco e todas as porções amadeiradas da planta (Nelson & Cox, 2011). A celulose é insolúvel na maioria dos solventes, inclusive solventes alcalinos fortes. A dificuldade em isolar celulose da madeira na forma pura ocorre porque a celulose está fortemente associada à lignina e hemicelulose. A hemicelulose (Figura 2) é uma mistura de polissacarídeos e contém glicose, manose, galactose, xilose, arabinose, resíduos de ácido 4-O-metilglucurônico e ácido galacturônico (Figura 4). Em madeira de angiospermas, ainda é possível encontrar traços de raminose. A hemicelulose pode estar presente em quantidades muito altas e anormais quando a planta se encontra em estresse. É solúvel em soluções alcalinas e facilmente hidrolisada por ácidos (Pettersen, 1984).



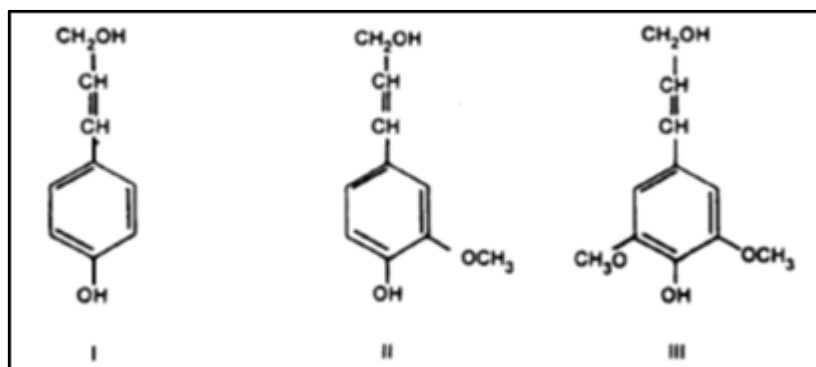
**Figura 1** - Estrutura química da celulose (Ferreira et al., 2009).



**Figura 2** - Estrutura química da hemicelulose (Ferreira et al., 2009).



**Figura 3** - Unidades que compõem a hemicelulose (Morais et al., 2005).



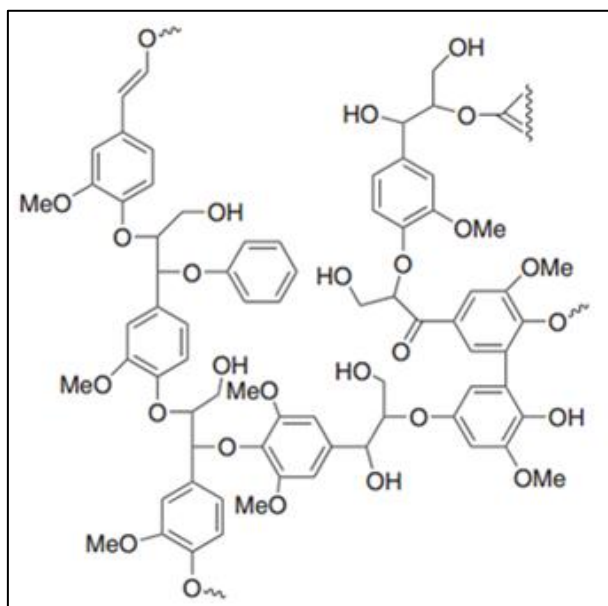
**Figura 4** - Precursores da lignina no processo de biossíntese (Pettersen, 1984).

O outro componente importante da madeira é a lignina, um dos biopolímeros mais abundantes do planeta, sendo muito importante para o sequestro de carbono (Mottiar et al., 2016). A lignina é uma substância fenólica composta por quantidades variadas de unidades de fenilpropano ligadas a hidroxi- e metoxi-. Os precursores da biossíntese de lignina são os álcoois *p*-cumarílico (Figura 4 –

I), coniferílico (Figura 4 –II) e sinapílico (Figura 4 – III). I é um precursor secundário de lignina em madeiras de angiospermas e gimnospermas, II é o precursor dominante na biossíntese de lignina em madeiras de gimnospermas e II e III são ambos precursores de lignina em madeiras angiospermas (Pettersen, 1984).

A lignina é então classificada de acordo com a quantidade relativa de monômeros derivados destes precursores. Os monômeros são hidroxifenila (H), guaiacila (G) e siringila (S), derivados dos álcoois *p*-cumarílico, coniferílico e sinapílico, respectivamente (Barbosa et al, 2008). A lignina de gimnospermas, por exemplo, é mais resistente por ser composta principalmente por subunidades G (Mottiar et al., 2016). Na produção de kraft de celulose, a relação de lignina siringil e guaiacil (S/G) apresentou uma alta correlação positiva com a eficiência de deslignificação e o rendimento da polpa kappa 17 (Gomes et al., 2008).

Estudos com *Eucalyptus* encontraram concentrações de lignina entre 18,7 e 27,4% (Carvalho, 2015; Rencoret et al., 2011), enquanto a proporção de lignina S/G correspondeu a 1,4 - 3,6; com uma média de 2,2 (Alves et al., 2011; Hou & Li, 2011; Lupoi et al., 2014; Ramadevi et al., 2016). *Eucalyptus pellita*, descrito em apenas um trabalho, apresentou proporção de lignina S/G entre 1,7 e 2,8 (Ramadevi et al., 2016). *E. benthamii*, por outro lado, teve sua proporção descrita pela primeira vez neste trabalho (2,34 a 4,20; média de 2,97 - Capítulo 3).



**Figura 5** - Estrutura química da lignina (Ferreira et al., 2009).

Os extrativos são uma variedade de classes de compostos orgânicos como gorduras, ceras, alcaloides, proteínas, fenóis simples e complexos, açúcares simples, pectinas, mucilagens, gomas, resinas, terpenos, amido, glicosídeos, saponinas e óleos essenciais. Vários destes têm funções de



intermediários no metabolismo da planta, como reserva de energia ou parte da defesa da planta contra ataques de microrganismos, e a maioria é solúvel em solventes neutros, tais como água e álcool. Contribuem para as propriedades da madeira como cor, odor e resistência à decomposição. A fuligem (cinzas), que é o resíduo inorgânico produzido quando a madeira é queimada em altas temperaturas, corresponde a menos de 1% em árvores de zonas temperadas, e um pouco mais de 1% em climas tropicais (Pettersen, 1984).

## **2.6. Infravermelho**

### **2.6.1. Infravermelho médio e infravermelho próximo**

A espectroscopia do infravermelho pode ser dividida em três regiões principais: infravermelho próximo (NIR), médio (MIR) e distante. O tipo de espectroscopia vibracional mais utilizado é a NIR, que compreende a extensão de 750 a 2.500 nm (Lupoi et al., 2014). Em análises de rotina no setor alimentar, a tecnologia já vem sendo utilizada à algum tempo, com componentes ópticos a base de silicone e detectores prontos para conexão com materiais de telecomunicação convencionais, o que reduz os custos na implementação deste tipo de sensor em comparação a sensores que utilizam a região do MIR (Haas & Mizaikoff, 2016). A instrumentação simples e a facilidade de lidar com as amostras ao se utilizar NIR têm permitido análises de grandes números de amostras de biomassa para identificar traços de interesse em diversos programas, como para produção eficiente de biocombustíveis (Robinson & Mansfield, 2009; Stackpole et al., 2011).

A espectroscopia do infravermelho médio (MIR) compreende a extensão de 2.500 a 25.000 nm (400 a 4.000  $\text{cm}^{-1}$ ) (Haas & Mizaikoff, 2016). Diferentemente da tecnologia NIR, que apresenta sinais de intensidade moderada, MIR caracteriza modos vibracionais fundamentais mais pronunciados das moléculas, associados a características espectrais distintas, que se relacionam diretamente com a estrutura molecular (Haas & Mizaikoff, 2016; Lupoi et al., 2014). Essa diferença na intensidade ocorre porque a energia utilizada na espectroscopia MIR é maior, portanto interage mais fortemente com as moléculas e gera respostas mais energéticas.

A espectroscopia MIR têm sido mais frequentemente utilizada para obter informações qualitativas e estruturais em análises de biomassa, e menos como uma ferramenta de quantificação para desenvolvimento de modelos multivariados de predição. Isso ocorre pois a água apresenta alta absorção, o que dificulta análises de amostras aquosas ou biológicas (Lupoi et al., 2014).

### 2.6.2. História do infravermelho

A radiação do infravermelho próximo foi descoberta por Sir William Hershel, que publicou dois artigos em 1800, detalhando seu estudo sobre o efeito de aquecimento do espectro de radiação solar. Suas descobertas foram feitas através do uso de um grande prisma de vidro que espalhava a luz solar. Ele posicionou três termômetros ao longo dos raios criados e percebeu que próximo da extremidade vermelha o efeito de aquecimento se tornava aparente e além do vermelho, onde não existia luz visível, ocorriam temperaturas maiores. Hershel descreveu este fenômeno como “calor radiante” e “espectro termométrico”, o que ele erroneamente classificou como sendo algo diferente da luz. Ampere então, em 1885, utilizando um termopar que fora recentemente inventado, provou que possuíam as mesmas características ópticas, portanto se tratava do mesmo fenômeno: luz (Burns & Ciurczak, 2008).

A ausência de aparelhos de detecção impediu que houvesse um progresso real do entendimento e uso de luz fora da parte visível do espectro. Em 1880, um importante passo foi tomado quando Abney e Festing notaram que a placa fotográfica possuía sensibilidade ao infravermelho próximo. Eles então foram capazes de medir o espectro de líquidos orgânicos de 1.000 a 1.200 nm, o que foi um trabalho de muita significância, sendo a primeira de uma série de medições de infravermelho próximo, como também as primeiras interpretações. Abner e Festing reconheceram tanto agrupamento atômico quanto a importância da ligação de hidrogênio no espectro NIR (Burns & Ciurczak, 2008).

Em 1905, Coblentz construiu um espectrômetro rudimentar que era capaz de detectar a região do infravermelho médio (2.500 a 50.000 nm) e descobriu que cada composto possui um “*fingerprint*” único, ou seja, dois compostos nunca teriam o mesmo espectro, mesmo que eles tivessem a mesma composição de elementos, sendo o caso de alguns isômeros como propan-1-ol e propano-2-ol. Coblentz também notou alguns padrões no espectro, como no caso de compostos com grupo OH que, sejam fenóis ou álcoois, absorviam na região 2.700 nm do espectro. A partir disso, muitos químicos, durante a primeira metade do século XX, utilizaram esta técnica para caracterizar grupos funcionais, estendendo a base de dados de compostos orgânicos e suas características e assinaturas espectrais. Porém, o uso de espectroscopia do infravermelho era muito limitado, possivelmente pela ausência de equipamentos apropriados (Burns & Ciurczak, 2008).

Na década de 30, o sulfato de chumbo (PbS), também conhecido como galena, estava sendo estudado como composto semicondutor e com os investimentos da Segunda Guerra Mundial, houve um estímulo para o desenvolvimento do PbS como um detector de infravermelho médio para

detecção de calor, que só se tornou disponível para aplicações comerciais a partir dos anos 50 (Burns & Ciurczak, 2008).

Nos anos 50 houve uma crescente demanda por determinações quantitativas de umidade, proteína e óleo que fossem rápidas. Karl Norris, trabalhando para o USDA (*United States Department of Agriculture* – Departamento de Agricultura dos Estados Unidos), foi encarregado de solucionar este problema para trigo. Ele então escolheu NIR, trabalhando com uma técnica primitiva comparada com as técnicas utilizadas atualmente (Burns & Ciurczak, 2008).

Durante os anos 80, o microprocessador estava integrado à maioria dos instrumentos, possibilitando aquisições de dados e manipulações mais sofisticadas. O escopo do tratamento de dados e a interpretação foram melhorados para incluir regressão linear múltipla (*multiple linear regression* - MLR), mínimos quadrados parciais (*partial least squares* – PLS), análise de componente principal (*principal component analysis* - PCA) e análise de *cluster* para que estes dados pudessem ser utilizados em modelos de predição. Outras empresas de software surgiram, oferecendo diversos tipos de tratamento, aumentando as opções dos usuários. No caso do infravermelho próximo, o avanço da tecnologia possibilitou a interpretação da região do NIR, pois as diferenças não são visuais como na região do infravermelho médio.

A tecnologia NIR evoluiu rapidamente desde 1970, sendo amplamente aceita atualmente e utilizado em vários setores como instrumento de medidas, sendo utilizada para medições contínuas e controles de processos produtivos, com diversas aplicações, incluindo produtos químicos, farmacêuticos, tabaco, comida, bebidas e demais produtos. A tecnologia utilizada atualmente nos equipamentos NIR depende inteiramente de computadores, tanto na obtenção e controle dos dados a partir do instrumento, quanto na calibração e análise dos dados. O desenvolvimento de computadores cada vez mais potentes e acessíveis tem contribuído para o surgimento de várias aplicações da NIR, em diferentes áreas, incluindo a área da saúde, farmacêutica e até nas mais tradicionais produções de alimentos (Blanco & Villarroya, 2002; Huang et al., 2008; Jamrógiewicz, 2012; Roggo et al., 2007).

Medidas indiretas de propriedades da madeira através da NIR têm sido utilizadas como uma forma de acelerar e reduzir os custos de fenotipagem de madeira (Schimleck et al., 2000). Essa técnica utiliza dados espectrais coletados em aparelhos NIR após leituras de amostras de madeira, que também são analisadas através de química analítica úmida (*wet chemistry*). O espectro é comparado às medidas encontradas na análise química e através de curvas de calibração utilizando modelos multivariados podemos prever a composição química das amostras semelhantes e desconhecidas que apenas tiveram seus espectros coletados. Nos últimos anos, a espectroscopia do infravermelho próximo (NIR) têm se tornado um dos métodos mais atrativos e utilizados para

análise, tanto qualitativa quanto quantitativa (Cozzolino, 2009). A NIR pode ser utilizada para análise de produtos naturais em plantas: proteína, açúcares, acidez e álcool (Batten, 1998), assim como para prever composição química em vários produtos orgânicos como cereais, carne, bebidas, frutas e vegetais (Ben-Gera & Norris, 1968; de Alencar Figueiredo et al., 2006; Huang et al., 2008; Lopez et al., 2013). Essa análise de composição pode estar relacionada com controle de qualidade de grãos, por exemplo (de Alencar Figueiredo et al., 2006). A NIR também tem sido utilizada para análise de composição de madeira, bambu, palha de milho e arroz, quanto à lignina (Uner, et al., 2009), celulose (Schwanninger et al., 2004; Sills & Gossett, 2012), hemicelulose (Xu et al., 2013) ou ambos (Li et al., 2015), e extrativos (Da Silva et al., 2013; Karlinasari et al., 2014).

A espectroscopia é uma técnica instrumental analítica que se baseia nas propriedades de absorção ou emissão de energia eletromagnética das moléculas. No caso da espectroscopia NIR, região localizada entre 750 e 2.500 nm, a resposta obtida é um espectro que só pode ser interpretado através de programas de análise estatística (análise multivariada). Quando o equipamento e os programas são calibrados com precisão, é possível prever várias propriedades da madeira e da polpa celulósica (Santos et al., 2009).

A NIR é uma técnica de análise não invasiva para materiais orgânicos, amplamente utilizada por empresas de manufatura, incluindo empresas do setor florestal, para controle de qualidade e monitoramento de processo produtivo (Baillères et al., 2002). Uma das vantagens dessa técnica, principalmente em comparação com métodos tradicionais, é ser uma técnica não destrutiva e rápida (Pasquini, 2003). Uma grande desvantagem do uso da NIR é que as sobreposições características desta técnica e sua complexidade podem dificultar a quantificação e interpretação dos dados produzidos (Burns & Ciurczak, 2008). Por outro lado, as bandas de sobreposição (*overtones*) mais largas podem diminuir a necessidade de utilizar um grande número de comprimento de ondas para calibração e análises de rotina (Cozzolino, 2009).

O infravermelho interage com o material através da absorção de um fóton com o momento dipolar da ligação, sendo que somente as vibrações que resultam em uma mudança no momento dipolo da molécula serão observadas na região do IV. As ligações químicas simétricas como H<sub>2</sub>, Cl<sub>2</sub>, e O<sub>2</sub>, por exemplo, não absorvem energia nesta região.

Ligações mais fortes e ligações entre átomos mais leves vibram em frequências maiores. Existem dois tipos de vibração molecular: o estiramento e o dobramento. No estiramento, ocorre vibração ao longo da ligação, podendo ocorrer de forma simétrica ou assimétrica. No dobramento, ocorre mudança no ângulo da ligação entre dois átomos, podendo ser simétrica no plano ou simétrica

fora do plano. O estiramento apresenta comprimentos de onda maiores porque a energia cinética é maior.

A possibilidade do uso da NIR para análise de componentes de biomassa depende da absorção seletiva de luz pelos sobretons e combinação de vibrações das ligações químicas orgânicas, que gera picos largos e muito sobrepostos (Pasquini, 2003). Recentemente, vários estudos resumiram algumas bandas características de absorção a grupos funcionais em amostras de madeira (Li et al., 2015; Xu et al., 2013), aumentando um pouco o entendimento sobre a relação entre os sinais espectrais e a composição química da biomassa.

Na identificação de espécies, exemplo de análise qualitativa com a NIR, o uso da técnica ainda está ligada à algum especialista para identificação correta das amostras e uso destas para calibração dos modelos, etapa essencial no desenvolvimento de modelos de classificação ou discriminação (Pastore *et al.*, 2011).

O espectro NIR pode ser afetado por fatores como ruído de fundo, espalhamento da luz e heterogeneidade da amostra. Nos casos onde se mede a refletância difusa, devem-se utilizar primeira e segunda derivadas, assim como métodos de correção multiplicativa de sinal, visto que diferenças na espessura ou textura do material podem gerar dispersão de luz e conseqüentemente, dados não acurados (Pasquini, 2003). Os métodos de tratamento ajudam a reduzir, eliminar ou padronizar os efeitos indesejáveis presentes no espectro, sem alterar a informação principal contida no espectro gerado (Siesler et al., 2002).

### **2.6.3. Uso da espectroscopia do infravermelho próximo em eucalipto**

Os pesquisadores pioneiros da técnica do infravermelho próximo em madeira foram Birkett e Gambino (1988, 1998), Easty et al. (1990) e Wright *et al.* (1990).

Diversos trabalhos que utilizam a NIR no estudo de espécies do gênero *Eucalyptus* podem ser encontrados na literatura. Na fenotipagem de eucalipto, o procedimento tem sido eficiente e confiável para predizer conteúdo de celulose, lignina, razão entre lignina siringil e guaiacil – S/G, composição de extrativos, densidade e rendimento da polpa (Costa e Silva et al., 2008; Hein et al., 2012; Raymond & Schimleck, 2002; Stackpole et al., 2011). A técnica foi utilizada para análises rápidas de conteúdo de extrativos, lignina e proporção lignina S/G (Baillères et al., 2002), assim como para analisar propriedades químicas e físicas de madeira de híbridos (n=1000) de *Eucalyptus* (*E. grandis*, *E. urophylla*, *E. camaldulensis* e híbridos *E. grandis* x *E. urophylla*) encontrando bons valores de correlação para os modelos, principalmente para as análises químicas como de lignina (solúvel, insolúvel e total) e proporção S/G (Lima, 2014). Em outro trabalho, podem-se diferenciar

indivíduos das espécies *E. nitens* (n=50) e *E. globulus* (n=258), a partir dos espectros de folhas moídas, utilizando apenas 20 amostras na validação (Castillo et al., 2008).

## 2.7. Quimiometria

As técnicas de análise estatísticas multivariadas aplicadas a dados químicos (quimiometria) são necessárias para extrair a informação presente no espectro NIR sobre a composição e definir um modelo de calibração (Ferreira et al, 1999).

O espectro NIR é composto por três regiões de sobretons (3ª: 700 a 1.150nm, 2ª: 1.050 a 1.600nm e 1ª: 1.500 a 2.050nm) e uma banda de combinação (2.000 a 2.500nm)(Xiaobo et al, 2010), especialmente importantes quando se está analisando uma composição química complexa como a de plantas, por exemplo. Outros fatores que também podem influenciar a interpretação do espectro são efeitos de espalhamento dependentes de comprimento de onda, heterogeneidade de tecidos, efeitos ambientais, ruídos do instrumento e outros efeitos que causam variações no espectro. Visando corrigir ou diminuir tais efeitos, tornam-se então fundamentais algumas técnicas multivariadas assim como transformação dos dados para a interpretação dos espectros.

Algumas das transformações utilizadas no pré-processamento de dados de espectroscopia são correção de dispersão multiplicativa (*multiplicative scatter correction* – MSC) e a transformação de variação normal padrão (*standard normal variate* – SNV). A transformação MSC requer um espectro referência ou “ideal”, portanto essa transformação utiliza o espectro médio de todo o set de calibração, utilizando fatores de correção tanto aditivos quanto multiplicativos, onde cada espectro é individualmente regredido ao espectro médio do set de calibração e o coeficiente de regressão linear fornece estimativas deste valor. A transformação MSC é dependente do set de calibração e a retirada de outliers altera o resultado desta transformação (Dhanoa et al, 1994). Por outro lado, a transformação SNV é aplicada a cada espectro individualmente, independente e sem referência a algum set. Essa transformação subtrai a média de cada espectro e depois o espectro é calculado pelo desvio padrão do próprio espectro. Dados espectrais transformados por SNV apresentam média igual a zero e variância igual a 1 (Dhanoa et al., 1994). Outras transformações utilizadas em análises de espectros de infravermelho são Savitzky-Golay (2ª derivada), Detrend, entre outras.

Algumas das análises multivariadas que são utilizadas nas análises de espectros são PCA (análise de componente principal), LDA (análise de discriminante linear), PLS (regressão de mínimos quadrados parciais), PLS-DA (análise discriminante de regressão de mínimos quadrados parciais), entre outras. Para avaliar os modelos, os parâmetros utilizados normalmente são o coeficiente de determinação ( $R^2$ ), erro padrão de predição (SEP) e os valores da razão da performance pelo desvio (RPD)(Karlinasari et al, 2014).

### **3. Objetivos**

#### **3.1. Objetivo geral**

Desenvolver modelos para a predição da proporção de lignina siringil/guaiacil da madeira de eucalipto, baseado em NIR e utilizando dois espectrômetros de infravermelho próximo (Viavi - portátil e FOSS - bancada), em uma população de *E. benthamii* (n=484) e uma população de *E. pellita* (n=736).

#### **3.2. Objetivos específicos**

- 3.2.1.** Desenvolver modelos de classificação de espécies utilizando espectrômetros portátil (MicroNIR 1700) e de bancada (Foss NIRSystems 5000) em duas populações grandes de eucaliptos (*E. benthamii* =484 e *E. pellita* =736) e testar a eficiência para a discriminação de espécies (capítulo 2).
- 3.2.2.** Desenvolver modelos da proporção de lignina siringil:guaiacil em duas populações de eucaliptos (*E. benthamii* e *E. pellita*) com uma amostragem de referência reduzida (n=60), a partir de pequenas quantidades de serragem nas análises químicas (Robinson & Mansfield, 2009)(capítulo 3).
- 3.2.3.** Comparar a calibração e validação de modelos de predição do espectrofotômetro portátil MicroNIR1700 em relação a um espectrofotômetro de bancada FOSS NIRSystems 5000 para as mesmas amostras (capítulo 2 e 3).

#### 4. Referências bibliográficas

- Alves, A., Simoes, R., Stackpole, D. J., Vaillancourt, R. E., Potts, B. M., Schwanninger, M., & Rodrigues, J. (2011). Determination of the syringyl/guaiacyl ratio of Eucalyptus globulus wood lignin by near infrared-based partial least squares regression models using analytical pyrolysis as the reference method. *Journal of Near Infrared Spectroscopy*, 19(5), 343–348. <https://doi.org/10.1255/jnirs.946>
- Andrade, E. N. (1961). *O eucalipto*. Retrieved from <http://www.sifloresta.ufv.br/handle/123456789/3558>
- Araújo, E. S. N. N., Gimenes, M. A., & Lopes, C. R. (2002). Phylogenetic relationships among genera Eucalyptus and Corymbia species based on rDNA internal transcribed spacers sequences. *Scientia Forestalis*, 62, 75–85.
- Associação Brasileira de Celulose e Papel - BRACELPA. (2014). Florestas plantadas 2014. Retrieved May 27, 2016, from <http://bracelpa.org.br/bra2/?q=node/134>
- Baillères, H., Davrieux, F., & Ham-pichavant, F. (2002). Near infrared analysis as a tool for rapid screening of some major wood characteristics in a eucalyptus breeding program. *Annals of Forest Sciences*, 59(5–6), 479–490. <https://doi.org/10.1051/forest:2002032>
- Baker, A., & Walker, S. (2006). Assessment of the relative amenability to vegetative propagation by leafy cuttings of 14 tropical and subtropical eucalyptus and corymbia species. *Plantation Technology in Tropical Forest Science*, 79–87. [https://doi.org/10.1007/4-431-28054-5\\_7](https://doi.org/10.1007/4-431-28054-5_7)
- Barbosa, L. C. A., Maltha, C. R., Silva, V. L., & Colodette, J. L. (2008). Determinação da relação siringila/guaiacila da lignina em madeiras de eucalipto por pirólise acoplada à cromatografia gasosa e espectrometria de massas (PI-CG/EM). *Quimica Nova*, 31(8), 2035–2041. <https://doi.org/10.1590/S0100-40422008000800023>
- Batten, G. D. (1998). Plant analysis using near infrared reflectance spectroscopy: the potential and the limitations. *Australian Journal of Experimental Agriculture*, 38(7), 697–706. <https://doi.org/10.1071/EA97146>
- Bayly, M. J. (2016). Phylogenetic studies of eucalypts: Fossils, morphology and genomes. *Proceedings of the Royal Society of Victoria*, 128(1), 12–24. <https://doi.org/10.1071/RS16002>
- Ben-Gera, I., & Norris, K. H. (1968). Direct Spectrophotometric Determination of Fat and Moisture in Meat Products. *Journal of Food Science*, 33(1), 64–67. <https://doi.org/10.1111/j.1365-2621.1968.tb00885.x>
- Benson, D. H. (1985). Aspects of the ecology of a rare tree species, Eucalyptus benthamii, at Bents Basin, Wallacia. *Cunninghamia*. Retrieved from [http://rbgsyd.nsw.gov.au/\\_\\_data/assets/pdf\\_file/0009/97632/Benson\\_2b.pdf](http://rbgsyd.nsw.gov.au/__data/assets/pdf_file/0009/97632/Benson_2b.pdf)
- Benson, D., & Mcdougall, L. (1998). Ecology of Sydney plant species Part 6 Dicotyledon family Myrtaceae. *Cunninghamia*, 5(4), 808–987.
- Birkett, M. D., & Gambino, M. J. T. (1988). Potential applications for Near Infrared Spectroscopy in the pulping industry. *Paper Southern Africa*, 11(12), 34–38.
- Birkett, M. D., & Gambino, M. J. T. (1998). Estimation of pulp kappa number with near infrared spectroscopy. *Tappi Journal*, 72(9), 193–197.
- Blanco, M., & Villarroya, I. (2002). NIR spectroscopy: A rapid-response analytical tool. *TrAC - Trends in Analytical Chemistry*, 21(4), 240–250. [https://doi.org/10.1016/S0165-9936\(02\)00404-](https://doi.org/10.1016/S0165-9936(02)00404-)



- Bremer, B., Bremer, K., Chase, M., Fay, M., Reveal, J., Soltis, D., & Stevens, P. (2009). An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG III. *Botanical Journal of the Linnean Society*, 161, 105–121. <https://doi.org/10.1111/j.1095-8339.2009.00996.x>
- Bristow, M. (2008). *Growth of Eucalyptus pellita in mixed species and monoculture plantations*. Southern Cross University, Lismore, NSW. Southern Cross University.
- Bristow, M., Vanclay, J. K., Brooks, L., & Hunt, M. (2006). Growth and species interactions of *Eucalyptus pellita* in a mixed and monoculture plantation in the humid tropics of north Queensland. *Forest Ecology and Management*, 233(2–3), 285–294. <https://doi.org/10.1016/j.foreco.2006.05.019>
- Burns, D. A., & Ciurczak, E. W. (2008). *Handbook Of Near-Infrared Analysis* (3rd ed.). CRC Press.
- Carvalho, D. M. de. (2015). *Study on the structure and properties of xylan extracted from eucalyptus , sugarcane bagasse and sugarcane straw*. KTH Royal Institute of Technology.
- Castillo, R., Contreras, D., Freer, J., Ruiz, J., & Valenzuela, S. (2008). Supervised pattern recognition techniques for classification of *Eucalyptus* species from leaves NIR spectra. *Journal of the Chilean Chemical Society*, 53(4), 1709–1713. <https://doi.org/10.4067/S0717-97072008000400016>
- Castro, C. A. de O., Resende, R. T., Bhering, L. L., & Cruz, C. D. (2016). Brief history of *Eucalyptus* breeding in Brazil under perspective of biometric advances. *Ciencia Rural*, 46(9), 1585–1593.
- Çetinkol, Ö. P., Smith-Moritz, A. M., Cheng, G., Lao, J., George, A., Hong, K., ... Holmes, B. M. (2012). Structural and Chemical Characterization of Hardwood from Tree Species with Applications as Bioenergy Feedstocks. *PLoS ONE*, 7(12). <https://doi.org/10.1371/journal.pone.0052820>
- Chaffey, N. (2002). Why is there so little research into the cell biology of the secondary vascular system of trees? *New Phytologist*, 153(2), 213–223. <https://doi.org/10.1046/j.0028-646X.2001.00311.x>
- CIB. (2008). *Guia do Eucalipto: oportunidades para um desenvolvimento sustentável*. Conselho de Informações sobre Biotecnologia. <https://doi.org/10.1111/j.1752-1734.2009.01350.x>
- Costa e Silva, J., Borralho, N. M. G., Araújo, J. A., Vaillancourt, R. E., Potts, B. M., Silva, J. C. E., & Araujo, J. A. (2008). Genetic parameters for growth, wood density and pulp yield in *Eucalyptus globulus*. *Tree Genetics & Genomes*, 5(2), 291–305. <https://doi.org/10.1007/s11295-008-0174-9>
- Cozzolino, D. (2009). Near infrared spectroscopy in natural products analysis. *Planta Medica*, 75(7), 746–756. <https://doi.org/10.1055/s-0028-1112220>
- Da Silva, A. R., Pastore, T. C. M., Braga, J. W. B., Davrieux, F., Okino, E. Y. A., Coradin, V. T. R., ... Do Prado, A. G. S. (2013). Assessment of total phenols and extractives of mahogany wood by near infrared spectroscopy (NIRS). *Holzforschung*, 67(1), 1–8. <https://doi.org/10.1515/hf-2011-0207>
- de Alencar Figueiredo, L., Davrieux, F., Fliedel, G., Rami, J., Chantreau, J., Deu, M., ... Mestres, C. (2006). Development of NIR equations for food grain quality traits through exploitation of a core collection of cultivated sorghum. *Journal of Agricultural and Food Chemistry*, 54, 8501–

8509. <https://doi.org/10.1255/nirn.1047>

- Dhanoa, M. S., Lister, S. J., Sanderson, R., & Barnes, R. J. (1994). The link between multiplicative scatter correction (MSC) and standard normal variate (SNV) transformations of NIR spectra. *Journal of Near Infrared Spectroscopy*, 2(1), 43–47. <https://doi.org/10.1255/jnirs.30>
- Easty, D. B., Berben, S. A., DeThomas, F. A., & Brimmer, P. J. (1990). Near-infrared spectroscopy for the analysis of wood pulp: quantifying hardwood softwood mixtures and estimating lignin content. *Tappi Journal*, 73(10), 257–261.
- Ferreira, C. A., & Silva, H. D. da. (2004). *Eucalyptus para a Região Amazônica, Estados de Rondônia e Acre*. Colombo.
- Ferreira, M. (1979). *Escolha de Espécies de Eucalipto*.
- Ferreira, M. M. C., Antunes, A. M., Melgo, M. S., & Volpe, P. L. O. (1999). Quimiometria I: calibração multivariada, um tutorial. *Química Nova*, 22(5), 724–731.
- Ferreira, V. F., Rocha, D. R., & Silva, F. C. (2009). Potencialidades e oportunidades na química da sacarose e outros açúcares. *Química Nova*, 32(3), 623–638.
- Food and Agriculture Organization of the United Nations – FAO. (2004). *Forest genetic resources conservation and management*. Rome.
- Gill, A. M., Brooker, M. I. H., & Moore, P. H. R. (1992). Seed weights and numbers as a function of fruit size and subgenus in some Eucalyptus species from South-western Australia. *Australian Journal of Botany*, 40(1), 103–111.
- Gomes, F. J. B., Gouvea, A. de F. G., Colodette, J. L., Gomide, J. L., Carvalho, A. M. M. L., Trugilho, P. F., ... Rosado, A. M. (2008). Influence of content and S/G relation of the wood lignin on kraft pulping performance. *O Papel*, 69, 95–105.
- Graça, M. E. C., Shimizu, J. Y., & Tavares, F. R. (1999). *Capacidade de rebrota de enraizamento de Eucalyptus benthamii*. *Boletim de Pesquisa Florestal* (Vol. 39). Colombo. <https://doi.org/10.1017/CBO9781107415324.004>
- Grattapaglia, D., Vaillancourt, R. E., Shepherd, M., Thumma, B. R., Foley, W., Külheim, C., ... Myburg, A. A. (2012). Progress in Myrtaceae genetics and genomics: Eucalyptus as the pivotal genus. *Tree Genetics and Genomes*, 8(3), 463–508. <https://doi.org/10.1007/s11295-012-0491-x>
- Haas, J., & Mizaikoff, B. (2016). Advances in Mid-Infrared Spectroscopy for Chemical Analysis. *Annu. Rev. Anal. Chem.*, 9, 45–68. <https://doi.org/10.1146/annurev-anchem-071015-041507>
- Harwood, C. E., Alloysius, D., Pomroy, P., Robson, K. W., & Haines, M. W. (1997). Early growth and survival of Eucalyptus pellita provenances in a range of tropical environments, compared with E-grandis, E-urophylla and Acacia mangium. *New Forests*, 14(3), 203–219.
- Hein, P. R. G., Bouvet, J. M., Mandrou, E., Vigneron, P., Clair, B., & Chaix, G. (2012). Age trends of microfibril angle inheritance and their genetic and environmental correlations with growth, density and chemical properties in Eucalyptus urophylla S.T. Blake wood. *Annals of Forest Science*, 69(6), 681–691. <https://doi.org/10.1007/s13595-012-0186-3>
- Higa, R. C. V. (1999). Aspectos ecológicos e silviculturais do Eucalyptus benthamii Maiden et Cabbage. *Boletim de Pesquisa Florestal*, 38(1971), 121–122.
- Higa, R. C. V., & Pereira, J. C. D. (2003). *Usos potenciais do Eucalyptus benthamii Maiden et Cabbage*. Colombo.
- Higa, R. C. V., Mora, A. L., & Higa, A. R. (2000). *Plantio de Eucalipto na Pequena Propriedade*

*Rural. Embrapa Florestas* (Vol. 54).

- Higashi, E. N., Silveira, R. L. V. A., & Gonçalves, A. N. (2000). *Propagação vegetativa de Eucalyptus: princípios básicos e a sua evolução no Brasil*.
- Hill, K. D., & Johnson, L. A. S. (1995). Systematic studies in the eucalypts. 7. A revision of the bloodwoods, genus *Corymbia* (Myrtaceae). *Telopea*, 6, 185–504.
- Hou, S., & Li, L. (2011). Rapid Characterization of Woody Biomass Digestibility and Chemical Composition Using Near-infrared Spectroscopy. *Journal of Integrative Plant Biology*, 53(2), 166–175. <https://doi.org/10.1111/j.1744-7909.2010.01003.x>
- Huang, H., Yu, H., Xu, H., & Ying, Y. (2008). Near infrared spectroscopy for on/in-line monitoring of quality in foods and beverages: A review. *Journal of Food Engineering*, 87(3), 303–313. <https://doi.org/10.1016/j.jfoodeng.2007.12.022>
- Huntley, S. K., Ellis, D., Gilbert, M., Chapple, C., & Mansfield, S. D. (2003). Significant increases in pulping efficiency in C4H-F5H-transformed poplars: Improved chemical savings and reduced environmental toxins. *Journal of Agricultural and Food Chemistry*, 51(21), 6178–6183. <https://doi.org/10.1021/jf034320o>
- IBÁ. (2017). *Anuário Estatístico da Indústria Brasileira de Árvores - 2017. Indústria Brasileira de Árvores*. <https://doi.org/10.1017/CBO9781107415324.004>
- Jamrógiewicz, M. (2012). Application of the near-infrared spectroscopy in the pharmaceutical technology. *Journal of Pharmaceutical and Biomedical Analysis*, 66, 1–10. <https://doi.org/10.1016/j.jpba.2012.03.009>
- Jovanovic, T., & Booth, T. (2002). Improved species climatic profiles. *A Report for the RIRDC/L&W Australia/FWPRDC/MDBC Joint Venture Agroforestry Program*, (2), 74.
- Karlinasari, L., Sabed, M., Wistara, I. N. J., & Purwanto, Y. A. (2014). Near infrared (NIR) spectroscopy for estimating the chemical composition of (*Acacia mangium* Willd.) wood. *Journal of the Indian Academy of Wood Science*, 11(2), 162–167. <https://doi.org/10.1007/s13196-014-0133-z>
- Li, X., Sun, C., Zhou, B., & He, Y. (2015). Determination of Hemicellulose, Cellulose and Lignin in Moso Bamboo by Near Infrared Spectroscopy. *Scientific Reports*, 5(October), 17210. <https://doi.org/10.1038/srep17210>
- Lima, A. M. N., Silva, I. R., Neves, J. C. L., Novais, R. F., Barros, N. F., Mendonça, E. S., ... Leite, F. P. (2006). Soil organic carbon dynamics following afforestation of degraded pastures with eucalyptus in southeastern Brazil. *Forest Ecology and Management*, 235(1–3), 219–231. <https://doi.org/10.1016/j.foreco.2006.08.331>
- Lima, B. M. De. (2014). *Bridging genomics and quantitative genetics of Eucalyptus: genome-wide prediction and genetic parameter estimation for growth and wood properties using high-density SNP data*. ESALQ - Universidade de São Paulo.
- Lopez, A., Arazuri, S., Garcia, I., Mangado, J., Jaren, C., & Accepted, J. (2013). A review on the application of Near-Infrared Spectroscopy for the analysis of potatoes. *Journal of Agricultural and Food Chemistry*, 61, 5413–5424. <https://doi.org/10.1021/jf401292j>
- Lupoi, J. S., Singh, S., Davis, M., Lee, D. J., Shepherd, M., Simmons, B. A., & Henry, R. J. (2014). High-throughput prediction of eucalypt lignin syringyl/guaiacyl content using multivariate analysis: a comparison between mid-infrared, near-infrared, and Raman spectroscopies for model development. *Biotechnology for Biofuels*, 7(93), 1–14. [23](https://doi.org/10.1186/1754-</a></p></div><div data-bbox=)

- Mizrachi, E., Mansfield, S. D., & Myburg, A. A. (2012). Cellulose factories: Advancing bioenergy production from forest trees. *New Phytologist*. <https://doi.org/10.1111/j.1469-8137.2011.03971.x>
- Morais, S. A. L. De, Nascimento, E. A. Do, & Melo, D. C. De. (2005). Análise da madeira de Pinus oocarpa parte I: estudo dos constituintes macromoleculares e extrativos voláteis. *Revista Árvore*, 29(3), 461–470. <https://doi.org/10.1590/S0100-67622005000300014>
- Mottiar, Y., Vanholme, R., Boerjan, W., Ralph, J., & Mansfield, S. D. (2016). Designer lignins: Harnessing the plasticity of lignification. *Current Opinion in Biotechnology*, 37, 190–200. <https://doi.org/10.1016/j.copbio.2015.10.009>
- Moura, V. P. G., & Guimarães, D. P. (2003). *Produção de mudas de Eucalyptus para o estabelecimento de plantios florestais. Comunicado Técnico - EMBRAPA* (Vol. 85).
- Müller, B. V., Pereira, M., Cunha, A. B., Klitzke, R. J., & Nicoletti, M. F. (2014). Avaliação das Principais Propriedades Físicas e Mecânicas da Madeira de Eucalyptus benthamii Maiden et Cambage Evaluation of the Main Physical and Mechanical Properties of Eucalyptus benthamii Maiden et Cambage Wood. *Floresta E Ambiente*, 21(4), 535–542.
- Myburg, A. A., Grattapaglia, D., Tuskan, G. A., Hellsten, U., Hayes, R. D., Grimwood, J., ... Schmutz, J. (2014). The genome of Eucalyptus grandis. *Nature*, 510(7505), 356–362. <https://doi.org/10.1038/nature13308>
- Nelson, D. L., & Cox, M. M. (2011). *Princípios de bioquímica de Lehninger*. Porto Alegre: Artmed.
- Pasquini, C. (2003). Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society*, 14(2), 198–219. <https://doi.org/10.1590/S0103-50532003000200006>
- Pereira, J. C. D., Sturion, J. A., Higa, A. R., Higa, R. C. V., & Shimizu, J. Y. (2000). *Características da madeira de algumas espécies de eucalipto plantadas no Brasil*. Colombo.
- Pettersen, R. C. (1984). The chemical composition of wood. *The Chemistry of Solid Wood*, 207, 57–126. <https://doi.org/10.1021/ba-1984-0207>
- Pryor, L. D., & Johnson, L. A. S. (1971). *A classification of the eucalypts (Australian)*. Canberra.
- Pryor, L. D., Williams, E. R., & Gunn, B. V. (1995). A morphometric analysis of Eucalyptus urophylla and related taxa with descriptions of two new species. *Australian-Systematic-Botany*, 8(1), 57–70. <https://doi.org/10.1071/SB9950057>
- Ramadevi, P., Hegde, D. V., Varghese, M., Kamalakannan, R., Ganapathy, S. P., & Gurumurthy, D. S. (2016). Evaluation of Lignin Syringyl/Guaiacyl Ratio in *Eucalyptus Camaldulensis* across Three Diverse Sites Based on near Infrared Spectroscopic Calibration Modelling with Five *Eucalyptus* Species and its Impact on Kraft Pulp Yield. *Journal of Near Infrared Spectroscopy*, 24(6), 529–536. <https://doi.org/10.1255/jnirs.1251>
- Raymond, C. A., & Schimleck, L. R. (2002). Development of near infrared reflectance analysis calibrations for estimating genetic parameters for cellulose content in Eucalyptus globulus. *Canadian Journal of Forest Research-Revue Canadienne De Recherche Forestiere*, 32(1), 170–176. <https://doi.org/10.1139/X01-174>
- Rencoret, J., Gutiérrez, A., Nieto, L., Jiménez-Barbero, J., Faulds, C. B., Kim, H., ... Del Río, J. C. (2011). Lignin composition and structure in young versus adult Eucalyptus globulus plants.

*Plant Physiology*, 155(2), 667–82. <https://doi.org/10.1104/pp.110.167254>

- Rezende, G. D. S. P., Resende, M. D. V. de, & Assis, T. F. de. (2014). Eucalyptus Breeding for Clonal Forestry. *Forestry Sciences*, 81, 393–424. <https://doi.org/10.1007/978-94-007-7076-8>
- Robinson, A. R., & Mansfield, S. D. (2009). Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and near-infrared reflectance-based prediction modeling. *Plant Journal*, 58(4), 706–714. <https://doi.org/10.1111/j.1365-313X.2009.03808.x>
- Roggo, Y., Chalus, P., Maurer, L., Lema-Martinez, C., Edmond, A., & Jent, N. (2007). A review of near infrared spectroscopy and chemometrics in pharmaceutical technologies. *Journal of Pharmaceutical and Biomedical Analysis*, 44(3 SPEC. ISS.), 683–700. <https://doi.org/10.1016/j.jpba.2007.03.023>
- Santos, Á. F. Do, Auer, C. G., & Grigoletti, A. (2001). *Doenças do eucalipto no sul do Brasil: identificação e controle*. Colombo.
- Santos, R. B. Dos, Gomide, J. L., & Sousa, L. C. De. (2009). Predição de qualidade da madeira e da polpa celulósica por técnica de espectroscopia de infravermelho próximo (NIRS). *Revista Árvore*, 33, 759–767. <https://doi.org/10.1590/S0100-67622009000400019>
- Schimleck, L. R., Raymond, C. A., Beadle, C. L., Downes, G. M., Kube, P. D., & French, J. (2000). Applications of NIR spectroscopy to forest research. *Appita Journal*, 53(5), 458–464.
- Schwanninger, M., Rodrigues, J. C., Pereira, H., & Hinterstoisser, B. (2004). Effects of short-time vibratory ball milling on the shape of FT-IR spectra of wood and cellulose. *Vibrational Spectroscopy*, 36(1), 23–40. <https://doi.org/10.1016/j.vibspec.2004.02.003>
- Serpe, E. L. (2015). *Efeitos de diferentes dosagens de adubação no crescimento inicial de Eucalyptus benthamii na região sul do estado do Paraná*. Universidade Estadual do Centro-Oeste.
- Siesler, H. W., Ozaki, Y., Kawata, S., & Heise, H. M. (2002). *Near-Infrared Spectroscopy*. Weinheim: Wiley-VCH.
- Sills, D. L., & Gossett, J. M. (2012). Using FTIR to predict saccharification from enzymatic hydrolysis of alkali-pretreated biomasses. *Biotechnology and Bioengineering*, 109(2), 353–362. <https://doi.org/10.1002/bit.23314>
- Stackpole, D. J., Vaillancourt, R. E., Alves, A., Rodrigues, J., & Potts, B. M. (2011). Genetic Variation in the Chemical Components of Eucalyptus globulus Wood. *G3*, 1(2), 151–9. <https://doi.org/10.1534/g3.111.000372>
- Stanturf, J., Vance, E. D., Fox, T. R., & Kirst, M. (2013). Eucalyptus beyond its native range: environmental issues in exotic bioenergy plantations. *International Journal of Forestry Research*, 1–5. <https://doi.org/10.1155/2013/463030>
- Sun, D., Dickinson, G. R., & Robson, K. J. (1996). Growth of Eucalyptus pellita and E. urophylla and effects on pasture production on the coastal lowlands of tropical northern Australia. *Australian Forestry*, 59(3), 136–141. <https://doi.org/10.1080/00049158.1996.10674679>
- The Plant List. (2013). Retrieved July 14, 2016, from [www.theplantlist.org](http://www.theplantlist.org)
- Uner, B., Karaman, İ., Tanriverdi, H., & Özdemir, D. (2009). Prediction of Lignin and Extractive Content of Pinus nigra Arnold. var. Pallasiana Tree Using Near Infrared Spectroscopy and Multivariate Calibration. *Journal of Wood Chemistry and Technology*, 29(1), 24–42. <https://doi.org/10.1080/02773810802607567>

- Vega-Sánchez, M. E., & Ronald, P. C. (2010). Genetic and biotechnological approaches for biofuel crop improvement. *Current Opinion in Biotechnology*. Elsevier Ltd. <https://doi.org/10.1016/j.copbio.2010.02.002>
- Vuong, Q. V, Chalmers, A. C., Bhuyan, D. J., Bowyer, M. C., & Scarlett, C. J. (2015). Botanical , Phytochemical , and Anticancer Properties of the Eucalyptus Species. *Chemistry & Biodiversity*, 12(6), 907–924. <https://doi.org/10.1002/cbdv.201400327>
- Willför, S., Sundberg, A., Pranovich, A., & Holmbom, B. (2005). Polysaccharides in some industrially important hardwood species. *Wood Science and Technology*. <https://doi.org/10.1007/s00226-005-0039-4>
- Wilson, P. G. (2011). Myrtaceae. *Springer Berlin Heidelberg*, 10, 212–271.
- Wilson, P. G., O'Brien, M. M., Gadek, P. A., & Quinn, C. J. (2001). Myrtaceae revisited: A reassessment of infrafamilial groups. *American Journal of Botany*, 88(11), 2013–2025. <https://doi.org/10.2307/3558428>
- Wright, J., Birkett, M. D., & Gambino, M. J. T. (1990). Prediction of pulp yield and cellulose content from wood using near infrared reflectance spectroscopy. *Tappi Journal*, 73(8), 164–166.
- Xiaobo, Z., Jiewen, Z., Povey, M. J. W., Holmes, M., & Hanpin, M. (2010). Variables selection methods in near-infrared spectroscopy. *Analytica Chimica Acta*, 667(1–2), 14–32. <https://doi.org/10.1016/j.aca.2010.03.048>
- Xu, F., Yu, J., Tesso, T., Dowell, F., & Wang, D. (2013). Qualitative and quantitative analysis of lignocellulosic biomass using infrared techniques: A mini-review. *Applied Energy*, 104, 801–809. <https://doi.org/10.1016/j.apenergy.2012.12.019>
- Yang, J. C., Chung, J. Der, & Chen, Z. Z. (1995). Vegetative propagation of adult *Eucalyptus grandis* X *urophylla* and comparison of growth between micropropagated plantlets and rooted cuttings. *Plant Cell Reports*, 15(3–4), 170–173. <https://doi.org/10.1007/BF00193713>
- Zanata, M., Freitas, M. L. M., Silva, M. T., Morais, E., Zanatto, A. C. S., & Sebbenn, A. M. (2010). Parâmetros genéticos e ganhos na seleção em teste de progênies de polinização aberta de *Eucalyptus pellita*, em Batatais - SP. *Revista Do Instituto Florestal*, 22(2), 233–242.

## Capítulo 2 – Desenvolvimento de modelo para discriminação de espécies utilizando espectroscopia do infravermelho próximo com equipamento portátil e de bancada.

### Introdução

A identificação de espécies vegetais sempre foi uma tarefa difícil, limitada a poucos especialistas. Para a indústria madeireira, a identificação correta das espécies e a avaliação da qualidade da madeira são muito importantes, devido ao valor comercial atribuído a cada tipo de madeira, atividade que depende da escassa disponibilidade de profissionais.

A espectroscopia do infravermelho (NIR), técnica baseada na absorbância do espectro na região 750 a 2.500 nm, permite a discriminação baseada nos compostos orgânicos mais abundantes nos órgãos e tecidos vegetais, entre eles os compostos encontrados em madeiras. Dessa forma, a NIR associada às análises multivariadas, surge como uma tecnologia rápida, barata e acurada, capaz de auxiliar na identificação das espécies lenhosas (Tsuchikawa, 2003).

O objetivo desse capítulo é comparar o desempenho de modelos de classificação de espécies a partir de espectros da serragem (n=1.220) de duas populações de *Eucalyptus* (*E. benthamii* e *E. pellita*) adquiridos em dois espectrômetros, um de bancada (FOSS NIRSystems 5000) e um portátil (Viavi MicroNIR 1700). Equipamentos de bancada são frequentemente equipados com interferômetros, apresentando maior sensibilidade e resolução, assim contribuirão para a geração de modelos mais robustos, além de apresentarem uma maior vida útil. Já os equipamentos portáteis são mais baratos, menores, fáceis de manusear, com interfaces simples e mais ergonômicos (Santos et al., 2013); além de apresentarem a opção de serem utilizados *in situ*.

As análises multivariadas utilizadas nos espectros foram a análise da componente principal (PCA), análise linear discriminante (LDA) e a análise discriminante por regressão de mínimos quadrados parciais (PLS-DA). Os parâmetros de avaliação dos modelos são: erro padrão de calibração, erro padrão de validação cruzada e coeficiente de determinação. Apesar dos indivíduos utilizados serem de duas espécies florestais comerciais e seus indivíduos serem facilmente diferenciados pela folha, os resultados encontrados reforçam o potencial da NIR como técnica para a identificação a partir da serragem da madeira dessas duas espécies, o que não seria evidente morfológicamente.

O texto a seguir foi convidado (09.08.2017) e enviado (20.10.2017) para ser publicado pela comissão organizadora do International Conference of Near Infrared Spectroscopy (ICNIRS 2017, Copenhagen) em revista com Identificador Digital de Objetos (DOI - Digital Object Identifier). Ele se encontra no formato sugerido pela comissão organizadora do ICNIRS 2017.

# Comparative performance of bench and portable near infrared spectrometers for measuring wood samples of two *Eucalyptus* species (*E. pellita* and *E. benthamii*)

C.P. Diniz<sup>a\*</sup>, D. Grattapaglia<sup>b</sup> and L.F. de Alencar Figueiredo<sup>a</sup>.

<sup>a</sup> Universidade de Brasília, Campus Universitário Darcy Ribeiro, Instituto de Ciências Biológicas, Departamento de Botânica, Asa Norte, CEP 70910-900, Brasília, DF, Brazil

<sup>b</sup> Embrapa Recursos Genéticos e Biotecnologia, Parque Estação Biológica, CP 02372, CEP 70770-900, Asa Norte, Brasília, DF Brazil.

Corresponding author's e-mail address: [carolina.pdiniz@gmail.com](mailto:carolina.pdiniz@gmail.com)

## Brief abstract of the paper

The well-established and ever growing applications of NIR spectroscopy for wood science and technology using bench spectrometers is unquestionable. However, the increasing interest in using portable spectrometers due to their perceived advantages raises the question of how good are they? In this work, we acquired spectra from wood sawdust samples of two *Eucalyptus* species, *E. pellita* and *E. benthamii*, totaling 1,200 individual trees sampled in experimental breeding populations. Spectra were acquired in parallel using a bench spectrometer (FOSS NIRSystems 5000) and a portable instrument (Viavi MicroNIR1700). Spectra acquisition on the bench spectrometer had a longer wavelength (1,100 to 2,500nm versus 908 to 1,700nm) and a shorter absorbance profile. Principal component analysis on data of both instruments was not able to clearly separate the two species suggesting some level of overlapping distributions of the chemical composition of their respective wood properties. Linear discriminant analysis however, had excellent accuracies with the bench spectrometer (~99%), showing significantly better discrimination than the one obtained with the portable one (~92%). Similarly, partial least square discriminant analysis showed correlations ( $r$ ) around 0.96 and standard errors of calibration ( $SEC$ ) and of cross validation ( $SEC_V$ ) lower than 0.155 on the bench instrument, while the portable spectrometer had  $r \sim 0.88$  with  $SEC$  and  $SEC_V$  below 0.235. Although the bench instrument showed a better performance, the parameters estimated with the portable spectrometer were very satisfactory given its intrinsic limitations in robustness and handling needs. Chemical analyses for lignin content are in progress which, combined with data transformation and selection of spectra regions, could put the two instruments on similar performance grounds. These results are particularly relevant for rapid and simple wood phenotyping applications in advanced tree breeding operations.

## Introduction

*Eucalyptus* is the largest genus in the Myrtaceae family, currently including 822 described species<sup>1</sup>. Most species from this genus are native to Australia<sup>2</sup> and some are widely used for commercial and industrial purposes. Commercial eucalypt species are model woody plants because of their high adaptability and fast growth rate in tropical climates<sup>3</sup>, especially *E. grandis* and *E. globulus* upon which breeding programs on tropical and temperate regions respectively are mainly based on. *Eucalyptus benthamii* and *E. pellita* on the other hand, despite their fast growth are yet to be more widely adopted species. While *E. benthamii* displays resistance to periodic frosts in subtropical regions<sup>4</sup>, *E. pellita* is particularly interesting for harsh tropical climates<sup>5</sup> due to its heat tolerance, and resistance to some tropical pest and diseases<sup>6</sup>. These two species used in hybrid breeding with *E. grandis* show great potential to expand high productivity plantations to areas currently considered marginal due to limiting environmental regimes.

*Eucalyptus* wood is mostly composed by cellulose (68 - 78%), lignin (18.7 - 22.6%) and hemicellulose (24 - 28%)<sup>7,8</sup>. Cellulose is mainly used to produce Kraft pulp and paper products<sup>9</sup>. The main use of highly lignified wood in the industry is for biomass, roundwood and wood-based panels' production<sup>10</sup>. *Eucalyptus* wood lignin contains syringyl (S) and guaiacyl (G) monomers<sup>11</sup> and the higher syringyl to guaiacyl (S/G) ratio, the cheaper it is to remove lignin from the pulp to produce high quality fiber and paper<sup>11,12</sup>. Besides high volume wood productivities, one of the main targets of eucalypt breeding programs is to increase the S/G ratio to reduce processing costs and enhance the quality of end-products<sup>13</sup>. Measuring wood properties in large numbers of individuals sampled in progeny and clonal trials along breeding programs is a laborious and expensive endeavor. Near infrared spectroscopy (NIR) for rapid and inexpensive wood phenotyping has therefore been an extremely useful tool in operational tree breeding<sup>14</sup>.

Bench spectrometers are often equipped with interferometers that have better optical properties than the portable ones. They are superior on signal-to-noise ratio, sensitivity and optical resolution<sup>15</sup>. Besides, they can be used for several years



allowing the production of more robust models. Although they are expensive, bench instruments are becoming smaller and less expensive<sup>16</sup>. On the other hand, portable instruments provide new applications, especially the possibility of *in situ* acquisition, offer a larger number of different technical specifications, are usually smaller, easier to use, have a simple user interface, are ergonomically designed, and cheaper<sup>15</sup>.

The objective of this work was to compare the performance of spectra acquired with a bench and a portable spectrometer of wood dust samples of two *Eucalyptus* species (*E. pellita* and *E. benthamii*). Performance was evaluated by the discrimination power using geometric and multivariate analysis. The spectra analysis will later be correlated with ongoing chemical analysis carried out on a sample subset selected based on the most representative spectra.

## Materials and Methods

### Sample collection and preparation

This study was carried out on progeny trials of *E. benthamii* and *E. pellita* that are part of a pilot genomic selection-based breeding program of EMBRAPA (Brazilian Agricultural Research Corporation)<sup>17</sup>. The *E. benthamii* progeny trial, established with 40 seed sources from wild Australian populations, and four bulked seed sources, was planted in May 2007 in Candói, PR, Brazil. The experiment was thinned following 25 heavy frosts recorded. The remaining 508 trees were ultimately sampled at age 56 months for growth and wood samples collected. The *E. pellita* breeding trial was composed of 24 open pollinated families derived from a second-generation clonal seed orchard located in Mareeba, Queensland, Australia planted in February 2010 in Rio Verde, GO, Brazil. Phenotypic evaluations were carried out at age 42 months for growth and wood samples collected. Sawdust samples from individual trees were collected using a power drilling machine (40cm long, ½ inch diameter), at diameter at breast height (1.3m), all in the same geographical orientation. The bark was removed and the sawdust was collected and stored in paper bags for drying. Afterwards, the sawdust samples were individually ground and sifted with a 32 mesh (500µm) screen and stored in capped Falcon tubes to prevent both contamination and humidity variation.

### Spectra acquisition

The 1,200 sawdust samples collected, were individually placed in ring cups and the spectra acquired in both bench and portable spectrometer on the same day. We used the bench instrument, FOSS NIRSystems 5000 (1,100 to 2,500nm, 2nm intervals) with two readings per sample and the portable spectrometer, Viavi (JDSU) MicroNIR™1700 (908 to 1,676nm, 10nm intervals) with three readings per sample.

### Data analysis

Data from both instruments were analyzed separately using The Unscrambler X® (version 10.2) software. First, a plotline from the raw data was generated to identify and remove any outlier, and the data was analyzed as follows. i) principal component analysis (PCA) performed with 11 components; ii) linear discriminant analysis (LDA) and; iii) partial least square for discriminant analysis (PLS-DA). In order to improve the results, some data transformations were applied to the spectral data. The main data treatment used were standard normal variate (SNV), multiple scatter correction (MSC), Detrend, first and second derivatives using Savitzky-Golay method (4<sup>th</sup>-order polynomial fit, 25 smoothing points)<sup>18</sup>. Spectra from the bench spectrometer were also analyzed for a shorter region (from 1,100 to 1,676nm), that overlaps with the range of the portable instrument and displays some variables found to show correlation to some wood traits like lignin (1672nm), cellulose (1428, 1488 and 1590nm)<sup>19</sup>, and extractives (1,434 to 1,477nm, 1,626 to 1,694nm and 1,000 to 2,500)<sup>20</sup>.

## Results and Discussion

### Spectra analysis

A total of 6,100 spectra were acquired, 2,440 with the bench and 3,660 with the portable one (Figures 1 and 2). Spectra appeared to be homogenous and did not separate the two species (data not shown).

The overlapping spectra were not surprising given the phylogenetic proximity of the two species that despite grouped in different sections, belong to the same subgenus (*Symphyomyrtus*)<sup>2</sup>. Wood species from more distant taxonomic units such as different families or genera usually present different spectra profiles<sup>21</sup>.

The spectra from the portable instrument showed a broader profile, with a smaller number of peaks, again not unsurprising given the larger range of the bench instrument (1,100 – 2,498nm) when compared to the one of the portable NIR (908 – 1,676nm). Besides, a larger human interference is likely in the portable Viavi instrument because of the necessary handling to acquire the closer reading of each sample.

### Principal Component Analysis (PCA)

The PCA for untransformed spectra from both the bench and the portable spectrometers explained around 99% of the variance, with the two first components explaining 94% and 5% and 89 and 9%, respectively. These PCA did not discriminate the wood samples of two species (Figures 3 and 4). The different treatments applied to the spectra reduced the explained variance for the two first components on the bench and portable spectrometers from 99 to 80% and from 98 to 86% respectively, for both SNV and MSC (data not shown), which indicates a higher complexity not revealed by all the spectra wavelengths at untreated spectra. However, these treatments still did not separate the species analyzed. Even using only spectra for regions corresponding to specific wood chemical components (lignin, syringyl to guaiacyl (S/G) ratio, cellulose, and extractives) it was not possible to separate these species, suggesting that the two species despite their contrasting ecological adaptation profiles, show relatively overlapping wood chemical properties although with clear extremes even in the PCA (Figures 3 and 4). The use of spectra regions for specific wood components did not improve the results of spectra profile, LDA, and PLS-DA.

### Linear Discriminant Analysis (LDA)

LDA on the other hand allowed good discrimination of the samples of the two *Eucalyptus* species. Data from the bench spectrometer showed the best results in the LDA with and without data transformation with accuracies higher than 98% (Table 1). The highest accuracy with the first derivative followed by Detrend resulted that among all 484 *E. pellita* and 736 *E. benthamii* samples analyzed, only eleven and sixteen respectively, were not correctly assigned to their respective species. For the untreated spectra, 31 samples were mismatched.

Portable spectrometer data analyses for treated spectra presented however a slightly lower accuracy (86.49% to 93.45%) although statistically significant (*t* Test,  $\alpha=5\%$ ). The untreated spectra presented 90.83% accuracy, with 112 samples not correctly assigned. In the data set transformed by SNV and MSC, slightly higher accuracies were obtained (91.9 and 92%), resulting in a mismatch of 98 and 100 samples, respectively. Even though data transformations increased slightly the accuracy, it was not statistically significant.

Furthermore, spectra for a shorter range (1,100 to 1,676nm) were collected with the bench spectrometer and LDA performed to approximate the range used by the portable spectrometer. The results were similar to those obtained with the whole spectra (1,100 to 2,498 nm) and were still better than those obtained from the portable spectrometer (data not shown). This result suggests that a higher wavelength will not result in a better LDA. Likewise, the slight increase on accuracies in both instruments after data treatment reinforces that the wavelength and handling to acquire the spectra were not responsible for the better performance of the bench instrument.

### Partial Least Square Regression for Discriminant Analysis (PLS-DA)

According to the parameters used to evaluate the PLS-DA, the bench spectrometer was statistically better (*t* Test, 5%) than the portable one (Table 2 and 3), except for the second derivative with Savitzky-Golay (S-G). PLS-DA for the bench spectrometer (Table 2) showed both a higher correlation and coefficient of determination ( $R^2$ ), and both a lower standard error of calibration (SEC) and cross validation (SECV). Spectra treated with the first derivative with S-G showed the best results ( $R^2_{cal}=0.942$  and  $R^2_{val}=0.939$ ), very close to SNV and MSC ( $R^2_{cal}=0.928$  and  $R^2_{val}=0.922$ ; for both). That spectra treatment was also the best with the LDA.

On the other hand, spectra acquired with the portable spectrometer (Table 3), showed a lower correlation and a larger standard error. The best results were from data transformed by SNV ( $R^2_{cal}=0.808$  and  $R^2_{val}=0.802$ ), MSC ( $R^2_{cal}=0.807$  and  $R^2_{val}=0.801$ ) and, surprisingly, raw data ( $R^2_{cal}=0.806$  and  $R^2_{val}=0.799$ ). Detrend and both first and second derivatives (Savitzky-Golay) showed lower correlations, with  $R^2_{cal}$  ranging from 0.771 to 0.774, while SEC and SECV ranged between 0.215 and 0.238.

### Conclusion

A comparative performance analysis of a bench and a portable near infrared spectrometer for measuring wood samples of two *Eucalyptus* species (*E. pellita* and *E. benthamii*) was carried out on a large set of 1,200 sawdust samples from field trees sampled in progeny trials. Although it is indisputable that the bench instrument had an overall better performance, the portable instrument had a very satisfactory performance given its intrinsic limitations and handling needs. PCA on

spectra data from both spectrometers was not able to discriminate the samples of the two species. The portable spectrometer had a higher scatter profile which likely reflects its lesser robustness and the necessary handling during spectra acquisition. LDA and PLS-DA on data from the bench spectrometer provided statistically better results than those obtained with data from the portable instrument. For most parameters estimated in this study, the bench spectrometer provided 7-10% higher accuracies and correlations when compared to the portable one. These results were observed with and without spectra transformation, and also using known spectra regions for specific wood chemical components. At the end, a slight improvement was observed on parameters analyzed (accuracies,  $r$ ,  $R^2$ , SEC, and SECV) in both instruments. The chemical analysis currently in progress will allow further and more in depth assessments on the comparative ability of the two instruments to effectively predict specific wood chemical traits relevant to tree breeding.

## Acknowledgement

We thank Dr. Tereza Pastore from the Laboratory of Forest Products and the Professor Dr. Jez Braga (University of Brasília, Chemistry Department) for the availability of the Viavi MicroNIR1700 used in this work acquired through Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), project number 473936/2013-5. We also thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior for the scholarship provided during the masters course for the first author and Fundação de Apoio à Pesquisa do Distrito Federal for the sponsorship to attend the ICNIRS 2018. We also acknowledge Paulo E.T. Santos and Estefano Paludzyszyn Filho of Embrapa Forestry Center for providing the raw wood samples for this study. This work was also supported by PRONEX-FAP-DF grant 2009/00106-8 'NEXTREE', CNPq grant 400663/2012-0 and EMBRAPA grant 03.11.01.007.00.00 to DG. DG had a research fellowship from CNPq.

## References

1. The Plant List, <http://www.theplantlist.org/> (2013, accessed 18 October 2017).
2. M.I.H. Brooker, "A new classification of the genus *Eucalyptus* L'hér. (Myrtaceae)", *Australian Systematic Botany* **13**, 79–148 (2000). <https://doi.org/10.1071/SB98008>
3. G.D.S.P. Rezende, M.D.V. de Resende and T. F. de Assis, "*Eucalyptus* breeding for clonal forestry", *Challenges and opportunities for the world's forests in the 21st Century*, *Forestry Sciences* **81**, 393–424 (2014). [http://doi.org/10.1007/978-94-007-7076-8\\_16](http://doi.org/10.1007/978-94-007-7076-8_16)
4. R. Arnold, B. Li, J. Luo, F. Bai and T. Baker, "Selection of cold-tolerant *Eucalyptus* species and provenances for inland frost-susceptible, humid subtropical regions of southern China", *Australian Forestry* **78**, 180–193 (2015). <http://dx.doi.org/10.1080/00049158.2015.1063471>
5. C.E. Harwood, D. Alloysius, P. Pomroy, K.W. Robson and M.W. Haines, "Early growth and survival of *Eucalyptus pellita* provenances in a range of tropical environments, compared with *E. grandis*, *E. urophylla* and *Acacia mangium*", *New Forests* **14**, 203–219 (1997). <https://doi.org/10.1023/A:1006524405455>
6. L.M.S. Guimarães, M. Titon, D. Lau, L.N. Rosse, L.S.S. Oliveira, C.C.G. Rosado, G.G.O. Christo and A.C. Alfenas, "*Eucalyptus pellita* as a source of resistance to rust, ceratocystis wilt and leaf blight", *Crop Breeding and Applied Biotechnology* **10**, 124–131 (2010).
7. Ö.P. Çetinkol, A.M. Smith-Moritz, G. Cheng, J. Lao, A. George, K. Hong, R. Henry, B.A. Simmons, J.L. Heazlewood and B.M. Holmes, "Structural and chemical characterization of hardwood from tree species with applications as bioenergy feedstocks", *PLoS ONE* **7**, 1–11 (2012). <http://doi.org/10.1371/journal.pone.0052820>
8. C.A. Nunes, C.F. Lima, L.C.A. Barbosa, J.L. Colodette and P.H. Fidêncio, "Determinação de constituintes químicos em madeira de eucalipto por pi-cg/em e calibração multivariada: comparação entre redes neurais artificiais e máquinas de vetor suporte", *Química Nova* **34**, 279–283 (2011).
9. A. van Heiningen, "Converting a kraft pulp mill into an integrated forest biorefinery", *Pulp & Paper Canada* **107**, T141-146 (2006).
10. Food and Agriculture Organization of the United Nations, *Statistical Yearbook – Latin America and the Caribbean food and agriculture*. FAO Publications, Santiago, Chile (2014).
11. J.C. del Río, A. Gutiérrez, I.M. Rodríguez, D. Ibarra and Á.T. Martínez, "Composition of non-woody plant lignins and cinnamic acids by Py-GC/MS, Py/TMAH and FT-IR", *Journal of Analytical and Applied Pyrolysis* **79**, 39–46 (2007). <http://doi.org/10.1016/j.jaap.2006.09.003>
12. P. Ramadevi, D.V. Hegde, M. Varghese, R. Kamalakannan, S.P. Ganapathy and D.S. Gurumurthy, "Evaluation of lignin syringyl/guaiacyl ratio in *Eucalyptus camaldulensis* across three diverse sites based on near infrared spectroscopic calibration modelling with five *Eucalyptus* species and its impact on Kraft pulp yield", *Journal of Near Infrared Spectroscopy* **24**, 529–536 (2016). <http://doi.org/10.1255/jnirs.1251>
13. H. Baillères, F. Davrieux and F. Ham-Pichavant, "Near infrared analysis as a tool for rapid screening of some major wood characteristics in a *Eucalyptus* breeding program", *Annals of Forest Science* **59**, 479–490 (2002). <http://doi.org/10.1051/forest:2002032>
14. C.A. Raymond and L.R. Schimleck, "Development of near infrared reflectance analysis calibrations for estimating genetic parameters for cellulose content in *Eucalyptus globulus*", *Canadian Journal of Forest Research* **32**, 170–176 (2002).
15. C.A.T. dos Santos, M. Lopo, R.N.MJ Páscoa, J.A. Lopes, "A review on the applications of portable near-infrared spectrometers in the agro-food industry", *Applied Spectroscopy* **67**, 1215–1233 (2013).
16. H. Mark and M. Bradley, "Review of new spectroscopic instrumentation for 2016", *Spectroscopy* **31**, 40–62 (2016).

17. B.S.F. Müller, L.G. Neves, J.E. de Almeida Filho, M.F.R. Resende Jr, P.R. Muñoz, P.E.T. dos Santos, E. Paludzyszyn Filho, M. Kirst and D. Grattapaglia, "Genomic prediction in contrast to a genome-wide association study in explaining heritable variation of complex growth traits in breeding populations of *Eucalyptus*", *BMC Genomics* **18**, 1–17 (2017). <http://doi.org/10.1186/s12864-017-3920-2>
18. J.S. Lupoi, S. Singh, B.A. Simmons and R.J. Henry, "Assessment of lignocellulosic biomass using analytical spectroscopy: an evolution to high-throughput techniques", *Bioenergy Research* **7**, 1–23 (2014). <http://doi.org/10.1007/s12155-013-9352-1%20>
19. A. Sandak, J. Sandak and M. Negri, "Relationship between near-infrared (NIR) spectra and the geographical provenance of timber", *Wood Science and Technology* **45**, 35–48 (2011). <http://doi.org/10.1007/s00226-010-0313-y>
20. A.M.M. Alves, R.F.S. Simões, C.A. Santos, B.M. Potts, J. Rodrigues and M. Schwanninger, "Determination of *Eucalyptus globulus* wood extractives content by near infrared-based partial least squares regression models: comparison between extraction procedures", *Journal of Near Infrared Spectrosc.* **20**, 275–285 (2012). <http://doi.org/10.1255/jnirs.946>
21. S. Tsuchikawa, K. Inoue, J. Noma and K. Hayashi, "Application of near-infrared spectroscopy to wood discrimination", *Journal of Wood Science* **49**, 29–35 (2003).

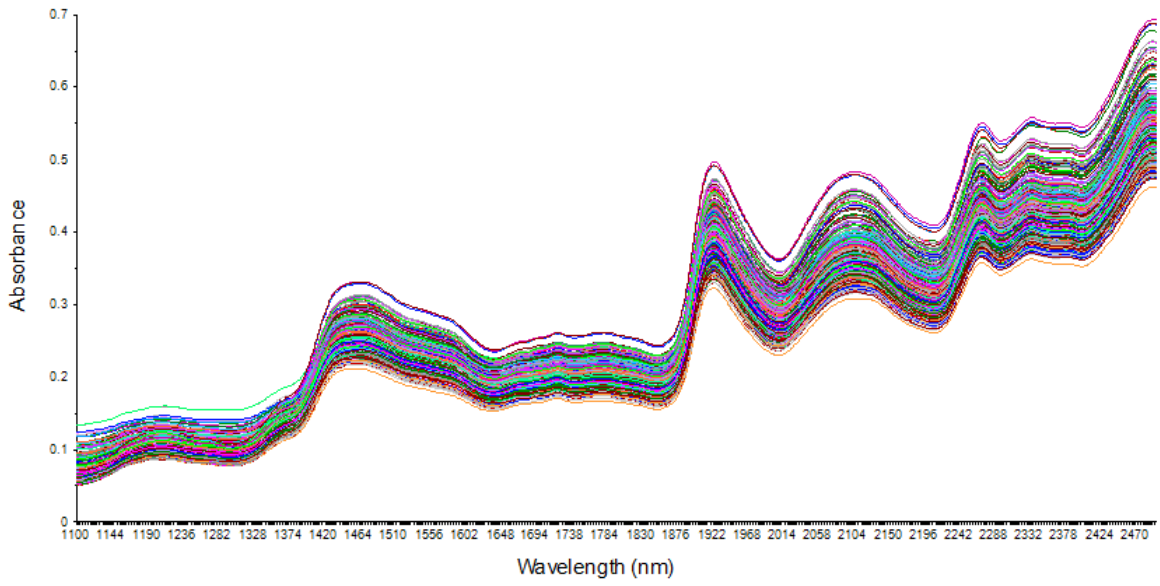


Figure 1. NIR spectra (n=2,440) of *Eucalyptus* species from bench spectrometer (FOSS NIRSystems 5000).

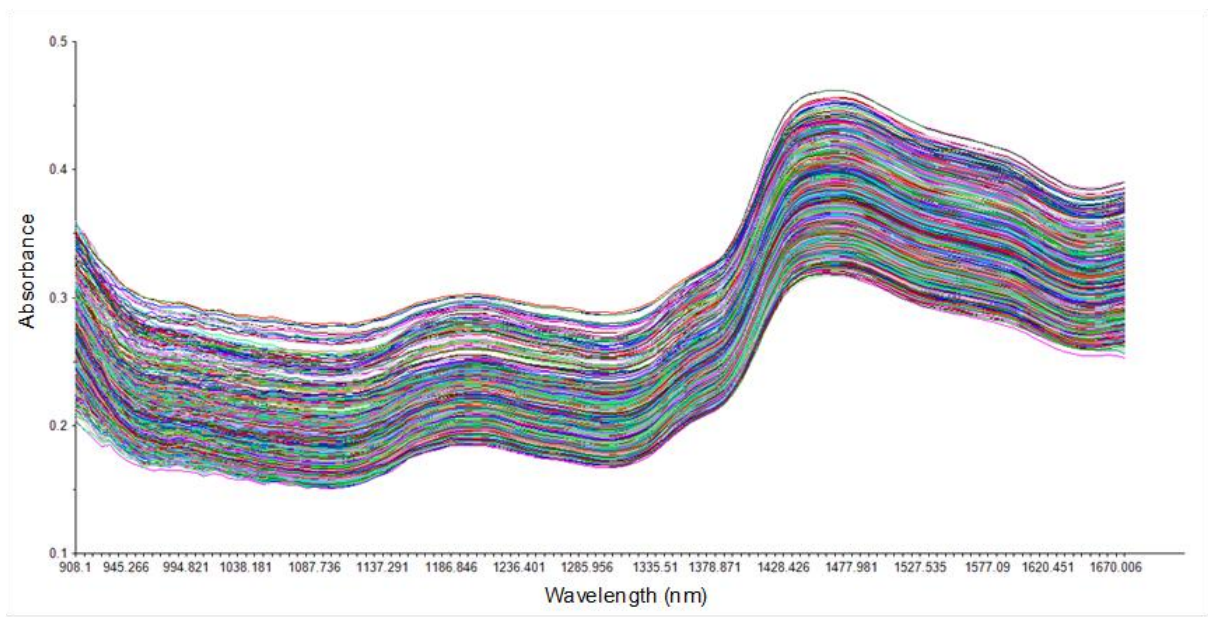


Figure 2. NIR spectra (n=3,660) of *Eucalyptus* species from portable spectrometer (Viavi MicroNIR1700).

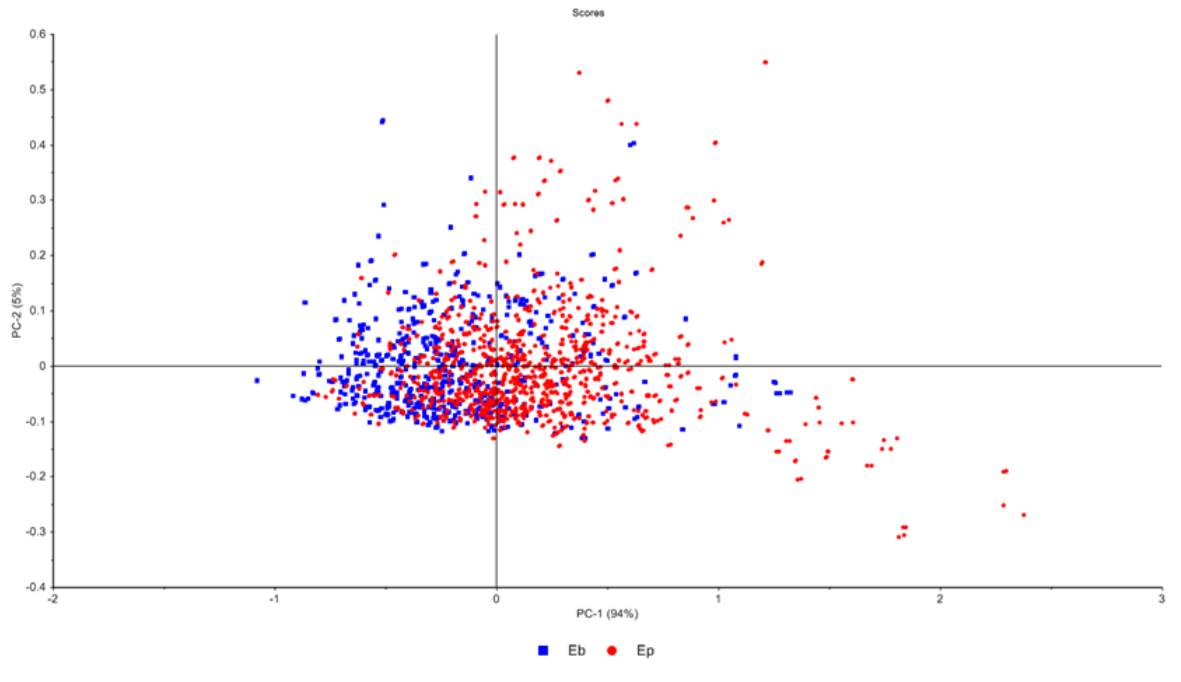


Figure 3. PCA from spectra (n=2,440) acquired on bench spectrometer (FOSS NIRSystems 5000) from *Eucalyptus benthamii* (Eb - blue squares) and *Eucalyptus pellita* (Ep - red circles).

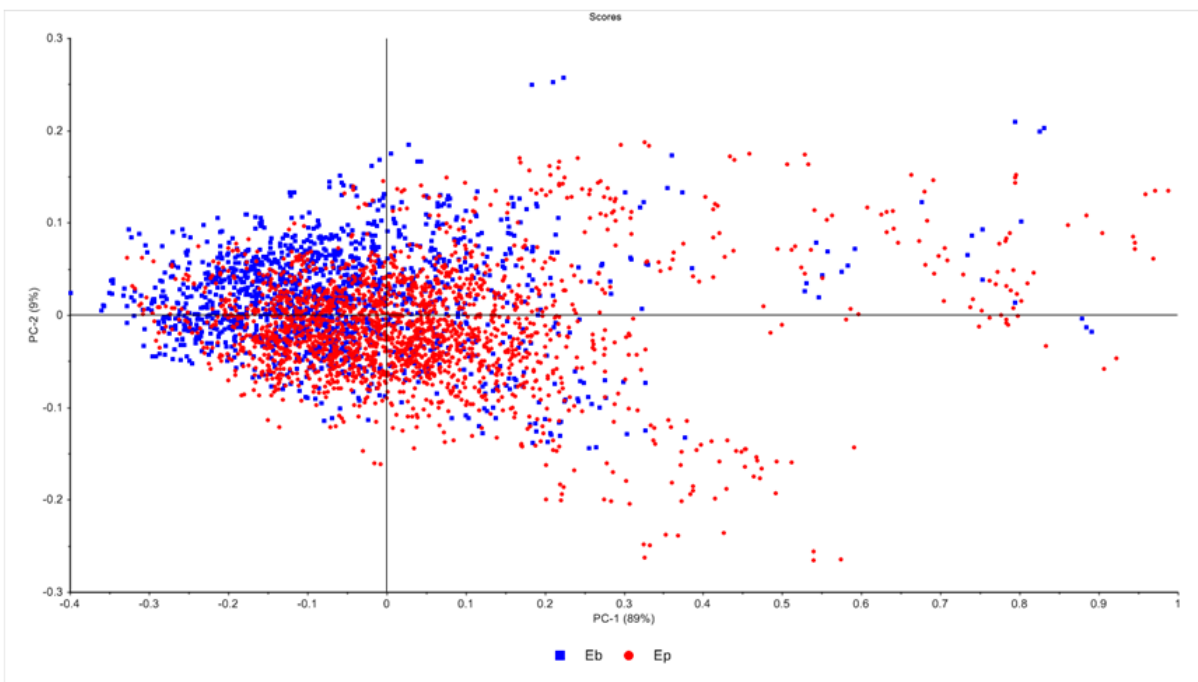


Figure 4. PCA from spectra acquired (n=3,660) in the portable spectrometer (Viavi MicroNIR1700) from *Eucalyptus benthamii* (Eb - blue squares) and *Eucalyptus pellita* (Ep - red circles).

**Table 1. LDA from spectral data acquired in bench (n=2,440) and portable (n=3,660) spectrometer from *E. pellita* and *E. benthamii*.**

Spectrometer	Raw	SNV	MSC	Detrend	1 <sup>st</sup> D (S-G)	2 <sup>nd</sup> D (S-G)
Bench	97.46a	98.61a	98.61a	98.69a	99.14a	95.57b
Portable	90.83c	91.86c	91.97b	91.24b	93.45b	86.49c

Accuracy is given in percentage. SNV: standard normal variate; MSC: multiple scatter correction; D: derivative; S-G: Savitzky-Golay. Different lowercase letters (a–c) indicate significant differences using Student *t* test ( $p < 0.05$ ).

**Table 2. PLS-DA from spectra (n=2,440) acquired on bench spectrometer (FOSS NIRSystems 5000) in classification of *E. pellita* and *E. benthamii*.**

Data transformation		Raw	SNV	MSC	Detrend	1 <sup>st</sup> D (S-G)	2 <sup>nd</sup> D (S-G)
Calibration	Correlation	0.951	0.964	0.963	0.958	0.970	0.957
	$R^2$	0.905	0.928	0.928	0.917	0.942	0.915
	SEC	0.151	0.131	0.131	0.141	0.118	0.143
Validation	Correlation	0.950	0.960	0.960	0.955	0.969	0.955
	$R^2$	0.903	0.922	0.922	0.912	0.939	0.913
	SECV	0.153	0.137	0.136	0.145	0.121	0.145

SNV: standard normal variate; MSC: multiple scatter correction; D: derivative; S-G: Savitzky-Golay;  $R^2$ : coefficient of determination; SEC: standard error of calibration; SECV: standard error of cross validation.

**Table 3. PLS-DA from spectra (n=3,660) acquired on portable spectrometer (Viavi MicroNIR1700) in classification of *E. pellita* and *E. benthamii*.**

Data transformation		Raw Data	SNV	MSC	Detrend	1 <sup>st</sup> D (S-G)	2 <sup>nd</sup> D (S-G)
Calibration	Correlation	0.898	0.897	0.898	0.880	0.879	0.878
	$R^2$	0.806	0.808	0.807	0.774	0.772	0.771
	SEC	0.216	0.215	0.215	0.233	0.233	0.234
Validation	Correlation	0.894	0.895	0.895	0.875	0.877	0.874
	$R^2$	0.799	0.802	0.801	0.765	0.770	0.763
	SECV	0.219	0.218	0.218	0.237	0.235	0.238

SNV: standard normal variate; MSC: multiple scatter correction; D: derivative; S-G: Savitzky-Golay;  $R^2$ : coefficient of determination; SEC: standard error of calibration; SECV: standard error of cross validation.

## Conclusões e Perspectivas

Pela primeira vez, duas grandes populações de *E. benthamii* (n=484) e *E. pellita* (n=736) tiveram os espectros da serragem de suas árvores adquiridas por um espectrômetro de bancada (FOSS NIRSystem 5000) e um portátil (Viavi MicroNIR 1700), na região do infravermelho próximo; de 1.100 a 2.500 nm e de 908 a 1.676 nm, respectivamente.

No gráfico de linhas e na PCA, as duas espécies não apresentaram separação, isso com ou sem tratamento dos espectros, em ambos espectrômetros. Os parâmetros considerados para a comparação entre os modelos na LDA são as acurácias e na PLS-DA são os:  $R^2$ , SEC e SECV. Na LDA, a classificação/discriminação das duas espécies foi muito boa no equipamento de bancada (acurácia = 99%) e levemente inferior no equipamento portátil (93%). Na PLS-DA, o equipamento de bancada também apresentou resultados superiores aos do equipamento portátil, com  $R^2$  de 0,94 e 0,80; respectivamente.

Considerando os dois equipamentos utilizados neste trabalho, mesmo com o desempenho em geral inferior, o equipamento portátil é uma boa escolha para projetos que visem a identificação de espécies, sobretudo utilizando a serragem da madeira. Esse equipamento é mais barato, menor e prático para aquisição dos espectros. Por outro lado, quando comparado com o equipamento de bancada, a sua amplitude na absorção dos espectros é maior e o seu comprimento de onda é menor. Esta última desvantagem impossibilita a aquisição de espectros para alguns compostos orgânicos de grande importância.

Inicialmente, os dois espectrômetros não mostraram grande diferença na classificação das espécies em estudo, baseado na LDA (acurácia de 99 versus 93%, respectivamente). De posse das análises da proporção de lignina S/G poderemos concluir sobre a diferença de ambos no desenvolvimento de modelos com dados quantitativos, que será discutido no próximo capítulo (Capítulo 3).



### Capítulo 3 – Desenvolvimento de modelos de predição da proporção siringil/guaiacil utilizando espectroscopia do infravermelho próximo.

#### Introdução

Os métodos tradicionais de quantificação de celulose, lignina e outros compostos na madeira são longos, onerosos e laboriosos (Raymond & Schimleck, 2002). Em programas de melhoramento, muitos descendentes são gerados e precisam ser analisados nas etapas de seleção dos indivíduos. A espectroscopia do infravermelho próximo (NIR) é uma tecnologia que associada aos métodos de análises químicas tradicionais pode fornecer predições precisas a um número elevado de amostras. Apesar de apresentar um alto custo na implementação, o custo da NIR diminui com o tempo, devido a uma menor demanda de trabalho, custos nas análises e manutenção dos equipamentos em longo prazo (Raymond & Schimleck, 2002).

Diversos trabalhos têm demonstrado que o uso de NIR é eficiente e confiável para fenotipagem da madeira (Hein et al., 2012; Robinson & Mansfield, 2009; Schimleck et al., 2000). Entre as principais análises estão a predição do conteúdo de lignina, proporção siringil/guaiacil, celulose, e extrativos (Baillères et al., 2002; Raymond & Schimleck, 2002; Stackpole et al., 2011). Na indústria de celulose, a proporção S/G é uma das características químicas mais relevantes, visto que quanto maior essa relação, menores são os custos da produção da polpa de celulose (Ramadevi et al., 2016).

Neste trabalho, a NIR foi utilizada com o objetivo de desenvolver modelos de predição da proporção de lignina S/G utilizando um espectrômetro portátil (Viavi MicroNIR 1700) e um de bancada (FOSS NIRSystems 5000). Os parâmetros considerados na avaliação dos modelos foram: coeficiente de determinação da validação ( $R^2$ ), razão da performance pelo desvio (RPD), desvio padrão relativo [média do coeficiente de variação dos valores de referência pelos preditos (RSD)] e correlação do ranqueamento de Sperman ( $r_s$ ).

As espécies utilizadas neste trabalho são *E. benthamii* e *E. pellita*. O destaque para *E. benthamii* está na resistência à geadas (Higa, 1999) e rápido crescimento (Graça et al., 1999), sendo muito utilizado para produção de celulose, lenha e carvão (Higa & Pereira, 2003), mas sua madeira não é recomendada para usos estruturais (Higa & Pereira, 2003; Müller et al., 2014). Já *E. pellita* apresenta uma madeira mais rica em lignina e extrativos, tornando uma madeira mais pesada (Pereira et al., 2000) sendo destinada para o uso em pisos, revestimento, painéis, soleiras, construções e estruturas (Zanata et al., 2010). O interesse em *E. pellita* está na facilidade de propagação (Baker & Walker, 2006), rápido crescimento inicial (Bristow et al., 2006; Sun et al., 1996) e resistência a doenças (Harwood et al., 1997).

Este artigo foi submetido à revista "Bioresource Technology".

Near-infrared based models for lignin syringyl/guaiacyl ratio of *Eucalyptus benthamii* and *E. pellita* using a streamlined thioacidolysis procedure as the reference method.

Carolina Pinto Diniz<sup>a</sup>, Dario Grattapaglia<sup>b,c</sup>, Shawn D. Mansfield<sup>d</sup>, Lúcio Flávio de Alencar Figueiredo<sup>a\*</sup>.

<sup>a</sup>Botany Department, Biology Institute, University of Brasília, CP: 04457, Brasília 70910-900, DF, Brazil. <sup>b</sup>Plant Genetics Laboratory, Genetic Resources and Biotechnology, Brasília, 70770-910, DF, Brazil. <sup>c</sup>Graduate Program in Genomic Sciences and Biotechnology, University Catholic of Brasília, SGAN 916 Modulo B, Brasília, 70790-160, DF, Brazil <sup>d</sup>Department of Wood Science, University of British Columbia, 4030-2424 Main Mall, Vancouver, BC, Canada, V6T 1Z4.

## Abstract

The lignin syringyl/guaiacyl (S/G) ratio is a key trait in the cellulose pulping industry. We used near-infrared (NIR) spectroscopy data collected with a bench and a portable instrument to develop prediction models for S/G ratios from sawdust samples of 1,220 trees of two contrasting *Eucalyptus* species. Based on the spectra data, samples were selected for S/G analysis using a streamlined thioacidolysis method with minimal sample input. NIR models were developed for each species separately and jointly. Instead of just using the coefficient of determination and the ratio of performance deviation, we taking into account the average of the relative standard deviation and the Spearman's rank correlation to evaluate the models. The bench spectrometer performance was superior. NIR models under single species were better for *E. benthamii*, while the joint model was better for *E. pellita*. These NIR models should prove useful for high throughput wood phenotyping in advanced breeding programs.

## Keywords:

Portable spectrometer; bench spectrometer; Spearman's rank correlation; relative standard deviation; NIR models; S/G ratio.

## 1. Introduction

Lignin syringyl (S) and guaiacyl (G) monolignols are the two most abundant lignin monomers in the secondary cell wall, with their appearance was a key event in the evolution of land plants (Weng and Chapple, 2010). The monomer S content is higher in angiosperms while the latter G prevails in

---

\*Corresponding author: lucioalencar@unb.br

gymnosperm. The proportion of syringyl to guaiacyl (S/G ratio) affects biomass pretreatment efficiency, especially for pulping. The G units are resistant to lignin depolymerization while S lignin is more easily removed (del Río et al, 2007). The higher the S/G ratio, the less alkali is required for pulp production, reducing the cost of pulping (Ramadevi et al., 2016). Understanding the available variation and developing robust methods for measuring syringyl/guaiacyl (S/G) ratios in particular wood sources has multiple applications both in advanced tree breeding programs that require large-scale sample screening efforts, as well in the industry to allow rapid on demand assessment of the input material in pulping.

Commercial species of *Eucalyptus* make up the largest proportions of planted forests in the tropics due to their high adaptability, rapid growth and versatile wood qualities. Cellulose pulp, energy, and solid wood products are the main products derived from these forests (Myburg et al., 2007). The genus *Eucalyptus* includes more than eight hundred species, most of them Australian native, out of which ten species are the most widely planted (Harwood et al., 1997). Among these, the temperate species *E. benthamii* has received increased attention from the industry due to its' resistance to frost and colder climates and good wood properties for cellulose production (Higa & Pereira, 2003). *E. pellita*, on the other hand, widely planted in hot equatorial climates, displays drought and pest resistance (Harwood et al, 1997). Although *E. pellita* has a heavier wood, which is more lignified and rich in extractives, and therefore not particularly adequate for pulp production (Pereira et al, 2000), it has been an important alternative for hybrid breeding, with introgression of its' stress tolerance into species with better wood properties for pulping.

Near-infrared spectroscopy (NIR) is a vibrational spectroscopy (750 to 2,500nm) in which the energy radiated in the sample interacts with the bonds in the organic molecules and the absorbance measured from each sample results in a spectra profile. It is a simple and rapid technology that requires little sample preparation. NIR has been widely used to develop prediction models for chemical wood traits in breeding programs and to reduce cost and time in the pulp delignification process. Reported *Eucalyptus* S/G ratio from NIR studies present averages from 2.0 to 4.0 (Table 1) and low variability (6 to 15%), except for Hodge et al. (2018) which ranged from 16 to 38%.

High throughput, low cost, wood chemical and physical phenotyping has become a critical component of advanced *Eucalyptus* breeding programs. The development of robust NIR spectroscopy models has been the best solution to reduce the cost and time to evaluate large populations. However, studies to date have only tackled the main planted species such as *E. grandis*, *E. urophylla* and *E. globulus*, with almost no information available for additional species of the genus of high breeding and industrial interest. In this study, we developed models to predict S/G

ratio using NIR technology for two increasingly important species of *Eucalyptus* using a streamlined thioacidolysis procedure as the reference method. We also compared the performance of the models developed using a bench and a portable instrument.

## 2. Materials and Methods

### 2.1 Sample collection and preparation

The 1,220 tree samples used in this study were obtained from a pilot genomic selection-based breeding program at Embrapa (Brazilian Agricultural Research Corporation). Population samples of *E. benthamii* (56 months) and *E. pellita* (42 months) were composed of 484 and 736 individual trees, respectively. Further details on the two breeding populations, planting location and experimental field trials were described by Müller et al. (2017).

### 2.2 Spectra analysis and development of models

NIR spectra were acquired from grounded and sifted (32 mesh) sawdust using a bench (FOSS NIRSystems 5000; 1,100 to 2,500nm, 2nm intervals) and a portable (Viavi MicroNIR™ 1700; 908 to 1,676nm, ~6 nm intervals) spectrometer in the same quartz ring cup. Two replicate spectra were acquired for each of the 1,220 samples with the bench instrument and three replicates with the portable instrument.

For each *Eucalyptus* population, a set of thirty trees was selected based on neighborhood Mahalanobis distance (NH) among the sawdust spectra of all samples for each species separately using the software WinISI II (v. 1.5). Before selection, spectra were processed with detrend correction followed by standard normal variate (SNV), 2<sup>nd</sup> derivative (five points and five points polynomial smoothing). A total of sixty trees, thirty for each species, were selected for wet lab S/G ratio analysis.

In order to improve the models, some pre-processing data treatments were applied to the spectral data, for a total of 2,440 spectra for the bench instrument and 3,660 spectra for the portable instrument. The main data treatments employed comprised standard normal variate (SNV), multiple scatter correction (MSC), detrend, 1<sup>st</sup> and 2<sup>nd</sup> derivatives (Savitzky-Golay, 4<sup>th</sup> polynomial order, both 25 and 11 smoothing points), normalize and extended multiple scatter correction (EMSC) (Lupoi et al., 2014; Ramadevi et al., 2016).

Calibrations were performed using a partial least square (PLS) regression in the software The Unscrambler X® (v10.2) with and without spectra processing. Models used all replications from each sawdust of tree. Calibrations were carried out for each species separately using spectra data from twenty trees leaving ten for external validation. A joint-species model was also calibrated using data

from forty trees and validated for each species with the same ten trees not included in this calibration set. A cross-validation was carried out with parameters set as random with twenty segments, but not shown in this study. Calibration and validation models, for both instruments, were calculated using the whole spectrometer range as well as specific spectra ranges suggested as more relevant regions for S/G ratio or lignin content measurements (Lupoi et al., 2014; Ramadevi et al., 2016).

The parameters used to evaluate the calibrations and external validations were: standard deviation (SD), standard error of prediction (SEP), the coefficient of determination of validation ( $R^2$ ) based on Pearson correlation. The ratio of performance deviation (RPD) was calculated as SD/SEP. The relative standard deviation (RSD) measures the average dispersion (percentage of coefficient of variation - CV) between the reference and the predicted values for each external validation. The Spearman's correlation ( $r_s$ ) was also calculated for the validation set.

### 2.3 S/G ratio analysis

The S/G ratio was measured by thioacidolysis using the method described by Robinson & Mansfield (2009) that uses 10mg of ground, extract-free oven-dried wood flour from each individual sample as substrate and 10ml of reaction mixture [2.5% boron trifluoride etherate and 10% ethanethiol, in recently distilled dioxane (v/v)]. Gas chromatography was performed on a Hewlett Packard 5890 series II instrument. The S/G ratio was calculated using the S and G areas.

## 3. Results and Discussions

We measured S/G ratio for wood samples of individuals of two *Eucalyptus* species in an attempt to capture the existing variation for the trait toward selective breeding applications. Although the proportion of samples for which the S/G ratio was measured was limited (~5%), these samples were selected based on NIR spectra diversity. S/G ratio showed a normal distribution. Overall, the average S/G ratio of *E. benthamii* was slightly higher and significantly different from that of *E. pellita* (2.97 versus 2.78); in line with previous estimates, where values varied from 2.0 to 4.0 in other species of *Eucalyptus* (Table 1). *E. pellita* showed a slightly higher variability than *E. benthamii* (14.7 versus 13.7%), also in the same range as in other studies (Table 1) for *Eucalyptus* (6.0 to 15.0%), except for Hodge et al. (2018) who worked with populations from 4 to 25 years of old. Our results indicate that the selection of the samples for wet lab analysis was adequate, covering the available variation.

The estimates of S/G ratio for *E. benthamii* are the first to be reported for this increasingly important species for breeding and plantation forestry given its' cold resistance and high-quality wood for pulping (Higa & Pereira, 2003), despite its known genetic vulnerability. For *E. pellita*, our S/G measurements had a higher range (2.0-3.5) than previous (1.7-2.8) estimates (Ramadevi et al.,

2016).

Models with spectra data treatment for the bench equipment showed more satisfactory parameters when compared to those without data treatment (Table 2), highlighting how treatments can enhance spectral features and reduce noise or unwanted sources (Zeaiter et al., 2005). These bench models were better (Figure 2) than those built with the portable spectrometer (Table 3, Figure 3).

Our models used the whole wavelength range for both spectrometers. The selection of specific wavelengths, which have been reported to improve S/G ratio prediction for some *Eucalyptus* models (Lupoi et al., 2014; Ramadevi et al., 2016), did not improve our models. Comparing the common wavelength range (1,100-1,700nm) for both instruments, the portable instrument had models similar or worse than with the full wavelength. For the bench spectrometer, models were similar. Treatments reduced the spectra scattering on the portable spectrometer and its qualitative models showed only slightly lower accuracy (linear discriminant analysis) and  $R^2$  (93%; 0.80) than with the bench spectrometer (99%; 0.94) (Diniz et al., 2018).

The best bench models for S/G ratio,  $R^2$  and RPD (Table 2) were not far from those reported by the literature for *Eucalyptus* species (Table 1). Specific models for *E. pellita* were similar compared to Ramadevi et al. (2016) calibrating only with *E. pellita*, and better models (Table 2) calibrating with samples of both species together. Using the SECV instead of SEP to calculate the RPD, as reported in other studies (Table 1), our models showed a significantly better  $R^2$  and RPD (0.97 and 5.42) for *E. benthamii*. However, this is not recommended (Williams, 2014). Our  $R^2$  and RPDs (Table 2) were similar or superior to those reported (0.21-0.96 and 0.30-5.20, respectively) by the literature (Table 1) who worked with large samples of *Eucalyptus* species. Very good models ( $R^2=0.96$  and RPD=5.2) were obtained for *E. globulus* with a limited variability but with a different method (pyrolysis) and a bench instrument (Alves et al., 2011).

The use of  $R^2$  and RPD to evaluate NIR models is questioned because these parameters depend on the SD (Esbensen et al., 2014; Fearn, 2014). However, taking into account the average of the relative standard deviation (RSD) with the best models, we had a low RSD for the bench spectrometer, almost twice that of the portable spectrometer (Table 4). The RSD of our S/G ratio model on the bench spectrometer is in the same range as that found in *E. grandis* (0.30 to 3.22) (Govender et al., 2009) and in line with the RSD found among triplicates for many other traits in different crops (Fontaine et al., 2004). We also calculated the Spearman's rank correlation ( $r_s$ ), which is a non-linear Pearson correlation coefficient between the ranked variables. The  $r_s$  for the bench spectrometer was considerably superior to the estimate for the portable spectrometer (Table 4). The

additional use of these two parameters (RSD and  $r_s$ ) suggests that poorly classified models based on RPD could, in fact, be still satisfactory. The use of  $r_s$  to compare NIR predictions is rarely recorded (Khaled et al., 2006). Figure 2 summarizes the performance of these four parameters for the models built for the bench and portable spectrometers. For wood traits with relatively low variability and high heritability like S/G ratio, these two parameters may, therefore, be used for additional evaluation of NIR models initially deemed poor by a low RPD and  $R^2 > 0.8$ .

A potential limitation of our study would be the relatively small samples size used to obtain the S/G ratio. To address this potential caveat of our results we took advantage of the access to a much larger sample size of S/G ratio data in a previous *Eucalyptus* study (Lima, 2014). A model was calibrated with 250 diverse samples and validated with 100, providing RPD=2.33,  $R^2=0.82$ ,  $r_s=0.90$ , and RSD=2.08. Despite the much larger sample size, again we ended up with what would be considered a relatively poor RPD (Williams, 2014). However, the  $r_s$  and RSD parameters indicated that the model could be robust. Thus, instead of just taking  $R^2$  and RPD as unique parameters to evaluate a model, we suggest that the  $r_s$  and RSD should also be taken into account.

Finally, the method employed to measure the S/G ratio (Robinson and Mansfield, 2009) has the great advantage of using a small amount of sawdust per analysis, with only 10mg necessary. Such a small sample input is convenient for rapid wood quality screening purposes in a breeding program involving large numbers of samples that cannot afford to cut down trees to produce sufficient sawdust as required for classical methods of wood chemical analysis. Although this S/G ratio measurement protocol was developed using a poplar hybrid (a temperate wood with lower extractives content than tropical *Eucalyptus* wood), it proved to be successful for the two *Eucalyptus* species despite their variable extractive content. Models were slightly better for *E. pellita*, a wood with higher extractives content than *E. benthamii* (Harwood et al., 1997).

#### 4. Conclusions

We used NIR data collected with bench and portable instruments to build prediction models for S/G ratio, a key trait in the cellulose pulping industry, for two contrasting *Eucalyptus* species. The thioacidolysis method was rapid, efficient and adequate for *Eucalyptus*. The bench spectrometer performance was significantly superior. The species model was better for *E. benthamii*, while the joint model for *E. pellita*. Spearman's rank correlation ( $r_s$ ) and the relative standard deviation (RSD) may be used to attest robust models even with less satisfactory  $R^2$  and RPD. These models should prove useful for high throughput wood phenotyping in breeding programs.

## Acknowledgments

We thank the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq - project n.473936/2013-5). The Coordenação de Aperfeiçoamento de Pessoal de Nível Superior for the scholarship for CPD. We also acknowledge Embrapa Forestry Center for providing the raw wood samples. This work was also supported by PRONEX-FAPDF grant 2009/00106-8 'NEXTREE', CNPq grant 400663/2012-0 and Embrapa grant 03.11.01.007.00.00 to DG. DG received a research fellowship from CNPq.

## References

- Alves, A., Simões, R., Stackpole, D.J., Vaillancourt, R.E., Potts, B.M., Schwanninger, M., Rodrigues, J., 2011. Determination of the syringyl/guaiacyl ratio of *Eucalyptus globulus* wood lignin by near infrared-based partial least squares regression models using analytical pyrolysis as the reference method. *J. Near Infrared Spectrosc.* 19, 343–348. <https://doi.org/10.1255/jnirs.946>
- del Río, J.C., Gutiérrez, A., Rodríguez, I.M., Ibarra, D., Martínez, Á.T., 2007. Composition of non-woody plant lignins and cinnamic acids by Py-GC/MS, Py/TMAH and FT-IR. *J. Anal. Appl. Pyrolysis* 79, 39–46. <https://doi.org/10.1016/j.jaap.2006.09.003>
- Diniz, C.P., Grattapaglia, D., de Alencar Figueiredo, L.F., 2018. Comparative performance of bench and portable near infrared spectrometers for measuring wood samples of two *Eucalyptus* species (*E. pellita* and *E. benthamii*), in: International Conference on Near Infrared Spectroscopy.
- Esbensen, K., Geladi, P., Larsen, A., 2014. The RPD myth. *NIR news* 25, 24. <https://doi.org/10.1255/nirn.1462>
- Fearn, T., 2014. The overuse of R<sup>2</sup>. *NIR news* 25, 2014. <https://doi.org/10.1255/nirn.1464>
- Fontaine, J., Hörr, J., Schirmer, B., 2004. Amino acid contents in raw materials can be precisely analyzed in a global network of near-infrared spectrometers: collaborative trials prove the positive effects of instrument standardization and repeatability files. *J. Agric. Food Chem.* 52, 701–708. <https://doi.org/10.1021/jf030555o>
- Govender, M., Bush, T., Spark, A., Bose, S.K., Francis, R.C., 2009. An accurate and non-labor intensive method for the determination of syringyl to guaiacyl ratio in lignin. *Bioresour. Technol.* 100, 5834–5839. <https://doi.org/10.1016/j.biortech.2009.06.009>
- Harwood, C.E., Alloysius, D., Pomroy, P., Robson, K.W., Haines, M.W., 1997. Early growth and survival of *Eucalyptus pellita* provenances in a range of tropical environments, compared with *E. grandis*, *E. urophylla* and *Acacia mangium*. *New For.* 14, 203–219.
- Hein, P.R.G., Lima, J.T., Chaix, G., 2010. Effects of sample preparation on NIR spectroscopic estimation of chemical properties of *Eucalyptus urophylla* S.T. Blake wood. *Holzforschung* 64, 45–54. <https://doi.org/10.1515/HF.2010.011>
- Higa, R.C.V., Pereira, J.C.D., 2003. Usos potenciais do *Eucalyptus benthamii* Maiden et Cabbage. Colombo.
- Hodge, G.R., Acosta, J.J., Unda, F., Woodbridge, W.C., Mansfield, S.D., 2018. Global near infrared spectroscopy models to predict wood chemical properties of *Eucalyptus*. <https://doi.org/10.1177/0967033518770211>
- Khaled, R.A.H., Duru, M., Decruyenaere, V., Jouany, C., Cruz, P., 2006. Using leaf traits to rank native grasses according to their nutritive value. *Rangel. Ecol. Manag.* 59, 648–654. <https://doi.org/10.2111/05-031R2.1>
- Lima, B.M. De, 2014. Bridging genomics and quantitative genetics of *Eucalyptus*: genome-wide



prediction and genetic parameter estimation for growth and wood properties using high-density SNP data. ESALQ - Universidade de São Paulo.

- Lupoi, J.S., Singh, S., Davis, M., Lee, D.J., Shepherd, M., Simmons, B.A., Henry, R.J., 2014. High-throughput prediction of eucalypt lignin syringyl/guaiacyl content using multivariate analysis: a comparison between mid-infrared, near-infrared, and Raman spectroscopies for model development. *Biotechnol. Biofuels* 7, 1–14. <https://doi.org/10.1186/1754-6834-7-93>
- Milagres, F.R., 2013. Espectroscopia de infravermelho próximo para predição de propriedades da madeira de híbridos de *Eucalyptus spp.* Universidade Federal de Viçosa.
- Müller, B.S.F., Neves, L.G., de Almeida Filho, J.E., Resende, M.F.R., Muñoz, P.R., dos Santos, P.E.T., Filho, E.P., Kirst, M., Grattapaglia, D., 2017. Genomic prediction in contrast to a genome-wide association study in explaining heritable variation of complex growth traits in breeding populations of *Eucalyptus*. *BMC Genomics* 18. <https://doi.org/10.1186/s12864-017-3920-2>
- Myburg, A.A., Potts, B.B., Marques, C.M., Kirst, M., Gion, J.M., Grattapaglia, D., Grima-Pettenati, J., 2007. *Eucalyptus*, in: Kole, C. (Ed.), *Forest Trees*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 115–160. <https://doi.org/10.1007/978-3-540-34541-1>
- Pereira, J.C.D., Sturion, J.A., Higa, A.R., Higa, R.C.V., Shimizu, J.Y., 2000. Características da madeira de algumas espécies de eucalipto plantadas no Brasil. Colombo.
- Ramadevi, P., Hegde, D. V., Varghese, M., Kamalakannan, R., Ganapathy, S.P., Gurumurthy, D.S., 2016. Evaluation of lignin syringyl/guaiacyl ratio in *Eucalyptus camaldulensis* across three diverse sites based on near infrared spectroscopic calibration modelling with five *Eucalyptus* species and its impact on kraft pulp yield. *J. Near Infrared Spectrosc.* 24, 529–536. <https://doi.org/10.1255/jnirs.1251>
- Robinson, A.R., Mansfield, S.D., 2009. Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and near infrared reflectance-based prediction modeling. *Plant J.* 58, 706–714. <https://doi.org/10.1111/j.1365-313X.2009.03808.x>
- Stackpole, D.J., Vaillancourt, R.E., Alves, A., Rodrigues, J., Potts, B.M., 2011. Genetic variation in the chemical components of *Eucalyptus globulus* wood. *G3* 1, 151–9. <https://doi.org/10.1534/g3.111.000372>
- Viana, L.C., Trugilho, P.F., Hein, P.R.G., da Silva, J.M., Lima, J.T., 2010. Modelos de calibração e a espectroscopia no infravermelho próximo para predição das propriedades químicas e da densidade básica da madeira de *Eucalyptus*. *Cienc. Florest.*
- Weng, J.-K., Chapple, C., 2010. The origin and evolution of lignin biosynthesis. *New Phytol.* 187, 273–285. <https://doi.org/10.1111/j.1469-8137.2010.03327.x>
- Williams, P., 2014. Tutorial: The RPD statistic: a tutorial note. *NIR news* 25, 22. <https://doi.org/10.1255/nirn.1419>
- Zeaiter, M., Roger, J.M., Bellon-Maurel, V., 2005. Robustness of models developed by multivariate calibration. Part II: The influence of pre-processing methods. *TrAC - Trends Anal. Chem.* 24, 437–445. <https://doi.org/10.1016/j.trac.2004.11.023>

## Figure Captions

**Figure 1** - Scatter plots of predicted and reference data for the validation models of S/G ratio for *E. benthamii* (Eb) and *E. pellita* (Ep) using bench spectrometer. The top plots are species-specific models and the bottom plots are joint-species models.

**Figure 2** - Comparison among the four statistical parameters for the best models built for S/G ratio for species-specific models and joint-species models using the bench (A) and the portable (B) spectrometers. Eb: *E. benthamii* and Ep: *E. pellita*.

**Figure 3** - Scatter plots of predicted and reference data for the validation models of S/G ratio for *E. benthamii* (Eb) and *E. pellita* (Ep) using the portable spectrometer. The top plots are species-specific models and the bottom plots are joint-species models.

**Table 1** – NIR prediction models at literature for S/G ratio based on *Eucalyptus* species.

Authors	<sup>a</sup> N <sub>spp</sub>	<sup>b</sup> Age	Nc/Nv	Min	Max	Mean	SD	CV	R <sup>2</sup>	<sup>c</sup> RPD <sub>EXT</sub>	<sup>d</sup> RPD <sub>CV</sub>
Hein et al., 2010	1	14	60	1.7	3.0	2.4	0.34	14.5	0.74-0.94	-	1.90-2.88
Viana et al., 2010	2	3	120	-	-	2.8	0.27	9.6	0.21-0.90	-	0.30-2.30
Alves et al., 2011	1	16	42/36	1.8	2.2	2.0	0.12	5.9	0.96	-	3.30-5.20
Stackpole et al., 2011	1	16	135/45	1.8	2.1	2.0	0.08	4.0	0.47	-	-
Milagres, 2013	4	5	139/46	2.4	4.4	3.3	3.30	14.1	0.75-0.93	-	2.00-2.74
Lima, 2014	3	5	250/100	1.9	4.2	2.9	0.43	14.8	0.87	-	-
Lupoi et al., 2014	13	10	75	1.4	3.0	2.2	0.08	3.7	0.57-0.83	-	-
Ramadevi et al., 2016	5	3	120/36	1.5-2.1	2.4-3.6	-	-	-	0.63-0.92	1.33-2.81	2.39
Hodge et al., 2018	5	4-25	186	-	-	4.0	0.5-1.5	15.5-37.5	0.46-0.90	-	1.28-3.30

<sup>a</sup>N<sub>spp</sub>: number of *Eucalyptus* species. <sup>b</sup>Age given in years. <sup>c</sup>RPD calculated using SEP. <sup>d</sup>RPD calculated using SECV.

**Table 2** - PLS models for S/G ratio with the bench spectrometer.

	Parameters	<sup>a</sup> Eb	<sup>b</sup> Ep	<sup>c</sup> Joint/Eb	<sup>c</sup> Joint/Ep
Non-treated spectra	<sup>d</sup> Nc/Nv	20/10	20/10	37/10	40/10
	R <sup>2</sup>	0.23	0.15	0.41	0.31
	SEP	0.49	0.44	0.29	0.57
	N factors	0	1	8	7
	RPD	0.72	0.94	1.23	0.72
Treated spectra	Treatment	2D+EMSC	SNV+2D	2D+EMSC	2D+MSC
	Nc/Nv	20/10	20/10	37/10	40/10
	R <sup>2</sup>	0.86	0.68	0.72	0.82
	SEP	0.20	0.31	0.20	0.20
	N factors	4	11	6	8
	RPD	1.81	1.32	1.80	2.08

<sup>a</sup>Eb: model calibrated and validated with *E. benthamii*.

<sup>b</sup>Ep: model calibrated and validated with *E. pellita*.

<sup>c</sup>Joint: model calibrated using data set for both species and validated on each separately.

<sup>d</sup>Nc/Nv: number of trees used in the calibration and in the validation.

**Table 3** - PLS models for S/G ratio with the portable spectrometer.

	Parameters	Eb	Ep	Joint/Eb	Joint/Ep
Non-treated spectra	Nc/Nv	20/10	20/10	40/10	40/10
	R <sup>2</sup>	0.14	0.23	0.46	0.21
	SEP	0.42	0.49	0.33	0.57
	N factors	0	0	4	0
	RPD	0.85	0.82	1.06	0.71
Treated spectra	Treatment	SNV+2D	SNV+2D	Normalize+2D	SNV+2D
	Nc/Nv	20/10	20/10	40/10	40/10
	R <sup>2</sup>	0.04	0.04	0.39	0.37
	SEP	0.52	0.48	0.31	0.58
	N factors	0	0	1	0
	RPD	0.69	0.85	1.12	0.69

**Table 4** – Relative standard deviation (RSD) and Spearman's rank correlation ( $r_s$ ) in the best NIR prediction models for S/G ratio.

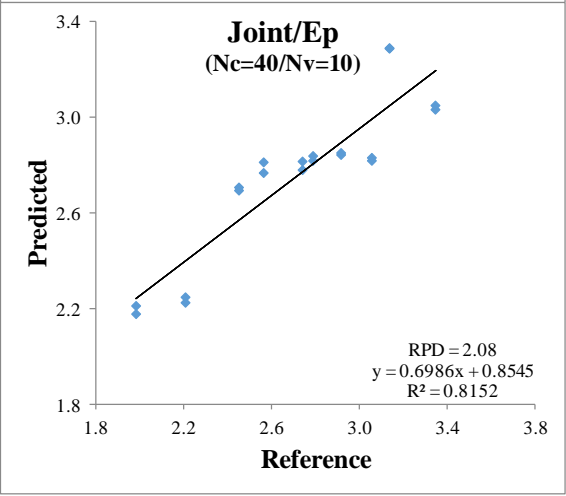
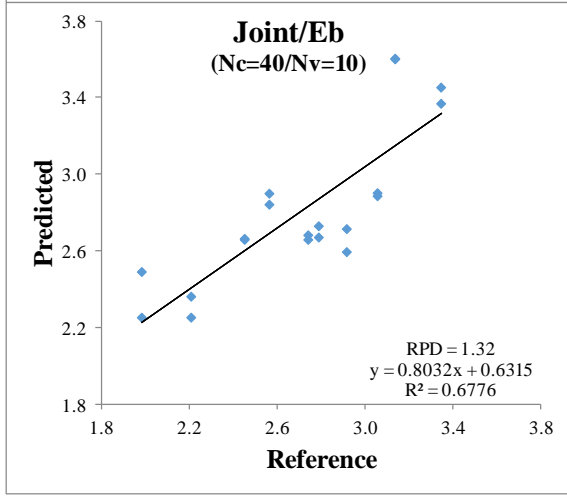
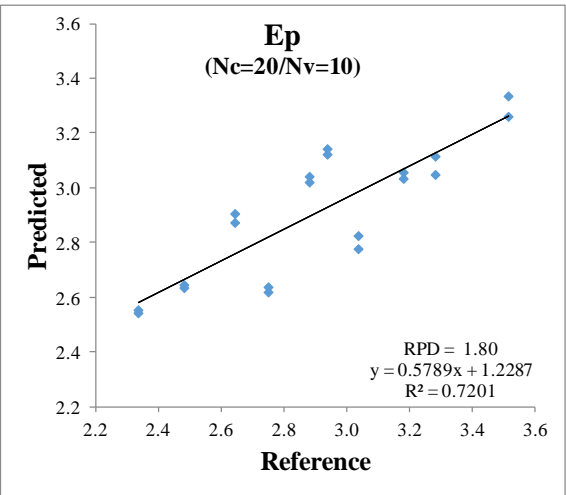
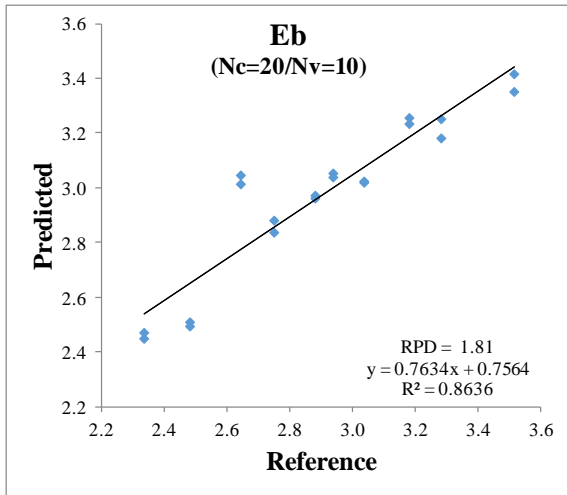
Model		<sup>a</sup> Eb		<sup>b</sup> Ep		<sup>c</sup> Joint/Eb		<sup>d</sup> Joint/Ep	
Spectrometer		Bench	Portable	Bench	Portable	Bench	Portable	Bench	Portable
RSD	Minimum	0.29	1.50	0.50	0.58	2.81	0.19	0.54	1.10
	Maximum	10.01	18.1	16.08	28.88	6.68	13.53	7.71	27.69
	Average	2.72	7.60	5.52	9.74	4.59	6.39	4.06	10.03
$r_s$		0.91	0.00	0.83	0.26	0.81	0.00	0.95	0.03

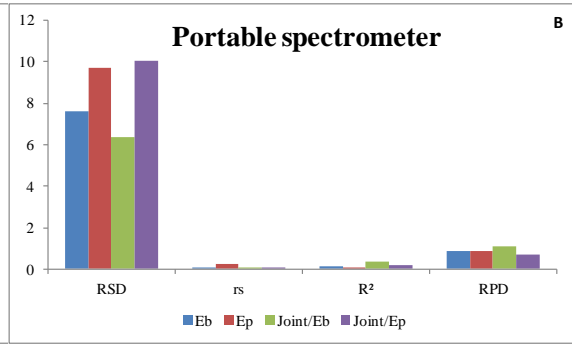
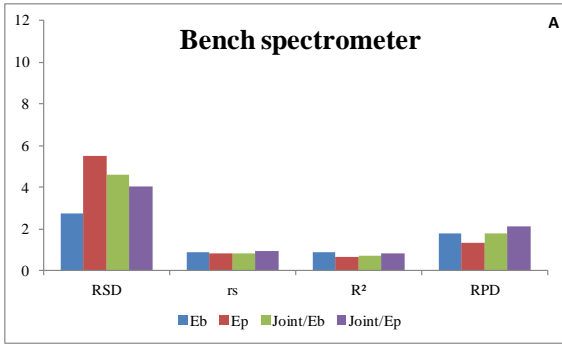
<sup>a</sup>Eb: 2D+EMSC treatment for the bench, and no data treatment for the portable spectrometer.

<sup>b</sup>Ep: SNV+2D treatment on both spectrometers.

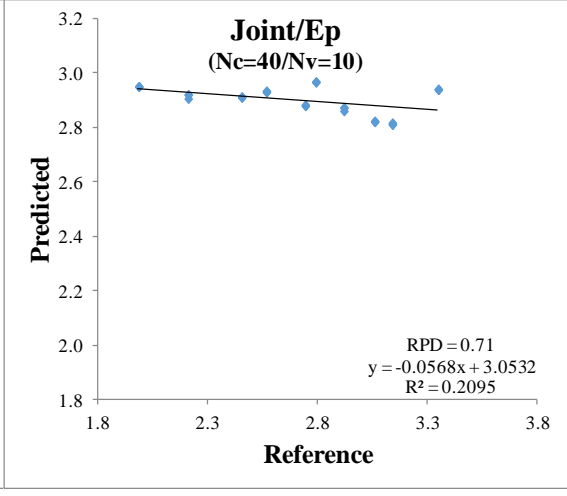
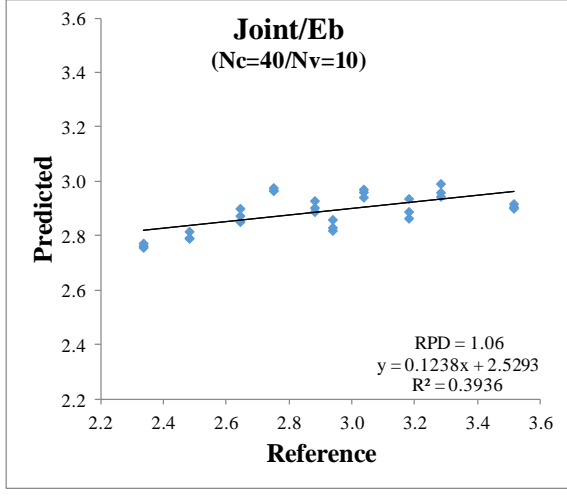
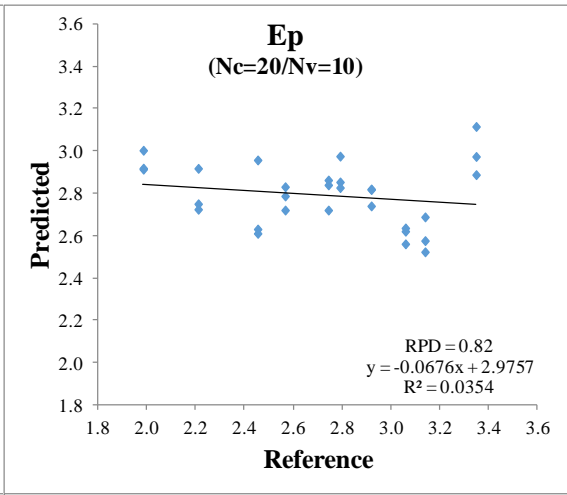
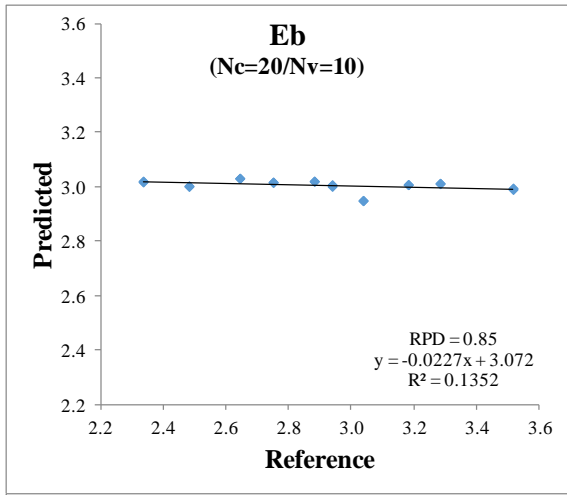
<sup>c</sup>Joint/Eb: 2D+EMSC treatment for the bench, and Normalize+2D for the portable spectrometer.

<sup>d</sup>Joint/Ep: 2D+MSC treatment for the bench and no data treatment for the portable spectrometer.









## Conclusões e Perspectivas

Pela primeira vez a proporção de lignina S/G foi determinada para *Eucalyptus benthamii*, que se destaca devido à sua tolerância ao frio. Já para *E. pellita*, espécie com maior tolerância às doenças tropicais, a determinação de S/G ocorreu pela segunda vez.

A tecnologia NIR com o equipamento de bancada se mostrou eficiente para a predição pela validação externa da proporção de lignina S/G nas duas populações de eucalipto (*E. benthamii* e *E. pellita*). Apesar dos valores baixos de RPD dos modelos (1,81 e 2,08), o alto valor de  $R^2$  (0,86 e 0,82), o baixo desvio médio relativo entre os valores de predição e referência (2,72 e 4,06), e a alta correlação de Spearman (0,91 e 0,95) reforçam a eficiência dos modelos. Já com a validação cruzada, a RPD aumentou para *E. benthamii* (3,27), mas não aumentou para *E. pellita* (1,63).

Por se encontrarem ainda armazenadas, as amostras das 1.220 árvores podem ser analisadas para outras características de interesse como conteúdo de celulose e lignina. Portanto, outros modelos preditivos poderão ser calibrados utilizando os dados espectrais já coletados, para a mesma população.

Os modelos de predição desenvolvidos nesse trabalho serão utilizados nos programas de melhoramento genético da Embrapa para as espécies em estudo, assim como poderão ser utilizados pela indústria, visando a redução dos custos e tempo no processo de deslignificação da madeira.

As mesmas amostras que foram analisadas neste trabalho estão sendo analisadas no espectrômetro FT-ATR (Bruker, 2.500 – 15.408 nm). Os dados serão analisados e comparados às regressões obtidas a partir dos dados do equipamento portátil e de bancada. Espectros de amostras de *Eucalyptus* de outros projetos (Genolyptus e International Paper Brazil) também foram adquiridos no mesmo equipamento de bancada e futuramente serão analisados, com objetivo do desenvolvimento de modelos globais para predição de diversas características como conteúdo de lignina, celulose, hemicelulose, extrativos e outros.