



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21c/2018/12.12.11.47-TDI

ESTRATIFICAÇÃO DA FLORESTA AMAZÔNICA COM BASE EM NUVENS DE PONTOS LIDAR

Mauro Lúcio Rodrigues de Assis

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada, orientada pelo Dr. Rafael Duarte Coelho dos Santos, aprovada em 20 de dezembro de 2018.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/3SCFP85>>

INPE
São José dos Campos
2019

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE
Gabinete do Diretor (GBDIR)
Serviço de Informação e Documentação (SESID)
CEP 12.227-010
São José dos Campos - SP - Brasil
Tel.:(012) 3208-6923/7348
E-mail: pubtc@inpe.br

**COMISSÃO DO CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO
DA PRODUÇÃO INTELECTUAL DO INPE (DE/DIR-544):****Presidente:**

Dr. Marley Cavalcante de Lima Moscati - Centro de Previsão de Tempo e Estudos
Climáticos (CGCPT)

Membros:

Dra. Carina Barros Mello - Coordenação de Laboratórios Associados (COCTE)

Dr. Alisson Dal Lago - Coordenação-Geral de Ciências Espaciais e Atmosféricas
(CGCEA)

Dr. Evandro Albiach Branco - Centro de Ciência do Sistema Terrestre (COCST)

Dr. Evandro Marconi Rocco - Coordenação-Geral de Engenharia e Tecnologia
Espacial (CGETE)

Dr. Hermann Johann Heinrich Kux - Coordenação-Geral de Observação da Terra
(CGOBT)

Dra. Ieda Del Arco Sanches - Conselho de Pós-Graduação - (CPG)

Silvia Castro Marcelino - Serviço de Informação e Documentação (SESID)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon

Clayton Martins Pereira - Serviço de Informação e Documentação (SESID)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Simone Angélica Del Ducca Barbedo - Serviço de Informação e Documentação
(SESID)

André Luis Dias Fernandes - Serviço de Informação e Documentação (SESID)

EDITORAÇÃO ELETRÔNICA:

Ivone Martins - Serviço de Informação e Documentação (SESID)

Murilo Luiz Silva Gino - Serviço de Informação e Documentação (SESID)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA, INOVAÇÕES E COMUNICAÇÕES
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21c/2018/12.12.11.47-TDI

ESTRATIFICAÇÃO DA FLORESTA AMAZÔNICA COM BASE EM NUVENS DE PONTOS LIDAR

Mauro Lúcio Rodrigues de Assis

Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada, orientada pelo Dr. Rafael Duarte Coelho dos Santos, aprovada em 20 de dezembro de 2018.

URL do documento original:

<<http://urlib.net/8JMKD3MGP3W34R/3SCFP85>>

INPE
São José dos Campos
2019

Dados Internacionais de Catalogação na Publicação (CIP)

Assis, Mauro Lúcio Rodrigues de.

A76e Estratificação da Floresta Amazônica com base em nuvens de pontos LiDAR / Mauro Lúcio Rodrigues de Assis. – São José dos Campos : INPE, 2019.

xix + 51 p. ; (sid.inpe.br/mtc-m21c/2018/12.12.11.47-TDI)

Dissertação (Mestrado em Computação Aplicada) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2018.

Orientador : Dr. Rafael Duarte Coelho dos Santos.

1. LiDAR. 2. Amazônia. 3. Estratificação. 4. Fuzzy C-Means. 5. Random Forest. I.Título.

CDU 004.81:164(811.3)



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

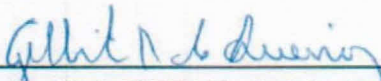
This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).

Aluno (a): **Mauro Lúcio Rodrigues de Assis**

Título: "ESTRATIFICAÇÃO DA FLORESTA AMAZÔNICA COM BASE EM NUVENS DE PONTOS LIDAR"

Aprovado (a) pela Banca Examinadora
em cumprimento ao requisito exigido para
obtenção do Título de **Mestre** em
Computação Aplicada

Dr. Gilberto Ribeiro de Queiroz



Presidente / INPE / São José dos Campos - SP

() Participação por Vídeo - Conferência

Aprovado () Reprovado

Dr. Rafael Duarte Coelho dos Santos



Orientador(a) / INPE / SJCampos - SP

() Participação por Vídeo - Conferência

Aprovado () Reprovado

Dra. Francisca Rocha de Souza



Membro da Banca / INPE / São José dos Campos - SP

() Participação por Vídeo - Conferência

Aprovado () Reprovado

Dr. Fábio Fagundes Silveira



Convidado(a) / UNIFESP / São José dos Campos - SP

() Participação por Vídeo - Conferência

Aprovado () Reprovado

Este trabalho foi aprovado por:

() maioria simples

() unanimidade

“Educação não é preparação para a vida; educação é a vida em si”.

JOHN DEWEY

*À minha mãe, **Dalva**, pelo Norte.
À minha mulher, **Bete**, por quase tudo o mais.
Ao meu filho, **Pedro**, pelo que faltava.*

AGRADECIMENTOS

Essa dissertação não existiria se não tivesse a colaboração de muitas pessoas, pessoal e profissionalmente. Várias delas conheci durante os anos em que estive envolvido no mestrado, gente brilhante e apaixonada pelo que faz.

Ao meu orientador, Rafael Duarte Coelho dos Santos, agradeço pelas horas de conversa sempre inspiradoras e pela paciência durante esse período.

Ao brilhante professor e amigo Fernando Espírito-Santo, que, dentre outras coisas, me acolheu não só na Universidade de Lancaster no período em que comecei os estudos para essa dissertação, mas também em sua própria casa.

Ao Dr. Jean Ometto, que me recebeu no CCST/INPE durante a duração do projeto EBA, o que permitiu que eu tivesse acesso aos dados usados neste estudo.

À amiga Roberta Cantinho, que entre outras coisas me abriu a oportunidade de trabalhar no projeto EBA, ajudou-me com a estada em Lancaster e com a publicação do artigo do mestrado.

Ao Prof. Dr. Eric Görgens, outro bom amigo, que me orientou com a sua generosidade costumeira em vários momentos desse trabalho.

À amiga Dra. Francisca Rocha de Souza Pereira, pela generosidade em estar sempre disponível para esclarecer dúvidas durante a implementação dessa dissertação.

Aos colegas de turma nas classes do mestrado, pela acolhida e ajuda com as disciplinas.

Aos membros da banca, Dr. Gilberto Ribeiro de Queiroz, Dr. Fabio Fagundes Silveira e Dra. Francisca Rocha de Souza Pereira, por aceitarem fazer parte dessa banca e pelo tempo dedicado à análise desse trabalho.

À CAPES (Coordenação Nacional de Ensino Superior) pela bolsa concedida para esse mestrado.

RESUMO

Sensores LiDAR (Light Detection and Ranging) produzem dados de alta resolução coletados ao longo do perfil vertical das florestas. Na Amazônia brasileira a biomassa varia entre 150 e 400 ton.ha⁻¹, mas pouco se sabe sobre estrutura vertical da floresta. Muitos dos processos de pesquisa e análise de dados são baseados na ideia de estratificação (agrupamento de regiões, populações ou outros dados com características semelhantes). Num bioma gigantesco e pouco explorado como o amazônico, uma das dificuldades é justamente encontrar dados abrangentes para estratificar a floresta em termos de sua estrutura vertical. O presente trabalho busca propor uma forma de estratificação da Amazônia baseada unicamente em dados LiDAR coletados por todo o bioma. Para isso foram identificadas métricas LiDAR que melhor se correlacionam com a estrutura vertical da floresta, aproveitando-se o potencial desse sensor em descrevê-la e a correlação importante dessa estrutura com outras características da floresta. Ao testar-se essa hipótese com o algoritmo de agrupamento Fuzzy C-Means com diferentes parâmetros não se encontrou indicação de agrupamento significativa. Da mesma forma, visualmente analisando-se a distribuição de valores não é possível indicar nenhum tipo de agrupamento. A partir daí, buscou-se encontrar uma indicação de outro caminho para se chegar a essa estratificação. Foi analisada a métrica skewness, obtida a partir das métricas LiDAR, que se mostrou promissora para que esse objetivo seja atingido.

Palavras-chave: LiDAR. Amazônia. Estratificação. Fuzzy C-Means. Random Forest.

AMAZON FOREST STRATIFICATION BASED ON LIDAR POINT CLOUDS

ABSTRACT

LiDAR (Light Detection and Ranging) sensors produce high-resolution data collected along the vertical structure of the forests. In the Brazilian Amazon biomass would be between 150 and 400 $\text{ton}\cdot\text{ha}^{-1}$, but little is known about the vertical structure of the forest. Many of the research and data analysis processes are based on the idea of stratification (grouping of regions, populations or other data with characteristics similar). In a gigantic biome and little explored as the Amazon, one of the difficulties is precisely to find comprehensive data to stratify the forest emtermos of its vertical structure. The present work seeks to propose a form of Amazon stratification based only on LiDAR data collected throughout the biome. For this, LiDAR metrics were identified that better correlate with the vertical structure of the forest, taking advantage of the potential of this sensor in describing it and the important correlation of this structure with other characteristics of the forest. When testing this hypothesis with the Fuzzy C-Means clustering algorithm with different parameters, no indication of significant clustering was found. Likewise, visually analyzing the distribution of values it's not possible to indicate any type of grouping. After that, an attempt was made to find an indication of another way to reach this stratification. The skewness metric, obtained from the LiDAR metrics, was shown to be promising for this goal to be achieved.

Keywords: LiDAR. Amazon. Stratification. Fuzzy C-Means. Random Forest.

LISTA DE FIGURAS

	<u>Pág.</u>
2.1 Bioma amazônico brasileiro.	4
2.2 Mapa de Vegetação da Amazônia	5
2.3 Estratificações do Bioma Amazônia	6
2.4 Imagem criada a partir dos dados coletados em um voo LiDAR. As cores são artificialmente geradas a partir da altitude do ponto	8
2.5 Região ampliada da imagem a partir do mesmo voo visto de cima (esq) e em perspectiva (dir)	8
2.6 Trecho de Modelo Digital do Terreno	9
2.7 Trecho de Modelo Digital de Superfície	9
2.8 Distribuição Weibull, variação parâmetro β	11
2.9 Distribuição Weibull, variação parâmetro η	12
2.10 Estimativas de biomassa da Amazônia conforme (a) (SAATCHI et. al., 2007); (b) (SAATCHI et. al., 2011); (c) (NOGUEIRA et. al., 2008); (d) (BACCINI et. al., 2012); (e) (BRASIL, MINISTÉRIO DA CIÊNCIA E TECNOLOGIA, 2010)	14
2.11 Transectos voados no projeto EBA e que foram usados nesse projeto . . .	15
3.1 Avião utilizado para coleta dos dados, onde se vê o sensor LiDAR instalado	17
3.2 Distribuição normal com diferentes valores de skewness	23
3.3 Valores shape e scale plotados para os transectos analisados.	24
3.4 Coeficientes em função do parâmetro C com m=2.	26
3.5 Coeficientes em função do parâmetro C com m=15.	27
3.6 Máscara do MODIS aplicada às métricas LiDAR.	28
4.1 Perfis de transectos (1).	33
4.2 Perfis de transectos (2).	34
4.3 Perfis de transectos (3).	35
4.4 Biomassa em função do skewness.	36
4.5 Mapa do skewness estimado para a área do Bioma Amazônia.	37
4.6 Histograma das células por classe de skewness.	39
4.7 Distribuição das classes de skewness no Bioma Amazônico.	40
4.8 Mapa de Biomassa da Amazônia (AVITABILE et al., 2016)	41

LISTA DE TABELAS

	<u>Pág.</u>
3.1 Parâmetros de voos LiDAR do projeto EBA	18
3.2 Parâmetros a serem validados nas nuvens de pontos.	20
3.3 Métricas LiDAR.	22
4.1 Importância relativa entre as variáveis usadas como entrada no modelo de estimativa do skewnes.	38
4.2 Classe de skewness usadas para elaboração do mapa de classes	39

SUMÁRIO

	<u>Pág.</u>
1 INTRODUÇÃO	1
1.1 Contexto e Motivação	1
1.2 Objetivo	2
2 REVISÃO BIBLIOGRÁFICA	3
2.1 O Bioma Amazônia	3
2.2 Tecnologia Airborne Laser Scanning (ALS)	7
2.3 LiDAR e a Estrutura Vertical da Floresta	10
2.4 Distribuição de probabilidade Weibull	10
2.5 Random Forest	12
2.6 O Projeto EBA	12
3 MATERIAL E METODOS	17
3.1 Dados utilizados	17
3.2 Formato dos dados LiDAR utilizados	18
3.3 Recebimento e controle de qualidade das nuvens de ponto LiDAR	19
3.4 Preparação dos dados para obtenção das métricas LiDAR	20
3.5 Geração das métricas LiDAR	21
3.6 Aplicação da Distribuição Weibull em relação à altura pontos LiDAR	23
3.7 Análise da métrica skewness para estratificação	28
4 RESULTADOS E DISCUSSÕES	31
4.1 Processos de controle de qualidade	31
4.2 Análise dos parâmetros da distribuição	32
5 CONCLUSÃO	43
REFERÊNCIAS BIBLIOGRÁFICAS	45

1 INTRODUÇÃO

1.1 Contexto e Motivação

A Amazônia responde por 14% da produção líquida de umidade no planeta e é o maior componente (66%) da variação anual dessa umidade (WAGNER et al., 2017).

Um outro dado que ilustra a importância do estudo desse bioma é que ele armazena cerca de 86 Pg de biomassa acima do solo (SAATCHI et al., 2007). O crescimento das árvores e da área florestada implica em retirada (fixação) de carbono da atmosfera, ao mesmo tempo em que a morte das árvores, seja de forma natural ou pela ação do homem, transfere carbono para a atmosfera terrestre (WILLIAMS et al., 2017). É sabido que o aumento da quantidade de CO₂ na atmosfera é um dos fenômenos causadores do aquecimento global, que é a elevação da temperatura média da atmosfera terrestre observado nas últimas décadas, com possíveis efeitos bastante negativos sobre a dinâmica da vida na Terra (DONOHOE et al., 2014).

Segundo o IBGE, aproximadamente 60% do Bioma Amazônia pertence ao território brasileiro, que compreende a Amazônia Legal¹. Os 40% restantes estão distribuídos entre os países da Bolívia, Colômbia, Equador, Guiana, Guiana Francesa, Peru, Suriname e Venezuela.

Uma das técnicas usadas quando nos propomos a estudar um fenômeno de grande alcance e complexidade é a estratificação, que é a divisão dos dados a respeito do fenômeno em camadas ou estratos semelhantes. Isso permite que, identificadas características de comportamento para amostras de um determinado estrato, essas características possam ser inferidas como sendo representativa de todos os indivíduos membros do estrato.

Em Ciências da Terra, a estratificação geralmente se dá separando-se a população de estudo por uma ou mais características dessa população. No caso de florestas, a estrutura vertical (forma com que a biomassa se distribui em função da altura em relação ao solo) se apresenta como uma boa forma de identificação de diferentes tipos de florestas e, conseqüentemente, de estratificação de maciços florestais.

O LiDAR (Light Detection And Ranging) aerotransportado (ALS) é um dos sensores que fornecem dados sobre a estrutura vertical da floresta, fornecendo informações

¹Em 1953, a Constituição Federal criou o conceito político de “Amazônia Legal”, que representa 61% do território brasileiro. Além dos sete estados da região Norte, inclui o estado do Mato Grosso e cerca de 79% do estado do Maranhão

precisas sobre o terreno abaixo da floresta, da rugosidade da copa das árvores bem como de outros estratos que compõem a estrutura da floresta (liana, madeira morta em pé e no solo, galhos, distribuição foliar etc.) (LEFSKY, 1997), (LOVELL et al., 2003).

1.2 Objetivo

O objetivo deste trabalho é investigar a possibilidade de uso de métricas obtidas a partir de nuvens de pontos LiDAR para se estratificar o Bioma Amazônia, de modo a dar suporte a futuros estudos que fizerem estimativas quantitativas relativas a todo esse bioma ou mesmo a parte dele tenham uma possibilidade de agrupamento que não tenha relação com dados de campo.

Mais especificamente a seguinte hipótese foi testada nesse trabalho: considerando que o sensor LiDAR coleta pulsos lasers de alta resolução dentro do dossel florestal (ao longo das alturas das árvores), o perfilhamento da estrutura vertical das florestas (distribuição de galhos, folhas e aberturas do dossel) pela distribuição de probabilidade Weibull ou outros índices derivados dos dados LiDAR permitirá estratificar a estrutura vertical das florestas e a agrupar os dados LiDAR com distribuições similares.

Para se testar essa hipótese foram utilizados dados LiDAR de 826 sobrevoos da Amazônia que cobriram cada um deles uma área de 400 m de largura por 12 km de comprimento, 3600 ha de área.

2 REVISÃO BIBLIOGRÁFICA

2.1 O Bioma Amazônia

Floresta é uma área medindo mais de 0,5 ha com árvores maiores que 5 m de altura e cobertura de copa superior a 10%, ou árvores capazes de alcançar estes parâmetros in situ. Não inclui terra que está predominantemente sob uso agrícola ou urbano (FAO, 2015). Floresta tropical, também conhecida como floresta tropical pluvial ou floresta úmida é um tipo de domínio fitogeográfico. caracterizado por receber uma grande quantidade de chuvas e apresentar elevados índices de biodiversidade e de evapotranspiração.

As florestas tropicais localizam-se em áreas equatoriais e em zonas situadas próximas aos trópicos. Sua ocorrência se dá nos continentes americano, africano e asiático. No Brasil, compreende a Floresta Amazônica, a Mata Atlântica e alguns grupos vegetacionais do sul do país, como a Mata de Araucária.

Bioma pode ser definido como “um agrupamento de fisionomia homogênea e independente da composição florística. Estende-se por uma área bastante grande e sua existência é controlada pelo macroclima. Na comunidade terrestre os biomas correspondem às principais formações vegetais naturais.”(COUTINHO, 2006).

Sessenta por cento do Bioma Amazônico está localizado no Brasil 2.1, se estendendo ainda por outros nove países da América do Sul. Quatro estados ou departamentos dessas nações chamam-se Amazonas devido à parcela importante de seus territórios ocupada pelo bioma.

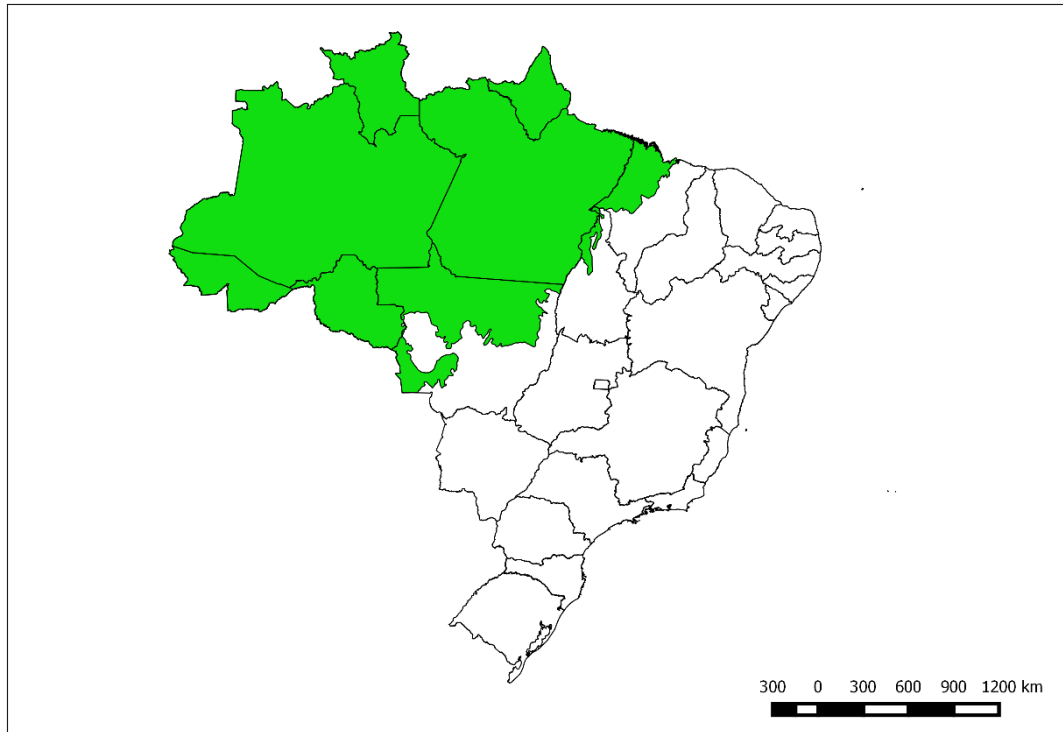
No Brasil o Bioma Amazônia ocupa cerca de 40% do território nacional (Figura 2.1).

Apesar de ser um maciço florestal contínuo, o Bioma Amazônico, por sua grande extensão, é composto por estratos, cada um com características que os diferenciam uns dos outros. Podemos por exemplo caracterizar diferentes tipos de florestas com relação à lâmina d'água que está presente acima do solo ao longo do ano. Nessa linha temos basicamente quatro tipos:

Mata de igapó: Composição vegetativa que ocorre em áreas de baixo relevo próxima a rios, permanecendo por essa razão, alagada.

Mata de várzea: Se estabelece em áreas mais elevadas em relação às matas de igapó, ficando inundadas somente no período das cheias dos rios.

Figura 2.1 - Bioma amazônico brasileiro.



Mata de terra firme: Ocorre em áreas fora do alcance das cheias. Geralmente são as florestas mais altas do bioma amazônico.

Floresta semi-úmida: Formada nas áreas de transição do bioma amazônico para outros biomas (FREITAS, 2018).

Outra estratificação que influencia a floresta são os diferentes tipos de vegetação que foram observados (Figura 2.2).

Ainda outras estratificações podem ser aplicadas ao Bioma Amazônia, como se pode ver na Figura 2.3 (SAATCHI et al., 2007).

Figura 2.2 - Mapa de Vegetação da Amazônia

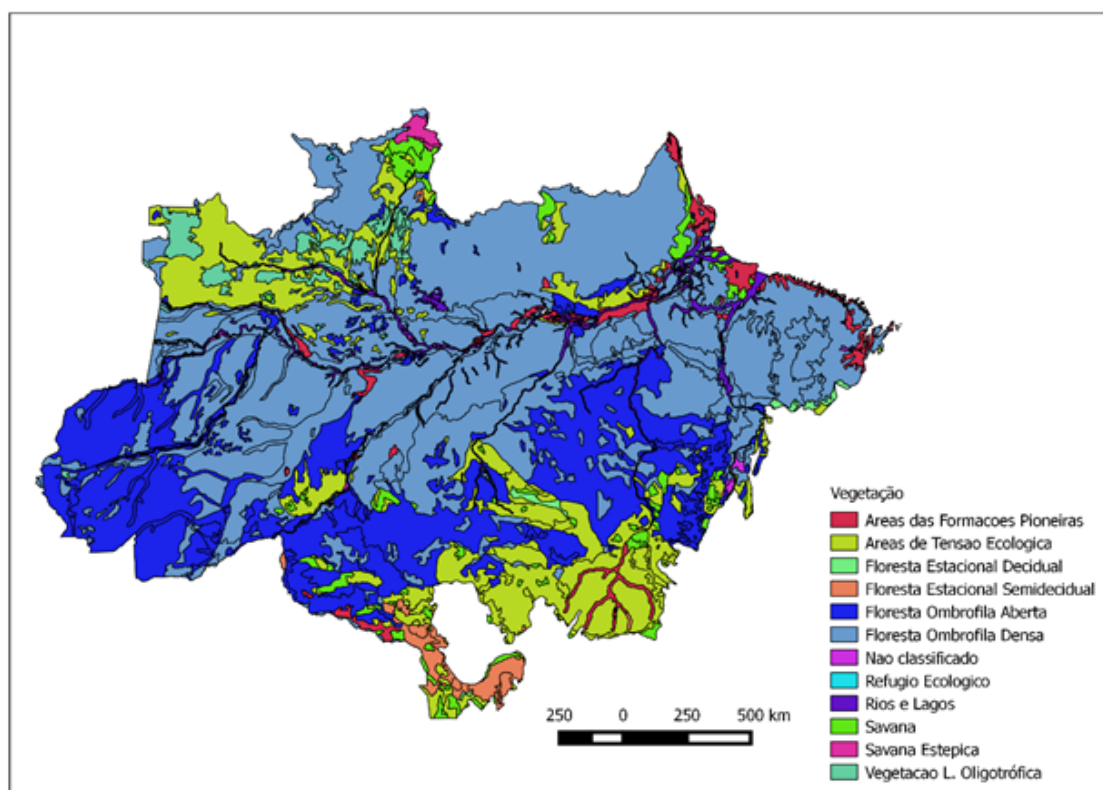
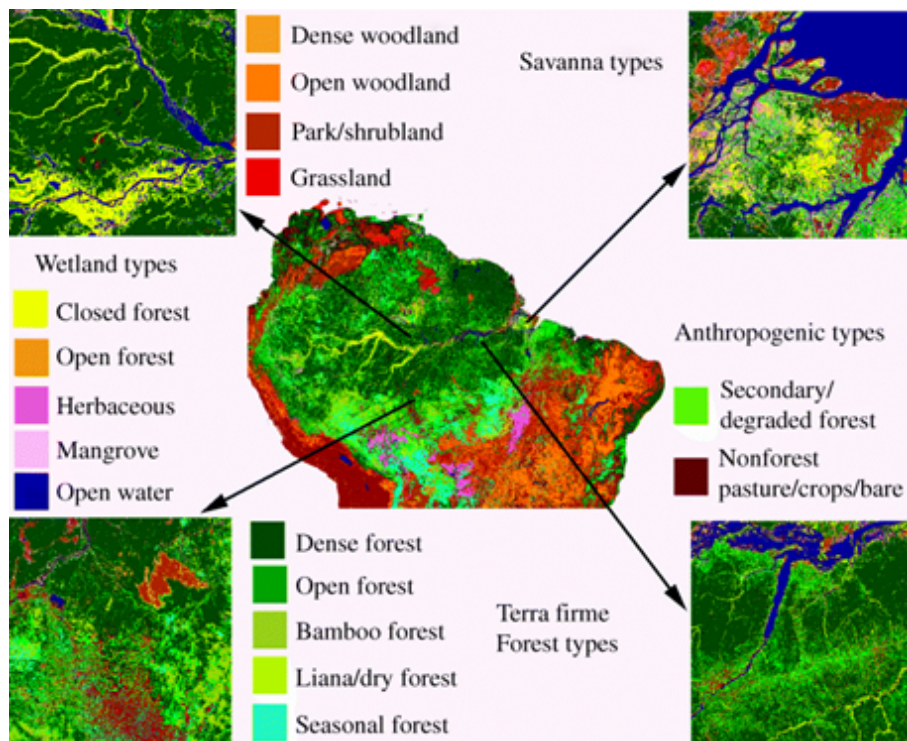


Figura 2.3 - Estratificações do Bioma Amazônia



2.2 Tecnologia Airborne Laser Scanning (ALS)

O LiDAR é um sensor ativo que emite pulsos laser em direção a um obstáculo a ser examinado e coleta a energia refletida pelo obstáculo, de maneira a medir a distância entre o sensor e esse obstáculo, normalmente através de um espelho que recebe o laser retroespalhado.

O sistema então mede o tempo entre a emissão do pulso e a captação do retorno, calculando então a distância entre a plataforma com o sensor e o objeto (Equação 2.1)

$$\Delta R = c\Delta t/2 \quad (2.1)$$

Onde:

$\Delta R(\text{m})$ é a distância entre o sensor e o objeto que refletiu o pulso.

$t(\text{s})$ é o intervalo de tempo entre a emissão do pulso (um ponto específico dessa emissão, por exemplo, a elevação da tensão de emissão a partir de um determinado limiar).

$c(\text{m.s}^{-1})$ é a velocidade da luz.

O comprimento de onda usado em sensores LiDAR varia do visível ao infravermelho próximo, dependendo do equipamento e do objeto a ser estudado.

O sinal de retorno varia em intensidade dependendo da energia inicial do pulso, tamanho do objeto interceptado e características de reflectância do alvo. A densidade de pontos depende da altitude do sensor, velocidade da plataforma, campo de visada do sistema e frequência de amostragem (AXELSSON, 1999).

Existem sensores LiDAR de retorno discreto e sensores de onda contínua (waveform). Nos sensores de onda contínua (full waveform), a energia retornada é coletada a intervalos de tempo fixo, medindo a diferença de fase entre o sinal emitido e o sinal recebido após este ter sido retroespalhado pela superfície (LIM et al., 2003). Apesar das nuvens de pontos objeto deste trabalho foram coletadas no formato full waveform, os dados analisados foram os retornos discretos.

Os dados usados neste estudo foram captados por um ALS (Airborne Laser Scanner), em que o sensor LiDAR é instalado em uma aeronave que, sobrevoando a área

de estudos, direciona os seus pulsos para o solo sobrevoado, coletando o retroespalhamento advindo dos obstáculos ao laser através de um espelho de varredura.

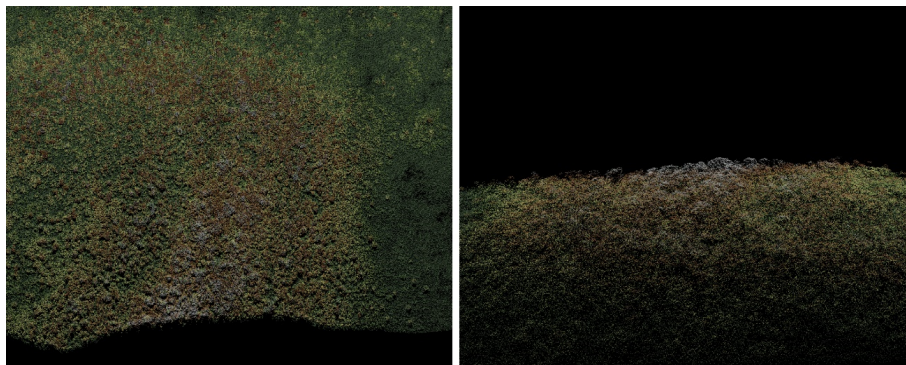
Usando-se uma Unidade de Medição Inercial (IMU) consegue-se, a partir do registro preciso da posição no espaço em que o sensor está, medir-se a distância entre sensor e o objeto responsável pelo retroespalhamento do pulso laser.

O produto gerado pelo sensor permite obter uma visão em perspectiva ou mesmo em 3D do objeto de observação, com o equipamento adequado (Figuras 2.4 e 2.5).

Figura 2.4 - Imagem criada a partir dos dados coletados em um voo LiDAR. As cores são artificialmente geradas a partir da altitude do ponto



Figura 2.5 - Região ampliada da imagem a partir do mesmo voo visto de cima (esq) e em perspectiva (dir)



Um dos usos que se pode fazer dos dados LiDAR coletados é produzir-se modelos digitais de superfície, tanto de solo quanto de copa das árvores.

Isso se faz classificando-se os pontos de solo e não-solo por software e em seguida traçando-se uma superfície que una os pontos próximos gerando uma figura que

representará, no caso de uma floresta, o relevo (DTM – Modelo Digital do Terreno, (Figura 2.6) ou a superfície formada pela copa das árvores DSM – Modelo Digital de Superfície (KRAUS; PFEIFER, 2001) (Figura 2.7).

Figura 2.6 - Trecho de Modelo Digital do Terreno

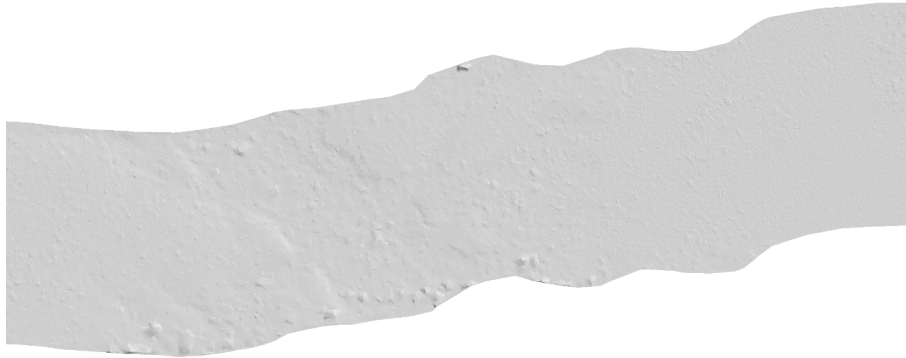


Figura 2.7 - Trecho de Modelo Digital de Superfície



Esse tipo de imagem é usado para, por exemplo, planejar-se exploração comercial de maciços de florestas plantadas, identificação e estimativas de cursos d'água, dentre outros usos.

Outro tipo de informação a partir das nuvens de pontos LiDAR é obtida dividindo-se

a nuvem segundo uma grade de quadrículas de tamanho fixo e computando-se métricas estatísticas para cada uma dessas quadrículas. Posteriormente ajusta-se modelos matemáticos a esses dados correlacionando-os algum outro tipo de informação que se tenha de solo ou usa-se as métricas para encontrar florestas semelhantes, processo que foi usado nesse estudo e será descrito em detalhes adiante.

2.3 LiDAR e a Estrutura Vertical da Floresta

Uma das formas de expressão das diferenças entre povoamentos florestais é a sua estrutura vertical, que descreve como os componentes acima do solo da floresta se distribuem ao longo de sua altura.

Esta diferenciação é determinada sobretudo pela resposta ao decréscimo da disponibilidade de luz ao longo do perfil vertical da floresta (SOUZA et al., 2004).

A penetração de luz na floresta é uma função da disposição e da altura dos indivíduos ao longo do perfil vertical (KOHYAMA, 1989). Tal disposição é a causa maior da formação de grupos de plantas com hábitos similares, que ocupam mesmos nichos ecológicos e desempenham funções semelhantes.

Por ser um sensor que coleta leituras distribuídas ao longo da estrutura vertical da floresta, já que os pulsos emitidos pelo sensor ultrapassam a superfície refletiva desta, o LiDAR se apresenta como opção para descrever essa estrutura. Com o barateamento das campanhas de coleta de dados LiDAR, é de se esperar que cada vez mais ele seja usado para o monitoramento florestal. Apesar disso, muito pouco se sabe sobre a assinatura vertical dos diferentes tipos de florestas (ASNER et al., 2012).

Existe uma forte correlação entre a estrutura vertical da floresta e características como biomassa acima do solo, diâmetro das copas, diâmetros do fuste (LIM et al., 2003), (ZIMBLE et al., 2003). Essas informações, por sua vez, são essenciais para a caracterização de maciços florestais, podendo ser usados tanto para planejamento de ações de preservação, acompanhamento de crescimento/regeneração ou mesmo planejamento de exploração econômica desses maciços.

2.4 Distribuição de probabilidade Weibull

Distribuição de probabilidade é um modelo matemático que relaciona um certo valor de uma variável com a sua probabilidade de ocorrência.

Existem vários tipos de funções de distribuição de probabilidade. Uma dessas funções é a distribuição Weibull, Essa distribuição leva o nome de quem a descreveu pela primeira vez, Waloddi Weibull em 1937 e a publicou pela primeira vez em 1951 (WEIHULL, 1951). Essa distribuição é usada para descrever vários fenômenos, desde tempos de falhas de equipamentos até fenômenos biológicos.

A distribuição de probabilidade Weibull tem se mostrado importante na descrição do perfil vertical de povoamentos florestais (COOPS et al., 2007); e (DEAN et al., 2009) e existe uma correlação positiva entre os parâmetros da Distribuição Weibull e parâmetros estruturais da floresta (COOPS et al., 2009) , o que a faz candidata a ser usada nessa investigação.

Pode-se descrever a distribuição Weibull com dois parâmetros:

- a) **Shape** (β): também chamado slope (inclinação), por ser a inclinação da linha em uma distribuição de probabilidade. Diferentes valores desse parâmetro tem uma grande influência na curva da distribuição (Figura 2.8).
- b) **Scale** (η): uma alteração nesse parâmetro tem o mesmo efeito de uma mudança na escala do eixo das abcissas, como podemos ver na Figura 2.9.

Figura 2.8 - Distribuição Weibull, variação parâmetro β

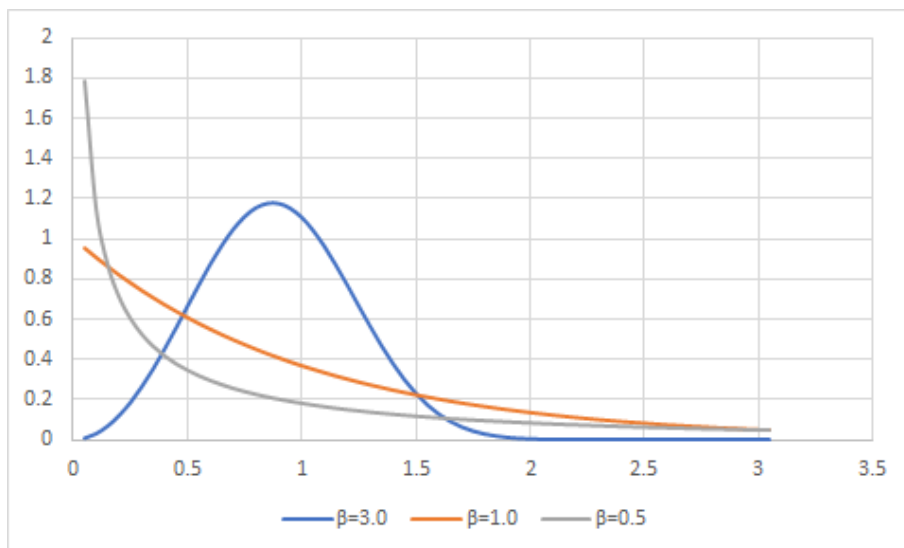
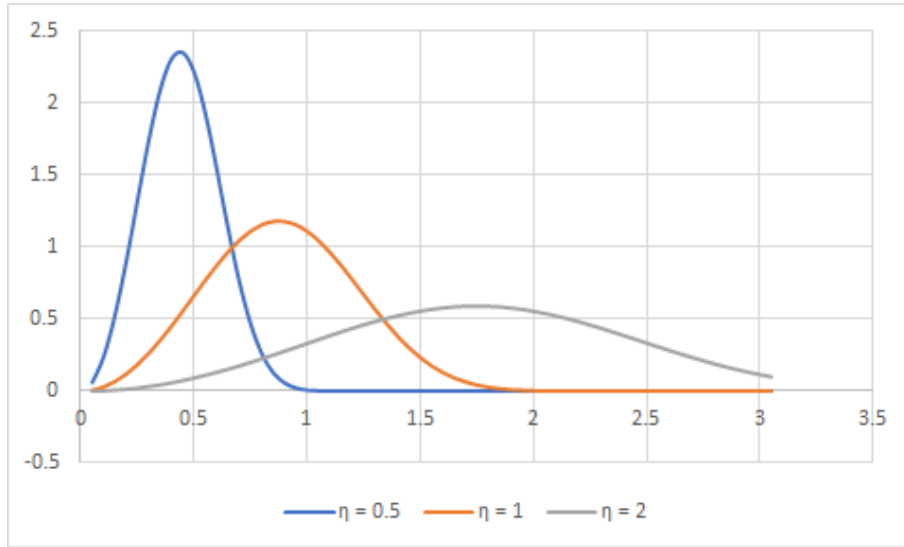


Figura 2.9 - Distribuição Weibull, variação parâmetro η



2.5 Random Forest

Classificação por combinação é o nome de uma abordagem de classificação que usa um conjunto de classificadores que trabalham juntos para identificar classes para instâncias não classificadas (FAWAGREH et al., 2014).

Random Forest é um algoritmo que é parte dessa abordagem. Nele são geradas combinações de árvores onde cada árvore cresce de acordo com um parâmetro aleatório. A predição final é obtida agregando-se o conjunto. O nome do algoritmo vem das duas ideias que o constituem, muitas árvores de decisão construídas a partir de números aleatórios ??, ??. Uma vantagem desse algoritmo é que ele pode ser usado tanto para classificação quanto para regressão, o que cobre grande parte das demandas em que se utiliza aprendizagem de máquina hoje.

2.6 O Projeto EBA

O Brasil, como participante de programas de redução de emissão de carbono como o REDD+ (Redução de Emissões de Desflorestamento e Degradação em Países em Desenvolvimento), tem que estabelecer uma referência para estimar os estoques de carbono na natureza, para a partir daí estabelecer objetivos de redução das emissões desse elemento. Essa estimativa também é importante para estabelecer as bases (preços) para o mercado de carbono, estabelecendo os fatores que determinam as

perdas e ganhos de carbono.

Quando uma floresta é removida, tem-se uma perda de carbono. Se a terra é dedicada a um novo uso, por exemplo, agricultura, tem-se um ganho de carbono, ainda que provavelmente menor que a perda pelo desflorestamento. Por outro lado, se a área é deixada para regeneração, ela atuará como um “sink” (tanque) de carbono, que vai se enchendo à medida que a floresta se desenvolve.

Existe uma relação direta entre as quantidades de biomassa e de carbono existentes na floresta. Pode-se então estimar-se o estoque de carbono de uma floresta estimando-se a quantidade da biomassa. Para estimar-se a biomassa de uma região florestal extensa são criados modelos matemáticos que estimam a biomassa com base em dados de campo e que depois são extrapolados usando-se dados coletados por satélites de observação da Terra.

Existem mapas de biomassa da Amazônia que foram criados usando-se esse método (BACCINI et al., 2012), (MITCHARD et al., 2014), (NOGUEIRA et al., 2008), (NOGUEIRA et al., 2015), (SAATCHI et al., 2007) e (SAATCHI et al., 2011).

Comparando-se as estimativas de biomassa de alguns desses mapas chegou-se à conclusão de que há uma diferença significativa entre eles (OMETTO et al., 2015) (Figura 2.10).

Assim sendo, tornou-se importante que se buscasse elaborar um novo mapa com um grau de incerteza menor. Foi criado então o projeto Environmental Monitoring via Satellite in The Amazon Biome- MSA/Amazon Fund -Subproject 7- **Estimating Biomass in the Amazon (EBA)**, coordenado pelo pesquisador Jean Ometto (CCST/INPE).

Nesse projeto foram coletados dados LiDAR de 1000 transectos de 400m de largura por 13000m de comprimento, transectos esses distribuídos por todo o bioma Amazônia. Noventa por cento desses transectos foram distribuídos aleatoriamente e dez por cento foram dispostos sobre áreas as quais se conhecem dados de campo de biomassa ou se tem modelos matemáticos já usados para estimar a biomassa da Amazônia. Os dados LiDAR serão responsáveis por diminuir a incerteza das estimativas de biomassa em relação aos mapas anteriormente elaborados.

Neste estudo foram usadas 836 nuvens de pontos (transectos) dentre as coletadas pelo projeto EBA 2.11.

Figura 2.10 - Estimativas de biomassa da Amazônia conforme (a) (SAATCHI et. al., 2007); (b) (SAATCHI et. al., 2011); (c) (NOGUEIRA et. al., 2008); (d) (BACCINI et. al., 2012); (e) (BRASIL, MINISTÉRIO DA CIÊNCIA E TECNOLOGIA, 2010)

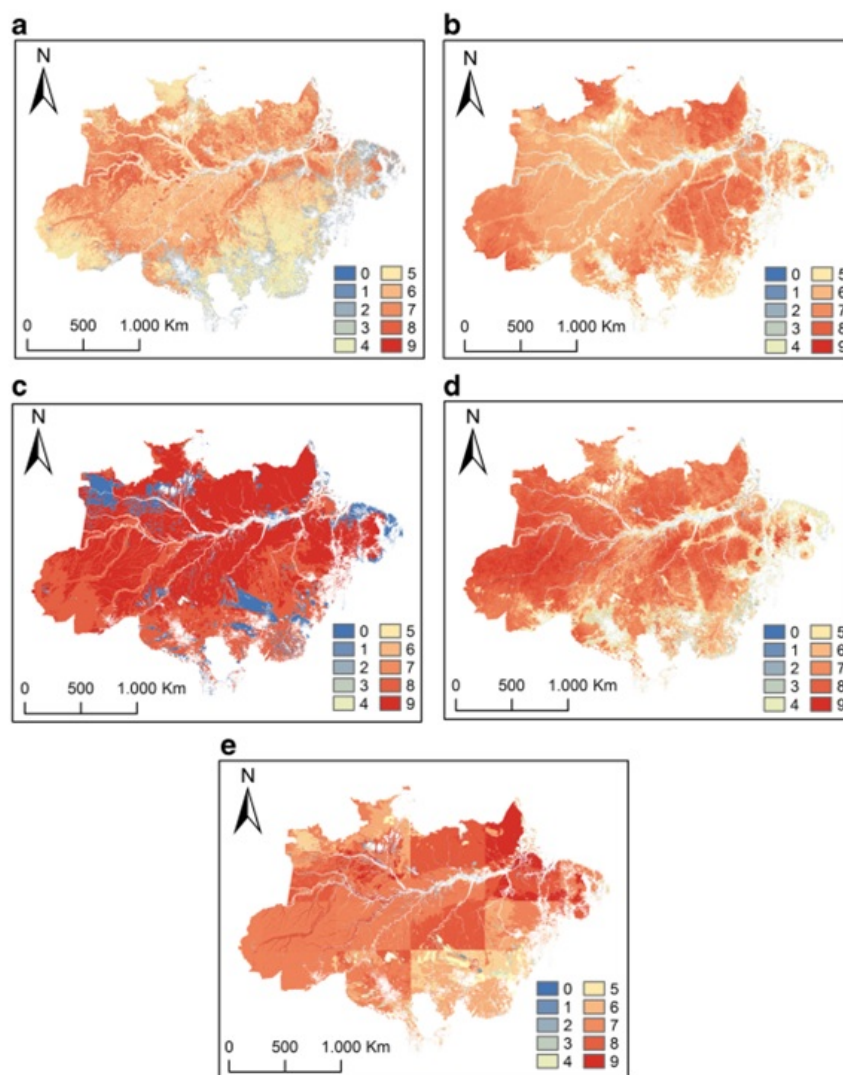
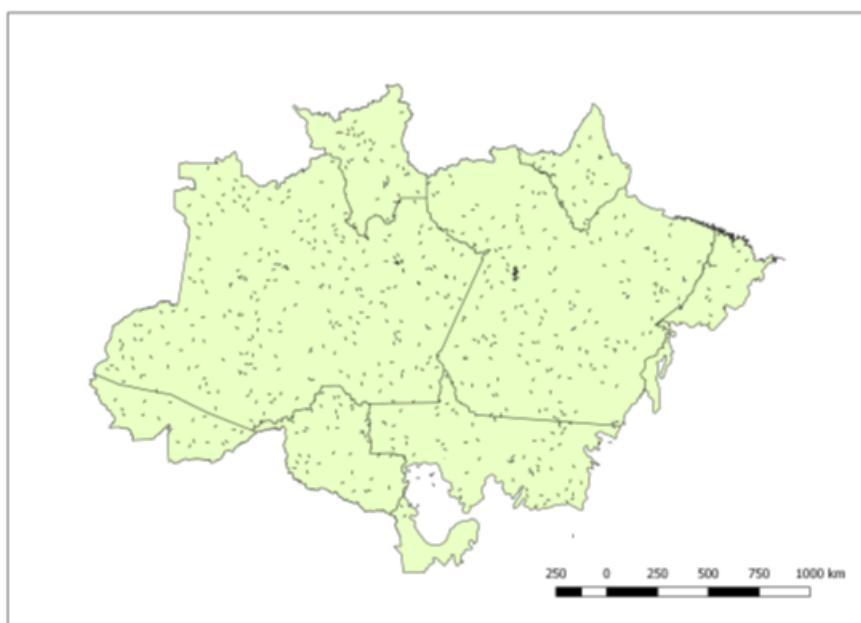


Figura 2.11 - Transectos voados no projeto EBA e que foram usados nesse projeto



3 MATERIAL E METODOS

3.1 Dados utilizados

Foram coletadas 836 nuvens de pontos LiDAR distribuídas sobre o Bioma Amazônico, entre março/2016 e abril/2018. Essa coleta de dados foi feita dentro do Projeto EBA (Estimativa de Biomassa da Amazônia), que visava mapear a distribuição de biomassa na Amazônia. Essas nuvens foram coletadas a partir de aeronave de asa fixa (CESSNA modelo 206), em cujo corpo estava acoplado um Sensor LiDAR HARRIER 68i com frequência de escaneamento de 200 Hz, capaz de coletar dados LiDAR full waveform e também uma IMU (Unidade de Medição Inercial) Dual frequency GNSS receiver L1 / L2 (Figura 3.1).

Figura 3.1 - Avião utilizado para coleta dos dados, onde se vê o sensor LiDAR instalado



A distribuição dessas nuvens no bioma pode ser vista na Figura 2.11.

A lista com os parâmetros das nuvens coletadas para esse estudo pode ser vista na Tabela 3.1.

Tabela 3.1 - Parâmetros de voos LiDAR do projeto EBA

Parâmetro	Valor
Ângulo total de escaneamento	45
Densidade mínima de pulsos	$4 m^{-2}$
Discreto/Full waveform	Discreto e full waveform
Footprint	0,30 m
Altura de voo	600 m
Largura de banda no chão	500 m
Versão do arquivo LAS	1.3

Quando da contratação de voos ALS, é necessário que sejam especificados os parâmetros de voo e os indicadores de qualidade da nuvem de pontos utilizada, dentre outras informações. O parâmetro mais importante a ser definido é a densidade de retornos por unidade de área. Como o objetivo da coleta dos dados analisados no âmbito do Projeto EBA é fazer um mapa de biomassa do Bioma Amazônia, era interessante que se coletasse o mais volume de dados possível devido ao tamanho da área de estudo. Assim, como o custo do voo é função a densidade de pontos, foi feito um estudo com o objetivo de determinar a mínima densidade de pontos que poderia ser utilizada sem que se perdesse precisão nas estimativas de biomassa a partir das nuvens coletadas. O estudo mostrou que com uma densidade de retornos acima de $2 m^{-2}$ esse objetivo seria atendido. Foi então definida uma densidade de retornos mínima a ser tolerada de $4 m^{-2}$, a ser mantida em todas as nuvens entregues.

3.2 Formato dos dados LiDAR utilizados

Dados LiDAR podem ser armazenados em diferentes formatos, sendo os mais comuns: i) LAS, definido e apoiado pela Sociedade Americana de Fotogrametria e Sensoriamento Remoto (ASPRS) e ii) LAZ, produto do software gratuito e de código aberto Rapidlasso GmbH laszip, que compacta os arquivos LAS sem alteração da informação. Usa-se menos frequentemente o formato CSV (valores separados por vírgula), porque normalmente os arquivos LiDAR contém milhões de registros, se beneficiando mais de um formato específico mais compacto.

Os dados usados no Projeto EBA foram entregues no formato LAS v1.3, conforme especificado na contratação dos voos. Segundo essa especificação, os dados LiDAR full waveform de uma nuvem de pontos são divididos em dois arquivos, com extensões LAS e WDP. No arquivo LAS encontra-se o cabeçalho que contém informações sobre como os dados estão organizados naquele arquivo, já que o formato permite

certa flexibilidade sobre isso e também os dados discretos coletados. No arquivo WDP encontram-se os dados full waveform correspondentes a cada registro discreto existente no arquivo LAS. Em cada registro discreto do arquivo LAS existe um apontador para a posição em bytes do arquivo discreto em que se encontram os correspondentes dados full waveform, bem como o número de registros full waveform coletados para o registro.

Os transectos analisados contém em torno de 5×10^7 registros de dados discretos cada. Cada registro de dados tem no mínimo 160 registros full waveform, que não foram objetos de análise neste trabalho.

3.3 Recebimento e controle de qualidade das nuvens de ponto LiDAR

Recebidos os dados no formato e no meio especificado em contrato, é necessário proceder à validação destes. Vários problemas podem ocorrer durante a coleta (voo) e no pós-processamento para geração da nuvem de pontos LiDAR no formato demandado que influenciariam os resultados obtidos a partir do processamento dessas nuvens. Exemplo destes problemas seriam: altura de voo excessiva em relação ao solo, voo muito rápido, grande variação da velocidade de voo, erro de posicionamento do voo (sobrevoo de área incorreta), falha do sensor presença de nuvens entre o sensor e a floresta.

Para isso, foi desenvolvido um conjunto de rotinas em Python que faz uma série de verificações das nuvens entregues, tanto com relação à integridade do formato dos dados quanto dos dados em si (??). Essas rotinas foram aplicadas em todas as nuvens processadas, de maneira a produzir um relatório com dados a respeito das nuvens coletadas, bem como listando eventuais inconsistências. A Tabela 3.2 indica as inconsistências avaliadas:

Tabela 3.2 - Parâmetros a serem validados nas nuvens de pontos.

Etapa de Validação	Descrição
Assinatura do arquivo	Os arquivos LAS começam com a assinatura LAS, o que indica que a nuvem foi processada com o software correto.
Versão do arquivo LAS/LAZ	Verifica se a versão do formato LAS do arquivo atual corresponde à especificada em contrato.
Número de retornos total	Quantidade de retornos na nuvem corresponde ao cabeçalho.
Quantidade de retornos por retorno	Quantidade de retornos por número de retorno existente na nuvem corresponde ao cabeçalho.
Valores mínimo e máximo	Máximos e mínimos valores das coordenadas X, Y e Z no arquivo e no cabeçalho.
Densidade global de retornos	Quantidade de retornos total dividida pela área efetivamente voada.
Máximo de células abaixo da densidade	Percentual de células (área da célula especificada na verificação) cuja densidade está abaixo da mínima.
Detecção de ruídos	Verifica se existe uma quantidade expressiva de outliers que indica problemas no sensor ou interferência de nuvens.

3.4 Preparação dos dados para obtenção das métricas LiDAR

Uma vez validadas as nuvens de pontos com relação à sua qualidade, veio a etapa de processamento, cujo objetivo é obter as métricas que serão usadas nas estimativas de estratificação. Para isso foi desenvolvido um conjunto de rotinas em Python usando-se multithreading para permitir o processamento de nuvens de pontos simultaneamente. Essas rotinas fazem chamadas ao software LASTools.

- a) **Filtragem de ruídos:** nessa etapa são removidos os pontos que tem “poucos” pontos em seu entorno. Para isso a rotina cria um cubo de 27 células (3 x 3 x 3) em torno do ponto e conta quantos pontos existem nessas células do entorno. Se o número de pontos nas células adjacentes for inferior a 5, o ponto é descartado. Foi usado o **lasnoise** para fazer isso.
- b) **Tiling:** para a execução da rotina de determinação da altura de cada ponto em relação ao solo existe uma limitação de cerca de 20×10^6 pontos por nuvem. Como as nuvens de pontos a serem usadas neste estudo em torno de 50×10^6 pontos, foi necessário dividir as nuvens para que seja possível processá-las. Para as nuvens processadas no nosso caso as nuvens foram divididas em tiles de 1000 x 1000 m, possíveis de serem processados para

a classificação de pontos. Foi mantido ainda um buffer de 100 x 100 metros em cada tile, de maneira a evitar o efeito borda quando processada a classificação dos pontos de solo. Usou-se o **lastile** para fazer essa divisão. A opção **-reversible** deve ser obrigatoriamente ativada nesse caso, porque nossa intenção é, ao fim do processamento de classificação de pontos descrito a seguir, reconstituir a nuvem de pontos.

- c) **Classificação dos pontos de solo:** dentro do registro de cada retorno no formato LAS tem um campo para classificação desse retorno. Essa etapa irá tentar identificar os pontos, classificando-os com o código 2 (ground) ou 1 (non-ground). O LASTools (**lasground**) usa uma versão modificada do algoritmo Axellson 2000 TIN para classificar os pontos (AXELSSON, 1999).
- d) **Normalização da altura:** nesta etapa os pontos que foram classificados como solo terão sua altura igualada a 0 e os pontos não-solo terão sua altura ajustada em relação ao “solo” (superfície definida a partir dos pontos de solo) Aqui foi usado o **lasheight** com a opção **-z** para que a altura fosse assim normalizada.
- e) **Classificação dos pontos não-solo em edificação e a vegetação:** a maioria dos transectos foi coletada sobre área de vegetação, mas existem algumas áreas antropizadas, então se faz necessário que os pontos advindos de estruturas construídas pelo homem sejam classificados, processo feito pelo **lasclassify**.
- f) **“Fusão” dos tiles na nuvem LiDAR:** geração de uma cópia da nuvem inteira normalizada e classificada, com a remoção dos buffers (**lastile** novamente).

3.5 Geração das métricas LiDAR

Nesse processo a nuvem de pontos é dividida em tiles de 50 x 50 m e são calculadas métricas para cada um desses tiles.

As métricas LiDAR geradas neste estudo são as encontradas na Tabela 3.

Para uso da distribuição Weibull foi usada a média (Avg) dos primeiros retornos de maneira a priorizar os pontos mais altos da quadrícula.

Tabela 3.3 - Métricas LiDAR.

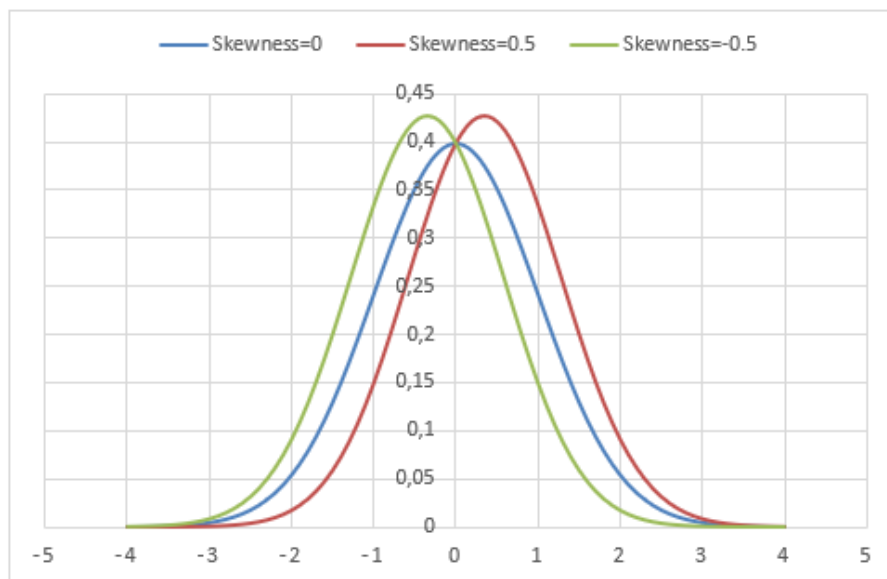
Métrica	Descrição
All	Número de retornos existentes.
Max	Altura mínima dentre os retornos.
Min	Altura máxima dentre os retornos.
Avg	Média das alturas dos retornos.
Std	Desvio padrão das alturas dos retornos.
Ske	Assimetria (skewness) da curva distribuição dos dados de altura.
Kur	Curtose da curva de distribuição dos dados de altura.
Perc	Altura que contém um determinado percentual de retornos em relação ao total. Foram calculados os percentis 1 5 10 20 25 30 40 50 60 70 75 80 90 95 99.
Bin	Altura que contém um determinado percentual de retornos em relação ao total. Foram calculados os percentis 1 5 10 20 25 30 40 50 60 70 75 80 90 95 99.
Perc	Percentual de retornos abaixo de uma determinada altura (m). Foram calculados os bincentis: 5 10 20 30 40 50 60 70 80 90.
Count	Conta os pontos existentes em faixas de alturas. Foram contados os retornos nas seguintes faixas: 0 1 2.5 5.0 7.5 10 15 20 25 30 35 40 45 50.
Dec	Razão entre a contagem acima e o número total de retornos.
Cov	Estimativa da cobertura da copa, definida como sendo o número de primeiros retornos dividido pelo total de retornos.

Dentre as métricas lidar, a skewness (assimetria) e a kurtosis (curtose) são as que tem a ver com a distribuição de altura, já que são calculadas em função da distribuição de pontos em relação ao eixo z.

Como o nome diz o valor de skewness expressa a assimetria de uma curva, no caso, de uma distribuição. Valor de skewness igual a zero indica uma assimetria perfeita em torno da média da distribuição (distribuição normal, por exemplo). Um skewness positivo indica que a curva se inclina para a direita do pico, enquanto que um skewness negativo indica o contrário. A Figura 3.2 mostra o efeito da variação do skewness sobre a distribuição normal (KIM, 2013).

Já a kurtosis é uma medida do “achatamento” de uma curva (em inglês usa-se “peakedness”) (KIM, 2013).

Figura 3.2 - Distribuição normal com diferentes valores de skewness



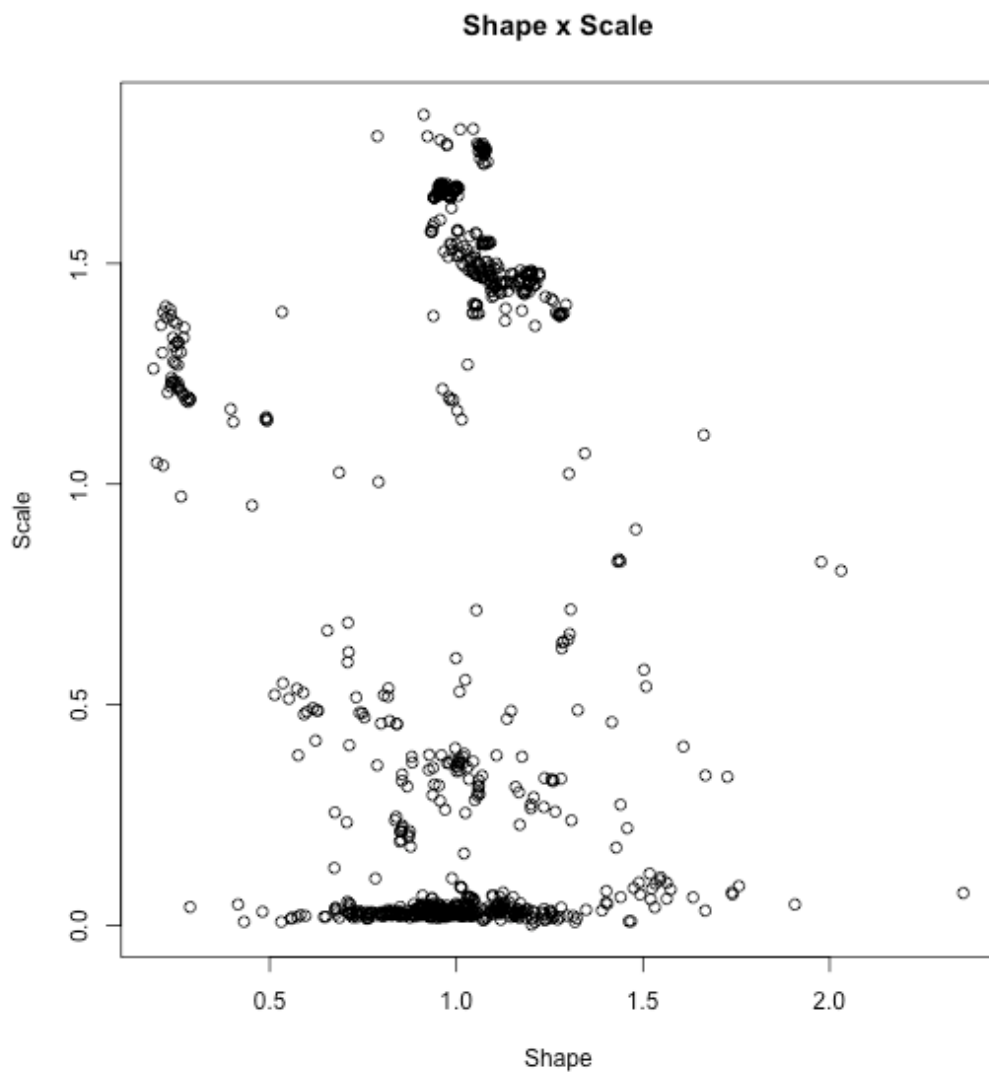
3.6 Aplicação da Distribuição Weibull em relação à altura pontos LiDAR

Para aplicar a distribuição Weibull em relação à altura dos pontos foi escolhida a métrica avg, que representa a altura média dos retornos LiDAR de um determinado pixel. Foram gerados os parâmetros shape e scale de cada um dos 836 transectos processados.

Usamos como base para gerar a distribuição Weibull a média dos pontos de primeiro retorno, de maneira a obter um dado de entrada que tivesse mais a ver com a estrutura vertical da floresta.

Em seguida foram usadas técnicas de agrupamento para tentar-se encontrar clusters a serem correlacionados com variáveis da floresta.

Figura 3.3 - Valores shape e scale plotados para os transectos analisados.



A Figura 3.3 foi obtida plotando-se os valores médios de shape em função scale

médio para cada transecto.

É possível verificar que alguns dos dados, representados por pontos no gráfico da Figura 3.3 aparentam estar agrupados, enquanto outros estão dispersos. Já pela análise do gráfico vemos que aparentemente existem apenas dois grupos, com outros dados dispersos. Mesmo assim procedeu-se à análise de agrupamento, tentando verificar se alguma correlação existiria.

Para a análise do agrupamento escolhemos o algoritmo Fuzzy C-Means (BEZDEK et al., 1984). Este algoritmo particiona o conjunto de dados em vários grupos, permitindo que um vetor de dados tenha diferentes pertinências a cada grupo, sendo, então, mais flexível do que o tradicionalmente usado K-Means (JAIN, 2010).

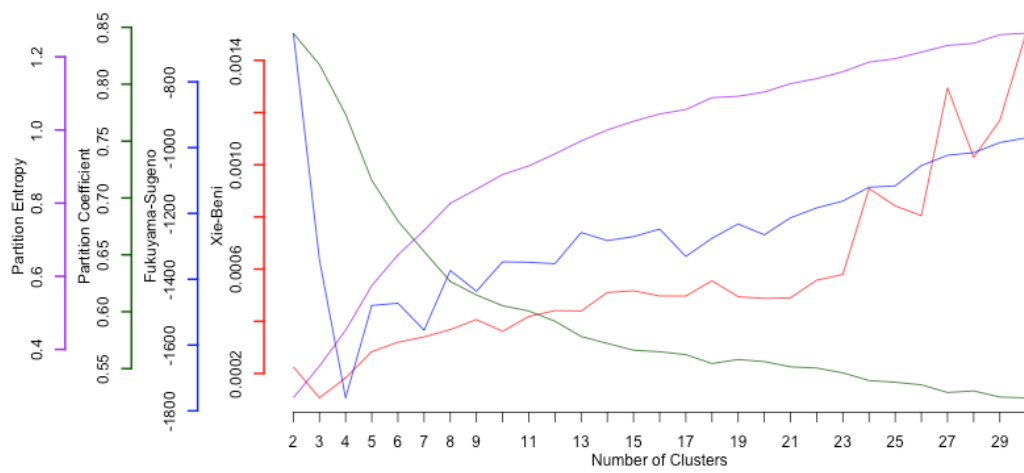
O algoritmo utiliza como entrada o conjunto de dados, um parâmetro C que indica o número de grupos que deve ser formado e um peso exponencial m que indica o quão espalhadas as pertinências de cada vetor de dados podem estar para cada grupo. m é um valor escolhido empiricamente, não devendo ser muito elevado (BEZDEK et al., 1984). Se $m=1$ os resultados do algoritmo Fuzzy C-Means são equivalentes ao do K-Means.

Outra vantagem do algoritmo Fuzzy C-Means é que é possível calcular várias métricas de qualidade do agrupamento obtido (PAL; BEZDEK, 1995). Estas métricas são úteis para ajudar a estabelecer um valor adequado para C , o número desejado de grupos. O algoritmo é executado para vários valores de C , as métricas para cada execução são calculadas, e seus mínimos (ou máximos, dependendo da métrica) são avaliadas para identificar bons candidatos para C .

Para identificar o melhor valor de C para os coeficientes Weibull executamos o algoritmo Fuzzy C-Means com os dados, C entre 2 e 40 e $m=2$. Avaliamos as métricas Entropia da Partição, Coeficiente da Partição, Fukuyama-Sugeno e Xie-Beni. Para identificar valores ideais de C devemos procurar o valor máximo para Entropia da Partição e o valor mínimo para as outras métricas (PAL; BEZDEK, 1995), (XIE; BENI, 1991).

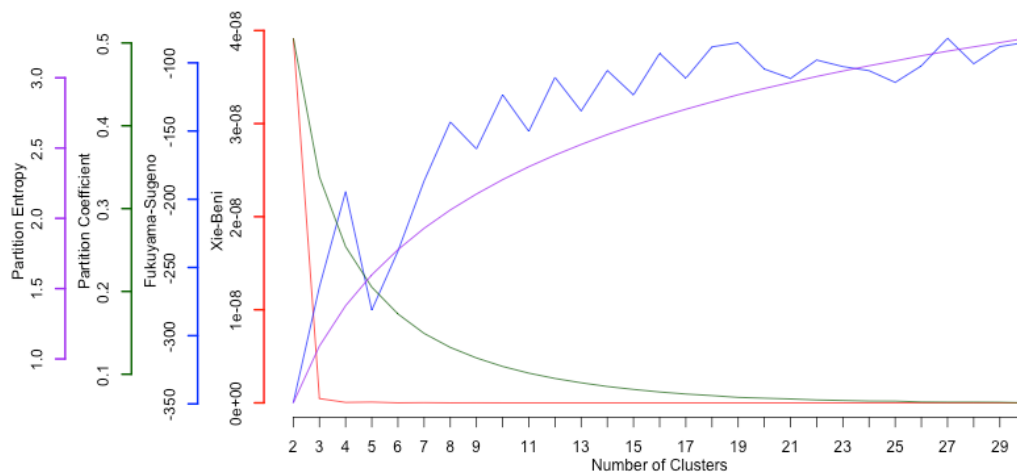
Foram variados os valores dos coeficientes para C entre 2 e 30, usando $m=2$ (Figura 3.4). Cada coeficiente foi calculado como a média dos valores obtidos de 10 execuções do algoritmo, para minimizar diferenças causadas pela aleatoriedade inicial usada pelo algoritmo.

Figura 3.4 - Coeficientes em função do parâmetro C com m=2.



Aumentar o valor de m causa agrupamentos mais "relaxados", com dados pertencentes a mais clusters com menores valores de pertinência. Um efeito colateral é que as curvas das métricas ficam ainda mais suaves, como na com $m=15$ (Figura 3.5).

Figura 3.5 - Coeficientes em função do parâmetro C com $m=15$.



As figuras mostram a dificuldade de se escolher o parâmetro m desse algoritmo.

Também mostram que identificar a presença de clusters no caso em estudo é muito difícil por esse algoritmo, já que estes podem variar de 3 a 5, que seriam pouco representativos dada a grande variação de tipos florestais no Bioma Amazônia.

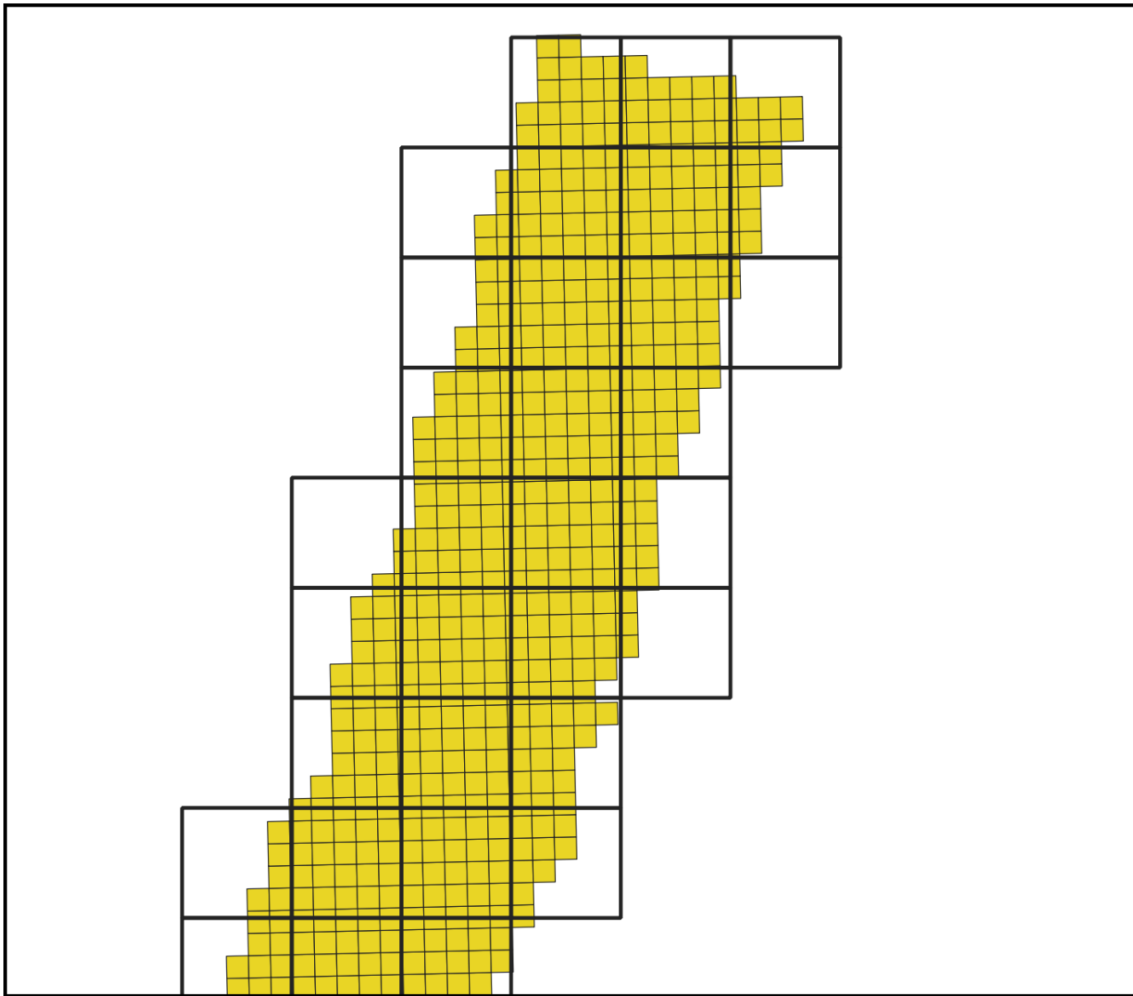
Dentre as métricas LiDAR duas se relacionam mais diretamente com a distribuição de altura: skewness e kurtosis, pela própria definição dessas métricas, calculadas em função da distribuição de pontos ao longo do eixo Z.

Como a métrica skewness tem a ver com a assimetria da curva, decidiu-se investigar a viabilidade desta com a estratificação da floresta.

3.7 Análise da métrica skewness para estratificação

Para analisar o skewness como uma alternativa de estratificação, gerou-se um a tabela com o skewness médio ponderado pela área sobre a máscara do MODIS (Figura 3.6). O pixel usado no MODIS tem 249,736 m² de lado.

Figura 3.6 - Máscara do MODIS aplicada às métricas LiDAR.



Essa aplicação foi feita para para testar uma prática diferente para estudo da estratificação da Amazônia baseada em métricas LiDAR.

A partir desses 14013 registros foi gerada um estimativa do skewness para todo o

bioma. Utilizou-se o algoritmo Random Forest para gerar um modelo de regressão com o skewness e as variáveis de sensoriamento remoto abaixo descritas (KARLSON et al., 2015), (BREIMAN, 2001). Com esse modelo gerou-se, também usando o Random Forest, o valor de skewness para todo o Bioma.

Foi usada como biblioteca para implementação da rotina que aplicou o Random Forest aos dados a plataforma H₂O (<https://www.h2o.ai/>). Esta plataforma foi selecionada porque tem implementação para Python e R, o algoritmo faz uso intensivo de paralelização podendo até ser rodado sob Hadoop, o que o torna especialmente adequado a ser usado no equipamento disponível (servidor com 16 núcleos e 32 Gb de RAM). Também é a mais utilizada das plataformas de machine-learning atuais, e possui uma versão free que poderia ser usada sem prejuízo aos resultados alcançados.

Foram usados como variáveis de entrada os seguintes dados:

EVI, NDVI: Dados do satélite MODIS/Terra com os produtos Vegetation Indices 16-Day L3 Global 250m MOD13Q1 - Enhanced Vegetation Index (EVI) e MODIS Normalized Difference Vegetation Index (NDVI) - 16-Day L3 Global 250m SIN Grid / imagens obtidas da série anual (jan até dez) temporal de 16 dias do ano de 2016. Para o NDVI e EVI foram geradas as imagens referentes às médias, medianas, mínimo, máximo, 1º quantil, 3º quantil, desvio padrão da série temporal de 2016.

PALSAR: polarizações HH e HV com resolução de 25m e obtidos no ano de 2015.

TRMM: Dados de precipitação: A série de precipitação pluviométrica (milímetros por mês) utiliza o produto 3B43 V6 do TRMM (Tropical Rainfall Measuring Mission). O valor de cada pixel é precipitação - acumulado mensal (mm) de jan/1998 a dez/2016, resolução de 0,25 graus (cerca de 27km).

Também foi gerada um agrupamento de skewness em 14 classes divididas igualmente e com a eliminação de 11% dos pixels nas duas extremidades da faixa de valores, considerados outliers. Essa seleção foi feita restringindo-se sucessivamente os limites superior e inferior considerados de maneira a que se obtivesse no mínimo 1000 ocorrências de cada classe.

Da mesma forma que no exercício anterior, foi gerado um mapa por classe de skewness, mostrando a distribuição dessas classes pelo bioma (Figura 4.7). O algoritmo usado foi também o Random Forest, desta vez como classificador e não como estimador.

4 RESULTADOS E DISCUSSÕES

4.1 Processos de controle de qualidade

As entregas das nuvens de pontos pela empresa contratada para coletar os dados eram feitas em lotes, que eram entregues assim que pré-processados. Esse pré-processamento, feito pela mesma empresa que coletou os dados, converte os dados do formato coletado pelo sensor e armazenado no próprio avião, que é um formato proprietário do fabricante do equipamento, para o formato LAS v1.3, especificado em contrato como padrão de recebimento.

O primeiro lote entregue, constituído por 148 transectos apresentou problemas em todos os arquivos ao passar pelo processo de controle de qualidade adotado no recebimento pelo INPE (ASSIS *et al.*, 2018), problemas esses que inviabilizariam o uso das informações full waveform coletadas. Apesar do projeto EBA prever usar somente a parte discreta da nuvem de pontos, como um dos objetivos do projeto era disponibilizar os dados a outros estudos que se interessassem em usá-los, resolveu-se buscar a correção desses problemas.

Foi então feita uma visita à empresa, que explicou que os problemas ocorreram porque o sensor era novo e estava sendo utilizado por eles pela primeira vez com aproveitamento dos dados full waveform. De toda forma, o que aconteceu dá a medida da importância de se avaliar a qualidade das nuvens de pontos recebidas sob contrato.

Os dados foram então novamente pré-processados e revalidados, passando nos testes de validação.

A partir daí, alinhadas as expectativas de qualidade das nuvem de pontos com o fornecedor, seguiram-se as entregas dos lotes voados, em intervalos de aproximadamente dois meses e em lotes menores do que o primeiro durante um período de 14 meses.

A cada entrega os respectivos lotes eram validados. Pequenos problemas, passíveis de correção pela própria equipe de qualidade, por exemplo, arquivos com nomenclatura fora do padrão, eram corrigidos pela própria equipe do projeto. Arquivos faltantes eram solicitados ao fornecedor.

Os problemas mais sérios encontrados vieram de da validação do número máximo de células com densidade abaixo da mínima especificada. A cada lote era enviado

um relatório à empresa contratada que tinha que responder ao documento com suas considerações. Em 86% dos casos levantados o problema deveu-se ao voo ter se dados sobre corpos de água, já que estes não retro-espalham os pulsos enviados pelo sensor. Ao se planejar os voos a locação dos transectos levou isso em consideração, mas pequenos rios e lagos, que podem ser inclusive sazonais, não puderam ser identificados previamente e foram sobrevoados.

Outra causa para a variação da densidade de pontos ao longo do transecto é a presença de vento de cauda ou de frente durante o voo, que provocam dificuldade para a uniformização da velocidade de voo durante a coleta. Algumas nuvens foram aceitas nessa condição, desde que o relatório de voo feito pelo piloto e entregue a cada lote indicasse essa ocorrência para aquele voo e que o número de células abaixo do percentual mínimo fosse pequeno, da ordem de 5%.

As áreas que após essas análises se apresentavam com problemas e para as quais não haviam explicações técnicas como as acima foram voadas novamente, no mesmo local ou em local diferente, conforme negociado entre equipe do projeto EBA e o fornecedor, de modo a favorecer a logística de deslocamento da aeronave sem interferir significativamente na distribuição dos voos. Isso foi necessário em 9 das 836 áreas em estudo.

4.2 Análise dos parâmetros da distribuição

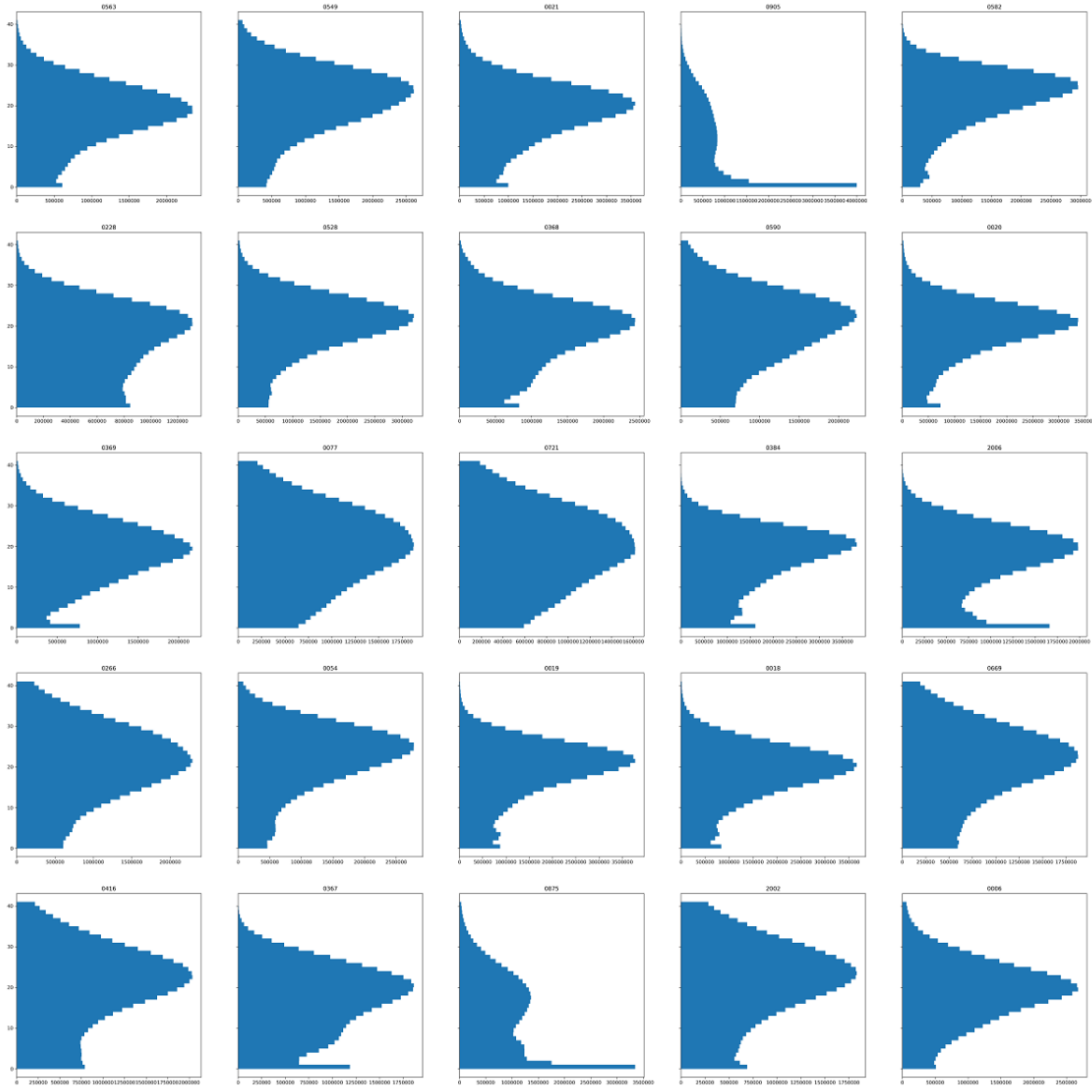
Para uma análise prévia dos perfis que poderiam ser encontrados, foram gerados gráficos que descrevessem visualmente esses perfis, correlacionando a contagem de retornos com a altura desses em relação ao solo.

Exemplos de perfis LiDAR em função da altura de alguns transectos estão mostrados nas figuras 4.1, 4.2, 4.3.

Pelas figuras dos perfis percebe-se que, apesar de existirem alguns perfis tipo “J invertido”, a maioria deles é sigmoidal. Isso já indica uma dificuldade de se achar alguma maneira de se distinguir matematicamente uma curva da outra.

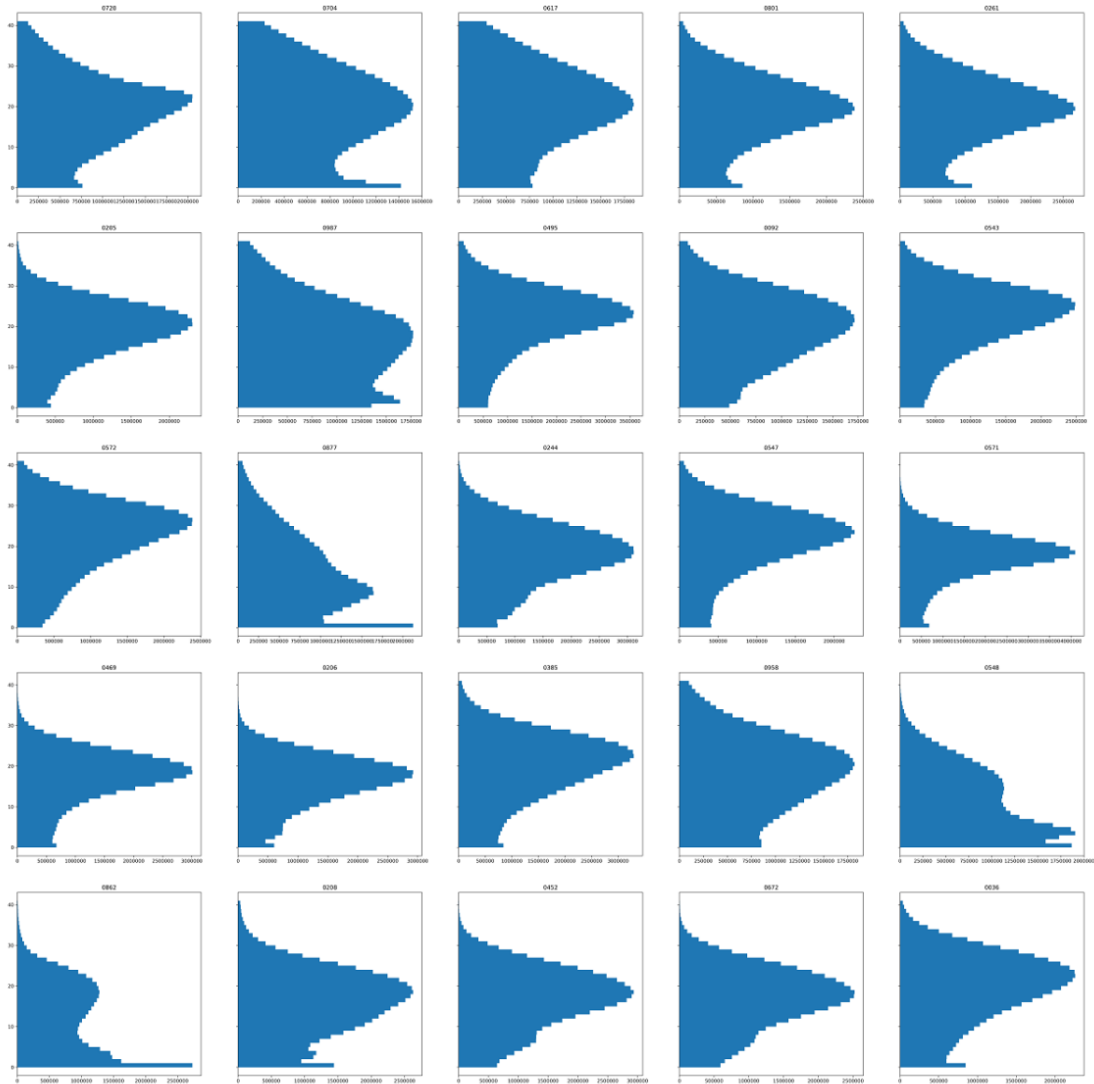
Já ao se observar visualmente a possibilidade de agrupamento das amostras (parâmetros Weibull dos transectos) vê-se a dificuldade de se encontrar, através dessa abordagem, um número de estratos que pudesse explicar diferentes classes de vegetação ou alguma variável contínua (biomassa, por exemplo) usando-se somente esses estratos.

Figura 4.1 - Perfis de transectos (1).



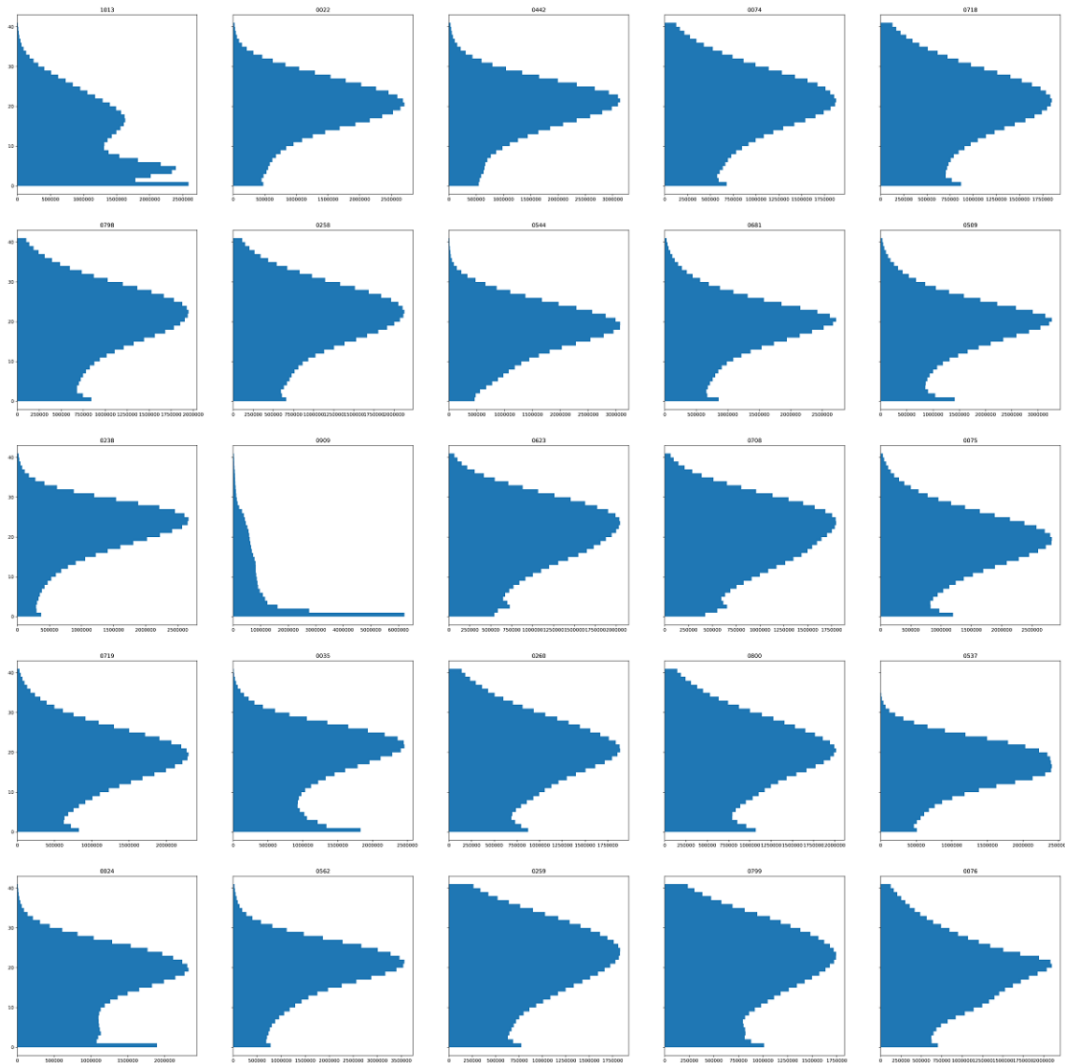
Eixo X: contagem de pontos, eixo Y a altura (m) destes pontos em relação ao solo.

Figura 4.2 - Perfis de transectos (2).



Eixo X: contagem de pontos, eixo Y a altura (m) destes pontos em relação ao solo.

Figura 4.3 - Perfis de transectos (3).



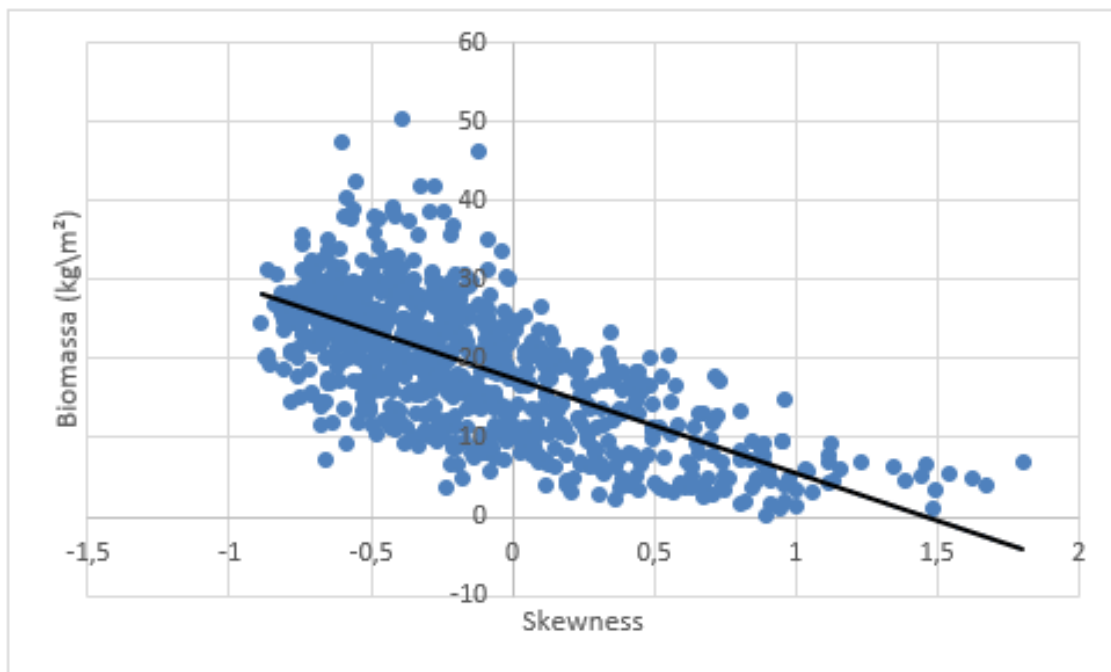
Eixo X: contagem de pontos, eixo Y a altura (m) destes pontos em relação ao solo.

Aplicando-se o algoritmo selecionado houve a indicação de que podem existir de 3 a 5 grupos com base nos parâmetros scale e shape, o que confirmou o que se viu na Figura 4.4. Mesmo com a variação do parâmetro C entre m o resultado se manteve, o que diz que o Fuzzy C-Means foi “assertivo” com relação aos grupos encontrados.

Concluimos então que, usando-se os parâmetros que descrevem a distribuição Weibull não foi possível identificar estratos entre os dois parâmetros da distribuição Weibull não conseguimos identificar estratos, ou seja, na abordagem que adotamos em com essas amostras, que são bastante abrangentes, não conseguimos estratificar de maneira satisfatória o Bioma Amazônico.

Correlacionando biomassa, que tem correlação com a estrutura da floresta com a skewness, achamos a correlação de Pearson $r(p) = -0,66$ (Figura 4.4).

Figura 4.4 - Biomassa em função do skewness.

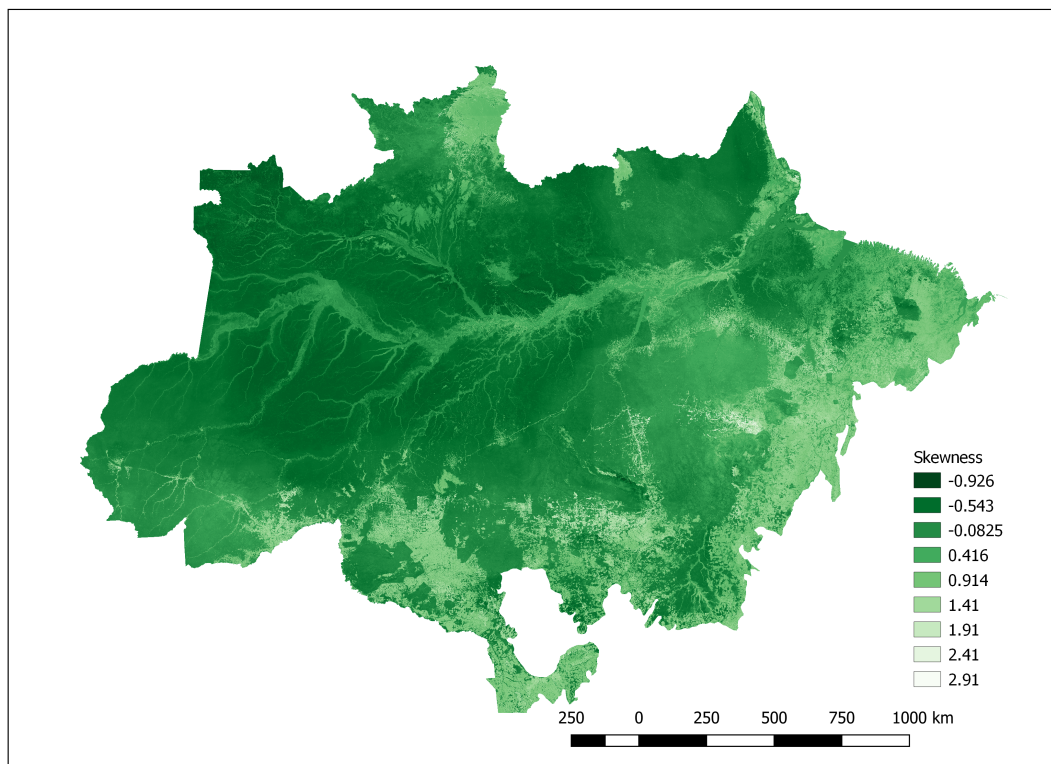


A biomassa dos transectos também é uma estimativa e foi feita usando-se a equação de descrita por Longo (LONGO et al., 2016). Como visto anteriormente, a biomassa tem uma correlação importante com a estrutura vertical da floresta. Como nessa

equação não entra a skewness, resolveu-se considerar essa estimativa de biomassa como a “verdade” para analisarmos a qualidade do uso dessa métrica para analisar a estrutura da floresta.

Essa estimativa deu origem a um mapa de skewness estimado (Figura 4.5).

Figura 4.5 - Mapa do skewness estimado para a área do Bioma Amazônia.



Esse mapa foi gerado partir dos valores de skewness calculados para pixels de 250m na mesma máscara a ser usada pelo mapa de biomassa feito no Projeto EBA, nas áreas sobrevoadas pelo LiDAR. Assim, para cada pixel de 250m foi calculado a média dos skewness estimados para os pixels de 50m coincidentes.

A partir daí foi gerado um modelo Random Forest usando-se os valores das variáveis obtidos de satélites descritos na seção 2.5 a partir da mesma máscara.

Tabela 4.1 - Importância relativa entre as variáveis usadas como entrada no modelo de estimativa do skewnes.

Variável	Importância relativa
y	0,348907
x	0,19922
srtm	0,157672
trmm_mean	0,150406
ndvi_q3	0,0273484
ndvi_mean	0,0236573
evi_q1	0,0208319
ndvi_q1	0,0188601
evi_q3	0,0172077
palsar_hv	0,0125773
evi_mean	0,0118392
palsar_hh	0,0114722

Na geração do mapa da Figura 4.5, o r^2 foi de 0,59.

Valores mais altos de skewness mostram os bosques florestais com a distribuição da altura com uma tendência de ter maiores valores (a curva da distribuição com pico mais para a direita) e os valores mais baixos de skewness mostram os bosques florestais que a distribuição da altura dos pontos lidar com uma tendência mais para os menores valores de altura.

O Random Forest fornece como resultado, além das estimativas, a importância de cada variável de entrada em relação às outras (Tabela 4.1).

A coordenada y, que é a coordenada geográfica da projeção SRID 5880 correspondente ao eixo Norte-Sul, “explica” em torno de um terço do valor da skewness encontrada pelo modelo. Uma provável explicação para isso é que, devido à grande extensão da floresta, a variação climática tenha uma influência importante na estrutura vertical da floresta, e aquela por sua vez tem a ver com a distância do Equador.

O resultado do agrupamento do skewness em classes pode ser visto na Figura 4.6.

Figura 4.6 - Histograma das células por classe de skewness.

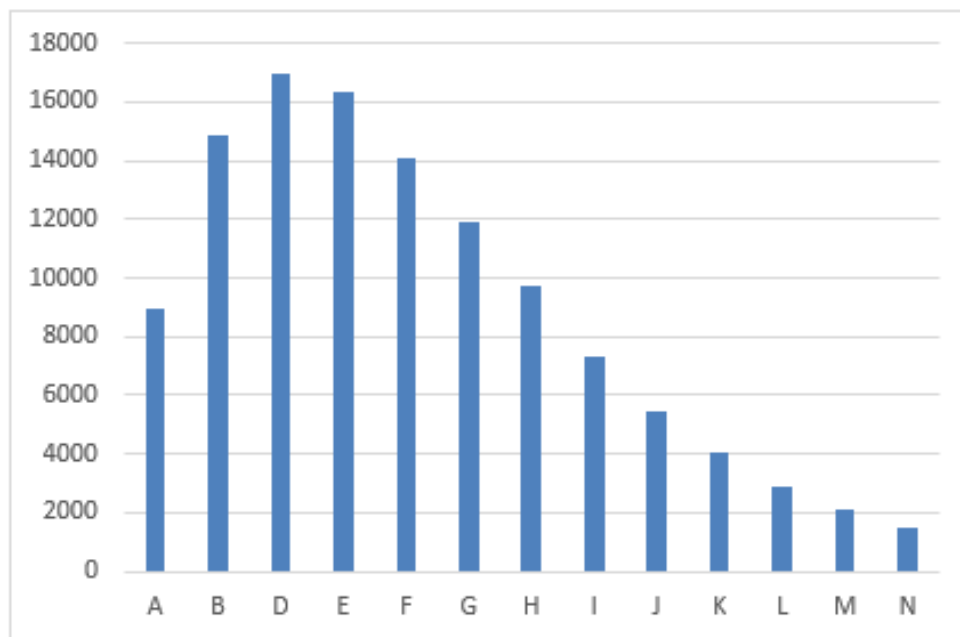
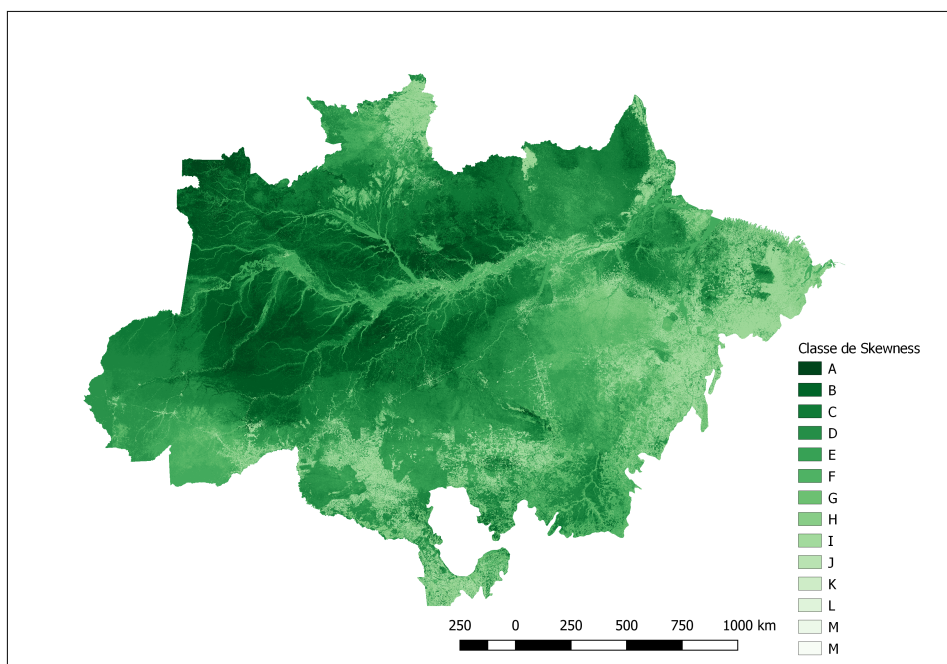


Tabela 4.2 - Classe de skewness usadas para elaboração do mapa de classes

Classe de skewness	Valor mínimo	Valor máximo
A	-0.909	-0.766
B	-0.766	-0.623
D	-0.623	-0.480
E	-0.480	-0.337
F	-0.337	-0.194
G	-0.194	-0.051
H	-0.051	0.092
I	0.092	0.235
J	0.235	0.378
K	0.378	0.521
L	0.521	0.664
M	0.664	0.807
N	0.807	0.950

Uma outra abordagem analisada foi dividir-se o skewness em 14 classes (representando as classes de vegetação mais significativas conforme a classificação do IBGE). Foram usadas as classes descritas na tabela 4.2

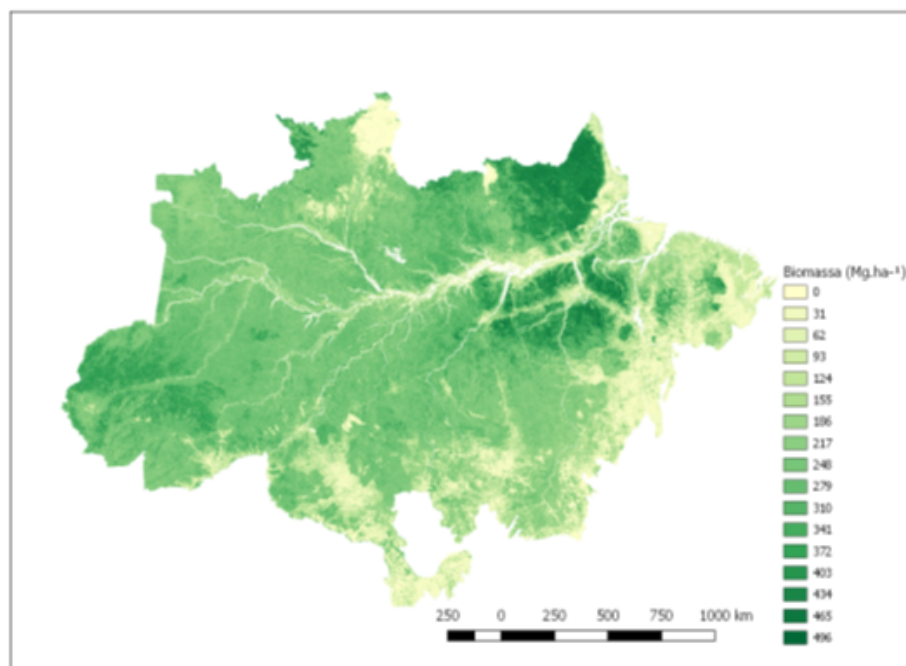
Figura 4.7 - Distribuição das classes de skewness no Bioma Amazônico.



Usando-se as classes a partir do skewness pode-se gerar um novo mapa, dessa vez usando-se o Random Forest para fazer uma estimativa de classificação de uma variável categórica, a classe gerada.

Comparando-se visualmente os mapas das figuras 4.5 e 4.7 não aparecem diferenças significativas, como era de se esperar, já que a relação entre eles é clara. Foram feitas as duas porque o algoritmo permite tanto o uso de variáveis contínuas de predição e, como se trata de estratificação, buscou-se ver se o fato de criarmos categorias a priori teria algum tipo de influência. Com isso concluímos que, caso opte-se por usar o skewness como classificador, pode-se optar por categorizar os dados depois da estimativa pelo Random Forest, já que o esforço computacional para o Random Forest fazer a predição com base nos dados já categorizados é bem maior. Isso se dá porque o Random Forest usa de um artifício para trabalhar com variável preditiva categórica que aumenta significativamente o esforço computacional.

Figura 4.8 - Mapa de Biomassa da Amazônia (AVITABILE et al., 2016)



Para termos outros parâmetros para analisarmos se seria possível a utilização do skewness como variável de estratificação da floresta, pode-se comparar o mapa gerado para as classes de skewness primeiramente com o mapa de biomassa feito por Avitabile et al. (AVITABILE et al., 2016) (Figura 4.8).

Para comparar melhor os dois mapas o mapa de biomassa foi cortado na mesma máscara do Bioma.

Comparando-se os dois visualmente pode-se inferir que existe uma correlação entre a distribuição dos valores de skewness pelo Bioma e a biomassa. Isso indica, pelo menos numa primeira análise, que realmente a skewness pode ser um parâmetro interessante para se descrever diferentes tipos de estrutura florestal, já que, como visto anteriormente, existe uma correlação forte entre biomassa e essa estrutura.

5 CONCLUSÃO

Esse trabalho teve como objetivo, como apresentado na introdução, encontrar um mecanismo que permitisse estratificar o Bioma Amazônia unicamente através da análise das métricas LiDAR.

Foi desenvolvido um conjunto de rotinas para tratamento das nuvens de ponto disponíveis, tanto para aferir a qualidade dos dados quanto para processar as métricas geradas.

O caminho adotado nessa dissertação, foi usar os parâmetros shape e scale da distribuição Weibull, implementada sobre os pontos distribuídos em função da altura. Usando-se os valores desses dois parâmetros para a identificação de grupos usando-se o algoritmo c-means, não foi possível identificar a formação de grupos distintos que pudesse ser usada em estratificação florestal. De toda forma, a própria análise visual dos grupo plotados já levava a essa conclusão.

Outra análise que foi feita foi usar uma das métricas LiDAR, a skewness, para se buscar a estratificação da floresta. Nesse processo foi usado o valor dessa métrica para 142013 pixels de 249,736 m², correspondentes às áreas cobertas pelos voos LiDAR.

Em seguida esses valores foram extrapolados para todo o bioma usando-se o algoritmo Random Forest. O r^2 obtido nesse processo foi de 0,59. Essa extrapolação foi feita de duas formas: uma regressão, em que os valores de skewness foram usados como calculados (números de ponto flutuante), outra classificação, onde os valores de skewness foram classificados em 13 classes (A a N) segundo o

A comparação visual entre o mapa gerado a partir do skewness e o mapa de biomassa indica que pode existir no estudo do skewness uma alternativa de descrição dos diferentes perfis verticais da floresta.

O principal produto desse trabalho foi evidenciar a importância do skewness de nuvens de pontos LiDAR como um descritor da variabilidade da estrutura florestal.

Uma sugestão de trabalho futuro é aprofundar a análise da possibilidade do uso da skewness para esse tipo de análise, bem como estudar seu uso em combinação com outras métricas LiDAR.

REFERÊNCIAS BIBLIOGRÁFICAS

ASNER, G. P.; MASCARO, J.; MULLER-LANDAU, H. C.; VIEILLEDENT, G.; VAUDRY, R.; RASAMOELINA, M.; HALL, J. S.; BREUGEL, M. V. A universal airborne lidar approach for tropical forest carbon mapping. **Oecologia**, v. 168, n. 4, p. 1147–1160, 2012. Disponível em:

<<https://doi.org/10.1016/j.foreco.2008.08.024>>. Acesso em: 8 fev. 2019. 10

ASSIS, M. L. R. de; OMETTO, J. P.; SANTOS, R.; GÖRGENS, E.; SATO, L. d. S.; PEREIRA, F. R.; CANTINHO, R. Z.; OLIVEIRA, P. V. de. Controle e garantia de qualidade de dados lidar para estimativa de biomassa florestal. In: **Anais Simpósio de Geotecnologia no Pantanal (GEOPANTANAL)**. [s.n.], 2018. p. 656–664. Disponível em: <<https://www.geopantanal.cnptia.embrapa.br/Anais-Geopantanal/pdfs/p96.pdf>>.

Acesso em: 18 fev. 2019. 31

AVITABILE, V.; HEROLD, M.; HEUVELINK, G. B. M.; LEWIS, S. L.; PHILLIPS, O. L.; ASNER, G. P.; ARMSTON, J.; S. ASHTON, P.; BANIN, L.; BAYOL, N.; J. BERRY, N.; BOECKX, P.; JONG, B. H. J. D.; DEVRIES, B.; GIRARDIN, C. J.; KEARSLEY, E.; LINDSELL, J. A.; LOPEZ-GONZALEZ, G.; LUCAS, R.; DERMALHI, Y.; MOREL, A.; MITCHARD, E.; NAGY, L.; QIE, L.; QUINONES, M. J.; RYAN, C. M.; FERRY, S. W.; SUNDERLAND, T.; LAURIN, G. V.; GATTI, R. C.; VALENTINI, R.; VERBEECK, H.; WIJAYA, A.; WILLCOCK, S. An integrated pan-tropical biomass map using multiple reference datasets. **Global Change Biology**, v. 22, n. 4, p. 1406–1420, 2016. Disponível em: <<https://onlinelibrary.wiley.com/doi/full/10.1111/gcb.13139>>.

Acesso em: 8 fev. 2019. xv, 41

AXELSSON, P. Processing of laser scanner data—algorithms and applications. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 54, n. 2-3, p. 138–147, 1999. Disponível em:

<[https://doi.org/10.1016/S0924-2716\(99\)00008-8](https://doi.org/10.1016/S0924-2716(99)00008-8)>. Acesso em: 8 fev. 2019. 7, 21

BACCINI, A.; GOETZ, S. J.; WALKER, W. S.; LAPORTE, N. T.; SUN, M.; SULLA-MENASHE, D.; HACKLER, J.; BECK, P. S. A.; DUBAYAH, R.; FRIEDL, M. A.; SAMANTA, S. Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. **Nature Climate Change**, v. 2,

n. 3, p. 182, 2012. Disponível em:

<<https://www.nature.com/articles/nclimate1354>>. Acesso em: 8 fev. 2019.

13

BEZDEK, J. C.; EHRLICH, R.; FULL, W. Fcm: the fuzzy c-means clustering algorithm. **Computers & Geosciences**, v. 10, n. 2-3, p. 191–203, 1984.

Disponível em: <<https://www.sciencedirect.com/science/article/pii/0098300484900207>>.

Acesso em: 8 fev. 2019. 25

BREIMAN, L. Random forests. **Machine Learning**, v. 45, n. 1, p. 5–32, 2001. 29

COOPS, N. C.; HILKER, T.; WULDER, M. A.; ST-ONGE, B.; NEWNHAM, G.; SIGGINS, A.; TROFYMOW, J. T. Estimating canopy structure of douglas-fir forest stands from discrete-return LIDAR. **Trees**, v. 21, n. 3, p. 295, 2007.

Disponível em:

<<https://link.springer.com/article/10.1007/s00468-006-0119-6>>. Acesso em: 8 fev. 2019. 11

COOPS, N. C.; VARHOLA, A.; BATER, C. W.; TETI, P.; BOON, S.; GOODWIN, N.; WEILER, M. Assessing differences in tree and stand structure following beetle infestation using LIDAR data. **Canadian Journal of Remote Sensing**, v. 35, n. 6, p. 497–508, 2009. Disponível em:

<<https://www.tandfonline.com/doi/abs/10.5589/m10-005>>. Acesso em: 8 fev. 2019. 11

COUTINHO, L. M. O conceito de bioma. **Acta Botanica Brasílica**, v. 20, n. 1, p. 13–23, 2006. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-33062006000100002&lng=en&nrm=iso>. Acesso em: 8 fev. 2019. 3

DEAN, T. J.; CAO, Q. V.; ROBERTS, S. D.; EVANS, D. L. Measuring heights to crown base and crown median with lidar in a mature, even-aged loblolly pine stand. **Forest Ecology and Management**, v. 257, n. 1, p. 126–133, 2009.

Disponível em: <<https://doi.org/10.1016/j.foreco.2008.08.024>>. Acesso em: 8 fev. 2019. 11

DONOHUE, A.; ARMOUR, K. C.; PENDERGRASS, A. G.; BATTISTI, D. S. Shortwave and longwave radiative contributions to global warming under increasing co₂. **Proceedings of the National Academy of Sciences**, v. 111, n. 47, p. 16700–16705, 2014. Disponível em:

<<https://www.pnas.org/content/111/47/16700.full>>. Acesso em: 8 fev. 2019. 1

FAWAGREH, K.; GABER, M. M.; ELYAN, E. Random forests: from early developments to recent advancements. **Systems Science & Control Engineering: An Open Access Journal**, v. 2, n. 1, p. 602–609, 2014. Disponível em: <<https://doi.org/10.1080/21642583.2014.956265>>. Acesso em: 8 fev. 2019. 12

FOOD AND AGRICULTURE ORGANIZATION (FAO). **FRA 2015**: terms and definitions. 2015. Disponível em: <<http://www.fao.org/docrep/017/ap862e/ap862e00.pdf>>. 3

FREITAS, E. de. Composição da floresta amazônica. **Brasil Escola**, 2018. Disponível em: <<https://brasilecola.uol.com.br/brasil/a-composicao-floresta-amazonica.htm>>. Acesso em: 28 jan. 2019. 4

JAIN, A. K. Data clustering: 50 years beyond k-means. **Pattern Recognition Letters**, v. 31, n. 8, p. 651–666, 2010. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0167865509002323>>. Acesso em: 8 fev. 2019. 25

KARLSON, M.; OSTWALD, M.; REESE, H.; SANOU, J.; TANKOANO, B.; MATTSSON, E. Mapping tree canopy cover and aboveground biomass in sudano-sahelian woodlands using landsat 8 and random forest. **Remote Sensing**, v. 7, n. 8, p. 10017–10041, 2015. Disponível em: <<https://www.mdpi.com/2072-4292/7/8/10017>>. Acesso em: 8 fev. 2019. 29

KIM, H.-Y. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. **Restorative Dentistry & Endodontics**, v. 38, n. 1, p. 52–54, 2013. Disponível em: <<https://synapse.koreamed.org/search.php?where=aview&id=10.5395/rde.2013.38.1.52&code=2185RDE&vmode=FULL>>. Acesso em: 8 fev. 2019. 23

KOHYAMA, T. Simulation of the structural development of warm-temperate rain forest stands. **Annals of Botany**, v. 63, n. 6, p. 625–634, 1989. Disponível em: <<https://academic.oup.com/aob/article-abstract/63/6/625/253055>>. Acesso em: 8 fev. 2019. 10

KRAUS, K.; PFEIFER, N. Advanced DTM generation from LIDAR data. **International Archives Of Photogrammetry Remote Sensing And**

Spatial Information Sciences, v. 34, n. 3/W4, p. 23–30, 2001. Disponível em: <<https://pdfs.semanticscholar.org/37b7/4836b625ef9b9720312a3d45888223ab0e5e.pdf>>. Acesso em: 8 fev. 2019. 9

LEFSKY, M. A. **Application of LIDAR remote sensing to the estimation of forest canopy and stand structure**. Thesis (PhD) — University of Virginia, Charlottesville VA, EUA, 1997. Disponível em: <<http://adsabs.harvard.edu/abs/1997PhDT.....212L>>. Acesso em: 8 fev. 2019. 2

LIM, K.; TREITZ, P.; WULDER, M.; ST-ONGE, B.; FLOOD, M. Lidar remote sensing of forest structure. **Progress in Physical Geography**, v. 27, n. 1, p. 88–106, 2003. Disponível em: <<https://journals.sagepub.com/doi/10.1191/0309133303pp360ra>>. Acesso em: 8 fev. 2019. 7, 10

LONGO, M.; KELLER, M.; SANTOS, M. N. dos; LEITOLD, V.; PINAGÉ, E. R.; BACCINI, A.; SAATCHI, S.; NOGUEIRA, E. M.; BATISTELLA, M.; MORTON, D. C. Aboveground biomass variability across intact and degraded forests in the Brazilian Amazon. **Global Biogeochemical Cycles**, v. 30, n. 11, p. 1639–1660, 2016. Disponível em: <<https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2016GB005465>>. Acesso em: 8 fev. 2019. 36

LOVELL, J.; JUPP, D. L.; CULVENOR, D.; COOPS, N. Using airborne and ground-based ranging LIDAR to measure canopy structure in australian forests. **Canadian Journal of Remote Sensing**, v. 29, n. 5, p. 607–622, 2003. Disponível em: <<https://www.tandfonline.com/doi/abs/10.5589/m03-026>>. Acesso em: 8 fev. 2019. 2

MITCHARD, E. T. A.; FELDPAUSCH, T. R.; BRIENEN, R. J. W.; LOPEZ-GONZALEZ, G.; MONTEAGUDO, A.; BAKER, T. R.; LEWIS, S. L.; LLOYD, J.; QUESADA, C. A.; GLOOR, M.; STEEGE, H. ter; MEIR1, P.; ALVAREZ, E.; ARAUJO-MURAKAMI, A.; ARAGÃO, L. E. O. C.; ARROYO, L.; AYMARD, G.; BANKI, O.; BONAL1, D.; BROWN, S.; BROWN, F. I.; CERÓN, C. E.; MOSCOSO, V. C.; CHAVE, J.; COMISKEY, J. A.; CORNEJO, F.; MEDINA, M. C.; COSTA, L. D.; COSTA, F. R. C.; FIORE, A. D.; DOMINGUES, T. F.; ERWIN, T. L.; FREDERICKSON, T.; HIGUCHI, N.; CORONADO, E. N. H.; KILLEEN, T. J.; LAURANCE, W. F.; LEVIS, C.; MAGNUSSON, W. E.; MARIMON, B. S.; JUNIOR, B. H. M.; POLO, I. M.;

MISHRA, P.; NASCIMENTO, M. T.; NEILL, D.; VARGAS, M. P. N.; PALACIOS, W. A.; PARADA, A.; MOLINA, G. P.; PEÑA-CLAROS, M.; PITMAN, N.; PERES, C. A.; POORTER, L.; PRIETO, A.; RAMIREZ-ANGULO, H.; CORREA11, Z. R.; ROOPSIND, A.; ROUCOUX, K. H.; RUDAS, A.; SALOMÃO, R. P.; SCHIETTI, J.; SILVEIRA1, M.; SOUZA, P. F. de; STEININGER, M. K.; STROPP, J.; TERBORGH, J.; THOMAS, R.; TOLEDO, M.; TORRES-LEZAMA, A.; ANDE, T. R. van; HEIJDEN, G. M. F. van der; VIEIRA, I. C. G.; VIEIRA, S.; VILANOVA-TORRE, E.; VOS, V. A.; WANG, O.; ZARTMAN, C. E.; MALHI, Y.; PHILLIPS, O. L. Markedly divergent estimates of a mazon forest carbon density from ground plots and satellites.

Global Ecology and Biogeography, v. 23, n. 8, p. 935–946, 2014. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1111/geb.12168>>.

Acesso em: 8 fev. 2019. 13

NOGUEIRA, E. M.; FEARNside, P. M.; NELSON, B. W.; BARBOSA, R. I.; KEIZER, E. W. H. Estimates of forest biomass in the brazilian amazon: new allometric equations and adjustments to biomass from wood-volume inventories.

Forest Ecology and Management, v. 256, n. 11, p. 1853–1867, 2008.

Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0378112708005689>>.

<<https://www.sciencedirect.com/science/article/pii/S0378112708005689>>.

Acesso em: 8 fev. 2019. 13

NOGUEIRA, E. M.; YANAI, A. M.; FONSECA, F. O.; FEARNside, P. M. Carbon stock loss from deforestation through 2013 in brazilian amazonia. **Global Change Biology**, v. 21, n. 3, p. 1271–1292, 2015. Disponível em:

<<https://onlinelibrary.wiley.com/doi/full/10.1111/gcb.12798>>. Acesso

em: 8 fev. 2019. 13

OMETTO, J. P.; AGUIAR, A. P.; ASSIS, T.; SOLER, L.; VALLE, P.; TEJADA, G.; LAPOLA, D. M.; MEIR, P. Amazon forest biomass density maps: tackling the uncertainty in carbon emission estimates. In: OMETTO, J. P.; JONAS, R. B.; NAHORSKI, Z. (Ed.). **Uncertainties in Greenhouse Gas Inventories**. [s.n.], 2015. p. 95–110. Disponível em:

<https://link.springer.com/chapter/10.1007/978-3-319-15901-0_7>.

Acesso em: 8 fev. 2019. 13

PAL, N. R.; BEZDEK, J. C. On cluster validity for the fuzzy c-means model.

IEEE Transactions on Fuzzy systems, v. 3, n. 3, p. 370–379, 1995. Disponível em: <<https://ieeexplore.ieee.org/abstract/document/413225>>. Acesso em:

8 fev. 2019. 25

SAATCHI, S. S.; HARRIS, N. L.; BROWN, S.; LEFSKY, M.; MITCHARD, E. T. A.; SALAS, W.; ZUTTA, B. R.; BUERMANN, W.; LEWIS, S. L.; HAGEN, S.; PETROVA, S.; WHITE, L.; SILMAN, M.; MOREL, A. Benchmark map of forest carbon stocks in tropical regions across three continents. **Proceedings of the National Academy of Sciences**, v. 108, n. 24, p. 9899–9904, 2011. Disponível em: <<https://www.pnas.org/content/108/24/9899.short>>. Acesso em: 8 fev. 2019. 13

SAATCHI, S. S.; HOUGHTON, R.; ALVALA, R. D. S.; SOARES, J. V.; YU, Y. Distribution of aboveground live biomass in the amazon basin. **Global Change Biology**, v. 13, n. 4, p. 816–837, 2007. Disponível em: <<https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2486.2007.01323.x>>. Acesso em: 8 fev. 2019. 1, 4, 13

SOUZA, D. R. d. et al. Estratificação vertical em floresta ombrófila densa de terra firme não explorada, amazônia oriental. **Revista Árvore**, v. 28, n. 5, 2004. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-67622004000500008>. Acesso em: 8 fev. 2019. 10

WAGNER, F. H.; HÉRAULT, B.; ROSSI, V.; HILKER, T.; MAEDA, E. E.; SANCHEZ, A.; LYAPUSTIN, A. I.; GALVÃO, L. S.; WANG, Y.; ARAGÃO, L. E. Climate drivers of the amazon forest greening. **PloS one**, v. 12, n. 7, p. e0180932, 2017. Disponível em: <<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0180932>>. Acesso em: 8 fev. 2019. 1

WEIHULL, W. A statistical distribution function of wide applicability. **J Appl Mech**, v. 18, p. 290–293, 1951. Disponível em: <<https://pdfs.semanticscholar.org/88c3/7770028e7ed61180a34d6a837a9a4db3b264.pdf>>. Acesso em: 8 fev. 2019. 11

WILLIAMS, R. G.; ROUSSENOV, V.; GOODWIN, P.; RESPLANDY, L.; BOPP, L. Sensitivity of global warming to carbon emissions: effects of heat and carbon uptake in a suite of earth system models. **Journal of Climate**, v. 30, n. 23, p. 9343–9363, 2017. Disponível em: <<https://journals.ametsoc.org/doi/10.1175/JCLI-D-16-0468.1>>. Acesso em: 8 fev. 2019. 1

XIE, X. L.; BENI, G. A validity measure for fuzzy clustering. **IEEE Transactions on Pattern Analysis & Machine Intelligence**, n. 8, p.

841–847, 1991. Disponível em:

<<https://www.computer.org/csdl/trans/tp/1991/08/i0841.pdf>>. Acesso em: 8 fev. 2019. 25

ZIMBLE, D. A.; EVANS, D. L.; CARLSON, G. C.; PARKER, R. C.; GRADO, S. C.; GERARD, P. D. Characterizing vertical forest structure using small-footprint airborne lidar. **Remote sensing of Environment**, v. 87, n. 2-3, p. 171–182, 2003. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0034425703001391>>.

Acesso em: 18 fev. 2019. 10

PUBLICAÇÕES TÉCNICO-CIENTÍFICAS EDITADAS PELO INPE

Teses e Dissertações (TDI)

Teses e Dissertações apresentadas nos Cursos de Pós-Graduação do INPE.

Manuais Técnicos (MAN)

São publicações de caráter técnico que incluem normas, procedimentos, instruções e orientações.

Notas Técnico-Científicas (NTC)

Incluem resultados preliminares de pesquisa, descrição de equipamentos, descrição e ou documentação de programas de computador, descrição de sistemas e experimentos, apresentação de testes, dados, atlas, e documentação de projetos de engenharia.

Relatórios de Pesquisa (RPQ)

Reportam resultados ou progressos de pesquisas tanto de natureza técnica quanto científica, cujo nível seja compatível com o de uma publicação em periódico nacional ou internacional.

Propostas e Relatórios de Projetos (PRP)

São propostas de projetos técnico-científicos e relatórios de acompanhamento de projetos, atividades e convênios.

Publicações Didáticas (PUD)

Incluem apostilas, notas de aula e manuais didáticos.

Publicações Seriadas

São os seriados técnico-científicos: boletins, periódicos, anuários e anais de eventos (simpósios e congressos). Contam destas publicações o Internacional Standard Serial Number (ISSN), que é um código único e definitivo para identificação de títulos de seriados.

Programas de Computador (PDC)

São a seqüência de instruções ou códigos, expressos em uma linguagem de programação compilada ou interpretada, a ser executada por um computador para alcançar um determinado objetivo. Aceitam-se tanto programas fonte quanto os executáveis.

Pré-publicações (PRE)

Todos os artigos publicados em periódicos, anais e como capítulos de livros.