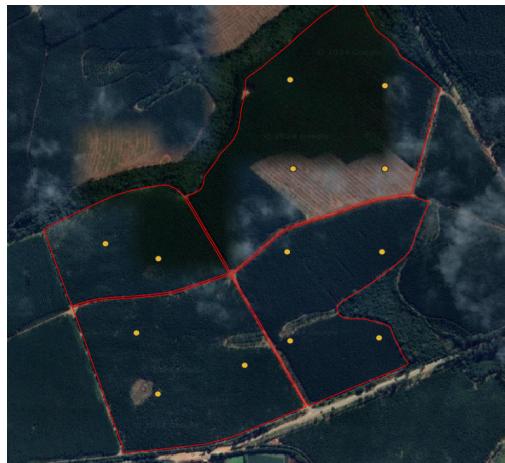


Descrição das etapas de processamento de dados de inventário realizado com o auxílio da tecnologia LiDAR

Otávio Magalhães Silva Souza



Piracicaba, SP – Data de Emissão: 05 de agosto de 2024

Contents

| | |
|---|-----------|
| 1 Pacotes utilizados no R (colocar breve descrição - já tem uma descriçaozinha no R passado em aula) | 3 |
| 1.1 Tidyverse (https://livro.curso-r.com/4-2-tidyverse.html) | 3 |
| 1.2 Sf | 3 |
| 1.3 Tidyterra | 4 |
| 1.4 Terra | 4 |
| 1.5 Stars | 4 |
| 1.6 Tools | 4 |
| 1.7 RColorBrewer | 4 |
| 1.8 Progress | 4 |
| 1.9 Reshape2 | 4 |
| 1.10 Mapview | 4 |
| 1.11 LidR | 4 |
| 1.12 RCSF | 4 |
| 1.13 Future | 4 |
| 2 Descrição da área | 5 |
| 3 Grid e parcelas já inventariadas | 6 |
| 4 Conceitos e definições da Dupla Amostragem | 7 |
| 5 Processamento da nuvem LiDAR | 7 |
| 5.1 Fluxograma e etapas Dupla amostragem | 9 |
| 6 Fluxograma e etapas Tripla amostragem | 24 |
| 7 Regressão linear simples | 25 |
| 8 Arborimetria preditiva | 25 |

1 Pacotes utilizados no R (colocar breve descrição - já tem uma descriçaozinha no R passado em aula)

1.1 Tidyverse (<https://livro.curso-r.com/4-2-tidyverse.html>)

O Tidyverse é um pacote guarda-chuva e contém diversas funções úteis para garantir o dinamismo no script, visualização, processamento e análise dos dados, modelagem etc.

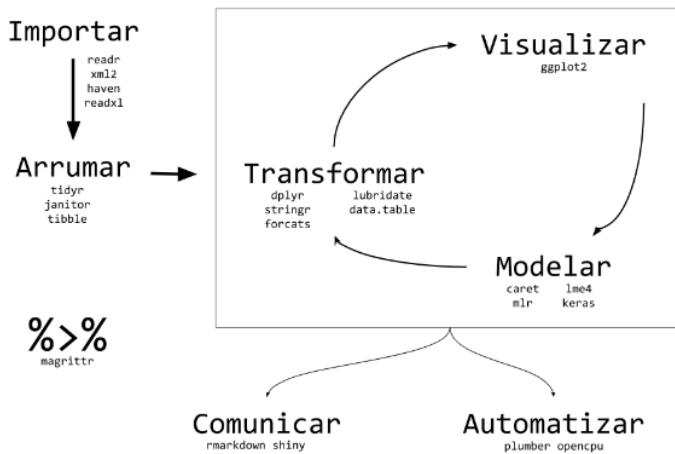


Figure 1: Tidyverse

1.2 Sf

Pacote utilizado para manipulação de objetos do mundo real. Descreve a forma com que esses objetos podem ser armazenados e importados e quais operações geométricas podem ser definidas por eles.

1.3 Tidyterra

1.4 Terra

1.5 Stars

1.6 Tools

1.7 RColorBrewer

1.8 Progress

1.9 Reshape2

1.10 Mapview

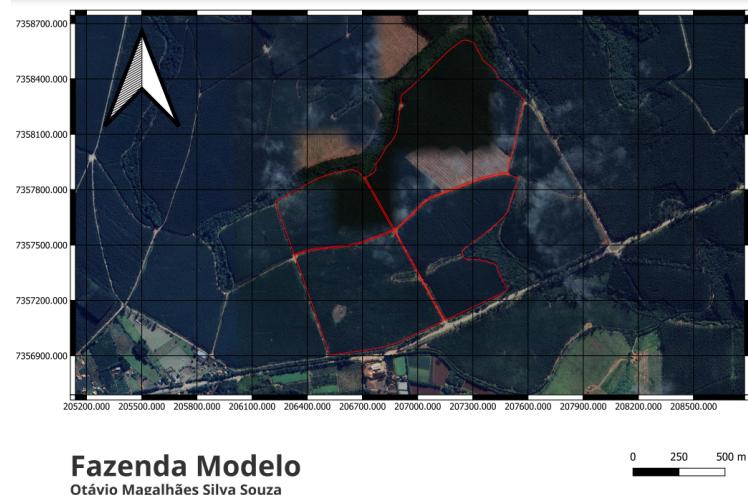
1.11 LidR

1.12 RCSF

1.13 Future

2 Descrição da área

A área a ser estudada como “Fazenda Modelo” localiza-se no município de São Miguel Arcanjo (SP), pode ser identificada pelas coordenadas (-23.86707°, -47.87772°) e possui 129,784 ha, que dividem-se em 4 subtalhões: 301a (18,933 ha), 301d (34,468 ha), 302a (47,602 ha) e 302c (28,781 ha).



Fazenda Modelo
Otávio Magalhães Silva Souza

Figure 2: Mapa da propriedade

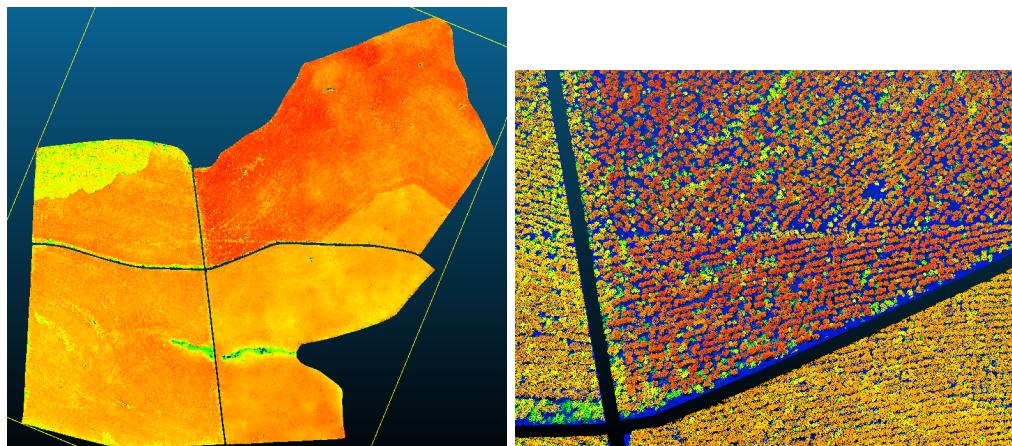


Figure 3: Nuvens LiDAR normalizadas

3 Grid e parcelas já inventariadas

A região foi dividida em 3454 parcelas, onde 2960 delas possuem 400m^2 , enquanto as outras são menores por estarem na borda e abrangerem áreas além da área de interesse. Além disso, 13 das parcelas possuem dados de inventário florestal e podem ser identificadas pelos seguintes Id's: 993, 1526, 1770, 1881, 3165, 3628, 3660, 3730, 5052, 5091, 5106 e 5122.



Figure 4: Parcelas com dados de inventário

4 Conceitos e definições da Dupla Amostragem

A dupla amostragem é composta por duas fases (s1 e s2) - s1: compreende uma gama de variáveis explanatórias para cada ponto pertencente a s1. As variáveis explanatórias derivam de informações auxiliares disponíveis em grande quantidade ao longo da área florestal;

- s2: constitui o inventário terrestre feito num número limitado de subamostras, onde todas pertencem à s1 e fornecem o valor das variáveis de interesse, ex. densidade local.

Exemplo prático e definições (<https://www.ipef.br/publicacoes/scientia/nr108/cap09.pdf>)

- O erro amostral (E%) é calculado a partir do desvio padrão S², que representa a variação de uma série de médias retiradas da população
- A dupla amostragem se divide em duas fases. Na primeira fase há uma maior intensidade amostral e se é medida a variável auxiliar, que deve ser de mensuração facilitada. Na segunda fase a intensidade amostral é menor e se mede a variável de interesse e a variável auxiliar, pois será calculado um estimador de regressão entre as duas. (existe uma diferença entre estimadores de regressão e de razão - estudar)
- A alta correlação entre métricas LiDAR e parâmetros biofísicos da floresta justificam a adoção da tecnologia na primeira fase da DA, além de reduzir a intensidade amostra (custo e trabalho) - Primeira fase: lançamento das parcelas e sobrevôo com o LiDAR
- Etapas (2a fase): mensuração de variáveis das árvores (ex. DAP), medição de algumas árvores para gerar modelo hipsométrico, que será usado para descobrir a altura das outras árvores, cálculo do volume também por modelo e extração do volume da parcela.
- O estimador de regressão (estudar) do VTCC deve relacionar a(s) variável(es) auxiliar(es) e também compreende a outros índices (constantes para cada região)
- O desvio padrão, intervalo de confiança etc são feitos depois

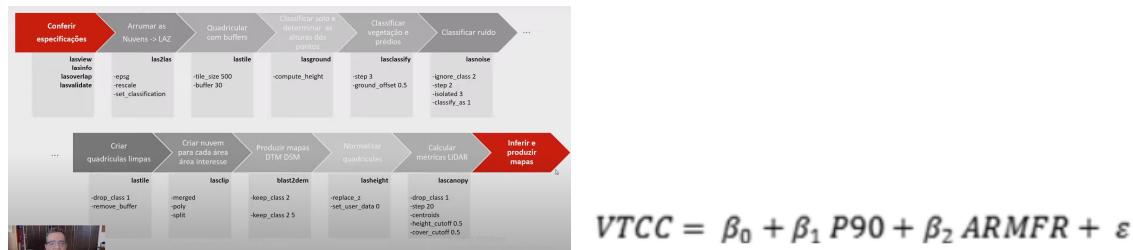


Figure 5: Fluxograma DA e estimador de regressão para VTCC

5 Processamento da nuvem LiDAR

1. Conferir especificações
 - i. Verificar se a nuvem gerada se enquadra no que é esperado: densidade de pontos, área escaneada, classificação de pontos etc.
2. Arrumar nuvens
 - i. Verificação do sistema de coordenadas utilizado e atributos de latitude, longitude e elevação
3. Quadricular
 - i. Repartição da nuvem em pedaços menores a fim de ganhar eficiência no processamento em paralelo

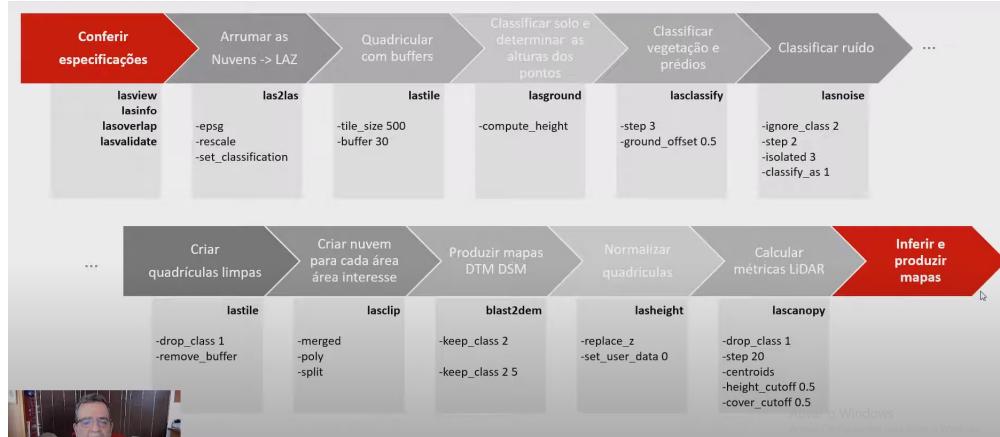


Figure 6: Fluxograma do processamento da nuvem LiDAR

4. Definir solo e determinar altura dos pontos
 - i. Assegura e certifica a classificação do solo por diferentes métodos
5. Classificar outros pontos e ruídos
 - i. Separar os pontos restantes por sua natureza (edificações, corpos d'água etc)
6. Quadrículas limpas
 - i. remoção dos ruídos após classificação
7. Produção de mapas DTM (elevação do solo), DSM (elevação do solo + vegetação) e DEM (elevação da vegetação)
8. Normalização da nuvem
 - i. Joga tudo pro mesmo plano
9. Cálculo de métricas LiDAR
 - i.
 - ii. O p90 (percentil 90) representa a altura em que 90% dos pontos da nuvem LiDAR se encontram abaixo desta. Cada quadrícula possui o seu próprio p90 e ele tende a apresentar boa correlação com a realidade de campo. A intenção de existirem diversas métricas LiDAR é a de ver quais delas possuem boa correlação com os atributos de interesse

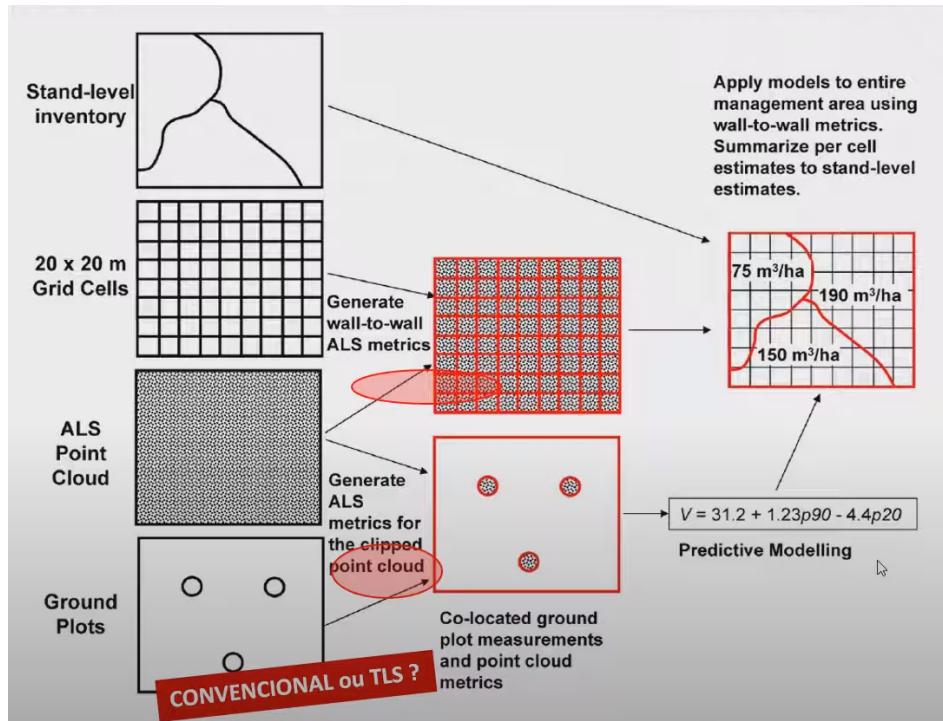


Figure 7: Esquematização do processo de inventário com ALS e informação terrestre

5.1 Fluxograma e etapas Dupla amostragem

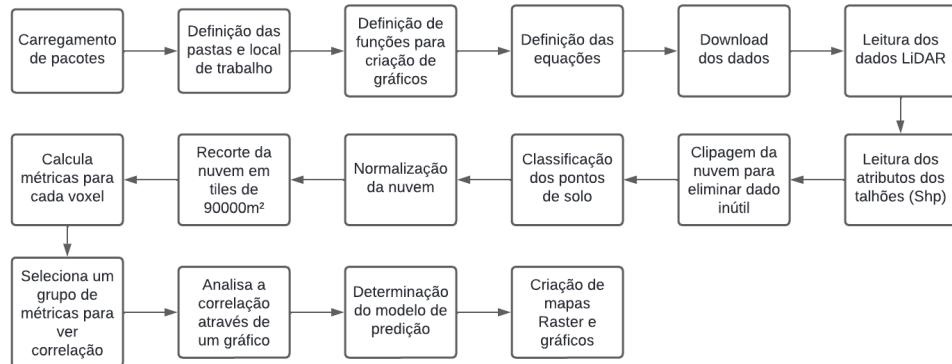


Figure 8: Fluxograma das etapas de processamento de dados LiDAR para fins de inventário florestal

1. Carregamento dos pacotes

- Diversos são os pacotes carregados. Os nomes e a utilidade de cada um estão descritos na primeira seção do documento.

2. Definição das pastas e local de trabalho

- GitHub
 - i. C: - pasta raiz
 - ii. GitRepo - diretório em que estão agrupados os arquivos a serem upados no GitHub
 - iii. PRJ_FAZENDAMODELO - pasta do projeto
 - iv. RMD - código e arquivos utilizados na redação do presente documento
 - v. RESULTADOS - plot da matriz de correlação
 - vi. BATCHR - arquivo tipo R com os scripts utilizados no pré e pós-processamento dos dados
 - vii. SAIDASSIG - arquivos gerados no QGIS
 - viii. SHAPES - arquivos de entrada para uso no SIG
- LiDAR
 - i. C: - pasta raiz
 - ii. LiDAR - agrupa todos as nuvens de pontos utilizadas no script
 - iii. PRJ_FAZENDAMODELO - pasta do projeto
 - iv. NUVENS - onde se localizam as nuvens de pontos
 - v. A13 - reúne as nuvens do ano de 2013
 - vi. TALHOES - nuvem segregada por talhão
 - vii. NoNORM - nuvens com solo classificado
 - viii. SiNORM - nuvens com solo classificado e normalizadas
 - ix. RSTR_qua - raster apresentando a estimativa da variável de interesse para cada talhão

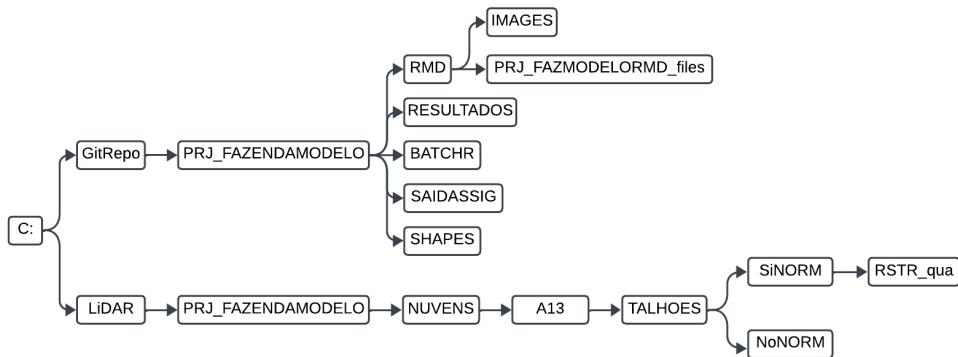


Figure 9: Organização dos diretórios

3. Definição das funções para criação dos gráficos

4. Definição das equações (estudar quais são)

5. Download e leitura dos dados LiDAR
 - i. Ao todo foram baixadas 6 nuvens de pontos LiDAR, que antes do processamento encontravam-se da seguinte maneira:

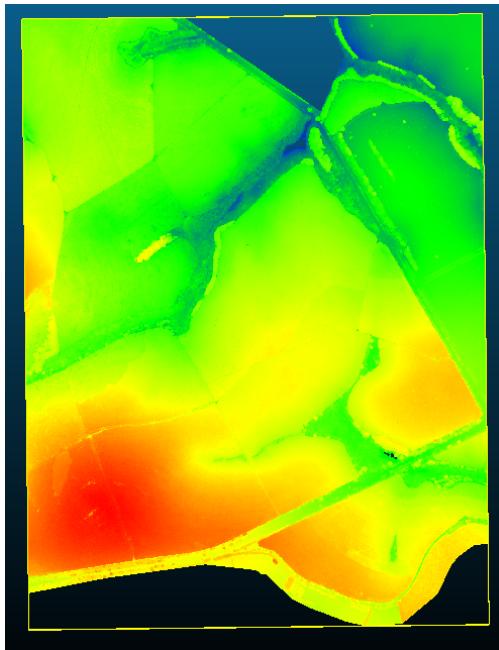


Figure 10: Nuvens de pontos LiDAR pré-processadas

- ii. As nuvens foram baixadas pelo seguinte link: https://github.com/FlorestaR/dados/blob/main/5_LIDARF/Modelo/CLOUDS/

- . Download e leitura dos dados em Shapefile
- i. Os shapes foram baixados pelo seguinte link: https://github.com/FlorestaR/dados/blob/main/5_LIDARF/Modelo/SHAPEs

| | |
|------------|---------------|
| Fazenda | A |
| ID_Talhao | 301 |
| CODTALHAO | 52I214 |
| CHAVE | 52I214_6 |
| CHAVE2 | P301410062 |
| NUMPARCELA | 6 |
| NUMMEDICAO | 2 |
| ANOREF | 2013 |
| DATAREALIZ | 4/18/2013 |
| IDINV | 3,70000 |
| AREAPARCEL | 399,73000 |
| LAT_UTM23S | 206632,09800 |
| LONG_UTM23 | 7357127,11200 |
| MHDOM | 19,37000 |
| VTCC | 199,85000 |
| VCCC | 196,85000 |
| SUBTALHAO | 301d |

Figure 11: Dados contidos nas parcelas inventariadas

. Clipagem da nuvem para eliminação de dados indesejados (mostrar nuvem antes e depois)

Antes da clipagem: 68mi pontos

Após clipagem: 18mi pontos

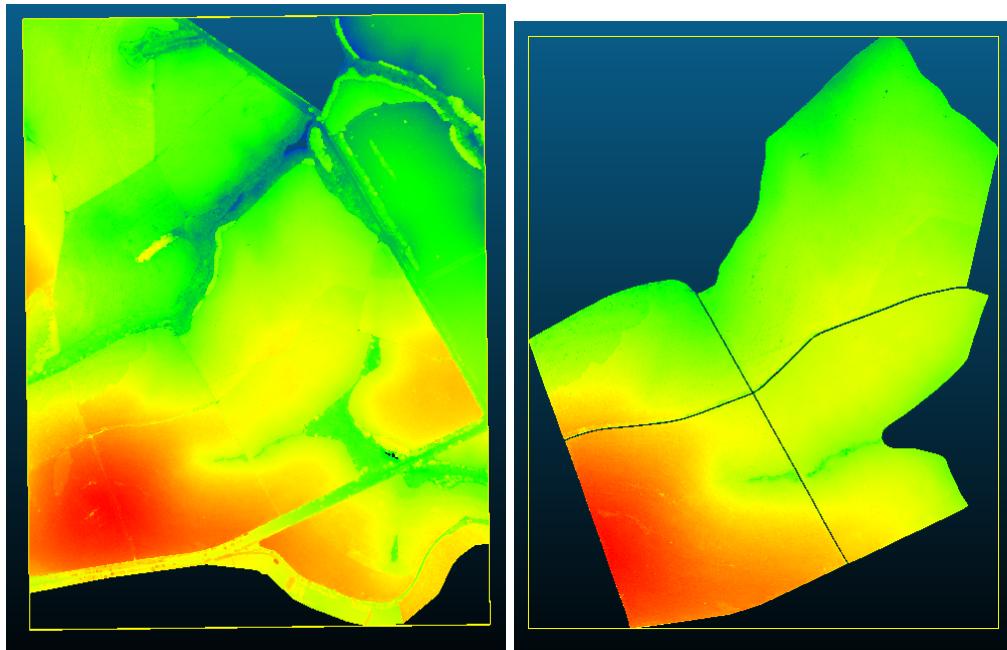


Figure 12: Comparativo entre as nuvens de pontos antes e após a clipagem

. Classificação

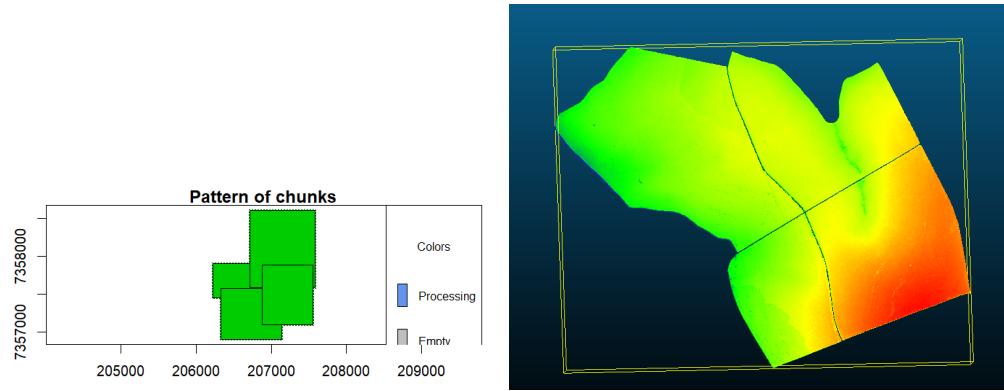


Figure 13: Processamento e resultado da classificação de solo das nuvens LiDAR

. Normalização i. A etapa de normalização tem por finalidade nivelar toda a nuvem e é um passo que está diretamente correlacionado à classificação do solo.

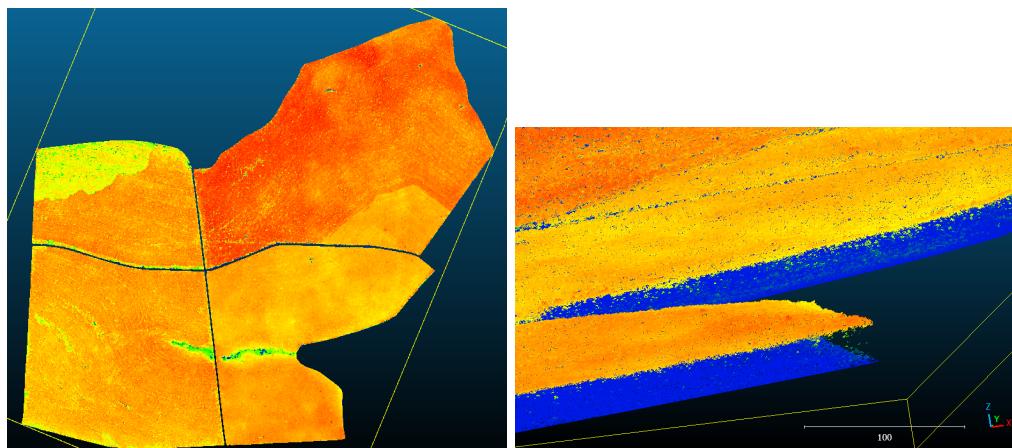


Figure 14: Resultado da normalização das nuvens

. Recorte da nuvem em tiles 300x300m O recorte da nuvem em tiles menores tme a função de facilitar o processamento dos dados, tornando-o mais rápido e dinâmico

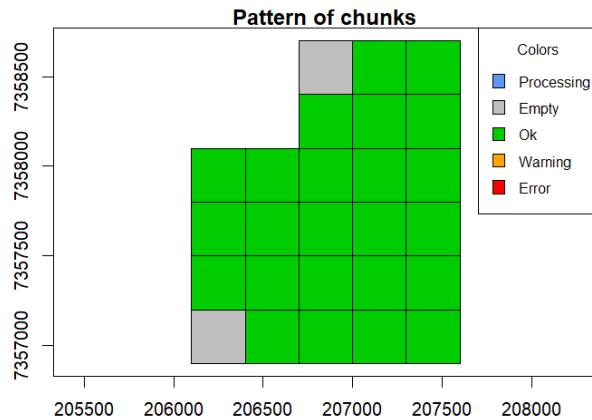


Figure 15: Retile da nuvem em quadrados de 300x300m

11. Cálculo de métricas para cada voxel (explicar voxel)

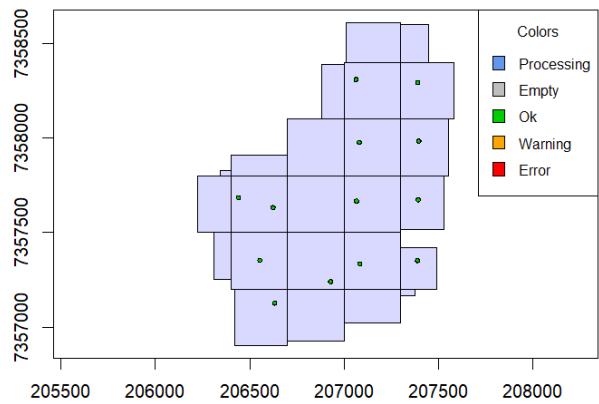


Figure 16: Cálculo das métricas para cada voxel

. Seleção de um grupo de métricas para estudo de correlação

| | VTCC | MHDOM | IDINV | Zmean | zq45 | zq75 | zq95 | zpcum2 | zpcum4 | zpcum6 | pzabovezmean | pzabove2 |
|----|------|-------|-------|-------|------|------|------|--------|--------|--------|--------------|----------|
| | <db> | <db> | <db> | <db> | <db> | <db> | <db> | <db> | <db> | <db> | <db> | <db> |
| 1 | 247. | 22.2 | 5.2 | 12.6 | 13.1 | 23.2 | 24.2 | 1.08 | 1.08 | 1.55 | 55.1 | 55.2 |
| 2 | 204. | 19.9 | 5.2 | 11.8 | 2.98 | 21.9 | 22.7 | 8.81 | 8.94 | 9.65 | 54.0 | 56.3 |
| 3 | 285. | 27.4 | 5.2 | 13.2 | 19.8 | 22.1 | 23.7 | 1.14 | 1.34 | 1.59 | 61.0 | 61.3 |
| 4 | 249. | 24.5 | 5.2 | 14.1 | 20.4 | 22.7 | 24.2 | 1.21 | 3.21 | 3.94 | 63.8 | 64.7 |
| 5 | 335. | 32.7 | 5.2 | 12.5 | 0 | 26.0 | 27.2 | 7.19 | 8.62 | 16.8 | 49.3 | 51.5 |
| 6 | 307. | 28.1 | 5.2 | 12.3 | 0.37 | 24.8 | 25.9 | 9.33 | 9.50 | 10.3 | 49.9 | 51.0 |
| 7 | 200. | 19.4 | 3.7 | 13.2 | 19.8 | 22.1 | 23.7 | 1.14 | 1.34 | 1.59 | 61.0 | 61.3 |
| 8 | 219. | 19.9 | 3.7 | 14.1 | 20.4 | 22.7 | 24.2 | 1.21 | 3.21 | 3.94 | 67.8 | 64.7 |
| 9 | 155. | 14.8 | 3.7 | 10.7 | 15.3 | 18.8 | 20.8 | 2.47 | 3.20 | 3.52 | 57.9 | 59.3 |
| 10 | 254. | 23.4 | 3.7 | 14.7 | 21.4 | 23.6 | 25.5 | 4.86 | 5.56 | 5.56 | 63.1 | 65.3 |
| 11 | 247. | 22.1 | 3.7 | 13.7 | 20.3 | 23.0 | 24.5 | 10.6 | 11.0 | 11.2 | 60.1 | 63.8 |
| 12 | 238. | 23.9 | 5.2 | 14.0 | 21.2 | 22.9 | 24.2 | 2.04 | 2.19 | 2.55 | 62.0 | 62.8 |
| 13 | 276. | 25.8 | 5.2 | 15.9 | 23.2 | 24.4 | 25.2 | 2.55 | 2.55 | 2.55 | 65.9 | 66.4 |

Figure 17: Tabela com as métricas escolhidas

. Análise da correlação por meio de gráfico (falar do gráfico)

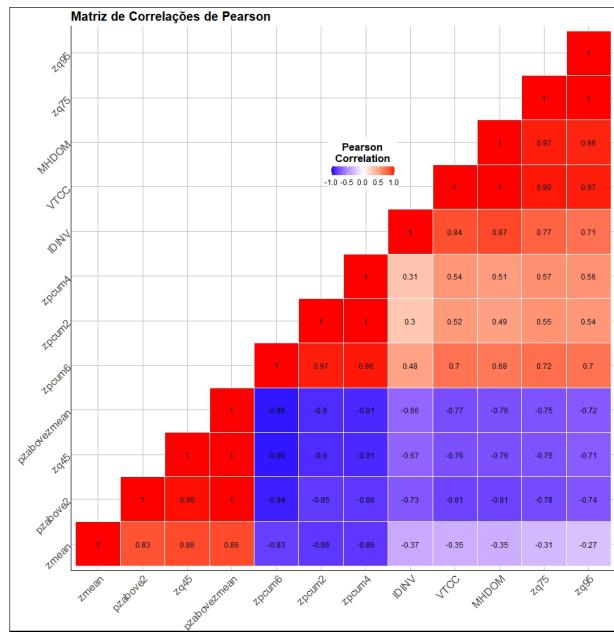


Figure 18: Resultado da análise de correlação

. Determinação do modelo de predição

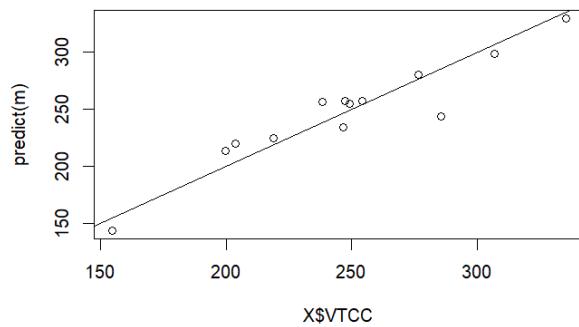


Figure 19: nao sei

15

. Criação de mapas raster e gráficos

6 Fluxograma e etapas Tripla amostragem

Composta por 3 fases, s0, s1 e s2:

- O princípio básico é o de que as variáveis explanatórias derivadas das informações auxiliares estão disponíveis em duas frequências diferentes. A fase s0 fornece informações sobre toda a área, enquanto s1 possui dados adicionais de amostras de s0. Logo, a partir da informação terrestre coletada de um número x de parcelas de campo (1 camada de informação) é possível aferir sobre informações adicionais para outras parcelas a partir do uso de preditores (2 camadas de informação) e, por fim, o LiDAR coleta dados sobre todas as parcelas (3 camadas de informação).

- Logo, a motivação por trás da TA é a de que a gama de informações de s1 adiciona alto poder preditivo às variáveis disponíveis para todas as parcelas da área (s0)

Journal of Statistical Software

5

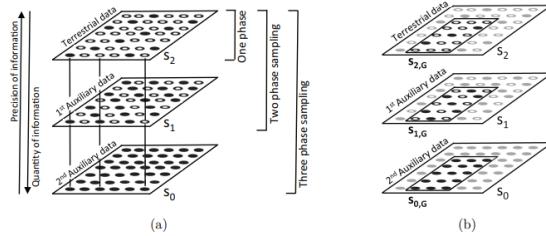


Figure 2: (a) Concept of multiphase sampling. The square represents the forest area for which an inventory is being conducted. The points denote the sample locations x . Filled points indicate available information. (b) Illustration of the small area estimation problem.

Figure 20: Esquema TA com Lidar

Estimativas de pequenas áreas (small area estimation)

- Para sub áreas da floresta onde há pouca informação terrestre, como em G na figura acima. Para essas áreas, o uso da amostragem multifásica pode ser mais eficiente, já que utilizam um número reduzido de parcelas em campo para chegar à mesma precisão da ACE e ACS. Por outro lado, a sub área em questão pode ser pequena demais para justificar a adoção de um modelo de regressão separado, uma vez que este pode resultar em um intervalo de confiança indesejadamente abrangente;

- A ideia, então, é a de utilizar toda a riqueza de informações presente nas amostras de S2 para ajustar o modelo (equação utilizada para o cálculo de uma variável de interesse) e aplicá-lo para a sub área em questão.

- O potencial viés que surge da aplicação do modelo na sub área é corrigido pelo uso de modelos residuais empíricos derivados da área amostrada em campo. (não entendi direito essa parte) - trecho no texto: The potential bias of applying that model in G is then corrected for by using the empirical model residuals derived from that small area.

- Caso não existam parcelas de campo na sub área, então deve-se aceitar o viés na estimativa e no modelo. São essas as estimativas sintéticas, mas apesar do viés é possível calcular a sua variação.

DESIGN-BASED VS MODEL-DEPENDENT APPROACH (ver dps)

- Model-dependent approach: as parcelas amostrais são fixas e as observações retiradas desses locais são assumidas como variáveis aleatórias, assim como a floresta assume o papel de ser o meio realizador desse processo estocástico (????? o que isso significa?). Embora os locais das parcelas possam ser arbitrariamente escolhidos, o modelo deve escrever adequadamente o processo estocástico, a fim de garantir resultados parciais. (????)

- Os estimadores do forestinventory se baseiam na design-based approach. Design-based approach baseia-se na randomização dos locais de amostra de forma uniforme e independente. Logo, a floresta por si só e qualquer valor de densidade local de x (pertencente a) F (floresta) são fixos e não um resultado de um processo estocástico (estudar). (????)

- Na estrutura do Model-dependent approach as parcelas são fixas e resultam em um processo estocástico GPT: what is the difference between design-based approach and model-dependent approach?

7 Regressão linear simples

Modelar a relação entre duas variáveis e se expressa na forma de uma reta no plano cartesiano.

$$y = \beta_0 + \beta_1 x + \varepsilon$$

Figure 21: Equação base da regressão linear simples

- Y = variável resposta ou variável dependente
- X = variável independente, regressora ou explicativa
- ε = erro. Pode ser que a Y seja explicada por X , mas falte algo que explique essa variabilidade. É a flutuação aleatória que ocorre ao tentar explicar Y por X . Seja por imperfeição do modelo, erros na mensuração ou outras variáveis fora de controle.

8 Arborimetria preditiva

Em florestas plantadas há grande relação entre o DAP e a altura da árvore. O procedimento de medição de campo é o seguinte: 1 - DAP de todas as árvores da parcela; 2 - altura total de uma amostra da parcela; 3 - estima-se a relação entre o DAP e a altura total da amostra; 4 - utiliza essa relação para prever a altura de todas as árvores da parcela. A seleção das árvores da amostra pode ser feita de forma aleatória ou sistemática. A seleção aleatória garante maior confiabilidade, enquanto a sistemática é mais simplificada (ex. 1 árvore a cada 15). A *relação hipsométrica é a relação média entre o DAP e a altura das árvores individuais. A curva sigmoidal é o melhor modelo para a relação média entre o DAP e a altura

Um modelo é linear quando cada termo é uma constante ou o produto de um parâmetro e uma variável preditora. Uma equação linear é construída adicionando os resultados para cada termo. Modelos lineares seguem esse padrão: Resposta = constante + parâmetro * preditor + ... + parâmetro * preditor. Caso contrário a equação é não-linear.

O postulado de Cotta define que o volume de uma árvore depende do seu diâmetro, altura e forma. Todas as árvores de mesmo diâmetro, altura e forma possuem o mesmo volume. Isso serve para o lenho de tronco único. O fator forma pode ser visto como um índice quantitativo que representa a razão entre o volume sólido e cilíndrico de uma árvore. O volume sólido de uma amostra da parcela pode ser medido por cubagem rigorosa. O fator forma é igual ao somatório dos volumes sólidos dividido pelo somatório dos volumes cilíndricos ($[\pi/4] * d^2 * h$).

O volume de uma dada árvore é o resultado da multiplicação entre o fator de forma e o volume cilíndrico do indivíduo. A seleção das árvores deve ser aleatória, sem preocupação em selecionar árvores de todas as classes de DAP e altura. **Vamos usar o LiDAR TLS para estimar o diâmetro e a forma?**

O volume pode ser calculado em função de diferentes variáveis preditoras, como a altura, DAP, altura até a primeira ramificação ou outro. Equações de dupla entrada ou equações padrão são aquelas que utilizam o DAP e a altura para cálculo do volume.

A relação entre o volume sólido e o volume cilíndrico tende à uma relação de proporcionalidade (a curva que os representa passa pela origem). Com a adoção do volume comercial, são consideradas as árvores a partir de um tal diâmetro, o que afasta a curva da origem e, portanto, não evidencia uma relação proporcional.

É possível modelar o fator de forma como uma função do DAP ao invés de encontrá-lo a partir da relação entre o DAP e a altura. É adequado para quando a forma da árvore muda de acordo com o DAP.

| Quadro 6.1 Modelos não lineares de relação hipsométrica | |
|---|--|
| Modelo | Forma preditiva e de estimativa |
| Chapman-Richards | $h = \beta_0 [1 - \exp(-\beta_1 \cdot d)]^{\beta_2} + \varepsilon$ |
| Weibull | $h = \beta_0 [1 - \exp(-\beta_1 \cdot d^{\beta_2})] + \varepsilon$ |
| Monomolecular | $h = \beta_0 [1 - \beta_1 \cdot \exp(-\beta_2 \cdot d)] + \varepsilon$ |
| Gompertz | $h = \beta_0 \exp[-\beta_1 \cdot \exp(-\beta_2 \cdot d)] + \varepsilon$ |
| Logístico | $h = \beta_0 / [1 + \beta_1 \cdot \exp(-\beta_2 \cdot d)] + \varepsilon$ |

h - altura total das árvores individuais;
d - diâmetro à altura do peito (DAP);
 $\beta_0, \beta_1, \beta_2$ - parâmetros a serem estimados, $\beta'_0 = \exp(\beta_0)$;
 ε - componente estocástico do modelo.

Figure 22: Quadro de modelos não lineares (esquerda) e lineares (direita)

| Quantificação de recursos florestais | |
|--|---|
| Quadro 6.2 Modelos lineares de relação hipsométrica. A forma da estimativa se refere à maneira utilizada para estimar os parâmetros dos modelos, enquanto que forma preditiva se refere à maneira como o modelo é aplicado para se prever a altura | |
| Modelo | Forma de estimativa |
| Hiperbólicos | |
| 1 | $h = \beta_0 + \beta_1 \cdot d + \varepsilon$ |
| 2 | $h = \beta_0 + \beta_1 \cdot d + \beta_2 \cdot d^2 + \varepsilon$ |
| 3 | $h = \beta_0 + \beta_1 \cdot \ln(d) + \beta_2 \cdot \ln(d)^2 + \varepsilon$ |
| 4 | $h = \beta_0 + \beta_1 \cdot d + \beta_2 \cdot d^2 + \beta_3 \cdot d^3 + \varepsilon$ |
| 5 | $h = \beta_0 + \beta_1 \cdot d + \beta_2 \cdot d^2 + \beta_3 \cdot d^3 + \varepsilon$ |
| Potência | |
| 1 | $h = \beta_0 + \beta_1 \cdot \ln(d) + \varepsilon$ |
| 2 | $h = \beta_0 + \beta_1 \cdot \ln(d)^2 + \varepsilon$ |
| 3 | $h = \beta_0 + \beta_1 \cdot \ln(d) \cdot \ln(d+1) + \varepsilon$ |
| Exponencial (Schumacher) | |
| 4 | $h = \beta_0 + \beta_1 \cdot \ln(d) + \varepsilon$ |
| Semilogarítmico | |
| 5 | $h = \beta_0 + \beta_1 \cdot \ln(d) + \varepsilon$ |

h - medida da altura total das árvores individuais;
 \hat{h} - altura total predita;
d - diâmetro à altura do peito (DAP);
 $\beta_0, \beta_1, \beta_2$ - parâmetros a serem estimados, $\beta'_0 = \exp(\beta_0)$;
 ε - componente estocástico do modelo;

In - logaritmo natural ou neperiano

Figure 23: Quadro de modelos não lineares (esquerda) e lineares (direita)

Quadro 6.3 Modelos de equação padrão de volume ou equação de dupla entrada. Modelos marcados com asterisco (*) são modelos não lineares. Todos os modelos são apresentados na sua forma de estimação

Modelos de variável combinada

| | | |
|-------|-----------|---|
| VC-1 | Spurr | $v_i = \beta_0 + \beta_1(d_i^2 \cdot h_i) + \varepsilon_i$ |
| VC-2 | log-Spurr | $\ln(v_i) = \beta_0 + \beta_1 \cdot \ln(d_i^2 \cdot h_i) + \varepsilon_i$ |
| VC-3* | | $v_i = \beta_0 + \beta_1(d_i^2 \cdot h_i)^{\beta_2} + \varepsilon_i$ |

Modelos de potência variável

| | | |
|-------|-------------------|--|
| PV-1 | Schumacher-Hall | $\ln(v_i) = \beta_0 + \beta_1 \cdot \ln(d_i) + \beta_2 \cdot \ln(h_i) + \varepsilon_i$ |
| PV-2* | | $v_i = \beta_0 \cdot d_i^{\beta_1} \cdot h_i^{\beta_2} + \varepsilon_i$ |
| PV-3* | | $v_i = \beta_{00} + \beta_0 \cdot d_i^{\beta_1} \cdot h_i^{\beta_2} + \varepsilon_i$ |
| PV-4 | Baden-Württemberg | $\ln(v_i) = \beta_0 + \beta_1 \cdot \ln(d_i) + \beta_2 \cdot \ln^2(d_i) + \beta_3 \cdot \ln(h_i) + \beta_4 \cdot \ln^2(h_i) + \varepsilon_i$ |
| PV-5* | | $v_i = \beta_0 \cdot d_i^{\beta_1+\beta_2-\ln(d_i)} \cdot h_i^{\beta_3+\beta_4+\ln(h_i)} + \varepsilon_i$ |

Modelos de harmonização DAP-altura

| | | |
|-----|------------------|---|
| H-1 | Stoate | $v_i = \beta_0 + \beta_1 \cdot d_i^2 + \beta_2(d_i^2 \cdot h_i) + \beta_3 \cdot h_i + \varepsilon_i$ |
| H-2 | Meyer | $v_i = \beta_0 + \beta_1 \cdot d_i + \beta_2 \cdot d_i^2 + \beta_3(d_i \cdot h_i) + \beta_4(d_i^2 \cdot h_i) + \beta_5 \cdot h_i + \varepsilon_i$ |
| H-3 | Meyer modificada | $v_i = \beta_0 + \beta_1 \cdot d_i + \beta_2 \cdot d_i^2 + \beta_3(d_i \cdot h_i) + \beta_4(d_i^2 \cdot h_i) + \varepsilon_i$ |
| H-4 | Näslund | $v_i = \beta_0 + \beta_1 \cdot d_i^2 + \beta_2(d_i^2 \cdot h_i) + \beta_3(d_i \cdot h_i^2) + \beta_4 \cdot h_i^2 + \varepsilon_i$ |

Modelos de fator forma

| | | |
|-------|--------|---|
| FF-1* | Takata | $v_i = (d_i^2 \cdot h_i) / [\beta_0 + \beta_1 \cdot d_i] + \varepsilon_i$ |
| FF-2* | Honer | $v_i = d_i^2 / [\beta_0 + \beta_1/h_i] + \varepsilon_i$ |

i – índice que indica as árvores na amostra destrutiva ($i = 1, 2, \dots, n$);

v_i – volume sólido (total ou comercial);

d_i – DAP;

h_i – altura total ou comercial;

\ln – logaritmo natural ou neperiano;

ε_i – componente estocástico do modelo.

Figure 24: Equação base da regressão linear simples