

Controle e Garantia de Qualidade de Dados LiDAR para Estimativa de Biomassa Florestal

Mauro Lúcio Rodrigues de Assis ¹

Jean Pierre Ometto ¹

Rafael Santos ²

Eric Görgens ³

Luciane Sato ¹

Francisca Rocha de Souza Pereira ¹

Roberta Zecchini Cantinho ⁴

Pedro Valle de Oliveira ⁵

¹ Instituto Nacional de Pesquisas Espaciais - INPE

Av. dos Astronautas, 1758 - Caixa Postal 515

12201-970 - São José dos Campos - SP, Brasil

assismauro64@gmail.com; jean.ometto@inpe.br, lucianesato@gmail.com, franrspereira@gmail.com

² Instituto Nacional de Pesquisas Espaciais - INPE

Av. dos Astronautas, 1758 - Caixa Postal 515

12201-970 - São José dos Campos - SP, Brasil

rafael.santos@inpe.br

³ Universidade Federal dos Vales do Jequitinhonha e Mucuri

Campus JK, Departamento de Engenharia Florestal

39100-000 - Diamantina - MG, Brasil

e.gorgens@gmail.com

⁴ Programa das Nações Unidas para o Desenvolvimento (PNUD),
Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC)

CEP 70720-610 - Brasília - DF, Brasil

rzcantinho@gmail.com

⁵ South Dakota State University - Geospatial Sciences Center of Excellence

1021 Medary Ave, Wecota Hall 225

Brookings, SD, USA - 57007

pedrovco@gmail.com

Resumo. Estimar a biomassa florestal de um bioma é importante para vários propósitos, como calcular o balanço de carbono, que é uma demanda aos países participantes do Programa das Nações Unidas Sobre a Redução do Desmatamento e Degradação Florestal em Países em Desenvolvimento (UN-REDD). Uma forma de se estimar biomassa de florestas é usar dados produzidos a partir da tecnologia LiDAR aerotransportado (ALS), usando-se para isso modelos matemáticos que correlacionam métricas calculadas a partir das nuvens de pontos LiDAR coletadas com a biomassa existente. Para maximizar o retorno no uso dessa técnica, é necessário que os dados LiDAR coletados apresentem características de qualidade de maneira a que imperfeições nos dados coletados inerentes ao próprio processo de coleta destes não impactem a qualidade das estimativas obtidas. Esse trabalho propõe: a) um conjunto de parâmetros a serem especificados quando da contratação de voos LiDAR para uso em estimativas de biomassa florestal b) um protocolo de validação dos dados recebidos a fim de confrontá-los com os parâmetros de voo e encontrar outras imperfeições significativas. Este protocolo é apresentado através de um conjunto de programas de computador disponível em <<https://github.com/ebacst/pyLiDARForest>>.

Palavras-chave: LiDAR, validação, controle de qualidade, biomassa florestal, Amazônia, ALS.

Abstract. Estimating the biomass of a forestry biome is important for a number of purposes, such as calculating the carbon balance demanded by countries participating in the United Nations Program on Reducing Deforestation and Forest Degradation in Developing Countries (UN-REDD). One of the forms used to estimate forest biomass is to use the data produced from the airborne LiDAR (ALS) technology, using mathematical models that correlate metrics calculated from the collected point clouds with the existing biomass. To maximize the return on the use of this technique, it is necessary that the collected LiDAR data present quality characteristics so that imperfections in the collected data inherent in the data collection process itself do not impact the quality of the estimates obtained. This work proposes: a) a set of parameters to be specified when contracting LiDAR flights for use in biomass estimations b) a validation protocol of the received data not only by comparing them with the flight parameters but also by helping to find other imperfections. This protocol is presented through a set of computer programs available at <<https://github.com/ebaccst/pyLiDARForest>>.

Keywords: LiDAR, quality assurance and control, forest biomass, Amazon, ALS.

1. Introdução

As florestas desempenham um importante papel no contexto das mudanças climáticas, uma vez que, durante o processo de fotossíntese, sequestram o gás carbônico da atmosfera segundo Barbosa et al. (2014). Neste sentido, a quantificação da biomassa e do carbono presentes nos diferentes compartimentos da floresta possibilita estimar o carbono estocado em sua estrutura conforme IPCC (2006). No entanto, as estimativas de biomassa florestal estão associadas a níveis elevados de incerteza que refletem nas estimativas de emissões de gases de efeito estufa por mudança de uso e cobertura da terra, em Ometto et al. (2014).

Neste contexto, um dos métodos que tem se mostrado efetivo é o uso da tecnologia de altimetria laser ou *Light Detection and Rangig* (LiDAR), descrita em Wehr, A.; Lohr, U (1999). O LiDAR é um sensor que permite medir a distância entre este e o alvo, através da multiplicação entre a velocidade da luz e o intervalo de tempo que o pulso Laser leva para alcançar o objeto, também segundo Wehr, A.; Lohr, U (1999). Devido à alta intensidade de pontos e precisão dessa tecnologia, ela tem sido amplamente usada na representação em três dimensões (3D) de paisagens, estimativa de biomassa florestal, conforme Sassan e Saatchi (2011), Longo et al. (2016), Chen et al. (2016), entre outros.

Um dos usos do LiDAR é o sistema de *Airborne Laser Scanning* (ALS), onde o sensor é instalado em um avião e apontado para o solo, emitindo pulsos laser e registrar seu retorno ao longo do voo da aeronave, indicado em Axelson, P. (1999). Durante o voo, a posição precisa do avião no espaço é coletada por um GNSS (do inglês Global Navigation Satellite System) e corrigido com base no deslocamento da aeronave por uma Unidade de Referência Inercial (do inglês *Inertial Measurement Unit*, IMU). Os dados coletados são posteriormente processados para se obter uma nuvem de pontos georreferenciados. Devido à alta densidade de pontos e precisão, esta tecnologia tem sido amplamente utilizada na estimativa de biomassa florestal.

Após o processamento dos dados LiDAR coletados, estes podem ser armazenados em diferentes formatos, sendo os mais comuns: i) LASer (LAS), definido e apoiado pela Sociedade Americana de Fotogrametria e Sensoriamento Remoto (do inglês American Society for Photogrammetry & Remote Sensing – ASPRS) e ii) LAZ, produto do software gratuito e de código aberto Rapidlasso GmbH, que compacta os arquivos LAS sem alteração da informação. As rotinas de validação apresentadas nesse estudo foram elaboradas para suportar ambos os formatos (LAS e LAZ), podendo ser facilmente adaptadas para outros formatos.

Geralmente, a aquisição de dados LiDAR a partir de ALS é efetuada por empresas privadas mediante contratos de prestação de serviços. Sendo assim, é necessário validar a qualidade dos dados recebidos.

O presente trabalho foi desenvolvido processando nuvens de pontos coletadas sobre a Amazônia. Contudo, o uso do LiDAR tem sido cada vez maior no estudo dos mais diversos biomas e, apesar desse sensor normalmente ter muito de sua energia absorvida por lâmina

d'água, acreditamos no seu potencial de uso no Pantanal pelo fato de que grande parte dessa região não permanece alagada por todo o ano.

2. Objetivo

O objetivo desse trabalho foi desenvolver uma rotina de recebimento e validação de dados (nuvens de pontos) LiDAR. Essa rotina deve garantir que os dados validados possam ser usados com segurança para os fins a que se destinarem, sem que eventuais imperfeições interfiram nos resultados obtidos.

3. Material e Métodos

Os dados utilizados neste estudo foram coletados no âmbito do projeto Monitoramento Ambiental por Satélite no Bioma Amazônia, subprojeto Estimativa de Biomassa da Amazônia (EBA)¹, conduzido pelo Centro de Ciência do Sistema Terrestre (CCST) do Instituto Nacional de Pesquisas Espaciais (INPE). O intuito do projeto EBA é gerar um mapa de biomassa e carbono para o bioma Amazônia e, para tanto, foram realizados voos ALS em 1000 transectos distribuídos aleatoriamente na região (**Figura 1**). Cada um desses transectos possui uma largura de aproximadamente 300m e comprimento de 12.500m (Figura 2). Os dados foram coletados por uma empresa terceirizada no formato full waveform e entregues em formato LAS versão 1.4. Para fins deste estudo, foram utilizados somente os dados dos primeiros 624 transectos entregues em sua forma discreta, suficientes para a estimativa de biomassa.



Figura 1 - Distribuição dos transectos ALS no bioma Amazônia (em vermelho). Fonte: projeto EBA.

1 Maiores informações sobre essa iniciativa podem ser encontradas em: <http://ebaccst.wixsite.com/blog> <http://www.fundoamazonia.gov.br/pt/home/> e <http://www.ccst.inpe.br/projetos/eba-estimativa-de-biomassa-na-amazonia/>

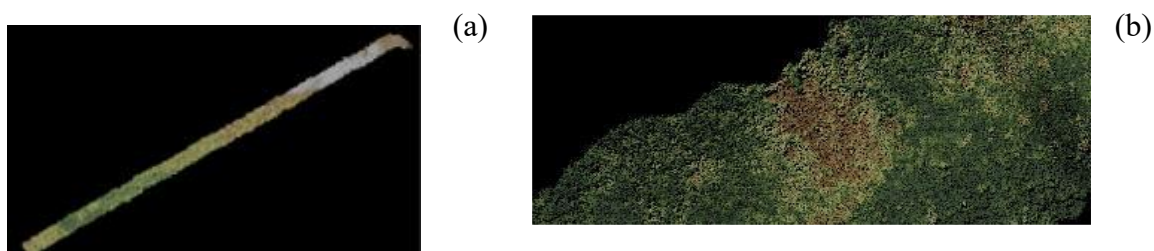


Figura 2 - (a) Imagem de nuvem LiDAR, com 12.500m de comprimento e 300m de largura, colorida em função da coordenada Z (altitude); (b) Zoom em parte do transecto.

Primeiramente, foram construídas, em linguagem Python, programas de computador que processassem as nuvens recebidas (estas poderiam estar em formato LAS ou LAZ) e gerassem um arquivo .csv² contendo os resultados da validação. As rotinas desenvolvidas usaram a biblioteca Laspy³ desenvolvida especificamente para ler, modificar e criar dados LiDAR. Essencialmente, a biblioteca laspy fornece uma classe para acessar um leitor de parâmetros LiDAR e obter uma estrutura de matriz tri-dimensional com os dados LiDAR a serem validados. O conjunto de rotinas também usa o **Fusion**, software livre disponibilizado pelo Serviço Florestal dos Estados Unidos⁴.

O processo de validação das nuvens de pontos LiDAR checkou oito diferentes itens:

- 1) Assinatura do arquivo ou LAS/LAZ: todos os arquivos LAS/LAZ devem possuir uma assinatura em seus primeiros quatro bytes;
- 2) Versão do arquivo LAS/LAZ: há quatro versões (1.1, 1.2, 1.3, 1.4) é importante checar se é a versão estabelecida em contrato, porque as versões mais antigas não contém informações que podem ser relevantes.
- 3) Número de retornos: verifica se o número de retornos especificado no cabeçalho LAS/LAZ corresponde ao número de retornos que realmente existe no arquivo;
- 4) Valores mínimo e máximo: compara os valores mínimo e máximo para as coordenadas X, Y e Z no cabeçalho do LAS/LAZ com os valores especificados nos registros dos retornos;
- 5) Densidade global de retornos: verifica se o número de retornos por unidade de área estabelecido em contrato foi alcançado. Esse é um parâmetro definido pelo usuário. Para o caso do projeto EBA, foi especificado um mínimo de 4 retornos por metro quadrado.
- 6) Máximo de células abaixo da densidade: Além da necessidade de atingir a densidade mínima de retornos, é importante garantir que a densidade de pontos seja uniforme ao longo da nuvem. Deve-se então especificar um percentual máximo da área voada em que a densidade de retornos abaixo da densidade global é aceitável. No caso do Projeto EBA, definiu-se que no máximo 20% da área do transecto poderia ter uma densidade menor que 4 pontos/m², sendo esse valor é um parâmetro informado pelo usuário.
- 7) Detecção de ruídos: seja por um mau funcionamento temporário do sensor, seja por obstáculos interpostos entre o avião e a floresta (nuvens baixas, pássaros), as nuvens de pontos LiDAR podem apresentar ruídos. Essa rotina identifica nuvens de pontos “suspeitas”, por terem subconjuntos de pontos com valores no eixo Z fora do esperado.

2 Arquivos .csv, do inglês *Comma-Separated Values* (CSV), são formatos simples de armazenamento, que agrupam as informações de arquivos de texto em planilhas; cada linha em um texto .csv representa uma linha em uma planilha.

3 Detalhes da biblioteca Laspy podem ser encontrados em <http://laspy.readthedocs.io/en/latest/>

4 Detalhes sobre o USFS Fusion:

https://www.fs.fed.us/eng/rsac/lidar_training/Fusion/story_html5.html?lms=1

Essas nuvens então devem ser validadas visualmente, visto que fatores como uma grande variação da altitude do terreno sobrevoado podem levar a resultados falso-positivos.

- 8) Verificação visual da imagem da nuvem de pontos: as nuvens de pontos que não passaram nas etapas anteriores de validação são analisadas visualmente, pois pode acontecer do problema ser causado por uma característica do local que impede a solicitação de novo voo para a empresa. Destaca-se que essa etapa é usada em conjunto com os passos descritos anteriormente.

O processo de verificação visual da nuvem LiDAR é bastante demorado, uma vez que as nuvens contém muitos registros (normalmente da ordem de dezenas de milhões). Neste sentido, foi criada uma aplicação em Python para gerar imagens em formato JPG ou PNG a partir dos arquivos LAS/LAZ. Neste processo, a classificação dos retornos é feita com base em seus valores de coordenada Z, e a imagem é gerada a partir de uma palheta de cores semelhantes à de uma floresta, que varia em função dos valores mais baixos (marrom) aos mais altos (verde) (**Figura 3**). Foi desenvolvida uma aplicação para essa finalidade que permite que sejam processadas várias nuvens de pontos simultaneamente, podendo o usuário especificar o número de processadores disponíveis para serem usados para essa tarefa.

4. Resultados e Discussão

O arquivo de saída do processo de validação proposto neste estudo é um arquivo texto em formato .csv, que contém os resultados para todos os arquivos LAS/LAZ processados. Este arquivo reporta algumas estatísticas, os resultados das validações e mensagens de erros encontrados.

Usando essas rotinas, foram processados 624 arquivos LiDAR. Os resultados são apresentados na **Tabela 1**.

Tabela 1 - Resultado da validação dos 624 transectos LiDAR.

Etapa de validação	Transectos aprovados	Transectos não aprovados
1. Assinatura do arquivo	624	0
2. Versão do arquivo LAS/LAZ	624	0
3. Número do retorno	624	0
4. Quantidade de retornos	191	433
5. Valores mínimo e máximo	624	0
6. Densidade global de retornos	620	4
7. Máximo de células abaixo da densidade	501	123
8. Detecção de ruídos	603	21

Como apresentado na **Tabela 1**, foram encontrados problemas com as etapas 4, 6 e 7. Com relação aos erros encontrados na etapa 4, como descrito anteriormente, foi realizada uma comparação com relação à quantidade de retornos armazenado no cabeçalho com a quantidade de retornos efetivamente existente para o mesmo ponto. No entanto, durante essa etapa, notou-se que muitos arquivos apresentavam um número de retornos registrado maior do que 5, o que não se refletia nas informações do cabeçalho. Essa questão foi então verificada com o fabricante do sensor, que informou que, muitas vezes, são registrados mais de 5 retornos por pulso, porém o formato do cabeçalho dos dados LiDAR possui capacidade para armazenar apenas 5 campos de retorno. Sendo assim, optou-se por não considerar essa etapa como uma não conformidade do projeto EBA, já que foi usado um artifício para entregar mais informação que o formato comportava.

Os desvios observados nas etapas 6 e 7 de validação podem estar relacionados a duas causas: erro durante a aquisição dos dados ou alguma superfície não reflexiva ao LiDAR, como a água. Nos dois casos é necessário proceder à verificação visual das nuvens ou de imagens geradas a partir delas.



Figura 3 – Imagem PNG gerada a partir de uma imagem LiDAR, colorida com uma palheta de cores “florestal”.

As **Figuras 4a e 4b** mostram imagens criadas durante o processo de validação, com pixels de 20 x 20 m, coloridas conforme a densidade de retornos: i) pixels amarelos mostram onde não há dados (fora da faixa de voo); ii) pixels vermelhos representam baixa densidade de retornos (abaixo de 4,0 retornos/m², neste caso); iii) pixels verdes englobam áreas com 4,0 a 8,0 retornos/m² e iv) pixels azuis representam valores com mais de 8,0 retornos/m².

Pode-se observar que a uma densidade de pontos é maior na região central do transecto do que nas bordas (**Figura 4a**). Isso acontece porque o sensor é posicionado no centro da aeronave e faz a varredura perpendicularmente à linha de voo. Sendo assim, o número de pontos que não retornam para o sensor é menor do que aqueles que estão mais próximos do eixo de voo.

Além disso, ainda na **Figura 4a**, nota-se que há uma distribuição uniforme de pixels vermelhos ao longo das bordas do transecto, enquanto há pixels verdes na região central e uma baixa quantidade de pixels azuis, o que indica que, provavelmente, a aeronave estava voando mais alto do que o planejado.

A etapa 6 de validação (Máximo de células abaixo da densidade) apresenta causas semelhantes à etapa 5. No transecto da **Figura 5**, apesar da densidade de retornos média estar de acordo com a especificação (5,63 retornos/m²), 28,77% dos pixels apresenta uma densidade abaixo da especificação de 4 retornos/m².

As faixas de pixels vermelhos ao longo da linha de voo (**Figura 5**) são ocasionadas devido às variações na velocidade de voo. De acordo com os pilotos das aeronaves, esse fenômeno é reflexo de ventos de alta velocidade que dificultam a manutenção da velocidade ao longo do voo. Neste caso, o transecto foi recusado e a empresa responsável pela coleta de dados precisou

sobrevoar a área novamente.

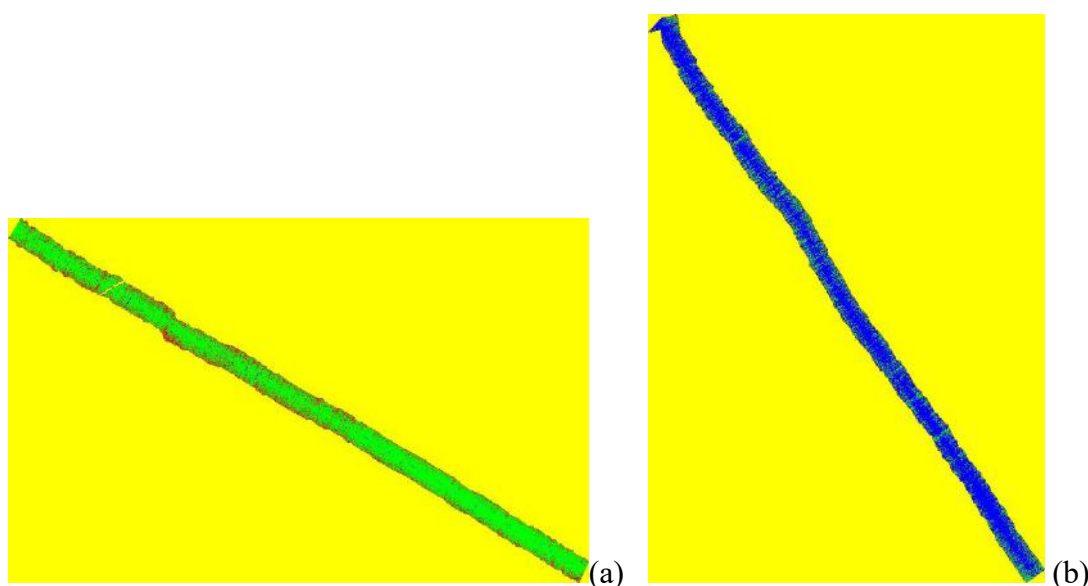


Figura 4 – Nuvem com (a) baixa densidade global de retornos (4,1 retornos/m²) e (b) alta densidade global de retornos (8,4 retornos/m²).

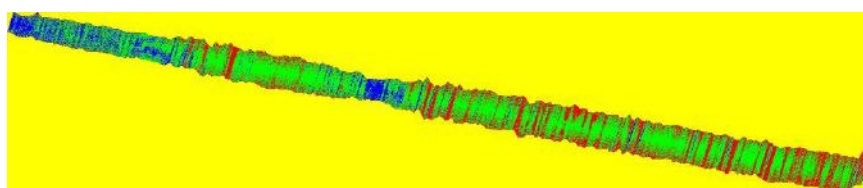


Figura 5 - Transecto com 28,77% dos pixels de 20x20m com densidade abaixo de 4 retornos/m².

No transecto apresentado na **Figura 6**, foi feita uma manobra com a aeronave sem que fosse desligado o sensor LiDAR, o que criou um efeito de “volta” ao final do transecto. Como para essa região específica a densidade de retornos estava abaixo da solicitada, o resultado da validação apontou que mais de 36,51% do transecto estava com densidade de retornos menor do que o solicitado. No entanto, para esse caso, toda a área do transecto estava coberta com a densidade de retornos mínima e, portanto, foi somente necessário cortar a nuvem LiDAR, não havendo necessidade de sobrevoar a área novamente.

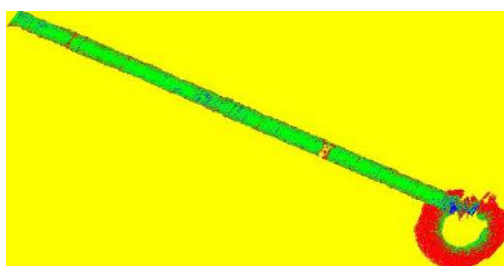


Figura 6 - Transecto com 36,51% of 20x20m pixels com densidade de menor que 4 retornos/m².

Para a etapa 8 de validação foi usada um script específico, denominado “findnoisyclouds.py”. O resultado desse processo é um arquivo .csv contendo os campos apresentados na **Tabela 2**.

Tabela 2 – Dados gerados pelo findnoisyclouds.py

Campo	Conteúdo
LAS_File	Nome do arquivo
number_of_returns	Número de retornos total
Groundavg	Media de altura pontos acima do limiar(*)
Groundstd	Desvio padrão da altura dos pontos acima do limiar
Allavg	Desvio padrão de todos os pontos
Stdallavg	Desvio padrão dos pontos de solo

(*) Consideram-se pontos altos os de altura em relação ao solo maior que 80 m, já que não se esperam muitas árvores com essa altura na Floresta Amazônica. Esse parâmetro é informado pelo usuário.

Em seguida, a partir dos resultados desse .csv, foram identificadas as nuvens com valores acima do limite de aceite. Estas foram então verificadas visualmente (**Figura 7**).

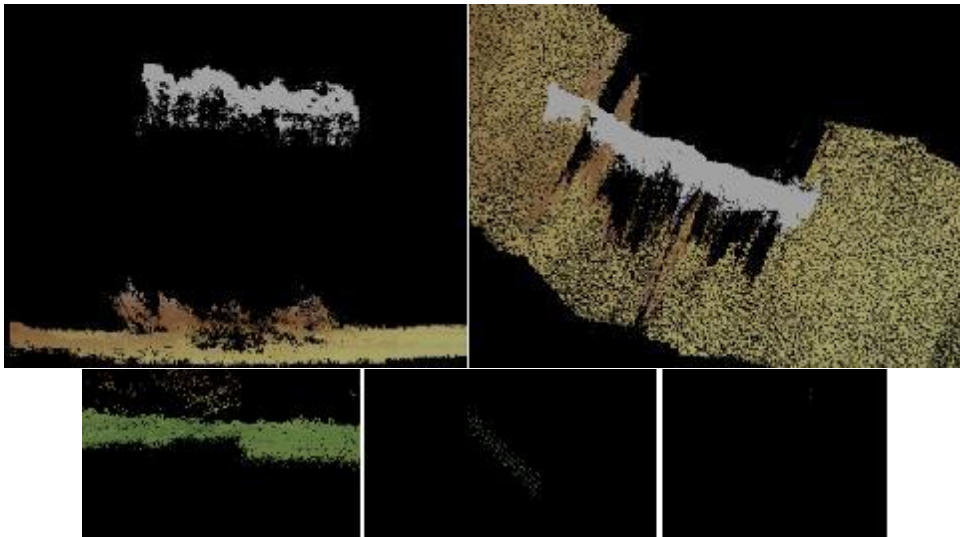


Figura 7 – Transectos com ruídos.

O algoritmo utilizado nessa etapa (“findnoisyclouds.py”) foi testado para regiões planas, como o bioma Amazônia. Em áreas mais íngremes, ele pode apontar falsos positivos (**Figura 8**). Sendo assim, para essas regiões, sugere-se utilizar parâmetros como desvio padrão ou até mesmo não utilizar esse método.



Figura 8 –Transecto falso positivo quando a presença de ruídos devido ao uso do algoritmo “findnoisyclouds.py” em regiões de relevo acentuado.

5. Conclusão

Os controles de qualidade dos dados de entrada de qualquer projeto de pesquisa são essenciais para o sucesso das estimativas feitas a partir dos produtos do sensoriamento remoto. O procedimento proposto, além de fornecer ferramentas para que esse controle seja feito em nuvens de pontos LiDAR, permitiu a otimização do tempo e redução da carga de trabalho para a validação de 624 nuvens de pontos LiDAR, que constituem cerca de 250 gigabytes de dados, à medida em que foram sendo entregues. A velocidade com que o processo foi feito atendeu tanto a expectativa da equipe do projeto, que tinha prontamente disponibilizados os dados entregues, quanto à do fornecedor, que teve a liberação de pagamento de suas faturas pelos serviços prestados dois ou três dias após a entrega de cada lote.

6. Agradecimentos

Gostaríamos de agradecer ao Fundo Amazônia, financiador do Projeto EBA, por propiciar a coleta dos dados aqui utilizados e sua disponibilização para uso nesse estudo.

7. Referências

- Barbosa, J. M.; Broadbent, E. N.; Bitencourt, M. D. Remote sensing of aboveground biomass in tropical secondary forests: a review. **International Journal of Forestry Research**, v. 2014, p. 1–14, 2014.
- IPCC Guidelines for National Greenhouse Gas Inventories. 2006, Japan: IGES, 2006. p. 673.
- Pan, Y.; Birdsey, R. A.; Fang, J.; Houghton, R.; Kauppi, P. E.; KURZ, W. A.; Phillips, O. L.; Shvidenko, A.; Lewis, S. L.; Canadell, J. G.; Ciais, P.; Jackson, R. B.; Pacala, S. W.; McGuire, A. D.; Piao, S.; Rautiainen, A.; Sitch, S.; Hayes, D. A large and persistent carbon sink in the world's forests. **Science** (New York, N.Y.), v. 333, n. 6045, p. 988–93, 2011.
- Pan, Y.; Birdsey, R. A.; Phillips, O. L.; Jackson, R. B. The structure, distribution, and biomass of the world's forests. **Annual Review of Ecology, Evolution, and Systematics**, v. 44, n. 1, p. 593–622, 2013.
- Wehr, A.; Lohr, U. Airborne laser scanning—an introduction and overview. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 54, n. 2-3, p. 68–82, jul. 1999.
- Rocha de Souza Pereira, F.; Kampel, M.; Gomes Soares, M.L.; Estrada, G.C.D.; Bentz, C.; Vincent, G. Reducing Uncertainty in Mapping of Mangrove Aboveground Biomass Using Airborne Discrete Return Lidar Data. **Remote Sensing**, v. 10, n. 637, p. 1-21, 2018.
- Vincent, G.; Sabatier, D.; Blanc, L.; Chave, J.; Weissenbacher, E.; Pélissier, R.; Fonty, E.; Molino, J.-F.; Coutron, P. Accuracy of small footprint airborne LiDAR in its predictions of tropical moist forest stand structure. **Remote Sensing of Environment**, v. 125, p. 23–33, 2012.
- Laser Scanner Data—Algorithms and Applications. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 54, n. 2-3, p. 138–147, jul. 1999.
- Baltsavias, E. . Airborne laser scanning: basic relations and formulas. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 54, n. 2-3, p. 199–214, jul. 1999.
- Sheng, Y. Quantifying the size of a lidar footprint: a set of generalized equations. **IEEE Geoscience and Remote Sensing Letters**, v. 5, n. 3, p. 419–422, jul. 2008.