

XXI ENMC
ENCONTRO NACIONAL DE
MODELAGEM COMPUTACIONAL

IX ECTM
ENCONTRO DE CIÊNCIA E
TECNOLOGIA DE MATERIAIS



08 a 11 de Outubro de 2018
Instituto Federal Fluminense
Búzios - RJ

CONTROLE E GARANTIA DE QUALIDADE DE DADOS LIDAR

Mauro Lúcio Rodrigues de Assis¹ - assismauro64@gmail.com

Rafael Santos¹ - rafael.santos@inpe.br

Jean Pierre Ometto¹ - jean.ometto@inpe.br

Eric Görgens² - e.gorgens@gmail.com

Luciane Sato¹ - lucianesato@gmail.com

Francisca Rocha de Souza Pereira¹ - franrspereira@gmail.com

Roberta Zecchini Cantinho³ - roberta.cantinho@inpe.br

Pedro Valle de Oliveira⁴ - pedrovco@gmail.com

¹ Instituto Nacional de Pesquisas Espaciais - INPE
Av. dos Astronautas, 1758 - Caixa Postal 515
12201-970 - São José dos Campos - SP, Brasil

² Universidade Federal dos Vales do Jequitinhonha e Mucuri
Campus JK, Departamento de Engenharia Florestal
39100-000 - Diamantina - MG, Brasil

³ Programa das Nações Unidas para o Desenvolvimento (PNUD),
Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC)
CEP 70720-610 - Brasília - DF, Brasil

⁴ South Dakota State University - Geospatial Sciences Center of Excellence
1021 Medary Ave, Wecota Hall 225
Brookings, SD, USA - 57007

Resumo. O uso de sensores Light Detect And Ranging (LiDAR) aerotransportados (ALS), tem se intensificado em uma série de aplicações que vão desde estudos quantitativos de estruturas florestais até a formação/atualização de cadastros de imóveis urbanos. 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 não impactem a qualidade das estimativas obtidas. Esse trabalho propõe: a) um conjunto de parâmetros a serem especificados quando da contratação de voos LiDAR 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/ebaccst/pyLiDARForest>.

Palavras-chave: LiDAR, Validação, Controle de qualidade, ALS.

1. INTRODUÇÃO

Concebido logo depois da invenção do laser, no início dos anos 1960, o uso da tecnologia de altimetrialaser ou *Light Detection and Ranging* (LiDAR) (Wehr, A.; Lohr, U, 1999) vem se disseminando em um grande conjunto de aplicações. 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 (Wehr, A.; Lohr, U, 1999). Devido à alta densidade de pontos e precisão dessa tecnologia, ela tem sido amplamente usada na representação em três dimensões (3D) de paisagens, batimetria, geração de modelos digitais de terreno, estimativa de biomassa florestal, 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 registrando seu retorno ao longo do voo da aeronave (AXELSSON, 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 resolução, 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.

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. As rotinas de validação apresentadas nesse estudo foram elaboradas para suportar ambos os formatos (LAS e LAZ), podendo ser facilmente adaptadas para outros formatos.

Os procedimentos aqui descritos foram testados em dados coletados na Amazônia. Como o LiDAR é uma tecnologia bastante adequada ao estudo da estrutura vertical de outros biomas, os procedimentos aqui propostos podem perfeitamente ser úteis ao processamento de dados LiDAR coletados em outros biomas brasileiros.

2. OBJETIVO

O objetivo desse trabalho foi desenvolver e apresentar um conjunto de procedimentos, implementados em scripts (programas de computador) escritos na linguagem Python que assegurassem a qualidade de dados LiDAR ALS a serem posteriormente usados em atividades de pesquisa científica, servindo também a outros usos.

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 624 transectos distribuídos aleatoriamente na região. Cada um desses transectos possui uma largura de aproximadamente 300m e comprimento de 12.500m (Fig. 1). Os dados foram coletados por uma empresa terceirizada no formato full waveform e entregues em formato LAS versão 1.4. O grande volume de dados coletados em uma das maiores campanhas ALS já feitas no mundo fez com que pudesse ser acumulado bastante conhecimento sobre os problemas de qualidade que podem aparecer em dados dessa natureza. Foram analisados cerca de $31,2 \times 10^9$ retornos LiDAR somando 500 Gb de dados.

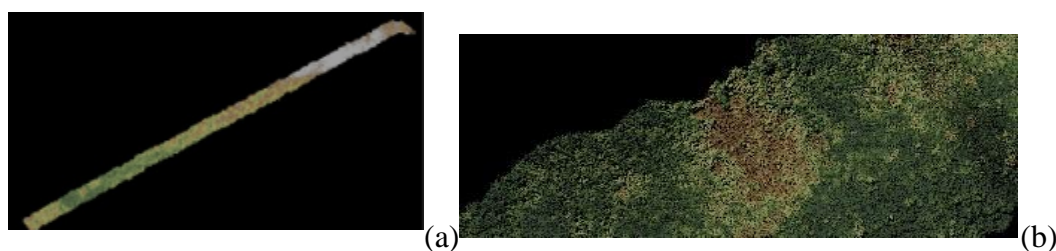


Figura 1 - (a) Imagem de nuvem LiDAR, com 12.500m de comprimento e 300m de largura, colorida em função da coordenada Z (altitude); (b) Zoom em parte do transecto.

Primeiramente, foram construídos, 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 LASF: 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

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/>

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

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. Essas nuvens então devem ser validadas visualmente, visto que fatores como uma grande variação da altitude do terreno sobrevoado podem levar a resultados falsos 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) (Fig. 2). 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.



Figura 2 – Imagem PNG gerada a partir de uma imagem LiDAR e colorida com uma palheta de cores “florestal”.

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 nas etapas 4, 6 e 7. Com relação aos erros da etapa 4, 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 no âmbito 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.

As Fig. s 3a e 3b 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 (Fig. 3a). 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 Fig. 3a, 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 Fig. 4, 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 (Fig. 4) 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 da aeronave ao longo do trajeto. Neste caso, o transecto foi recusado e a empresa responsável pela coleta de dados precisou sobrevoar a área novamente.

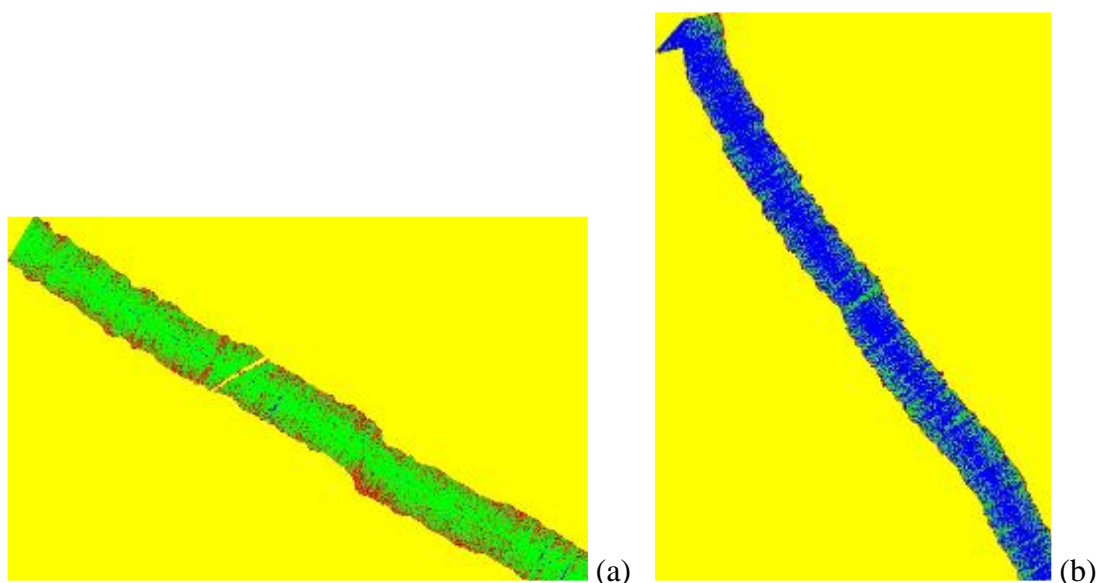


Figura 3 – Nuvem com (a) baixa densidade global de retornos (4,1 retornos/m²) e (b) alta densidade global de retornos (8,4 retornos/m²).

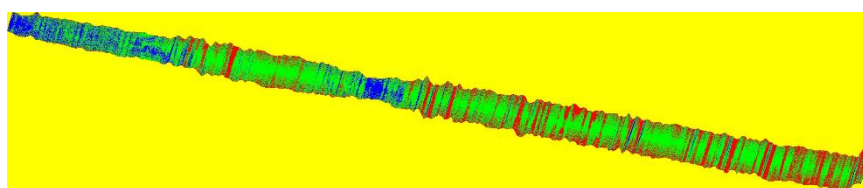


Figura 4 - Transecto com 28,77% dos pixels de 20x20m com densidade abaixo de 4 retornos/m².

No transecto apresentado na Fig. 5 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.

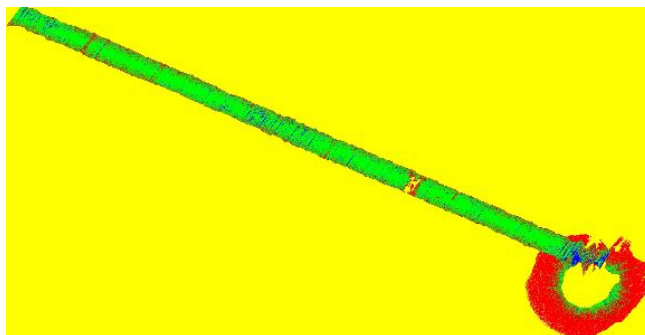


Figura 5 - Transecto com 36,51% of 20x20m pixels com densidade de retornos menor que 4/m².

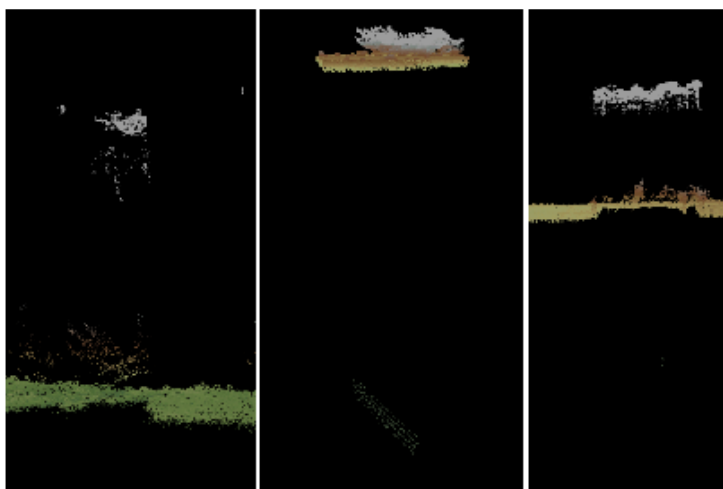
Para a etapa 8 de validação foi usada um script específico, denominado “findnoisyclouds.py”. O resultado desse processo é um arquivo formato Comma-Separated Values (CSV) contendo os campos apresentados na Tabela 2.

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 Amazônia. Esse parâmetro é informado pelo usuário.

Em seguida, a partir dos resultados do arquivo CSV gerado foram identificadas as nuvens com com valores acima do limite de aceite. Estas foram então verificadas visualmente (Fig. 6).



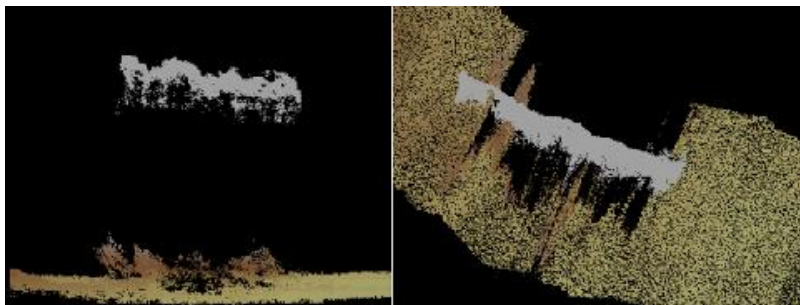


Figura 6 – Transectos com ruídos.

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 (Fig. 7). Sendo assim, para essas regiões, sugere-se utilizar parâmetros como desvio padrão ou até mesmo não utilizar esse método.

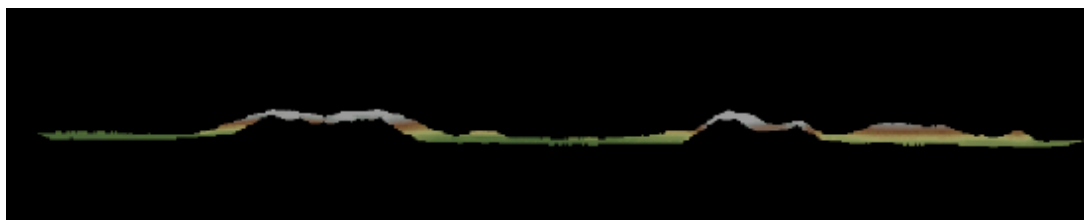


Figura 7 –Transecto com grande variação de altitude do solo, o que gerou um “falso positivo” (indicação de ruídos) devido ao uso do algoritmo “findnoisyclouds.py”.

5. CONCLUSÃO

O controle de qualidade dos dados de entrada de qualquer projeto de pesquisa é essencial para o sucesso das estimativas feitas. 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 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. Usando-se esse

REFERÊNCIAS

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, 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.
Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0924271699000155>.

Barbosa, J. M.; Broadbent, E. N.; Bitencourt, M. D. *Journal of Photogrammetry and Remote Sensing*, v. 54, n. 2-3, p. 138–147, jul.1999.

Laser Scanner Data—Algorithms and Applications. *ISPRS Journal of Photogrammetry and Remote Sensing*, v. 54, n. 2-3, p. 138–147, jul.1999.
Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S0924271699000088>

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.
Disponível em: <http://www.annualreviews.org/doi/abs/10.1146/annurev-ecolsys-110512-135914>.

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.

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.
Disponível em: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?>

QUALITY CONTROL AND ASSURANCE OF LIDAR DATA

Abstract. *The use of airborne LiDAR (ALS) sensors has intensified in a number of applications ranging from quantitative studies of forest structures to the formation / updating of urban real estate registries. In order 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 data inherent in the 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 b) a validation protocol of the data received in order to confront them with flight parameters and find other significant imperfections. This protocol is presented through a set of computer programs available at <https://github.com/ebaccst/pyLiDARForest>*