

TreeX (TREE EXTRACTOR) – UMA FERRAMENTA PARA ANÁLISE DE DOSEL FLORESTAL E CONTAGEM DE ÁRVORES A PARTIR DE DADOS LiDAR

TreeX (TREE EXTRACTOR) – A TOOL FOR FOREST CANOPY ANALYSIS AND TREE COUNTING WITH LiDAR DATA

Rodrigo de Campos Macedo ¹, João Roberto dos Santos ², João Viane Soares ³

Instituto Nacional de Pesquisas Espaciais - INPE, Av. dos Astronautas, 1758, 12.227-010 - São José dos Campos, SP, Brasil, ¹macedo@dsr.inpe.br ²jroberto@dsr.inpe.br ³viane@dsr.inpe.br

RESUMO

Os meios mais utilizados para obter dados dendrométricos são os inventários e levantamentos de campo, mas desde 1993, sensores ópticos ativos, conhecidos como laserscanners, começaram a ser utilizados especificamente no meio florestal. A hipótese deste artigo é que a utilização das abordagens de máximas locais e crescimento por regiões tornará possível a compatibilização entre a estimativa de parâmetros dendrométricos com dados obtidos via laserscanner aerotransportado e proveniente de inventário florestal tradicional. Os objetivos foram: implementar o algoritmo TreeX - desenvolvido para delimitar árvores e estimar parâmetros dendrométricos - e, validar estas estimativas em relação às referências de campo, para três parâmetros: quantidade de árvores, altura arbórea e área de copa, com erro máximo aceitável de 5%. A área de estudo é localizada no município de Igaratá/SP e possui 145,46ha de plantio de eucaliptos com 4 anos de idade. Foi utilizado o Optech ALTM 2050, com footprint de 0,25m, adquiridos em um vôo de 1.000m de altura. Considerando todas as parcelas medidas em campo, atingiu-se um percentual de erro de -3,52% para quantidade de árvores, -2,26% para altura, +19,36% para área de copa. A partir da análise dos dados foi possível verificar que apesar do algoritmo subestimar a quantidade de árvores e a altura, as estimativas apresentaram erro aceitável, porém apresentou superestimativa da área de copa, com erro acima do limite tolerável.

Palavras-chave: Inventário Florestal; LiDAR; Local Maxima.

ABSTRACT

Forest inventory survey are commonly used to obtain dendrometric data. but since 1993, optical active sensors, known as laserscanners, started to be specifically used in forest applications. The hypothesis of this paper is that the utilization of local maxima and region growing approaches will make it possible the estimative of dendrometric parameters with LiDAR data, compatible with the field measurement. The aims are: to implement the TreeX algorithm - developed to delimitate the trees and to estimate the dendrometric parameters - and, to validate these estimatives in relation to field references, for three parameters: amount of trees, tree height and crown area, with acceptable maximum error of 5%. The study area is located in the municipality of Igaratá (Sao Paulo State) and owns a four-year old eucalypts plantation, distributed over an area of 145.46 ha. The Optech ALTM 2050 has been employed, with a 0.25 m footprint, acquired by an average flight height of 1,000 m. Considering all plots regarded in this analysis, the error share was 3.52% for the total number of trees; 2.26% for tree height; 19.36% for crown area. The algorithm underestimated the number of trees and their height and overestimated the crown area, but the estimate is though acceptable.

Keywords: Forest Inventory; LiDAR; Local Maxima.

INTRODUÇÃO

A mensuração dos principais parâmetros dendrométricos é essencial para a modelagem e geração de prognósticos de volume madeireiro. Em povoamentos florestais homogêneos, a altura total, justamente por

ser mais influenciada por fatores genotípicos que ambientais, varia menos que a área basal. Sendo assim, torna-se factível realizar estimativas volumétricas apenas com os parâmetros relacionados ao dossel florestal – área de copa, quantidade de árvores e, principalmente, altura total.

A altura total é uma propriedade biofísica importante para indicar características da cobertura florestal que está sendo observada, tais como condições estruturais, grau de perturbação, condições de sítio, nível de produção florestal e outros aspectos úteis para o monitoramento e modelagem (LEITE e ANDRADE, 2003). Nos últimos anos, a geração e aquisição de dados hipsométricos com o uso de sensores remotos vem se intensificando, principalmente com LiDAR (Light Detection and Ranging), considerado um sensor com maior acurácia para obtenção de altura de alvos (LEWIS e HANCOCK, 2007). Já a área de copa pode ser relacionada com área basal e, conseqüentemente, substituí-la em inventário florestal com dados remotos. Ao estimar essas variáveis, torna-se necessário conhecer o erro da estimativa, principalmente através de validação, que consiste em comparar estatisticamente os valores das referências adotadas com os valores estimados.

Além da obtenção das alturas das árvores, é necessário delimitá-las, visando obter os demais parâmetros dendrométricos. Ainda não há aplicativos comerciais disponíveis para a delimitação das árvores e extração de variáveis dendrométricas. BALTSAVIAS (1999) comenta que os programas comerciais limitam-se a converter, classificar pontos e gerar modelos de superfície, sendo que parâmetros como contagem de árvores, cálculo de altura e área de copa, restringem-se a soluções locais e não-comerciais.

A delimitação das árvores pode ser realizada a partir das copas, tais como segmentação em imagens ópticas (ST-ONGE *et al.*, 2004) e/ou a partir das alturas. Neste artigo não serão abordadas técnicas que utilizam fotografias ou imagens orbitais de alta resolução como co-variável (através de classificação orientada a objeto, principalmente), mas sim técnicas que utilizam como insumo básico um MDA (Modelo Digital de Altura) – uma grade regular com os valores das alturas contidos nos pixels, porém sem a delimitação dos indivíduos arbóreos.

Uma abordagem pioneira para algoritmos que utilizam o MDA como insumo principal para a individualização das árvores é a adaptação de técnicas de delimitação de bacias (GOUGEON, 1998). Outra abordagem é a busca pelos valores mínimos locais. Através da ligação destes pixels há a formação de polígonos que representam as copas, sendo que as áreas dos polígonos representam as áreas das copas e os centróides representam a altura total (DIEDERSHAGEN *et al.*, 2000). Visando evitar a tendência de se formar copas justapostas, é possível iniciar a busca pelas máximas locais (BRANDTBERG *et al.*, 2003). FRIEDLAENDER e KOCH (2000) expressam a importância da lógica de máximas locais para a identificação do topo de uma árvore. PERSSON *et al.* (2002), em florestas boreais, determinaram os topos das árvores através de máximas locais, validando os resultados com a posição da árvore no terreno. Após a delimitação do topo das árvores, é possível estimar as copas através de um raio de busca com relação hipsométrica (KINI e POPESCU, 2004). Em povoamentos clonais, devido à seleção genotípica, as alturas tendem a uma certa homogeneidade, enquanto para outros parâmetros biofísicos não tão afetados pelo genótipo – copa e área basal – tal homogeneidade não ocorre. Por isto, as relações hipsométricas são muito baixas, limitando a utilização deste algoritmo. Outra forma de se delimitar as copas é através de algoritmos de crescimento por região, definindo-se limiares, tais como altura mínima, diâmetro máximo, diâmetro mínimo.

A hipótese deste artigo é que através de máximas locais é possível detectar os pontos referentes aos topos das árvores, enquanto que a área de copa pode ser modelada a partir de crescimento de regiões, possibilitando a estimativa de parâmetros dendrométricos a partir de dados LiDAR, compatíveis com as mensurações de campo. Os objetivos foram: i) implementar o algoritmo TreeX (Tree eXtractor), cuja finalidade é delimitar as árvores e permitir a estimativa dos parâmetros dendrométricos e, ii) validar estas estimativas em relação à referência de campo, para três parâmetros: quantidade de árvores, altura total e área de copa, com erro máximo aceitável de 5%.

MATERIAL E MÉTODOS

Área de Estudo

A área de estudo é localizada em Igaratá, Estado de São Paulo, na região conhecida como Vale do Paraíba. Possui 267,98ha, sendo 145,46ha ocupados com eucaliptos plantados em maio de 2004. Esta área foi escolhida por ser uma plantação florestal controlada, com a existência de inventário florestal contínuo e um censo de quantidade de árvores em três de seus nove talhões. (PIMENTEL, 2008). A Figura 1 apresenta a localização da área de estudo e uma visão geral da cobertura florestal.

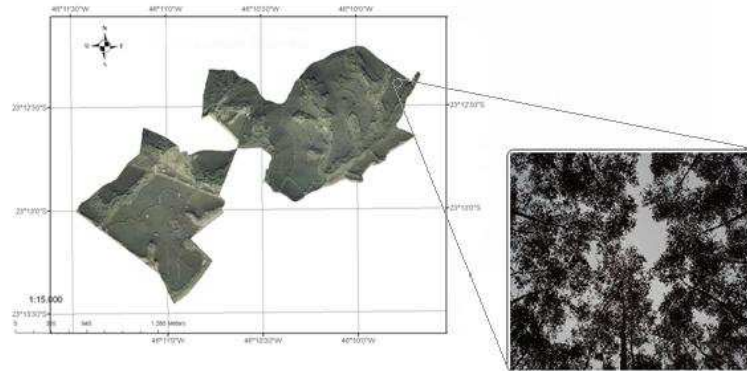


Figura 1. Localização da área de estudo e visão geral da cobertura florestal.

Figure 1. Study area localization and general view of forest cover.

Procedimentos Gerais

A Figura 2 apresenta o esquema metodológico utilizado, diferenciando os insumos (dados primários e derivados), processos e produtos, salientando que a validação do algoritmo consiste em uma comparação estatística entre os parâmetros dendrométricos estimados e os mensurados nas parcelas de campo, utilizados como referência.



Figura 2. Fluxograma metodológico adotado.

Figure 2. Methodological flowchart used.

O levantamento de campo foi realizado em junho de 2008 e foram medidas 20 parcelas retangulares de 60 árvores (espaçamento de 3x2m), aleatoriamente distribuídas, evitando-se proximidade com estradas e aceiros, visando minimizar os efeitos de borda. A intensidade amostral foi de 1 parcela a cada 7,5ha, totalizando 7713m² (0,53% da área plantada) e 1210 árvores. Foram medidos os perímetros das parcelas e mensurado o DAP de todos os indivíduos arbóreos presentes em seu interior. As variáveis altura total, diâmetro de copa e densidade foliar foram medidas apenas nas dez primeiras árvores e nas cinco dominantes. A determinação da altura foi efetuada com hipsômetro de Haaga. Os diâmetros de copa foram medidos na projeção vertical, com trena. Para conhecer as coordenadas dos vértices das parcelas, foi realizada uma poligonal aberta através de estação total até algum ponto com boa abertura – em geral, estradas ou pastagem na vizinhança. Nestes locais abertos, foram fixados 2 pontos (rovers) DGPS (Differential Global Position System) de 1 frequência para posterior cálculo de azimute e transporte de coordenadas absolutas para as coordenadas da estação total. Simultaneamente aos rovers, outro equipamento DGPS L1 foi fixado em um ponto de referência (sub-base). Posteriormente, esta coordenada sub-base foi corrigida a partir de um DGPS de dupla frequência utilizado como base. Esta coordenada, por sua vez foi corrigida a partir da antena da Escola Politécnica da USP (pertencente à RBMC (Rede Brasileira de Monitoramento Contínuo) e adotada como referência principal). STUMPF (1993), afirma que a incompatibilidade entre a precisão posicional da parcela ou do transecto amostral e a resolução de uma imagem, é a maior fonte de erro durante o processamento dos resultados. Utilizou-se 15° de máscara e valor 10 de DOP (Diluição da Precisão).

O aerolevanteamento foi realizado em abril de 2008, a uma altura de 1.000m, com frequência de varredura de 58,7Hz, ângulo de varredura de 15° e largura do feixe de 25cm. A densidade de pontos adquiridos foi de 5 por m². O laserscanner utilizado foi o Optech ALTM 2050 (50kHz), que atua na região espectral de 1064nm, registra 2 retornos (primeiro e último pulsos) e possui sistema de varredura com espelho oscilante (Z-Scan). O SMI (Sistema de Medição Inercial) utilizado foi o Applanix (200Hz).

As parcelas de campo foram utilizadas como dados de entrada no TreeX e os parâmetros dendrométricos extraídos foram validados utilizando-se as referências de campo e – especificamente para a quantidade de indivíduos arbóreos – o censo digital em fotografia aérea de alta resolução. Este censo foi considerado mais preciso que as extrapolações dos dados medidos em parcelas, devido ao fato de ter sido realizada contagem árvore-a-árvore de três talhões inteiros.

Visando auxiliar na comparação entre as mensurações de campo e as estimativas do algoritmo, foi calculado o RMSE (Root Mean Square Error) como medida estatística da magnitude do erro, expresso na Equação 1.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2} \quad (1)$$

onde: N = População; xi = valores individuais

Além do RMSE, foram calculados o erro amostral, o intervalo de confiança e analisados os resíduos, como pode ser visto nas Equações 2 e 3.

$$E = t.S_{\bar{y}} \quad (2)$$

onde E=Erro Amostral; t=student tabelado; S =Desvio da estimativa da média

Para o valor t, considerou-se 1,96 (95%) em todos os cálculos.

$$IC = \left\{ \bar{y} - t.S_{\bar{y}} \leq \mu \leq \bar{y} + t.S_{\bar{y}} \right\} \quad (3)$$

onde IC=Intervalo de Confiança; =estimativa da média

RESULTADOS E DISCUSSÃO

As principais variáveis dendrométricas medidas e a distribuição hipsométrica podem ser visualizadas na Tabela 1 e Figura 3, respectivamente.

Tabela 1. Variáveis dendrométricas por talhão/parcela.

Table 1. Dendrometric variables per stand/plot.

PARCELA	ÁREA PARCELA (m ²)	FUSTES/ha	ALTURA (m)	ÁREA COPA (m ²)
1	395,60	1517	21,15	5,31
2	390,60	1485	24,90	6,84
3	383,25	1566	22,62	6,86
4	384,56	1560	20,49	5,39
5	372,13	1585	22,19	7,14
6	382,95	1567	22,90	4,65
7	364,00	1648	22,05	7,56
8	421,80	1351	23,49	6,50
9	374,00	1578	22,64	4,79
10	363,81	1649	21,15	6,77
11	361,92	1630	21,26	4,97
12	387,45	1574	22,12	6,84
13	388,24	1520	23,32	5,48
14	387,66	1599	19,45	7,19
15	417,36	1414	21,97	5,75
16	371,85	1533	21,09	6,49
17	366,00	1585	21,41	5,58
18	367,20	1634	22,65	3,69
19	422,73	1419	23,74	1,63
20	409,92	1464	23,24	1,86
TOTAL/ MÉDIA	7713,03	1543,90	22,19	5,56

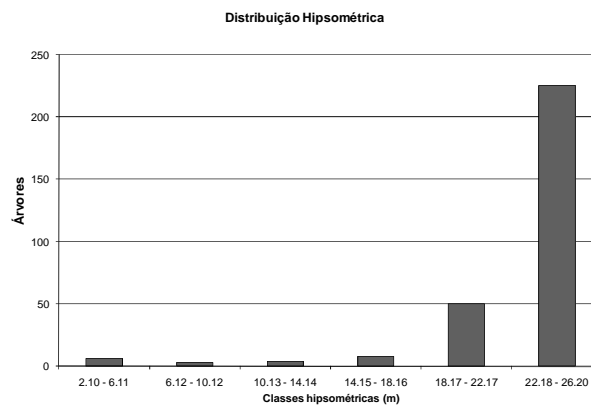


Figura 3. Distribuição hipsométrica.

Figure 3. Hypsometric distribution.

O aspecto geral do povoamento é de cerca de 1500 fustes por ha, com copas de 5m² e homogeneização de altura com predominância de indivíduos acima de 20m.

Subtraindo-se o MDT do MDS, foi gerado o MDA. Na Figura 4 é possível visualizá-los.

Os valores máximos deste MDA são relacionados às parcelas com aprox. 20 anos de idade. Os valores negativos são dados espúrios referentes a pequenos corpos de água.

Foram extraídas as quantidades de árvores através das máximas locais e extrapoladas por ha para tornar a medida comparável com outros conjuntos de dados. A comparação entre as quantidade de árvores contadas nas parcelas de campo (média de 1544 árvores por ha) e as estimadas pelo TreeX (média de 1598 árvores por ha) apresentou RMSE de 124 árvores por ha (3,65% de resíduos), com superestimativa influenciada diretamente pelo efeito do filtro passa-baixa, tamanho e deslocamento da janela móvel. Foram plotados os resíduos por parcela, sem ajustes, expressando apenas as diferenças (Figura 5A).

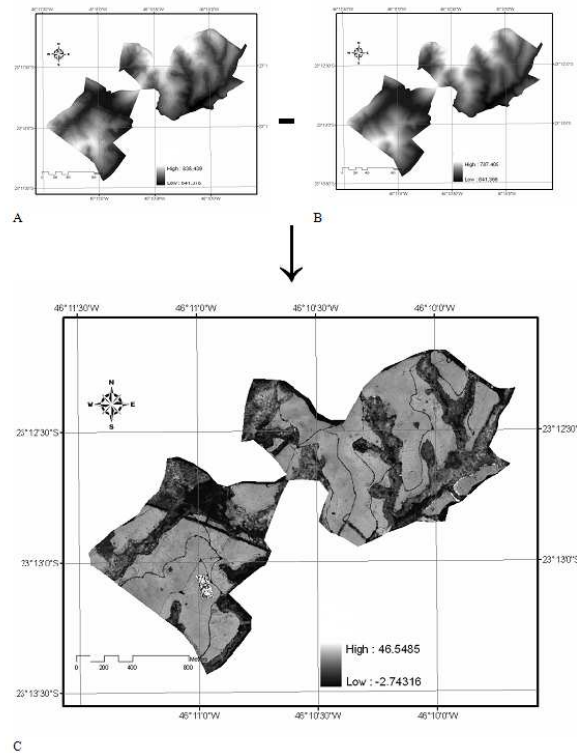


Figura 4. Modelo Digital de Superfície (A), Modelo Digital de Terreno (B) e o resultado da subtração de ambos, Modelo Digital de Alturas (C).

Figure 4. Digital Surface Model (A), Digital Terrain Model (B) and the result of the subtraction from both, Canopy Height Model (C).

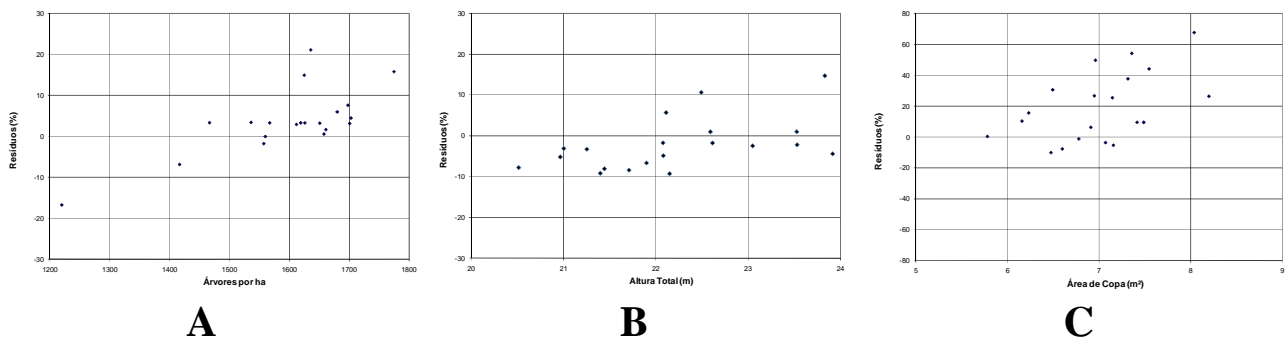


Figura 5. Análise de resíduos de regressão da estimativa via TreeX, para quantidade de árvores (A), altura total (B) e área de copa (C).

Figure 5. Regression analysis residuals of the TreeX estimative, for amount of trees (A), total height (B), crown area (C).

Através da plotagem dos resíduos é perceptível a tendência de superestimação na quantidade de árvores. Visando auxiliar na comparação entre os dois tipos de medidas (campo x algoritmo), foram calculados o erro amostral e os intervalos de confiança (95%) para os dois conjuntos de dados (Tabela 2).

Os erros de estimativa para os dados de campo e para o algoritmo foram de 2,49 e 5,29, respectivamente ($t = 1,96$), com média de resíduos de 3,65%. Apesar dessa superestimação, os intervalos de confiança são coincidentes, qualificando a estimativa como aceitável.

Tabela 2. Intervalo de Confiança (95%) para os parâmetros dendrométricos mensurados e estimados.
 Table 2. Confidence interval (95%) for the measured and estimated dendrometric parameters.

	Intervalo de Confiança (95%)	Referência (parcela de campo)	Algoritmo (TreeX)	Referência (Censo Digital)	Algoritmo (TreeX)
Quantidade de árvores por ha	Mínimo	1480,78	1504,34	1214,71	1188,64
	Média	1543,90	1598,44	1259,44	1212,80
	Máximo	1607,02	1692,53	1304,17	1236,96
Altura Total	Mínimo	21,83	21,45	-	-
	Média	22,78	22,21	-	-
	Máximo	23,73	22,97	-	-
Área de Copa	Mínimo	5,28	6,53	-	-
	Média	6,01	7,00	-	-
	Máximo	6,75	7,47	-	-

Visando incrementar a análise, foi utilizada como referência a informação de quantidade de árvore obtida via censo digital para três talhões. O resultado da comparação entre este censo digital (média de 1259 árvores por ha) com a estimativa do algoritmo (média de 1213 árvores por ha) apresentou RMSE de 80 árvores por ha. Diferentemente da comparação com as parcelas de campo, não houve viés positivo. Com erro de -3,52% e RMSE calculado em 80 árvores por ha houve considerável melhoria nos resultados. Na Tabela 2 é possível verificar os intervalos de confiança (95%) para os dois conjuntos de dados. A sobreposição dos intervalos de confiança expressa a aceitação da estimativa em relação ao censo.

Os erros de estimativa para o censo e para o algoritmo foram de 3,22 e 4,06, respectivamente ($t=1,96$). A diferença entre as referências – censo por fotografia aérea e a extrapolação da contagem via parcelas (ou estimação da contagem via inventário das parcelas) – pode ser explicada por diversos motivos, mas principalmente pelo fato de a área efetivamente plantada ser menor que a área cadastrada (base vetorial). Tal diferença pode ser ocasionada por: estradas mais largas que as plotadas na base cartográfica; presença de parcelas experimentais em áreas indicadas como “efetivo plantio”; redução de área plantada devido a afloramentos rochosos ou mesmo presença de galhada (proveniente do ciclo anterior) no momento do plantio; influência da declividade, etc. Como esta diferença entre área efetivamente plantada e área cadastrada na base vetorial é muito difícil de ser controlada, optou-se por adotar o censo via fotografia aérea como referência.

Realizou-se a comparação entre altura total mensurada em campo (média de 22,78m) e estimada via TreeX (média de 22,21m), com RMSE calculado em 1,48m e média dos resíduos em -2,26%, sendo que sua distribuição pode ser visualizada na Figura 5B. Através da plotagem dos resíduos, é perceptível a assertividade quanto à altura total. Foram calculados o desvio da média, o erro amostral e o intervalo de confiança (95%) para os dois conjuntos de dados. A Tabela 2 mostra os resultados. Os erros de estimativa de altura total mensurada a partir de dados de campo e através do algoritmo foram de 2,54 e 2,08, respectivamente ($t=1,96$).

Também se comparou a área de copa mensurada em campo (média de 6,01m²) e estimada via TreeX (média de 7,00m²), com RMSE de 1,49m² e média dos resíduos em 19,36%, efeito direto dos limiões do crescimento de região. A distribuição pode ser visualizada na Figura 5C. Através da plotagem dos resíduos ficou perceptível a necessidade de melhorias na estimativa de área de copa. Somente a área de copa, a altura e o contato com a próxima copa (variáveis de entrada), talvez não sejam suficientes para gerar um polígono compatível com a copa real. Foram calculados o desvio da média, o erro amostral e o intervalo de confiança (95%) para os dois conjuntos de dados (Tabela 2). Os erros de estimativa de área de copa mensurada a partir de dados de campo e através do algoritmo foram respectivamente ($t=1,96$), de 13,46 e 4,83.

CONCLUSÕES

Diante dos resultados apresentados, aceita-se a hipótese inicial, ou seja, a utilização do algoritmo TreeX apresentou resultados compatíveis com a acurácia necessária à gestão florestal. Ressalta-se o fato de que povoamentos clonais apresentam baixa variabilidade em altura. Analisando-se os intervalos de confiança, é possível perceber semelhança estatística para a estimação da quantidade de árvores, altura total e da área de

copa, porém, não é recomendada a aceitação da estimativa de área de copa, devido aos altos valores dos resíduos e do erro de estimativa.

Visando melhorar a estimativa de área de copa, sugere-se a utilização de imagens de alta resolução espacial e os processos de classificação orientada a objetos são extremamente vantajosos, principalmente na delimitação das copas.

Recomenda-se o aprimoramento e otimização do TreeX e o estudo de outras abordagens para a extração de variáveis dendrométricas a partir de dados LiDAR, incluindo povoamentos de diversas idades e de distintos índices de sítio.

Devido à elevada acurácia na medida hipsométrica obtida via LiDAR, aliada ao bom desempenho do algoritmo implementado, conjectura-se a possibilidade de realizar estimativas volumétricas com base em quantidade de árvores e altura total, desconsiderando-se a área basal, propiciando a realização de inventário estritamente remoto.

REFERÊNCIAS

- BALTSAVIAS, E.P. Airborne laser scanning: existing systems and firms and other resources. *ISPRS Journal of Photogrammetry & Remote Sensing*, v. 54, n. 2/3. p.164–198. 1999.
- BRANDTBERG, T.; WARNERA, T.; LANDENBERGERB, R.E.; MCGRAWB, J.B. Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America. *Remote Sensing of Environment*, v. 85. p. 290–303. 2003.
- DIEDERSHAGEN, O.; KOCH, B.; WEINACKER, H. Automatic Segmentation And Characterisation of Forest Stand Parameters Using Airborne Lidar Data, Multispectral and Fogis Data. *International Archives of Photogrammetry, Remote Sensing and spatial Information Sciences*, v. XXXVI. Amsterdam. 2000.
- FRIEDLAENDER, H.; KOCH, B. First Experience In The Application of Laserscanner Data for the Assessment of Vertical and Horizontal Forest Structures. *IAPRS*, v. XXXIII, Amsterdam, p. 693. 2000.
- GOUGEON, F. Automatic individual tree crown delineation using a valley-following algorithm and a rule-based system. *Proceedings of the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry*. Pacific Forestry Centre, Victoria B. C., 1998.
- KINI, A.; POPESCU, S.C. TREEVAW: A VERSATILE TOOL FOR ANALYZING FOREST CANOPY LIDAR DATA: A PREVIEW WITH AN EYE TOWARDS FUTURE. CD-ROM Proceedings. *ASPRS 2004 Fall Conference*, Kansas City, Missouri. 2004.
- LEITE, H.G.; ANDRADE, V.C.L. de. Importância das variáveis altura dominante e altura total em equações hipsométricas e volumétricas. *Revista Árvore*, Viçosa-MG, v.27, n.3, p.301-310, 2003.
- LEWIS, P.; HANCOCK, S. *LiDAR for vegetation applications*. UCL, Gower St, London, UK. January 23, 2007.
- PERSSON, A.; HOLMGREEN, J.; SÖDERMAN, U. Detecting and measuring individual tree using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing*, v. 68. p. 925-932, 2002.
- PIMENTEL, V.T.F. Uso de técnicas de interpretação de imagens na contagem das copas das árvores em plantios de eucalipto. Relatório final de estágio. Universidade Federal de Lavras. Lavras/MG. 2008. 21p.
- ST-ONGE, B.; JUMELET, J.; COBELLO, M.; VÉGA, C. Measuring individual tree height using a combination of stereophotogrammetry and lidar. *Canadian Journal for Remote Sensing*, v. 34. p. 2122–2130. 2004.
- STUMPF, K.A. The estimation of forest vegetation cover descriptions using a vertical densitometer. In: JOINT INVENTORY AND BIOMETRICS WORKING GROUPS SESSION. In: *The SAF National Convention*, Indianapolis, IN. 1993.