

Universidade Federal de São Carlos
Centro de Ciências Biológicas e da Saúde
Programa de Pós-Graduação em Ciências Ambientais

Gabriel Guariglia Perez

Uso de imagens do Sentinel 2 na estimativa de parâmetros
biofísicos da vegetação em áreas de Mata Atlântica

São Carlos - SP

2018

Universidade Federal de São Carlos
Centro de Ciências Biológicas e da Saúde
Programa de Pós-Graduação em Ciências Ambientais

Gabriel Guariglia Perez

Uso de imagens do Sentinel 2 na estimativa de parâmetros
biofísicos da vegetação em áreas de Mata Atlântica

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Ambientais da Universidade Federal de São Carlos como requisito parcial para a obtenção do título de Mestre em Ciências Ambientais.

Orientador: Prof. Dr. Vandoir Bourscheidt

Co-orientador: Prof. Dr. Luciano Elsinor Lopes

São Carlos - SP

2018

Guariglia Perez, Gabriel

Uso de imagens do Sentinel 2 na estimativa de parâmetros biofísicos da
vegetação em áreas de Mata Atlântica / Gabriel Guariglia Perez. -- 2018.

71 f. : 30 cm.

Dissertação (mestrado)-Universidade Federal de São Carlos, campus São
Carlos, São Carlos

Orientador: Vandoir Bourscheidt

Banca examinadora: Fábio Marcelo Breunig, Thiago Sanna Freire Silva
Bibliografia

Ficha catalográfica elaborada pelo Programa de Geração Automática da Secretaria Geral de Informática (SIn).

DADOS FORNECIDOS PELO(A) AUTOR(A)

Bibliotecário(a) Responsável: Ronildo Santos Prado – CRB/8 7325



UNIVERSIDADE FEDERAL DE SÃO CARLOS


Centro de Ciências Biológicas e da Saúde
Programa de Pós-Graduação em Ciências Ambientais

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado do candidato Gabriel Guariglia Perez, realizada em 22/08/2018:



Prof. Dr. Vandoir Bourscheidt
UFSCar



Prof. Dr. Luciano Elsinor Lopes
UFSCar



Prof. Dr. Thiago Sanna Freire Silva
UNESP

Prof. Dr. Fabio Marcelo Breunig
UFSCar

Certifico que a defesa realizou-se com a participação à distância do(s) membro(s) Fabio Marcelo Breunig e, depois das arguições e deliberações realizadas, o(s) participante(s) à distância está(ão) de acordo com o conteúdo do parecer da banca examinadora redigido neste relatório de defesa.



Prof. Dr. Vandoir Bourscheidt

Aos meus avós

AGRADECIMENTOS

Agradeço primeiramente a toda minha família, sei que nunca estarei sozinho enquanto tiver vocês por perto. Em especial, agradeço à minha mãe e ao meu pai por todo o suporte, dedicação e amor. Sem vocês eu não seria ninguém.

Também agradeço ao meu irmão por me aguentar por todos esses anos hahaha. Agradeço aos meus primos e primas, tios e tias, aos meus avós (aos que já se foram e à minha vó que ainda está aqui), e aos cães e gatos, muito obrigado por fazerem parte da minha vida!

Agradeço aos meus amigos pela companhia nos bons e maus momentos da vida. Amigos são a família que escolhemos para nós. Aos amigos que estão longe, a distância e o tempo não diminuem o carinho que eu sinto por todos vocês!

Muito obrigado aos amigos de infância da rua, aos meus amigos da Educativa, do Sapiens e do prédio em São Paulo, amizades antigas, mas que duram até hoje. Agradeço aos amigos da graduação, às novas amizades desses últimos anos no mestrado e ao pessoal da USP de Ribeirão Preto. Muito obrigado a todos por tornarem minha vida mais feliz.

Agradeço muito ao Luciano e ao Vandoir, meus orientadores, por todo o tempo dedicado ao trabalho, pela amizade e pelas noites de violão. Agradeço ao Vandoir por todas as paçocas e pipocas que me deram energia para continuar e ao Luciano por sempre ter o cuidado de perguntar se me acordou. Agradeço também à Paty pela amizade e pelas festas na chácara!

Muito obrigado a todo o pessoal do LEAP. Apesar do calor, do frio, da chuva e das picadas de mosquito, as coletas de campo foram uma experiência muito boa graças à companhia de vocês. Agradeço especialmente à Ju pela amizade e por fornecer os dados de campo (como você diz, é nois na estrutura da vegetação!). Sem essa contribuição esse trabalho não seria metade do que é hoje. Também agradeço o pessoal que participou do conecta, foi muito divertido fazer o jogo com vocês (temos que terminar ele agora hahaha).

Também agradeço muito aos professores Fábio, Thiago, Veraldo e Luiz Eduardo, que aceitaram serem membros da minha banca de defesa. Aos que

participaram da banca de qualificação do texto, muito obrigado por todos os comentários. O trabalho melhorou muito com a ajuda de vocês.

Agradeço a ambas as bancas de qualificação da aula, ao Victor, Andreia, Felipe, Hildo, Reinaldo e à Patrícia, por me ajudarem a melhorar nas apresentações. Apesar do susto, no fim deu tudo certo!

Também agradeço a toda a organização e aos palestrantes do evento em Lages. O que aprendi me ajudou bastante no desenvolvimento deste trabalho.

Agradeço às pessoas do projeto paisagens sustentáveis da EMBRAPA por fornecerem os dados de LiDAR, outra contribuição essencial para o desenvolvimento do trabalho.

Ao PPGCAm pela oportunidade de cursar o mestrado. Meu muito obrigado aos amigos do departamento, por todas as conversas, piadas, grupos de estudo e pela oportunidade de jogar futebol com os craques de São Carlos! Agradeço a todos os professores pelas aulas, conselhos, empréstimos de livro e pela amizade. Em especial ao Victor pela ajuda com os modelos *Random Forest* e ao Felipe, que me orientou durante o PESCD. Também agradeço ao pessoal da secretaria, especialmente o Vinícios, pela paciência, por me lembrar de prazos importantes e por sempre estar à disposição para ajudar.

Por fim, agradeço pelo financiamento recebido para a participação em eventos científicos. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001

RESUMO

Informações sobre parâmetros biofísicos da vegetação podem ser usadas para as mais diversas aplicações, e o sensoriamento remoto vem se mostrando uma boa ferramenta para obtê-las. Neste estudo foi investigada a possibilidade de usar imagens do Sentinel 2 para estimar parâmetros biofísicos da vegetação medidos em campo e com LiDAR (Light Detection and Ranging). O trabalho foi feito em áreas de Mata Atlântica no estado de São Paulo (Brasil) usando três imagens do Sentinel 2. Adicionalmente, foi utilizada uma imagem Landsat-8/OLI por proximidade temporal com os dados LiDAR, testado o efeito da aplicação de correção topográfica nas imagens e feita uma análise de espectros de reflectância de folhas em laboratório. As variáveis de campo analisadas foram altura, DAP (Diâmetro à Altura do Peito), porcentagem de cobertura e número de indivíduos. As variáveis LiDAR foram altura dos primeiros ecos, altura dos últimos ecos e quantidade de ecos por pulso. Um total de 26 variáveis foram extraídas para comparação com as imagens em modelos de regressão OLS (*Ordinary Least Squares*) e RF (*Random Forest*). Essas comparações foram feitas com as bandas de forma individual, com os índices de vegetação RVI (*Ratio Vegetation Index*), NDVI (*Normalized Difference Vegetation Index*), SAVI (*Soil Adjusted Vegetation Index*), EVI (*Enhanced Vegetation Index*), NDWI (*Normalized Difference Water Index*), NDI45 (*Normalized Difference Index B4 and B5*), IRECI (*Inverted Red-Edge Chlorophyll Index*) e S2REP (*Sentinel 2 Red Edge Position*), e com todas as razões possíveis entre duas bandas. Os resultados mostram que muitos dos parâmetros biofísicos tem relação com as imagens (r^2 de até 0,62), e a correção topográfica parece ter efeito positivo nas estimativas, principalmente das variáveis derivadas de LiDAR. Os melhores modelos gerados para os dados de campo foram regressões múltiplas entre bandas do Sentinel 2 e a altura das árvores, cobertura do dossel e biomassa. Para os dados LiDAR, os resultados foram melhores que os de campo, principalmente em regressões múltiplas entre bandas do Sentinel 2 e altura e cobertura do dossel. Dentre as imagens, a que apresentou melhores relações com os parâmetros biofísicos foi a de 26 de dezembro de 2016 do Sentinel 2. Os resultados da validação dos modelos LiDAR mostram que eles podem ser usados em áreas diferentes da mesma imagem em que foram treinados e também podem ser usados em imagens diferentes desde que seja aplicada uma correção atmosférica apropriada. De maneira geral, índices de vegetação não tiveram resultados melhores que as bandas individuais. Dentre as bandas, destaca-se o papel de B5 (705 nm, red-edge) no sucesso dessas estimativas, o que é confirmado pelos resultados da análise de espectros de reflectância feita em laboratório. Para trabalhos futuros, recomenda-se investigar melhor o comprimento de onda próximo a 705 nm e o potencial da banda 5 do Sentinel 2. Para a criação de modelos que possam ser usados em outras imagens, recomenda-se a utilização de imagens do Sentinel 2 em nível 2A (reflectância de superfície), que estarão disponíveis globalmente até o fim de 2018. Por fim, recomenda-se considerar outras variáveis, como classes de estágio sucessional, e a comparação das imagens com outras fitofisionomias.

Palavras-chave: LiDAR, Sensoriamento remoto, Estrutura da vegetação.

ABSTRACT

Information about vegetation biophysical parameters can be used in many applications, and remote sensing is showing itself to be a good tool to obtain it. In this study we investigated the possibility of using Sentinel 2 images to estimate vegetation biophysical parameters measured in field and with LiDAR (Light Detection and Ranging). The study was conducted in areas covered by Atlantic Forest in the state of São Paulo (Brazil) using three Sentinel 2 images. Additionally, we used a Landsat-8/OLI image for temporal proximity with the LiDAR data, tested the effect of topographic correction on the images and made an analysis of leaf reflectance in lab. The analyzed field variables were height, DBH (Diameter at breast height), percentage of canopy cover, and number of individuals. The LiDAR variables were height of the first returns, height of the last returns and number of returns per pulse. A total of 26 variables were extracted for comparisons with the images through OLS (Ordinary Least Squares) and RF (Random Forest) regression models. These comparisons were made with single bands, the vegetation indices RVI (Ratio Vegetation Index), NDVI (Normalized Difference Vegetation Index), SAVI (Soil Adjusted Vegetation Index), EVI (Enhanced Vegetation Index), NDWI (Normalized Difference Water Index), NDI45 (Normalized Difference Index B4 and B5), IRECI (Inverted Red-Edge Chlorophyll Index) and S2REP (Sentinel 2 Red Edge Position), and with all the possible ratios between two bands. The results show that many biophysical parameters are related to the images (r^2 up to 0.62), and the topographic correction seems to have a positive effect in the estimates, especially for LiDAR derived variables. The best models generated with field data were regressions between multiple Sentinel 2 bands and tree height, canopy cover and biomass. For the LiDAR data, the results were in general better than with field data and also involved regressions with multiple Sentinel 2 bands in comparison with canopy height and cover. Among the images, the one that presented better relations with the biophysical parameters were Sentinel 2 image of December 26, 2016. The validation results of the LiDAR models show that they work in different areas of the same image in which they were trained but can only be applied in different images after an appropriate atmospheric correction. In general, vegetation indices did not show better results than individual bands. Among the bands, we highlight the role of B5 (705 nm, red-edge) for the success of many estimates, which is confirmed by the results of the analysis of leaf reflectance made in lab. For future studies, we recommend better investigation of wavelengths close to 705 nm and the potential of Sentinel 2 band 5. For the creation of models that can be used in different images, we recommend the use of level 2A Sentinel 2 images (surface reflectance), that will be available globally by the end of 2018. Finally, we also recommend to consider other variables such as classes of successional stages and the comparison of Sentinel 2 imagery with other phytophysiognomies.

Key-words: LiDAR, Remote Sensing, Vegetation structure.

LISTA DE FIGURAS

- Figura 1 – Localização da área de estudo com as parcelas de campo (vermelho) e as áreas de coleta dos dados LiDAR (ciano). As parcelas de campo não estão em escala.27
- Figura 2 – Bandas das imagens do Sentinel 2 de acordo com sua resolução espacial. A imagem também mostra a largura de banda e a posição no espectro eletromagnético. Fonte: modificado de ESA.29
- Figura 3 – Imagens do Sentinel 2 utilizadas no estudo com as respectivas parcelas de campo e áreas de coleta dos dados LiDAR. Nenhuma das áreas utilizadas se encontra em áreas cobertas por nuvens, embora estas possam ter influências na radiometria por conta de efeitos de espalhamento.31
- Figura 4 – Composição RGB da imagem do Landsat-8/OLI utilizada no estudo com a localização das áreas Cantareira 1 e 2 em que foram medidos os dados LiDAR. Apenas a área Cantareira 1 (esquerda) foi utilizada por conta da presença de nuvens na Cantareira 2 (direita).33
- Figura 5 – Exemplo de parcela utilizada na coleta dos dados de campo. A área em cinza representa a área de coleta (3 vértices de 25 por 2 metros, ou 150 m²). Os pontos com cruz são os locais da coleta de dados de cobertura da vegetação.39
- Figura 6 – Recortes da composição RGB (bandas 2, 3 e 4) das imagens do Sentinel 2 com detalhe de alguns dos buffers gerados (vermelho) com 25 metros de raio para representar as parcelas de campo. As imagens da esquerda são de 8 de janeiro de 2016 e as da direita de 23 de dezembro de 2016.42
- Figura 7 – Localização e delimitação (em vermelho) das áreas Cantareira 1 e 2.....43
- Figura 8 – Subtração da imagem de elevação dos ecos do solo (imagem central) da elevação dos primeiros ecos (imagem à direita), resultando em uma imagem que pode ser usada como um modelo de altura do dossel (CHM) (imagem à esquerda). As imagens correspondem à área Cantareira 1.45
- Figura 9 – Perfis topográficos passando pelo centro das áreas Cantareira 1 e 2 nos sentidos Norte-Sul e Oeste-Leste. As linhas azuis representam os modelos digitais de superfície (DSM) gerados a partir dos ecos classificados como solo e as linhas vermelhas os DSM gerados a partir dos últimos ecos. As setas vermelhas no canto superior direito de cada gráfico representam o sentido e a posição em que foram

traçados os perfis topográficos em relação às áreas Cantareira 1 e 2 (em cinza), que podem ser vistas na Figura 7.	46
Figura 10 – Histograma para as áreas Cantareira 1 (esquerda) e 2 (direita) mostrando a distribuição dos pixels da imagem de altura dos últimos ecos de acordo com seus valores nas áreas de vegetação. Os pixels negativos correspondem a menos de 1% do total de pixels para ambas as áreas e não tem valores menores que -1,3 m.....	46
Figura 11 – Imagem de altura dos primeiros ecos (CHM), altura dos últimos ecos e número de ecos por pulso para as áreas Cantareira 1 e 2.	47
Figura 12 – Um dos métodos de reamostragem da imagem de altura dos primeiros ecos, ou CHM (0,5 m) para a mesma resolução das imagens do Sentinel 2 (10 m). Nesse método, tirou-se primeiro o valor máximo dos pixels em cada quadrante de 5 m (imagem à esquerda) e depois calculou-se a média desses valores para determinar o valor final do pixel (imagem à direita).....	48
Figura 13 – Exemplo do cálculo da porcentagem de cobertura. Os pixels do CHM (imagem da esquerda) considerados como cobertos por vegetação tem valor 1 e os pixels considerados como clareiras tem valor 0. Na reamostragem para a mesma resolução do Sentinel foi feita uma média aritmética dos valores, resultando em 0,75 no exemplo (75% de cobertura vegetal no pixel). A imagem é uma simplificação, já que um pixel do Sentinel 2 (10 m) contém um total de 400 pixels do CHM.	49
Figura 14 – Reamostragem e alinhamento para pixels de 10 m do Sentinel 2 seguido de recorte delimitando as áreas de vegetação natural. Também foram gerados buffers aleatórios (em vermelho) dentro dessas áreas para calcular as médias dentro deles como método de amostragem alternativo. A área representada na figura é a Cantareira 1.....	51
Figura 15 – Representação esquemática da assinatura espectral de diferentes camadas de folha.....	56
Figura 16 – <i>Boxplots</i> mostrando a dispersão de algumas das variáveis de campo com os valores de média e coeficiente de variação (CV) para todas as 34 parcelas amostradas. BSAST (Biomassa Seca Acima do Solo total - kg), C (porcentagem de cobertura do dossel - %), DAPT (Diâmetro à Altura do Peito Total - cm), AM (Altura Média - m).	58
Figura 17 – Valores de r para a correlação (Spearman) das variáveis derivadas dos dados de campo. (Siglas – ver quadros 3 ou 8).....	59

Figura 18 – Gráficos de dispersão com o valor de r e p para a correlação (Spearman) entre as variáveis AB (Area Basal - m ² /ha) e BSAST (Biomassa Seca Acima do Solo - kg) (esquerda) e AT (Altura Total - m) e DAPT (Diâmetro à Altura do Peito Total - cm) (direita).	59
Figura 19 – Valores de r ² para a regressão linear (OLS) das bandas do Sentinel 2 apenas com correção atmosférica (eixo x) e dos parâmetros biofísicos da vegetação derivados dos dados de campo (eixo y). (Siglas – ver quadros 3 ou 8)	60
Figura 20 – Valores de r ² para a regressão linear (OLS) das bandas do Sentinel 2 com aplicação de correção topográfica e atmosférica (eixo x) e das variáveis derivadas dos dados de campo (eixo y). (Siglas – ver quadros 3 ou 8).	61
Figura 21 – Banda 5, NDVI e razão da banda 11 pela banda 3 com as imagens de janeiro (esquerda) e dezembro (direita) do Sentinel 2 com aplicação de correção topográfica e atmosférica. Em vermelho, a localização das parcelas de campo (fora de escala). Apesar da grande quantidade de nuvens nas imagens, as 34 parcelas selecionadas para análise não estão em áreas nubladas.	62
Figura 22 – Valores de r ² para a regressão linear (OLS) dos índices de vegetação (eixo x) derivados das bandas do Sentinel 2 com aplicação de correção topográfica e atmosférica e das variáveis derivadas dos dados de campo (eixo y). (Siglas – ver quadros 3 ou 8)	62
Figura 23 – Valores de r ² para a regressão linear (OLS) entre a variável AMed (altura mediana) obtida em campo e a razão entre as duas bandas do Sentinel 2 representadas no eixo y e x (respectivamente o numerador e denominador da razão). Imagens com aplicação de correção topográfica e atmosférica.	63
Figura 24 – Valores de r ² para a regressão linear (OLS) entre a variável AM (altura media) obtida em campo e a razão entre as duas bandas do Sentinel 2 representadas no eixo y e x (respectivamente o numerador e denominador da razão). Imagens com aplicação de correção topográfica e atmosférica.	64
Figura 25 – Gráficos de dispersão para a relação entre as razões de B11 e B3 (eixo x) corrigidas atmosfericamente e topograficamente e as variáveis de campo AM (altura media), à esquerda, e AMed (altura mediana), à direita (eixo y) para todas as 34 parcelas de campo.	64
Figura 26 – Valores de r da correlação (Speraman) entre bandas do Sentinel 2 com aplicação de correção topográfica e atmosférica.	65

Figura 27 – Valores de r^2 para a regressão linear (OLS) da comparação pixel a pixel das bandas do Sentinel 2 (2016) apenas com correção atmosférica DOS1 (eixo x) e das variáveis obtidas por LiDAR (eixo y).....	67
Figura 28 – Valores de r^2 para a regressão linear (OLS) da comparação pixel a pixel das bandas do Sentinel 2 (2016) com aplicação de correção topográfica e atmosférica DOS1 (eixo x) e das variáveis obtidas por LiDAR (eixo y).	67
Figura 29 – Valores de r^2 para a regressão linear (OLS) da comparação pixel a pixel das bandas do Landsat-8/OLI (2014) para a Cantareira 1 com aplicação de correção topográfica e atmosférica DOS1 (eixo x) e das variáveis de estrutura vegetal obtidas por LiDAR (eixo y).	68
Figura 30 – Valores de r^2 para a regressão linear (OLS) da comparação pixel a pixel das bandas do Sentinel 2 (2015) com aplicação de correção topográfica e atmosférica DOS1 (eixo x) e das variáveis de estrutura vegetal obtidas por LiDAR (eixo y).....	69
Figura 31 - Valores de r^2 para a regressão linear (OLS) da comparação por buffers das bandas do Sentinel 2 (2016) com aplicação de correção topográfica e atmosférica DOS1 (eixo x) e das variáveis de estrutura vegetal obtidas por LiDAR (eixo y).....	69
Figura 32 – <i>Boxplots</i> mostrando a dispersão de algumas das variáveis derivadas de LiDAR com os valores de média e coeficiente de variação (CV) para todos os 338 buffers gerados.	71
Figura 33 – Valores de r para a correlação (Spearman) das variáveis derivadas dos dados de LiDAR.	72
Figura 34 – Valores de r^2 para a regressão linear (OLS) da comparação por buffers dos índices de vegetação derivados das imagens do Sentinel 2 (2016) com aplicação de correção topográfica e atmosférica DOS1 (eixo x) e das variáveis de estrutura vegetal obtidas por LiDAR (eixo y).	73
Figura 35 – Valores de r^2 para a regressão linear (OLS) da comparação por buffers das razões das bandas do Sentinel 2 com aplicação de correção topográfica e atmosférica DOS1 (eixos x e y) e da variável densidade.	74
Figura 36 – Gráfico de dispersão da relação entre B5 (705 nm) com aplicação de correção topográfica e atmosférica (eixo x) e as variáveis derivadas de LiDAR (eixo y) altura máxima 5 (esquerda) e altura stdev (direita) com a linha do modelo de regressão linear (OLS) desenhada. Acima do gráfico, os valores de r^2 , p , coeficiente de	

<p>multiplicação e constante a ser adicionada na fórmula do modelo ($y = 50,94 + x * -486,71$).....</p>	75
<p>Figura 37 – Mapa de erros do modelo de regressão da relação entre B5 e a altura máxima 5 para as áreas de floresta. Pixels mais azulados representam subestimativas e pixels mais avermelhados superestimativas. Cores mais intensas representam erros maiores.....</p>	75
<p>Figura 38 – Valores das importâncias das bandas do Sentinel 2 usadas para gerar os modelos <i>Random Forest</i> para as variáveis Altura máxima 5, Cobertura 10, Densidade, Altura máxima 10, Altura média e Profundidade.</p>	78
<p>Figura 39 – Histograma mostrando a distribuição dos valores de RMSE para todos os cinco mil modelos OLS gerados com as bandas do Sentinel 2 (23/12/2016) para validação em áreas aleatórias para seis das variáveis derivadas de LiDAR.....</p>	80
<p>Figura 40 – Comparação entre as predições (eixo x) com bandas do Sentinel 2 (23/12/2016) de modelos de regressão OLS (<i>ordinary least squares</i>) e RF (<i>Random Forest</i>) e os valores reais (eixo y) em metros das variáveis Altura média, Altura máxima 5 e Altura máxima 10 para as áreas usadas como teste. Os modelos OLS tem o erro (RMSE) igual à mediana da distribuição de erros dos cinco mil modelos testados.....</p>	81
<p>Figura 41 – Comparação entre as predições (eixo x) com bandas do Sentinel 2 (23/12/2016) de modelos de regressão OLS (<i>ordinary least squares</i>) e RF (<i>Random Forest</i>) e os valores reais (eixo y) das variáveis Cobertura 10, Profundidade (m) e Densidade (e/m) para as áreas usadas como teste. Os modelos OLS tem o erro (RMSE) igual à mediana da distribuição de erros dos cinco mil modelos testados... </p>	82
<p>Figura 42 – Comparação entre as predições (eixo x) com bandas do Sentinel 2 (08/01/2016) de modelos de regressão OLS (<i>ordinary least squares</i>) e RF (<i>Random Forest</i>) treinados na imagem de 23 de dezembro de 2016 e os valores reais (eixo y) em metros das variáveis Altura máxima 5, Altura máxima 10 e Altura média.....</p>	83
<p>Figura 43 – Comparação entre as predições (eixo x) com bandas do Sentinel 2 (08/01/2016) de modelos de regressão OLS (<i>ordinary least squares</i>) e RF (<i>Random Forest</i>) treinados na imagem de 23 de dezembro de 2016 e os valores reais (eixo y) das variáveis Cobertura 10, Profundidade (m) e Densidade (e/m).....</p>	84
<p>Figura 44 – Comparação entre as predições (eixo x) com bandas do Sentinel 2 (08/01/2016) de modelos de regressão OLS (<i>ordinary least squares</i>) e RF (<i>Random</i></p>	

<i>Forest</i>) treinados na imagem de 23 de dezembro de 2016 e os valores reais (eixo y) das variáveis Cobertura 10, Profundidade (m) e Densidade (e/m).....	85
Figura 45 – Comparação entre as previsões (eixo x) com bandas do Sentinel 2 (08/01/2016) de modelos de regressão OLS (<i>ordinary least squares</i>) e RF (<i>Random Forest</i>) treinados na imagem de 23 de dezembro de 2016 e os valores reais (eixo y) das variáveis Cobertura 10, Profundidade e Densidade . Imagens em nível 2A (Reflectância de superfície) com o uso do sen2cor para correção atmosférica.	86
Figura 46 – Gráficos mostrando as reflectâncias (eixo x) das diferentes camadas de folhas observadas ao longo do espectro eletromagnético (eixo y).....	87
Figura 47 – Gráficos mostrando a diferença de reflectância (eixo y) entre 1 e 4 camadas de folhas ao longo do espectro eletromagnético em nanômetros (eixo x) com a posição das bandas 3, 5, 8A, 10, 11 e 12. Valores positivos (vermelho) significam aumento da reflectância com o número de folhas para o dado comprimento de onda, valores negativos (azul) significam o inverso. Abaixo, um recorte do início do <i>red-edge</i> mostrando a posição e largura da banda 5 (cinza) em detalhe.....	88
Figura 48 – Processo de construção de gráficos da comparação pixel a pixel (esquerda) e da comparação por buffers (direita) evidenciando uma das fontes de ruído presente no primeiro método.	94

LISTA DE QUADROS

Quadro 1 – Valores de elevação e azimute solar para as imagens do Sentinel 2 usadas no estudo.....	32
Quadro 2 – Índices de vegetação, fórmulas com bandas do Sentinel 2 e suas respectivas referências. B2: 490 nm, B4: 665 nm, B5: 705 nm, B6: 740 nm, B7: 783 nm, B8: 842 nm, B11: 1610 nm.....	38
Quadro 3 – Variáveis de campo utilizadas e suas respectivas descrições. Cada parcela amostrada tem um valor para cada uma das variáveis abaixo.....	41
Quadro 4 – Informações sobre os dados LiDAR usados no estudo para as áreas Cantareira 1 e 2.....	44
Quadro 5 – Número e percentual de pontos classificados como solo para as áreas Cantareira 1 e 2.....	44
Quadro 6 – Variáveis LiDAR utilizadas e suas respectivas descrições para cada pixel de 100 m ²	51
Quadro 7 – Especificações do espectrômetro FieldSpec 4 Hi-Re da ASD.	55
Quadro 8 – Valores de média, desvio padrão (Std) e coeficiente de variação (CV) para as variáveis derivadas dos dados de campo de todas as 34 parcelas. AM (Altura média das árvores amostradas), C (proporção de cobertura da parcela), DAPM (diâmetro à altura do peito médio das árvores amostradas), N (número de árvores amostradas na parcela), DAPT (diâmetro à altura do peito total das árvores amostradas), AT (altura total das árvores amostradas), AB (área basal total das árvores da parcela pela área de coleta). BSASM (Biomassa seca acima do solo média), BSAST (biomassa seca acima do solo total), AM > 5 (altura média das árvores maiores de cinco metros), N > 5 (número de árvores maiores de cinco metros), AMed (altura mediana das árvores amostradas), DAPMed (diâmetro à altura do peito mediano das árvores amostradas).	57
Quadro 9 – Fórmula, valor de r ² ajustado, RMSE e RRSE de alguns dos melhores modelos de regressão múltipla encontrados entre variáveis de estrutura coletadas em campo e imagens do Sentinel 2. A descrição das variáveis pode ser vista no Quadro 3 (seção 2.4). B1: 0,443 nm, B2: 0,490 nm, B3: 0,560 nm, B4: 0,665 nm, B5: 0,705 nm, B6: 740 nm, B7: 783 nm, B8: 842 nm, B8A: 865 nm, B9: 945 nm, B10: 1375 nm, B11: 1610 nm, B12: 2190 nm.....	66

Quadro 10 – Valores de média, desvio padrão (Std) e coeficiente de variação (CV) para as variáveis derivadas dos dados de campo de todos os 338 buffers aleatórios.	71
Quadro 11 – Lista contendo a fórmula e valor de r^2 ajustado, RMSE e RRSE de alguns dos melhores modelos de regressão OLS múltipla encontrados para variáveis de estrutura derivadas de dados LiDAR. A descrição das variáveis pode ser vista no Quadro 6 (seção 2.5). B1: 0,443 nm, B2: 0,490 nm, B3: 0,560 nm, B4: 0,665 nm, B5: 0,705 nm, B6: 740 nm, B7: 783 nm, B8: 842 nm, B8A: 865 nm, B9: 945 nm, B10: 1375 nm, B11: 1610 nm, B12: 2190 nm.....	76
Quadro 12 – Modelos de regressão <i>Random Forest</i> para cada variável derivada dos dados LiDAR com as quatro bandas mais importantes e os valores de r^2 <i>out of bag</i> , RMSE e RRSE. A descrição das variáveis pode ser vista no Quadro 6 (seção 2.5). B1: 0,443 nm, B2: 0,490 nm, B3: 0,560 nm, B4: 0,665 nm, B5: 0,705 nm, B6: 740 nm, B7: 783 nm, B8: 842 nm, B8A: 865 nm, B9: 945 nm, B10: 1375 nm, B11: 1610 nm, B12: 2190 nm.	77

LISTA DE ABREVIATURAS E SIGLAS

AB: Área Basal

AGB: *Above Ground Biomass* (Biomassa Acima do Solo)

AM: Altura Média

AMed: Altura Mediana

AM > 5: Altura média dos indivíduos maiores que 5 metros

AT: Altura Total

AVIRIS: *Airborne Visible / Infrared Imaging Spectrometer* (Espectrômetro Aerotransportado de Imageamento Visível / Infravermelho)

BSASM: Biomassa Seca Acima Do Solo Média

BSAST: Biomassa Seca Acima Do Solo Total

BOA: Bottom Of Atmosphere (Fundo da Atmosféra)

C: Porcentagem de cobertura

CASI-1500: *Compact Airborne Spectrographic Imager-1500* (Imageador Compacto Espectrográfico Aerotransportado-1500)

CBERS: *China-Brazil Earth-Resources Satellite* (Satélite Sino-Brasileiro de Recursos Terrestres)

CHM: *Canopy Height Model* (Modelo de Altura do Dossel)

CV: Coeficiente de variação

DAP: Diâmetro à altura do peito

DAPM: DAP Médio

DAPMed: DAP Mediano

DAPT: DAP Total

DOS1: *Dark Object Subtraction 1* (Subtração do Objeto Escuro 1)

EMBRAPA: Empresa Brasileira de Pesquisa Agropecuária

EMPLASA: Empresa Paulista de Planejamento Metropolitano

EO-1: *Earth Observing-1* (Observando a Terra-1)

ESA: *European Space Agency* (Agência Espacial Européia)

EVI: *Enhanced Vegetation Index* (Índice de Vegetação Melhorado)

GPS: *Global Positioning System* (Sistema de Posicionamento Global)

IRECI: *Inverted Red-Edge Chlorophyll Index* (Índice de Clorofila do Limite Vermelho Invertido)

L1TP: Level-1 Precision and Terrain (Nível-1 Precisão e Terreno)

LAI: *Leaf Area Index* (Índice de Área Foliar)

LCC: Leaf Chlorophyll Concentration (Concentração de Clorofila Foliar)

LiDAR: *Light Detection And Ranging* (Detecção de Luz e Telemetria)

DSM: *Digital Surface Model* (Modelo Digital de Superfície)

MGRS: *Military Grid Reference System* (Sistema de Referência de Grade Militar)

MSI: *Multispectral Instrument* (Instrumento Multiespectral)

N: Número de indivíduos

$N > 5$: Número de indivíduos maiores que 5 metros

NASA: *National Aeronautics and Space Administration* (Administração Nacional da Aeronáutica e Espaço)

NDI45: *Normalized Difference Index B4 and B5* (Índice de Diferença Normalizada B4 e B5)

NDVI: *Normalized Difference Vegetation Index* (Índice de Vegetação por Diferença Normalizada)

NDWI: *Normalized Difference Water Index* (Índice de Água por Diferença Normalizada)

NEdI: *Noise Equivalent Radiance* (Radiância Equivalente ao Ruído)

NIR: Near Infrared (Infravermelho Próximo)

OPALS: Orientation and Processing of Airborne Laser Scanning data (Orientação e Processamento de dados de Scanner Laser Aerotransportado)

RE1: red-edge 1, mais próximo do vermelho

RE2: red-edge 2, mais próximo do NIR

REP: *Red-edge Position* (Posição da Borda do Vermelho)

RF: *Random Forest* (Floresta Aleatória)

RVI: *Ratio Vegetation Index* (Índice de Vegetação por Razão Simples)

S2REP: *Sentinel 2 Red-edge Position* (Posição da Borda do Vermelho do Sentinel 2)

SAGA: *System for Automated Geoscientific Analysis* (Sistema para Análises Geocientíficas Automatizadas)

SAVI: *Soil Adjusted Vegetation Index* (Índice de Vegetação Ajustado ao Solo)

SCP: *Semi-Automatic Classification Plugin* (Complemento de Classificação Semi-Automática)

SWIR: Shortwave Infrared (Infravermelho de Ondas Curtas)

TIR: *Thermal Infrared* (Infravermelho Termal)

TOA: *Top Of Atmosphere* (Topo da Atmosfera)

VANT: Veículo Aéreo Não Tripulado

VIS: Espectro Visível

SUMÁRIO

1	INTRODUÇÃO	22
2	METODOLOGIA.....	27
2.1	ÁREA DE ESTUDO.....	27
2.2	IMAGENS UTILIZADAS	28
2.2.1	Sentinel 2.....	28
2.2.2	Landsat-8/OLI	32
2.3	ÍNDICES CALCULADOS	34
2.4	DADOS DE CAMPO	38
2.5	DADOS LIDAR.....	43
2.6	ANÁLISE DOS DADOS.....	52
2.7	ANÁLISE EM LABORATÓRIO: ESPECTROS DE REFLECTÂNCIA.....	55
3	RESULTADOS	57
3.1	ANÁLISE COM DADOS DE CAMPO	57
3.1.1	Análise geral dos dados	57
3.1.2	Relações com o Sentinel 2	60
3.2	ANÁLISE COM DADOS LIDAR	66
3.2.1	Análise geral dos dados	66
3.2.2	Relações com o Sentinel 2	73
3.2.3	Teste de modelos	79
3.3	ESPECTROS DE REFLECTÂNCIA.....	87
4	DISCUSSÃO	89
5	CONCLUSÃO E RECOMENDAÇÕES.....	101
	REFERÊNCIAS BIBLIOGRÁFICAS	103
	APÊNDICE A – CONSTRUÇÃO DAS VARIÁVEIS LIDAR	116
	APÊNDICE B – GRÁFICOS E QUADROS COMPLEMENTARES	118

1 INTRODUÇÃO

Pode-se dizer muito sobre o que acontece em um ambiente natural observando a vegetação que o cobre. As características e organização das plantas podem influenciar em diversas condições, recursos e processos naturais, tais como o ciclo hidrológico (BALBINOT et al. 2008, LIMA, 2017), velocidade do vento (SCHINDLER; BAUHUS; MAYER, 2012, ZHU et al., 2004), luminosidade (CANHAM et al. 1994) e diversidade de espécies e suas interações (TAKATA, 2017). Dentre as possíveis aplicações para esse conhecimento, estão a identificação do estágio sucessional de florestas (LONGHI et al. 2006), práticas de restauração e manejo florestal (ALVES; METZGER, 2006, GONÇALVES; SANTOS, 2008, SOUZA et al. 2004) e cultivos de espécies arbóreas como os de eucalipto (SILVEIRA; MOREIRA; HIGASHI, 2004).

Em muitos casos utiliza-se metodologias de campo para medir parâmetros biofísicos da vegetação como altura, DAP (Diâmetro à Altura do Peito), LAI (Índice de Área Foliar – *Leaf Area Index*), área basal, cobertura e formato do dossel, dentre outros (GUILHERME; MORELLATO; ASSIS, 2004, MORENO; NASCIMENTO; KURTZ, 2003, PRATA; ASSIS; JOLY, 2001, PRETZSCH et al., 2015, SOARES, 1999). De maneira geral, o trabalho de campo fornece informações bastante precisas sobre um local específico e proporciona uma compreensão que vai além dos dados medidos. Suas maiores desvantagens, no entanto, são o custo e a escala de cobertura. O trabalho de campo normalmente necessita de bastante esforço e acaba se tornando inviável em áreas muito grandes dependendo dos recursos disponíveis. Nesse sentido, pode-se contar com outra fonte de informações: o sensoriamento remoto.

Nas últimas décadas, as tecnologias de sensoriamento remoto avançaram significativamente, e hoje essas informações possuem considerável confiabilidade e ampla disponibilidade. Existe uma grande variedade de dados obtidos remotamente que podem ser úteis para o estudo da vegetação, e várias pesquisas têm sido desenvolvidas na área mostrando bons resultados nesse sentido (ABREU; COUTINHO, 2014; LI; XU; GUO, 2014; ROCCHINI; CHIARUCCI; LOISELLE, 2004, ROCCHINI; RICOTTA; CHIARUCCI, 2007; TAN, 2008). Dados de alta resolução espacial, por exemplo, podem ser muito úteis na estimativa de biomassa vegetal (LEBOEUF et al. 2007) e na delimitação de copas de árvores através de análises visuais ou algoritmos de segmentação (KE; QUACKENBUSH, 2011). Imagens

hiperespectrais, por sua vez, podem ser usadas para identificar as espécies presentes em uma área por meio de suas assinaturas espectrais. Carlson et al. (2007) e Longhitano, Amaral e Almeida (2009), por exemplo, mostram que é possível estimar a riqueza de espécies arbóreas em seus estudos utilizando imagens hiperespectrais aéreas obtidas com os sensores AVIRIS (3,6 metros, 224 bandas) e CASI-1500 (0,75 metros, 288 bandas), respectivamente, combinando alta resolução espacial e espectral para atingir bons resultados.

Outra tecnologia bastante promissora para aplicações florestais e com ampla utilização atualmente são os sensores LiDAR (*Light Detection and Ranging* - Detecção e Telemetria por Laser), que além de poderem ser aerotransportados, podem ser usados em terra para a coleta de dados de campo (DASSOT; CONSTANT; FOURNIER, 2011). Os aparelhos emitem vários pulsos de laser por segundo e detectam a distância da reflexão desses pulsos, criando uma nuvem de pontos tridimensional localizada espacialmente por um GPS. Além da posição e altura de cada ponto, outras informações podem ser armazenadas, como a intensidade da reflexão e o ângulo de incidência do laser (LEFSKY et al., 2002, LIM et al., 2003). O uso desses sensores permite a criação de modelos tridimensionais de alta precisão que podem ser aplicados em muitas situações, inclusive para áreas de floresta.

Estudos mostram que informações como altura, cobertura de dossel, número de árvores, biomassa, LAI e muitos outros parâmetros biofísicos podem ser medidos usando essa tecnologia (BOEHM; LIESENBERG; FRANK, 2010, CHEN et al. 2007, GETZIN et al. 2017, KHOSRAVIPOUR et al. 2014, OLIVEIRA et al. 2012, POPE; TREITZ, 2013, SABOL; PATOCKA; MIKITA, 2014, TSENG; LIN; WANG, 2016, ZHAO; POPESCU, 2009). Alonzo et al. (2015) por exemplo, mostra que estimativas de LAI feitas com LiDAR geram resultados parecidos aos de estimativas feitas com fotografias hemisféricas tiradas em campo (r^2 0,82). NÆSSET e ØKLAND também mostram bons resultados na estimativa de altura, com precisão de até 1,49 metros para árvores com média de 20 metros em áreas de floresta boreal. Além de estimar parâmetros biofísicos individuais da vegetação, é possível usar essa tecnologia na identificação de características estruturais de florestas, permitindo, por exemplo, a identificação de estágios de sucessão com acurácia superior a 95% com o uso de modelos RF (Floresta Aleatória - *Random Forest*) (FALKOWSKI et al., 2009). Por fim, é possível identificar relações entre imagens de sensores passivos e LiDAR,

possibilitando a realização de estimativas desses dados (PASCUAL et al. 2010; VIERLING et al. 2002).

Os exemplos de sensores citados acima são capazes de gerar informações de alta qualidade, mas embora a tecnologia tenha se tornado mais acessível com o passar dos anos, ainda é difícil obter essas informações para qualquer região do globo sem algum custo. Imagens multiespectrais, no entanto, podem ser obtidas facilmente de forma gratuita pela internet. Apesar de possuírem resolução espacial, temporal e espectral limitadas, estes dados normalmente possuem cobertura global e fazem parte de vastas coleções disponíveis para uso. Essas imagens passam por várias etapas de processamento antes de chegarem ao usuário final, o que facilita sua utilização. Alguns programas que disponibilizam imagens gratuitamente são o CBERS (Satélite Sino-Brasileiro de Recursos Terrestres - *China-Brazil Earth Resources Satellite*), uma parceria entre o Brasil e a China, o programa Landsat da NASA e o programa Sentinel da ESA (Agência Espacial Europeia – *European Space Agency*).

Pode-se usar essas imagens, por exemplo, para detectar mudanças na cobertura de vegetação ao longo do tempo (ENCINA, 2015), fazer estimativas com precisão razoável para riqueza de espécies vegetais (FOODY; CUTLER, 2006) e de concentração de clorofila (MAIRE et al. 2012, FRAMPTON et al. 2013, VERRELST et al. 2015). Já com relação a parâmetros biofísicos da vegetação propriamente ditos, a maioria dos estudos na literatura dizem respeito ao uso dessas imagens para a estimativa de variáveis ligadas à quantidade de folhas, como LAI (DELEGIDO et al. 2011, SUGAWARA et al. 2009). Poucos estudos, no entanto, exploram as relações com fatores como altura e DAP. Embora essas variáveis não causem alterações diretas na reflectância das plantas, estudos indicam a existência de relações entre elas e outras variáveis que afetam o espectro refletido pela vegetação, como LAI (SABOL; PATOCKA; MIKITA, 2014, XIAO et al. 2006) e biomassa, que pode ser calculada a partir de variáveis como altura e DAP (CHAVE et al., 2005). Assim, embora que com menor precisão, o uso de imagens multiespectrais para a estimativa de variáveis que afetam indiretamente a resposta espectral das plantas, como DAP e a altura da vegetação, é possível. Accioly e colaboradores (2002), por exemplo, observaram relações com r^2 superior a 0,3 para modelos de regressão entre índices de vegetação e altura medida em campo com imagens do Landsat-5/TM. Pascual et al. (2010) também mostram relações entre altura e imagens multiespectrais para

florestas de pinheiros em Madrid (Espanha). No estudo, a altura de dosséis derivada de dados LiDAR foi comparada a índices de vegetação calculados com imagens do Landsat-7/ETM+, gerando modelos com r^2 de até 0,76.

Estimativas de parâmetros biofísicos da vegetação usando imagens multiespectrais não tem a mesma precisão das feitas com LiDAR ou metodologias de campo, mas o baixo custo destas abordagens pode compensar para o uso em grandes áreas dependendo da precisão necessária. Dentre as imagens gratuitas que poderiam ser usadas para isso, uma alternativa com razoável resolução temporal, espacial e espectral são as imagens do Sentinel 2. O programa consiste de dois satélites lançados pela ESA e diversos estudos mostram seu potencial na obtenção de informações sobre parâmetros biofísicos da vegetação (DELEGIDO et al., 2011, FRAMPTON et al., 2013, MAJASALMI, RAUTIAINEN, 2016) e, mais recentemente, na classificação de estágios sucessionais (SOTHE et al., 2017). Embora não tenham resolução hiperespectral, os satélites do Sentinel 2 carregam sensores MSI (Instrumento Multiespectral – *Multispectral Instrument*), com quatro bandas estreitas de resolução espacial de 20 metros posicionadas na região entre o vermelho e NIR (Infravermelho Próximo – *Near Infrared*) no espectro eletromagnético (ESA, 2015). Essa faixa espectral, chamada de *red-edge* (680nm~ a 800nm~), mostra relação com LAI e LCC (Concentração de Clorofila nas Folhas – *Leaf Chlorophyll Concentration*) (DELEGIDO et al., 2011; FRAMPTON et al., 2013) e ainda pode ser relevante na discriminação de espécies vegetais (GONG; PU; YU, 1997, SOBHAN, 2007).

A importância dessa região do espectro vem de um comportamento de reflectância bastante característico da vegetação, que reflete pouco na região do vermelho, devido à absorção da clorofila, e muito na região do NIR, por conta do espalhamento causado pela estrutura celular das folhas (PONZONI; SHIMABUKURO; KUPLICH, 2012). A faixa *red-edge* (borda do vermelho) é onde se observa essa transição de baixa para alta reflexão, e variações nesse comportamento podem corresponder a diferenças nas características da vegetação observada (FERREIRA et al., 2013). Essa diferença na reflexão do NIR e do vermelho já foi muito explorada na forma de índices de vegetação como o NDVI (Índice de Vegetação por Diferença Normalizada – *Normalized Difference Vegetation Index*) (ROUSE et al., 1974), SAVI (Índice de Vegetação Ajustado ao Solo – *Soil Adjusted Vegetation Index*) (HUETE,

1988) e EVI (Índice de Vegetação Melhorado – *Enhanced Vegetation Index*) (HUETE et al., 2002).

Índices de vegetação são uma forma simples de utilizar a informação contida em várias bandas de forma conjunta. Estudos mostram que a utilização de índices de vegetação pode ter várias aplicações, tais como estimativas de biomassa vegetal (GOSWAMI et al., 2015), discriminação entre vegetação nativa e áreas de cultivo (RISSO et al., 2009, SILVA et al., 2015) e estimativas de LAI (DELEGIDO et al., 2011). Alguns índices são desenvolvidos especificamente para um sensor, como o S2REP (Posição da Borda do Vermelho para o Sentinel 2 – *Sentinel 2 Red Edge Position*) e o IRECI (Índice de Clorofila do Limite Vermelho Invertido – *Inverted Red-Edge Chlorophyll Index*), ambos criados por Frampton e colaboradores (2013) para a utilização com bandas do Sentinel 2 em estimativas de LCC e LAI.

Por ter imagens gratuitas, cobertura global, alta resolução temporal e uma boa resolução espacial e espectral na faixa *red-edge*, o Sentinel 2 é um dos programas mais promissores para estudos sobre vegetação da atualidade. Considerando esse potencial e a importância dessas informações, este estudo explora a possibilidade do uso de imagens do Sentinel 2 para estimar parâmetros biofísicos da vegetação em áreas de mata atlântica, buscando responder às seguintes perguntas:

1. Imagens do Sentinel 2 podem fornecer boas estimativas de parâmetros biofísicos da vegetação medidos em campo e com LiDAR?
2. As relações entre imagens do Sentinel 2 e parâmetros biofísicos da vegetação são influenciadas pelos efeitos da topografia?
3. Quais parâmetros biofísicos da vegetação podem ser melhor estimados usando imagens do Sentinel 2?
4. Quais índices de vegetação e combinações de bandas do Sentinel 2 apresentam relações mais fortes com parâmetros biofísicos da vegetação?

Adicionalmente, foi feito um experimento em laboratório medindo o efeito da reflectância da vegetação em diferentes camadas de folha com o objetivo de melhor explicar os resultados obtidos no estudo. Os procedimentos e dados utilizados para responder a essas perguntas são descritos em detalhes na seção 2.

2 METODOLOGIA

2.1 ÁREA DE ESTUDO

O estudo foi conduzido no estado de São Paulo (Brasil), entre as serras da Cantareira e da Mantiqueira. A região é caracterizada por diversos tipos de uso da terra, incluindo monoculturas, áreas urbanas, pasto, silvicultura e vegetação natural. Ademais, a área abrange uma parte do sistema Cantareira, que tem grande importância no abastecimento de água da região metropolitana de São Paulo e outros municípios do entorno (WHATELY, CUNHA, 2007). A região está inserida no bioma Mata Atlântica e compreende muitos remanescentes florestais. A área de estudo (Figura 1) envolve 34 parcelas com raio de 25 m nas quais foram coletadas informações sobre parâmetros biofísicos da vegetação em campo (TAKATA, 2017) e duas áreas em que foram coletadas informações através de sensores LiDAR aerotransportados (EMBRAPA, 2016) que serão descritas em detalhes respectivamente nas seções 2.4 e 2.5.

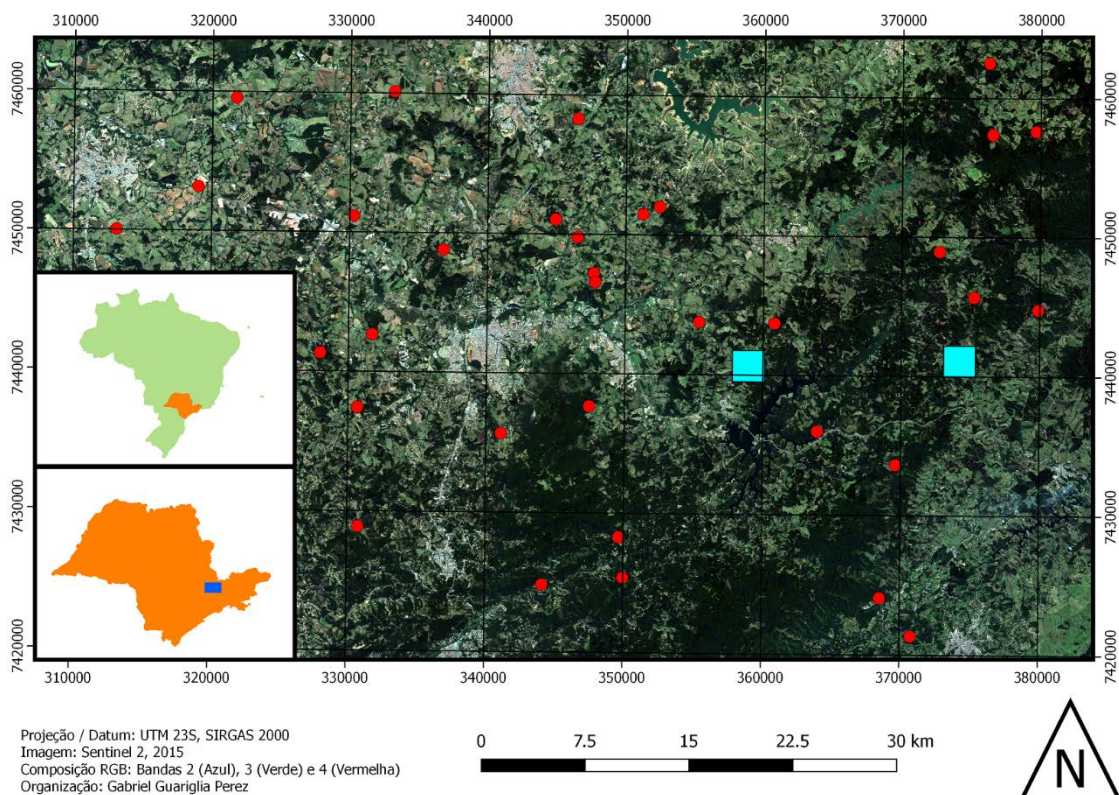


Figura 1 – Localização da área de estudo com as parcelas de campo (vermelho) e as áreas de coleta dos dados LiDAR (ciano). As parcelas de campo não estão em escala.

Vale destacar que como os dados de campo e os dados LiDAR foram obtidos de fontes diferentes, não existem parcelas de campo que correspondam exatamente às áreas com dados LiDAR, apesar de ambos os tipos de dados estarem localizados na mesma região. Assim, não foi possível a comparação direta entre essas duas informações de controle no presente estudo. Além disso, a metodologia de coleta de campo utilizada por Takata (2017) resultou em informações bastante diferentes das variáveis derivadas de LiDAR geradas, já que os dados de campo correspondem a plantas de diversos estratos arbóreos e os dados de LiDAR, em sua maioria, a informações sobre o dossel da floresta.

2.2 IMAGENS UTILIZADAS

2.2.1 Sentinel 2

A maioria das imagens utilizadas no estudo são dos satélites Sentinel 2 (SPOTO et al., 2012). Os satélites (A e B) carregam o sensor Multispectral Instrument (MSI), com 4 bandas de 10 metros, 6 bandas de 20 metros e 3 bandas de 60 metros de resolução espacial, conforme detalhado na Figura 2. As bandas do sensor estão localizadas aproximadamente entre 450 e 2200 nm e tem larguras variáveis. As imagens são disponibilizadas no chamado nível 1-C de processamento, com correções geométricas e radiométricas já aplicadas, de modo que as imagens são disponibilizadas ortoretificadas e em valores de reflectância TOA (Topo da Atmosfera – *Top of Atmosphere*). A resolução temporal é de 5 dias contando o tempo de passagem dos dois satélites da constelação (ESA, 2015).

O Sentinel 2 é especialmente promissor para se obter informações sobre vegetação por conta de sua resolução espectral, possuindo 4 bandas (bandas 5, 6, 7 e 8a) na faixa do *red-edge*, com 20 metros de resolução espacial cada. Além disso, as 4 bandas de 10 metros, que correspondem ao azul, verde, vermelho e NIR, (respectivamente B2, B3, B4 e B8) permitem o cálculo de alguns índices de vegetação em uma resolução espacial mais refinada. Com a combinação de resolução espacial e espectral do Sentinel 2, espera-se que os dados provenientes de seus sensores tenham melhores relações com parâmetros biofísicos da vegetação se comparados a

outros satélites com imagens multiespectrais. Sua maior desvantagem, no entanto, é que os satélites foram lançados recentemente e por isso não existem imagens anteriores a junho de 2015. Apesar disso, seu tempo curto de revisita deverá garantir uma boa disponibilidade de imagens no futuro.

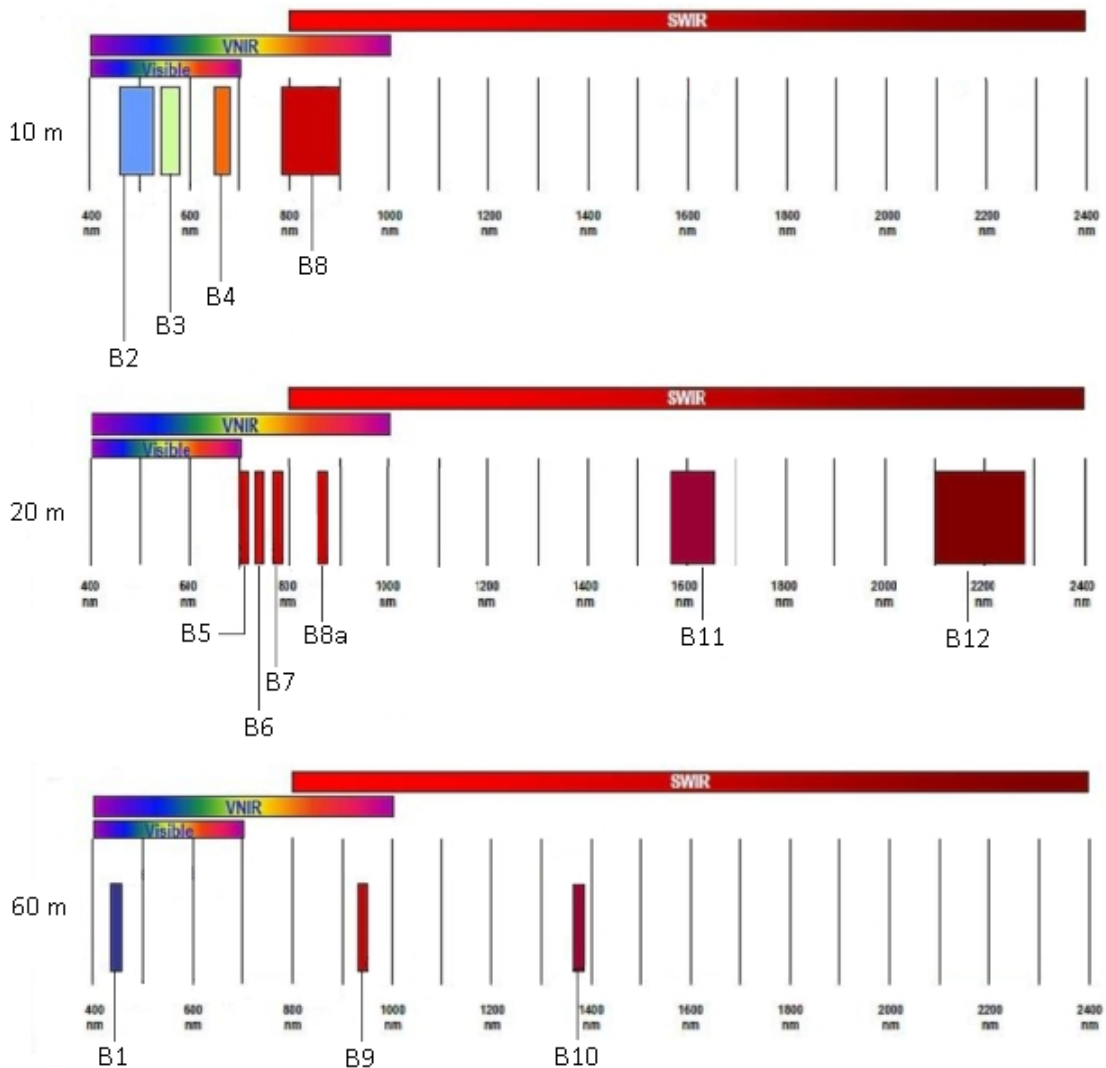


Figura 2 – Bandas das imagens do Sentinel 2 de acordo com sua resolução espacial. A imagem também mostra a largura de banda e a posição no espectro eletromagnético. Fonte: modificado de ESA.

As três imagens do Sentinel 2 utilizadas no estudo (Figura 3) estão localizadas na região 23KLQ do MGRS (Sistema de Referência de Grade Militar – *Military Grid Reference System*) e correspondem às datas 11 de Agosto de 2015, 08 de Janeiro de 2016 e 23 de Dezembro de 2016. Essas imagens foram escolhidas principalmente pela proximidade temporal com os dados de controle, mas por conta disso, as imagens obtidas têm uma grande quantidade de nuvens. Embora isso possa influenciar na radiometria das imagens por conta de efeitos de espalhamento, métodos de correção

atmosférica podem ser usados. Assim, para melhor atender aos objetivos do trabalho, as imagens passaram por processos de correção atmosférica e topográfica, já que o relevo e as condições atmosféricas podem causar alterações significativas na resposta espectral da vegetação percebida por um sensor orbital (GAIDA et al. 2016, MARCELLO et al. 2016).

Para importar as imagens para o QGIS (QGIS Development Team, 2018) e obter valores de reflectância de superfície (BOA – *Bottom Of Atmosphere*), foi usado o complemento SCP (Complemento de Classificação Semi-Automática - *Semi-Automatic Classification Plugin*) (CONGEDO, 2018), que aplica o método simplificado DOS1 (MORAN et al. 1992) de correção atmosférica ao importar as imagens para o programa. Esse método de correção foi escolhido por sua simplicidade e por estar disponível no complemento do QGIS mencionado. Além disso, o método não necessita de informações adicionais sobre a atmosfera do local, o que facilita bastante sua aplicação. Apesar de possuir algumas limitações quanto a correção de bandas em comprimento de ondas maiores (CHAVEZ, 1996), o método apresentou bons resultados em estudos anteriores (SANCHES et al., 2011, SONG et al., 2001) e espera-se que este seja suficiente para os propósitos deste estudo. Para verificar isso, também foi utilizado o complemento *sen2cor* do software SNAP desenvolvido pela ESA (MÜLLER-WILM; DEVIGNOT; PESSIOT, 2018), que foi desenvolvido especificamente para correção atmosférica de imagens do Sentinel 2 e aplica diversos métodos de correção. As imagens com a correção do *sen2cor* foram usadas somente em comparação com o método DOS1 para fazer a validação de alguns modelos gerados com imagens do Sentinel 2 e dados de LiDAR quando o objetivo foi o de testar a aplicação de modelos em imagens diferentes das usadas para treino.

A correção topográfica foi feita usando o algoritmo *Topographic correction* do SAGA (Sistema para Análises Geocientíficas Automatizadas - *System for Automated Geoscientific Analysis*) (CONRAD et al., 2015), também realizado dentro do QGIS, que utiliza o azimute e a elevação solar (Quadro 1) e um modelo de elevação para aplicar a correção. Como um dos parâmetros de entrada para o algoritmo utilizou-se um modelo digital de superfície com resolução de 5 metros disponibilizado pela EMPLASA (Empresa Paulista de Planejamento Metropolitano) (2010). O método de correção escolhido dentro do algoritmo de correção foi *Minnaert with slope* (LAW, NICHOL, 2004), com um coeficiente de Minnaert de 0,4 para as quatro primeiras

bandas e 0,5 para as demais. A escolha para esses parâmetros foi baseada na análise visual das imagens resultantes e na capacidade de redução na variância em índices de vegetação calculados com as imagens aplicando cada método de correção disponível no algoritmo (PEREZ; BOURSCHIEDT, 2018).

Ambos os conjuntos de imagens corrigidas e não corrigidas topograficamente foram usadas para comparação com os dados de estrutura da vegetação, com o objetivo de determinar se a correção topográfica melhora a relação entre as variáveis de uma maneira significativa.

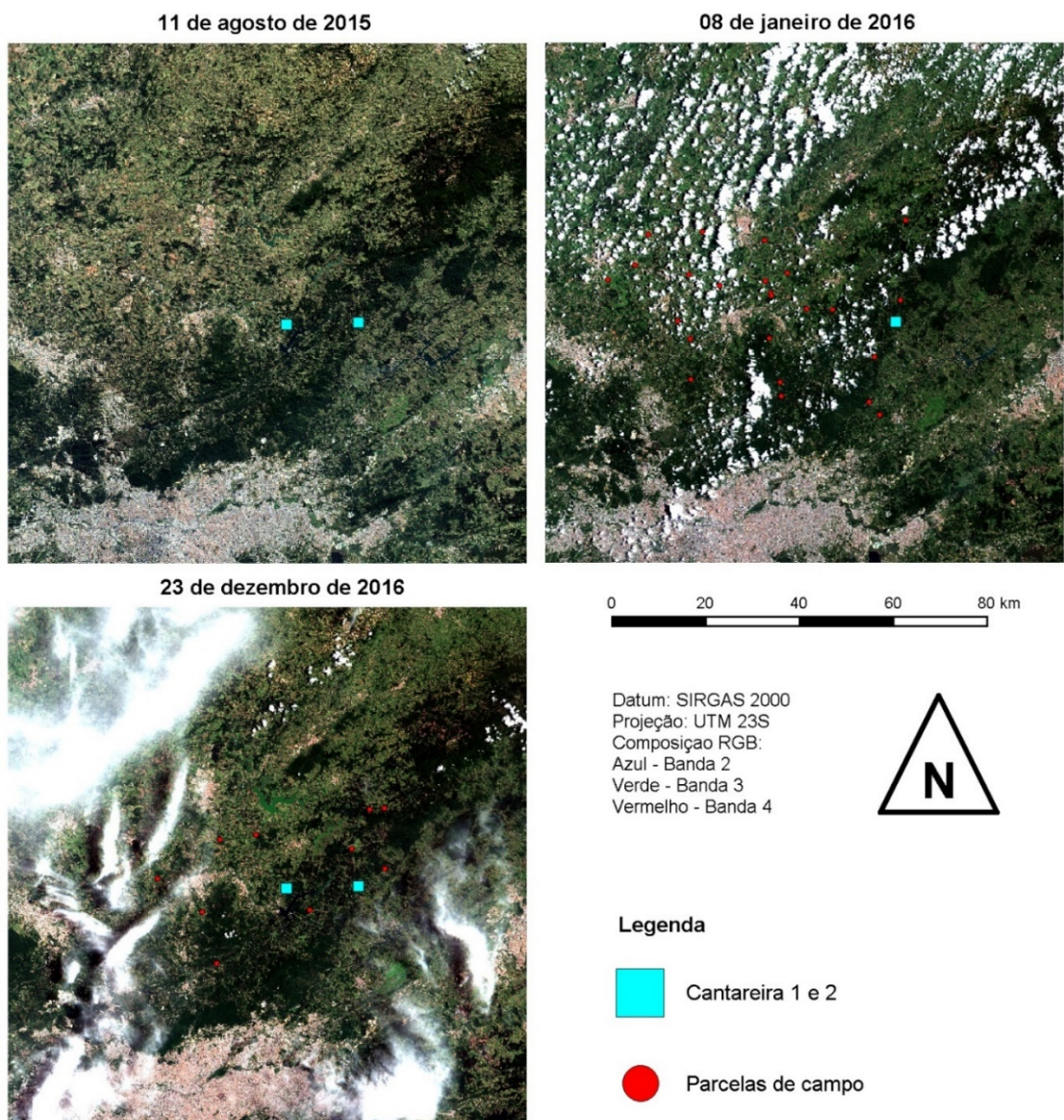


Figura 3 – Imagens do Sentinel 2 utilizadas no estudo com as respectivas parcelas de campo e áreas de coleta dos dados LiDAR. Nenhuma das áreas utilizadas se encontra em áreas cobertas por nuvens, embora estas possam ter influências na radiometria por conta de efeitos de espalhamento.

Quadro 1 – Valores de elevação e azimute solar para as imagens do Sentinel 2 usadas no estudo.

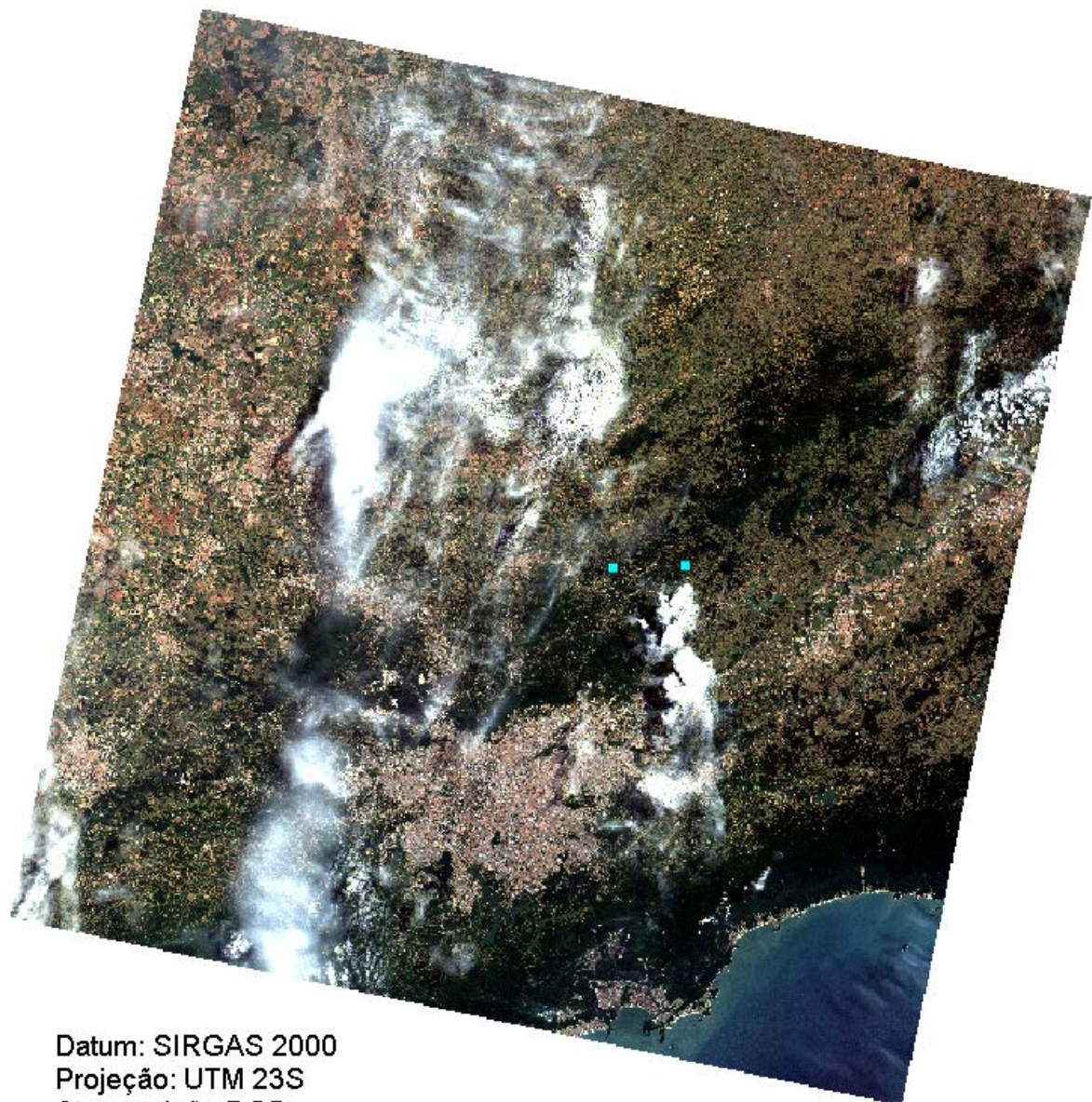
Data de Aquisição	Azimute solar (graus)	Elevação solar (graus)
11 de agosto de 2015	38,220	42,839
08 de janeiro de 2016	88,044	64,659
23 de dezembro de 2016	96,068	65,597

2.2.2 Landsat-8/OLI

Os dados de LiDAR (coletados em 30 de setembro e 04 de outubro de 2014) usados para comparação com as imagens do Sentinel 2 são anteriores a 2015, data em que o primeiro satélite do Sentinel 2 foi lançado. Assim, foi utilizada uma imagem do Landsat-8/OLI (USGS, 2016) de 06 de outubro de 2014 (Figura 4) para verificar se a variação temporal tem efeitos relevantes na resposta espectral da vegetação, já que esta imagem está bem próxima temporalmente das coletas dos dados de LiDAR. Por conta da presença de nuvens na Cantareira 2, apenas a área Cantareira 1 foi utilizada.

O tempo de revisita do Landsat-8/OLI é de 16 dias e suas imagens são compostas por 11 bandas. A primeira delas se posiciona em uma faixa de comprimento inferior ao azul, as bandas 2, 3 e 4 na faixa do visível, a banda 5 na região do NIR, as bandas 6, 7 e 9 na região do SWIR, as bandas 10 e 11 na faixa do TIR e uma banda pancromática abrangendo a região visível (banda 8). Todas as bandas têm resolução espacial de 30 m, com exceção das duas bandas termais (100 m) e da banda pancromática (15 m). Para o estudo, utilizou-se somente as bandas localizadas na região do visível, NIR e SWIR, com 30 m de resolução espacial. Assim como para as imagens do Sentinel 2, o complemento SCP foi usado para aplicar a correção atmosférica DOS1 (MORAN et al. 1992).

Além da correção atmosférica, foi aplicada uma correção topográfica em um processo similar ao que foi feito para as imagens do Sentinel 2, mas utilizando um coeficiente de Minnaert de 0,5 para todas as bandas. Os valores de elevação e azimute solar usados foram 58,086 e 60,099 respectivamente. A imagem disponibilizada pela USGS é de nível L1TP e, portanto, são disponibilizadas radiometricamente calibradas e ortoretificadas.



Datum: SIRGAS 2000
Projeção: UTM 23S
Composição RGB:
Azul - Banda 2
Verde - Banda 3
Vermelho - Banda 4

0 30 60 80 120 km



Cantareira 1 e 2



Figura 4 – Composição RGB da imagem do Landsat-8/OLI utilizada no estudo com a localização das áreas Cantareira 1 e 2 em que foram medidos os dados LiDAR. Apenas a área Cantareira 1 (esquerda) foi utilizada por conta da presença de nuvens na Cantareira 2 (direita).

2.3 ÍNDICES CALCULADOS

Com as bandas das imagens descritas, foram calculados os índices RVI (Índice de Vegetação por Razão – *Ratio Vegetation Index*), NDVI (Índice de Vegetação por Diferença Normalizada – *Normalized Difference Vegetation Index*), SAVI (Índice de Vegetação Ajustado ao Solo – *Soil Adjusted Vegetation Index*), EVI (Índice de Vegetação Melhorado – *Enhanced Vegetation Index*), NDWI (Índice de Água por Diferença Normalizada – *Normalized Difference Water Index*), NDI45 (Índice de Diferença Normalizada B4 e B5 – *Normalized Difference Index B4 and B5*), IRECI (Índice de Clorofila do Limite Vermelho Invertido – *Inverted Red-Edge Chlorophyll Index*) e S2REP (Posição da Borda do Vermelho para o Sentinel 2 – *Sentinel 2 Red Edge Position*). Esses índices foram escolhidos com base em um ou mais dos três critérios abaixo:

1. Ampla documentação na literatura e uso em diversas aplicações
2. Relação com um ou mais parâmetros biofísicos da vegetação
3. Índices específicos para uso com bandas do Sentinel 2

O RVI (JORDAN, 1969; PEARSON; MILLER, 1972) é o mais antigo dentre os índices escolhidos, além de ser um dos mais simples de ser calculado. O índice realça a diferença entre as bandas do vermelho e do infravermelho próximo, e por ser uma razão reduz interferências de mesma intensidade nas duas faixas. O RVI apresenta relações com abertura do dossel (SASAKI et al., 2008) e pode ser usado para mapear áreas de floresta (GOETZ et al., 2003), além de apresentar boas relações com LAI usando bandas simuladas do Sentinel 2 (bandas 4 e 8) (MAJASALMI; RAUTIAINEN, 2016). Como mostra a equação 1, o índice consiste de uma razão simples entre o Infravermelho próximo (NIR) e o vermelho (R).

$$RVI = \frac{NIR}{R} \quad (1)$$

O NDVI foi proposto por Rouse e colaboradores (1974) e é bastante utilizado até hoje para diversas finalidades, como detecção de áreas cobertas por vegetação e estimativas de biomassa (GONZÁLEZ-ALONSO et al., 2006), verificação do vigor e produção de culturas, estimativas de cobertura vegetal, dentre outras (ABREU, COUTINHO, 2014; MESQUITA et al. 2013; PARUELO et al. 2000). Muitos índices existentes utilizam princípios parecidos, como o NDI45 (DELEGIDO et al., 2011), que

substituí a banda NIR por bandas na faixa do RE e parece ter melhores relações com LAI usando imagens simuladas do Sentinel 2. Assumindo que NIR corresponde à faixa espectral do infravermelho próximo, RE à do *red-edge* e R à do vermelho, o NDVI e o NDI45 são definidos pelas equações 2 e 3 respectivamente.

$$NDVI = \frac{NIR - R}{NIR + R} \quad (2)$$

$$NDI45 = \frac{RE - R}{RE + R} \quad (3)$$

Uma das variações do NDVI que foi usada no estudo é o SAVI, proposto por Huete (1988). O índice acrescenta um fator de correção para a influência do solo, o que resulta em uma representação mais fiel da vegetação se comparado ao NDVI. Embora o índice tenha sido desenvolvido para uso em áreas de vegetação esparsa, o SAVI foi incluído principalmente por ser frequente utilizado na literatura (ARAUJO; SANTOS; SHIMABUKURO, 2000, SILVA, 2015, XUE; SU, 2017) e por ser aparentemente uma melhor alternativa se comparado ao NDVI para uso em áreas de floresta densa, mostrando menos problemas de saturação (HUETE et al., 1997). Além disso, um fator de correção (L), que pode variar de 0 (áreas com muita vegetação) a 1 (áreas com pouca vegetação) pode ser usado para essas áreas. Com $L = 0,25$, o cálculo do SAVI pode ser visto na equação 4.

$$SAVI = \frac{NIR - R}{NIR + R + L} * (1 + L) \quad (4)$$

Além do SAVI, utilizou-se o EVI, que compensa pelos efeitos da atmosfera na banda do vermelho utilizando a banda do azul pra correção. Foi desenvolvido por Huete e colaboradores (2002), e diferentemente do NDVI e do SAVI ele não é normalizado. O índice foi desenvolvido para uso com o sensor MODIS (*Moderate Resolution Imaging Spectroradiometer*), mas como foi desenvolvido a partir de imagens Landsat TM 4 e 5 degradadas, tem sido utilizado com sucesso para essas imagens (GINCIENE, BITENCOURT, 2011). Neste estudo, o EVI foi calculado de forma similar ao que foi feito por Majasalmi e Rautiainen (2016, p. 431) com bandas simuladas do Sentinel 2, mas com uso da banda 8 no lugar da banda 7 para manter uma melhor resolução espacial. Dentre as aplicações do EVI, podemos citar a

diferenciação de áreas de floresta, agricultura e solo (RISSO et al. 2009) e estimativas de LAI (ROBERTS, ROTH, PERROY, 2011). A equação 5 descreve o cálculo do EVI, com os coeficientes de correção sendo C1 igual a 6, C2 igual a 7,5 e a banda do azul representada por B.

$$EVI = \frac{NIR - R}{NIR + C1 * R - C2 * B + L} \quad (5)$$

Outro índice que foi utilizado é o NDWI, um índice criado para a detecção de água em plantas. O índice é normalizado e resulta em valores entre 1 e -1. Sua fórmula é semelhante à do NDVI, mas ao invés de utilizar a faixa do vermelho o NDWI utiliza o SWIR. O nome foi criado por Gao (1996), mas o índice foi utilizado pela primeira vez por Kimes et al. (1981) com o nome de ND45. A diferença entre o NDWI e o ND45 também está na faixa espectral do SWIR que utilizam (1250 nm para o NDWI e 1600 nm para o ND45). Neste estudo utilizou-se o ND45 por conta das bandas disponíveis, mas o nome será mantido como NDWI para evitar confusões com o NDI45. Para tal, fez-se uso de B11 (1610 nm) como SWIR e B8 (842 nm) como NIR para imagens do Sentinel 2. A equação 6 mostra o cálculo do índice.

$$NDWI = \frac{NIR - SWIR}{NIR + SWIR} \quad (6)$$

O IRECI, criado por Frampton e colaboradores (2013), foi desenvolvido especialmente para a aplicação em imagens do Sentinel 2 usando as bandas do vermelho, *red-edge* e do NIR. No estudo, o IRECI foi o de melhor desempenho dentre os índices testados para estimar LAI e a quantidade de clorofila nos dosséis, com r^2 superior a 0,8 na maioria dos casos para ambas as variáveis. Considerando que RE1 e RE2 correspondem a faixas do red-edge mais próxima do R e do NIR respectivamente, sua expressão é dada pela equação 7.

$$IRECI = \frac{NIR - R}{\frac{RE1}{RE2}} \quad (7)$$

O último índice considerado é o REP. O índice visa mostrar no espectro eletromagnético onde se encontra o ponto de maior inclinação entre a menor e a maior reflectância nas regiões do vermelho e NIR, respectivamente. Existem várias maneiras de se calcular o REP, desde o uso de modelos lineares simples até modelos

de maior complexidade (BARANOSKI, 2005). No estudo de Frampton e colaboradores (2013) o índice, chamado de S2REP, não apresentou resultados tão bons quanto o IRECI de maneira geral, mas teve uma correlação de 91% para LCC (Concentração de Clorofila nas Folhas – *Leaf Chlorophyll Concentration*) em uma das áreas do estudo. Por ser um índice especialmente adaptado para a aplicação em dados do Sentinel 2, o S2REP foi utilizado no estudo de acordo com a equação 8, que pode ser vista a seguir.

$$S2REP = 705 + 35 * \frac{\frac{NIR - R}{2} - 5}{RE2 - RE1} \quad (8)$$

Todas as equações dos índices descritos podem ser vistas de forma adaptada para as bandas do Sentinel 2 no Quadro 1. Em adição aos índices descritos, foram calculadas todas as razões possíveis entre duas bandas, de maneira similar ao que foi feito por Majasalmi e Rautiainen (2016). As razões foram calculadas de maneira automática por meio de um algoritmo criado utilizando linguagem de programação científica *python* através do ambiente de desenvolvimento integrado (IDE) *spyder*. Cada razão foi nomeada como "X por Y", onde X é o nome da banda numeradora e Y é o nome da banda denominadora. Por exemplo, a razão entre B7 e B5 foi nomeada de "B7 por B5". Não foram calculadas as razões para as mesmas bandas (ex. B1 por B1) e para os inversos de pares de bandas (ex. B5 por B7). Assim, excluindo combinações entre as mesmas bandas (13) e entre pares repetidos (78) de todas as combinações possíveis (169), 78 combinações distintas foram obtidas de acordo com a equação 9 (n = número de bandas a serem combinadas, 13 no caso do Sentinel 2). Usando a mesma equação, no caso do Landsat-8/OLI foram obtidas um total de 15 combinações distintas (já que foram utilizadas apenas 6 bandas).

$$\text{número de combinações} = \frac{n^2 - n}{2} \quad (9)$$

Quadro 2 – Índices de vegetação, fórmulas com bandas do Sentinel 2 e suas respectivas referências. B2: 490 nm, B4: 665 nm, B5: 705 nm, B6: 740 nm, B7: 783 nm, B8: 842 nm, B11: 1610 nm.

Índice	Fórmula	Referência
NDVI	$(B8 - B4) / (B8 + B4)$	Rouse et al. (1974)
NDI45	$(B5 - B4) / (B5 + B4)$	Delegido et al. (2011)
RVI	$B8 / B4$	Jordan (1969)
SAVI	$((B8 - B4) / (B8 + B4 + 0,25)) * (1 + 0,25)$	Huete (1988)
EVI	$G * (B8 - B4) / (B8 + C1 * B4 - C2 * B2 + L)$	Huete et al. (2002)
NDWI	$(B8 - B11) / (B8 + B11)$	Kimes et al. (1981)
S2REP	$705 + 35 * ((B7 + B4) / 2 - B5) / (B6 - B5)$	Frampton et al. (2013)
IRECI	$B7 - B4 / (B5 / B6)$	Frampton et al. (2013)

2.4 DADOS DE CAMPO

Todos os dados de campo utilizados foram disponibilizados por Takata (2017) e foram coletados no verão entre os anos de 2014, 2015, 2016 e 2017. As informações foram coletadas originalmente para observar a influência da estrutura da vegetação nas interações de plantas e polinizadores. As coletas foram realizadas em diversos fragmentos de mata atlântica na área de estudo (Figura 1 – seção 2.1). Os dados foram coletados em um total de 43 parcelas, evitando locais com clareiras, mas apenas 34 foram utilizadas neste estudo por conta da presença de nuvens nas imagens. Destas 34 parcelas, 11 foram amostradas entre 2014 e 2015, 13 entre 2015 e 2016 e 10 entre 2016 e 2017. Os pontos centrais das parcelas, registrados com aparelhos GPS (Gramin, eTrex 10, 20 e 30), foram instalados a aproximadamente 75 metros para o interior dos fragmentos com o intuito de minimizar efeitos de borda (HARPER et al., 2005). Os aparelhos GPS utilizados possuem precisão próxima de 3 metros em áreas abertas (GRAMIN, 2015), mas a cobertura das florestas pode aumentar a imprecisão. Embora não existam muitos estudos sobre esse efeito em áreas de mata atlântica para os aparelhos utilizados em específico, Rodríguez-Pérez, Álvarez e Sanz-Ablanedo (2007) mostram que mesmo modelos eTrex mais antigos não apresentam acurácia inferior a 10 metros sob o dossel de florestas de *pinus*.

O desenho amostral de Takata (2017) consistiu no estabelecimento de parcelas onde foram medidos diversos parâmetros biofísicos e de estrutura da vegetação. As parcelas estabelecidas têm um formato hexagonal, com 25 metros do ponto central até cada um dos vértices, conforme mostrado na Figura 5. As medidas de estrutura

foram todas obtidas ao longo de três vértices em um raio de um metro, totalizando uma área de aproximadamente 150m² amostrada por hexágono. Nas parcelas foram obtidas medidas de altura total das árvores, DAP (Diâmetro à Altura do Peito) e proporção de cobertura do dossel. Todas as medidas foram tiradas de plantas com um perímetro maior que 5 centímetros a uma altura de 1,3 metros, as de perímetro menor sendo consideradas como parte do sub-bosque. O DAP de cada planta foi calculado a partir de seu perímetro (perímetro / π) medido com uma trena a uma altura de 1,3 metros seguindo a metodologia descrita por Newton (2007). A partir das medidas de DAP foi calculada a área basal (m²) de cada indivíduo amostrado usando a equação 9 (REID; STEPHEN, 1999 apud TAKATA, 2017).

$$\text{área basal (m}^2\text{)} = \pi * \left(\frac{\text{DAP}}{200}\right)^2 \quad (9)$$

A altura foi estimada de forma visual tendo como base uma referência medindo 1 metro e corresponde à altura total da árvore (da base até o ponto mais alto). As estimativas sempre foram feitas pela mesma pessoa para minimizar erros. A proporção de cobertura do dossel foi feita com o uso de um densiômetro esférico côncavo. Os valores obtidos são uma média das medidas feitas em sete pontos da parcela (Figura 5) no fim da tarde, com quatro medidas feitas por ponto em diferentes orientações (Norte, Sul, Leste e Oeste). As medidas de cobertura também foram feitas a 1,3 metros do solo e pela mesma pessoa para minimizar erros nas estimativas.

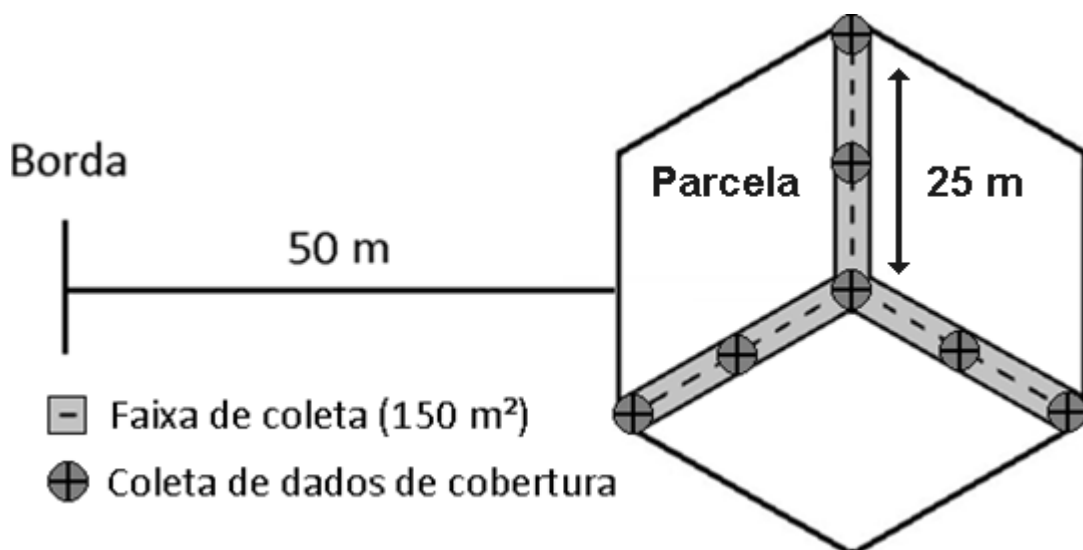


Figura 5 – Exemplo de parcela utilizada na coleta dos dados de campo. A área em cinza representa a área de coleta (3 vértices de 25 por 2 metros, ou 150 m²). Os pontos com cruz são os locais da coleta de dados de cobertura da vegetação.

As variáveis utilizadas no presente estudo foram todas derivadas das quatro informações descritas acima (DAP, área basal, altura e proporção de cobertura). Ao todo foram calculadas treze variáveis (Quadro 2): AM (altura média), AMed (altura mediana), AM > 5 (altura média dos indivíduos maiores que 5 metros), AT (altura total), DAPM (DAP médio), DAPMed (DAP mediano), DAPT (DAP total), N (número de indivíduos), N > 5 (número de indivíduos maiores que 5 metros), C (porcentagem de cobertura), AB (área basal por área amostrada), BSASM (biomassa seca acima do solo média), BSAST (biomassa seca acima do solo total).

As variáveis AM, AMed e AT correspondem respectivamente à média, mediana e somatória dos valores de altura de todos os indivíduos amostrados por parcela em metros. Da mesma maneira o DAPM, DAPMed e DAPT correspondem respectivamente à média, mediana e somatória dos valores de DAP de cada árvore amostrada por parcela em centímetros. A variável N é o número de indivíduos amostrados por parcela e a variável C é a porcentagem de cobertura do dossel. O termo “> 5” no nome das variáveis simboliza que a medida foi feita apenas para árvores com mais de cinco metros de altura. A variável AB é a somatória das áreas basais das árvores amostradas em uma parcela (m²) dividida pela área de coleta (ha) e corresponde à área ocupada pelo tronco das árvores a uma altura de 1,3 metros na área de coleta (m²/ha).

Por fim, as variáveis BSASM e BSAST correspondem respectivamente aos valores totais e médios estimados de biomassa seca acima do solo para as árvores amostradas (kg). A biomassa foi calculada usando a equação 10 (CHAVE, 2005) para estimativas de BSAS (biomassa seca acima do solo) em áreas de floresta tropical húmida, testada com sucesso para áreas de mata atlântica em específico (OLIVEIRA et al., 2013, VIEIRA et al., 2008).

$$BSAS = \exp(-2,977 + \ln(p * DAP^2 * A)) \quad (10)$$

O cálculo envolve a altura (A), o DAP e a densidade da madeira (p) de cada indivíduo. Vale destacar que o valor de densidade da madeira não foi amostrado, mas foi usado um valor fixo de 0,7 kg/m³ de acordo com valores encontrados na literatura para a aplicação em modelos alométricos de estimativa de biomassa em áreas de mata atlântica (CHAVE et al., 2006, NETO et al., 2013). Como valores específicos de

densidade de madeira para os indivíduos amostrados são considerados de grande importância no cálculo da biomassa, as medidas obtidas não devem ser usadas como valores de referência para áreas de mata atlântica e são incluídas aqui apenas como medidas comparativas entre as parcelas amostradas.

Quadro 3 – Variáveis de campo utilizadas e suas respectivas descrições. Cada parcela amostrada tem um valor para cada uma das variáveis abaixo.

Variável	Descrição	unidade
AM	Média aritmética das alturas das árvores amostradas.	m
AMed	Mediana das alturas das árvores amostradas.	m
AM > 5	Média aritmética das alturas das árvores maiores que 5 metros amostradas.	m
AT	Somatória das alturas das árvores amostradas.	m
DAPM	Média aritmética do DAP das árvores amostradas.	cm
DAPMed	Mediana do DAP das árvores amostradas.	cm
DAPT	Somatória do DAP das árvores amostradas.	cm
N	Número de árvores amostradas.	-
N > 5	Número de árvores maiores que 5 metros amostradas.	-
C	Proporção de cobertura do dossel medido pela entrada de luz.	-
AB	Área basal total das árvores amostradas dividida pela área de coleta.	m ² /ha
BSASM	Biomassa seca acima do solo média para todas as árvores amostradas	kg
BSAST	Biomassa seca acima do solo total para todas as árvores amostrados	kg

Para verificar se as imagens do Sentinel 2 têm relação com as variáveis medidas, elas foram comparadas aos dados de campo descritos acima. Para buscar a correspondência espacial das parcelas de campo nas imagens, os pontos centrais dos hexágonos foram importados para o QGIS e, para cada ponto, foi feito um *buffer* de 25 metros de raio (Figura 6). Os *buffers* em áreas de nuvens foram descartados e os demais foram usados para o recorte das imagens das bandas e índices. Essas imagens recortadas foram usadas em comparação com dados de campo utilizando linguagem de programação científica *python* através do IDE *spyder*.

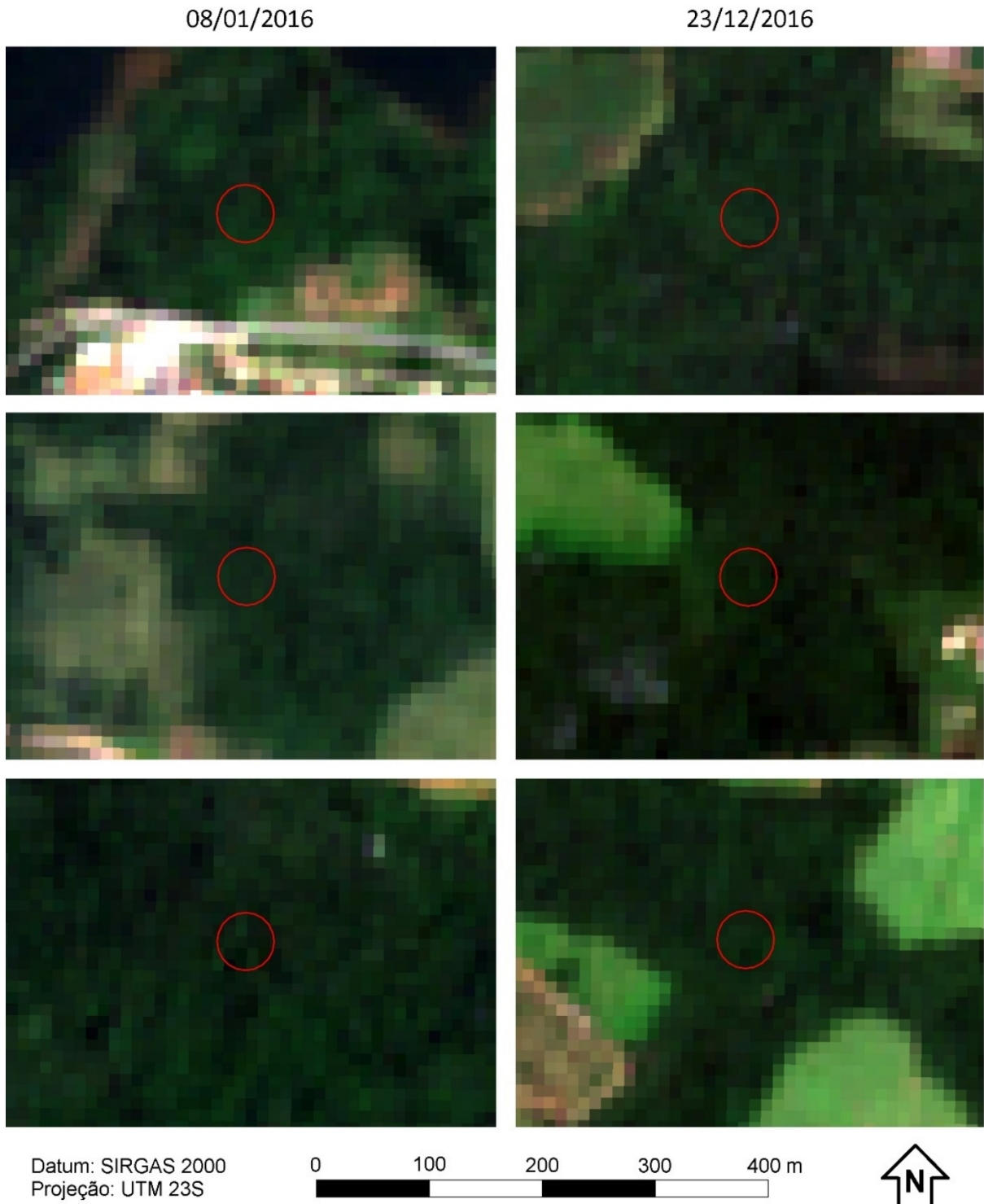


Figura 6 – Recortes da composição RGB (bandas 2, 3 e 4) das imagens do Sentinel 2 com detalhe de alguns dos buffers gerados (vermelho) com 25 metros de raio para representar as parcelas de campo. As imagens da esquerda são de 8 de janeiro de 2016 e as da direita de 23 de dezembro de 2016.

2.5 DADOS LIDAR

Os dados LiDAR utilizados no estudo foram disponibilizados pelo projeto Paisagens Sustentáveis, uma parceria entre a EMBRAPA e o serviço florestal dos EUA (EMBRAPA, 2016). Os dados podem ser obtidos no endereço eletrônico <https://www.paisagenslidar.cnptia.embrapa.br/webgis/> mediante requisição de acesso à plataforma. A coleta de dados foi realizada nos meses de setembro e outubro de 2014 e correspondem a duas áreas, denominadas de Cantareira 1 e 2 (Figura 7). As áreas possuem ocupação diversa, incluindo pastagens, agricultura, silvicultura e fragmentos de vegetação arbórea nativa. As áreas não correspondem ao mesmo local das parcelas de campo, embora estejam localizadas na mesma região, conforme já apresentado na seção 2.1. Informações gerais sobre os dados LiDAR de ambas as áreas podem ser vistas no quadro 4.

Usando metodologias similares às aplicadas a dados LiDAR na literatura para estimativas de altura, LAI e outros parâmetros biofísicos (CHEN et al., 2007, FALKOWSKI et al. 2009, GETZIN et al. 2017, ZHAO; POPESCU, 2009), foram geradas inúmeras variáveis relacionadas a parâmetros biofísicos da vegetação para a comparação com as imagens do Sentinel 2, que serão descritas a seguir.

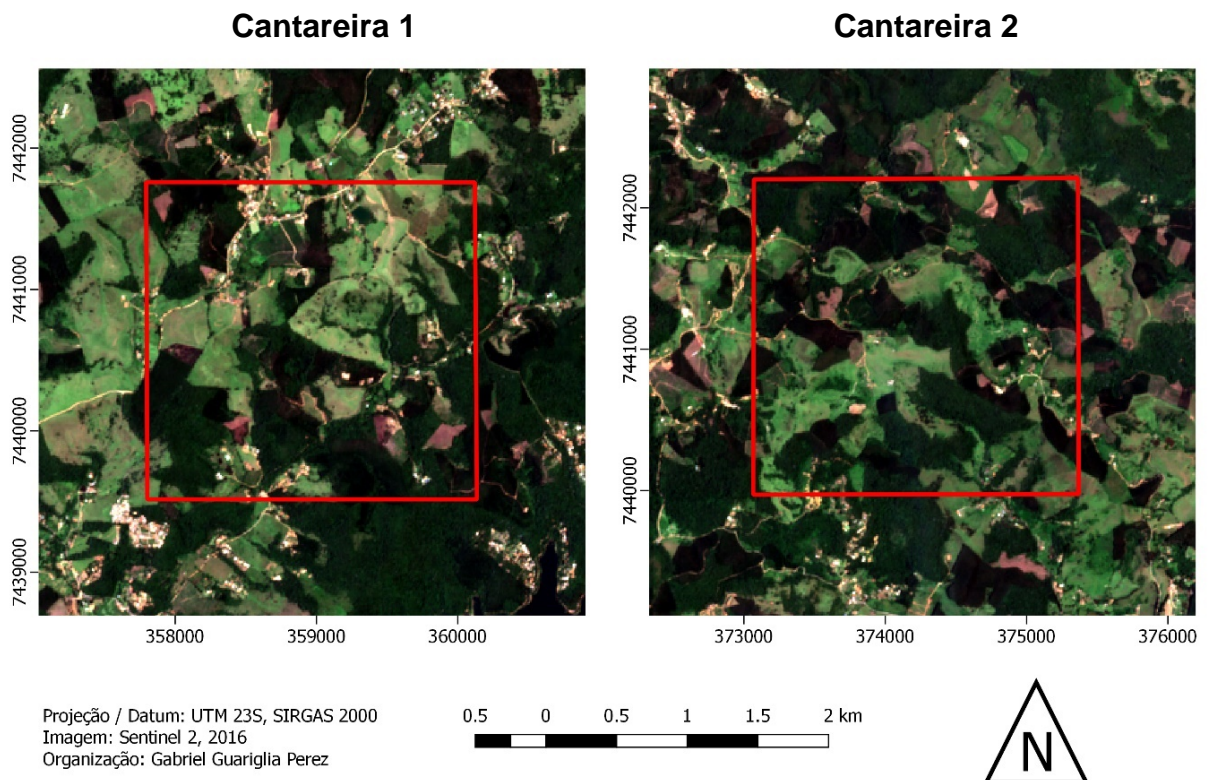


Figura 7 – Localização e delimitação (em vermelho) das áreas Cantareira 1 e 2

Quadro 4 – Informações sobre os dados LiDAR usados no estudo para as áreas Cantareira 1 e 2.

Nome	Cantareira 1	Cantareira 2
Data de aquisição	04/10/2014	30/09/2014 e 04/10/2014
Localização	Nazaré Paulista (SP)	
Datum	SIRGAS 2000	
Projeção	UTM 23 S	
Área	499 ha	
Total de pontos	172613355	197680865
Densidade	34,62 pontos / m ²	39,62 pontos / m ²
Densidade (1 ^{os} retornos)	17,6 pontos / m ²	20,01 pontos / m ²
Altura média de voo	853 m	
Campo de visão	10°	
Scanner	OPTECH, ORION, 09SEN243	
Frequência do scanner	67,5 Hz	
Instrumento GNSS	APPLANIX, 09SEN243	
Frequência GNSS	5Hz	
Instrumento IMU	LITTON, 413996	
Frequência IMU	100 kHz	
Sobreposição de voo	65%	

Primeiramente, a nuvem de pontos LiDAR foi convertida em um CHM (modelo de altura de dossel). O procedimento consistiu na utilização do software OPALS (PFEIFER et al., 2014) para fazer um DSM (Modelo Digital de Superfície – *Digital Surface Model*) considerando apenas os pontos LiDAR classificados como solo (os dados foram disponibilizados com os pontos já classificados) e um DSM considerando apenas os pontos referentes aos primeiros ecos. A quantidade e porcentagem de pontos classificados como solo podem ser vistas no Quadro 5.

Quadro 5 – Número e percentual de pontos classificados como solo para as áreas Cantareira 1 e 2.

Pontos	Cantareira 1	Cantareira 2	Total
Total	172613355	197680865	370294220
Pontos de solo	13250870	11575173	24826043
Pontos de solo (%)	7,68	5,86	6,70

Os dois modelos gerados foram carregados no QGIS e usou-se a ferramenta calculadora raster para subtrair a elevação dos ecos de solo da elevação dos primeiros ecos, gerando uma imagem correspondente à altura da superfície relativa ao solo, o que equivale à altura dos dosséis em áreas com vegetação (Figura 8). O mesmo procedimento foi adotado para o DSM gerado a partir da elevação dos últimos ecos.

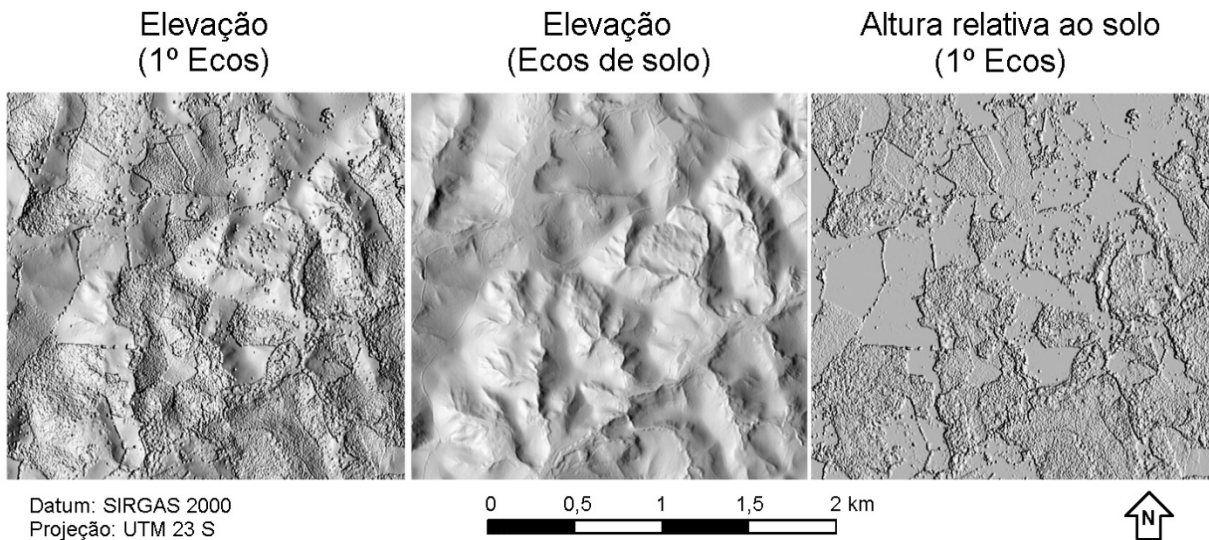


Figura 8 – Subtração da imagem de elevação dos ecos do solo (imagem central) da elevação dos primeiros ecos (imagem à direita), resultando em uma imagem que pode ser usada como um modelo de altura do dossel (CHM) (imagem à esquerda). As imagens correspondem à área Cantareira 1.

Usando o complemento *Terrain profile* do aplicativo QGIS, foram traçados perfis topográficos nos sentidos Norte – Sul e Oeste – Leste passando pela região central das áreas Cantareira 1 e 2. Os perfis foram traçados para os DSMs gerados a partir dos ecos de solo e a partir dos últimos ecos, comparados na Figura 9. Nota-se que o DSM dos ecos de solo estão sempre abaixo ou tem o mesmo valor dos DSM dos últimos ecos, e que mesmo áreas com vegetação (em vermelho) parecem não afetar o DSM gerado a partir dos ecos de solos. Ademais, foram feitos histogramas das imagens de altura dos últimos ecos (DSM dos últimos ecos - DSM dos ecos de solo) para identificar a proporção de pixels com valores negativos nas áreas de vegetação estudada (Figura 10). Esses dois procedimentos tiveram o objetivo de verificar a qualidade do DSM dos ecos de solo e tem grande importância já que é a referência de elevação utilizada para o cálculo do CHM e da altura dos últimos ecos.

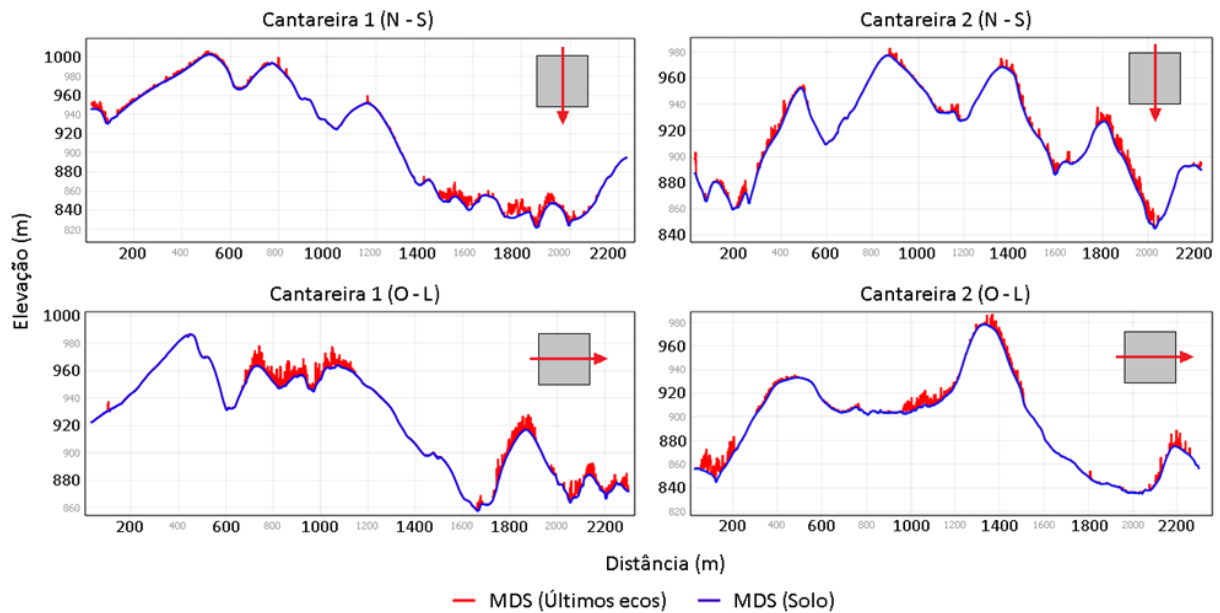


Figura 9 – Perfis topográficos passando pelo centro das áreas Cantareira 1 e 2 nos sentidos Norte-Sul e Oeste-Leste. As linhas azuis representam os modelos digitais de superfície (DSM) gerados a partir dos ecos classificados como solo e as linhas vermelhas os DSM gerados a partir dos últimos ecos. As setas vermelhas no canto superior direito de cada gráfico representam o sentido e a posição em que foram traçados os perfis topográficos em relação às áreas Cantareira 1 e 2 (em cinza), que podem ser vistas na Figura 7.

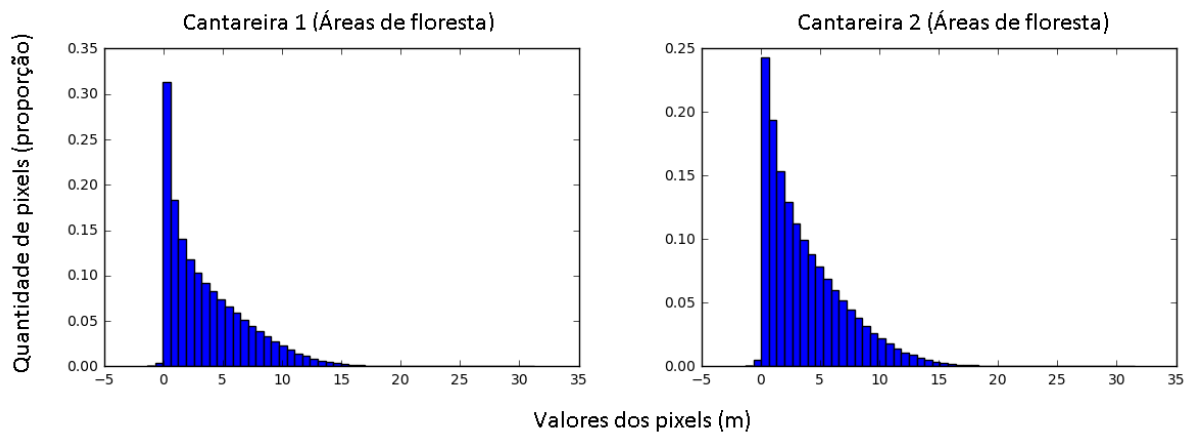


Figura 10 – Histograma para as áreas Cantareira 1 (esquerda) e 2 (direita) mostrando a distribuição dos pixels da imagem de altura dos últimos ecos de acordo com seus valores nas áreas de vegetação. Os pixels negativos correspondem a menos de 1% do total de pixels para ambas as áreas e não tem valores menores que -1,3 m.

Além das imagens de altura dos primeiros (CHM) e últimos ecos, foram geradas imagens do número de ecos por pulso de laser emitido, sendo essas também geradas usando o OPALS. Assim, as três imagens principais obtidas foram altura dos primeiros ecos, altura dos últimos ecos e quantidade de ecos, todas com 0,5 metros de

resolução espacial (Figura 11). Destas três imagens foram geradas todas as variáveis derivadas de LiDAR usadas no estudo, descritas em detalhe a seguir.

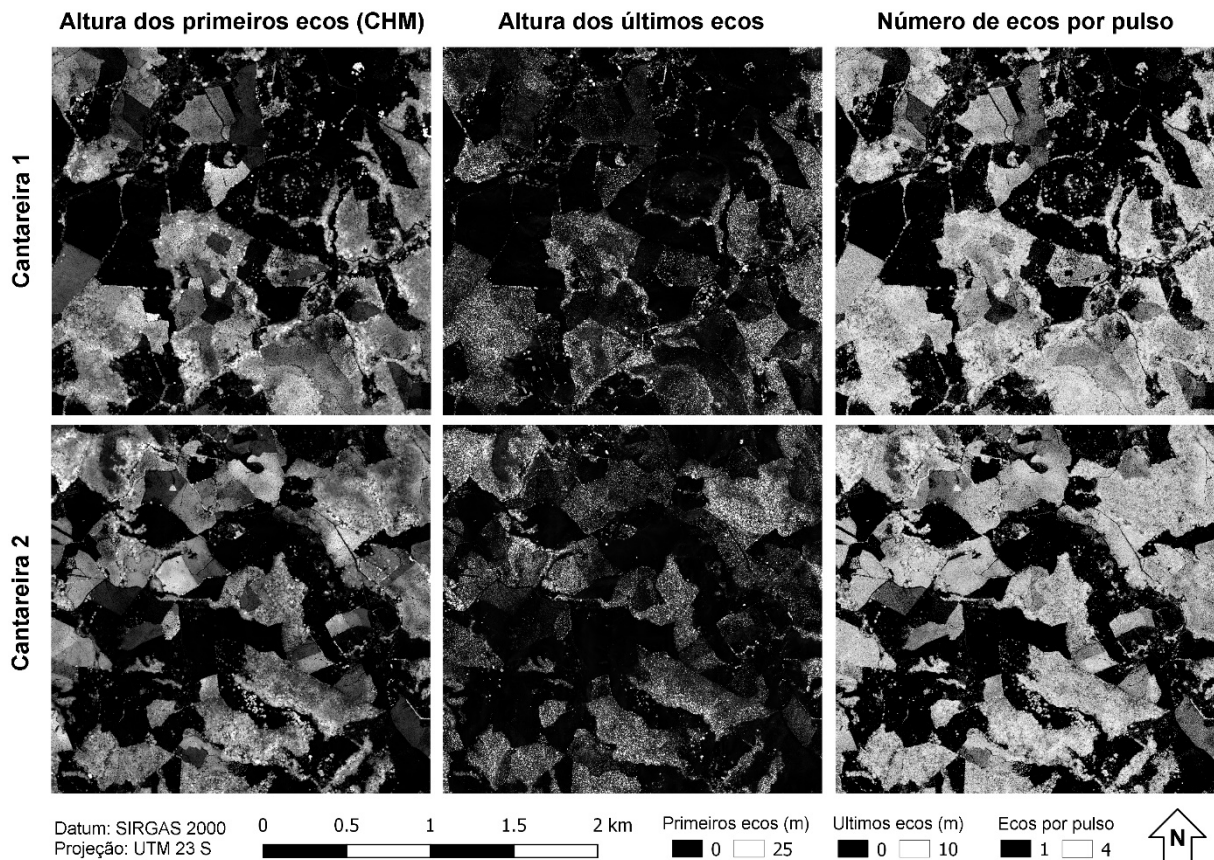


Figura 11 – Imagem de altura dos primeiros ecos (CHM), altura dos últimos ecos e número de ecos por pulso para as áreas Cantareira 1 e 2.

A partir da imagem de altura dos primeiros ecos, foram geradas quatro imagens correspondentes à altura da vegetação dentro de um pixel de 100 m² do Sentinel 2. Essas variáveis são variações de um grupo de métricas classificadas por Chen et al., (2007) e posteriormente por Zhao e Popescu, (2009) como “métricas relacionadas à altura” para a estimativa de parâmetros biofísicos e estruturais como LAI, área basal e volume de troncos. Para isso, a imagem de altura dos primeiros ecos foi reamostrada e alinhada às imagens do Sentinel 2 de quatro maneiras distintas utilizando a ferramenta “alinhar imagem” do QGIS. A primeira delas consistiu de uma média aritmética dos pixels da imagem dos primeiros ecos dentro de um pixel do Sentinel 2 (10 x 10 m) e corresponde a uma média da altura do dossel dessa área. A segunda é a altura máxima da imagem dos ecos dentro do mesmo pixel, e corresponde à altura do maior ponto do dossel na área. A terceira (Figura 12) é uma média das alturas

máximas em 4 quadrantes de 25 m² dentro de um pixel de 100 m² do Sentinel 2, considerando as alturas máximas de 4 pontos distintos do dossel.

Esse terceiro procedimento para a altura dos dosséis teve como base os resultados da utilização do algoritmo *meanshift segmentation* do *Orfeo Tool Box* (OTB) disponível como ferramenta no QGIS para delinear os dosséis das árvores nas áreas automaticamente. O procedimento teve inspiração no estudo realizado por Ke e Quackenbush (2011) sobre o uso de algoritmos de segmentação para delimitação de dosséis de árvores de florestas decíduas e de coníferas usando imagens de alta resolução. O algoritmo utilizado aqui foi capaz de delinear os dosséis com uma precisão razoável se comparado com o delimitação manual, chegando a uma média de aproximadamente 100 m² de área por árvore. Os resultados dessa análise mostraram que em aproximadamente 60% dos casos as árvores delimitadas tiveram áreas inferiores a 100 m², indicando uma probabilidade razoável de haver mais de uma árvore contida em um pixel do Sentinel 2. Assim, o objetivo do terceiro procedimento descrito no parágrafo anterior é considerar o valor máximo de mais de um indivíduo de árvore por pixel.

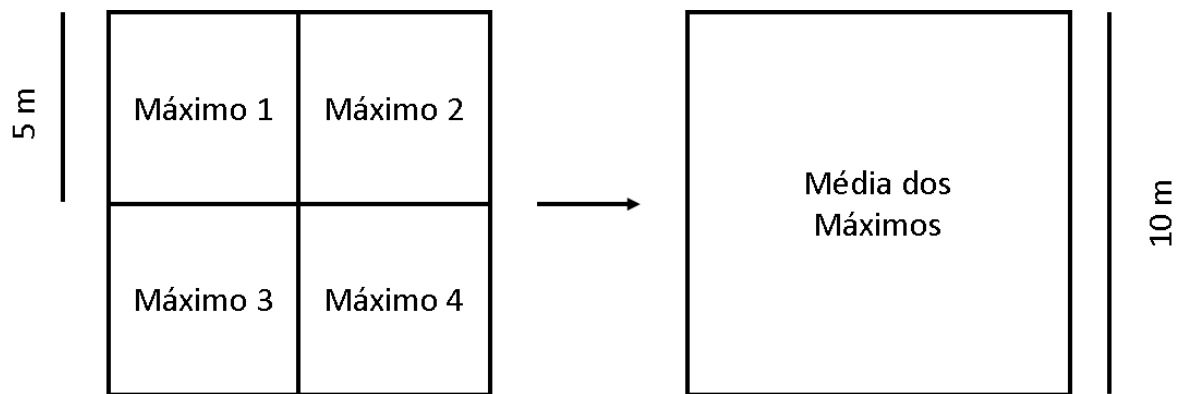


Figura 12 – Um dos métodos de amostragem da imagem de altura dos primeiros ecos, ou CHM (0,5 m) para a mesma resolução das imagens do Sentinel 2 (10 m). Nesse método, tirou-se primeiro o valor máximo dos pixels em cada quadrante de 5 m (imagem à esquerda) e depois calculou-se a média desses valores para determinar o valor final do pixel (imagem à direita).

Por fim, de maneira similar ao feito por Chen (2007), foi feita uma quarta imagem de altura, chamada de “altura stdev”, para corresponder ao desvio padrão dos pixels do CHM dentro da área de um pixel do Sentinel 2 com o objetivo de verificar se variações de altura poderiam interferir na reflectância das imagens.

Além do cálculo das imagens de altura, foram feitas imagens da porcentagem de cobertura de floresta, ou da quantidade de clareiras na região de maneira similar à variável *Gap fraction* utilizada por Getzin et al., (2017), considerando áreas abaixo de dez metros como clareiras. Para isso, dividiu-se o total de pixels do CHM considerados como cobertos por vegetação pelo total de pixels do CHM dentro de um pixel do Sentinel 2 (Figura 13). O processo foi feito de três maneiras distintas usando três definições diferentes de “clareira”, considerando uma região como coberta quando esta possuía vegetação de altura superior a um, cinco e dez metros. A variável de um metro foi feita para aproximar as estimativas de LiDAR da variável C obtida em campo. Já as de cinco e dez metros tiveram o objetivo de detectar uma variação maior de clareiras em outros estratos arbóreos. Assim, foram geradas três imagens da porcentagem de cobertura da região de acordo com cada definição de clareira.

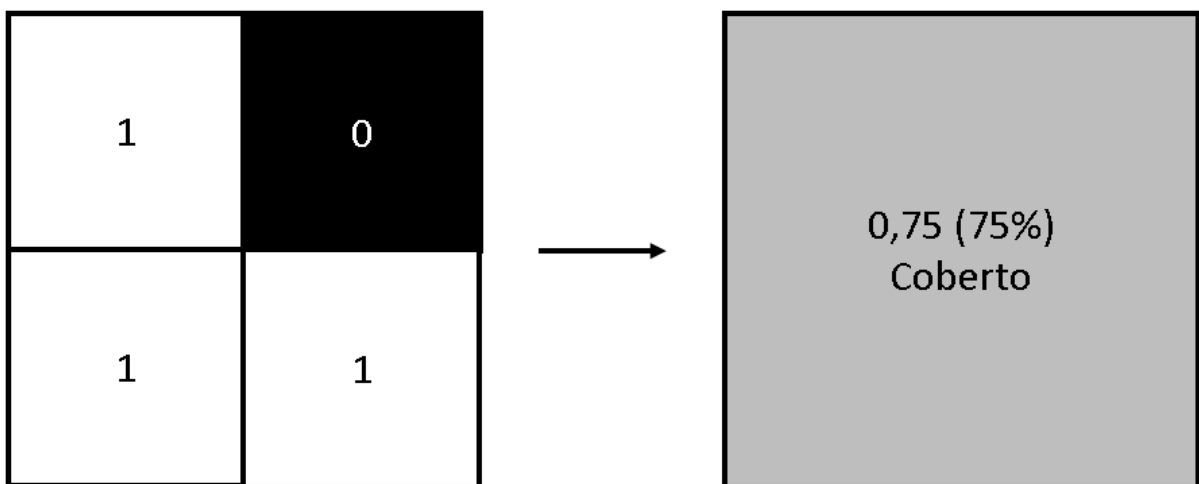


Figura 13 – Exemplo do cálculo da porcentagem de cobertura. Os pixels do CHM (imagem da esquerda) considerados como cobertos por vegetação tem valor 1 e os pixels considerados como clareiras tem valor 0. Na reamostragem para a mesma resolução do Sentinel foi feita uma média aritmética dos valores, resultando em 0,75 no exemplo (75% de cobertura vegetal no pixel). A imagem é uma simplificação, já que um pixel do Sentinel 2 (10 m) contém um total de 400 pixels do CHM.

Adicionalmente foram geradas imagens correspondentes à altura e cobertura das árvores relativas ao seu entorno com o objetivo de detectar influências de áreas próximas, como efeitos de sombra. O procedimento para gerar essas imagens foi o cálculo de uma imagem de altura média em um raio de 10 m de cada pixel do CHM dos primeiros ecos usando a ferramenta *simple filter* do SAGA dentro do QGIS. Com a calculadora raster a imagem de média gerada foi subtraída do CHM original dos primeiros ecos, resultando em uma imagem que mostra a altura relativa de cada pixel

aos seus vizinhos em um raio de 10 metros. A imagem de cobertura relativa foi feita da mesma forma, mas considerando todas as alturas menores que a média em um raio de 10 metros como 0 (clareira) e todas as alturas maiores como 1 (área coberta). As duas imagens foram reamostradas pela média para a resolução do Sentinel 2, conforme descrito para as outras variáveis.

Por fim, foram geradas imagens cujo objetivo é refletir a densidade da vegetação levando em conta a quantidade de ecos e a capacidade dos pulsos laser de penetrar em áreas mais baixas do dossel. Para isso, mais três imagens foram geradas: Ecos, Profundidade e Densidade. A imagem de Ecos é obtida basicamente pela reamostragem por média aritmética da imagem original de quantidade de ecos por pulso laser para a resolução das bandas do Sentinel 2, sendo similar à variável " $r(in +.grd)/pulse$ " utilizada por Zhao e Popescu (2009, p.1633). A imagem de profundidade é a altura dos primeiros ecos menos a variável de altura dos últimos ecos, que mostra a distância que o pulso laser percorreu dentro do dossel. Essa variável é uma modificação da variável "*mean free path*" (KUSAKABE et al., 2005 apud ZHAO; POPESCU, 2009) e tem o intuito de medir a densidade da vegetação, já que o aumento da densidade de folhas e galhos deve diminuir a distância que o pulso laser consegue percorrer sem ser bloqueado por um obstáculo. Por fim, a variável Densidade é a variável Ecos dividida pela Profundidade, e também foi pensada para representar a densidade de folhas em um dossel. Vale destacar que essa variável não corresponde à densidade de ecos por área, mas sim a uma relação entre o número de ecos e a distância entre o primeiro e último eco de um pulso.

Após a geração das variáveis descritas (Quadro 6), as áreas de vegetação nativa foram identificadas e separadas por meio de interpretação visual das imagens do Sentinel 2, das próprias imagens das variáveis geradas e da série temporal de imagens de alta resolução do Google Earth. A separação dessas áreas foi feita usando um polígono delimitado manualmente como máscara de recorte usando a ferramenta Clip do QGIS. Dentro delas, também foram gerados 338 *buffers* aleatórios de 25 metros de raio para tirar as médias como método alternativo de amostragem (Figura 14). As imagens do Sentinel 2 foram recortadas da mesma maneira para posterior comparação. Os procedimentos usados para obter as variáveis descritas podem ser vistos em mais detalhes no Apêndice A (Figuras A1 e A2).

Além das variáveis de estrutura, foram geradas imagens de declividade, aspecto e sombreamento recortadas para as áreas de floresta usando o DSM de pontos de solo para posteriores comparações com o relevo.

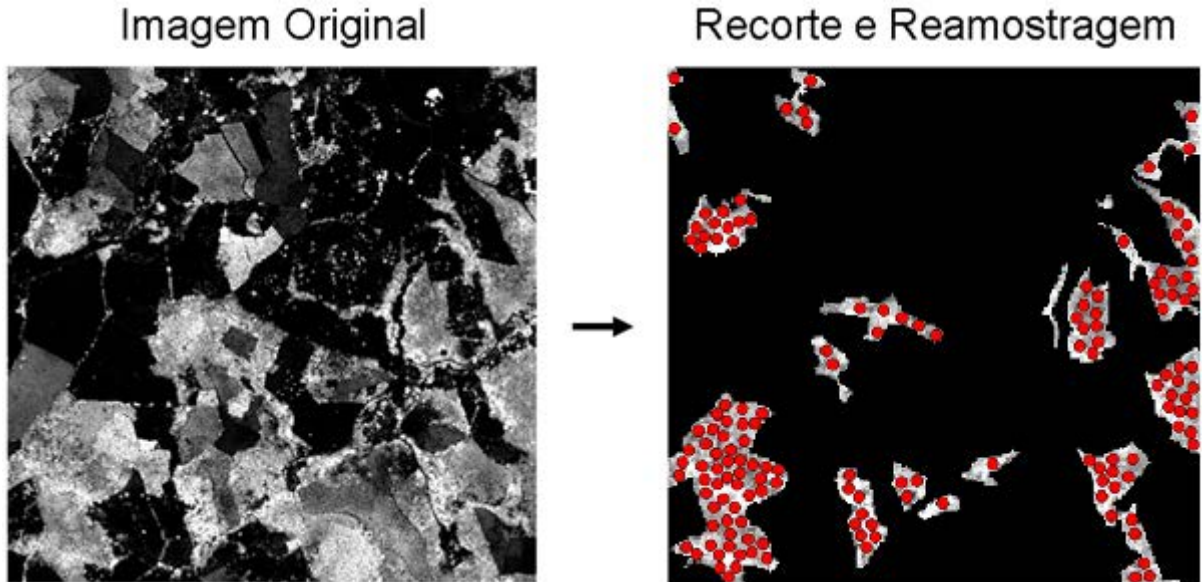


Figura 14 – Reamostragem e alinhamento para pixels de 10 m do Sentinel 2 seguido de recorte delimitando as áreas de vegetação natural. Também foram gerados buffers aleatórios (em vermelho) dentro dessas áreas para calcular as médias dentro deles como método de amostragem alternativo. A área representada na figura é a Cantareira 1.

Quadro 6 – Variáveis LiDAR utilizadas e suas respectivas descrições para cada pixel de 100 m².

Altura mínima	Média aritmética da imagem de altura dos últimos ecos.	m
Altura média	Média aritmética da imagem de altura dos primeiros ecos.	m
Altura máxima 10	Valor máximo da imagem de altura dos primeiros ecos.	m
Altura máxima 5	Média do valor máximo da imagem de altura dos primeiros ecos em quatro quadrantes de 25 m ² .	m
Altura relativa	Altura do pixel menos a média dos valores de altura dos pixels vizinhos em um raio de 10 m.	m
Altura stdev	Desvio padrão da imagem de altura dos primeiros ecos.	m
Cobertura 1	Porcentagem de áreas cobertas por vegetação, sendo consideradas clareiras áreas com menos de 1 m de altura.	%
Cobertura 5	Porcentagem de áreas cobertas por vegetação, sendo consideradas clareiras áreas com menos de 5 m de altura.	%
Cobertura 10	Porcentagem de áreas cobertas por vegetação, sendo consideradas clareiras áreas com menos de 10 m de altura.	%
Cobertura relativa	Porcentagem de pixels mais altos em um raio de 10m.	%
Ecos	Quantidade de retornos por pulso laser emitido.	ecos
Profundidade	Altura média menos altura mínima.	m
Densidade	Razão de ecos por profundidade.	ecos/m

2.6 ANÁLISE DOS DADOS

Com o intuito de localizar as parcelas de campo nas imagens, os pontos centrais dos hexágonos foram importados para o *software* QGIS. Para cada ponto foi feito um buffer de 25 m de raio. Os buffers em áreas de nuvens foram descartados e os restantes foram usados como mascaras para o recorte das imagens de bandas e índices. Em seguida, os dados de campo e as imagens recortadas foram importadas para o IDE *spyder*¹, onde foram feitos modelos de regressão linear para a relação entre as médias dos pixels dentro de cada buffer recortado e os dados de campo.

O procedimento foi feito de maneira similar para os dados LiDAR, com a importação de todas as imagens para o *spyder*, mas os dados foram importados de duas maneiras distintas. Uma delas consistiu em importar as imagens inteiras como arquivos separados, possibilitando fazer uma comparação pixel a pixel das imagens de estrutura derivadas de LiDAR e das imagens do Sentinel 2 e Landsat-8/OLI. A outra foi utilizar buffers aleatórios com 25 m de raio dentro das áreas de vegetação para calcular a média das imagens para cada um desses buffers, de modo a obter resultados parecidos aos da comparação com dados de campo. Como o uso dos buffers serviu para minimizar erros causados pelas resoluções diferentes das bandas do Sentinel 2 e de alinhamento dos pixels entre as imagens de satélite e os dados LiDAR, a maioria dos resultados apresentados no estudo para as variáveis LiDAR são referentes a essa metodologia. Além disso, foram feitas comparações do mesmo tipo com imagens de declividade, aspecto e sombreamento recortados para a área com o intuito de verificar o efeito dessas variáveis apesar da correção topográfica aplicada.

Primeiramente, medidas de média, desvio padrão e coeficiente de variação (CV – razão entre o desvio padrão e a média) dos dados de campo e dos dados LiDAR foram geradas, além de correlações (Spearman) entre essas variáveis para observar seu comportamento conjunto. Depois, para verificar se as imagens têm relação com as variáveis derivadas de campo e de LiDAR, elas foram comparadas por meio de modelos de regressão linear OLS (Mínimos Quadrados Ordinários – *Ordinary Least Squares*) da biblioteca *statsmodels* (SEABOLD; PERKTOLD, 2010) escrita em *python* usando todos os dados obtidos para treino por meio de valores de r^2 , RMSE (*Root*

¹ Os scripts usados no *spyder* podem ser acessados através do link:
<https://1drv.ms/f/s!Audlk83d5sO0gddvSD3y67Ot-4gtSw>

Mean Square Error) e RRSE (*Root Relative Square Error*), O RMSE sendo a média da magnitude dos erros do modelo e o RRSE o valor de RMSE dividido pelo desvio padrão dos dados de teste, este último podendo ser usado para comparações entre modelos com variáveis distintas. Quanto menores os valores de RMSE e RRSE menor o erro dos modelos e, portanto, melhor sua capacidade de predição (HOFF et al., 2013, GUO et al., 2015).

Adicionalmente, foram realizadas análises de regressão múltipla envolvendo mais de uma banda do Sentinel 2 para ambos os dados de campo e os dados LiDAR. Os melhores modelos de regressão múltipla foram selecionados com base no valor de r^2 ajustado, AIC, e na quantidade de bandas envolvidas no modelo (penalidade de 0,03 no valor de r^2 ajustado para cada variável explicativa adicional). Após feitos os modelos de regressão, foram gerados gráficos de dispersão mostrando as relações entre as variáveis observadas e gráficos no estilo *heatmap* mostrando as melhores relações pelos valores de r^2 ajustado. As equações resultantes dos melhores modelos foram aplicadas nas imagens do Sentinel 2 para verificar o erro dos modelos espacialmente. Esta análise geral não envolveu o processo de validação (com separação de amostras para treino e validação dos modelos), o que foi feito apenas para os modelos mais significativos.

Além dos modelos OLS, modelos de regressão RF (*random forest*) de aprendizado de máquina (BREIMAN, 2001) com o algoritmo *RandomForestRegressor* da biblioteca *scikit-learn* (PEDREGOSA et al., 2011) em python foram gerados. Os modelos foram feitos usando cinco mil árvores de decisão, já que números maiores que estes não resultaram em melhoras significativas para nenhuma das variáveis. Destes modelos, foram obtidos valores de pseudo r^2 a partir de subgrupos de dados *out-of-bag* (BREIMAN, 1996), RMSE, RRSE, e o valor de importância de cada variável. É importante ressaltar que o valor de pseudo r^2 *out-of-bag* em modelos RF é calculado de maneira diferente do r^2 ajustado obtido para os modelos OLS. No entanto, como esses valores tem um significado semelhante, foram usados como comparação entre os dois tipos de modelos em conjunto com o RMSE e o RRSE. Assim como para os modelos OLS, foram construídos quadros com os valores obtidos. Por conta da menor quantidade de amostras de dados de campo, esses modelos só foram gerados para os dados derivados de LiDAR.

Para a construção de modelos com validação, foram escolhidas as variáveis com as relações mais fortes observadas nos resultados das análises anteriores. A escolha das bandas inclusas nesses modelos foi baseada na minimização da quantidade de bandas sem causar redução significativa dos valores de r^2 ajustado para os modelos OLS e r^2 *out of bag* para os modelos RF. Para ambos os modelos OLS e RF, a separação das áreas de treino e teste foi feita de duas maneiras diferentes. Primeiro, usando a imagem de 23 de dezembro de 2016, os buffers recortados foram separados em grupos aleatórios com proporção de 75% dos dados para treino dos modelos e 25% para teste. Para os modelos OLS em específico, a aleatorização das áreas de treino e teste foi feita de cinco mil formas diferentes. Essas cinco mil repetições possibilitaram a realização de análises da distribuição de erros (RMSE) dos modelos OLS através da construção de histogramas, obtendo-se um modelo de valor mediano para os erros na área. Para os modelos RF, por conta do tempo de processamento, apenas uma repetição foi realizada, com a inclusão de cinco mil árvores de decisão no modelo. Foram gerados gráficos de dispersão entre os valores preditos pelos modelos OLS de valor mediano e os modelos RF com as bandas do Sentinel 2 e os valores reais das variáveis de controle. Esse primeiro método de teste possibilitou verificar se os modelos obtidos poderiam ser usados em outras áreas na mesma imagem em que foram treinados. Mais uma vez, por conta da menor quantidade de amostras disponíveis de dados de campo, esse processo foi realizado apenas para os dados derivados de LiDAR.

Por fim, para testar se os modelos OLS e RF gerados poderiam ser usados em outras imagens, eles foram treinados usando todos os dados para a imagem de 23 de dezembro de 2016 e testados na imagem de 08 de janeiro de 2016 com os dados correspondentes à Cantareira 2 (a presença de nuvens impossibilitou o uso dos dados da Cantareira 1 nessa imagem). Da mesma forma, gráficos de dispersão mostrando as relações entre os valores preditos por alguns modelos selecionados e os valores reais foram construídos para as variáveis com relação mais forte de acordo com as análises anteriores. Ambas as imagens podem ser vistas na Figura 3 (seção 2.2.1).

2.7 ANÁLISE EM LABORATÓRIO: ESPECTROS DE REFLECTÂNCIA

Como dito anteriormente, foi feita uma análise da assinatura espectral de diferentes camadas de folhas usando um espectrorradiômetro. Os resultados obtidos no experimento foram usados para melhor explicar o porquê de algumas bandas do Sentinel 2 se relacionarem melhor com variáveis de estrutura vegetal do que outras. Em especial, o procedimento teve como enfoque a faixa espectral próxima à 705 nm, já que este é o comprimento de onda em que a banda 5 das imagens do Sentinel 2 é centrada, tendo em vista que essa é a banda que apresentou melhores resultados, como será demonstrado e discutido posteriormente nas seções 3 e 4.

Todas as folhas usadas no experimento foram coletadas de um mesmo indivíduo da espécie *Annona squamosa* e medidas em um intervalo de 30 minutos após a coleta com o objetivo de evitar variações na concentração de clorofila, umidade e estrutura interna das folhas. O aparelho utilizado foi o FieldSpec 4 Hi-Res da ASD, com suas especificações mostradas no Quadro 7.

Quadro 7 – Especificações do espectrômetro FieldSpec 4 Hi-Re da ASD.

Faixa espectral medida	350 a 2500 nm
Resolução	VNIR: 3 nm para 700 nm SWIR 1 e 2: 8 nm para 1400 e 2100 nm
NEdl	VNIR: 1.0×10^{-9} W/cm ² /nm/sr – 700 nm SWIR 1: 1.4×10^{-9} W/cm ² /nm/sr – 1400 nm SWIR 2: 2.2×10^{-9} W/cm ² /nm/sr – 2100 nm
Dimensões: Altura x largura x profundidade	12.7 x 36.8 x 29.2 cm
Peso (sem a bateria)	5.44 kg (12 lbs)

O procedimento consistiu em gerar a assinatura espectral de diferentes números de camadas de folhas (Figura 15). Após a calibração do sensor (utilizando uma referência branca), quatro folhas sobrepostas foram fixadas em um suporte acima de uma amostra de solo e o aparelho foi direcionado a elas para realizar as medidas. Valores de reflectância foram tomados mantendo a folha no topo fixa e removendo as folhas inferiores de forma sequencial. A folha do topo não foi retirada para evitar diferenças no ângulo e distância das folhas em relação ao sensor e à fonte de luz, muitas vezes responsáveis por variações maiores do que o próprio efeito do número de camadas que se queria observar.

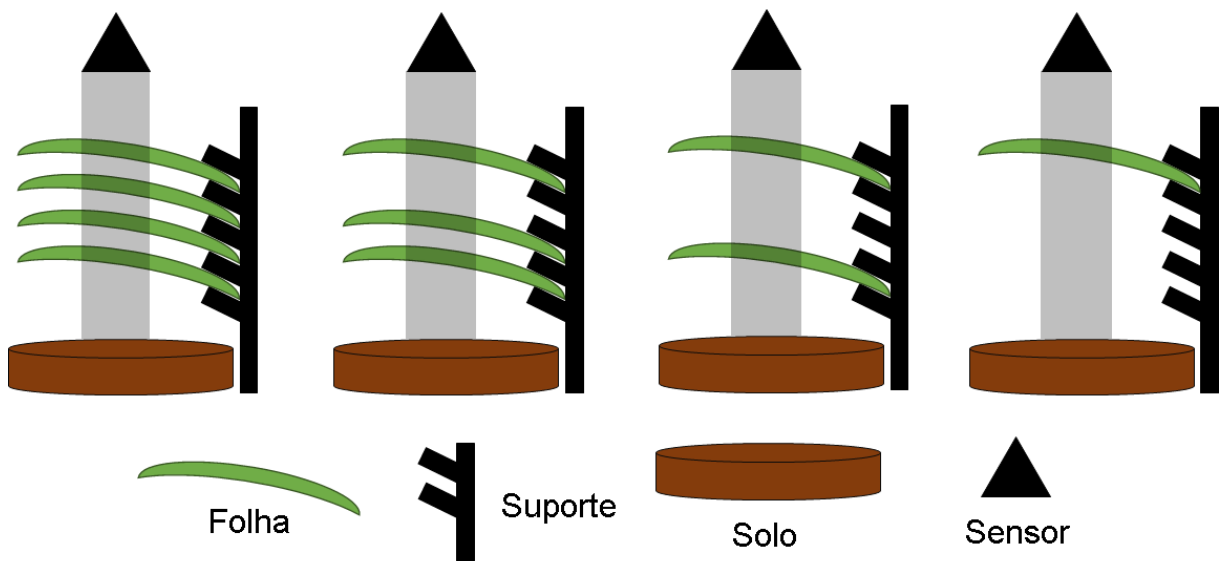


Figura 15 – Representação esquemática da assinatura espectral de diferentes camadas de folha.

Os valores de reflectância foram obtidos para os diferentes números de camadas de folha, e um gráfico da resposta espectral dessas medidas foi gerado. Além disso, os valores de reflectância de uma folha foram subtraídos dos valores de reflectância de quatro folhas, a partir dos quais construiu-se um gráfico de diferenças de reflectância entre uma e quatro camadas de folhas. O objetivo de realizar esse procedimento foi observar como o aumento ou diminuição de camadas de folhas modifica a resposta espectral da vegetação, buscando simular, de forma simplificada, o comportamento de florestas com diferentes estratos ou densidades de folhas.

Ressalta-se que esse experimento não permite considerar aspectos direcionais da reflectância, bem como efeitos causados pelo relevo, solo, atmosfera, presença de flores, frutos, galhos e posicionamento das folhas, que são de grande importância no sensoriamento remoto da vegetação e podem causar alterações significativas nos valores de reflectância de uma floresta real (BREUNIG et al., 2012, PONZONI; SHIMABUKURO; KUPLICH, 2012). Esses efeitos não foram incluídos no experimento pois, buscou-se apenas verificar o comportamento do efeito aditivo das camadas de folhas, principalmente para a faixa espectral próxima a 705 nm mencionada.

3 RESULTADOS

3.1 ANÁLISE COM DADOS DE CAMPO

3.1.1 Análise geral dos dados

Como descrito na seção 2.6, primeiramente foi feita uma análise geral das variáveis de campo das 34 parcelas comparando seus coeficientes de variação (CV – razão entre a média e o desvio padrão) e valores de correlação de spearman. É possível observar pela comparação de CV no Quadro 8 e pela Figura 16 que as variáveis de campo apresentaram dispersões bem diferentes umas das outras. Dentre elas, a que apresentou menor variabilidade foi a variável C (porcentagem de cobertura do dossel; CV = 0,0058). Também pode-se observar variações baixas nas variáveis AM (altura média) e AM > 5 (altura média dos indivíduos maiores do que 5 metros), com CVs iguais a 0,14 e 0,1. Diferente das outras variáveis de altura, as distribuições de AMed (altura mediana) e AT (Altura Total) apresentaram CV maior, de 0,32 e 0,38 respectivamente. Já as variáveis ligadas ao DAP (diâmetro à altura do peito) e ao número de indivíduos apresentaram variações mais altas, com valores de CV desde 0,31 (DAPM – DAP médio) até 0,76 (BSASM – Biomassa Seca Acima do Solo Média).

Quadro 8 – Valores de média, desvio padrão (Std) e coeficiente de variação (CV) para as variáveis derivadas dos dados de campo de todas as 34 parcelas. AM (Altura média das árvores amostradas), C (proporção de cobertura da parcela), DAPM (diâmetro à altura do peito médio das árvores amostradas), N (número de árvores amostradas na parcela), DAPT (diâmetro à altura do peito total das árvores amostradas), AT (altura total das árvores amostradas), AB (área basal total das árvores da parcela pela área de coleta). BSASM (Biomassa seca acima do solo média), BSAST (biomassa seca acima do solo total), AM > 5 (altura média das árvores maiores de cinco metros), N > 5 (número de árvores maiores de cinco metros), AMed (altura mediana das árvores amostradas), DAPMed (diâmetro à altura do peito mediano das árvores amostradas).

Variável	Unidade	Média	Std	CV
AM	m	4,62	0,64	0,14
C	%	99,17	0,58	0,0058
DAPM	cm	7,82	2,41	0,31
N	-	55,76	22,40	0,40
DAPT	cm	411,34	144,85	0,35
AT	m	252,75	95,08	0,38
AB	m ² /ha	44,61	27,23	0,61
BSASM	kg	45,41	34,52	0,76
BSAST	kg	2286,97	1558,85	0,68
AM > 5	m	7,06	0,73	0,10
N > 5	-	17,47	9,15	0,52
AMed	m	4,14	0,74	0,18
DAPMed	cm	4,41	1,43	0,32

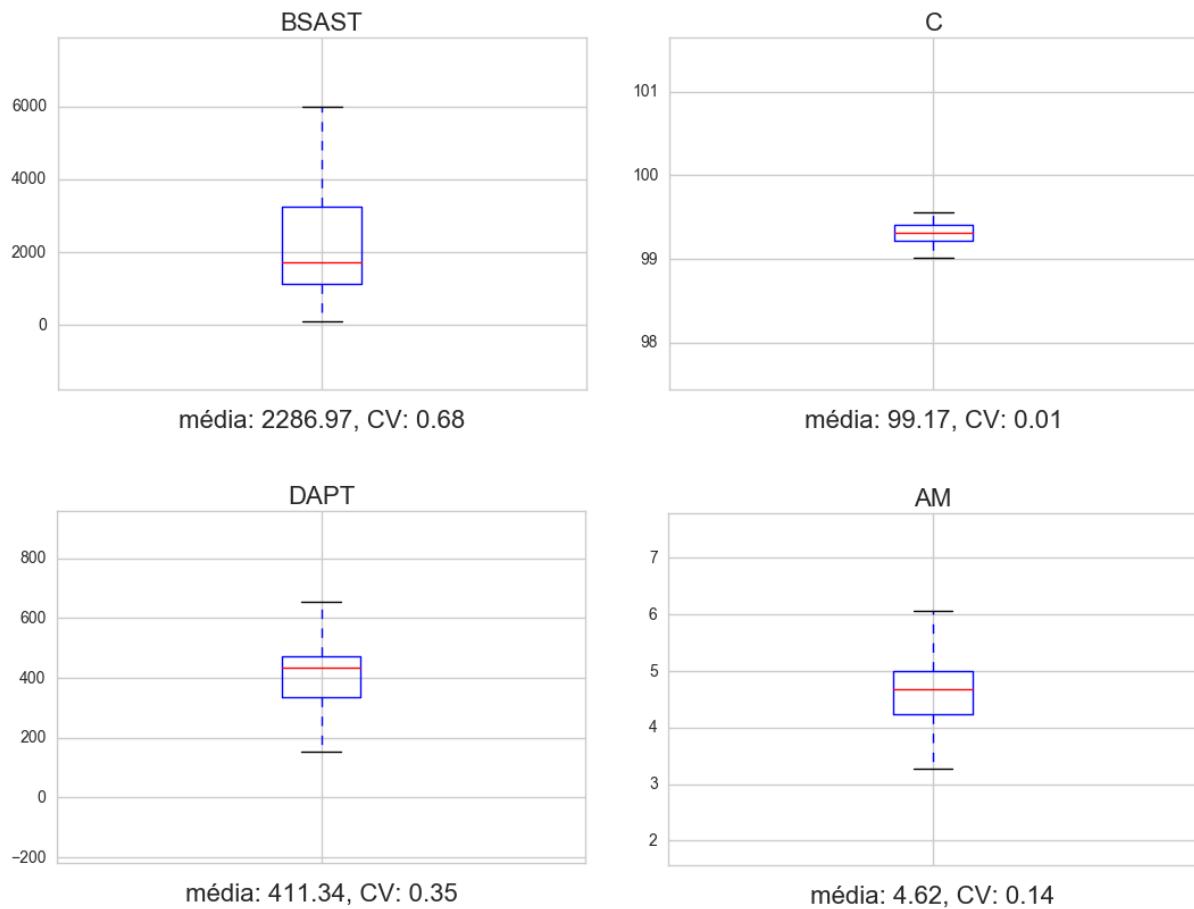


Figura 16 – *Boxplots* mostrando a dispersão de algumas das variáveis de campo com os valores de média e coeficiente de variação (CV) para todas as 34 parcelas amostradas. BSAST (Biomassa Seca Acima do Solo total - kg), C (porcentagem de cobertura do dossel - %), DAPT (Diâmetro à Altura do Peito Total - cm), AM (Altura Média - m).

As relações entre essas variáveis podem ser observadas pelas correlações (Spearman) apresentadas na Figura 17. Algumas variáveis apresentam correlações fortes entre si ($r > 0,8$) e em sua maioria positiva. As variáveis com correlação mais forte foram BSAST (Biomassa Seca Acima do Solo Total) e AB (Área Basal) ($r = 0,95$). Esse resultado é coerente, já que ambas utilizam DAP (Diâmetro à altura do peito) em seus cálculos e são medidas de densidade da vegetação em uma área. Do mesmo modo, a variável DAPT (DAP total) apresentou correlação positiva com ambas BSAST e AB ($r = 0,7$ e $0,78$ respectivamente). DAPM (DAP médio) e DAPMed (DAP mediano) não apresentaram boas correlações com essas três variáveis, já que correspondem a medidas de centralidade dos indivíduos amostrados, e não a valores totais para a parcela como as variáveis anteriores. Outros valores altos de correlação foram encontrados entre as variáveis N (número de indivíduos) e AT ($r = 0,88$), AT e $N > 5$ (número de indivíduos maiores do que cinco metros) ($r = 0,8$), AT e DAPT ($r = 0,79$,

Figura 18) e BSASM (BSAS média) e AM > 5 ($r = 0,61$). A variável C (porcentagem de cobertura dos dosséis) não apresentou correlação forte com nenhuma outra variável de campo.

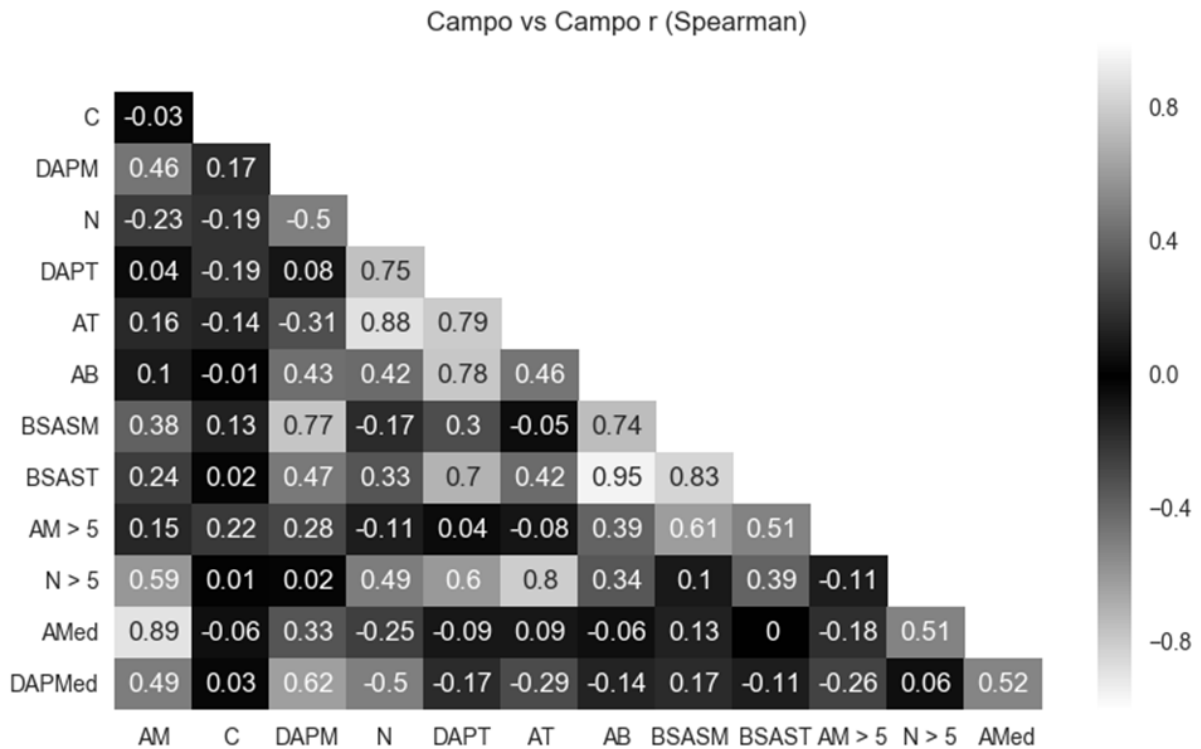


Figura 17 – Valores de r para a correlação (Spearman) das variáveis derivadas dos dados de campo. (Siglas – ver quadros 3 ou 8)

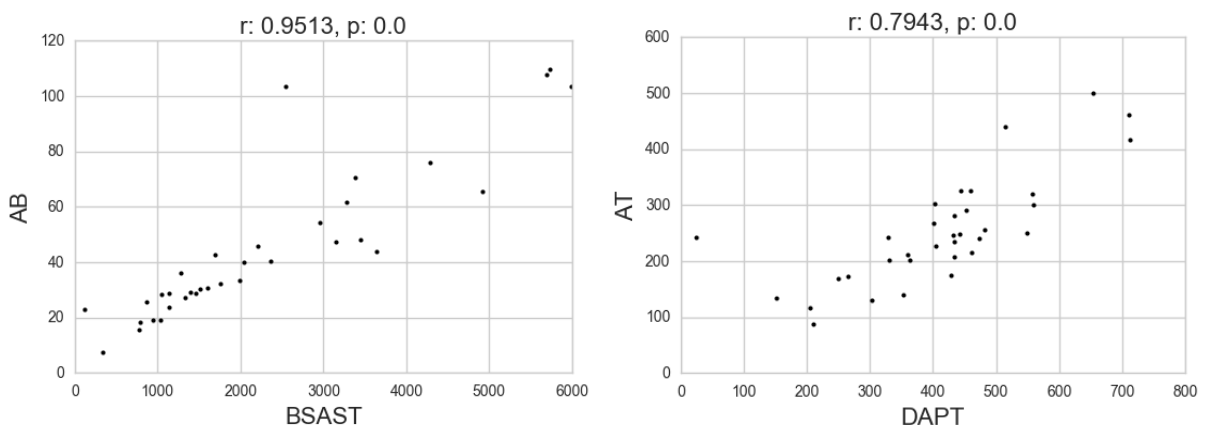


Figura 18 – Gráficos de dispersão com o valor de r e p para a correlação (Spearman) entre as variáveis AB (Area Basal - m^2/ha) e BSAST (Biomassa Seca Acima do Solo - kg) (esquerda) e AT (Altura Total - m) e DAPT (Diâmetro à Altura do Peito Total - cm) (direita).

3.1.2 Relações com o Sentinel 2

Os resultados mostram que as relações encontradas na regressão das bandas das imagens do Sentinel 2 com as variáveis derivadas dos dados de campo, mostradas anteriormente no Quadro 8, tiveram valores de r^2 baixos, próximos a 0,2 nos melhores casos. Nota-se ainda, na comparação entre imagens não corrigidas e corrigidas topograficamente (Figuras 19 e 20, respectivamente) que a correção não causa muitas alterações nos resultados. No entanto, as mudanças observadas são de forma geral positivas, de modo que as análises posteriores são referentes às imagens topograficamente corrigidas.

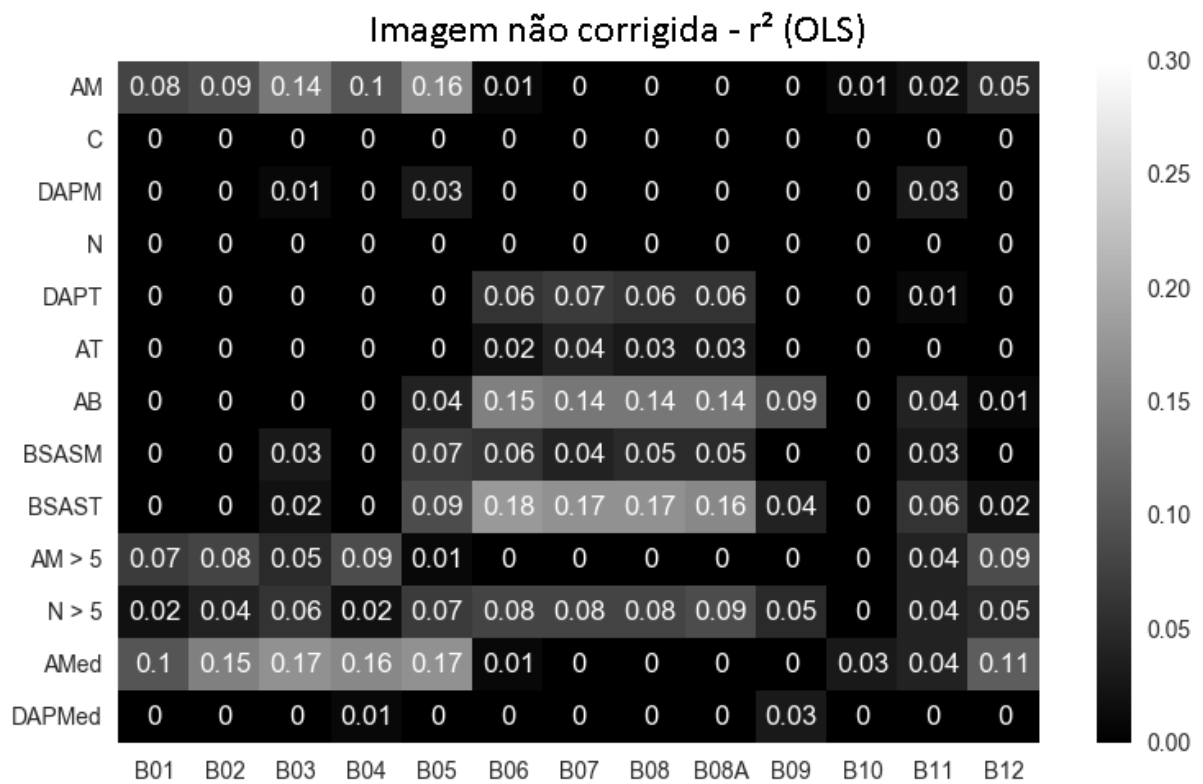


Figura 19 – Valores de r^2 para a regressão linear (OLS) das bandas do Sentinel 2 apenas com correção atmosférica (eixo x) e dos parâmetros biofísicos da vegetação derivados dos dados de campo (eixo y). (Siglas – ver quadros 3 ou 8)

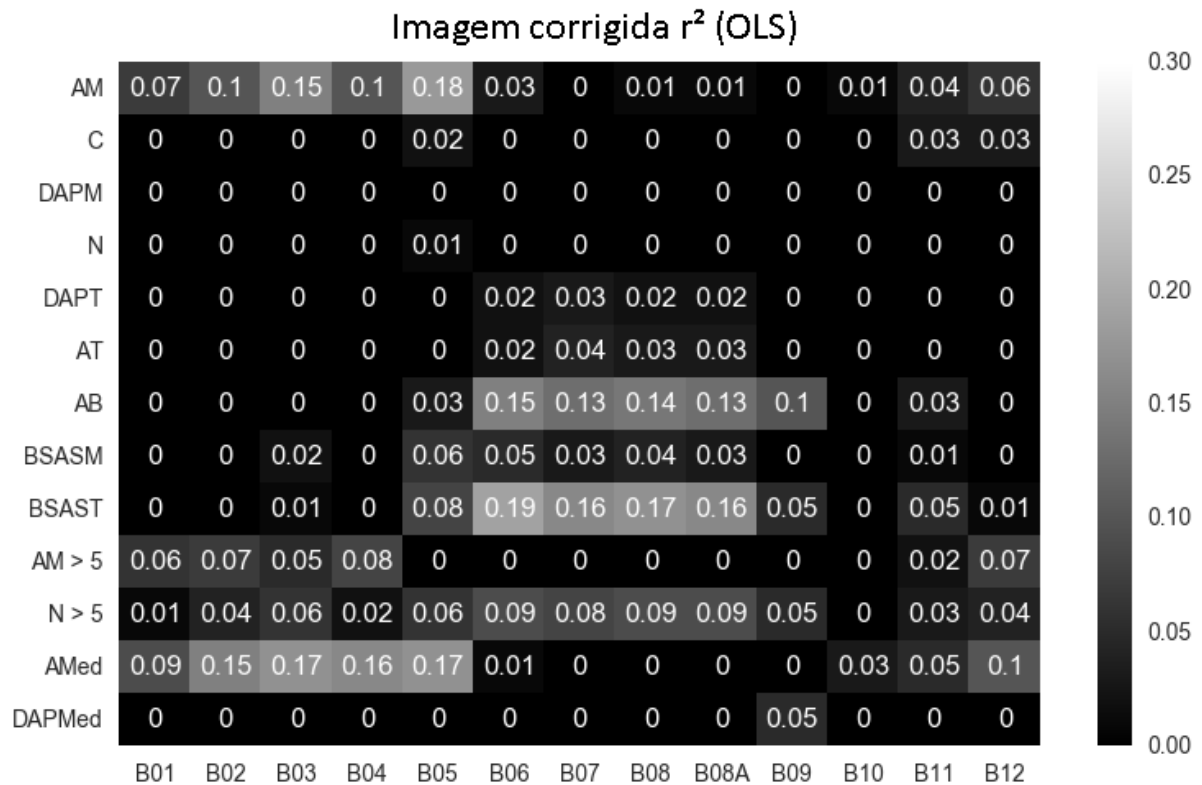


Figura 20 – Valores de r^2 para a regressão linear (OLS) das bandas do Sentinel 2 com aplicação de correção topográfica e atmosférica (eixo x) e das variáveis derivadas dos dados de campo (eixo y). (Siglas – ver quadros 3 ou 8).

Dentre as relações observadas com imagens do Sentinel 2, as variáveis altura média e mediana (AM e AMed) apresentaram relações significativas com as bandas da região visível e início do *red-edge*, chegando a um r^2 de 0,18 para a regressão entre B5 (705 nm) (Figura 21) e AM. Já as variáveis BSAST e AB apresentaram resultados melhores com as bandas do NIR e *red-edge* (r^2 0,19 e 0,15 respectivamente para B6 – 740 nm). O restante das variáveis não apresentou relações com as bandas do Sentinel 2, incluindo o DAP, número de indivíduos, cobertura do dossel e BSASM (Biomassa Seca Acima do Solo Média). Quanto aos índices de vegetação, os valores de r^2 das regressões OLS com as variáveis de campo foram bastante similares. Assim como para as bandas individuais, as melhores relações foram observadas em sua maioria para as variáveis AMed, com os índices NDVI (Figura 21) e RVI ($r^2 = 0,18$ e $0,16$), e as variáveis AB e BSAST, que tiveram relações melhores com EVI ($r^2 = 0,16$ e $0,18$) e SAVI ($r^2 = 0,16$ e $0,18$). Os valores de r^2 para todos os índices podem ser vistos na Figura 22.

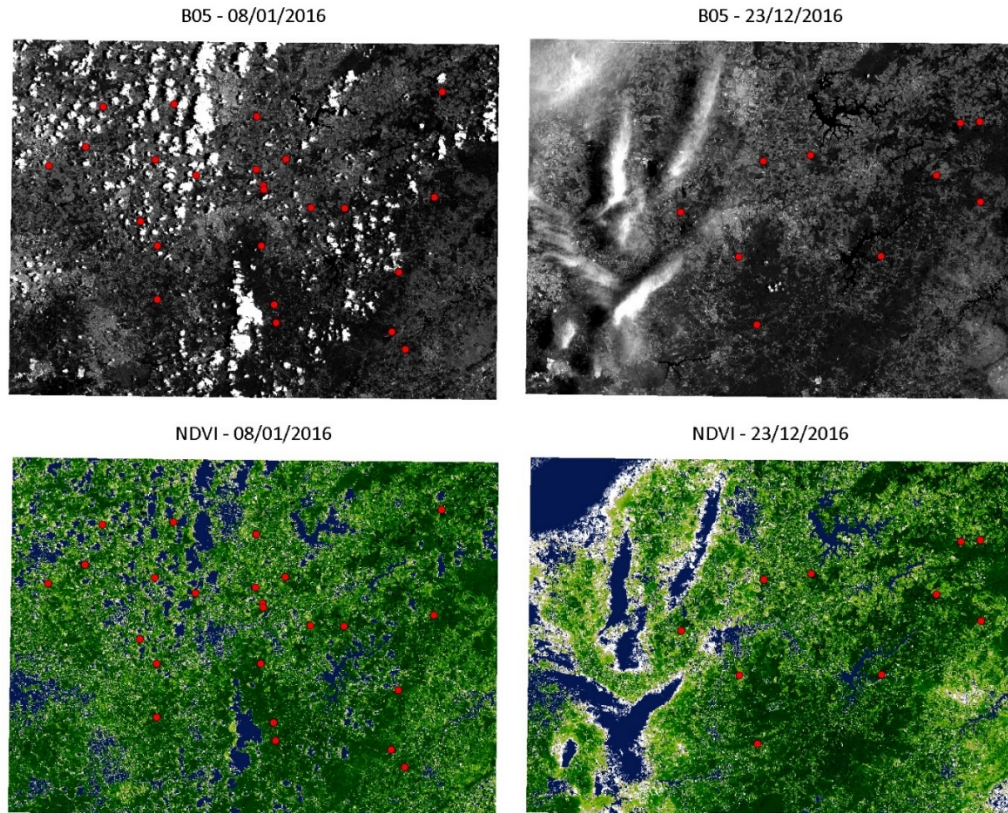


Figura 21 – Banda 5, NDVI e razão da banda 11 pela banda 3 com as imagens de janeiro (esquerda) e dezembro (direita) do Sentinel 2 com aplicação de correção topográfica e atmosférica. Em vermelho, a localização das parcelas de campo (fora de escala). Apesar da grande quantidade de nuvens nas imagens, as 34 parcelas selecionadas para análise não estão em áreas nubladas.

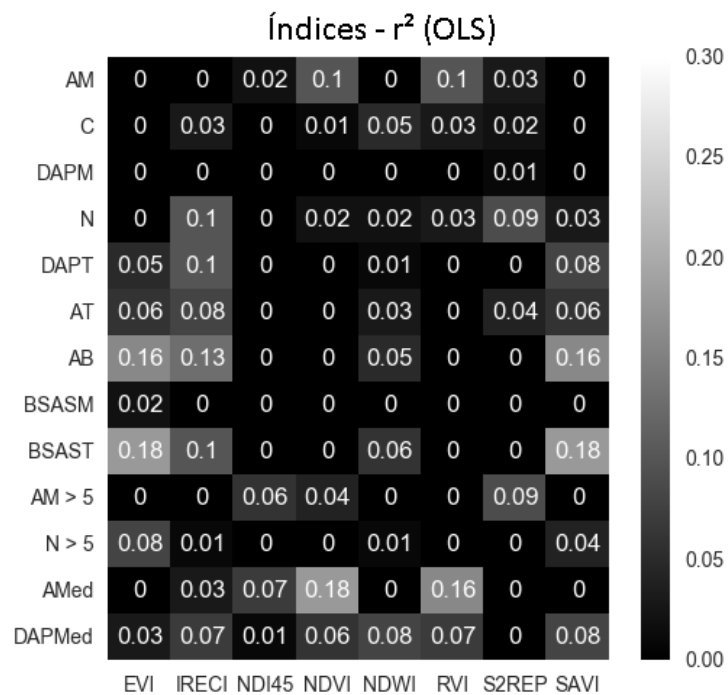


Figura 22 – Valores de r^2 para a regressão linear (OLS) dos índices de vegetação (eixo x) derivados das bandas do Sentinel 2 com aplicação de correção topográfica e atmosférica e das variáveis derivadas dos dados de campo (eixo y). (Siglas – ver quadros 3 ou 8)

Para as razões entre bandas calculadas (Figuras 23, 24 e Apêndice B – Figuras B2 a B7), os melhores resultados estão nas relações entre AM e AMed e as razões entre as bandas B3 (560 nm) e B11 (1610 nm) ($r^2 = 0,24$ e $0,26$, respectivamente - Figura 25). B3 (560 nm) ainda apresentou resultados parecidos em razões com outras bandas para essas duas variáveis de campo, superando os valores de r^2 encontrados para os índices de vegetação usados e para as bandas consideradas individualmente. Além disso, B4 (665 nm) e B5 (705 nm) apresentaram relações com essas variáveis em razões com outras bandas do NIR e SWIR para ambas a média e a mediana da altura (Figuras 23 e 24).

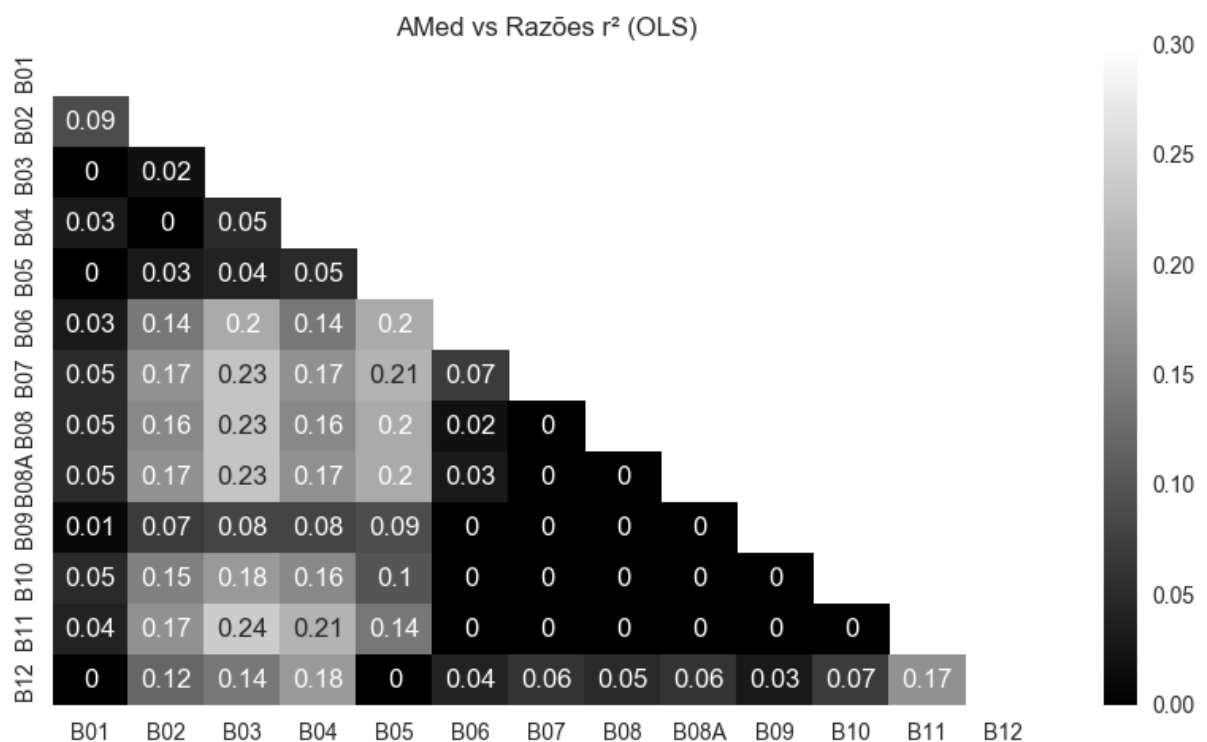


Figura 23 – Valores de r^2 para a regressão linear (OLS) entre a variável AMed (altura mediana) obtida em campo e a razão entre as duas bandas do Sentinel 2 representadas no eixo y e x (respectivamente o numerador e denominador da razão). Imagens com aplicação de correção topográfica e atmosférica.

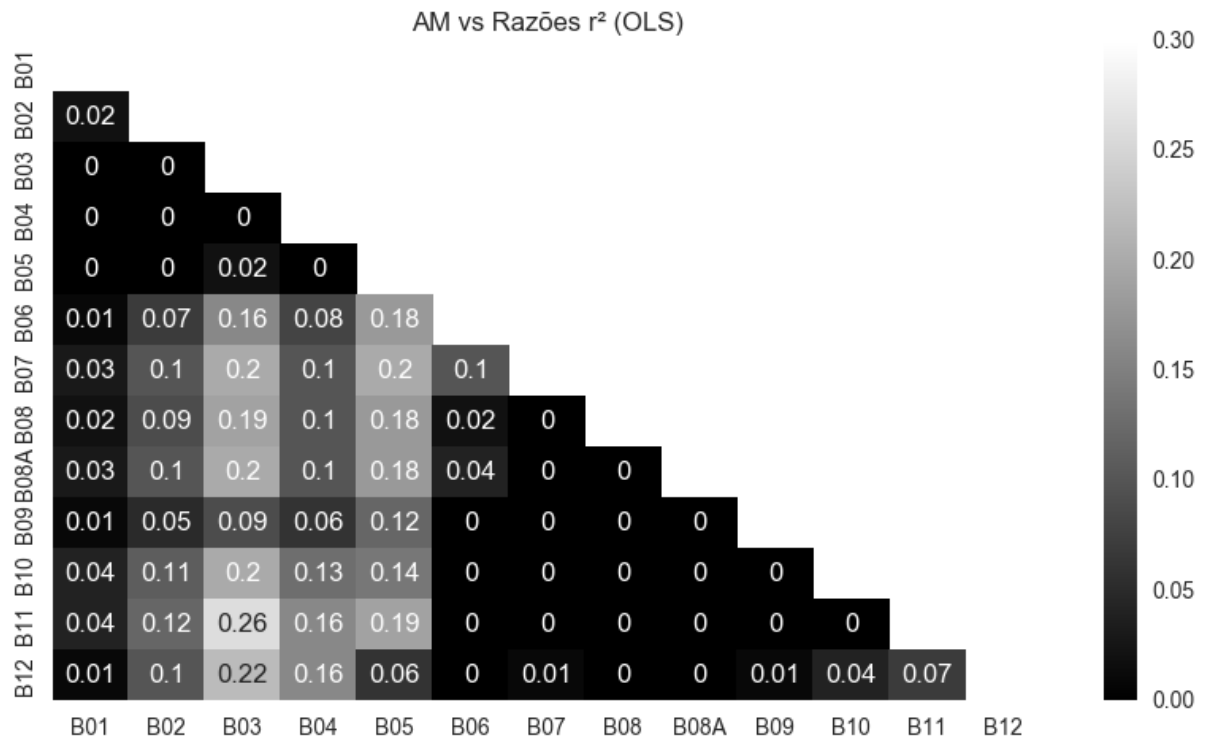


Figura 24 – Valores de r^2 para a regressão linear (OLS) entre a variável AM (altura média) obtida em campo e a razão entre as duas bandas do Sentinel 2 representadas no eixo y e x (respectivamente o numerador e denominador da razão). Imagens com aplicação de correção topográfica e atmosférica.

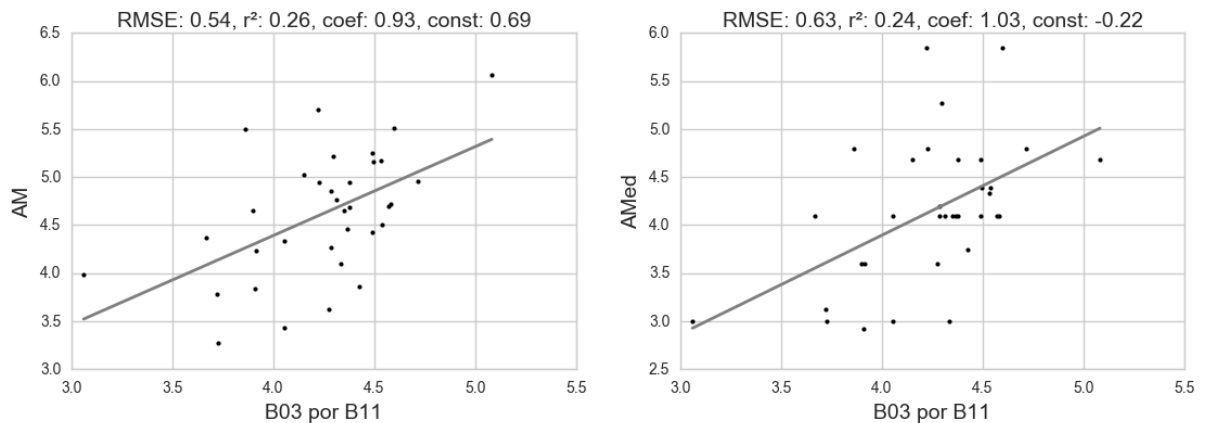


Figura 25 – Gráficos de dispersão para a relação entre as razões de B11 e B3 (eixo x) corrigidas atmosféricamente e topograficamente e as variáveis de campo AM (altura média), à esquerda, e AMed (altura mediana), à direita (eixo y) para todas as 34 parcelas de campo.

Os resultados de campo mostram que, no geral, as razões entre bandas do visível e NIR e do visível e SWIR apresentam relações razoáveis com a altura, mas razões entre NIR e SWIR não. Isso é coerente com o melhor desempenho das bandas do visível nas análises de bandas individuais se comparadas às de outras faixas do espectro. Razões entre bandas correspondentes à mesma faixa espectral também

não apresentaram bons resultados, o que provavelmente está relacionado à similaridade entre as respostas das bandas em faixas de comprimento de ondas mais próximos (Figura 26).

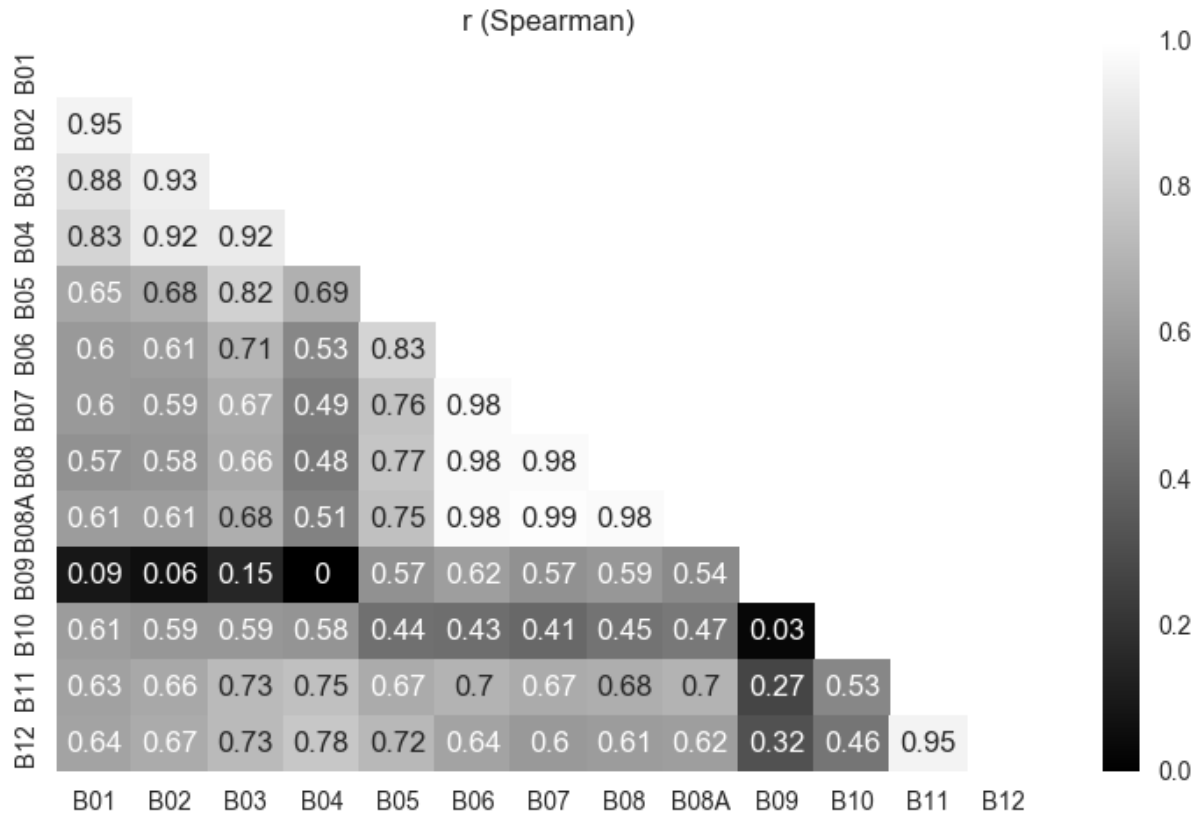


Figura 26 – Valores de r da correlação (Spearman) entre bandas do Sentinel 2 com aplicação de correção topográfica e atmosférica.

Nas análises de regressão múltipla envolvendo mais de uma banda do Sentinel 2 para cada variável de estrutura, pôde-se observar uma melhora significativa nos valores de r^2 (Quadro 9). Conjuntos envolvendo principalmente B3 (560 nm), B4 (665 nm), B5 (705 nm), B7 (783 nm), B8 (842 nm) e B8A (865 nm) apresentaram os melhores resultados. Dentre as variáveis de campo observadas, as que apresentaram melhor relação nas regressões múltiplas foram AM, AMed e C. Observa-se também que apesar de BSAST e AB terem apresentado relações melhores com algumas bandas individuais e índices de vegetação, o uso de modelos de regressão múltipla não melhorou muito as respostas dessas variáveis. Resultados complementares desta análise estão disponíveis no Apêndice B.

Quadro 9 – Fórmula, valor de r^2 ajustado, RMSE e RRSE de alguns dos melhores modelos de regressão múltipla encontrados entre variáveis de estrutura coletadas em campo e imagens do Sentinel 2. A descrição das variáveis pode ser vista no Quadro 3 (seção 2.4). B1: 0,443 nm, B2: 0,490 nm, B3: 0,560 nm, B4: 0,665 nm, B5: 0,705 nm, B6: 740 nm, B7: 783 nm, B8: 842 nm, B8A: 865 nm, B9: 945 nm, B10: 1375 nm, B11: 1610 nm, B12: 2190 nm.

Variável	Equação	r^2	RMSE	RRSE
AM	$6,06 + B01 * 119,91 + B03 * -371,56 + B04 * 210,73 + B08 * 15,19$	0,38	0,47 m	0,74
C	$104,12 + B04 * 173,82 + B05 * -137,54 + B06 * 69,3 + B07 * 88,64 + B08A * -110,76 + B12 * -79,23$	0,45	0,39 %	0,67
DAPM	$4,53 + B06 * -309,64 + B08A * 184,28 + B09 * 167,22$	0,14	2,12 cm	0,88
N	$21,33 + B03 * -2204,22 + B05 * 3482,5 + B07 * -3123,13 + B08A * 2480,59$	0,23	18,4	0,82
DAPT	$653,11 + B05 * 6885,96 + B07 * -17520,07 + B08A * 13707,62$	0,13	128 cm	0,89
AT	$529,3 + B05 * 4525,13 + B07 * -1714,64$	0,08	88,6 m	0,93
AB	$205,32 + B03 * 1436,98 + B06 * -852,06$	0,2	23,7 m ² /ha	0,87
BSASM	$147,2 + B01 * 2835,74 + B06 * -5671,13 + B07 * 3091,68 + B09 * 3085,1$	0,26	27,9 kg	0,81
BSAST	$12284,41 + B01 * 90508,34 + B06 * -48062,78$	0,23	1325 kg	0,85
AM > 5	$8,68 + B05 * -75,58 + B07 * 12,58 + B11 * -83,55 + B12 * 192,62$	0,21	0,61 m	0,83
N > 5	$57,1 + B01 * 1214,34 + B03 * -2755,53 + B04 * 1545,36$	0,13	8,11	0,89
AMed	$5,07 + B01 * 124,27 + B03 * -325,53 + B04 * 146,12 + B08 * 16,08$	0,31	0,57 m	0,78
DAPMed	$4,32 + B02 * 425,63 + B03 * -485,91 + B07 * -89,11 + B08 * 118,72$	0,25	1,16 cm	0,81

3.2 ANÁLISE COM DADOS LIDAR

3.2.1 Análise geral dos dados

Conforme detalhado na seção 2.6, as variáveis derivadas de LiDAR foram comparadas às imagens do Sentinel 2 e do Landsat-8/OLI de várias maneiras distintas, considerando imagens corrigidas e não corrigidas topograficamente, fazendo comparações por pixel ou pela média em buffers aleatórios e utilizando imagens de datas diferentes para verificar se a proximidade com a coleta de dados tem efeito nos resultados.

Primeiramente, assim como feito para os dados de campo, as variáveis derivadas de LiDAR foram comparadas a imagens corrigidas e não corrigidas topograficamente, mas diferentemente dos resultados dos dados de campo, pode-se notar melhoras expressivas após a correção (Figuras 27 e 28).

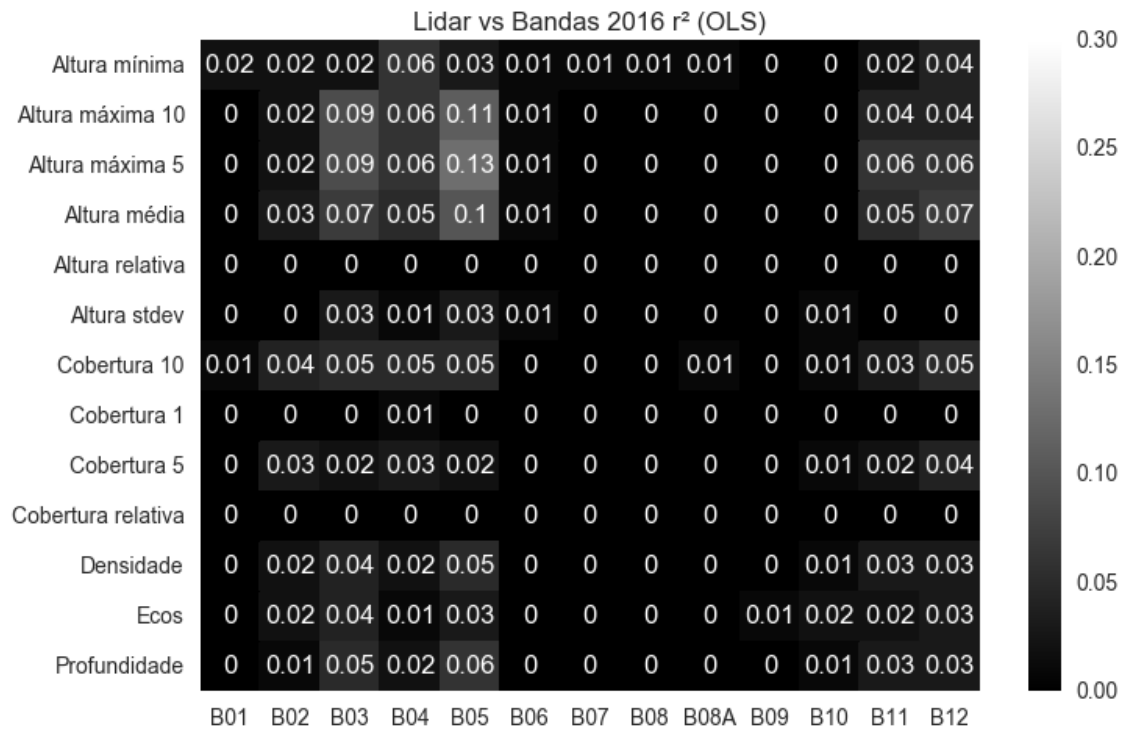


Figura 27 – Valores de r^2 para a regressão linear (OLS) da comparação pixel a pixel das bandas do Sentinel 2 (2016) apenas com correção atmosférica DOS1 (eixo x) e das variáveis obtidas por LiDAR (eixo y).

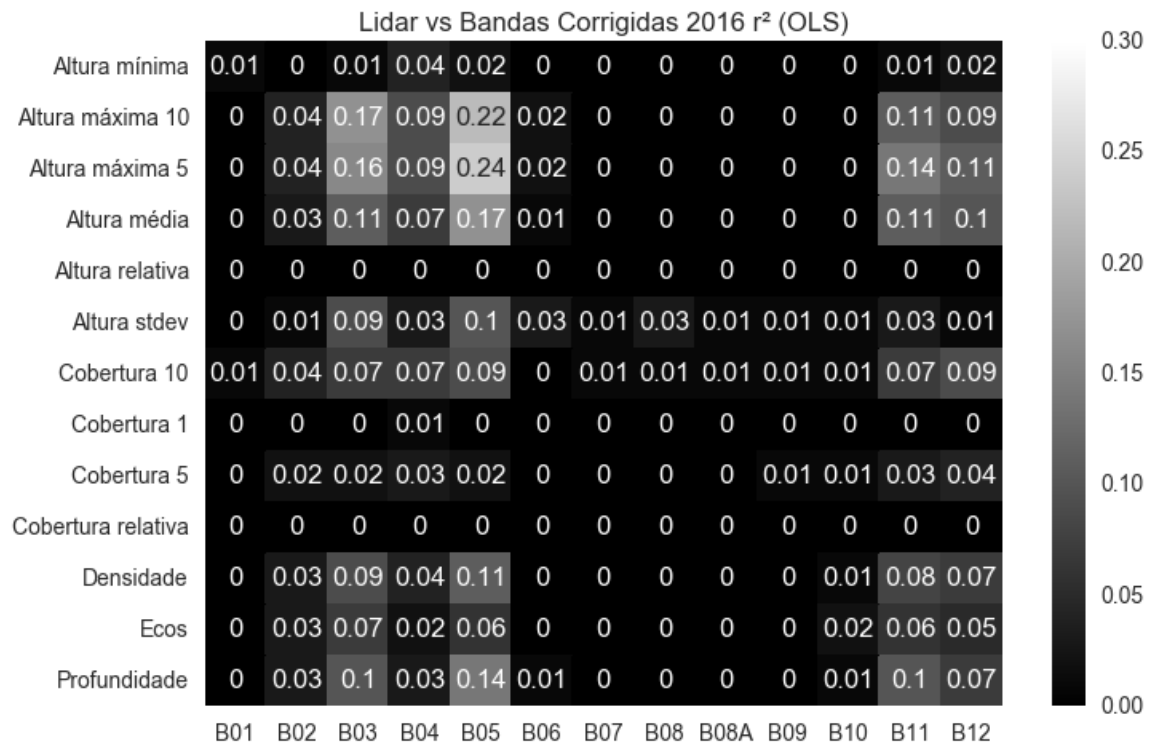


Figura 28 – Valores de r^2 para a regressão linear (OLS) da comparação pixel a pixel das bandas do Sentinel 2 (2016) com aplicação de correção topográfica e atmosférica DOS1 (eixo x) e das variáveis obtidas por LiDAR (eixo y).

Os dados LiDAR também foram comparados a imagens de datas diferentes para verificar se a distância temporal das coletas teria alguma influência nos resultados. As análises mostram que os resultados das imagens do Landsat-8/OLI de 2014 apenas para a Cantareira 1 (Por conta da presença de nuvens na área Cantareira 2) e do Sentinel 2 de 2015 para ambas as áreas (Figuras 29 e 30 respectivamente), mesmo estando mais próximas temporalmente à coleta dos dados LiDAR em 2014, foram bastante inferiores aos resultados da imagem do Sentinel 2 de 2016 (Figura 28). Possíveis razões para estas variações são discutidas na seção 5.

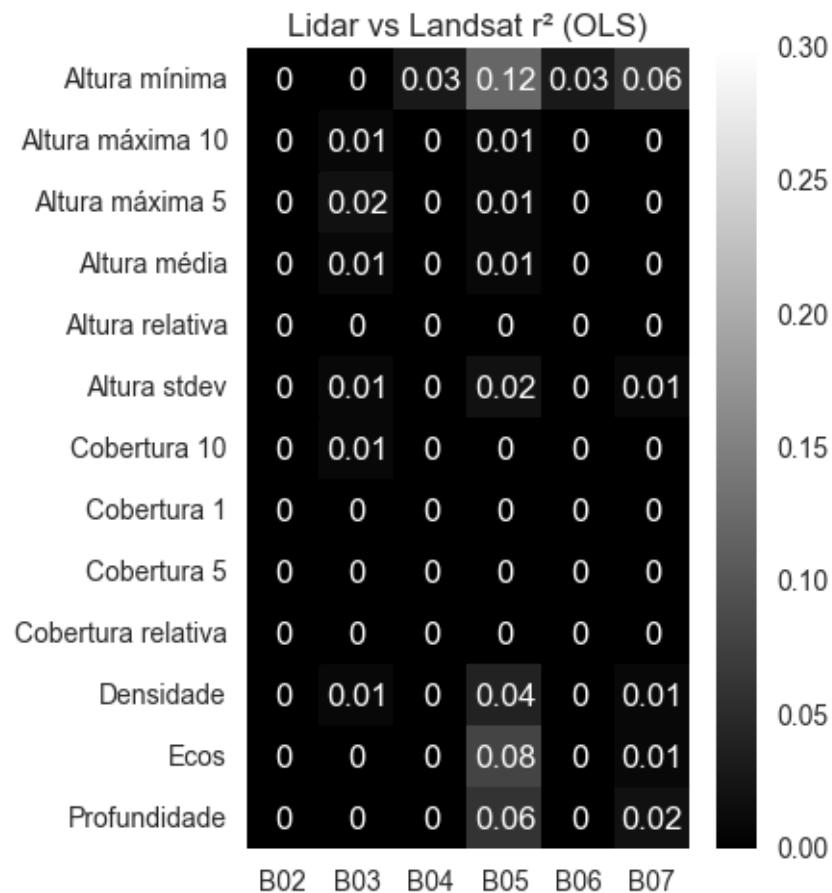


Figura 29 – Valores de r^2 para a regressão linear (OLS) da comparação pixel a pixel das bandas do Landsat-8/OLI (2014) para a Cantareira 1 com aplicação de correção topográfica e atmosférica DOS1 (eixo x) e das variáveis de estrutura vegetal obtidas por LiDAR (eixo y).

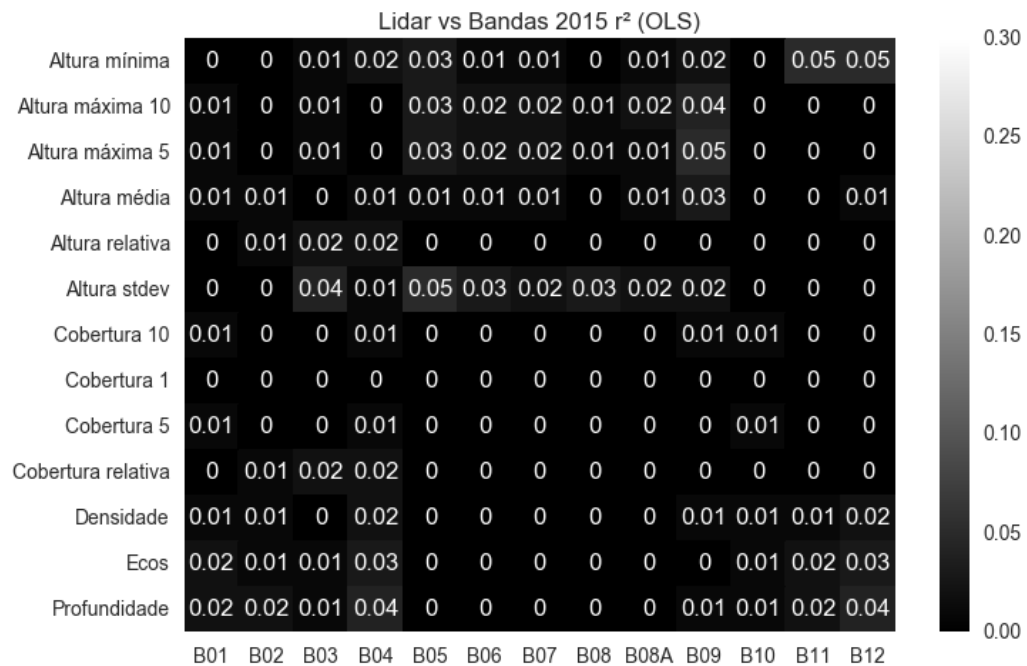


Figura 30 – Valores de r^2 para a regressão linear (OLS) da comparação pixel a pixel das bandas do Sentinel 2 (2015) com aplicação de correção topográfica e atmosférica DOS1 (eixo x) e das variáveis de estrutura vegetal obtidas por LiDAR (eixo y).

Ademais, é possível verificar que os resultados obtidos com o uso dos 338 buffers aleatórios como forma de amostragem (Figura 31) foram bastante superiores aos dos obtidos com a comparação pixel a pixel (Figura 28).

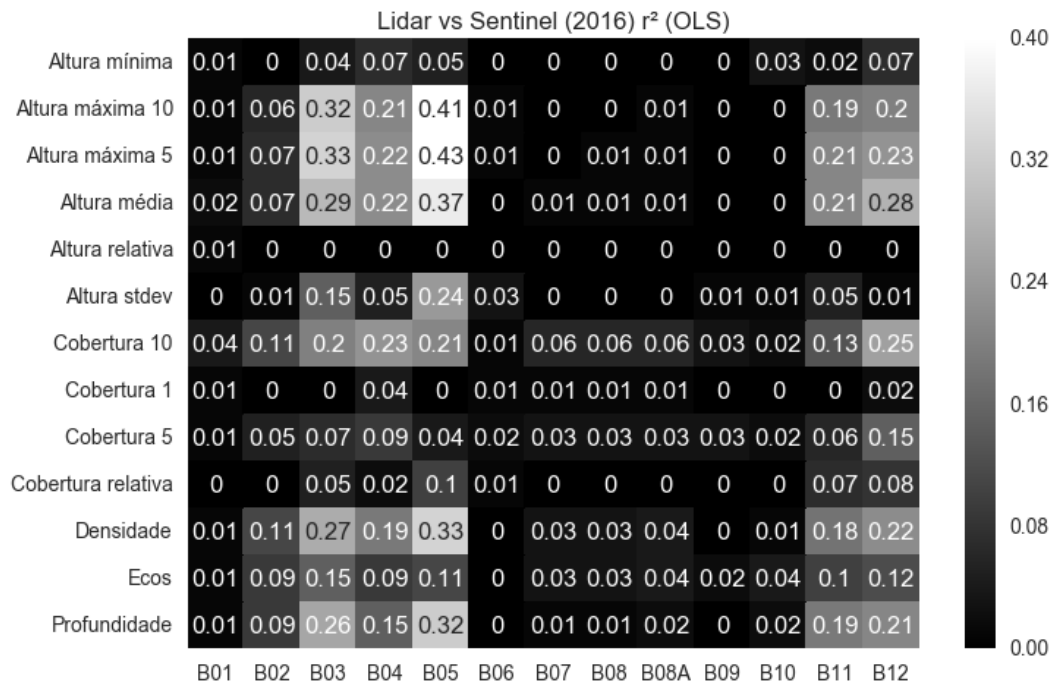


Figura 31 - Valores de r^2 para a regressão linear (OLS) da comparação por buffers das bandas do Sentinel 2 (2016) com aplicação de correção topográfica e atmosférica DOS1 (eixo x) e das variáveis de estrutura vegetal obtidas por LiDAR (eixo y).

Além das imagens citadas acima, a imagem de 08 de janeiro de 2016 do Sentinel 2 foi utilizada, mas apenas para testes de alguns dos modelos gerados, como foi mostrado na seção 3.2.3. Como existem nuvens na área Cantareira 1 nessa imagem, os testes foram feitos apenas na área Cantareira 2.

Considerando todas as informações fornecidas até o momento nesta seção, os resultados apresentados a seguir são referentes à comparação da amostragem com buffers das variáveis derivadas de dados LiDAR e da imagem do Sentinel 2 de 23 de dezembro de 2016 com a aplicação da correção topográfica.

Inicialmente, pode-se observar no Quadro 10 e na Figura 32 que assim como as variáveis de campo, as variáveis LiDAR diferiram bastante em termos de variabilidade. Dentre elas, a variável Cobertura 1 (proporção de áreas com mais de 1 m de altura dentro do buffer) foi uma das que apresentou menor variação ($CV = 0,0075$). Essa variável foi pensada para ser equivalente à porcentagem de cobertura coletada em campo (C), e ambas apresentam variações baixas de magnitude similar ($CV = 0,0058$ para C). Variáveis como Altura relativa (Altura relativa aos pixels vizinhos), Cobertura relativa (Proporção de áreas maiores que a Altura relativa dentro do buffer), Cobertura 5 (Áreas com menos de 5 metros de altura consideradas como clareiras) e Ecos (Número de ecos por pulso) também apresentaram baixa variabilidade ($CV = 0,06$ e $0,0676$ respectivamente). Já as variáveis Altura máxima 10 (Altura máxima na área de um pixel do Sentinel 2), Altura máxima 5 (Média da altura máxima em quatro quadrantes de 5 por 5 metros dentro de um pixel do Sentinel 2) e Altura média (média da imagem de altura dos primeiros ecos no buffer) apresentaram variações mais altas ($CV = 0,17$, $0,18$ e $0,21$ respectivamente). A variável com maior CV ($0,31$) foi a Altura mínima (média da imagem de altura dos últimos ecos no buffer), porém, mesmo as maiores variações encontradas para as variáveis derivadas do LiDAR não se comparam às das variáveis de campo, que em geral foram maiores.

Quadro 10 – Valores de média, desvio padrão (Std) e coeficiente de variação (CV) para as variáveis derivadas dos dados de campo de todos os 338 buffers aleatórios.

Variável	Unidade	Média	Std	CV
Altura mínima	m	3,78	1,17	0,31
Altura máxima 10	m	18,55	3,16	0,17
Altura máxima 5	m	16,77	2,96	0,18
Altura média	m	12,92	2,68	0,21
Altura stdev	m	2,9	0,68	0,23
Cobertura 10	-	0,73	0,2	0,27
Cobertura 1	-	1	0,01	0,0075
Cobertura 5	-	0,95	0,06	0,0600
Cobertura relativa	-	0,54	0,04	0,0677
Densidade	ecos/m	0,37	0,08	0,21
Ecos	-	3,17	0,21	0,0676
Profundidade	m	9,14	2,4	0,26

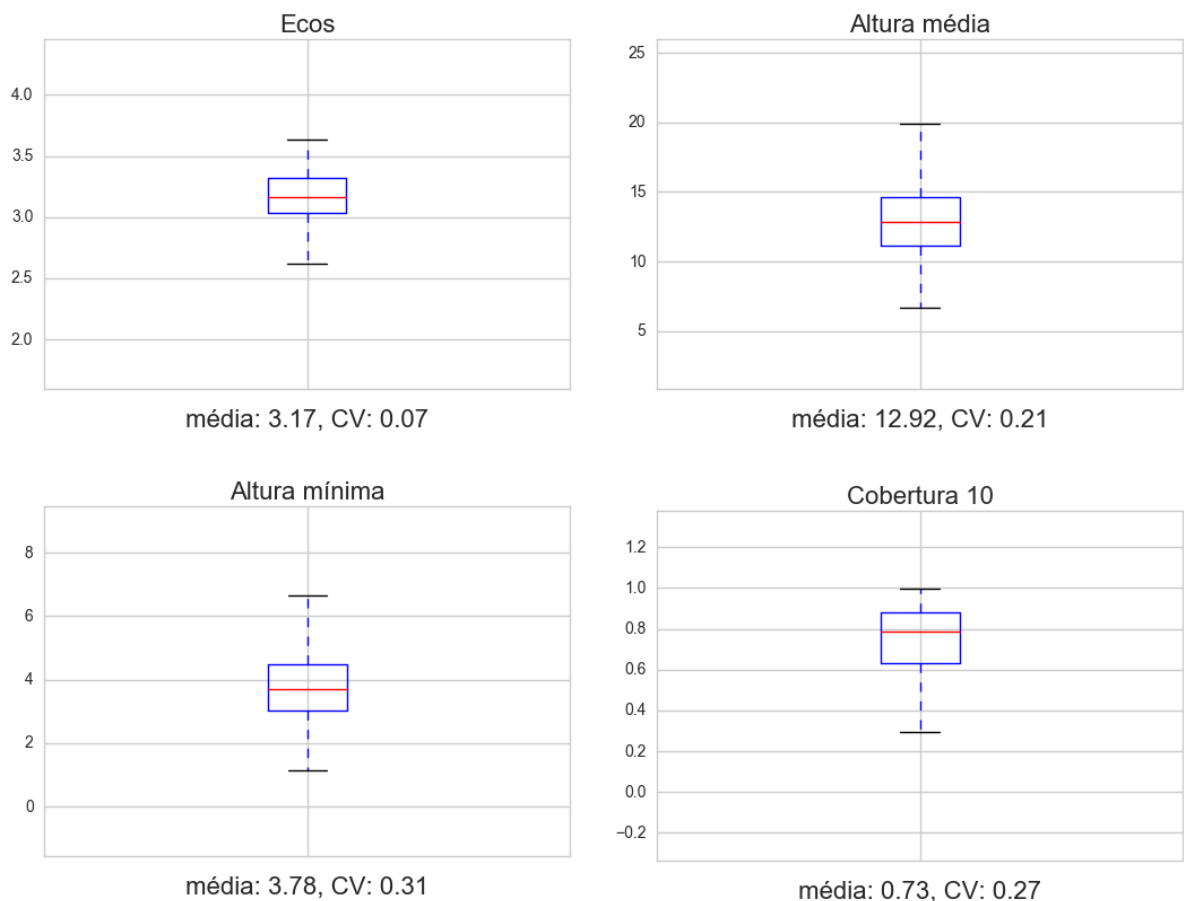


Figura 32 – *Boxplots* mostrando a dispersão de algumas das variáveis derivadas de LiDAR com os valores de média e coeficiente de variação (CV) para todos os 338 buffers gerados.

Quanto às correlações de Spearman entre as variáveis derivadas de LiDAR (Figura 33), nota-se que a maioria das variáveis derivadas da altura dos primeiros ecos (CHM) tem relações fortes e positivas entre si ($r > 0,8$). As exceções nesse caso são as Coberturas 1 e 5, a Altura relativa e a Densidade (Número de ecos por profundidade), que possuem correlação positivas, mas em menor magnitude ($r < 0,4$). Também é possível observar fortes correlações negativas entre a variável Densidade e as variáveis Altura máxima 10 ($r = -0,87$), Altura máxima 5 ($r = -0,91$) e Altura média ($r = -0,91$), já que todas tem relações com a variável Profundidade (Altura média menos Altura mínima) (r entre 0,81 e 0,89), usada como denominador no cálculo da Densidade. Já a variável Ecos (Número médio de ecos por pulso do LiDAR) teve correlações razoáveis com as alturas derivadas do CHM (r entre 0,53 e 0,67) e com a Cobertura 10 ($r = 0,7$). Ecos teve uma correlação forte e negativa com Densidade ($r = -0,76$), pois apesar de ser o numerador em seu cálculo, também tem correlação com as variáveis de altura, conforme mencionado e, conseqüentemente, com a própria profundidade ($r = 0,86$). Assim, maiores valores de Ecos significam maiores valores de Profundidade, aparentemente em maior proporção, reduzindo os valores de Densidade com o aumento de Ecos.

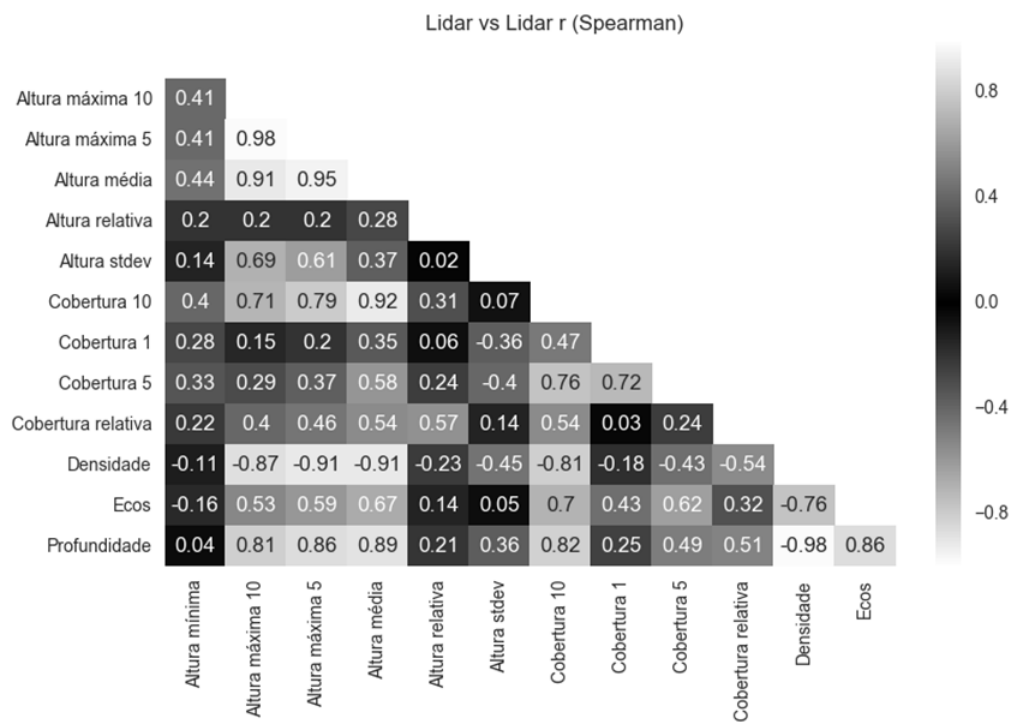


Figura 33 – Valores de r para a correlação (Spearman) das variáveis derivadas dos dados de LiDAR.

3.2.2 Relações com o Sentinel 2

Nas regressões OLS entre parâmetros biofísicos e imagens do Sentinel 2, assim como para os dados de campo, as medidas de altura apresentaram resultados melhores se comparadas a outras variáveis utilizadas. As variáveis Densidade, Profundidade, Cobertura 10 e Altura stdev apresentaram bons resultados em alguns casos, mas como foi visto, com exceção de Altura Stdev, elas apresentam correlações fortes com variáveis de altura e podem estar mais relacionadas a ela do que à densidade da vegetação. Na comparação de bandas individuais (Figura 31 – seção 3.2.1), os resultados mostram que B5 (705 nm) foi a banda mais bem-sucedida, seguida de B3 (560 nm) e das bandas B11 (1610 nm) e B12 (2190 nm) na comparação com as variáveis LiDAR. Já nas regressões OLS com índices de vegetação, o índice que apresentou a melhor relação foi o S2REP, seguido do NDWI em alguns casos (Figura 34). Novamente pôde-se observar uma relação melhor com variáveis de altura, com as variáveis densidade e profundidade apresentando boas relações. Diferente das análises com os dados de campo, o NDVI, o RVI, o SAVI e o EVI não apresentaram resultados melhores que os outros índices de vegetação.

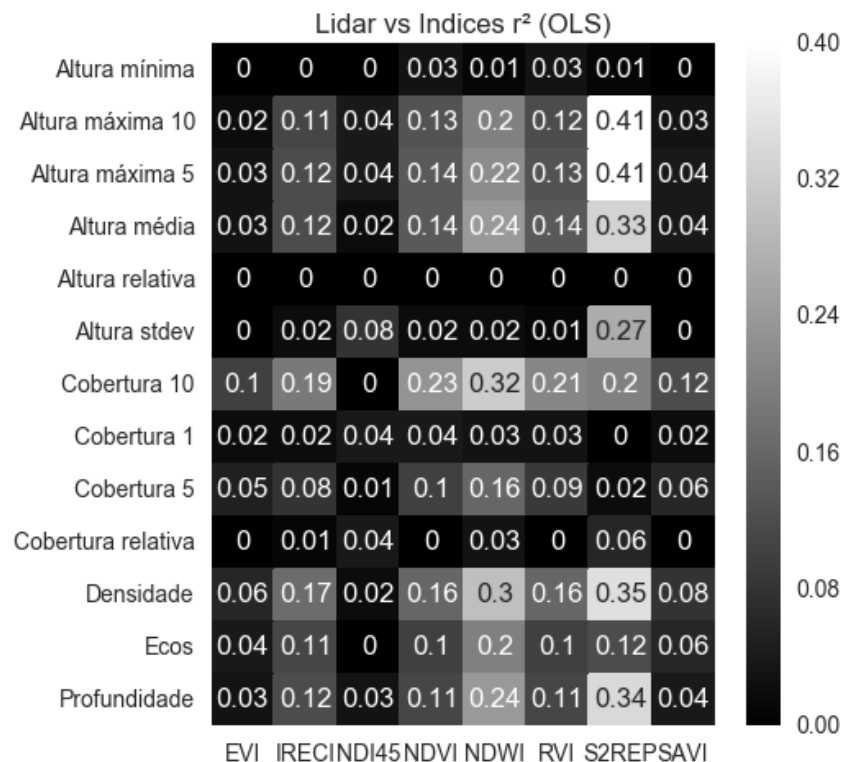


Figura 34 – Valores de r^2 para a regressão linear (OLS) da comparação por buffers dos índices de vegetação derivados das imagens do Sentinel 2 (2016) com aplicação de correção topográfica e atmosférica DOS1 (eixo x) e das variáveis de estrutura vegetal obtidas por LiDAR (eixo y).

A relação entre as razões das bandas e as variáveis de altura não mostrou muita melhora se comparada com as relações das bandas individuais ou índices de vegetação. Apesar disso, as razões das bandas B5 (705 nm) e bandas do NIR (B7 - 783 nm, B8 – 842 nm e B8A 865 nm) pareceram melhorar os resultados da comparação das imagens com a variável Densidade (Figura 35), com aumentos de até 0,05 em comparação com os melhores modelos de regressão da variável com bandas individuais e índices de vegetação (Densidade vs S2REP – $r^2 = 0,35$ e Densidade vs B5 – $r^2 = 0,33$).

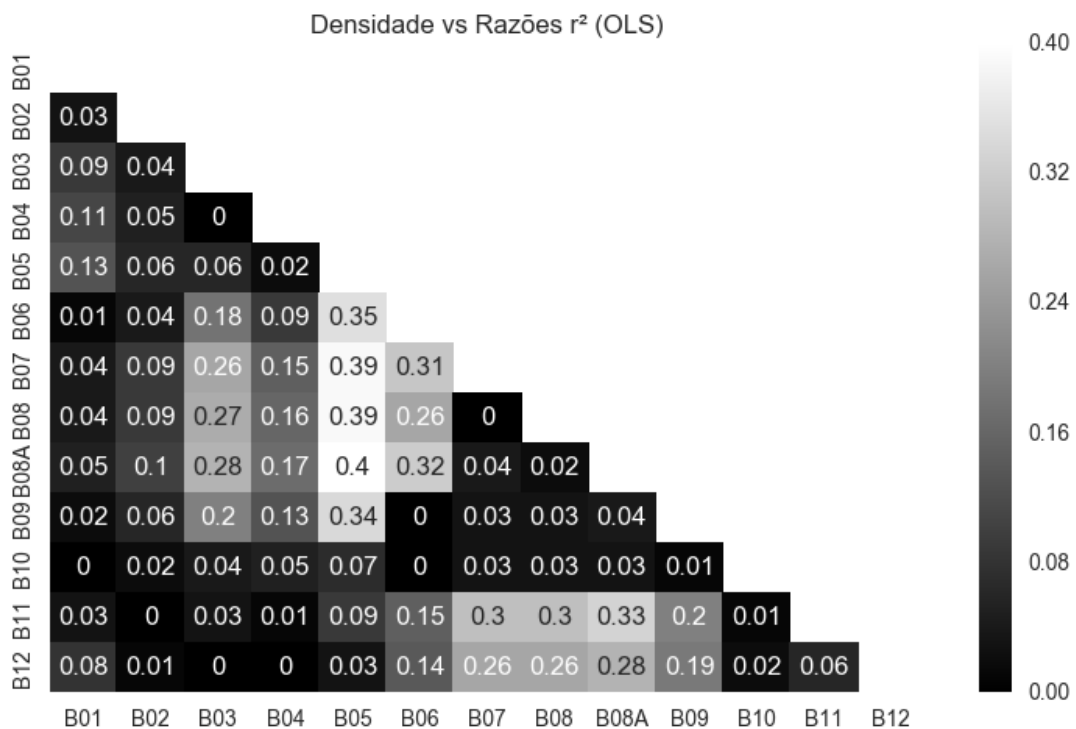


Figura 35 – Valores de r^2 para a regressão linear (OLS) da comparação por buffers das razões das bandas do Sentinel 2 com aplicação de correção topográfica e atmosférica DOS1 (eixos x e y) e da variável densidade.

Para ilustrar melhor as relações entre a variável Altura máxima 5 e Altura stdev com a banda 5 (B5) indicadas na Figura 27 no início da seção 3.2.1, o gráfico de dispersão entre estas variáveis é apresentado na Figura 36. Com base no modelo de regressão OLS resultante da análise destas variáveis, foram criadas imagens feitas com a predição das variáveis LiDAR através da aplicação da equação nas bandas do Sentinel 2. A comparação entre essas estimativas e os valores das variáveis de controle resultou em mapas de erros em que é possível observar que o ruído não é distribuído aleatoriamente na imagem e encontra-se concentrado em algumas regiões

específicas, o que sugere a existência de um ou mais componentes espaciais importantes causando parte das variações observadas. A Figura 37 mostra o mapa de erros da regressão entre B5 (705 nm) e Altura máxima 5 como exemplo disso.

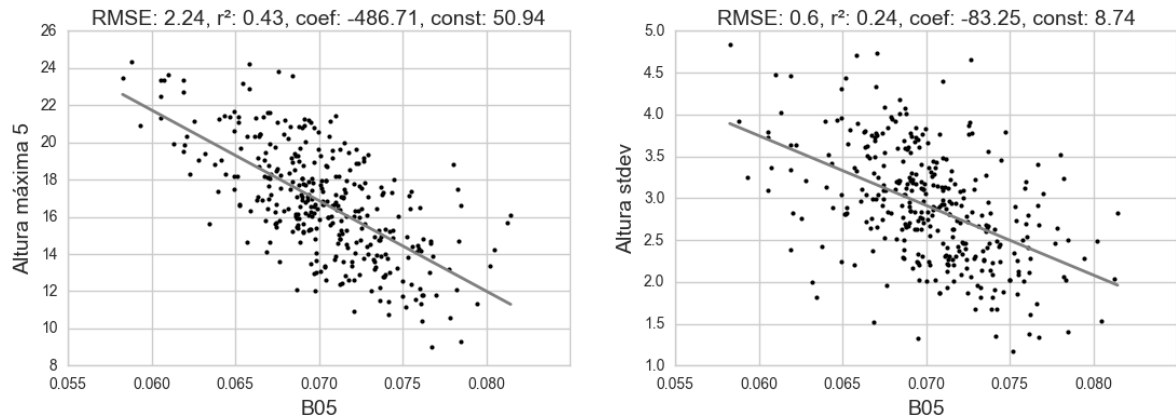


Figura 36 – Gráfico de dispersão da relação entre B5 (705 nm) com aplicação de correção topográfica e atmosférica (eixo x) e as variáveis derivadas de LiDAR (eixo y) altura máxima 5 (esquerda) e altura stdev (direita) com a linha do modelo de regressão linear (OLS) desenhada. Acima do gráfico, os valores de r^2 , p, coeficiente de multiplicação e constante a ser adicionada na fórmula do modelo ($y = 50,94 + x * -486,71$).

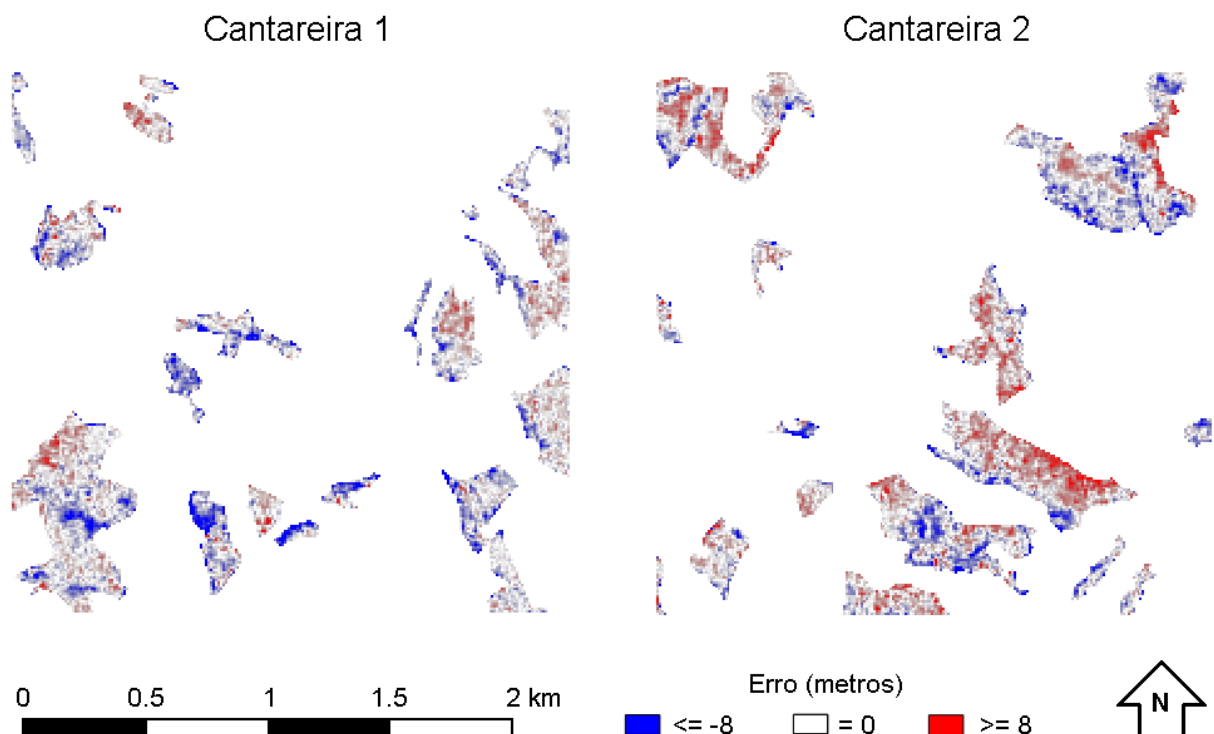


Figura 37 – Mapa de erros do modelo de regressão da relação entre B5 e a altura máxima 5 para as áreas de floresta. Pixels mais azulados representam subestimativas e pixels mais avermelhados superestimativas. Cores mais intensas representam erros maiores.

Os resultados das regressões múltiplas (OLS) envolvendo mais de uma banda do Sentinel 2 por variável de estrutura também apresentaram melhoras no r^2 , especialmente para alturas no geral, profundidade e densidade. Alguns resultados particularmente interessantes envolvem B5 (705 nm), B8 (842 nm), B8A (865 nm) e B11 (1610 nm) em comparação com essas variáveis (Quadro 11). Adicionalmente, incluir parâmetros fisiográficos (aspecto, declividade e sombreamento) nos modelos de regressão múltipla não melhorou muito as estimativas (r^2), provavelmente por conta da correção topográfica já aplicada.

Quadro 11 – Lista contendo a fórmula e valor de r^2 ajustado, RMSE e RRSE de alguns dos melhores modelos de regressão OLS múltipla encontrados para variáveis de estrutura derivadas de dados LiDAR. A descrição das variáveis pode ser vista no Quadro 6 (seção 2.5). B1: 0,443 nm, B2: 0,490 nm, B3: 0,560 nm, B4: 0,665 nm, B5: 0,705 nm, B6: 740 nm, B7: 783 nm, B8: 842 nm, B8A: 865 nm, B9: 945 nm, B10: 1375 nm, B11: 1610 nm, B12: 2190 nm.

Variável	Equação	r^2	RMSE	RRSE
Altura mínima	$-1,15 + B05 * -100,58 + B10 * 515,86 + B11 * 166,45 + B12 * -318,28$	0,26	1,00 m	0,86
Altura máxima 10	$23,91 + B05 * -612,82 + B08A * 64,44 + B10 * 1125,92$	0,55	2,10 m	0,67
Altura máxima 5	$48,06 + B05 * -415,72 + B08A * 51,38 + B011 * -123,31$	0,55	1,98 m	0,67
Altura média	$40,27 + B05 * -339,58 + B08 * 56,8 + B11 * -125,51$	0,51	1,86 m	0,7
Altura relativa	$99,8 + B01 * 16,43 + B05 * -9,17 + B06 * 8,16 + B07 * -4,67 + B10 * 13,64 + B12 * -3,41$	0,04	0,12 m	0,97
Altura stdev	$3,74 + B05 * -105,53 + B08A * 14,87 + B10 * 206,01 + B11 * -74,1 + B12 * 158,52$	0,4	0,52 m	0,77
Cobertura 10	$1,98 + B05 * -17,54 + B08 * 4,34 + B12 * -19,83$	0,45	0,14	0,74
Cobertura 1	$1,0 + B02 * 1,67 + B03 * -1,59 + B04 * -1,59 + B05 * 0,95 + B09 * -0,28 + B11 * 0,44 + B12 * -0,92$	0,1	0,01	0,94
Cobertura 5	$0,92 + B01 * 8,28 + B03 * -16,23 + B04 * 16,08 + B06 * 1,52 + B08 * 1,89 + B08A * -2,0 + B09 * 1,73 + B11 * 5,47 + B12 * -18,08$	0,31	0,05	0,82
Cobertura relativa	$0,67 + B01 * 4,57 + B02 * -8,71 + B03 * 9,77 + B04 * 10,13 + B05 * -10,22 + B06 * 2,21 + B07 * -1,21 + B09 * 1,5 + B12 * -3,25$	0,22	0,03	0,87
Densidade	$-0,3 + B05 * 8,98 + B08A * -1,73 + B11 * 3,95$	0,53	0,05 e/m	0,68
Ecos	$4,12 + B03 * -56,99 + B04 * 57,96 + B08A * 5,74 + B11 * -13,74$	0,34	0,17	0,81
Profundidade	$24,3 + B04 * 460,94 + B05 * -366,44 + B08A * 64,25 + B11 * -148,88$	0,52	1,66 m	0,69

Os resultados dos modelos de regressão RF (*Random Forest*) utilizando as bandas do Sentinel 2 para comparação com as variáveis derivadas de LiDAR foram similares aos das regressões OLS, como pode ser visto pelos valores de r^2 (*out of bag*), RMSE e RRSE (Quadro 12). Ademais, as variáveis que mais se relacionam com as bandas do Sentinel 2 foram Altura máxima 5, Altura máxima 10, Altura média, Cobertura 10, Densidade e Profundidade. Para essas variáveis se observa uma grande importância de B3 (560 nm), B10 (1375 nm), B11 (1610 nm), B12 (2190) e principalmente B5 (705 nm) nas relações entre as bandas e as variáveis de um modo geral (Figura 38). Especificamente para a Cobertura 10, no entanto, B4 (665 nm) parece ter mais importância. Diferentemente dos modelos OLS apresentados até o momento, os valores de RMSE e RRSE gerados aqui foram obtidos em áreas de teste não utilizadas para montar os modelos.

Os resultados da validação de modelos, tanto OLS quanto RF, são mostrados na seção 3.2.3.

Quadro 12 – Modelos de regressão *Random Forest* para cada variável derivada dos dados LiDAR com as quatro bandas mais importantes e os valores de r^2 *out of bag*, RMSE e RRSE. A descrição das variáveis pode ser vista no Quadro 6 (seção 2.5). B1: 0,443 nm, B2: 0,490 nm, B3: 0,560 nm, B4: 0,665 nm, B5: 0,705 nm, B6: 740 nm, B7: 783 nm, B8: 842 nm, B8A: 865 nm, B9: 945 nm, B10: 1375 nm, B11: 1610 nm, B12: 2190 nm.

Variável	Valores de importância das variáveis	r^2	RMSE	RRSE
Altura mínima	B10 (0,15), B12 (0,13), B09 (0,11), B04 (0,09)	0,27	0,93 m	0,86
Altura máxima 10	B05 (0,35), B11 (0,09), B10 (0,08), B03 (0,08)	0,54	2,04 m	0,64
Altura máxima 5	B05 (0,34), B03 (0,10), B11 (0,09), B10 (0,06)	0,56	1,87 m	0,62
Altura média	B05 (0,22), B03 (0,16), B12 (0,14), B04 (0,04)	0,50	1,85 m	0,66
Altura relativa	B01 (0,11), B03 (0,10), B09 (0,09) B11 (0,09)	-0,09	0,12 m	1,02
Altura stdev	B05 (0,29), B10 (0,13), B01 (0,09), B12 (0,06)	0,32	0,46 m	0,76
Cobertura 10	B04 (0,23), B12 (0,18), B05 (0,07), B07 (0,07)	0,51	0,15	0,7
Cobertura 1	B04 (0,19), B01 (0,15), B10 (0,09), B08 (0,08)	-0,15	0,01	0,99
Cobertura 5	B12 (0,25), B10 (0,12), B04 (0,12), B09 (0,10)	0,26	0,05	0,8
Cobertura relativa	B05 (0,17), B12 (0,13), B09 (0,09), B03 (0,08)	0,10	0,04	0,95
Densidade	B05 (0,22), B03 (0,10), B12 (0,10), B11 (0,09)	0,51	0,06 e/m	0,67
Ecos	B12 (0,17), B10 (0,12), B11 (0,10), B03 (0,09)	0,26	0,2	0,86
Profundidade	B05 (0,20), B03 (0,14), B12 (0,13), B11 (0,09)	0,43	1,88 m	0,72

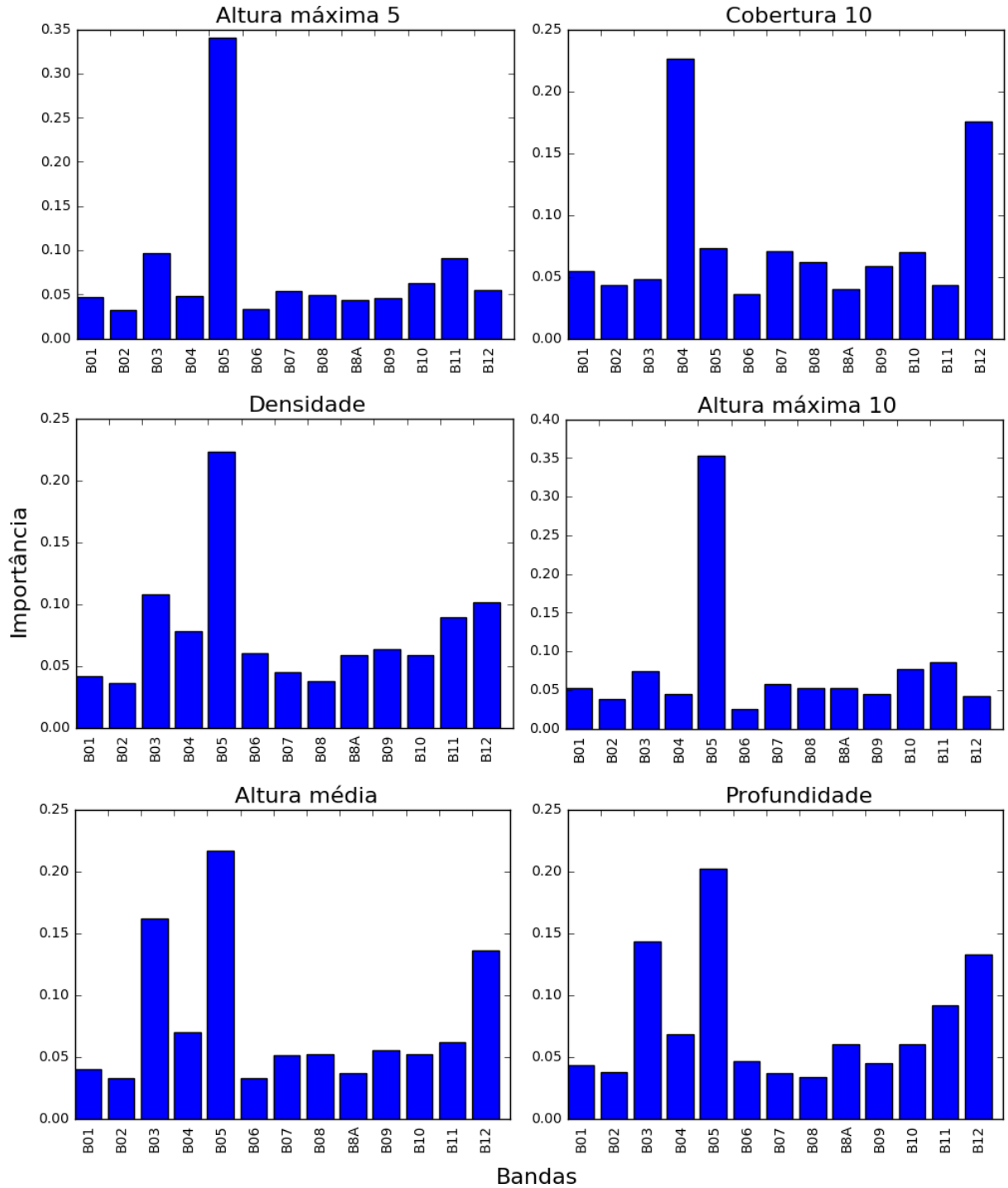


Figura 38 – Valores das importâncias das bandas do Sentinel 2 usadas para gerar os modelos *Random Forest* para as variáveis Altura máxima 5, Cobertura 10, Densidade, Altura máxima 10, Altura média e Profundidade.

3.2.3 Teste de modelos

Para as variáveis derivadas dos dados LiDAR foi possível testar tanto os modelos OLS quanto os RF por conta da maior quantidade de amostras e da melhor performance de maneira geral. As bandas incluídas nos modelos apresentados a seguir foram selecionadas manualmente com a ajuda de algoritmos baseados nos valores de r^2 para os modelos OLS e na importância das bandas para modelos RF. É importante ressaltar que as bandas selecionadas para os modelos OLS e RF para uma mesma variável não são exatamente as mesmas, já que os dois tipos de modelo são diferentes. Como as variáveis que apresentaram melhores relações com as bandas do Sentinel 2 na seção 3.2.2 foram Altura máxima 5, Altura máxima 10, Altura média, Cobertura 10, Profundidade e Densidade, os resultados apresentados nesta seção são com enfoque nessas variáveis. Nesta etapa, para testar a confiabilidade dos modelos gerados com imagens do Sentinel 2, a validação foi feita de duas maneiras distintas (detalhes na seção 2.6):

- Separando os dados da imagem de 12 de dezembro de 2016 para as áreas Cantareira 1 e 2 em grupos aleatórios de treino (75% dos dados) e teste (25% dos dados).
- Utilizando todos os dados da imagem de 12 de dezembro de 2016 como treino e os dados correspondentes somente à Cantareira 2 da imagem de 01 de janeiro de 2016 como validação, por conta da presença de nuvens na Cantareira 1 nessa imagem.

Os resultados do primeiro método de validação se mostraram satisfatórios. Dentre as cinco mil combinações aleatórias de áreas de treino e teste, foram obtidas a distribuição e a mediana dos valores de RMSE para todos os modelos OLS gerados. Com base na figura 39, pode-se observar que as distribuições dos valores de RMSE dos cinco mil modelos gerados para cada variável são bem próximas a uma distribuição normal, permitindo a escolha de modelos com valores de RMSE mediano. As previsões feitas usando esses modelos para as variáveis são condizentes com os valores reais por todo o gráfico de dispersão (Figura 40), com erros bem parecidos aos de modelos feitos com todos os dados sem validação apresentados nas seções anteriores. Já para os modelos RF, por conta da maior complexidade e tempo de processamento, não foram geradas repetições na aleatorização dos dados usados

para treino e teste. Ao invés disso, foi gerado um modelo para cada variável incluindo cinco mil árvores de decisão como margem de segurança, já que mais de mil árvores não alteravam os resultados de forma significativa (RMSE e distribuição das predições vs valores reais) dos modelos, conforme mencionado na seção 2.6. Mais uma vez, os modelos OLS e RF tiveram resultados parecidos, mostrando que as relações entre as imagens do Sentinel 2 e das variáveis citadas anteriormente podem ser relativamente bem explicadas por modelos mais simples (Figuras 40 e 41).

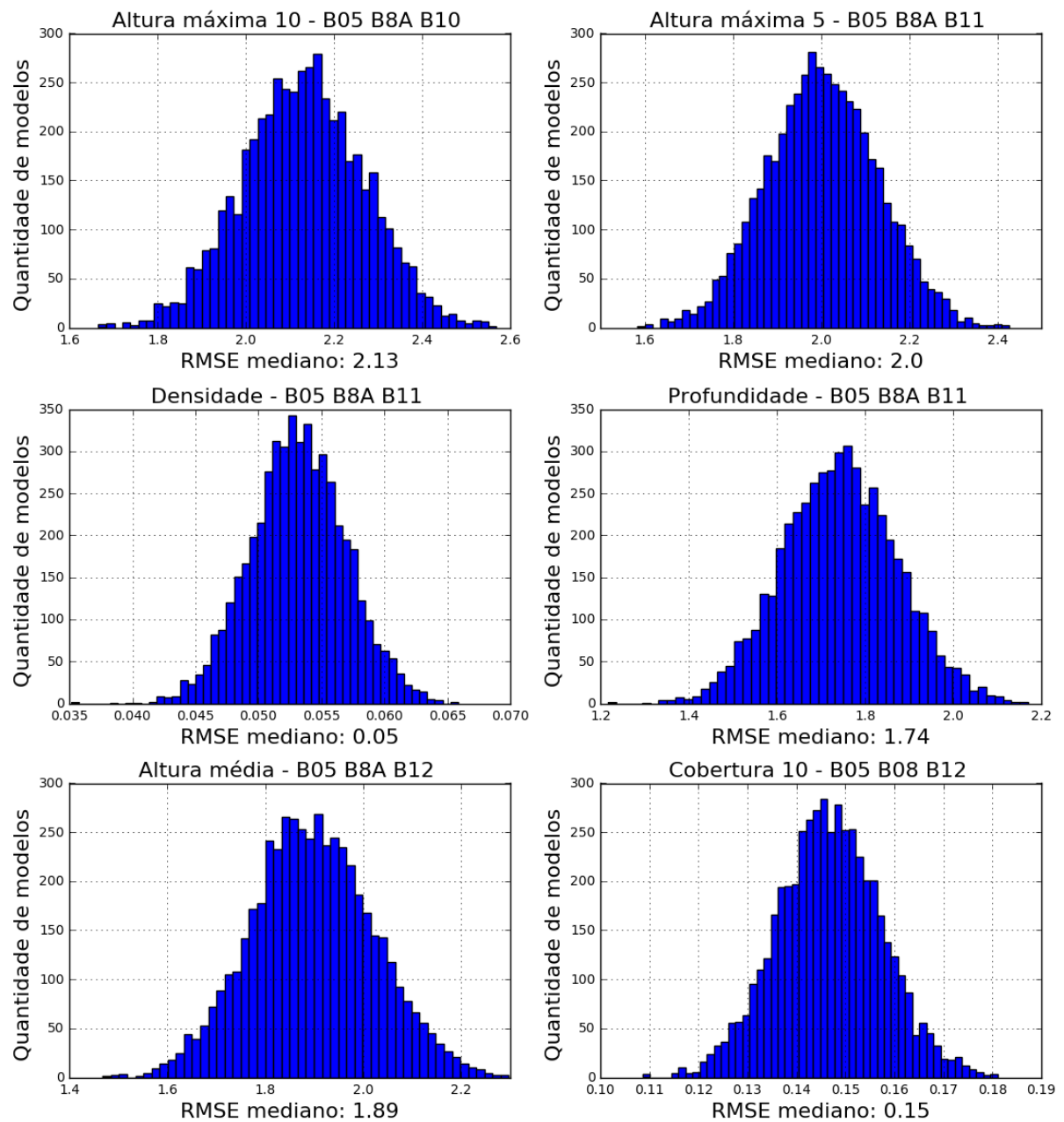


Figura 39 – Histograma mostrando a distribuição dos valores de RMSE para todos os cinco mil modelos OLS gerados com as bandas do Sentinel 2 (23/12/2016) para validação em áreas aleatórias para seis das variáveis derivadas de LiDAR.

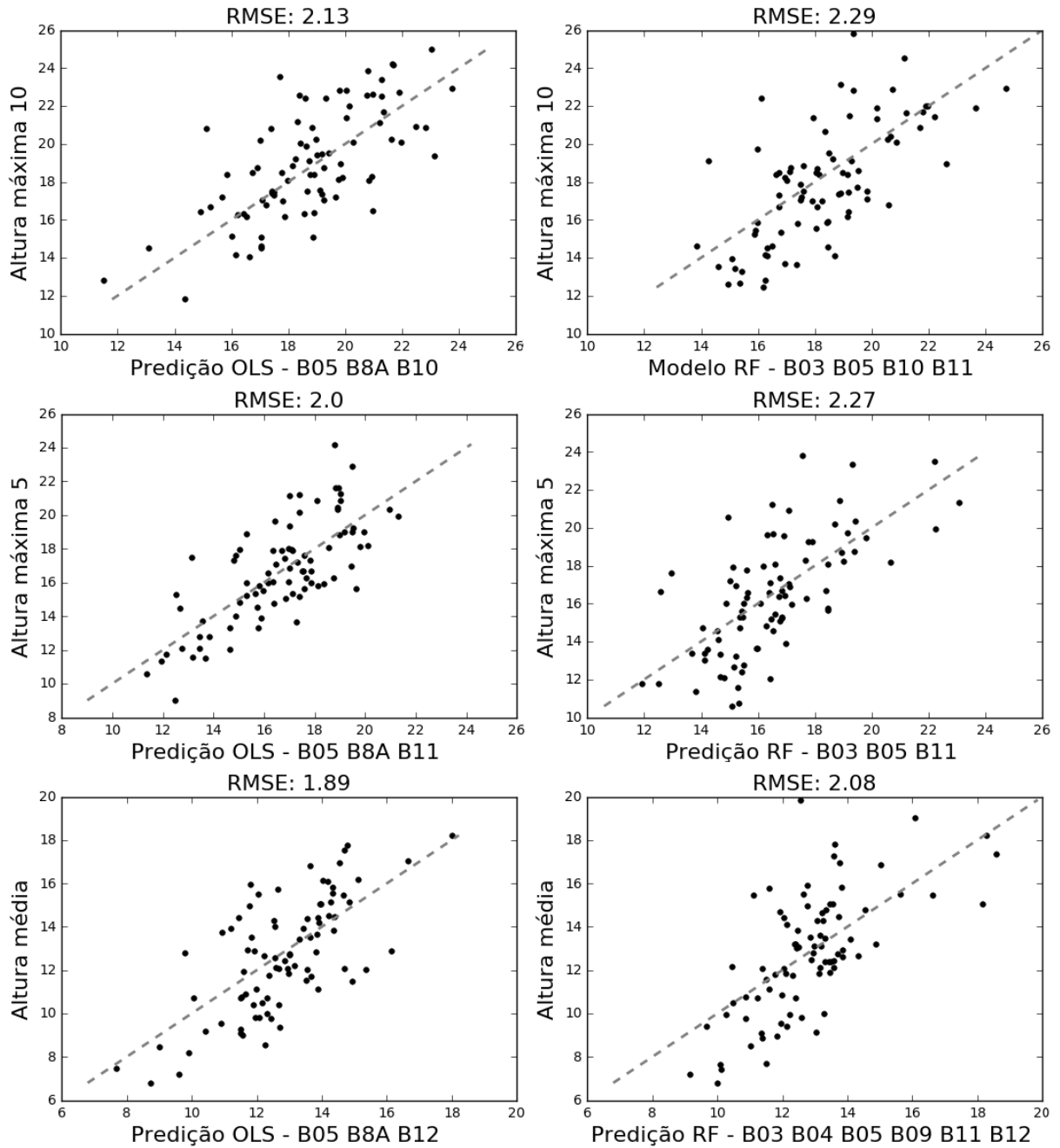


Figura 40 – Comparação entre as previsões (eixo x) com bandas do Sentinel 2 (23/12/2016) de modelos de regressão OLS (*ordinary least squares*) e RF (*Random Forest*) e os valores reais (eixo y) em metros das variáveis *Altura média*, *Altura máxima 5* e *Altura máxima 10* para as áreas usadas como teste. Os modelos OLS tem o erro (RMSE) igual à mediana da distribuição de erros dos cinco mil modelos testados.

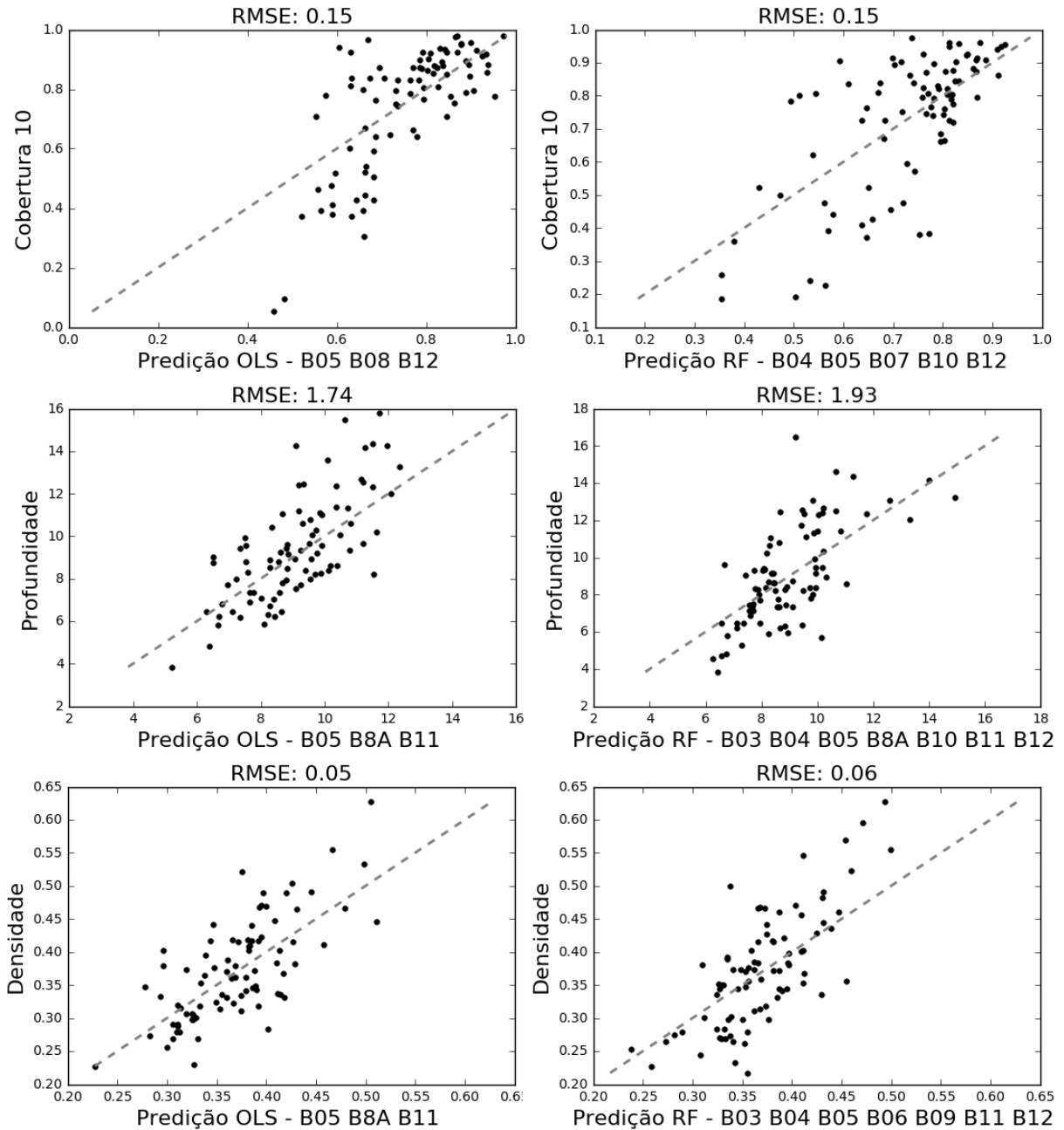


Figura 41 – Comparação entre as previsões (eixo x) com bandas do Sentinel 2 (23/12/2016) de modelos de regressão OLS (*ordinary least squares*) e RF (*Random Forest*) e os valores reais (eixo y) das variáveis Cobertura 10, Profundidade (m) e Densidade (e/m) para as áreas usadas como teste. Os modelos OLS tem o erro (RMSE) igual à mediana da distribuição de erros dos cinco mil modelos testados.

Para o segundo método de validação, usando a imagem de 23 de dezembro de 2016 para treino nas áreas Cantareira 1 e 2 e a imagem de 08 de janeiro de 2016 para testes na área Cantareira 2 os resultados diferiram bastante. É possível observar que, mesmo para os melhores modelos, a predição não é capaz de reproduzir os valores reais, mostrando tendências bastante características (Figuras 42 e 43). Muitos dos

modelos, tanto OLS quanto RF, parecem superestimar valores mais baixos e subestimar valores mais altos para a imagem de teste na comparação com os valores reais, o que indica diferenças nos valores de reflectância das imagens usadas para treino e teste.

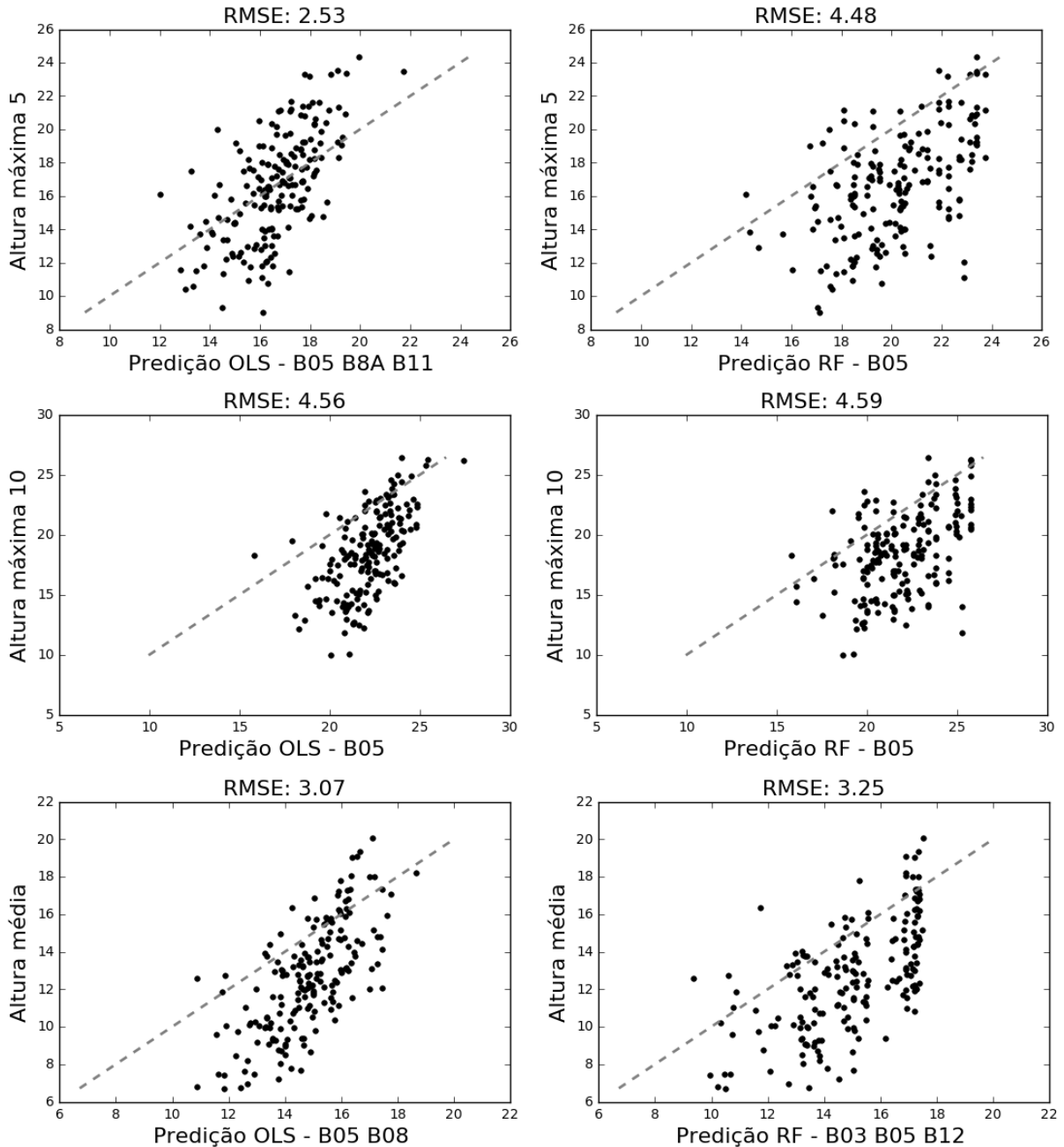


Figura 42 – Comparação entre as previsões (eixo x) com bandas do Sentinel 2 (08/01/2016) de modelos de regressão OLS (*ordinary least squares*) e RF (*Random Forest*) treinados na imagem de 23 de dezembro de 2016 e os valores reais (eixo y) em metros das variáveis Altura máxima 5, Altura máxima 10 e Altura média.

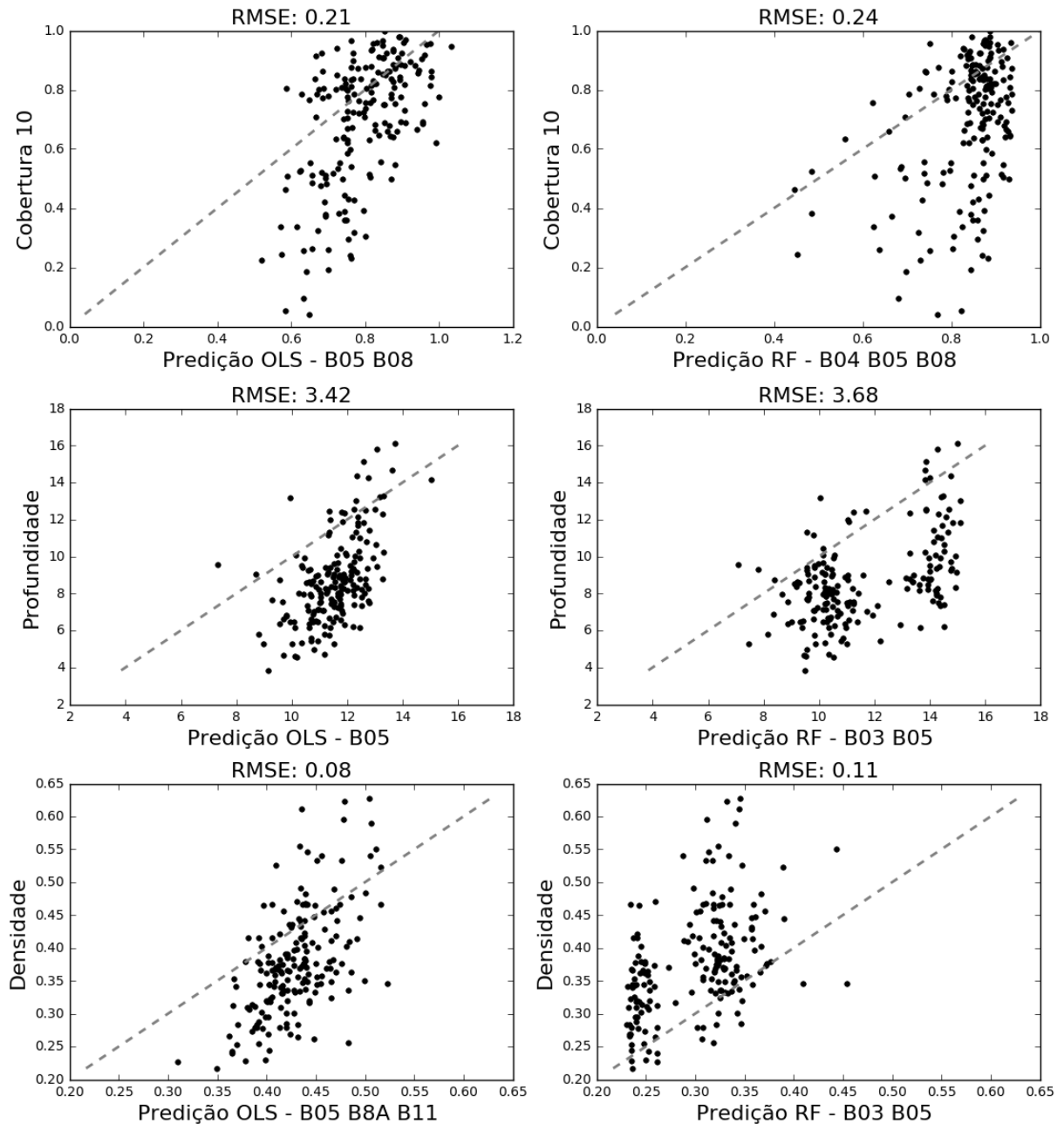


Figura 43 – Comparação entre as previsões (eixo x) com bandas do Sentinel 2 (08/01/2016) de modelos de regressão OLS (*ordinary least squares*) e RF (*Random Forest*) treinados na imagem de 23 de dezembro de 2016 e os valores reais (eixo y) das variáveis Cobertura 10, Profundidade (m) e Densidade (e/m).

Também é possível observar pelos gráficos das variáveis Altura máxima 5, Cobertura 10 e Profundidade (Figura 44) que modelos com mais bandas envolvidas tem um desempenho bastante inferior aos modelos com poucas bandas apresentados nas figuras anteriores. Além disso, é possível notar que os modelos RF são capazes de manter os valores preditos em uma escala mais próxima à dos valores reais se comparados aos modelos OLS.

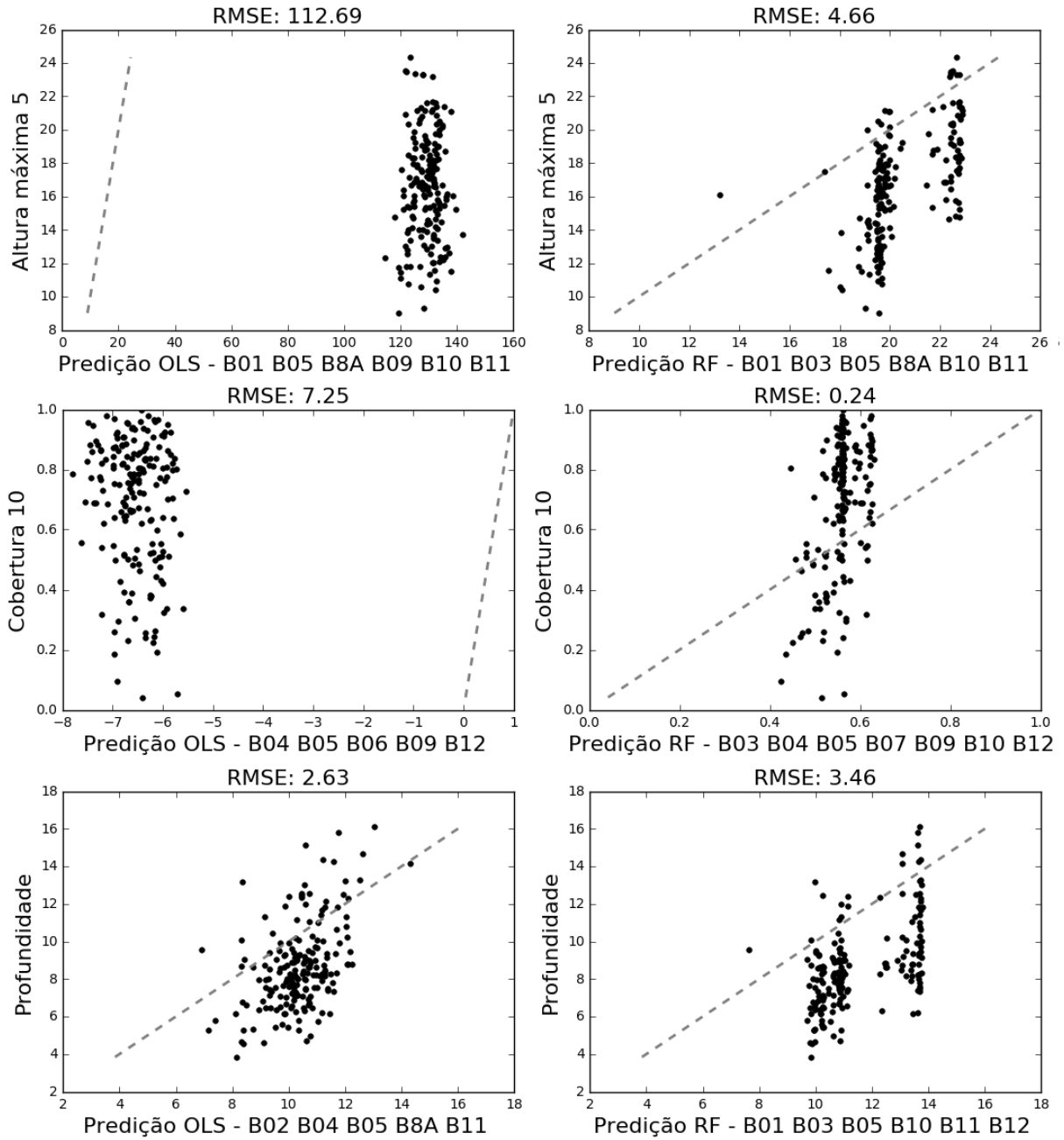


Figura 44 – Comparação entre as previsões (eixo x) com bandas do Sentinel 2 (08/01/2016) de modelos de regressão OLS (*ordinary least squares*) e RF (*Random Forest*) treinados na imagem de 23 de dezembro de 2016 e os valores reais (eixo y) das variáveis Cobertura 10, Profundidade (m) e Densidade (e/m).

Uma possível razão para as diferenças no desempenho dos modelos na estimativa de parâmetros biofísicos da vegetação testados nas imagens de 08 de janeiro de 2016 e 23 de dezembro de 2016 são os efeitos atmosféricos, como será melhor detalhado na seção 4. Para compensar por esses efeitos, o mesmo teste foi feito com imagens de nível 2A (Reflectância de Superfície) obtidas através do

complemento sen2cor da ESA (diferentemente dos testes feitos anteriormente com imagens corrigidas pelo método DOS1). Os resultados podem ser vistos na Figura 45 e mostram que o sen2cor é mais apropriado que o DOS1 para esse tipo de aplicação. Mesmo modelos com muitas bandas não são afetados pela mudança de imagens 2A, com os erros se distribuindo de maneira uniforme nos gráficos, similarmente ao ocorrido no primeiro método de validação.

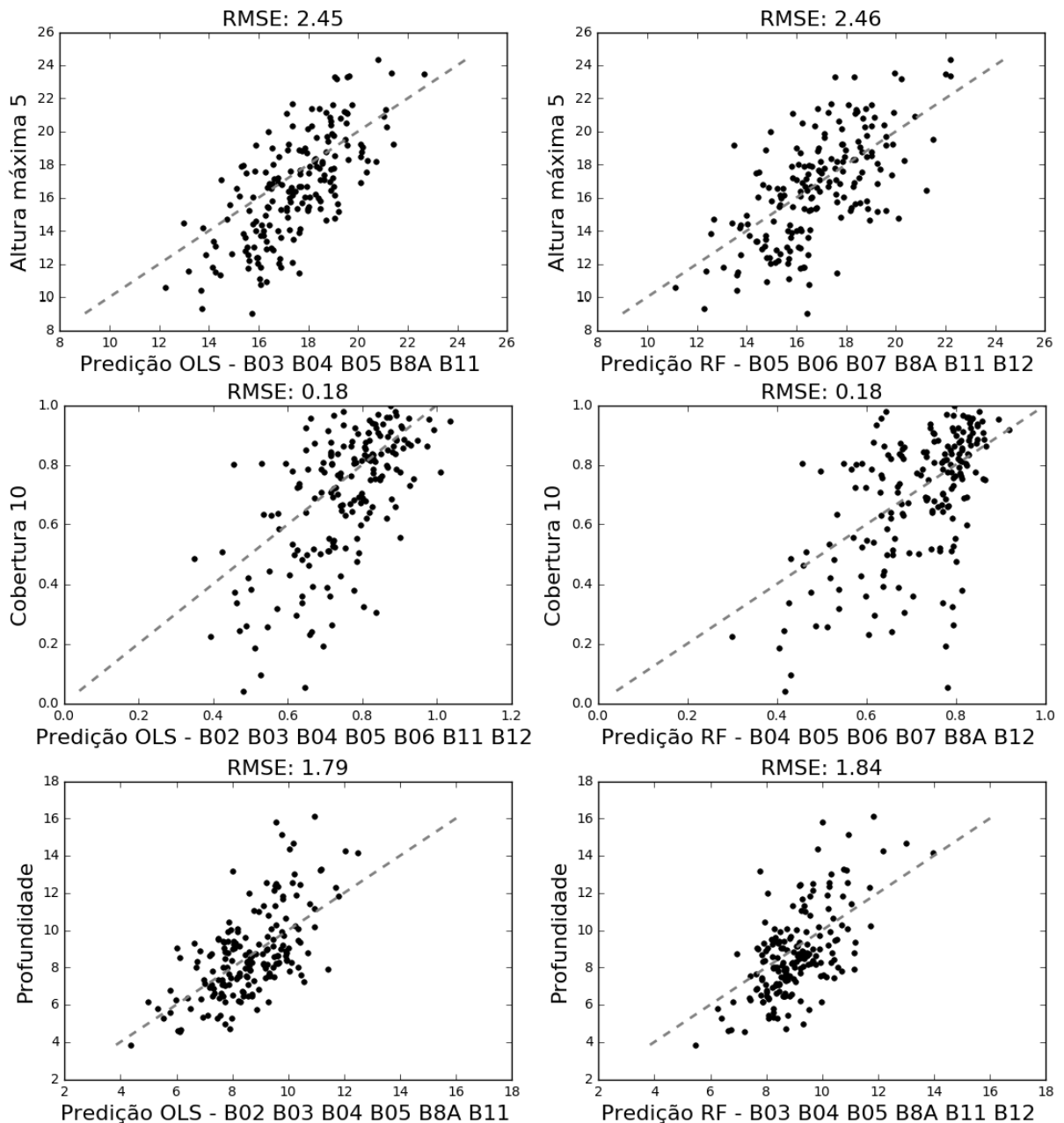


Figura 45 – Comparação entre as previsões (eixo x) com bandas do Sentinel 2 (08/01/2016) de modelos de regressão OLS (*ordinary least squares*) e RF (*Random Forest*) treinados na imagem de 23 de dezembro de 2016 e os valores reais (eixo y) das variáveis Cobertura 10, Profundidade e Densidade. Imagens em nível 2A (Reflectância de superfície) com o uso do sen2cor para correção atmosférica.

3.3 ESPECTROS DE REFLECTÂNCIA

No experimento de espectros de reflectância foram obtidas as assinaturas espectrais em reflectância de 1 e 4 camadas de folhas (Figura 46). Embora a reflectância entre camadas com número de folhas diferente tenha sido parecida, é possível que a maior quantidade de folhas presentes em um dossel real intensifique as diferenças. De acordo com o esperado, a reflectância no espectro visível diminui com o aumento do número de camadas de folha, mas aumenta na região do NIR e varia bastante na região do SWIR dependendo do comprimento de onda.

Espectros de reflectância para diferentes camadas de folhas

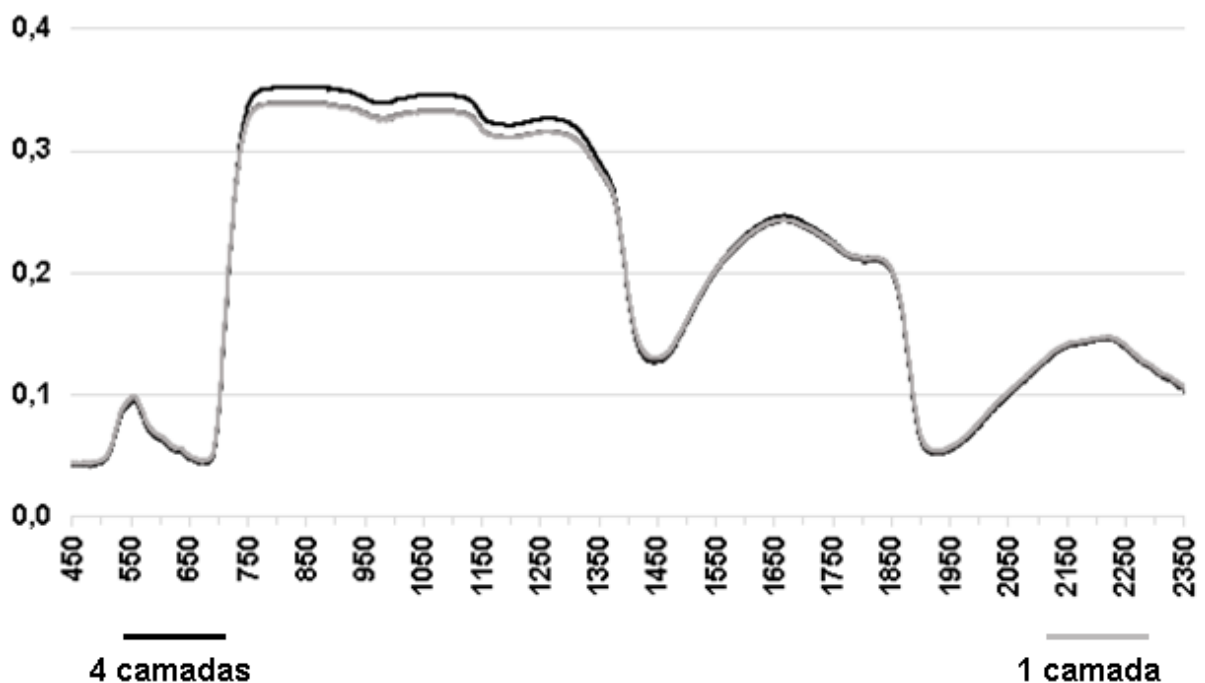


Figura 46 – Gráficos mostrando as reflectâncias (eixo x) das diferentes camadas de folhas observadas ao longo do espectro eletromagnético (eixo y).

O gráfico de variação entre 1 e 4 camadas de folhas gerado (Figura 47) mostra que a maior diferença ocorre na região do NIR, além de variações expressivas ocorrendo na região do verde, SWIR, e em uma pequena faixa do red-edge (690 a 720 nm). Essa faixa do red-edge que apresentou mais variação (698 a 716 nm) está bem próxima à B5 do Sentinel 2, que é centrada em 705 nm. A faixa do verde que apresentou mais variações também corresponde muito bem à B3, centrada em 560 nm. Já para a região do NIR como um todo, grandes variações podem ser observadas, e a banda que corresponde à maior variação é a 8A, centrada em 865 nm. Por fim, as

bandas 10, 11 e 12 do Sentinel 2 não parecem corresponder às regiões de maior variação do SWIR, que parece ocorrer próximo a 1880 nm.

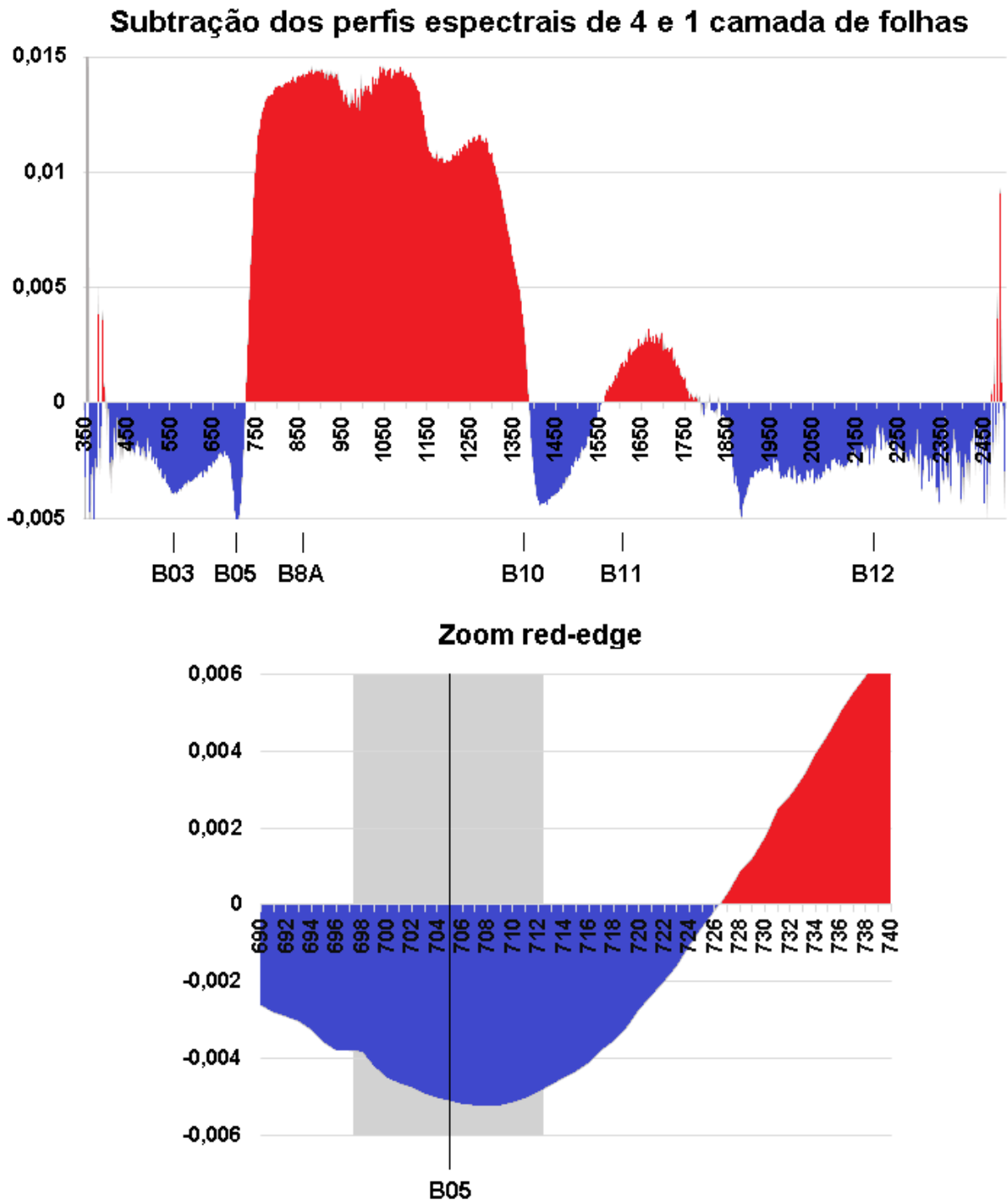


Figura 47 – Gráficos mostrando a diferença de reflectância (eixo y) entre 1 e 4 camadas de folhas ao longo do espectro eletromagnético em nanômetros (eixo x) com a posição das bandas 3, 5, 8A, 10, 11 e 12. Valores positivos (vermelho) significam aumento da reflectância com o número de folhas para o dado comprimento de onda, valores negativos (azul) significam o inverso. Abaixo, um recorte do início do *red-edge* mostrando a posição e largura da banda 5 (cinza) em detalhe.

4 DISCUSSÃO

Os resultados encontrados mostram relações claras entre alguns dos parâmetros biofísicos estudados e bandas do Sentinel 2, evidenciando o grande potencial dessas imagens para o estudo da vegetação. As relações encontradas com bandas individuais, índices de vegetação e razões entre bandas, no entanto, apresentam valores de coeficiente de determinação (r^2) próximo a 0,4 nos melhores casos, o que indica variações significativas nos modelos gerados. Utilizar mais de uma banda simultaneamente em regressões múltiplas parece melhorar as estimativas ($r^2 \cong 0,55$), mas não o suficiente para explicar uma proporção expressiva da variação dos modelos. Isso indica que os parâmetros biofísicos estudados aqui não podem ser preditos totalmente pelas imagens com as técnicas utilizadas. Quanto ao poder preditivo dos modelos, medido pelos valores de RMSE, os resultados são mais animadores. Os erros encontrados, principalmente nas previsões de algumas variáveis derivadas de LiDAR são pequenos se comparados aos seus valores de média e desvio padrão. Como exemplo disso, a variável Altura Máxima 5 varia aproximadamente entre 9 e 24 m, e seu RMSE é de 2,45 m. Considerando o intervalo de 15 m, o erro fica próximo a 16,3%, ou 14,4% se considerarmos a média de 17 m. Um valor razoável para aplicações que não necessitam de grande precisão.

Vale destacar que algumas das variáveis de controle utilizadas apresentaram baixa variabilidade e suas diferenças podem ser sutis, o que amplifica a influência de erros. No geral, variáveis derivadas das parcelas de campo apresentaram maiores variabilidades do que as variáveis derivadas dos dados LiDAR. Uma possível explicação para isso é a diferença na distribuição dessas variáveis no espaço. Apesar dos dados LiDAR terem sido coletados em uma área maior, eles se concentram nas regiões Cantareira 1 e 2, enquanto que as parcelas de campo estão espalhadas por toda a área de estudo, em alguns casos separadas por dezenas de quilômetros de distância. Parcelas distribuídas em locais diferentes aumentam as chances de mudanças em condições que podem afetar a organização estrutural da vegetação, como altitude (ALVES et al., 2010), composição do solo e clima (QUESADA et al., 2012), aumentando a variabilidade dos dados.

Outro fator que pode ter contribuído para a maior variabilidade das parcelas de campo é o tamanho da área amostrada, já que foi demonstrado na literatura que

valores de CV tendem a diminuir com o aumento da área amostrada para variáveis como AGB (Biomassa Acima do Solo – *Above Ground Biomass*), área basal e número de indivíduos amostrados (WAGNER et al., 2010). Além disso, Ruiz et al., (2014) mostram que a área mínima para estimar parâmetros como volume, biomassa e área basal deve ser de pelo menos 500 m², e de 300 m² para a cobertura do dossel. Como a área amostrada nas parcelas de campo foi de 150 m², é possível que essas variáveis não tenham chegado a valores estáveis para a região em que foram coletadas, o que pode aumentar o erro dos modelos construídos para o campo.

Também houveram variáveis com baixa variabilidade. Dentre elas, podemos destacar as de cobertura do dossel em específico, representada aqui pela variável C medida em campo e as variáveis Cobertura 1, 5 e 10 derivadas dos dados LiDAR. Os valores encontrados em média para essas variáveis, principalmente C (99%) e Cobertura 1 (100%), que são medidas em pontos de vista bem próximas do solo (1,3 e 1 metros respectivamente), foram altos. Isso por si só explica em parte essa baixa variação, já que os valores estão sempre muito perto de 100%. Esses valores são maiores que alguns dos valores encontrados na literatura medidos com densiômetro em florestas úmidas, com médias em torno de 90% (SUGANUMA et al., 2008, SUGANUMA et al., 2013), mas é importante levar em conta a subjetividade envolvida nas medidas feitas com densiômetro (COOK et al., 1995) para comparar esses valores aos de outros estudos na literatura. Além disso, é importante lembrar que para a cobertura medida em campo, áreas de clareira foram ativamente evitadas no estabelecimento das parcelas, aumentando os valores de cobertura e reduzindo ainda mais a variabilidade dessa variável. Ressalta-se que, para as variáveis de cobertura medidas com LiDAR, a resolução do CHM inicial gerado (0,5 metros) pode limitar a variabilidade, principalmente da Cobertura 1, já que aberturas no dossel com áreas menores do que 0,25 m² se misturaram com os valores de altura adjacentes e não aparecem como clareiras nessa variável se o valor médio for maior que 1 metro.

Outro fator importante é a variabilidade dos parâmetros biofísicos dentro de cada parcela ou buffer observados. As variáveis derivadas de LiDAR Altura stdev, Altura relativa e Cobertura relativa tiveram como objetivo representar essa variabilidade, mostrando a influência que as diferenças de altura teriam na resposta espectral das imagens. Embora Altura e Cobertura relativas não tenham relações muito boas com imagens do Sentinel 2 segundo os modelos de regressão gerados,

para Altura stdev em particular foram encontradas relações com r^2 razoavelmente alto (Seção 3.2.2 - Quadros 11 e 12), com principal influência da banda 5. Isso mostra que variabilidades de altura podem ter um efeito importante na resposta espectral das florestas, já que os efeitos de sombra causados por essas diferenças de altura podem reduzir os valores de reflectância do dossel. Araujo, Santos e Shimabukuro (2000) elencam esse efeito como a provável causa da redução de valores de SAVI em áreas de Floresta Amazônica madura se comparado a estágios sucessionais menos avançados, já que estas apresentam mais variação no dossel e geram mais áreas sombreadas.

Um fator que poderia contribuir para uma melhora nos modelos é considerar as relações entre os parâmetros biofísicos estudados, que embora tenham sido mostradas nas análises de correlação (Spearman) realizadas, não foram levadas em conta nas análises de regressão. Uma possível alternativa para isso é o uso de estágios sucessionais ao invés de variáveis biofísicas isoladas, como feito por Sothe et al. (2017) em áreas de Mata Atlântica usando imagens do Landsat-8/OLI e do Sentinel 2 com algoritmos de aprendizado de máquina. O estudo também mostra que métricas texturais usando as bandas do SWIR e do red-edge podem ser bastante importantes como parâmetro de entrada para modelos *Random Forest* com imagens do Sentinel 2. Quanto aos dados de controle para realizar esse tipo de comparação, Falkowski et al. (2009) mostram que é possível identificar estágios sucessionais com o uso de dados de LiDAR através do número de ecos em diferentes classes de altura. Assim, locais onde predominam ecos mais baixos, dominados por vegetação de menor porte, podem ser classificados como estágios sucessionais iniciais, enquanto a predominância de ecos mais altos mostraria a presença de árvores maiores, que é um indicador de estágios sucessionais mais avançados (DEWALT; MALIAKAL; DENSLOW, 2003).

Ainda sobre os ruídos dos modelos, podemos atribuir uma importância menor a efeitos da topografia, já que se espera que este tenha sido compensado pela correção topográfica. Além disso, a adição de imagens de relevo, declividade e sombreamento nos modelos de regressão múltipla não representaram melhorias significativas nos valores de r^2 na maioria dos casos. Efeitos da atmosfera, no entanto, podem ter causado problemas, especialmente considerando a quantidade de nuvens presentes nas imagens. Esses efeitos podem ter influência principalmente nas

comparações feitas com variáveis derivadas de dados de campo, já que para essa análise foram utilizadas duas imagens de datas distintas.

O método de correção atmosférica DOS1 (MORAN et al. 1992) foi aplicado para todas as imagens, mas apesar de este ser um método de fácil aplicação e capaz de, por exemplo, reduzir a variabilidade do NDVI em diferentes imagens (SANCHES et al., 2011) e melhorar a classificação de imagens Landsat-5/TM em séries temporais (SONG et al., 2001), é possível que ele não tenha sido suficiente para o uso conjunto das imagens utilizadas, já que as diferenças de reflectância encontradas dentro da área de estudo são pequenas e variações mínimas podem significar grandes mudanças. Como explica Chavez (1996), métodos baseados na subtração de objetos escuros, como o DOS1, são úteis principalmente na redução de efeitos aditivos causados pelo espalhamento na atmosfera, mas não compensam por efeitos multiplicativos da transmitância causados em conjunto pelo espalhamento e a absorção. Por conta disso, esse método de correção perde a eficácia em comprimentos de onda maiores, o que o torna menos apropriado para uso com bandas na faixa do NIR e SWIR (SANCHES et al., 2011). Uma possível alternativa nesse sentido é o uso do *sen2cor* (MÜLLER-WILM; DEVIGNOT; PESSIOT, 2018), um complemento disponível para o *software* SNAP da ESA que combina várias técnicas de correção atmosférica para a correção de imagens do Sentinel 2 em específico, transformando imagens de nível 1C (TOA) em 2A (BOA). A ESA espera disponibilizar imagens 2A para download no mundo todo até o fim de 2018, o que facilitará bastante a obtenção dessas imagens¹.

Um fator que pode indicar interferências da atmosfera são os resultados dos testes de modelos LiDAR treinados na imagem de 23 de dezembro de 2016 e testados na imagem de 08 de janeiro de 2016, que mostram a existência de diferenças nessas duas imagens. A principal evidência disso são os gráficos das previsões realizadas, que mostram erros que superestimam ou subestimam os valores reais das variáveis de maneira não aleatória, principalmente para modelos que envolvem uma maior quantidade de bandas. Como exemplo, para as variáveis de altura, os modelos superestimam os valores mais baixos das variáveis de controle no início e subestimam os valores no fim quando aplicados às imagens de teste. Entretanto, para

¹ Imagens nível 2A: <https://earth.esa.int/web/sentinel/user-guides/Sentinel%20-2-MSI/product-types/level-2a>

comparações feitas na mesma imagem, além dos valores de RMSE dos modelos serem praticamente reduzidos pela metade, as predições passam a se distribuir de maneira quase uniforme no entorno dos valores reais das variáveis de controle, já que efeitos atmosféricos são reduzidos dentro de uma mesma imagem. Novamente sobre o teste dos modelos em imagens diferentes das que foram usadas para treino, como foi demonstrado na Figura 45 (Seção 3.2.3), a utilização de imagens em nível 2A obtidas com o *sen2cor* teve valores de RMSE muito mais baixos se comparados com os valores obtidos com imagens corrigidas pelo método DOS1. Ademais, a distribuição dos erros ficou mais uniforme para as imagens de nível 2A, mostrando que essas imagens são mais adequadas para esse tipo de aplicação do que imagens corrigidas atmosféricamente pelo método DOS1.

Outro fator de ruído a ser considerado é o posicionamento correto das informações utilizadas no estudo. As coordenadas centrais das parcelas de campo foram registradas com um GPS com erro de até aproximadamente 10 metros pelo fato de as medidas terem sido registradas sob o dossel de florestas densas (RODRÍGUEZ-PÉREZ; ÁLVAREZ; SANZ-ABLANEDO, 2007), o que pode causar deslocamentos consideráveis, já que as parcelas têm um raio de 25 metros. Erros de posicionamento também podem afetar a comparação com os dados LiDAR. Todas as imagens utilizadas, incluindo variáveis LiDAR e imagens multiespectrais, foram alinhadas, convertidas para o mesmo sistema de projeção e tiveram seus pixels verificados visualmente para garantir o posicionamento correto das informações, mas as diferenças de resolução dificultaram a realização dessa tarefa, o que pode ter causado algumas imprecisões. Isso é confirmado ao se observar que as análises de comparação pixel a pixel tiveram um resultado bastante inferior às dos buffers comparando os valores de r^2 ajustado. Uma possível explicação para isso é de que os buffers suavizam os erros de posicionamento ao considerar uma média de vários pixels ao invés de apenas um, o que adicionalmente minimiza erros causados pelas diferenças de resolução, como explicado a seguir.

As bandas do Sentinel 2 têm resoluções espaciais de 10, 20 e 60 metros. Embora todas tenham sido convertidas para resoluções de 10 metros para possibilitar o cálculo de índices e comparações entre bandas e as imagens de LiDAR geradas, bandas de 20 e 60 metros ainda guardam menos informações que bandas de 10 metros. Essa diferença pode causar ruídos, já que um pixel de 20 ou 60 metros do

Sentinel 2 acaba correspondendo a vários pixels das imagens LiDAR nos modelos criados. Quando se utiliza buffers ao invés de pixels individuais, no entanto, todas as bandas do Sentinel 2 e imagens LiDAR passam a ter a mesma quantidade de informação, eliminando esse efeito (Figura 48). Nesse sentido, a comparação por buffers seria equivalente a comparar imagens com a mesma resolução.

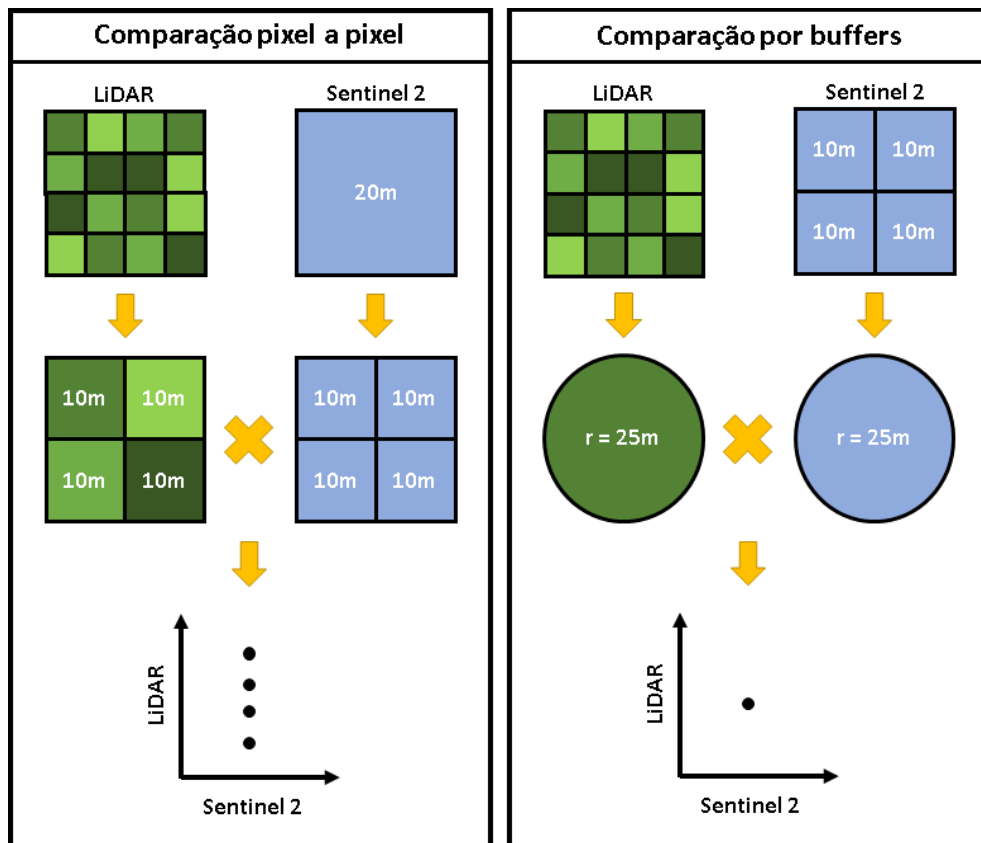


Figura 48 – Processo de construção de gráficos da comparação pixel a pixel (esquerda) e da comparação por buffers (direita) evidenciando uma das fontes de ruído presente no primeiro método.

No caso dos dados derivados de LiDAR, podemos considerar a variação temporal como fonte de ruído, já que as imagens do Sentinel 2 de 2015 (11/08/2015) e do Landsat-8/OLI de 2014 (06/10/2014) utilizadas para comparação com o LiDAR tiveram resultados bastante inferiores aos da imagem de 2016 (08/01/2016). Isso provavelmente se deve ao fato dos anos de 2014 e 2015 terem sido caracterizados por um período atípico de seca na região (MARENGO et al. 2015), além de a imagem de 2016 mencionada ser correspondente ao período de verão, mais chuvoso, enquanto que as imagens de 2015 e 2014 correspondem ao período de inverno, tipicamente mais seco na área de estudo.

Embora os dados de LiDAR usados no estudo tenham sido coletados nesse período de seca atípica mencionado e em uma época do ano naturalmente mais seca

(início da primavera), por corresponderem diretamente a características estruturais da vegetação, provavelmente não tenham sofrido grandes mudanças com a seca. A resposta espectral das imagens usadas, por outro lado, é dependente de características que influenciam diretamente a resposta espectral da vegetação, como concentração de pigmentos fotossintetizantes, quantidade de água e estrutura interna das folhas, que podem sofrer alterações com o estresse hídrico. De fato, Carter e Knapp (2001) demonstram que uma tendência geral de diferentes plantas submetidas a diversos tipos de estresse, incluindo estresse hídrico, é a redução da concentração da clorofila nas folhas, o que causa mudanças na resposta espectral da vegetação principalmente em comprimentos de onda próximos a 700 nm. Essa faixa corresponde justamente à banda 5 do Sentinel 2 (centrada 705 nm), que como pode ser visto na seção 3, é responsável pelos melhores resultados encontrados neste estudo. Isso indica que o estresse hídrico pode ter sido um fator fundamental no pior desempenho dos modelos gerados com a imagem do Sentinel 2 de 2015 e a imagem do Landsat-8/OLI. Assim, mesmo que a imagem de 2016 esteja mais distante temporalmente dos dados LiDAR, o fato de ela corresponder a um período mais chuvoso pode explicar seu maior sucesso.

Outro fator importante a ser considerado são as variações de características da vegetação não medidas no estudo. É sabido que características como estrutura interna e concentração de clorofila nas folhas desempenham papéis importantes na resposta espectral da vegetação (GATES et al., 1965, KNIPLING, 1970, SEAGER et al., 2005). Informações desse tipo não podem ser obtidas observando apenas parâmetros biofísicos como altura, DAP e cobertura vegetal, e podem estar relacionadas com as variações espaciais observadas no mapa de erro gerado entre a banda 5 e a variável Altura máxima 5 (Seção 3.2.2 - Figura 37). As regiões localizadas com maior erro podem corresponder a locais de vegetação com comportamento de reflectância distinto dos demais, já que fatores como disponibilidade de água (GAO, 1996, CARTER; KNAPP, 2001) e diferenças na composição de espécies podem causar variações significativas na resposta espectral de uma floresta (CARLSON et al., 2007, FERREIRA et al., 2016).

Uma das variáveis não medidas diretamente neste estudo que tem grande importância é o LAI (Índice de Área Foliar – *Leaf Area Index*). Essa variável apresenta boa relação com a altura da vegetação (ZHAO; POPESCO, 2009), o que

provavelmente se deve à maior quantidade de camadas de folhas associados ao aumento da altura dos indivíduos, o que pode ser observado analisando o processo de sucessão de uma floresta (CARVALHO; NASCIMENTO; BRAGA, 2007, CHAZDON, 2014). A relação entre altura e LAI pode ser uma boa explicação para o maior sucesso das estimativas de variáveis de altura se comparadas às outras variáveis de estrutura medidas, chegando a um r^2 ajustado próximo a 0,6 nos melhores casos. Delegido et al. (2011) por exemplo, mostram que LAI pode ser estimado com bandas do Sentinel 2 simuladas pela diferença normalizada de B4 (665 nm) e B5 (705 nm), embora no presente estudo não tenha apresentado boas relações com nenhum dos parâmetros biofísicos estudados (NDI45). Já Majasalmi e Rautiainen (2016) encontraram melhores relações ($r^2 = 0,93$) com LAI usando a razão entre B4 (665 nm) e B8 (842 nm) (RVI), também para imagens simuladas do Sentinel 2.

Para comparações da literatura com os parâmetros biofísicos medidos neste estudo, podem ser citados os resultados de Accioly et al. (2002), que mostra relações entre RVI e a altura em áreas de cerrado, cerradão e mata húmida medida em campo e imagens do Landsat-5/TM ($r^2 = 0,36$). No presente estudo o RVI foi um dos índices que apresentou melhores relações com a altura medida em campo, embora com um valor de r^2 inferior ($r^2 = 0,16$). Já para dados derivados de LiDAR, os resultados de Pascual et al. (2010) mostram relações entre a altura de dosséis derivadas de LiDAR e imagens do Landsat-7/ETM+, com modelos chegando a um r^2 próximo a 0,7 para índices de vegetação como o NDVI, mas com valores de RMSE similares aos obtidos no presente estudo, variando entre 1,9 e 2,3 m. Staben, Lucieer e Scarth (2018) também mostram resultados parecidos com imagens Landsat-5/TM e Landsat-7/ETM+, com valores de r^2 de 0,53 e RMSE de 2,8 m nos melhores casos. No presente estudo, os valores de r^2 encontrados em regressões múltiplas entre as bandas do Sentinel 2 e variáveis LiDAR de altura foram bastante parecidos com o encontrado no primeiro trabalho, mas os valores de RMSE foram mais parecidos com os de Pascual et al. (2010), próximos de 2 m, como mostrado na seção 3.2.3.

É importante destacar que além de sensores orbitais diferentes, ambos os trabalhos estudam outras fitofisionomias, e as diferenças de resultados encontradas em relação a este estudo podem ser principalmente derivadas disso. Quanto às outras variáveis avaliadas, a Profundidade, Densidade e Cobertura 10 estimadas com os dados de LiDAR apresentaram bons resultados em regressões com múltiplas bandas.

No entanto, as variáveis mencionadas devem ser interpretadas com cuidado, já que elas incluem a altura em seus cálculos e apresentam correlações fortes com esta (Seção 3.2.2 - Figura 33). Assim, é bem possível que elas estejam mais relacionadas à altura das árvores do que à densidade de folhas nos dosséis, que era o propósito inicial dessas variáveis.

Já a variável BSAST foi uma das que apresentou melhores resultados com os dados de campo, principalmente com as bandas do fim do *red-edge*, NIR e com o índice EVI. A biomassa acima do solo não diz respeito diretamente à densidade de folhas, principalmente no caso de BSAST, que é calculada a partir da altura e do DAP, mas variáveis desse tipo podem ter relação e até mesmo serem estimadas por LAI (MADUGUNDU; NIZALAPUR; JHA, 2008). Além disso, diversos estudos mostram relações entre biomassa e índices de vegetação com resultados variados (ARAUJO; SANTOS; SHIMABUKURO, 2000, GONZÁLEZ-ALONSO et al., 2006, GOSWAMI et al., 2015, MADUGUNDU; NIZALAPUR; JHA, 2008). Mais especificamente para imagens do Sentinel 2 (simuladas a partir de dados hiperespectrais), Majasalmi e Rautiainen (2016) identificaram relações mais fortes entre AGB e razões entre bandas do NIR e SWIR, especialmente B11 (1610 nm) e B10 (1375 nm) ($r^2 \cong 0,23$). Os resultados encontrados no presente estudo são similares, com razões entre bandas NIR e B10 (1375 nm) atingindo valores de r^2 próximos a 0,2 em regressões OLS com BSAST (Apêndice B, Figura B5). Resultados parecidos também podem ser observados para a variável AB, que tem forte correlação com BSAST (Figura 17, seção 3.1.1), já que ambas as variáveis são calculadas a partir do DAP.

Para a porcentagem de cobertura medida em campo, é importante destacar que esta não apresentou boas relações com bandas separadas, razões entre bandas e índices de vegetação, mas foi uma das melhores variáveis de campo nos modelos de regressão múltipla ($r^2 = 0,48$). Isso indica que a cobertura não está relacionada a uma faixa espectral em específico, mas sim a um conjunto de respostas em diferentes regiões do espectro. Esses resultados, no entanto, devem ser interpretados com cautela. Como a quantidade de amostras de campo é pequena ($n = 34$) e essa variável apresentou uma das menores variações, estes resultados podem ser um efeito do acaso. Além disso, a variável derivada de LiDAR correspondente (Cobertura 1) não teve resultados tão bons ($r^2 = 0,1$) com uma amostragem muito maior ($n = 338$).

Dentre as bandas analisadas, as que tiveram melhores resultados quando relacionadas às variáveis de controle foram B3 (560 nm) e B5 (705 nm), principalmente quando correlacionadas à altura e a medidas de densidade e cobertura da vegetação medidas por LiDAR. Os resultados obtidos em laboratório corroboram esses resultados. Ao observar a diferença entre a reflectância para uma ou 4 camadas de folhas, as maiores diferenças negativas ocorreram nas faixas do verde e início do red-edge, as mesmas faixas de B3 (560 nm) e B5 (705 nm) do Sentinel 2. Uma possível explicação para isso pode estar na absorção das folhas, que ocorre de forma mais forte na região do visível, mas é menor na faixa do verde e no início do red-edge (SEAGUER et al., 2005). Essa absorção mais fraca nessas duas regiões permite que elas respondam à quantidade de vegetação sem serem totalmente absorvidas nas primeiras camadas de folhas de um dossel, como pode acontecer para regiões do espectro onde ocorre maior absorção, gerando problemas de saturação que são comuns principalmente em índices como o NDVI (HUETE et al., 1997). Assim, as bandas localizadas na região do vermelho e do azul atingem um valor mínimo de reflectância rapidamente com o aumento da vegetação, enquanto bandas correspondentes ao verde ou ao início do red-edge conseguem responder a uma maior quantidade de vegetação sem sofrer tanto esse efeito. Além do mais, como dito anteriormente, variações na clorofila afetam de maneira mais forte justamente a faixa próxima a 700 nm em que B5 está localizada (CARTER; KNAPP, 2001).

Quanto às bandas do infravermelho e do final do red-edge, nenhuma obteve bons resultados em relações com a altura, apesar de ser na faixa NIR que encontram as maiores variações positivas com o aumento do número de folhas segundo a análise de reflectância feita em laboratório. Uma possível explicação para isso pode estar na diferença dos valores de reflectância observados em dosséis e em folhas isoladas. Apesar de existir um aumento na reflectância causado pela refração das múltiplas camadas de folhas, esse efeito é compensado pelo sombreamento mútuo entre as folhas de um dossel, diminuindo a reflectância se comparado com uma folha isolada em ambos o espectro visível e infravermelho (PONZONI; SHIMABUKURO; KUPLICH, 2012). Além disso, a reflectância de um dossel obtida por plataformas orbitais está ligada a outros fatores, como efeitos atmosféricos, relevo, solo, presença de galhos, flores, frutos, posicionamento e ângulo das folhas e efeitos direcionais. (BREUNIG et al., 2012, PONZONI; SHIMABUKURO; KUPLICH, 2012). Esses efeitos não puderam

ser considerados na análise com o espectrorradiômetro, já que ela foi feita em um ambiente controlado e usando folhas alinhadas de um mesmo indivíduo com o propósito de evitar esse tipo de variação.

Apesar de não terem mostrado bons resultados na previsão de variáveis de altura quando consideradas individualmente, as bandas na faixa do infravermelho são essenciais para gerar muitos dos índices de vegetação existentes (seção 2.3). Bandas do NIR e SWIR se mostraram úteis em conjunto com bandas do visível e *red-edge* em regressões múltiplas OLS e RF, melhorando consideravelmente as previsões de variáveis de altura e cobertura da vegetação. Além disso, como dito anteriormente, as melhores relações com BSAST aconteceram com as bandas do NIR, o que reforça a importância dessa região do espectro. Já bandas do SWIR tiveram resultados intermediários individualmente. No experimento realizado em laboratório é possível notar que nenhuma das bandas do Sentinel 2 na faixa do SWIR está localizada nos picos de variação de reflectância entre diferentes camadas de folhas, e isso pode estar relacionado ao desempenho relativamente inferior dessas bandas. No entanto, assim como para o NIR, a faixa do SWIR também foi muito importante na melhora da relação entre as variáveis e do poder preditivo de modelos de regressão múltipla, embora também pareça ser mais afetada por efeitos atmosféricos do que comprimentos de ondas menores (VIS e NIR).

Quanto aos índices de vegetação estudados, nota-se que estes não apresentaram resultados melhores do que para as bandas analisadas individualmente em comparação com ambas as medidas de campo e com dados LiDAR. Os índices que se destacaram também foram diferentes em cada caso. Para os dados de campo o NDVI e RVI foram os melhores em comparação com a altura mediana e IRECI para biomassa. Já para os dados LiDAR, o que se destacou mais foi o S2REP para as variáveis de altura e para profundidade. Esses resultados mostram a diferença entre os dois tipos de dados de controle. Para as variáveis de altura, por exemplo, é importante destacar que os dados LiDAR correspondem à altura dos dosséis, enquanto que os dados de campo são referentes à altura dos indivíduos amostrados, muitos deles com alturas abaixo do dossel, mesmo considerando o uso das variáveis de campo calculadas apenas com indivíduos com mais de cinco metros. Quanto às razões entre bandas, algumas apresentaram resultados melhores que os dos índices de vegetação, e em alguns casos foram superiores aos resultados de bandas

individuais, mas no geral tiveram resultados parecidos. Isso mostra que o uso de índices de vegetação nem sempre pode ser considerada a melhor alternativa para a análise de parâmetros biofísicos da vegetação, o que já foi observado anteriormente na literatura (LAWRENCE; RIPPLE, 1998).

Quanto aos modelos gerados, B5 (705 nm) parece ser a banda que melhor explica a maioria dos parâmetros biofísicos medidos, e sua mistura com outras bandas através de modelos de regressão múltipla, tanto OLS quanto RF, se mostrou uma melhor alternativa se comparado ao uso dos índices de vegetação estudados. Apesar disso, mesmo que modelos de regressão múltipla considerando várias bandas tenham reduzido bastante os valores de RMSE dos modelos testados em locais diferentes da mesma imagem em que foram treinados, um grande número de bandas em um modelo parece estar associado a maiores erros quando este é aplicado em outras imagens, como é demonstrado no fim da seção 3.2.3 (Figura 44). Nesse sentido, modelos com poucas bandas apresentam maior segurança para aplicações em imagens diferentes segundo os resultados obtidos neste estudo. Como dito anteriormente, este efeito pode estar associado a interferências da atmosfera, que tem maior influência quando se consideram múltiplas imagens. No entanto, com a correção atmosférica adequada, uma quantidade maior de bandas no modelo parece não causar nenhum tipo de problema.

Ainda, as melhores relações encontradas parecem ser lineares, já que modelos OLS (Mínimos Quadrados Ordinários - *Ordinary Least Squares*) e RF (Floresta Aleatória - *Random Forest*) tiveram resultados bastante similares. Isso pode ser explicado principalmente por conta da baixa variação das variáveis observadas no estudo. Reforça-se que todos os dados são correspondentes a áreas de mata atlântica preservada, e por isso podem não variar o suficiente para serem influenciados por efeitos de saturação que poderiam tornar essas relações não-lineares (HUETE; LIU; LEEUWEN, 1997). Dito isto, para as previsões em áreas de mata atlântica deste trabalho, modelos lineares parecem representar bem as relações encontradas. Estas relações, no entanto, podem mudar se uma escala maior de variação dos dados for considerada. Um exemplo disso seria a inclusão de dados de fitofisionomias distintas ou de áreas degradadas em um mesmo modelo de previsão. Neste contexto é bem possível que a maior variação dos dados permita a observação de relações não-lineares entre as imagens e parâmetros biofísicos.

5 CONCLUSÃO E RECOMENDAÇÕES

Em suma, os resultados encontrados mostram que imagens do Sentinel 2, principalmente a banda 5 (705 nm), podem ser utilizadas em estimativas de altura e cobertura vegetal com precisão razoável em áreas de Mata Atlântica. O uso de várias bandas em regressões múltiplas aumenta o poder preditivo dos modelos e permite estimativas de cobertura do dossel, mas o aumento do número de variáveis pode não compensar dependendo do objetivo. Regressões múltiplas com bandas tiveram resultados melhores que índices de vegetação, que podem não ser a melhor alternativa quando o objetivo é a comparação de áreas de floresta. É importante lembrar que, embora tenham sido validados com sucesso em áreas de teste na mesma imagem, os modelos de regressão, principalmente os que envolvem muitas bandas, não foram capazes de estimar com confiabilidade parâmetros biofísicos em imagens diferentes com uma correção atmosférica simples. Assim, modelos desse tipo só devem ser utilizados em outras imagens se uma correção atmosférica adequada for aplicada. Além disso, as relações entre imagens e parâmetros biofísicos discutidas são todas referentes à vegetação de Mata Atlântica no estado de São Paulo (Brasil), e podem diferir para outros tipos de vegetação.

Quanto às aplicações, a conclusão é de que imagens do Sentinel 2 podem ser úteis principalmente para detectar diferenças entre áreas para as aplicações que não necessitam de uma precisão tão grande. Vale destacar que o uso das imagens é recomendado para realizar estimativas de altura, cobertura e densidade da vegetação, mas não para as outras variáveis medidas no estudo, pois as variações podem ser muito altas e os resultados gerados teriam baixa confiabilidade. Uma boa alternativa pode ser usar pontos de controle em que se conhece a estrutura da vegetação e fazer estimativas dos arredores com as imagens. Para estudos em que uma precisão maior é necessária, outras metodologias devem ser consideradas, como o uso de dados LiDAR, imagens de alta resolução espacial e espectral ou coletas de campo feitas em grandes áreas. Para tal, a disponibilidade de verba na realização de pesquisas é essencial, já que essas metodologias são bem mais onerosas que o uso de imagens do Sentinel 2 como estimativa.

Assim, para cada pergunta proposta conclui-se que:

1. É possível usar imagens do Sentinel 2 para estimar altura e cobertura medidas em campo e por LiDAR com precisão razoável, mas os modelos demonstrados possuem algumas limitações de precisão.
2. A correção topográfica não pareceu influenciar muito as estimativas dos resultados obtidos em campo, mas melhorou significativamente as estimativas dos dados LiDAR e no geral deve ser aplicada.
3. As melhores estimativas foram as de altura, cobertura e densidade da vegetação. De maneira geral, as estimativas de dados LiDAR foram mais bem sucedidas que as de dados medidos em campo.
4. As bandas com melhores resultados foram B5 (705 nm) e B3 (560 nm). Os índices de vegetação com melhor resultado foram o RVI e NDVI para os dados de campo e S2REP para dados LiDAR. Modelos de regressão múltipla OLS e RF envolvendo mais de uma banda apresentaram os melhores resultados.

Para estudos futuros, recomenda-se considerar:

- Testar a utilização das imagens do Sentinel 2 para estimar parâmetros biofísicos em diferentes fitofisionomias.
- Comparar imagens do Sentinel 2 com parâmetros biofísicos que influenciem diretamente a resposta espectral da vegetação, como LAI e biomassa verde, e utilizar outras variáveis para comparação, como métricas texturais.
- Comparar imagens do Sentinel 2 com informações que envolvam mais de um parâmetro biofísico, como os estágios sucessionais de florestas.
- Testar o uso de imagens do Sentinel 2 na classificação de diferentes variáveis categóricas, como tipos de vegetação, cultivos e outros usos da terra.
- Uso de imagens 2A (sen2cor) ou outros métodos de correção atmosférica mais avançados que o DOS1 para as imagens do Sentinel 2.
- Realizar amostragens de campo em áreas maiores que possibilitem a criação de modelos mais confiáveis para esse tipo de variável.
- Estudar mais a fundo a reflectância da vegetação na faixa de 705 nm, tendo em vista o maior sucesso dos modelos com a banda 5 do Sentinel 2 e os resultados do experimento de espectros de reflectância feito em laboratório.
- Explorar as relações de parâmetros biofísicos da vegetação com dados de outros programas, como o Sentinel 1, o Rapideye e o PlanetScope.

REFERÊNCIAS BIBLIOGRÁFICAS

ABREU, K. M. P.; COUTINHO, L. M. Sensoriamento remoto aplicado ao estudo da vegetação com ênfase em índice de vegetação e métricas da paisagem. **Revista Vértices**, v. 16, n. 2010, p. 173–198, 2014. Disponível em: <http://www.gnresearch.org/doi/10.5935/1809-2667.20140012>

ACCIOLY, L. J. O. et al. Relações empíricas entre a estrutura da vegetação e dados do sensor TM/LANDSAT. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 6, n. 3, p. 492-498, 2002. Disponível em: <http://www.scielo.br/pdf/%0D/rbeaa/v6n3/v6n3a19.pdf>

ALONZO, M. et al. Mapping urban forest leaf area index with airborne lidar using penetration metrics and allometry. **Remote Sensing of Environment**, v. 162, p. 141-153, 2015. Disponível em: <http://www.sciencedirect.com/science/article/pii/S003442571500084X>

ALVES, L. F.; METZGER, J. P. A regeneração florestal em áreas de floresta secundária na Reserva Florestal do Morro Grande, Cotia, SP. **Biota Neotropica**, v. 6, n. 2, 2006. Disponível em: <http://www.redalyc.org/html/1991/199114291004/>

ALVES, L. F. et al. Forest structure and live aboveground biomass variation along an elevational gradient of tropical Atlantic moist forest (Brazil). **Forest ecology and management**, v. 260, n. 5, p. 679-691, 2010. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0378112710002926>

ARAUJO, L. S.; SANTOS, J. R.; SHIMABUKURO, Y. E. Relationship between SAVI and biomass data of forest and Savanna Contact Zone in the Brazilian Amazonia. **International Archives of Photogrammetry and Remote Sensing**, v. 33, n. B7/1; PART 7, p. 77-81, 2000. Disponível em: http://mtc-m12.sid.inpe.br/col/sid.inpe.br/iris@1912/2005/07.21.01.30.58/doc/2000_araujo.pdf

BALBINOT, R. et al. O papel da floresta no ciclo hidrológico em bacias hidrográficas. **AMBIÊNCIA**, v. 4, n. 1, p. 131-149, 2009. Disponível em: <http://revistas.unicentro.br/index.php/ambiencia/article/view/294>

BARANOSKI, G. V. G.; ROKNE, J. G. A practical approach for estimating the red edge position of plant leaf reflectance. **International Journal of Remote Sensing**, v. 26, n. 3, p. 503-521, 2005. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/01431160512331314029>

BOEHM, H.; LIESENBERG, V.; FRANK, J. Relating tree height variations to peat dome slope in Central Kalimantan, Indonesia using small-footprint airborne LiDAR data. In: SILVILASER 2010, 10., 2010, Freiburg, Anais... p. 216-228. Disponível em: <http://www.isprs.org/proceedings/2010/SilviLaser2010.pdf>

BREIMAN, L. Out-of-bag-estimation. **Berkeley**, 1996, p.13. Disponível em: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.45.3712&rep=rep1&type=pdf>

BREIMAN, Leo. Random forests. **Machine learning**, v. 45, n. 1, p. 5-32, 2001. Disponível em: <https://link.springer.com/article/10.1023/A:1010933404324>

BREUNIG, F. M. et al. Avaliação dos efeitos direcionais em fragmentos florestais: estudo de caso na região sul do Brasil. 2012, p. 466-459. Disponível em: https://www.researchgate.net/publication/235765817_AVALIACAO_DOS_EFEITOS_DIRECIONAIS_EM_FRAGMENTOS_FLORESTAIS_ESTUDO_DE_CASO_NA_REGIAO_SUL_DO_BRASIL

CANHAM, C. D. et al. Causes and consequences of resource heterogeneity in forests: interspecific variation in light transmission by canopy trees. **Canadian Journal of Forest Research**, v. 24, n. 2, p. 337-349, 1994. Disponível em: <http://www.nrcresearchpress.com/doi/pdf/10.1139/x94-046>

CARLSON, K. M. et al. Hyperspectral remote sensing of canopy biodiversity in Hawaiian lowland rainforests. **Ecosystems**, v. 10, n. 4, p. 536-549, 2007. Disponível em: <https://link.springer.com/article/10.1007/s10021-007-9041-z>

CARTER, G. A.; KNAPP, A. K. Leaf optical properties in higher plants: linking spectral characteristics to stress and chlorophyll concentration. **American journal of botany**, v. 88, n. 4, p. 677-684, 2001. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.2307/2657068>

CARVALHO F. A.; NASCIMENTO, M. T.; BRAGA, J. M. A. Estrutura e composição florística do estrato arbóreo de um remanescente de Mata Atlântica submontana no município de Rio Bonito, RJ, Brasil (Mata Rio Vermelho). **Revista Árvore**, v. 31, n. 4, 2007. Disponível em: <http://www.redalyc.org/html/488/48831417/>

CHADWICK, J. Integrated LiDAR and IKONOS multispectral imagery for mapping mangrove distribution and physical properties. **International journal of remote sensing**, v. 32, n. 21, p. 6765-6781, 2011. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/01431161.2010.512944>

CHAVE, J. et al. Regional and phylogenetic variation of wood density across 2456 neotropical tree species. **Ecological applications**, v. 16, n. 6, p. 2356-2367, 2006. Disponível em: <https://esajournals.onlinelibrary.wiley.com/doi/full/10.1890/1051-0761%282006%29016%5B2356%3ARAPVOW%5D2.0.CO%3B2>

CHAVE, J. et al. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. **Oecologia**, v. 145, n. 1, p. 87-99, 2005. Disponível em: <https://link.springer.com/article/10.1007/s00442-005-0100-x>

CHAVEZ, P. S. et al. Image-based atmospheric corrections-revisited and improved. **Photogrammetric engineering and remote sensing**, v. 62, n. 9, p. 1025-1035, 1996. Disponível em: <https://www.unc.edu/courses/2008spring/geog/577/001/www/Chavez96-PERS.pdf>

CHAZDON, R. L. **Second growth: the promise of tropical forest regeneration in an age of deforestation**. University of Chicago Press, 2014.

CHEN, Q. et al. Estimating basal area and stem volume for individual trees from lidar data. **Photogrammetric Engineering & Remote Sensing**, v. 73, n. 12, p. 1355-1365, 2007. Disponível em: <http://www.ingentaconnect.com/content/asprs/pers/2007/00000073/00000012/art00004>.

CONGEDO, L. Semi-automatic classification plugin documentation. p. 198, 2018. Disponível em: <http://dx.doi.org/10.13140/RG.2.2.29474.02242/1>

CONRAD, O. et al. System for automated geoscientific analyses (SAGA) v. 2.1.4. **Geoscientific Model Development**, v. 8, n. 7, p. 1991-2007, 2015. Disponível em: <https://www.geosci-model-dev.net/8/1991/2015/gmd-8-1991-2015.html>

COOK, J. G. et al. Spherical densimeters produce biased estimates of forest canopy cover. **Wildlife Society Bulletin**, p. 711-717, 1995. Disponível em: https://www.jstor.org/stable/3783003?seq=1#page_scan_tab_contents

DASSOT, M.; CONSTANT, T.; FOURNIER, M. The use of terrestrial LiDAR technology in forest science: application fields, benefits and challenges. **Annals of Forest Science**, v. 68, n. 5, p. 959-974, 2011. Disponível em: <https://link.springer.com/article/10.1007/s13595-011-0102-2>

DELEGIDO, J. et al. Evaluation of Sentinel 2 red-edge bands for empirical estimation of green LAI and chlorophyll content. **Sensors**, v. 11, n. 7, p. 7063-7081, 2011. Disponível em: <http://www.mdpi.com/1424-8220/11/7/7063/htm>

DEWALT, S. J.; MALIAKAL, S. K.; DENSLOW, J. S. Changes in vegetation structure and composition along a tropical forest chronosequence: implications for wildlife. **Forest Ecology and Management**, v. 182, n. 1-3, p. 139-151, 2003. Disponível em: <https://www.sciencedirect.com/science/article/pii/S037811270300029X>

EMPLASA. Empresa Paulista de Planejamento Metropolitano S.A. "Modelo Digital de superfície (DSM) do estado de São Paulo." São Paulo: EMLASA, 2010.

EMBRAPA, *Paisagens Sustentáveis*, 2016. Disponível em: <https://www.embrapa.br/busca-de-solucoes-tecnicas/-/produto-servico/3862/paisagens-sustentaveis>

ENCINA, César Claudio Cáceres. **Análise da estrutura vegetal de fragmentos de cerrado através de sensoriamento remoto**. 2015. Dissertação de Mestrado. Disponível em: <http://repositorio.ufms.br:8080/jspui/handle/123456789/2421>

ESA, Sentinel 2 MSI Introduction, 2015. Disponível em: <https://earth.esa.int/web/sentinel/user-guides/Sentinel%202-msi>

ESA, *Spatial resolution*. Disponível em <https://earth.esa.int/web/sentinel/user-guides/Sentinel%202-msi/resolutions/spatial>

FALKOWSKI, M. J. et al. Characterizing forest succession with lidar data: An evaluation for the Inland Northwest, USA. **Remote Sensing of Environment**, v. 113, n. 5, p. 946-956, 2009. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0034425709000145>

FERREIRA, M. P. et al. Analyzing the spectral variability of tropical tree species using hyperspectral feature selection and leaf optical modeling. **Journal of Applied Remote Sensing**, v. 7, n. 1, p. 073502-073502, 2013. Disponível em: <https://www.spiedigitallibrary.org/journals/Journal-of-Applied-Remote-Sensing/volume-7/issue-1/073502/Analyzing-the-spectral-variability-of-tropical-tree-species-using-hyperspectral/10.1117/1.JRS.7.073502.full?SSO=1>

FOODY, G. M.; CUTLER, M. E. J. Mapping the species richness and composition of tropical forests from remotely sensed data with neural networks. **Ecological Modelling**, v. 195, n. 1-2, p. 37-42, 2006. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0304380005005740>

FRAMPTON, W. J. et al. Evaluating the capabilities of Sentinel 2 for quantitative estimation of biophysical variables in vegetation. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 82, p. 83-92, 2013. International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). Disponível em: <http://dx.doi.org/10.1016/j.isprsjprs.2013.04.007>

GAIDA, W. et al. Variações da reflectância e dos índices de vegetação em função dos parâmetros da modelagem topográfica no Parque Estadual do Turvo, Rio Grande do Sul, Brasil. **Investigaciones Geográficas, Boletín del Instituto de Geografía**, v. 2016, n. 91, p. 105-123, 2016. Disponível em: http://www.scielo.org.mx/scielo.php?pid=S0188-46112016000300105&script=sci_arttext

GAO, B. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. **Remote sensing of environment**, v. 58, n. 3, p. 257-266, 1996. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0034425796000673>

GATES, D. M. et al. Spectral properties of plants. **Applied optics**, v. 4, n. 1, p. 11-20, 1965. Disponível em: <https://www.osapublishing.org/ao/abstract.cfm?uri=ao-4-1-11>

GINCIENE, B. R.; BITENCOURT, M. D. Utilização do EVI (Enhanced Vegetation Index) para maior sensibilidade na detecção de mudanças temporais em fragmentos de floresta estacional semidecidual. SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, **Anais...** Curitiba, PR, p. 1620-1627, 2011. Disponível em: https://www.researchgate.net/profile/Bruno_Ginciene/publication/267429884_Utilizac_ao_do_EVI_Enhanced_Vegetation_Index_para_maior_sensibilidade_na_deteccao_d_e_mudancas_temporais_em_fragmentos_de_floresta_estacional_semidecidual/links/562e0d6308aef25a244333f1.pdf

GETZIN, S. et al. Using airborne LiDAR to assess spatial heterogeneity in forest structure on Mount Kilimanjaro. **Landscape Ecology**, v. 32, n. 9, p. 1881-1894, 2017. Disponível em: <https://link.springer.com/article/10.1007/s10980-017-0550-7>

GOETZ, S. J. et al. IKONOS imagery for resource management: Tree cover, impervious surfaces, and riparian buffer analyses in the mid-Atlantic region. **Remote sensing of environment**, v. 88, n. 1-2, p. 195-208, 2003. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0034425703002414>

GONÇALVES, W. et al. Classificação de estratos florestais utilizando redes neurais artificiais e dados de sensoriamento remoto/Classification of forest types using artificial neural networks and remote sensing data. *Revista Ambiente & Água*, v. 11, n. 3, p. 612, 2016. Disponível em: http://www.scielo.br/scielo.php?pid=S1980-993X2016000300612&script=sci_arttext

GONÇALVES, F. G.; SANTOS, JR dos. Composição florística e estrutura de uma unidade de manejo florestal sustentável na Floresta Nacional do Tapajós, Pará. **Acta Amazonica**, v. 38, n. 2, p. 229-244, 2008. Disponível em: <http://www.scielo.br/pdf/aa/v38n2/v38n2a06>

GONG, P.; PU, R.; YU, B. Conifer species recognition: an exploratory analysis of in situ hyperspectral data. **Remote Sensing of Environment**, v. 62, n. 2, p. 189-200, 1997. Disponível em: <https://nature.berkeley.edu/~penggong/PDFpapers/GongRSE1997.pdf>

GONZÁLEZ-ALONSO, F. et al. Forest biomass estimation through NDVI composites. The role of remotely sensed data to assess Spanish forests as carbon sinks. **International Journal of Remote Sensing**, v. 27, n. 24, p. 5409-5415, 2006. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/01431160600830748>

GOSWAMI, S. et al. Relationships of NDVI, Biomass, and Leaf Area Index (LAI) for six key plant species in Barrow, Alaska. **PeerJ PrePrints**, 2015. Disponível em: <https://peerj.com/preprints/913/>

GRAMIN, **eTrex manual do utilizador Modelos: 10, 20, 20x, 30, 30x**. 2015. Taiwan. Disponível em: https://static.garmin.com/pumac/eTrex_10_20x_30x_OM_PT.pdf

GWENZI, D. et al. Predictions of tropical forest biomass and biomass growth based on stand height or canopy area are improved by Landsat-scale phenology across Puerto Rico and the US Virgin Islands. **Remote Sensing**, v. 9, n. 2, p. 123, 2017. Disponível em: <http://www.mdpi.com/2072-4292/9/2/123/htm>

GUILHERME, F. A. G.; MORELLATO, L. P. C.; ASSIS, M. A. Horizontal and vertical tree community structure in a lowland Atlantic Rain Forest, Southeastern Brazil. **Braz. J. Bot.**, v. 27, n. 4, p. 725-737, 2004. Disponível em: http://www.scielo.br/scielo.php?pid=S0100-84042004000400012&script=sci_arttext&tlng=es

GUO, H. et al. An ensemble learning for predicting breakdown field strength of polyimide nanocomposite films. **Journal of Nanomaterials**, v. 2015, p. 7, 2015. Disponível em: <https://dl.acm.org/citation.cfm?id=2930533>

HARPER, K. A. et al. Edge influence on forest structure and composition in fragmented landscapes. **Conservation Biology**, v. 19, n. 3, p. 768-782, 2005. Disponível em: <http://onlinelibrary.wiley.com/doi/10.1111/j.1523-1739.2005.00045.x/full>

HOFF, T, E. et al. Reporting of irradiance modeling relative prediction errors. *Progress in Photovoltaics: Research and Applications*, v. 21, n. 7, p. 1514-1519, 2013. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1002/pip.2225>

HUETE, A. R. A soil-adjusted vegetation index (SAVI). **Remote sensing of environment**, v. 25, n. 3, p. 295-309, 1988. Disponível em: <https://www.sciencedirect.com/science/article/pii/003442578890106X>

HUETE, A. R. et al. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. **Remote sensing of environment**, v. 59, n. 3, p. 440-451, 1997. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0034425796001125>

HUETE, A. R. et al. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. **Remote sensing of environment**, v. 83, n. 1-2, p. 195-213, 2002. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0034425702000962>

HUETE, A. R.; LIU, H.; LEEUWEN., W. J. D. The use of vegetation indices in forested regions: issues of linearity and saturation. In: **Geoscience and Remote Sensing, 1997. IGARSS'97. Remote Sensing-A Scientific Vision for Sustainable Development., 1997 IEEE International**. IEEE, 1997. p. 1966-1968. Disponível em: <https://ieeexplore.ieee.org/abstract/document/609169>

JORDAN, C. F. Derivation of leaf-area index from quality of light on the forest floor. **Ecology**, v. 50, n. 4, p. 663-666, 1969. Disponível em: <https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.2307/1936256>

KE, Y.; QUACKENBUSH, L. J. A comparison of three methods for automatic tree crown detection and delineation from high spatial resolution imagery. **International Journal of Remote Sensing**, v. 32, n. 13, p. 3625-3647, 2011. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/01431161003762355>

KHOSRAVIPOUR, A. et al. Generating pit-free canopy height models from airborne lidar. **Phot. Eng. & Rem. Sens.**, v. 80, n. 9, p. 863-872, 2014. Disponível em: <http://www.ingentaconnect.com/content/asprs/pers/2014/00000080/00000009/art00003>

KIMES, D. S. et al. Temporal relationships between spectral response and agronomic variables of a corn canopy. **Rem. Sens. Environ.**, v. 11, p. 401-411, 1981. Disponível em: <https://www.sciencedirect.com/science/article/pii/0034425781900377>

KNIPLING, E. B. Physical and physiological basis for the reflectance of visible and near-infrared radiation from vegetation. **Rem. Sens. Environ.**, v. 1, n. 3, p. 155-159, 1970. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0034425770800219>

LAW, K. H.; NICHOL, J. Topographic correction for differential illumination effects on IKONOS satellite imagery. **Int. Arch. Photogramm. Remote Sens. Spat. Inform. Sci.**, v. 35, p. 641-646, 2004. Disponível em: <http://www.isprs.org/proceedings/XXXV/congress/comm3/papers/347.pdf>

LAWRENCE, R. L.; RIPPLE, W. J. Comparisons among vegetation indices and bandwise regression in a highly disturbed, heterogeneous landscape: Mount St. Helens, Washington. **Remote Sensing of environment**, v. 64, n. 1, p. 91-102, 1998. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0034425797001715>

LEBOEUF, A. et al. A shadow fraction method for mapping biomass of northern boreal black spruce forests using QuickBird imagery. **Rem. Sens. Environ.**, v. 110, n. 4, p. 488-500, 2007. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0034425707001794>

LEFSKY, M. A. et al. Lidar remote sensing for ecosystem studies. **AIBS Bulletin**, v. 52, n. 1, p. 19-30, 2002. Disponível em: <https://academic.oup.com/bioscience/article/52/1/19/291259>

LI, Z.; XU, D.; GUO, X. Remote sensing of ecosystem health: opportunities, challenges, and future perspectives. **Sensors (Basel, Switzerland)**, v. 14, n. 11, p. 21117–21139, 2014. Disponível em: <https://www.mdpi.com/1424-8220/14/11/21117/htm>

LIM, K. et al. LiDAR remote sensing of forest structure. **Progress in physical geography**, v. 27, n. 1, p. 88-106, 2003. Disponível em: <http://journals.sagepub.com/doi/abs/10.1191/0309133303pp360ra>

LIMA, R. S. **Fatores de variação das estimativas de precipitação interna em uma floresta secundária**. 2017. Dissertação (Mestrado em Ciências Ambientais) - Programa de Pós-Graduação em Ciências Ambientais, Universidade Federal de São Carlos, São Carlos. Disponível em: <https://repositorio.ufscar.br/handle/ufscar/9150?show=full>

LONGHI, S. J. et al. Classificação e caracterização de estágios sucessionais em remanescentes de Floresta Ombrófila Mista na FLONA de São Francisco de Paula, RS, Brasil. **Ciência Florestal**, v. 16, n. 2, 2006. Disponível em: <https://periodicos.ufsm.br/cienciaflorestal/article/view/1893>

LONGHITANO, G. A.; AMARAL, C. H.; ALMEIDA, T. I. R. Detecção da biodiversidade de coberturas vegetais secundárias pela resposta espectral de pigmentos e água foliar em imagens do sensor hiperespectral CASI-1500. XIV SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO. **Anais...** p.1135–1142, 2009. Natal. Disponível em: <http://marte.sid.inpe.br/attachment.cgi/dpi.inpe.br/sbsr@80/2008/11.17.20.17/doc/1135-1142.pdf>

MADUGUNDU, R.; NIZALAPUR, V.; JHA, C. S. Estimation of LAI and above-ground biomass in deciduous forests: Western Ghats of Karnataka, India. **International**

Journal of Applied Earth Observation and Geoinformation, v. 10, n. 2, p. 211-219, 2008. Disponível em:

<https://www.sciencedirect.com/science/article/pii/S0303243407000773>

MAIRE, G. et al. Calibration of a species-specific spectral vegetation index for leaf area index (LAI) monitoring: example with MODIS reflectance time-series on eucalyptus plantations. **Remote Sensing**, v. 4, n. 12, p. 3766-3780, 2012. Disponível em:

<http://www.mdpi.com/2072-4292/4/12/3766/htm>

MAJASALMI, T.; RAUTIAINEN, M. The potential of Sentinel 2 data for estimating biophysical variables in a boreal forest: a simulation study. **Remote Sensing Letters**, v. 7, n. 5, p. 427-436, 2016. Disponível em:

<https://www.tandfonline.com/doi/abs/10.1080/2150704X.2016.1149251>

MARCELLO, J. et al. Assessment of atmospheric algorithms to retrieve vegetation in natural protected areas using multispectral high resolution imagery. **Sensors**, v. 16, n. 10, p. 1624, 2016. Disponível em: <https://www.mdpi.com/1424-8220/16/10/1624/htm>

MARENGO, J. A. et al. A seca e a crise hídrica de 2014-2015 em São Paulo. **Revista USP**, n. 106, p. 31-44, 2015. Disponível em:

<https://www.revistas.usp.br/revusp/article/view/110101>

MESQUITA, F. O. et al. Uso do sensoriamento remoto para avaliação da distribuição espacial e quantificação de Macrófitas na barragem Umari – Upanema , RN. **Agropecuária Científica no Semiárido**, v. 9, n. 2, p. 102–109, 2013. Disponível em:

<http://revistas.ufcg.edu.br/acsa/index.php/ACSA/article/view/255>

MORAN, M. S. et al. Evaluation of simplified procedures for retrieval of land surface reflectance factors from satellite sensor output. **Remote Sensing of Environment**, v. 41, n. 2-3, p. 169-184, 1992. Disponível em:

<https://www.sciencedirect.com/science/article/pii/003442579290076V>

MORENO, M. R.; NASCIMENTO, M. T.; KURTZ, B. C. Estrutura e composição florística do estrato arbóreo em duas zonas altitudinais na mata atlântica de encosta da região do imbé, RJ. **Acta Botanica Brasilica**, v. 17, n. 3, p. 371–386, 2003. Disponível em:

<http://www.scielo.br/pdf/%0D/abb/v17n3/17706.pdf>

MÜLLER-WILM, U.; DEVIGNOT, O.; PESSIOT, L. Sen2Cor Configuration and User Manual. 2018. Disponível em: <http://step.esa.int/main/third-party-plugins-2/sen2cor/>

NÆSSET, E.; ØKLAND, T. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. **Rem. Sens. Environ.**, v. 79, n. 1, p. 105-115, 2002. Disponível em:

<http://www.sciencedirect.com/science/article/pii/S0034425701002437>

NETO et al. Densidade Básica Da Madeira Para Quantificação De Biomassa Em Vegetação De Mata Atlântica. In: IV CONEFLOR – III SEEFOR, 2013, Vitória da Conquista. **Anais...** 2013. p. 435–440. Disponível em:

http://www.uesb.br/eventos/seeflor/publicacoes/2013/458_PDFsam_Anais_IV_CONEFLOR_III_SEEFOR.pdf

NEWTON, A. C. **Forest Ecology and Conservation: A Handbook of Techniques**. Oxford University Press. 2007. p. 480 Disponível em: https://www.researchgate.net/publication/258911672_Forest_Ecology_and_Conservation_A_Handbook_of_Techniques

OLIVEIRA, L. T. et al. Application of LIDAR to forest inventory for tree count in stands of Eucalyptus sp. **Cerne**, v. 18, n. 2, p. 175-184, 2012. Disponível em: http://www.scielo.br/scielo.php?pid=S0104-77602012000200001&script=sci_arttext&tlnq=es

OLIVEIRA et al. Quantificação De Biomassa E Carbono De Fragmento De Mata Atlântica Localizada Na Apa Do Pratigi. In: IV CONEFLORE – III SEEFLORE, 2013, Vitória da Conquista. **Anais...** 2013. p. 435–440. Disponível em: http://www.uesb.br/eventos/seeflor/publicacoes/2013/1121_PDFsam_Anais_IV_CONEFLORE_III_SEEFLORE.pdf

PARUELO, J. et al. Estimation of primary products of subhumid range lands from remote sensing data. **Applied Vegetation Science**, v. 3, n. 2, p. 189–195, 2000. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.2307/1478997>

PASCUAL, C. et al. Relationship between LiDAR-derived forest canopy height and Landsat images. **Int. Journ. Rem. Sens.**, v. 31, n. 5, p. 1261-1280, 2010. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/01431160903380656>

PEARSON, R. L.; MILLER, L. D. Remote mapping of standing crop biomass for estimation of the productivity of the shortgrass prairie. EIGHTH INTERNATIONAL SYMPOSIUM ON REMOTE SENSING OF ENVIRONMENT. **Anais...** p. 1355, 1972.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. **Journal of machine learning research**, v. 12, n. Oct, p. 2825-2830, 2011. Disponível em: <http://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf>

PEREZ, G. G.; BOURSCHEIDT, V. Avaliação de métodos de correção topográfica aplicados a índices de vegetação. In: V JORNADA DE GESTÃO E ANÁLISE AMBIENTAL: ÁREAS NATURAIS PROTEGIDAS, 5., 2018, São Carlos. **Anais...**

PFEIFER, N. et al. OPALS—A framework for Airborne Laser Scanning data analysis. **Computers, Environment and Urban Systems**, v. 45, p. 125-136, 2014. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0198971513001051>

PONZONI, F. J.; SHIMABUKURO, Y. E.; KUPLICH, T. M. **Sensoriamento remoto da vegetação**. 2. ed. Oficina de Textos, 2012.

PRATA, E. M. B.; ASSIS, M. A.; JOLY, C. A. Composição florística e estrutura da comunidade arbórea na transição da Floresta Ombrófila Densa das Terras Baixas - Floresta Ombrófila Densa Submontana do Núcleo Picinguaba / PESM, Ubatuba, sudeste do Brasil. **Biota Neotrop**, v. 11, n. 2, p. 285–299, 2011. Disponível em: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1676-06032011000200029

PRETZSCH, H. et al. Crown size and growing space requirement of common tree species in urban centres, parks, and forests. **Urban For. Urban Green.**, v. 14, n. 3, p. 466-479, 2015. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1618866715000473>

POPE, G.; TREITZ, P. Leaf area index (LAI) estimation in boreal mixedwood forest of Ontario, Canada using light detection and ranging (LiDAR) and WorldView-2 imagery. **Remote sensing**, v. 5, n. 10, p. 5040-5063, 2013. Disponível em: <http://www.mdpi.com/2072-4292/5/10/5040/htm>

QGIS Development Team. QGIS Geographic Information System. **Open Source Geospatial Foundation**. 2018. Disponível em: https://www.qgis.org/pt_BR/site/

RISSO, J. et al. Potencialidade dos índices de vegetação EVI e NDVI dos produtos MODIS na separabilidade espectral de áreas de soja. SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 14., 2009, Natal. **Anais...** p. 379-386, 2009. Disponível em: <http://marte.sid.inpe.br/attachment.cgi/dpi.inpe.br/sbsr@80/2008/11.18.00.07/doc/379-386.pdf>

ROBERTS, D. A.; ROTH, K. L.; PERROY, R. L. 14 Hyperspectral Vegetation Indices. **Hyperspectral remote sensing of vegetation.**, p. 309-327, 2016.

ROCCHINI, D.; CHIARUCCI, A.; LOISELLE, S. A. Testing the spectral variation hypothesis by using satellite multispectral images. **Acta Oecologica**, v. 26, n. 2, p. 117-120, 2004. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1146609X04000311>

ROCCHINI, D.; RICOTTA, C.; CHIARUCCI, A. Using satellite imagery to assess plant species richness : The role of multispectral systems. **Applied Vegetation Science**, v. 10, n. 2002, p. 325-331, 2007. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1654-109X.2007.tb00431.x>

RODRÍGUEZ-PÉREZ, J. R.; ÁLVAREZ, M. F.; SANZ-ABLANEDO, E. Assessment of low-cost GPS receiver accuracy and precision in forest environments. **Journal of Surveying Engineering**, v. 133, n. 4, p. 159-167, 2007. Disponível em: https://www.researchgate.net/publication/238179576_Assessment_of_Low-Cost_GPS_Receiver_Accuracy_and_Precision_in_Forest_Environments

ROUSE JR, J_W et al. Monitoring vegetation systems in the Great Plains with ERTS. GODDARD SPACE FLIGHT CENTER 3D ERTS-1 SYMP. v. 1, p. 309-317, 1974.

SABOL, J.; PATOCKA, Z.; MIKITA, T. Usage of LiDAR data for leaf area index estimation. **GeoScience Engineering**, v. 60, n. 3, p. 10-18, 2014. Disponível em: <https://content.sciendo.com/view/journals/gse/60/3/article-p10.xml>

SANCHES, L. D. et al. Análise comparativa de três métodos de correção atmosférica de imagens Landsat-5/TM-TM para obtenção de reflectância de superfície e NDVI. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 15., 2011, Curitiba.

Anais... São José dos Campos: INPE, 2011., 2011. Disponível em: <https://www.alice.cnptia.embrapa.br/handle/doc/902348>

SASAKI, T. et al. Estimation of leaf area index and canopy openness in broad-leaved forest using an airborne laser scanner in comparison with high-resolution near-infrared digital photography. **Landscape and Ecological Engineering**, v. 4, n. 1, p. 47-55, 2008. Disponível em: <https://link.springer.com/article/10.1007/s11355-008-0041-8>

SCHINDLER, D.; BAUHUS, J.; MAYER, H. Wind effects on trees. **European Journal of Forest Research**, v.131, n.1, p. 159-163, 2012. Disponível em: <https://link.springer.com/article/10.1007/s10342-011-0582-5>

SEABOLD, S.; PERKTOLD, J. Statsmodels: Econometric and statistical modeling with python. In: **Proceedings of the 9th Python in Science Conference**. Austin: SciPy society, 2010. p. 61. Disponível em: <https://www.statsmodels.org/stable/index.html>

SEAGUER, S. et al. Vegetation's red edge: a possible spectroscopic biosignature of extraterrestrial plants. **Astrobiology**, v. 5, n. 3, p. 372-390, 2005. Disponível em: <https://www.liebertpub.com/doi/abs/10.1089/ast.2005.5.372>

SILVA, M. V. R. et al. Aplicação do índice de vegetação ajustado ao solo-SAVI para a identificação de fragmentos de caatinga em cultivos de Agave sisalana Perrine na região Semiárida do Brasil. In: XVII SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO – SBSR, 17., 2015, João Pessoa. **Anais...** p. 5850-5857 Disponível em: <http://marte2.sid.inpe.br/attachment.cgi/sid.inpe.br/marte2/2015/06.15.16.03.47/doc/p1202.pdf>

SILVEIRA, R. L. V. A.; MOREIRA, A.; HIGASHI, E. N.. Crescimento e sobrevivência de mudas de eucalipto sob doses de boro cultivadas em condições de viveiro e de campo. 2004. Disponível em: https://www.researchgate.net/profile/Adonis_Moreira/publication/291197438_Crescimento_e_sobrevivencia_de_mudas_de_eucalipto_sob_doses_de_boro_cultivadas_e_m_condicoes_de_viveiro_e_de_campo/links/56b8802d08ae3c1b79b2d5b7.pdf

SOARES, M. L. G. Estrutura vegetal e grau de perturbação dos manguezais da Lagoa da Tijuca, Rio de Janeiro, RJ, BRASIL. **Revista Brasileira de Biologia**, v. 59, n. 3, p. 503–515, 1999. Disponível em: http://www.scielo.br/scielo.php?pid=S0034-71081999000300016&script=sci_abstract&tlng=pt

SOBHAN, M. I. **Species discrimination from a hyperspectral perspective**. 2007. 176 f. Tese - International Institute for Geoinformation Science and Earth Observation, Wageningen University, Netherlands. 2007.

SOTHE, C. et al. Evaluating Sentinel 2 and Landsat-8 data to map successional forest stages in a subtropical forest in Southern Brazil. **Remote Sensing**, v. 9, n. 8, p. 838, 2017. Disponível em: <http://www.mdpi.com/2072-4292/9/8/838/htm>

SOUZA, D. R. et al. Ciclo de corte econômico ótimo em floresta ombrófila densa de terra firme sob manejo florestal sustentável, Amazônia Oriental. **Revista Arvore**, v. 28, n. 5, 2004. Disponível em: <http://www.redalyc.org/html/488/48828507/>

SPOTO, F. et al. Overview of Sentinel-2. 2012 GEOSCIENCE AND REMOTE SENSING SYMPOSIUM (IGARSS), 2012 IEEE International. IEEE, 2012. p. 1707-1710. Disponível em: <https://ieeexplore.ieee.org/document/6351195>

STABEN, G.; LUCIEER, A.; SCARTH, P. Modelling LiDAR derived tree canopy height from Landsat TM, ETM+ and OLI satellite imagery—A machine learning approach. **International Journal of Applied Earth Observation and Geoinformation**, v. 73, p. 666-681, 2018. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0303243418303970>

SUGANUMA, M. S. et al. Comparando metodologias para avaliar a cobertura do dossel e a luminosidade no sub-bosque de um reflorestamento e uma floresta madura. **Revista Árvore**, v. 32, n. 2, 2008. Disponível em: <http://www.redalyc.org/html/488/48813383020/>

SUGANUMA, M. S. et al. Ecossistemas de referência para restauração de matas ciliares: existem padrões de biodiversidade, estrutura florestal e atributos funcionais? **Revista Árvore**, p. 835-847, 2013. Disponível em: <https://repositorio.unesp.br/handle/11449/109453>

SUGAWARA, L. M. et al. Avaliação de três métodos de estimativa de índice de área foliar aplicados à cana-de-açúcar. **Simpósio Brasileiro de Sensoriamento Remoto**, v. 14, p. 409-506, 2009. Disponível em: <http://marte.sid.inpe.br/col/dpi.inpe.br/sbsr@80/2008/11.18.00.14/doc/499-506.pdf>

TAKATA, J. T. **Abelhas visitantes florais da Mata Atlântica: efeitos de fatores locais e de paisagem**. 2017. 72p. Monografia (graduação em Ciências Biológicas) – Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2017.

TAN, S. Y. Satellite biomodelling: use of remote sensing for predicting plant biodiversity. INTERNATIONAL CONFERENCE - Studying, Modeling and Sense Making of Planet Earth. **Anais...** p.8, 2008. Disponível em: <http://www.aegean.gr/geography/earth-conference2008/papers/papers/B02ID164.pdf>

TSENG, Y.; LIN, L.; WANG, C. Mapping CHM and LAI for Heterogeneous Forests Using Airborne Full-Waveform LiDAR Data. Terrestrial, **Atmospheric & Oceanic Sciences**, v. 27, n. 4, 2016. Disponível em: <http://tao.cgu.org.tw/index.php/articles/archive/space-science/item/1374-2016012904isrs>

USGS, Landsat-8/OLI History, 2016. Disponível em: <https://landsat.usgs.gov/landsat-8-history>

VERRELST, J. et al. Experimental Sentinel 2 LAI estimation using parametric, non-parametric and physical retrieval methods—A comparison. **ISPRS Journal of Photogrammetry and Remote Sensing**, v. 108, p. 260-272, 2015. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0924271615001239>

VIEIRA, S. A. et al. Estimation of biomass and carbon stocks: the case of the Atlantic Forest. **Biota Neotropica**, v. 8, n. 2, p. 0-0, 2008. Disponível em: http://www.scielo.br/scielo.php?pid=S1676-06032008000200001&script=sci_arttext

VIÉRLING, Lee A. et al. Relationships among airborne scanning lidar, high resolution multispectral imagery, and ground-based inventory data in a ponderosa pine forest. IN: GEOSCIENCE AND REMOTE SENSING SYMPOSIUM, 2002. IGARSS'02. 2002 IEEE INTERNATIONAL. IEEE, 2002. p. 2912-2914. Disponível em: <https://ieeexplore.ieee.org/abstract/document/1026819>

WHATELY, M.; CUNHA, P. Cantareira. um olhar sobre o maior manancial de água da Região Metropolitana de São Paulo: resultados do diagnóstico socioambiental participativo do sistema Cantareira. **São Paulo: Instituto Socioambiental**, 2007. Disponível em: <http://www.bibliotecadigital.abong.org.br/bitstream/handle/11465/1200/10289.pdf?sequence=1>

XIAO, C. et al. Variation of specific leaf area and upscaling to leaf area index in mature Scots pine. **Trees**, v. 20, n. 3, p. 304, 2006. Disponível em: <https://link.springer.com/article/10.1007/s00468-005-0039-x>

XIAO, Q.; MCPHERSON, E. G. Tree health mapping with multispectral remote sensing data at UC Davis, California. **Urban Ecosystems**, v. 8, n. 3, p. 349-361, 2005. Disponível em: <https://link.springer.com/article/10.1007%2Fs11252-005-4867-7?LI=true>

XUE, J.; SU, B. Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. **Journal of Sensors**, v. 2017, 2017. Disponível em: <https://www.hindawi.com/journals/js/2017/1353691/abs/>

ZHAO, K.; POPESCU, S. Lidar-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA. **Rem. Sens. Environ.**, v. 113, n. 8, p. 1628-1645, 2009. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0034425709000893>

ZHU, J. et al. Wind profiles in and over trees. **J. For. Res.**, v. 15, n. 4, p. 305-312, 2004. Disponível em: <https://link.springer.com/article/10.1007/BF02844959>

APÊNDICE A – CONSTRUÇÃO DAS VARIÁVEIS LIDAR



Figura A1: Fluxogramas mostrando os procedimentos utilizados para o cálculo das variáveis LiDAR.

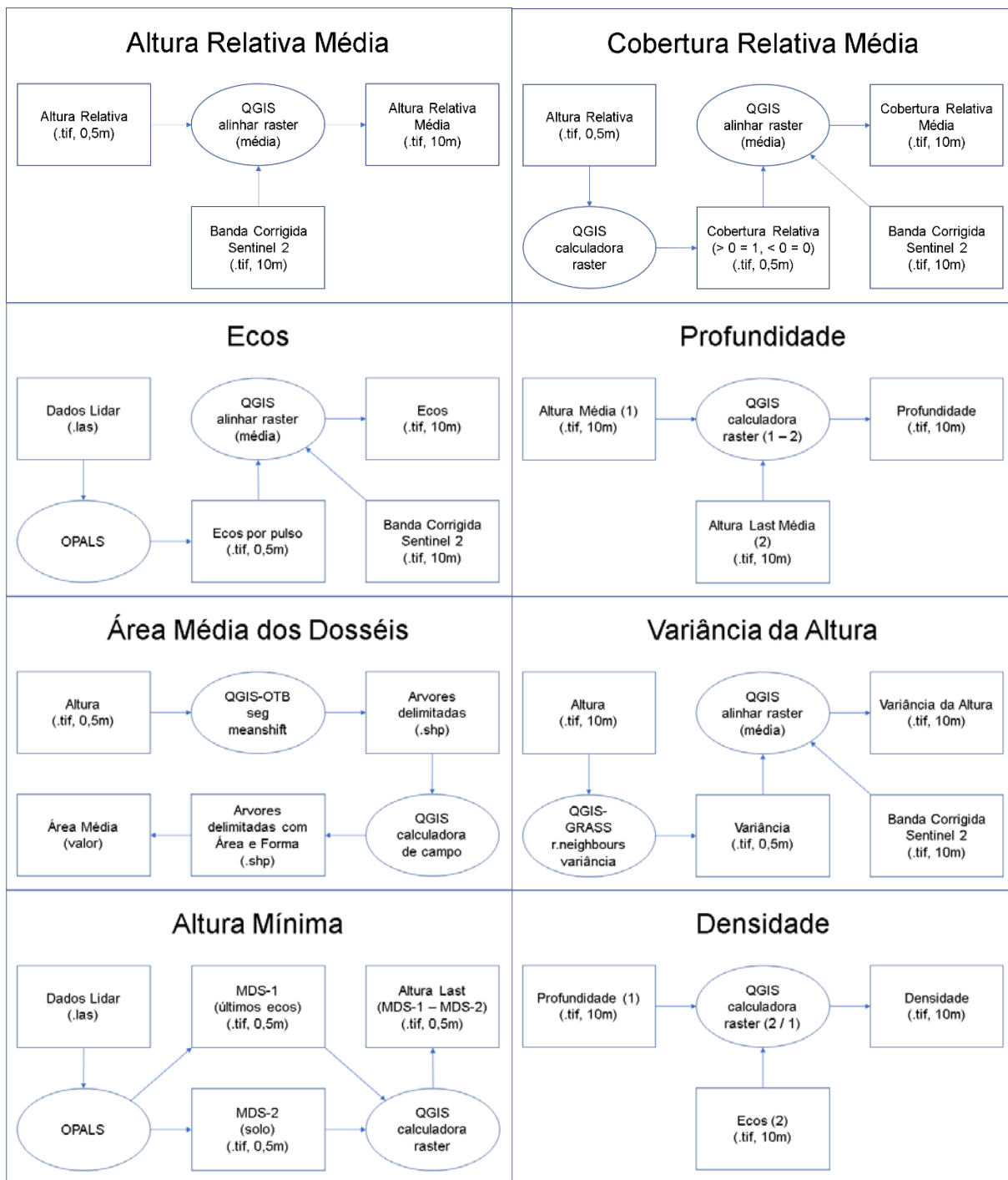


Figura A2: Fluxogramas mostrando os procedimentos utilizados para o cálculo das variáveis LiDAR.

APÊNDICE B – GRÁFICOS E QUADROS COMPLEMENTARES

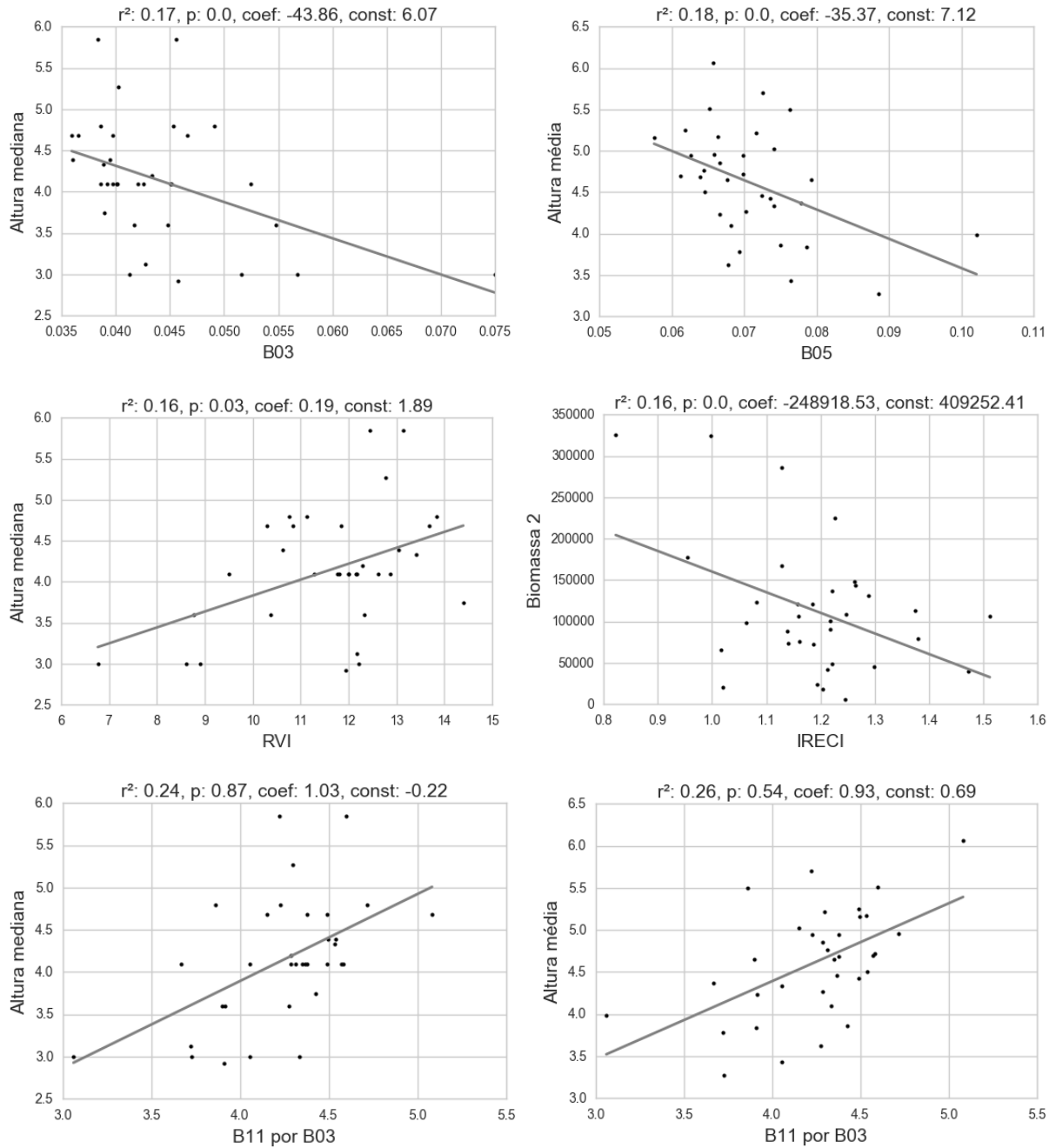


Figura B1: Gráficos de pontos com a reta do modelo de regressão (OLS) para algumas das melhores relações entre variáveis de campo e imagens do Sentinel 2.

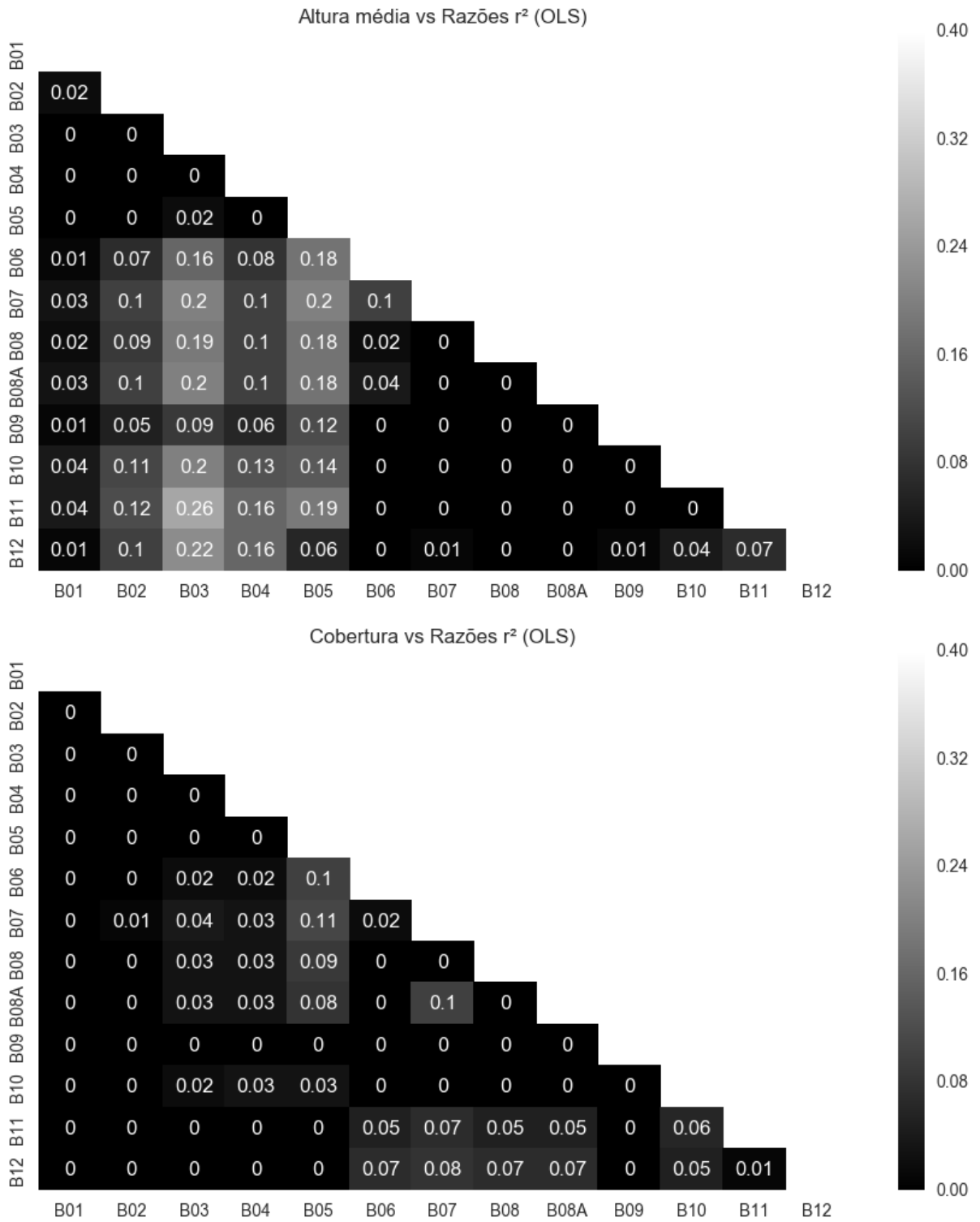


Figura B2: Gráficos mostrando os valores de r^2 ajustado para modelos de regressão (OLS) entre as variáveis de campo e razões entre bandas do Sentinel 2.

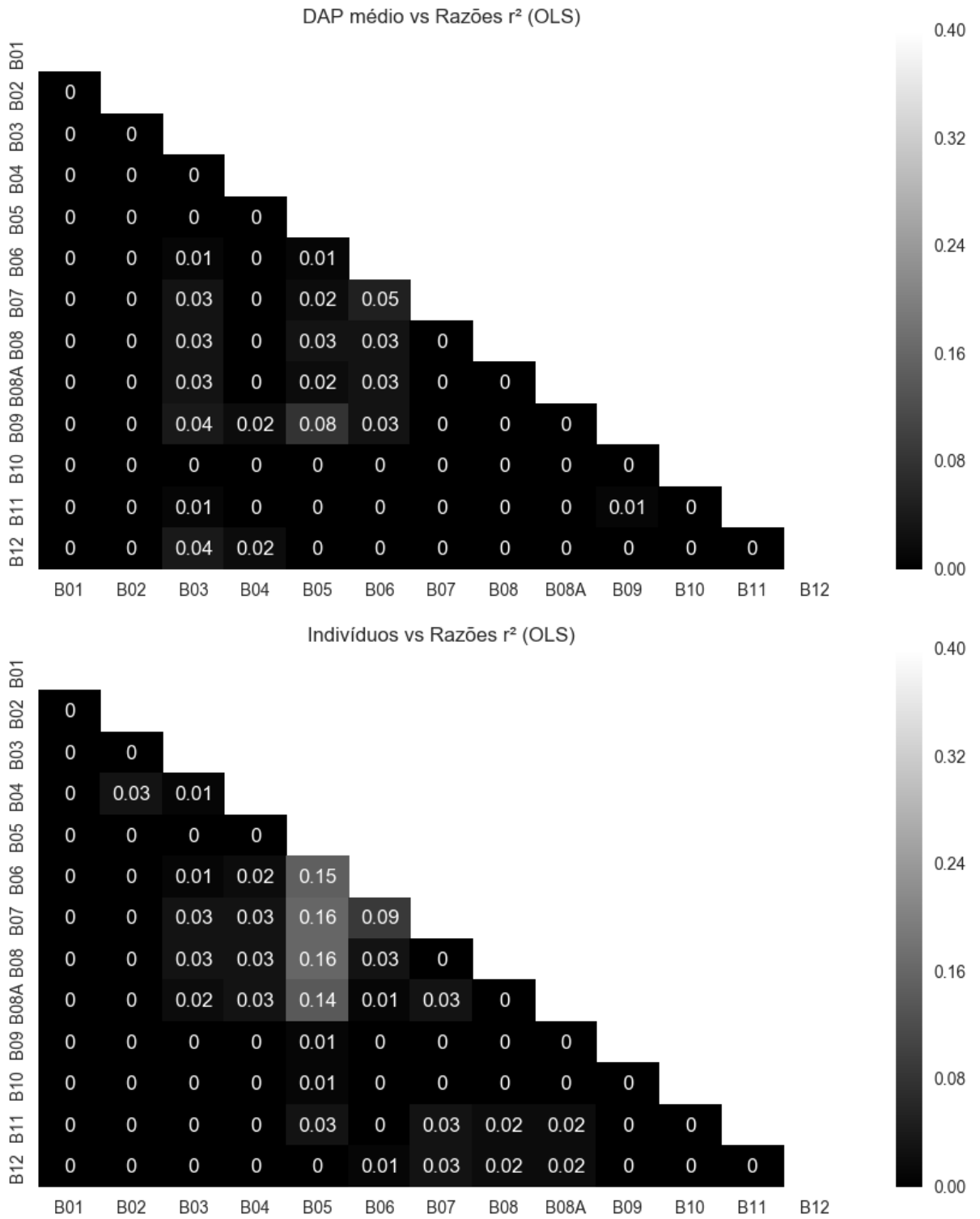


Figura B3: Gráficos mostrando os valores de r² ajustado para modelos de regressão (OLS) entre as variáveis de campo e razões entre bandas do Sentinel 2.

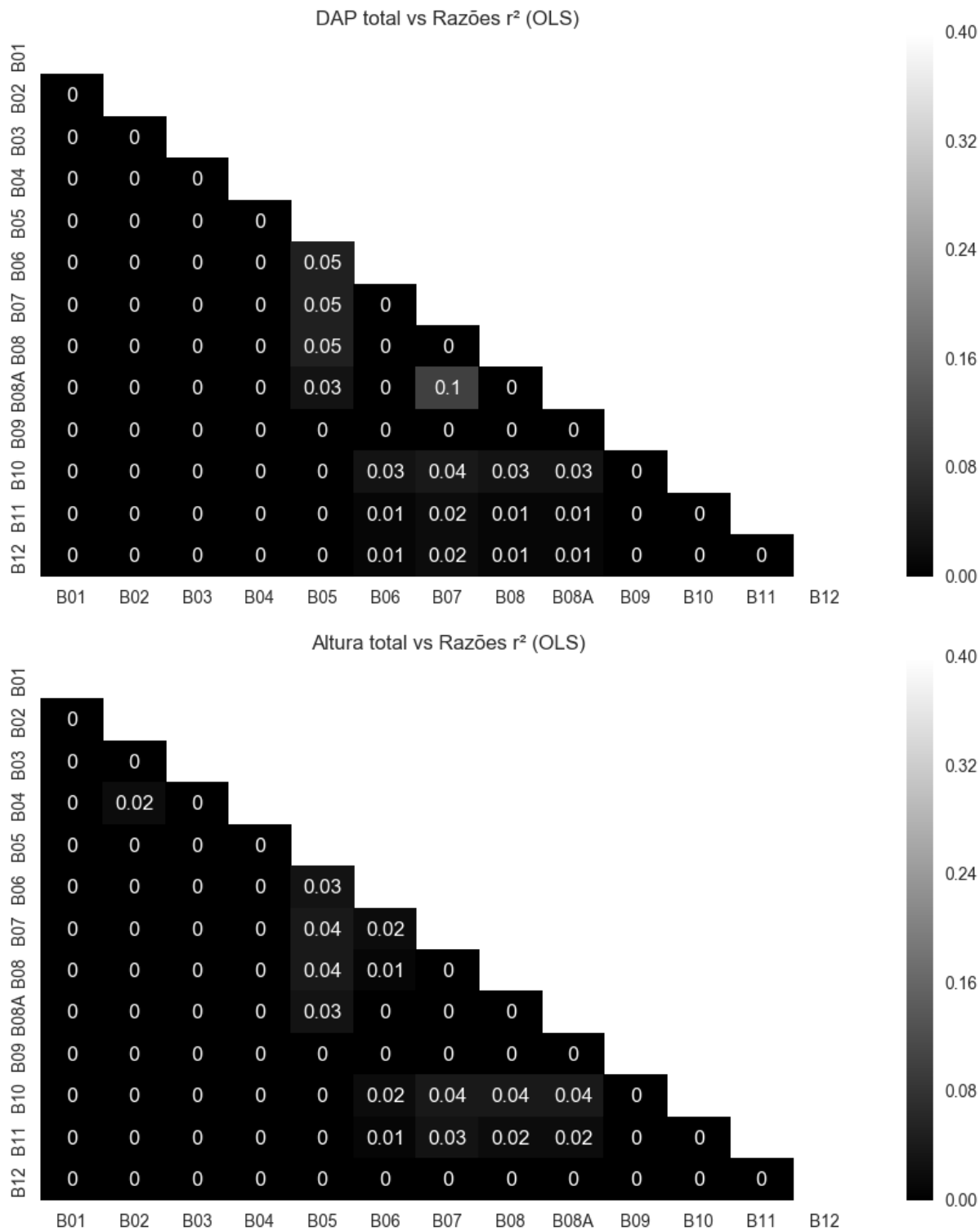


Figura B4: Gráficos mostrando os valores de r^2 ajustado para modelos de regressão (OLS) entre as variáveis de campo e razões entre bandas do Sentinel 2.

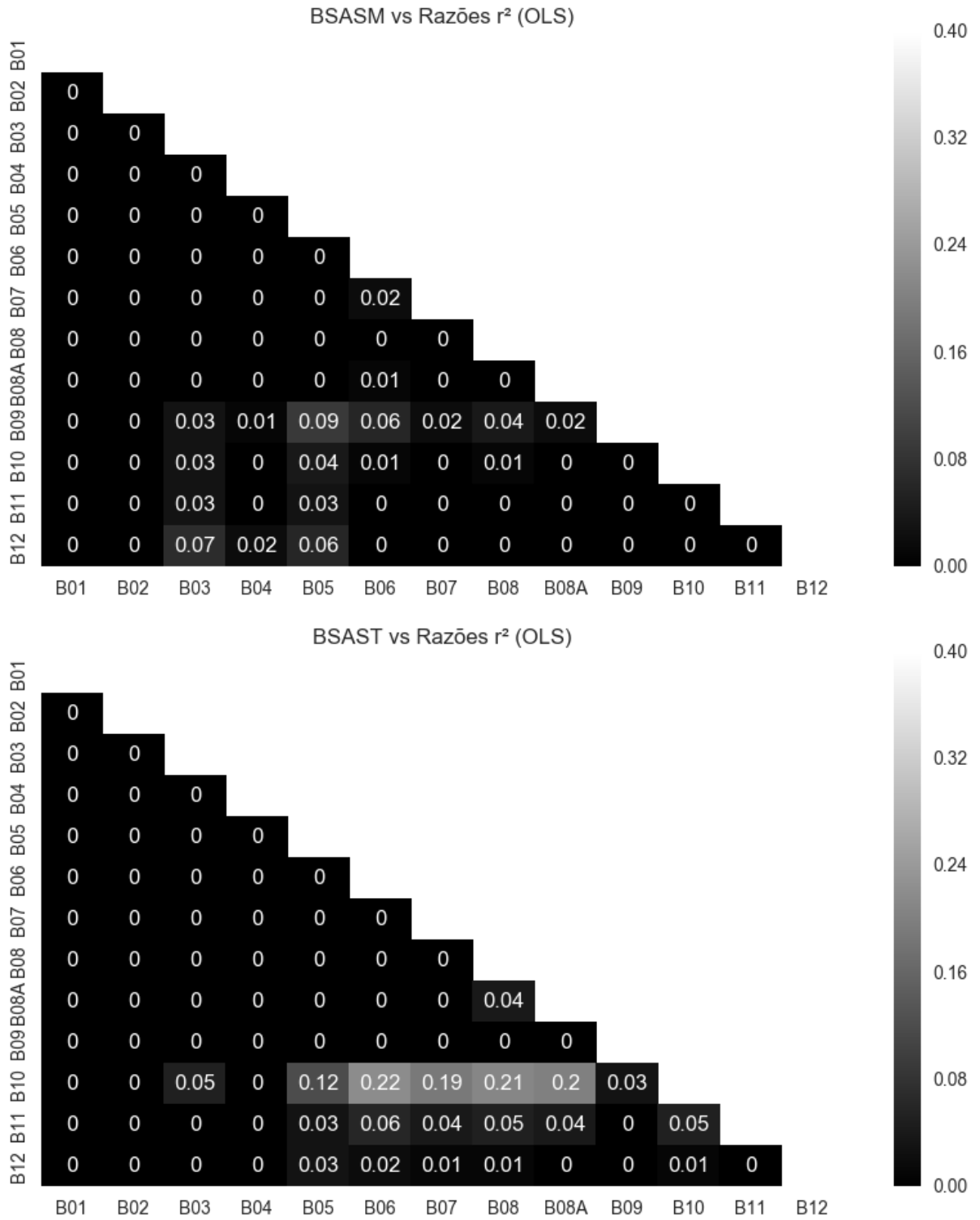


Figura B5: Gráficos mostrando os valores de r^2 ajustado para modelos de regressão (OLS) entre as variáveis de campo e razões entre bandas do Sentinel 2.

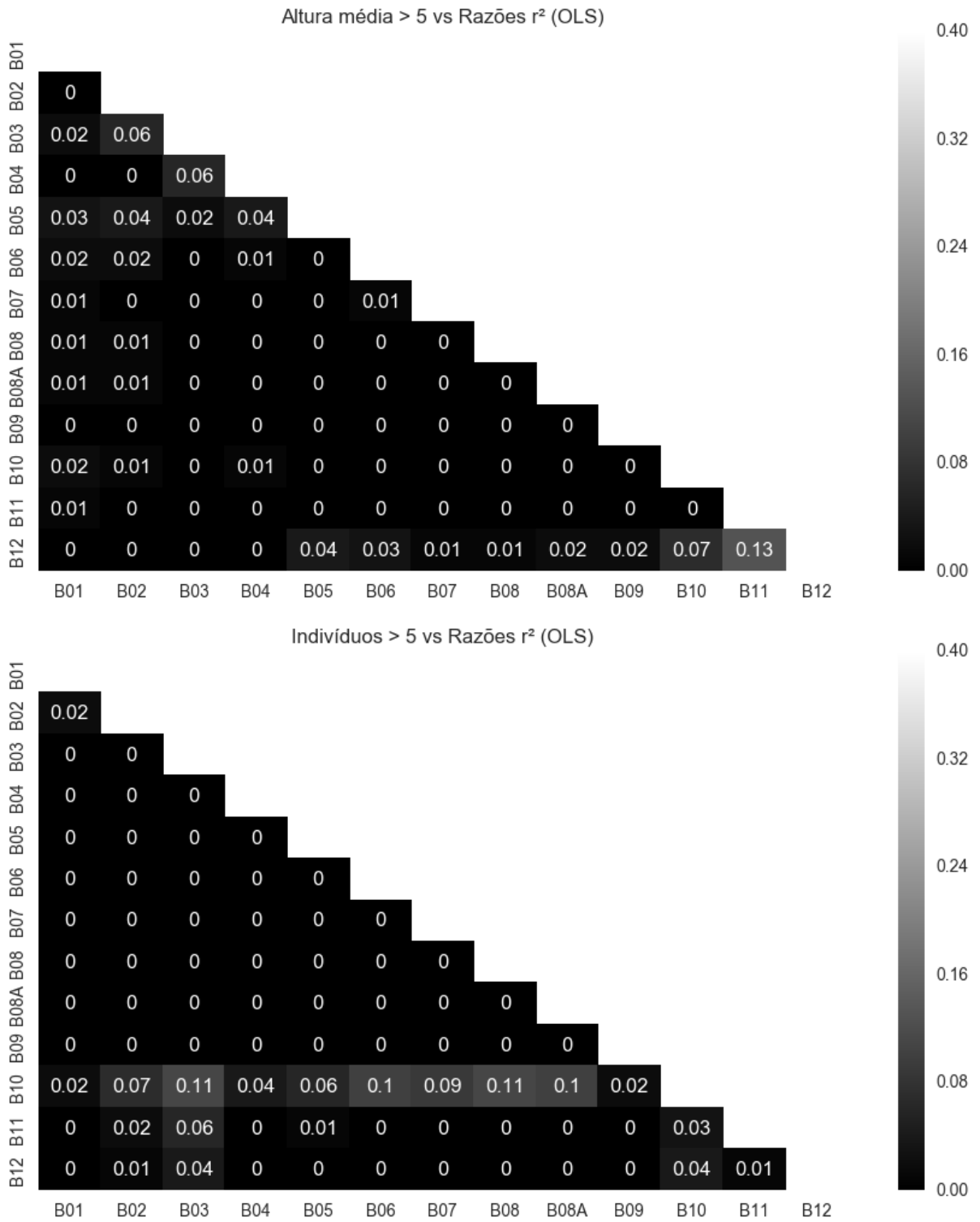


Figura B6: Gráficos mostrando os valores de r² ajustado para modelos de regressão (OLS) entre as variáveis de campo e razões entre bandas do Sentinel 2.

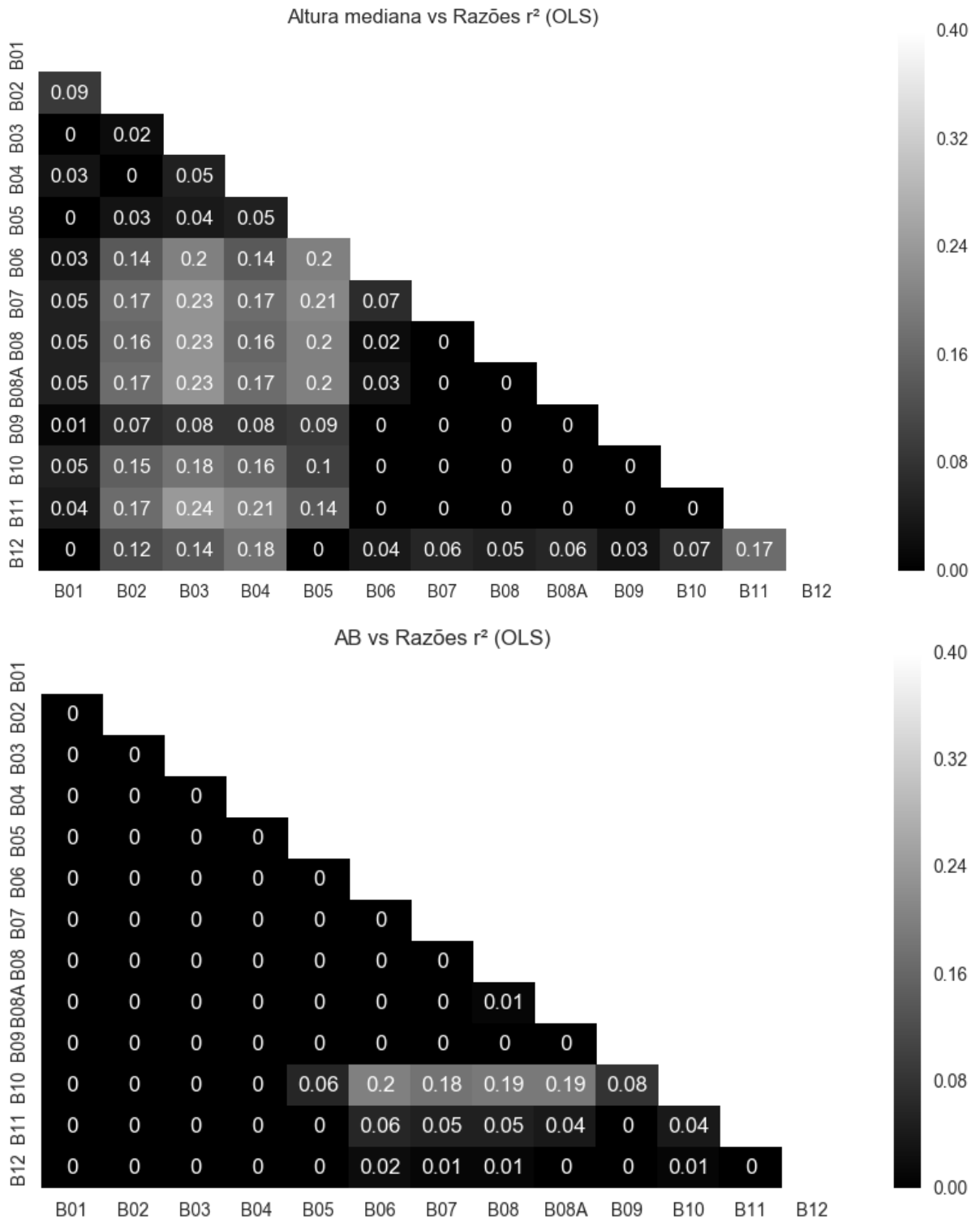


Figura B7: Gráficos mostrando os valores de r^2 ajustado para modelos de regressão (OLS) entre as variáveis de campo e razões entre bandas do Sentinel 2.

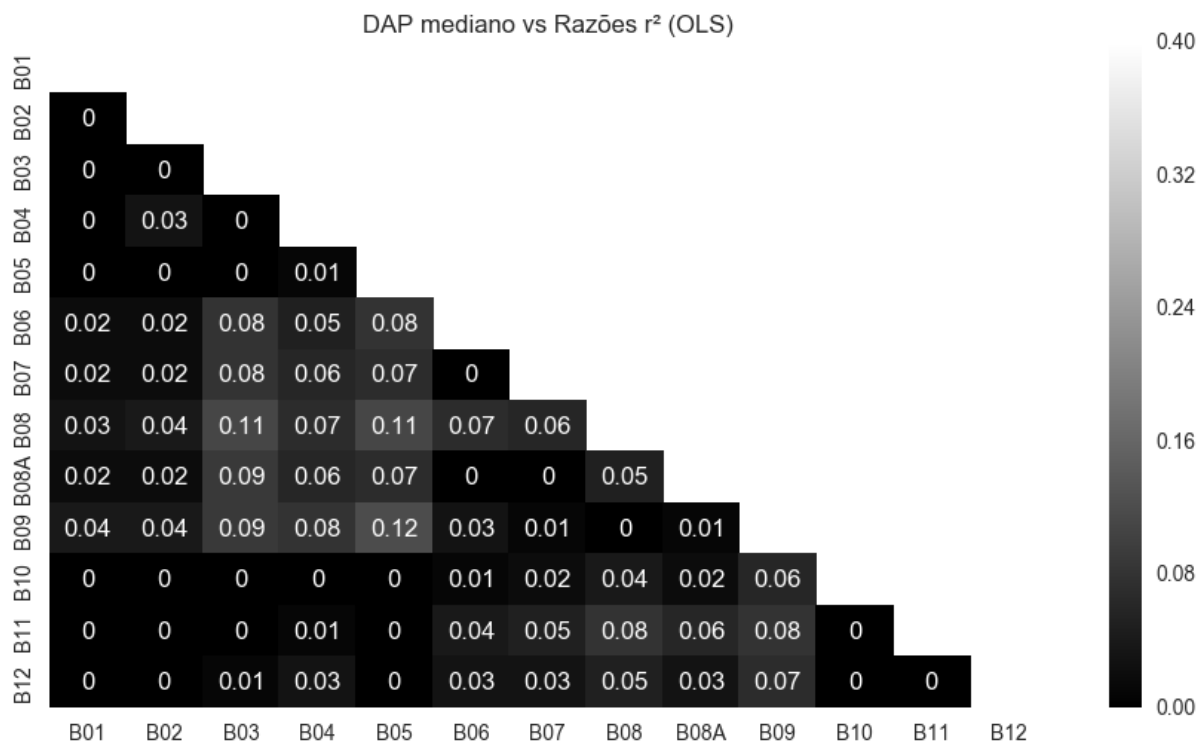


Figura B8: Gráficos mostrando os valores de r^2 ajustado para modelos de regressão (OLS) entre as variáveis de campo e razões entre bandas do Sentinel 2.

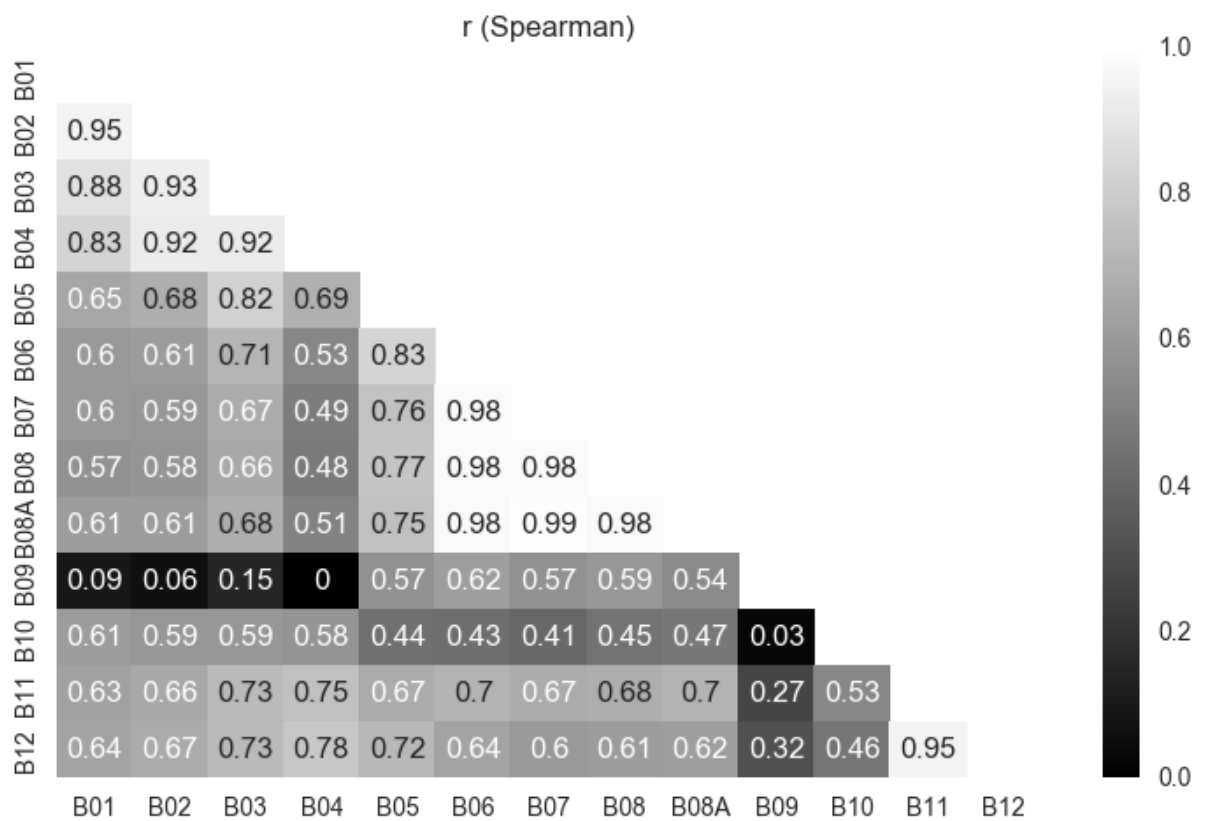
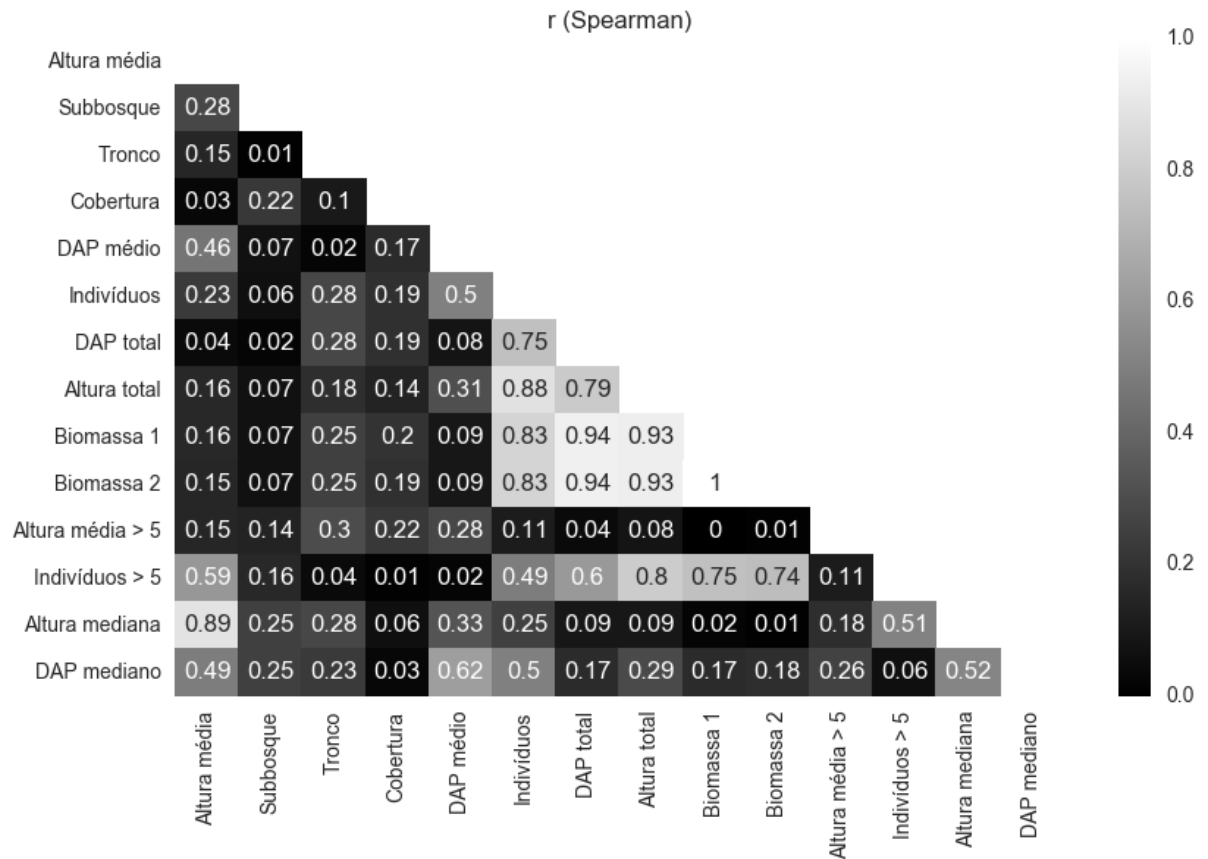


Figura B9: Gráficos mostrando os valores de r (Spearman) para correlação entre variáveis de campo (imagem de cima) e entre bandas (imagem de baixo).

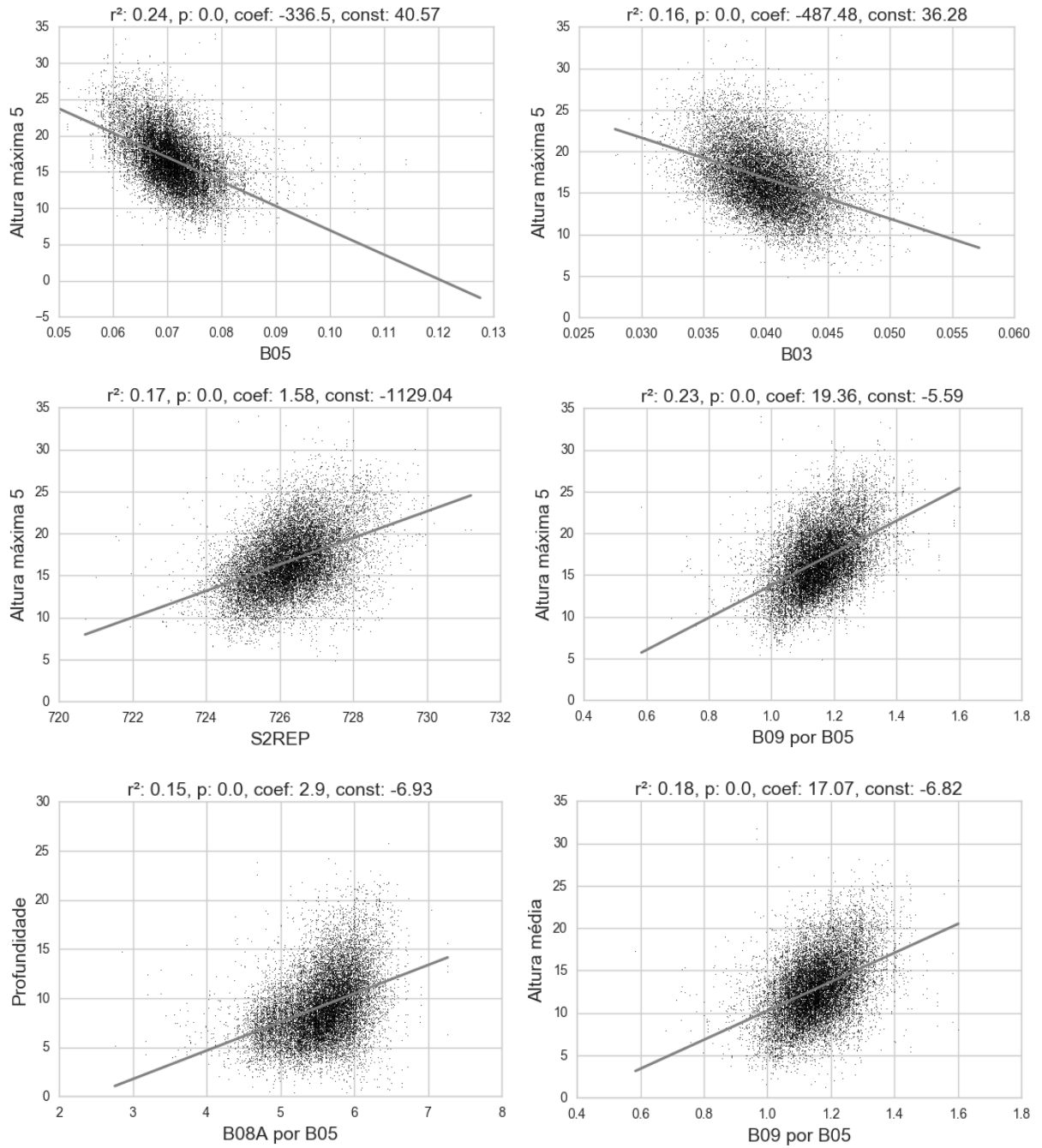


Figura B10: Gráficos de pontos com a reta do modelo de regressão (OLS) para algumas das melhores relações entre variáveis LiDAR e imagens do Sentinel 2 da comparação pixel a pixel.

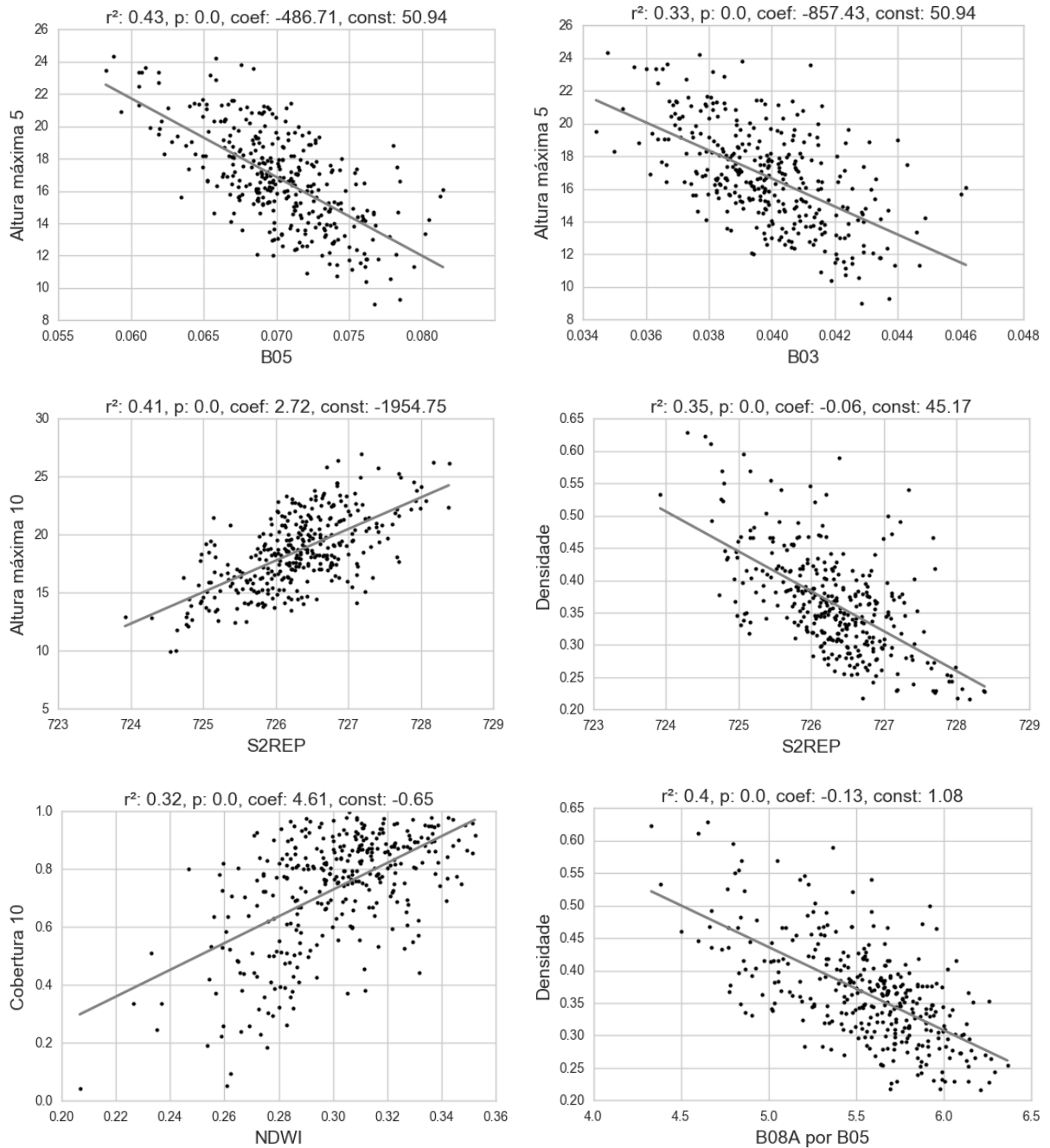


Figura B11: Gráficos de pontos com a reta do modelo de regressão (OLS) para algumas das melhores relações entre variáveis LiDAR e imagens do Sentinel 2 da comparação com buffers.

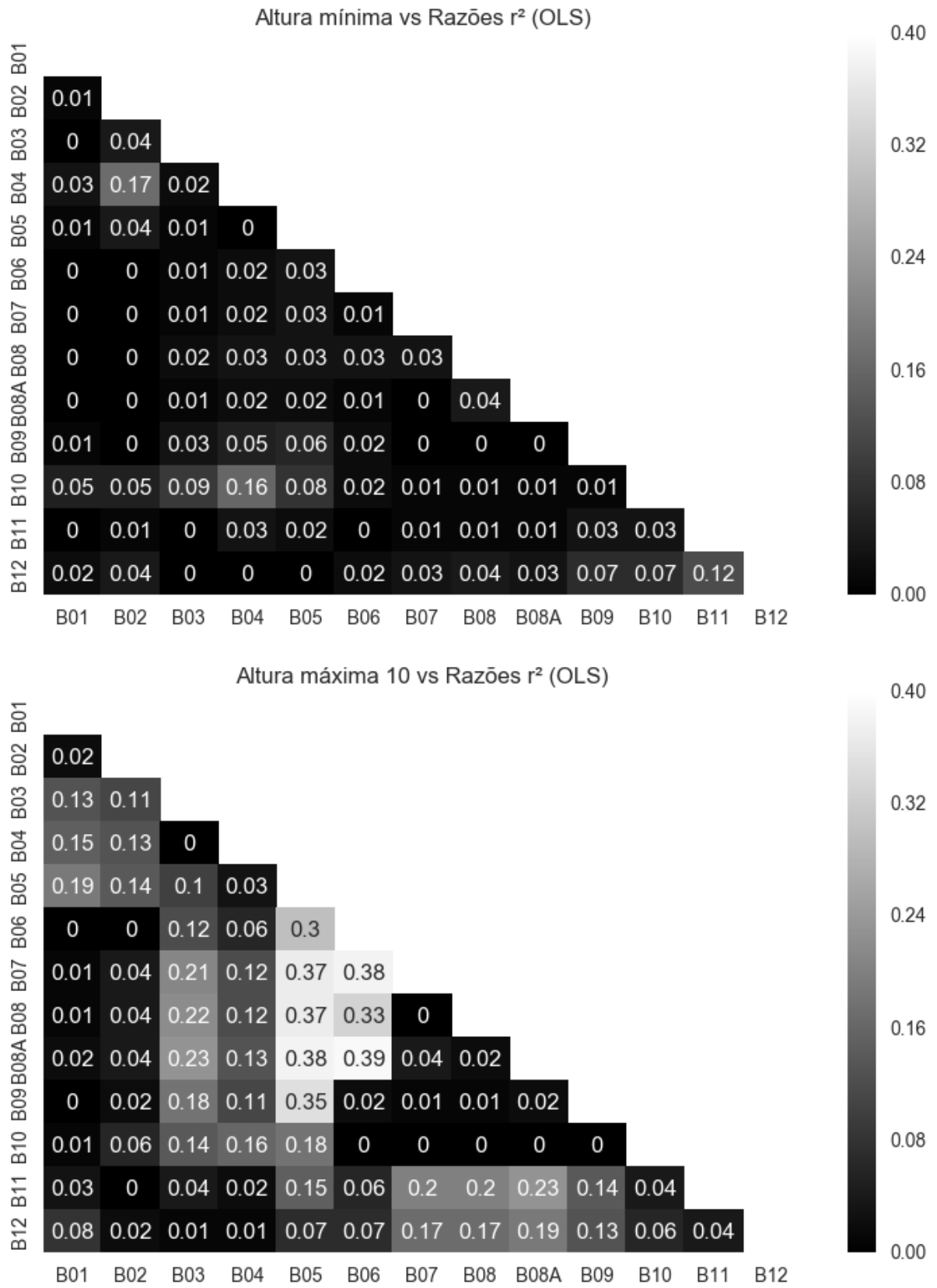


Figura B12: Gráficos mostrando os valores de r² ajustado para modelos de regressão (OLS) entre as variáveis LiDAR e razões entre bandas do Sentinel 2.

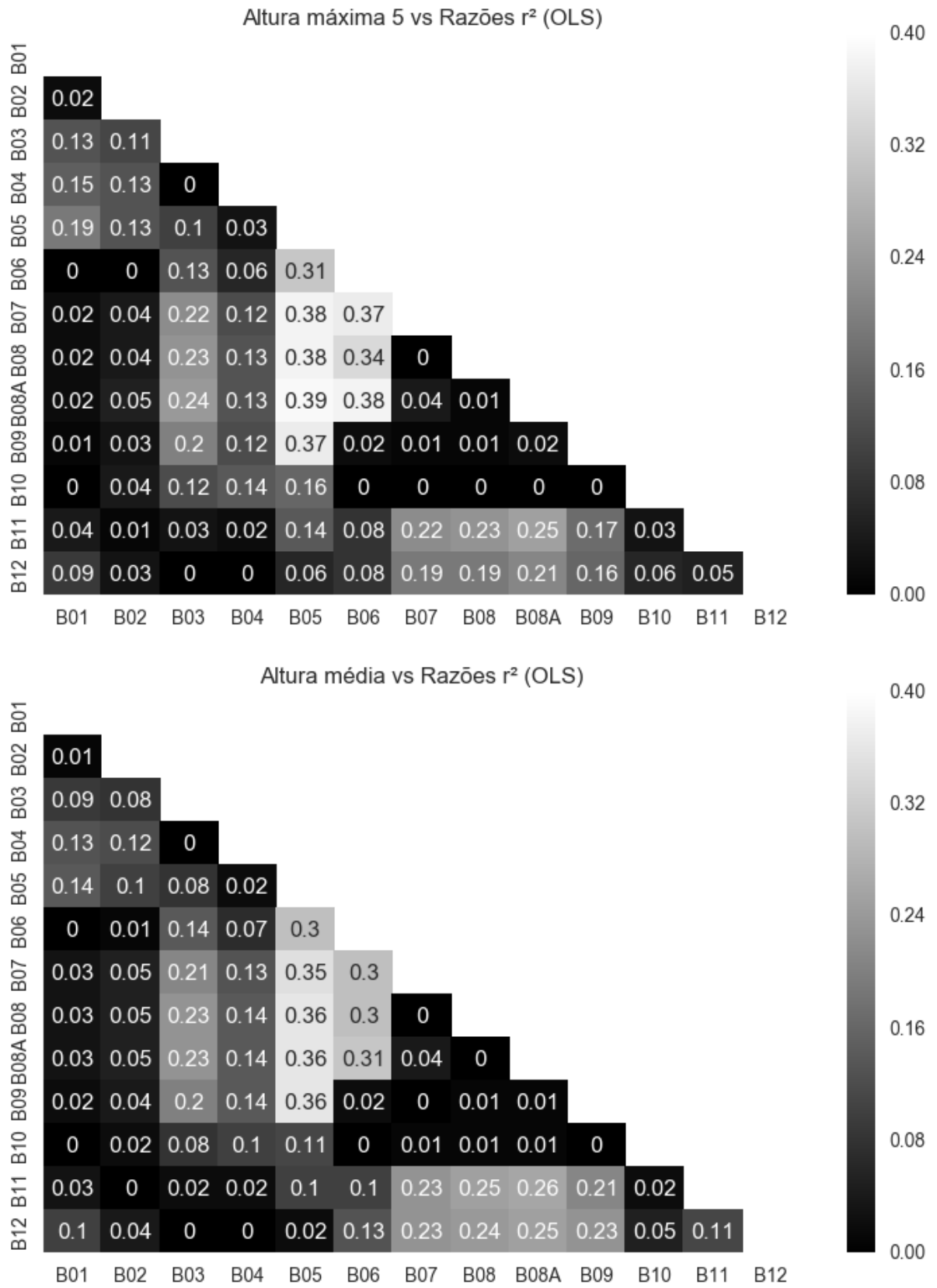


Figura B13: Gráficos mostrando os valores de r^2 ajustado para modelos de regressão (OLS) entre as variáveis LiDAR e razões entre bandas do Sentinel 2.

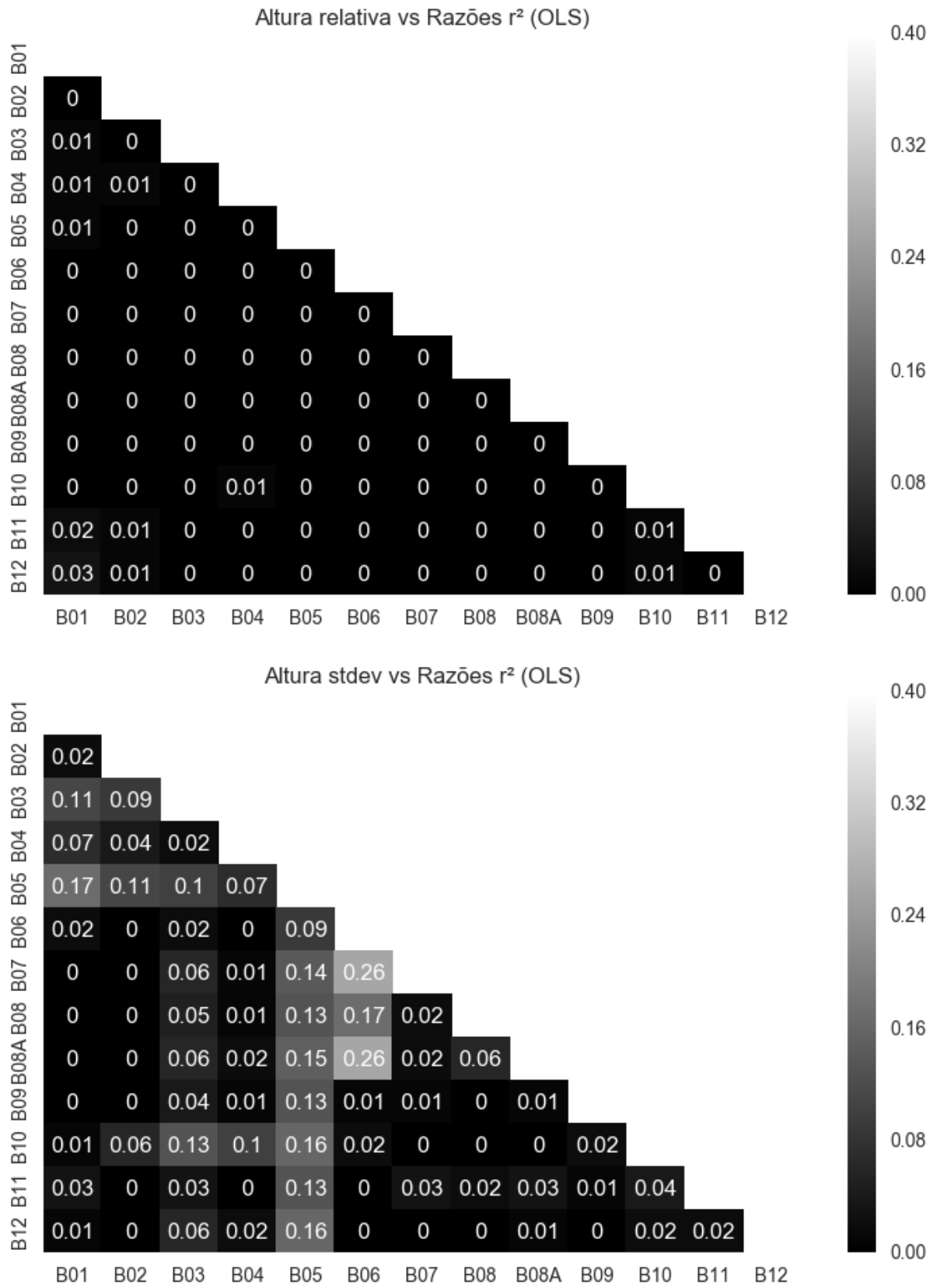


Figura B14: Gráficos mostrando os valores de r^2 ajustado para modelos de regressão (OLS) entre as variáveis LiDAR e razões entre bandas do Sentinel 2.

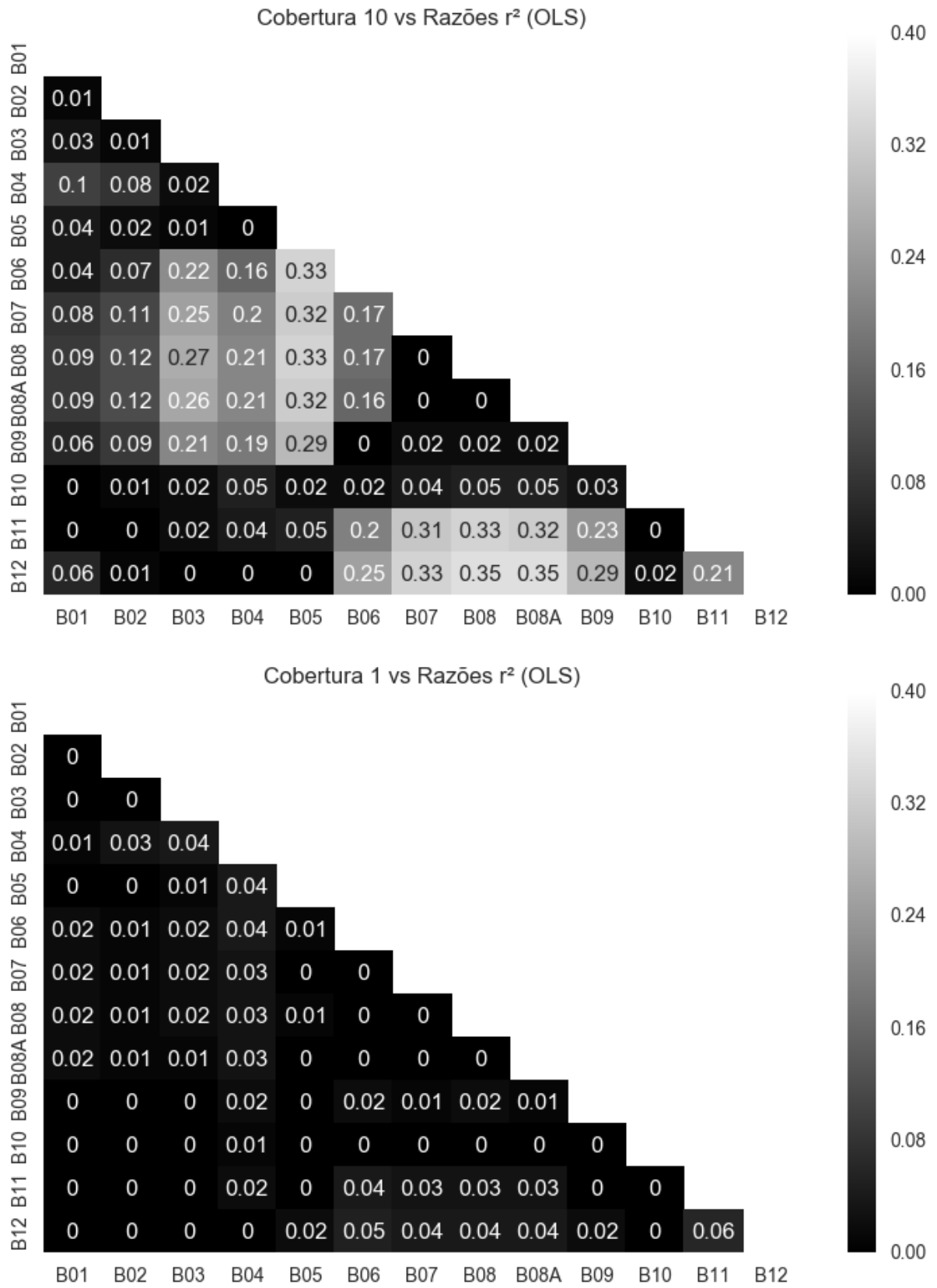


Figura B15: Gráficos mostrando os valores de r^2 ajustado para modelos de regressão (OLS) entre as variáveis LiDAR e razões entre bandas do Sentinel 2.

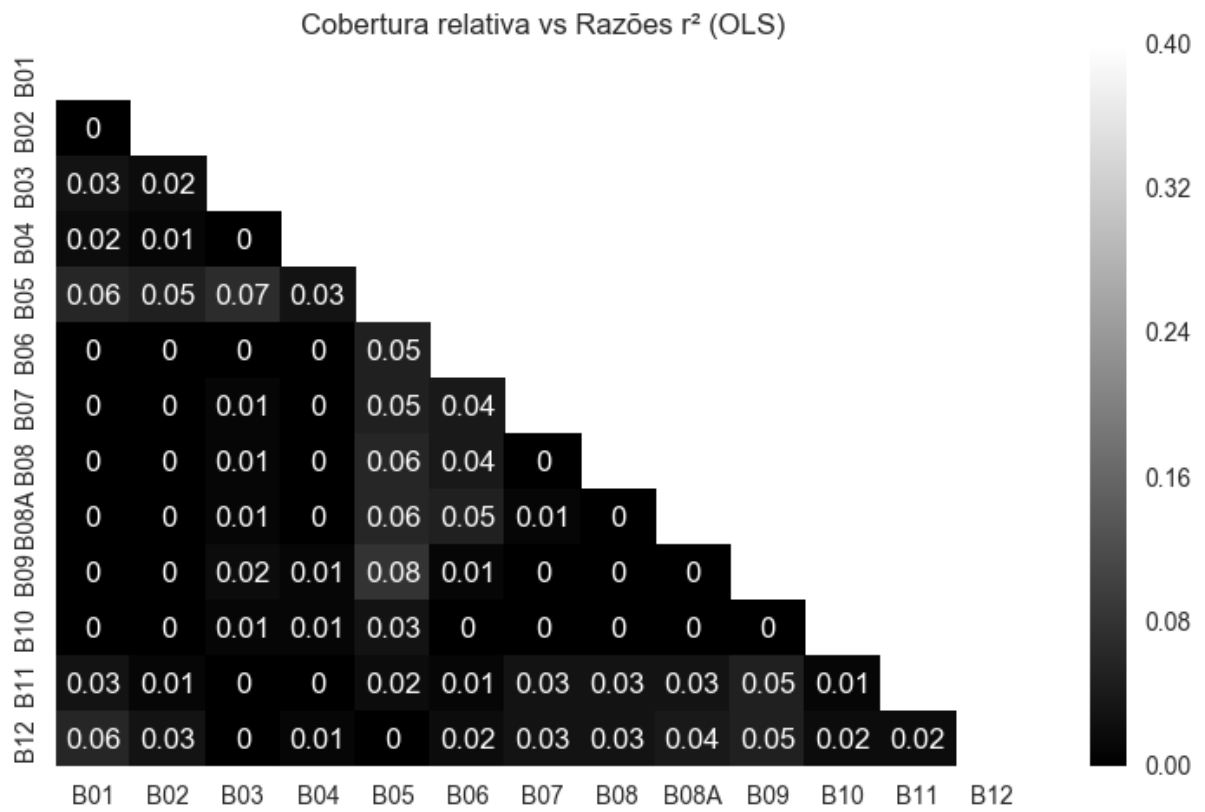
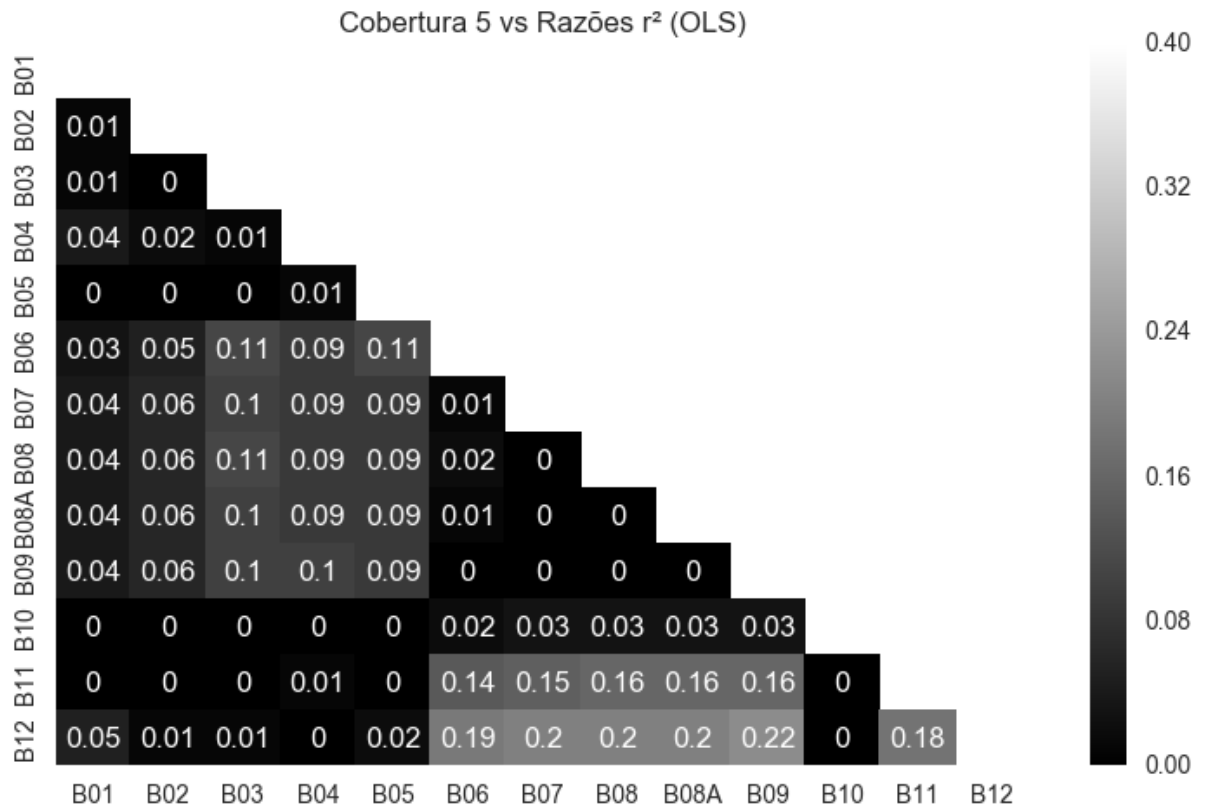


Figura B16: Gráficos mostrando os valores de r^2 ajustado para modelos de regressão (OLS) entre as variáveis LiDAR e razões entre bandas do Sentinel 2.

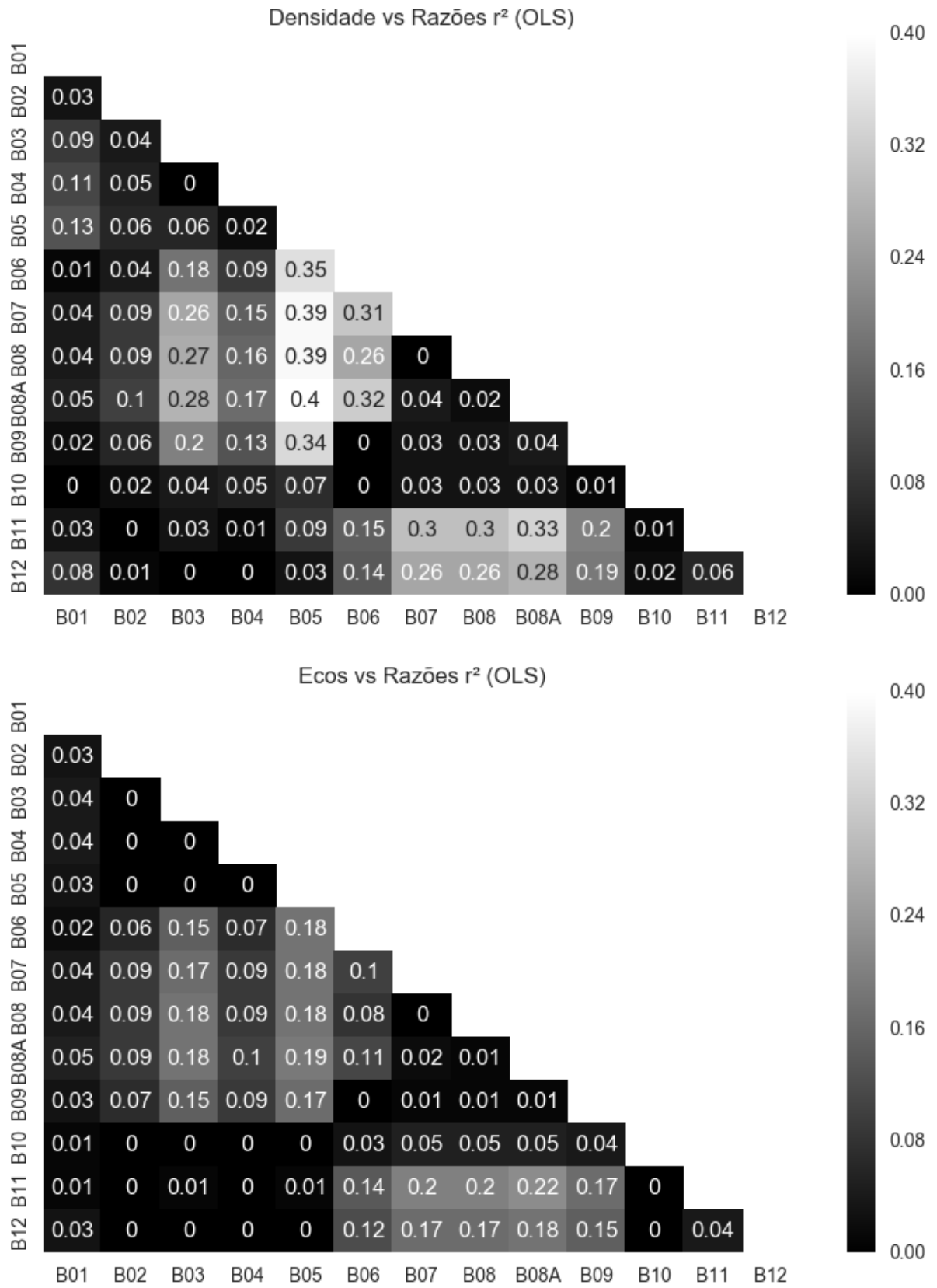


Figura B17: Gráficos mostrando os valores de r^2 ajustado para modelos de regressão (OLS) entre as variáveis LiDAR e razões entre bandas do Sentinel 2.

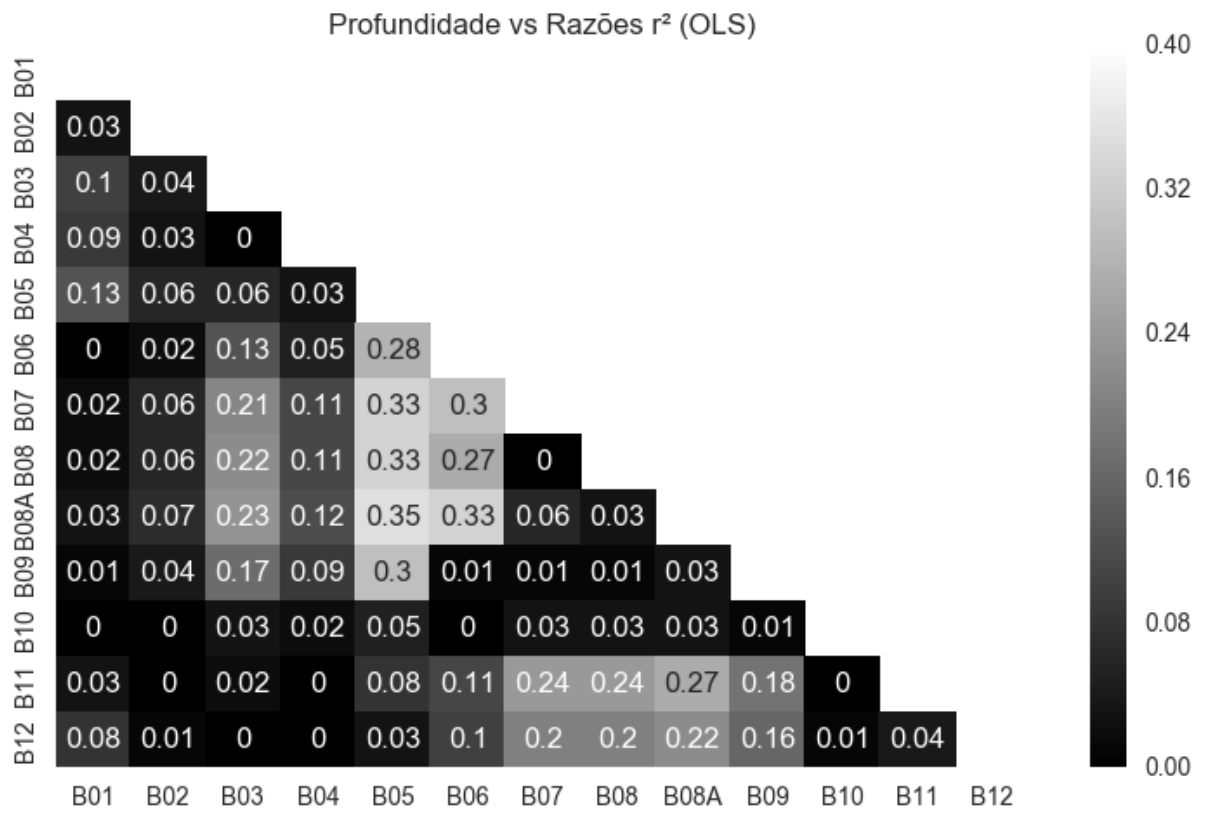


Figura B18: Gráficos mostrando os valores de r^2 ajustado para modelos de regressão (OLS) entre as variáveis LiDAR e razões entre bandas do Sentinel 2.

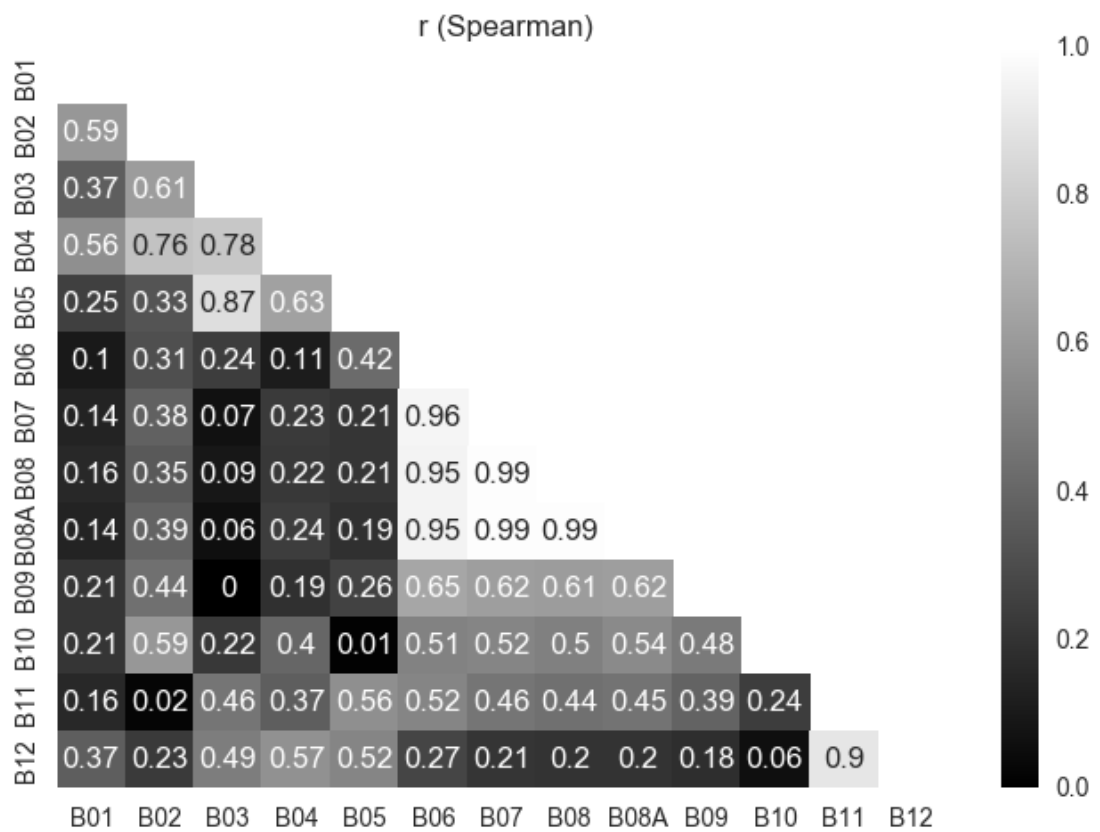
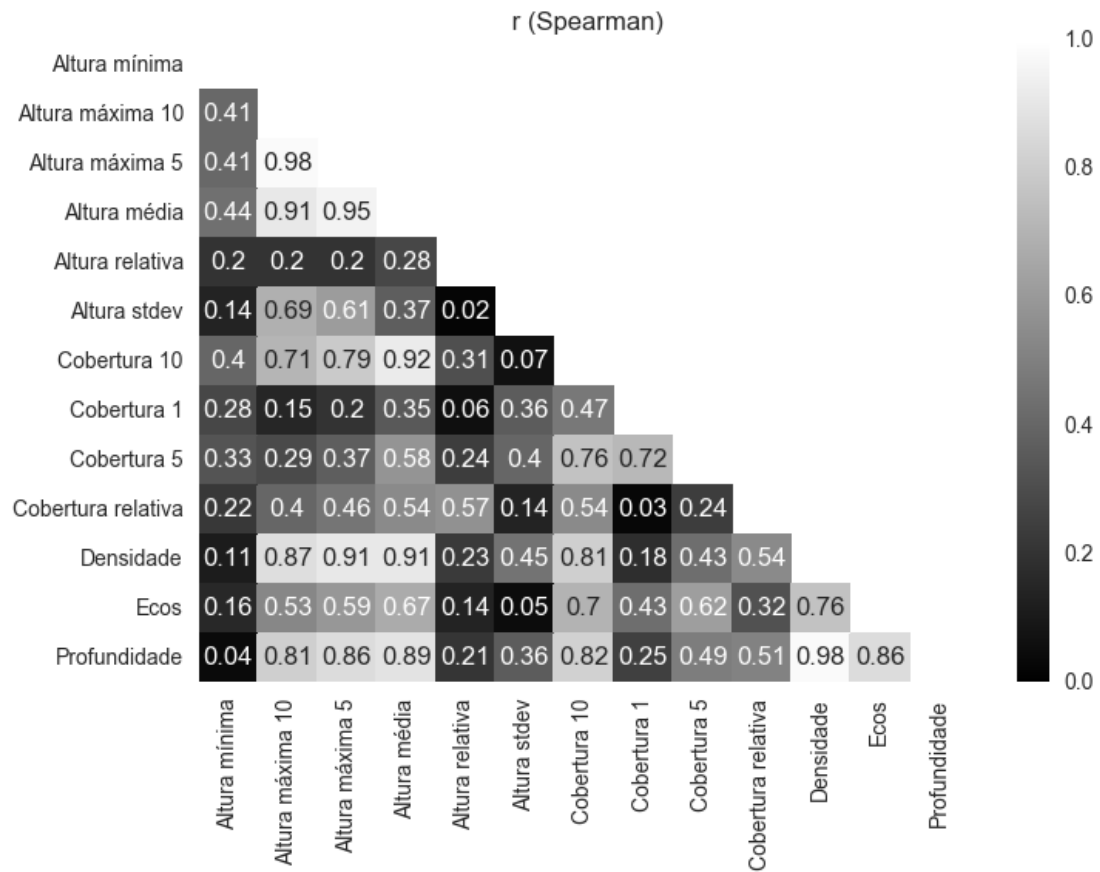


Figura B19: Gráficos dos valores de r (Spearman) para a correlação entre as variáveis LiDAR (imagem de cima) e entre bandas do Sentinel 2 (imagem de baixo).

Quadro B1: Melhores modelos de regressão (OLS) entre múltiplas bandas do Sentinel 2 e variáveis de estrutura vegetal medidas em campo segundo o valor de AIC com r^2 ajustado, AIC e RMSE.

Variável	Fórmula	r^2	AIC	RMSE
AM	$6,06 + B01 * 119,91 + B03 * -371,56 + B04 * 210,73 + B08 * 15,19$	0,38	55,62	0,47
C	$104,2 + B01 * -53,41 + B04 * 226,13 + B05 * -143,91 + B06 * 67,83 + B07 * 86,19 + B08A * -104,82 + B12 * -88,88$	0,48	45,02	0,37
DAPM	$4,53 + B06 * -309,64 + B08A * 184,28 + B09 * 167,22$	0,14	155,71	2,12
N	$22,59 + B03 * -3905,6 + B05 * 5046,31 + B07 * -3187,26 + B08A * 2642,0 + B09 * -1056,09$	0,26	304,07	17,7
DAPT	$872,01 + B05 * 11540,58 + B07 * -22564,6 + B08A * 18702,37 + B11 * -4076,64$	0,16	434,62	125
AT	$529,3 + B05 * 4525,13 + B07 * -1714,64$	0,08	407,4	88,6
AB	$205,32 + B03 * 1436,98 + B06 * -852,06$	0,2	317,52	23,7
BSASM	$147,2 + B01 * 2835,74 + B06 * -5671,13 + B07 * 3091,68 + B09 * 3085,1$	0,26	332,87	27,9
BSAST	$12284,41 + B01 * 90508,34 + B06 * -48062,78$	0,23	591,37	1325
AM > 5	$8,68 + B05 * -75,58 + B07 * 12,58 + B11 * -83,55 + B12 * 192,62$	0,21	72,86	0,61
N > 5	$-4,69 + B01 * 1539,81 + B03 * -5296,02 + B04 * 2058,09 + B05 * 1609,92 + B09 * -518,66 + B10 * 6498,65$	0,2	246,45	7,38
AMed	$4,41 + B01 * 194,24 + B02 * -187,17 + B03 * -286,54 + B04 * 213,42 + B08 * 17,38$	0,33	68,47	0,56
DAPMed	$11,96 + B02 * 453,26 + B03 * -475,21 + B07 * -109,45 + B08 * 139,06 + B10 * -772,7$	0,27	116,46	1,12

Quadro B2: Melhores modelos de regressão (OLS) entre múltiplas bandas do Sentinel 2 e variáveis de cobertura vegetal medidas em campo segundo o valor de r^2 ajustado com r^2 ajustado, AIC e RMSE.

Variável	Fórmula	r^2	AIC	RMSE
AM	$6,48 + B01 * 134,35 + B03 * -367,65 + B04 * 257,01 + B05 * -38,3 + B07 * 70,31 + B08 * 45,6 + B08A * -94,65 + B11 * 48,64 + B12 * -81,63$	0,43	56,18	0,41
C	$104,2 + B01 * -53,41 + B04 * 226,13 + B05 * -143,91 + B06 * 67,83 + B07 * 86,19 + B08A * -104,82 + B12 * -88,88$	0,48	45,02	0,37
DAPM	$11,22 + B04 * -434,12 + B06 * -229,49 + B08A * 175,07 + B11 * -311,87 + B12 * 767,01$	0,16	156,9	2,04
N	$24,37 + B01 * 2233,69 + B03 * -6195,02 + B05 * 5659,42 + B07 * -3274,37 + B08A * 2692,46 + B09 * -1042,06$	0,27	304,53	17,3
DAPT	$452,43 + B01 * 18485,48 + B03 * -32503,79 + B05 * 27126,18 + B06 * -10198,27 + B07 * -18784,19 + B08A * 20790,97$	0,18	435,26	119
AT	$577,34 + B01 * 16080,16 + B03 * -50676,66 + B04 * 17269,12 + B05 * 22337,62 + B07 * -5511,86 + B08 * 5388,43 + B09 * -7487,3$	0,13	409,32	78,7
AB	$187,72 + B02 * -5788,95 + B03 * 6973,77 + B06 * -2749,49 + B07 * 1274,18$	0,21	318,96	22,7
BSASM	$155,33 + B01 * 4427,43 + B05 * -2138,22 + B06 * -4395,85 + B07 * 2334,81 + B09 * 3433,57$	0,27	333,17	27,2
BSAST	$11802,83 + B01 * 133043,46 + B06 * -152996,32 + B07 * 64172,32 + B09 * 60146,73$	0,24	592,89	1277
AM > 5	$8,7 + B05 * -79,54 + B07 * 69,9 + B08A * -53,98 + B11 * -72,89 + B12 * 186,11$	0,23	73,04	0,59
N > 5	$9,11 + B01 * 1694,35 + B03 * -7538,24 + B04 * 4539,71 + B05 * 1304,1 + B06 * 923,39 + B08A * -392,37 + B09 * -842,7 + B10 * 7382,79 + B12 * -614,1$	0,22	247,8	6,90
AMed	$4,25 + B01 * 235,66 + B02 * -175,44 + B03 * -370,17 + B04 * 311,97 + B07 * 48,96 + B08 * 50,81 + B08A * -79,8 + B11 * 71,53 + B12 * -141,8$	0,36	69,96	0,50
DAPMed	$11,96 + B02 * 453,26 + B03 * -475,21 + B07 * -109,45 + B08 * 139,06 + B10 * -772,70$	0,27	116,46	1,12

Quadro B3: Melhores modelos de regressão (OLS) entre múltiplas bandas do Sentinel 2 e variáveis de cobertura vegetal medidas por LiDAR segundo o valor AIC.

Variável	Fórmula	r ²	AIC	RMSE
Altura mínima	$-5,21 + B01 * 132,37 + B02 * 278,25 + B03 * -193,52 + B04 * -319,91 + B08 * 56,06 + B08A * -59,91 + B09 * 57,27 + B10 * 472,18 + B11 * 172,92 + B12 * -299,69$	0,34	933,53	0,96
Altura máxima 10	$23,23 + B01 * 370,42 + B02 * -368,13 + B05 * -430,45 + B06 * -81,3 + B08A * 109,23 + B09 * 114,73 + B10 * 1212,3 + B11 * -104,41$	0,6	1434,4	1,97
Altura máxima 5	$26,29 + B01 * 314,16 + B02 * -344,84 + B04 * 249,54 + B05 * -524,72 + B08A * 62,12 + B09 * 121,19 + B10 * 888,54 + B11 * -125,6$	0,61	1381,7	1,82
Altura média	$22,7 + B01 * 343,41 + B02 * -337,75 + B04 * 468,93 + B05 * -458,4 + B08 * 51,96 + B09 * 129,78 + B10 * 448,2 + B12 * -312,94$	0,57	1350,1	1,74
Altura relativa	$99,8 + B01 * 16,43 + B05 * -9,17 + B06 * 8,16 + B07 * -4,67 + B10 * 13,64 + B12 * -3,41$	0,04	-465,5	0,12
Altura stdev	$4,21 + B02 * -80,57 + B03 * 176,01 + B04 * -163,72 + B05 * -142,81 + B07 * 20,73 + B08 * -47,08 + B08A * 32,76 + B09 * 18,08 + B10 * 302,89 + B11 * -96,9 + B12 * 222,8$	0,47	499,34	0,49
Cobertura 10	$0,97 + B01 * 25,58 + B02 * -31,96 + B04 * 39,52 + B05 * -25,89 + B08 * 4,11 + B09 * 7,89 + B10 * 25,52 + B11 * 7,83 + B12 * -40,54$	0,5	-364,33	0,14
Cobertura 1	$1,0 + B02 * 1,67 + B03 * -1,59 + B04 * -1,59 + B05 * 0,95 + B09 * -0,28 + B11 * 0,44 + B12 * -0,92$	0,1	-2380	0,01
Cobertura 5	$0,92 + B01 * 8,28 + B03 * -16,23 + B04 * 16,08 + B06 * 1,52 + B08 * 1,89 + B08A * -2,0 + B09 * 1,73 + B11 * 5,47 + B12 * -18,08$	0,31	-1091	0,05
Cobertura relativa	$0,67 + B01 * 4,57 + B02 * -8,71 + B03 * 9,77 + B04 * 10,13 + B05 * -10,22 + B06 * 2,21 + B07 * -1,21 + B09 * 1,5 + B12 * -3,25$	0,22	-1355	0,03
Densidade	$-0,03 + B01 * -8,39 + B02 * 22,43 + B03 * -8,22 + B04 * -18,85 + B05 * 15,39 + B08A * -2,0 + B09 * -2,75 + B10 * -12,32 + B11 * 5,04$	0,61	-1078	0,05
Ecos	$4,95 + B02 * -27,23 + B03 * -48,6 + B04 * 78,85 + B08A * 5,28 + B10 * -28,12 + B11 * -16,45$	0,36	-226,4	0,17
Profundidade	$27,05 + B01 * 216,42 + B02 * -612,02 + B03 * 205,98 + B04 * 713,35 + B05 * -431,25 + B07 * -78,56 + B08A * 126,01 + B09 * 68,95 + B11 * -176,18$	0,56	1284,2	1,57

Quadro B4: Melhores modelos de regressão (OLS) entre múltiplas bandas do Sentinel 2 e variáveis de estrutura vegetal medidas por LiDAR segundo o valor de r^2 ajustado.

Variável	Fórmula	r^2	AIC	RMSE
Altura mínima	$-5,67 + B01 * 142,15 + B02 * 247,84 + B03 * -147,67 + B04 * -324,22 + B06 * -18,33 + B08 * 60,64 + B08A * -54,4 + B09 * 61,61 + B10 * 471,19 + B11 * 178,25 + B12 * -311,45$	0,35	934,31	0,93
Altura máxima 10	$23,33 + B01 * 364,77 + B02 * -440,45 + B03 * 200,21 + B05 * -529,4 + B06 * -59,7 + B08 * -43,55 + B08A * 132,51 + B09 * 123,24 + B10 * 1219,26 + B11 * -104,18$	0,6	1436,3	1,96
Altura máxima 5	$26,42 + B01 * 340,3 + B02 * -449,58 + B03 * 225,8 + B04 * 208,38 + B05 * -603,76 + B08A * 59,2 + B09 * 131,65 + B10 * 879,29 + B11 * -125,52$	0,62	1381,9	1,81
Altura média	$21,52 + B01 * 353,17 + B02 * -347,56 + B04 * 430,59 + B05 * -436,29 + B07 * -85,29 + B08 * 57,29 + B08A * 72,05 + B09 * 127,95 + B10 * 494,85 + B12 * -312,36$	0,57	1351,1	1,73
Altura relativa	$99,8 + B01 * 16,43 + B05 * -9,17 + B06 * 8,16 + B07 * -4,67 + B10 * 13,64 + B12 * -3,41$	0,04	-465,5	0,12
Altura stdev	$4,21 + B02 * -80,57 + B03 * 176,01 + B04 * -163,72 + B05 * -142,81 + B07 * 20,73 + B08 * -47,08 + B08A * 32,76 + B09 * 18,08 + B10 * 302,89 + B11 * -96,9 + B12 * 222,8$	0,47	499,34	0,49
Cobertura 10	$1,05 + B01 * 26,25 + B02 * -32,38 + B04 * 40,94 + B05 * -32,21 + B06 * 5,45 + B07 * -5,73 + B08 * 6,6 + B09 * 7,97 + B10 * 26,56 + B11 * 6,61 + B12 * -37,14$	0,5	-364,1	0,14
Cobertura 1	$0,99 + B02 * 1,46 + B03 * -1,82 + B04 * -1,47 + B05 * 0,99 + B08 * 0,05 + B09 * -0,35 + B10 * 0,8 + B11 * 0,39 + B12 * -0,81$	0,1	-2379	0,01
Cobertura 5	$0,92 + B01 * 8,28 + B03 * -16,23 + B04 * 16,08 + B06 * 1,52 + B08 * 1,89 + B08A * -2,0 + B09 * 1,73 + B11 * 5,47 + B12 * -18,08$	0,31	-1091	0,05
Cobertura relativa	$0,65 + B01 * 5,19 + B02 * -9,25 + B03 * 9,19 + B04 * 11,58 + B05 * -10,11 + B06 * 1,83 + B08A * -0,94 + B09 * 1,38 + B11 * 1,14 + B12 * -5,44$	0,22	-1355	0,03
Densidade	$-0,04 + B01 * -8,56 + B02 * 22,54 + B03 * -8,41 + B04 * -20,44 + B05 * 18,11 + B06 * -2,16 + B07 * 1,89 + B08A * -2,5 + B09 * -2,63 + B10 * -13,15 + B11 * 4,93$	0,61	-1077	0,05
Ecos	$4,69 + B02 * -26,08 + B03 * -40,08 + B04 * 83,19 + B05 * -13,04 + B06 * 6,91 + B07 * -7,45 + B08A * 7,74 + B09 * 5,15 + B10 * -25,38 + B11 * -16,14$	0,36	-224,2	0,17
Profundidade	$27,05 + B01 * 216,42 + B02 * -612,02 + B03 * 205,98 + B04 * 713,35 + B05 * -431,25 + B07 * -78,56 + B08A * 126,01 + B09 * 68,95 + B11 * -176,18$	0,56	1284,2	1,57

Quadro B5: Melhores modelos simplificados de regressão (OLS) entre múltiplas bandas do Sentinel 2 e variáveis de cobertura vegetal medidas em campo.

Variável	Fórmula	r ²	AIC	RMSE
AM	$6,06 + B01 * 119,91 + B03 * -371,56 + B04 * 210,73 + B08 * 15,19$	0,38	55,62	0,47
C	$104,12 + B04 * 173,82 + B05 * -137,54 + B06 * 69,3 + B07 * 88,64 + B08A * -110,76 + B12 * -79,23$	0,45	45,89	0,39
DAPM	$4,53 + B06 * -309,64 + B08A * 184,28 + B09 * 167,22$	0,14	155,72	2,12
N	$21,33 + B03 * -2204,22 + B05 * 3482,5 + B07 * -3123,13 + B08A * 2480,59$	0,23	304,51	18,4
DAPT	$653,11 + B05 * 6885,96 + B07 * -17520,07 + B08A * 13707,62$	0,13	434,71	128
AT	$529,3 + B05 * 4525,13 + B07 * -1714,64$	0,08	407,4	88,6
AB	$205,32 + B03 * 1436,98 + B06 * -852,06$	0,2	317,62	23,7
BSASM	$147,2 + B01 * 2835,74 + B06 * -5671,13 + B07 * 3091,68 + B09 * 3085,1$	0,26	332,87	27,9
BSAST	$12284,41 + B01 * 90508,34 + B06 * -48062,78$	0,23	591,37	1325
AM > 5	$8,68 + B05 * -75,58 + B07 * 12,58 + B11 * -83,55 + B12 * 192,62$	0,21	72,86	0,61
N > 5	$57,1 + B01 * 1214,34 + B03 * -2755,53 + B04 * 1545,36$	0,13	246,84	8,11
AMed	$5,07 + B01 * 124,27 + B03 * -325,53 + B04 * 146,12 + B08 * 16,08$	0,31	68,6	0,57
DAPMed	$4,32 + B02 * 425,63 + B03 * -485,91 + B07 * -89,11 + B08 * 118,72$	0,25	116,58	1,16

Quadro B6: Melhores modelos simplificados de regressão (OLS) entre múltiplas bandas do Sentinel 2 e variáveis de cobertura vegetal medidas com LiDAR.

Variável	Fórmula	r ²	AIC	RMSE
Altura mínima	$-1,15 + B05 * -100,58 + B10 * 515,86 + B11 * 166,45 + B12 * -318,28$	0,26	968,8	1,00
Altura máxima 10	$23,91 + B05 * -612,82 + B08A * 64,44 + B10 * 1125,92$	0,55	1470,2	2,10
Altura máxima 5	$34,15 + B05 * -474,22 + B08A * 65,69 + B10 * 747,67 + B11 * -100,48$	0,58	1403,0	1,90
Altura média	$40,27 + B05 * -339,58 + B08 * 56,8 + B11 * -125,51$	0,51	1388,5	1,86
Altura relativa	-	-	-	-
Altura stdev	$3,74 + B05 * -105,53 + B08A * 14,87 + B10 * 206,01 + B11 * -74,1 + B12 * 158,52$	0,4	532,6	0,52
Cobertura 10	$1,98 + B05 * -17,54 + B08 * 4,34 + B12 * -19,83$	0,45	-340,7	0,14
Cobertura 1	-	-	-	-
Cobertura 5	$1,17 + B08 * 0,82 + B12 * -6,62$	0,22	-1054,4	0,05
Cobertura relativa	$0,8 + B05 * -2,03 + B12 * -1,7$	0,12	-1322,7	0,03
Densidade	$-0,3 + B05 * 8,98 + B08A * -1,73 + B11 * 3,95$	0,53	-1024,8	0,05
Ecos	$4,12 + B03 * -56,99 + B04 * 57,96 + B08A * 5,74 + B11 * -13,74$	0,34	-217,9	0,17
Profundidade	$24,3 + B04 * 460,94 + B05 * -366,44 + B08A * 64,25 + B11 * -148,88$	0,52	1311,4	1,66

Quadro B7 - Valores das variáveis de estrutura para cada parcela de campo. AM (Altura média das árvores amostradas), C (proporção de cobertura da parcela), DAPM (diâmetro à altura do peito médio das árvores amostradas), N (número de árvores amostradas na parcela), DAPT (diâmetro à altura do peito total das árvores amostradas), AT (altura total das árvores amostradas), AB (área basal total das árvores da parcela pela área de coleta).

Parcela	AM (m)	C (%)	DAPM (cm)	N	DAPT (cm)	AT (m)	AB (m ² /ha)
1	4,24	99,34	9,05	51	461,72	216,00	103,41
13	5,49	99,38	7,33	55	403,38	302,15	33,37
15	4,10	99,45	6,38	71	452,68	291,00	47,98
1p10	4,71	99,23	7,98	45	359,01	212,14	28,59
1p25	5,22	99,03	12,33	17	209,61	88,69	30,37
1p26	3,78	99,42	9,30	46	427,89	173,89	42,88
1p35	6,07	99,36	15,38	23	353,63	139,52	43,84
1p41	5,17	99,31	9,20	48	441,55	247,92	65,52
1p5	4,76	99,25	8,93	54	482,00	257,04	61,64
1p7	4,69	99,30	7,24	60	434,34	281,39	47,41
1p9	5,25	99,17	8,09	88	711,48	461,57	109,63
20	3,28	96,60	5,61	127	712,10	416,46	36,11
22	3,44	97,80	5,61	59	330,73	202,79	27,27
26	4,66	99,25	8,19	25	204,78	116,40	18,93
29	3,99	99,91	6,97	62	431,85	247,36	39,87
2p2	4,37	99,55	10,12	30	303,54	131,04	28,95
2p25	5,71	99,23	8,53	47	400,72	268,16	23,76
2p28	4,43	99,11	5,79	113	654,04	500,29	45,88
2p5	4,69	99,42	8,75	64	559,71	300,33	54,17
2p7	4,94	99,40	8,86	41	363,34	202,64	29,11
2p8	4,50	99,20	5,07	30	152,19	135,14	7,39
3	3,63	98,07	7,60	57	433,28	206,88	103,41
3p1	4,65	99,45	6,32	52	328,85	242,05	18,19
3p16	5,17	99,21	7,29	63	459,14	325,61	40,21
3p3	5,02	99,36	10,99	50	549,65	251,21	107,58
3p6	5,51	99,25	6,43	80	514,48	440,51	30,65
3p60	4,34	99,29	5,92	75	443,63	325,55	32,21
3p69	3,84	99,41	5,67	44	249,61	168,83	15,57
4	4,26	99,51	0,43	57	24,54	243,00	22,97
4p3	4,86	99,36	8,45	66	557,57	320,58	76,11
4p4	4,46	99,43	8,76	54	472,99	241,02	70,54
6	4,94	99,30	8,78	46	404,04	227,40	28,55
8	4,95	99,41	7,56	35	264,62	173,40	19,04
9p28	3,86	99,01	7,10	61	432,98	235,52	25,55

Quadro B7 - Valores das variáveis de estrutura por parcela de campo. BSASM (Biomassa seca acima do solo média), BSAST (biomassa seca acima do solo total), AM > 5 (altura média das árvores maiores de cinco metros), N > 5 (número de árvores maiores de cinco metros), AMed (altura mediana das árvores amostradas), DAPMed (diâmetro à altura do peito mediano das árvores amostradas).

Parcela	BSASM (kg)	BSAST (kg)	AM > 5 (m)	N > 5	AMed (m)	DAPMed (cm)
1	117,37	5986,09	7,35	12	3,60	3,34
13	36,17	1989,44	7,22	27	4,80	4,46
15	48,61	3451,60	7,27	18	3,00	3,18
1p10	32,36	1456,31	6,79	16	4,10	4,87
1p25	89,13	1515,21	6,81	10	5,27	5,51
1p26	36,85	1695,32	6,38	11	3,12	6,04
1p35	158,37	3642,46	9,09	11	4,68	7,96
1p41	102,46	4918,16	8,00	18	4,39	4,46
1p5	60,62	3273,55	7,35	18	4,10	3,39
1p7	52,52	3151,35	7,33	15	4,10	4,01
1p9	65,10	5729,22	6,90	43	4,68	4,20
20	10,03	1273,81	6,09	14	3,00	3,82
22	22,52	1328,81	7,00	9	3,00	2,87
26	41,40	1034,99	7,50	8	3,60	4,62
29	32,89	2039,08	9,27	11	3,00	3,18
2p2	37,80	1134,01	7,09	8	4,10	6,02
2p25	24,09	1132,24	6,71	28	5,85	7,48
2p28	19,49	2201,82	6,51	31	4,10	3,66
2p5	46,24	2959,45	7,02	22	4,10	4,78
2p7	33,97	1392,72	6,76	16	4,68	6,37
2p8	11,24	337,07	6,11	9	4,39	3,58
3	44,52	2537,79	7,37	7	3,60	3,50
3p1	15,12	786,05	6,68	17	4,10	3,79
3p16	37,45	2359,18	7,86	23	4,33	3,38
3p3	113,76	5688,16	7,11	20	4,68	4,78
3p6	20,02	1601,38	6,96	46	5,85	4,14
3p60	23,34	1750,24	6,23	20	4,10	4,39
3p69	17,58	773,60	7,87	9	2,93	3,53
4	1,99	113,53	6,65	12	4,20	0,05
4p3	64,87	4281,27	7,02	24	4,10	3,89
4p4	62,59	3379,71	6,87	16	4,10	4,01
6	22,67	1042,66	6,11	22	4,80	6,27
8	26,76	936,62	6,76	15	4,80	5,10
9p28	14,17	864,08	5,85	8	3,74	5,25