

Articles

Evaluation of digital height models of vegetation derived from LiDAR data

Avaliação de modelos digitais de altura da vegetação derivados de dados LiDAR

Leticia da Cunha Mose Ferreira¹ 
Cristiane Nunes Francisco¹ 
Pedro José Farias Fernandes¹ 

¹Universidade Federal Fluminense , Niterói, RJ, Brazil

ABSTRACT

Vegetation Digital Height Models (DHM) are generated by subtracting the Digital Terrain Model (DTM) from the Digital Surface Model (DSM), both of which can be constructed through the interpolation of three-dimensional coordinates obtained from Light Detection and Ranging (LiDAR) data. However, these models may contain errors associated with terrain topography and the interpolation method employed. Therefore, this study aimed to evaluate DHM generated using four interpolation methods applied to LiDAR data through statistical analyses designed to identify differences among the models and assess the influence of topography. The models were generated using Inverse Distance Weighting, Spline, Natural Neighbor, and Ordinary Kriging interpolators for the Jacaré River basin, located in Niterói, Rio de Janeiro, Brazil. The study area covers approximately 6 km², with an elevation range of up to 400 m, and is predominantly covered by dense tropical rainforest. The results showed that the DHM exhibited similar data distributions, with no statistically significant differences among the models according to ANOVA. Furthermore, no significant linear relationship was observed between vegetation height variance and the analyzed topographic variables (slope and elevation). However, the Spline interpolator showed a greater occurrence of extreme values, whereas Ordinary Kriging exhibited greater surface smoothing and lower data variability. Thus, under the environmental conditions of the study area, characterized by dense forest cover and rugged terrain, the Natural Neighbor and Inverse Distance Weighting interpolators are recommended for generating DHM from LiDAR data, as they showed a lower tendency to under- or overestimate values.

Keywords: Laser scanning; Spatial interpolation; Deterministic methods; Geostatistics; Forest structure

RESUMO

Os Modelos Digitais de Altura (MDA) da vegetação são gerados pela subtração entre o Modelo Digital de Superfície (MDS) e o Modelo Digital de Terreno (MDT), que podem ser construídos pela interpolação das coordenadas tridimensionais obtidas por sistemas de perfilamento a laser (LiDAR). No entanto, os modelos interpolados podem apresentar erros provenientes tanto da topografia quanto do tipo de interpolador utilizado. Desta forma, o artigo objetiva avaliar os MDA gerados por quatro métodos de interpolação aplicados a dados LiDAR, através de análises estatísticas, a fim de verificar as diferenças entre os modelos, bem como a influência da topografia. Foram gerados modelos com os interpoladores Inverso da Distância Ponderada, Spline, Vizinho Natural e Krigagem Ordinária para a bacia do rio Jacaré, situada no município de Niterói, RJ, com 6 km², amplitude altimétrica de 400 m e predominância de floresta ombrófila densa. Os resultados indicaram que os MDA apresentaram similaridade quanto à distribuição dos dados, sem diferenças estatisticamente significativas entre os modelos segundo a ANOVA. Também não foi encontrada relação linear significativa entre a variância da altura da vegetação e as variáveis topográficas analisadas (declividade e hipsometria). No entanto, o interpolador Spline apresentou maior ocorrência de valores extremos, enquanto a Krigagem Ordinária apresentou maior suavização da superfície, com redução da variabilidade dos dados. Dessa forma, para as condições da área estudada, floresta densa em relevo acidentado, os interpoladores Inverso da Distância Ponderada e Vizinho Natural são recomendados para a geração de MDA com dados LiDAR, por apresentarem menor tendência à subestimação ou superestimação dos dados.

Palavras-chave: Perfilamento a laser; Interpolação espacial; Métodos determinísticos; Geoestatística; Estrutura florestal

1 INTRODUCTION

The development of remote sensing, especially through airborne active sensors, has enabled the collection of three-dimensional terrain data, allowing the generation of information for various applications, such as mapping topography, land cover, and forest variables (Locks; Matricardi, 2019; Giongo et al., 2010). In this context, ALS (Airborne Laser Scanning) systems stand out for their speed and precision in data acquisition, as they integrate, on an airborne platform, the Global Navigation Satellite System (GNSS), responsible for the aircraft's positioning; the Inertial Navigation System (INS), which determines its attitude angles; and the Light Detection and Ranging (LiDAR) system (Brandalize, 2002).

LiDAR is a high-precision data acquisition technology based on the emission of laser pulses and the recording of the return time of the reflected signal. The distance

between the sensor and the surface is estimated from the difference in time between the emission of the pulse and the return of the reflected signal after it is intercepted by the ground or by objects above it, such as vegetation and buildings (Brandalize, 2002; Verma et al., 2016). In areas with dense vegetation cover, the first LiDAR return corresponds, in most cases, to the highest points of the canopy, while the last return is highly likely to correspond to the ground surface, although it may still contain records of low-strata vegetation (Stonge et al., 2003).

When mounted on airborne platforms, this sensor generates millions of points with X, Y, and Z coordinates, forming a point cloud that represents the terrain's topography and the natural and artificial features present on the surface surveyed (Giongo et al., 2010; Brandalize, 2002). Spatial interpolation is then applied to this point cloud to estimate the elevation of unsampled locations based on values measured at neighboring points, generating continuous surfaces (Burrough; McDonnell, 1998).

Interpolators can be divided into two groups based on the method used in the estimation: geostatistical methods, which model spatial variability based on the autocorrelation among sampled points, and deterministic methods, which produce surfaces using mathematical functions based on the proximity between samples (Camargo et al., 2003). The choice of interpolation method, therefore, has a direct influence on the quality of the generated models.

The models derived from these procedures include the Digital Elevation Model (DEM), which corresponds to the digital representation of the elevation of the Earth's surface; the Digital Terrain Model (DTM), generated from points corresponding to the bare terrain surface; and the Digital Surface Model (DSM), which includes, in addition to elevation, natural and artificial objects superimposed on the terrain (Burrough; McDonnell, 1998). Thus, by subtracting the DTM from the DSM, the Digital Height Model (DHM) is obtained, which represents the height of surface features present on the terrain surface (Hyypä et al., 2001).

Because they represent vegetation height, DHM allow the characterization of the horizontal and vertical structure of forests (Garcia-Gutierrez et al., 2014; Wezyk, 2012), which contributes to studies aimed at volume estimation, species classification, individual tree identification, and forest growth monitoring (Wang; Glenn, 2008; Andersen et al., 2006). Thus, they constitute an important alternative for reducing the costs of forest inventories, especially in large areas, due to the better cost-benefit ratio in terms of time and resources invested (Giongo et al., 2010).

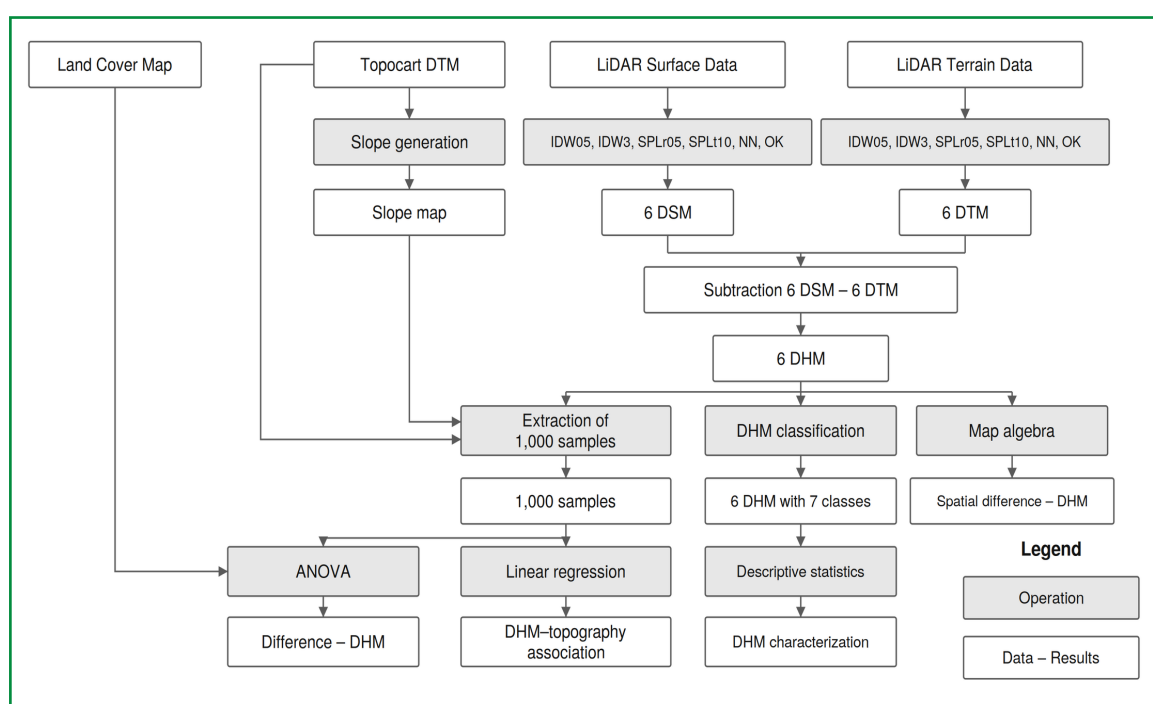
However, methodological challenges related to LiDAR data modeling persist, particularly regarding filtering algorithms and interpolation methods. In areas with gentle terrain, the results tend to be satisfactory; however, on sloped or vegetated terrain, limitations are observed in the separation between the surface and the ground, particularly due to the influence of the lower vegetation layer (Stereńczak et al., 2016; Schimalski et al., 2009). Furthermore, the interpolation method ranks among the primary sources of error in LiDAR-derived models (Căţeanu; Ciubotaru, 2019; Luo et al., 2015; Hodgson; Bresnahan, 2004), which underscores the importance of understanding its influence on DHM in areas with rugged topography and dense vegetation cover.

Thus, based on the hypothesis that the variations observed in the DHM values obtained from LiDAR data result both from the interpolation method used and from the topographic and land cover characteristics of the study area, this study aims to analyze the DHM generated by models derived from four interpolation methods in order to identify differences among the models and assess the influence of topography and land cover on the results.

2 MATERIALS AND METHODS

The methodological procedures are represented in the flowchart (Figure 1). The first step consisted of collecting LiDAR data to generate DSM and DTM using four interpolation methods with different parameterizations. Subsequently, six DHM were

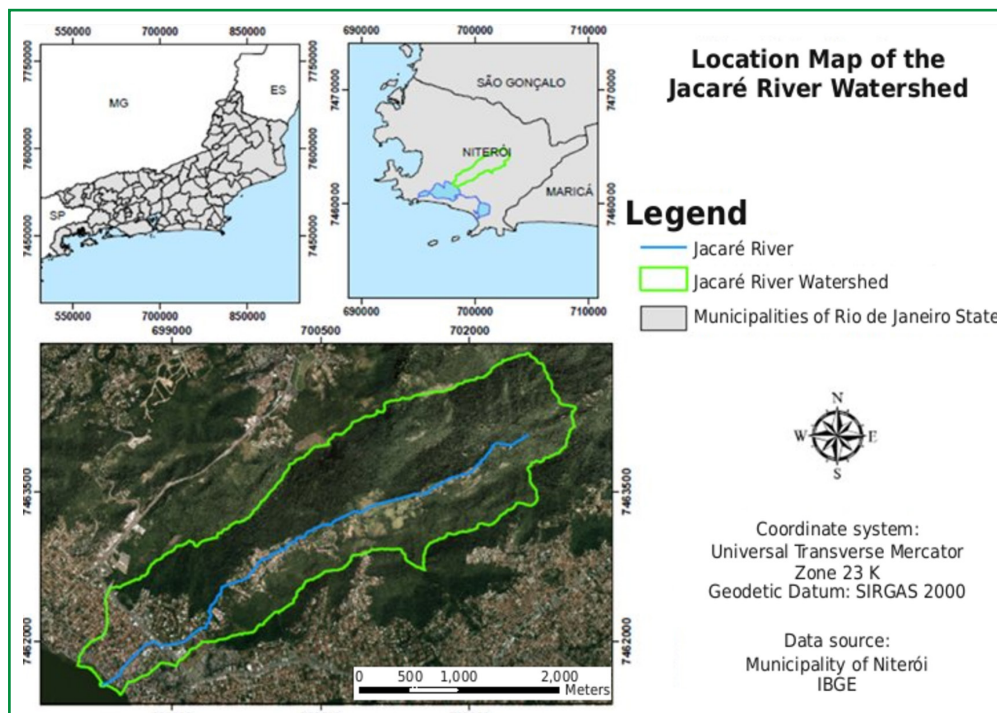
generated by subtracting the DTM from the DSM corresponding to each interpolator. In the next step, statistical analyses of the DHM were performed, including descriptive statistics to characterize the data distribution, analysis of variance (ANOVA) to detect differences among the models, map algebra for the spatial analysis of differences among the DHM, and, finally, linear regression to assess the association between the DHM and topographic variables.



Source: Authors (2024)

The study area corresponds to the Jacaré River watershed (JRW), covering an area of 6 km², located in the Coastal Region of the municipality of Niterói, in the state of Rio de Janeiro (Figure 2). The area is predominantly covered by dense secondary rainforest situated on steep slopes, with an elevation range of up to 400 m. The high-density urban area is concentrated on the plain located in the lower reaches of the basin near the river mouth; however, moving inland along the main channel, human occupation becomes sparse until rural properties predominate in the upper reaches (BAPTISTA, 2009).

Figure 2 – Location of the study area, Jacaré River watershed, Niterói, RJ



Source: Authors (2024)

2.1 LiDAR Data Survey

The LiDAR data used in this study were acquired and processed by TOPOCART, the company responsible for the survey conducted in the municipality of Niterói in September 2014. The flight covered an area of approximately 140 km², with an average altitude of 1,900 m, an average speed of 120 km/h, a density of 5 points m⁻², and a scan angle ranging between 10° and 15°, depending on terrain conditions and flight altitude (TOPOCART, 2014).

The raw data, consisting of three-dimensional coordinates, were processed using the Terrascan-Terrasolid system on the Microstation V8 platform, where filtering procedures were also performed to separate terrain points from points corresponding to surface targets (TOPOCART, 2014). Based on the two files derived from this operation, the DTM and DSM were generated with a spatial resolution of 0.5 m using four interpolation methods available in the ArcGIS 3D Analyst extension: Inverse Distance Weighting, Spline, Natural Neighbor, and

Ordinary Kriging. The selection of these interpolators aimed to compare different deterministic methods with a geostatistical method widely used in the modeling of continuous surfaces from altimetric data (Stereńczak et al., 2020; Căţeanu; Ciubotaru, 2020; Chen; Li, 2019; Santos et al., 2017; Tan; Xu, 2014).

2.2 Model Generation

The Inverse Distance Weighting method, a deterministic approach, is based on spatial proximity, assuming that the closer the points are, the greater their influence on the estimate, with greater weight assigned to the nearest samples (Li; Heap, 2014, 2008). The main parameter to be configured is the power value, which controls the influence of neighboring samples on value estimation. To compare different levels of influence, two power values were applied: 0.5 (IDW05) and 3 (IDW3).

Spline (SPL) is a deterministic method that estimates values using a mathematical function that minimizes the total curvature of the surface, resulting in a smooth surface (Li; Heap, 2014, 2008). There are two formulations that influence smoothing: regularized (r), which creates a smooth and gradually varying surface, and tensioned (t), which creates a rougher surface (ESRI, 2024). To compare surfaces with different smoothing levels, the regularized configuration with power 0.5 (SPLr05) and the tensioned configuration with power 10 (SPLt10) were used. In accordance with the system recommendations, a lower value was adopted for the regularized model because it results in a smoother surface, whereas a higher value was used for the tensioned model because it produces a rougher surface (ESRI, 2024).

Natural Neighbor is a deterministic method that selects the subsets of samples closest to the point to be estimated (natural neighbors), based on the construction of Voronoi or Delaunay diagrams, to calculate the weights of the influence areas (Li; Heap, 2014, 2008; Watson, 1999). Since this interpolator does not require additional parameterization, the system's default configuration was adopted, in which estimation is automatically performed based on the natural neighborhood of the sample points.

Ordinary Kriging, a geostatistical method, assumes that the distance and direction between samples reflect the spatial correlation used to explain surface

variation. The investigation of the spatial behavior of the phenomenon is carried out through samples represented in a semivariogram, from which the most appropriate estimation model is selected to generate the output surface (Camargo et al., 2003). Semivariogram analysis indicated that the exponential model provided the best fit and was therefore considered the most suitable model for interpolating altimetric data compared to other kriging models (Pavlova, 2017).

Finally, based on the difference between the DSM and the DTM generated by the interpolation methods described above and summarized in Table 1, six DHM were obtained and subsequently subjected to statistical analyses.

Table 1 – Parameters of the DTM and DSM interpolation models

Method	Abrev.	Type	Parameter	Result
Inverse Distance Weighting	IDW05	-	P = 0.5	Lower weighting of nearby points
	IDW3	-	P = 3	Greater influence of nearby points.
Spline	SPLr05	Regularized	P = 0.5	Smoother surface
	SPLt10	Tension	P = 10	Rougher surface
Natural Neighbor	NN	-	-	-
Kriging	OK	Ordinary	Exponential	Best model fit

Source: Authors (2024)

2.3 Statistical Analyses

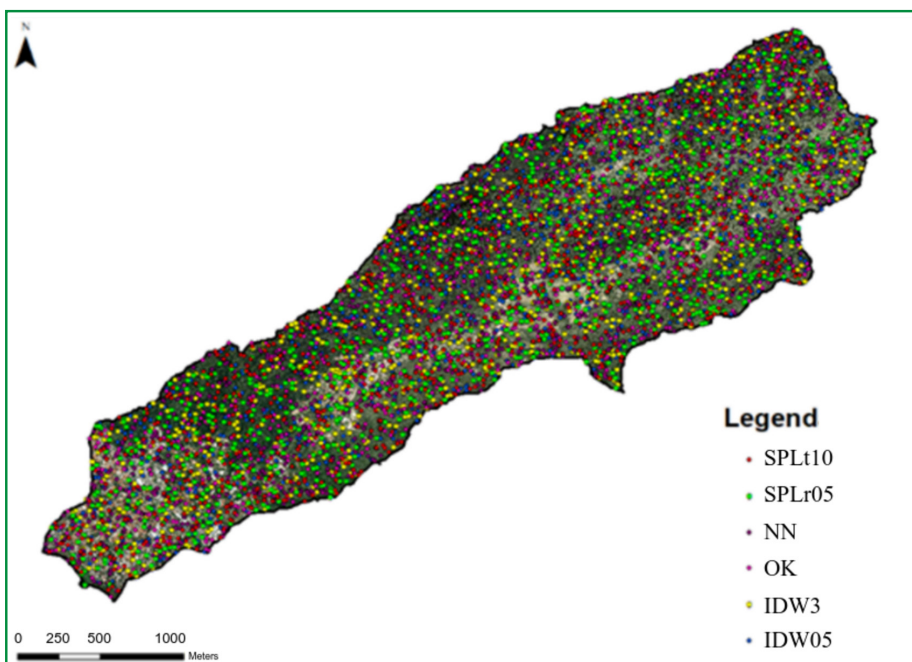
To calculate descriptive statistics and create histograms, seven classes were defined in the DHM, associated with the land cover types identified in high-resolution images: 0 m – bare soil; 0–0.5 m – herbaceous vegetation; 0.5–1 m – herbaceous and subshrub vegetation; 1–5 m – subshrub and shrub vegetation; 5–12 m – low tree vegetation; 12–20 m – medium-height tree vegetation; and > 20 m – tall tree vegetation.

Subsequently, to analyze the spatial discrepancies among the estimated values, the pixel-by-pixel variance among the six DHM models was calculated, considering the values estimated by each interpolator in each cell. The result was a variance image classified into seven elevation classes representing the variability among the models.

The next step involved an ANOVA to verify the existence of significant differences among the DHM means and, consequently, identify discrepancies in elevation values among the six interpolated models. A stratified ANOVA was also performed for the five land cover classes based on the mapping proposed by Bohrer and Charmont (2016): tall vegetation, low-to-medium vegetation, ground vegetation, urban area, and bare soil.

For the ANOVA, statistical inference based on probability sampling was used, employing simple random sampling of 1,000 height measurements from each DHM (Figure 3), ensuring equal probability of selection for all population members. Samples were selected independently for each model to avoid direct overlap between groups. The independence of observations was assumed based on random sampling, although the possibility of residual spatial autocorrelation is acknowledged. Residual normality was considered satisfied due to the large sample size, according to the Central Limit Theorem. Furthermore, ANOVA is considered robust to moderate deviations from homoscedasticity, particularly given the large number of random samples and equal sample size among groups.

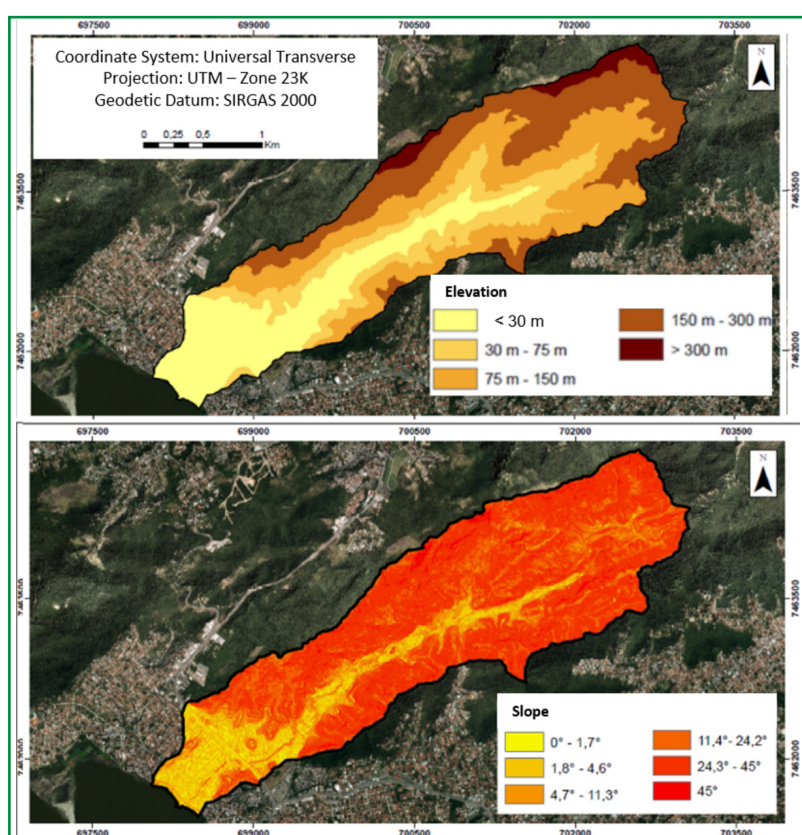
Figure 3 – Distribution of sample points from the interpolation methods, Jacaré River basin, Niterói, RJ



Source: Authors (2024)

Finally, to verify the influence of topography on DHM values, a simple linear regression analysis was performed between vegetation height variance and the slope and hypsometric variables (Figure 4). These datasets were generated from the LiDAR-derived DTM with a spatial resolution of 0.5 m (TOPOCART, 2014). A total of 1,000 sample points were extracted, excluding urban areas to avoid the influence of building height values. These analyses were also performed in the R statistical environment.

Figure 4 – Hypsometric and slope map of the Jacaré River basin, Niterói, RJ



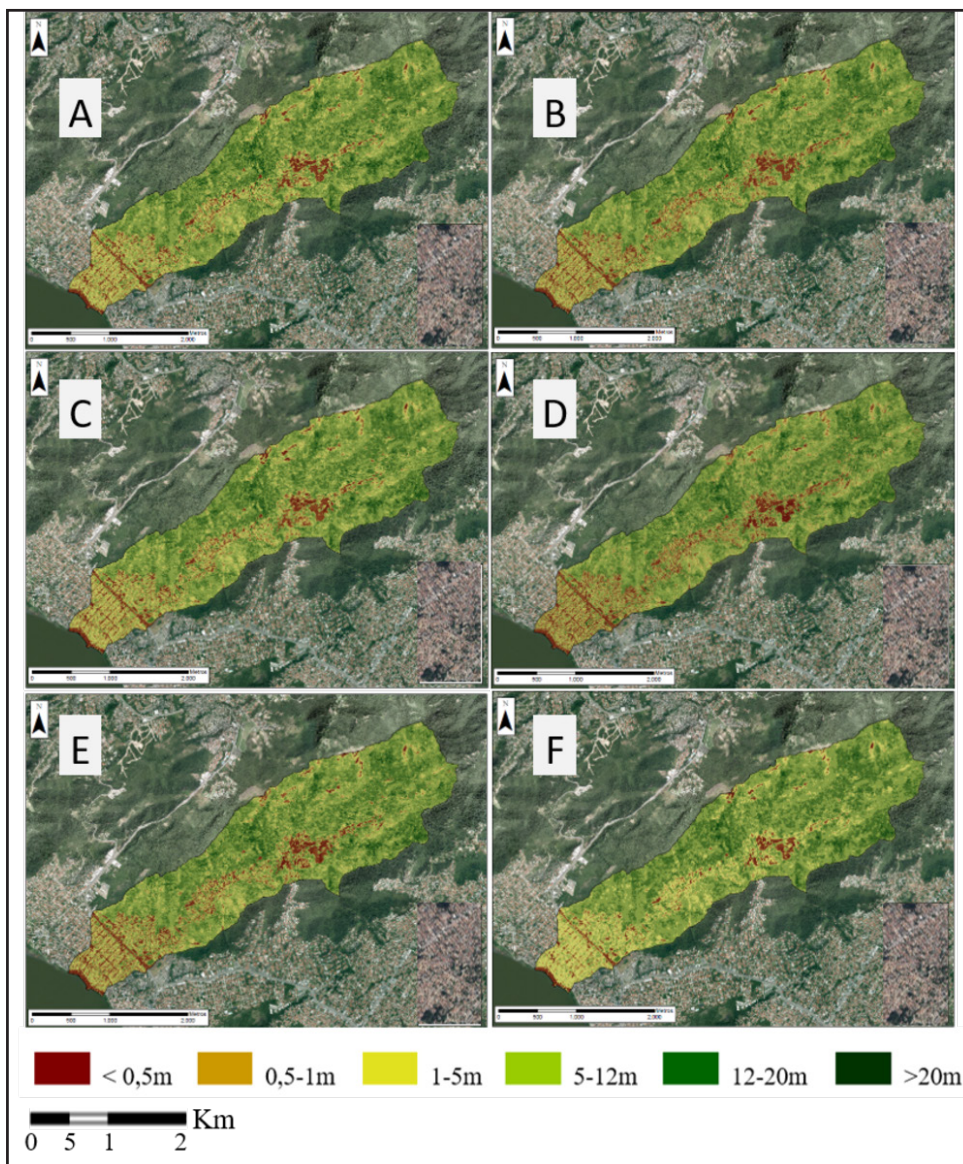
Source: Authors (2024)

3 RESULTS

A visual analysis of the six DHM, derived from the subtraction of the models generated using the Inverse Distance Weighting, Spline, Natural Neighbor, and Ordinary Kriging interpolators, reveals a similar pattern in the distribution of height classes (Figure 5). The lowest values correspond to roads and herbaceous vegetation located along the

main channel and on the plain, followed by buildings in the urban area (up to 5 m). On forested slopes, classes ranging from 5 to 20 m predominate, whereas the tallest areas, exceeding 20 m, correspond to tree vegetation located on the best-preserved slopes of the Jacaré River basin. However, the Ordinary Kriging model shows a distinct pattern in the two lowest height classes: while the class up to 0.5 m occupies a smaller area, the 1–5 m class occupies a larger area compared to the other models.

Figure 5 – DHM generated using different interpolation methods in the Jacaré River basin, Niterói, RJ

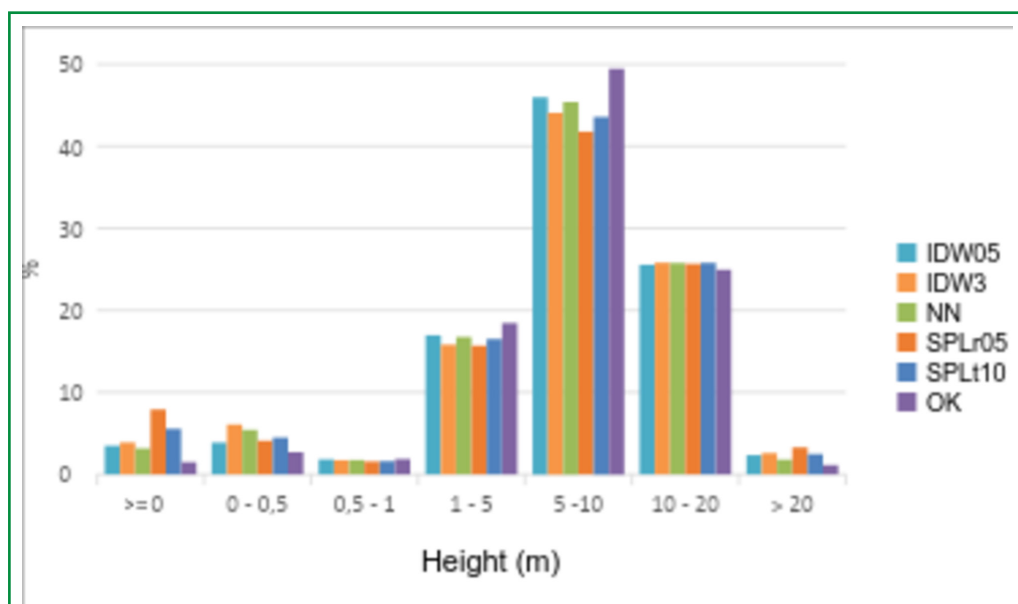


Source: Authors (2024)

In where: (A) IDW05; (B) IDW3; (C) NN; (D) SPLr05; (E) SPLt10; and (F) OK.

Analysis of the histogram corroborates the previous observations, indicating that the DHM exhibit similar frequency distributions, with a predominance of intermediate height classes (Figure 6). Vegetation cover exceeding 5 m in height predominates in all models, representing more than 70% of the study area, followed by the 1–5 m class, which accounts for approximately 20%, whereas the remaining classes (up to 1 m and greater than 20 m) together account for approximately 10%. However, some specific characteristics can be observed: (a) the data distribution in the Inverse Distance Weighting and Natural Neighbor models is highly similar, with only minor redistributions among the lower classes; (b) the Spline models show a greater tendency to generate extreme values, particularly in the regularized configuration; and (c) the Ordinary Kriging model exhibits greater surface smoothing, with a concentration of values in the intermediate classes.

Figure 6 – Histogram of DHM according to interpolation methods

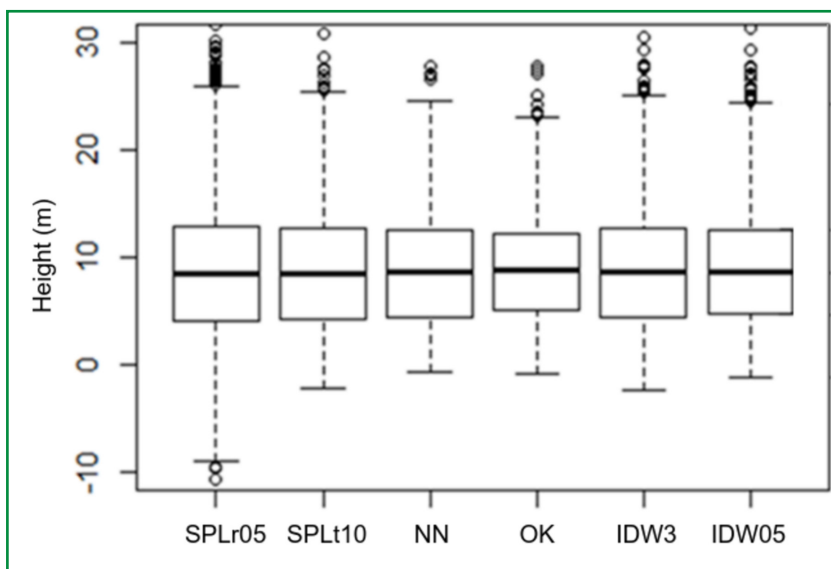


Source: Authors (2024)

The box plot analysis also supports the previous observations. The data distribution in the DHM is similar, with a median close to 9 m, a first quartile around 4 m, and a third quartile around 13 m (Figure 7). However, Ordinary Kriging shows values

more concentrated within the intermediate intervals of the distribution, resulting in a lower standard deviation (4.9 m). Regarding the lower and upper limits, with the exception of the Regularized Spline 0.5 model, which exhibited a greater number of maximum outliers (reaching 424 m) and was the only model presenting minimum outliers (reaching -300 m), the remaining models presented values close to 0 m and 25 m, respectively. Kurtosis further reinforces this difference: in the Regularized Spline 0.5 model, the distribution was leptokurtic (kurtosis = 32.5), indicating heavy tails and a greater occurrence of outliers, which increased the standard deviation to 6.4 m, whereas in the remaining models this value was approximately 5.5 m.

Figure 7 – Box plot of DHM according to interpolation method



Source: Authors (2024)

The comparison of DHM using ANOVA confirms the previous analyses, since the p-value of 0.35, which is higher than the adopted significance level (0.05), indicates no statistically significant difference among the representative sample groups of each model. The ANOVA applied to land cover classes—tall vegetation ($p = 0.85$), short- to medium-height vegetation ($p = 0.90$), and bare soil ($p = 0.46$)—also indicated no significant differences among the models. However, the low-growing vegetation class showed a significant difference ($p = 0.008$), whereas the urban area presented a borderline value ($p = 0.056$), indicating that the models differ more substantially within these classes.

This result is likely associated with the differences observed in the lower elevation classes, particularly those below 0.5 m, especially in the Ordinary Kriging DHM, which exhibited lower frequency values, and in the Spline models, which showed a higher concentration of values within these classes. The greater occurrence of outliers in the Spline models within these classes of limited vertical range may also have contributed to this result.

An analysis of the differences among the six DHM reveals that the smallest discrepancies were obtained using the Inverse Distance Weighting, Natural Neighbor, and Ordinary Kriging interpolators, with differences ranging from 22 to 30 m (Table 2). The largest differences are associated with the Spline interpolator, reaching values between 93 and 420 m, particularly in the Regularized Spline model, which may be related to the outliers generated by this interpolator. These extreme values reflect differences resulting from the extrapolative behavior of the Spline method, since they occur exclusively in operations involving this interpolator.

Table 2 – Height differences among the DHM according to interpolation method

Interpolators	IDW3 (m)	NN (m)	SPLr05 (m)	SPLt10(m)	OK (m)
IDW05	22,3	26,5	417,4	97,0	28,3
IDW3		25,6	417,2	96,5	22,3
NN			419,8	93,5	30,6
SPLr05				393,2	419,6
SPLt10					95,5

Source: Authors (2024)

The DHM variance map (Figure 8) complements the results by enabling analysis of the spatial pattern of variability among the models. Values below 2 m predominate, accounting for approximately 70% of the study area. The highest differences, exceeding 25 m and representing less than 0.1% of the area, are located on forested slopes and concentrated around points that possibly correspond to trees with heights overestimated by the Spline interpolator. Considering the measures of central tendency of the DHM, with a median of 9 m and first and third quartiles

of 4 m and 13 m, respectively, variance values below 2 m indicate low dispersion among the models relative to the structural range of vegetation in the study area. Therefore, the predominance of this magnitude can be considered expected and methodologically acceptable.

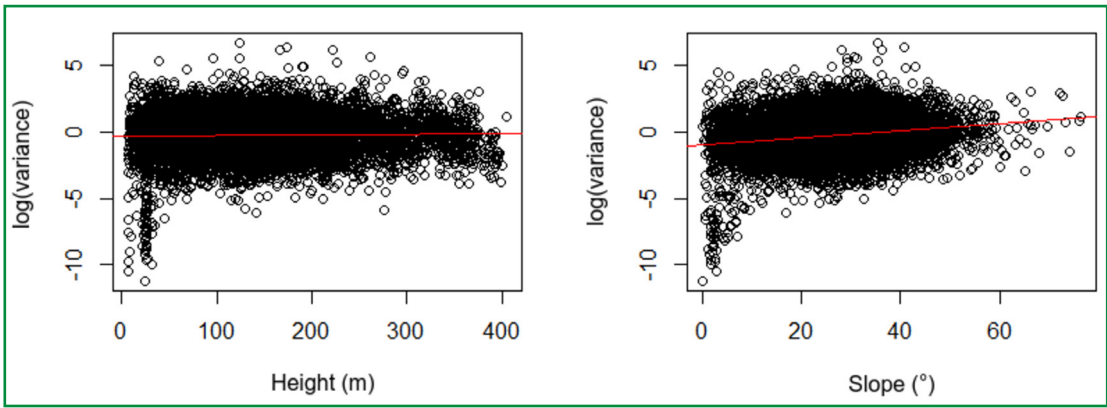
Figure 8 – Variance map of the DHM generated by the IDW, SPL, NN, and OK interpolators, Jacaré River basin, Niterói, RJ



Source: Authors (2024)

Finally, the regression between the variance of the DHM and the topographic variables is presented in the scatter plots (Figure 9). No linear relationship was observed between model variance and hypsometry or slope, since the coefficients of determination (R^2) were very low, corresponding to 0.036 and 0.0009, respectively. This result indicates that these topographic variables have low explanatory power regarding vegetation height variance, demonstrating the absence of a significant linear relationship between these variables.

Figure 9 – Scatter plots of DHM variance in relation to hypsometry and slope in the Jacaré River basin, Niterói, RJ



Source: Authors (2024)

Based on the obtained results, a summary table (Table 3) was prepared highlighting the main advantages and limitations of the evaluated interpolators considering their performance in generating DHM.

Table 3 – Summary of the main advantages and limitations of the evaluated interpolators

Interpolador	Vantagens	Limites
Inverse Distance Weighting	MDA similar to that of the other interpolators; estimates constrained to the range of observed values.	-
Natural Neighbor		
Ordinary Kriging	MDA similar to that of the other interpolators.	Tendency to produce a smoother surface, reducing spatial variability.
Spline		Estimates may fall outside the range of observed values, generating extreme values and a greater range of differences compared to the other models.

Source: Authors (2024)

4 DISCUSSIONS

The interpolators produced models whose behavior was consistent with their respective theoretical and methodological assumptions. The Spline models exhibited

greater data dispersion, as evidenced by the leptokurtic distribution, higher standard deviation values, larger differences in relation to the other interpolators, and a greater number of minimum and maximum outliers. Because this interpolator extrapolates beyond the limits of the input dataset (ESRI, 2024), it is more prone to generating extreme values, which explains the discrepancies observed among the models. Thus, although it produces smoother surfaces, the Spline method exhibits lower adherence to the observed values, making it more suitable for the rapid generation of maps intended to represent general characteristics of the variable (LI; HEAP, 2008). In the present study, this behavior was evident in the outliers that exceeded the upper limit of the DHM by more than 30 times.

The Ordinary Kriging model, in turn, exhibited a lower standard deviation and a higher frequency of values concentrated within the intermediate intervals of the distribution, which is consistent with its premise of minimizing variance. This characteristic results in lower data variability and a tendency to underestimate maximum values while overestimating minimum values, consequently concentrating the data within intermediate classes (OLIVER; WEBSTER, 2014).

Studies evaluating interpolation methods applied to LiDAR data for DTM generation have reported results similar to those obtained in the present study, indicating that, in general, interpolators exhibit a high degree of similarity (Stereńczak et al., 2020; Căţeanu; Ciubotaru, 2020; Santos et al., 2017). In similar contexts, these studies were conducted in mountainous areas covered by dense vegetation using LiDAR datasets with densities ranging from 5 to 22 points m^{-2} . The convergence among interpolation methods may be associated with the high point density of the LiDAR data, as highlighted by Gonçalves (2006), since higher sampling densities tend to reduce the influence of the interpolator on the final surface.

Stereńczak et al. (2020), based on the analysis of DTM generated using different interpolators and filtering methods with a density of 6 points m^{-2} , obtained RMSE values ranging from 0.19 to 0.23 m and concluded that the evaluated interpolation

methods are capable of producing very similar error levels after appropriate calibration. Similarly, Căţeanu and Ciubotaru (2020), when testing nine interpolators, including those evaluated in the present study, obtained RMSE (Root Mean Square Error) values ranging from 0.11 to 0.28 m in models with a spatial resolution of 0.5 m and concluded that, despite their relatively similar performance, Spline-based models are visually more appealing due to their smoothing effect.

Santos et al. (2017) observed very small differences among Kriging, Natural Neighbor, and Delaunay Triangulation, with differences of less than 5 m across most of the study area, suggesting Delaunay Triangulation for interpolation due to its lower computational requirements. Chen and Li (2019) further state that simple interpolation methods, such as Inverse Distance Weighting and Natural Neighbor, show strong potential for generating DTM from large datasets such as LiDAR data due to their lower computational and memory costs. Conversely, in the study conducted by Ismail et al. (2016), although the interpolator results were similar and applied to an area covered by tree vegetation, Kriging exhibited greater accuracy than Inverse Distance Weighting and Spline.

The ANOVA results indicated no significant differences among the DHM, except for the low-growing vegetation and urban classes. This result is also associated with the interaction between interpolation methods and the characteristics of these land cover classes. In low-growing vegetation areas, the reduced vertical amplitude favors more abrupt transitions with adjacent classes, which become amplified by the subtraction between the DTM and the DSM. In urban areas, the presence of abrupt surface discontinuities, such as buildings and walls, also contributes to the amplification of intraclass differences.

Thus, the observed ANOVA results may be associated with the greater influence of small altimetric variations within these classes, amplified by the interpolation process. In this context, because Spline estimates values beyond the sample range, it tends to amplify differences, whereas Inverse Distance Weighting and Kriging, by

assigning weights based on the spatial relationship between samples and estimated points, tend to constrain estimates within the observed range, thereby reducing the magnitude of differences among the models in these more sensitive classes (ISMAIL et al., 2016).

Likewise, the absence of a significant linear relationship between DHM variance and the hypsometric and slope variables indicates that, in the study area, topography did not exert a significant influence on vegetation height variance among the models. This result partially contrasts with the literature, which generally indicates an association between more complex terrain and greater interpolation uncertainty. Ismail et al. (2016), in tests conducted with LiDAR data across three slope classes, observed an increase in DTM RMSE values with increasing slope steepness, ranging from 0.5 m to 1.4 m between the 0–5° and 11–15° classes, respectively, for the three interpolators evaluated—the same interpolators analyzed in the present study. Similarly, Pavlova (2017) also reported greater accuracy in flat areas compared to sloped terrain, with RMSE values below 0.6 m in a study conducted using data derived from topographic maps at scales of 1:100,000 and 1:50,000.

Stereńczak et al. (2016) state that the results obtained from the analysis of more than 100 DTM do not differ from previous studies demonstrating that interpolation errors tend to increase with slope steepness. However, this study, conducted in a mountainous area covered by dense vegetation, concluded that understory vegetation height exerts a greater influence than slope in locations where it occurs in significant amounts. Since the Jacaré River basin is predominantly covered by dense rainforest, the low correlation observed with topographic variables may be associated with the fact that understory structure exerts a greater influence on model variability than the topographic variables analyzed in the present study.

5 CONCLUSIONS

This study generated six DHM derived from different interpolation methods applied to the generation of DTM and DSM in a forested area with rugged terrain. Data analysis indicated that the DHM generated by the four interpolation methods—Inverse Distance Weighting, Natural Neighbor, Spline, and Ordinary Kriging—showed a high degree of similarity in data distribution, since, in addition to presenting similar descriptive statistics, the analysis of variance did not indicate significant differences among the representative sample groups of each model. Furthermore, no significant linear relationship was observed between vegetation height variance and the slope and hypsometric variables of the study area, suggesting that, under the evaluated conditions, the DTM results were not significantly influenced by these topographic variables.

However, despite the similarity indicated by the statistical measures, the Spline interpolator exhibited a greater occurrence of extreme values, whereas Ordinary Kriging showed greater surface smoothing, resulting in reduced data variability. Considering that the DHM corresponds to the difference between two interpolated models, namely the DSM and the DTM, it is recommended to use interpolators with a lower tendency to underestimate or overestimate values.

Thus, considering the conditions of the Jacaré River basin—dense forest cover, rugged terrain, and high sampling density—Natural Neighbor and Inverse Distance Weighting are the most suitable interpolators for generating DTM from LiDAR data. In addition to avoiding excessive smoothing and extrapolation effects, these methods produce continuous and consistent estimated surfaces while requiring lower methodological complexity and fewer computational resources.

As a methodological limitation, it should be noted that the analysis was conducted in a single study area characterized by dense rainforest, rugged terrain, and high LiDAR point density. Although these conditions are representative of applications in mountainous tropical environments, the results should not be automatically generalized to areas with different vegetation cover patterns, topographic conditions, or sampling densities. Therefore, future studies conducted under different physiographic conditions are recommended.

REFERÊNCIAS

- BAPTISTA, J. V.; FERNANDES, V. F. Alterações ambientais em decorrência do processo de urbanização acelerada na bacia hidrográfica do Rio Jacaré, Niterói, RJ. In: SEMINÁRIO DE RECURSOS HÍDRICOS DA BACIA HIDROGRÁFICA DO PARAÍBA DO SUL, 2., 2009, Taubaté. **Proceedings...** Taubaté: IPABHi, 2009. p. 537–544.
- BRANDALIZE, A. A. Perfilamento a laser: comparação com métodos fotogramétricos. In: SIMPÓSIO BRASILEIRO DE GEOMÁTICA, 1., 2002, Presidente Prudente. **Proceedings...** São Paulo, 2002.
- BOHRER, C. B. A.; CHARMONT, R. **Mapeamento da cobertura vegetal e do solo. In: Plano estratégico para a renaturalização do Rio Jacaré.** v. II. Niterói: Prefeitura Municipal de Niterói, 2016. p. 327–345.
- BURROUGH, P. A.; MCDONNELL, R. A. **Principles of geographical information systems.** Oxford: Oxford University Press, 1998.
- CAMARGO, E.; FUCKS, S.; CÂMARA, G. Análise espacial de superfícies. In: FUCKS, S. et al. (org.). **Análise espacial de dados geográficos.** São José dos Campos: INPE, 2003.
- CHEN, C.; LI, Y. A fast global interpolation method for digital terrain model generation from large LiDAR-derived data. **Remote Sensing**, v. 11, n. 11, p. 1324, 2019.
- CĂȚEANU, M.; CIUBOTARU, A. Accuracy of ground surface interpolation from airborne laser scanning (ALS) data in dense forest cover. **ISPRS International Journal of Geo-Information**, v. 9, n. 4, 2020.
- ESRI. **How spline works.** ArcGIS Resource Center. Available at: https://help.arcgis.com/en/arcgisdesktop/10.0/help/index.html#/How_Spline_works/009z00000078000000/. Accessed in: 20 set. 2024.
- GARCIA-GUTIERREZ, J.; GONZALEZ-FERREIRO, E.; RIQUELME-SANTOS, J. C.; MIRANDA, D.; DIEGUEZ-ARANDA, U.; NAVARRO-CERRILLO, R. M. Evolutionary feature selection to estimate forest stand variables using LiDAR. **International Journal of Applied Earth Observation and Geoinformation**, v. 26, p. 119–131, 2014.
- GLONGO, M. et al. LiDAR: princípios e aplicações florestais. **Pesquisa Florestal Brasileira**, v. 30, n. 63, p. 231–240, 2010.
- GONÇALVES, G. Analysis of interpolation errors in urban digital surface models created from LiDAR data. In: INTERNATIONAL SYMPOSIUM ON SPATIAL ACCURACY ASSESSMENT IN NATURAL RESOURCES AND ENVIRONMENT SCIENCES, 7., 2006, Lisboa. **Proceedings...** Lisboa, 2006. p. 160–168.
- HODGSON, M. E.; BRESNAHAN, P. Accuracy of airborne LiDAR-derived elevation: empirical assessment and error budget. **Photogrammetric Engineering & Remote Sensing**, v. 70, n. 3, p. 331–339, 2004.
- HYYPPÄ, J. et al. A segmentation-based method to retrieve stem volume estimates from 3D tree height models produced by laser scanners. **IEEE Transactions on Geoscience and Remote Sensing**, v. 39, n. 5, p. 969–975, 2001.
- ISMAIL, Z. et al. Evaluating error of LiDAR-derived DEM interpolation for vegetation area. **International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences**, v. XLII-4/W1, p. 141–150, 2016.

JENSEN, J. R. **Sensoriamento remoto do ambiente: uma perspectiva em recursos terrestres**. São José dos Campos: Parêntese, 2009.

LI, J.; HEAP, A. D. Spatial interpolation methods applied in the environmental sciences: a review. **Environmental Modelling & Software**, v. 53, p. 173–189, 2014.

LI, J.; HEAP, A. D. **A review of spatial interpolation methods for environmental scientists**. Geoscience Australia, Record 2008/23, 2008.

LOCKS, C. J.; MATRICARDI, E. A. T. Estimativa de impactos da extração seletiva de madeiras na Amazônia utilizando dados LiDAR. **Ciência Florestal**, v. 29, n. 3, p. 481–495, 2019.

LUO, S. et al. Estimation of wetland vegetation height and leaf area index using airborne laser scanning data. **Ecological Indicators**, v. 48, p. 550–559, 2015.

OLIVER, M. A.; WEBSTER, R. A tutorial guide to geostatistics: computing and modelling variograms and kriging. **CATENA**, v. 113, p. 56–69, 2014.

PACHECO, A. P. et al. Classificação de pontos LiDAR para a geração do MDT. **Boletim de Ciências Geodésicas**, v. 17, n. 3, p. 438–471, 2011.

PAVLOVA, A. I. Analysis of elevation interpolation methods for creating digital elevation models. **Optoelectronics, Instrumentation and Data Processing**, v. 53, n. 2, p. 171–177, 2017.

SANTOS, E. G. et al. Avaliação de modelo digital de terreno (MDT) e declividade utilizando diferentes interpoladores a partir de dados LiDAR. In: CONGRESSO BRASILEIRO DE CARTOGRAFIA, 27., 2017, Rio de Janeiro. **Proceedings...** Rio de Janeiro: SBC, 2017. p. 721–725.

STEREŃCZAK, K. et al. Comparison of various algorithms for DTM interpolation from LiDAR data in dense mountain forests. **European Journal of Remote Sensing**, v. 49, p. 599–621, 2016.

STONGE, B.; TREITZ, P.; WULDER, M. A. Tree and canopy height estimation with scanning lidar. **Remote Sensing of Forest Environments: Concepts and Case Studies**. 2003.

TAN, Q.; XU, X. Comparative analysis of spatial interpolation methods: an experimental study. **Sensors & Transducers**, v. 165, n. 2, p. 155–163, 2014.

TOPOCART. **Levantamento aerofotogramétrico e perfilamento a laser: Niterói-RJ**. Relatório técnico. Niterói: Secretaria Municipal de Urbanismo e Mobilidade, 2014.

VERMA, N. K. et al. Comparison of canopy volume measurements derived from high-resolution imagery and LiDAR. **Remote Sensing**, v. 8, p. 388, 2016.

WANG, C.; GLENN, N. F. A linear regression method for tree canopy height estimation using airborne LiDAR data. **Canadian Journal of Remote Sensing**, v. 34, n. 2, p. 217–227, 2008.

WATSON, D. The natural neighbor series manuals and source codes. **Computers & Geosciences**, v. 25, n. 4, p. 463–466, 1999.

WEZYK, P. The integration of terrestrial and airborne laser scanning technologies in forest parameter retrieval. **Ambiência**, v. 8, p. 533–548, 2012.

Authorship Contribution

1 Leticia da Cunha Mose Ferreira

Geographer, MSc. in Biosystems Engineering

<https://orcid.org/0009-0001-6634-0919> • leticiamose@gmail.com

Contribution: Investigation; Data curation; Formal analysis; Validation; Writing – original draft

2 Cristiane Nunes Francisco

Geographer, PhD. in Geosciences, Professor

<https://orcid.org/0000-0002-8688-9810> • cristianenf@id.uff.br

Contribution: Conceptualization; Methodology; Supervision; Writing – review & editing

3 Pedro José Farias Fernandes

Geographer, PhD. in Geography

<https://orcid.org/0000-0002-1862-2401> • pj_fernandes@id.uff.br

Contribution: Data curation; Formal analysis; Methodology; Validation

How to quote this article

FERREIRA, L. C. M.; FRANCISCO, C. N.; FERNANDES, P. J. F. Evaluation of digital height models of vegetation derived from LiDAR data. **Ciência Florestal**, Santa Maria, v. 36, e89194, p. 1-23, 2026. DOI 10.5902/1980509889194. Available from: <https://doi.org/10.5902/1980509889194>. Accessed in: day month abbr. year.

Data Availability Statement:

Datasets related to this article will be available upon request to the corresponding author.

Evaluators in this article:

Prof. Dr. Fernando Coelho Eugenio, *Section Editor*

Editorial Board:

Prof. Dr. Cristiane Pedrazzi, *Editor-in-Chief*

Prof. Dr. Dalton Righi, *Associate Editor*

Miguel Favila, *Managing Editor*