COMMENTARY

Revised: 10 July 2023

Which metrics derived from airborne laser scanning are essential to measure the vertical profile of ecosystems?

W. Daniel Kissling^{1,2} Vifang Shi^{1,2}

¹Institute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, Amsterdam, The Netherlands

²LifeWatch ERIC, Virtual Laboratory and Innovations Centre (VLIC), University of Amsterdam Faculty of Science, Amsterdam, The Netherlands

Correspondence

W. Daniel Kissling, Institute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, Amsterdam, The Netherlands. Email: w.d.kissling@uva.nl

Funding information

European Commission, Grant/Award Number: 101003553 and 101060639; Netherlands eScience Center, Grant/ Award Number: ASDI.2016.014

Editor: Christine Howard

Abstract

In a recent perspective (Diversity and Distributions, 29, 39-50), '10 variables' were proposed to measure vegetation structure from airborne laser scanning (ALS) for assessing species distributions and habitat suitability. We worry about this list because the variables predominantly represent variation in vegetation height, the vertical variability of vegetation biomass is insufficiently captured, and variables of vegetation cover are ill-defined or not ecosystem agnostic. We urge for a better defined, more comprehensive and more balanced list, and for assessing which information from ALS point clouds is truly essential to measure the major dimensions of 3D vegetation structure within and across ecosystems and animal habitats. We think that the currently proposed 'list of 10 ALS metrics' is premature and that researchers and stakeholders should be cautious in adopting this list.

KEYWORDS

Ecosystem structure, Essential Biodiversity Variable, Habitat condition, LiDAR metrics, Point clouds, Structural complexity

Quantifying the 3D structure of vegetation is of fundamental importance for understanding, modelling, monitoring and predicting biodiversity and ecosystems (Bakx et al., 2019; Davies & Asner, 2014; MacArthur & MacArthur, 1961; Pereira et al., 2013). In particular, Light Detection And Ranging (LiDAR) point clouds from ALS surveys - typically conducted at national, regional or landscape scales have emerged as crucial datasets for monitoring and modelling biodiversity (Bakx et al., 2019; Müller & Vierling, 2014; Valbuena et al., 2020). To extract relevant information on the vertical profile of ecosystems, the raw LiDAR point clouds from ALS surveys need to be further processed, for example, by calculating statistical properties (i.e. features, or metrics) of the point cloud within infinite square cells or cubes (Meijer et al., 2020). Examples are the mean height of vegetation points within a grid cell or the density of vegetation points within a particular horizontal layer (e.g. within cubes). Since the number of LiDAR metric names, metric definitions and calculation methods used in ecological papers is vast and often confusing (see e.g. Appendix 3 of Bakx et al., 2019), we welcome the idea from Moudrý et al. (2023) to start a deeper discussion about which ALS metrics of vegetation structure should be easily accessible to researchers and stakeholders. Here, we aim to contribute to this discussion by highlighting that the proposed 'list of 10 ALS metrics' has several flaws.

First, many of the metrics proposed by Moudrý et al. (2023) for example, all metrics of vegetation height (maximum height, mean height, and height percentiles) - are highly correlated with each other (see Figure 1a for an example from a country-wide dataset of the Netherlands). This multi-collinearity extends beyond the height metrics themselves because several of the other variables proposed by Moudrý et al. (2023), such as those related to cover and vertical variability, are also highly correlated with height (Figure 1a). For instance, variables measuring vegetation cover and density (i.e. the number of laser returns) in specific layers (e.g. above 3 m or within 5-20 m) or some of the vertical variability

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. Diversity and Distributions published by John Wiley & Sons Ltd.



FIGURE 1 Covariation among 25 metrics of vegetation structure derived from a country-wide, 10 m resolution airborne laser scanning dataset across the whole Netherlands. (a) Correlation matrix (Spearman's Rank correlation coefficients r) of metrics grouped into vegetation height, cover and vertical variability. Coloured boxes behind metric abbreviations show congruence with variables from Moudrý et al. (2023) (see legend). All height metrics (upper left corner) are highly correlated (r>0.8), and several metrics of vegetation cover and vertical variability are also highly correlated with vegetation height. (b) Axes from a Principal Component Analysis (PCA) explaining in total ~75% of variation among the 25 metrics. PCA axis 1 (Dim1, explaining 55% variation) is mainly characterized by vegetation height, with percentiles (e.g. Hp75 and Hp95), averages (Hmean and Hmedian) and maximum values (Hmax) of height making the strongest contributions. PCA axis 3 (Dim3, explaining 8% variation) is mainly characterized by vertical variability, with skewness (Hskew), coefficient of variation (Coeff_var_z) and kurtosis (Hkurt) making the strongest contributions. Note that the standard deviation of vegetation height (Hsd) and the foliage height diversity (Entropy_z) indicated with the two red arrows are grouped by Moudrý et al. (2023) into the vertical variability class, even though they strongly correlate with vegetation height (Dim1) and hence do not represent vertical variability (Dim3). See figure legend for metric abbreviations. Metric calculations are explained in Kissling et al. (2023), and additional details of the PCA are provided in the Appendix A of Kissling et al. (2022).

= density of vegetation points above 3 m

BR above

metrics (e.g. foliage height diversity based on the Shannon-Wiener index) are highly correlated with vegetation height metrics (for an example, see Figure 6a,b in Kissling et al., 2022). Hence, the list of ALS metrics from Moudrý et al. (2023) predominantly represents variation in vegetation height whereas it is short in variables that capture other dimensions of vegetation structure (see following paragraphs below). We believe that a more comprehensive set of variables is needed to better reflect the multidimensionality of the environmental (i.e. ecosystem structure) space, especially with variables that can deviate from simple statistical or allometric scaling relationships with vegetation height (West et al., 2009). This may be particularly important in the context of essential biodiversity variables (EBVs) where a minimum set of measurements, complementary to one another, is needed to capture the major dimensions

of biodiversity change (Pereira et al., 2013; Schmeller et al., 2017). We suggest that the list of ALS metrics needs to be more balanced to better capture the key dimensions of ecosystem structure (Valbuena et al., 2020). Testing a larger range of metrics and their co-variation, independence, deviation and unexplained (residual) variance relative to vegetation height can be a first step (Figure 1). Moreover, identifying the relative contributions of independent variables and how metrics capture the concentration of the vertical point distribution (i.e. dispersion relative to location) may be particularly informative (Valbuena et al., 2017). Applying dimensionality reduction methods such as a Principal Component Analysis (PCA) can further help to identify which metrics represent different dimensions of ecosystem structure in a particular dataset (see e.g. Kissling et al., 2022).

Second, the list of 10 ALS metrics from Moudrý et al. (2023) proposes only two variables for measuring the vertical variability of vegetation structure, namely the standard deviation of vegetation height (SD) and the foliage height diversity (FHD) based on the Shannon-Wiener index (sometimes referred to as entropy). However, both variables are highly collinear with height metrics (Figure 1b) and hence do not represent independent variation of the vertical variability of vegetation structure. The SD is often correlated with the mean, while the coefficient of variation (CV) - also known as relative standard deviation (i.e. the ratio of the standard deviation to the mean) - is not. The CV of vegetation height is thus a better choice than the SD of vegetation height (Figure 1b). In addition, metrics such as the skewness or kurtosis of vegetation height were not included in the list of Moudrý et al. (2023), but they can capture important ecological aspects of the vertical variability of vegetation (Figure 1b, see also Figure 6 in Kissling et al., 2022). Skewness and kurtosis (as well as the CV) of vegetation height can therefore represent variation of vertical variability that is independent from vegetation height and cover (see Dim3 in Figure 6a of Kissling et al., 2022). Previous analyses with L-moments from ALS data (i.e. statistics summarizing the vertical probability density distribution of point clouds) also confirm that CV and skewness are independent metrics that contain information about ecosystem structure in Boreal forests (e.g. uneven tree size classes), i.e. variation that is not already captured by vegetation height (Valbuena et al., 2017). We therefore recommend that ALS metrics of vertical variability of vegetation should not be restricted to SD and FHD, but also include skewness, kurtosis and CV of vegetation height. Moreover, some authors emphasize that FHD is not designed to describe continuous variables, and hence suggest replacing FHD with Lorenz curves and Gini coefficients (Valbuena et al., 2021). Whether these alternative metrics do indeed provide a better description of the vertical variability of vegetation than FHD needs to be tested with a range of datasets and in different ecosystems, not only in forests.

Third, several definitions of vegetation cover metrics from Moudrý et al. (2023) are not ecosystem agnostic or their definitions are potentially misleading, which can introduce confusion and ambiguities in metric calculations. For instance, three of the cover variables (i.e. cover of the herbaceous layer, cover of the shrub layer, and cover of the tree layer) imply a forest vegetation with herb, shrub and tree layers. However, for different ecosystems (e.g. grasslands, shrublands and forests), the top vegetation layer is not necessarily represented by trees (Figure 2a). Similarly, a middle vegetation layer is not necessarily represented by shrubs in all ecosystems. The definition and calculation of cover in specific vegetation layers therefore needs to be more explicit in terms of upper and lower height boundaries, rather than defining it as 'lowest, middle and top vegetation layer', as proposed by Moudrý et al. (2023). Our suggestion to define concrete upper and lower height boundaries makes such metrics ecosystem agnostic, i.e. independent of a particular ecosystem. While Moudrý et al. (2023) in their discussion suggest the use of at least 10 height bins for the cover variable 'Density proportions (%)', they do not suggest this for the three variables of herbaceous, shrub

Diversity and Distributions –WILEY

and tree layer cover, where it would be desirable. We therefore recommend being explicit about the definition of height bins for variables in the vegetation cover class. Recently published data papers which provide LiDAR vegetation cover metrics from country-wide ALS surveys are indeed more explicit than Moudrý et al. (2023), for example, providing vegetation cover for nine height bins across the Netherlands (Kissling et al., 2023) or 24 height bins across Denmark (Assmann et al., 2022).

Fourth, the calculation and definition of vegetation cover metrics as proposed by Moudrý et al. (2023) can introduce biases in representing the vertical distribution of vegetation biomass. For instance, density proportions can be calculated from point clouds in several ways depending on whether non-vegetation points are included or not (Figure 2b). Moudrý et al. (2023) suggest to calculate density proportions as the 'proportion of returns in a certain bin to the total number of returns', that is, using the number of all points (including non-vegetation points) in each height bin divided by the total number of points (also including non-vegetation points). Assmann et al. (2022) calculated vegetation proportions as the ratio of vegetation returns to total returns, that is, using the number of vegetation points (excluding non-vegetation points) in each height bin divided by the total number of points (also including non-vegetation points). Both ways of calculating density proportions may not correctly represent vegetation cover, with the former being potentially more biased than the latter (Figure 2c). We suggest that vegetation cover metrics for specific layers such as 'density proportions' (Moudrý et al., 2023) or 'vegetation proportions by height bin' (Assmann et al., 2022) should be calculated as the number of vegetation points in a height bin relative to the total number of vegetation points, that is, explicitly excluding non-vegetation points (Kissling et al., 2023). This will better represent the vertical distribution of vegetation density (rather than the amount of points relative to all points) and represent an independent dimension of vegetation structure compared to other vegetation metrics (e.g. Dim2 in Figure 6a of Kissling et al., 2023).

In conclusion, we urge researchers and stakeholders to be cautious in adopting the list of Moudrý et al. (2023) because it (1) is biased towards metrics representing vegetation height, (2) lacks metrics that can capture independent information on vertical variability of vegetation structure and (3) contains ambiguous information on the definition and calculation of vegetation cover metrics. We agree with Moudrý et al. (2023) that a systematic testing is needed. We therefore urge for comprehensive assessments of a large range of vegetation metrics from multiple ALS datasets to quantify their ecological relevance, statistical redundancy and independent contribution for measuring the key dimensions of vegetation structure, namely ecosystem height, ecosystem cover and ecosystem structural complexity (Valbuena et al., 2020). Due to the scale-dependence of ecological patterns and processes, LiDAR metrics might need to be calculated at different spatial resolutions (Atkins et al., 2023), or their fine-scale heterogeneity (i.e. horizontal variability) needs to be aggregated at a coarser resolution (Graham et al., 2019). The time for such assessments is ripe because (1) a large





Vegetation points / total points

Vegetation points / all vegetation points



FIGURE 2 Important aspects of defining and calculating vegetation cover metrics from airborne laser scanning (ALS) point clouds. (a) The definition of layers for vegetation cover metrics (e.g. density of vegetation points in specific height bins) should be explicit in terms of their upper and lower height boundaries because the definition of a top layer (or a middle vegetation layer) can differ among ecosystems (e.g. forest, shrubland and grassland). (b) Point cloud of a 30×30m plot showing the vertical distribution of vegetation and non-vegetation points (buildings and ground) within specific height bins. (c) Density proportions in height bins may not correctly represent vegetation cover if they are calculated as the number of all points (including non-vegetation points) in each height bin divided by the total number of points (left) or as the number of vegetation points in each height bin divided by the total number of points (right).

number of country-wide ALS datasets is now openly accessible (see overviews in Kissling et al., 2022; Moudrý et al., 2023; Stereńczak et al., 2020), (2) user friendly, free and open source software has been developed to calculate a large range of ALS metrics (Meijer et al., 2020; Roussel et al., 2020) and (3) high-throughput (reproducible and open source) workflows are now available to perform the efficient, scalable, distributed and standardized processing of multiterabyte LiDAR point clouds into ALS metrics (Kissling et al., 2022). Until such assessments are performed, proposing a list of 10 ALS metrics may seem premature.

ACKNOWLEDGEMENTS

The development of the data product used in this paper was funded by the Netherlands eScience Center (https://www.esciencecenter.nl), grant number ASDI.2016.014, through the project 'eScience infrastructure for Ecological applications of LiDAR point clouds' (eEcoLi-DAR). The development of geospatial data products derived from LiDAR is supported by LifeWatch ERIC (https://www.lifewatch.eu/), a European research infrastructure consortium with focus on biodiversity and ecosystem research. W.D.K. acknowledges funding from the European Union for the EuropaBON project (grant agreement No. 101003553) and the MAMBO project (grant agreement No. 101060639) which both aim to support the development of an EUwide framework for monitoring biodiversity and ecosystems, including the use of LiDAR for measuring and monitoring vegetation structure and habitat condition.

FUNDING INFORMATION

European Union (grant agreement 101003553 and 101060639), Netherlands eScience Center (ASDI.2016.014) and LifeWatch ERIC (no grant number available).

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

PEER REVIEW

The peer review history for this article is available at https:// www.webofscience.com/api/gateway/wos/peer-review/10.1111/ ddi.13760.

DATA AVAILABILITY STATEMENT

Data Accessibility Statement: The country-wide data of ecosystem structure (10 m resolution raster layers available as GeoTIFF files) from the third Dutch airborne laser scanning survey are openly accessible from the Zenodo repository (https://zenodo.org/recor d/6421381). The details of this dataset are described in a recent data paper (Kissling et al., 2023) and its accompanying processing workflow (Kissling et al., 2022). The raw data (LiDAR point clouds) from the third Dutch national airborne laser scanning flight campaign (AHN3) are openly accessible from the repository of the PDOK webservices (https://app.pdok.nl/ahn3-downloadpage/).

DATA ACCESSIBILITY STATEMENT

The dataset used in this paper has been described and published in a recent data paper (Kissling et al., 2023) and its accompanying processing workflow (Kissling et al., 2022). It provides country-wide data of ecosystem structure (10m resolution raster layers available as GeoTIFF files) from the third Dutch airborne laser scanning survey. The raster files from this data paper (Kissling et al., 2023) are openly accessible from the Zenodo repository (https://zenodo.org/ record/6421381). The raw data (LiDAR point clouds) from the third Dutch national airborne laser scanning flight campaign (AHN3) are openly accessible from the repository of the PDOK web services (https://app.pdok.nl/ahn3-downloadpage/).

ORCID

W. Daniel Kissling D https://orcid.org/0000-0002-7274-6755

REFERENCES

Assmann, J. J., Moeslund, J. E., Treier, U. A., & Normand, S. (2022). EcoDes-DK15: High-resolution ecological descriptors of vegetation and terrain derived from Denmark's national airborne laser scanning data set. Earth System Science Data, 14(2), 823–844. https://doi.org/10.5194/essd-14-823-2022

- Atkins, J. W., Costanza, J., Dahlin, K. M., Dannenberg, M. P., Elmore, A. J., Fitzpatrick, M. C., Hakkenberg, C. R., Hardiman, B. S., Kamoske, A., LaRue, E. A., Silva, C. A., Stovall, A. E. L., & Tielens, E. K. (2023). Scale dependency of lidar-derived forest structural diversity. *Methods in Ecology and Evolution*, 14, 708–723. https://doi.org/10.1111/2041-210X.14040
- Bakx, T. R. M., Koma, Z., Seijmonsbergen, A. C., & Kissling, W. D. (2019). Use and categorization of light detection and ranging vegetation metrics in avian diversity and species distribution research. *Diversity and Distributions*, 25, 1045–1059. https://doi.org/10.1111/ ddi.12915
- Davies, A. B., & Asner, G. P. (2014). Advances in animal ecology from 3D-LiDAR ecosystem mapping. *Trends in Ecology & Evolution, 29*(12), 681-691. https://doi.org/10.1016/j.tree.2014.10.005
- Graham, L. J., Spake, R., Gillings, S., Watts, K., & Eigenbrod, F. (2019). Incorporating fine-scale environmental heterogeneity into broadextent models. *Methods in Ecology and Evolution*, 10(6), 767–778. https://doi.org/10.1111/2041-210x.13177
- Kissling, W. D., Shi, Y., Koma, Z., Meijer, C., Ku, O., Nattino, F., Seijmonsbergen, A. C., & Grootes, M. W. (2023). Country-wide data of ecosystem structure from the third Dutch airborne laser scanning survey. *Data in Brief*, 46, 108798. https://doi.org/10.1016/j. dib.2022.108798
- Kissling, W. D., Shi, Y., Koma, Z., Meijer, C., Ku, O., Nattino, F., Seijmonsbergen, A. C., & Grootes, M. W. (2022). Laserfarm – A high-throughput workflow for generating geospatial data products of ecosystem structure from airborne laser scanning point clouds. *Ecological Informatics*, 72, 101836. https://doi.org/10.1016/j.ecoinf.2022.101836
- MacArthur, R., & MacArthur, J. W. (1961). On bird species diversity. Ecology, 42(3), 594–598.
- Meijer, C., Grootes, M. W., Koma, Z., Dzigan, Y., Gonçalves, R., Andela, B., van den Oord, G., Ranguelova, E., Renaud, N., & Kissling, W. D. (2020). Laserchicken—A tool for distributed feature calculation from massive LiDAR point cloud datasets. *SoftwareX*, 12, 100626. https://doi.org/10.1016/j.softx.2020.100626
- Moudrý, V., Cord, A. F., Gábor, L., Laurin, G. V., Barták, V., Gdulová, K., Cord, A. F., Gábor, L., Laurin, G. V., Barták, V., Gdulová, K., Malavasi, M., Rocchini, D., Stereńczak, K., Prošek, J., Klápště, P., & Wild, J. (2023). Vegetation structure derived from airborne laser scanning to assess species distribution and habitat suitability: The way forward. *Diversity and Distributions*, *29*, 39–50. https://doi. org/10.1111/ddi.13644
- Müller, J., & Vierling, K. (2014). Assessing biodiversity by airborne laser scanning. In M. Maltamo, E. Næsset, & J. Vauhkonen (Eds.), Forestry applications of airborne laser scanning: Concepts and case studies (pp. 357–374). Springer Netherlands.
- Pereira, H. M., Ferrier, S., Walters, M., Geller, G. N., Jongman, R. H. G., Scholes, R. J., Bruford, M. W., Brummitt, N., Butchart, S. H. M., & Wegmann, M. (2013). Essential biodiversity variables. *Science*, 339(6117), 277–278. https://doi.org/10.1126/science.1229931
- Roussel, J.-R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R. H., Meador, A. S., Auty, D., Coops, N., Tompalski, P., Goodbody, T. R. H., Meador, A. S., Bourdon, J.-F., de Boissieu, F., & Achim, A. (2020). lidR: An R package for analysis of airborne laser scanning (ALS) data. *Remote Sensing of Environment*, 251, 112061. https://doi. org/10.1016/j.rse.2020.112061
- Schmeller, D. S., Mihoub, J.-B., Bowser, A., Arvanitidis, C., Costello, M. J., Fernandez, M., Mihoub, J.-B., Bowser, A., Arvanitidis, C., Costello, M. J., Fernandez, M., Geller, G. N., Hobern, D., Kissling, W. D., Regan, E., Saarenmaa, H., Turak, E., & Isaac, N. J. B. (2017). An operational definition of essential biodiversity variables. *Biodiversity and Conservation*, 26(12), 2967–2972. https://doi.org/10.1007/s1053 1-017-1386-9

WILEY – Diversity and Distributions

6

- Stereńczak, K., Laurin, G. V., Chirici, G., Coomes, D. A., Dalponte, M., Latifi, H., & Puletti, N. (2020). Global airborne laser scanning data providers database (GlobALS)—A new tool for monitoring ecosystems and biodiversity. *Remote Sensing*, 12(11), 1877.
- Valbuena, R., Adnan, S., Maltamo, M., Mehtätalo, L., Ammaturo, R., & Lovejoy, T. (2021). Moving on from foliage height diversity: Determining maximum entropy in 3-dimensional variables. In Paper presented at the SilviLaser conference 2021. Technische Universität Wien.
- Valbuena, R., Maltamo, M., Mehtätalo, L., & Packalen, P. (2017). Key structural features of boreal forests may be detected directly using L-moments from airborne lidar data. *Remote Sensing of Environment*, 194, 437–446. https://doi.org/10.1016/j.rse.2016.10.024
- Valbuena, R., O'Connor, B., Zellweger, F., Simonson, W., Vihervaara, P., Maltamo, M., O'Connor, B., Zellweger, F., Simonson, W., Vihervaara, P., Maltamo, M., Silva, C. A., Almeida, D. R. A., Danks, F., Morsdorf, F., Chirici, G., Lucas, R., Coomes, D. A., & Coops, N. C. (2020). Standardizing ecosystem morphological traits from 3D information sources. *Trends in Ecology & Evolution*, 35, 656–667. https://doi. org/10.1016/j.tree.2020.03.006
- West, G. B., Enquist, B. J., & Brown, J. H. (2009). A general quantitative theory of forest structure and dynamics. *Proceedings of the National Academy of Sciences*, 106(17), 7040–7045. https://doi.org/10.1073/ pnas.0812294106

BIOSKETCHES

The authors are interested in biodiversity research and ecological applications of remote sensing, particularly in light detection and ranging (LiDAR) and how LiDAR-derived metrics can inform biodiversity monitoring, the mapping of animal habitats and ecosystem change detection.

Author contributions: W.D.K. and Y.S. contributed to conceptualization, writing—review and editing and visualization; Y.S. contributed to formal analysis; W.D.K. contributed to writing original draft, supervision and funding acquisition.

How to cite this article: Kissling, W. D., & Shi, Y. (2023). Which metrics derived from airborne laser scanning are essential to measure the vertical profile of ecosystems? *Diversity and Distributions*, 00, 1–6. https://doi.org/10.1111/ddi.13760