Systematic synthesis of the notion of scale levels and digital representation in the application of 3D geospatial technology for virtual forests

Arnadi Murtiyoso[1,2], Normand Overney[1], and Deni Suwardhi[3]

[1] Forest Resources Management, Institute of Terrestrial Ecosystems, Department of Environmental Systems Science, ETH Zurich, Switzerland – (arnadi.murtiyoso, normand.overney)@usys.ethz.ch
[2] Université de Strasbourg, INSA Strasbourg, CNRS, ICube Laboratory UMR 7357, Photogrammetry and Geomatics Group, 67000 Strasbourg, France – arnadi.murtiyoso@insa-strasbourg.fr
[3] Spatial System and Cadastre Research Group, Faculty of Earth Sciences and Technology, Institut Teknologi Bandung, Bandung 40132, Indonesia – deni.suwardhi@itb.ac.id

Keywords: virtual forest, 3D, LoD, scale, CityGML, definitions, data format, sensors.

Abstract

In recent years, novel 3D reconstruction methods have been developed to enhance traditional image-based point cloud generation techniques, leading, among other things, to the emergence of the concept of virtual forests. However, within the forestry domain, this concept is still somewhat ambiguous and may carry different meanings depending on the context of its implementation. In this paper, we identify the need to consolidate existing notions of virtual forests and aim to present a unified definition from a geomatics perspective. To this end, we introduce two concepts: a sensor-focused Level of Scale (LoS) and a data-focused Level of Detail (LoD). For the latter, which draws inspiration from existing definitions in CityGML, an additional data format dimension has been incorporated. A numerical analysis was conducted to highlight the importance of data representation and scale. While the results in this paper are not intended to establish a standard—achieving such would require much greater effort in the future—they hopefully provide a foundation for establishing a standard in 3D vegetation mapping and representation that can be practically applied within the forestry domain.

1 Introduction

The concept of virtual forests has received increasing attention in the literature for good reason. Virtual representations of trees and vegetation enable the creation of a geometrically and semantically reliable 3D digital twin of forest environments (Hejtmánek et al., 2022; Huang et al., 2021; Murtiyoso et al., 2023). Such digital twins have numerous important applications, particularly in forest inventories (Lopez Serrano et al., 2022), forest information systems (Holopainen et al., 2020), and environmental modeling (Fabrika et al., 2018).

However, the term *virtual forests* itself can take on different meanings depending on various factors. It may refer to a 3D representation of the forest "as is," a visualization of an ideal forest state, an immersive 3D environment (such as in VR, AR, or MR), or even a non-3D digital representation of the forest. This diversity in definitions has led to varied interpretations and approaches as the concept has evolved and been applied more broadly.

One area that remains unclear, especially within the virtual forest sector, is the relationship between scale and detail and their dependence on the choice of sensors. In this paper, we aim to systematically synthesize these concepts, providing a structured understanding of the different methods used to create virtual forests from a geospatial perspective. Additionally, we propose an extension to the existing CityGML geospatial data model, incorporating our conceptual schema as a case study for its practical implementation. Thus, the main contribution of this paper is the introduction of *levels of scale* and *levels of detail* for virtual forests, based on a synthesis of the current state of the field.

To demonstrate the practical value of the conceptual ideas presented in this paper, we propose a small extension to the CityGML model to incorporate these concepts. CityGML, which is based on the Geography Markup Language (GML), offers a standardized framework for representing and exchanging 3D city models and includes various thematic modules for urban features, such as vegetation.

Within the CityGML framework, the *SolitaryVegetationObject* represents individual vegetation elements like trees, shrubs, or other single plants within a city model. These objects belong to the vegetation thematic module and can be represented at different levels of detail (Figure 3). CityGML also supports *Application Domain Extensions* (ADE), which provide a structured approach to extending the CityGML Conceptual Model (CM) for specialized applications or domains.

An ADE is constructed using a UML conceptual model and allows new classes, attributes, and relationships to be added while maintaining compatibility with CityGML's core structure. This enables the integration of custom data that meets the semantic requirements of specific applications. Using this functionality, we will extend the *SolitaryVegetationObject* feature to include both point cloud and mesh representations.

Further analysis will be conducted to add a numerical assessment of the importance of the LoS (Level of Scale) and LoD (Level of Detail) concepts. This will involve comparing datasets of the same forest object, each with varying LoS and LoD levels. Two numerical analyses are planned: (1) calculating the projected canopy area and (2) calculating the object's volume. The projected canopy area serves as a useful indicator of crown diameter and has potential applications in environmental modeling, while the volume measurement is essential for estimating forest biomass.

Through these experiments, the paper aims to emphasize the significance of scale in applying 3D geospatial technologies within the forestry sector.

2 State of the Art

When discussing scale in the context of geospatial technologies for vegetation mapping, two distinct paradigms can be considered: (1) sensor-focused classification and (2) data-focused classification.

The sensor-focused classification offers future users guidance on selecting the most suitable sensor for their specific needs—a framework that, to our knowledge, does not yet exist in the forestry domain. Traditionally, forestry applications often approach sensor selection in reverse, starting with available sensors and exploring what can be achieved with them (Allen et al., 2023; Béland et al., 2014; Weiser et al., 2022; Yusup et al., 2023). Researchers commonly acquire tools such as Terrestrial Laser Scanners (TLS) or Mobile Laser Scanners (MLS) and then test the limitations of these technologies in forestry settings.

While this approach is effective for research, it may be unsustainable for large-scale industrial or commercial applications where considerations like time, cost, and resource constraints are paramount. Instead, a specification-centric approach, similar to that used in traditional mapping and surveying, could better support commercial applications. To address this, we propose a sensor-focused classification system known as Levels of Scale (LoS), which categorizes various geospatial techniques used in forestry based on suitability. This concept draws from the work of Nex and Remondino (2014) and builds upon prior studies (Murtiyoso et al., 2023, 2024) that classified geospatial techniques by their intended outcomes.

The second paradigm focuses on the data representation phase rather than data acquisition. This approach, known as Levels of Detail (LoD), provides a systematic framework for classifying data in an intuitive way. While LoD concepts are well established in the OGC CityGML standard, they were developed primarily for urban contexts, with limited applicability to vegetation, especially forests (Suwardhi et al., 2022; Ambarwari et al., 2024). Consequently, vegetation LoD classifications are less standardized, with researchers often defining LoD according to their unique requirements (Fu et al., 2024). In this paper, we aim to synthesize prominent vegetation LoD classifications into a more generalized framework, without attempting to redefine existing standards.

An additional, though equally important, consideration in LoD discussions is data format. While solid object models are common in 3D GIS and Building Information Models (BIM), the irregular nature of forestry data may favour alternative formats such as 3D meshes or point clouds. Point clouds, for example, may be preferable for applications like data visualization and species identification. In contrast, forest growth models may not require realistic representations, making simpler models sufficient. Thus, multiple 3D data representations—point clouds, meshes, and solid models—are required to meet the diverse needs of forestry applications

3 Proposed Solution and Discussions

3.1 Theoretical Definition

The proposed concept of LoS, presented in Figure 1, considers two primary factors: scene size and scene complexity. Under this schema, future users with specific project requirements can consult the LoS classification to identify the optimal sensor for their needs. Although this paper introduces the initial schema, we suggest that, with further refinement, the LoS concept could be highly beneficial for both new and experienced users of geospatial technology in the forestry sector. This is particularly relevant for 3D technologies, which are still relatively novel in forestry; recent studies have begun evaluating which sensors are best suited to practical applications, such as national forest inventories (Kükenbrink et al., 2022; Liang et al., 2018).

Figure 2, heavily influenced by Ortega-Córdova (2018), illustrates the proposed LoD concept. Complementary to the LoS concept, the LoD classification is crucial as it helps determine the most suitable type of data representation based on user needs. This concept, like the term "virtual forest," encompasses a wide range of interpretations, underscoring the need for a unified, if not standardised, definition of LoD for forestry applications.

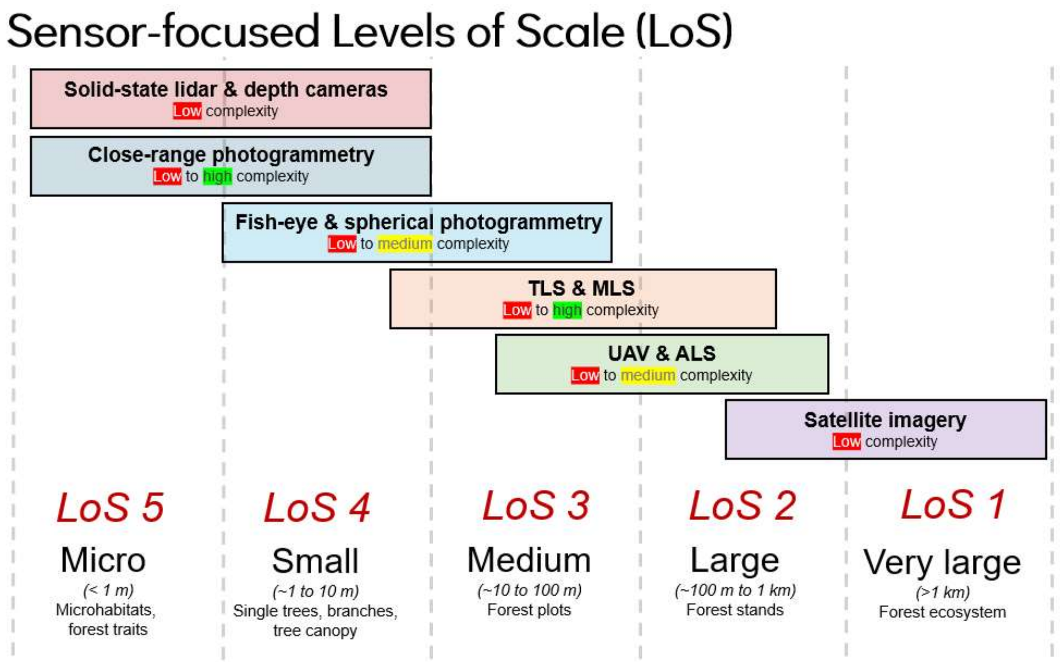


Figure 1. The proposed schema for the sensor-focused Levels of Scale (LoS) based on the required resolution of the end data. The aim of the schema is to propose the best-suited sensors for users. Based on Murtiyoso et al. (2024).

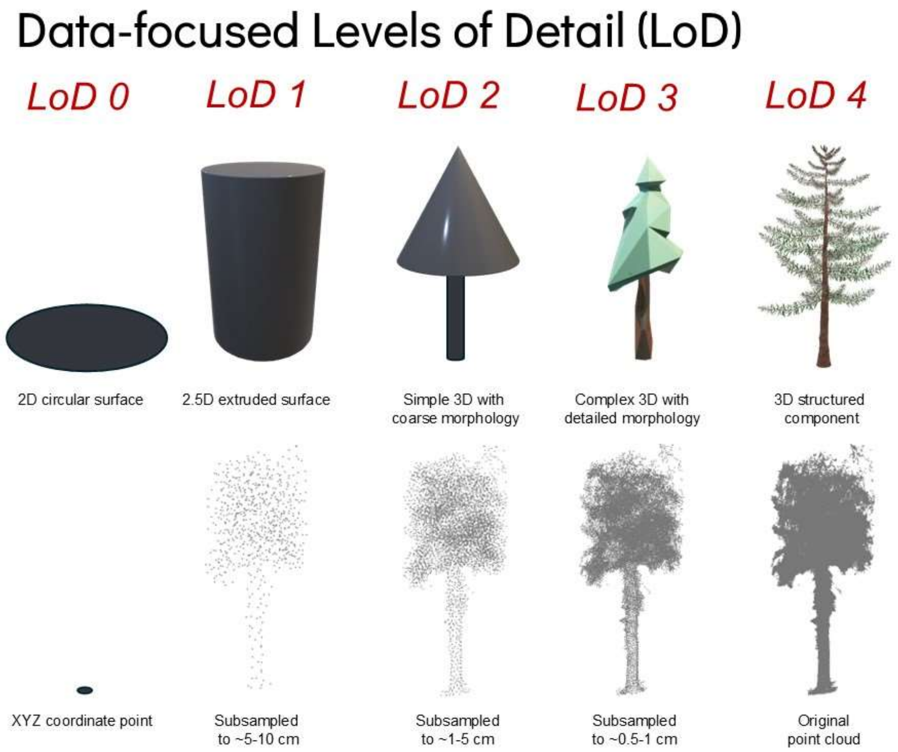


Figure 2. A data-focused definition of the Levels of Detail (LoD) concept derived from (Ortega-Córdova, 2018) and extended to the point cloud format.

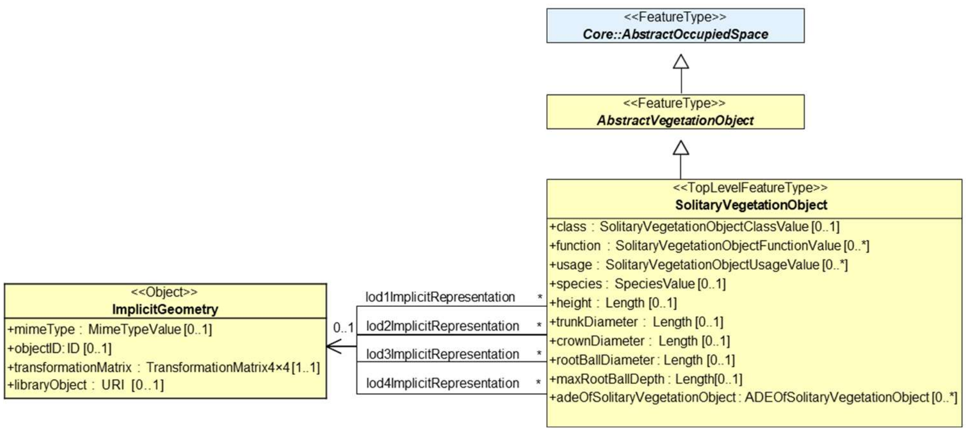


Figure 3. Class Diagram for a part of the Vegetation module in CityGML highlighting the possibility to extend the SolitaryVegetationObject to include custom data formats through the ADE method.

In addition to the question of scale and complexity in the LoD concept, we propose another layer that addresses the data format. The use of point clouds is well known in forestry applications and, in many cases, they are sufficient, even though meshes and solid models are also important, especially when discussing forest modelling. To this end, we also attempt to present point clouds as an alternative format to 3D models.

3.2 Technical Implementation

A more technical implementation of Figure 2 was developed in CityGML. While the LoD concept already exists in CityGML, an extension to include different data formats is not yet common. We therefore propose a conceptual model, as shown in Figure 3, where the LoD should also be extended to the representation of point clouds. Using this approach, a user may choose how to represent the data not only according to what is available to them, but also tailored to the specific requirements of their application.

The 3D data preparation begins with LAS files, which are transformed into a format compatible with visualisation software like Cesium by converting to glTF (glb). This model is then translated into CityGML format, enabling the model to be queryable in a database with searchable attributes associated with each feature. Using CityGML as a base, the model is then converted into 3D tiles, allowing for spatial visualisation within the Cesium environment.

To determine optimal LoD, we utilized CloudCompare to analyse subsampling ranges from 0.04 to 0.2 Figure 4. The results were plotted on a graph, with the x-axis representing subsampling levels and the y-axis showing the point count generated from the LAS files.

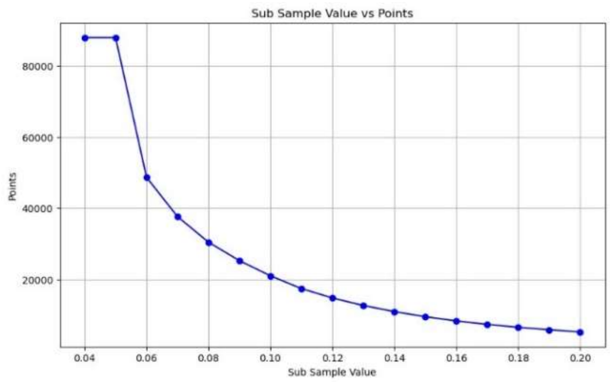


Figure 4. Sub-sampling process of the example point cloud to determine the appropriate resolution of each LoD.

Figure 5. Visualisation of the LoD, including the point cloud, in Cesium.

Based on this data, four quantiles were identified to define LoD levels 1 through 4, enabling an evaluation of the mean subsampling method for creating a 3D object in CityGML that is subsequently visualized in Cesium Figure 5.

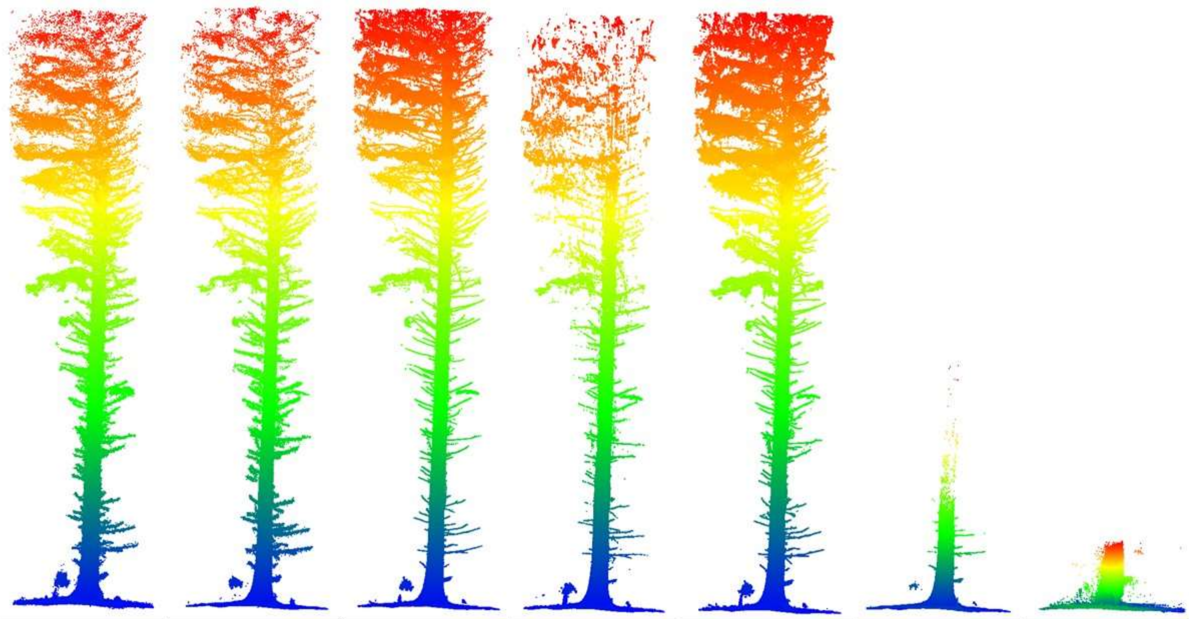
While this study provides a model for a specific tree species, we recommend a broader application of this methodology to evaluate LoD3 and LoD4 across various species, given the density requirements necessary for detailed point clouds.

3.3 Numerical Assessment

A numerical assessment was also conducted to demonstrate the necessity of both the LoS and LoD concepts. While the two concepts are closely linked, as one represents the data acquisition aspect and the other the data representation aspect, they are nonetheless distinct definitions.

Table 1 shows the comparison of the same tree captured using seven different sensors with varying levels of point cloud resolution. The data was sourced from the Silvilaser 2021 Benchmark dataset (Hollaus & Chen, 2023). Several standard point cloud parameters were extracted from all datasets and presented.

In Table 1, the data generated by TLS provided the highest density in terms of point cloud resolution, while MLS corresponds to the lower levels of TLS density. This came at the expense of a larger file size and higher point cloud count.



	GeoSlam ZEB Horizon	LiBackpack DGC50	Riegl VZ400i	Faro Focus 3D X330	Leica RTC360	GoPro (fish-eye photogrammetry)	iPad Lidar (Sitescape)
Size	16.5 Mb	7.5 Mb	52.4 Mb	94.7 Mb	46.3 Mb	70.3 Mb	3.8 Mb
Point count	651 466	226 753	1 342 191	3 731 735	1 836 442	2 668 872	149 343
Avg. density	9 pts/cm³	2 pts/cm³	24 pts/cm³	203 pts/cm³	11 pts/cm³	61 pts/cm³	16.3 pts/cm³
Family	MLS	MLS	TLS	TLS	TLS	Photogrammetry	SSL
LoS	LoS 3	LoS 3	LoS 4	LoS 5	LoS 4	LoS 5	LoS 4

Table 1. A comparison table of seven sensors used to scan the same tree. The tree point cloud was taken from the Silvilaser 2021 Benchmark dataset (Hollaus & Chen, 2023).

The point cloud generated from fish-eye photogrammetry (GoPro) also produced very dense results, though at the expense of capturing the upper parts of the canopy. A similar phenomenon was observed in the iPad dataset.

These observations show several interesting points:

1. Range and cost of sensors: The GoPro was shown to be capable of reaching an LoS 5 in our definition, though it is very limited in terms of range. TLS can also achieve LoS 4 and even 5, but cost considerations are important. This raises the question of *whether a more expensive tool is justifiable when a lower-cost sensor can deliver a comparable quality.*
2. Data size: With higher density, the data size also increases. While visualisation purposes may benefit from higher-density data for increased realism, this may not hold true for other applications, e.g. forest growth modelling. This raises the question of *whether higher-density data is necessary when lower-density data meets the application's requirements.*
3. Quality of data: It is also worth noting that data quality, in terms of noise, is an important factor in choosing a sensor. TLS sensors tend to produce less noise compared to MLS and other technologies, even with similar point cloud resolutions. This raises the question of *whether geometric quality is a critical factor in the application.*

From these three points, the answers tend to converge towards the same idea: the choice of sensors should depend on the requirements of the application, or, in other words, the expected output. The decision should be based on the expected geometric quality, the scale of the scene, and the cost. In this regard, Figure 1 can be used as a starting point to assist future users in making their choice.

	LOD 1	LOD 2	LOD 3	LOD 4	Reference
Canopy surface (m^2)	15.34	11.22	20.51	24.56	24.38
Volume (m^3)	129.61	22.36	58.94	59.91	61.30

Table 2. Surface and volume of the 3D models generated with different LoDs.

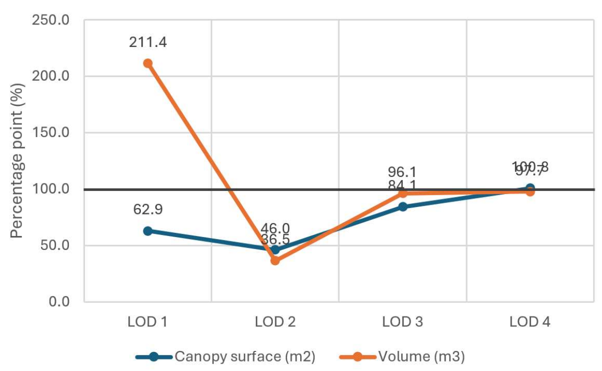


Figure 6. Percentage of each LoD canopy surface area and volume with regards to the reference data.

To demonstrate the necessity of having a data-focused LoD as described in Figure 2, two numerical analyses were performed on the 3D models generated from four LoDs (1 to 4). The first analysis involves computing the surface area of the canopy when projected onto the ground. The second analysis calculates the volume of the 3D model from the four LoDs, which serves as a useful indicator for biomass estimation. As a reference for comparison, the canopy surface area and volume of the 3D mesh of the original point cloud were also provided. The results of the two analyses are summarised in Table 2, while Figure 6 shows the percentage of the values in relation to the reference.

From Figure 6, it can be inferred that LoD 4 provides the most accurate representation of the object, with LoD 3 following closely behind. However, it is worth noting that the file size of LoD 4 is significantly larger than that of LoD 3. This suggests that LoD 3 may be more suitable for modelling purposes where complex computations require the input data to be as lightweight as possible. This further supports the underlying hypothesis already proposed as a response to the challenges of LoS: the choice of LoD depends on the specific requirements of the application. For instance, to measure or store attribute information on tree height, LoD 1 is already sufficient and offers a much smaller file size. When diameter at breast height (DbH) is required, LoD 2 is necessary, but LoDs 3 and 4 are excessive.

4 Conclusions

In this paper, a synthesis of the existing concepts of 3D data scale and representation within the context of a virtual forest was discussed. Based on the literature and prior work, a diagram was presented to summarise the concept of Levels of Scale. This diagram was proposed with the hope that it may assist future virtual forest users in determining the correct sensor required for their respective projects. A second diagram, describing the concept of Levels of Detail, was also presented. While heavily influenced by existing definitions within the CityGML standard, this diagram considers various aspects of the forestry domain, including an extension to encompass multiple formats of 3D data, notably the addition of point cloud representation. The paper implemented the proposed theoretical framework within CityGML and Cesium, with several numerical analyses also performed. Key takeaways from these operations include:

1. Output-centred approach: The considerations for using geospatial data in forestry should shift from a sensor-centred to an output-centred perspective. The choice of sensor and acquisition strategy should depend on the desired output, particularly the expected scale, scope, and geometric resolution. In this regard, the proposed LoS concept can serve as a starting point for future users.
2. Inclusion of multiple data formats: The representation of 3D forest data in information systems should accommodate multiple data formats. While traditional GIS in its 3D form relies on 3D models, there is a growing demand for alternative representations using point clouds. The use of the LoD concept is thus essential and should be extended to include these formats.
3. LoS and LoD choices: Ultimately, the selection of the appropriate LoS and LoD for each project depends largely on the cost and the specific nature of the project.

In the future, a more concrete implementation of the concepts presented in this paper will be investigated and developed. The overarching objective of this work is to promote a more sustainable use of geospatial data in forestry, moving beyond the raw exploitation of point clouds towards a more integrated information system. In the long term, this approach could greatly benefit forest management practices.

References

Allen, M. J., Grieve, S. W. D., Owen, H. J. F., & Lines, E. R. 2023. Tree species classification from complex laser scanning data in Mediterranean forests using deep learning. Methods in Ecology and Evolution, 14(7), 1657–1667. doi.org/10.1111/2041-210X.13981

Ambarwari, A., Suwardhi, D., Rani, M. S., Husni, E., Junaidy, D. W., Agirachman, F. A., Murtiyoso, A., & Griess, V. C. 2024. Conceptual Model of Graph-based Individual Tree and Its Utilization in Digital Twin and Metaverse of Urban Forest. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *48*, 7-12.

Béland, M., Baldocchi, D. D., Widlowski, J.-L., Fournier, R. A., & Verstraete, M. M. 2014. On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial LiDAR. Agricultural and Forest Meteorology, 184, 82–97. doi.org/10.1016/j.agrformet.2013.09.005

Fabrika, M., Valent, P., & Scheer, L. 2018. Thinning trainer based on forest-growth model, virtual reality and computer-aided virtual environment. Environmental Modelling and Software, 100, 11–23. https://doi.org/10.1016/j.envsoft.2017.11.015

Fu, R., Pađen, I., & García-Sánchez, C. 2024. Should we care about the level of detail in trees when running urban microscale simulations? Sustainable Cities and Society, 101, 105143. doi.org/10.1016/J.SCS.2023.105143

Hejtmánek, L., Hůla, M., Herrová, A., & Surový, P. 2022. Forest digital twin as a relaxation environment: A pilot study. Frontiers in Virtual Reality, 3, 1033708. doi.org/10.3389/frvir.2022.1033708

Hollaus, M., & Chen, Y.-C. 2023. SilviLaser 2021 Benchmark Dataset - Terrestrial Challenge (1.1) [Data set]. TU Wien. doi.org/10.48436/kndye-egv02

Holopainen, J., Mattila, O., Pöyry, E., & Parvinen, P. 2020. Applying design science research methodology in the development of virtual reality forest management services. Forest Policy and Economics, 116, 102190.

Huang, J., Lucash, M. S., Scheller, R. M., & Klippel, A. 2021. Walking through the forests of the future: using data-driven virtual reality to visualize forests under climate change. International Journal of Geographical Information Science, 35(6), 1155–1178.

Kükenbrink, D., Marty, M., Bösch, R., & Ginzler, C. 2022. Benchmarking laser scanning and terrestrial photogrammetry to extract forest inventory parameters in a complex temperate forest. International Journal of Applied Earth Observation and Geoinformation, 113(March), 102999. doi.org/10.1016/j.jag.2022.102999

Liang, X., Hyyppä, J., Kaartinen, H., Lehtomäki, M., Pyörälä, J., Pfeifer, N., Holopainen, M., Brolly, G., Francesco, P., Hackenberg, J., Huang, H., Jo, H.-W., Katoh, M., Liu, L., Mokroš, M., Morel, J., Olofsson, K., Poveda-Lopez, J., Trochta, J., … Wang, Y. 2018. International benchmarking of terrestrial laser scanning approaches for forest inventories. ISPRS Journal of Photogrammetry and Remote Sensing, 144, 137–179. doi.org/10.1016/j.isprsjprs.2018.06.021

Lopez Serrano, F. R., Rubio, E., García Morote, F. A., Andrés Abellán, M., Picazo Córdoba, M. I., García Saucedo, F., Martínez García, E., Sánchez García, J. M., Serena Innerarity, J., Carrasco Lucas, L., García González, O., & García González, J. C. 2022. Artificial intelligence-based software (AID-FOREST) for tree detection: A new framework for fast and accurate forest inventorying using LiDAR point clouds. International Journal of Applied Earth Observation and Geoinformation, 113, 103014. doi.org/https://doi.org/10.1016/j.jag.2022.103014

Murtiyoso, A., Cabo, C., Singh, A., Obaya, D. P., Cherlet, W., Stoddart, J., Fol, C. R., Beloiu Schwenke, M., Rehush, N., Stereńczak, K., Calders, K., Griess, V. C., & Mokroš, M. 2024. A Review of Software Solutions to Process Ground-based Point Clouds in Forest Applications. Current Forestry Reports. doi.org/10.1007/s40725-024-00228-2

Murtiyoso, A., Holm, S., Riihimäki, H., Krucher, A., Griess, H., Griess, V. C., & Schweier, J. 2023. Virtual forests: a review on emerging questions in the use and application of 3D data in forestry. International Journal of Forest Engineering, 35(1), 29–42. doi.org/10.1080/14942119.2023.2217065

Nex, F., & Remondino, F. 2014. UAV: platforms, regulations, data acquisition and processing. In F. Remondino & S. Campana (Eds.), 3D Recording and Modelling in Archaeology and Cultural Heritage: Theory and Best Practices (pp. 73–86). Archaeopress.

Ortega-Córdova, L. M. 2018. Urban Vegetation Modeling 3D Levels of Detail. TU Delft.

Suwardhi, D., Fauzan, K. N., Harto, A. B., Soeksmantono, B., Virtriana, R., & Murtiyoso, A. 2022. 3d modeling of individual trees from lidar and photogrammetric point clouds by explicit parametric representations for green open space (gos) management. *ISPRS International Journal of Geo-Information*, *11*(3), 174.

Weiser, H., Schäfer, J., Winiwarter, L., Krašovec, N., Fassnacht, F. E., & Höfle, B. 2022. Individual tree point clouds and tree measurements from multi-platform laser scanning in German forests. Earth System Science Data, 14(7), 2989–3012. doi.org/10.5194/essd-14-2989-2022

Yusup, A., Halik, U., Keyimu, M., Aishan, T., Abliz, A., Dilixiati, B., & Wei, J. 2023. Trunk volume estimation of irregular shaped Populus euphratica riparian forest using TLS point cloud data and multivariate prediction models. Forest Ecosystems, 10, 100082. doi.org/10.1016/j.fecs.2022.100082