SPONSORED BY THE

Federal Ministry of Education and Research

VirtualForests

Co-funded by the Erasmus+ Programme of the European Union

LiDAR data pre-processing

Viet Nguyen
Faculty of Forest and Environment
University for Sustainable Development
Eberswalde, Germany

UVa

Hochschule für nachhaltige Entwicklung Eberswalde

cesefor

IEFC

HUS VNU UNIVERSITY OF SCIENCE

AgroParisTech Talents for a sustainable planet

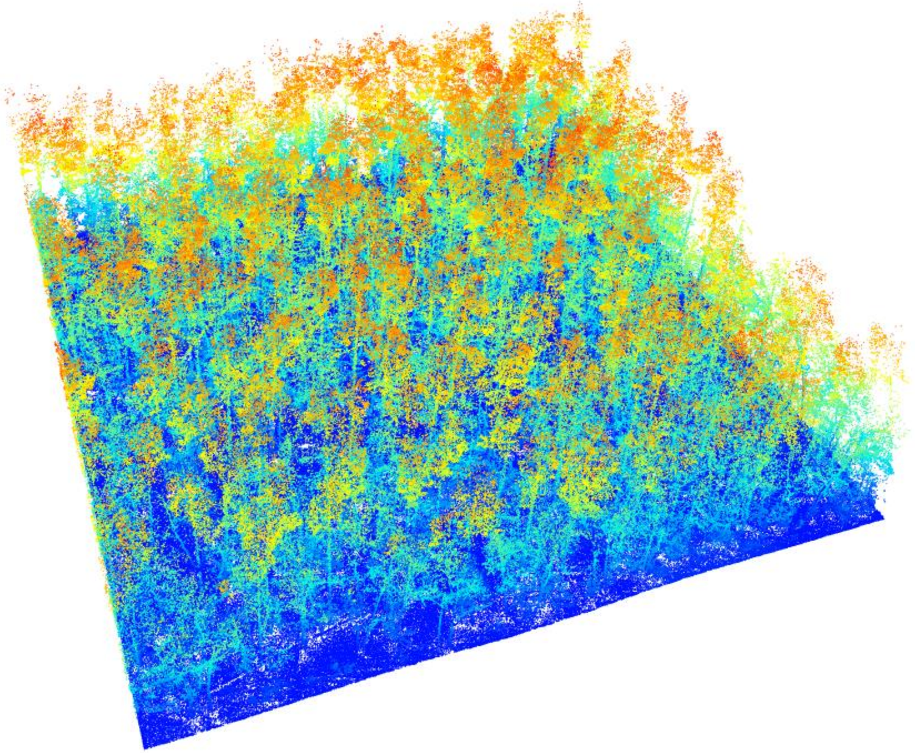
Introduction

- Raw point cloud data should be pre-processed before being used for further analysis.
- The basic preprocessing procedures including clipping, filtering, thinning, and normalization.
- R programming language was used in this tutorial.
- Recommended materials for the overview of LiDAR analysis for forestry: White et al., 2016, Kelly & Tommaso, 2015, Wulder et al., 2012, van Leeuwen & Nieuwenhuis, 2010

Content

1. **Tools for LiDAR processing**
2. **Dataset**
3. **Point cloud pre-processing**

- Clipping
- Thinning
- Filtering
- Normalization



Open-source tools

- There are several tools for processing LiDAR data, such as LAStools, FUSION, or Cloudcompare. In particular for forestry applications, there are also some R libraries, e.g. rLiDAR, rGEDI, treeIS, etc.
- Recommended material: Open-Source tools in R for forestry and forest ecology (Atkins et al., 2022)

Open-source tools

- FUSION was one of the first lidar analysis platforms designed for forestry. The software capabilities focus on area-based approach.
- CloudCompare is a commonly used software nowadays but it doesn't have functions particularly for forestry.
- LAStools is also commonly used for large-data processing, but the fee is applied to work with large data.

Open-source tools

Package	Brief Description	Comment
PDAL	Point Data Abstraction Library in C/C++. Designed for translating and processing point cloud data. PDAL Contributors (2018)	Generic and multi-purpose C++ library
PCL	Point Cloud Library in C++. Cross-platform and designed for 2D/3D image and point cloud processing. Rusu and Cousins (2011)	
CloudCompare	3D point cloud and triangular mesh processing software. It was originally designed to perform comparisons between two dense 3D point clouds	
LAStools	A suite of LiDAR processing tools widely known for their very high speed and high productivity. They combine robust algorithms with efficient I/O and clever memory management to achieve high throughput for datasets containing billions of points.	A few modules for LAS files manipulation are open source and cross-platform
Whitebox GAT	GIS capabilities for ALS including basic and advanced tools for ALS processing	
FUSION/LDV	A collection of task-specific command line programs (FUSION) and a viewer (LDV) McGaughey (2015)	Sources available on request only. Not cross-platform.
GRASS GIS	FOSS Geographic Information System software suite used for geospatial data management. It supports basic and advanced ALS data processing and analysis.	
SPDlib	A set of open source software tools for processing laser scanning data (i.e., LiDAR), including data captured from airborne and terrestrial platforms. Bunting et al. (2011, 2013)	

Source: Jean-Romain Roussel et al., 2020 https://doi.org/10.1016/j.rse.2020.112061

lidR package

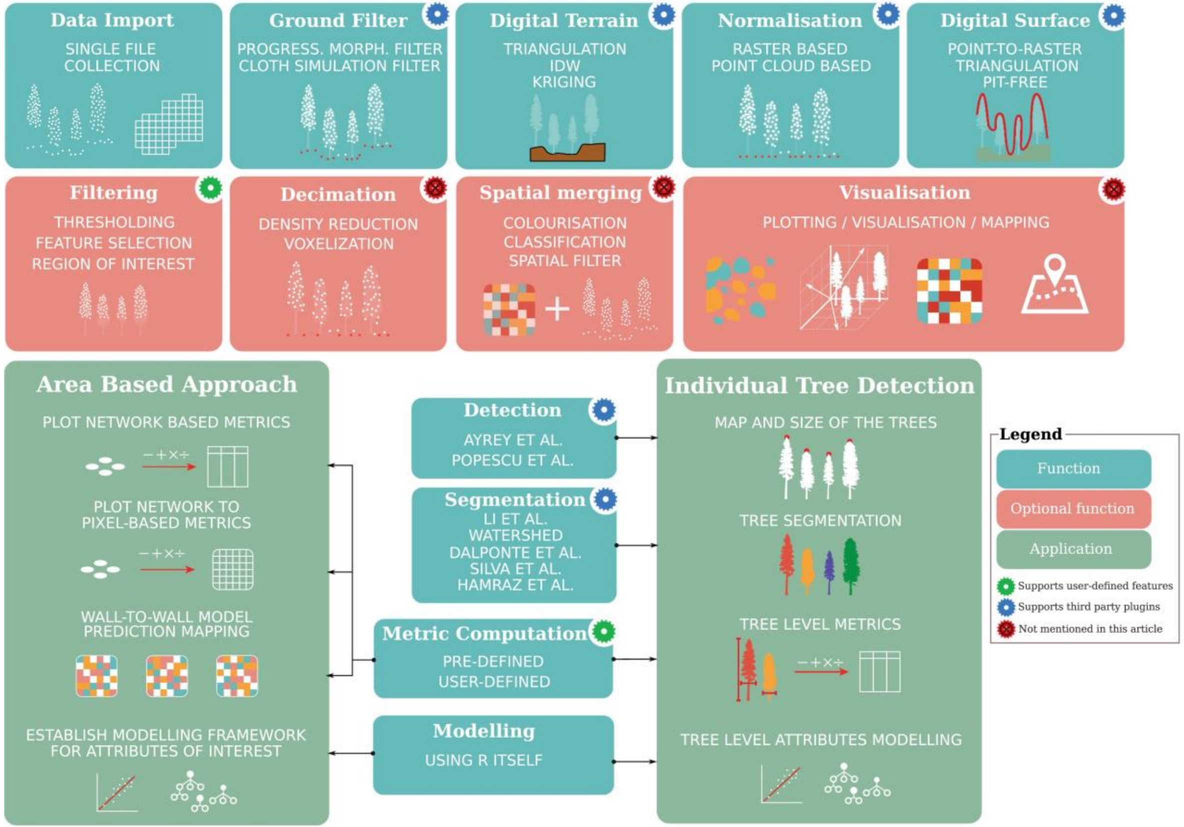
- Among others, the lidR package in R released in 2020 rising as one of the most commonly used tools for lidar processing, which is able to conduct many common procedures to analyze LiDAR data.
- Recommended material: lidR: An R package for analysis of Airborne Laser Scanning (ALS) data (Roussel et al., 2020)

lidR package

- In forestry, there are two common approaches have been developed to derive forest attributes: the area-based (ABA) and the individual tree crown (ITC) approach (Eysn et al., 2015).
- In ABA, tree attributes are estimated in grid cells based on metrics that summarise the distribution of the point cloud within each cell
- ITS allows classifying individual trees and then calculating tree attributes at a single tree level

lidR package

Overview of key functions in the lidR package



Source: Jean-Romain Roussel et al., 2020 https://doi.org/10.1016/j.rse.2020.112061

Möllergrab marteloscope point cloud

Site description:

*100*100m plot*

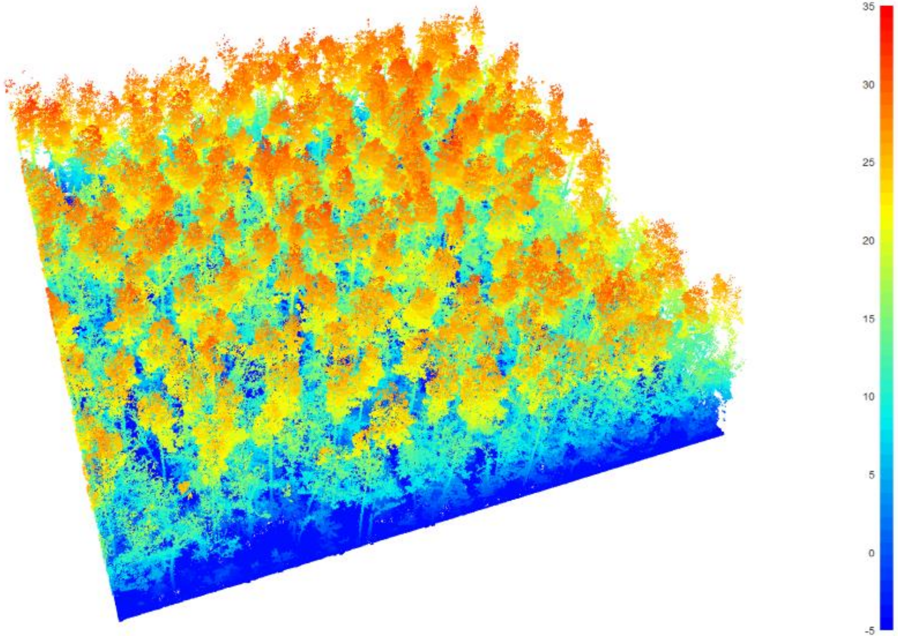
Pine with understory Beech forest

Eberswalde, Germany

Data collection:

Personal laser scanning (ZEB-HORIZON)

January 2022



VirtualForests

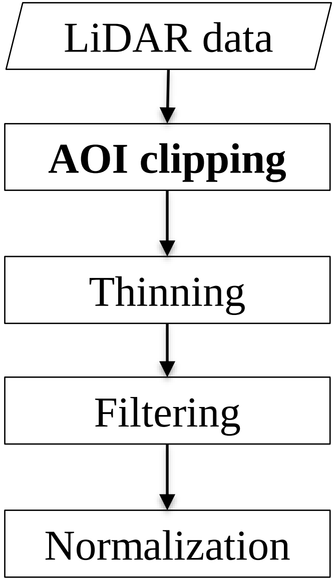
Co-funded by the Erasmus+ Programme of the European Union

Möllergrab marteloscope inventory

Count	Tree_No	Species	DBH_cm	H_m	QMP	BA_m2	V_m3	x	y
1	1	RBU	96	89	1	0,01	0,02	4192702656	585165519
2	2	RBU	168	140	1	0,02	0,09	4192716639	5851655842
3	3	RBU	136	144	1	0,01	0,05	4192720241	5851655921
4	4	RBU	115	87	1	0,01	0,03	4192712534	5851658264
5	5	GKI	395	306	1	0,12	1,63	4192740596	5851657152
6	6	RBU	8	82	1	0,01	0,01	4192750769	5851656713
7	7	GKI	406	304	1	0,13	1,71	4192791476	5851661781
8	8	GKI	323	280	1	0,08	1,01	4192826191	5851662044
9	9	GFI	76	76	1	0	0,02	4192832216	585166229
10	10	RBU	176	176	1	0,02	0,11	4192879064	5851661639

Inventory from 2021 with 743 trees to validate tree attribute estimation

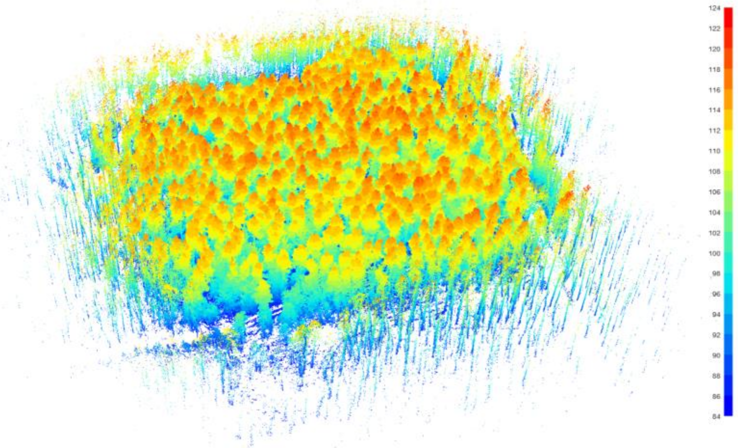
Clipping



The original point cloud was first clipped to the marteloscope area (vector file). As only the trees inside the marteloscope were interested, this clipping process facilitated reducing significant processing time and data storage.

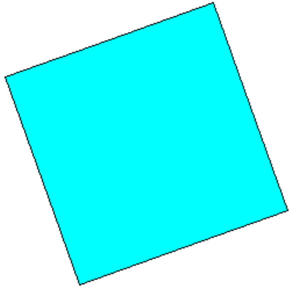
Clipping

LiDAR data
↓
AOI clipping
↓
Thinning
↓
Filtering
↓
Normalization

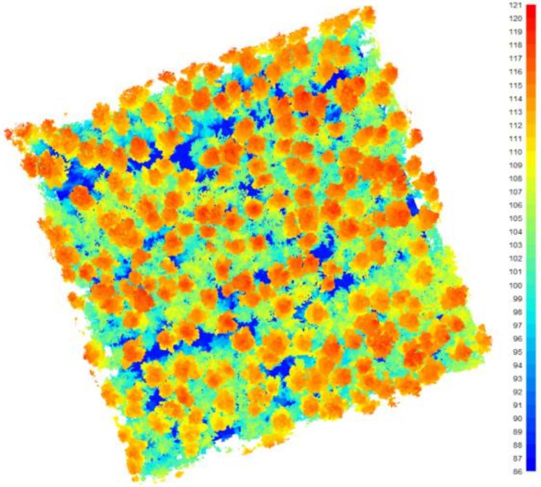


+

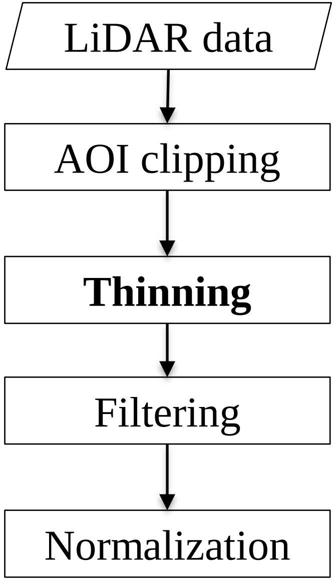
Mollergrab Boundary



=

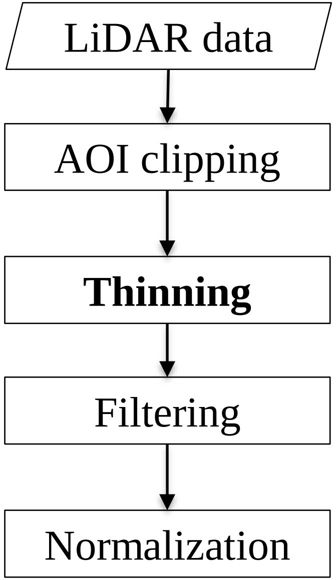


Thinning



The thinning step commonly used with point clouds from airborne laser scanning (ALS) systems. ALS systems usually cannot acquire data in a large area by a single flight, instead, the aircraft or UAV need to fly over back and forth to record the area. To make sure there is no gap in the point cloud due to the long distance between flight paths, the flight path should stay close.

Thinning



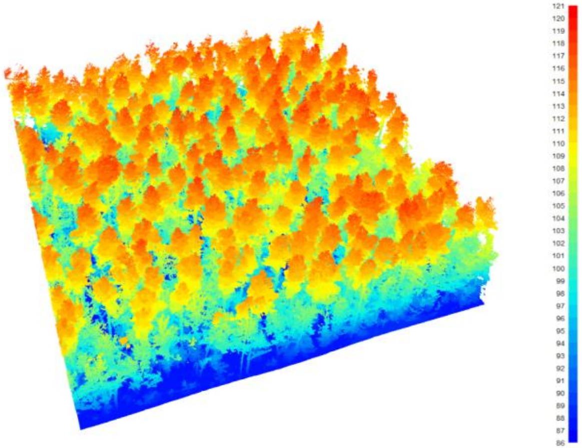
As a result, there are overlapping areas in the acquired data that have a higher point density than the other. This leads to problems with further processing.

For our data, the thinning step intended to reduce large amounts of point cloud while preserving essential information about the object, further processing would be much more straightforward with less computing time and storage

Thinning

The *decimate_points* function in the *lidR* package removes points randomly to reach a homogeneous density of 100 points/m²

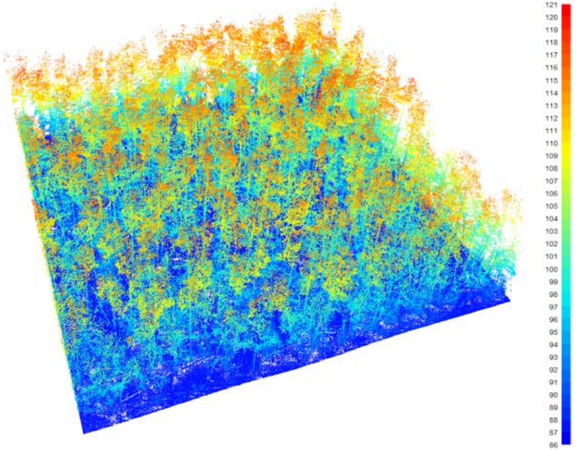
LiDAR data → AOI clipping → **Thinning** → Filtering → Normalization



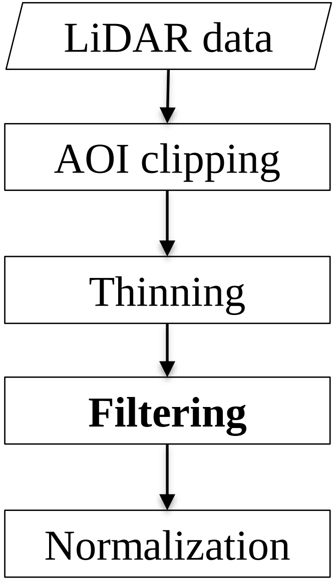
≈ 30000 points/m²

→

100 points/m²

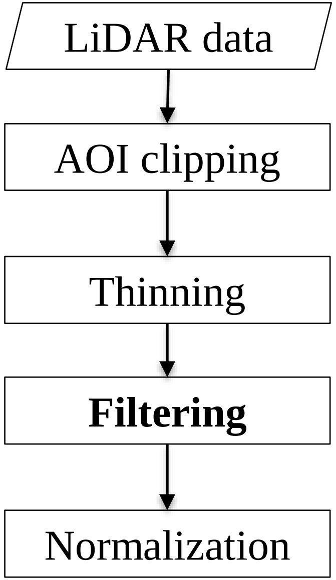


Filtering



The raw point cloud often contains noises or outliers due to the limitation of sensors, the inherent noise of the acquisition equipment, lighting, the reflective nature of the surface, and artifact in the scene e.g. flying birds (Han et al., 2017). Hence, filtering procedures on raw point clouds are required to generate reliable data for further processing.

Filtering

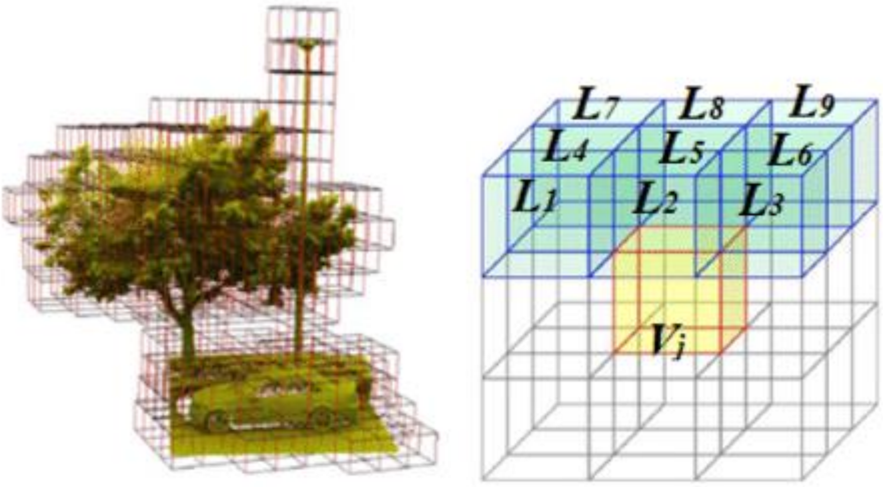


The isolated voxels filter (IVF) algorithm builds 3 x 3 x 3 = 27 voxels surrounding every point in the point cloud and counts how many points are there in the voxels. If the number of points is lower than a predefined number, then the point will be removed.

Filtering

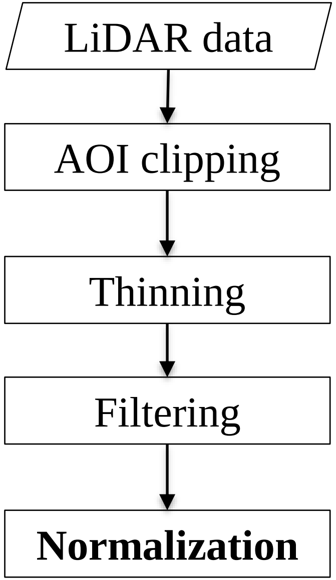
Isolated voxels filter (IVF) algorithm

LiDAR data
↓
AOI clipping
↓
Thinning
↓
Filtering
↓
Normalization



Source: Ye et al., 2022 https://doi.org/10.1109/TITS.2020.3028033

Normalization



The common following step in point cloud preprocessing is normalization. Subtracting the terrain surface from all the point clouds simplifies and facilitates the following procedures, by removing the influence of terrain on above-ground measurements.

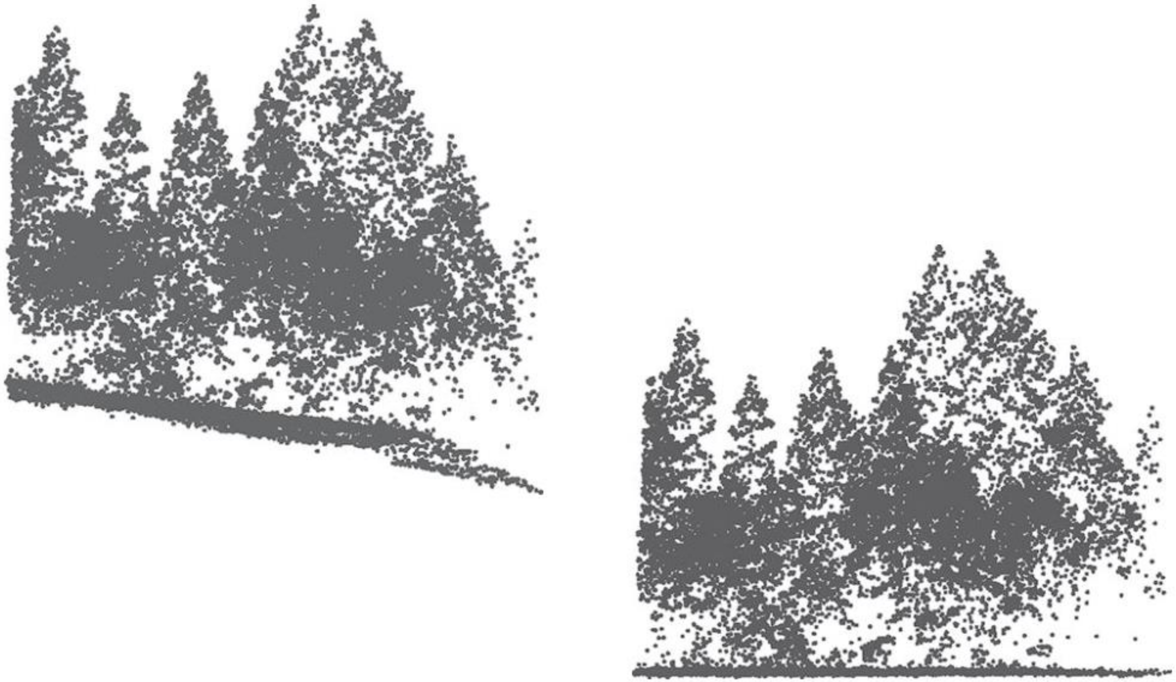
Specifically, in forest application, it allows comparing tree height relatively instead of absolute tree height.

Co-funded by the Erasmus+ Programme of the European Union

Normalization

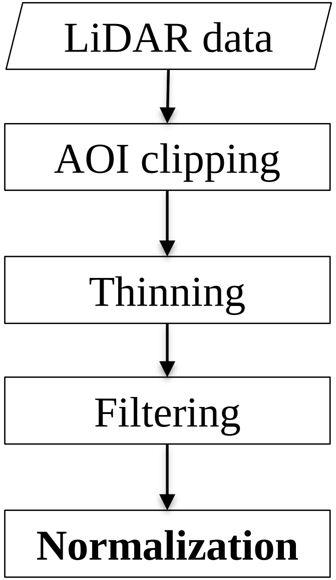
Source: Li et al., 2012 https://doi.org/10.14358/PERS.78.1.75

LiDAR data
↓
AOI clipping
↓
Thinning
↓
Filtering
↓
Normalization



Before *After*

Normalization

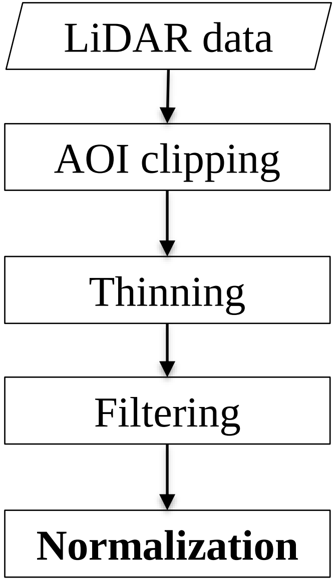


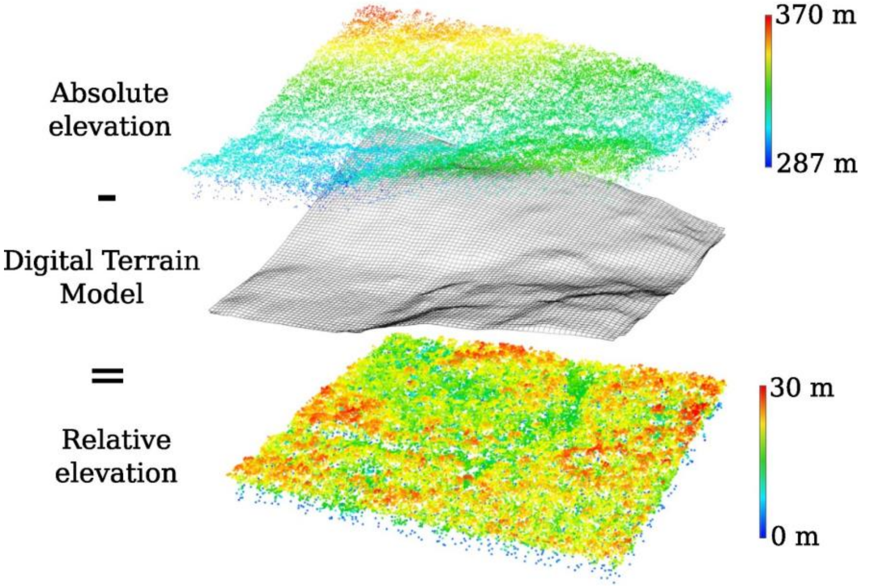
1st raster-based approach

Point cloud normalization based on the raster DTM approach has been widely used due to its explicit. For each point in the data, the algorithm first detects the corresponding DTM pixel underneath and then subtracts this pixel value from the elevation value of the point. However, the method has potential errors. This is because the fact of discrete nature of the raster DTM was created using regularly spaced points, which do not match the actual position of the ground points (Roussel et al., 2020)

Normalization

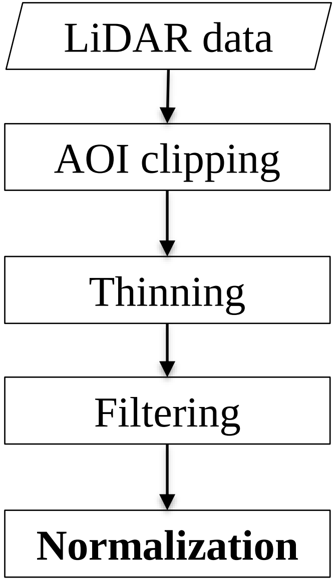
1st raster-based approach





Source: Jean-Romain Roussel et al., 2020
https://doi.org/10.1016/j.rse.2020.112061

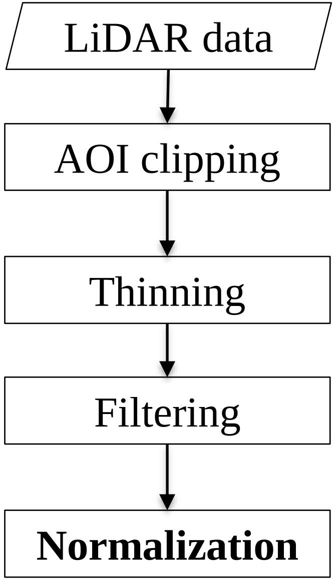
Normalization



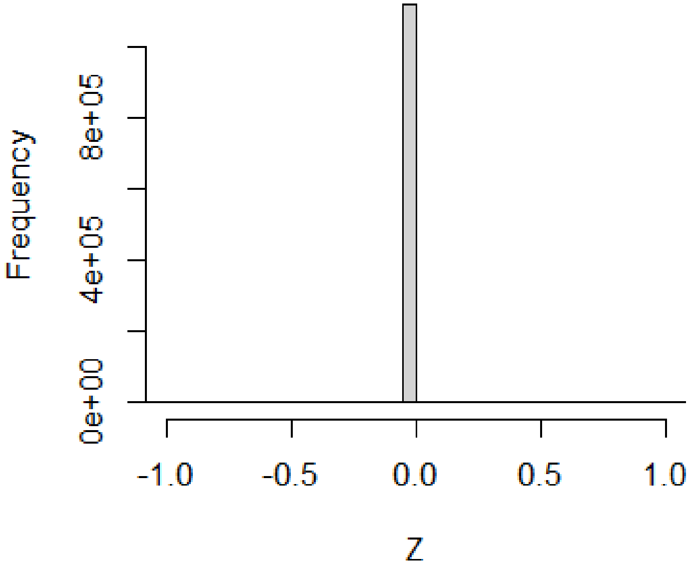
2nd point cloud-based approach

Point cloud-based normalization utilizes all returns, interpolating each ground point to its exact position beneath the non-ground return (Khosravipour et al., 2014). Hence, using this approach, all ground points after interpolating are at exactly 0 height value (figure in the next slide)

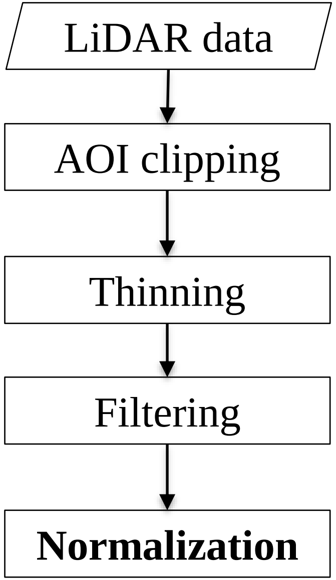
Normalization



2nd point cloud-based approach



Normalization

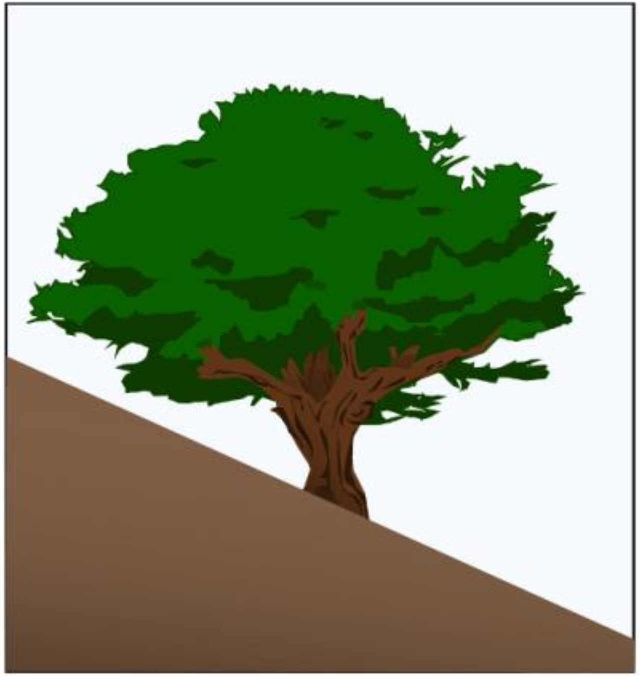


Note!: Normalization has a drawback of distorting the shape of the object especially on high slopes (picture in the next slide). Therefore, based on the application, some prefer to not normalize the point cloud to preserve the geometry of the tree canopy (Roussel et al., 2020).

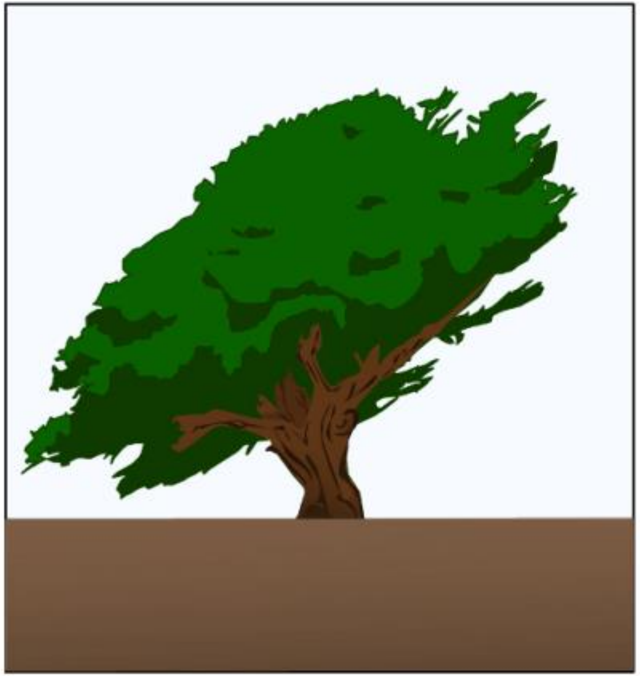
Normalization

Source: Jean-Romain Roussel et al., 2020 https://doi.org/10.1016/j.rse.2020.112061

LiDAR data → AOI clipping → Thinning → Filtering → **Normalization**



Before



After

VirtualForests

Co-funded by the Erasmus+ Programme of the European Union

¡Thank you for your reading!

UVa

Hochschule für nachhaltige Entwicklung Eberswalde

cesefor

IEFC

HUS VNU UNIVERSITY OF SCIENCE

AgroParisTech Talents for a sustainable planet

References

- Atkins, J. W., Stovall, A. E. L., & Alberto Silva, C. (2022). Open-Source tools in R for forestry and forest ecology. Forest Ecology and Management, 503, 119813. https://doi.org/10.1016/j.foreco.2021.119813
- Eysn, L., Hollaus, M., Lindberg, E., Berger, F., Monnet, J.-M., Dalponte, M., Kobal, M., Pellegrini, M., Lingua, E., Mongus, D., & Pfeifer, N. (2015). A Benchmark of Lidar-Based Single Tree Detection Methods Using Heterogeneous Forest Data from the Alpine Space. Forests, 6(5), Article 5. https://doi.org/10.3390/f6051721
- Han, X.-F., Jin, J. S., Wang, M.-J., Jiang, W., Gao, L., & Xiao, L. (2017). A review of algorithms for filtering the 3D point cloud. Signal Processing: Image Communication, 57, 103–112. https://doi.org/10.1016/j.image.2017.05.009
- Kelly, M., & Tommaso, S. D. (2015). Mapping forests with Lidar provides flexible, accurate data with many uses. California Agriculture, 69(1). https://doi.org/10.3733/ca.v069n01p14
- Khosravipour, A., Skidmore, A. K., Isenburg, M., Wang, T., & Hussin, Y. A. (2014). Generating Pit-free Canopy Height Models from Airborne Lidar. Photogrammetric Engineering & Remote Sensing, 80(9), 863–872. https://doi.org/10.14358/PERS.80.9.863

References

- Li, W., Guo, Q., Jakubowski, M. K., & Kelly, M. (2012). A New Method for Segmenting Individual Trees from the Lidar Point Cloud. *Photogrammetric Engineering & Remote Sensing*, *78*(1), 75–84. https://doi.org/10.14358/pers.78.1.75
- Nikolai, Knapp. (2021, October). *Introduction to lidar data processing* [Paper presentation]. Module: Collection & analysis of lidar data, Eberswalde, Brandenburg, Germany.
- Roussel, J.-R., Auty, D., Coops, N. C., Tompalski, P., Goodbody, T. R., Meador, A. S., Bourdon, J.-F., Boissieu, F. de, & 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
- van Leeuwen, M., & Nieuwenhuis, M. (2010). Retrieval of forest structural parameters using LiDAR remote sensing. European Journal of Forest Research, 129(4), 749–770. https://doi.org/10.1007/s10342-010-0381-4
- White, J. C., Coops, N. C., Wulder, M. A., Vastaranta, M., Hilker, T., & Tompalski, P. (2016). Remote Sensing Technologies for Enhancing Forest Inventories: A Review. Canadian Journal of Remote Sensing, 42(5), 619–641. https://doi.org/10.1080/07038992.2016.1207484

References

- Wulder, M. A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N. C., Hilker, T., Bater, C. W., & Gobakken, T. (2012). Lidar sampling for large-area forest characterization: A review. Remote Sensing of Environment, 121, 196–209. https://doi.org/10.1016/j.rse.2012.02.001
- Ye, C., Zhao, H., Ma, L., Jiang, H., Li, H., Wang, R., Chapman, M. A., Junior, J. M., & Li, J. (2022). Robust Lane Extraction From MLS Point Clouds Towards HD Maps Especially in Curve Road. *IEEE Transactions on Intelligent Transportation Systems*, *23*(2), 1505–1518. https://doi.org/10.1109/tits.2020.3028033