EGU2020-11674
https://doi.org/10.5194/egusphere-egu2020-11674
EGU General Assembly 2020

Unlocking modern nation-scale LiDAR datasets with FOSS – the Laserchicken framework

Meiert W. Grootes[1], Christiaan Meijer[1], Zsofia Koma[2], Bouwe Andela[1], Elena Ranguelova[1], and W. Daniel Kissling[2]

[1]Netherlands eScience Center, Science Park 140, 1098 XG Amsterdam, The Netherlands
[2]Institute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, P.O. Box 94240, 1090 GE Amsterdam, The Netherlands

LiDAR as a remote sensing technology, enabling the rapid 3D characterization of an area from an air- or spaceborne platform, has become a mainstream tool in the (bio)geosciences and related disciplines. For instance, LiDAR-derived metrics are used for characterizing vegetation type, structure, and prevalence and are widely employed across ecosystem research, forestry, and ecology/biology. Furthermore, these types of metrics are key candidates in the quest for Essential Biodiversity Variables (EBVs) suited to quantifying habitat structure, reflecting the importance of this property in assessing and monitoring the biodiversity of flora and fauna, and consequently in informing policy to safeguard it in the light of climate change an human impact.

In all these use cases, the power of LiDAR point cloud datasets resides in the information encoded within the spatial distribution of LiDAR returns, which can be extracted by calculating domain-specific statistical/ensemble properties of well-defined subsets of points.

Facilitated by technological advances, the volume of point cloud data sets provided by LiDAR has steadily increased, with modern airborne laser scanning surveys now providing high-resolution, (super-)national scale datasets, tens to hundreds of terabytes in size and encompassing hundreds of billions of individual points, many of which are available as open data.

Representing a trove of data and, for the first time, enabling the study of ecosystem structure at meter resolution over the extent of tens to hundreds of kilometers, these datasets represent highly valuable new resources. However, their scientific exploitation is hindered by the scarcity of Free Open Source Software (FOSS) tools capable of handling the challenges of accessing, processing, and extracting meaningful information from massive multi-terabyte datasets, as well as by the domain-specificity of any existing tools.

Here we present Laserchicken a FOSS, user-extendable, cross-platform Python tool for extracting user-defined statistical properties of flexibly defined subsets of point cloud data, aimed at enabling efficient, scalable, and distributed processing of multi-terabyte datasets. Laserchicken can be seamlessly employed on computing architectures ranging from desktop systems to distributed clusters, and supports standard point cloud and geo-data formats (LAS/LAZ, PLY,

GeoTIFF, etc.) making it compatible with a wide range of (FOSS) tools for geoscience.

The Laserchicken feature extraction tool is complemented by a FOSS Python processing pipeline tailored to the scientific exploitation of massive nation-scale point cloud datasets, together forming the Laserchicken framework.

The ability of the Laserchicken framework to unlock nation-scale LiDAR point cloud datasets is demonstrated on the basis of its use in the eEcoLiDAR project, a collaborative project between the University of Amsterdam and the Netherlands eScience Center. Within the eEcoLiDAR project, Laserchicken has been instrumental in defining classification methods for wetland habitats, as well as in facilitating the use of high-resolution vegetation structure metrics in modelling species distributions at national scales, with preliminary results highlighting the importance of including this information.

The Laserchicken Framework rests on FOSS, including the GDAL and PDAL libraries as well as numerous packages hosted on the open source Python Package Index (PyPI), and is itself also available as FOSS (https://pypi.org/project/laserchicken/ and https://github.com/eEcoLiDAR/).