**LiDAR processing steps**

Identifying fine-scale habitat preferences of threatened butterflies using airborne laser scanning

**Jan Peter Reinier de Vries1, Zsófia Koma1, Michiel F. WallisDeVries2,3 & W. Daniel Kissling1,\***

1Institute of Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, Amsterdam, the Netherlands.

2De Vlinderstichting / Dutch Butterfly Conservation, Wageningen, the Netherlands

3Plant Ecology and Nature Conservation Group, Wageningen University, Wageningen, the Netherlands

\*Corresponding author: E-mail address: wdkissling@gmail.com (W.D. Kissling).

Analyses are performed in the following order:

Data preparation:

* A: Step1\_prepdownload.R : create polygons for the area of interest based on transect locations
* B: Step2\_prepdownloadtiles.R : create a list of tiles which needs to be downloaded from geodata.nationaalgeoregister.nl
* C: Download the tiles
* D: Step3\_download\_and\_intersect.R: extract area of interest from the downloaded AHN3 data

Metrics calculation:

* E: Step4\_calc\_metrics\_25m : metrics calculation at 25 m radius around the observation points
* F: Step5\_calc\_metrics\_100m : metrics calculation at 100 m radius around the observation points

This documentation only aims to demonstrate the LiDAR processing workflow, so only a small set of transects is used. All derived metrics across the Netherlands for the selected butterfly species are provided under the Data preparation directory as the following files:

Butterfly\_lidarmetrics\_v6\_25m.csv

Butterfly\_lidarmetrics\_v6\_100m.csv

***Data preparation***

**A:** Step1\_prepdownload.R

* Input:
	+ a csv file with the spatial location of transects where the butterflies were observed. This has four attributes: Transect, Tr\_sec, x and y. The Transect is the id of the transect, Tr\_sec is the section within the transect, the x and y are the spatial coordinates of the observations in the RD\_New coordinate system
	+ ahn3 shapefile contains the spatial location (polygon) and name of the tiles
* An area of interest polygon is created based on the butterfly observation points and a 250 m buffer around them.The area of interest polygon is extracted for every measured transect. Moreover, the original csv file with the transects is converted into a shapefile.
* Output:
	+ shapefile of area of interest polygon
	+ shapefile of transects

**B:** Step2\_prepdownloadtiles.R

* Input:
	+ shapefile of area of interest polygon
	+ ahn3 shapefile contains the spatial extent and name of the tiles
* Based on the area of interest polygons the ahn3 tiles for download are identified and downloaded. Then a txt file is created with URL addresses to download the data from geodata.nationaalgeoregister.nl database
* Output: a txt file with the list of URL adresses

**C:** Download the tiles

* Using the URL addresses, the files are downloaded using curl in bash (while read p; do curl -o ${p:62:73} "${p%?}";done < ahn3list.txt) or just using the links in the txt file (manually)

 **D:** Step3\_download\_and\_intersect.R

* Input:
	+ shapefile of area of interest polygon
	+ ahn3 laz files organized in the directory
* The point cloud related to the area of interest polygon is extracted
* Output: laz files named with the transect id

***Metrics calculation***

**E:** Step4\_calc\_metrics\_25m

* Input:
	+ Shapefile of transects with Transect, Tr\_sec, x, and y attributes
	+ Area of interest laz files named by transect in the working directory
* Metrics are calculated with 25 m radius.
* Output: csv file with the derived metrics linked to the transects and transect sections.

**D:** Step5\_calc\_metrics\_100m

* Input:
	+ Shapefile of transects with Transect, Tr\_sec, x, and y attributes
	+ Area of interest laz files named by transect in the working directory
* Metrics are calculated with 100 m radius.
* Output: csv file with the derived metrics linked to the transects and transect sections.