EUROPEAN COMMISSION
DIRECTORATE-GENERAL
JOINT RESEARCH CENTRE
Directorate D – Sustainable Resources
Bio-Economy Unit

Measuring *Euclidean Distance.*

Available in the free software *GuidosToolbox*. This is a living document, please check for the latest version at: **https://forest.jrc.ec.europa.eu/en/activities/lpa/gtb/**

Contact: **peter.vogt@ec.europa.eu**, October 2020

1. Background

Patch-based metrics usually describe only certain patch aspects and do not account for outside patch features. This type of metrics suffers from intrinsic limitations valid for all summary indices: a) they provide a single average value only and do not account for the variability of the feature under evaluation, and b) they cannot provide any spatially explicit information. In addition to typical patch aspects, such as patch area and edge length, the assessment on the shape of the patch, including the number and size of potential openings within a given patch, provides important information on intactness of forest patches, continuous interior area as well as the spatial configuration of forest cover. When evaluating a binary map, the foreground objects (i.e. forest patches) should not be analyzed on their own but instead they must be seen within the spatial context of the complementary surrounding background objects (non-forest land). Only a contemporary analysis of foreground and background together guarantees a holistic analysis of the entire system. This document outlines an approach circumventing typical patch index limitations by measuring the Euclidean distance distribution on a binary map.

Scope of this report and for consideration:

Illustration of an analysis methodology - Euclidean Distance and Hypsometric Curve (HMC) - to measure the degree of intactness, shape and spatial arrangement of patches on a given binary map. The methodology outlined in this document has been developed with the goal to

a) bypass typical limitations of existing patch-based metrics.
b) be based on geometric concepts providing a generic framework, which can be applied to any kind of data type and at any scale.
c) provide quantitative measures in a map product, which can be directly related to the degree of fragmentation, intactness, inter-patch distance and spatial configuration.
d) provide a concise summary of foreground as well as background patch attributes: the total number of patches, shape and size distribution and accounting for perforations.

Euclidean Distance and HMC could be used as a state variable to describe forest configuration and fragmentation. The quantitative map product and its summary can also be used to locate and directly measure changes in fragmentation and cover change over time. The outlined methodology may be useful for landscape planning, monitoring and biodiversity studies.

2. Euclidean Distance and Hypsometric Curve

2.1 The methodology

The presence of roads and/or different land cover types generates landscape patterns. On a binary map, i.e. forest – non-forest, there will be a certain number of forest patches of various area, shape (elongated or compact, perforated or intact, etc.), and spatial configuration (clumped in one area of the image or spread all over the image). The various forest patch attributes and their spatial inter-patch distance can be visualized by calculating the Euclidean Distance using a morphological wave front propagation algorithm. Here, we measure the distance from the forest – non-forest boundary into both, the forest patches as well as non-forest areas. The result can be interpreted as a pseudo-elevation map where positive values are associated with land (forest), negative values with sea (non-forest) and a value of zero corresponds to the coast line (forest– non-forest boundary). On any location of this map, the pixel value displays the shortest distance to the nearest forest boundary. Red/brown/dark-green colored pixels indicate compact/intact forest patches while purple colored pixels show areas which are far away from forest patches, see Figure 1. Because every pixel has a distance value, the entire map can be summarized as a distance histogram (Figure 2), showing the frequency of distance values for both, foreground and background. Alternatively, this histogram can be displayed by showing the normalized, cumulative frequency distribution, which is known as the Hypsometric Curve (HMC), a well-established concept in geomorphology and hydrology (Figure 3). The outline and shape of the HMC contemporarily summarizes the spatial attributes and configuration of forest patches as well as their position in the surrounding background. A relatively flat shape of the HMC for the foreground pixels is obtained for few, large and intact forest patches while the presence of many small patches will generate a steeper HMC. Similarly, the shape for the background pixels of the HMC shows the inter-patch distance distribution in the image under study. Both parts of the HMC, foreground and background, can then be summarized with standard summary indices, such as area under the curve, maximum distance and number of objects in the foreground/background, etc.

2.2 The input data:

The input image for the Euclidean Distance analysis must be a binary raster map. It must have the values 1 byte for background (non-forest), 2 byte for foreground (forest) and it may have the value 0 byte to mask missing data (i.e. clouds, ocean or no-data area).

2.3 The output data:

The Euclidean Distance analysis provides:

1. a color-coded map showing distance ranges into and outside of the forest patches
2. a map showing the actual distance values into and outside of the forest patches
3. Euclidean distance bar plot summary image and statistics in a txt-file
4. HMC bar plot summary image and statistics in txt and csv-formatted files

Note: The distance in the resulting products is measured in units of pixels. They can be converted into actual distance [meters] by multiplication with the spatial pixel resolution, if this information is available.

3. Example application

This section illustrates example images and respective results to illustrate the features of Euclidean Distance and Hypsometric Curve.
All image processing was done in GuidosToolbox. The following results can be reproduced by using the example data from the workshop material provided for GuidosToolbox.

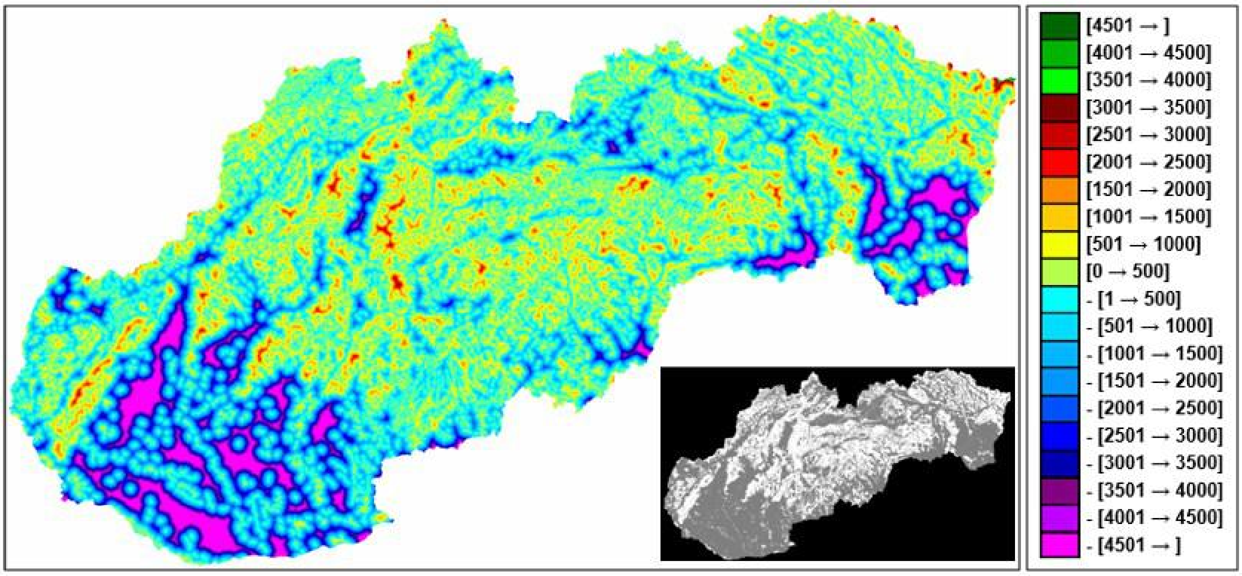
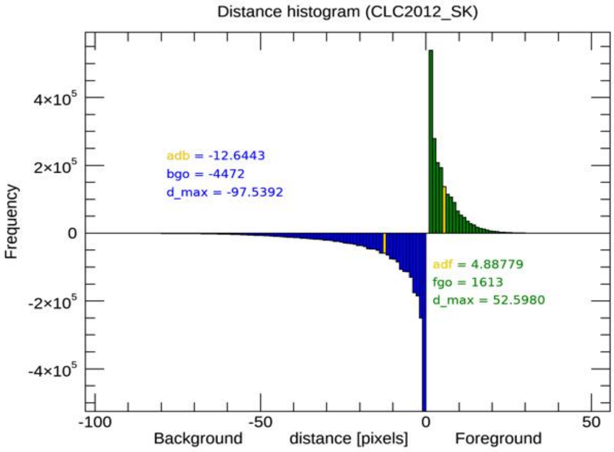


Figure 1: Example of a forest map (gray-colored inlay) in Slovakia (CORINE, 100m resolution) showing the Euclidean Distance measured from the forest boundary into the forest (yellow, red, green) and outside the forest (blue, purple). The color table to the right shows distance range groups in meters, denoted positive inside forest and negative outside forest).

Figure 1 shows the Euclidean distance distribution for a forest map derived from CORINE land cover map in 2012 having a spatial resolution of 100m per pixel. The image in Figure 1 is designed to provide a visual overview by using color-coded distance ranges. This enables to quickly pinpoint the locations and extent of the most compact forest areas as well as those non-forest areas which are far away from any forest cover. A second map output (not shown here) provides the same image as shown in Figure 1 but with actual distance values, which can be used for additional post-processing. For example, a user could be interested to detect compact interior forest cover, i.e. defined as those forest pixels that are more than 500m away from the forest boundary. Then such a mask of compact interior forest is easily obtained by applying the threshold value of > 500m to the distance map, visually corresponding to the colors yellow, red and green in Figure 1.
In a similar way, a user interested in investigating forest edge habitat of 500m can quickly obtain such a map by applying a threshold < 500m to the distance map, visually corresponding to the light-green color in Figure 1.

In addition to the two spatial map products described above, the Euclidean Distance analysis provides statistical summary results in form of:

- a bar plot and a txt-file listing the Euclidean distance frequency distribution (Figure 2)
- a bar plot of the Hypsometric Curve (HMC) (Figure 3) and a txt and csv-file listing the HMC summary attributes (Figure 4).




```
Euclidean distance result using file: CLC2012_SK
Background: average distance, number of objects, maximum distance:
-12.6443  -4472  -97.5392
Foreground: average distance, number of objects, maximum distance:
 4.88779   1613   52.5980
Distance histogram (rounded to nearest integer, negative values for background)
      bin ID      frequency
        -98      -1.00000
        -97      -19.0000
        -96      -61.0000
          ⋮          ⋮
         -4      -175166.
         -3      -184565.
         -2      -250062.
         -1      -525387.
          0       0.00000
          1       539140.
          2       278831.
          3       208230.
          4       193318.
          ⋮          ⋮
         49       12.0000
         50       6.00000
         51       3.00000
         52       3.00000
         53       1.00000
```

Figure 2: Euclidean distance histogram bar plot diagram (left) and txt-file summary (right). Background attributes are shown in blue color and foreground attributes in green color.

Figure 2 shows the distance histogram summary of the Euclidean distance map in Figure 1.

1	1	1	1	1
1	2	2	2	1
1	2	3	2	1
1	2	2	2	1
1	1	1	1	1

The small image to the left shows a single square-shaped forest patch of length 5 pixels. Starting from the forest boundary, there are 16 pixels of distance 1, 8 pixels of distance 2 and 1 pixel of distance 3. Figure 2 shows the same analysis scheme but conducted for all forest pixels and all non-forest pixels in Figure 1.

The average distance value for all background/foreground pixels is denoted adb/adf (here: 12.6/4.9 pixels) and shown in yellow in the bar plot image. The total number of background/foreground objects is denoted bgo/fgo (here: 4472/1613) and d_max denotes the maximum distance found (here: 98/53); the txt-file provides the counts per distance value.

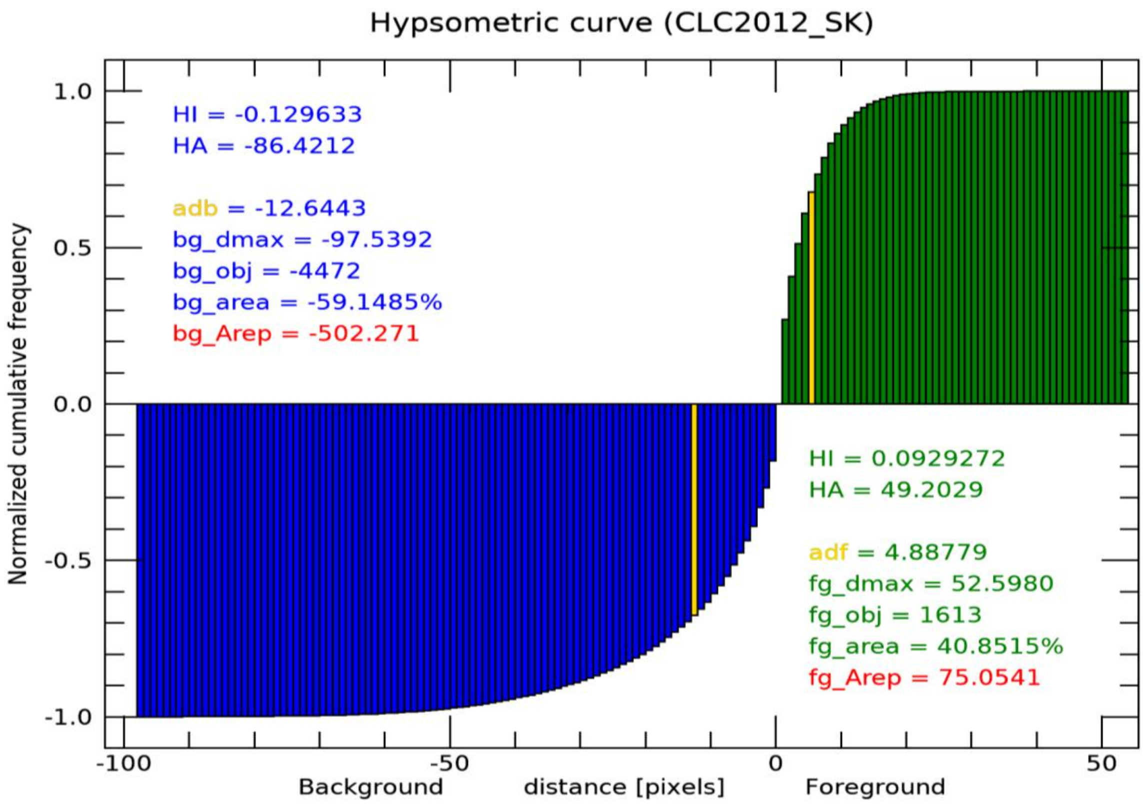


Figure 3: Hypsometric curve of the Euclidean distance map. Background attributes are shown in blue color and foreground attributes in green color.

Figure 3 shows the HMC of the Euclidean distance map in Figure 1. The HMC is derived from Figure 2 by building the cumulative and normalized distance. For example, and using the foreground pixels (green), we can find that ~12% of all forest pixels are border pixels (distance = 1 pixel), or ~50% of all forest pixels are within 3 pixels distance to the forest border. Like Figure 2, the HMC provides additional information for both, background and foreground:

- HI: hypsometric index: adb/bg_dmax or adf/fg_dmax
- HA: hysometric area: The integral area below the HMC outline
- adb/adf: The Average Distance in all background pixels and all foreground pixels
- bg_dmax/fg_dmax: The maximum distance found in all background pixels and all foreground pixels; i.e. fg_dmax corresponds to the value of the forest pixel at the most compact/intact forest area and bg_dmax shows the distance value of a non-forest pixels which is furthest away from a forest patch.
- bg_obj/fg_obj: The total number of objects/patches in the image
- bg_area/fg_area: The total area of background/foreground (here: 59% and 41%)
- bg_Arep/fg_Arep: The Representative Area of background/foreground

The value for the Average Distance in the foreground could be used as an indicator for fragmentation. For example, if a landscape is composed of large patches then the Average Distance value will be larger compared to a landscape having small forest patches.
The Representative Area is proportional to the Average Distance and simply another form of summarizing the distance distribution: The Representative Area is equal to the area of a circle with a radius of adf/adb. In other words, and for the example shown here: if the entire forest cover of Slovakia is equally distributed into circular-shaped patches having the average radius adf, then each of these patches has the Representative Area of 75.05 hectare (using the pixel resolution of 100m ⇔ pixel area of 1 hectare). As with the Average Distance, changes in the Representative Area could be used as a measure for fragmentation processes.

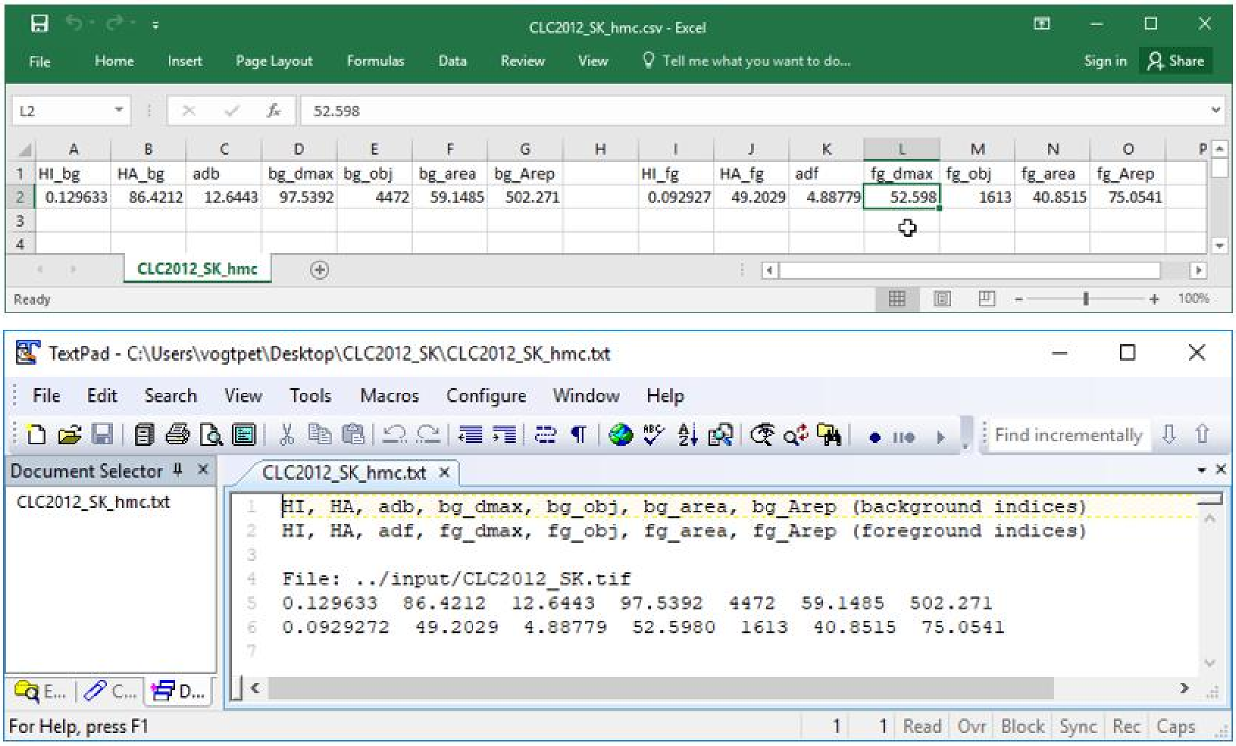


Figure 4: Hypsometric curve summary statistics in csv- and txt-format.

Figure 4 shows HMC summary statistics extracted from Figure 3 in csv and txt-format.

Motivation/Purpose: A user or policy maker may be interested to detect geographic areas where forest patches are of larger extent (more intact) and/or located relatively close to each other. Such areas could then be selected as focal points in political directives targeted at landscape conservation and restoration. When looking at temporal trends the described methodology may be suitable to investigate the impact of land cover change on the degree of fragmentation of existing forest cover: the division of larger patches, and especially the introduction of holes into intact forest patches, will lead to a decrease in the Average Distance value or the Representative Area. In biodiversity and habitat suitability studies, the assessment of the impact of habitat isolation on population survival is of high interest. Here, the Average Distance of the background can provide a comprehensive summary index addressing the distance distribution between existing habitats, i.e. an increasing value implies more isolation. The contemporary analysis of both, foreground and background, as well as inter-patch distance and patch area and shape is also of importance for connectivity studies: larger patches provide more inner-patch connectivity and shorter inter-patch distance facilitates the movement between patches.
The provision of geospatial maps in addition to simple summary indices is of key relevance for any kind of land cover planning: on individual images, it allows locating hotspots of highly fragmented or dispersed patches. When analyzing time series, it is possible to delineate areas of no change from areas where changes are manifold. This fact may also be useful to investigate if a given landscape planning measure, applied at a given region, in fact had a quantifiable influence on the specific land cover type in that region. Finally, the geometric nature of this simple, yet comprehensive analysis scheme, permits the application to any kind of land cover map (forest, wetlands, urban green-space, etc.) and at any scale (local, regional, continental). Using distance from the forest border is intuitive, easy to understand and hence easy to communicate to end-users or the general public. The potential to locate and quantify changes over time, directly compare different sites and to *measure* the progress and outcome of political directives may also be beneficial for the design of policy programs or landscape planning projects targeted at conservation, restoration, or investigating the potential impact of land cover changes on and near protected natural parks.

4. Discussion

This document describes a methodology for the contemporary assessment of patch area, shape and inter-patch distance distribution. The approach is based on measuring the Euclidean distance into and away from the foreground boundary. It accounts for the number and relative size of all objects (foreground and background), including their morphometric aspects such as perforations and elongated versus compact form. While the maps provide information on the location and configuration of the spatial setup, the Hypsometric Curve provides a concise summary of the distance distribution as well as statistical details and summary indices designed to provide key messages on the mutual inter-dependence of foreground and background and to measure changes in the count, area and relative contribution of the foreground patches in the image.

The Euclidean distance and HMC assessment scheme provide:

- **Generic concept:** The methodology is based on geometric principles only. As such, it can be applied to any kind of forest raster maps, independent of the definition of forest and the spatial resolution of the forest map. In contrast to many existing fragmentation schemes, the outlined methodology provides quantitative measures and is not a descriptive measure for a specific faunal species.

- **Distance statistics:** The outlined methodology provides a summary of the entire landscape configuration and detailed statistics, including frequency of occurrence, area, shape of each object in the image.

- **Distance maps:** Providing spatially explicit maps allows for highlighting hotspots of highly fragmented areas or those dominated by undisturbed natural land cover. The spatial information of the map product may be of high importance for monitoring, planning and risk assessment. For example, the feature to quantify changes across time series is an essential requisite to measure progress or the overall effectiveness of political directives.

- **Communication:** The simple, yet intuitive analysis scheme is easy to communicate and can be related to a variety of land cover planning measures, change in time series of biodiversity or other environmental indicators and direct comparisons with results from other regions. The assessment of distance contributes to answering questions like:
  - **State Analysis:** How high is the degree of fragmentation or intactness in different administrative units/counties or ecological regions of the country?
  - **Trend Analysis:** Where and how much have compact forest patches changed over the past decades? How are the trends, and what can be expected for the future?
  - **Monitoring & Assessment:** How strong was the impact of a specific political directive or planning program in the targeted area? How big is the change outside of the monitored region? Does the result of the program merit the money spent? With these findings, what are the implications for future planning and which areas should be targeted first?

5. Conclusions

The Euclidean Distance & HMC analysis scheme is available in the free JRC software *GuidosToolbox*. The software can be downloaded and used by anybody and for any kind of analysis. All data analysis schemes in GuidosToolbox are based on geometric principles, which permits processing categorical maps of any kind and spatial resolution.

With the outlined setup, and the availability in a free software, each user can test the described methodology and its settings on their own data. In addition to Euclidean Distance, GuidosToolbox provides dedicated routines for spatial mapping and quantification of pattern, fragmentation, connectivity and other aspects derived from land cover maps.
Further information is available in related product sheets available on the GuidosToolbox homepage.

References:

Vogt P., Riitters, K., 2017. GuidosToolbox: universal digital image object analysis. European Journal of Remote Sensing 50:1, 352-361, DOI: 10.1080/22797254.2017.1330650.

Kozak, J. et al., 2018. Forest Cover Increase Does Not Trigger Forest-Fragmentation Decrease: Case Study from the Polish Carpathians. Sustainability 2018, 10, 1472, DOI: 10.3390/su10051472.