Graphab 2.8
User Manual

Céline Clauzel, Jean-Christophe Foltête, Xavier Girardet, Gilles Vuidel

2022-05-11

Contents

1 Introduction

1.1 About Graphab

Graphab is a software application for modeling ecological networks using landscape graphs. It is composed of four modules for:

- constructing graphs, including loading initial landscape data and identifying patches and links (Euclidean distances or least-cost paths)

- computing connectivity metrics from graphs

- integrating graph-based connectivity metrics into species distribution models

- visual and cartographic interfacing

1.1.1 Authors

Graphab has been developed by Gilles Vuidel and Jean-Christophe Foltête at ThéMA laboratory (University of Franche-Comté – CNRS). Funding has been provided by the French Ministry of Ecology, Energy, Sustainable Development and the Sea (ITTECOP Program). The Graphab logo was designed by Gachwell.

1.1.2 Terms of use

Graphab is distributed in open source, under the GPL license. Users must cite one of the following references [Foltête et al.(2012a), Foltête et al.(2021)] in their publications:
Foltête J.-C., Vuidel G., Savary P., Clauzel C., Sahraoui Y., Girardet X., Bourgeois M. 2021. Graphab: an application for modeling and managing ecological habitat networks. Software Impacts.8: 100065.
Foltête J.C., Clauzel C., Vuidel G., 2012. A software tool dedicated to the modelling of landscape networks, Environmental Modelling & Software, 38: 316-327.

1.2 System requirements

Graphab runs on any computer supporting Java 8 or later (PC under Linux, Windows, Mac, etc.). However, when dealing with very large datasets, the amount of RAM memory in the computer will limit the maximum number of nodes and links that can be processed in a single run with Graphab. In addition, for some complex metrics, processing power (CPU) will determine the speed of computing. For details, see section Processing capabilities and limitations below and the journal article [Foltête et al.(2012a)].

1.3 Installing the software and launching a project

Graphab can be downloaded from https://sourcesup.renater.fr/www/graphab.

- Download and install Java 8 or later - adoptium.net. It is best to install the 64-bit version of Java.

- Download graphab-2.8.jar

- Launch graphab-2.8.jar

After launching graphab-2.8.jar, the File menu provides access to four sections:

- File / New project: to create a new project in which all data and results are saved automatically.

- File / Open project: to open an existing project.

- File / Preferences: to change certain software parameters: English/French; maximum amount of memory to use; number of processors to use. It is recommended to adjust the amount of memory and number of processors to suit your computer (see section Processing capabilities and limitations).

- File / Log window: to display the event log.

2 Starting a Graphab project

New projects are created from the File / New project menu. The user must complete a series of windows to identify the project, import a landscape map, and create a link set. Each project is associated with a single landscape map but may contain several link sets. After the start phase, the project is the medium for creating multiple graphs and for computing connectivity metrics.
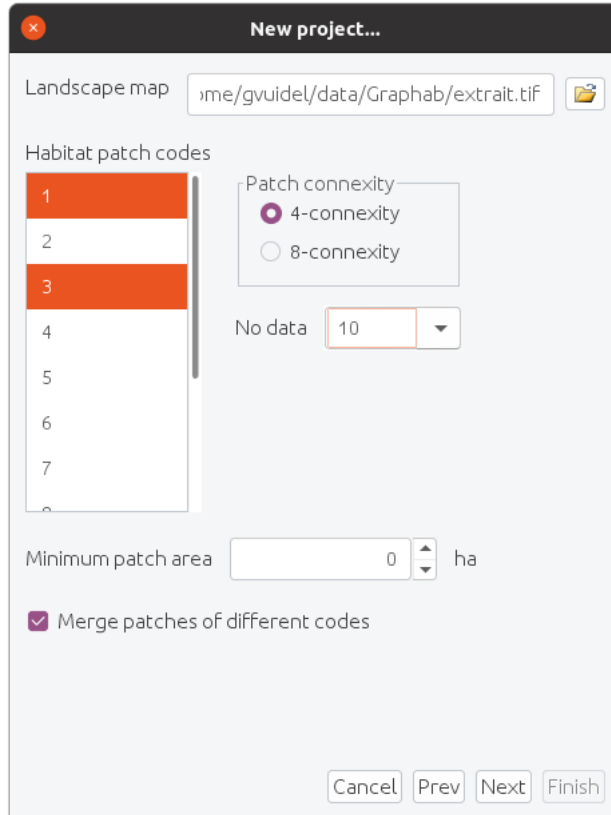
2.1 Identifying a project

In the first window, the user must enter a project name and specify the folder in which it is to be created.
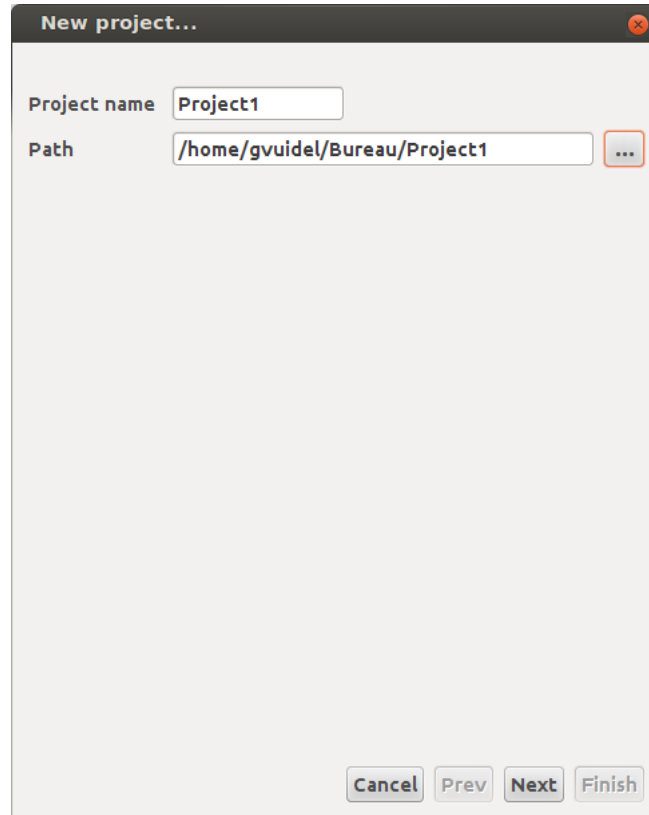
2.2 Importing landscape maps and defining nodes

The second window is for importing the landscape map. It must be a raster file (*.tif, *.asc, *.rst) in which the value of each pixel corresponds to a category (land cover or other classification).

If the raster format is *.tif without a Geotiff extension, the file must be associated with a world file for geolocation (*.tfw) structured as follows:

Example	
10.0	Pixel size in the X-direction
0.0	Rotation about X-axis
0.0	Rotation about Y-axis
-10.0	Pixel size in the Y-direction
821755.0	X coordinate of the center of the upper-left pixel
2342995.0	Y coordinate of the center of the upper-left pixel

If the raster format is *.rst, the file must be associated with a georeferencing file (*.rdc) generated by Idrisi software.

The units of the image coordinate system must be meters. If not, the areal and distance units will be incorrect. The image can be reprojected in a metric projection (UTM, Lambert93) using GIS software.

No data: pixel value representing the absence of data in the raster file.

Habitat patch codes: pixel values assigned to the habitat category used to define habitat patches. Several values can be selected by using the `Ctrl` key.

Minimum patch area: minimum area in hectares for a habitat patch to become a graph node.

Patch connexity:

- 4-connexity: a habitat patch consists of the central pixel with its four neighbors if they are of the same value;

- 8-connexity: a habitat patch consists of the central pixel with its eight neighbors if they are of the same value.

2.3   Creating link sets

The third window is for creating a link set for which several parameters must be selected: topology and link weighting. Creating a link set is the final step in starting a Graphab project. However, users may create new link sets within the same project via the Graphs / Create link set menu.

2.3.1   Link topology

Two topologies are available:

- planar: only links that form a minimal planar graph are considered. This topology is set up through Voronoi polygons around each habitat patch. These polygons are defined from the edges of patches in Euclidean distance.

- complete: all the links between patches are potentially taken into account.

Max distance: this option specifies a threshold distance. Links that exceed this distance are no longer created. This limits the number of links created and so accelerates the creation of the link set. The unit of the distance depends on the type of distance: in meters for the euclidean distance and in cost for the least cost distance.

Ignore links crossing patch: this option means that a link between two patches (A and C in the figure below) which crosses an intermediate patch (B) is not created. It is recommended for calculating the betweenness centrality metric (BC) to take into account how often a patch lies on the shortest path between all pairs of patches in the graph. If the option is unchecked, a link is created between two patches (A and C) crossing an intermediate patch, representing the complete true distance between A and C.

Checked option : links crossing patches are removed   Unchecked option : links crossing patches are kept
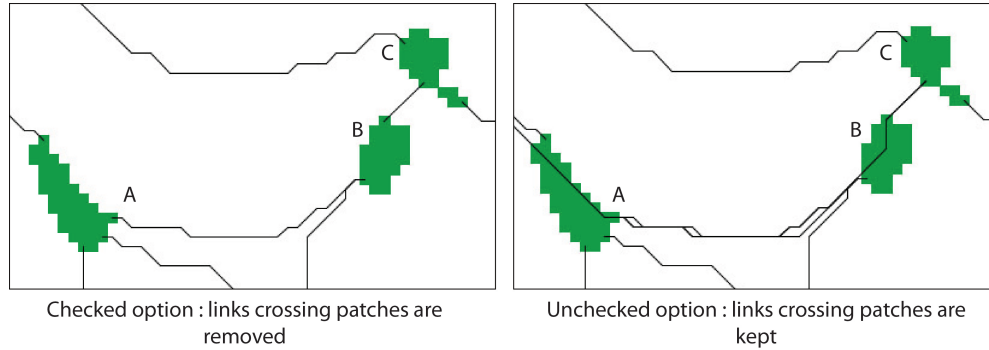
Save real path:

- checked box: links are saved as paths representing the actual route of the link between two patches.

- unchecked box: links are saved in topological form only. In this case, display of links with the realistic view is unavailable. This is recommended for graphs with very many links (e.g. a non-thresholded complete graph) so as to limit the use of memory. Unless paths are saved, intra-patch distances cannot be included in the computation of metrics.

2.3.2 Distance (or link impedance)

Distances are calculated from edge to edge between patches. Two main types of distance are available: Euclidean distances and least-cost distances.

- Euclidean distance: links are defined in Euclidean distances (distance as the crow flies between patches), meaning the matrix is considered to be uniform.

- Least-cost distance: links are defined in cost distances. This distance is equal to the sum of the costs of all the pixels along the least-cost path. Matrix heterogeneity is taken into account by assigning a resistance value (friction) to each landscape category. The user can activate this option in either of two ways:

    1. either by specifying different costs for the landscape map categories in the table,

    2. or from an external raster file (*.tif, *.asc or *.rst) in which each pixel has a resistance value.

For each link created, its metric length (`DistM`) and its cost-unit distance (`Dist`) are saved and available in "link properties" (see Object properties).

Resistance values previously defined can be weighted by slope [Clauzel et al.(2015b)]. To compute slope, the DEM should be imported from the menu Data | Set DEM raster.

The parameter coef (c) can adjust the importance of the slope (p) weighting. For a given pixel, the resulting resistance value (r_{final}) is computed given the formula:

$$r_{final} = r * (1 + c.p)$$

with the slope $p = \frac{h}{l}$ ratio of the height (h) and the length (l).
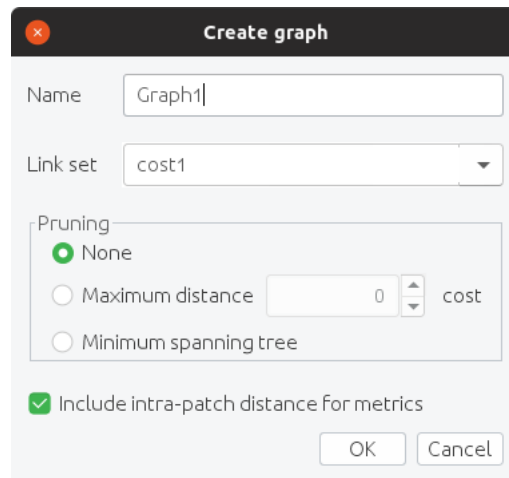$p = 0$ if the slope is flat, $p = 1$ for a slope of 100% ($h = l$).

When $c = 1$ the resistance value is doubled for a slope of 100%.
When $c = 10$ the resistance value is doubled for a slope of 10%.

3 Graphs

3.1 Creating graphs

A Graphab project may entail the creation of several graphs. Each graph is created from a given link set: either the link set defined in the initial project, or a new link set defined from the Graph / Create link set menu.

Graphs are created from the Graph / Create graph menu.

First, the new graph must be named.

The user must select one of the link sets created previously (cf. Creating link sets) and then select the type of pruning:

- None: all links between patches are validated, regardless of length.

- Maximum distance: the selected links are less than or equal to the selected maximum distance.

- Minimum spanning tree: graph connecting all the patches in which the total weight of links is minimal.

"Include intra-patch distances for metrics" option: if the box is checked, the computation of metrics includes the distances between and across patches (recommended). If the box is unchecked, only the distances between (but not across) patches are taken into account.

For a pruned graph with maximum distance, the unit of the distance depends on the type of distance used in creating the link set. If the link set is created using Euclidean distances, the maximum distance is in meters. If the link set is created using cost distances and the impedance is cumulative cost, the maximum distance is given as a cumulative cost. The unit of the distance (cost or meter) is shown just after the "max dist" field.

An approximation of the distance metric (DistM) expressed as a cumulative cost (Dist) can be obtained by displaying the double log plot of the link set and using the equation below to perform the conversion:

$$Dist = e^{intercept+slope.log(DistM)}$$
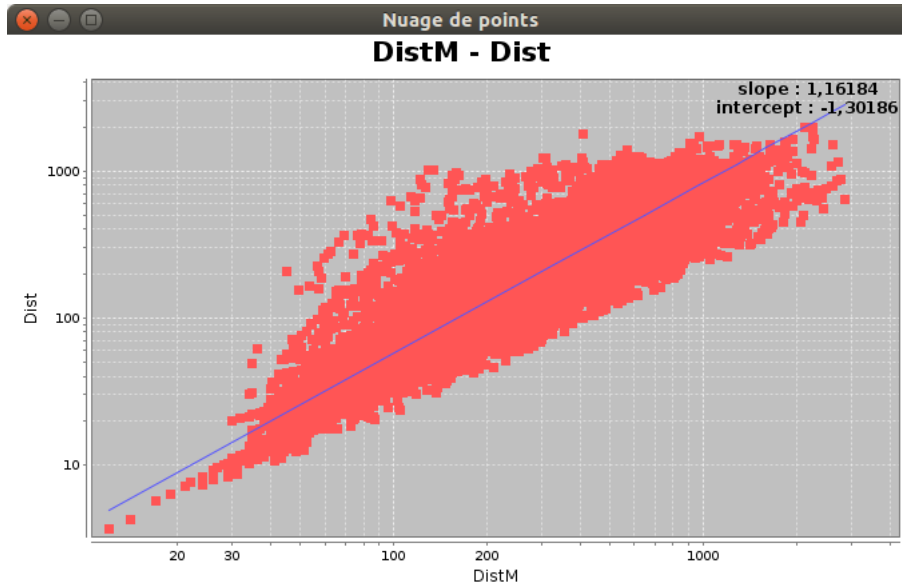
The estimation can be performed directly in Graphab from the menu "Dist. conversion" available by right-clicking on a link set name.
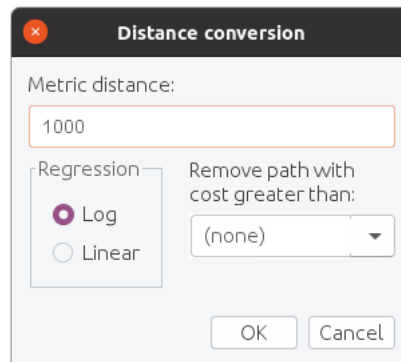
The parameter "Remove path with cost greater than" allows to exclude from the regression the paths crossing pixels with cost greater or equal to the given value. It is useful to exclude links crossing barrier elements from the distance conversion. This parameter cannot be used with linkset using external cost raster.

To perform a multiscale analysis, it is often necessary to create a series of graphs in which increasing maximum distances are defined. Users can create this series manually. But if the objective is to analyze the behavior of a metric according to the maximum distance, the Metrics / Batch graphs menu can be used (see Batch graph).

3.2   Graph clustering

Graphab implements a clustering algorithm that maximises the modularity index [Newman(2006)]. Modularity is a measure of the quality of the clustering of the graph nodes. The underlying principle is that a good clustering involves a large number of links within groups and a small number of inter-group links. The algorithm used is based on the greedy algorithm, followed by a local optimization [Brandes et al.(2008)].

Modularity is calculated with a weight (w_{ij}) defined for each link of the graph:

$$w_{ij} = (a_i a_j)^{\beta} e^{-\alpha d_{ij}}$$

The two parameters β et α are used to define respectively the importance of the patch capacity $(a_i a_j)$ and the importance of the distance (d_{ij}) for the weight of the link w_{ij}. If $\alpha = \beta = 0$, then the weights are identical: $w_{ij} = 1$.

The graph clustering is available by right-clicking on the name of the graph. In the first window, the user defines the parameters α et β. The parameter α is not defined directly, it is determined from two other parameters: a distance d and a probability p (cf. Parameters of weighted metrics). To define α to 0, p must be equal to 1.

The right part of the window allows to select which clustering to keep in memory. By default, all clustering are kept in memory to display them after. These settings are useful to avoid memory overflow with large graphs (more than 5,000 nodes). In this case, graphab can keep only the best clustering (ie. with the best modularity) or keep all clustering in a given size interval between 1 and the number of nodes.
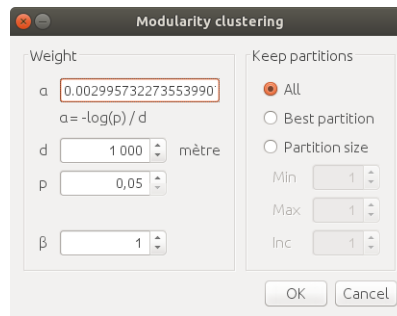
After validating this window, the calculation starts until a clustering maximizing modularity is obtained. The result is displayed as a new layer of the graph. Two other windows complete this result.
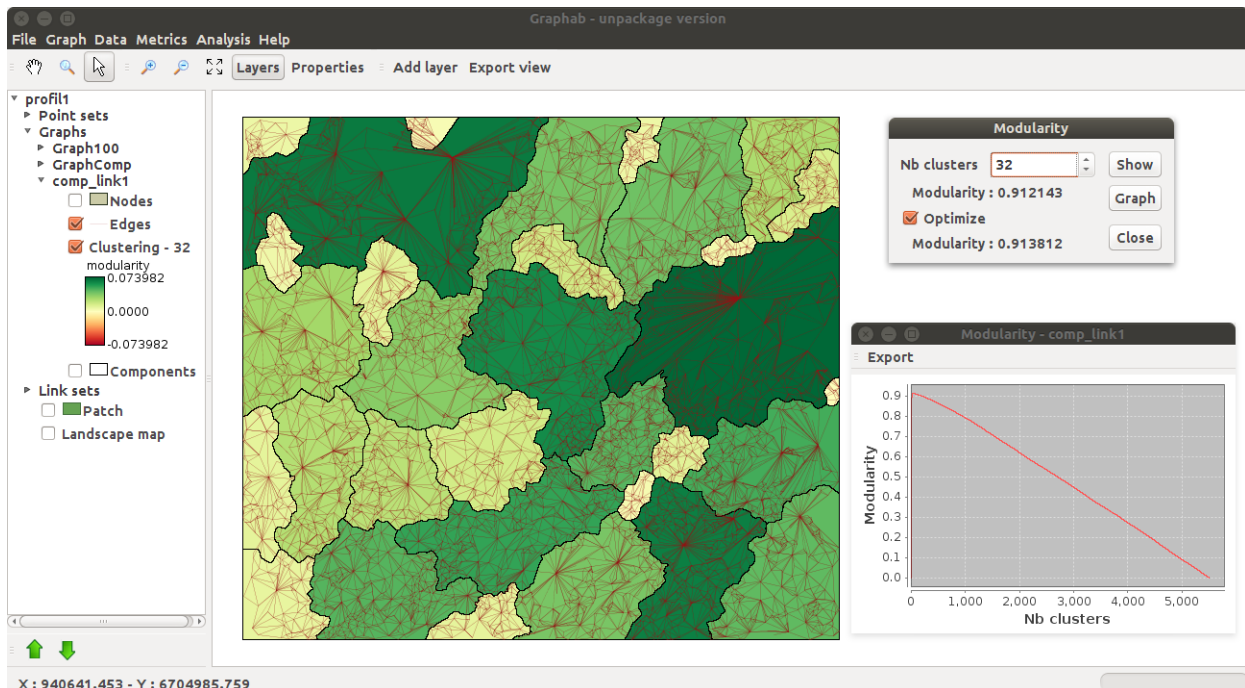
The new layer shows the elements of the spatial partition as polygons. Since a given element corresponds to a set of patches, it is represented by a polygon resulting from the aggregation of the Voronoï's polygons of all patches of the cluster. The modularity value is used to assign a color to the polygon. The overall modularity is the sum of the modularity values of all clusters.

In a first window, a graph shows how modularity varies according to the number of clusters. In the previous figure, the maximal modularity is reached with a small number of clusters.

The second window displays the number of clusters that maximizes modularity and the value of modularity. Several functions are available from this window:
- the number of clusters can be changed to show sub-optimal partitions. The result may be visualized by clicking on the button "Show". The new layer of partition is added to the graph.
- the button "Graph" allows you to create a new graph including only the intra-cluster links.
- the checkbox "Optimization" achieves a local optimization of the clustering after the use of the greedy algorithm, that significantly improves the modularity value. This option is enabled by default.

4   Corridors

Graphab allows to calculate the corridors representing, for a given maximum distance d_{max}, the area that can be traversed between 2 patches habitat *ie.* the area representing the set of possible paths connecting 2 patches and having a distance less than the distance d_{max}.

Corridors can be calculated only from a linkset that has been defined from costs. To use this feature, just right click on a linkset and click on "Corridor". The maximum cost distance d_{max} must be entered before launching the calculation.
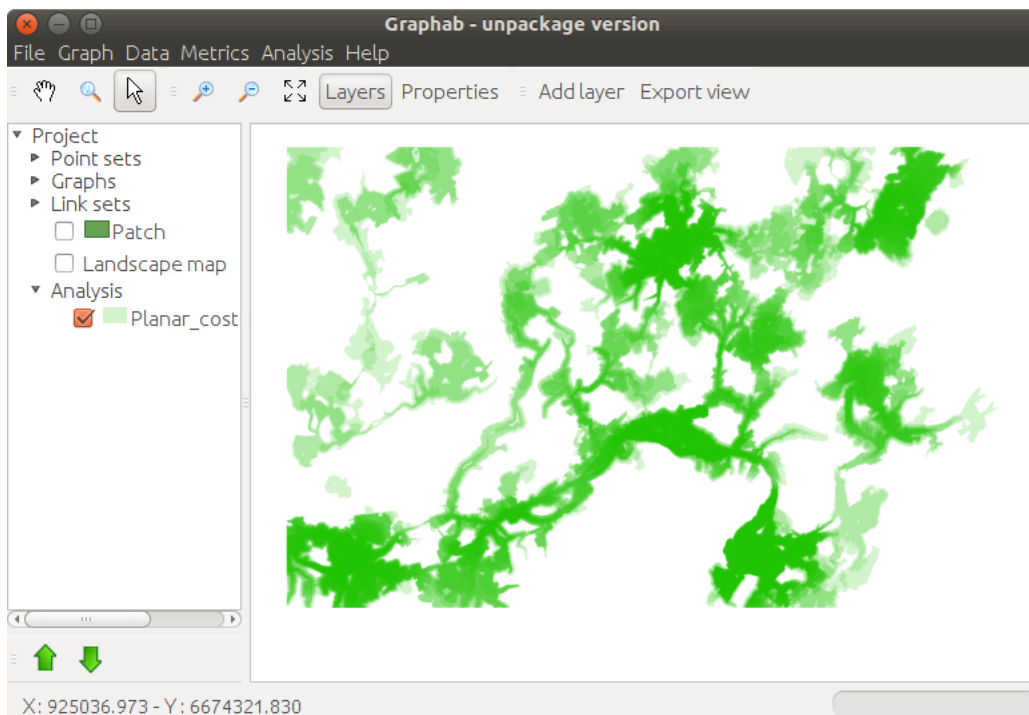
The result is displayed in a new vector or raster layer depending on the selected output. In vector output, the layer contains for each link of the selected linkset a polygon representing its corridor. Links with a distance greater than d_{max} will have no corridor. In raster output, each pixel contains the number of corridors passing through it.

For vector layer, the default view uses transparency to show corridor overlays.

5 Patch capacity

The capacity of a patch reflects its intrinsic quality as an indicator of its demographic potential. A patch with a high capacity can accommodate a large population and vice versa. Capacity is included directly in the calculation of some area connectivity metrics and weighted connectivity metrics (see Metrics family and computing level).

When the project is first created the patch capacity is equal by default to the patch area in m². However, users may replace area by any other quality indicator. In some cases, species presence is related not to patch size but to the area of other types of land cover around the patch. For example, the presence of amphibians in a breeding pond does not depend on the pond size but on the amount of terrestrial habitat surrounding the pond.

The Data / Set patch capacity menu can be used to modify the capacity of all patches by two methods.

5.1 Capacity as a function of the neighborhood

The first method can be used to define patch capacity as a function of the neighborhood composition and to calculate it directly from Graphab.

Users must define three parameters: type of distance, maximum distance, landscape categories.

Cost: This is the spatial metric (Euclidean or cost distance) corresponding to one link set available in the project. The use of costs in this procedure amounts to defining an anisotropic neighborhood around patches which may differ greatly from a buffer function. For consistency, it is recommended to use the same type of distance as was used in creating the links of the graph. For a link set created with Euclidean distance, the user must select "all costs = 1".

Max cost: Like the graph maximum distance, the unit of this maximum distance depends on the type of distance used in creating the link set (Euclidean or cost distance).

Codes included: the user may select one or more landscape categories, other than the habitat category, to be included in calculating capacity.

The "cost weight" option introduces a weighting with distance to the patch through a negative exponential function. In this way, the areas selected have greater weight if they are close to the patch and vice versa.

The capacity values calculated replace the patch area for all subsequent computations. But users can return to the initial parameter via the Data / Set patch capacity menu and by selecting "Patch area".

5.2   Capacity defined from external data

The second method is for importing a data table (*.csv) describing all the patches of the project and containing capacity values defined in advance by the user. The patch identifiers in the table must be the same as the patch identifiers in the Graphab project.

The capacity values in the imported table replace patch area values for all subsequent computations after importing. But users can restore the initial parameter via the Data / Set patch capacity menu and by selecting "Patch area".

5.3   Patch removing

After setting patches capacity, it can be useful to remove patches whose capacity is too low. The entry "Remove patches" from the Data menu, lets you create a new project identical to the current project and remove patches with a capacity below a given threshold.

After entering the required minimum capacity, the new project is created in a subdirectory of the current project and directly loaded.

6   Calculating connectivity metrics

6.1   Metrics family and computing level

Each graph in a project can be used to compute different connectivity metrics. The details of how they are computed and references are listed in the Annex. Computations are made at several levels corresponding to major sections in the Metrics menu (table 1):

- Metrics /Global metrics: describe the entire graph.
- Metrics /Component metrics: describe connectivity within each component (or sub-graph).
- Metrics /Local metrics: describe the connectivity of each graph element (node or link).

- Metrics /Delta metrics: also describe each graph element, but using a specific computing method. Using the removal method (remove nodes or remove links), the relative importance of each graph element is assessed by computing the rate of variation in the global metric induced by each removal. The result of a delta-metric is at a local level but by reference to the global level.

After selecting one of these four computing methods, three families of metrics are available in the new window:
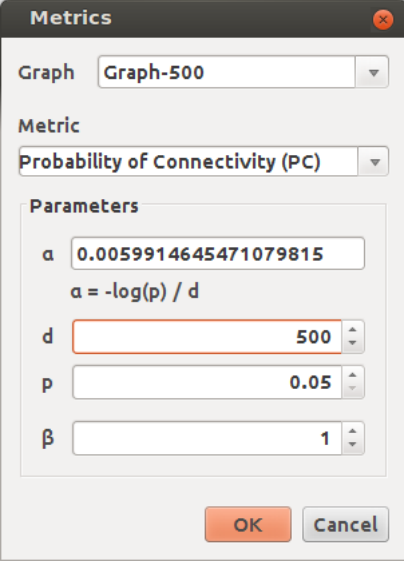
- weighted metrics are based on criteria of distance and patch capacity. They have to adjusted to suit the reference species. These metrics involve computing paths in a graph via Dijkstra's algorithm. After selecting one of these metrics, the user must specify the desired adjustment.

- area metrics are based primarily on the area criterion. If capacity corresponds to a criterion other than patch area, these metrics can be computed and they are expressed in the unit of the criterion used.

- topological metrics are derived from graph theory and they do not require adjustment.

Whichever the selected level, the user must first specify the graph on which the calculation will be made and then select the connectivity metric.

Family	Connectivity metrics	Code	Computing level			Delta metrics
			Global	Comp.	Local	
Weighted	Flux	F	×	×	×	
	Equivalent Connectivity	EC	×	×		×
	Probability of Connectivity	PC	×	×		×
	Interaction Flux (replace FPC)	IF			×	
	Fractions of delta Probability of Connectivity	dPC				×
	Betweenness Centrality index	BC			×	
	Integral Index of Connectivity	IIC	×	×		×
	Current Flow	CF			×	
Area	Mean Size of the Components	MSC	×			
	Size of the Largest Component	SLC	×			
	Class Coincidence Probability	CCP	×			
	Expected Cluster Size	ECS	×			
Topological	Node Degree	Dg			×	
	Clustering Coefficient	CC			×	
	Closeness Centrality	CCe			×	
	Eccentricity	Ec			×	
	Connectivity Correlation	CCor			×	
	Number of Components	NC	×			
	Graph Diameter	GD	×	×		×
	Harary Index	H	×	×		×
	Wilks' Lambda	W	×			

Table 1: Connectivity metrics and computing level

6.2 Parameters of weighted metrics

6.2.1 Alpha parameter

Several metrics include a weighting in their calculation which converts the distance between patches into the probability of movement. These metrics are F, IF, PC, EC, BC... The weighting is based on an exponential function:

$$p = e^{-\alpha d}$$

where p is the probability of movement between two patches, d the distance between these patches, and α a parameter defining the rate of decline in probability as distance increases. As it is not easy to determine the value of the α parameter, Graphab calculates it from the other two parameters. Users must specify the distance corresponding to a certain value of probability, e.g.:

- the maximum dispersal distance of species corresponding to a small value of p (0.05 or 0.01).

- the average dispersal distance of species corresponding to a median value of p (0.5).

The value of α is automatically obtained from the formula:

$$\alpha = -\log(p)/d$$

6.2.2 Beta parameter

The metrics F, IF, CF, and BC are controlled by the β parameter. This parameter is the exponent applied to patch capacity. It adjusts the relative balance between the weight of distances and the weight of patch capacity in the weighting of metrics. Taking the example of the metric F in local computation, whose generic form is:

$$F = \sum a^\beta e^{-\alpha d}$$

- a value of $\beta = 0$ means that the patch capacity plays no part in the weighting.

- a value of $\beta = 1$ means that the patch capacity acts linearly in the weighting.

- a value of $\beta = 2$ means that the patch capacity is squared in the weighting.

- a value of $\beta = 0.5$ means that the square root of the patch capacity features in the weighting.

- a value of $\beta = -1$ means that the patch capacity acts in an inversely proportional way in the weighting.

In addition to these few examples, any weighting values are possible.

6.3 Calculating batch metrics

Every metric compatible with the global level can be calculated following the variation of the scale of distances. This variation may concern either graph pruning (6.3.1), or metric adjustment (6.3.2). The type of distance used depends on the type of distance used in creating the link set (Euclidean, least-cost distance, or least-cost path).

6.3.1 Batch graph

The Metrics / Batch graph menu allows a series of pruned graphs to be created from a given link set and a metric to be calculated for each graph at global level. The maximum distance of successive graphs are increasingly defined in either of two ways:

- distance: a fixed increment of distance is defined between successive graphs. The metric values are therefore calculated inregular intervals of distance.

- number of links: a fixed number of links is defined between successive graphs. This number of links is automatically converted into distance used to prune the graph. These distances may be unevenly spread.

Graphs are defined following three criteria selected by users:

- min: smallest distance used for the first graph in the series. By default, this minimum is 0, corresponding to the total absence of links.

- max: maximum distance used for the final graph in the series. By default, this maximum corresponds to the maximum distance or to the link number of the selected link set.

- increment: distance value added between each new graph.

Once the calculation is completed, the software opens a new window displaying the curve of the selected metric versus the maximum distance. The values of this curve can be saved with the "Export" button by selecting text format.

6.3.2 Batch parameter

The Metrics / Batch parameter menu is used to calculate a series of metrics from a given graph. This procedure applies to the weighted metrics only. It is divided into two entries: local metrics or global metrics.

Batch parameter for local metrics

A local weighted metric is calculated in series according to the variation of one of its parameters. The user must select the graph, the metric, and the parameter to be varied. The variation of computation is defined by:

- min: minimum value of the parameter,

- max: maximum value of the parameter,

- increment: interval value between two metric computations.

Once the calculation is completed, the patches (and in some cases the links) of the graph are characterized by a series of additional metrics.

Batch parameter for global metrics

For a given graph, a global weighted metric is calculated in series according to the variation of one of its parameters. As previously, this variation is defined between a minimum value (min), a maximum value (max), and with an interval (increment).

The procedure ends with the opening of a new window displaying the curve of the selected metric versus the parameter.

Table 2 summarizes possible metrics calculations.

Family	Connectivity metrics	Code	Patch capacity	Intra-patch distance	Parameters		Batch graph	Batch param
					α	β		
Weighted	Flux	F	×	×	×	×	×	×
	Equivalent Connectivity	EC	×	×	×		×	×
	Probability of connectivity	PC	×	×	×		×	×
	Interaction Flux	IF	×	×	×	×		×
	Fractions of delta Probability of connectivity	dPC	×	×	×			
	Betweenness centrality index	BC	×	×	×	×		×
	Integral index of connectivity	IIC	×				×	
	Current flow	CF	×			×		×
Area	Mean size of the components	MSC	×				×	
	Size of the largest component	SLC	×				×	
	Class coincidence probability	CCP	×				×	
	Expected cluster size	ECS	×				×	
Topological	Node Degree	Dg						
	Clustering coefficient	CC						
	Closeness centrality	CCe		x				
	Eccentricity	Ec		x				
	Connectivity correlation	CCor						
	Number of components	NC					×	
	Graph diameter	GD		×			×	
	Harary Index	H					×	
	Wilks' Lambda	W					×	

Table 2: Possible connectivity metrics calculations

6.4 Interpolating metrics

The Analysis / Metric interpolation menu is used to create raster layers from local metrics calculated at patch level. This transformation is based on a specific spatial interpolation which assigns connectivity values of patches to each cell of a grid, using a decreasing weighting function from the patch edge (weight of 1). Overall, the farther cells are away from the graph, the lower their connectivity values.

The weighting is a negative exponential function as $p = e^{-\alpha d}$ for which the user selects a distance (d) corresponding to a certain probability (p) and the software deduces the value of the α parameter. In theory, this adjustment must be consistent with the choice of reference graph or of any weighted metrics, using the same value of d.

The Multi connection option allows several patches to be included in the calculation of metrics at the point level. The calculation is based on a sum or a weighted mean of values of all patches in the vicinity of the points, up to the specified Maximum distance.

The distance used in these calculations depends on the reference graph. If it is based on least-cost distance, the spatial interpolation uses the same distance and not Euclidean distance.

The metric interpolation is used automatically in calculating species distribution models; see Species distri-

bution model.

7 Patch addition

The function "Add patches" is available in the menu "Analyses". It is used to test a set of locations to create new habitat patches according to the connectivity gain they provide. Graphab calculates first the selected global connectivity metric, then it adds virtually a patch and the links between this patch and existing patches (only if their distance is less than the maximum distance of the graph in the case of a pruned graph) and calculates the global metric again. Once all locations are tested, the software validates the one for which patch addition produces the biggest increase in the metric value. The process is repeated until the desired number of additional patches is reached by including elements already added at each step.

Several parameters must be defined:

- Graph: the graph used for computing the connectivity metric and adding new patches. The graph must not include intra-patch distance and the underlying link set must be in complete topology.

- Global metric: the global metric to maximize and his settings (Parameters button)

- Nb Patches: number of patches (which become new nodes of the graph) to be added.

The definition of the patch set to test can be done in two ways:

- by automatically applying a grid of a given resolution (field "cell size" to define). In this case, Graphab tests each centroid of the cells by adding a virtual new patch. By default, all centroids have the same capacity value (=1). Nevertheless, it is possible to import a raster layer in which each pixel of the lansdcape matrix has its own capacity value,

- by importing a shapefile of points or polygons. In this case, the user must indicate which column of the attribute table contains the patch capacity.

Multi-patches option : this option can be used to test the addition of several patches simultaneously. The additionnal patches are searched in a neighbourhood window of x cells (parameter to define) around the first patch.

At the end of the calculation, a folder named "addpatch_nX-GraphX_..." is created and saved in the project directory. It contains:

- a shapefile containing the added nodes. The attribute table indicates the step at which each node was validated and the corresponding value of the global metric,

- two shapefiles containing all the links of the graph, including those connecting the new nodes in topological form ("topo_links_graphXXX") or in realistic form ("links_graphXX"),

- a raster "landuse.tif" containing the modified land use,

- a folder named "detail" containing the shapefile of the tested locations and the corresponding value of the global metric for each step.

References : [Clauzel et al.(2015a), Foltête et al.(2014)]

8   Meta-patch

From the Graph menu, the "Create metapatch project" entry creates a new project where the nodes of the graph correspond to sets of interconnected patches. This feature allows to better take into account the nested scales of ecological processes, especially the functional definition of a habitat patch that is dependent on the scale considered.

To use this feature, it is necessary to create an initial project and a graph pruned at the desired distance (eg daily distance). Each component of this pruned graph correspond to a meta-patch. It is therefore constituted by all interconnected patches.

By default, the metapatch capacity is defined as the sum of the capacity of the patches composing the metapatch. The other option, "Weighted distance", is used to reflect the removal between patches. For this option, you must define the parameter α by giving a distance d and an associated probability p (see Parameters of weighted metrics). The resulting capacity (C) of a metapatch is:

$$C = \frac{1}{n} \sum_{i}^{n} \sum_{j}^{n} c_j e^{-\alpha d_{ij}}$$

The "Min. capacity" option keeps only the metapatch that have a capacity greater than or equal to the given capacity.

The new project is saved in a subdirectory of the original project and is automatically loaded at the end of the process.

Reference : [Clauzel et al.(2015b)]

9 Connecting graphs and point data

The main part of the software is for graph construction and computation of connectivity metrics. But it is often useful to connect these elements with external data. Graphab allows graph data to interact with a points data set.

9.1 Importing points sets

Point data can be imported via the Data / Import point set menu. These data may contain several attributes but only binary attributes (presence/absence) are taken into account in certain procedures (see Species distribution model).
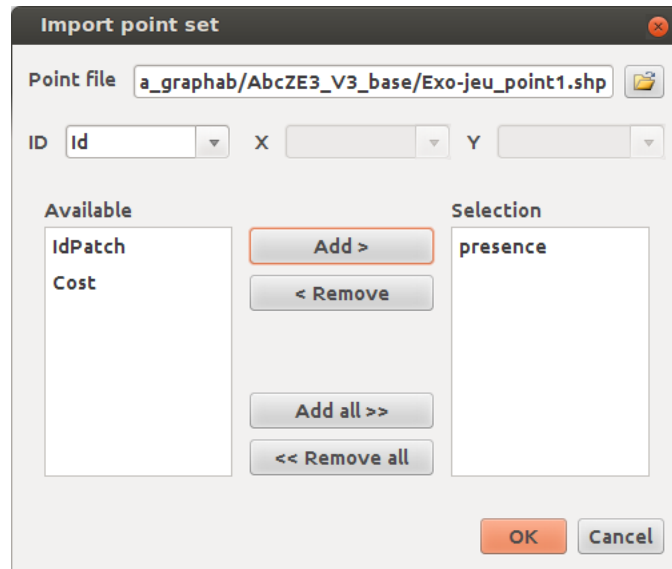
The imported file may be either in shapefile format (*.shp), or in table format (*.csv). For files in table format, the user must specify the columns corresponding to identifiers and to the XY coordinates of points in the table. The attributes to be considered must be selected from the list of attributes available.

If point data do not contain an absence attribute, they cannot be used in species distribution models. If the user wants to set up a species distribution model, a set of pseudo-absence points can be generated by the Data / Generate random point menu.

9.2 Inter-point distance matrix

Point data imported to Graphab can be used to calculate the inter-point distance matrix by right-clicking on the name of the point data and selecting "Distance matrix".
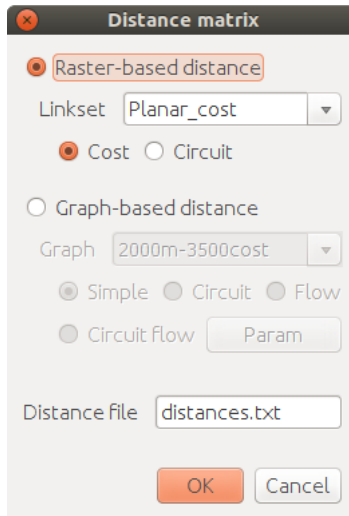
Two types of distance are available: raster-based distance and graph-based distance.

9.2.1 Raster-based distance

In raster mode, the distances are calculated based on the selected link set. This link set can not be euclidean.

According to the selected link set, the resulting distances may be the cumulative cost distance or the length of the least cost path. In both cases, the calculation includes costs assigned to the landscape map categories, as defined when creating the link set. The result is a distance matrix which is independent of the graph; this matrix corresponds to the calculation provided by the Geographicinformation Systems.

The Circuit option allows to calculate another distance between each point by calculating the resistance distance, i.e. the total resistance of an electrical circuit of resistors instead of calculating the shortest path. This method considers all possible paths and not only the shortest one (cf. [McRae et al.(2008)]).

9.2.2 Graph-based distance

In graph mode, the distances are calculated based on the selected graph. This graph must be based on the link set used for point set import.

- Least cost : calculates distances according to the shortest path in the selected graph. The type of distance is the same as that used in creating the link set of the selected graph. At both ends of a given path, the calculation includes the distance between each point and the nearest patch. Depending on the choice made when creating the graph, the calculation may or may not include intra-patch distances.

- Circuit : calculates the resistance distance between the closest patch of each point. The graph is seen as an electrical circuit where each link of the graph represents a resistor. This option considers all possible paths, not just the shortest one.

9.3 Generating random points

The Data / Generate random points menu can be used to generate a set of pseudo-absence points based on a set of presence points.
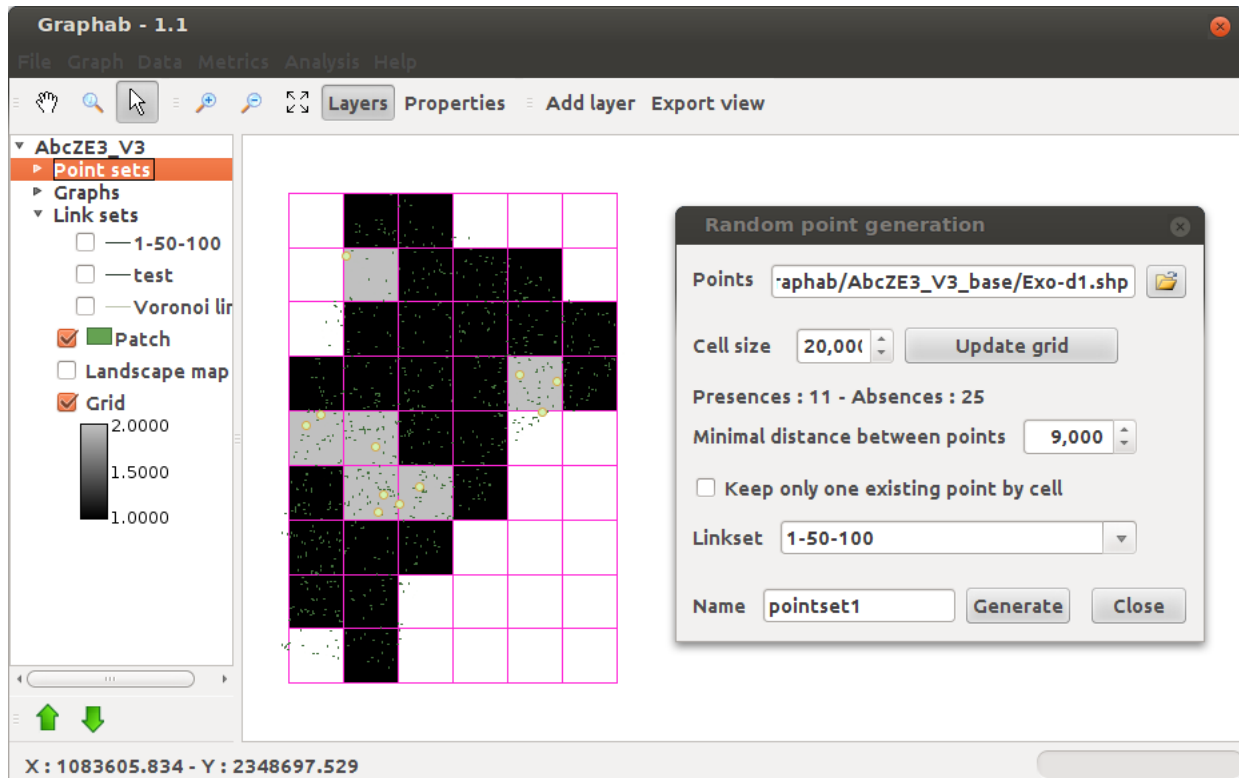
The user must load a file of presence points and specify the name of the set of presence/pseudo-absence points to be created.

Several parameters must be defined to randomly sample absence points:

- Cell size of the grid (in meters) to define the size of cells from which absence points will be potentially sampled. The "update grid" button can be used to display the grid according to the selected cell size.

- Minimum distance between points: this function reduces the effects of spatial autocorrelation by specifying a minimum distance in meters to be observed between the generated absence points and between these points and presence points.

- Type of distance: the unit of the minimum distance between points depends on the type of the distance used in the link set selected.

- Keep only one existing point by cell: this option (checked box) retains only one presence point in each cell thereby reducing the effects of spatial autocorrelation.

9.4   Species distribution model

If a set of presence/absence points has been defined, the software can use the connectivity metrics calculated from a graph as predictors in a species distribution model from the Analysis / Species distribution model menu. Such modeling is possible even if points are not located in habitat patches, by means of a spatial extrapolation of the values of metrics. The logistic regression model is a based on minimizing the AIC criterion.
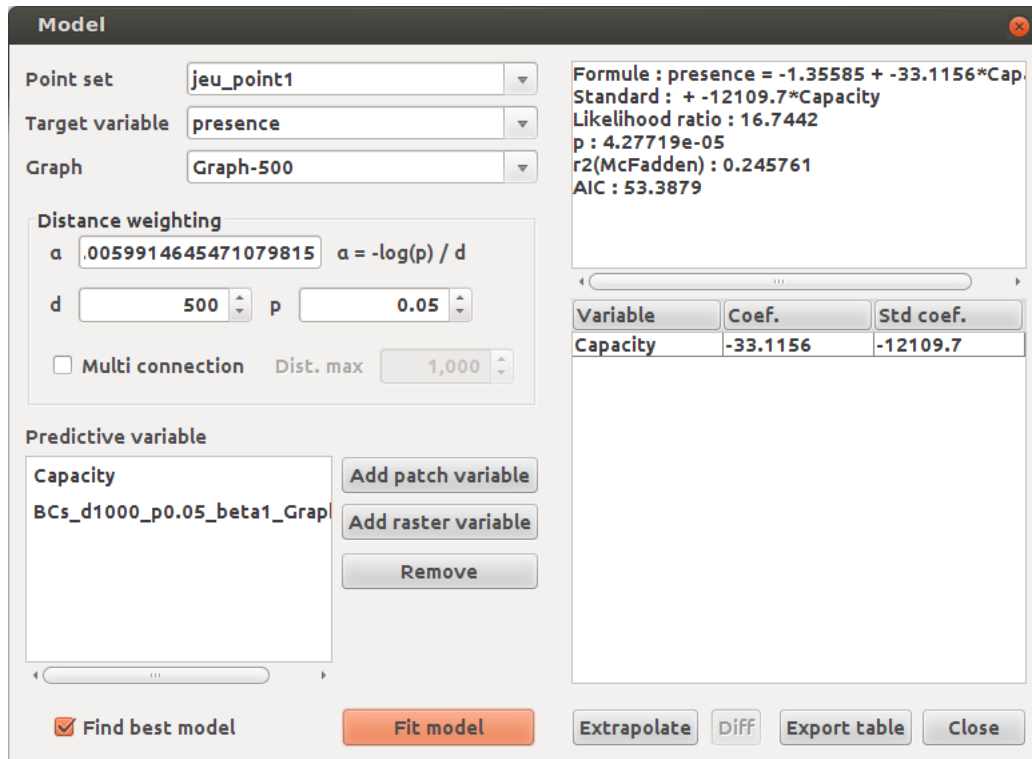
First, the user must specify:

- the set of point data to use,

- the target variable in the predictive model,

- the reference graph for the use of connectivity metrics.

9.4.1 Weighting for extrapolating metrics to points

The values of metrics are calculated for any point by a spatial interpolation. This interpolation is based on values being weighted by a decreasing function from patch edges (weight of 1). The weight decreases as the same negative exponential function as the one used for weighted metrics (see Parameters of weighted metrics), the adjustment is therefore identical.

The user selects a distance (d) corresponding to a certain probability (p) and the software deduces the value of the α parameter. In principle, this adjustment must be consistent with the choice of reference graph or of any weighted metrics included in the model, using the same value of d.

The Multi connection option allows several patches to be included in calculating metrics at points. Calculation is based on the weighted mean of values of all patches surrounding points, up to the specified Maximum distance.

Details of this weighting are given in [Foltête et al.(2012b)].

9.4.2 Estimating the model

The selection of a graph to perform the model displays all available connectivity metrics among predictive variables.

Metrics from another graph can be added by clicking on the Add patch variable button.

External variables can also be added by clicking on the Add external variable button.

Once the predictor variables have been selected, the Fit model button can be used to calculate the coefficients of the logistic regression. The results are displayed on the right-hand side of the window. The Find best model option tests all possible combinations of variables and selects the one that minimizes the AIC criterion.

9.4.3   Using the model

A predictive model which is considered to be valid can be used in several ways.

The Export table button can be used to export a table (*.csv format) with all statistical variables involved in the regression.

The Extrapolate button provides an estimation of the probability of the species presence in all cells of a grid.
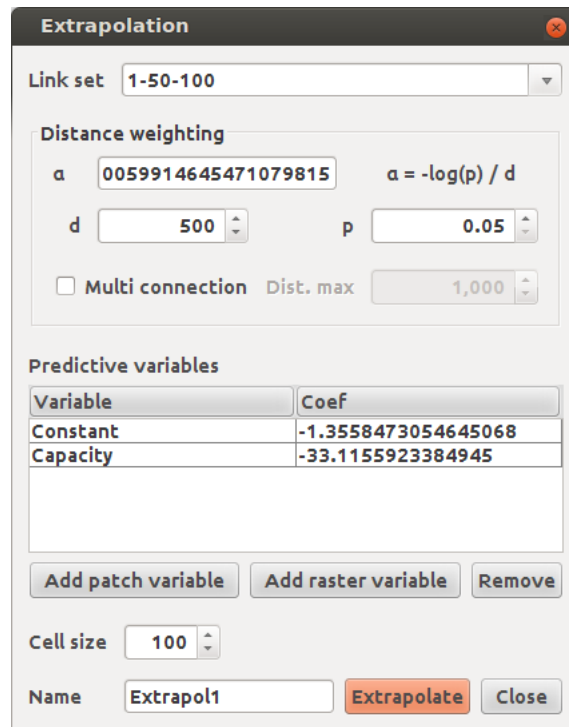
In the new window which opens, the user finds the model parameters as described previously. The cell size of the grid (in meters) indicates the level of spatial accuracy of the extrapolation. This parameter has a significant consequence on the computing time required to obtain the result. The result is saved as a raster layer in *.tif format and is displayed in the main window.

References : [Foltête et al.(2012b), Foltête et al.(2012a), Girardet et al.(2013), Clauzel et al.(2013)]

10   Display

10.1   Graph properties

Properties of a graph are available by right-clicking on the name of the graph. Two ways for viewing graphs are available:

- The topologic view displays a simplified view of the graph in which nodes are represented by dots and links by straight lines between centroids.

- The realistic view displays habitat patches according to their actual boundaries and links are represented by least-cost paths between two patches.
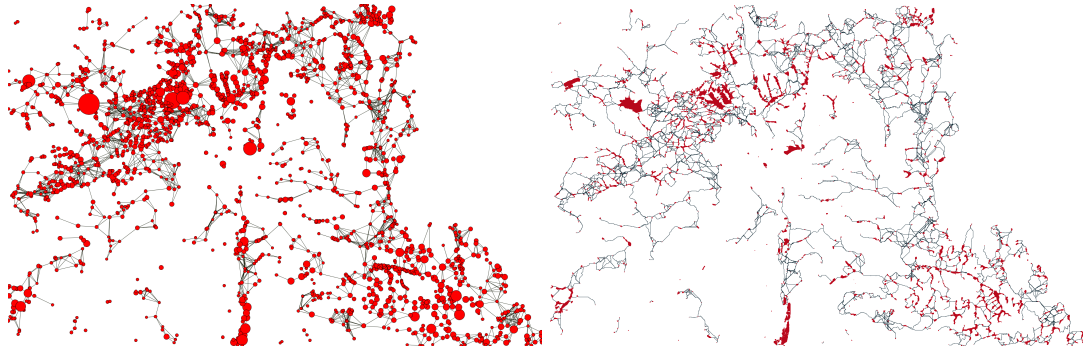
Left: Topologic view - Right: Realistic view

The Remove entry can be used to remove the selected graph. The project is saved automatically.

Properties displays the parameters used in constructing the graph: graph name, graph type with the possible maximum distance used, and the number of links.

The OD Matrix (Origin–Destination Matrix) entry creates a table with the distance between each pair of nodes for the given graph. The unit of distance depends on the type of distance used in the graph. The absence of any connection between two nodes is noted NaN (Not a Number). This matrix is saved in the project file in text format named: "graph name- odmatrix.txt ".

10.2   Object properties

The properties of link sets, graph elements (nodes, links, and components), and point data are available by right-clicking on each of them.

The Style menu includes the display parameters for objects: color, line width, label, symbol size (for nodes only). Objects can be represented in the same way (single symbol) or according to some attribute. A discretization method can be applied to classify objects according to the values of the selected attribute. By default, the legend of objects is displayed in the table of contents. It can be masked by unchecking the Legend button.

The Export menu can be used to export objects to a shapefile (*.shp), a geopackage (.gpkg) or a text file (*.txt).

The Statistic menu displays the distribution of the values of one or more attributes:

- scatter plot: values of two attributes are plotted on a two-dimensional graph,

- histogram: the bar chart of the values of an attribute is generated.

It is also possible to display the values of a given object by selecting it with the white arrow. After selection, the values of attributes are displayed in a new right-hand column named "feature properties". This column can be closed by clicking on the Properties menu in the top bar.
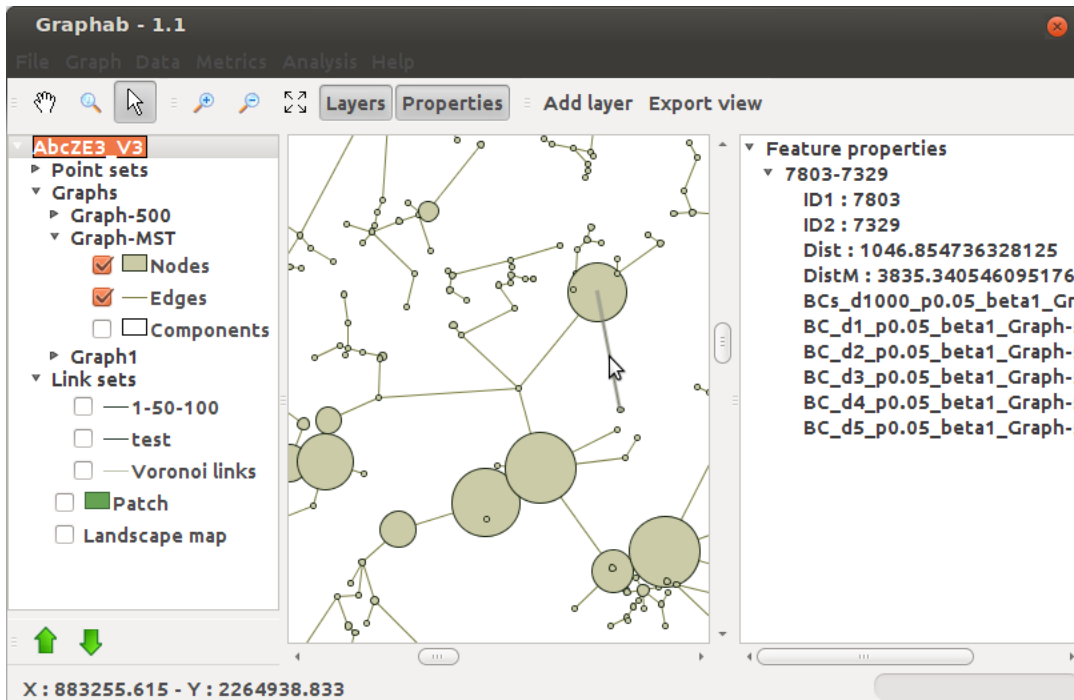
11 Processing capabilities and limitations

Graph-based methods provide an efficient modeling framework, but they can raise a question of computing capacity. Two specific points have received particular attention in Graphab: (1) calculation of link sets, (2) calculation of connectivity metrics. All these computations have been optimized by parallelization. This development mode improves computational efficiency by using a multi-processor architecture, a quad-core processor being theoretically four times faster than a single-core processor.

In [Foltête et al.(2012a)], several tests were conducted to measure the computational capacity of Graphab in different configurations. Three configurations were compared for these tests: (1) one core (3 Go RAM) corresponding to a current desktop computer, (2) four cores (6 Go RAM) corresponding to a workstation, and (3) 20 cores (15 Go RAM) corresponding to a server. The landscape map used was a grid of 14000*18000 pixels (252 millions of pixels) representing the landscape elements of the region of Franche-Comté (France) at a spatial resolution of 10 m. The landscape map contained 22,634 habitat patches.

Topology	Distance	Current desktop	Workstation	Server
Complete	Euclidean	1927s (32 min)	516s (8 min)	133s (2 min)
	Least-cost	19252s (5h 21 min)	4301s (1h 11 min)	1037s (17 min)
Planar	Euclidean	43s	12s	2.6s
	Least-cost	1080s (18 min)	295s (5 min)	82s (1 min)

Table 3: Computation times (seconds) required for calculating several link sets

The memory used by the software plays an important role. If there is not enough RAM, computation will be slower or may fail (OutOfMemoryError or GC Overhead message). The File/ Preferences / Memory menu can be used to adjust the memory allocated to Graphab. If you have a 32-bit version of Java, Graphab will be limited to about 2 Go (2000 Mo) of memory. If your computer has more than 2 Go of RAM memory, it is highly recommended you install the 64-bit version of Java to use the available memory beyond 2 Go.

12 Metrics

Family	Connectivity metrics	Code	Computing level			Delta metrics
			Global	Comp.	Local	
Weighted	Flux	F	×	×	×	
	Equivalent Connectivity	EC	×	×		×
	Probability of Connectivity	PC	×	×		×
	Interaction Flux (replace FPC)	IF			×	
	Fractions of delta Probability of Connectivity	dPC				×
	Betweenness Centrality index	BC			×	
	Integral Index of Connectivity	IIC	×	×		×
	Current Flow	CF			×	
Area	Mean Size of the Components	MSC	×			
	Size of the Largest Component	SLC	×			
	Class Coincidence Probability	CCP	×			
	Expected Cluster Size	ECS	×			
Topological	Node Degree	Dg			×	
	Clustering Coefficient	CC			×	
	Closeness Centrality	CCe			×	
	Eccentricity	Ec			×	
	Connectivity Correlation	CCor			×	
	Number of Components	NC	×			
	Graph Diameter	GD	×	×		×
	Harary Index	H	×	×		×
	Wilks' Lambda	W	×			

Table 4: Summary table of metrics in Graphab

Terms	Meaning
n	Number of patches
nc	Number of components
n_k	Number of patches in the component k
N_i	All patches close to the patch i
a_i	Capacity of the patch i (generally the surface area)
ac_k	Capacity of the component k (sum of the capacity of the patches composing k)
A	Area of the study zone
d_{ij}	Distance between the patches i and j (generally the least-cost distance between them)
$e^{-\alpha d_{ij}}$	Probability of movement between the patches i and j
α	Brake on movement distance
β	Exponent to weight more or less capacity

Table 5: Mathematical terms used

12.1 Weighted metrics

12.1.1 Flux

Flux (F)	Formula	Meaning
Global level	$$S\#F = \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} a_j^{\beta} e^{-\alpha d_{ij}}$$	For the entire graph: sum of potential dispersions from all patches.
Local level	$$F_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} a_j^{\beta} e^{-\alpha d_{ij}}$$	For the focal patch i : sum of capacity of patches other than i and weighted according to their minimum distance to the focal patch through the graph. This sum is an indicator of the potential dispersion from the patch i or, conversely to the patch i.
Values	Values depend on the definition of a. If a represents an area, F expresses an area. Minimum value: 0 Maximum value: Total area of habitat	
Comment	The path used in the graph is the one that maximizes $e^{-\alpha d}$, i.e. the one that minimizes the distance d (or the cost) between the patches i and j. This metric is called Area Weighted Flux (AWF) in some publications. However in Graphab, a is more general because it represents patch capacity, which may be their area or some other criterion chosen by the user. Similarly, the weighting is variable depending on the β parameter. In CS22, AWF is calculated only from patches directly connected to the focal patch, while Graphab takes into account indirectly connected patches.	
References	[Urban and Keitt(2001)] [Saura and Torné(2009)] [Foltête et al.(2012b)]	

12.1.2 Equivalent Connectivity

Equivalent Connectivity (EC)	Formula	Meaning
Global level Component level Delta	$EC = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j e^{-\alpha d_{ij}}}$	Square root of the sum of products of capacity of all pairs of patches weighted by their interaction probability.
Values	Unit corresponds to the capacity unit of the patches. Minimum value: 0 Maximum value: sum of capacities	
Comment	For each pair of patches, the path of the graph used is the one that maximizes $e^{-\alpha d}$, i.e. the one that minimizes the distance d (or the cost) between the patches i and j.	
Reference	[Saura et al.(2011)]	

12.1.3 Probability of Connectivity

Probability of Connectivity (PC)	Formula	Meaning
Global level Component level Delta	$PC = \dfrac{1}{A^2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j e^{-\alpha d_{ij}}$	Sum of products of capacity of all pairs of patches weighted by their interaction probability, divided by the square of the area of the study zone. This ratio is the equivalent to the probability that two points randomly placed in the study area are connected.
Values	Values correspond to a probability. Minimum value: 0 Maximum value: 1	
Comment	For each pair of patches, the path of the graph used is the one that maximizes $e^{-\alpha d}$, i.e. the one that minimizes the distance d (or the cost) between the patches i and j. This metric is not available when a_i does not represent patch area.	
Reference	[Saura and Pascual-Hortal(2007)]	

12.1.4 Interaction Flux (replace FPC)

Interaction Flux (IF)	Formula	Meaning
Local level	$$IF_i = \sum_{j=1}^{n} a_i^{\beta} a_j^{\beta} e^{-\alpha d_{ij}}$$	Sum of products of the focal patch capacity with all the other patches, weighted by their interaction probability.
Values	Minimum value: 0 Maximum value: sum of capacities	
Comment	For each pair of patches, the path of the graph used is one that maximizes $e^{-\alpha d}$, i.e. one that minimizes the distance d (or the cost) between the patches i and j. This metric corresponds to the local contribution of a patch in the PC index, since $PC = \frac{1}{A^2} \sum_i IF_i$, with $\beta = 1$. It is the equivalent of $dPC_{flux} + dPC_{area}$ not divided by the global value of PC. However, the IF metric is obtained faster because it is not calculated on the basis of patch removal (delta mode).	
References	[Foltête et al.(2014)],[Sahraoui et al.(2017)]	

12.1.5 Fractions of delta Probability of Connectivity

Fractions of delta Probability of Connectivity (dPC, dPC_{area}, dPC_{flux}, $dPC_{connector}$)		
	Formula	Meaning
Delta	$$dPC_i = \frac{(PC - PC_i')}{PC}$$ $$dPC_i = dPC_{area} + dPC_{flux} + dPC_{connector}$$ $$dPC_{area} = \frac{a_i^2}{A^2 PC}$$ $$dPC_{flux} = \frac{IF_i}{PC} - dPC_{area}$$	Rate of variation between the value of PC index and the value of PC' corresponding to the removal of the patch i. The value of dPC is decomposed into three parts: - dPC_{area} is the variation induced by the area lost after removal; - dPC_{flux} is the variation induced by the loss of interaction between the patch i and other patches; - $dPC_{connector}$ is the variation induced by the modification of paths connecting other patches and initially routed through i.
Values	Minimum value: 0 Maximum value: 1	
Comment	If a does not represent patch area, dPC_{area} does not express a loss of area but a loss of capacity.	
Reference	[Saura and Rubio(2010)]	

12.1.6 Betweenness Centrality index

Betweenness Centrality index (BC)	Formula	Meaning
Local level	$$BC_i = \sum_j \sum_k a_j^\beta a_k^\beta e^{-\alpha d_{jk}}$$ $$j,k \in \{1..n\}, k < j, i \in P_{jk}$$	Sum of the shortest paths through the focal patch i, each path is weighted by the product of the capacities of the patches connected and of their interaction probability. Pjk represents all the patches crossed by the shortest path between the patches j and k.
Values	Values depend on the configuration. They correspond to a weight of potential transit. Minimum value: 0 Maximum value: square of the total area of habitat.	
Comment	With an adjustment of $\alpha = 0$ and $\beta = 0$ (uniform weighting of paths), the BC index is the same as that used in other types of graphs. An adjustment of $\alpha = 0$ and $\beta = 1$ gives paths a weight proportional to the product of the capacities of the patches that they connect, whatever their distance. In [Foltête et al.(2012a), Foltête et al.(2012b)], the BC_l index was proposed so as to give greater weight to paths exceeding a given criterion (e.g. dispersal distance). But tests showed that this index was strongly correlated with the weighted BC adjusted with $\alpha = 0$. In [Bodin and Saura(2010)], the BC_{pc} is the weighted BC with d equal to the dispersal distance, α as $e^{-\alpha d} = 0.05$ and $\beta = 1$.	
Reference	[Bodin and Saura(2010)] [Foltête et al.(2012a)]	

12.1.7 Integral Index of Connectivity

Integral Index of Connectivity (IIC)	Formula	Meaning
Global level Component level Delta	$$IIC = \frac{1}{A^2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{a_i a_j}{1 + nl_{ij}}$$	For the entire graph: product of patch capacities divided by the number of links between them, the sum is divided by the square of the area of the study zone. IIC is built like the PC index but using the inverse of a topological distance rather than a negative exponential function of the distance based on the link impedance.
Values	Minimum value: 0 Maximum value: 1	
Reference	[Pascual-Hortal and Saura(2006)]	

12.1.8 Current Flow

Current Flow (CF)	Formula	Meaning
Local level	$$CF_i = \sum_j^n c_i^j$$	Sum of currents passing through the patch i. c_i^j represents the current through the patch i when currents are sent from all patches (except j) to the patch j. The patch j is connected to the ground.
Values	Minimum value : 0 Maximum value : $(n-1)(n-2)$ if $\beta = 0$ $(n-2)\sum_i^{n-1} a_i$ if $\beta = 1$	
Comment	The CF metric uses the electrical circuit theory. Each link of the graph corresponds to a resistor, the current sources and the ground are attached to the patches. If $\beta = 0$, each patch emits a current of 1, if $\beta = 1$, each patch emits a current equals to its capacity. This metric can be seen as an equivalent of BC metric (with $\alpha = 0$ and $\beta = 0$) that considers all possible paths, not just the shortest one.	
Reference	[Girardet et al.(2015)]	

12.2   Area metrics

12.2.1   Mean Size of the Components

Mean Size of the Components (MSC)	Formula	Meaning
Global level	$$MSC = \frac{1}{nc}\sum_{k=1}^{nc} ac_k$$	Mean of the component capacities.
Values	Minimum value: minimum capacity Maximum value: SLC	

12.2.2   Size of the Largest Component

Size of the Largest Component (SLC)	Formula	Meaning
Global level	$$SLC = max\{ac_k\}$$	Largest capacity of components.
Values	Minimum value: minimum capacity Maximum value: maximum capacity	

12.2.3 Class Coincidence Probability

Class Coincidence Probability (CCP)	Formula	Meaning
Global level	$$CCP = \sum_{k=1}^{nc} \left(\frac{ac_k}{\sum_l ac_l} \right)^2$$	Probability that two points randomly placed on the graph belong to the same component.
Values	Minimum value: $\rightarrow 0$ (as many components as patches and regular capacities) Maximum value: 1 (only one component)	
Reference	[Pascual-Hortal and Saura(2006)]	

12.2.4 Expected Cluster Size

Expected Cluster Size (ECS)	Formula	Meaning
Global level	$$ECS = \frac{1}{\sum_k ac_k} \sum_{k=1}^{nc} ac_k^2$$	Expected size of a component
Values	Minimum value: minimum capacity (as many components as patches and regular capacities) Maximum value: sum of capacities (only one component)	
Reference	[O'Brien et al.(2006)]	

12.3 Topological metrics

12.3.1 Wilks' Lambda

Wilks' Lambda (W)	Formula	Meaning
Global level	$$W_{Lambda} = \frac{|W|}{|T|}$$	Ratio between within-classes (within-components) (co)variance $|W|$ and total (co)variance $|T|$
Values	Minimum value: 0 the patches belonging to the same component are identical for all the variables (*i.e.* perfect partition) Maximum value: 1 patches belonging to the same component differ as much as with the other patches (*i.e.* meaning an irrelevant partition)	
Reference	[Everitt and Dunn(2001)], [Foltête and Vuidel(2017)]	

12.3.2   Harary Index

Harary Index (H)	Formula	Meaning
Global level Component level Delta	$$H = \frac{1}{2} \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{1}{nl_{ij}}$$	Sum of the inverse of the number of links between all pairs of patches.
Values	Minimum value: 0 for a graph with no edges Maximum value: $\frac{n(n-1)}{2}$ for a complete graph	
Comment	For pairs of patches not connected by a path, we have : $nl_{ij} = +\infty$	
Reference	[Ricotta et al.(2000)]	

12.3.3   Graph Diameter

Graph Diameter (GD)	Formula	Meaning
Global level Component level Delta	$$GD = \max_{ij} d_{ij}$$ $$GD = \max_{i} Ec_i$$	Greatest distance between two patches of the graph.
Values	Minimum value: 0 Maximum value: $+\infty$	
Comment	When the nodes i and j are not related $d_{ij} = 0$ This metric is the global version of the metric Ec_i	
Reference	[Urban and Keitt(2001)]	

12.3.4   Number of Components

Number of Components (NC)	Formula	Meaning
Global level	$$NC = nc$$	Number of components of the graph.
Values	Minimum value: 1 Maximum value : n	
Reference	[Urban and Keitt(2001)]	

12.3.5 Node Degree

Node Degree (Dg)	Formula	Meaning
Local level	$$Dg_i = |N_i|$$	Number of edges connected to the node i *ie.* number of patches connected directly to the patch i
Values	Minimum value: 0 Maximum value: n	
Comment	There is an equivalence between the node degree and the number of nearest patches because graphs are not oriented and do not contain any loops.	
Reference	Freeman, 1979	

12.3.6 Clustering Coefficient

Clustering Coefficient (CC)	Formula	Meaning
Local level	$$CC_i = \frac{1}{|N_i|\,(|N_i| - 1)} \sum_{j \in N_i} |N_i \cap N_j|$$	Ratio of the number of nodes close to i which are neighbors to each other over the possible total.
Values	Minimum value: 0 Maximum value: 1	
Comment	Si $|N_i| \leq 1 \rightarrow CC_i = 0$	
Reference	[Ricotta et al.(2000)]	

12.3.7 Closeness Centrality

Closeness Centrality (CCe)	Formula	Meaning
Local level	$$CCe_i = \frac{1}{n_k - 1} \sum_{\substack{j=1 \\ j \neq i}}^{n_k} d_{ij}$$	Mean distance from the patch i to all other patches of its component k.
Values	Minimum value: 0 Maximum value: $+\infty$	
Comment	Si $n_k = 1 \rightarrow CCe_i = 0$	
Reference	Freeman, 1979	

12.3.8   Eccentricity

Eccentricity (Ec)	Formula	Meaning
Local level	$$Ec_i = \max_j d_{ij}$$	Maximum distance from the patch i to another patch of its component k.
Values	Minimum value: 0 Maximum value: $+\infty$	
Reference	[Urban and Keitt(2001)]	

12.3.9   Connectivity Correlation

Connectivity correlation ($CCor$)	Formula	Meaning
Local level	$$CCor_i = \frac{|N_i|^2}{\sum_{j \in N_i} |N_j|}$$	Ratio between the degree of the node i and the degree of its neighboring patches j
Values	Minimum value: 0 Maximum value: $|N_i|$	
Comment	If $|N_i| = 0 \rightarrow CCor_i = 0$	
Reference	[Minor and Urban(2008)]	

References

[Bodin and Saura(2010)] Orjan Bodin and Santiago Saura, 2010. Ranking individual habitat patches as connectivity providers: Integrating network analysis and patch removal experiments. *Ecological Modelling*, 221(19) 2393 − 2405.

[Brandes et al.(2008)] U. Brandes, D. Delling, M. Gaertler, R. Gorke, M. Hoefer, Z. Nikoloski, and D. Wagner, 2008. On modularity clustering. *Knowledge and Data Engineering, IEEE Transactions on*, 20(2) 172–188.

[Clauzel et al.(2015a)] Celine Clauzel, Cyrielle Bannwarth, and Jean-Christophe Foltete, 2015a. Integrating regional-scale connectivity in habitat restoration: An application for amphibian conservation in eastern france. *Journal for Nature Conservation*, 23 98 − 107.

[Clauzel et al.(2015b)] Celine Clauzel, Deng Xiqing, Wu Gongsheng, Patrick Giraudoux, and Li Li, 2015b. Assessing the impact of road developments on connectivity across multiple scales: Application to yunnan snub-nosed monkey conservation. *Biological Conservation*, 192 207 − 217.

[Clauzel et al.(2013)] Céline Clauzel, Xavier Girardet, and Jean-Christophe Foltête, 2013. Impact assessment of a high-speed railway line on species distribution: Application to the european tree frog (hyla arborea) in franche-comté. *Journal of Environmental Management*, 127 125 − 134.

[Everitt and Dunn(2001)] Brian S Everitt and Graham Dunn, 2001. *Applied multivariate data analysis*, volume 2. Wiley Online Library.

[Foltête and Vuidel(2017)] Jean-Christophe Foltête and Gilles Vuidel, 2017. Using landscape graphs to delineate ecologically functional areas. *Landscape Ecology*, 32(2) 249–263.

[Foltête et al.(2021)] Jean-Christophe Foltête, Gilles Vuidel, Paul Savary, Céline Clauzel, Yohan Sahraoui, Xavier Girardet, and Marc Bourgeois, 2021. Graphab: an application for modeling and managing ecological habitat networks. *Software Impacts*, 8 100065.

[Foltête et al.(2012a)] Jean-Christophe Foltête, Céline Clauzel, and Gilles Vuidel, 2012a. A software tool dedicated to the modelling of landscape networks. *Environmental Modelling and Software*, 38 316 − 327.

[Foltête et al.(2012b)] Jean-Christophe Foltête, Céline Clauzel, Gilles Vuidel, and Pierline Tournant, 2012b. Integrating graph-based connectivity metrics into species distribution models. *Landscape Ecology*, 27(4) 557–569.

[Foltête et al.(2014)] Jean-Christophe Foltête, Xavier Girardet, and Céline Clauzel, 2014. A methodological framework for the use of landscape graphs in land-use planning. *Landscape and Urban Planning*, 124 140 − 150.

[Girardet et al.(2015)] Xavier Girardet, Géraldine Conruyt-Rogeon, and Jean-Christophe Foltête, 2015. Does regional landscape connectivity influence the location of roe deer roadkill hotspots? *European Journal of Wildlife Research*, 61(5) 731–742.

[Girardet et al.(2013)] Xavier Girardet, Jean-Christophe Foltête, and Céline Clauzel, 2013. Designing a graph-based approach to landscape ecological assessment of linear infrastructures. *Environmental Impact Assessment Review*, 42 10 − 17.

[McRae et al.(2008)] Brad H. McRae, Brett G. Dickson, Timothy H. Keitt, and Viral B. Shah, 2008. Using circuit theory to model connectivity in ecology, evolution, and conservation. *Ecology*, 89(10) 2712 − 2724.

[Minor and Urban(2008)] Emily Minor and Dean Urban, 2008. A graph-theory framework for evaluating landscape connectivity and conservation planning. *Conservation biology*, 22 297–307.

[Newman(2006)] M. E. J. Newman, 2006. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23) 8577–8582.

[O'Brien et al.(2006)] Dan O'Brien, Micheline Manseau, Andrew Fall, and Marie-Josée Fortin, 2006. Testing the importance of spatial configuration of winter habitat for woodland caribou: An application of graph theory. *Biological Conservation*, 130(1) 70 − 83.

[Pascual-Hortal and Saura(2006)] Lucía Pascual-Hortal and Santiago Saura, 2006. Comparison and development of new graph-based landscape connectivity indices: towards the priorization of habitat patches and corridors for conservation. *Landscape Ecology*, 21(7) 959–967.

[Ricotta et al.(2000)] C. Ricotta, A. Stanisci, G.C. Avena, and C. Blasi, 2000. Quantifying the network connectivity of landscape mosaics: a graph-theoretical approach. *Community Ecology*, 1(1) 89–94.

[Sahraoui et al.(2017)] Yohan Sahraoui, Jean-Christophe Foltête, and Céline Clauzel, 2017. A multi-species approach for assessing the impact of land-cover changes on landscape connectivity. *Landscape Ecology*, 32(9) 1819–1835.

[Saura et al.(2011)] Santiago Saura, Christine Estreguil, Coralie Mouton, and Mónica Rodríguez-Freire, 2011. Network analysis to assess landscape connectivity trends: Application to european forests (1990–2000). *Ecological Indicators*, 11(2) 407 − 416.

[Saura and Pascual-Hortal(2007)] Santiago Saura and Lucía Pascual-Hortal, 2007. A new habitat availability index to integrate connectivity in landscape conservation planning: Comparison with existing indices and application to a case study. *Landscape and Urban Planning*, 83(2–3) 91 − 103.

[Saura and Rubio(2010)] Santiago Saura and Lidón Rubio, 2010. A common currency for the different ways in which patches and links can contribute to habitat availability and connectivity in the landscape. *Ecography*, 33(3) 523–537.

[Saura and Torné(2009)] Santiago Saura and Josep Torné, 2009. Conefor sensinode 2.2: A software package for quantifying the importance of habitat patches for landscape connectivity. *Environmental Modelling and Software*, 24(1) 135 − 139.

[Urban and Keitt(2001)] Dean Urban and Timothy Keitt, 2001. Landscape connectivity: A graph-theoretic perspective. *Ecology*, 82(5) 1205–1218.