



Graphab 2.0

Manuel d'utilisation

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

13/09/2017

Table des matières

1	Introduction	2
1.1	A propos de Graphab	2
1.2	Configuration requise	2
1.3	Installation et lancement du programme	2
2	Démarrer un projet Graphab	3
2.1	Identification du projet	3
2.2	Importation de la carte du paysage et définition des nœuds	4
2.3	Création d'un jeu de liens	5
3	Graphes	7
3.1	Création d'un graphe	7
3.2	Partitionnement de graphe	8
4	Capacité des taches	10
4.1	Capacité calculée comme une fonction du voisinage	10
4.2	Capacité définie à partir de données externes	11
4.3	Suppression de taches	11
5	Calcul des métriques de connectivité	12
5.1	Famille de métriques et niveaux de calcul	12
5.2	Paramétrage des métriques pondérées	14
5.3	Calcul des métriques par lot	15
5.4	Interpolation des métriques	17
6	Ajout de taches	18
7	Méta-taches	20
8	Liens entre graphes et données ponctuelles	20
8.1	Import de données ponctuelles	21
8.2	Matrices de distance inter-points	21
8.3	Générer des points aléatoires	22
8.4	Modèle de distribution d'espèce	23
9	Affichage	26
9.1	Propriétés des graphes	26
9.2	Propriétés des objets	26
10	Capacités de traitement et limitations	27
11	Annexes	28
11.1	Détail de calcul des métriques	28

1 Introduction

1.1 A propos de Graphab

Le programme Graphab est un outil de modélisation des réseaux écologiques fondé sur les graphes paysagers. C'est un outil intégré composé de 4 modules :

- création de graphes à partir d'une carte de paysage avec identification des taches d'habitat et des liens (distance euclidienne ou chemin de moindre coût)
- calcul d'un grand nombre de métriques de connectivité
- intégration des métriques de connectivité dans un modèle de distribution d'espèce
- interface de géo-visualisation

La version 2.0 apporte plusieurs nouvelles fonctionnalités, dont :

- l'ajout de taches,
- la création de méta-taches,
- le partitionnement de graphe,
- l'intégration du calcul de pente dans la création des liens,
- la théorie des circuits électriques.

Attention, les projets ne sont pas compatibles entre les version 1.x et 2.x de Graphab !

1.1.1 Auteurs

Le programme Graphab a été développé par Gilles Vuidel et Jean-Christophe Foltête au laboratoire [ThéMA \(Université de Franche-Comté – CNRS\)](#). Le développement du logiciel a été financé par le ministère de l'écologie, du développement durable, des transports et du logement dans le cadre du programme [ITTECOP](#). Le logo de Graphab a été conçu par [Gachwell](#).

1.1.2 Conditions d'utilisation

Le programme Graphab est disponible librement sous licence GPL. Les utilisateurs de Graphab sont invités à citer la référence [[Foltête et al.\(2012a\)](#)] dans leurs travaux :

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 Configuration requise

Graphab fonctionne sur n'importe quel ordinateur supportant Java 7 ou supérieur (PC sous Linux, Windows, Mac...). Toutefois, lorsqu'il s'agit de données très volumineuses, la quantité de mémoire vive (RAM) de l'ordinateur peut limiter le nombre maximum de nœuds et de liens qui peuvent être traités en une seule fois avec Graphab. De plus, pour certaines métriques complexes, la puissance de votre processeur va déterminer la vitesse de leur calcul. Vous trouverez plus de détails à ce sujet dans la section [10](#) et dans [[Foltête et al.\(2012a\)](#)].

1.3 Installation et lancement du programme

Graphab est téléchargeable à cette adresse : <http://thema.univ-fcomte.fr/productions/graphab>.

- Télécharger et installer Java 7 ou + ([java.com](#)). Installer de préférence la version 64 bits de Java.
- Télécharger graphab-2.0.jar
- Lancer graphab-2.0.jar

Une fois graphab-2.0.jar lancé, le menu fichier donne accès à 4 rubriques :

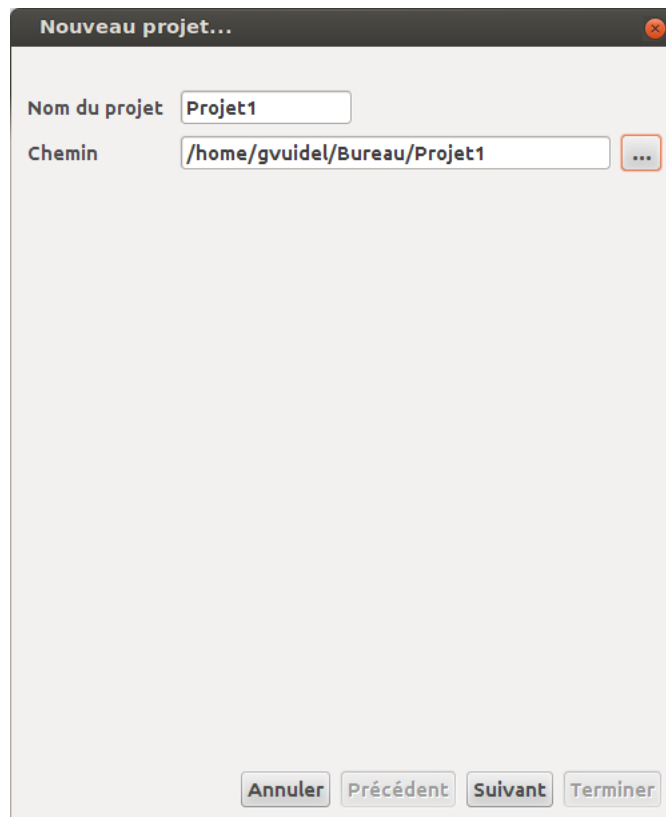
- Fichier / Nouveau projet : pour commencer la création d'un projet dans lequel toutes les données et résultats sont automatiquement enregistrés.
- Fichier / Ouvrir un projet : ouvre un projet déjà créé auparavant.
- Fichier / Préférences : pour changer certains paramètres du programme : langue anglais/français; quantité maximale de mémoire utilisable; nombre de processeurs à utiliser. Il est recommandé d'adapter la quantité de mémoire et le nombre de processeurs aux capacités de l'ordinateur utilisé et aux traitements prévus (voir à ce sujet la section [Capacités de traitement et limitations](#)).
- Fichier / Journal : affiche le journal des événements.

2 Démarrer un projet Graphab

La mise en place d'un nouveau projet Graphab s'effectue à partir du menu : Fichier / Nouveau projet. L'utilisateur doit renseigner plusieurs fenêtres successives : identification du projet, importation de la carte paysagère, création d'un jeu de liens. Un projet correspond à l'exploitation d'une seule carte paysagère initiale, mais il peut comporter plusieurs jeux de liens. Au terme de la phase de démarrage, le projet sera le support pour la création de multiples graphes et calculs de métriques de connectivité.

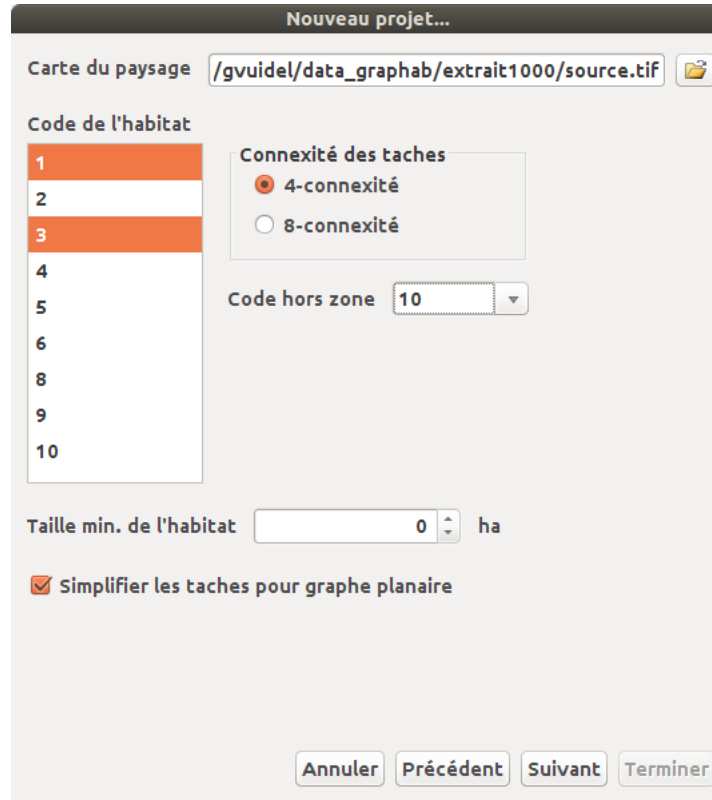
2.1 Identification du projet

Dans la première fenêtre, l'utilisateur doit nommer le projet et indiquer dans quel dossier celui-ci sera créé.



2.2 Importation de la carte du paysage et définition des nœuds

La deuxième fenêtre concerne l'importation de la carte du paysage. Celle-ci doit être un fichier de type raster (*.tif, *.rst) dans lequel la valeur de chaque pixel correspond à une catégorie (occupation du sol ou autre type de classification).



S'il est en format *.tif, sans extension Geotiff, le fichier doit être associé à un fichier de géoréférencement (*.tfw) structuré comme suit :

Exemple	
10	Taille du pixel en X
0	Rotation sur les lignes (0 pour les bonnes images)
0	Rotation sur les colonnes (0 pour les bonnes images)
-10	Taille du pixel en Y
821755	Coordonnée en X du pixel haut gauche
2342995	Coordonnée en Y du pixel haut gauche

S'il est en format *.rst, le fichier doit être associé à un fichier de géoréférencement (*.rdc) suivant la structure proposée dans le logiciel Idrisi.

Important ! L'unité du système de coordonnées de l'image doit être en mètre. Si ce n'est pas le cas, les unités de surface et de distance seront erronées. Il est possible de reprojecter l'image dans une projection métrique (Lambert93, UTM, ...) à l'aide d'un SIG.

Code hors zone : code des pixels correspondant à l'absence de valeurs dans le fichier raster.

Code de l'habitat : valeurs des pixels correspondant à l'habitat utilisé pour la définition des taches. Plusieurs valeurs peuvent être sélectionnées en maintenant la touche **Ctrl**.

Taille min. de l'habitat : surface minimale en hectare pour qu'une tache d'habitat devienne un nœud du graphe.

Connexité des taches :

- 4-connexité : une tache est constituée du pixel central et de ses 4 voisins s'ils ont la même valeur ;
- 8-connexité : une tache est constituée du pixel central et de ses 8 voisins s'ils ont la même valeur.

Simplifier les taches pour graphe planaire : cette option, lorsqu'elle est cochée, permet d'accélérer la création d'un graphe planaire, en simplifiant les limites des polygones des taches. Ce processus de simplification n'étant pas déterministe, la création de deux graphes planaires pour une même carte du paysage peut engendrer des limites de polygones légèrement différentes. Par conséquent, cette option ne doit pas être cochée dans le cas d'une comparaison de graphes planaires, par exemple quand on analyse les conséquences de changements d'occupation du sol sur la connectivité.

2.3 Création d'un jeu de liens

La troisième fenêtre concerne la création d'un jeu de liens, ce qui implique la sélection de plusieurs paramètres : topologie et pondération des liens. La création de ce jeu de liens constitue la dernière étape dans la mise en place initiale du projet. Cependant, il est possible de créer à tout moment de nouveaux jeux de liens dans le même projet, par le menu : Graphes / Créer un jeu de liens.

Code	Cost
1	1
2	1
3	1
4	1
5	1
6	1
8	1
9	1

2.3.1 Topologie des liens

Deux topologies sont disponibles :

- planaire : seuls les liens formant un graphe planaire minimal sont pris en compte. Cette topologie est mise en place par l'intermédiaire de polygones de Voronoï autour de chaque tache d'habitat. Ces polygones sont définis à partir des bords des taches, en distance euclidienne.
- complet : tous les liens entre les taches sont potentiellement pris en compte.

Distance max : Cette option disponible pour la topologie complète permet de préciser un seuil de distance au-delà duquel les liens ne sont plus créés. Elle permet ainsi de limiter le nombre de liens créés et d'accélérer la création des liens en topologie complète.

Supprimer les liens traversant les taches : cette option, décochée par défaut, permet de ne pas créer de lien entre deux taches (A et C sur la figure suivante) passant par une tache intermédiaire (B). Ceci est recommandé pour le calcul de la métrique de centralité intermédiaire (BC) pour prendre en compte la fréquence avec laquelle chaque tache se trouve sur les plus courts chemins liant les paires de nœuds sur le graphe. A noter : cette option fournit une approximation de la distance intra-taches entre A et C. Si la case est décochée, un lien entre A et C passant par B sera créé représentant la distance complète réelle entre les deux taches.

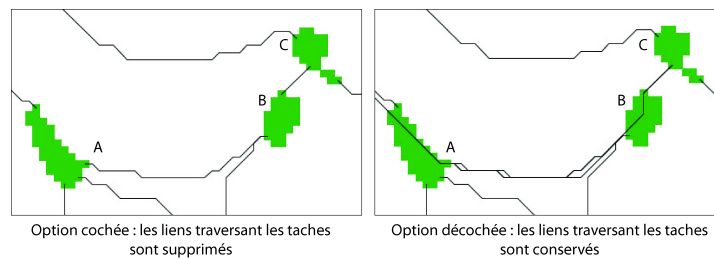


FIGURE 1 – Illustration de l'option « Supprimer les liens traversant les taches »

Enregistrer les chemins :

- si la case est cochée : les liens sont enregistrés sous forme de chemins représentant le tracé réel du lien entre deux taches.
- si la case est décochée : les liens sont enregistrés sous forme topologique uniquement. Dans ce cas, la visualisation des liens en vue réaliste n'est plus possible. Ceci est recommandé lorsque le graphe contient de nombreux liens (dans le cas d'un graphe complet non seuillé par exemple) afin de limiter l'utilisation de la mémoire. A noter : si les chemins ne sont pas enregistrés, il n'est pas possible d'inclure les distances intra-taches dans le calcul des métriques.

2.3.2 Distance (ou impédance des liens)

Les distances sont calculées de bord à bord entre les taches. Deux principaux types de distance sont possibles : distance euclidienne et distance de moindre coût.

- Distance euclidienne : les liens sont calculés en distance euclidienne (parcours à vol d'oiseau entre les taches), ce qui revient à considérer la matrice comme uniforme.
- Distance de moindre coût : les liens sont calculés en distance coût. L'hétérogénéité de la matrice est prise en compte en assignant des valeurs de résistance (ou friction) aux classes paysagères. L'utilisateur peut activer cette possibilité de deux façons :
 1. soit en indiquant les coûts correspondant aux catégories de la carte du paysage dans le tableau ;
 2. soit à partir d'un fichier raster externe (format *.tif ou *.rst) contenant pour chaque pixel une valeur de résistance.

L'usage des distances de moindre coût donne accès à deux types d'impédance calculés à partir des mêmes chemins :

- coût cumulé : l'impédance est égale à la somme des coûts de tous les pixels du chemin parcouru.
- longueur du chemin de moindre coût : l'impédance est égale à la longueur métrique du chemin parcouru.

Pour chaque lien créé, sa distance métrique et sa distance en unité de coût sont enregistrées et consultables dans les propriétés (cf. [Propriétés des objets](#)).

Les valeurs de résistance définies précédemment peuvent être pondérées par la pente [[Clauzel et al.\(2015b\)](#)]. Pour que la pente puisse être calculée, il faut tout d'abord charger un MNT à partir du menu Données | Importer un MNT.

Le paramètre coef (c) permet d'ajuster l'importance de la pondération par la pente (p). Pour un pixel donné, la résistance résultante (r_{final}) est calculée comme suit :

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

avec $p = \frac{h}{l}$ la pente exprimée par le rapport hauteur(h) sur longueur (l).
 $p = 0$ pour une pente nulle et $p = 1$ pour une pente de 100% ($h = l$).

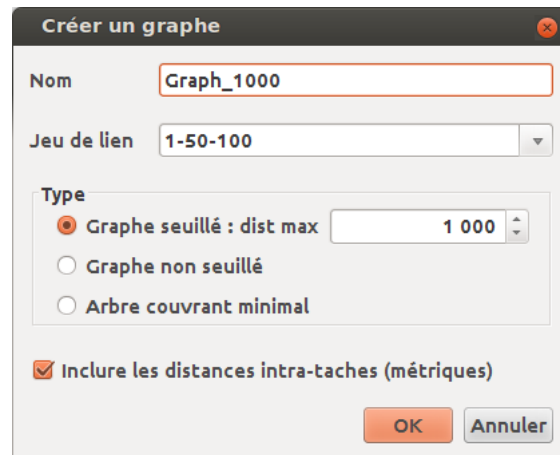
Pour $c = 1$ la valeur de résistance est doublée avec une pente de 100%.
Pour $c = 10$ la valeur de résistance est doublée avec une pente de 10%.

3 Graphes

3.1 Création d'un graphe

Un projet Graphab peut donner lieu à la création de plusieurs graphes. Chaque graphe étant construit à partir d'un jeu de lien, soit on utilise le jeu de lien défini lors de la mise en place initiale du projet, soit on définit un nouveau jeu de lien par le menu : Graphes / Créer un jeu de liens.

La création d'un graphe s'effectue par le menu : Graphe / Créer un graphe.



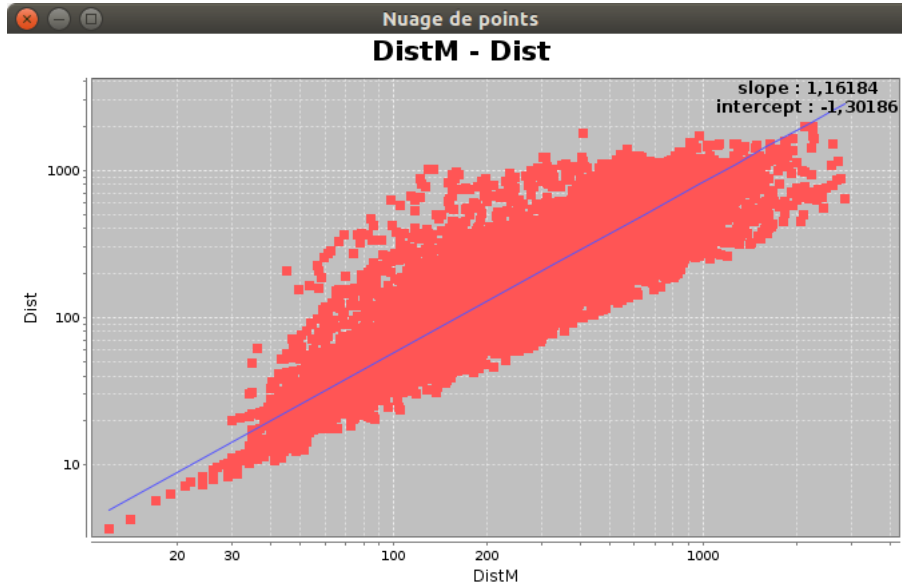
Le graphe à créer doit tout d'abord être nommé.

L'utilisateur doit ensuite sélectionner un des jeux de liens créés précédemment (cf. [Création d'un jeu de liens](#)), puis choisir un type de graphe :

- graphe seuillé : les liens retenus sont inférieurs ou égaux au seuil de distance choisi.
- graphe non seuillé : tous les liens sont créés entre les taches quelle que soit leur distance.
- arbre couvrant minimal : graphe reliant toutes les taches dont le poids total des liens est minimal.

Option « Inclure les distances intra-taches » : si la case est cochée, le calcul des métriques à l'étape 3 tient compte des distances intra-taches (recommandé); dans le cas contraire, seules les distances inter-taches sont prises en compte.

Pour un graphe seuillé, l'unité de la distance de seuillage dépend du type de distance utilisé dans la création du jeu de liens. Si le jeu de lien a été créé en distance euclidienne, la distance de seuillage du graphe est exprimée en mètres. Si la création du jeu de lien a été réalisée en distance-coût, la distance de seuillage du graphe est exprimée en coûts cumulés.



Il est possible d'obtenir une approximation de la distance métrique (appelée DistM) exprimée en coûts cumulés (appelée Dist) en affichant le nuage de points en double logarithmique correspondant au jeu de lien utilisé, et en utilisant l'équation ci-dessous pour effectuer cette conversion :

$$Dist = e^{intercept + slope \cdot \log(DistM)}$$

L'estimation peut être réalisée directement dans le logiciel à partir du menu Conversion dist. présent dans le menu contextuel d'un jeu de lien (clic droit sur le jeu de lien).

Pour réaliser une analyse multiscalaire, il est souvent nécessaire de créer une série de graphes seuillés dont le seuil est défini de façon croissante. L'utilisateur peut créer manuellement cette série de graphes, mais si l'objectif est d'analyser le comportement d'une métrique de connectivité en fonction du seuil, il est préférable d'utiliser le menu : Métriques / Graphes par lot (cf. 5.3.1).

3.2 Partitionnement de graphe

Graphab implémente un algorithme de partitionnement de graphe qui maximise l'indice de modularité [Newman(2006)]. La modularité est une mesure de la qualité d'un partitionnement des noeuds d'un graphe. Le principe sous-jacent est qu'un bon partitionnement d'un graphe implique un nombre de liens à l'intérieur des groupes important et un nombre de liens inter-groupes faible. Le calcul utilisé est basé sur l'algorithme glouton, suivi d'une optimisation locale [Brandes et al.(2008)].

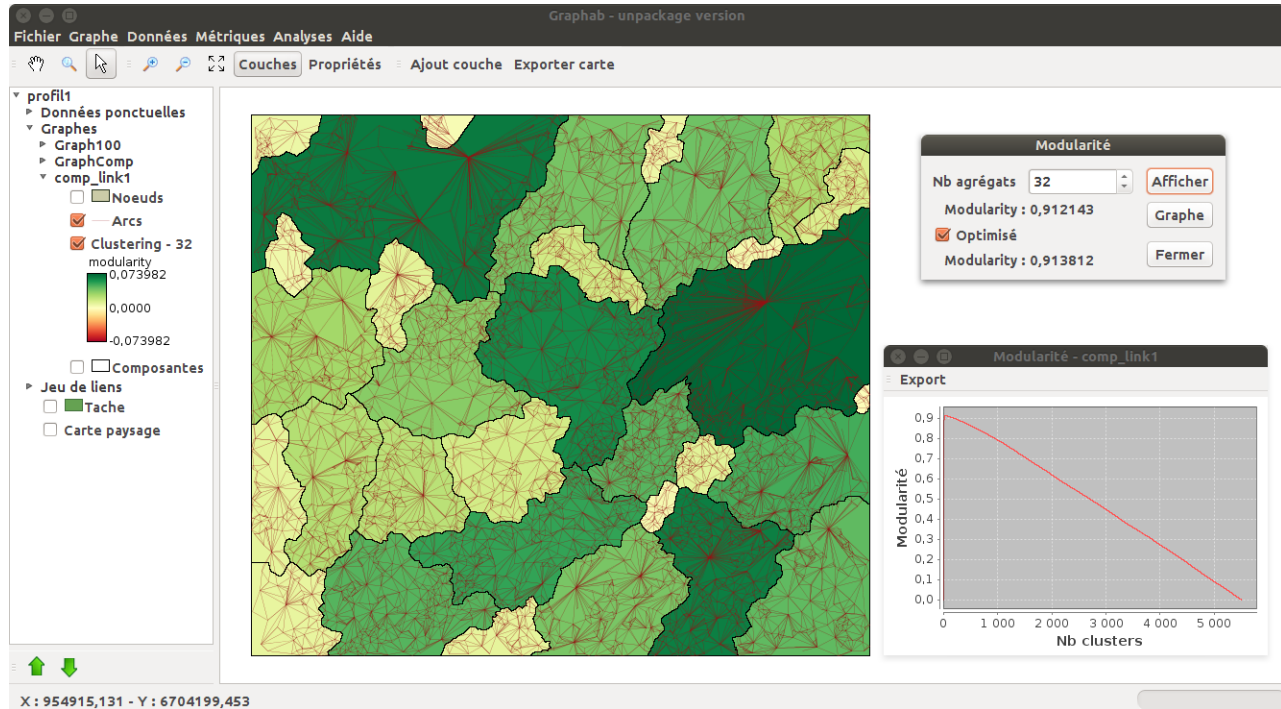
La modularité est calculée à partir d'un poids (w_{ij}) défini pour chaque lien du graphe :

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

Les 2 paramètres β et α permettent de définir respectivement l'importance de la capacité des tâches ($a_i a_j$) et l'importance de la distance (d_{ij}) pour le poids du lien w_{ij} . Si $\alpha = \beta = 0$ alors les poids sont tous identiques : $w_{ij} = 1$.

Le partitionnement est accessible depuis le menu contextuel d'un graphe (clic droit sur le nom d'un graphe). La première fenêtre demande de renseigner les paramètres α et β . Le paramètre α n'est pas renseigné directement, il est déterminé à partir de deux autres paramètres : une distance d et une probabilité p (cf. [Paramétrage des métriques pondérées](#)). Pour définir α à 0, il faut mettre p à 1.

Après validation de cette fenêtre, le calcul se lance jusqu'à obtenir un partitionnement maximisant la modularité. Le résultat s'affiche sur une nouvelle couche du graphe. Deux autres fenêtres viennent compléter ce résultat.



La nouvelle couche montre les groupes de la partition sous forme de polygones. Comme un groupe correspond à un ensemble de taches, il est représenté par un polygone agrégeant les polygones de Voronoï des taches affectées à celui-ci. La couleur d'un groupe correspond à sa valeur de modularité ; la modularité totale étant la somme des modularités de tous les groupes.

Une première fenêtre contient un graphique montrant l'évolution de la modularité en fonction du nombre de groupes (clusters). Dans la figure précédente, on voit que la modularité maximale est atteinte pour un nombre groupes relativement faible.

La deuxième fenêtre affiche le nombre de groupes (Nb agrégats) maximisant la modularité ainsi que la valeur de modularité. Plusieurs fonctionnalités sont disponibles sur cette fenêtre :

- le nombre de groupes peut être modifié pour visualiser des partitionnements sous optimaux. Le résultat peut être visualisé en cliquant sur le bouton "Afficher". Une nouvelle couche de partitionnement est ajouté au graphe.
- le bouton "Graphe" permet de créer un nouveau graphe ne contenant que les liens intra-groupes.
- la case à cocher "Optimisation" réalise une optimisation locale du partitionnement après l'algorithme glouton permettant d'améliorer sensiblement la valeur de modularité. Cette option est active par défaut.

4 Capacité des tâches

La capacité d'une tâche traduit sa « qualité intrinsèque », considérée comme un indicateur de son potentiel démographique. Ainsi une tâche avec une forte capacité pourra accueillir une forte population et inversement. La capacité intervient directement dans le calcul de certaines métriques de connectivité « de surface » et « pondérées » (voir la section 5).

Par défaut (lors de la création initiale du projet), la capacité d'une tâche est égale à sa superficie en m². Cependant, il est possible de remplacer la superficie par n'importe quel autre indicateur de qualité. En effet, dans certains cas, la présence d'une espèce n'est pas liée à la superficie de la tâche mais à la superficie d'autres types d'occupation du sol autour de la tâche. Par exemple la présence d'amphibiens dans un étang de reproduction ne dépend souvent pas de la taille du plan d'eau mais de la superficie de l'habitat terrestre autour de ce plan d'eau.

Avec le menu Données / Définir la capacité des tâches, il est possible de modifier la capacité de l'ensemble des tâches du projet par deux méthodes.

4.1 Capacité calculée comme une fonction du voisinage

La première méthode est de définir la capacité des tâches comme une fonction de la composition de leur voisinage et de procéder à ce calcul directement dans Graphab.

The screenshot shows the 'Capacité des tâches' dialog box. It features two radio buttons: 'Surface de la tâche (défaut)' (unselected) and 'Superficie du voisinage' (selected). Under 'Superficie du voisinage', there is a 'Coûts' dropdown menu set to 'JeulienCout', a 'Coût max.' input field with '100', and a checkbox 'Pondération par le coût' which is unchecked. To the right is a 'Codes inclus' list with items 1, 2, 3, and 4; items 2 and 3 are highlighted in orange. Below this is another radio button 'Import d'un fichier CSV' (unselected). Under it is a 'Fichier capacité' input field with a file selection icon. At the bottom are 'Id tâche' and 'Capacité' dropdown menus, and 'OK' and 'Annuler' buttons.

L'utilisateur doit sélectionner plusieurs critères : type de distance, distance maximale, catégories paysagères.

Coûts : il s'agit de la métrique spatiale (euclidienne ou ensemble de coûts) correspondant à un des jeux de liens disponibles dans le projet. Il est à noter que l'usage des coûts dans cette procédure revient à définir un voisinage anisotropique autour des tâches, ce qui peut différer fortement d'une fonction de type buffer. Pour des raisons de cohérence, il est conseillé d'utiliser le même type de distance que celui qui a permis de définir les liens du graphe. Dans le cas d'un jeu de lien créé avec une distance euclidienne, il faut choisir « all cost = 1 ».

Coût max. : Comme pour la distance de seuillage du graphe, l'unité de cette distance maximale de voisinage dépend du type de distance utilisé dans la création du jeu de liens (euclidien ou distance-coût).

Codes inclus : l'utilisateur peut choisir une ou plusieurs catégories paysagères, autre que la catégorie « habitat », à prendre en compte dans le calcul de la capacité.

L'option « pondération par la distance » permet d'introduire une pondération en fonction de l'éloignement à la tâche, par le biais d'une fonction exponentielle négative. De cette façon, les surfaces sélectionnées compteront plus si elles sont proches de la tâche et inversement.

Les valeurs de capacité ainsi calculées remplacent les surfaces des tâches pour tous les calculs qui suivront, mais il est possible de revenir à l'état initial par le menu : Données / Définir la capacité des tâches et en cochant l'option Surface de la tâche.

4.2 Capacité définie à partir de données externes

L'autre possibilité est d'ouvrir un tableau de données au format *.csv, décrivant toutes les tâches du projet et contenant comme attribut les valeurs de capacité définies au préalable par l'utilisateur. Les identifiants du tableau importé doivent être ceux qui ont été définis lors de la création initiale du projet.

Capacité des tâches

Surface de la tâche (défaut)

Superficie du voisinage

Coûts JeulienCout Codes inclus

Coût max. 0

Pondération par le coût

1

2

3

4

Import d'un fichier CSV

Fichier capacité .1984/profil1/patches.csv

Id tâche Id

Capacité Capacity

OK Annuler

Les valeurs de capacité du tableau importé remplacent les surfaces des tâches pour tous les calculs qui suivront cette importation, mais il est possible de revenir à l'état initial en cochant l'option Surface de la tâche.

4.3 Suppression de tâches

Après avoir défini une capacité différente de la surface de la tâche, il peut être intéressant de supprimer les tâches dont la capacité est trop faible. L'entrée "Supprimer des tâches" du menu Données, permet de créer un nouveau projet identique au projet actuel en supprimant les tâches ayant une capacité inférieure à un seuil donné.

Après avoir renseigné la capacité minimale voulue, un nouveau projet est créé dans un sous répertoire du projet actuel et chargé à la fin du traitement.

5 Calcul des métriques de connectivité

5.1 Famille de métriques et niveaux de calcul

Chaque graphe créé dans un projet peut être utilisé pour calculer différentes métriques de connectivité, dont le détail de calcul et les références bibliographiques sont mentionnés en annexe. Ces calculs s’effectuent à plusieurs niveaux correspondant aux grandes entrées du menu Métriques (tableau 1) :

- Métriques / Métrique globale : les métriques caractérisent le graphe entier.
- Métriques / Métrique par composante : les métriques caractérisent la connectivité interne de chaque composante (ou sous-graphe).
- Métriques / Métrique locale : les métriques caractérisent la connectivité de chaque élément du graphe, nœud ou lien.
- Métriques / Delta-métriques : les métriques caractérisent également chaque élément du graphe, mais à partir d’un mode de calcul spécifique. En utilisant la méthode dite de la « suppression » (suppression de nœud ou suppression de lien), l’importance relative de chaque élément du graphe est évaluée en calculant le taux de variation que sa suppression occasionne sur une métrique globale. Le résultat d’une delta-métrique est donc local, mais en référence au niveau global.

Dans la fenêtre de dialogue qui s’ouvre suite au choix d’un de ces quatre modes de calcul, trois familles de métriques sont disponibles :

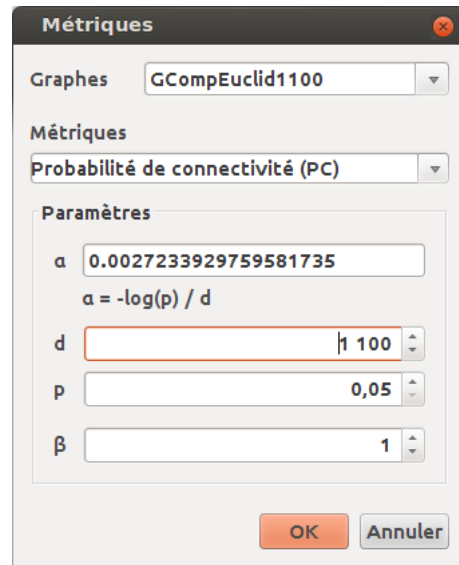
- les métriques pondérées, qui sont fondées sur des critères de distance et de capacité des taches et qui nécessitent un paramétrage adapté à l’espèce de référence. Ces métriques nécessitent des calculs de chemin sur les graphes, par le biais de l’algorithme de Dijkstra. En choisissant ces métriques, l’utilisateur doit indiquer le réglage souhaité.
- les métriques de surface, qui sont fondées en priorité sur un critère de surface. Dans le cas où la capacité ne correspond pas à la surface des taches mais à un autre critère, ces métriques peuvent être calculées et elles s’expriment dans l’unité du critère considéré.
- les métriques topologiques, qui sont issues de la théorie des graphes et qui ne nécessitent pas de paramétrage.

Famille	Métrique de connectivité	Code	Niveau de calcul			Delta métriques
			Global	Comp.	Local	
Métriques pondérées	Flux	F	×	×	×	
	Probabilité de connectivité	PC	×	×		×
	Probabilité de connectivité flux	FPC			×	
	Fractions de delta Probabilité de connectivité	dPC				×
	Indice de centralité intermédiaire	BC			×	
	Indice intégral de connectivité	IIC	×	×		×
	Flux circuit	CF			×	
Métriques de surface	Taille moyenne des composantes	MSC	×			
	Taille de la plus grande composante	SLC	×			
	Probabilité de coïncidence de classe	CCP	×			
	Taille de composante attendue	ECS	×			
Métriques topologiques	Degré	Dg			×	
	Coefficient de groupement	CC			×	
	Centralité de proximité	CCe			×	
	Excentricité	Ec			×	
	Corrélation de connectivité	CCor			×	
	Nombre de composantes	NC	×			
	Diamètre	GD	×	×		×
	Indice d'Harary	H	×	×		×

TABLE 1 – Liste des métriques de connectivité et de leur niveau de calcul. Les codes sont indiqués suivant le nom anglais des métriques.

Quel que soit le niveau choisi, l'utilisateur doit d'abord indiquer le graphe sur lequel le calcul va s'appliquer, puis choisir la métrique de connectivité.

5.2 Paramétrage des métriques pondérées



5.2.1 Paramètre alpha

Plusieurs métriques intègrent dans leur calcul une pondération où les distances entre les taches sont converties en probabilité de déplacement. Ces métriques sont : F, FPC, PC, BC. La pondération est basée sur une fonction exponentielle de la forme :

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

où p est la probabilité de déplacement entre deux taches, d la distance entre ces taches et α un paramètre contrôlant la vitesse avec laquelle la probabilité diminue quand la distance augmente. Comme la valeur du paramètre α n'est pas facile à déterminer, le logiciel propose de le calculer à partir des deux autres paramètres. L'utilisateur doit indiquer la distance correspondant à une certaine valeur de probabilité, par exemple :

- la distance de dispersion maximale de l'espèce correspondant à une faible valeur de p (0.05 ou 0.01).
- la distance de dispersion moyenne de l'espèce correspondant à une valeur médiane de p (0.5).

La valeur α est automatiquement obtenue à partir de la formule :

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

Dans le cas d'un graphe seuillé, il est supposé que la distance d utilisée dans ce paramétrage soit en cohérence avec celle qui a été utilisée pour le seuillage du graphe.

5.2.2 Paramètre bêta

Les mêmes métriques F, PC, FPC, BC sont contrôlées par le paramètre β . Ce paramètre est l'exposant appliqué à la capacité des taches. Il joue sur l'équilibre relatif entre le poids des distances et le poids des capacités des taches dans la pondération des métriques. En prenant l'exemple de la métrique F en calcul local, dont la forme générique est la suivante :

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

- une valeur de $\beta = 0$ signifie que la capacité des taches ne joue pas de rôle dans la pondération.
- une valeur de $\beta = 1$ signifie que la capacité des taches joue de façon linéaire dans cette pondération.

- une valeur de $\beta = 2$ signifie que la capacité des tâches est portée à la puissance 2.
- une valeur de $\beta = 0.5$ signifie que la racine carrée de la capacité des tâches intervient dans la pondération.
- une valeur de $\beta = -1$ signifie que la capacité des tâches joue de façon inversement proportionnelle.

Au-delà de ces quelques exemples, toutes les valeurs de pondération sont possibles.

5.3 Calcul des métriques par lot

Chaque métrique compatible avec le niveau global peut être calculée en chaîne suivant la variation de l'échelle des distances. Cette variation peut concerner soit le seuillage du graphe (5.3.1), soit le paramétrage des métriques elles-mêmes (5.3.2). Le type de distance considéré pour le seuillage est celui qui a été défini pour le jeu de lien choisi, à savoir distance euclidienne, distance de moindre coût, chemin de moindre coût.

5.3.1 Graphes par lot

Le menu Métriques / Graphes par lot permet de créer une série de graphes seuillés à partir d'un jeu de liens donné, et de calculer pour chacun une métrique de niveau global. Les seuils des graphes successifs sont définis de façon croissante, suivant deux modes :

- distance : un intervalle de distance fixe est déterminé entre les graphes successifs. Les valeurs de la métrique sont donc calculées à des intervalles réguliers de distance.
- nombre de liens : un nombre de liens fixe est défini entre les graphes successifs. Ce nombre de liens est automatiquement transformé en distance utilisée pour seuiller le graphe. Ces distances seuil peuvent s'échelonner de façon irrégulière.

Les graphes sont définis suivant trois critères donnés par l'utilisateur :

- minimum : plus petit seuil, utilisé pour le premier graphe de la série. Par défaut, ce minimum est à 0, ce qui correspond à l'absence totale de liens.
- maximum : seuil maximal, utilisé pour le dernier graphe de la série. Par défaut, ce maximum correspond à la distance maximale, ou le nombre de lien, du jeu de lien sélectionné.
- incrément : valeur de distance ajoutée à chaque nouveau graphe.

Une fois le calcul terminé, le logiciel ouvre une nouvelle fenêtre dans laquelle la courbe de la métrique choisie est affichée en fonction du seuil de distance. Les valeurs de cette courbe peuvent être enregistrées par le bouton export, en sélectionnant le format texte.

5.3.2 Paramètres par lot

Le menu Métriques / Paramètres par lot permet de calculer en série des métriques à partir d'un graphe donné. Cette possibilité s'applique exclusivement aux métriques pondérées. Elle se subdivise en deux entrées : métriques locales ou globales.

Paramètres par lot pour métriques locales

Une métrique pondérée de niveau local est calculée en série suivant la variation d'un de ses paramètres. L'utilisateur doit choisir le graphe, la métrique et le paramètre à faire varier. La variation du calcul est déterminée par :

- min : valeur minimale du paramètre
- max : valeur maximale du paramètre
- incrément : valeur d'intervalle entre deux calculs de la métrique

Une fois le calcul terminé, les taches (et dans certains cas les liens) du graphe sont caractérisées par une série de métriques supplémentaires.

Paramètres par lot pour métriques globales

Pour un graphe donné, une métrique pondérée de niveau global est calculée en série suivant la variation d'un de ses paramètres. Comme précédemment, cette variation est définie entre une valeur minimale (min), une valeur maximale (max) et un intervalle (incrément).

La procédure se termine par l'ouverture d'une nouvelle fenêtre dans laquelle la courbe de la métrique choisie est affichée en fonction du paramètre.

Le tableau 2 récapitule les possibilités de calcul des métriques.

Type	Métrique de connectivité	Code	Capacité des taches	Distance intra-taches	Paramètres		Graphe par lot	Paramètre par lot
					α	β		
Métriques pondérées	Flux	F	×	×	×	×	×	×
	Probabilité de connectivité	PC	×	×	×	×	×	×
	Probabilité de connectivité flux	FPC	×	×	×	×	×	
	Fractions de delta Probabilité de connectivité	dPC	×	×	×	×		
	Centralité intermédiaire	BC	×	×	×	×		×
	Indice Intégral de Connectivité	IIC	×				×	
	Flux circuit	CF	×				×	×
Métriques de surface	Taille moyenne des composantes	MSC	×				×	
	Taille de la plus grande composante	SLC	×				×	
	Probabilité de coïncidence de classe	CCP	×				×	
	Taille de composante attendue	ECS	×				×	
Métriques topologiques	Degré	Dg						
	Coefficient de groupement	CC						
	Centralité de proximité	CCe		x				
	Excentricité	Ec		x				
	Corrélation de connectivité	CCor						
	Nombre de composantes	NC					×	
	Diamètre	GD		×			×	
	Indice d'Harary	H					×	

TABLE 2 – Possibilités de calcul des métriques de connectivité.

5.4 Interpolation des métriques

Le menu "Analyses / Interpoler une métrique" permet de générer des couches raster à partir des métriques locales calculées au niveau des taches. Cette transformation est fondée sur une interpolation spatiale spécifique, qui permet d'attribuer les valeurs de connectivité des taches à chaque cellule d'une grille, en utilisant une fonction de pondération décroissante à partir de la bordure des taches (poids de 1). Globalement, plus les cellules sont situées à l'écart du graphe, plus elles obtiennent des valeurs faibles de connectivité.

The screenshot shows a dialog box titled "Interpoler une métrique". It contains the following fields and controls:

- Graphe:** A dropdown menu with "comp_link1" selected.
- Métrique:** A dropdown menu with "FPC_d100_p0.05_beta1_comp_link1" selected.
- Pondération:**
 - alpha:** A text input field containing "0.029957322735539908". To its right is the formula $\alpha = -\log(p) / d$.
 - d:** A spinner box with the value "100".
 - p:** A spinner box with the value "0,05".
 - Connexion multiple:** A checked checkbox.
 - Dist. max:** A spinner box with the value "1 000".
 - Aggregation:** Two radio buttons, "Somme" (selected) and "Moyenne".
- Résolution:** A spinner box with the value "10".
- Nom:** A text input field containing "Interp_FPC_d100_p0.05_beta1_comp_link1".
- Buttons:** "Interpoler" (highlighted in orange) and "Fermer".

La pondération est une fonction exponentielle négative de la forme $p = e^{-\alpha d}$ pour laquelle l'utilisateur doit choisir une distance (d) correspondant à une certaine probabilité (p) et le logiciel en déduit la valeur du paramètre α . En principe, ce réglage est à rendre cohérent avec le choix du graphe de référence ou le choix des éventuelles métriques pondérées qui sont à mobiliser, en utilisant la même valeur de d.

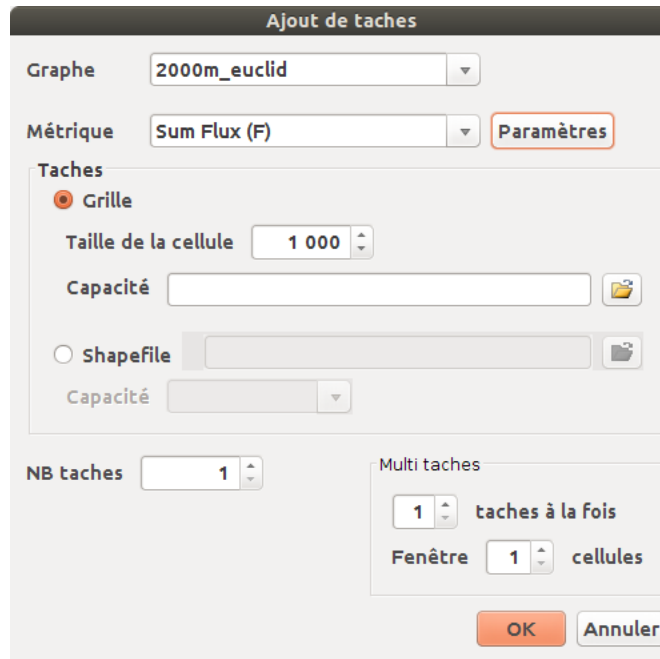
L'option Connexion multiple permet de faire intervenir plusieurs taches dans le calcul des métriques au niveau des points. Le calcul est fondé sur une somme ou une moyenne pondérée des valeurs issues de toutes les taches comprises dans le voisinage des points, jusqu'à la Distance maximale indiquée.

Il est à noter que la distance considérée dans ces calculs dépend du graphe de référence. Si celui-ci est fondé sur une distance de moindre coût, l'interpolation spatiale utilise cette même distance et non une simple distance euclidienne.

L'interpolation des métriques est aussi utilisée automatiquement dans le calcul des modèles de distribution d'espèce (voir [Modèle de distribution d'espèce](#)).

6 Ajout de taches

La fonction "Ajout de taches" est accessible à partir du menu Analyses. Elle permet de tester un ensemble d'emplacements pour créer de nouvelles taches d'habitat en fonction du gain de connectivité qu'elles procurent. Le logiciel calcule d'abord la métrique globale sélectionnée, puis ajoute virtuellement une tache et les liens entre cette tache et les taches existantes (seulement si leur distance est inférieure au seuillage du graphe dans le cas d'un graphe seuillé) et recalcule la métrique globale. Une fois tous les emplacements testés, le logiciel valide l'emplacement qui procure le gain maximal de connectivité. La procédure est répétée jusqu'à l'ajout du nombre souhaité de nouvelles taches.



Plusieurs paramètres sont à définir :

- Graphe : le graphe utilisé pour le calcul de la métrique de connectivité et l'ajout de nouvelles tâches. Le jeu de lien sous-jacent doit être en topologie complète.
- Métrique : la métrique globale que l'on cherche à maximiser ainsi que son paramétrage (bouton Paramètres)
- NB tâches : nombre de tâches (qui deviendront de nouveaux nœuds du graphe) que l'on souhaite ajouter.

La définition du jeu de tâches à tester peut se faire de deux manières :

- En appliquant automatiquement une grille d'une certaine résolution (champs "taille de la cellule" à renseigner). Dans ce cas, le logiciel teste chaque centroïde de cellule en ajoutant virtuellement une nouvelle tâche. Par défaut, tous les centroïdes ont une valeur de capacité identique (=1). Il est néanmoins possible d'importer une couche raster où chaque pixel de la matrice paysagère a une valeur de capacité propre.
- En important un shapefile de points ou de polygones. Il faut dans ce cas indiquer à quelle colonne de la table attributaire de ce shapefile correspond la capacité des tâches.

Option "multi-tâches" : cette option permet de tester l'ajout de plusieurs tâches en simultanément. Les tâches supplémentaires sont recherchées dans une fenêtre de x cellules (paramètre à définir) autour de la première tâche.

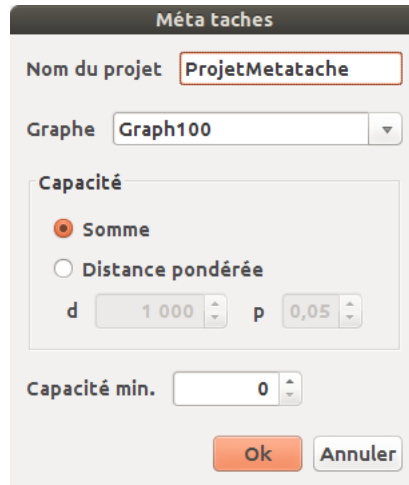
A la fin du calcul, un répertoire nommé "addpatch_nX-GraphX_..." est créé et enregistré dans le répertoire du projet. Il contient :

- un shapefile contenant les nœuds ajoutés. La table attributaire indique l'étape à laquelle a été validé chacun des nœuds et la valeur de la métrique globale correspondante.
- deux shapefiles contenant l'ensemble des liens du graphe, y compris ceux reliant les nouveaux nœuds, en version topologique ("topo_links_graphXXX") ou en version réaliste "links_graphXX"
- un raster "landuse.tif" contenant l'occupation du sol modifiée
- un répertoire "détail" contenant pour chaque étape le shapefile des emplacements testés et la valeur de la métrique globale correspondant.

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

7 Méta-taches

À partir du menu Graphe, l'option "Créer un projet de méta-tache" permet de créer un nouveau projet où les noeuds du graphe correspondent à des ensembles de tâches connectées à une certaine distance. Cette fonction permet de mieux prendre en compte l'emboîtement des échelles dans les processus écologiques et notamment le fait que la définition fonctionnelle d'une tâche d'habitat varie selon l'échelle considérée (Theobald, 2006; Zetterberg et al., 2010).



Pour utiliser cette fonction, il est nécessaire de créer un projet initial et un graphe seuillé à la distance souhaitée (par exemple à la distance quotidienne). Chaque composante de ce graphe seuillé correspondra à une méta-tache. Celle-ci est donc constituée de l'ensemble des tâches d'habitat connectées entre elles.

La capacité des méta-taches est définie par défaut comme la somme de la capacité des tâches composant la méta-tache. L'autre option "Distance pondérée" permet de tenir compte de l'éloignement des tâches entre elles. Pour cette option, il faut définir le paramètre α en donnant une distance d et une probabilité p associée (cf. [Paramétrage des métriques pondérées](#)). La capacité résultante (C) d'une méta-tache s'écrit :

$$C = \frac{1}{n} \sum_i^n \sum_j^n c_j e^{-\alpha d_{ij}}$$

L'option "capacité minimale" permet de ne garder pour le nouveau projet que les méta-taches qui ont une capacité supérieure ou égale au seuil défini.

Le nouveau projet est enregistré dans un sous répertoire du projet initial et est chargé automatiquement à la fin du traitement.

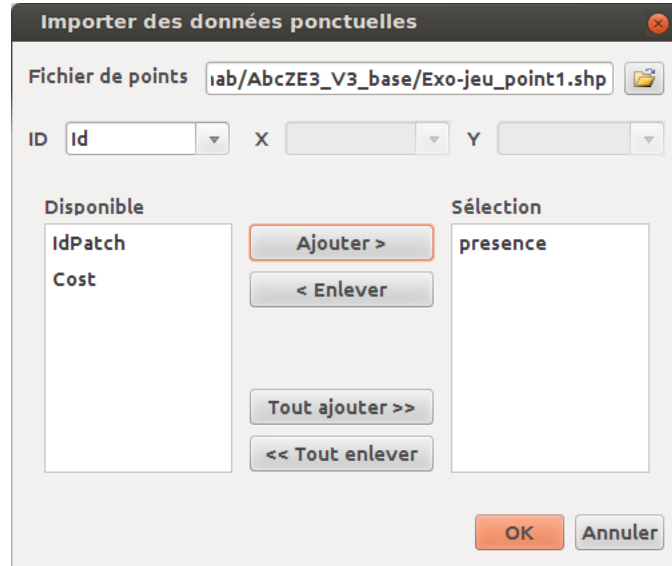
Références : [[Clauzel et al.\(2015b\)](#)]

8 Liens entre graphes et données ponctuelles

Le corps principal du logiciel est la mise en place des graphes et le calcul de métriques de connectivité, mais il est souvent très utile de relier ces éléments à des données externes. Le logiciel permet de faire interagir les données issues des graphes avec un jeu de données de points.

8.1 Import de données ponctuelles

Des données ponctuelles peuvent être importées par le menu : Données / Importer des données ponctuelles. Ces données peuvent contenir plusieurs attributs, mais seuls les attributs binaires (présence/absence) seront pris en compte dans certaines procédures (voir Modèles de distribution d'espèce au 6.3).

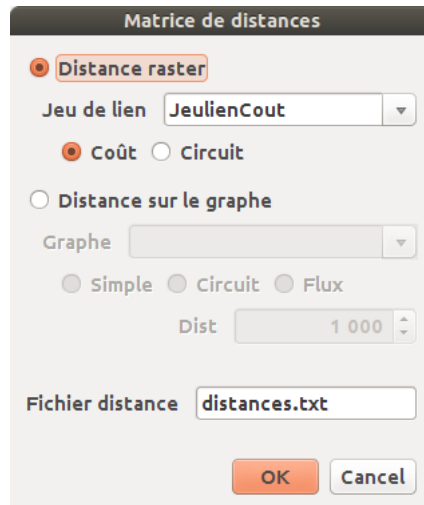


Le fichier importé peut être soit sous forme de couche vectorielle au format shapefile (*.shp), soit sous forme de tableau (*.csv). Dans le second cas, il faut indiquer les colonnes du tableau qui correspondent aux identifiants et aux coordonnées en X et en Y des points. Les attributs à prendre en compte doivent être sélectionnés à partir de la liste des attributs disponibles.

Si les données ponctuelles ne contiennent pas d'attribut d'absence, elles ne peuvent pas être utilisées dans des modèles de distribution d'espèce. Si l'utilisateur envisage de mettre en place un modèle de distribution d'espèce, l'utilisateur doit créer un jeu de données de pseudo-absence avec Données / Générer des points aléatoires.

8.2 Matrices de distance inter-points

Les données ponctuelles importées dans Graphab peuvent être à l'origine du calcul de matrices de distance inter-points par le menu Matrice de distance du menu contextuel d'un jeu de points (clic droit sur le nom d'un jeu de points).



Les distances peuvent être calculées de 2 manières : directement sur le raster ou bien à travers un graphe.

8.2.1 Distance raster

En mode raster, les distances sont calculées en fonction d'un jeu de lien de référence. Ce jeu de lien ne peut pas être en distance euclidienne.

Selon le jeu de lien choisi, il peut s'agir de la distance coût cumulée ou de la longueur du chemin de moindre coût. Dans les deux cas, le calcul intègre les coûts associés aux catégories de la carte de paysage, tels qu'ils ont été définis dans le jeu de liens. Le résultat est une matrice de distance indépendante du graphe; cette matrice correspond au calcul offert par les Systèmes d'Information Géographique.

L'option Circuit permet de calculer la "distance de résistance" entre chaque point, c'est-à-dire la résistance équivalente d'un réseau de résistances électriques au lieu du calcul du chemin le plus court. Cette méthode permet de tenir compte de l'ensemble des chemins possibles et non pas uniquement du plus court (cf. [McRae et al.(2008)]).

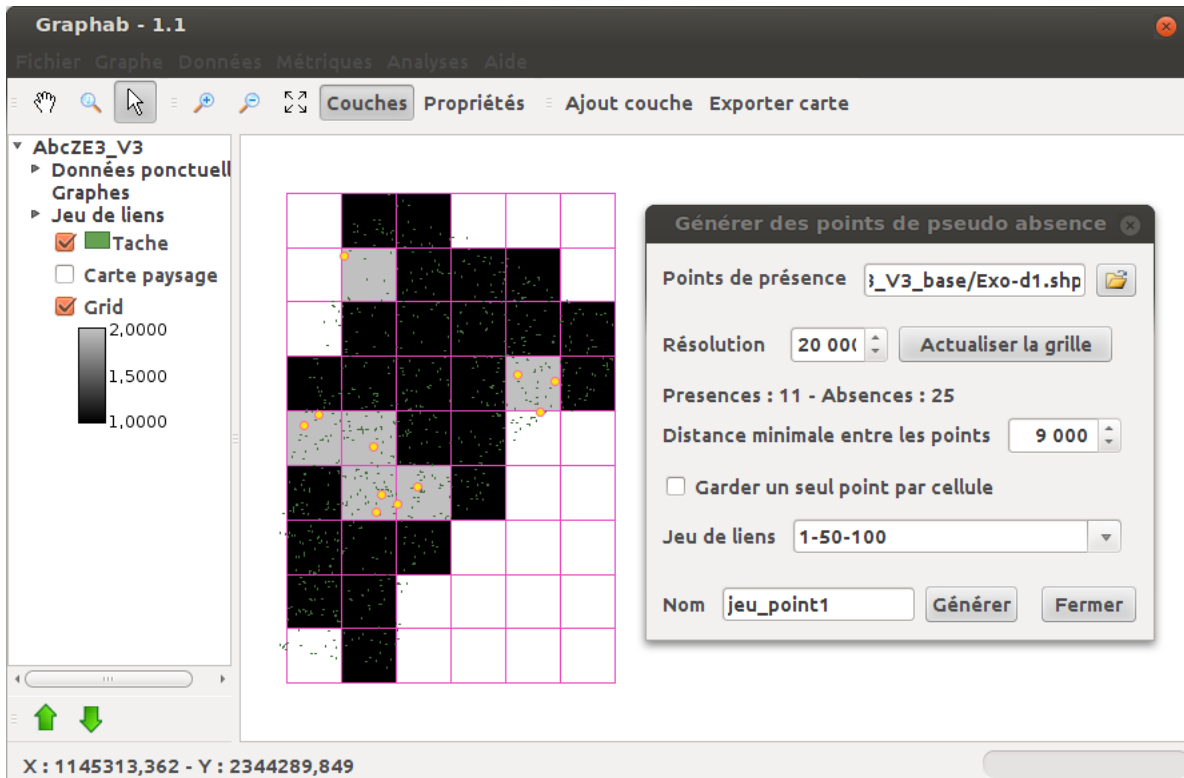
8.2.2 Distance sur le graphe

En mode graphe, les distances sont calculées en fonction du graphe sélectionné. Ce graphe doit être basé sur le jeu de lien utilisé à l'import du jeu de point.

- Simple : distance calculée suivant le plus court chemin sur le graphe de référence. Le type de distance correspond à celui qui a été utilisé pour la définition du jeu de lien qui a servi à définir le graphe de référence. Aux deux extrémités du chemin considéré, le calcul intègre la distance entre chaque point et la tache la plus proche. Suivant le choix qui a été fait au moment de la création du graphe, le calcul intègre éventuellement les distances intra-taches.
- Circuit : calcule la "distance de résistance" entre les taches les plus proches de chaque point. Le graphe est vu comme un réseau électrique où chaque lien du graphe représente une résistance. Cette option permet de tenir compte de l'ensemble des chemins possibles et non pas uniquement du plus court.

8.3 Générer des points aléatoires

Le menu Données / Générer des points aléatoires permet de créer un ensemble de points de pseudo-absence en s'appuyant sur un ensemble de points de présence.



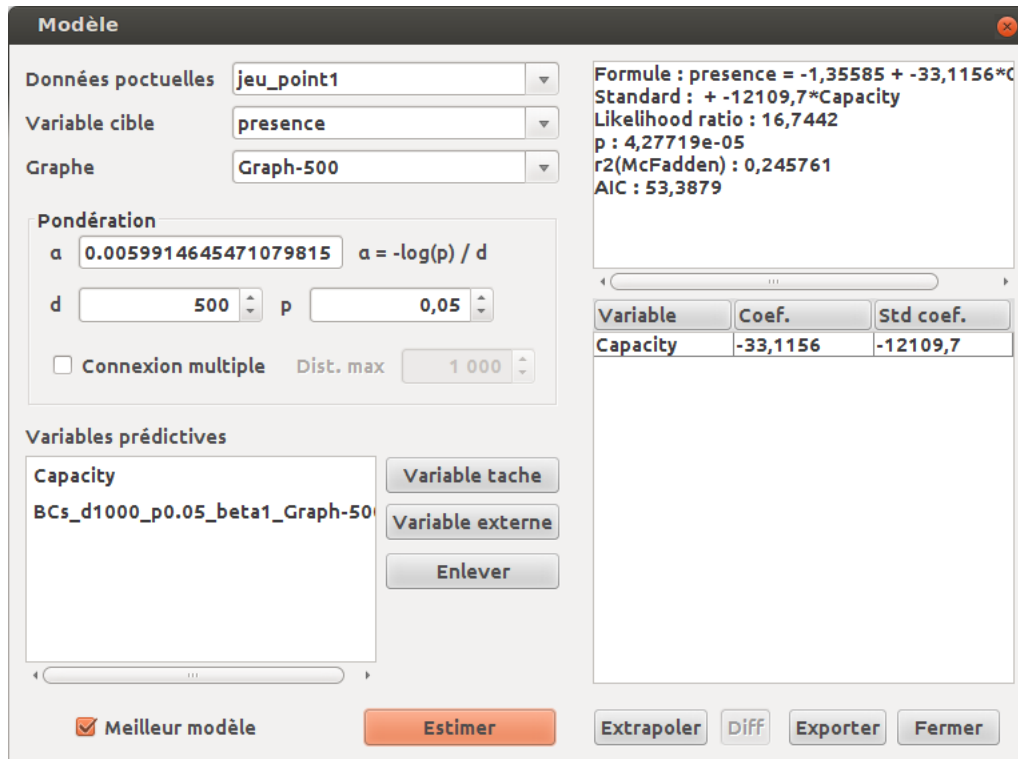
L'utilisateur doit charger un fichier de points de présence et indiquer le nom du jeu de points de présence / absence à créer.

Plusieurs paramètres sont à définir pour effectuer un tirage aléatoire des points d'absence :

- Résolution de la grille (en mètres) pour définir la taille des cellules dans lesquelles seront éventuellement tirés des points d'absence. Le bouton « actualiser la grille » permet d'afficher la grille selon la résolution choisie
- Distance minimale entre les points (en mètres) : cette fonction permet d'atténuer les effets d'autocorrélation spatiale en indiquant une distance minimale à respecter entre les points d'absence générés et entre ces points et les points de présence.
- Type de distance : l'unité de la distance minimale entre les points dépend du type de distance utilisé dans le jeu de liens sélectionné.
- Garder un point par cellule : cette option permet de ne garder qu'un point de présence dans chaque cellule pour atténuer les effets d'autocorrélation spatiale.

8.4 Modèle de distribution d'espèce

Si un jeu de points de présence / absence a été défini, le logiciel permet d'utiliser les métriques de connectivité calculées à partir d'un graphe comme variables prédictives dans un modèle de distribution d'espèce, suivant le menu Analyse / Estimation du modèle. Cette modélisation est possible même si les points ne sont pas situés dans les taches d'habitat, par le biais d'une extrapolation spatiale des valeurs des métriques. Le modèle est une régression logistique fondé sur la minimisation du critère AIC.



Au préalable, l'utilisateur doit indiquer :

- le jeu de données ponctuelles à analyser ;
- la variable cible dans le modèle prédictif ;
- le graphe de référence pour l'utilisation des métriques de connectivité.

8.4.1 Pondérations pour l'extrapolation des métriques au niveau des points

Les valeurs des métriques sont calculées pour chaque point d'analyse par le biais d'une interpolation spatiale. Celle-ci est fondée sur une pondération des valeurs par une fonction décroissante à partir du bord des taches (poids de 1). Le poids décroît selon une fonction exponentielle négative de la même forme que celle qui est utilisée pour les métriques pondérées (voir 5.2.1), c'est pourquoi on retrouve le même réglage.

L'utilisateur choisit une distance (d) correspondant à une certaine probabilité (p) et le logiciel en déduit la valeur du paramètre α . En principe, ce réglage est à rendre cohérent avec le choix du graphe de référence ou le choix des éventuelles métriques pondérées qui sont à mobiliser, en utilisant la même valeur de d .

L'option Connexion multiple permet de faire intervenir plusieurs taches dans le calcul des métriques au niveau des points. Le calcul est fondé sur une moyenne pondérée des valeurs issues de toutes les taches comprises dans le voisinage des points, jusqu'à la Distance maximale indiquée.

Les détails de cette forme de pondération sont donnés dans [Foltête et al.(2012b)].

8.4.2 Estimation du modèle

Le choix d'un graphe pour appliquer le modèle fait apparaître toutes les métriques de connectivité disponibles parmi les variables prédictives.

Des métriques dérivées d'autres graphes peuvent être ajoutées en cliquant sur le bouton Variable tache.

Des variables provenant d'autres sources peuvent aussi être ajoutées, en cliquant sur le bouton Variable externe.

Une fois que les variables prédictives que l'utilisateur veut intégrer sont en place, le bouton Estimer permet de calculer les coefficients de la régression logistique. Les résultats s'affichent sur la partie droite de la fenêtre. L'option Meilleur modèle consiste à tester toutes les combinaisons de variables et de sélectionner celle qui minimise le critère AIC.

8.4.3 Utilisation du modèle

Un modèle prédictif considéré comme valide peut être utilisé de plusieurs façons.

Suivant le bouton Exporter, le tableau contenant toutes les variables statistiques impliquées dans la régression peut être exporté au format *.csv.

Le bouton Extrapoler permet d'estimer la probabilité de présence de l'espèce dans toutes les cellules d'une grille.

Extrapolation

Jeu de lien: 1-50-100

Pondération

α : 0.0059914645471079815 $\alpha = -\log(p) / d$

d: 500 p: 0,05

Connection multiple Dist. max: 1 000

Variables prédictives

Variable	Coef
Constant	-1.3558473054645068
Capacity	-33.1155923384945

Attribut tache Image externe Enlever

Résolution: 100

Nom: Extrapol1 Extrapoler Fermer

Dans la nouvelle fenêtre qui s'affiche, l'utilisateur retrouve les paramètres du modèle tels qu'ils ont été décrits précédemment. La résolution de la grille (en mètres) indique le niveau de précision spatiale selon lequel cette extrapolation est effectuée. Ce critère a une forte incidence sur le temps de calcul nécessaire pour obtenir le résultat. Le résultat est archivé sous forme de couche raster au format *.tif et affiché dans la fenêtre principale.

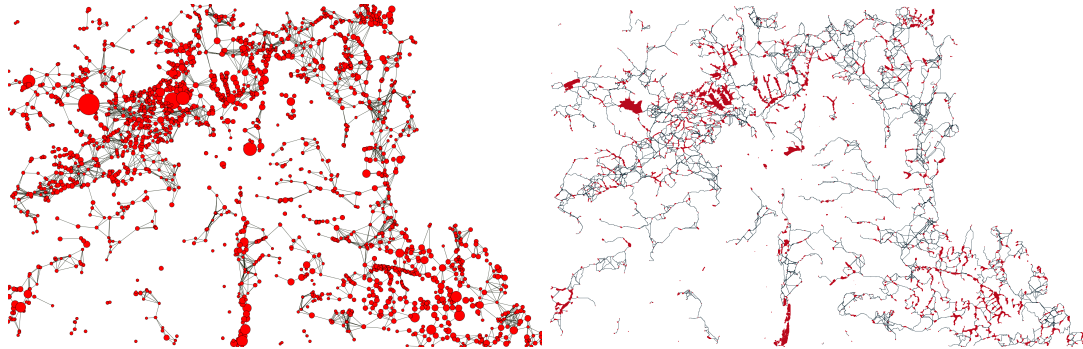
Références : [Foltête et al.(2012b), Foltête et al.(2012a), Girardet et al.(2013), Clauzel et al.(2013)]

9 Affichage

9.1 Propriétés des graphes

Les propriétés des graphes sont accessibles par un clic droit sur le nom du graphe. Deux modes de visualisation des graphes sont disponibles :

- La vue topologique affiche une vue simplifiée du graphe où les nœuds sont représentés par des cercles et les liens par des lignes droites de centroïde à centroïde.
- La vue réaliste affiche les taches d’habitat selon leurs limites réelles et les liens sont représentés par les chemins de moindre coût entre deux taches.



Gauche : vue topologique - Droite : vue réaliste

L’entrée Supprimer permet de supprimer le graphe sélectionné. Le projet est enregistré automatiquement après.

L’entrée Propriétés rappellent les paramètres utilisés pour la construction du graphe : le nom du graphe, le jeu de liens sélectionné, le type de graphe avec éventuellement le seuil de distance utilisé et le nombre de liens.

L’entrée Matrice OD (matrice Origine-Destination) permet de créer un tableau indiquant la distance entre chaque paire de nœuds pour le graphe sélectionné. L’unité de la distance dépend du type de distance utilisé pour ce graphe. L’absence de connexion entre deux nœuds est notée NaN (Not a Number). Cette matrice est archivée dans le dossier du projet sous la forme d’un fichier texte nommé « nom du graphe-odmatrix.txt ».

9.2 Propriétés des objets

Les propriétés de chaque jeu de liens, des éléments d’un graphe (nœuds, liens ou composantes) et des données ponctuelles sont accessibles par un clic droit sur chacun d’entre eux.

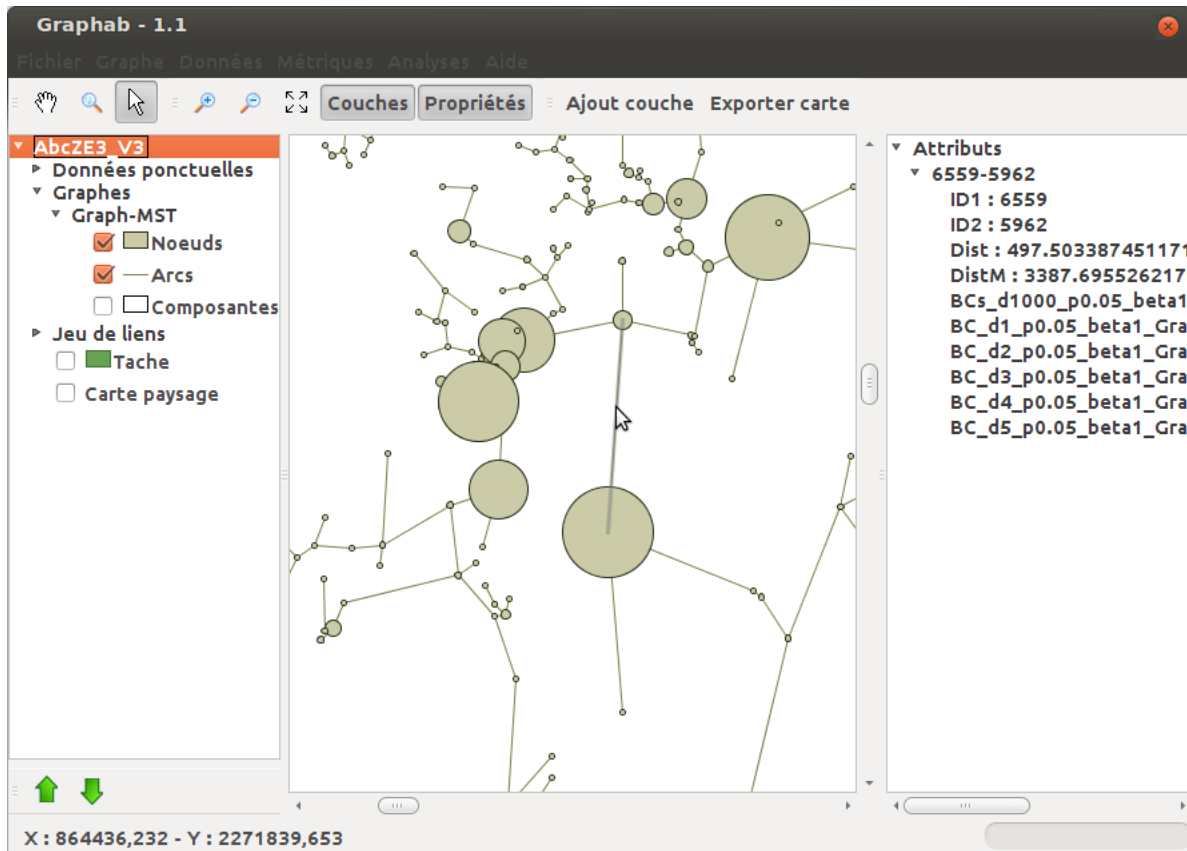
Le menu Style donne accès aux paramètres d’affichage des objets : couleur, épaisseur des contours, étiquettes, taille des symboles (pour les nœuds uniquement). Les objets peuvent être représentés de la même manière (symbole unique) ou en fonction d’un attribut. Il est possible d’appliquer une méthode de discrétisation pour classer les objets en fonction des valeurs de l’attribut sélectionné. Par défaut, la légende des objets est affichée dans la table des matières. Il est possible de la cacher en décochant le bouton Légende.

Le menu Export permet d’exporter les objets dans un fichier vecteur (*.shp) ou un fichier texte (*.txt).

Le menu Statistiques permet d’afficher la distribution des valeurs d’un ou plusieurs attributs :

- nuage de point : affiche un nuage de point mettant en relation les valeurs de deux attributs
- histogramme : affiche l’histogramme des valeurs d’un attribut.

Il est également possible d'afficher les valeurs d'un objet particulier en le sélectionnant à l'aide de la flèche. Après sélection, la valeur des attributs s'affiche dans une nouvelle colonne à droite. Il est possible de fermer cette colonne en cliquant sur le menu Propriétés dans la barre du haut.



10 Capacités de traitement et limitations

Les méthodes basées sur les graphes sont connues pour fournir un cadre de modélisation efficace, mais peuvent poser une question de capacité de calcul. Deux points spécifiques ont reçu une attention particulière dans Graphab : (1) le calcul des jeux de liens, (2) le calcul des métriques de connectivité. Tous ces calculs ont été optimisés par parallélisation. Ce mode de développement améliore l'efficacité de calcul en utilisant une architecture multiprocesseur, un processeur quatre-cœur permettant théoriquement un calcul quatre fois plus rapide.

Dans [Foltête et al. (2012a)], plusieurs tests ont été effectués afin de mesurer la capacité de calcul de Graphab dans différentes configurations. Trois configurations ont été comparées : 1) un cœur (3Go RAM) correspondant à un ordinateur standard ; 2) quatre cœurs (6 Go RAM) correspondant à une station de travail ; 3) 20 cœurs (15 Go RAM) correspondant à un serveur. La carte d'occupation du sol était un fichier raster de 14000*18000 pixels (252 millions de pixels) représentant l'occupation du sol de la région Franche-Comté (France) à une résolution spatiale de 10 mètres. La carte contenait 22 634 taches d'habitats. Le tableau 3 donne quelques temps de calcul pour la création de liens.

Topologie	Distance	Ordinateur standard	Station de travail	Serveur
Complète	Euclidienne	1927s (32 min)	516s (8 min)	133s (2 min)
	Moindre coût	19252s (5h 21 min)	4301s (1h 11 min)	1037s (17 min)
Planaire	Euclidienne	43s	12s	2.6s
	Moindre coût	1080s (18 min)	295s (5 min)	82s (1 min)

TABLE 3 – Temps de calcul (en seconde) pour la création de différents jeux de liens.

La mémoire allouée au programme joue aussi un rôle important. Si il n'y a pas assez de mémoire vive, le calcul sera plus lent voire même ne pourra pas aboutir (message `OutOfMemoryError` ou `GC Overhead`). Le menu Fichier / Préférences / Mémoire permet de régler la mémoire allouée pour Graphab. Si vous avez une version 32 bits de Java, Graphab sera limité à environ 2 Go (2000 Mo) de mémoire. Si votre ordinateur a plus que 2 Go de mémoire vive il est fortement conseillé d'installer une version 64 bits de Java pour utiliser la mémoire disponible au-delà des 2 Go.

11 Annexes

11.1 Détail de calcul des métriques

Famille	Métrique de connectivité	Code	Niveau de calcul			Delta métriques
			Global	Comp.	Local	
Métriques pondérées	Flux	F	×	×	×	
	Probabilité de connectivité	PC	×	×		×
	Probabilité de connectivité flux	FPC			×	
	Fractions de delta Probabilité de connectivité	dPC				×
	Indice de centralité intermédiaire	BC			×	
	Indice intégral de connectivité	IIC	×	×		×
	Flux circuit	CF			×	
Métriques de surface	Taille moyenne des composantes	MSC	×			
	Taille de la plus grande composante	SLC	×			
	Probabilité de coïncidence de classe	CCP	×			
	Taille de composante attendue	ECS	×			
Métriques topologiques	Degré	Dg			×	
	Coefficient de groupement	CC			×	
	Centralité de proximité	CCe			×	
	Excentricité	Ec			×	
	Corrélation de connectivité	CCor			×	
	Nombre de composantes	NC	×			
	Diamètre	GD	×	×		×
	Indice d'Harary	H	×	×		×

TABLE 4 – Tableau récapitulatif des métriques présentes dans Graphab

Terme	Signification
n	Nombre de taches
nc	Nombre de composantes
n_k	Nombre de taches de la composante k
N_i	Ensemble des taches voisines de la tache i
a_i	Capacité de la tache i (en général sa surface)
ac_k	Capacité de la composante k (somme des capacités des taches qui la composent)
A	Surface de la zone d'étude
d_{ij}	distance entre les taches i et j (en général la distance de moindre coût qui les sépare)
$e^{-\alpha d_{ij}}$	Probabilité de déplacement entre les taches i et j
α	Frein de la distance vis-à-vis des déplacements
β	Exposant permettant de pondérer plus ou moins la capacité

TABLE 5 – Termes mathématiques utilisés

11.1.1 Métriques pondérées

Flux (F)		
	Formule	Signification
Niveau global	$S\#F = \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n a_j^\beta e^{-\alpha d_{ij}}$	Pour le graphe entier : somme des dispersions potentielles à partir de toutes les taches.
Niveau local	$F_i = \sum_{\substack{j=1 \\ j \neq i}}^n a_j^\beta e^{-\alpha d_{ij}}$	Pour la tache focale i : somme des capacités des taches différentes de i et pondérées en fonction de leur distance minimale à la tache focale en parcourant le graphe. Cette somme est un indicateur de la dispersion potentielle à partir de la tache i , ou inversement à destination de la tache i .
Valeurs	Les valeurs dépendent de la définition de a . Si a représente une surface, F exprime une surface. Valeur minimale : 0 Valeur maximale : surface totale de l'habitat	
Commentaire	Le chemin du graphe utilisé est celui qui maximise $e^{-\alpha d}$, c'est-à-dire celui qui minimise la distance d (ou le coût) entre les taches i et j . Cette métrique est appelée Area Weighed Flux (AWF) dans certaines références. Dans Graphab cependant, a est plus général car il représente la capacité des taches, qui peut être leur surface ou un autre critère choisi par l'utilisateur. De même la pondération est variable en fonction du paramètre β . Dans CS22, AWF est calculé seulement à partir des taches reliées directement à la tache focale, alors que Graphab tient compte des taches indirectement connectées.	
Références	[Urban and Keitt(2001)] [Saura and Torné(2009)] [Foltête et al.(2012b)]	

Probabilité de Connectivité (PC)		
	Formule	Signification
Niveau global Niveau composante Delta	$PC = \frac{1}{A^2} \sum_{i=1}^n \sum_{j=1}^n a_i^\beta a_j^\beta e^{-\alpha d_{ij}}$	Somme des produits de la capacité de tous les couples de taches pondérées par leur probabilité d'interaction, divisée par le carré de la zone d'étude. Ce rapport équivaut à la probabilité que deux individus tirés au hasard dans la zone d'étude parviennent à entrer en contact.
Valeurs	Les valeurs correspondant à une probabilité. Valeur minimale : 0 Valeur maximale : 1	
Commentaire	Pour chaque couple de taches, le chemin du graphe utilisé est celui qui maximise $e^{-\alpha d}$, c'est-à-dire celui qui minimise la distance d (ou le coût) entre les taches i et j . Dans CS22, la pondération des capacités est fixée à 1 ; dans Graphab elle peut être modifiée. Si a ne représente pas la surface des taches, le résultat n'est plus une probabilité.	
Référence	[Saura and Pascual-Hortal(2007)]	

Probabilité de Connectivité Flux (FPC)		
	Formule	Signification
Niveau local	$FPC_i = \frac{1}{A^2} \sum_{j=1}^n a_i^\beta a_j^\beta e^{-\alpha d_{ij}}$	Somme des produits de la capacité de la tache focale avec toutes les autres taches, pondérées par leur probabilité d'interaction, divisée par le carré de la surface de la zone d'étude.
Valeurs	Valeur minimale : 0 Valeur maximale : 1	
Commentaire	Pour chaque couple de taches, le chemin du graphe utilisé est celui qui maximise $e^{-\alpha d}$, c'est-à-dire celui qui minimise la distance d (ou le coût) entre les taches i et j . Cette métrique est simplement la contribution locale d'une tache à l'indice PC , puisque $PC = \sum_i FPC_i$. Elle équivaut à $dPC_{flux} + dPC_{area}$ non divisé par la valeur globale du PC . Cependant la métrique FPC est obtenue plus rapidement car son calcul n'est pas fondé sur la suppression itérative des taches (mode delta).	

Fractions de delta Probabilité de Connectivité ($dPC, dPC_{area}, dPC_{flux}, dPC_{connector}$)		
	Formules	Signification
Delta	$dPC_i = \frac{(PC - PC'_i)}{PC}$ $dPC_i = dPC_{area} + dPC_{flux} + dPC_{connector}$ $dPC_{area} = \frac{a_i^{2\beta}}{A^2 PC}$ $dPC_{flux} = \frac{FPC_i}{PC} - dPC_{area}$	<p>Taux de variation entre la valeur de l'indice PC et la valeur PC' représentant la suppression de la tâche i.</p> <p>La valeur de dPC se décompose en trois parties :</p> <ul style="list-style-type: none"> - dPC_{area} correspond à la variation due à la perte de la surface suite à la suppression de la tâche i ; - dPC_{flux} correspond à la variation due à la perte des interactions entre la tâche i et les autres tâches ; - $dPC_{connector}$ correspond à la variation due à la modification des chemins reliant les autres tâches et passant initialement par i.
Valeurs	Valeur minimale : 0 Valeur maximale : 1	
Commentaire	Dans CS22, la pondération des capacités est fixée à 1 ; dans Graphab elle peut être modifiée. Si a ne représente pas la surface des tâches, le résultat n'est plus une probabilité et dPC_{area} n'exprime pas une perte de surface mais de capacité.	
Référence	[Saura and Rubio(2010)]	

Indice de centralité intermédiaire (BC)		
	Formule	Signification
Niveau local	$BC_i = \sum_j \sum_k a_j^\beta a_k^\beta e^{-\alpha d_{jk}}$ $j, k \in \{1..n\}, k < j, i \in P_{jk}$	<p>Somme des plus courts chemins passant par la tâche focale i, chaque chemin étant pondéré par le produit des capacités des tâches reliées et de leur probabilité d'interaction.</p> <p>P_{jk} représente l'ensemble des tâches traversées par le plus court chemin entre les tâches j et k.</p>
Valeurs	Les valeurs dépendent du paramétrage. Elles correspondent à un poids de transit potentiel. Valeur minimale : 0 Valeur maximale : surface totale de l'habitat au carré.	
Commentaire	<p>Avec un paramétrage de $\alpha = 0$ et $\beta = 0$ (pondération uniforme des chemins), on retrouve l'indice BC tel qu'il est utilisé dans d'autres types de graphes.</p> <p>Un paramétrage de $\alpha = 0$ et $\beta = 1$ donne aux chemins un poids égal au produit des capacités des tâches qu'ils relient, quelle que soit leur distance.</p> <p>Dans [Foltête et al.(2012a), Foltête et al.(2012b)], l'indice BC_l a été proposé pour pondérer plus fortement les chemins supérieurs au critère choisi (distance de dispersion par exemple). Mais des tests ont ensuite montré que cet indice était très fortement corrélé au BC pondéré avec $\alpha = 0$.</p> <p>Dans [Bodin and Saura(2010)], le BC_{pc} correspond au BC pondéré avec d égal à la distance de dispersion, α tel que $e^{-\alpha d} = 0.05$ et $\beta = 1$.</p>	
Références	[Bodin and Saura(2010)] [Foltête et al.(2012a)]	

Indice Intégral de Connectivité (<i>IIC</i>)		
	Formule	Signification
Niveau global Niveau composante Delta	$IIC = \frac{1}{A^2} \sum_{i=1}^n \sum_{j=1}^n \frac{a_i a_j}{1 + nl_{ij}}$	Produit des capacités des taches divisé par le nombre de liens qui les sépare, la somme étant divisée par le carré de la surface de la zone d'étude. IIC est construit comme l'indice PC, mais en utilisant l'inverse d'une distance topologique plutôt qu'une fonction exponentielle négative de la distance basée sur l'impédance des liens.
Valeurs	Valeur minimale : 0 Valeur maximale : 1	
Référence	[Pascual-Hortal and Saura(2006)]	

Flux circuit (<i>CF</i>)		
	Formule	Signification
Niveau local	$CF_i = \sum_j^n c_i^j$	Somme des courants traversant la tache focale i . c_i^j représente le courant à travers la tache i lorsque un courant est émis depuis toutes les taches vers la tache j (la tache j est reliée à la terre)
Valeurs	Valeur minimale : 0 Valeur maximale : $(n - 1)(n - 2)$ si $\beta = 0$ $(n - 2) \sum_i^{n-1} a_i$ si $\beta = 1$	
Commentaire	La métrique <i>CF</i> utilise la théorie des circuits électriques. Chaque lien du graphe correspond à une résistance électrique, les sources de courants et la terre sont quant à elles attachées aux taches. Si $\beta = 0$ chaque tache émet un courant unitaire, si $\beta = 1$ chaque tache émet un courant égal à sa capacité. Cette métrique peut être vue comme un équivalent de la métrique BC (avec $\alpha = 0$ et $\beta = 0$) qui tient compte de l'ensemble des chemins possibles et non pas uniquement du chemin le plus court.	
Références	[Girardet et al.(2015)]	

11.1.2 Métriques de surface

Taille moyenne des composantes (MSC)		
	Formule	Signification
Niveau global	$MSC = \frac{1}{nc} \sum_{k=1}^{nc} ac_k$	Moyenne des capacités des composantes.
Valeurs	Valeur minimale : capacité minimale Valeur maximale : SLC	

Taille de la plus grande composante (SLC)		
	Formule	Signification
Niveau global	$SLC = \max\{ac_k\}$	Plus grande capacité des composantes.
Valeurs	Valeur minimale : capacité minimale Valeur maximale : capacité maximale	

Probabilité de coïncidence de classe (CCP)		
	Formule	Signification
Niveau global	$CCP = \sum_{k=1}^{nc} \left(\frac{ac_k}{\sum_l ac_l} \right)^2$	Probabilité que deux points choisis au hasard sur le graphe appartiennent à la même composante.
Valeurs	Valeur minimale : $\rightarrow 0$ (autant de composantes que de tâches et capacités uniformes) Valeur maximale : 1 (une seule composante)	
Référence	[Pascual-Hortal and Saura(2006)]	

Taille de composante attendue (ECS)		
	Formule	Signification
Niveau global	$ECS = \frac{1}{\sum_k ac_k} \sum_{k=1}^{nc} ac_k^2$	Taille attendue d'une composante
Valeurs	Valeur minimale : capacité minimale (autant de composantes que de taches et capacités uniformes) Valeur maximale : somme des capacités (1 seule composante)	
Référence	[O'Brien et al.(2006)]	

11.1.3 Métriques topologiques

Indice d'Harary (H)		
	Formule	Signification
Niveau global Niveau composante Delta	$H = \frac{1}{2} \sum_{i=1}^n \sum_{\substack{j=1 \\ j \neq i}}^n \frac{1}{nl_{ij}}$	Somme de l'inverse du nombre de liens séparant tous les couples de taches.
Valeurs	Valeur minimale : 0 pour un graphe sans lien (autant de composantes que de taches) Valeur maximale : $\frac{n(n-1)}{2}$ pour un graphe complet	
Commentaire	Pour les couples de taches non connectées par un chemin on a : $nl_{ij} = +\infty$	
Référence	[Ricotta et al.(2000)]	

Diamètre (GD)		
	Formule	Signification
Niveau global Niveau composante Delta	$GD = \max_{ij} d_{ij}$ $GD = \max_i Ec_i$	Distance la plus grande entre deux taches du graphe
Valeurs	Valeur minimale : 0 Valeur maximale : $+\infty$	
Commentaire	Quand les nœuds i et j ne sont pas connexes $d_{ij} = 0$ Cette métrique est la version globale de la métrique Ec_i	

Nombre de composantes (NC)		
	Formule	Signification
Niveau global	$NC = nc$	Nombre de composantes du graphe
Valeurs	Valeur minimale : 1 Valeur maximale : n	

Degré du noeud (Dg)		
	Formule	Signification
Niveau local	$Dg_i = N_i $	Nombre de liens connectés au noeud i <i>ie.</i> nombre de taches voisines de la tache i
Valeurs	Valeur minimale : 0 Valeur maximale : n	
Commentaire	Il y a équivalence entre le degré d'un nœud et le nombre de nœuds voisins car les graphes ne sont pas orientés et ne contiennent pas de boucle.	

Coefficient de groupement (CC)		
	Formule	Signification
Niveau local	$CC_i = \frac{1}{ N_i (N_i - 1)} \sum_{j \in N_i} N_i \cap N_j $	Rapport du nombre de nœuds voisins de i voisins entre eux sur le total possible.
Valeurs	Valeur minimale : 0 Valeur maximale : 1	
Commentaire	Si $ N_i \leq 1 \rightarrow CC_i = 0$	

Centralité de proximité (CCe)		
	Formule	Signification
Niveau local	$CCe_i = \frac{1}{n_k - 1} \sum_{\substack{j=1 \\ j \neq i}}^{n_k} d_{ij}$	Distance moyenne de la tâche i vers toutes les autres tâches de sa composante k .
Valeurs	Valeur minimale : 0 Valeur maximale : $+\infty$	
Commentaire	Si $n_k = 1 \rightarrow CCe_i = 0$	

Excentricité (Ec)		
	Formule	Signification
Niveau local	$Ec_i = \max_j d_{ij}$	Distance maximale de la tâche i vers une autre tâche de sa composante.
Valeurs	Valeur minimale : 0 Valeur maximale : $+\infty$	

Corrélation de connectivité (CCor)		
	Formule	Signification
Niveau local	$CCor_i = \frac{ N_i ^2}{\sum_{j \in N_i} N_j }$	Rapport entre le degré du nœud i et le degré de ses nœuds voisins j
Valeurs	Valeur minimale : 0 Valeur maximale : $ N_i $	
Commentaire	Si $ N_i = 0 \rightarrow CCor_i = 0$	
Référence	[Minor and Urban(2008)]	

Références

- [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. Dellinger, M. Gaertler, R. Gorke, M. Hofer, 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.
- [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-Rogéon, 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 prioritization 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.
- [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.