

## Analyse Spatiale - Examen Final

### Exercice : Mesures de Réseau et Résilience

La description complète d'un réseau suppose la donnée d'une grande quantité d'information, puisque la moitié de la matrice d'adjacence est nécessaire dans le cas non-dirigé (matrice complète dans le cas dirigé). L'établissement de typologies, c'est à dire de sortes de classes d'équivalence topologiques au sens large, est une façon de voir les enjeux de la recherche actuelle sur les réseaux : existe-t-il des formes typiques de réseau, et comment les représenter dans une dimension réduite ? Relativement importante est la dimension épistémologique de cette interrogation fondamentale, qui peut sous certaines hypothèses être ramenée à l'opposition d'un réductionisme à une vision intrinsèque de la complexité. Si une classification systématique réductrice existe pour tout système complexe, alors les niveaux d'émergence supérieurs n'ont pas de signification propre. Il est paradoxal d'observer dans ce cas la position ambiguë de certains travaux de physiciens qui tiennent ce réductionisme comme un dogmatisme mais prétendent s'attaquer à des problèmes complexes typiques des systèmes socio-techniques.

Les mesures de réseau globales comme nous avons vu en cours sont une façon de répondre partiellement à cette question de réduction dimensionnelle. En pratique, on veut être également capable de relier ces mesures à des propriétés pratiques du réseau, qui seront cruciales pour le design et le management de réseaux réels (par exemple réseaux techniques : transport d'électricité, internet ; réseaux de transport ; réseaux sociaux ; réseaux de villes ; etc.) : par exemple le coût, la résilience, la performance. Cet exercice propose de survoler ces deux questions fondamentales pour différents exemples de réseaux réels et synthétiques : illustrer dans un premier temps la signification concrète de différentes valeurs en relation à la topologie "apparente" du réseau ; dans un second temps explorer des liens potentiels entre mesures et propriétés afin de donner une idée d'une manière de caractériser la résilience.

Ce qui est demandé pourra paraître relativement simple mais est central pour une appréhension juste de la complexité des processus géographiques tels qu'ils occurrent dans toute leur réalité. La table 1 donne les valeurs d'indicateurs, pour certains types de réseaux. On fournit leur valeur moyenne et leur écart-type dans le cas de réseaux synthétiques aléatoires pour lesquels les mesures auront alors été estimées sur  $b = 500000$  répétitions des réseaux aléatoires.<sup>1</sup> Pour les réseaux réels ou synthétiques non variables, une valeur seule est donnée, sachant que l'estimation de paramètres moyens dans des situations réelles est directement liée à la stationnarité spatio-temporelle des processus<sup>2</sup>, ce qui est également une question ouverte concernant les systèmes spatiaux complexes. Les figures 1 et 2 illustrent des exemples des réseaux considérés. Les questions suivantes portent sur une interprétation simple des mesures. Parmi les réseaux considérés, on étudie des réseaux routiers réels, dont la localisation spatiale est présentée en figure 3, ainsi que des réseaux synthétiques.

<sup>1</sup>qu'on fixe comme arbitraire. Le problème du nombre de répétitions nécessaires pour une convergence raisonnable des indicateurs statistiques et hors du champ de ce devoir, et d'ailleurs un problème ouvert pour la plupart des modèles de simulation complexes puisque qu'on sait établir des intervalles de confiance soit sous certaines hypothèses théoriques de distribution statistique soit par simulation (*bootstrap*) ce qui ne réduit pas la complexité intrinsèque.

<sup>2</sup>et donc à leur ergodicité, i.e. à l'équivalence entre moyenne spatiale et moyenne temporelle

Les réseaux étudiés sont donc les suivants<sup>3</sup> :

- Réseau routiers réels (simplifiés à une résolution de 100m) : Paris, Ile-de-France, La Courtine (Creuse), Grand Lyon, London Metropolitan Area, Randstad
- Aléatoire (probabilité fixe d'établir un lien entre chaque paire de noeud)
- Attachement préférentiel (type Barabasi-Albert : les liens sont établis itérativement avec une probabilité proportionnelle au degré des noeuds)
- Grille perturbée (grille régulière dont on retire une proportion fixée de liens)
- Arbre (au nombre de feuilles par branche fixe)

Les mesures calculées pour un réseau  $N = (V, E)$  sont :

- Statistiques descriptives : nombre de noeuds  $|V|$  et nombre de liens  $|E|$
- Densité  $\gamma$
- Degré moyen  $\bar{d}$
- Diamètre<sup>4</sup>  $\delta$
- Centralité d'intermédiarité  $b$  : s'agissant d'une mesure locale (associée ici aux liens), on considère sa moyenne  $\langle b \rangle$  et son niveau de hiérarchie<sup>5</sup>  $\alpha[b]$
- Centralité de proximité  $c$  : de même on calcule  $\langle c \rangle$  et  $\alpha[c]$
- Efficacité  $e = \frac{2}{n \cdot (n-1)} \sum_{i < j} \frac{1}{d_{ij}}$  avec  $d_{ij}$  distance topologique entre  $i$  et  $j$
- Coefficient de clustering  $t$ , qui donne la probabilité que les voisins d'un noeud soient connectés
- Modularité  $\mu$  qui donne une mesure plus générale de la structure en communauté du graphe

**Formes Urbaines et indicateurs de réseaux** → Commentez qualitativement les différentes formes des systèmes territoriaux observées. On pourra par exemple se poser la question de la polycentricité d'une Mégarégion Urbaine. Reliez ces observations aux valeurs prises par les indicateurs que vous jugez pertinents.

---

<sup>3</sup>chaque générateur synthétique a des paramètres propres, pour lesquels nous choisissons les valeurs par défaut suivantes : probabilité aléatoire d'Erdos-Renyi  $p = 0.005$  ; proportion de liens de la grille conservés 65% ; attachement préférentiel : nouveau liens  $m = 10$ , exposant  $\alpha = 1$ , exposant de vieillissement  $\beta = -2$ , pas de vieillissement 100 ; nombre de feuille par branche de l'arbre  $f = 3$ .

<sup>4</sup>pour toutes les mesures liées au plus courts chemins, les distances topologiques et non pondérées ont été prises en compte, afin de permettre la comparabilité des réseaux réels et des réseaux synthétiques. Pour une comparabilité des réseaux réels ayant des couvertures géographiques d'étendue significativement différentes, il faut normaliser par le diamètre.

<sup>5</sup>donné par la pente de la regression linéaire d'un fit brutal d'une loi rang-taille :  $\log b_i = \beta + \alpha \cdot \log i$  où les  $b_i$  sont triés par ordre décroissants.

Réseau	$ V $	$ E $	$\gamma$	$\bar{d}$	$\delta$	$\langle b \rangle$
Aléatoire	1000	$2498 \pm 50$	$0.005 \pm 1 \cdot 10^{-4}$	$5 \pm 0.1$	$9 \pm 0.59$	$0.0018 \pm 5.4 \cdot 10^{-5}$
Att. Préf.	1000	$4579 \pm 21$	$0.0092 \pm 4.3 \cdot 10^{-5}$	$9.2 \pm 0.043$	$7 \pm 0.18$	$0.00084 \pm 8.6 \cdot 10^{-6}$
Grille	$499 \pm 3.7$	624	$0.005 \pm 7.5 \cdot 10^{-5}$	$2.5 \pm 0.019$	$53 \pm 4.5$	$0.03 \pm 0.0027$
Arbre	1000	999	0.002	2	12	0.0099
Ile-de-France	$ V $	$ E $	$\gamma$	$\bar{d}$	$\delta$	$\langle b \rangle$
Paris	$ V $	$ E $	$\gamma$	$\bar{d}$	$\delta$	$\langle b \rangle$
Grand Lyon	$ V $	$ E $	$\gamma$	$\bar{d}$	$\delta$	$\langle b \rangle$
La Courtine	$ V $	$ E $	$\gamma$	$\bar{d}$	$\delta$	$\langle b \rangle$
Randstad	$ V $	$ E $	$\gamma$	$\bar{d}$	$\delta$	$\langle b \rangle$
London	$ V $	$ E $	$\gamma$	$\bar{d}$	$\delta$	$\langle b \rangle$
Réseau	$\alpha [b]$	$\langle c \rangle$	$\alpha [c]$	$e$	$t$	$\mu$
Aléatoire	$-0.29 \pm 0.0092$	$0.46 \pm 0.68$	$-0.076 \pm 0.032$	$0.24 \pm 0.003$	$0.005 \pm 0.0011$	$0.45 \pm 0.0072$
Att. Préf.	$-0.63 \pm 0.015$	$0.26 \pm 0.0021$	$-0.062 \pm 0.0022$	$0.28 \pm 0.0018$	$0.079 \pm 0.0028$	$0.66 \pm 0.0082$
Grille	$-1.2 \pm 0.077$	$7.3 \pm 3.5$	$-0.99 \pm 0.31$	$0.072 \pm 0.0032$	0	$0.87 \pm 0.0058$
Arbre	-1	0.1	-0.094	0.11	0	0.93

Table 1: Valeurs de mesures de réseaux pour différents exemples typiques réels et synthétiques

Réseau	$ V $	$ E $	$\gamma$	$\bar{d}$	$\delta$	$\langle b \rangle$	$\alpha [b]$	$\langle c \rangle$	$\alpha [c]$	$e$	$t$	$\mu$
Aléatoire	$ V $	$ E $	$\gamma$	$\bar{d}$	$\delta$	$\langle b \rangle$	$\alpha [b]$	$\langle c \rangle$	$\alpha [c]$	$e$	$t$	$\mu$
Attachement Préférentiel	$ V $	$ E $	$\gamma$	$\bar{d}$	$\delta$	$\langle b \rangle$	$\alpha [b]$	$\langle c \rangle$	$\alpha [c]$	$e$	$t$	$\mu$
Grille Perturbée	$ V $	$ E $	$\gamma$	$\bar{d}$	$\delta$	$\langle b \rangle$	$\alpha [b]$	$\langle c \rangle$	$\alpha [c]$	$e$	$t$	$\mu$

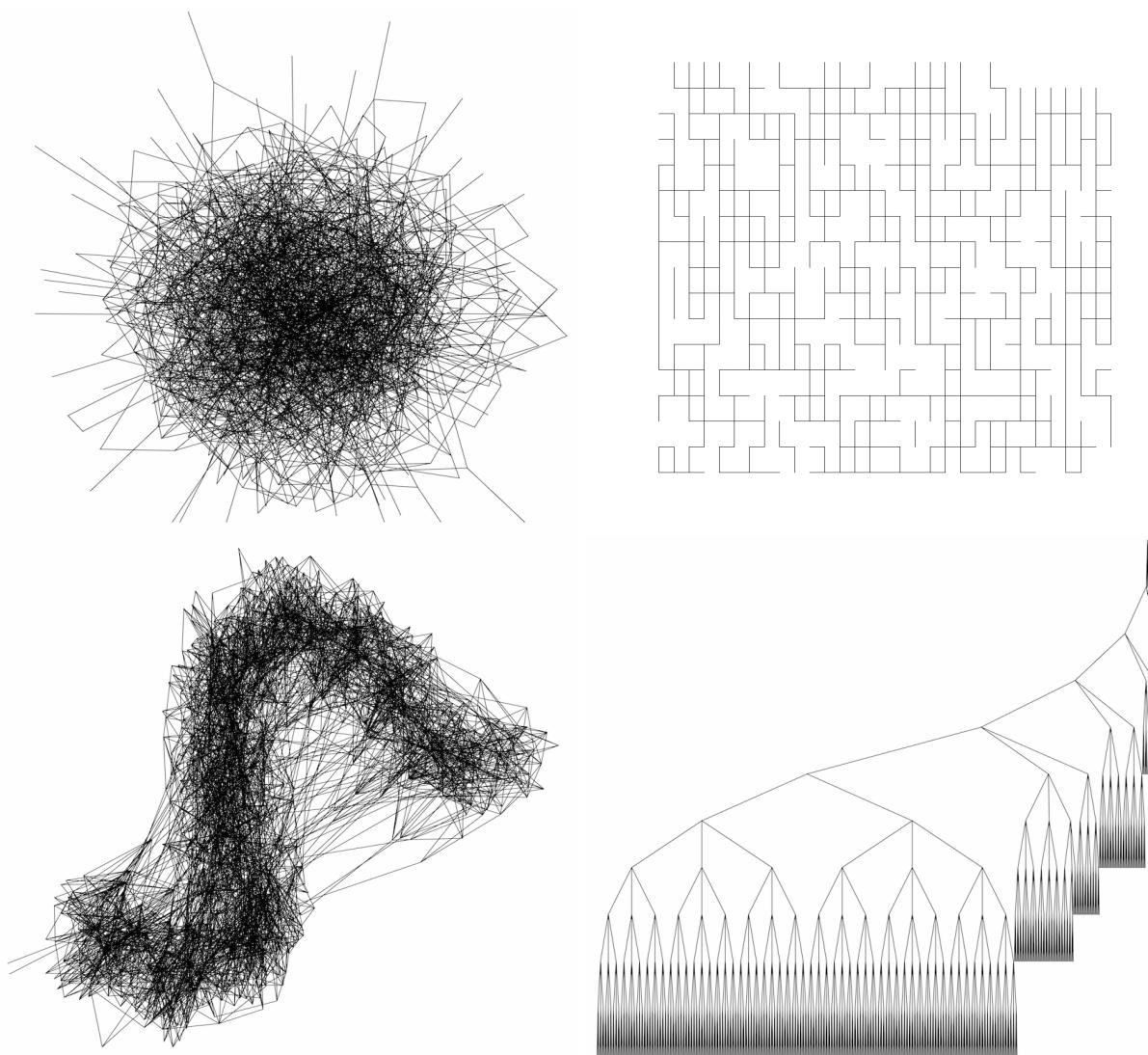
Table 2: Correlations estimées

**Interprétation des Indicateurs** → Pour les réseaux synthétiques, commentez les valeurs prises par les indicateurs. Lesquelles étaient intuitivement attendues ? Dans quelle mesure est-on capable de caractériser et discriminer chaque type de réseau. De quel réseau synthétique s'attend-on à ce que les réseaux réels soient les plus proches ? Peut-on le confirmer par les indicateurs ?

Nous proposons à présent de tenter une caractérisation de la résilience des réseau. La définition utilisée prend en compte la capacité du réseau à rester performant face à la rupture de lien, comme proposé par [Ash and Newth, 2007]. On considère la rupture aléatoire d'une proportion fixée  $\alpha$  de liens, et on note  $N_\alpha$  le réseau résultant de la suppression à partir du réseau  $N$ . L'indicateur de résilience au niveau  $\alpha$  est alors défini par  $r = \mathbb{E}[\Delta_\alpha e] = \mathbb{E}[e(N_\alpha) - e(N)]$ . On peut définir de même les variations des autres indicateurs. La structure de covariance des différentes variations, et en particulier la correlation avec la résilience, est un moyen indirect de caractériser la résilience, au sens de quel type de propriété permet au réseau d'augmenter sa résilience. La table 2 donne pour chaque réseau aléatoire les correlations estimées  $\rho[X] = \hat{\rho}[\Delta_\alpha e, \Delta_\alpha X]$  où l'estimateur  $\hat{\rho}$  est calculé en pratique sur une plage de valeurs pour  $\alpha$  (de 0.5 à 0.95 par 0.05) et sur un nombre de répétitions fixé ( $b =$ , en faisant l'hypothèse que répéter sur le réseau est équivalent à répéter sur la suppression des liens).

**Caractérisation de la résilience** → Commentez à partir de la table des corrélations, pour les différents types de réseau, les facteurs influençant la résilience. Quel enseignement peut-on en tirer pour la conception de réseaux techniques par exemple ?

**Résilience Dynamique et Topologie (Question bonus)** Les approches à la notion de résilience sont diverses et complémentaires. L'aspect dynamique, au sens par exemple du temps nécessaire au système



**Document 1:** Représentation d'instances des exemples synthétiques de réseaux. Dans l'ordre de haut en bas et de droite à gauche : réseau aléatoire, grille perturbée, attachement préférentiel, arbre.

pour retrouver son état initial après perturbation, est particulièrement intéressant pour les systèmes urbains. Dans le cas de réseaux où les noeuds ont leur dynamique propre, les solutions pour une définition robuste et universelle sont très récentes, comme celle proposée par [Gao et al., 2016].

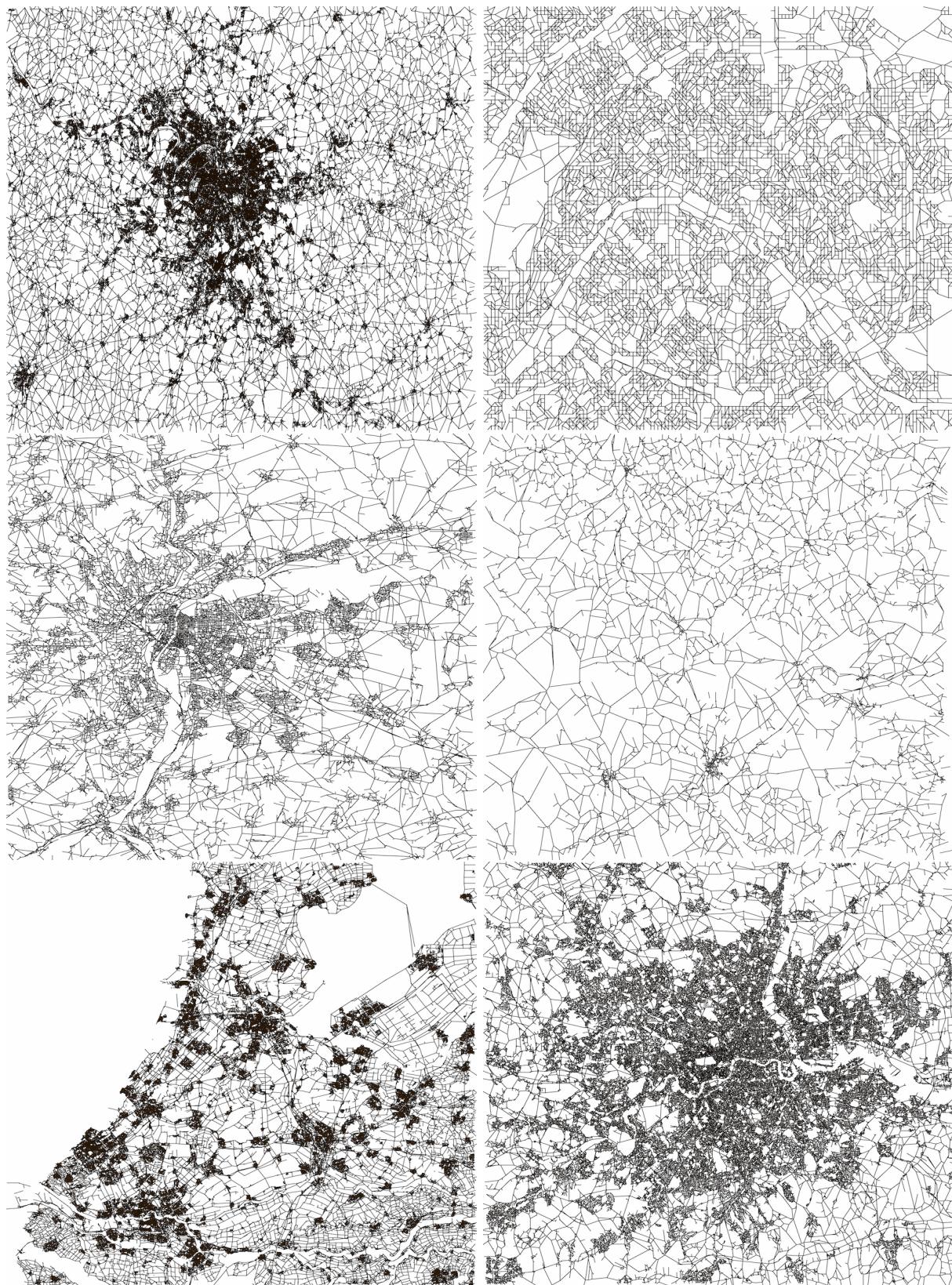
→ Vous semble-t-il simple, dans le cas de dynamiques couplées au sein d'un réseau, d'isoler la contribution de la topologie du réseau de celle des dynamiques propres à la dynamique générale ? Dans quelle mesure serait-il alors complexe de quantifier la résilience dynamique dans une dimension réduite ?

★      ★

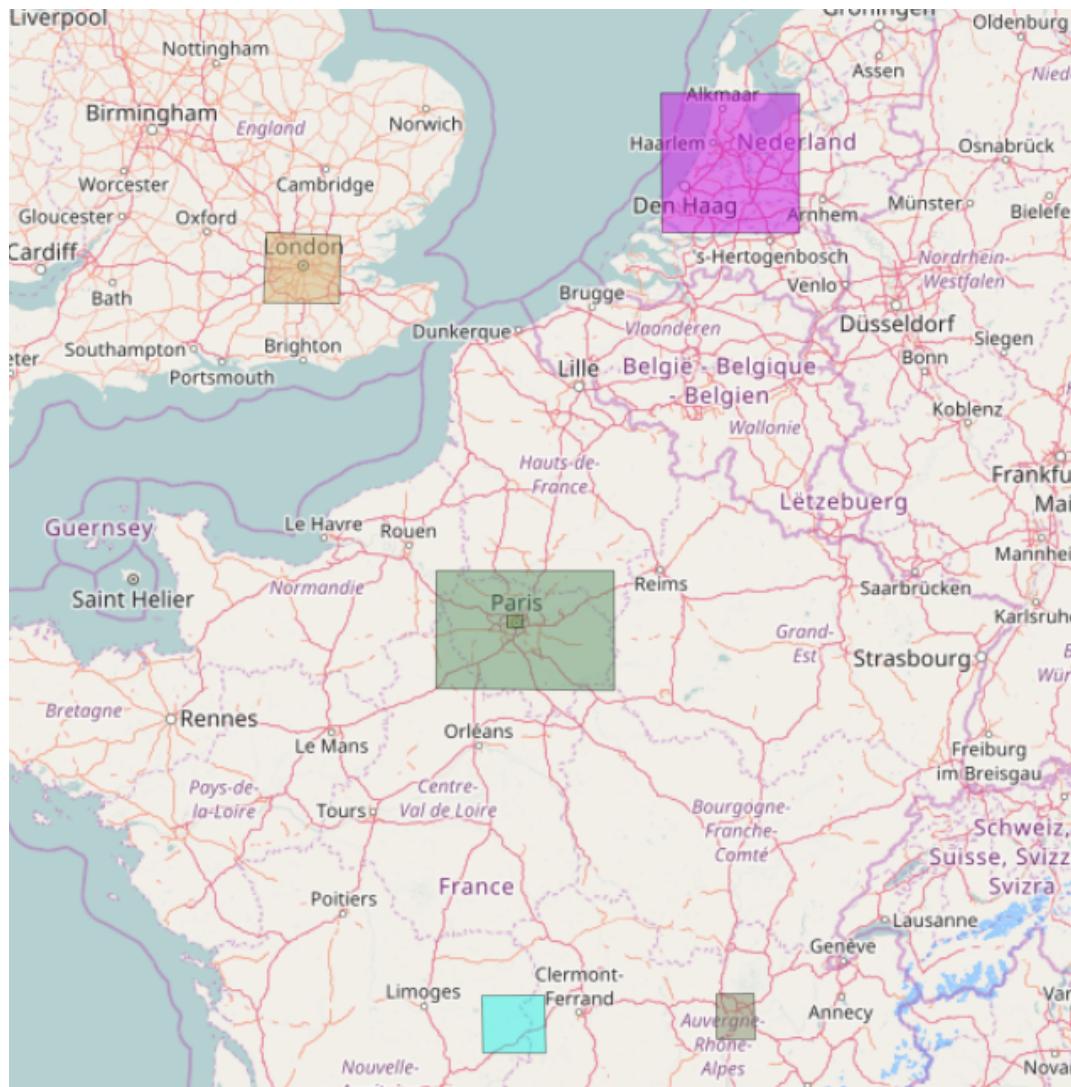
★

## References

- [Ash and Newth, 2007] Ash, J. and Newth, D. (2007). Optimizing complex networks for resilience against cascading failure. *Physica A: Statistical Mechanics and its Applications*, 380:673–683.



**Document 2:** Réseaux routiers réels étudiés. Dans l'ordre de haut en bas et de droite à gauche : Ile-de-France, Paris, Grand Lyon, La Courtine, Randstad, London Metropolitan Area



**Document 3:** Localisation Géographique des réseaux réels étudiés.

[Gao et al., 2016] Gao, J., Barzel, B., and Barabási, A.-L. (2016). Universal resilience patterns in complex networks. *Nature*, 530(7590):307–312.