

# M(A,A) en considérant la population

Julie Gravier

14 juin 2022

## 1 Introduction

J'ai initialement envisagé de considérer des disques de rayon  $r$  à partir des points  $P_i$ . Puis d'intersecter ces différents disques selon 1) la répartition de l'activité à chaque date; 2) la taille des disques selon  $r$ . L'objectif, après intersection, était d'établir des estimations de populations pour chaque polygone minimal (résultat de l'intersection) et les quartiers parisiens (données de population connues) au fil du temps. Dès lors, chaque polygone minimal aurait pu être envisagé comme contenant une population estimée en tant qu'un ensemble de clients potentiels pour chaque  $P_i$ , à condition – évidemment – de diviser la population estimée de chaque polygone minimal par le nombre de  $P_i$  auxquels co-appartient tel ou tel polygone minimal. Hélas, la gestion des intersections de disques constitués à partir de points est complexe dès lors que l'on a en entrée des points trop proches dans l'espace (globalement on obtient un ensemble de données spatiales constitué en *Geometry Collections*; voir par exemple, la discussion de cette issue Github). En bref, la gestion des intersections pour ce type de configurations spatiales n'est pas fonctionnelle dans la library GEOS.

La solution pour pallier ce problème a été d'établir en amont une grille et d'estimer les populations dans des disques de rayon  $r$  sans considérer la concurrence entre  $P_i$  d'une même activité. Plus précisément, les étapes mises en place sont :

- construction d'une grille hexagonale pour tout Paris (dans sa taille post-1860), choix initial de 50 mètres de côté (variation de la taille et de la forme de la grille évidemment possible);
- pour chaque date où l'on connaît la population, interpolation spatiale de la population des quartiers dans cette grille hexagonale. Dit autrement, la population est inférée pour chaque élément de la grille  $\in$  Paris, selon une pondération en fonction de la surface de l'hexagone  $\in$  dans les quartiers (à partir des algo de [Prener and Revord, 2019]);
- construction des disques de rayon  $r$  à partir des point  $P_i$ , i.e. un équivalent tout simple de buffer,  $r \leftarrow 100\dots i\dots 2,000$  mètres, tous les 50 m.;
- interpolation spatiale pour chaque disque de  $P_i$  de la population à partir de celle estimée dans la grille hexagonale.

Soit, au final :

$$M_{pi} = \frac{n_i/pop_i}{N_i/pop_{tot}} \quad (1)$$

Avec  $n_i$  le nombre d'activités dans le disque de rayon  $r$  de  $P_i$  et  $pop_i$  la population estimée dans ce disque ;  $N_i$  étant le nombre d'activités totale et  $pop_{tot}$  la population totale. Le  $M(A,A)$  en considération la population est donc :

$$M(A,A)_{pi} = \frac{1}{N_i} \sum_{i=1}^{N_i} M_{pi} \quad (2)$$

## 2 $M(A,A)$ de Marcon et Puech et $M(A,A)$ en considérant la population

Construction d'un espace théorique de 100 sur 100. Si la population est répartie de manière uniforme et que l'on construit une grille hexagonale pour laquelle on interpole la population, alors les effets de bords sont (logiquement) notables – comme on le note dans figure 1.

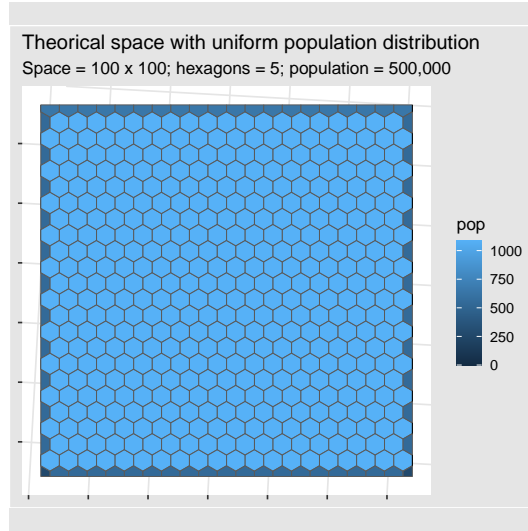


FIGURE 1 – Population uniforme dans une maille hexagonale

Le second cas pour l'instant testé est celui d'une population distribuée de manière aléatoire dans la grille hexagonale.

Un ensemble de 100 points, répartis de quatre manières a été considéré systématiquement :

- de façon aléatoire,
- en deux clusters de taille différentes (deux disques avec répartition aléatoire à l'intérieur de chacun d'eux),

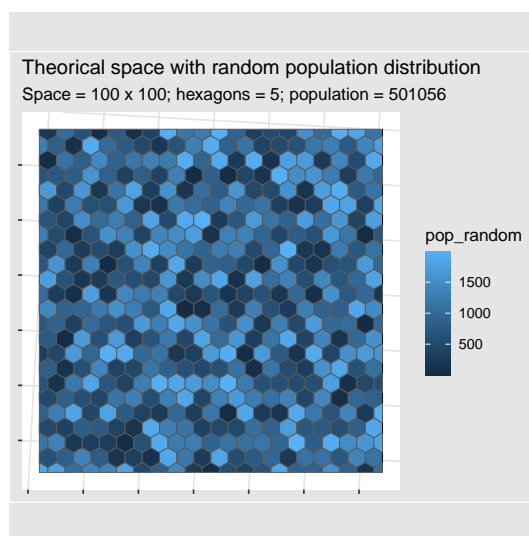


FIGURE 2 – Population distribuée aléatoirement

- équidistants les uns des autres,
- sous une forme plus complexe – pour reprendre les mots de Marcon et Puech – à savoir en cluster et aléatoirement.

Pour chacune des distributions de points, la comparaison entre le  $M(A,A)$  de Marcon et Puech – correspondant aux distributions de la fonction  $M$  de droite des figures suivantes – et le  $M(A,A)$  en considérant la population – graphiques de gauches – a été faite. De manière générale, quelle que soit la distribution des points, on identifie 1) que les ordres de grandeurs de  $M$  sont identiques et 2) que le  $M(A,A)$  en considérant la population a tendance à varier plus fortement, en particulier dans le cas de points initialement répartis de manière aléatoire ou de façon plus complexe. Cette variation de  $M$  résulte du découpage initial de l'espace dans une maille, d'autant que la variation du rayon  $r$  considéré<sup>1</sup> (de 1 à chaque fois) est plus petite que la taille de chacune des maille (chaque hexagone mesure 5 de côté).

J'ai par ailleurs effectué un premier test sur les relieurs. On identifie en particulier que l'activité apparaît bien moins concentrée si l'on considère la population et non uniquement l'espace, bien que  $M$  soit largement supérieur à 1 dans les rayons les plus petits, surtout pour la première moitié du 19e s.

---

1. Notons que je me suis trompée dans les légendes des graphiques suivants. Il ne s'agit évidemment pas du "radius on network" mais bien d'un "radius" tout court en  $x$ .

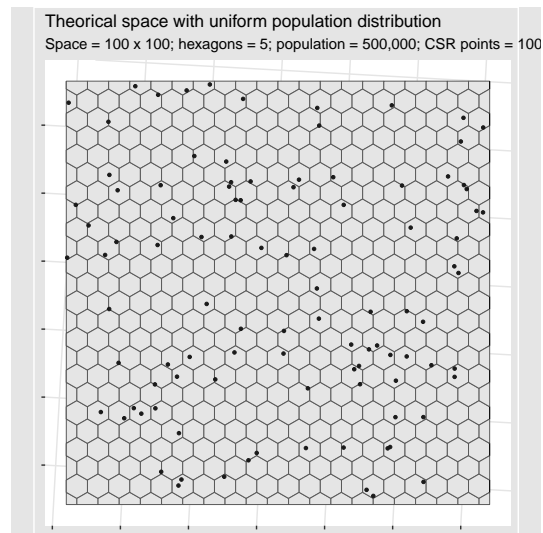


FIGURE 3 – Points aléatoires

## Références

[Prener and Revord, 2019] Prener, C. G. and Revord, C. K. (2019). areal : An R package for areal weighted interpolation. *Journal of Open Source Software*, 4(37).

CSR points on uniform population

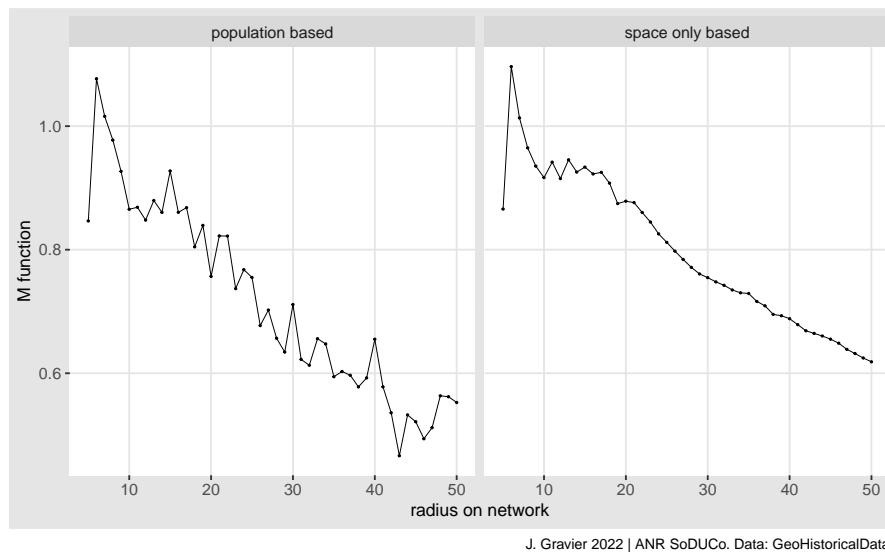


FIGURE 4 – Points aléatoires et population uniforme

CSR points on random population

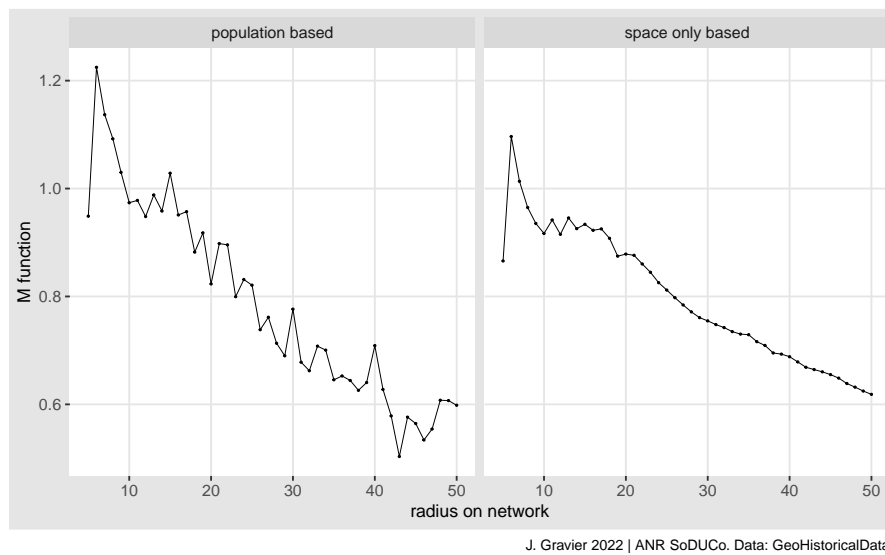


FIGURE 5 – Points et population aléatoires

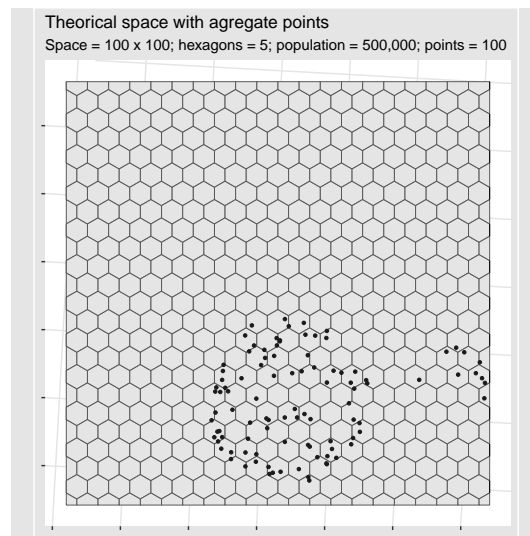


FIGURE 6 – Points clusterisés

Agregate points on uniform population

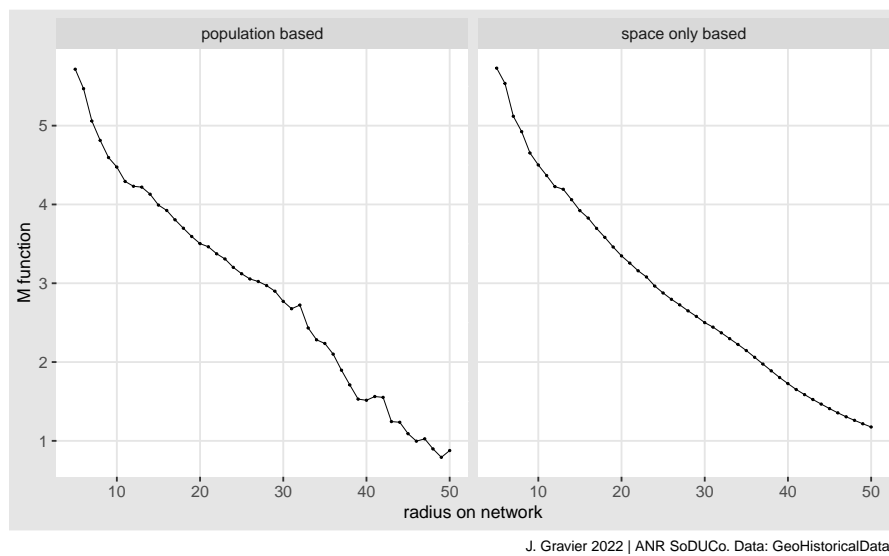


FIGURE 7 – Points clusterisés et population uniforme

Agregate points on random population

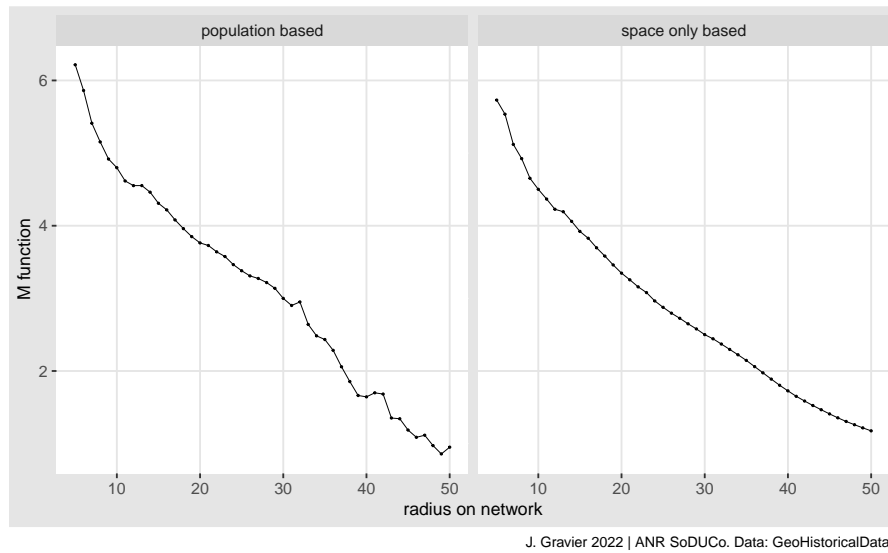


FIGURE 8 – Points clusterisés et population aléatoire

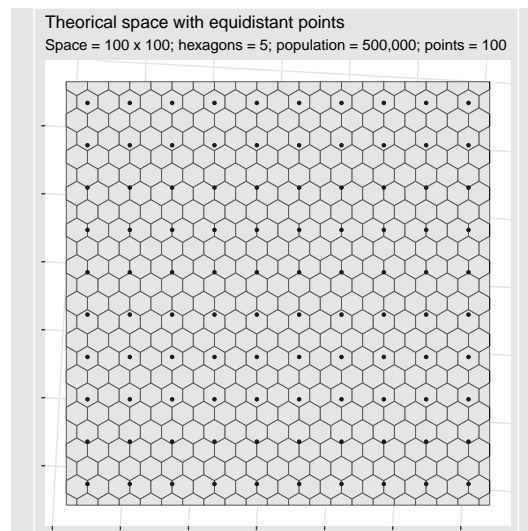


FIGURE 9 – Points équidistants

Equidistant points on uniform population

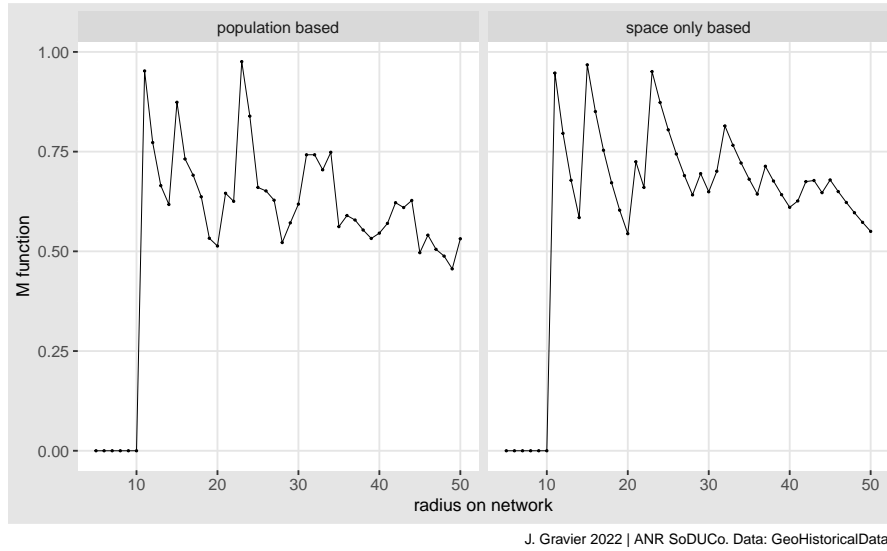


FIGURE 10 – Points équidistants et population uniforme

Equidistant points on random population

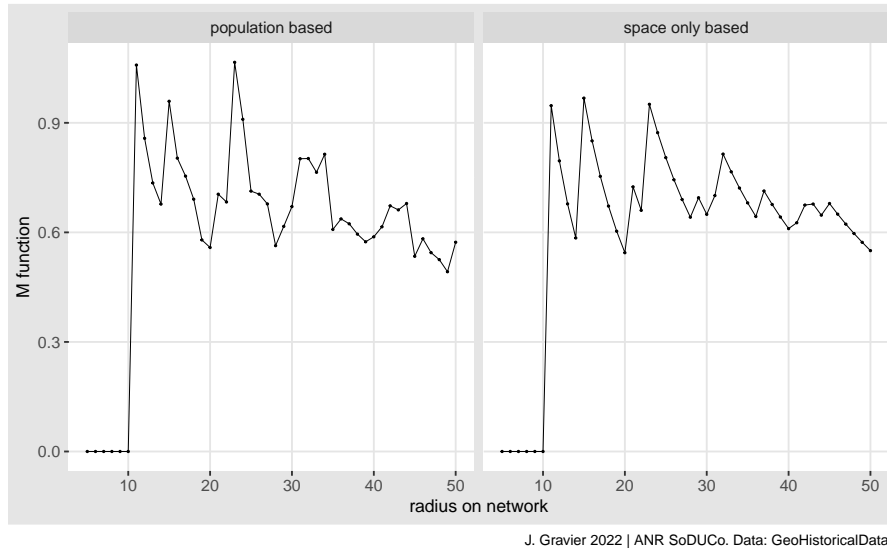


FIGURE 11 – Points équidistants et population aléatoire



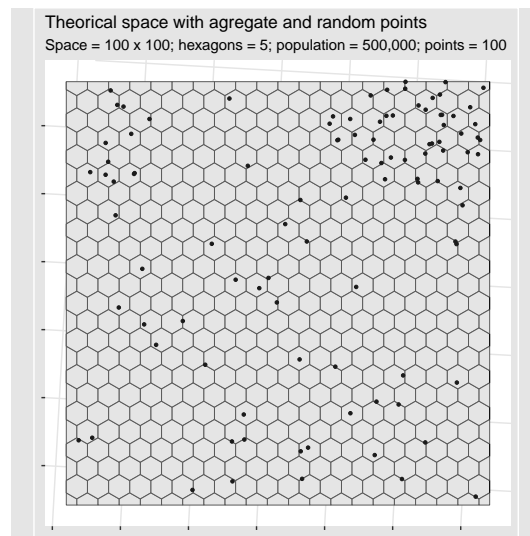


FIGURE 12 – Points ayant une distribution spatiale plus complexe

Agregate points on random population

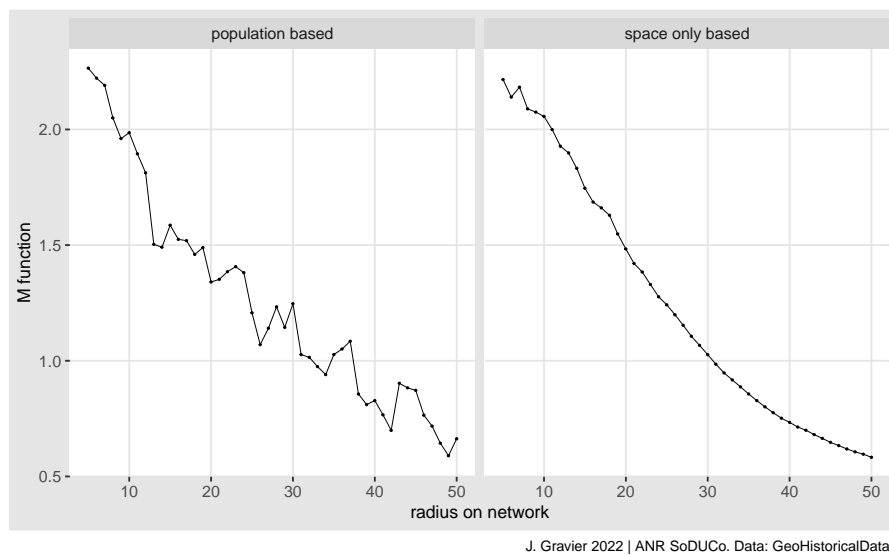


FIGURE 13 – Avec une population uniforme

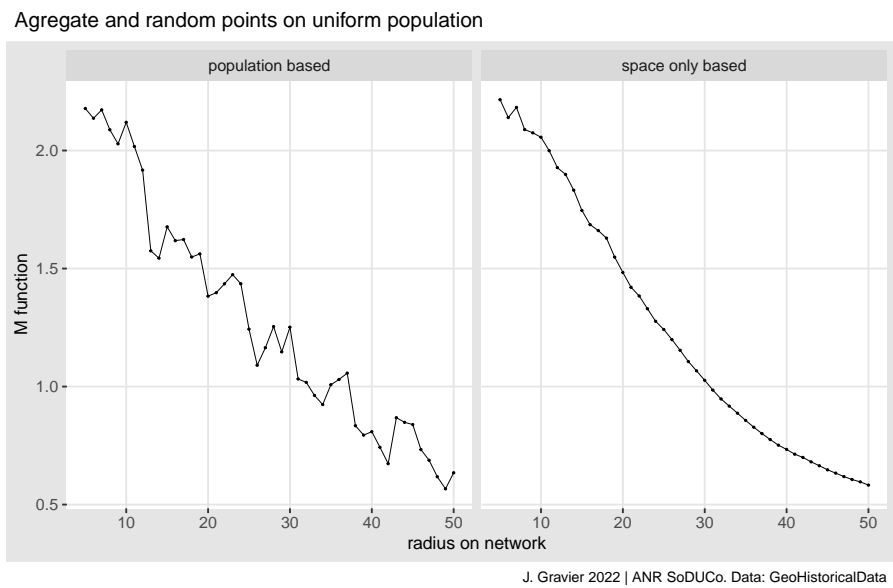


FIGURE 14 – Avec une population aléatoire

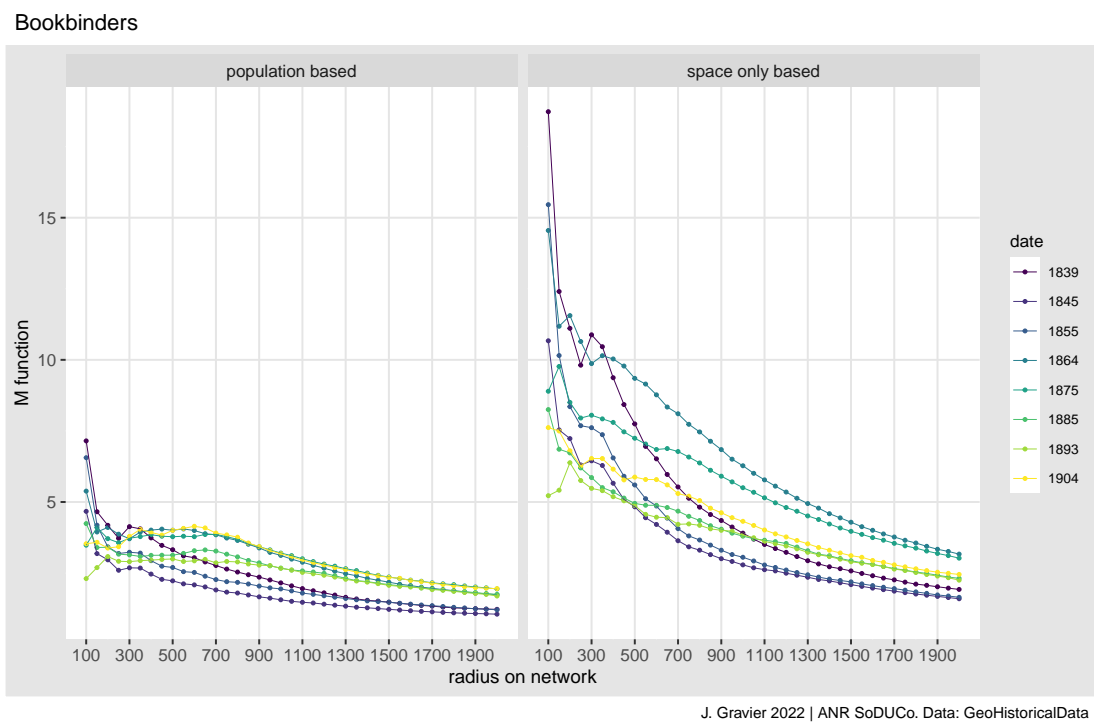


FIGURE 15 – Comparaison des  $M(A,A)$  des relieurs