

## **Annexe A. 2. La construction d'une matrice de voisinage spatial (*k* plus proches voisins)**

Cette annexe porte sur la construction d'une matrice de voisinage devant permettre de réduire une matrice de flux, dans l'objectif de les cartographier en fonction d'un critère d'éloignement discret. Elle intéresse le chapitre 7.

Cette matrice de voisinage concerne le *plus proche voisin*, elle est élaborée à partir d'un fond de carte ponctuel et non surfacique comme dans le chapitre 7. L'objectif est de montrer la variation du *graphe de voisinage* obtenu en fonction du type d'implantation spatiale considéré : ponctuel ou linéaire (voir section 7.2 du Chapitre 7).

La mesure du voisinage selon le *plus proche voisin* associe donc ici la localisation (x,y) des barycentres des unités spatiales, avec une mesure de proximité géographique continue.

### **A.2.1 Présentation de l'approche**

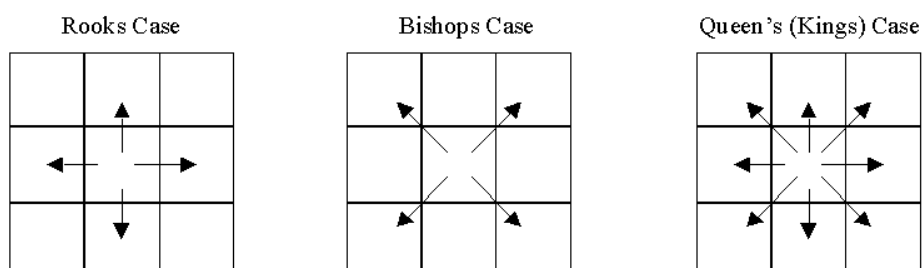
L'approche mobilise une *matrice des k-contiguïtés*, parfois qualifiée dans la littérature de matrice de *poids spatiaux*. Cette matrice, généralement utilisée dans les analyses de régression spatiale, sert à déterminer l'importance et à la portée spatiale d'une mesure d'auto corrélation spatiale des données, c'est-à-dire de l'extension dans l'espace du degré de ressemblance d'une unité spatiale donnée à ses voisins. On définit généralement à partir de cette *matrice des k-contiguïtés*, une *matrice de distances* pouvant être discrète ou continue (par exemple de  $1/D_{ij}$  afin de considérer la décroissance de l'effet de la distance), ou une *matrice de pondération spatiale* ( $W_{ij}$ ), qui définit la part en pourcentage du poids du lien dans la somme marginale en ligne :  $W_{ij}=W_{ij}/O_i$ , ou encore le total général. Cette matrice servira à tester l'auto corrélation spatiale.

Fondée sur la notion de *voisinage*, la mesure *des k-contiguïtés* est donc composée de deux étapes. Elle nécessite de définir le critère de voisinage à partir duquel deux unités spatiales seront considérées comme voisines ou proches. Le résultat est stocké dans une matrice de voisinage qui prend la forme d'une matrice d'adjacence : les couples de voisins prennent la valeur '1' sinon 0. La seconde étape vise à définir le poids attribué à chacun des liens de cette matrice de voisinage, compte tenu de leur proximité au regard de la thématique du flux, correspondant à la *matrice de poids spatiaux*.

Seule la première étape est nous intéresse ici dans cette approche cartographique de sélection de valeurs de flux intégrant le voisinage des lieux.

Le *voisinage* d'un lieu peut en effet être défini de deux manières : en termes de *contiguïté* des zones selon qu'elles partagent une frontière commune ( $C_{ij}=1$ ) ou un point. Dans ce cas, elle tient également compte de la distance à un nombre *k* de *plus proche voisins*. Le premier cas correspond ainsi, sous certaines conditions, à la matrice ( $C_{ij}$ ) dont il a été question dans la section 7.2.1. Il s'agit de l'exemple de droite représenté sur la Figure A.2-30., qui illustre le déplacement de la Reine sur un échiquier.

**Figure A.2-30. Méthodes d'estimation du voisinage spatial fondé sur la contiguïté des zones**



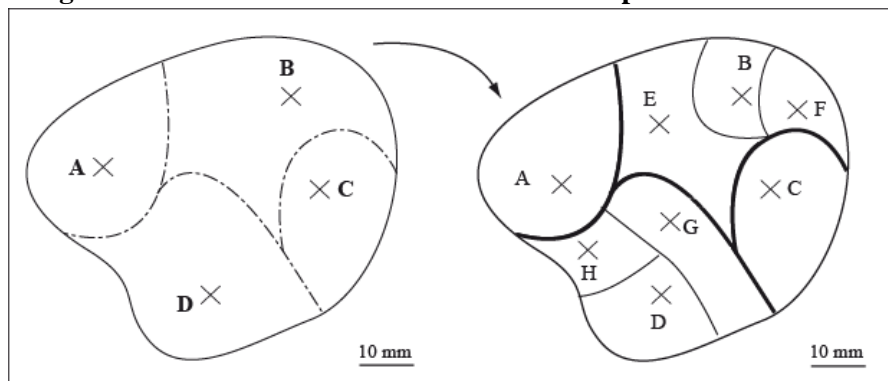
La Figure A.2-30 présente les trois possibilités théoriques de sélection du plus proche voisin (j), en fonction de l'orientation de (i). Elle est souvent envisagée, dans la littérature, en référence aux déplacements de la Tour, du Fou et de la Reine sur un échiquier.

La différence par rapport à l'approche présentée dans la section 7.2.1 réside dans le fait que la recherche de *plus proche(s) voisin(s)* considère une implantation ponctuelle de la zone d'étude — la matrice des points de coordonnées (x, y) — et non surfacique comme dans le cas de (Cij=1) et une mesure d'éloignement continu. Le graphe de voisinage spatial correspond, en effet, à une matrice où la valeur d'une cellule (i,j) est soit celle d'un nombre *k* de (j) voisins géographiques considérés dans un périmètre donné au regard d'une mesure de distance (continue, généralement (Dij) ou discrète, correspondant à une distance géodésique, par exemple (CKij) telle que présentée dans la section 7.2.2) ; le choix de la mesure d'éloignement étant spécifique à la thématique. On peut, par exemple, recourir à une distance sociale ou géographique pour déterminer le nombre (*k*) de voisins, en considérant un ordre où une portée de distance : par exemple le nombre (*k*) de voisins situés dans un rayonnage de  $\alpha$  km ou inférieur au premier quartile ( $D_{ij} < Q_1$ ).

Ainsi, de la même façon que la matrice des *k* voisins avec ( $k=1$ ) ne correspondait à (Cij=1) (1 frontière) que dans le cas de la Reine (voir Figure A.2-30), on observe des différences importantes lorsque l'on considère des portées discrètes plus importantes ( $k=1, 2, 3, \dots, n$ ). Pour illustrer ces différences, nous rappelons la formation de notre matrice théorique d'exemple local à huit lieux.

La Figure A.2-31 illustre la construction de ces deux matrices dont l'ordre de voisinage est finalement d'ordre 1, la valeur de *k* variant fondamentalement entre elles deux.

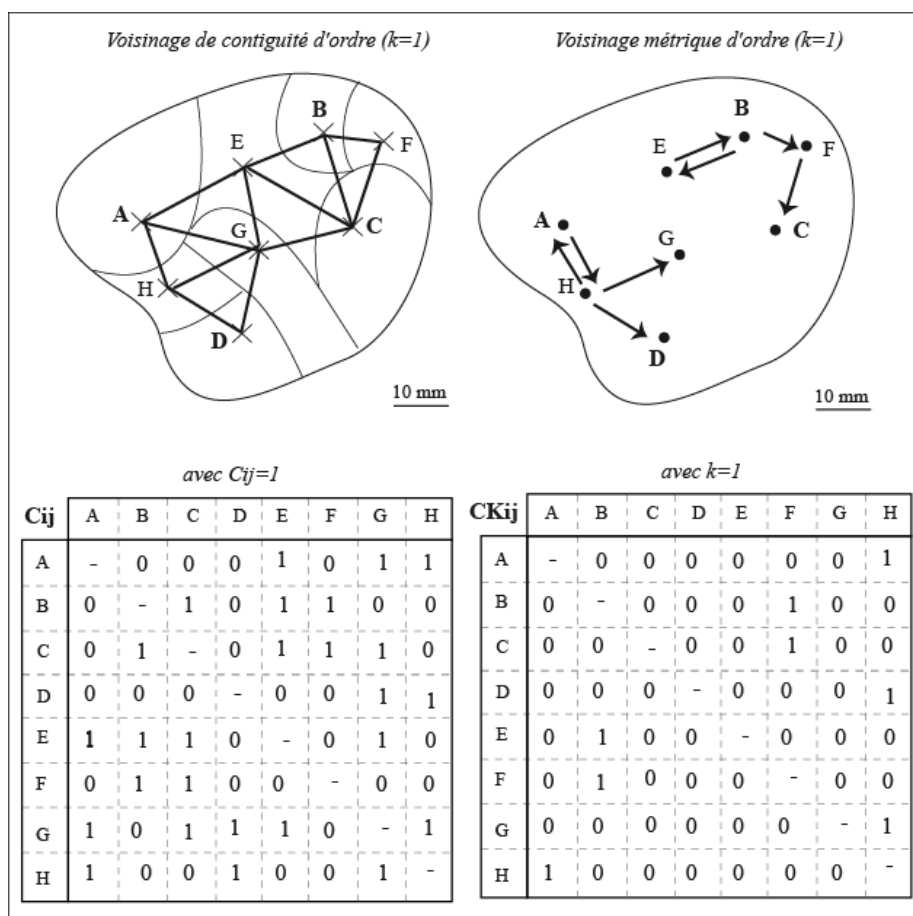
**Figure A.2-31. Présentation de la matrice exemple local à huit lieux**



Pour bien comprendre les différences entre les matrices ( $C_{ij} \neq CK_{ij}^{k=1}$ ) qui sont finalement toutes deux d'ordre 1 ( $k=1$ ), nous allons supposer que les 4 points de notre matrice d'exemple local correspondent, en réalité, aux chefs-lieux d'un découpage administratif territorial de niveau régional que nous mobiliserons dans la prochaine section. Cela nous permet dans l'immédiat de considérer une zone d'étude au niveau local qui compte huit lieux. C'est cette modification de l'exemple local qui est présentée sur la Figure 7-12.

Sur la Figure A.2-32, l'image de gauche correspond à la matrice (Cij) alors que celle de droite à (CKij) avec ( $k=1$ ), présentée dans le chapitre 7. La première s'appuie sur une mesure binaire (l'existence d'une frontière) mobilisant ainsi un maillage spatial discret ; la seconde requiert une mesure de distance, que nous avons supposé être exprimée en (kilo)mètres (mais plusieurs mesures sont envisageables), pour identifier l'ordre de voisinage considéré. Compte tenu de ces différences, nous proposons d'utiliser les expressions respectives de *voisinage de contiguïté* et de *voisinage métrique* (dans les deux cas : *d'ordre k=1*) pour distinguer littéralement ces deux matrices.

**Figure A.2-32. Illustration du voisinage de contiguïté versus voisinage métrique. Comparaison des méthodes d'estimation des  $k=1$  (plus proches) voisins**



Trois différences majeures sont notables sur le graphe de voisinage issu de ces deux matrices (voir Figure A.2-32), lesquelles expliquent par ailleurs la forte dissemblance de leurs motifs spatiaux et de leur sémiologie graphique. La première tient au nombre de liens considérés pour chacun des lieux ; la seconde à leur caractère symétrique ou non par rapport à la diagonale principale et la troisième, à l'élimination de certains lieux de l'analyse.

La matrice ( $CK_{ij}$ ) avec ( $k=1$ ), située sur la partie droite de la Figure A.2-32 ne retient qu'un seul voisin par lieu : le plus proche en distance ; il s'agit du *premier plus proche voisin* qui n'est représenté que par un lien, pour chacun des lieux. A l'inverse, pour la matrice ( $C_{ij}$ ), un même lieu peut avoir plusieurs voisins avec lesquels il partage une frontière, tout dépend de la morphologie des zones d'étude, *a fortiori* du type de maillage utilisé pour l'analyse : l'utilisation d'un maillage régulier ou théorique, tel un pavage de Voronoï, ne conduira pas au même résultat que l'utilisation d'un maillage hétérogène.

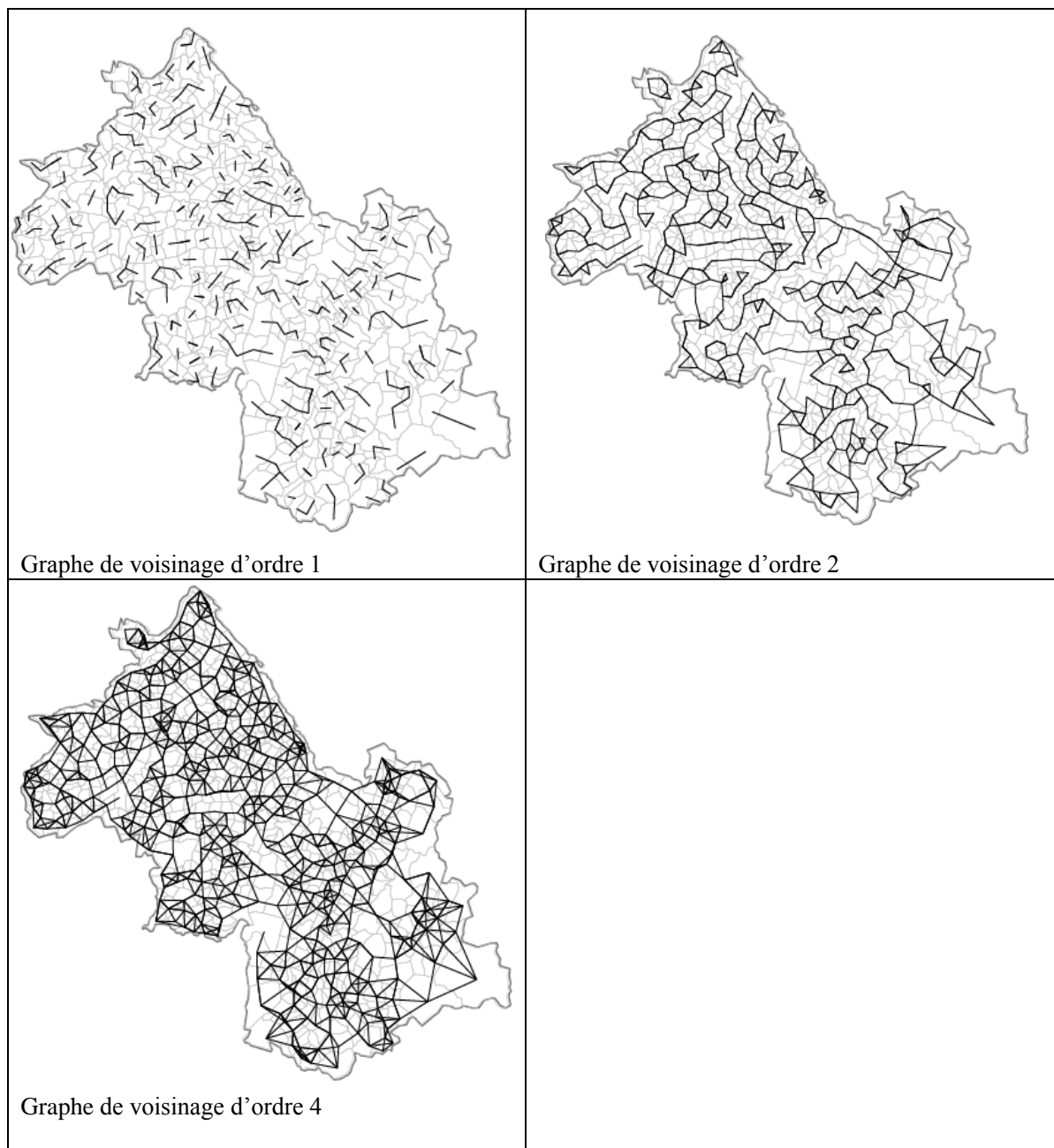
La deuxième différence et non la moindre entre ces deux mesures de proximité de voisinage d'ordre 1, réside dans la non symétrie de la matrice ( $CK_{ij}$ ), la matrice ( $C_{ij}$ ) étant quant à elle symétrique. On observe, en effet, sur la Figure A.2-32, que le plus proche voisin de (E) est (B) tandis que le plus proche voisin de (B) est (F). De la même façon, certains lieux, tels que (C) ou (G) par exemple, ne sont le plus proche voisin de personne, même s'il a nécessairement un voisin le plus proche. Cela s'explique par le caractère asymétrique de cette matrice.

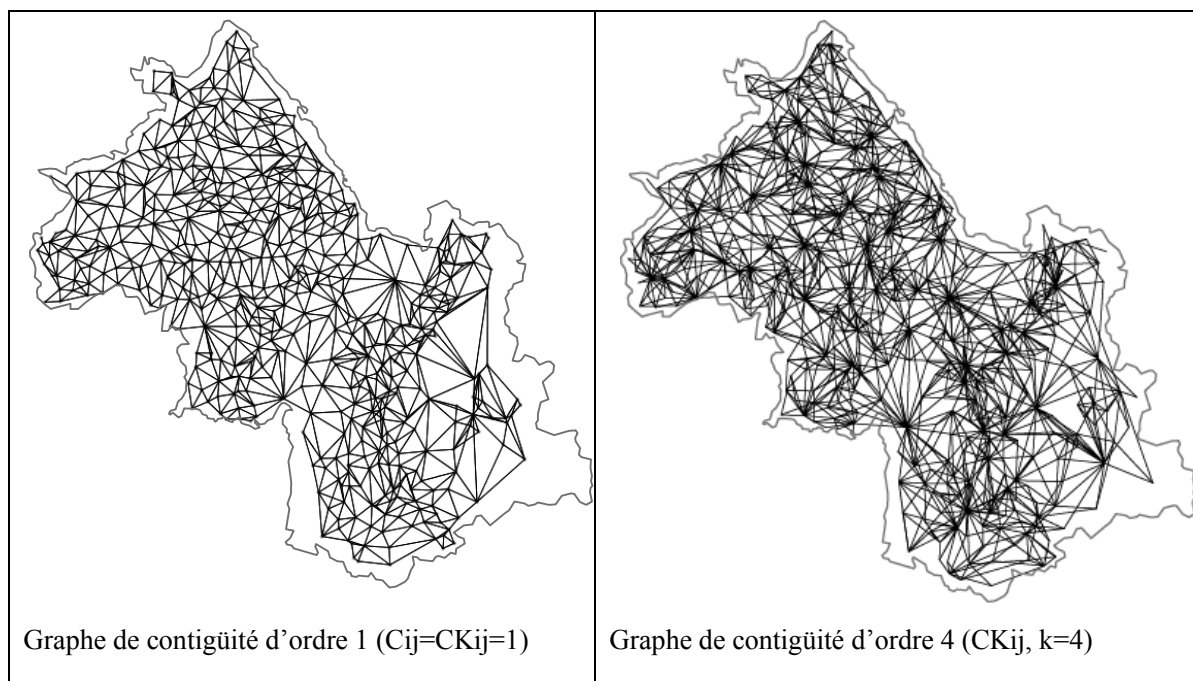
### A.2.1 Comparaison des différents graphes de voisinage de l'Isère

La Figure A.2-33 applique la procédure de la Figure A.2-3 au fond de carte des communes de l'Isère, pour  $(k=1)$ ,  $(k=2)$  et  $(k=4)$ , elle présente les graphes de voisinage d'ordres 1, 2 et 4.

La matrice  $(CK_{ij})$  avec  $(k=1)$  étant *asymétrique*, sa représentation graphique nécessite le recours à la flèche afin d'identifier sans confusion le lieu depuis lequel le plus proche voisin est considéré. C'est la raison pour laquelle la sémiologie des motifs de la spatialisation de ces deux *graphes de voisinage* diffère.

**Figure A.2-33. Graphes de voisinage et de contiguités d'ordres  $k=(1)$ ,  $k=(2)$  et  $k=(4)$  du fond de carte de l'Isère**





D'autres différences peuvent être mentionnées, parmi lesquelles le choix de la direction en degrés de recherche du plus proche voisin ; celui d'un maillage théorique<sup>243</sup> permettant d'identifier le (ou les) voisin.s concerné.s ou la métrique utilisée. Dans le cas de ( $C_{ij}$ ), nous avons vu que la métrique est binaire et mobilisée seule, alors que dans le cas de ( $CK_{ij}$ ) avec ( $k=1$ ), la mesure est binaire, ordinale et toujours associée à une métrique d'éloignement ( $CK_{ij}, D_{ij}$ ).

L'utilisation de la matrice ( $CK_{ij}$ ) ouvre aussi la voie vers la construction de matrices des deux plus proches voisins jusqu'à ( $k=n$ ) plus proches voisins.

Le programme pour Rstudio© permettant la construction des matrices présentées est disponible en annexe, ainsi que sur le site du groupe de travail ElementR.

<sup>243</sup> Les travaux menés en écologie, sur des implantations ponctuelles, mobilisent des maillages et différentes conditions. Par exemple : deux points seront considérés comme voisins s'ils partagent une frontière dans le pavage de Delaunay et si aucun autre point ne se trouve à l'intérieur du diamètre qu'ils définissent (Gabriel et Sokal, 1969), cités par <http://pbil.univ-lyon1.fr/R/pdf/ter4.pdf> (page 5)