



Rédactrice/Rédacteur: Monika Ferster, Daniel Assoulin, METH, BFS, juillet 2019

Note méthodologique

Statistique de la superficie: document de référence pour les formules

1 Formalisation: notations et exemples

Si l'on quadrille la surface du territoire suisse par des lignes distantes de 100 m les unes des autres, on obtient env. 4.1 millions de points d'intersection (points d'intersection des coordonnées de la carte nationale suisse). Nous identifions ces points par leurs coordonnées (i, j) , avec $(i, j) \in K$, K étant l'ensemble des coordonnées des points d'intersection situés en Suisse. Pour pouvoir déterminer par exemple la commune dont fait partie un point d'intersection, il faut connaître l'état des communes suisses. Ordinairement, on utilise pour chaque année considérée l'état des communes le plus récent, y compris pour les années passées, sans que cela soit toujours précisé expressément.

Pour chaque point, on relève au maximum une valeur par année. Deux relevés sont généralement distants de plusieurs années. Le survol d'un territoire se fait normalement dans les mois d'été, époque où la couverture du sol est la plus visible. L'année est donc utilisée comme unité de temps des séries chronologiques. Pour chaque point d'intersection, les données qui lui sont associées, constituent une série chronologique. Leur ensemble forme un espace de séries temporelles de grande dimension.

NOAS04 est la nomenclature standard de la couverture et de l'utilisation du sol en Suisse. Elle comporte quatre degrés d'agrégation: le premier à 4 catégories, le deuxième à 17 catégories, le troisième à 27 catégories et le quatrième à 72 catégories (catégories de base). Le premier degré d'agrégation, le plus grossier, est constitué par les surfaces d'habitat et d'infrastructure (s), les surfaces agricoles (l), les surfaces boisées (w) et les surfaces improductives (u). Dans les explications qui suivent, nous utiliserons souvent, à titre d'illustration, ces quatre catégories les plus générales. Chaque surface appartient à une seule catégorie: il n'y a pas de chevauchements.

Notations supplémentaires

$y_t(i, j)$: catégorie de couverture et d'utilisation du sol au point (i, j) pour l'année t .

$f_k(i, j)$: année du k^e survol du point (i, j) . Il y a eu jusqu'ici quatre périodes de survol: 1979-1985 ($k = 1$), 1992-1997 ($k = 2$), 2004-2009 ($k = 3$) et 2013-2018 ($k = 4$).

$y_{f_k(i, j)}(i, j)$: catégorie de couverture et d'utilisation du sol du point (i, j) lors de la k^e période de survol de ce point (i, j) .

$y_{f_k}(i, j)$: forme abrégée de l'expression précédente.

d : une zone géographique (pas nécessairement d'un seul tenant). Ces zones sont basées sur les polygones des communes, des districts et des cantons. Au niveau le plus élevé, d correspond à l'ensemble de la Suisse ($d = CH$).

$(i, j) \in d$: le point (i, j) se trouve dans la zone d .

Par exemple, $\sum_{(i,j) \in BL} I_{\{y_{2010}(i,j)=s\}}$ exprime le nombre de points de la catégorie surfaces d'habitat et d'infrastructure s dans le canton de Bâle-Campagne en 2010. La fonction indicatrice I , est définie par:

$$I_{\{condition\}} = \begin{cases} 1 & \text{si } condition \text{ remplie} \\ 0 & \text{sinon.} \end{cases}$$

Applications

Ces définitions étant posées, on peut pour l'année 2009, dernière année de la période de survol 2004-2009, exprimer de la manière suivante les grandeurs qui nous intéressent.

- a1) Nombre de points d'intersection de la catégorie surfaces d'habitat et d'infrastructure s dans le canton de Zurich, $d = ZH$, en 2009:

$$\sum_{(i,j) \in ZH} I_{\{y_{2009}(i,j)=s\}}.$$

- a2) Proportion de points de la catégorie surfaces d'habitat et d'infrastructure s dans le canton de Zurich en 2009:

$$\frac{1}{\sum_{(i,j) \in ZH} I} \sum_{(i,j) \in ZH} I_{\{y_{2009}(i,j)=s\}}.$$

Remarque : I désigne une fonction constante égale à 1, de sorte que sommer les valeurs de I revient à compter le nombre de termes dans la somme, c'est-à-dire, dans notre exemple, le nombre de points dans le canton de Zurich.

- a3) Nombre de points de la catégorie surfaces d'habitat et d'infrastructure s en Suisse, $d = CH$, en 2009:

$$\sum_{(i,j) \in CH} I_{\{y_{2009}(i,j)=s\}}.$$

- a4) Proportion de points de la catégorie surfaces d'habitat et d'infrastructure s en Suisse en 2009:

$$\frac{1}{\sum_{(i,j) \in CH} I} \sum_{(i,j) \in CH} I_{\{y_{2009}(i,j)=s\}}.$$

Les termes qui sont définis par a1 - a4 peuvent être considérés comme des paramètres cible, qui sont estimés dans le cadre de la statistique de la superficie:

- b1) En 2008 (pendant la troisième période de survol), le canton de ZH (d) a été entièrement survolé ($f_3 = 2008$). On a établi le nombre de points d'intersection qui se trouvent dans la catégorie surfaces d'habitat et d'infrastructure:

$$\sum_{(i,j) \in ZH} I_{\{y_{f_3}(i,j)=s\}} = S_{f_3, ZH}.$$

Pour l'année 2009 (a1), on utilise les photographies aériennes de la période de survol 2004-2009. On considère que les changements survenus entre 2008 et 2009 sont négligeables.

- b2) On obtient une estimation de la proportion des surfaces d'habitat et d'infrastructure s du canton de Zurich en 2009 par la formule suivante :

$$\frac{1}{\sum_{(i,j) \in ZH} I} \sum_{(i,j) \in ZH} I_{\{y_{f_3}(i,j)=s\}}.$$

- b3) Nombre de points de la catégorie surfaces d'habitat et d'infrastructure s en Suisse, $d = CH$, en 2009:

$$\sum_{(i,j) \in CH} I_{\{y_{f_3}(i,j)=s\}}.$$

A noter que la troisième période de survol s'étend sur les années 2004-2009 et que $f_3(i, j)$ se rapporte aux années de cette période. Ici aussi, on néglige les changements survenus pendant la période de survol.

- b4) La formule suivante donne une estimation de la proportion de points appartenant à la catégorie surfaces d'habitat et d'infrastructure s en Suisse en 2009:

$$\frac{1}{\sum_{(i,j) \in CH} I} \sum_{(i,j) \in CH} I_{\{y_{f_3}(i,j)=s\}}.$$

Complément d'information: dans les cas b3 et b4, on utilise pour chaque point (i, j) la valeur observée pendant la troisième période de survol (la plus récente) comme estimation $\hat{y}_{2009}(i, j)$ de la catégorie de couverture et d'utilisation du sol en 2009:

$$\hat{y}_{2009}(i, j) = y_{f_3(i,j)}(i, j). \quad (1)$$

Dans le meilleur des cas, on a la situation $f_3(i, j) = 2009$, c'est-à-dire

$$\hat{y}_{2009}(i, j) = y_{2009}(i, j);$$

dans le pire des cas, on a $f_{3(i,j)} = 2004$, c'est-à-dire

$$\hat{y}_{2009}(i, j) = y_{2004}(i, j).$$

Dans le cas le plus favorable, nous avons $2009 - f_3(i, j) = 0$ année d'écart entre l'année de référence pour les résultats 2009 et l'année où a été prise la photographie aérienne. Dans le cas le plus défavorable, l'écart est de $2009 - f_3(i, j) = 2009 - 2004 = 5$ années. D'après ce qui précède, on pourrait songer à indiquer, comme donnée complémentaire à l'estimation de la couverture et d'utilisation du sol d'une région d , l'âge moyen $\bar{a}_{d,z}$ des photographies aériennes utilisées pour l'année z (ici $z = 2009$):

$$\bar{a}_{d,2009} = \frac{1}{\sum_{(i,j) \in d} I} \sum_{(i,j) \in d} (2009 - f_3(i, j)). \quad (2)$$

2 Changements

Nous décrivons maintenant, comment, dans le cadre de ces définitions, on peut exprimer les changements survenus entre deux périodes de survol. Pour plus de clarté, ces questions seront abordées ici à travers ces deux situations.

1. Supposons que la troisième période de survol ait donné, pour une zone donnée d , une surface d'habitat et d'infrastructure $S_{f_3,d}$. Supposons que d ait été survolé à nouveau au cours de la quatrième période et que l'on ait obtenu alors une surface d'habitat et d'infrastructure $S_{f_4,d}$. Dans le cas le plus simple, la zone d a été survolée chaque fois entièrement en l'espace d'une même année, p. ex. en 2005 et 2014. Supposons enfin qu'on observe un accroissement des surfaces d'habitat et d'infrastructure de $S_{f_4,d} - S_{f_3,d} = 1800$ points d'intersection, soit 1800 hectares¹. En admettant une croissance linéaire, on aurait une augmentation annuelle de 200 hectares. Pour être complet, il faudrait indiquer sur quelle période cet accroissement a eu lieu, p. ex. en indiquant une période d'observation, qui serait ici de 9 ans.
2. Mais qu'en est-il si, au cours de la troisième période de survol, la zone d a été survolé en 2005 et en 2006 et si, au cours de la quatrième période, il a été survolé entièrement pendant l'année 2014? Par quoi faut-il diviser $S_{f_4,d} - S_{f_3,d}$ pour obtenir l'accroissement annuel? Et quelle période d'observation faut-il indiquer?

Approche retenue: Calculer pour chaque point d'intersection un changement annuel et une période d'observation correspondante.

Le changement pour une ou plusieurs catégories de couverture et d'utilisation du sol, par exemple pour la catégorie surfaces d'habitat et d'infrastructure s , entre la 3^e période de survol (2004-2009) et la 4^e (2013-2018) est décrite par la fonction suivante:

$$\Delta(i, j, s, f_3, f_4) = \begin{cases} 1 & \text{si } y_{f_3}(i, j) \neq s, y_{f_4}(i, j) = s \\ -1 & \text{si } y_{f_3}(i, j) = s, y_{f_4}(i, j) \neq s \\ 0 & \text{sinon.} \end{cases} \quad (3)$$

¹ Voir les explications données dans le chapitre suivant.

Le changement annuel (*pa* pour «per annum» ou «pro anno») pour les points d'intersection (i, j) se calcule alors de la manière suivante:

$$\Delta_{pa}(i, j, s, f_3, f_4) = \frac{\Delta(i, j, s, f_3, f_4)}{f_4(i, j) - f_3(i, j)} \quad (4)$$

On obtient le changement annuel total pour tous les points d'intersection de la zone d par la formule:

$$\Delta_{pa,a}(d, s, f_3, f_4) = \sum_{(i,j) \in d} \Delta_{pa}(i, j, s, f_3, f_4) \quad (5)$$

Le deuxième indice a indique ici la zone, c'est-à-dire l'area, pour distinguer la formule (5) de la formule (4).

Exemple : Pour calculer $\Delta_{pa,a}$, dans le cas simple où la totalité de la zone a a été survolée les deux fois au cours d'une même année (situation 1), on a

$$\begin{aligned} \Delta_{pa,a}(d, s, f_3, f_4) &= \sum_{(i,j) \in d} \Delta_{pa}(i, j, s, f_3, f_4) \\ &= \sum_{(i,j) \in d} \frac{\Delta(i, j, s, f_3, f_4)}{f_4(i, j) - f_3(i, j)} \\ &= \frac{S_{f_4,d} - S_{f_3,d}}{9}, \end{aligned}$$

ce qui correspond au changement annuel moyen au cours de ces 9 ans.

Interprétation : Si, pour une zone d , la formule (5) donne 6.4, on peut considérer ce chiffre comme l'accroissement annuel moyen des surfaces d'habitat et d'infrastructure, soit 6.4 hectares par an.

Si l'on veut comparer le changement annuel dans plusieurs zones de tailles différentes, on peut rapporter le changement annuel, par exemple pour les surfaces d'habitat et d'infrastructure s , à la surface totale de la zone d . On obtient la différence entre la proportion de s sur la zone d entre la première et la deuxième période de survol. Comme la proportion est un pourcentage, le changement est en points de pourcentage.

$$\bar{\Delta}_{pa,a}(d, s, f_3, f_4) = \frac{\Delta_{pa,a}(d, s, f_3, f_4)}{\sum_{(i,j) \in d} I}. \quad (6)$$

Interprétation : Si (6) fournit par exemple une valeur de 0.015 dans une zone d , cela correspond à une croissance annuelle moyenne de la surface d'habitat de 1.5% par rapport à la surface totale de d , respectivement d'une croissance annuelle moyenne de la part de surface d'habitat de 1.5 point de pourcentage.

Pour le changement annuel total (5) ou le changement du pourcentage (6), on peut, par analogie avec (2), indiquer la durée moyenne de la période d'observation.

$$\bar{a}(d, f_3, f_4) = \frac{\sum_{(i,j) \in d} (f_4(i, j) - f_3(i, j))}{\sum_{(i,j) \in d} I}.$$

En transformant légèrement cette formule, on obtient, comme durée moyenne de la période d'observation, la différence entre les années de vol moyennes de la zone d :

$$\bar{a}(d, f_3, f_4) = \frac{\sum_{(i,j) \in d} f_4(i, j)}{\sum_{(i,j) \in d} I} - \frac{\sum_{(i,j) \in d} f_3(i, j)}{\sum_{(i,j) \in d} I} = \bar{f}_{4,d} - \bar{f}_{3,d}. \quad (7)$$

La formule (6) exprime le changement annuel par rapport à la superficie totale de la zone d . Si on veut exprimer le changement par rapport aux surfaces d'habitat et d'infrastructure pendant la troisième période de survol, on obtient, avec b1) et (5),

$$\bar{\Delta}_{pa,f_3}(d, s, f_3, f_4) = \frac{\Delta_{pa,a}(d, s, f_3, f_4)}{\sum_{(i,j) \in d} I_{\{y_{f_3}(i,j)=s\}}}. \quad (8)$$

Interprétation : Si par exemple, sur la zone d , (8) donne la valeur de 0.1, soit 10 %, ce chiffre est l'accroissement annuel moyen (appelé aussi *taux de changement annuel*) des surfaces d'habitat et d'infrastructure par rapport aux surfaces d'habitat et d'infrastructure de la troisième période de survol.

Dans nos exemples, les formules se rapportent aux changements entre la troisième et la quatrième période de survol et à la catégorie surfaces d'habitat et d'infrastructure s . On peut les appliquer par analogie à d'autres catégories de couverture et d'utilisation du sol et à d'autres périodes de survol.

Propriété d'invariance de la surface

La propriété d'invariance des surfaces ci-dessous découle de la formule (3) :

Si, sur une zone donnée qu'il s'agisse d'un point d'intersection unique ou de l'ensemble de la Suisse on fait la somme de toutes les catégories, alors tous les changements s'annulent et on obtient toujours zéro.

Cela vaut également pour les formules (4), (5) et (6). Comme le dénominateur de la formule (8) dépend de la catégorie de couverture et d'utilisation du sol, l'invariance des surfaces ne vaut pas pour la formule (8).

3 Du nombre de points à la superficie

Les coordonnées hectométriques de la carte nationale de la Suisse forment un quadrillage dans lequel chaque carré représente un hectare. Chaque hectare (100m x 100m) est associé au point d'intersection inférieur gauche du carré. Ainsi, chaque point du quadrillage correspond à un hectare unique.

La méthode d'estimation de la couverture et d'utilisation du sol consiste à identifier la catégorie de couverture et d'utilisation du point d'intersection avec celle du carré correspondant. Dans cette approche, le nombre de points correspondant à une catégorie de couverture et d'utilisation du sol s'interprète directement comme estimation du nombre d'hectares de cette catégorie de couverture et d'utilisation du sol. Par exemple, pour l'estimation, en hectares, de la surface d'habitat et d'infrastructure du canton de Zurich en 2009, on aurait:

$$\sum_{(i,j) \in ZH} I_{\{y_{2009}(i,j)=s\}}.$$

L'incertitude liée au passage du point à la superficie est présentée, dans la documentation actuelle sur la statistique de la superficie ([1]), comme une erreur d'échantillonnage. Elle est exprimée par des intervalles de confiance sur la base d'une approximation de la distribution binomiale à l'aide de la distribution normale. Dans ce contexte, les points d'intersection de la statistique de la superficie sont souvent appelés points d'échantillonnage.

Références

- [1] David Altwegg und Sektion GEO (2015). Die Bodennutzung in der Schweiz - Auswertungen und Analysen / Bundesamt für Statistik. Neuchâtel, 2015. - BFS Publikation. (<https://www.bfs.admin.ch/bfs/de/home/statistiken/raum-umwelt/bodennutzungsbedeckung.assetdetail.349272.html>)