



AutorInnen: Monika Ferster, Daniel Assoulin, METH, BFS, Juli 2019

Methodologische Notiz

Arealstatistik: Referenzdokument zu den Formeln

1 Formalisierung: Notationen und Beispiele

Ein 1-Hektar Gitter über die Schweiz gelegt gibt ca. 4.1 Millionen Gitterschnittpunkte (Schnittpunkte der 100m-Koordinaten der Landeskarte der Schweiz). Wir indexieren diese Gitterschnittpunkte durch ihre Koordinaten (i, j) mit $(i, j) \in K$, wobei K die Gesamtheit aller Koordinaten der in der Schweiz liegenden Gitterschnittpunkte bezeichnet. Um die Gitterschnittpunkte zum Beispiel Gemeinden zuordnen zu können, braucht es zu K auch einen Gemeindestand. Üblicherweise wird dafür bei den interessierenden Jahren der Gemeindestand des aktuellsten Jahres genommen und die zurückliegenden Jahre auf diesen bezogen. Dies wird manchmal implizit vorausgesetzt und deshalb nicht explizit erwähnt.

Pro Gitterschnittpunkt wird maximal ein Wert pro Jahr erhoben, wobei zwischen 2 Erhebungen normalerweise mehrere Jahre liegen. Meist wird ein Gebiet in den Sommermonaten überflogen, da dann die Bodenbedeckung am sichtbarsten ist. Aufgrund dieser Situation wurden die Jahre als Zeiteinheit für die betrachteten Zeitreihen verwendet. Jeder Gitterschnittpunkt mit seinen Daten stellt eine Zeitreihe dar. In der Gesamtheit haben wir also einen hochdimensionalen Zeitreihenraum.

NOAS04 heisst die Standardnomenklatur für die Bodenbedeckung und -nutzung. Sie hat vier Aggregationsstufen: 1-4 Stufen, 1-17 Stufen, 1-27 Stufen oder 1-72 Grundkategorien. Die grösste Kategorisierung entspricht den Siedlungsflächen (s), Landwirtschaftsflächen (l), bestockten Flächen (w) und den unproduktiven Flächen (u). In den folgenden Ausführungen verwenden wir aus Anschaulichkeitsgründen oft diese grösste Einteilung. Die Kategorie einer Fläche ist eindeutig, d.h. es kommen keine Überlappungen vor.

Weitere Bezeichnungen

$y_t(i, j)$: Bedeckungs- und Nutzungskategorie des Gitterschnittpunktes (i, j) im Jahr t .

$f_k(i, j)$: Jahr des k -ten Ueberflugs des Gitterschnittpunktes (i, j) . Bisher gab/gibt es vier Überflugsperioden: 1979-1985 ($k = 1$), 1992-1997 ($k = 2$), 2004-2009 ($k = 3$) und 2013-2018 ($k = 4$).

$y_{f_k(i, j)}(i, j)$: liefert die Bedeckungs- und Nutzungskategorie für den Gitterschnittpunkt (i, j) in der k -ten Überflugsperiode eben dieses Gitterschnittpunktes (i, j) .

$y_{f_k}(i, j)$: Kurzform der obigen Bezeichnung.

d : Ein geografisches Gebiet, nicht notwendigerweise zusammenhängend. Die relevanten Gebiete basieren auf Gemeinde-, Bezirks- und Kantonspolygonen, respektive entsprechenden Aggregaten. Maximal entspricht d der ganzen Schweiz ($d = CH$).

$(i, j) \in d$: Gitterschnittpunkt (i, j) liegt im Gebiet d .

Mit $\sum_{(i, j) \in BL} I_{\{y_{2010}(i, j)=s\}}$ würde damit die Anzahl Gitterschnittpunkte der Bedeckungs- und Nutzungskategorie Siedlungsfläche s innerhalb des Kantons Basel Land im Jahr 2010 ausgedrückt. Für die Indikatorfunktion I gilt:

$$I_{\{Bedingung\}} = \begin{cases} 1 & \text{Bedingung ist erfüllt} \\ 0 & \text{sonst.} \end{cases}$$

Anwendungen

Mit den obigen Definitionen lassen sich für das Referenzjahr 2009, das letzte Jahr der Überflugsperiode 2004-2009, interessierende Grössen wie folgt ausdrücken.

- a1) Anzahl Gitterschnittpunkte der Kategorie Siedlungsfläche s des Gebietes Kanton Zürich, $d = ZH$, im Jahr 2009:

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

- a2) Anteil Punkte der Kategorie Siedlungsfläche s innerhalb des Kanton Zürich im Jahr 2009:

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

Bemerkung: I entspricht einer konstanten Funktion mit dem Wert 1 und die Summe der Werte von I somit der Anzahl Summanden in der Summe. Im Beispiel ist dies die Anzahl Gitterschnittpunkte im Kanton ZH.

- a3) Anzahl Gitterschnittpunkte der Kategorie Siedlungsfläche s der ganzen Schweiz, $d = CH$, im Jahr 2009:

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

- a4) Anteil Punkte der Kategorie Siedlungsfläche s innerhalb der Schweiz im Jahr 2009:

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

Die Terme in a1 - a4 können als theoretische Parameter betrachtet werden, welche im Rahmen der Arealstatistik geschätzt werden:

- b1) Im Jahr 2008 (gehört zur dritten Überflugsperiode) wird der Kanton ZH (d) vollständig überflogen ($f_3 = 2008$) und es wird festgehalten, wieviele Gitterschnittpunkte in die Kategorie Siedlungsfläche fallen:

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

Für das Jahr 2009 (a1) verwendet man die Luftbilder aus der Überflugsperiode 2004-2009 und geht implizit davon aus, dass eine allfällige Veränderung zwischen 2008 und 2009 vernachlässigbar ist.

- b2) Analog erhält man die Schätzung des Anteils Siedlungsfläche s des Kantons Zürich im Jahr 2009 mittels

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

- b3) Anzahl Punkte der Kategorie Siedlungsfläche s der ganzen Schweiz, $d = CH$, im Jahr 2009:

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

Zu beachten ist, dass sich die dritte Überflugsperiode über die Jahre 2004-2009 erstreckt und $f_3(i, j)$ entsprechende Jahre in dieser Periode reflektiert. Auch hier werden Entwicklungen innerhalb der Überflugsperiode vernachlässigt.

- b4) Analog erhält man die Schätzung für den Anteil Punkte der Kategorie Siedlungsfläche s innerhalb der Schweiz im Jahr 2009 mittels:

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

Eine etwas andere Betrachtungsweise: In den Fällen b3 und b4 verwendet man für jeden Punkt (i, j) den Wert aus der dritten (aktuellsten) Überflugsperiode als Schätzung $\hat{y}_{2009}(i, j)$ für die Bedeckungs- und Nutzungskategorie des Jahres 2009:

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

Bestenfalls hat man die Situation $f_3(i, j) = 2009$, das heisst

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

und im schlechtesten Fall $f_3(i, j) = 2004$, das heisst

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

Im besten Fall haben wir $2009 - f_3(i, j) = 0$ Jahre Differenz zwischen dem Referenzjahr für die Resultate 2009 und dem Jahr des Luftbildes. Im ungünstigsten Fall beträgt diese Differenz $2009 - f_3(i, j) = 2009 - 2004 = 5$ Jahre. Aufgrund obiger Betrachtungen könnte man sich überlegen, als Zusatzangabe zur Schätzung der Bedeckungs- und Nutzungskategorien einer Region d das durchschnittliche Alter $\bar{a}_{d,z}$ der entsprechenden, im Jahr z verwendeten Luftbilder anzugeben (hier $z = 2009$):

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

2 Veränderungen

Nun kommen wir zur Frage, wie sich Veränderungen zwischen zwei Überflugsperioden mit diesen Definitionen ausdrücken lassen. Aus Verständnisgründen werden im Folgenden solche Fragen und Antworten exemplarisch aufgeführt.

1. Die dritte Flugperiode lieferte für ein gewisses Gebiet d eine Siedlungsfläche von $S_{f_3,d}$. In der vierten Flugperiode sei d schon überflogen worden und man konnte eine Siedlungsfläche von $S_{f_4,d}$ berechnen. Im einfachsten Fall wird das Gebiet d jeweils in einem Jahr vollständig überflogen, z.B. 2005 und 2014. Angenommen man hätte eine Siedlungsflächen-Zunahme von $S_{f_4,d} - S_{f_3,d} = 1800$ Gitterschnittpunkten respektive 1800 Hektaren¹. Unter Annahme eines linearen Wachstums hätte man eine jährliche Zunahme von 200 Hektaren an Siedlungsfläche. Der Vollständigkeit halber müsste man auch noch angeben, aus welchem Zeitraum dieses jährliche Wachstum stammt, z.B. durch die Angabe einer Betrachtungsperiode von hier 9 Jahren.
2. Wie sieht es aber aus, wenn das Gebiet d in der dritten Überflugsperiode in den Jahren 2005 und 2006 überflogen wird und in der vierten immer noch vollständig im Jahr 2014?
Wodurch soll $S_{f_4,d} - S_{f_3,d}$ dividiert werden, um das jährliche Wachstum anzugeben?
Welche Betrachtungsperiode müsste dazu angegeben werden?

Ansatz: Für jeden einzelnen Gitterschnittpunkt eine jährliche Veränderung samt zugehöriger Betrachtungsperiode berechnen.

Die Veränderung punkto einer oder mehrerer Bedeckungs- und Nutzungskategorien - hier wird als Beispiel die Siedlungsfläche s genommen - der vorherigen 3. Überflugsperiode 2004-2009 zur aktuellen 4. Überflugsperiode 2013-2018 wird durch folgende Funktion beschrieben:

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

Damit lässt sich die jährliche (pa für "per annum" oder "pro anno") Veränderung des Gitterschnittpunktes (i, j) wie folgt berechnen:

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

und man erhält die totale jährliche Veränderung aller Gitterschnittpunkte im Gebiet d mit

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

¹ Erklärung siehe nächstes Kapitel.

Der zweite Index a steht hier für Gebiet respektive area, um Formel (5) von Formel (4) zu unterscheiden.

Beispiel: Berechnet man $\Delta_{pa,a}$ im einfachen Fall, in welchem das gesamte Gebiet in demselben Überflugsjahr erhoben wird (Situation 1.), ergibt sich

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

die mittlere jährliche Veränderung dieser 9 Jahre.

Interpretation: Falls (5) in einem Gebiet d einen Wert von 6.4 liefert, kann dies als mittlerer jährlicher Zuwachs der Siedlungsfläche von 6.4 Hektaren bezeichnet werden.

Möchte man unterschiedlich grosse Gebiete bezüglich jährlicher Veränderungen vergleichen, so kann die jährliche Veränderung zum Beispiel der Bedeckungs- und Nutzungskategorie s in Beziehung zur Gesamtfläche eines Gebiets d gesetzt werden. Es resultiert dabei die Differenz zwischen dem Anteil s im Gebiet d in den beiden Überflugsperioden. Da der Anteil eine prozentuale Angabe ist, ist die Veränderung ein Wert in Prozentpunkten.

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

Interpretation: Falls (6) in einem Gebiet beispielsweise einen Wert von 0.015 liefert, entspricht dies einem mittleren jährlichen Zuwachs der Siedlungsfläche von 1.5% relativ zur Gesamtfläche von d , respektive um einen mittleren jährlichen Zuwachs des Anteils Siedlungsfläche von 1.5 Prozentpunkten.

Zur totalen jährlichen Veränderung (5), respektive der Veränderung des Anteils (6), kann in Analogie zu (2) auch noch die mittlere Länge der Betrachtungsperiode angegeben werden.

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

Wird diese Formel etwas umgeformt, so erhält man als mittlere Länge der Betrachtungsperiode die Differenz der gemittelten Flugjahre des Gebietes d :

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

In (6) wird die jährliche Veränderung im Verhältnis zur totalen Bodenfläche des Gebiets d ausgedrückt. Möchte man auch Veränderungen relativ zur Fläche der Bedeckungs- und Nutzungskategorie der dritten Überflugsperiode ausdrücken, erhält man mit b1) und (5)

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

Interpretation: Falls (8) in einem Gebiet d beispielsweise einen Wert von 0.1, respektive 10 % liefert, kann dies als mittlerer jährlicher Zuwachs (oder auch als mittlere jährliche *Veränderungsrate*) der Siedlungsfläche im Verhältnis zur Siedlungsfläche gemäss dritter Überflugsperiode bezeichnet werden.

Zu Illustrationszwecken beziehen sich die Formeln hier auf die Veränderungen zwischen dritter und vierter Überflugsperiode und die Bedeckungs- und Nutzungskategorie s . Sie lassen sich analog auf andere Bedeckungs- und Nutzungskategorien und andere Überflugsperioden anwenden.

Eigenschaft Flächeninvarianz

Formel (3) besitzt die Eigenschaft Flächeninvarianz:

Wird in einem Gebiet - das kann ein einziger Gitterschnittpunkt oder auch die ganze Schweiz sein - über alle Kategorien summiert, so heben sich die gegenseitigen Veränderungen auf und es gibt immer Null.

Für die darauf aufbauenden Formeln (4), (5) und (6) gilt diese Eigenschaft ebenso. Weil der Nenner der Formel (8) von der Bedeckungs- und Nutzungskategorie abhängt, gilt für Formel (8) die Flächeninvarianz jedoch nicht.

3 Von der Anzahl Punkte zur Fläche

Die 100m-Koordinaten der Landeskarte der Schweiz überziehen die Schweiz mit einem Hektaren-Netz. Jedem Hektar (100m x 100m Quadrat) ist der Gitterschnittpunkt links unten zugeordnet. Insofern beschreibt ein Gitterschnittpunkt eindeutig einen Hektar.

Die für die Flächenschätzung verwendete Methode identifiziert die Bedeckungs- und Nutzungskategorie eines Gitterquadrates mit der Bedeckungs- und Nutzungskategorie des Gitterschnittpunktes. Mit diesem Vorgehen lassen sich die Anzahl Punkte einer gewissen Bedeckungs- und Nutzungskategorie direkt als geschätzte Anzahl Hektaren dieser Bedeckungs- und Nutzungskategorie interpretieren. Beispielsweise ergäbe sich für die geschätzte Siedlungsfläche des Kantons Zürich im Jahr 2009 in Hektaren:

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

Die mit dem Übergang von Punkt zur Fläche verbundene Unsicherheit wird in der aktuellen Dokumentation der Arealstatistik ([1]) oft als Stichprobenfehler bezeichnet und mittels Konfidenzintervallen basierend auf einer Normalverteilungsapproximation der Binomialverteilung ausgedrückt. Vor diesem Hintergrund werden die Gitterschnittpunkte in der Arealstatistik oft auch als Stichprobenpunkte bezeichnet.

Literatur

- [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/bodennutzung-bedeckung.assetdetail.349272.html>)