

VII. Bestimmung der Stichprobengröße - Wie viele Felder sind auf dem Raster zu zählen?

Sally Soria-Dengg, GEOMAR Helmholtz-Zentrum für Ozeanforschung Kiel
Düsternbrookerweg 20, 24105 Kiel, Deutschland

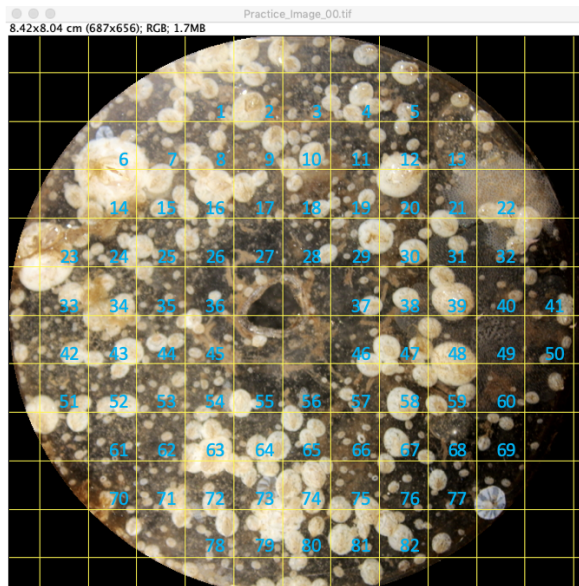
Eine kritische Frage bei der quantitativen Bildanalyse ist, wie repräsentative Teilproben einer Stichprobe ermittelt werden können. Dafür sind zwei Aspekte wichtig:

- die Auswahl der Teilproben sollte nach dem Zufallsprinzip erfolgen, um systematische Fehler zu vermeiden, und
- die Anzahl der Proben muss ausreichend groß sein, um statistisch signifikante Ergebnisse zu ermöglichen.

Für die Zählung von Individuen in Fotos der VIRTUE-Platten werden hier einige Ansätze vorgestellt.

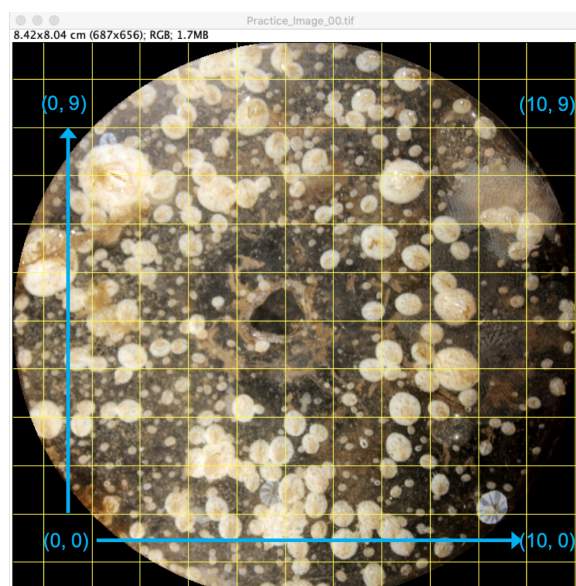
Erstellen zufälliger Teilproben:

1. Legen Sie mit dem Bildanalyseprogramm Fiji/ImageJ ein Gitter über Ihr Foto. (Vergl. Tutorial „I. Zählen von Bewuchs-Organismen auf einer Platte mit ImageJ (manuell)“.)
2. Die Größe der Quadrate hängt von der Größe und der Verteilung der auf den Platten wachsenden Organismen ab. Bei großen Organismen sollten die Quadrate größer sein, bei kleineren Organismen können sie kleiner sein. Wenn die Organismen gleichmäßig auf der Platte verteilt sind, kann die Gittergröße klein sein, aber wenn sie in Gruppen wachsen, sollten die Quadrate des Gitters größer sein, um sicherzustellen, dass alle Gruppen vertreten sind.
3. Speichern Sie das resultierende Bild als Referenz und zur späteren Verwendung in einem anderen Programm (siehe unten). Eine Möglichkeit, dies zu tun, besteht darin, **Image - Overlay - Flatten** zu verwenden, wodurch eine Kopie des Bildes (einschließlich des Rasters) erstellt wird, die Sie dann in einem Format Ihrer Wahl speichern können. Dies scheint jedoch nicht sehr zuverlässig zu funktionieren, so dass es besser sein kann, stattdessen einen Screenshot zu machen.
4. Ordnen Sie den Quadraten Referenznummern zu. Es gibt 2 Optionen:
 - a. Fügen Sie z.B. in PowerPoint oder OpenOffice nummerierte Beschriftungen in das Raster ein (von links oben beginnen nach rechts und den Vorgang in der nächsten Zeile wiederholen). Das Ergebnis wird eine Anzahl von N Gitterzellen sein; in unserem Beispiel $N=82$.
Dies ist vor allem dann nützlich, wenn Sie sich für alle Ihre Bilder auf dasselbe Gitter beziehen können, da die Nummerierung der Quadrate etwas mühsam ist.



Nummeriertes Gitter über dem Foto der Platte.

- b. Alternative: markieren Sie eine Gitterzelle unten links als Referenzpunkt ($x=0, y=0$) oder kurz $(0, 0)$ Ihres „Koordinatensystems“. Diese Zelle sollte so gewählt werden, dass sie die äußerste linke Spalte und die unterste Zeile markiert, die noch „vollständige“ Zellen enthalten (siehe Erklärung unten). Von dort aus können Sie später eine Gitterkoordinate (x, y) bestimmen, indem Sie x Zellen nach rechts und y Zellen nach oben zählen. Die letzte Zelle oben rechts hätte dann die Koordinaten (X_{\max}, Y_{\max}) , wobei X_{\max} die äußerste rechte Spalte und Y_{\max} die oberste Zeile mit vollständigen Gitterzellen wäre. In unserem Beispiel wären $X_{\max} = 10$ und $Y_{\max} = 9$. *Diese Methode ist schneller als a) und kann daher auf jedes Bild einzeln angewendet werden.*

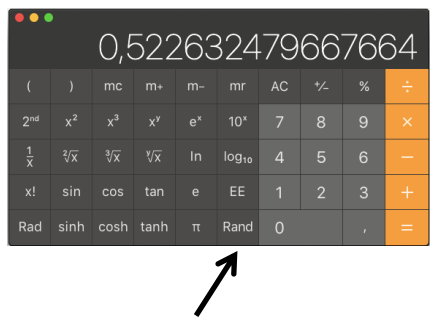


Koordinatengitter der Größe
 $(X_{\max}, Y_{\max}) = (10, 9)$

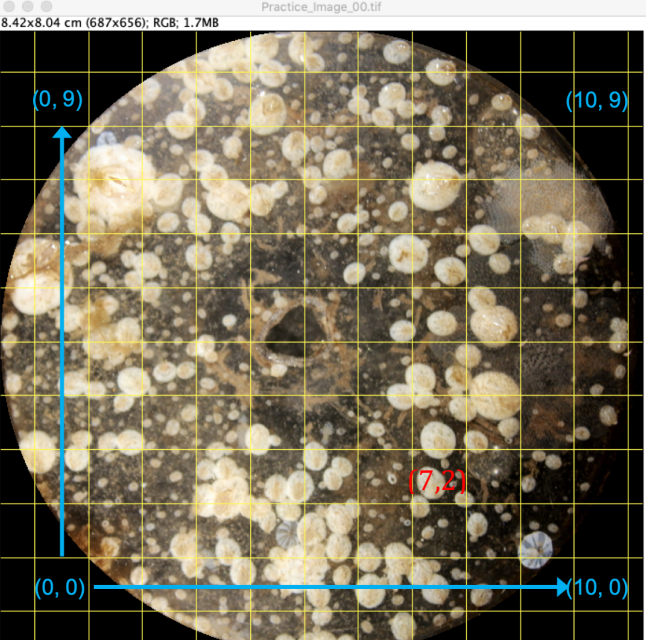
Bei beiden Methoden sollten nur „vollständige“ Zellen oder Quadrate einbezogen werden, um sicherzustellen, dass die Plattenfläche in jeder Stichprobe gleich ist. Dies sind Zellen, die keine Kanten oder das Loch der Platte enthalten.

5. Wählen Sie zufällige Quadrate zum Zählen. Generieren Sie dazu Zufallszahlen (Funktion: Rand) auf Ihrem Taschenrechner oder in einer Tabellenkalkulation. Der Rechner wird eine Zahl zwischen 0 und 1 erzeugen.

- a. Für den obigen Fall a) multiplizieren Sie diese Zufallszahl mit der Gesamtzahl N Ihrer Quadrate. Das (gerundete) Ergebnis ergibt das auf Ihrem Gitter zu zählende Quadrat.

<p>Beispiel: Zufallszahl: 0,52</p> <p>$0,52 \times 82 = 42,85$</p> <p>Zähle Quadrat Nr. 43</p>	
---	--

- b. Für Fall b) erzeugen Sie 2 Zufallszahlen in Ihrem Rechner (siehe oben) und multiplizieren diese mit Ihren X_{\max} - bzw. Y_{\max} -Koordinaten.

<p>Beispiel:</p> <p>Zufallszahlen: 0,74 und 0,21</p> <p>$x = 0,74 \times 10 = 7,4$ $y = 0,21 \times 9 = 1,9$</p> <p>Zähle Gitterzelle mit den Koordinaten: $x=7$ und $y=2$.</p>	
---	--

Beachten Sie, dass Sie in beiden Fällen auf die **nächste** ganze Zahl (ohne Dezimalstellen) runden müssen, wenn Sie alle Zellen einbeziehen wollen. (Runden Sie nicht einfach ab! Beispiel: Wenn $X_{\max}=10$, können Sie Spalte 10 nicht ansprechen, indem Sie eine Zufallszahl von 9,6 auf 9 statt auf 10 runden).

Bestimmung der Stichprobengröße:

(Das folgende Beispiel bezieht sich auf Methode b), aber Methode a) funktioniert auf genau dieselbe Weise).

- Erzeugen Sie Zufallskoordinatenpaare für **mindestens** 10% der Gesamtzahl der Quadrate in Ihrem Gitter. In unserem Beispiel würde die Gesamtzahl 82 Quadrate betragen. (Denken Sie daran: Nur „vollständige“ Quadrate einbeziehen.) Um auf der sicheren Seite zu bleiben, haben wir hier 14 gültige Zufallskoordinatenpaare x und y generiert.
- Zählen Sie die Anzahl der Individuen der interessierenden Art in den entsprechenden Quadraten. Beispielwerte für Seepocken, die von unserer Musterplatte gewonnen wurden, sind unten dargestellt. (Siehe Tutorial „I. Zählen von Bewuchs-Organismen auf einer Platte mit ImageJ (manuell)“.).

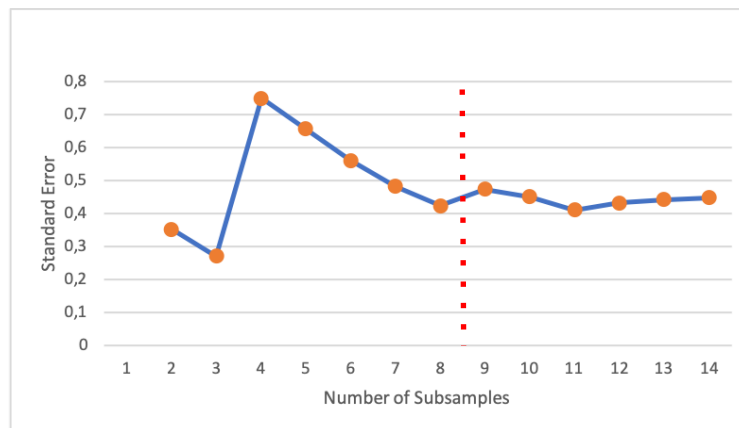
Teilprobe- Nummer, n	Zufalls- x-Koord.	Zufalls- y-Koord.	Anzahl (Individuen)
1	6	6	6
2	7	2	5
3	8	8	5
4	9	6	2
5	3	1	3
6	7	7	5
7	4	7	4
8	5	0	4
9	8	1	7
10	6	4	3
11	7	8	4
12	5	8	7
13	8	4	2
14	2	2	7

8. Berechnen Sie die kumulativen gleitenden Mittelwerte, Standardabweichungen und Standardfehler ($SE = s/\sqrt{n}$; s = Standardabweichung, n = Probengröße) der Werte, beginnend mit den ersten 2, dann den ersten 3 usw., bis wenigstens $n=10\%$ der Gesamtzahl der Quadrate in Ihrem Gitter.

Teilprobe- Nummer, n	Anzahl	Kumulatives Mittel	
1	6		
2	5	5,50	Mittel der Teilproben 1 & 2
3	5	5,33	Mittel der Teilproben 1, 2 & 3
4	2	4,50	Mittel der Teilproben 1, 2, 3 & 4
5	3	4,20	etc.
6	5	4,33	
7	4	4,29	
8	4	4,25	
9	7	4,56	
10	3	4,40	
11	4	4,36	
12	7	4,58	
13	2	4,38	
14	7	4,57	Mittel der Teilproben 1 bis 14

Teilprobe- Nummer, n	Anzahl	Kumulatives Mittel	Kumulative Standard- abweichung (s)	Kumulativer Standard- fehler $SE = s/\sqrt{n}$
1	6	6,00		
2	5	5,50	0,50	0,35
3	5	5,33	0,47	0,27
4	2	4,50	1,50	0,75
5	3	4,20	1,47	0,66
6	5	4,33	1,37	0,56
7	4	4,29	1,28	0,48
8	4	4,25	1,20	0,42
9	7	4,56	1,42	0,47
10	3	4,40	1,43	0,45
11	4	4,36	1,37	0,41
12	7	4,58	1,50	0,43
13	2	4,38	1,60	0,44
14	7	4,57	1,68	0,45

9. Plotten Sie eine Kurve des Standardfehlers gegen die Anzahl der Teilproben. Wenn sich die Änderung des Standardfehlers einer horizontalen Linie nähert, d.h. der Standardfehler mit zunehmendem Stichprobenumfang nicht weiter abnimmt, dann ist der Stichprobenumfang groß genug. (Streng genommen sollte die Steigung der Linie weniger als 20 % der ursprünglichen Steigung betragen, aber dies würde einen weiteren Berechnungsschritt hinzufügen).



In unserem Beispiel hätte das Zählen von 8 oder 9 Quadraten ausgereicht, um eine ziemlich repräsentative Stichprobengröße zu erhalten.

10. Wenn es immer noch große Änderungen des Standardfehlers gibt, sollten mehr Quadrate gezählt werden.

