

# GIS Einführung mit QGIS

Christoph Rinne

23. Juni 2023

## Inhaltsverzeichnis

<b>1</b>	<b>Raum &amp; Distanz</b>	<b>1</b>
1.1	Einleitung . . . . .	1
1.2	Umkreisanalyse . . . . .	1
1.3	Thiessen-Polygon, Voronoi-Diagram und Interpolation . . . . .	7
1.4	Thiessen-Polygone und triangulierte Distanzen . . . . .	7
	<b>Literatur</b>	<b>9</b>

## 1 Raum & Distanz

### 1.1 Einleitung

Analysen zu archäologischen Fundplätzen und ihrem Umfeld hat es auch vor der *Computer-Revolution* in der Archäologie in den 1980ern und den modernen GIS, ab den 1990ern, gegeben. Ein schönes Beispiel für klassische Techniken der Umfeldanalyse ist die Arbeit von Linke ([Linke, 1976](#)). Die Ursprünge in unserem Fach liegen in der **Siedlungsarchäologie**, in der Fundstellen, das Umfeld und Informationen zur Vegetation oder das Klima verknüpft und möglichst quantitativ in ihrer Bedingtheit analysiert werden. Hier gibt es eine umfangreiche Literatur, von der ich nur eines nennen will ([Jankuhn, 1977](#)). Die uns heute geläufigen theoretischen Grundlagen (zum Beispiel zur Zentralität), die Methoden und auch die Technik kommen aber aus der Geographie, insbesondere der Kultur- und Sozialgeographie ([Christaller, 1933](#); [Haggett, 1973](#)). Mit Blick auf Raum und Distanz möchte ich aus der “Findungsphase” im Fach mit Bezug auf das Thema “Raum um die Fundplätze” auf Arbeiten von Zimmermann verweisen ([Zimmermann u. a., 2004](#); [Zimmermann und Wendt, 2003](#)) und für Analysen auf der Basis der Distanz auf etwas ältere Arbeiten ([Siegmund, 1992](#); [Zimmermann, 1992](#)). Die Verfügbarkeit von Hard- und Software hat zu einem großen Spektrum diverser Interpolationsverfahren geführt, die Punktinformationen in die Fläche übertragen. Eine gute Zusammenfassungen zu diesem Themenkomplex finden Sie bei Herzog ([Herzog, 2007](#)). Daneben gibt es zahlreiche spezielle Entwicklungen und Phasen spezifischer Fragestellungen im Fach (z.B. **Aktivitätsareale** , **Archäoprognose**, **Least cost**), auf die ich hier nur hinweisen möchte ([Cziesla, 1990](#); [Herzog, 2013](#); [Kunow u. a., 2007](#)). Nach einer Etablierung der Technik im Fach ist meiner Ansicht nach aktuell wieder eine Beschleunigung zu erkennen, die durch bessere technischen Möglichkeiten, neue Methoden der Informationsverarbeitung und mit der Programmiersprache R der Archäologie neue Perspektiven eröffnet ([Nakoinz und Knitter, 2016a](#)). Dieser extrem verkürzte Abriss zu GIS in der Archäologie blendet zahlreiche Fragestellungen und vor allem die internationale, oft englischsprachige Literatur aus. Letztere finden Sie aber mit etwas Routine und spezifischen Stichworten in den Repositorien der Fachliteratur.

### 1.2 Umkreisanalyse

#### 1.2.1 Vorbemerkungen

Der Gedanke, das Umfeld einer Siedlung als Aktivitäts- und Nutzungsareal für soziale und ökonomische Bedürfnisse zu betrachten, liegt auf der Hand. Damit ergibt sich als erstes die Frage nach dem Radius um die Siedlung. Weitere Fragen ergeben sich aus dem verfügbaren Raum und den möglichen Nachbarn: Wie viel Platz habe ich eigentlich? Welche Siedlung ist gleichzeitig? Diese und weitere, eher archäologische Fragen werden von Linke diskutiert ([Linke, 1976, S. 24](#)). Weitere Fragen ergeben sich aus der Geographie,

z.B. durch Berge und Flüsse, die unsere Mobilität beeinflussen, diese beschleunigen und begrenzen können. Diese Fragen sind nicht neu und inzwischen oft behandelt worden (u.a. [Herzog, 2013](#)). Die jeweilige Antwort beeinflusst als grundlegende Entscheidung das Ergebnis der Analyse. Diese Voraussetzungen (Parameter) sollten anhand der jeweils vorliegenden Daten und der eigenen Fragestellung angemessen als auch nachvollziehbar dargelegt werden.

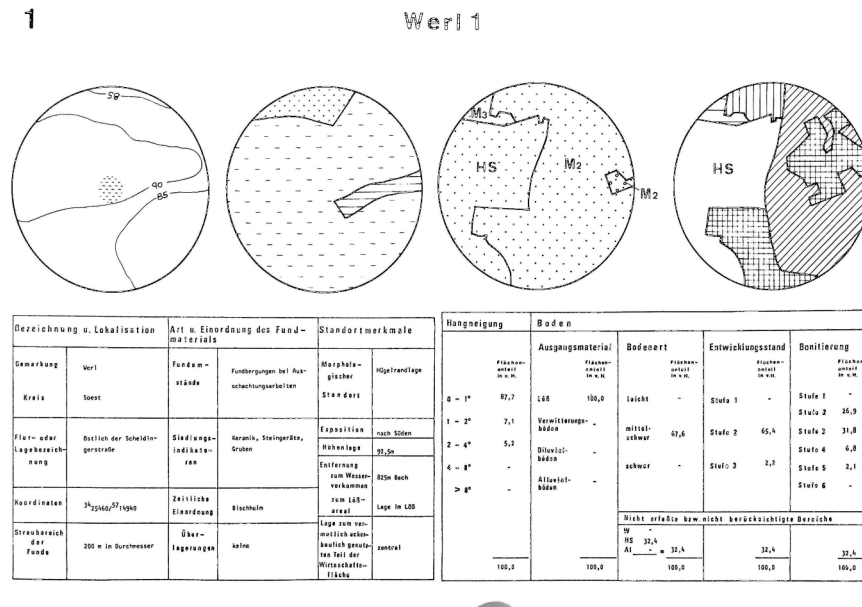


Abbildung 1: Umkreisanalyse von Linke mit quantitativer Auswertung (1976)

Um die Technik der Umkreisanalyse von Linke auf unsere Daten anzuwenden, will ich die Frage zur Lage der Talaiots im Geländere relief betrachten: ist diese exponiert, im mittleren Hangbereich oder eher in Tallage? Mein Ansatz ist der Vergleich der Höhenlage des Talaiot zum Umfeld. Dabei sind folgende Angaben zu ermitteln: 1. Die Reliefenergie im Umfeld und 2. die Lage des Talaiot zum Umfeld. Für die Frage der Kreisgröße berücksichtige ich als erstes die Auflösung unseres DGM von 200 m je Rasterzelle und entscheide mich, mangels einfach zu begründender Maße für einen Start mit  $r = 1$  km.

## 1.2.2 Umsetzen in QGIS

Als Daten verwende ich den originalen Rasterdatensatz des DGM 'PNOA\_MDT200\_ETRS89\_HU31\_Baleares.asc' und aus der SpatiaLite Datenbank die erstellte *view* mit den gefilterten talaiots `CREATE VIEW "talaiots" AS select * from sites where Tipo_yacim like 'Talaiot%'`.

Die Berechnung und Darstellung des Umkreises kann 1. in QGIS erfolgen, wahlweise in eine temporäre Datei oder in eine neue shp-Datei, oder 2. in SpatiaLite. Ersteres ist schnell und einfach, letzteres kann wegen der veränderten Geometrie (Punkt zu Polygon) nicht mehr als *spatial view* ausgeführt werden, sondern führt mit `create table <name> as select ...` ebenfalls zu redundanten Daten. Wesentlicher Vorteil: die Erstellung ist mit der SQL-Anweisung unmissverständlich nachvollziehbar. Die Analyse ist ansonsten identisch.

**1.2.2.1 Puffer erstellen** Im Menü von QGIS wählen Sie "Vektor -> Geometrieverarbeitung -> Puffer" oder suchen in der *toolbox* nach 'puffer' womit Sie auch weitere Varianten finden. Wir nutzen den einfachen 'Puffer', setzen Sie hier folgende Parameter: "Eingabelayer": talaiots, "Abstand": 1000 (Meter), belassen Sie den Rest auf den Vorgaben und erstellen einen temporären Layer. Wiederholen Sie den Vorgang, setzen Sie diesmal aber Segmente auf 100. Vergleichen Sie das Ergebnis bei einer großen Auflösung und lesen Sie dazu die Erläuterung rechts im Fenster zum Algorithmus. Starten Sie zum Vergleich auch nochmals die Funktion 'Vektoren puffern' aus der GDAL *toolbox*. Beachten Sie die ergänzenden Optionen "Nach Attribut Auflösen", d.h. Überlappende Puffer werden nicht pauschal aufgelöst sondern nur bei identischem Attribut verschmolzen, und "Erzeuge Objekt ... Geometriesammlung", wodurch bei Multi-Objekten (z.B. jeden Pfosten des Objektes Haus) ein eigener Puffer gezeichnet wird. Die Segmentierung des Kreises wird hier nicht erwähnt, erfolgt aber auch.

---

## Anmerkung

---

Trotz identischen Namens bieten und liefern die Tools der diversen Pakete Unterschiede. Und die angewendeten Parameter sind alle wichtig wenn das Ergebnis nachvollziehbar bleiben soll.

Das Feld **“Gemometriespaltenname”** wird **pauschal ausgefüllt** und muss eigenständig korrekt ausgefüllt werden.

In QGIS Version 3.18 erkennt dieses Tool aktuell die Geometrie in der SQLite und GPKG nicht und kann nicht ausgeführt werden.

---

**1.2.2.2 Rasterstatistik je Vektor (Kreis)** Im Ergebnis haben wir viele Kreise. Suchen Sie in der *toolbox* nach **‘Zonenstatistik’** (zonal stat) und starten Sie diese. Wählen Sie folgende Parameter: “Eingabelayer (ehem. Zonenvektorlayer)”: der Puffer, “Rasterlayer”: das Geländemodell (DGM 200 alias mdt 200), “Rasterkanal”: Kanal 1, “Ausgabespaltenpräfix”: s\_ (die Vergabe ist frei wählbar, sollte kurz sein und nicht mit einer Zahl starten) und bei “Zu berechnende Statistiken”: Wählen Sie bitte alle aus, bis auf ‘Summe’. Im Ergebnis haben Sie bei der Attributtabelle des Vektollayers für jede gewählte Statistik eine neue Spalte. Hier werden in der QGIS-Version (3.16 und 3.18) die englischen Begriffe verwendet, wodurch auch der Unterschied der zuvor doppelt angegebenen “Varianz” ersichtlich wird. Die **Reliefenergie** ist per Definition die Höhendifferenz in einer Flächeneinheit, Sie ist damit unabhängig von der absoluten Höhe über dem Meer. Zu den ermittelten Werten nachfolgend eine sehr knappe Erläuterung:

- Anzahl / count: Anzahl der betroffenen Zellen
- Summe (abgewählt): Summe der Zellwerte
- Mittelwert / mean: der Durchschnitt der gültigen Zellwerte (Zentralwert)
- Median / median: Der Wert in der Mitte der aufsteigend sortierten Zellwerte (Zentralwert)
- Standardabweichung / stdev: Der Durchschnitt der Differenz jedes Zellwertes zum Mittelwert aller Zellwerte (Streuungsmaß).
- Minimalwert / min: der kleinste Wert in einer Zelle
- Maximalwert / max: der größte Wert in einer Zelle
- Bereich / range: Differenz zwischen den beiden vorgenannten Min. und Max. (Streuungsmaß, die Reliefenergie nach Definition)
- Minderheit / minority: am seltensten vorkommende Wert
- Mehrheit / majority: am häufigsten auftretende Wert
- Varianz / variance: das Quadrat der Standardabweichung (Durchschnitt der quadrierten Differenzen jedes Wertes zum Mittelwert, Teil der Berechnung der Standardabweichung, Streuungsmaß).
- Eindeutige Werte / variety: Anzahl der unterschiedlichen Werte

Es gibt hier zahlreiche Werte, die einen Hinweis auf die Reliefenergie geben, vor allem:

- Bereich: Die Differenz zwischen dem kleinsten und größten Wert, wobei ein einzelner Gipfel das Ergebnis deutlich beeinflusst (z.B. der Kilimandscharo). In so einem Fall liegen der Mittelwert und der Median weiter auseinander.
- Standardabweichung: Gibt es viele Hügel und Täler sind viele Werte über ein breites Mittelfeld verteilt, die Standardabweichung ist demnach groß bzw. breit.
- Eindeutige Werte (variety), Vorausgesetzt wir haben Meterangaben oder sogar klassifizierte Werte in Einheiten von 5 Metern ist eine Ebene eher “langweilig”, eine leicht schräge Steilwand aber hoch dramatisch. Eine hügeliges Bergvorland, mit vielen Tälern und einem konstanten Anstieg bietet auch viele Unterschiede.

Wollten wir uns einen Überblick über ein Attribut, also eine der zuvor erläuterten Zahlen, von allen Umkreisen verschaffen, können wir aus dem Menü mit “Vektor -> Analyse-Werkzeuge -> Grundstatistik ..” grundlegende statistische Angaben für das jeweils gewählte Attribut abfragen. Das Ergebnis wird auf Englisch im aktiven Fenster unter dem Reiter “Protokoll” ausgegeben, dazu auf Deutsch in einer hier verlinkten HTML-Datei. Hier begegnen uns wie zuvor die “üblichen Verdächtigen”, dazu einige weitere Angaben:

- Anzahl (count): Anzahl der betroffenen Datensätze
- Eindeutige Werte (unique): Anzahl der unterschiedlichen Werte
- Fehlende Leerwerte (NULL) (empty): Anzahl von Zeilen ohne Information
- Minimalwert (min): der kleinste Wert in einer Zeile
- Maximalwert (max): der größte Wert in einer Zeile

- Bereich (range): Differenz zwischen den beiden vorgenannten Min. und Max
- Summe (sum): Summe der Werte
- Mittelwert (mean): der Durchschnitt der gültigen Werte
- Median (median): der Wert in der Mitte der aufsteigend sortierten Werte
- Standardabweichung (std\_dev): der Durchschnitt der Differenz jedes Wertes zum Mittelwert aller Werte
- Variationskoeffizient (cv): Standardabweichung dividiert durch den Mittelwert. Durch diese Normierung ist dieses Maß unabhängig von der Größe der Werte. Ist der Wert größer als 1 ist die Streuung größer als der Mittelwert.
- Minderheit (minority): am seltensten vorkommender Wert
- Mehrheit (majority): am häufigsten auftretender Wert
- Erstes Viertel (firstquartile): der 25%-Grenzwert der aufsteigend sortierten Werte.
- Drittes Viertel (thirdquartile): der 75%-Grenzwert der aufsteigend sortierten Werte.
- Interquartilabstand (IQR): Abstand zwischen dem 2. und dem 3. Quartil (Streuungsmaß).

**1.2.2.3 Visualisierung & Überblick** Kartieren Sie einige der vorgenannten Werte (Bereich, Standardabweichung, Variationskoeffizient und Interquartilabstand). Für die Werte von “Bereich” (range) wähle ich die Farbskala “Spectral” und fünf Klassen mit gleicher Anzahl (Quantile). Es verwundert nicht, dass die sehr hohen Werte in der Tramuntana zu finden sind und die niedrigen Werte in den zentralen Ebenen von Mallorca.

Der Wechsel auf die Werte der Standardabweichung (stdev) bei Beibehaltung der grafischen Parameter verändert das Bild nur ein wenig, am ehesten noch im mittleren Höhenbereich. Das Problem sind die Klassengrenzen. Betrachten Sie die Histogramme für beide Attribute und verschieben Sie die Grenzwerte nach ihrer Interpretation der “Gipfel” und “Täler” in der Verteilung. Mit Blick auf die vorangehende Erläuterung scheint Folgendes sichtbar: einzelne Gipfel verzerren das Bild. Aus diesem Grund tendiere ich zur Standardabweichung als Indikator einer allgemeinen Reliefenergie, die allgemein üblichen Maximaldifferenz scheint mir für unsere Frage weniger geeignet.

### 1.2.3 Exponiert oder im Tal?

Ich vergleiche als erstes den Höhenwert der Position des Talaiot mit dem Median des Umfeldes ( $z\_m - s\_median$ ). Auch für diesen Wert erstelle ich eine abgestufte Symbologie mit fünf Klassen, kontrolliere direkt das Histogramm und verschiebe die Grenzwerte entsprechend den erkennbaren Abschnitten der Verteilung. Die daraus resultierende Karte zeigt keine räumliche Differenzierung, lediglich die Talaiots im Mittelfeld ( $\pm 6$  m zum Median) liegen sicher aufgrund der fehlenden Reliefenergie in den Ebenen.

Als nächstes möchte ich die Höhenlage des Talaiot mit dem oberen Bereich der umliegenden Höhenwerte vergleichen. Eigentlich würde ich gerne das 3. Quartil (75%) als Schwellwert nehmen, aber der wird mir von der Zonenstatistik leider nicht geboten. Wir sollten also einen Weg finden, ohne das *tool* Zonenstatistik rechnen zu können. Ich helfe mir, indem ich die Standardabweichung zum Mittelwert als dem zugehörigen Zentralwert addiere ( $mean + stdev$ ) und diese Summe dann vom Höhenwert des Talaiot abziehe ( $z\_m - (mean + stdev)$ ). Auch hier wähle ich zur Darstellung in der Karte eine abgestufte Symbologie mit fünf Klassen die ich mir im Histogramm direkt ansehe und die Grenzen entsprechend der sichtbaren Verteilung anpasse. Im Ergebnis sehe ich in der Karte ein breit gestreutes Mittelfeld und die “niedrigen” Lagen finden sich aufgrund der Topographie natürlich in der Tramuntana. Eher unerwartet ist die recht gleichmäßige Verteilung der exponierten Talaiots, die oberhalb der oberen Standardabweichung liegen. Hier sollten wir zuerst das statistische Ergebnis zur Lage im Gelände individuell prüfen (so viele sind es nicht) und uns dann die weiteren Eigenschaften dieser Talaiots ansehen. Liegen hier besonders monumentale Konstruktionen vor, die eine Deutung als Standorte von zentraler Bedeutung bestärken?

### 1.2.4 Puffer in SpatiaLite

Leider ist es nicht möglich, dynamische Geometrien in einer *spatial view* zu verarbeiten. Da hierbei neue Geometrien erstellt werden, kann für diese natürlich nicht auf ein originäres Objekt wie bei den *spatial views* verwiesen werden. Für die Puffer zu unseren Punkten müssen wir einen anderen Weg beschreiten. Die mit der SQL-Anweisung zusammengestellten Daten werden einfach in eine neue Tabelle geschrieben. Die Syntax hierzu ähnelt derjenigen der zuvor erstellten Sichten (*views*).

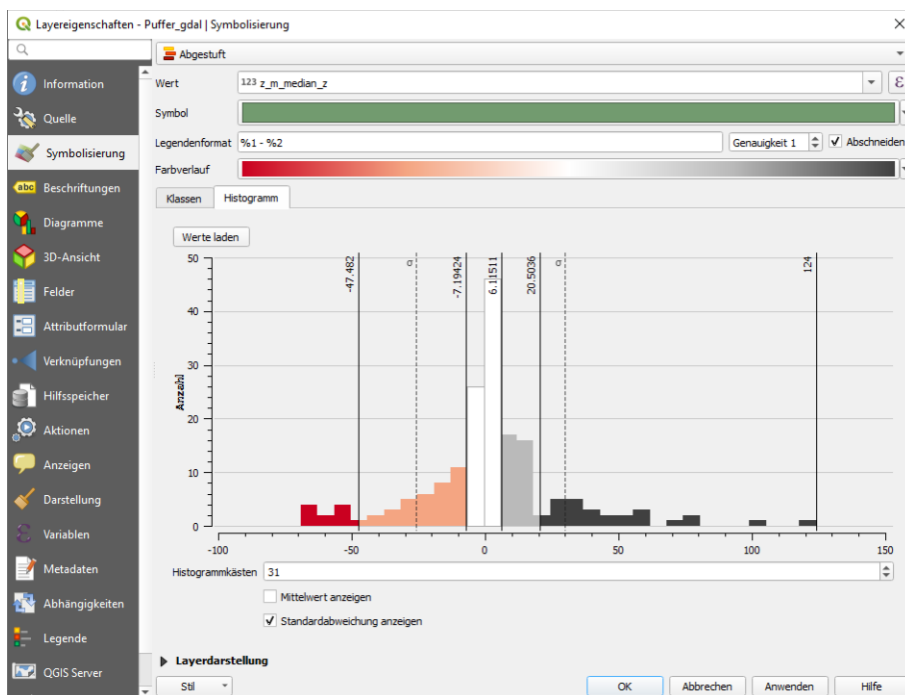


Abbildung 2: Histogramm der Differenz von der Lage des Talaiot zum Median des Umfeld ( $r = 1$  km)

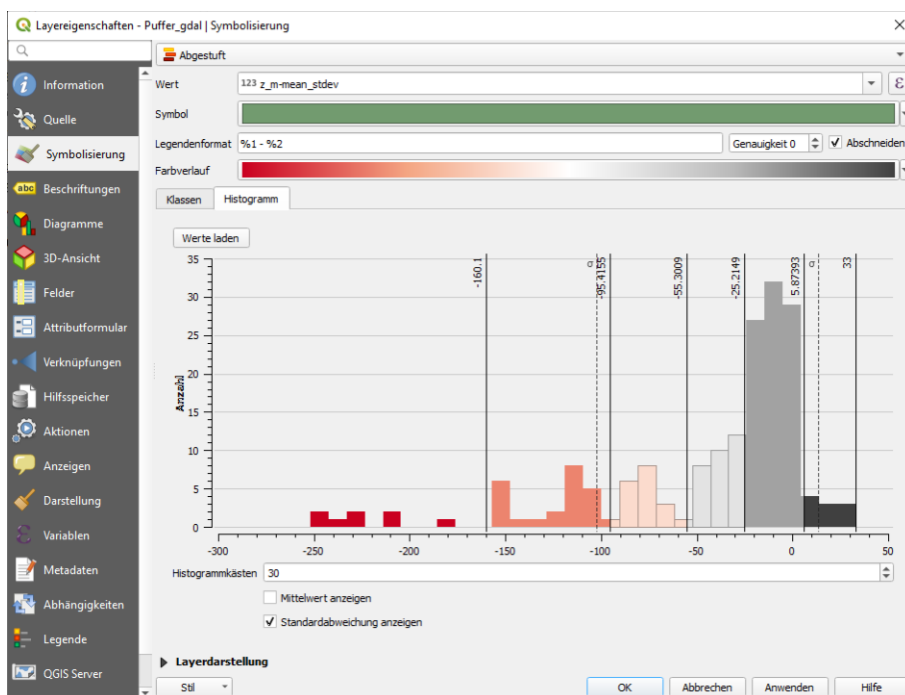


Abbildung 3: Histogramm der Differenz von Siedlungslage zur oberen Standardabweichung des Umfeld ( $r=1$ km)

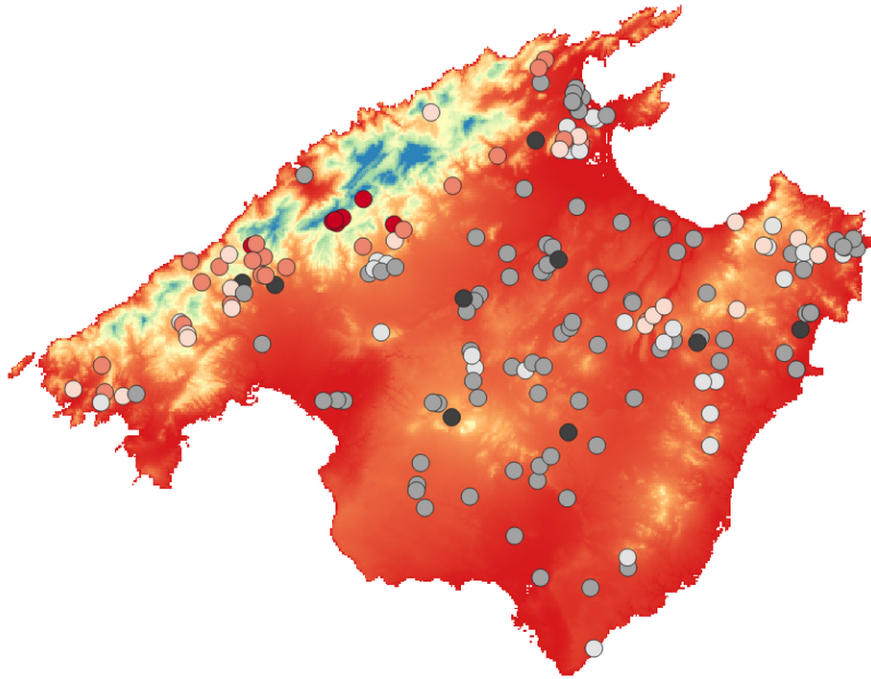


Abbildung 4: Karte der Differenz von Siedlungslage zur oberen Standardabweichung des Umfeld ( $r=1\text{km}$ )

```
create table talaiots_1km as
SELECT t.pk_uid, t.nr, t.nombre_yac, t.z_m,
ST_Buffer(t.geom, 1000) as geom
FROM "talaiots" as t;
```

Im Ergebnis haben Sie eine Tabelle mit einer Geometriespalte, die aber noch nicht als Geometrie in der Datenbank registriert ist. Dies ist in der SpatialLite GUI aber sehr einfach möglich. Öffnen in der Tabelle links im Hierarchiebaum das Kontextmenü zur Geometriespalte “r1km” und wählen Sie “recover geometry column”. Im folgenden Fenster müssen Sie die Parameter der Geometrie angeben: SRID (EPSG-Code), Geometrietyp und Dimension. Für den soeben erstellten Puffer sind dies: 25831, POLYGON, XY. Danach stehen die Daten als Geometrietabelle zur Verfügung.

Woher Sie diese Informationen bekommen? Die SRID entspricht natürlich der dem Ausgangsdatensatz. Dann gibt es mehrere Optionen:

- Sie fragen die Geometrie als Text ab und prüfen eigenständig.
- Sie fragen die Informationen direkt aus der Geometrie ab.

Führen Sie dazu die folgenden Abfragen einzeln aus.

```
select st_astext(geom) from talaiots_1km limit 1;
select geometrytype(geom) from talaiots_1km limit 1;
select coorddimension(geom) from talaiots_1km limit 1;
select srid(geom) from talaiots_1km limit 1;
```

Oder Sie wählen aus dem Kontextmenü zum Feld “geom” die Option “Check geometries”, wodurch folgende etwas komplexere Abfrage ausgeführt wird:

```
SELECT Count(*), GeometryType("geom"), Srid("geom"), CoordDimension("geom")
FROM "talaiots_1km"
GROUP BY 2, 3, 4
```

Das ist soweit nachvollziehbar, interessant ist die Abkürzung bei der Gruppierung durch den Zähler für die entsprechenden Spalten der Abfrage.

Schauen Sie sich die Rundung der Puffer bei einem sehr kleinen Maßstab an und vergleichen Sie diese mit den bereits erstellten. Führen Sie nachfolgend die Zonenstatistik in QGIS erneut aus.

### 1.3 Thiessen-Polygon, Voronoi-Diagramm und Interpolation

Thiessen-Polygone (alternativ Delaunay-Triangulation) sind eine Dreiecksvermaschung, bei der innerhalb des Umkreises zum gebildeten Dreieck kein weiterer Punkt liegt. Die Fundplätze bilden die Knoten (Punkte) der Kanten (Linien). Um die Fundplätze in das Zentrum einer Fläche zu bringen und so die Informationen des Punktes in die umliegende Fläche zu übertragen, wird auf jeder Kante die Mittelsenkrechte gebildet und bis zum Schnittpunkt mit einer anderen Mittelsenkrechten verlängert. Es entsteht ein dualer Graph, der durch erneutes Bilden der Mittelsenkrechten wieder zum Ausgangsmodell führt.

Bei diesem Verfahren gibt es methodische Probleme, zu denen noch spezifische der Archäologie zu ergänzen sind. Das grundlegende archäologische Problem gleich vorweg: Im Unterschied zu strategisch verteilten (Zufall, Raster etc) und erhobenen Messwerten in der Geographie sind archäologische Nachweis-Daten durch viele Faktoren beeinflusst: Aktivität von Ehrenamtlichen, Erhaltung, Vegetation, Landschaftsnutzung etc. Wir müssen also mit dem Fehlen relevanter Informationen rechnen.

Ein methodisches Problem der Triangulation ist der Randeffect des Netzes. Da hier die jeweiligen Nachbarn fehlen, werden besonders lange Strecken gebildet. Auch ignoriert die geradlinige Verbindung zwischen den Punkten weitere beeinflussende Faktoren der unterstellten räumlichen Nähe: Gebirgszüge, Steilhänge, Flüsse etc. Wird in dem abgeleitete Voronoi-Diagramm mit dem Punkt im Zentrum die Information unverändert vom Punkt auf die Fläche übertragen, entstehen an den Grenzen zu den benachbarten Feldern Sprünge in den Information statt fließender Übergänge. Aufgrund dieser Nachteile sind zahlreiche weitere Interpolationsverfahren entwickelt worden. Ein weit verbreitetes ist das *inverse distance weighting* (IDW). Die hierbei gewählte Gewichtung muss angemessen sein und kann durch eine Potenz variiert werden. Das in der Geographie ebenfalls gerne genutzte *kriging* bietet noch mehr Justierungsmöglichkeiten, wodurch sich z.B. bei Höhenmodellen sehr gute Approximationen an die Wirklichkeit erzielen lassen. Diese letztgenannten Berechnungen werden nicht nur rechnerisch komplexer, sondern vor allem im Rahmen einer auf archäologischen Argumenten beruhenden Modellbildung schwieriger begründbar. Für die Anwendung in der Archäologie lesen Sie hierzu zum Einstieg bitte (Herzog, 2007). Jüngere Darstellung mit komplexen, interdisziplinär ausgearbeiteten und begründeten Modellierungen von Informationen in die Fläche finden Sie u.a. bei Daniel Knitter (Knitter u. a., 2019; Nakoinz und Knitter, 2016b).

### 1.4 Thiessen-Polygone und triangulierte Distanzen

#### 1.4.1 Vorgehen in QGIS

Wie zu Beginn dieses Kapitels erwähnt sind die triangulierte Distanzen ein schlichtes aber durchaus interessantes Maß (Siegmond, 1992; Zimmermann, 1992). Wir können dies bei unserem vorliegenden Datensatz zu den Talaiots einmal durchspielen. Dabei haben wir aus der archäologischen Perspektive bestimmte Vorteile. Durch die klare Begrenzung des Untersuchungsgebietes als Insel sind Kanteneffekte zwar zu erwarten, entsprechen aber der unabdingbaren Realität auch wenn die “Wege” an einige Stellen über das Wasser führen. Der monumentale Charakter und der relativ späte Landesausbau mit dem vorhandenen Bewusstsein für diesen Denkmaltyp hat zu einer erwartbar guten Kenntnis (nicht zwingend Erhaltung) geführt. Wir können zudem mehrstufig Arbeiten und die gesamte Insel mit den Landschaftsgebieten (*comarcas*) oder der zuvor herausgestellten Lage vergleichen.

Der Weg ist mehrstufig. Wenn Sie ausschließlich in QGIS arbeiten wollen, müssen Sie folgende Schritte durchlaufen:

1. Suchen Sie in der *tool box* die “Delaunay-Triangulation”, der Eingabelayer ist “talaiots” und speichern Sie auch zukünftig das jeweilige Ergebnis temporär. Anmerkung: In einer älteren Version von QGIS wurde die *spatial view* nicht erkannt, alternativ konnte die Funktion *v.delaunay* der Grass-Erweiterung genutzt werden, bei der wiederum die Option Linien ausgeben” scheitert.
2. Polygone zu Linien: Im Ergebnis haben Sie eine Polylinie um jedes Dreieck.
3. Linien sprengen: Im Ergebnis haben Sie einzelne Linien. Fast jede Linie im Netz ist aber doppelt vorhanden, für jedes angrenzende Dreieck einmal.
4. Doppelte Geometrien löschen: Sie erhalten alle Distanzen einmalig.
5. Ergänzen Sie mit dem Feldrechner eine Spalte für die Linienlänge (\$length).
6. Exportieren Sie alle Längenmaße für eine statistische Analyse, z.B. in R.



Für eine erste Übersicht zu den eben ermittelten Streckenlängen öffnen Sie die Symbologie, wählen die abgestufte Darstellung, bestätigen [Klassifizieren] für eine Farbzuzuweisung der Bereiche, wechseln zum Reiter “Histogramm” und wählen abschließend [Werte laden]. Mit der Zahl hinter “Histogrammkästchen” können Sie die Anzahl der Balken für eine feinere Auflösung erhöhen. Einige Längen des Randbereiches sind sehr groß und beeinträchtigen die Darstellung. Wenn Sie nur den unteren Bereich der Längen darstellen wollen wählen sie bei “Wert” das [E] und fügen dort folgende Syntax ein: `if ("1"<20000, "1", '')` (Wenn die Länge < 20 km ist dann gib die Länge zurück, sonst Nichts). Wiederholen Sie [Klassifizieren] und betrachten erneut das Histogramm nach [Werte laden].

Wir sehen einen ersten Gipfel der Verteilung um 1,5 km, einen zweiten um 3 km und dann noch Bereiche von 4 km bis 8 km und um 10 km. Danach löst sich die Verteilung langsam auf. Wirklich Plakativ ist das Ergebnis nicht, doch die Dopplung von 1,5 km zu 3 km ist auffällig und alles jenseits von 11 km scheint eine “besonders” große Distanz zu sein. Sehen Sie sich die Distanze in der Karte an scheinen die beiden kurzen Distanzbereiche um 1,5 km und um 3 km eine Gruppe zu beschreiben. Die Distanzen von 4 bis 8 km trennen die Talaiots dieser Gruppen von den Talaiots in den anderen Gruppen. An diesem Punkt könnten diverse Hypothese formuliert und geprüft werden. Als erstes natürlich ob es neben der räumlichen Nähe innerhalb der Distanzgruppen auch einen weiteren Zusammenhang gibt: Datierung, Funde, Konstruktion etc.

#### 1.4.2 Vorgehen in SpatiaLite

Mit SQL wird der Weg wegen der komplexen Syntax in zwei Schritten durchgeführt. Nutzen Sie die SpatiaLite GUI und führen Sie folgende SQL-Anweisung aus:

```
create table talaiot_dist as
select st_DelaunayTriangulation(st_collect(geom),1) as geom
from talaiots;
```

Die Rückgabe der Abfrage ist eine Tabelle mit einem Objekt vom Typ *Multilinestring*. Die verwendete Funktion `st_DelaunayTriangulation` erwartet ein Geometrieobjekt, dieses wird mit der Funktion `st_collect()` aus allen Geometrien in der Spalte ‘geom’ von ‘talaiots’ erstellt. Die Funktion `st_DelaunayTriangulation` gibt mit der Option *edges\_only* die Kanten statt der Polygone aus. Die Übergabe dieses Parameters erfolgt ausschließlich durch 0 (falsch) oder 1 (wahr).

Die neue Tabelle ‘talaiot\_dist’ enthält genau eine Geometrie (geom) vom Typ *Multilinestring*, also eine Geometriekollektion, für das wir die Anzahl der enthaltenen Elemente abfragen.

Die einfache Variante nutzt die implementierte virtuelle Tabelle *ElementaryGeometries* (s. [SpatiaLite](#)).

```
select st_length(geometry) from ElementaryGeometries
where f_table_name = 'talaiot_dist' and origin_rowid=1;
```

Die umständlichere Alternative arbeitet mit Rekursion in SQLite. Das ist auf den ersten Blick irritierend aber nicht wirklich kompliziert. Zuerts eine Abfrage für die Anzahl der Strecken in der Multi-Geometrie.

```
select NumGeometries(geom)
from talaiot_dist;
```

Die nachfolgende Abfrage wird als Sicht gespeichert und hat zwei Abschnitte, die mit *WITH RECURSIVE* zusammengefasst werden. Im ersten Abschnitt gibt es zudem noch eine Unterabfrage.

```
create view talaiot_dist_length as
WITH RECURSIVE
cnt(x) AS (
  SELECT 1
  UNION ALL
  SELECT x+1 FROM cnt
  LIMIT (select NumGeometries(geom) from talaiot_dist)
)
SELECT st_length(GeometryN(geom, x)) as l FROM cnt, talaiot_dist;
```

Der erste Teil erzeugt einen Zähler ‘cnt’ (count) für ein Element ‘x’, diesem wird mit *as* eine Schleife in der Form einer *UNION ALL*-Abfrage zugewiesen. Das erste Element dieser *UNION-ALL* wird mit ‘Select



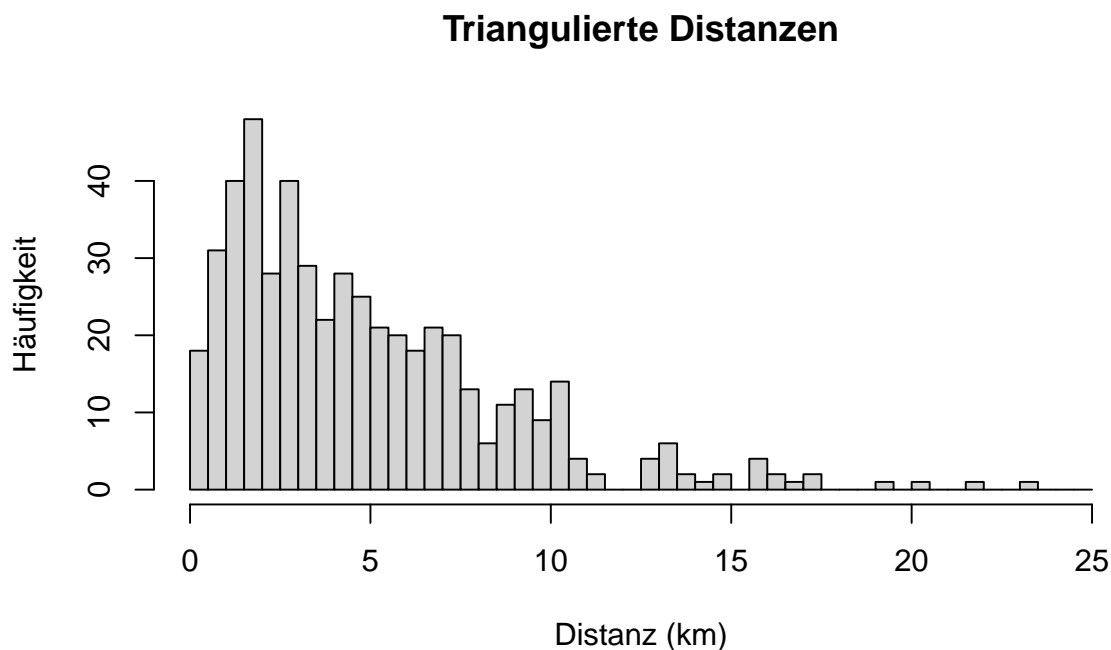
1' gesetzt, eine weitere *select*-Abfrage addiert für jedes weitere x stets +1. Mit *LIMIT* wird ein Grenzwert gesetzt der diesen Vorgang beendet. Dieses *LIMIT* wird durch eine Unterabfrage ermittelt, die mit der Funktion *NumGeometries()* die Anzahl der Elemente in dem Multistring der Triangulation abfragt (s.o.).

Das zweite Element ist eine Schlichte Abfrage über zwei Tabellen, den erzeugten Zähler cnt mit dem Element x und die Tabelle mit der einen Geometrie vom Typ *Multilinestring*. Abgefragt wird nun die Länge (*st\_length()*) für jedes x-te Element des *Multilinestring* (*GeometryN()*). Nach der Aktualisierung der Datenbank (*refresh*) steht die Sicht zur Verfügung, kann über das Kontextmenü z.B. in eine CSV-Datei exportiert und für eine statistische Analyse verwendet werden.

Wenn Ihnen das zu kompliziert ist und Sie schnell zum Ziel kommen wollen können Sie natürlich die Geometrie für QGIS mit "recover geometry column" (s.o.) lesbar machen, dort einfügen und mit der Funktion "Linien sprengen" einzelne Linien erzeugen.

### 1.4.3 Deutung des Ergebnisses

```
hist((1/1000)[l<25000], breaks=seq(0,25,0.5),
     main='Triangulierte Distanzen', xlab = 'Distanz (km)', ylab = 'Häufigkeit')
```



Der vorangehende Code rechnet die Meter-Distanzen in Kilometer um (1/1000), zuvor wird auf Distanzen unter 25000 m gefiltert ( $[l < 25000]$ ). Die Schrittweite (*breaks*) wird als Sequenz (*seq()*) von 0 bis 25 km mit Intervallen von 0,5 km festgelegt. Abschließend werden noch der Titel und die Beschriftungen der x- und y-Achse hinzugefügt. Vergrößern wir das Intervall auf 1 km entfällt die visuelle Trennung zwischen 1,5 km und 3 km und es bleibt eine leichter Absatz im Verlauf mit geringer Häufung von Distanzen bei 5 km. Es folgend unscharfe Häufungen bei den Vielfachen von 5 km. Das Histogramm weist auf Agglomerationen von Talaiots in einem Abstand von bis zu 3 km hin, diese könnten sich dann in Abständen von ca. 5 km wiederholen. Dazwischen klaffen aber auch Lücken, die wegen des Vielfachen von 5 km möglicherweise durch fehlende Gruppen verursacht worden sein könnten. Letzteres ist lediglich aus der Wiederholung der Distanz abgeleitet und kann auch andere Gründe haben.

## Literatur

Christaller, W., 1933. Die zentralen Orte in Süddeutschland : eine ökonomisch-geographische Untersuchung über die Gesetzmäßigkeit der Verbreitung und Entwicklung der Siedlungen mit städtischen Funktionen. G. Fischer, Jena.

- Cziesla, E., 1990. Siedlungsdynamik auf steinzeitlichen Fundplätzen. Methodische Aspekte zur Analyse latenter Strukturen, *Studies in Modern Archaeology*. Bonn.
- Haggett, P., 1973. Einführung in die kultur- und sozialgeographische Regionalanalyse. de Gruyter, Berlin.
- Herzog, I., 2013. Berechnung von optimalen Wegen am Beispiel der Zeitstraße. *Archäologische Informationen* 31, 67–96. <https://doi.org/10.11588/ai.2008.1&2.11119>
- Herzog, I., 2007. Simulationsexperimente zur Siedlungsdichte, in: Burmeister, H., Derks, S., Richthofen, J. von (Hrsg.), *Zweiundvierzig. Festschrift für Michael Gebühr zum 65. Geburtstag*, Internationale Archäologie : Studia honoraria. Marie Leidorf, Rahden/Westf., S. 3–14.
- Jankuhn, H., 1977. Einführung in die Siedlungsarchäologie. Berlin.
- Knitter, D., Günther, G., Hamer, W.B., Keßler, T., Seguin, J., Unkel, I., Weiberg, E., Duttmann, R., Nakoinz, O., 2019. Land use patterns and climate change - a modeled scenario of the Late Bronze Age in Southern Greece. *Environmental Research Letters* 14, 125003. <https://doi.org/10.1088/1748-9326/ab5126>
- Kunow, J., Müller, J., Schopper, F. (Hrsg.), 2007. Archäoprognose Brandenburg II, *Forschungen zur Archäologie im Land Brandenburg*. Brandenburgisches Landesamt für Denkmalpflege und Archäologisches Landesmuseum, Wünsdorf.
- Linke, W., 1976. Frühes Bauerntum und geographische Umwelt. Eine historisch-geographische Untersuchung des Früh- und Mittelneolithikums westfälischer und nordhessischer Bördenlandschaften, *Bochumer Geographische Arbeiten*. Paderborn.
- Nakoinz, O., Knitter, D., 2016a. *Modelling Human Behaviour in Landscapes: Basic Concepts and Modelling Elements, Quantitative Archaeology and Archaeological Modelling*. Springer International Publishing, Cham.
- Nakoinz, O., Knitter, D., 2016b. Regression and Interpolation, in: Nakoinz, O., Knitter, D. (Hrsg.), *Modelling Human Behaviour in Landscapes: Basic Concepts and Modelling Elements, Quantitative Archaeology and Archaeological Modelling*. Springer International Publishing, Cham, S. 87–105. [https://doi.org/10.1007/978-3-319-29538-1\\_5](https://doi.org/10.1007/978-3-319-29538-1_5)
- Siegmund, F., 1992. Triangulation als Methode zur Aufdeckung frühgeschichtlicher Siedlungsmuster und zur Schätzung von Siedlungsdichte. *Archäologische Informationen. Mitteilungen zur Ur- und Frühgeschichte* 15, 113–116.
- Zimmermann, A., 1992. Tesselierung und Triangulation als Techniken zur Bestimmung archäologischer Funddichte. *Archäologische Informationen. Mitteilungen zur Ur- und Frühgeschichte* 15, 107–112.
- Zimmermann, A., Richter, J., Frank, T., Wendt, K.-P., 2004. Landschaftsarchäologie II - Überlegungen zu Prinzipien einer Landschaftsarchäologie. *Bericht der Römisch-Germanischen Kommission* 35, 37–95.
- Zimmermann, A., Wendt, K.-P., 2003. Wie viele Bandkeramiker lebten 5.060 v. Chr.? Techniken Geographischer Informationssysteme zum Schätzen von Bevölkerungsdichten. *Archäologische Informationen* 26, 491–497.