# Optimale Datenerhebung mit minimalem Aufwand

Jörg Linstädter, Jürgen Richter, Anja Linstädter

Archäologische Informationen 25, 1-22

2001

**Universität zu Köln**
**Institut für Ur- und Frühgeschichte**
Weyertal 125
50923 Köln
www.ufg.uni-koeln.de

# Optimale Datenerhebung mit minimalem Aufwand

*Jörg Linstädter, Jürgen Richter und Anja Linstädter*

## Einführung

*Archäologische Datensammlung – ein Kosten-Nutzen-Problem*

In der vorliegenden Studie geht es um ein einfaches Verfahren, den Aufwand archäologischer Datenerhebungen erheblich zu verringern, ohne die Validität der möglichen Aussagen zu beeinträchtigen.

Optimale Datenerhebung mit geringstmöglichem Aufwand ist ein alter Archäologentraum. Um ihn zu verwirklichen, werden in der Regel repräsentative Stichproben gezogen. Repräsentativität ist, vereinfacht gesagt, dann gegeben, wenn die Auswertung einer Stichprobe (zum Beispiel einer begrenzten Anzahl archäologischer Objekte aus einem großen Inventar) dieselbe Aussage ergibt wie die Auswertung aller Objekte des gesamten Inventars. Sobald eine Stichprobe so umfangreich ist, daß sie repräsentativ für die Grundgesamtheit ist, kann die Aufnahme weiterer Daten unterbleiben.

In dem hier vorgestellten Verfahren wird schon während der Datenaufnahme der optimale Zeitpunkt ermittelt, die Datenaufnahme abzuschließen. Die Grundidee ist dabei, mit der Datenaufnahme einer kleinen Stichprobe zu beginnen und diese sofort auszuwerten, dann weitere Daten hinzuzufügen und die vergrößerte Stichprobe ein weiteres Mal auszuwerten – so lange, bis die Ergebnisse keine Veränderung mehr zeigen. Während der Datenaufnahme findet quasi parallel ein ständiger Auswertungsprozeß statt, dessen Verlauf jederzeit per Tastendruck abrufbar ist. Die ständige Auswertung liefert die Argumente zur Stichprobenbegrenzung, die am Ende anzeigen, wann eine Analyse weiterer Objekte nicht mehr nötig ist.

Die grundsätzlichen Vorteile der Stichprobenbildung sind in der Archäologie bereits seit Jahrzehnten bekannt und akzeptiert (z.B. IHM et al. 1978, 8-10; ZIMMERMANN 2001, 132-135). Die Bildung von Stichproben verhindert, daß – aus statistischer Sicht – unnötige Datenmengen aufgenommen werden.

Das Anliegen dieses Artikels ist nicht die Art und Weise der Stichprobengewinnung (Sampling-*Strategie*), sondern vielmehr die Frage, wie viele Datensätze für ein Studiendesign notwendig sind (Sampling-*Intensität*; TRAXLER 1997, 58 ff.). Dabei wird die Größe einer Stichprobe nicht anhand allgemeiner Richtwerte festgelegt, sondern empirisch und für den jeweiligen Einzelfall, das heißt für eine konkrete Fragestellung an eine konkrete Datensammlung. Dazu ist eine Beispiel-Datenbank erstellt worden, die auf eine einfache Merkmalsaufnahme an archäologischen Steinartefakten ausgerichtet ist, aber leicht anderen Inhalten angepaßt werden kann. Die für das Verfahren der Stichprobenbegrenzung notwendigen statistischen Methoden sind dieser Beispiel-Datenbank als Auswertungsroutinen zugeordnet. Mit ihrer Hilfe erfolgt dann eine automatisierte Überprüfung der Stichprobengröße parallel zur Datenerfassung. Hinter dem Verfahren steckt also die Vorstellung eines "selbstreferentiellen Datenaufnahme-Roboters", der das zeitaufwändige Aufnahme- und Auswertungsverfahren für den Archäologen optimiert.

*Anwendungsbeispiel*

Die Idee, während der Datenaufnahme parallel auszuwerten und damit die Stichprobengröße zu begrenzen, wird am Kölner Institut für Ur- und Frühgeschichte schon seit einigen Jahren verfolgt. Bei ersten Versuchen mit einem rund 8.000 Artefakte umfassenden Inventar einfacher Steintechnologie, das aus dem Nordwest-Sudan um 3000 v.Chr. stammt (Wadi Shaw 83/120), nahmen einige erhobene Merkmale schon nach wenigen Dutzenden untersuchter Objekte stabile Mittelwerte an. Weitere Versuche mit ähnlichen Ergebnissen mündeten in die Überlegung, eine Methodik zu entwickeln, die ein geregeltes und wiederholbares Verfahren der parallelen Datenauswertung gestattet.

Die arbeitstechnischen Vorteile eines solchen Verfahrens zeigen auch Untersuchungen von Werner Schön im Wadi el Akhdar (Gilf Kebir, Südwest-Ägypten) im Rahmen des DFG-Projektes B.O.S. (*"Besiedlungsgeschichte der Ostsahara"*) (SCHÖN 1996, 56). Schön überprüfte die Aussagefähigkeit einiger Inventar-Stichproben durch den Vergleich mit Inventaren, von denen Vollerhebungen vorlagen, und die er nachträglich in Stichproben zerlegte. Er kam dabei zu dem Ergebnis, daß eine zehnprozentige Stichprobe zu ver-

gleichbaren Ergebnissen geführt hätte wie die entsprechende Vollerhebung. Als im Rahmen des Sonderforschungsbereiches 389 "ACACIA" (*"Arid Climate, Adaptation and Cultural Innovation in Africa"*) an der Universität zu Köln die Bearbeitung der umfangreichen Steininventare aus dem Wadi Bakht (Gilf Kebir, Südwest-Ägypten) anstand, ergab sich die Gelegenheit, die bisherigen Ansätze zu überprüfen und statistisch abzusichern. Das Ziel der merkmalanalytischen Aufnahme von Steinartefakten im Wadi Bakht ist die qualitative und quantitative Beschreibung und gegebenenfalls Gliederung aller ausgegrabenen Inventare. Aus den Erfahrungen der bisherigen Bearbeitung erwies sich das Abgreifen einiger bestimmter Variablen an den Grundformen als besonders hilfreich (Tabelle). Eine Möglichkeit zur Trennung technologisch unterschiedlicher Komplexe böte im Rahmen einer explorativen Datenanalyse beispielsweise eine bimodale Verteilung metrisch skalierter Merkmale.

Allein die Fundstelle Wadi el Bakht 82/21 lieferte allerdings 30.000 Steinartefakte (LINSTÄDTER 1999), und schon die Untersuchung einer zehnprozentigen Stichprobe hätte hier einen erheblichen Aufwand bedeutet. Das Prinzip der schrittweisen Stichprobenerweiterung bei paralleler Auswertung sollte deshalb die gesamte Datenerhebung im Wadi Bakht steuern. Hierzu wurde auf ACCESS-Basis eine Datenbank-Konstellation erstellt. Das statistische Prinzip, das der hier vorgestellten Methode zugrundeliegt, wird – wie sich herausstellte – in der Ökologie als Methode der "Laufenden Mittelwerte" seit langem angewandt und verdankt seine Übertragung auf archäologische Fragestellungen somit auch der interdisziplinären Struktur des Sonderforschungsbereiches 389 ACACIA, in dem das Autorenteam aus zwei Archäologen und einer Vegetationsökologin zusammenfand.

## Prinzip des Sampling

### *Sampling-Strategie*

Bei der Erhebung empirischer Daten stehen dem Wissenschaftler grundsätzlich drei Wege offen: (1) die Vollerhebung, (2) die subjektive Auswahl und (3) die randomisierte zufallsgesteuerte Auswahl.

Eine Vollerhebung aller zugänglichen Daten ist nur in den wenigsten Fällen möglich und sinnvoll. Bei der subjektiven Auswahl wird der Studienaufwand gegenüber der Vollerhebung verringert, allerdings sind die Daten einer so erhobenen Stichprobe nicht zwangsläufig repräsentativ für die Grundgesamtheit. Erst bei der zufälligen Auswahl von Objekten aus der Grundgesamtheit kann mittels schließender (analytischer) Statistik von den Kennwerten der Stichprobe auf die entsprechenden Parameter der Grundgesamtheit geschlossen werden (MÜHLENBERG 1993, 240 ff.; ZÖFEL 2001, 59 ff.). Als Grundgesamtheit bezeichnet man dabei die fiktive Menge aller theoretisch untersuchbaren Objekte, die ein gemeinsames Merkmal aufweisen.

Bei einer archäologischen Datenerhebung treten meist zwei Zeitpunkte auf, zu denen implizit oder explizit eine Stichprobe gezogen wird. Schon vor und während der Ausgrabung geschieht dies meist bei der Festlegung der Lage und des Umrisses der Grabungsfläche (vgl. die Diskussion in ZIMMERMANN 2001, 131 ff.) und später, zu Beginn der Bearbeitung des Inventars, bei einer Auswahl aus dem vorliegenden Material. Häufig wird übersehen, daß bereits die Auswahl der Grabungsfläche mit dem Ziehen einer Stichprobe gleichzusetzen ist. Auch eine Vollerhebung des geborgenen Materials garantiert daher keinesfalls repräsentative Daten. Vielmehr ist es in jedem Fall sinnvoll, die Aussagekraft der Daten mit statistischen Methoden zu überprüfen.

Dem hier vorzustellenden Verfahren zur statistischen Überprüfung und zur Begrenzung der Stichprobengröße geht eine Materialaufnahme voraus, die sich standardisierter Methoden der Archäologie bedient: Zu Beginn jeder Fundplatzbearbeitung wird das gesamte Material nach Stückzahl und Rohmaterialien pro Grabungseinheit erfaßt und kartiert. Anschließend wird eine Sampling-Strategie zur Materialaufnahme ausgearbeitet. In unserem Beispiel werden Steinartefaktinventare aus je einer Viertelquadratmeter-Fläche als Stichproben betrachtet. Da nicht die Sampling-Strategie, sondern die Sampling-Intensität Gegenstand unseres Beitrags sein soll, haben wir uns entschlossen, die Flächen nach dem Zufallsprinzip auszuwählen (randomisierte Stichprobenauswahl). Die hier gewählte Sampling-Strategie ist nicht die einzig sinnvolle oder mögliche. So wird man sicherlich eine geschichtete Zufallsstichprobe bevorzugen, sobald deutlich unterscheidbare Konzentrationen, Befunde, Bereiche oder Horizonte innerhalb eines Fundplatzes charakterisiert werden sollen. Ganz unabhängig davon, welche Sampling-Strategie gewählt wurde, steht der Archäologe im nächsten Schritt vor der Aufgabe, die optimale Sampling-Intensität zu finden. Um diesen Arbeitsschritt geht es im folgenden.

### *Sampling-Intensität: Laufende Mittelwerte*

Die Sampling-Intensität wird durch die Größe der Stichprobe bestimmt. Die Stichprobe soll so klein wie möglich, aber so groß wie nötig sein.

Die Bestimmung einer optimalen Stichprobengröße kann auf einer Analyse der "Laufenden Mittel-

| TEILMENGEN | | | | |
|---|---|---|---|---|
| **Alle Grundformen** | **Alle Abschläge und Klingen** | **Vollständig erhaltene Abschläge** | **Vollständig erhaltene Klingen** | **Vollständig erhaltene Abschläge und Klingen** |
| 1. Anteil vollständig erhaltener Stücke | 5. Verhältnis von Abschlägen zu Klingen | 8. Mittelwert und Standardabweichungen der Längen | 12. Mittelwert und Standardabweichungen der Längen | 16. Mittelwert und Standardabweichungen der Längen des Schlagflächenrestes |
| 2. Anteil der Stücke mit Kortex | 6. Häufigkeit eines Accident Siret | 9. Mittelwert und Standardabweichung der Breite | 13. Mittelwert und Standardabweichung der Breite | 17. Mittelwert und Standardabweichungen der Breiten des Schlagflächenrestes |
| 3. Anteil der Abschläge | 7. Häufigkeit dorsaler Reduktion | 10. Mittelwert und Standardabweichung der Dicke | 14. Mittelwert und Standardabweichung der Dicke | 18. Mittelwert und Standardabweichung des QS-Indexes[4] |
| 4. Anteil der Klingen | | 11. Längen-Breiten-Index | 15. Längen-Breiten-Index | |

**Tabelle** Für das Verfahren der "laufenden Mittelwerte" herangezogene Kennwerte. Die "Teilmenge" bezeichnet diejenigen Einzelstücke aus dem Inventar der (nicht-modifizierten) Grundformen, die zur Ermittlung der jeweiligen Kennwerte herangezogen werden. Die Daten zu den Kennwerten 1-7 sind auf einem nominalen Skalenniveau erhoben, die Daten zu den Kennwerten 8-18 auf einem metrischen Skalenniveau.

werte" basieren. Diese Methode wurde unter anderem in der Ökologie für die Festlegung einer sinnvollen Anzahl von Untersuchungsflächen entwickelt (vgl. MUELLER-DOMBOIS & ELLENBERG 1974). Hinter dem Prinzip der Ermittlung laufender Mittelwerte steckt die Idee, so lange zu messen, bis die neu hinzukommenden Mittelwerte den Mittelwert der vorherigen Werte nicht mehr wesentlich verändern (TRAXLER 1997, 70). Diese Methode haben wir in unserer Beispiel-Datenbank direkt umgesetzt.

Die Größe der Stichprobe (d.h. die Sampling-Intensität) hängt prinzipiell von der gewünschten Genauigkeit ab. Je größer die Variabilität der Einzelergebnisse einer Stichprobe ist, desto mehr Einzeldaten müssen erhoben werden, um ein gleich exaktes Ergebnis zu erhalten (MÜHLENBERG 1993, 246 f.). Beeinflußt wird die Schwankung der Kennwerte und damit die Stichprobengröße z.B. von der Funddichte eines Fundplatzes, von der Fundverteilung (heterogen oder homogen) und der Standardisierung der Erfassungsmethode. *"Die Zahl der Stichproben wird immer ein Kompromiß zwischen den zur Verfügung stehenden Mitteln (Zeit, Kosten) und der gewünschten Sensitivität* [...] *sein"* (MÜHLENBERG 1993, 247). Das bedeutet, daß beim Ziehen einer Stichprobe in jedem Fall eine Kosten-Nutzen-Abwägung nötig ist. So muß bei der Methode der "Laufenden Mittelwerte" die Grenze, ab der keine Veränderung des Mittelwertes mehr wahrgenommen wird, vom Bearbeiter willkürlich gewählt werden. Eine solche Grenze kann zum Beispiel so gesetzt werden, daß die Veränderung des Mittelwertes weniger als 5% der nächstgrößeren Teststichprobe (d.h. hier konkret des nächsten Viertelquadrat-Inventars) ausmacht (TRAXLER 1997, 70).

## Fundplatzbearbeitung

### *Datenbankstruktur*

Datenbanken sind komplexe Werkzeuge, die der Datenverwaltung und -analyse dienen und deren Struktur an die Fragestellung und die fachlichen Bedürfnisse einer Arbeit angepaßt werden müssen (BREUER 1997, 91 ff.). Eine ACCESS-Datenbank besteht grundsätzlich aus Formaten zur Dateneingabe und -verwaltung ("Tabellen"), zur Datenanalyse ("Abfragen") sowie zur Automatisierung von Auswertungsroutinen ("Makros"). Auf die genaue Struktur der für die Datenaufnahme und -auswertung entwickelten Datenbank soll im Rahmen dieses Artikels nicht eingegangen werden. Die Datenbank-Struktur einschließlich der Daten aus dem hier vorgestellten Fallbeispiel und der Sampling-Routine können zusammen mit einer ausführlichen Dokumentation kostenlos aus dem Internet unter www.csaro.de heruntergeladen werden.

### *Datenaufnahme*

Sinnvollerweise erfolgt die Datenaufnahme bei einer Fundplatzauswertung in drei aufeinander aufbauenden Schritten. Diese Vorgehensweise ist auch in unserem

**Tabelle: Viertelquadrate (Definition der Grabungseinheiten)**

| Name des Datenfeldes | Inhalt | Beispiel |
|---|---|---|
| **ID** | Laufende Nummer der Grabungseinheit | 31 |
| **Fundplatz** | Bezeichnung des Fundplatzes | 198202101 |
| **X** | Koordinate des Viertelquadratmeter-Mittelpunktes in m: Rechtswert | 50,75 |
| **Y** | Koordinate des Viertelquadratmeter-Mittelpunktes in m: Hochwert | 52,75 |
| **Position** | Laufende Nummer des Arbeitsschrittes während der Ausgrabung | 9 |

**Tabelle: Inventar (Gruppenaufnahme der Grundformen, geordnet nach Rohmaterialklassen)**

| Name des Datenfeldes | Inhalt | Beispiel |
|---|---|---|
| **ID** | Laufende Nummer der Grabungseinheit | 31 |
| **RM** | Zuordnung zu Rohmaterialklasse, Code-Nummer | 15 |
| **Abschläge** | Anzahl der Abschläge | 5 |
| **Klingen** | Anzahl der Klingen | 2 |
| **..... usw.** | (weitere Grundformen) | ... |

**Tabelle: Einzelaufnahme (Individualaufnahme der Formmerkmale, Erhaltungsmerkmale und technischen Merkmale)**

| Name des Datenfeldes | Inhalt | Beispiel |
|---|---|---|
| **FUND** | Laufende Nummer des Fundstückes | 214 |
| **ID** | Laufende Nummer der Grabungseinheit | 31 |
| **RM** | Zuordnung zu Rohmaterialklasse, Code-Nummer | 15 |
| **GF** | Grundform: Abschlag=1; Klinge >15mm=2; Trümmer=3; Klinge<15mm=4; Abspliss=5; | 2 |
| **EH** | Erhaltung: Keine Aussage=99; vollständig=1; Basalfragment=2; Medialfragment=3; Distalfragment=4; Accident Siret=5; | 3 |
| **Kortex** | Kortexanteil: keine Aussage=99; keine Kortex=0; Kortexoberfläche < 25%=1; Kortexoberfläche 25-75%=2; Kortexoberfläche >75%=3; Kortexoberfläche vollständig=4; | 0 |
| **L** | Länge des Artefakts (mm) | 52 |
| **B** | Breite des Artefakts (mm) | 26 |
| **D** | Dicke des Artefakts (mm) | 8 |
| **SFR_L** | Schlagflächenrest-Länge (mm) | 4 |
| **SFR_B** | Schlagflächenrest-Breite (mm) | 2 |
| **..... usw.** | (weitere Merkmale) | ... |

**Abb. 1** Datenbankstruktur mit drei beteiligten Tabellen. In die Tabelle "Viertelquadrate" werden alle Grabungseinheiten des Fundplatzes eingetragen. In die Tabelle "Inventar" werden alle Fundobjekte, gruppiert nach ihrer Rohmaterial- und Grabungseinheiten-Zugehörigkeit eingetragen. In der Tabelle "Einzelaufnahme", die detaillierte Angaben zu den Merkmalen der einzelnen Objekte enthält, bildet je ein Objekt einen Datensatz. Es werden nur so viele Objekte (Datensätze) eingetragen, wie nötig sind, um statistisch repräsentative Ergebnisse zu erzielen. Die optimale Begrenzung dieser Tabelle "Einzelaufnahme" ist das Ziel der vorgeschlagenen Methode zur Bestimmung der Sampling-Intensität.

Beispiel vorgesehen (Abb. 1).

(1) Im ersten Schritt (Tabelle "Viertelquadrate") werden alle Grabungseinheiten einer Grabungsfläche eingegeben. In unserem Fall sind das die Flächen der Viertelquadratmeter jeweils eines stratigraphischen Niveaus. Zu jedem Datensatz gehört z.B. das entsprechende Fundplatzkürzel[1] sowie die exakten Koordinaten des Viertelquadrates innerhalb der Fundfläche.[2]

(2) Es folgt die quantitative Erfassung des kompletten Inventars einer Fundstelle (Tabelle "Inventar"). Im Fallbeispiel dieses Artikels werden nur Steinartefakte durchgezählt.[3] Hierbei werden spezielle Kategorien wie Werkabfälle, Sekundärmodifikationen und Geräte isoliert. Nach Fertigstellung der Tabelle "Inventar" wird unter Berücksichtigung der Fragestellung und der Fundplatz-Konstellationen eine Sampling-Strategie zur weiteren Materialaufnahme erarbeitet (vgl. Abschnitt 2.1). Dabei kann eine Kartierung der Fundverteilungen sinnvoll sein, um befundorientierte Teilmengen des Inventars abzugliedern.

(3) Nun werden aus der gewählten Gesamtheit der zu beprobenden Artefakte (hier: aller nicht-modifizierten Grundformen) Teststichproben gezogen. Bei heterogenen Situationen kann evtl. eine entsprechende Differenzierung in sinnvolle Teilmengen des Inventars und das Ziehen von Teststichproben aus diesen Teil-

mengen erfolgen. Das "Ziehen einer Teststichprobe" bedeutet konkret die zufallsgesteuerte Auswahl von kleinen Test-Mengen der zu bearbeitenden Artefakte (hier: Viertelquadratmeter-Inventare; s.u.) und die anschließende Einzelaufnahme, d.h. die Aufnahme eines bearbeiterabhängigen Merkmalskataloges an jedem einzelnen Artefakt (Tabelle "Einzelaufnahme"). In der Datenbank werden die Merkmalausprägungen (Maße und codierte Qualitäten) der aufgenommenen archäologischen Objekte in Datensätzen der Tabelle "Viertelquadrate" abgespeichert.

Die drei Datentabellen (Abb. 1) der Beispiel-Datenbank sind untereinander verknüpft. Durch diese Struktur wird erreicht, daß keine Mehrfacheingaben vorgenommen werden müssen (z.B. die Fundplatznummer bei der Inventarisierung der Einzelaufnahme). Die Datentabelle "Einzelaufnahme" bildet das Herzstück für die Ermittlung einer geeigneten Stichprobengröße und für sämtliche weitere vorinstallierten Auswertungsmöglichkeiten. Über die vorinstallierten Tabellen, Abfragen und Makros zur Ermittlung der geeigneten Stichprobengröße hinausgehend hat der Benutzer alle Möglichkeiten der eigenen Datenanalyse.

*Überwachung der Stichprobengröße*

Die Überwachung der Stichprobengröße erfolgt weitgehend automatisiert. Die einzige Handlung, die der Bearbeiter des Artefaktmaterials zusätzlich zur Dateneingabe durchführen muß, ist ein Mausklick nach der Eingabe von jeweils ca. 50 Datensätzen in die Tabelle "Einzelaufnahme". Dies wird möglich, weil zur Datenbank eine Auswertungsroutine (Sampling-Routine) gehört, die zur Ermittlung des geeigneten Stichprobenumfangs in einem iterativen Prozeß immer größere Teststichproben analysiert. Da bereits während der Ausgrabung auf der Basis von Viertelquadraten gearbeitet wurde und ein typisches Viertelquadrat-Inventar in unserem Beispiel ca. 50 ± 20 Einzelstücke umfaßt, bildet das Inventar eines Viertelquadrats für die schrittweise Vergrößerung der Teststichprobe im vorliegenden Beispiel eine sinnvolle Einheit. Um zu überprüfen, wann eine geeignete Stichprobengröße erreicht ist, werden nach jedem aufgenommenen Viertelquadrat-Inventar die "laufenden Mittelwerte" metrisch skalierter Daten und die "laufenden Prozent-Anteile" nominal skalierter Daten herangezogen (siehe Tabelle). Mittelwerte der metrisch skalierten Variablen und Prozent-Anteile der nominal skalierten Variablen bezeichnen wir im folgenden der Kürze wegen gelegentlich auch zusammenfassend als "Kennwerte".

Zu Beginn der Auswertung wird in unserem Fallbeispiel eine Teststichprobe von einem zufällig ausgewählten Viertelquadrat-Inventar gezogen. Alle Einzelfunde dieses Viertelquadrates werden entsprechend den Vorgaben in der Tabelle "Einzelaufnahme" beschrieben. Anschließend werden die Mittelwerte oder die Prozent-Anteile der in der Tabelle aufgelisteten Merkmale mittels einer Sampling-Routine berechnet. Bei jeder Ausführung der Sampling-Routine werden die Ergebnisse als neuer Datensatz in einer weiteren Tabelle ("Laufende Kennwerte") angefügt und bleiben hier für weitere Anwendungen verfügbar. Nachdem sich die Stichprobe durch ein weiteres aufgenommenes Viertelquadrat-Inventar vergrößert hat, werden erneut Mittelwerte und Prozent-Anteile errechnet. Die Tabelle "Laufende Kennwerte" wird wieder aktualisiert.

Die Tabelle mit den laufenden Kennwerten stellt den wichtigsten Auswertungsbaustein der Datenbank dar. Auf der Basis der laufenden Kennwerte läßt sich eine sichere Entscheidung fällen, wann der Stichprobenumfang ausreichend ist, und wann die Datenaufnahme aus statistischer Sicht beendet werden kann. Zu diesem Zeitpunkt liegt ein optimales Kosten-Nutzen-Verhältnis vor.

Um den Überblick zu erleichtern, können die Daten der "Laufenden Kennwerte" als Liniendiagramm dargestellt werden. Wegen der unterschiedlichen Skalenniveaus, auf denen die Daten erhoben worden sind (sowohl das hohe metrische Niveau für quantitative Daten als auch das niedrige nominale für qualitative Daten), werden dazu jeweils Liniendiagramme für die laufenden Mittelwerte der quantitativen Daten und für die laufenden Prozent-Anteile der qualitativen Daten erstellt (Abb. 2 und Abb. 3). Die Entwicklung eines jeden Merkmals läßt sich so schrittweise, also bei jeder Vergrößerung der Teststichprobe beobachten. Nach dem Einpendeln der Kennwerte kann auf eine weitere Datenaufnahme verzichtet werden.

Der entscheidende nächste Schritt besteht deshalb darin, festzustellen, wann ein solches "Einpendeln" – und zwar jedes einzelne Merkmal betreffend – eingetreten ist. Dazu muß eine sinnvolle Obergrenze gesetzt werden, bis zu der eine Veränderung eines Mittelwertes oder eines Prozentanteils als unwesentlich angesehen wird. In der vorliegenden Datenbank wird davon ausgegangen, daß es sich bei einer zweimaligen Veränderung des Mittelwertes um weniger als 5% oder eines Anteils um weniger als 20% erübrigt, die Teststichprobe weiter zu vergrößern. Die Toleranzgrenze von 5% für metrisch skalierte Daten wird allgemein als sinnvoll erachtet (TRAXLER 1997, 70). Durch die Verschärfung der Kriterien um die Bedingung, daß die Toleranzgrenze durch zwei aufeinanderfolgende Teststichproben nicht überschritten werden

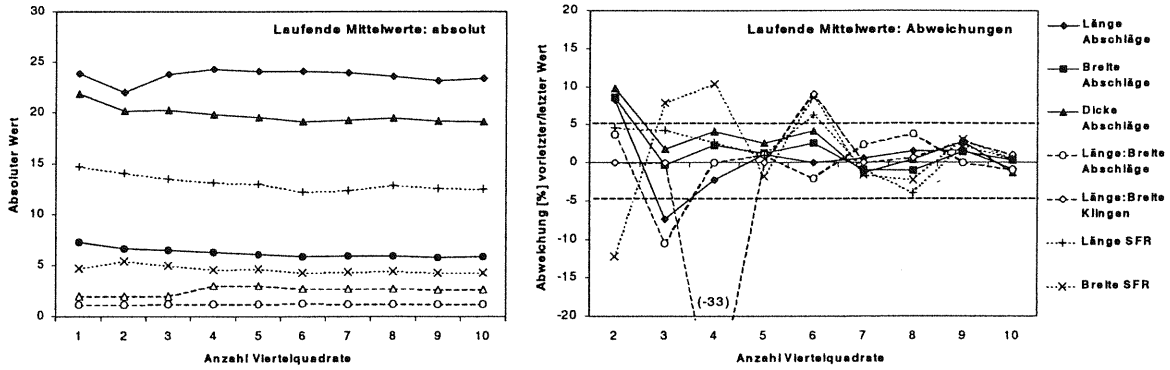


**Abb. 2** Laufende Mittelwerte ausgewählter, metrisch skalierter Fundmerkmale auf der Basis von zehn aufgenommenen Teststichproben (Viertelquadrat-Inventaren) der Fundstelle Wadi Bakht 82/1. Das Liniendiagramm links gibt die absoluten Mittelwerte, das rechte Diagramm dagegen die Abweichungen zwischen vorletztem und letztem Mittelwert (daher hier nur neun Werte). Dort ist zusätzlich die gewählte Toleranzgrenze von 5% Abweichung eingetragen.

darf, soll sichergestellt werden, daß die Abweichungen tatsächlich dauerhaft unter 5% bleiben. Für Prozentanteile gibt es in der Literatur noch keine entsprechende Empfehlung, da in der Ökologie selten mit nominal skalierten Daten gearbeitet wird.

Aus dem Vergleich der einpendelnden "laufenden Anteile" mit den laufenden Mittelwerten (Datenbasis: drei Fundplätze im Wadi Bakht mit jeweils zehn Teststichproben) haben wir pragmatisch eine Toleranzgrenze von 20 % abgeleitet. Es soll an dieser Stelle noch einmal betont werden, daß beide Toleranzgrenzen nicht mathematisch zwingend sind, sondern sich in der Praxis für die Festlegung einer hinreichend großen Stichprobe bewährt haben. Es bleibt somit auch jedem Bearbeiter überlassen, für seine eigenen Fragestellungen nötigenfalls schärfere oder auch weniger scharfe Kriterien festzulegen. Das (willkürliche) Festlegen einer erwünschten Genauigkeit in einem Sampling-Design läßt sich prinzipiell nicht durch eine statistische Routine ersetzen (LINSTÄDTER & LINSTÄDTER in Vorb.).

Eine Umsetzung der laufenden Kennwerte in Liniendiagramme (Abb. 2 links u. Abb. 3 links) zeigt, daß die laufenden Kennwerte selbst nur schlecht geeignet sind, um zu entscheiden, wann die Datenerhebung abgeschlossen werden kann. Dies hat seine Ursache in den unterschiedlichen Dimensionen der Merkmale. Abhilfe schafft hier eine Standardisierung. Da es uns nicht um die absolute Höhe der Werte, sondern um ihre relative Veränderung geht, erfolgt die Standardisierung so, daß jeweils die prozentuale Abweichung zwischen letztem und vorletztem Wert gezeigt wird. Zusätzlich kann zur Orientierung die gewählte Toleranz von 20% für die nominal skalierten Variablen (Abb. 3 rechts) und von 5% für die metrisch skalierten Variablen (Abb. 2 rechts) eingetragen werden.

## Ergebnisse

Die Entscheidung, wann die Datenaufnahme abgeschlossen werden darf, kann aus dem Verlauf der Abweichungen direkt abgelesen werden (vgl. Abb. 2 rechts u. Abb. 3 rechts). Zur Veranschaulichung und zum Vergleich untereinander stellen wir sowohl Originalwerte als auch einige aus ihnen abgeleitete Quotienten gemeinsam in einer Grafik dar. So findet sich im Diagramm der "Laufenden Anteile" neben dem "Anteil der Abschläge" und dem "Anteil der Klingen" auch der Quotient der Abschläge und Klingen und im Diagramm der "Laufenden Mittelwerte" neben der "Länge der Abschläge" und der "Breite der Abschläge" auch der Quotient der Länge und Breite der Abschläge.

Die Prozentanteile der nominal skalierten Merkmale (Abb. 3 rechts) vom neolithischen Fundplatz 82/21-1 überschreiten nach der Aufnahme von sieben Stichproben (hier identisch mit Viertelquadrat-Inventaren) letztmalig die festgelegte Abweichung von 20 Prozent. Von der achten Stichprobe an bleiben sie innerhalb der Toleranzgrenze. Nach den oben festgelegten Kriterien soll dies für zwei nacheinander folgende Stichproben erfüllt sein, bevor die Datenaufnahme beendet werden kann. Die Datenaufnahme der nominal skalierten Merkmale darf also mit der neunten Stichprobe beendet werden. Entscheidend waren hierbei die Merkmale, die sich am unruhigsten verhielten – der Anteil der Klingen am gesamten aufgenommenen Inventar und das damit eng zusammenhängende Verhältnis zwischen Abschlägen und Klingen. Wenn diese beiden Merkmale außer acht gelassen würden, wären schon seit der vierten Stichprobe stabile Verhältnisse eingetreten, und die Aufnahme hätte mit der sechsten Stichprobe beendet werden können.

Die Abweichungen der laufenden Mittelwerte

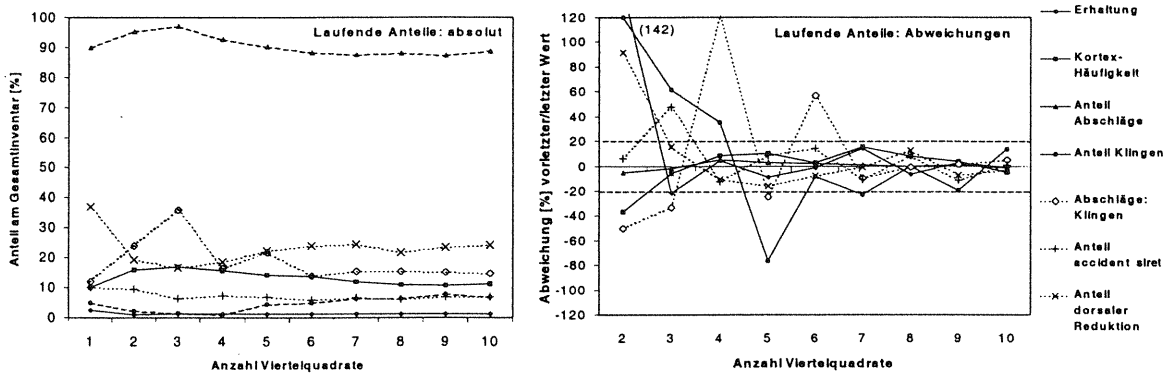


**Abb. 3** Laufende Anteile ausgewählter, nominal skalierter Fundmerkmale auf der Basis von zehn aufgenommenen Teststichproben (Viertelquadrat-Inventaren) der Fundstelle Wadi Bakht 82/1. Das Liniendiagramm links gibt die absoluten Anteile, das rechte Diagramm dagegen die Abweichungen zwischen vorletztem und letztem Prozentwert (daher hier nur neun Werte). Dort ist zusätzlich die gewählte Toleranzgrenze von 20% Abweichung eingetragen.

(Abb. 2 rechts) überschreiten nach der Aufnahme von sechs Stichproben letztmalig die festgelegte Toleranz von 5%. Die Datenaufnahme der metrisch skalierten Merkmale darf also mit der achten Stichprobe beendet werden. Würden auch hier wieder die unruhigsten Merkmale außer acht gelassen, also die Schlagflächenrest-Länge und -Breite sowie das Längen-Breiten-Verhältnis der Klingen, dann wären schon seit der vierten Stichprobe stabile Verhältnisse eingetreten, und die Datenaufnahme hätte auch hier mit der sechsten Stichprobe beendet werden können.

## Diskussion

Während der Fortschreibung der Ergebnis-Diagramme wird die Fundplatz-Fläche in unserem Fallbeispiel gleichsam Schritt für Schritt (Stichprobe für Stichprobe) durchwandert. Es wird deutlich, welche Variablen sich – eher unabhängig von den Flächenteilen – stabil, und welche sich – eher abhängig von den Flächenteilen – instabil verhalten. Variablen, die chorologisch empfindlich sind, wie in unserem Beispiel der Klingenanteil (Abb. 3), geben sich klar zu erkennen. Bei Merkmalsausprägungen, deren Anteile sich relativ schnell einpendeln, kann auf eine sehr homogene Verteilung in der Grabungsfläche geschlossen werden.

Bei den metrischen Merkmalen können besonders instabile Verläufe ebenfalls chorologisch bedingt sein. Zum Beispiel könnte in einem Flächenteil ein besonders kleinstückiges Rohmaterial benutzt worden sein. Dieses Ergebnis ist durch anfängliche Kartierungen des Inventares nicht zu erlangen. Als weitere Ursache für instabile Verläufe metrischer Werte kommt die verfügbare Meßgenauigkeit hinzu. Die meisten Maße wurden mit der Schieblehre genommen, die mit einer Genauigkeit von 1-1,5 mm abgelesen wird. Bei einer Klinge, deren Länge 5 cm beträgt, entspricht dies einer Genauigkeit von 2-3%. Bei der Dicke eines Abschlages oder bei der Länge und Breite des Schlagflächenrestes liegen die Maße aber meist bei wenigen Millimetern. Die Schieblehre liefert dann oft Ablese-Genauigkeiten von unter 50%.

Die soweit benannten Schwierigkeiten "chorologische Empfindlichkeit" und "Meßgenauigkeit" werden innerhalb der vorgestellten Konstellation zwar besonders deutlich, sie betreffen aber nicht speziell das hier vorgeschlagene Verfahren, sondern jede archäologische Datenerhebung. Ein demgegenüber einschlägiger Nachteil besteht darin, daß die Anzahl der berücksichtigten Individuen in die Beurteilung der einzelnen Variablen nicht unmittelbar einfließt. So basieren die Mittelwerte der Klingenmaße in unserem Beispiel auf weit weniger Einzelerhebungen als die Mittelwerte der Abschlagmaße. Es empfiehlt sich daher, bei sehr unruhigen laufenden Kennwerten die Größe der tatsächlich ausgewerteten Teilmenge des Artefaktmaterials zu überprüfen. Hierfür liefert die Tabelle "Laufende Kennwerte" für jeden Kennwert die entsprechende Stichprobengröße. Für metrische Daten ist aus statistischen Gründen z.B. eine Mindestgröße der Stichprobe von 30 Einzeldaten angebracht (TRAXLER 1997, 63). In unserem Beispiel könnte überlegt werden, ob eine selektive Nacherhebung der Grundform "Klinge" sinnvoll wäre.

Weiterhin ist von Fall zu Fall anzuraten, die aus der Methode der "Laufenden Mittelwerte" abgeleiteten Entscheidungen durch ein anderes Verfahren unabhängig zu prüfen. Für metrisch skalierte Daten steht hierfür eine statistische Methode zur Ermittlung der "Minimal Notwendigen Stichprobengröße" (BONHAM 1989, 43 ff.) zur Verfügung, die auf der Grundlage einer festzulegenden Vertrauenswahr-

scheinlichkeit und Genauigkeit die Anzahl der notwendigen Einzelbeobachtungen errechnet. Dieses Verfahren löst ein Problem, das bei allen Stichprobenverfahren besteht. Es wird deshalb in einem gesonderten Aufsatz behandelt (LINSTÄDTER & LINSTÄDTER in Vorb.).

Die hier vorgestellte Methodik hat ihren Wert in der Praxis der archäologischen Datenaufnahme. Sie gibt einen ständigen Überblick über die Entwicklung der einzelnen Kennwerte, ermöglicht erste Arbeitshypothesen über den Charakter der einzelnen Artefaktgruppen und ihrer Merkmale, erleichtert die Planung der weiteren Analyse-Strategie und informiert laufend über das Kosten-Nutzen-Verhältnis während der archäologischen Merkmalerhebung. Nicht zuletzt gestaltet sie den gesamten Prozeß der Merkmalerhebung interessanter und gibt denjenigen, die sie ausführen, einen ständigen Bezug zu den Inhalten und Zielen ihres Tuns.

## Anmerkungen

1 Im Gegensatz zur sonst üblichen Schreibweise werden sowohl die Fundplatznummer als auch die Bezeichnung des Viertelquadrates in eine Form überführt, die ihre computergestützte Verarbeitung wesentlich erleichtern. Wichtigster Vorteil ist ihr ausschließlich numerischer Charakter; beispielsweise wird die Bezeichnung 82/21-1 (das bedeutet: Fläche 1 der Fundstelle 21 des Jahres 1982) übersetzt in die Zahlenfolge 198202101.

2 Die Lage der Viertelquadrate innerhalb des Meter-Koordinatensystems der Ausgrabung wird durch die Zentimeter-Koordinaten des Mittelpunktes der Viertelquadratmeter-Flächen statt der früher üblichen "Viertelquadrat"-Bezeichnung aus Meter-Koordinate und einem Buchstaben (a,b,c,d) wiedergegeben (zum Beispiel statt 50/50b jetzt x=5075, y=5075). Durch diese Transformation sind die Daten für jedes Kartierungsprogramm verwendbar.

3 Im vorliegenden Fall wird das gesamte Steinartefakt-Inventar komplett erfaßt, und zwar auf der Basis von Viertelquadraten nach vorkommenden Rohmaterialien getrennt. Es werden die jeweiligen Anzahlen für die einzelnen Grundformen (Abschläge, Klingen, Trümmer sowie Absplisse), Werkabfälle (Kerne, Präparationsreste), Sekundärmodifikationen (Kerbreste, Stichelabfälle) und Geräte des gesamten Inventars aufgenommen.

4 Beim QS-Index handelt es sich um ein Auswertungskonzept, das den Vergleich verschiedener Fundstellen auf der Basis gleichartiger Angaben ermöglicht. Hierzu werden die Mittelwerte der Maße Breite, Dicke, Schlagflächenrestlänge und Schlagflächenrestbreite vollständiger Klingen und Abschläge in einem Quotienten zusammengefaßt (SCHÖN 1996, 64 ff.).

## Literatur


BONHAM, C.D. (1989) Measurements for terrestrial vegetation. New York 1989.

BREUER, M. (1997) Über den Aufbau einer vegetationskundlichen Datenbank. *EcoSys 6, 1997, 91-102.*

IHM, P., LÜNING, J. & A. ZIMMERMANN (1978) Statistik in der Archäologie. *Archaeo-Physika 9.* Köln 1978.

LINSTÄDTER, A. & J. LINSTÄDTER (in Vorb.) Wie groß müssen Stichproben sein? Berechnung der minimal notwendigen Stichprobengröße in der Archäologie.

LINSTÄDTER, J. (1999) Leben auf der Düne. Der mittelneolithische Fundplatz Wadi Bakht 82/21 im Gilf Kebir (Südwest-Ägypten). *Arch. Inf. 22/1, 1999, 115-124.*

MUELLER-DOMBOIS, D. & H. ELLENBERG (1974) Aims and methods of vegetation ecology. New York 1974.

MÜHLENBERG, M. (1993) Freilandökologie. Heidelberg 1993 [3., überarbeitete Aufl.].

SCHÖN, W. (1996) Ausgrabungen im Wadi el Akhdar, Gilf Kebir (SW-Ägypten). *Africa Praehistorica 8.* Köln 1996.

TRAXLER, A. (1997) Handbuch des vegetationsökologischen Monitorings. Methoden, Praxis, angewandte Projekte. Teil A: Methoden. Wien 1997.

ZIMMERMANN, A. (2001) Ist die politische Forderung nach der "beispielhaften Ausgrabung" aus fachlicher Sicht immer unerfüllbar? Zum Aspekt der Repräsentativität von Ausgrabungsergebnissen. *Arch. Nachrichtenbl. 6, 2000, 131-137.*

ZÖFEL, P. (2001) Statistik verstehen. Ein Begleitbuch zur computergestützten Anwendung. München 2001.



*Jörg Linstädter M.A.*
*Universität zu Köln*
*Institut für Ur- und Frühgeschichte*
*Forschungsstelle Afrika*
*D - 50923 Köln*

*Prof. Dr. Jürgen Richter*
*Universität zu Köln*
*Institut für Ur- und Frühgeschichte*
*D - 50923 Köln*

*Dr. Anja Linstädter*
*Universität zu Köln*
*Botanisches Institut*
*D - 50923 Köln*