

Die folgenden fünf Artikel sind Beiträge der AG Archäologie zur Tagung der Gesellschaft für Klassifikation, die vom 9.-11.3. in Dortmund stattfand

Seriation mit bedingter Korrespondenzanalyse: Simulationsexperimente

Michel van de Velden, Patrick J.F. Groenen und Jeroen Poblome

Zusammenfassung – Die Korrespondenzanalyse (CA) ist ein wohlbekanntes Verfahren, um geschlossene Fundkomplexe zu serieren. Ein Problem dabei ist, daß man mit der CA nur die relative Chronologie der Fundkomplexe erhält. Doch die Publikationen von GROENEN & POBLOME (2003) sowie POBLOME & GROENEN (2003) stellen ein Verfahren zur bedingten Korrespondenzanalyse vor, bei dem Zusatzinformationen in Form von Gleichungen und Ungleichungen für die Datierungen der Fundkomplexe berücksichtigt werden. Damit kann nicht nur das CA-Ergebnis verbessert werden, es ist auch möglich, absolute Datierungen für die Fundkomplexe zu errechnen. Der Aufsatz untersucht die Genauigkeit der bedingten CA mit Hilfe von Simulationsexperimenten. Konkret gesagt prüfen wir, wie gut die normale wie auch die bedingte CA in der Lage sind, die zeitliche Reihenfolge und die absoluten Datierungen zu rekonstruieren. Indem wir die Zahl der Zusatzinformationen und ihre Art (Gleichungen, Ungleichungen) variieren, erforschen wir, wie viele solcher Zusatzbedingungen notwendig sind, um eine zuverlässige zeitliche Einordnung zu erhalten. Die Simulationsexperimente orientieren sich an einem Datensatz aus Sagalassos (SW-Türkei) mit lokal hergestellter Keramik.

Schlüsselwörter – Korrespondenzanalyse, Simulation

Abstract – Correspondence analysis (CA) is a well-known technique for the seriation of archaeological assemblages. One problem with the CA seriation solution is that only a relative ordering of the assemblages is obtained. However, GROENEN & POBLOME (2003) and POBLOME & GROENEN (2003) proposed a constrained CA approach that incorporates additional information, in the form of equality and inequality constraints concerning the time points of the assemblages, in order to improve the usual CA solution. Using such constraints, explicit dates can be assigned to the seriation solution. In this paper, we study the quality of the constrained CA solution more closely using a simulation study. To be more precise, we study how well ordinary and constrained CA are able to recover the time order and the explicit dates. Moreover, by varying the amount and type (equality, inequality) of constraints, we investigate how much extra information is required to obtain a satisfactory reproduction of the true time frame. The simulation study is set up in such a way that it mimics the data of a series of ceramic assemblages consisting of the locally produced tableware from Sagalassos (SW Turkey).

Keywords – correspondence analysis, simulation

Einführung

Ein wichtiges Problem in der archäologischen Forschung ist die Sortierung von geschlossenen Fundkomplexen aufgrund der zugehörigen Funde, um die chronologische Reihenfolge zu rekonstruieren. Grundlage einer solchen Sortierung ist die Annahme, daß die Produktion der Objekte eines Fundtyps unimodal ist, d.h. jeder Fundtyp wurde nur eine begrenzte Zeit produziert, dabei wurde er langsam immer populärer bis nach dem Höhepunkt die Modewelle wieder abebbte. Eine weit verbreitete statistische Methode um das Sortierproblem zu lösen ist die Korrespondenzanalyse (grundlegend: GREENACRE 1984).

Mit Hilfe der Korrespondenzanalyse (engl. *correspondence analysis*, im folgenden abgekürzt mit CA) erhält man nicht nur eine Sortierung der Fundkomplexe, sondern auch die zugehörige Reihenfolge der darin vertretenen Fundtypen. Archäologische Fundkomplexe

mit zwei verschiedenen Datentypen können mit der CA analysiert werden: Besetzungs- und Häufigkeitsdaten. Bei Besetzungsdaten wird nur notiert, ob ein bestimmter Fundtyp in dem Fundkomplex vertreten ist, dabei wird das Vorhandensein eines Fundtyps mit einer 1 und das Fehlen mit einer 0 angegeben. Bei Häufigkeitsdaten wird nicht nur das Vorhandensein oder Fehlen, sondern die genaue Zahl der Objekte eines Fundtyps im jeweiligen Fundkomplex berücksichtigt. In diesem Aufsatz beschäftigen wir uns mit Häufigkeitsdaten.

In der Archäologie ist CA eine populäre Seriationsmethode, da sie relativ leicht durchgeführt und interpretiert werden kann (SHENNAN 1988). Aber das Ergebnis der CA ist nur die relative chronologische Ordnung, d.h. ohne Zusatzinformationen ist eine absolutchronologische Aussage über die Fundkomplexe nicht möglich. In der Regel gibt es jedoch Zusatzinformationen über die Sortierrichtung allgemein oder auch explizite Datierungsangaben für einzelne Fundkomple-

xe. Hinzu kommen die stratigraphischen Beziehungen zwischen den Fundkomplexen. Bei der klassischen CA ist es jedoch weder möglich, vorgegebene Jahreszahlen noch die stratigraphischen Beziehungen zu berücksichtigen.

GROENEN & POBLOME (2003) stellten ein Verfahren zur bedingten Korrespondenzanalyse vor, das die genannten Zusatzinformationen einbezieht. Bei einer Anwendung dieser Methode auf einen Geschirrdatensatz aus Sagalassos im Südwesten der Türkei konnten plausible Ergebnisse erzielt werden. Doch war es nicht möglich, eine Aussage über die Genauigkeit der Methode zu machen, d.h. anzugeben, wie genau die bedingte CA in der Lage ist, die Datierungsangaben für die Fundkomplexe zu errechnen. In diesem Aufsatz untersuchen wir dieses Problem mit Simulationsexperimenten. Das heißt, indem wir die Eigenschaften von archäologischen Daten nachahmen, können wir feststellen, wie genau die bedingte CA von GROENEN & POBLOME die Datierungen errechnen kann.

In den nächsten beiden Abschnitten folgt eine kurze Einführung in die CA und die bedingte CA. Danach beschreiben wir das Simulationsverfahren, das auf die Daten aus Sagalassos zugeschnitten ist. Anschließend stellen wir die Ergebnisse der Simulationsuntersuchung vor, gefolgt von einer abschließenden Diskussion der Resultate.

Korrespondenzanalyse

Bei der CA werden den Fundkomplexen und den Fundtypen Positionen auf einer Achse so zugeordnet, daß die Abstände zwischen den einzelnen Fundkomplexen bzw. zwischen den Fundtypen möglichst genau wiedergegeben werden. Es gibt mehrere exzellente Einführungen in die Methode, z.B. GREENACRE (1984). Eine Beschreibung der CA für Archäologen findet sich bei SHENNAN (1988).

In der Schreibweise der Matrizenrechnung kann man die CA als folgendes Minimierungsproblem formulieren:

$$\min_{\mathbf{r}, \mathbf{c}, \lambda} \left\| \mathbf{D}_r^{-1/2} (\mathbf{F} - \mathbf{f}_{++}^{-1} \mathbf{r} \mathbf{c}' - \lambda \mathbf{D}_r \mathbf{r} \mathbf{c}' \mathbf{D}_c) \mathbf{D}_c^{-1/2} \right\|^2, \quad (1)$$

wobei $\|\mathbf{X}\|^2$ die Summe der quadrierten Elemente der Matrix \mathbf{X} bezeichnet, \mathbf{F} ist die $n \times p$ Datenmatrix, bei der in Zeile i , Position j die Anzahl der Objekte von Fundtyp j in Fundkomplex i eingetragen ist, \mathbf{f}_{++} ist die Summe aller Elemente von \mathbf{F} , \mathbf{D}_r und \mathbf{D}_c sind Diagonalmatrizen, wobei die Diagonalelemente durch \mathbf{f}_{++} bzw. \mathbf{f}_{++} vorgegeben sind, mit \mathbf{f}_{++} und \mathbf{f}_{++} den Spaltensummen von \mathbf{F} . Die Vektoren \mathbf{r} und \mathbf{c} sind die gesuchten Positionen für die Fundkomplexe und die Fundtypen. Der Faktor λ ist ein Singulärwert und kann als Maß für die Güte der

Anpassung betrachtet werden. Die zeitliche Reihenfolge der Fundkomplexe ergibt sich aus den zugehörigen Positionen in Vektor \mathbf{r} . Doch wenn \mathbf{r} und \mathbf{c} das obige Minimierungsproblem lösen, dann auch $-\mathbf{r}$ und $-\mathbf{c}$. Das heißt, die CA ermittelt nicht die Richtung der Achse. Die Vektoren \mathbf{r} und \mathbf{c} sind bis auf das Vorzeichen und ihre Länge eindeutig bestimmt, deswegen werden sie standardisiert, so daß $\mathbf{r}' \mathbf{D}_r \mathbf{r} = \mathbf{c}' \mathbf{D}_c \mathbf{c} = 1$.

Als Ergebnis der CA erhält man nur eine relative chronologische Reihenfolge, Jahreszahlen für Fundkomplexe und Fundtypen können nicht in einfacher Weise aus den Achsenpositionen abgeleitet werden.

Bedingte Korrespondenzanalyse

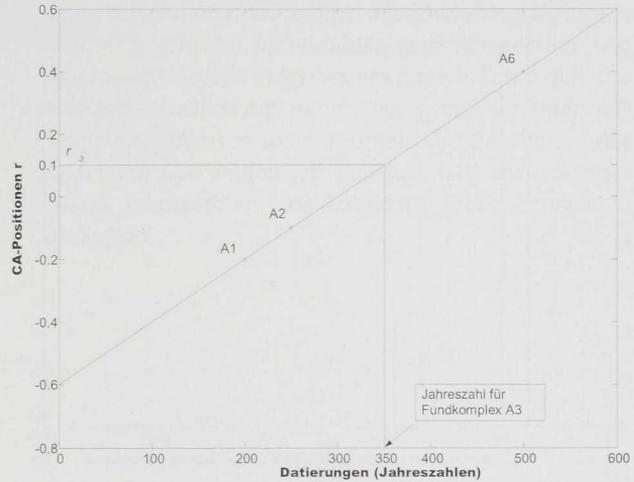
GROENEN & POBLOME (2003) stellten eine Methode vor, die die Berechnung von Datierungen der Fundkomplexe aufgrund von chronologischen Zusatzbedingungen erlaubt. Wenn die Zusatzinformationen in Bezug auf die Datierungen ausgewählter Fundkomplexe berücksichtigt werden, gibt es einen einfachen Zusammenhang zwischen den von der CA errechneten Positionen und den Jahreszahlen. Wir unterscheiden drei verschiedene Arten von Zusatzbedingungen:

1. Bei einigen Fundkomplexen sind die genauen Datierungen bekannt.
2. Manchmal weiß man, daß eine kleine Gruppe von Fundkomplexen zur gleichen Zeit entstanden ist.
3. Bei manchen Fundkomplexen ist die zeitliche Reihenfolge ihres Entstehens bekannt.

GROENEN & POBLOME (2003) zeigen, wie diese drei verschiedenen Arten von Zusatzinformationen bei der CA berücksichtigt werden können, indem entsprechende Gleichungen und Ungleichungen aufgestellt werden. Bei den Zusatzinformationen erster Art gehen wir folgendermaßen vor: Sei y_i die zum Fundkomplex i gehörige Datierung (Jahreszahl). Dann stellen wir sicher, daß die von der CA errechnete Position von Fundkomplex i sich linear aus dieser Datierung ergibt. Das heißt, $r_i = b_0 + b_1 y_i$, wobei b_0 eine unbekannte Konstante und b_1 die Steigung der Geraden ist. Sowohl die Konstante b_0 als auch die Steigung b_1 müssen bei der bedingten CA errechnet werden.

Anhand dieses linearen Zusammenhangs können für alle Fundkomplexe absolute Datierungen ermittelt werden. Gehen wir beispielsweise davon aus, daß die Datierungen der Fundkomplexe A1, A2 und A6 bekannt sind. Dann hängen die bei der bedingten CA errechneten Positionen dieser Fundkomplexe linear mit den vorgegebenen Jahreszahlen zusammen. Damit liegen diese drei Punkte auf einer Geraden. Abbildung 1 zeigt anhand des Beispiels von Fundkomplex A3, wie aufgrund der mit der bedingten CA errechneten Posi-

Abb. 1 Schätzung der Jahreszahlen.



tion r_3 die absolute Datierung geschätzt werden kann.

Bei den Zusatzinformationen zweiter Art, d.h. gleiche Datierungen verschiedener Fundkomplexe, müssen die Positionen dieser Fundkomplexe gleich sein (z.B. $r_i = r_j$). Die beiden ersten Arten von Zusatzinformationen können in der Matrixgleichung $\mathbf{Hb} = \mathbf{r}$ zusammengefaßt werden, wobei \mathbf{H} eine aus den Jahreszahlen, Einsen und Nullen konstruierte Matrix ist, \mathbf{b} ist der Vektor der Geradenparameter b_0 und b_1 sowie weiterer Koeffizienten, die sich aus den Zusatzbedingungen für gleiche Datierungen ergeben.

Das Problem kann durch den sogenannten Kernraum-Ansatz (engl. *null-space approach*) vereinfacht werden (BÖCKENHOLT & TAKANE 1994). Dabei berechnen wir den Kernraum von \mathbf{H}' , d.h. die Matrix \mathbf{H}_0 mit $\mathbf{H}'\mathbf{H}_0 = \mathbf{0}$. Damit kann die Bedingung $\mathbf{Hb} = \mathbf{r}$ auf beiden Seiten von links mit \mathbf{H}_0' multipliziert werden, und so erhalten wir: $\mathbf{H}_0'\mathbf{r} = \mathbf{0}$. Durch diese Umformung haben wir den Vorteil, daß die Gleichung nun unabhängig von dem unbekanntem Parametervektor \mathbf{b} ist.

Die Zusatzbedingungen dritter Art sind die Ungleichungen. Falls bekannt ist, daß Fundkomplex i später datiert als Fundkomplex j , so können wir sicherstellen, daß die von der bedingten CA errechneten Positionen auch diese Bedingung erfüllen (d.h. $r_i \geq r_j$). Diese Ungleichungen kann man umformen, so daß für eine entsprechend konstruierte Matrix \mathbf{G} gilt: $\mathbf{Gr} \geq \mathbf{0}$. Dabei entspricht jede Ungleichung einer Zeile der Matrix, für jeden Fundkomplex, der an mindestens einer Ungleichung beteiligt ist, hat Matrix \mathbf{G} eine eigene Spalte. Für den älteren Fundkomplex wird -1 , für den jüngeren $+1$ eingetragen, die restlichen Elemente der Zeile sind 0.

Das Ziel der bedingten CA ist es, den Ausdruck von (1) zu minimieren, und zwar unter den Nebenbedingungen $\mathbf{H}_0'\mathbf{r} = \mathbf{0}$ und $\mathbf{Gr} \geq \mathbf{0}$. Diese Aufgabe kann durch ein alternierendes Verfahren der kleinsten Quadrate (engl. *alternating least-squares*) gelöst werden.

Genauere Details hierzu sind in GROENEN & POBLOME (2003) ausgeführt.

Bei der Anwendung von bedingter CA auf die Geschirrrdaten von Sagalassos stimmten die Rechnergebnisse mit dem archäologischen Kenntnisstand überein (POBLOME & GROENEN 2003). Da jedoch die tatsächlichen Datierungen der Fundkomplexe nicht bekannt sind, ist es schwierig festzustellen, wie gut die Resultate dieser neuen Methode sind. Auch wenn durch die linearen Zusammenhänge zwischen den Datierungen eine Jahreszahl für jeden Fundkomplex angegeben werden kann, steht nicht fest, wie genau diese Jahreszahl für die Fundkomplexe ist, für die keine Datierung bekannt war.

Simulationsexperimente

Um die Leistungsfähigkeit der bedingten CA zu überprüfen, führten wir Simulationsexperimente durch, bei denen künstliche archäologische Daten erzeugt wurden. Die Idee, archäologische Fundkomplexe zu simulieren, ist nicht neu (GRAHAM et al. 1976; HERZOG & SCOLLAR 1988; LOCKYEAR 1991). Zwei dieser Anwendungen von Simulationsverfahren beschränken sich auf vorgegebene Befundarten, nämlich Münzdepots (LOCKYEAR 1991) bzw. Gräber (GRAHAM et al. 1976). Der Ansatz von HERZOG & SCOLLAR (1988) ist allgemeiner und kann auch Häufigkeitsdaten erzeugen. Diese Methode ist Teil des Bonn Archaeological Software Package (WINBASP, <http://www.uni-koeln.de/~al001/basp.html>). Um alle Fäden bei den Simulationsexperimenten in der Hand zu behalten, entschieden wir uns, ein eigenes Simulationsverfahren zu entwickeln. Da es sich um einen ersten Schritt zur Überprüfung der bedingten CA handelt, wurden die Simulationen so einfach wie möglich gestaltet.

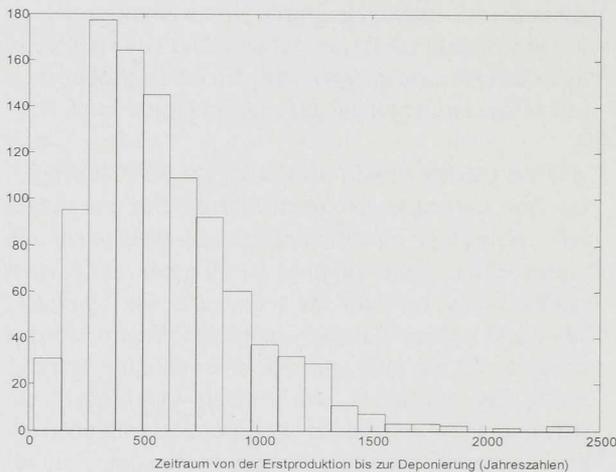


Abb. 2 Histogramm aus 1000 Jahreszahlen, die angeben, wie viel Zeit verstrichen ist von der Produktion eines ersten Objektes dieses Fundtyps bis zur Deponierung eines beliebigen Objektes dieses Typs.

Bei den hier vorgestellten Simulationsexperimenten wählten wir die Parameter und Einstellungen so, daß die auf diese Weise erzeugten Daten dem Geschirrdatensatz aus Sagalassos möglichst ähnlich sind. Wir betrachten einen Zeitraum von 1000 Jahren, der in fünfzig Abschnitte gleicher Länge eingeteilt wird. Jedem dieser Zeitabschnitte ordnen wir eine zufällige Anzahl von Fundkomplexen zu. Genauer gesagt unterliegt die Anzahl in jedem Zeitabschnitt einer Gleichverteilung mit Werten zwischen 1 und 5. In ähnlicher Weise ermitteln wir die Zahl der neu hinzu kommenden Fundtypen im betrachteten Zeitabschnitt, diese Zahl ist gleichverteilt zwischen 1 und 10. Dann wird für jeden Fundtyp festgelegt, wie viele Objekte davon in den Fundkomplexen erhalten sind, diese Anzahl ist auch gleichverteilt und liegt zwischen 5 und 1000. Die Deponierungszeit eines Fundobjektes ergibt sich aus der Summe der Anfangszeit des zugehörigen Fundtyps und einer Gamma-verteilten Zufallsvariablen mit Parametern 3 und 200. Stichproben dieser Gamma-Verteilung werden in Abbildung 2 in Form eines Histogramms dargestellt. Ausgehend von diesem Deponierungszeitpunkt wird der zugehörige Fundkomplex ausgewählt. Sind dem betreffenden Zeitabschnitt mehrere Fundkomplexe zugeordnet, wählen wir einen zufällig aus.

Nachdem wir alle Fundobjekte den Fundkomplexen zugeordnet haben, erhalten wir eine Datenmatrix, in der wir sowohl für alle Fundkomplexe als auch für alle Fundtypen positive Eintragungen haben. Diese Situation ist nicht sehr realistisch, weil in der Regel die Zahl der Fundkomplexe stärker beschränkt ist. Deshalb wählen wir aus den simulierten Fundkomplexen 20 zufällig aus, dazu 80 Fundtypen, und tragen nur die zugehörigen Informationen in die Datenmatrix ein.

Bei den simulierten Fundtypen und Fundkomplexen sind die exakten Datierungen bekannt. Mit Hilfe dieser

Informationen werden die Zusatzbedingungen erzeugt. Um zu ermitteln, wie die Zahl der Zusatzbedingungen das Ergebnis der bedingten CA verbessert, betrachten wir zwei Fälle:

1. Drei Datierungen sind bekannt, dazu 10 Ungleichungen.
2. Vier Datierungen sind bekannt, dazu 20 Ungleichungen.

Gleichungen für übereinstimmende Datierungen werden überall dort eingeführt, wo dies möglich ist: Wenn nach dem Auswahlprozess zwei Fundkomplexe zum gleichen Zeitabschnitt gehören, werden die zugehörigen Datierungen gleich gesetzt. Durch den Auswahlprozess kann es vorkommen, daß sich gar keine Gleichungen ergeben. Doch erhielten wir durchschnittlich zwei Gleichzeitigkeitsbedingungen pro Simulationsexperiment.

Für die beiden oben genannten Fälle führten wir jeweils 2000 Simulationsexperimente durch.

Ergebnisse

Um die Ergebnisse zwischen der bedingten und der normalen CA zu vergleichen, berechneten wir den Spearman'schen Rangkorrelationskoeffizienten zwischen der richtigen und den durch die beiden Verfahren ermittelten Reihenfolgen. In beiden Fällen wird die richtige chronologische Reihenfolge der Fundkomplexe gut rekonstruiert. Dabei erhielten wir mit der bedingten Methode etwas bessere Ergebnisse als mit der normalen CA. Die genauen Werte werden in Tabelle 1 aufgelistet.

	Spearmansche Rangkorrelations- koeffizient
Normale CA	0.9687
Bedingte CA mit 10 Ungleichungen	0.9724
Bedingte CA mit 20 Ungleichungen	0.9850

Tab. 1 Der Spearmansche Rangkorrelationskoeffizient zwischen der mit CA berechneten und der richtigen Reihenfolge.

Ein wichtiger Vorteil der neuen Methode ist die Berechnung der absoluten Datierungen für die Fundkomplexe. Da dies nicht mit der normalen CA möglich ist, kann man hier keine direkten Vergleiche ziehen. Doch mit dem Simulationsverfahren werden solche Datierungen erzeugt, die wir dann mit den errechneten Datierungen vergleichen können. In Abbildung 3 zeigen wir für ein ausgewähltes Simulationsexperiment die Gerade mit den rekonstruierten Jahreszahlen und dazu Punkte an den richtigen Werten. Der vertikale Abstand zwischen einem Punkt und der Geraden entspricht der Abweichung zwischen dem richtigen und dem rekonstruierten Wert. Dabei ist zu beachten, daß in diesem Fall vier Datierungen als bekannt vorausgesetzt wurden, die natürlich auf der Gerade liegen müssen. Bei diesem Simulationsbeispiel streuen die Punkte dicht um die Gerade, also kommen die mit der bedingten CA errechneten Datierungen den tatsächlichen Werten sehr nahe.

Um eine allgemeine Vorstellung über die Leistungsfähigkeit der bedingten CA zu erhalten, ermittelten wir zusätzlich Mittelwert und Median der Abweichungen zwischen richtiger und errechneter Datierung. Dabei zeigte es sich, daß die Erhöhung der Zahl der Ungleichungen nicht zu besseren Abweichungswerten führten. In beiden Fällen betrug der Mittelwert 94 und der Median 64 Jahre. Dieses Ergebnis bedeutet, daß es zwar einige wenige sehr hohe Abweichungen gab, doch die Mehrzahl war recht klein. Das Gleiche zeigt auch Abbildung 4, ein Histogramm, das wiedergibt, wie sich die mittleren Abweichungen für alle Simulationsexperimente verteilen.

Die Fundkomplexe, für die die Jahreszahlen als bekannt vorgegeben sind, werden bei den Simulationsexperimenten zufällig ausgewählt. Daher kann es passieren, daß die vorgegebenen Datierungen sehr nah beieinander liegen. In einem solchen Fall stützen sich die berechneten Datierungen nur auf ein kleines Intervall und sind möglicherweise weniger genau als bei einem größeren Intervall. In Abbildung 5 untersuchen wir diese Hypothese, indem wir die mitt-

lere Abweichung der Länge des Datierungsintervalls den vorgegebenen Jahreszahlen gegenüberstellen. Die Regressionsgerade zeigt den erwarteten Trend, d.h. die mittlere Abweichung nimmt ab, wenn die Intervalllänge vergrößert wird. Somit sind die Ergebnisse der bedingten CA tendenziell genauer, falls die vorgegebenen Jahreszahlen eine möglichst große Zeitspanne abdecken.

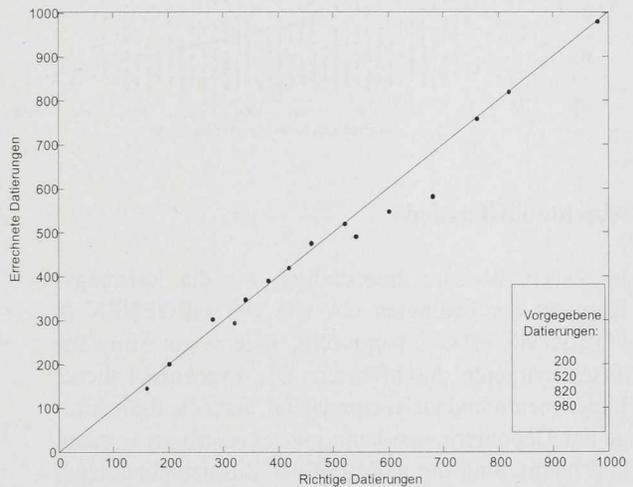


Abb. 3 Vergleich zwischen den richtigen und den berechneten Datierungen.

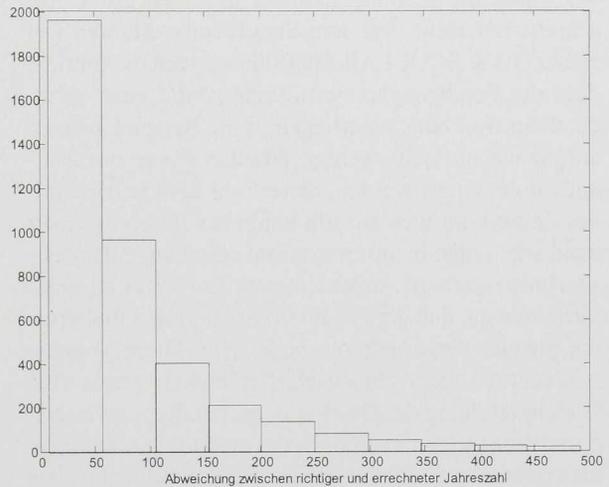


Abb. 4 Histogramm der mittleren Abweichung zwischen richtigen und simulierten Datierungen.

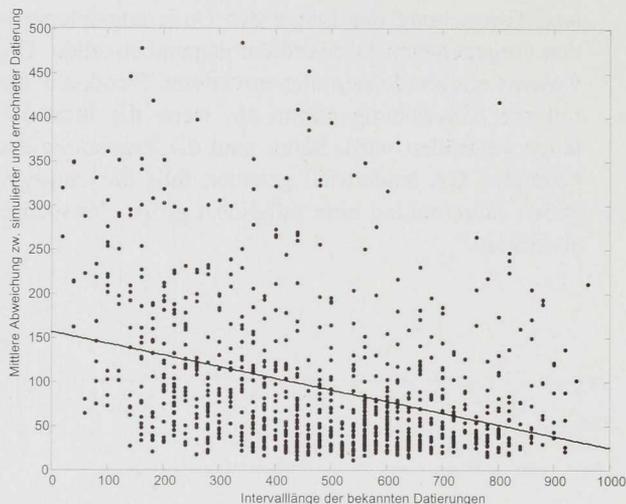


Abb. 5 Streudiagramm der mittleren Abweichungen im Vergleich zu der Spannweite der vorgegebenen exakten Datierungen.

Abschlußdiskussion

In diesem Aufsatz untersuchten wir die Leistungsfähigkeit der bedingten CA wie von GROENEN & POBLOME (2003) vorgestellt, indem wir Simulationsexperimente durchführten. Die Ergebnisse dieser Experimente sind vielversprechend, denn die Reihenfolge der Deponierungen kann gut rekonstruiert werden, und häufig sind die aufgrund der Zusatzbedingungen errechneten absoluten Datierungen hervorragend.

Die Simulationsexperimente waren recht einfach aufgebaut und können in Zukunft noch weiter verbessert werden. Dabei möchten wir die folgenden Aspekte berücksichtigen. Bei den hier vorgestellten Experimenten wurde die gesamte Laufzeit in gleich große Abschnitte unterteilt. Bei dem Simulationsverfahren von HERZOG & SCOLLAR (1988) kann auch die mittlere Zahl der Fundkomplexe pro Zeitabschnitt einer zeitlichen Entwicklung unterliegen, zum Beispiel können anfänglich nur sehr wenige Fundkomplexe pro Zeiteinheit deponiert werden, diese Zahl könnte dann mit der Zeit erhöht werden. Ein ähnliches Ergebnis erhält man, wenn man in unserem Ansatz die Länge der Zeitabschnitte variiert. Außerdem nehmen wir in unserem Verfahren an, daß der Produktionsdauer der Fundtypen die gleiche Verteilung zugrunde liegt. Diese Vorgabe ist eventuell nicht realistisch. Schließlich wurde eine Gleichverteilung der Objekte eines Fundtyps zwischen 5 und 1000 angenommen. In der archäologischen Praxis erscheint eine solche Gleichverteilungsannahme als wenig realistisch. Normalerweise gibt es viele Fundtypen mit relativ wenigen deponierten Objekten und eine kleine Anzahl mit großer Fundhäufigkeit. Mit diesen Aspekten möchten wir uns in zukünftigen Simulationsexperimenten beschäftigen.

Anmerkung

Die Autoren danken Irmela Herzog für die Übersetzung des englischen Textes ins Deutsche sowie wertvolle Hinweise und Kommentare.

Literatur

- BÖCKENHOLT, U. & Y. TAKANE (1994) Linear constraints in correspondence analysis. In: GREENACRE, M. & J. BLASIUS (eds.) *Correspondence analysis in the social sciences*. London 1994, 112-127.
- GRAHAM, I., GALLOWAY, P. & I. SCOLLAR (1976) Model studies in computer seriation. *Journal of Archaeological Science* 3, 1-30.
- GREENACRE, M.J. (1984) Theory and applications of correspondence analysis. New York 1984.
- GROENEN, P.J.F. & J. POBLOME (2003) Constrained correspondence analysis for seriation in archaeology applied to Sagalassos ceramic tablewares. In: SCHWAIGER, M. & O. OPITZ (eds.) *Exploratory Data Analysis in Empirical Research*. Berlin 2003, 90-97.
- HERZOG, I. & I. SCOLLAR (1988) A mathematical basis for simulation of seriatable data. In: RAHTZ, S.P.Q. (ed.) *Computer and Quantitative Methods in Archaeology 1988*. *British Archaeological Reports, International Series 446*. Oxford 1988, 53-62.
- LOCKYEAR, K. (1991) Simulating coin hoard formation. In: LOCKYEAR, K. & S.P.Q. RAHTZ (eds.) *Computer Applications and Quantitative Methods in Archaeology 1990*. *British Archaeological Reports, International Series 565*. Oxford 1991, 195-206.

POBLOME, J. & P. J.F. GROENEN (2003) Constrained correspondence analysis for seriation of Sagalassos tablewares. In: DOERR, M. & A. SARRIS (eds.) *Computer Applications and Quantitative Methods in Archaeology. Hellenic Ministry of Culture. Athens 2003, 301-306.*

SHENNAN, St. (1988) *Quantifying archaeology.* Edinburgh 1988.

*Michel van de Velden
Econometric Institute
Erasmus University
NL - Rotterdam
vandevelde@few.eur.nl*

*Patrick J.F. Groenen
Econometric Institute
Erasmus University
NL - Rotterdam
groenen@few.eur.nl*

*Jeroen Poblome
Sagalassos Archaeological Research Project
Catholic University Leuven
B - Leuven
jeroen.poblome@arts.kuleuven.ac.be*