Berliner Beiträge zur Archäometrie	Band 21	Seite 115-132	2008

Über Identifikationsmethoden, dargestellt am Beispiel römischer Baukeramik[1]

Hans-Georg Bartel,[2] Hans-Joachim Mucha[3] und Jens Dolata[4]

Abstract:

Cluster analysis is the most commonly used multivariate technique in archaeometry [MOMMSEN, H. et al. (1988), PAPAGEORGIOU, I. et al. (2001)]. Once a grouping of objects is obtained by clustering or in some other way, often new or unseen objects are to be assigned to these reference groups [MOMMSEN, H. (1981)]. This task of identification can be done by several statistical methods. Having a specific application (bricks and tiles in the north-western provinces of the Roman Empire) in mind, some appropriate identification methods are presented in action. Here the measurement of 19 chemical elements was done with X-ray fluorescence analysis (XRF). The main aim of the clustering step of this application was both to confirm military brickyards and to find places of those ones that are not yet identified. The identification task of two new (unseen) objects into the established reference groups is restricted to methods that can deal with quite different sizes of groups ranging from small groups containing four objects at least until large ones with 123 objects at most.

1. Anliegen und Aufgabe

Es besteht im Rahmen der Exploration von Daten nicht selten die Aufgabe, dass ein Objekt, von dem ein bestimmtes Merkmalsmuster bekannt ist, einer Klasse[5] zugeordnet werden soll, die zu einer bereits gegebenen Klassenzerlegung einer umfangreicheren Objektmenge gehört. Für die Objekte dieser Menge müssen eben die Merk-

---

1 Dieser Beitrag war ursprünglich im Herbst 2001 vom Herausgeber der Zeitschrift *Archaeometrica Acta* zur Veröffentlichung in deren 2. Jahrgang angeregt worden, um den auf dem Gebiet der Archäometrie tätigen Forschern einen Einblick in für sie nützliche mathematische Methoden zu geben. Da dieses Publikationsorgan aber bereits nach der Herausgabe des 1. Jahrgangs sein Erscheinen eingestellt hat, sind die Autoren Herrn Prof. Dr. JOSEF RIEDERER außerordentlich für die freundlich gewährte Möglichkeit dankbar, diesen Aufsatz in den *Berliner Beiträgen zur Archäometrie* publizieren zu dürfen und somit dem an sie herangetragenen Wunsch doch noch Rechnung tragen zu können. Diese und viele andere durch Herrn Prof. Dr. RIEDERER dankbar erfahrene Förderungen möchten die Autoren zum Anlass nehmen, ihm diesen Aufsatz anlässlich seines 65. Geburtstages im Dezember 2004 zu widmen.

2 Humboldt-Universität zu Berlin, Institut für Chemie, Brook-Taylor-Straße 2, D-12489 Berlin.

3 Weierstraß-Institut für Angewandte Analysis und Stochastik, Mohrenstraße 39, D-10117 Berlin.

4 Generaldirektion Kulturelles Erbe Rheinland-Pfalz, Direktion Archäologie, Arbeitsstelle Mainz, Große Langgasse 29, D-55116 Mainz.

5 Für den Terminus 'Klasse' kann im Rahmen dieses Aufsatzes synonym 'Cluster' verwendet werden. Ebenso verhält es sich mit den Begriffen 'Klassenzerlegung' und 'Partition'.

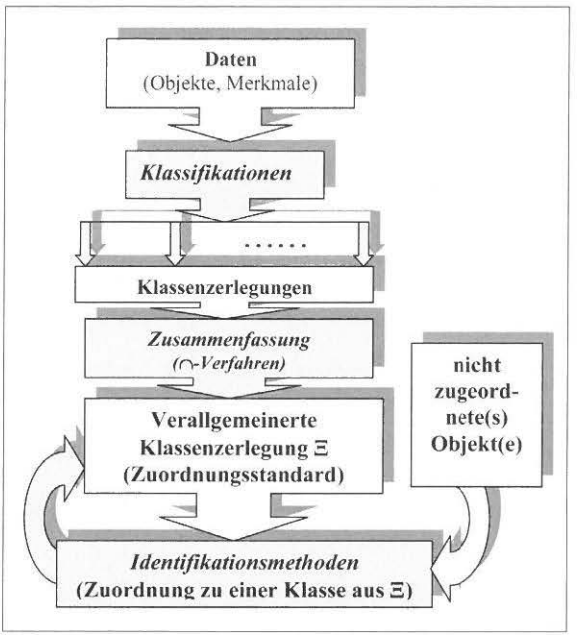

Abb. 1: Schema der Objektzuordnung zu einer verallgemeinerten Klassenzerlegung

male bekannt sein, die das Muster des untersuchten Objekts ausmachen. Der vorliegende Aufsatz diskutiert einige Lösungsmöglichkeiten dieser Aufgabe im Zusammenhang mit einem archäometrisch bzw. archäologisch relevanten Problem, in welchem die Objekte Artefakte sind, denen durch Analyse, Messung oder Beobachtung chemische, physikalische usw. Eigenschaften in einem Merkmalsmuster zugewiesen wurden. Zugleich ist angenommen worden, dass mehr als eine Klassenzerlegung vorliegt. Diese Annahme kommt den in der Praxis vorzufindenden Verhältnissen insofern nahe, als Klassifikationen oft unter verschiedenen Gesichtspunkten vorgenommen werden. In Abbildung 1 ist schematisch die Vorgehensweise dargestellt, wie sie im Folgenden beschrieben werden wird.[6]

2. Standardpartition

Als spezielles Beispiel sind für die Untersuchungen und Darlegungen im genannten Sinne römische Ziegel aus Obergermanien gewählt worden [Bartel, H.-G. et al. (2000, 2001, 2002, 2003), Dolata, J. (1994, 1996, 1997-2001, 1998, 1999, 2000a, 2000b, 2001a, 2001b, 2002, 2003, 2004), Dolata, J. et al. (1998/1999, 2001, 2003a, 2003b, 2004), Mucha, H.-J. et al. (2000/2001, 2003a, 2003b, 2005a, 2005b), Swart (2005), Swart et al. (2004), Werr, U. (1998)].

Entsprechend dem Schema in Abbildung 1 müssen zuerst die in den vorgegeben Daten auftretenden Objekte mindestens auf eine Weise nach den mit ihnen verknüpften

[6] Im Prinzip wäre natürlich auch eine erneute Clusteranalyse möglich, bei welcher das zu identifizierende Objekt der gegebenen Menge hinzugefügt ist. Auf diese Weise könnte aber diese kleine Ursache des Hinzufügens eine ziemlich große und unerwünschte Wirkung hervorrufen, indem sie die gefundenen Klassenzuordnungen in der ursprünglichen Partition erheblich verändert. Das mit der letzteren verbundene und als zuverlässig angesehene Modell wäre gewissermaßen unbegründet merklich modifiziert worden. Eine solche Schaffung eines neuen Modells ist erst dann vernünftig, wenn die ihm zugrundeliegende Menge durch Kenntnis einer größeren Anzahl von Objekten erweitert werden kann. Dieses bedeutendere Informationsvolumen rechtfertig sodann die eventuelle Schaffung eines neuen Modells durch Anwendung der Clusteranalyse.

Merkmalen klassifiziert werden. In dem zur Demonstration benutzten Beispiel handelt es sich um 602 Baukeramikproben, für welche mit Hilfe der RFA die jeweiligen Masseanteile an 19 chemischen Elementen als Merkmalsmuster ermittelt wurden.[7]

Mit dem Ziel ihrer Zuordnung zu bestimmten Heeresziegeleien, in denen die zugehörigen Ziegel produziert worden waren, sind für sie zwei Klassenzerlegungen ermittelt worden [DOLATA, J. (2000b), DOLATA, J. et al. (2003a), MUCHA, H.-J. et al. (2000/2001)]:

P1 Mit dem WARD-Verfahren der agglomerativen hierarchischen Clusteranalyse [BARTEL, H.-G. (1996), BOCK, H. H. (1974), MUCHA, H.-J. (1992), SPÄTH, H. (1975)][8] ist eine Partition erhalten worden, wobei sich auf mathematischem Wege die Anzahl von neun Klassen als günstig erwies. Ihre Bezeichnung erfolgt durch die Clusternummer $i = 1, \ldots, 9$. Diese Klassenzerlegung stützt sich auf die mit den Daten gegebene chemische und die mit der Berechnung verbundene mathematisch-statistische Information.[9] Als Abstandsmaß zwischen den Objekten wurde die quadrierte gewichtete euklidische Distanz zugrunde gelegt. Die hier benutzte Merkmalsgewichtung mit den inversen quadrierten Mittelwerten macht die Elementgehalte untereinander vergleichbar (insbesondere die Oxid- und Spurenelementgehalte) und erlaubt erst eine sinnvolle Anwendung der euklidischen Distanz bei unterschiedlichen Maßeinheiten.

P2 Bei der Überformung dieses Ergebnisses *P1* durch das Wissen des Archäologen wird ein Resultat erhalten, das zusätzlich zu der genannten auf archäologischer Information basiert. Die Bezeichnung der so ermittelten acht Cluster richtet sich nach den in Frage kommenden Standorten *A* der Heeresziegeleien: *F* – Frankfurt-Nied, *G* – Groß-Krotzenburg, *M* – Mainz, *R* – Rheinzabern, *S* – Straßburg-Königshofen, *W* – Worms, *X* – 'Unbekannt (hinsichtlich der Provenienz) 1', *Y* – 'Unbekannt 3'.[10]

Es ist notwendig, aber auch zweckmäßig von einer einzigen Klassenzerlegung auszugehen. Eine solche kann beispielsweise berechnet werden, indem man den mengentheoretischen Durchschnitt[11] aller Klassen einer Zerlegung mit allen Klassen einer

---

7 Die Messungen wurden von Herrn Priv.-Doz. Dr. GERWULF SCHNEIDER (Freie Universität Berlin, Arbeitsgruppe Archäometrie) durchgeführt. Die in die Berechnungen einbezogenen Oxid- und Spurenelementgehalte können der Tabelle 2 entnommen werden.

8 Die Angabe dieser Literaturauswahl macht eine nähere Beschreibung der agglomerativen hierarchischen Clusteranalyse überflüssig, zumal gerade sie in der Archäometrie seit längerer Zeit bekannt und etabliert ist. Andererseits ist die Auswahl der Literaturzitate so angelegt, dass sie von der allgemeinen mathematisch-statistischen bis zu einer rein archäologisch-anwendungsorientierten Darlegung reicht.

9 Zu Modellen und Anwendungen der statistischen Clusteranalyse in der Archäometrie siehe u.a. MOMMSEN, H. et al. (1988) sowie PAPAGEORGIOU, I. et al. (2001).

10 Ein ursprünglich mit 'Unbekannt 2' bezeichneter Standort hat sich inzwischen als Worms identifizieren lassen.

11 Der Durchschnitt (engl. *intersection*, russ. пересечение, Symbol $\cap$) zwischen zwei Mengen M und N ist die Menge aller (abstrakt mathematischen) Elemente, die sowohl in M als auch in N vorhanden sind. Beispielweise enthält der Durchschnitt zwischen der Menge der hier in den Ziegeln betrachten Spurenelemente $S = \{V, Cr, Ni, Zn, Rb, Sr, Y, Zr, Nb, Ba\}$ und den (chemischen) Elementen der 2. Hauptgruppe $H_{II} = \{Be, Mg, Ca, Sr, Ba, Ra\}$ nur zwei Elemente: $S \cap H_{II} = \{Sr, Ba\}$, das ist die Menge der zu den Spurenelementen der Ziegel gehörigen Erdalkalielemente.

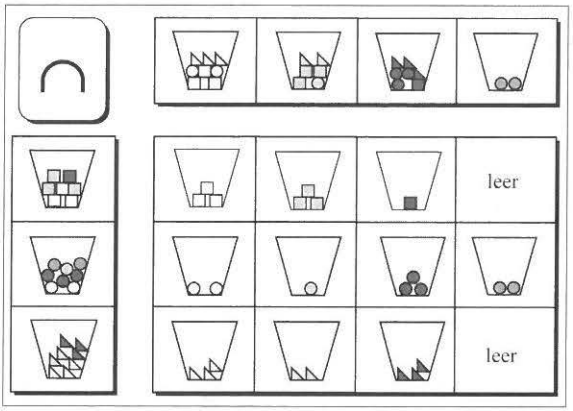

Abb. 2: Zur Erzeugung einer verallgemeinerten Klassenzerlegung mit dem ∩-Verfahren

zweiten bildet. Wie sich mathematisch zeigen lässt, stellen die nichtleeren Durchschnitte eine neue Klassenzerlegung dar, die als verallgemeinerte bezeichnet werden soll. Mit dieser ergeben die nichtleeren Durchschnitte einer eventuell weiterhin vorhandenen dritten Klassenzerlegung wiederum eine verallgemeinerte Klassifikation, etc.[12] Anstatt sukzessive vorzugehen, führt die Einbeziehung aller Informationen der einzelnen Partitionen durch eine direkte Kreuzung aller K (K>1) Partitionen zum selben Ergebnis, das sich dann anstelle einer Kreuztabelle als K-dimensionaler Datenwürfel vorstellen lässt. Eine so gewonnene verallgemeinerte Klassenzerlegung verdichtet und vereinigt gewissermaßen die Informationen aller Zerlegungen, aus der sie entstanden ist bzw. erzeugt wurde.

In Abbildung 2 ist dieses Verfahren illustriert. Hier liegen zwei Klassenzerlegungen ($K = 2$) einer Menge von verschieden gefärbten (vier Grautöne) und gestalteten (Quadrat, Kreis, Dreieck) Figuren vor, einmal nach der Färbung, zum anderen nach der Form. Die paarweise Schnittbildung ergibt zehn Klassen, die nunmehr sowohl nach Gestalt und als auch Farbe unterschieden sind.

Die Anwendung des ∩-Verfahren auf die beiden oben erwähnten Klassenzerlegungen *P1* und *P2* führt zu einer verallgemeinerten Klassifikation, die zwölf Klassen umfasst und in Abbildung 3 aufgezeigt ist. Für die allgemeine Bezeichnung ihrer Cluster ist das Symbol *Ai* gewählt worden, wobei *i* für die Clusternummer der Partition *P1* ($i = \{1, \ldots, 9\}$) und *A* für den Standort der Heeresziegeleien gemäß Partition *P2* ($A = \{F, G, M, R, S, W, X, Y\}$, s.o.) stehen. Die verallgemeinerte Klassenzerlegung

$$P_{\text{Standard}} = P1 \cap P2 = \{F1, F2, F3, G6, M8, R4, R8, R9, S7, W4, X9, Y5\},$$

dient im Weiteren als die Standardpartition, auf welche sich die Zuordnung der hinsichtlich ihrer Klassenzugehörigkeit unbekannten Ziegelproben bezieht. Es sei noch einmal betont, dass sie chemische, mathematisch-statistische und archäologische Information zugleich berücksichtigt. Eine Übersicht über die sehr unterschiedliche Anzahl von Objekten in den einzelnen Klassen[13] vermittelt Tabelle 1, wobei die Werte jeweils nach steigenden Klassenstärken sortiert sind.

---

12 Wegen der Kommutativität $M \cap N = N \cap M$ und Assoziativität $(M \cap N) \cap P = M \cap (N \cap P)$ des Durchschnitts ist die Reihenfolge bei dieser fortgesetzten Ermittlung einer verallgemeinerten Klassenzerlegung ohne Relevanz.

13 Diese Anzahlen werden auch Klassen- bzw. Clusterstärken oder Massen genannt.

P1		*P2*		P_{Standard}	
Cluster	Masse	Cluster	Masse	**Cluster**	**Masse**
5	7	M	4	**M8**	**4**
2	17	Y	7	**Y5**	**7**
3	27	W	19	**F2**	**17**
6	63	G	63	**W4**	**19**
4	68	X	67	**R9**	**20**
9	87	S	113	**F3**	**27**
1	93	F	137	**R4**	**49**
7	113	R	192	**G6**	**63**
8	127			**X9**	**67**
				F1	**93**
				S7	**113**
				R8	**123**

Tabelle 1: Übersicht über Clusterstärken (Massen)

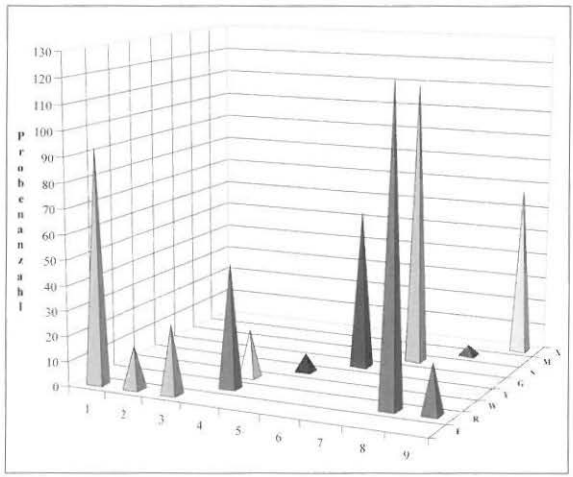

Abb. 3: Klassenstärken und die zwölf Vergleichsklassen

3. Zu identifizierende Objekte

Als Objekte, von denen die Herkunft aus einer Heeresziegelei und somit ihre Zuordnung zu einer der Klassen der Standardpartition *P*Standard unbekannt ist, wurden die zwei in Abbildung 4 gezeigten Ziegel gewählt, die folgendermaßen charakterisiert sind:

Beschreibung der zu identifizierenden Ziegel				
Bezeichnung der Probe	*Aufbewahrungsort*	*Inv.-Nr.*	*Fundort*	*Typ*
?Mainz	Direktion Archäologie	FM 99-012	Mainz	Putzhaftziegel, nicht gestempelt
?H507[14]	Antikensammlung der SMPK zu Berlin	TC 5780a	Niederbieber, Komplex 1858	*later*, Stempel anepigraphisch.

[14] Die beibehaltene Bezeichnung 'H507' geht auf die hier nicht weiter interessierende Analysensystematik von G. SCHNEIDER zurück.

In Tabelle 2 sind ihre mit der RFA ermittelten chemischen Merkmalsmuster aufgeführt, die in den im Folgenden zu beschreibenden Identifikationsmethoden benutzt wurden:

Tabelle 2: Chemische Zusammensetzung der zu identifizierenden Proben

Masseanteile der Oxide [%]									
Probe	SiO_2	TiO_2	Al_2O_3	Fe_2O_3	MnO	MgO	CaO	Na_2O	K_2O
?H507	67,63	0,48	10,67	3,86	0,09	2,51	11,57	0,54	2,46
?Mainz	66,55	0,58	12,64	3,96	0,08	2,16	9,98	1,02	2,84

Masseanteile der Spurenelemente [ppm]										
Probe	V	Cr	Ni	Zn	Rb	Sr	Y	Zr	Nb	Ba
?H507	62	92	49	61	97	255	22	123	12	349
?Mainz	72	77	37	60	109	326	31	246	10	597

Abb. 4: Der Putzhaftziegel *?Mainz* (links) und der *later ?H507* (rechts)

4. Einige Identifikationsmethoden

Im Folgenden werden einige der wichtigsten, im Prinzip bekannten Identifikationsverfahren[15] kurz erläutert und unter Zuhilfenahme des Demonstrationsbeispiels mit Hinblick auf ihre Wirksamkeit diskutiert. Das Ziel dieser Darlegungen besteht in erster Linie darin, diese Eignung abzuwägen und zumindest partiell Ausblicke zu eröffnen und aufzuzeigen, die zu deren Verbesserung führen können. Dass diese Untersuchungen aufgrund der Verwendung des speziellen Beispiels damit ebenfalls zu einer kleinen interessanten Detailaussage für die provinzialrömische Archäologie führt, ist ein ebenso gewünschtes wie erfreuliches Aperçu.

4.1. Identifikation mit der Diskriminanzanalyse

Das erste Verfahren der Zuordnung basiert auf der Diskriminanzanalyse (DA), die als Trennverfahren auch in der Archäologie und Archäometrie seit längerer Zeit durchaus etabliert ist [Bock, H. H. (1974), Ihm, P. (1978), Lebart, L. et al. (1984), Mucha, H.-J. (1992)]. Anstelle umfangreicherer, hier deswegen wohl nicht notwendiger Aus-

[15] Siehe z.B. Mommsen, H. (1981).

führungen sei nur eine treffende Definition wiederholt: *„Die Diskriminanzanalyse ist ein auf R. A. FISHER* [16] *zurückgehendes statistisches Analyseverfahren zur Trennung gewisser Objekte und Zuordnung dieser Objekte zu zwei oder mehreren Kollektiven. Diese Trennung erfolgt durch Erfassung einer bestimmten Anzahl von messbaren (zufälligen) Merkmalen dieser Objekte und durch Aufstellung einer sog. Trennfunktion (Trennformel, Diskriminanzfunktion). Es wird dabei angestrebt, durch diese Trennfunktion Entscheidungen über Zuordnungen der betrachteten Objekte zu bestimmten Kollektiven herbeizuführen."* [MÜLLER, P. H. (1970)]

In diesem Aufsatz wird die lineare Diskriminanzanalyse angewendet, die vom Modell mit gleicher Kovarianzstruktur in den Klassen ausgeht.[17] Diese Kovarianzstruktur wird durch Mittelung der Kovarianzmatrizen in den einzelnen Klassen geschätzt. Aus diesen Modellannahmen ergibt sich die sogenannte MAHALANOBIS-Distanz $d_M^2(\vec{r}_i, \vec{S}_C)$ zwischen einem Objekt i mit dem Merkmalsmuster $\vec{r}_i$ und einer Klasse C mit dem Schwerpunkt[18] $\vec{S}_C$ zu $d_M^2(\vec{r}_i, \vec{S}_C) = (\vec{r}_i - \vec{S}_C)^T V^{-1} (\vec{r}_i - \vec{S}_C)$, wobei V die geschätzte gemittelte Kovarianzmatrix ist.

Zu unterscheiden ist weiterhin die Diskriminanzanalyse mit und ohne Berücksichtigung der *a priori* Wahrscheinlichkeiten (DA^{mit} bzw. DA^{ohne}, vgl. [MUCHA, H.-J. (1992), MUCHA, H.-J. et al. (1996)]). Bei der ersteren beeinflussen zusätzlich zur MAHALANOBIS-Distanz die Massen (Probenanzahlen) der einzelnen Klassen die Klassifikationsregeln aufgrund der Schätzung der *a priori* Wahrscheinlichkeiten je Klasse durch das Verhältnis ihrer Masse zur Gesamtmasse (Anzahl aller Objekte), während ein solcher Einfluss bei der DA^{ohne} nicht vorhanden ist.

Um die Güte der Zuordnung unbekannter Objekte zu den mit den $n = 602$ Proben römischer Ziegel ermittelten Trennfunktionen in den beiden genannten Fällen (DA^{mit} bzw. DA^{ohne}) hinsichtlich der vorhandenen Unsicherheit abzuschätzen, wurde eine sogenannte *Cross-Validation* durchgeführt. Hierbei wird so vorgegangen, dass jedes der n Objekte (Proben) sukzessive hinsichtlich seiner Zuordnung zu einer Klasse als unbekannt vorausgesetzt wird. Die Trennfunktion wird sodann mit der Menge T der jeweils übrigen n-1 Objekten als Trainingssatz berechnet, um dann das Testobjekt auf dieser Grundlage einer Klasse zuzuordnen. Es kann so für jedes einzelne Objekt entschieden werden, ob die zuvor mit der Gesamtmenge von n Objekten gegebene Zuordnung wiedergefunden wird oder nicht. Die Fehlerquote ergibt sich nach der Durchführung dieser n Berechnungen als das Verhältnis der nicht verifizierten Wiederzuordnungen zur Gesamtanzahl n.

---

16 RONALD AYLMER FISHER (1890-1966), Professor an den Universitäten von London und Cambridge, arbeitete hauptsächlich auf dem Gebiet der Biometrie.

17 Die allgemeinere Annahme spezifischer Kovarianzmatrizen in den einzelnen Klassen verbietet sich hier wegen der geringen Objektanzahl in einigen Klassen und der gleichzeitig hohen Merkmalsanzahl.

18 Unter der Voraussetzung, dass alle Objekte eines Clusters C dasselbe Gewicht besitzen, berechnet sich sein Schwerpunkt oder Centroid zu $\vec{S}_C = \frac{1}{n}\sum_{j=1}^{n} \vec{r}_j$, wenn n die Clusterstärke von C und $\vec{r}_j$ der Merkmalsvektor des j-ten Objekts von C ($j = 1, \ldots, n$) sind.

Für den hier vorliegenden Fall der zwölf Klassen der Standardpartition wurde bei $n = 602$ auf diese Weise eine sehr gute Schätzung des wahren Fehlers der Zuordnung ermittelt. Sie beträgt für die DA^{mit} nur $^{21}/_{602}$, d.h. etwa 3,5%, für die DA^{ohne} $^{33}/_{602}$, also ca. 5,5% Fehler.[19] Daher kann gesagt werden, dass die Zuordnung eines zu identifizierenden Objekts mit Hilfe der DA auf jeden Fall mit einer Unsicherheit erfolgt, die höchstens 5,5% beträgt. Da für die beiden zu identifizierenden Proben sich die Zuordnung zu einer Klasse der Standardpartition mit der DA unabhängig von der Verwendung der Variante erwies, ist die Unsicherheit somit nicht größer als 3,5%. Diese Methode kann also unabhängig von der gewählten Variante als relativ zuverlässig angesehen werden.[20]

Abbildung 5 zeigt die Auftragung der zwölf Klassen der Standardpartition im Diagramm der mit der Diskriminanzanalyse berechneten 1. und 2. kanonischen Variablen und verdeutlicht die auf dieser Grundlage ermittelte Identifikation der beiden Proben *?Mainz* und *?H507*. Gefunden wurden somit sowohl mit der DA^{mit} als auch mit der DA^{ohne} mit einer statistisch durch *Cross-Validation* berechneten mittleren Sicherheit von mindestens 96,5% die Zuordnungen:

?Mainz → M8 und *?H507* → X9.

Ein Signifikanztest betreffs Klassenzugehörigkeit für jede der beiden Proben ist wegen der nicht unabhängig von den Daten erzeugten Standardpartition hier nicht angegeben und auch nicht sinnvoll. Interessant ist das auch aufgrund archäologischer Annahmen wahrscheinliche Ergebnis für *?Mainz*, da es trotz der extrem kleinen Masse von M8 sogar mit der Variante DA^{ohne} gefunden wurde.

4.2. Methode der „*k* nächsten Nachbarn"

Der nichtparametrischen kNN-Methode (*k Nearest Neighbours*) [BARTEL, H.-G. (1996), MUCHA, H.-J. (1992), SPÄTH, H. (1975)] zur Identifikation von Objekten hinsichtlich deren Klassenzugehörigkeit liegt ein einfacher Gedanke zugrunde: Die Objekte einer gegebenen Menge T, von der eine Klassenzerlegung $Z(T)$ bekannt ist, seien im Raum der Merkmale, nach denen die entsprechende Klassifikation durchgeführt wurde, angeordnet gedacht. Das zu identifizierende Objekt x wird gleichfalls in

---

19 Diese Fehlerraten sind indes im vorliegenden Fall etwas geschönt, da dieselben Daten zur Clusteranalyse benutzt worden sind und die daraus resultierende Partition *P1* in die Erzeugung der Standardpartition eingegangen ist, die hier für die Diskriminanzanalyse verwendet wurde.

20 Von den weiterhin noch gebräuchlichen Methoden zur Fehlerabschätzung sollte die *Resubstitution* genannte nicht herangezogen werden, da sie die auftretenden Fehler unterschätzt. Bei diesem Abschätzungsverfahren wird nur einmalig und zwar mit allen n Objekten des Trainingssatzes T durch die DA eine Trennfunktion berechnet. Mit dieser wird jedes Objekt der Menge T gewissermaßen noch einmal auf seine Mitgliedschaft zu der gerade ermittelten Klassenzugehörigkeit überprüft. Beträgt die Gesamtanzahl der Abweichungen f, so ergibt sich der Fehler zu $^{f}/_{n}$. Im vorliegenden Beispiel sind hiermit erhalten worden: $^{13}/_{602}$ bzw. 2,2% mit der DA^{mit} und $^{25}/_{602}$ bzw. 4,2% mit der DA^{ohne}. FAHRMEIR, L. et al. (1996) betrachtet die genannten und weitere Methoden der Fehlerabschätzung in der Diskriminanzanalyse.

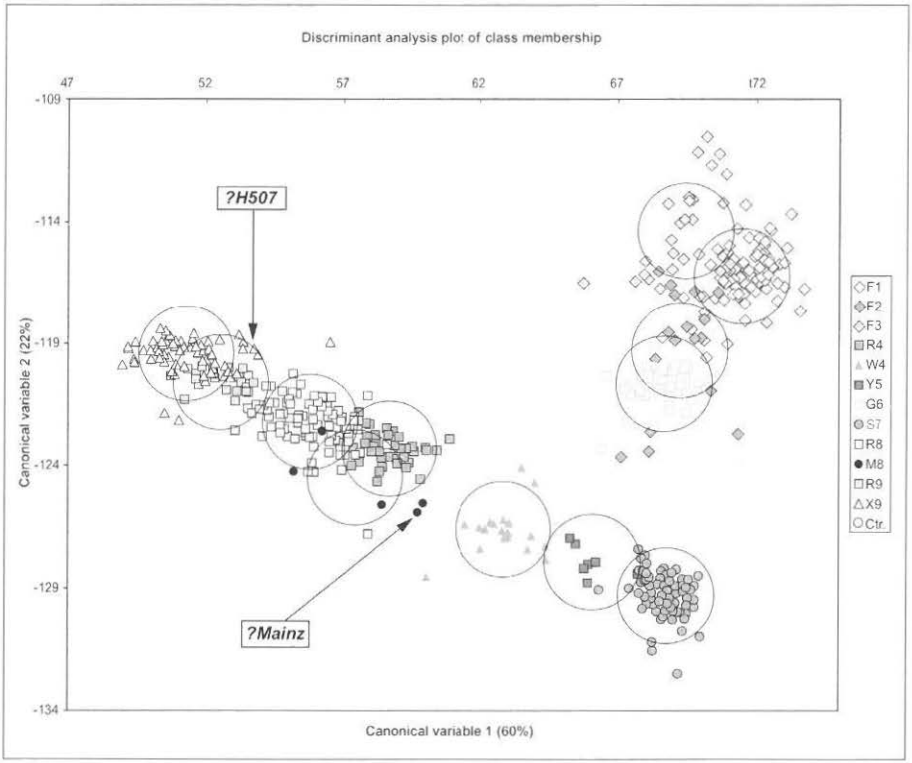

Abb. 5: Zur Zuordnung mit Hilfe der Diskriminanzanalyse

diesen Raum eingeführt, und seine hinsichtlich eines vereinbarten Distanzmaßes nächsten k (Anzahl) Nachbarn N_x werden ermittelt. Aus den Häufigkeiten, mit welchen die Klassen von $Z(T)$ bei den Objekten der Nachbarschaftsmenge N_x vorkommen, kann auf die Klassenzugehörigkeit von x geschlossen werden, wobei eine größere Häufigkeit für eine erhöhte Wahrscheinlichkeit der Zuordnung votiert.[21]

Diese Verhältnisse sind in Abbildung 6 schematisch dargestellt. Im zweidimensionalen Raum der Merkmale p und q ist die Zerlegung einer Menge von 22 Objekten (Kreise) in drei Klassen (unterschieden durch Graustufen) eingetragen. Das zu identifizierende Objekt ist durch ein Quadrat gekennzeichnet. Weiterhin sind seine fünf kleinsten euklidischen

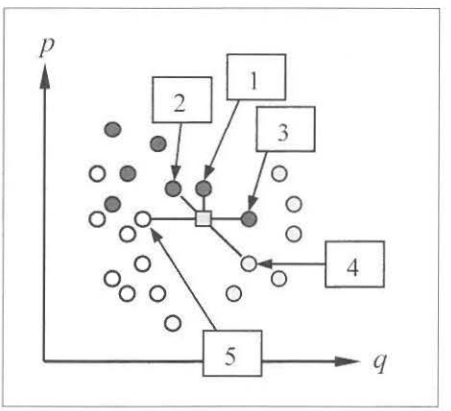

Abb. 6: Zur kNN-Methode

[21] Im Sinne einer eindeutigen Entscheidungsfindung nach dem Mehrheitsprinzip ist es zweckmäßig, für k eine ungerade Zahl zu wählen: $k = 1, 3, 5, \ldots$.

Tabelle 3: Auswertung von Abbildung 6			
nächste Nachbarn			
Anzahl	*Häufigkeit der Zugehörigkeit zur Klasse*		
(k)	● (1)	○ (2)	○ (3)
1	1	-	-
3	3	-	-
5	3	1	1

Abstände zu Objekten der als Partition gegebenen Menge in aufsteigender Folge eingezeichnet und bezeichnet. Daraus ergibt sich die Übersicht in Tabelle 3. Danach würde für $k = 1$ und $k = 3$ eine 100%ige Zugehörigkeit zur Klasse (1) gefolgert werden können, während sich für fünf nächste Nachbarn ($k = 5$) eine Zuordnungswahrscheinlichkeit zu dieser Klasse von nur 60% und für die beiden anderen Klassen (2) und (3) von jeweils 20% ablesen lässt. Auch hier wäre noch eine Zuordnung zur Klasse (1) vertretbar.

Im Demonstrationsbeispiel muss man sich die 602 Ziegelproben als Punkte im 19-dimensionalen Raum der in Tabelle 2 genannten Massenanteile der Oxide und Spurenelemente vorstellen, die gemäß der Standardpartition P_{Standard} markiert oder gefärbt sind. Für die Feststellung der kleinsten Abstände wurde wieder, wie zur Berechnung der Partition *P1* (s.o.), das gewichtete euklidische Maß verwendet. Es sind für die beiden zu identifizierenden Proben auf diese Weise die in Tabelle 4 wiedergegeben Ergebnisse erhalten worden, aus denen sich die folgenden Zuordnungen ablesen ließen:

Tabelle 4: Resultate der kNN-Methode für Ziegelproben					
Probe	*Klassenzuordnung in der Reihenfolge der kleinsten Abstände*				
	1	2	3	4	5
?Mainz	R8	R8	R8	R8	R8
?H507	X9	X9	X9	R8	X9

?Mainz → R8 und *?H507* → X9.

Die Identifikation des Ziegels ?*Mainz* mit denen, die zu der Klasse R8 gehören, ist bei dieser Methode wohl in erster Linie deren besonderer Masse geschuldet (s. auch 4.4.).

4.3. Minimaldistanzmethode

Bei der ***M***inimal***d***istanzmethode (MD) wird die gegebene Partition, deren Klassen für die Einordnung des unbekannten Objekts in Frage kommen, durch die Schwerpunkte (Zentroide) dieser Klassen vertreten. Das zu identifizierende Objekt wird der Klasse zugeordnet, zu deren Schwerpunkt es (bei festgelegtem Distanzmaß) minimalen Abstand, d.h. maximale Ähnlichkeit aufweist.[22]

Zur Identifizierung der beiden Ziegel unbekannter Herstellungsprovenienz wurde auch im Rahmen der MD wiederum die gewichtete euklidische Distanz verwendet, womit folgendes Resultat erhalten wurde:

?Mainz → M8 und *?H507* → R8,

das sich aus Tabelle 5 (s.u.) ablesen lässt.

[22] Ein kleinerer Abstand (Distanz) bedeutet kleinere Unähnlichkeit und somit größere Ähnlichkeit.

Die Klassenstärke wurde hierbei im Sinne der Berücksichtigung von *a priori* Wahrscheinlichkeiten nicht in die Entscheidung einbezogen. Natürlich könnte auch hier eine Berücksichtigung der Klassenstärken erfolgen, wie oben bei der Diskriminanzanalyse ausgeführt. Wie schon das Ergebnis mit der kNN-Methode angezeigt hat (vgl. für $k = 5$), weist die Probe *?H507* auch dort eine gewisse Ähnlichkeit zu der Rheinzaberner Klasse R8 auf. Möglicherweise besitzt dieser Ziegel an sich wegen der durch die DA wahrscheinlicher gemachten Zuordnung zu X9 (s. 4.1.) eine größere Ähnlichkeit zu dieser Klasse. Da die Masse von X9 nur etwa die Hälfte der Masse von R8 ausmacht, erhält das mit der DA^{mit} ermittelte Resultat X9 noch mehr Glaubwürdigkeit.

4.4. Sukzessive Entscheidung auf der Basis der hierarchischen Clusteranalyse

Ein weiteres Verfahren (DCA) basiert auf folgender Idee des ***d***ivisiv-alternativen Durchlaufens eines mit der agglomerativen hierarchischen ***C***luster***a***nalyse gewonnenen Dendrogramms:

Die Klassen der Standardzerlegung werden selbst als Objekte einer Menge angesehen,[23] die einer agglomerativen hierarchischen Clusteranalyse unterworfen wird, wobei der Schwerpunkt einer jeden Klasse das Merkmalsmuster darstellt, das sie als Objekt charakterisiert. Das Ergebnis lässt sich wie gewöhnlich mit dem Vorteil der größeren Anschaulichkeit als Dendrogramm wiedergeben, eines Graphen mit – wie der Name sagt –[24] baumartiger Struktur.

Alle Knoten dieses Baumes stellen ebenfalls Cluster dar, denen – wie den Blättern – sich ihr Schwerpunkt als Merkmalsmuster zuordnen lässt. Zu allen diesen Knoten bzw. Clustern (außer der Wurzel) wird nun der Abstand zu dem Objekt berechnet, dessen Zugehörigkeit zu einer der Klassen der Standardzerlegung, d.h. den Blättern, ermittelt werden soll. Auf diese Weise lassen sich die Knoten mit diesen Distanzen bewerten.

Die Aufgabe der Identifikation des hinsichtlich seiner Klassenzugehörigkeit unbekannten Objektes zu einem der Cluster der Standardzerlegung kann nun als dessen „Entlangfahren" auf dem Dendrogramm in divisiver Richtung aufgefasst werden, also von der Wurzel bis zu einem der Blätter, welches dann die ihm ähnlichste Klasse anzeigen sollte. Der Weg dorthin wird so ermittelt, indem in jedem erreichten Knoten – mit der Wurzel beginnend – als dessen Nachfolger derjenige ausgewählt wird, der mit dem kleineren Distanzwert, d.h. der größeren Ähnlichkeit zum unbekannten Objekt, markiert ist. Auf diese Weise wird eines der Blätter erreicht.[25]

---

23 Bei einer Menge von Mengen, die hier durch die Klassen gegeben sind, spricht man auch von einem Mengensystem.

24 τὸ δένδρον (griech.) Baum. In der Sprechweise der Graphentheorie ist ein 'Baum' ein solcher (zusammenhängender) Graph, der nur Verzweigungen und keine Zyklen enthält. Denkt man sich die Kanten des Dendrogrammgraphen mit der Richtung des fortschreitenden Agglomerationsprozesses versehen, so mögen die Knoten, aus denen jeweils nur eine Kante herausläuft, 'Blätter' und derjenige, in den nur Kanten einlaufen, die 'Wurzel' heißen.

25 Bei der Verwendung von Daten, die aus der Praxis stammen, wird die beschriebene Vergehensweise zu einem eindeutigen Resultat führen.

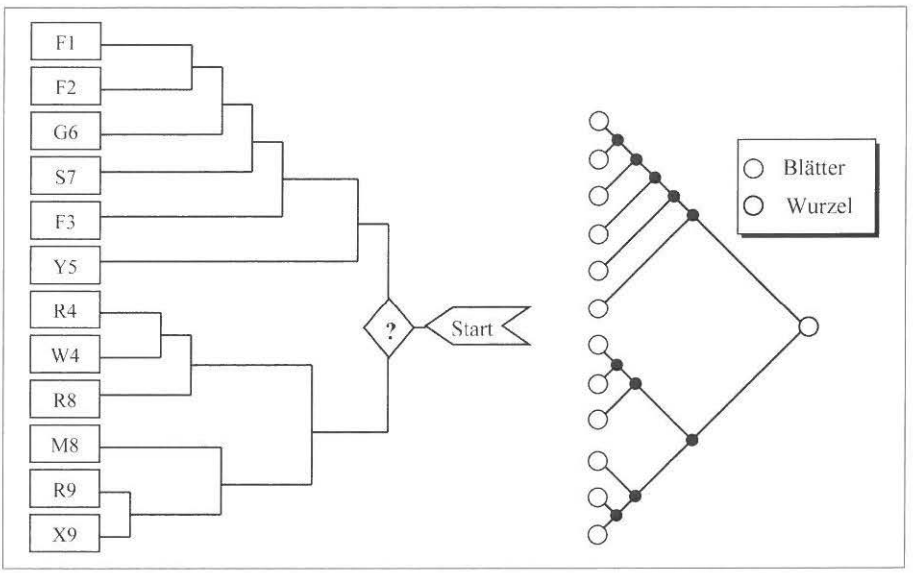

Abb. 7: Dendrogramm und zugehöriger Graph der Clusteranalyse der Standardpartition

Im Demonstrationsbeispiel stellen die zwölf Cluster der Standardpartition P_{Standard} die Blätter dar. Das bei Verwendung der agglomerativen hierarchischen Clusteranalyse nach WARD erhaltene Dendrogramm und die ihm entsprechende Darstellung, wie sie in der Graphentheorie gebräuchlicher ist, zeigt Abbildung 7.

Durchläuft man das Dendrogramm nach Markierung mit den entsprechenden Abstandswerten in der geschilderten Weise mit den zu identifizierenden Objekten *?H507* bzw. *?Mainz*, so werden die in Abbildung 8 eingezeichneten Wege durchlaufen. Der Übersichtlichkeit halber wurden nur die jeweils interessierenden Distanzwerte eingetragen.[26] Wie aufgezeigt, erreichen beide Proben dieselbe archäologisch zu Rheinzabern gehörige Klasse R8. Mit dem Blick auf die Ergebnisse anderer Identifikationsverfahren, insbesondere das in höherem Grade zuverlässigere, das auf der Diskriminanzanalyse (DA^{mit}, DA^{ohne}) basiert, ist insbesondere die Zuordnung von *?Mainz* zu R8 fragwürdig. In gewissem Umfang trifft das auch für *?H507* zu, denn die Zuordnung dieser Probe zu R8 steht in Konkurrenz mit derjenigen zu X9, die aufgrund anderer Identifikationsresultate die größere Glaubwürdigkeit besitzt.[27]

Zur Unterstützung dieser Aussagen sei Tabelle 5 betrachtet. In ihr sind die sechs Cluster bzw. Blätter, die nach der ersten Entscheidung für den weiterhin zu durchlaufenden Zweig bzw. Weg allein in Frage kommen (s. Abb. 8), und ihre Abstände zu den zu identifizierenden Objekten aufgeführt. Danach hat *?Mainz* zu M8 die kleinste Distanz auf diesem Zweig, und sie ist es auch im absoluten Sinne. Das entspricht

[26] Es wurde das 10^3-fache des mit standardisierten Merkmalswerten berechneten quadrierten euklidischen Abstands angegeben.

[27] S. hierzu auch die Bemerkungen zur MD im Abschnitt 4.3.

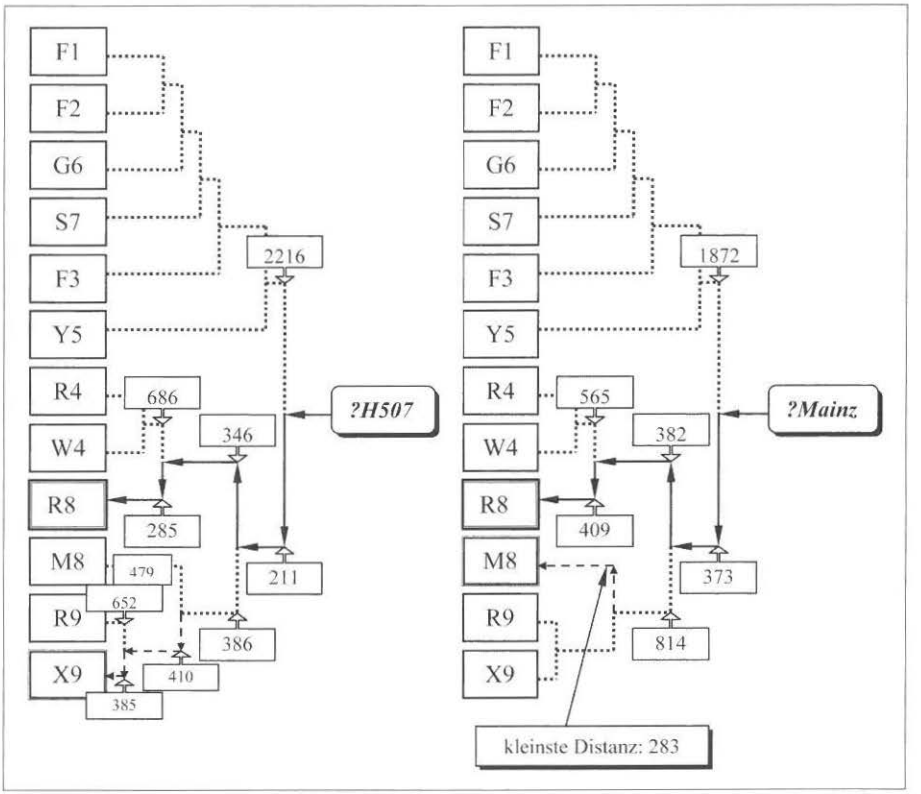

Abb. 8: Zuordnungswege im Dendrogramm

Tabelle 5: Abstände der zu identifizierenden Objekte zu Clusterschwerpunkten

?H507	
Cluster	***Distanz***
R8	285
X9	385
M8	479
R4	633
R9	652
W4	949

?Mainz	
Cluster	***Distanz***
M8	283
R8	409
W4	538
R4	624
X9	881
R9	976

nicht nur den Methoden DA^{mit}, DA^{ohne} und MD, sondern auch der archäologischen Hypothese, dass Fundort und Herstellungsort des zur Probe gehörigen Ziegels Mainz sind. Im Falle von *?H507* hat der Abstand zu R8 die geringere Distanz als zu X9, das aber durch die anderen Identifikationsmethoden als wahrscheinlicher gelten darf.

Offensichtlich „irrten" sich beide Proben auf ihrem Weg zu den Blättern an derselben Stelle, wie die Verhältnisse von Abbildung 8 lehren. Es ist dies gerade dort, wo die Klasse mit der größten Masse R8 ins Spiel kommt, gegen welche die mittelgroße von

X9 und erst recht die extrem kleine von M8 (s. Tabelle 1) deutlich abfallen. In Abbildung 8 ist zu erkennen, dass die alternative Entscheidung an dieser Verzweigung zu dem jeweils befriedigenderen Zuordnungsresultat geführt hätte. Die Erfahrung dieses Beispiels ist es somit, dass das Entscheidungskriterium der kleineren Distanz in bestimmten Knoten versagen kann, in diesem Sinne also zu grob ist.

Trotz ihres offensichtlich einsichtigen Grundgedankens führt die DCA-Methode in der hier geschilderten Form – wie das Beispiel lehrt – zu Resultaten, die nur mit Vorsicht in die Diskussion um die Identifikation der Klassenzugehörigkeit von Objekten aufgenommen werden dürfen. Gerade aber wegen dieser anschaulichen Klarheit verdient das Verfahren die Mühe des Erarbeitens einer zukünftigen Modifizierung, die diesen Mangel beseitigt oder zumindest auf ein hohes Maß einschränkt. Dabei wird die Berücksichtigung der Clusterstärken (Massen), etwa durch die Wahl eines geeigneten Distanzmaßes,[28] ebenso betrachtet werden müssen wie die Schaffung über die lokale Betrachtung hinausgehender oder gewissermaßen „vorausschauender" Entscheidungskriterien.

5. Zusammenfassung und archäologische Anmerkungen

In Tabelle 6 sind die Identifikationsergebnisse mit den in den Abschnitten 4.1. bis 4.4. erläuterten Verfahren zusammengestellt.

Tabelle 6: Identifikationsresultate an zwei römischen Ziegeln

Probe	*Zuordnung durch Verfahren*						
	DA^{mit}	DA^{ohne}	MD	kNN mit k =			DCA
				1	3	5	
?Mainz	M8	M8	M8	R8	3 R8	5 R8	R8
?H507	X9	X9	R8	X9	3 X9	3 X9, R8, X9	R8

Auf den ersten Blick ist erkenntlich, dass bei Annahme einer Gleichberechtigung der Methoden für beide Proben bzw. Ziegel kein eindeutiges Ergebnis der Zuordnung zu einer der Standardpartitionen gefunden werden kann. Wie aber im Abschnitt 4 dargelegt wurde, besteht eine derartige Gleichwertigkeit im absoluten Sinne nicht. So ist von den untersuchten Verfahren der DA, insbesondere der DA^{mit}, die größte Zuverlässigkeit zuzuschreiben. Die recht große Sicherheit lässt sich für diese Methoden auch numerisch ausdrücken Die mit den anderen Verfahren gewonnenen Ergebnisse hängen dagegen mehr oder weniger von den Massen der wegen ihrer Ähnlichkeit zu den zu identifizierenden Objekten in Betracht kommenden Klassen ab. Das hat sich deutlich bei der DCA[29] gezeigt und im Falle von *?Mainz* wegen der extremen Kleinheit von M8 auch bei der kNN-Methode.

[28] Der genannte Einfluss lässt sich beispielsweise daran zeigen, dass die Reihenfolge der in Tabelle 5 für *?Mainz* aufgeführten Werte bei Verwendung des MAHALANOBIS-Maßes verändert wird: 1. zu M8 (22, 48), 2. zu W4 (26,94).

[29] Die Unzulänglichkeiten, welche die von ihrem Grundgedanken eigentlich verständliche und sowohl praktische wie elegante Methode in der dargestellten ursprünglichen Form offensichtlich besitzt, durch entsprechende Abänderungen zu beseitigen, wird – gerade wegen dieses Vorteils der Anschaulichkeit – eine zukünftige Arbeitsaufgabe sein.

Die Darlegungen lehren somit, dass nur die Benutzung mehrerer unabhängiger Identifikationsmethoden und der Vergleich sowie die kritische Auswertung der erhaltenen Resultate zu einer einigermaßen gesicherten Aussage über die Klassenzugehörigkeit des untersuchten Objektes führen. Für die hier betrachteten Proben kann auf diese Weise folgende Zuordnung bzw. Identifikation vorgeschlagen werden:

Probe	*Zuordnung zu Ziegelei*	
	primär	*weiterhin zu berücksichtigen*
?Mainz	in Mainz (Klasse M8)	in Rheinzabern (Klasse R8)
?H507	unbekannter Provenienz № 1 (Klasse X9)	in Rheinzabern (Klasse R8)

Der Putzhaftziegel *?Mainz* weist keinen Ziegelstempel auf, der ihn als Produkt einer Heeresziegelei ausweisen würde. Ungestempelte Baukeramik ist indes sehr häufig.[30] Neben der Annahme, dass etwa ein Dutzend gleicher Ziegel vom selben Fundort und selben Befund zufällig allesamt zum nicht gestempelten Anteil einer sonst durch eine Militäreinheit produzierten Ziegelmenge stammen, ist durchaus auch die Herstellung am Verbauungsort Mainz in Betracht zu ziehen. Erst jüngst konnte durch epigraphische Studien und geochemische Reihenuntersuchungen gestempelter Ziegel die Annahme einer Heeresziegelei für Mainz ausgeschlossen werden. Die mathematische Zuweisung eines Ziegels an Mainz war aus archäologischer Sicht also unerwartet; macht aber durchaus Sinn: nur geringe Quantitäten dieser Bekleidziegel für bemalte Fachwerkwände und -decken wurden für kleinere Bauten benötigt und könnten deshalb durchaus jeweils vor Ort hergestellt worden sein. Erstmals wurde somit römische Ziegelherstellung für MOGONTIACUM (Mainz) wahrscheinlich. Offen bleibt, ob die qualitativ guten Produkte in privatwirtschaftlichem Kontext hergestellt wurden oder eine quasi staatliche Beschaffung durch Einheiten des Heeres und auf Befehl des Statthalters erfolgte.[31]

Auch die mathematische Zuweisung des anepigraphischen Ziegelstempels aus Niederbieber als Produkt der bislang nicht lokalisierten Herstellungsprovenienz 'Unbekannt 1' ist für die Militärarchäologie des späten 2. oder frühen 3. Jahrhunderts n. Chr. von Relevanz. Grundsätzlich lag keine bestimmte archäologische Erwartung vor. Der untersuchte Fundkomplex von 1858 umfasst ein für den Fundort repräsentatives Ensemble von Ziegelstempeln der 8. und 22. Legion sowie der 4. Vindelikerkohorte. Die historischen Rahmendaten für den Militärstandort Niederbieber sind die Gründung unter COMMODUS (161-192, Kaiser 180-192) bald nach 185 n. Chr. und die Zerstörung des Platzes, die um 260 n. Chr. erfolgte. Alle Baumaterialien sind dort regulär innerhalb dieses Zeitansatzes beschafft und verbaut worden. Für Ziegel der Herstellungsprovenienz 'Unbekannt 1' liegt bislang lediglich ein epigraphischer Datierungsansatz für die Stempeltypen der ANT-Gruppe der 22. Legion zwischen 198 und 222 n. Chr. vor. Dieser wird durch einen Fundkomplex vom Eisgrubweg in Mainz gestützt, bei dem die Stempeltypen 7 und 8 der Eisgrubgruppe (geochemisch als

30 Vom Fundort Mainz sind auch Stempel auf solchen seltenen Spezialziegeln vorhanden.

31 Ausführlicher zur historischen Relevanz: DOLATA, J. et al. (2001a) 107-108. Am Fundort (Mainz, Dagobertstraße 2, Handwerkskammer Rheinhessen) wurde eine Vitrine mit den untersuchten Ziegeln und ausführlicher Erläuterungstafel installiert.

Produkte der Herstellungsprovenienz 'Unbekannt 1' erwiesen) mit Groß-Krotzenburger Produkten vergesellschaftet sind. Für den speziellen sekundären Baubefund wird ein *terminus post quem* 213 n. Chr. angenommen.[32] Aus dem untersuchten Niederbieberer Fundkomplex wurden Ziegelstempel der 22. Legion geochemisch analysiert; unter anderen auch der Stempeltyp 8 der Eisgrubgruppe. Der nunmehr hier neu zugeordnete anepigraphische Ziegelstempel aus Niederbieber erweitert unsere Kenntnis der späten Ausbau- und Reparaturphase des obergermanisch-rätischen Limes. Schritt für Schritt gewinnen wir durch Klärung der Produkt- und Stempelspektren der einzelnen Ziegeleiorte ein differenziertes Verständnis der Produktionszusammenhänge antiker Heereslogistik im Arbeitsgebiet.

Die letzten Bemerkungen verdeutlichen zum wiederholten Male den Inhalt und die Möglichkeiten mathematisch-statistischer Datenanalysen: Sie liefern wichtige und zugleich fundierte Hinweise für Entscheidungsfindungen, die Bildung oder Unterstützung von Hypothesen etc., die – so wertvoll und eigentlich unentbehrlich sie auch sind – doch keine Beweise im strengen Sinn darstellen. Die eigentlich verbindliche Aussage muss der Experte, hier der Archäologe, treffen. Aber sie wird ihm durch die Nutzung mathematischer Methoden und durch die Zusammenarbeit mit Fachleuten der explorativen Datenanalyse auf jeden Fall erleichtert, wenn nicht überhaupt ermöglicht.

Literatur

BARTEL, H.-G. (1996): *Mathematische Methoden in der Chemie*. (Heidelberg – Berlin), S. 330-340.

BARTEL , H.-G., DOLATA, J., MUCHA, H.-J. (2000): Klassifikation gestempelter römischer Ziegel aus Obergermanien. *Archäometrie und Denkmalpflege*, 86-88.

BARTEL , H.-G., MUCHA, H.-J., DOLATA, J. (2001): Parametrische und nichtparametrische Identifikationsmethoden, dargestellt am Beispiel römischer Baukeramik aus Obergermanien. *Archäometrie und Denkmalpflege*, 104-106.

BARTEL , H.-G., MUCHA, H.-J., DOLATA, J. (2002): Automatische Klassifikation in der Archäometrie: Berliner und Mainzer Arbeiten zu oberrheinischen Ziegeleien in römischer Zeit. *Berliner Beiträge zur Archäometrie* 19, 31-62; 20 (2003), 201-202.

BARTEL , H.-G., MUCHA, H.-J., DOLATA, J. (2003): Über eine Modifikation eines graphentheoretisch basierten partitionierenden Verfahrens der Clusteranalyse. *Match* 48, 209-223.

BOCK, H. H. (1974): *Automatische Klassifikation*. (Göttingen).

DOLATA, J. (1994): Custos Castelli Figlinarum. *Mainzer Arch. Zeitschr.* 1, 67-72.

DOLATA, J. (1996): Hin zu einer archäologischen Nutzanwendung geochemischer Analytik römischer Baukeramik. *Mainzer Arch. Zeitschr.* 3, 105-125.

DOLATA, J. (1997-2001): Inventarisation und Forschung zu archäologischem Quellenmaterial: Römische Ziegelstempel aus Mainz – Erreichtes und Perspektiven. *Denkmalpflege in Rheinland-Pfalz* 52-56, 488-494.

[32] Die komplexe Methode wird am betreffenden Beispiel umrissen: DOLATA, J. (1998). Ausführlich: DOLATA, J. (2000b).

DOLATA, J. (1998): Archäologische und archäometrische Untersuchung an römischer Baukeramik und Ziegelstempeln. *Archäometrie und Denkmalpflege*, 93-95.

DOLATA, J. (1999): Ingenieurtechnische Untersuchung an antiken Ziegelsteinen aus Mainz: Interdisziplinäre Erforschung römischer Baukeramik und Ziegelstempel. *Ziegel Zeitschrift* 6, 421-423.

DOLATA, J. (2000a): Antike Ziegelsteingeschichten aus Mogontiacum. *Mainz* 1, 95-100.

DOLATA, J. (2000b): *Römische Ziegelstempel aus Mainz und dem nördlichen Obergermanien – Archäologische und archäometrische Untersuchungen zu chronologischem und baugeschichtlichem Quellenmaterial*. (Inauguraldissertation, Johann Wolfgang Goethe-Universität, Frankfurt/M.).

DOLATA, J. (2001a): Römische Ziegelstempel der sogenannten Flörsheimer Gruppe. *Flörsheimer Geschichtshefte* 3, 4-7.

DOLATA, J. (2001b): Die spätantike Heeresziegelei von Worms – Ein Beitrag zur Geschichte der legio XXII Primigenia aufgrund ihrer Ziegelstempel. *Der Wormsgau* 20, 43-77.

DOLATA, J. (2002): Zum Fund eines spätantiken Militärziegelstempels in einer römischen Villa in Weinsheim, Kreis Bad Kreuznach. *Archäologie in Rheinland-Pfalz*, 107-109.

DOLATA, J. (2003): Technisch versiertes Handwerk aus dem römischen Mainz. *Archäologie in Rheinland-Pfalz*, 47-48.

DOLATA, J. (2004): Eine sehr große Dachplatte mit Stempel der 30. Legion in Mainz, Zu Abmessungen von Dachziegeln des 1. bis 4. Jahrhunderts am nördlichen Oberrhein. *Archäologisches Korrespondenzblatt* 34, 519-529.

DOLATA, J., WERR, U. (1998/1999): Wie gleich ist derselbe? – Homogenität eines römischen Ziegels und Aussagegrenzen geochemischer Analytik aufgrund von Messtechnik und Materialvarietät. *Mainzer Arch. Zeitschr.* 5/6, 129-147.

DOLATA, J., MUCHA, H.-J., BARTEL, H.-G. (2001): Untersuchungsperspektiven zum kombinierten archäologischen und materialanalytischen Nachweis römischer Ziegelherstellung in Mainz und Worms. *Archäometrie und Denkmalpflege*, 107-109.

DOLATA, J., MUCHA, H.-J., BARTEL, H.-G. (2003a): Archäologische und mathematisch-statistische Neuordnung der Orte römischer Baukeramikherstellung im nördlichen Obergermanien. *Xantener Berichte* 13, 381-409.

DOLATA, J., BARTEL, H.-G., MUCHA, H.-J. (2003b): Statistische Untersuchung zur Aufklärung der Binnenstruktur römischer Ziegelproduktion von Frankfurt-Nied. *Archäometrie und Denkmalpflege*, 40-42.

DOLATA, J., MUCHA, H.-J., BARTEL, H.-G. (2004): Eine Anwendung der hierarchischen Clusteranalyse: „Unbekannt 1" als Provenienz einer bekannten obergermanischen Heeresziegelei? *Archäometrie und Denkmalpflege*, 28-30.

FAHRMEIR, L., HAMERLE, A., TUTZ, G. (Hrsg.) (1996): *Multivariate statistische Verfahren*. (Berlin), S. 369-371.

IHM, P. (1978): *Statistik in der Archäologie*. (= Archaeo-Physika Band 9). (Köln – Bonn), S. 452-463.

LEBART, L., MORINEAU, A., FÉNELON, J.-P. (1984): *Statistische Datenanalyse. Methoden und Programme*. (Berlin).

MOMMSEN, H. (1981): Filters to sort out pottery samples of the same provenience from a data bank of neutron activation analyses. *Archaeometry* 23 (2), 209-215.

MOMMSEN , H., KREUSER, A., WEBER, J. (1988): A method for grouping pottery by chemical composition. *Archaeometry* 30 (1), 47-57.

MUCHA, H.-J. (1992): *Clusteranalyse mit Mikrocomputern*. (Berlin), S. 95-118.

MUCHA, H.-J., BOCK, H. H. (1996): Classification and multivariate graphics: models, software and applications. *WIAS Report* 10 (Berlin).

MUCHA, H.-J., DOLATA, J., BARTEL, H.-G. (2000/2001):[33] Klassifikation von 613 Proben als Referenzen für die Herstellungsprovenienzen römischer Baukeramik im nördlichen Obergermanien. *Mainzer Arch. Zeitschr.* 7/8, 275-300, im Druck.

MUCHA, H.-J., DOLATA, J., BARTEL, H.-G. (2001): Validation of Results of Cluster Analysis of Roman Bricks and Tiles. In: W. GAUL, G. RITTER (eds.): *Classification, Automation, and New Media*. (Proceedings of the 24th Annual GfKl Conference, Passau, March 2000), (Berlin - Heidelberg - New York), S. 471-478.

MUCHA, H.-J., BARTEL, H.-G., DOLATA, J. (2003a): Core-Based Clustering Techniques: Methods, Software, and Applications. In: M. SCHADER, W. GAUL, M. VICHI (eds.): *Between Data Science and Applied Data Analysis*. (Proceedings of the 26th Annual Conference of the Gesellschaft für Klassifikation, Mannheim, July 2002), (Berlin – Heidelberg – New York), S. 74-82.

MUCHA, H.-J., BARTEL, H.-G., DOLATA, J. (2003b): Modellbasierte Clusteranalyse römischer Ziegel aus Worms und Rheinzabern. *Archäologische Informationen* 26 (2), 471-480.

MUCHA, H.-J., BARTEL, H.-G., DOLATA, J. (2005a): Modelbased Cluster Analysis of Roman Bricks and Tiles from Worms and Rheinzabern. In: C. WEIHS, W. GAUL (eds.): *Classification: The Ubiquitous Challenge*. (Proceedings of the 28th Annual GfKl Conference, Dortmund, March 2004), (Berlin – Heidelberg – New York), 317-324.

MUCHA, H.-J., BARTEL, H.-G., DOLATA, J. (2005b): Techniques of Rearrangements in Binary Trees (Dendrograms) and Applications. *Match* 54, 561-582.

MÜLLER, P. H. (Hrsg.) (1970): *Lexikon Wahrscheinlichkeitsrechnung und mathematische Statistik*. (Berlin), S. 32.

PAPAGEORGIOU, I., BAXTER, M. J., CAU, M. A. (2001): Modelbased cluster analysis of artefact compositional data. *Archaeometry* 43 (4), 571-588.

SPÄTH, H. (1975): *Cluster-Analyse-Algorithmen zur Objektklassifizierung und Datenreduktion*. (München - Wien), S. 162-189.

SWART, C. (2005): *Analyse römischer Ziegel mit ICP- OES /-MS: Methodenvergleich zwischen RFA und ICP*. (Inauguraldissertation, Freie Universität, Berlin).

SWART, C., PAZ, B., DOLATA, J., SCHNEIDER, G., SIMON, J., BARTEL, H.-G., MUCHA, H.-J. (2004): Analyse römischer Ziegel mit ICP-MS/-OES: Methodenvergleich zwischen RFA und ICP. *Archäometrie und Denkmalpflege*, 25-27.

WERR, U. (1998): Grenzen der Aussagekraft chemischer Analytik für römische Baukeramik. *Archäometrie und Denkmalpflege*, 96-98.

[33] Aktuelle Reihenfolge der Autoren: BARTEL, H.-G., DOLATA, J., MUCHA, H.-J.