

5 Statistische Auswertung der Ergebnisse

5.1 Theoretische Grundlagen¹

Bei großen Probenserien mit vielen gemessenen Größen ist eine Auswertung von Hand nicht mehr möglich. Deshalb wurden verschiedene statistische Methoden entwickelt, die eine differenzierte Datenauswertung ermöglichen, indem aus der Gesamtheit der vorliegenden Datenmenge einzelne Gruppen mit ähnlichen Merkmalen wie Farbe, Elementkonzentration etc. unterschieden werden können. Hier soll nur auf die in dieser Arbeit verwendeten Methoden näher eingegangen werden. Alle Methoden setzen voraus, dass die jeweiligen Variablen Werte in der gleichen Größenordnung aufweisen. Deshalb lassen sich Hauptelemente, die üblicherweise in Prozent angegeben werden, und Spurenelemente, die in $\frac{\text{mg}}{\text{kg}}$ (ppm) gemessen werden, nicht ohne weiteres vergleichen. Um dennoch alle Elemente in der Statistik verwenden zu können, müssen die Ergebnisse standardisiert werden.

Hierarchische Clusteranalyse

In der hierarchischen Clusteranalyse werden agglomerative und divisive Verfahren unterschieden. Die agglomerativen Verfahren gehen von den Beobachtungen als einzelne Klassen aus und bilden dann je nach Ähnlichkeit der Fälle (hier Proben) Gruppen. Die divisiven gehen den umgekehrten Weg von der trivialen Gruppe zu den einzelnen Fällen. In dieser Arbeit wurden u.a. die Methoden des „Average Linkage“ und der hierarchischen Clusteranalyse nach Ward als agglomerative Verfahren verwendet. Im ersten Schritt dieser Verfahren werden zwei Fälle zu einem Cluster vereinigt. Im zweiten Schritt wird dann je nach Distanz dem Cluster ein weiterer Fall hinzugefügt oder zwei Fälle werden zu einem neuen Cluster vereinigt. In den weiteren Schritten werden dann entweder den Clustern neue Fälle hinzugefügt oder auch Cluster, deren Abstand der Mittelwerte geringer ist als der Abstand zu einer noch ungruppierten Probe, vereinigt. So entsteht eine Hierarchie der Clusterbildung abhängig von den Distanz- und Ähnlichkeitsmaßen der Objekte. Als Abstandsmaß wird der quadrierte euklidische Abstand verwendet.

$$\text{EUCLID}(x, y) = \sqrt{\sum_i (x_i - y_i)^2}$$

x, y: einzelne Proben

i: Variablen der Probe

In so genannten Dendrogrammen kann die schrittweise Vereinigung der Proben und Cluster nachvollzogen werden. Eine sinnvolle Gruppeneinteilung lässt sich im Dendrogramm an der Stelle erkennen, an der die Distanzkoeffizienten bei der Vereinigung der Cluster sprunghaft ansteigen.

Bei der hierarchischen Clusteranalyse nach Ward basiert die Zuordnung einer Probe zu einer Gruppe auf der geringsten Erhöhung der Fehler-Quadratsumme. Ein Nachteil aller hierarchischen Methoden besteht darin, dass die Zuordnung zu einem Cluster endgültig ist und im Verlauf der Analyse nicht überprüft wird².

Clusterzentrenanalyse

Die Clusterzentrenanalyse nach der K-means Methode geht von einer vorgegebenen Anzahl an Clustern aus und ordnet diesen die Fälle zu. Die sinnvolle Anzahl der Gruppen lässt sich aus einem Dendrogramm der hierarchischen Clusteranalyse entnehmen. Dabei wird zu Beginn jeder Fall einem Cluster zugeordnet. Danach werden die Zentroide der Cluster und der quadratische euklidische Abstand der Fälle zum Zentroid berechnet, ebenso wie deren Abstände zu den anderen Gruppen. Jeder Fall wird dann dem Cluster zugeordnet, zu dessen Mittelpunkt er den geringsten Abstand aufweist. Danach werden Zentroide und Abstände neu berechnet. Dieses Vorgehen wird bis zu einem definierten Abbruchkriterium wiederholt.

Two-Step Clustering³

SPSS bietet neben den üblicherweise verwendeten hierarchischen Clustermethoden sowie der Clusterzentrenanalyse noch die Möglichkeit einer zweistufigen Klassifizierung. Im ersten Schritt wird die im Datensatz enthaltene sinnvolle Anzahl der Klassen ermittelt. Dies geschieht anhand des BIC (Bayesian Information Criterion). Im zweiten Schritt wird die Anzahl der Cluster und die Zuordnung der Proben zu den Clustern in jedem Schritt der hierarchischen Clusteranalyse durch die größtmögliche Änderung des Abstandes zwischen zwei benachbarten Clustern bestimmt.

Der Vorteil der Clusterzentrenanalyse und des Two-Step Clusterings besteht darin, dass die Proben im Laufe des Verfahrens auch die Gruppen wechseln können und somit eine bessere Zuordnung als bei den hierarchischen Methoden möglich ist.

Diskriminanzanalyse

Die Diskriminanzanalyse berechnet bei bekannter Gruppenzugehörigkeit der Fälle aus den Merkmalsvariablen durch Linearkombination eine oder mehrere Diskriminanzfunktionen, an Hand derer weitere ungruppierte Fälle einer Gruppe zugeordnet werden können.

$$Y = b_0 + b_1 \cdot X_1 + b_2 \cdot X_2 + \dots + b_J \cdot X_J$$

Die Diskriminanzfunktion ordnet nun jedem Individuum (und jeder Gruppe) einen Wert auf der Diskriminanzvariable Y zu. Um eine gute Trennung zu erreichen, werden die Parameter b_j der Diskriminanzfunktion so bestimmt, dass sich die Gruppen auf der Diskriminanzvariable Y maximal unterscheiden. Dabei soll die Streuung der Verteilung in den Gruppen möglichst klein sein. Die gesamte Streuung der Diskriminanzwerte setzt sich zusammen aus der Streuung zwischen den Gruppen und der Streuung innerhalb der Gruppen. Die Diskriminanzkoeffizienten werden derart bestimmt, dass der Quotient aus den Streuungen maximiert wird:

$$\frac{SAQ_{\text{zwischen}}}{SAQ_{\text{innerhalb}}} = \max$$

SAQ: Summe der Abweichungsquadrate vom Mittelwert

Als Maß für die Güte der Trennung in Gruppen wird Wilks' Lambda verwendet:

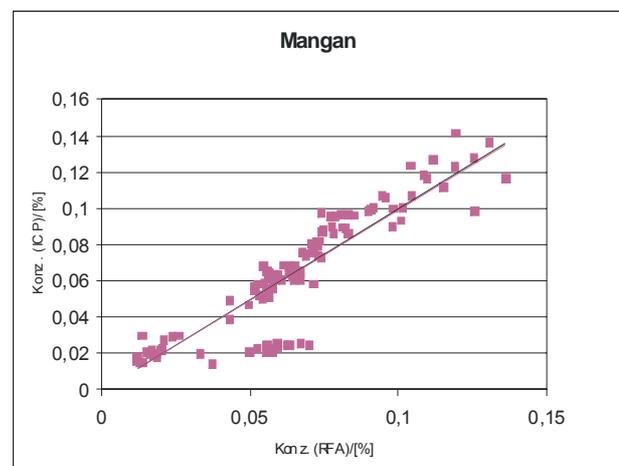
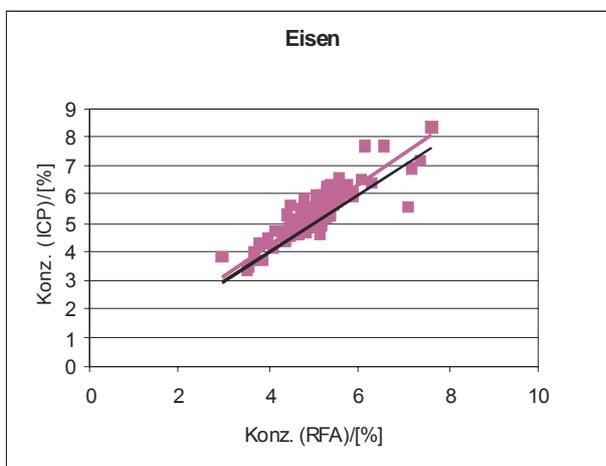
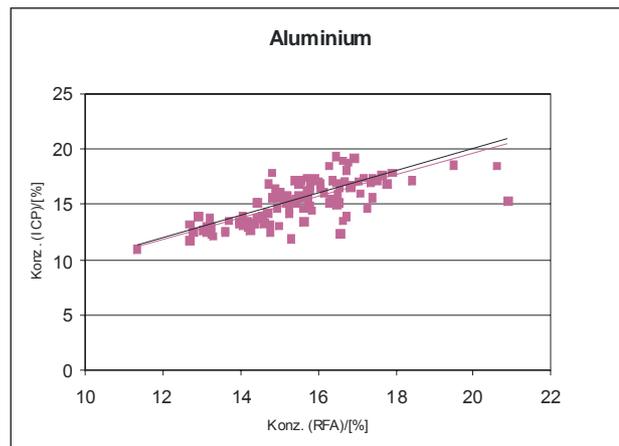
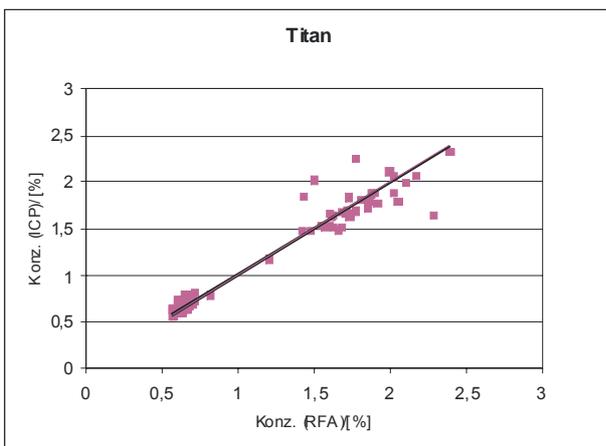
$$\lambda = \frac{SAQ_{\text{innerhalb}}}{SAQ_{\text{total}}} = \frac{\text{nicht erklärte Streuung}}{\text{gesamte Streuung}}$$

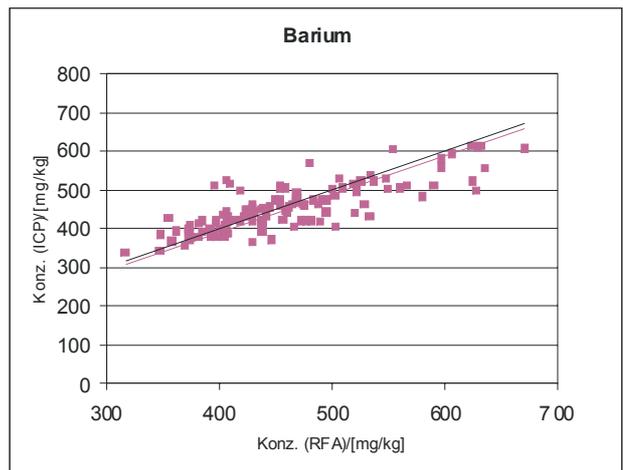
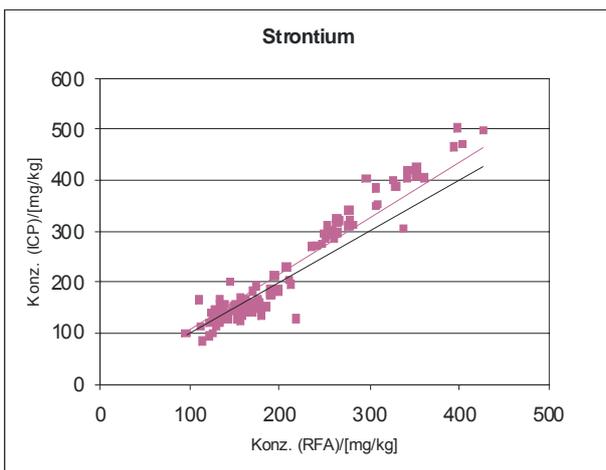
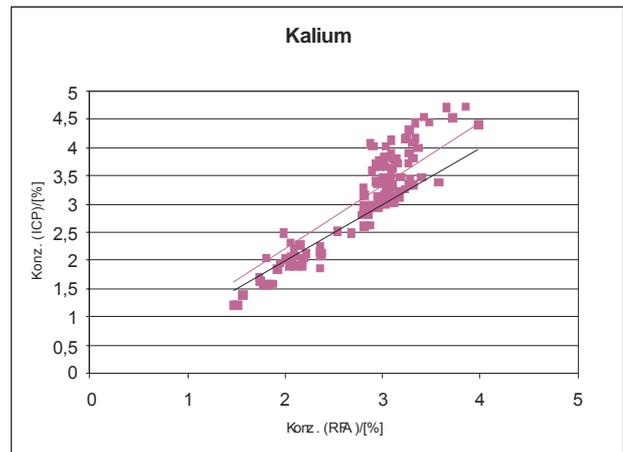
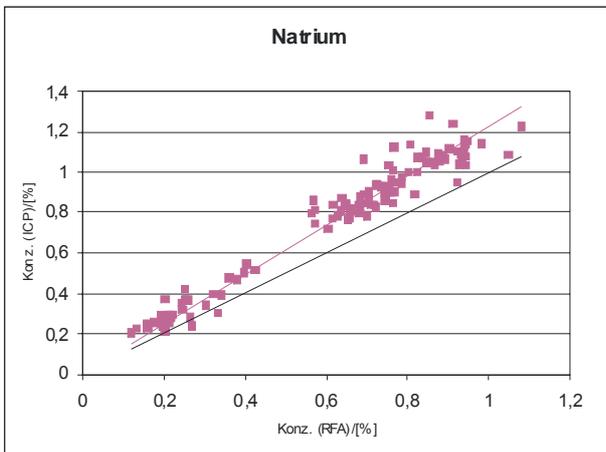
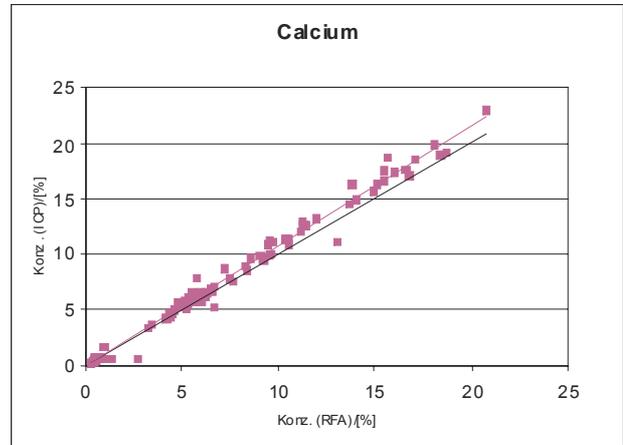
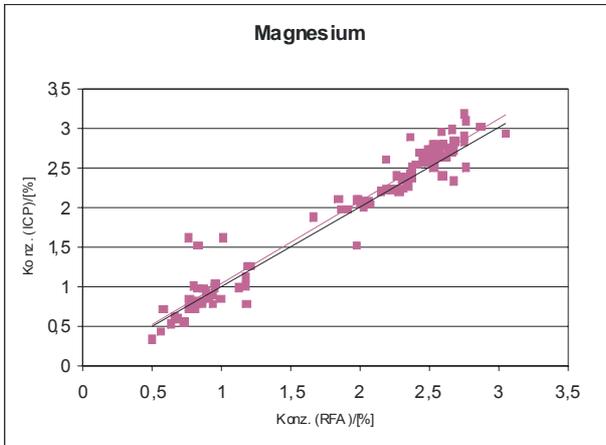
SAQ_{total}: Summe der gesamten im Datensatz enthaltenen Streuung

Je kleiner der Wert von λ desto besser ist die Trennung der Gruppen durch die Diskriminanzfunktion. In dieser Arbeit wird nach erfolgter Clusterzentrenanalyse zur besseren Trennung eine Diskriminanzanalyse durchgeführt und die errechneten Funktionen gegeneinander aufgetragen (Funktion1 ↔ Funktion2).

5.2 Vergleichbarkeit der Analysen von RFA und ICP

Nachdem alle Proben sowohl mit ICP als auch mit RFA vermessen waren, wurden die Analyseergebnisse von ICP und RFA für alle Elemente gegeneinander aufgetragen, um die Messmethoden zu vergleichen (Abb. 5.1). Dabei zeigte es sich, dass die ICP Ergebnisse systematisch von den RFA Werten abweichen. Um bei der statistischen Auswertung eine Vergleichbarkeit zu gewährleisten, wurden die ICP Ergebnisse mit einem aus den Graphen ermittelten Faktor multipliziert. Diesen Faktor erhält man, indem man die Abweichung der Ausgleichsgeraden aus der Auftragung von ICP gegen RFA Ergebnisse (violette Gerade in Abb. 5.1) von der Geraden, die man erhalten würde, wenn RFA und ICP Ergebnisse die gleichen Werte hätten (schwarze Gerade in Abb. 5.1), berechnet. Dabei wurden die ICP Werte vorher um den Glühverlust korrigiert, der bei der Herstellung der Schmelztabletten für die RFA entsteht. Dies ist auch deshalb erforderlich, um der unterschiedlichen Fähigkeit der einzelnen Ziegel, Wasser und CO₂ aufzunehmen, Rechnung zu tragen.





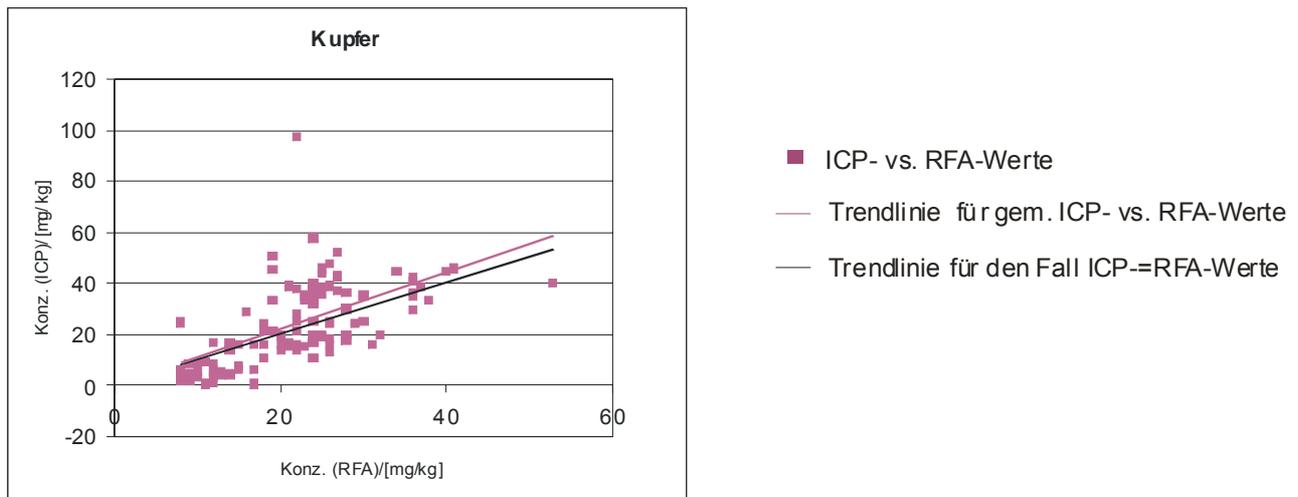


Abb. 5.1: Vergleich von ICP und RFA Ergebnissen für jedes Element zur Bestimmung der Korrekturfaktoren für die statistische Auswertung.

Um die Faktoren zu bestimmen, wurden Werte, die stark von der linearen Regression der ICP Ergebnisse abwichen, aus der Berechnung ausgeschlossen. Dann wurde eine Regressionsgerade durch die Messpunkte und den Ursprung gelegt. Die Steigung dieser Geraden beschreibt die durchschnittliche Abweichung der ICP Ergebnisse von den RFA Werten. Das Bestimmtheitsmaß der Regressionsgeraden gibt dabei das Maß für die Streuung der ICP Werte um diese Trendlinie an. Da in der Graphik die ICP gegen die RFA Ergebnisse aufgetragen wurden, wird diese Streuung allerdings nicht nur durch die Schwankungen in den ICP Messungen sondern auch durch die in den RFA Ergebnissen verursacht, die aber kleiner sind als die der ICP.

Tab. 5.1: Faktoren, mit denen die ICP Ergebnisse korrigiert wurden, um einen statistischen Vergleich mit RFA zu ermöglichen, und das Bestimmtheitsmaß der zur Korrektur verwendeten Regressionsgeraden.

Element	Korrekturfaktor	Bestimmtheitsmaß
Ti	1,0562	0,9978
Al	1,0359	0,5024
Fe	1,0991	0,7691
Mn	1,0947	0,7955
Mg	1,0601	0,9497
Ca	1,0778	0,9893

Element	Korrekturfaktor	Bestimmtheitsmaß
Na	1,3049	0,954
K	0,978	0,7987
Sr	1,1093	0,9121
Ba	0,975	0,5318
Cu	1,2016	0,4854

Nach der Korrektur wurden die Messergebnisse mit Hilfe statistischer Methoden gruppiert.

5.3 Chemische Zusammensetzung der Proben

Die Proben lassen sich schon an Hand der Elementzusammensetzung ohne statistische Berechnung in zwei Gruppen teilen. Die Ziegel aus Frankfurt-Nied zeigen einen deutlich höheren Gehalt an TiO_2 und Cr, wohingegen die Alkali- und Erdalkalikonzentrationen deutlich niedriger sind als bei den Proben aus Rheinzabern und Worms. Ebenso sind die Gehalte an Mn niedriger. In der Clusterzentrenanalyse ergaben sich vier Gruppen. In den bivariaten Plots und der Diskriminanzanalyse liegen die sieben extra gruppierten Proben (hier als unbekannt bezeichnet) immer in Worms oder Rheinzabern. Darauf wird in Abschnitt 5.6.1 näher eingegangen.

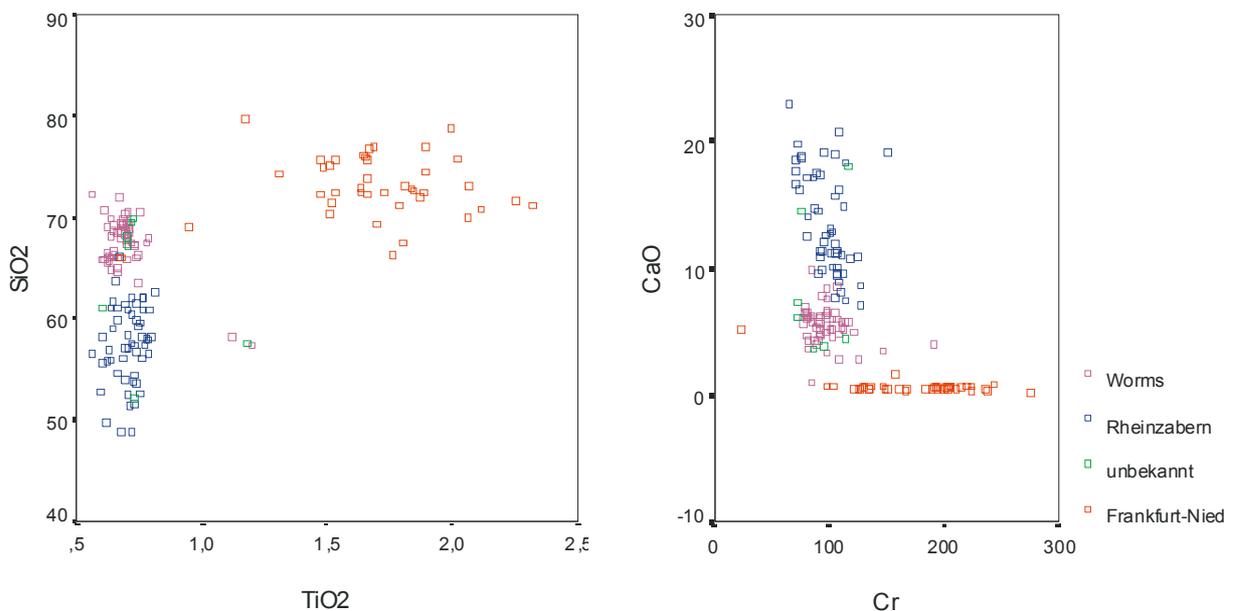


Abb. 5.2: Bivariate Plots der Zusammensetzung aller mit ICP untersuchten Ziegel. Die Zuordnung zu den Gruppen erfolgte über eine Clusterzentrenanalyse. **Links:** SiO_2 vs. TiO_2 . **Rechts:** CaO vs. Cr.

Worms und Rheinzabern, zwei kalkreichere Tone als der Ton aus Frankfurt-Nied, erscheinen hier kaum getrennt. Tatsächlich sind diese Gruppen chemisch sehr ähnlich. Sie unterscheiden sich hauptsächlich im Mangangehalt, der in Worms etwas niedriger ist, sowie in den Silizium- und Kalziumkonzentrationen, die in Worms etwas höher sind als in Rheinzabern. Auch ist die Konzentration an Alkaliionen in Rheinzabern niedriger als in Worms.

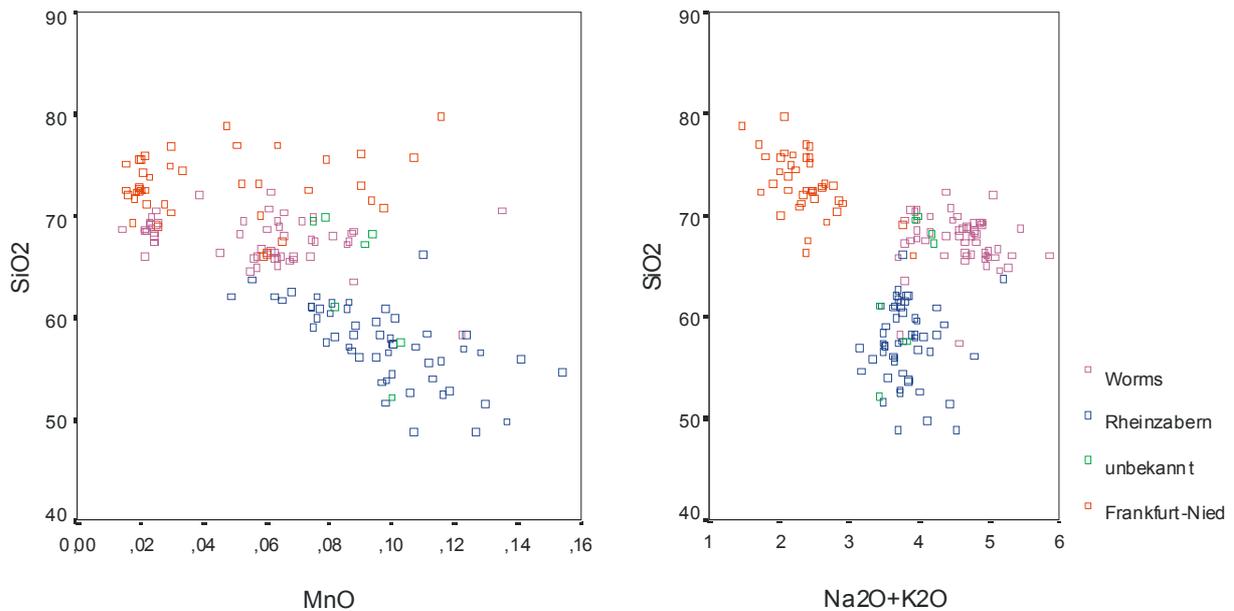


Abb. 5.3: Bivariate Plots der Zusammensetzung aller mit ICP untersuchten Ziegel. Die Zuordnung zu den Gruppen erfolgte über eine Clusterzentrenanalyse. **Links:** SiO₂ vs. MnO. **Rechts:** SiO₂ vs. K₂O+Na₂O.

Deutlicher als in zweidimensionalen bivariaten Plots lässt sich die Trennung der Gruppen anhand von bestimmten Elementen mit einer Dichteverteilungsfunktion zeigen (Abb. 5.4). Dazu wird aufgetragen, wie oft eine gewisse Konzentration dieser Elemente im Datensatz enthalten ist. In der Auftragung SiO₂ vs. TiO₂ (linke Graphik) sieht man, dass die Gruppe mit dem Herstellungsort Frankfurt-Nied sich durch ein deutliches Tal von denen mit den Herstellungsorten Worms und Rheinzabern abtrennen lässt. In der Darstellung (Na₂O+K₂O) vs. SiO₂ (rechte Graphik) trennen sich auch die Proben mit den Herstellungsorten Worms und Rheinzabern durch ein kleineres Tal von einander. Auch hier ist zu sehen, dass sich die Elementzusammensetzung der Ziegel aus den Produktionen von Worms und Rheinzabern so ähnlich ist, dass sich die Gruppen nicht anhand von bivariaten Elementplots trennen lassen. Hier ist eine multivariate Analyse zur genauen Zuordnung der Proben zu den Provenienzen erforderlich.

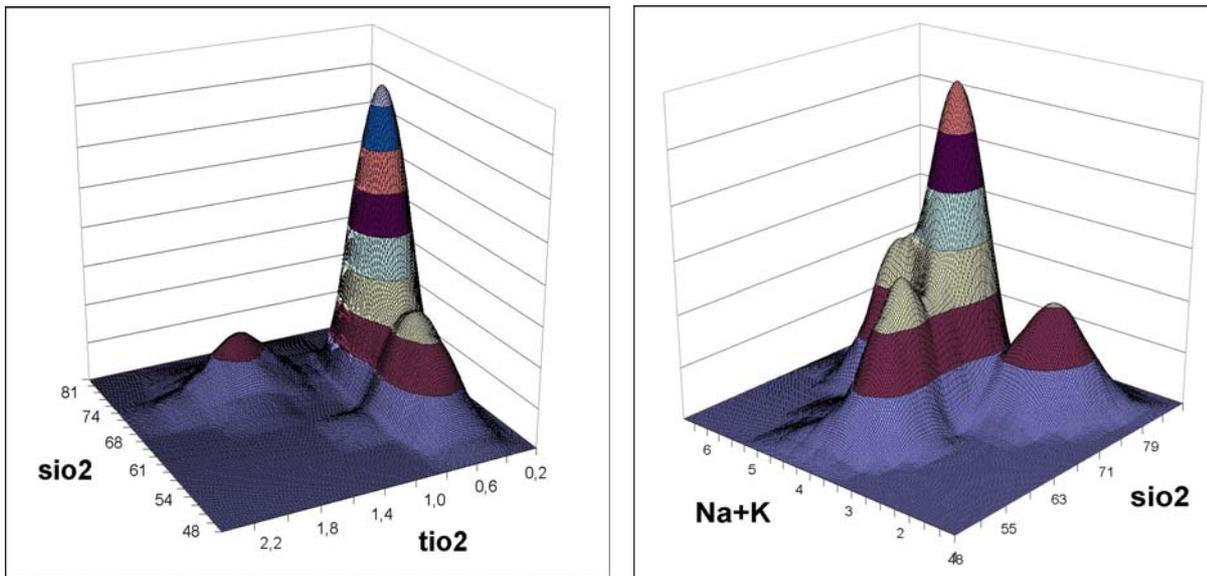


Abb. 5.4: Dichteverteilung der Ziegelproben in den einzelnen Elementgehalten; **links:** SiO_2 vs TiO_2 ; **rechts:** $\text{Na}_2\text{O}+\text{K}_2\text{O}$ vs SiO_2 .

5.4 Vergleich der Gruppierung der Messungen von ICP und RFA

Die Zusammensetzung der Proben lässt anhand von bivariaten Plots eine Einteilung in drei Gruppen vermuten, wie in vorigem Abschnitt gezeigt wurde. Im Folgenden soll überprüft werden, ob sich die Ziegel ausgehend von den RFA bzw. ICP Ergebnissen den gleichen statistischen Gruppen zuordnen lassen und ob die Verwendung weiterer Spurenelementen zusätzlich zum vorhandenen Datensatz die Gruppierung beeinflusst.

In diesem Abschnitt werden alle Proben bis auf 644-720 berücksichtigt. Auf die letztgenannten Proben wird in Abschnitt 5.5 näher eingegangen. Für die in diesem Abschnitt dargestellten Klassifizierungen wurden die Messergebnisse vor der Gruppierung nach der Two-Step Clustermethode logarithmiert. Zur besseren Darstellung wurde anschließend eine Diskriminanzanalyse durchgeführt.

Wie in Abb. 5.5 dargestellt ist, lassen sich für die RFA und die korrigierten ICP Ergebnisse die gleichen Gruppen bilden. Für beide statistischen Auswertungen wurden die Konzentrationen der gleichen Elemente und Elementoxide gewählt: SiO_2 , TiO_2 , Al_2O_3 , Fe_2O_3 , MnO , MgO , CaO , Na_2O , K_2O , Cr , Ni , Cu , Zn , Sr , Nb , Ba , La , Ce , Pb und Th . Zr und Y wurden nicht zur Klassifizierung verwendet, da ihre Wiederfindungsraten und die Reproduzierbarkeit der Ergebnisse nach dem Aufschluss zu gering waren. Zwar weichen die Werte der Diskriminanzfunktionen für RFA und ICP leicht von einander ab, dennoch werden alle Proben der gleichen Herstellungsprovenienz zuge-

ordnet. Obwohl Zr, das bisher in der statistischen Analyse der Ziegel als Gruppen bildendes Element galt, hier nicht berücksichtigt wurde, erfolgt die Klassifizierung der Proben mit den Ergebnissen aus den RFA Analysen in gleicher Weise wie unter Verwendung dieses Elements. Wie schon anhand der bivariaten Plots gezeigt wurde, lässt sich die Herstellungsprovenienz Frankfurt-Nied deutlich von Worms und Rheinzabern abtrennen, während die beiden letztgenannten sehr eng beieinander liegen. Bei der Two-Step Clusteranalyse der RFA Ergebnisse kommt es zu einer Fehlklassifizierung. Eine Probe, die nach der Diskriminanzanalyse zur Provenienz Worms gehört, wird Frankfurt-Nied zugeordnet. Auf die Leistungsfähigkeit der Klassifizierungsmethoden soll in Abschnitt 5.6 näher eingegangen werden.

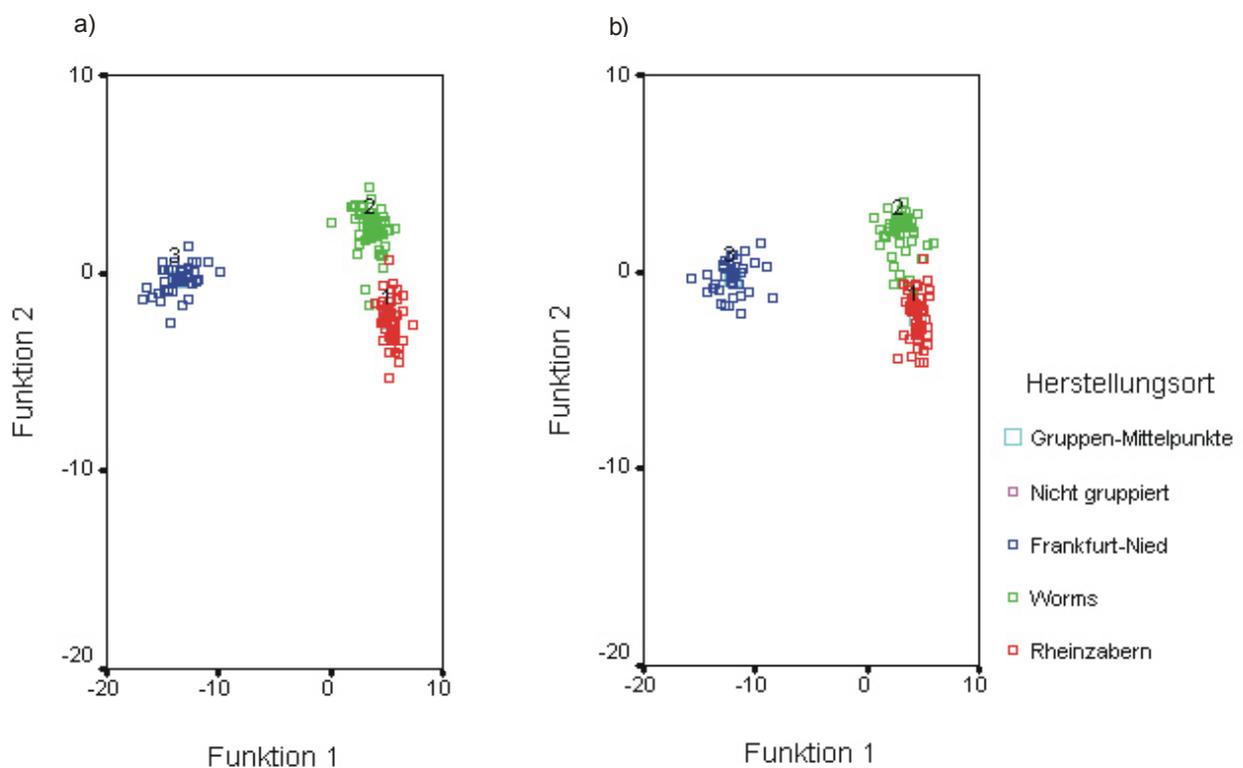


Abb. 5.5: Vergleich der Gruppierung der Ziegelproben für: a) Zusammensetzung gemessen mit RFA; b) Zusammensetzung gemessen mit ICP und anschließende Korrektur der Daten um den Glühverlust und die oben ermittelten Korrekturfaktoren (Auftragung der ersten beiden Diskriminanzfunktionen).

Der Glühverlust wird hauptsächlich von dem im Ton enthaltenen Wasser und dem als Carbonate gebundenen CO_2 verursacht. Die Fähigkeit des Ziegels, Wasser und CO_2 aufzunehmen, ist von der Porosität des Ziegels und der Brenntemperatur abhängig. Geht man davon aus, dass die Ziegelproduktion in einer Ziegelei stets nach demselben Verfahren ablief, sollten die Ziegel einer Pro-

duktion ähnliche Eigenschaften aufweisen. Deshalb sollten sich auch bei Vernachlässigung des Glühverlustes die gleichen Provenienzen ergeben. Der Glühverlust weist eine gewisse Varianz innerhalb einer Gruppe auf. Daher streuen die Ziegelproben innerhalb einer Gruppe bei einer Klassifizierung der original gemessenen ICP Ergebnisse ohne die Berücksichtigung des Glühverlustes und des Korrekturfaktors im Vergleich zu den RFA Ergebnissen (Abb. 5.6) breiter als dies in Abb. 5.5 der Fall war. Die Ziegelproben werden aber weiterhin den gleichen Provenienzen zugewiesen.

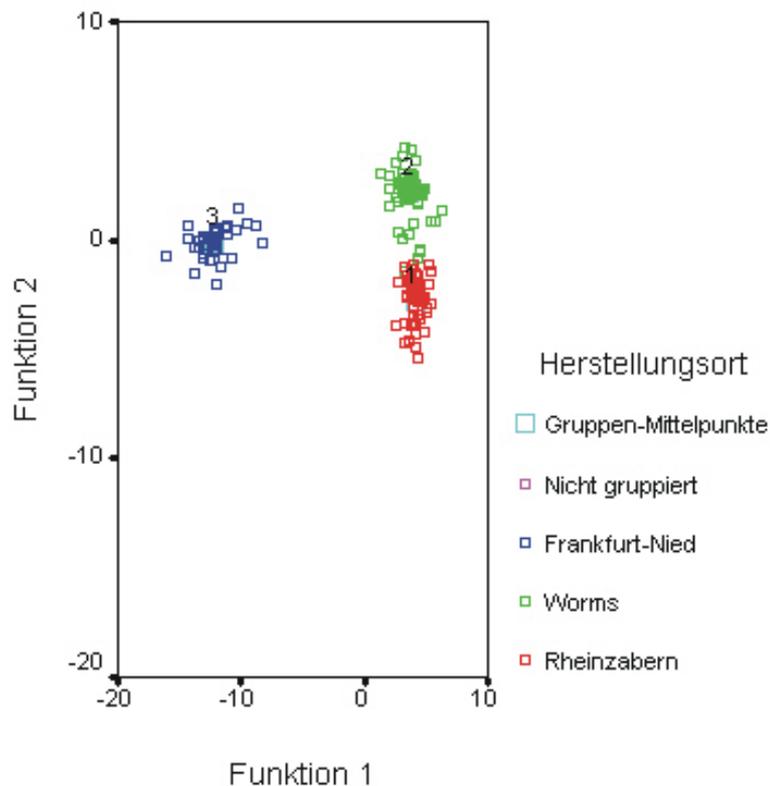


Abb. 5.6: Klassifizierung der original gemessenen ICP Ergebnisse ohne Berücksichtigung des Glühverlustes und der Korrekturfaktoren (Auftragung der ersten beiden Diskriminanzfunktionen).

Die Klassifizierungen der Ergebnisse der RFA sowie der korrigierten und der originalen Messwerte der ICP Messungen ergeben in der Statistik ein ähnliches Resultat. Fügt man nun in der statistischen Auswertung weitere Spurenelemente aus den ICP-MS Messungen hinzu, ändert sich die Gruppierung sichtbar (Abb. 5.7). Als weitere Spurenelemente wurden Ge, Pr, Nd, Sm, Gd, Dy, Er und Yb verwendet. Die übrigen noch vermessenen Spurenelemente lagen für die meisten Proben unter der Nachweisgrenze und konnten deshalb nicht für die statistische Auswertung verwendet werden. Die Werte für die Ziegel mit Herstellungsort Frankfurt-Nied streuen über einen größeren Bereich. Gleichzeitig rücken die Gruppen von Worms und Rheinzabern näher zusammen. In die-

sem Bereich kommt es auch zu einem Wechsel der Zuordnung der Ziegel H911, H922 und H938 von Worms nach Rheinzabern. Der Grund liegt in der Verteilung mancher Spurenelemente wie z.B. den seltenen Erden in der Tonlagerstätte. Sowohl Worms als auch Rheinzabern liegen im Rheintal, in dem die Lagerstätten durch Ablagerung von Sedimentgestein aus dem Rhein⁴ entstanden sind. Einige Spurenelemente sind statistisch über diese Ablagerungen verteilt und variieren deshalb nicht auf Grund der Lagerstätte. Diese Elemente „stören“ in der statistischen Auswertung, da sie zum „Verschmieren“ der Gruppen führen. Eine weitere Aufspaltung der Provenienz Frankfurt-Nied anhand der Spurenelemente ist nicht möglich.

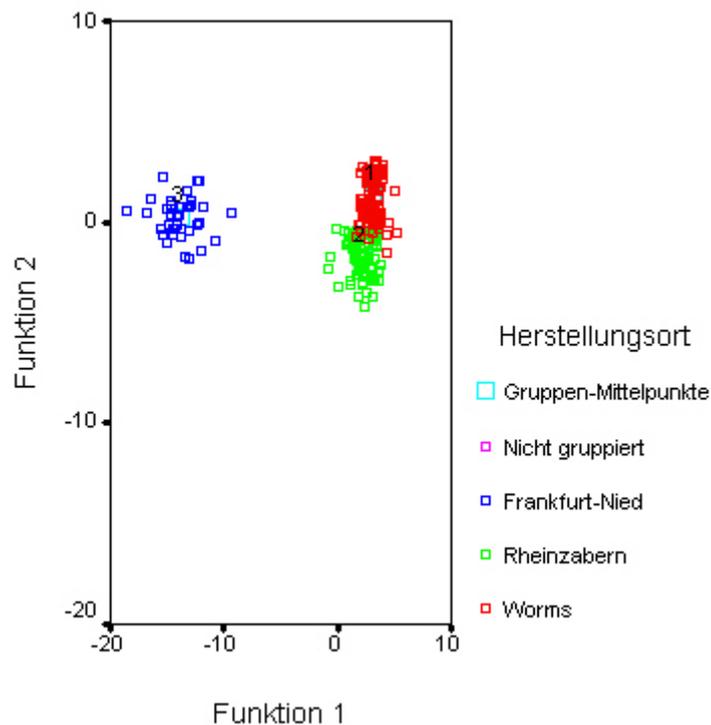


Abb. 5.7: Klassifizierung der Ziegelproben anhand der original gemessenen ICP Ergebnisse ohne Berücksichtigung des Glühverlustes und der Korrekturfaktoren unter Einbeziehung weiterer Spurenelemente (Auftragung der ersten beiden Diskriminanzfunktionen).

Ein Ziel dieser Arbeit war es zu zeigen, dass sich die Ergebnisse der Messungen mit ICP in die Gruppierung der Daten, die mit RFA erhalten wurden, einfügen lassen, da auch in Zukunft auf die umfangreiche Datenbank von G. Schneider zurückgegriffen werden soll. Deshalb war es nicht nur erforderlich zu prüfen, ob sich mit den Daten aus RFA und ICP Messungen die gleichen Gruppen erhalten lassen, sondern auch, ob man die gleiche Zuordnung erhält, wenn die ICP und RFA Ergebnisse in der gleichen statistischen Auswertung verwendet werden. Es wurden einmal die korrigierten Daten aus den Messungen mit ICP zusammen mit den Ergebnisse aus den Analysen

mit RFA in einer Klassifizierung (Abb. 5.8a) und einmal die original gemessenen Werte aus den ICP Messungen zusammen mit den RFA Daten (Abb. 5.8b) eingesetzt. Dementsprechend sind in beiden Gruppierungen alle Ziegelproben zweimal enthalten und sollten beide der gleichen Gruppe zugewiesen werden. Abweichungen ergaben sich nur bei sechs Proben. Ausnahmen sind die RFA Ergebnisse der Ziegel H867, H870, H877, die in der Klassifizierung zusammen mit den unkorrigierten ICP Ergebnissen von Worms nach Rheinzabern wechselten, und die RFA Ergebnisse der Ziegel G152, H323, H324, die in der gleichen Auswertung von Rheinzabern nach Worms verschoben wurden. D.h. bei 310 Analysen, die in die Statistik eingegangen sind, kam es zu 6 Fehlklassifizierungen und somit zu einer Abweichung von 1,9%. Mit einer Fehlerquote in dieser Größenordnung muss bei statistischen Auswerteverfahren gerechnet werden. Bei der gemeinsamen Klassifizierung der RFA Ergebnisse und der korrigierten ICP Werte erfolgte die Einordnung der Proben in die erwarteten Gruppen.

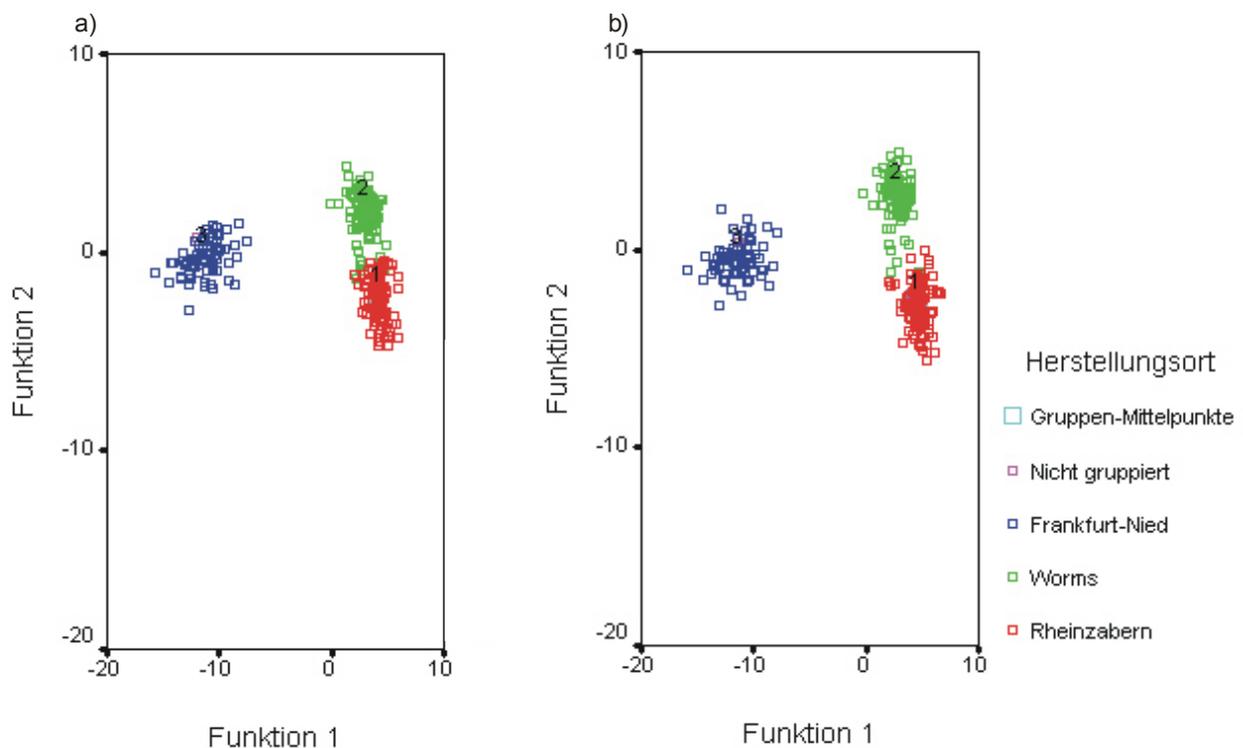


Abb. 5.8: Gemeinsame Klassifizierung von Ergebnissen aus ICP und RFA Messungen: a) RFA Ergebnisse mit korrigierten Werten aus den ICP Messungen; b) RFA Ergebnisse mit den Originalwerten aus den ICP Messungen (Auftragung der ersten beiden Diskriminanzfunktionen).

Auch hier ist zu sehen, dass die Gruppen bei Verwendung der korrigierten ICP Ergebnisse kompakter sind, als bei Verwendung der Originaldaten. Die Gründe sind hier die gleichen wie bei der alleinigen Klassifizierung der korrigierten ICP Ergebnisse. Es empfiehlt sich, in Zukunft sowohl den

Glühverlust als auch den Korrekturfaktor vor der statistischen Auswertung in die ICP Daten einzubeziehen. Bei den Korrekturfaktoren bleibt allerdings zu prüfen, ob sie nur für die Matrix der Ziegelsteine gültig sind, oder ob sie eine systematische Abweichung der Ergebnisse auf Grund der Verwendung unterschiedlicher Messmethoden darstellen.

5.5 Zuordnung der Proben aus dem Bühnentheater in Mainz zu den Ziegeleien

Wie schon in Kapitel 2 erwähnt wurde, war das Bühnentheater in Mainz das größte Theater dieser Art nördlich der Alpen. Architekturgeschichtliche und baumaterialtechnische Aspekte weisen auf den Bau in spätrömischer Zeit hin. Besonders interessant ist in diesem Zusammenhang eine Ziegelplatte mit einem Stempel der 22. Legion (Abb. 2.5), die die betreffende Bauphase des Theaters in konstantinische (308-310 n. Chr., t.p.q.) oder iulianische (355-360 n. Chr.) Zeit datiert. Nach konventioneller Meinung wird die Errichtung oder gründliche Erneuerung eines großen zivilen Repräsentationsbauwerks in Mainz zu dieser Spätzeit nicht erwartet⁵. Neben diesem gestempelten Ziegel wurden noch mehrere ungestempelte Ziegel in den Ziegeldurchschussslagen gefunden. Durch eine chemische Analyse sollte nun der Herstellungsort dieser Ziegel bestimmt und somit die späte Bauzeit des Theaters überprüft werden.

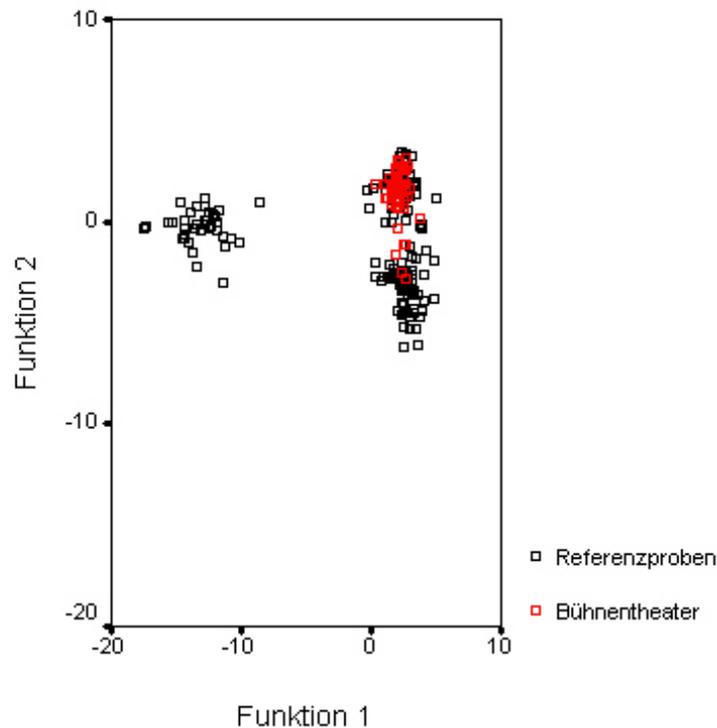


Abb. 5.9: Zuordnung der Ziegelproben aus dem Bühnentheater in Mainz zu den Herstellungsorten. Die schwarzen Quadrate beinhalten alle in den vorhergehenden Abschnitten besprochenen Ziegel als Referenzproben für die Ziegeleien, die roten sind die Proben 644, 663-720 aus dem Theater (Auftragung der ersten beiden Diskriminanzfunktionen).

Der größte Teil der Proben, unter ihnen der Ziegel mit dem Stempel der 22. Legion aus der Flörsheimer Gruppe (Boppard TypI), lässt sich eindeutig der Ziegelei in Worms zuordnen (Abb. 5.9). Dieses Ergebnis erhärtet die archäologische Annahme, dass die Bauzeit des Theaters in der Mitte des 4. Jhdts. nach Chr. anzusetzen ist. Auf Grund der chemischen Ähnlichkeit der Ziegel, die der Produktion in Worms zugeordnet werden können, mit dem gestempelten Ziegel, ist davon auszugehen, dass diese Ziegel zur gleichen Zeit hergestellt wurden. Es handelt sich also bei dem gestempelten Ziegel nicht um ein Einzelstück, das bei einer Ausbesserungsarbeit in das Theater gelangte, sondern steht im Zusammenhang mit einer größeren Baumaßnahme. Die Annahme einer Bauzeit im 4. Jhd. n. Chr. wird auch durch die Datierung des Caementitiums in diesem Teil des Theaters durch C. Goedicke bekräftigt, die für die Verarbeitung des Cementitiums eine Zeit von 424–85 n. Chr. ergab⁵.

Je nach verwendetem Statistikverfahren werden die Proben 703, 705, 706 und 716 Worms oder Rheinzabern zugewiesen. Hier sollte geprüft werden, ob noch eine weitere Ziegelei mit einem ähn-

lichen Elementmuster in dieser Zeit im Rheinbecken existierte. Denkbar wäre auch, dass während der Arbeiten eine neue Tonschicht innerhalb der Tonlagerstätte in Worms angeschnitten wurde. Die drei Proben 644, 704 und 720 werden statistisch eindeutig Rheinzabern zugeordnet. Es bleibt zu prüfen, ob es sich hier um Einzelstücke handelt, die im Zuge einer Reparaturmaßnahme in das Theater gelangten, oder ob ein Teil der verwendeten Ziegel in Rheinzabern produziert wurde.

5.6 Vergleich von verschiedenen Klassifizierungsmethoden unter Verwendung des Statistikprogramms SPSS

In diesem Abschnitt sollen Unterschiede in den Ergebnissen der Klassifizierung mittels verschiedener in SPSS angebotener Verfahren aufgezeigt und ihre Gründe diskutiert werden. Außerdem wird auf das Problem der Standardisierung von Messwerten mit unterschiedlichen Einheiten wie $\frac{\text{mg}}{\text{kg}}$ (ppm) und % eingegangen werden.

5.6.1 Verwendung verschiedener Statistikverfahren

Zu Beginn des Kapitels wurden die verschiedenen statistischen Methoden vorgestellt, die üblicherweise in den Naturwissenschaften zur Anwendung kommen. Es sollen hier die im Statistikprogramm SPSS angebotenen Klassifizierungsverfahren nach Ward, Two-Step Clustering und die Clusterzentrenanalyse miteinander verglichen werden. Als Datensatz wurden die Originaldaten der ICP Analysen aller Proben mit Ausnahme der Proben des Bühnentheaters 644-720 verwendet. Zur Berechnung wurden SiO_2 , TiO_2 , Al_2O_3 , Fe_2O_3 , MnO , MgO , CaO , Na_2O , K_2O , Cr , Ni , Cu , Zn , Sr , Nb , Ba , La , Ce , Pb und Th gewählt. Zur besseren Darstellung wurde anschließend ebenfalls eine Diskriminanzanalyse durchgeführt. Die Elementkonzentrationen wurden vor der Klassifizierung logarithmiert.

In Abb. 5.10 ist die Klassifizierung aufgetragen, die man mit dem hierarchischen Verfahren nach Ward (a) und mit der Clusterzentrenanalyse (b) erhält. Das Ergebnis der Two-Step Clusteranalyse wurde bereits in Abb. 5.7 gezeigt.

Anhand eines Ward-Dendrogramms wurde ermittelt, dass sich aus den gemessenen Daten drei Gruppen ergeben. Dies hat sich auch bei den anschließenden Clusterverfahren als die stabilste Lösung erwiesen. Verwendet man die Two-Step Clusteranalyse, ohne eine feste Anzahl an Klassen vorzugeben, erhält man vier Klassen, die allerdings bei jeder neuen Klassifizierung eine abweichende Zuordnung der Proben zu den Gruppen liefert. Erst bei einer Vorgabe von drei Klassen werden die Ziegel bei jeder neuen Berechnung wieder in die gleiche Klasse eingeordnet. Auch bei

der Clusterzentrenanalyse liefert nur die Vorgabe von drei Gruppen eine reproduzierbare Klassifizierung. Allerdings lässt sich die Klassifizierung nicht in einem Schritt erhalten. Um eine klare Trennung von Worms und Rheinzabern zu erreichen, muss der Datensatz zuerst in zwei Klassen getrennt werden. Da die Proben mit Herstellungsort Frankfurt-Nied so weit von den Ziegeln aus den Produktionen in Worms und Rheinzabern abweichen, lassen sich letztere Proben erst nach Entfernen der Daten aus Frankfurt-Nied aus dem Datensatz in zwei Gruppen trennen. Versucht man die Klassifizierung in drei Klassen in einem Schritt, so wird die dritte Klasse nur von 6 Proben gebildet, die bei einer Klassifizierung in zwei Schritten als Ausreißer erkannt werden. Diese Proben bildeten in den bivariaten Plots die vierte Gruppe, wurden aber bei der Diskriminanzanalyse den Herstellungsorten Worms und Rheinzabern zugewiesen. Im Gegensatz dazu kann die Klassifizierung nach Ward in einem Schritt ausgeführt werden. Bei der Two-Step Clusteranalyse sind die zwei Schritte, die Findung der Klassenanzahl und die anschließende Zuordnung der Proben, schon in der Routine enthalten. Beachten muss man auch, dass die einzelnen Methoden von unterschiedlichen Proben als Startwert ausgehen und deshalb verschiedene Gruppennumerierungen ergeben. Dies hat aber keinen Einfluss auf die Zuordnung der Proben zu den einzelnen Gruppen.

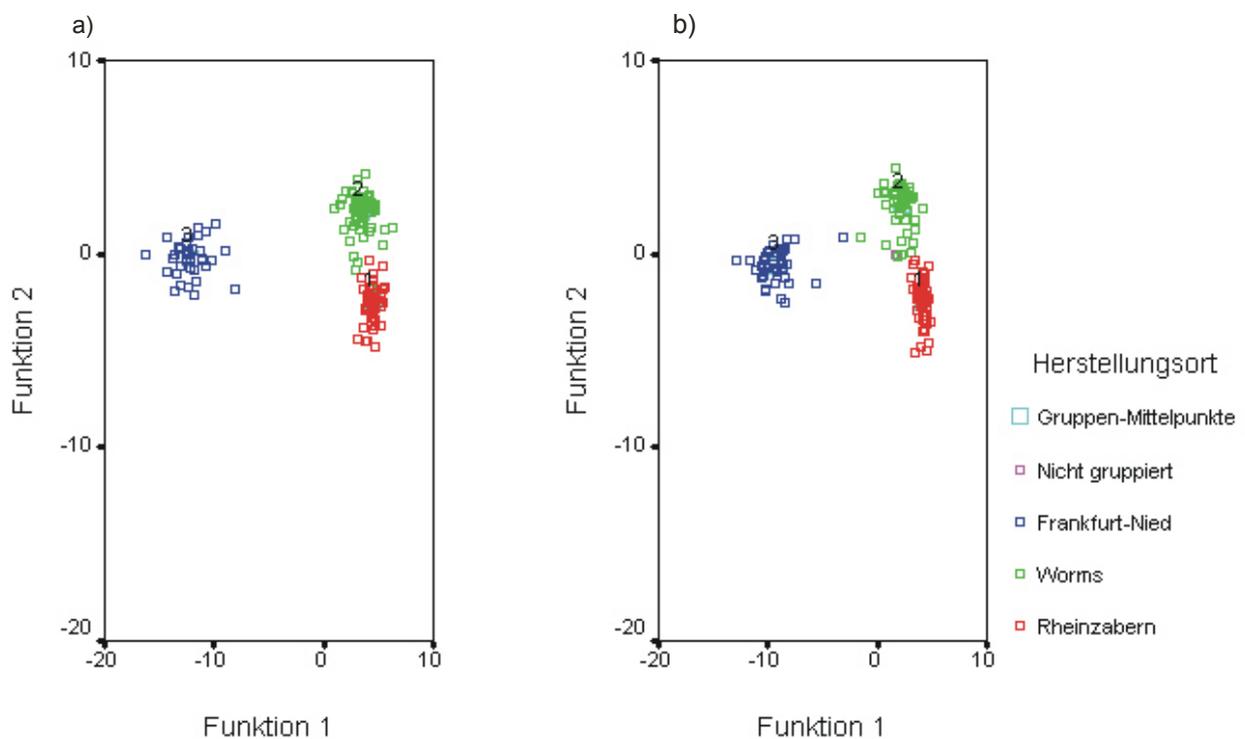


Abb. 5.10: Klassifizierung der Originaldaten aus den ICP Messungen mit zwei verschiedenen statistischen Verfahren. a) hierarchische Clusteranalyse nach Ward; b) Clusterzentrenanalyse (Auftragung der ersten beiden Diskriminanzfunktionen).

Am deutlichsten erfolgt die Trennung der Gruppen mit der Two-Step Clusteranalyse. Sie zeigt ein ähnliches Bild wie man es mit der Klassifizierung nach Ward erhält, da der erste Schritt der Two-Step Clusterbildung eine Analyse der Daten nach Ward darstellt. Durch die anschließende Überprüfung der Zuordnung in jedem Schritt der Klassifizierung erfolgt aber eine bessere Trennung auch sehr ähnlicher Gruppen wie die der Provenienzen Worms und Rheinzabern. Auch geht dieses statistische Verfahren nicht wie die Clusterzentrenanalyse von gleich großen Gruppen aus. Deshalb kommt es beim Two-Step Clustering auch nicht zu einer räumlichen Verzerrung der Gruppen mit größerer Probenanzahl (hier der Provenienzen Worms und Rheinzabern).

Bei den Proben aus dem Bereich F105-135, G006-175 und H178-H833 handelt es sich um Proben aus einem Satz von mehr als 600 Proben, die von G. Schneider mit RFA analysiert und klassifiziert wurden. Diese Proben wurden so ausgewählt, dass sie repräsentative Proben für die Ziegeleien in Frankfurt-Nied, Worms und Rheinzabern darstellen, d.h. sie stammen nicht aus den Randbereichen der Cluster und können damit eindeutig einem der Cluster zugewiesen werden. Bei allen drei verwendeten statistischen Verfahren ist dies auch der Fall. Bei den Proben H908-948 war die Provenienz nicht bekannt. Auch diese Proben wurden sowohl mit RFA als auch mit ICP untersucht. Das Ergebnis der Klassifizierung ist in Abschnitt 5.4 sowie in Anhang 3 aufgeführt. Die Klassifizierung dieser Ziegelproben ist nicht in jedem Fall stabil. So werden die Probe H909, H922, H923 und H942 in der hierarchischen Clusteranalyse nach Ward und dem Two-Step Clustering der Provenienz Worms zugeordnet, bei der Clusterzentrenanalyse aber Rheinzabern. Da die Probe H923 in Worms gefunden wurde, liegt hier Worms als Herstellungsort allerdings nahe. Die Ziegel H913 und H918 stammen laut Ward-Verfahren und Clusterzentrenanalyse aus Rheinzabern, dem Two-Step Verfahren nach aber aus Worms. Bei H914 ist es genau umgekehrt: Ward-Verfahren und Clusterzentrenanalyse weisen dieser Probe eine Herstellung in Worms zu, das Two-Step Verfahren aber eine Produktion in Rheinzabern. Denkbar wäre auch, dass diese Proben aus einer weiteren Ziegelei im Rheinbecken stammen, die hier auf Grund zu geringer Probenanzahl nicht als eigene Gruppe erkannt wird. Bemerkenswert ist, dass die Zuweisung dieser Proben in den jeweiligen statistischen Verfahren stabil ist. So ändert sich die Zuordnung auch nicht, wenn die Reihenfolge der Proben verändert wird, oder wenn nur ein statistisch zufällig ausgewählter Teil der Proben zur Klassifizierung verwendet wird (Anhang 3). Auch bei der hierarchischen Klassifizierung nach Ward ist, anders als erwartet, die Zuordnung der Proben in keinem Fall von der Reihenfolge der Proben abhängig. Allerdings ändert sich die Zuordnung der oben genannten Proben, wenn weitere Spurenelemente der Berechnung hinzugefügt werden. Man erhält allerdings auch dann keine einheitliche Zuordnung der oben genannten Proben zu den drei Herstellungsorten.

In dieser Arbeit wurde das Two-Step Clusterverfahren zur Klassifizierung benutzt, da es sich in keinem Fall als von der Reihenfolge abhängig erwiesen hat und in einem Schritt durchgeführt werden kann.

5.6.2 Verwendung verschiedener Standardisierungen der Messergebnisse

Um die Konzentrationen der Hauptelemente in % zusammen mit den Konzentrationen der Neben- und Spurenelemente in $\frac{\text{mg}}{\text{kg}}$ (ppm) in der statistischen Auswertung verwenden zu können, müssen die Messwerte auf eine vergleichbare Größe standardisiert werden. Des Weiteren gehen die statistischen Verfahren alle von einer Normalverteilung der Werte aus⁵. Bei Elementverteilungen in der Natur ist diese Voraussetzung meistens nicht erfüllt. Die Standardisierung soll also auch dazu führen, annähernd normalverteilte Werte zu erzeugen.

Zur Standardisierung stehen verschiedene Möglichkeiten zur Verfügung. Auf drei dieser Möglichkeiten soll in diesem Abschnitt näher eingegangen werden. So können die Ausgangsdaten entweder auf einen Mittelwert von 1 standardisiert oder logarithmiert werden. Durch die Transformation:

$$\text{Elementkonzentration}[X] \rightarrow \ln \left(\frac{\text{Wert}[X]}{\text{Mittelwert}[\bar{X}] + 1} \right)$$

können negative Werte, wie man sie sonst beim Logarithmieren erhält, vermieden werden, da diese bei manchen Auswerteverfahren stören können. Routinemäßig angeboten wird von SPSS nur die Standardisierung der Werte auf den Mittelwert 1. In Abb. 5.11 ist die Klassifizierung aufgetragen, die man bei einer Standardisierung auf den Mittelwert 1 (a) bzw. bei einer Standardisierung nach oben genannter Formel (b) erhält. Das Ergebnis der Two-Step Clusteranalyse nach dem Logarithmieren der Daten wurde bereits in Abb. 5.7 gezeigt.

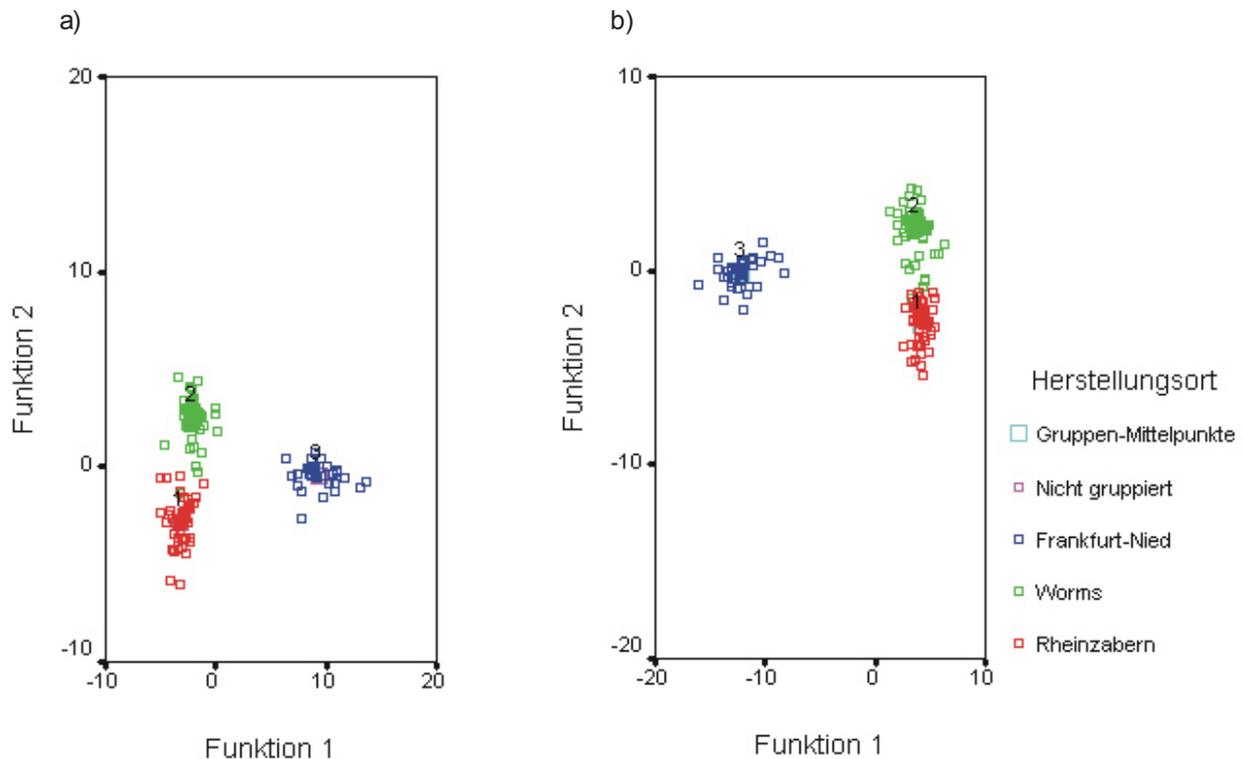


Abb. 5.11: Klassifizierung der Originaldaten aus den ICP Messungen nach zwei verschiedenen Methoden der Standardisierung. a) auf den Mittelwert 1; b) Umrechnen der Werte gemäß oben genannter Formel (Auftragung der ersten beiden Diskriminanzfunktionen).

Bei der Klassifizierung unter Verwendung der Standardisierung auf den Mittelwert 1 streuen die Proben innerhalb eines Clusters viel stärker als bei einer der Standardisierungen, in denen der Logarithmus verwendet wurde. Es ist hier zu beachten, dass sich die Skalierung in Abb. 5.11a von der in Abb. 5.11b und Abb. 5.7 unterscheidet. Die Werte spiegeln mehr von der ursprünglich in den Daten enthaltenen Streuung wieder. Dadurch liegen zwar die Gruppen-Mittelpunkte weiter auseinander, allerdings ist die Streuung innerhalb der Klassen auch deutlich größer als bei den beiden anderen Standardisierungsverfahren. Da innerhalb einer Gruppe davon auszugehen ist, dass diese Streuung von Inhomogenitäten im ursprünglichen Ton und von der Reproduzierbarkeit der Messungen stammt, erschwert diese Streuung eher die Klassifizierung. Auch enthält die Klassifizierung nach dieser Standardisierung mehrere Proben, die soweit von den anderen Proben entfernt liegen, dass sie vor einer Gruppierung erst aus dem Datensatz entfernt werden müssen, um eine sinnvolle Klassifizierung zu ermöglichen. Es handelt sich dabei um die Ziegelproben F133-135, G003 und G009, die in der Berechnung mit der Standardisierung auf den Mittelwert 1 eine eigene Gruppe bilden würden. Diese Proben werden bei allen anderen statistischen Berechnungen der Provenienz

Frankfurt-Nied zugewiesen. Diese Gruppe streut auf Grund der inhomogenen Zusammensetzung der Ziegel deutlich stärker als das bei den Ziegeln mit den Provenienzen Worms und Rheinzabern der Fall ist.

Die beiden Standardisierungsverfahren, die einen Logarithmus verwenden, liefern eine sehr ähnliche Gruppenbildung. Allerdings ist die Gruppe Rheinzabern bei der alleinigen Verwendung des natürlichen Logarithmus kompakter und näher an dem Ideal des sphärischen Clusters als bei Verwendung obiger Formel, die den Mittelwert einbezieht.

Unabhängig vom Standardisierungsverfahren werden alle Proben aus dem Bereich F105-135, G006-175 und H178-H833 den erwarteten Herstellungsorten zugewiesen. Allerdings kommt es bei den neu analysierten Proben H908-948 bei drei Proben zu unterschiedlichen Einordnungen. Es handelt sich dabei um die Ziegel H913, H918 und H923, die auch schon bei dem Vergleich der einzelnen statistischen Methoden nicht eindeutig zugewiesen werden konnten. Nach Standardisierung auf den Mittelwert 1 bzw. bei Verwendung des Logarithmus unter Einbeziehen des Mittelwertes werden diese Proben Rheinzabern zugeordnet. Werden die Daten nur logarithmiert ergibt sich als Herstellungsort Worms. Um diese Proben richtig zuordnen zu können, müsste der Datensatz noch um weitere Proben ergänzt werden, evtl. auch mit Ziegeln aus anderen Ziegeleien als den drei in dieser Arbeit betrachteten.

Da das einfache Logarithmieren als Standardisierungsmethode kompakte, gut getrennte Gruppen ohne zu große Streuungen ergab und eine einfache Methode auch für große Datensätze darstellt, wurde diese Methode in dieser Arbeit für die Datenauswertung verwendet.

Literaturverzeichnis

- [1] Janssen, J.; Laatz, W.; Statistische Datenanalyse mit SPSS für Windows; Springer-Verlag; Berlin, **2003**.
- [2] http://visor.unibe.ch/SS02/statistik3/Cluster_2.pdf.
- [3] <http://www.spss.com/pdfs/S115AD8-1202A.pdf>.
- [4] Schumann, Walter; Steine und Mineralien sammeln, S. 128ff; BLV; München, **1994**.
- [5] Dolata, J.; Goedicke, C.; Archäometrie und Denkmalpflege 2004, Jahrestagung; S. 107-109; Mannheim, **2004**.
- [6] Mucha, Hans-Joachim; Clusteranalyse mit Mikrocomputern; Akademie Verlag; Berlin, **1992**.