

5.6 Anwendung multivariater Methoden zur Ergebnisauswertung

Die in den vorangegangenen Kapiteln 5.2 bis 5.4 beschriebenen Auswertungen der Magen-, Kolon- und Rektumproben haben gezeigt, daß durch die Komplexität und Vielzahl der Variablen, d.h. den Meßgrößen (wie z.B. den einzelnen Elementgehalten), Eigenschaften oder Parametern (wie der Art der Gewebeprobe, Alter und Geschlecht der Probanden), für einige Elemente eine klare Zuordnung der hinsichtlich einer Akkumulation in einer der beiden Gewebearten (Normal- oder Tumorgewebe) nicht möglich ist. Aus diesem Grund werden neben der klassischen statistischen Auswertung in einem zweiten Schritt multivariate Methoden angewandt. Sie sollen die Identifizierung weiterer Gemeinsamkeiten und Charakteristika einzelner Elemente im Hinblick auf die Beteiligung an den Stoffwechselprozessen ermöglichen. Dabei kommen bei der vorliegenden Problemstellung die Clusteranalyse und die mehrdimensionale Hauptkomponentenanalyse als struktur-entdeckende Verfahren zum Einsatz.

Einige interessante Veröffentlichungen, die sich mit chemometrischen Fragestellungen beschäftigen und für die vorliegende Arbeit von Bedeutung sind, sind die Publikationen der Arbeitsgruppe R. HENRION und G. HENRION [120] - [122]. Erste Ergebnisse der Anwendung multivariater Verfahren zur Charakterisierung der Humangewebeproben aus der vorliegenden Arbeit wurden bereits veröffentlicht [123]¹⁵.

5.6.1 Anwendung der Clusteranalyse

Zur Ergebnisauswertung der Elementbestimmungen in den humanen Gewebeproben kommt ein hierarchisches Clusterverfahren zur Anwendung, welches den agglomerativen Algorithmus nach WARD verwendet. Das Ziel des WARD-Verfahrens besteht darin, jeweils diejenigen Objekte (Gruppen) zu vereinigen, die ein vorgegebenes Heterogenitätsmaß am wenigsten vergrößert, d.h. die die Streuung (Varianz) in einer Gruppe möglichst wenig erhöhen. Dadurch werden möglichst homogene Cluster gebildet. Als Proximitätsmaß wird dem WARD-Verfahren die quadrierte Euklidische Distanz zu-

¹⁵In [123] wurde die Auswertung der in Tab. 4.4 aufgeführten Kolongewebeproben mit Hilfe der 3-dimensionalen Hauptkomponentenanalyse beschrieben.

grunde gelegt. Als Heterogenitätsmaß wird das Varianzkriterium verwendet, welches auch als Fehlerquadratsumme bezeichnet wird [89].

Zur Clusteranalyse werden die Elementgehalte jeder Probenart – nach Ausschluß der Ausreißer – herangezogen. Mit Hilfe der Statistik-Software „STATGRAPHICS Plus“ (Version 1.4, Fa. Manugistics Inc., Rockville, USA) wird zunächst als Vorgabe die Bildung von zwei Clustern festgelegt. Dies bedeutet, daß die Gehalte der Normalgewebe *Cluster 1* und die Gehalte der Tumorgewebe *Cluster 2* zugeteilt werden. Unter Anwendung des WARD-Algorithmus erfolgt im nächsten Schritt die Berechnung der Anzahl der Objekte pro Cluster und des Mittelwertes („Centroid“) des Clusters. (Die Einzelwerte der in den Abbildungen 5.30 - 5.32 dargestellten Ergebnisse der Clusteranalyse für die Magen-, Kolon- und Rektumproben sind in Tab. A.14 im Anhang, Seite 143 aufgelistet).

Die Abbildungen 5.30 - 5.32 zeigen die oben angegebenen Clusteraufteilungen für die Magen-, Kolon- und Rektumgewebeproben. Eine „optimale“ Zuordnung der Elementgehalte in die beiden Cluster „Normalgewebe“ (Cluster 1) und „Tumorgewebe“ (Cluster 2) entspricht einer Aufteilung der analysierten Objekte für jedes Element von 50% zu 50% (d.h. daß alle Gehalte eines Elementes einer Gewebeprobe entsprechend ihrer Klassifizierung „Normalgewebe“ bzw. „Tumorgewebe“ korrekt zugeordnet werden).

Eine solche „optimale“ 50/50-Prozent-Zuordnung der Elementgehalte zu den beiden Clustern ist nur in wenigen Fällen zu beobachten. In den Magenproben werden nur die Elemente P (50/50), S (53/47) und Sr (41/59) korrekt klassifiziert. Bei allen übrigen Elementen stimmt die Zuordnung zu den beiden Clustern „Normalgewebe“ und „Tumorgewebe“ nicht. Die Kolonproben zeigen ein differenziertes Bild: S (46/54), K (39/61) und Cu (50/50) werden hier treffend eingeordnet. Für die Rektumproben sind die Elemente Se (mit einer Clusteraufteilung von 47% zu 53%) und Rb (mit einer 48/52-Zuordnung) gut, die Elemente P (63/37), S (60/40) und K (60/40) etwas schlechter als die der vorgenannten Probenarten Magen und Kolon klassifiziert.

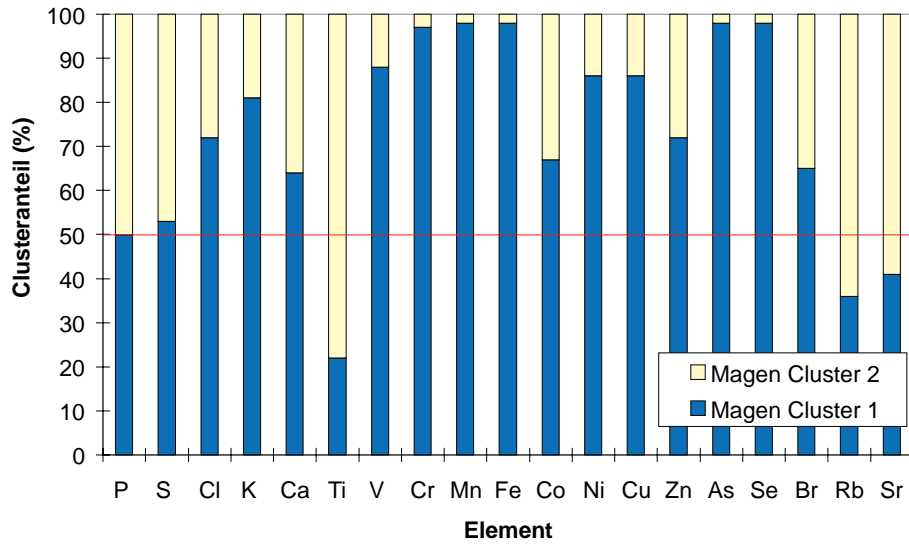


Abbildung 5.30: Clusteranalyse der Magenproben: Einordnung der Elementgehalte in zwei Cluster (WARD-Algorithmus), Cluster 1: Normalgewebe, Cluster 2: Tumorgewebe

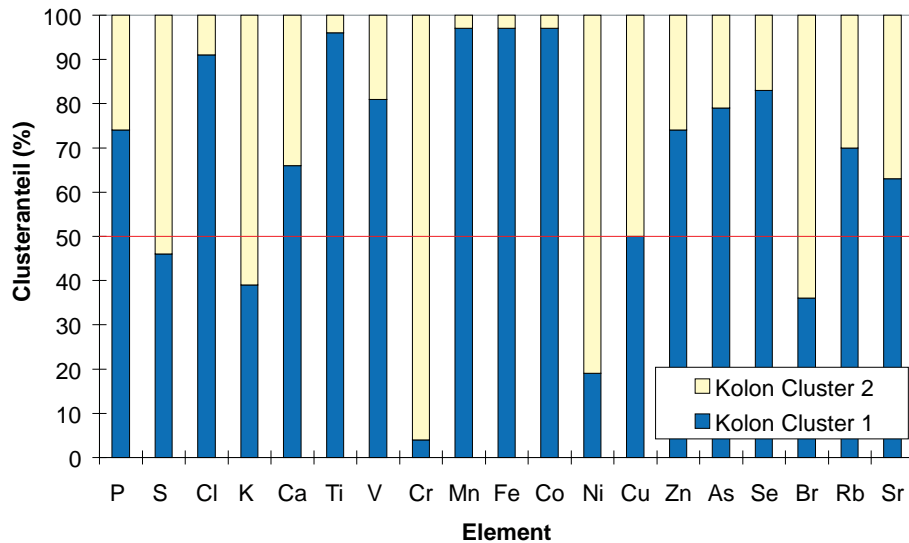


Abbildung 5.31: Clusteranalyse der Kolonproben: Einordnung der Elementgehalte in zwei Cluster (WARD-Algorithmus), Cluster 1: Normalgewebe, Cluster 2: Tumorgewebe

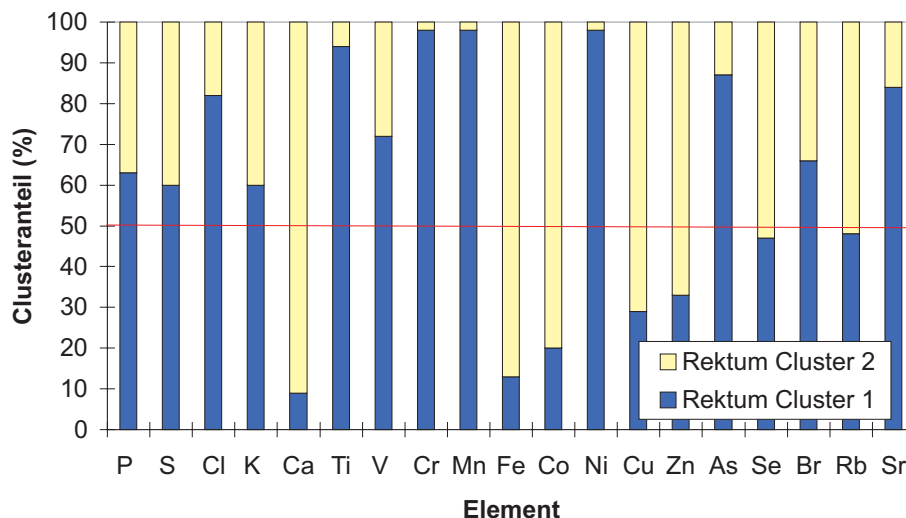


Abbildung 5.32: Clusteranalyse der Rektumproben: Einordnung der Elementgehalte in zwei Cluster (WARD-Algorithmus), Cluster 1: Normalgewebe, Cluster 2: Tumorgewebe

Unter Berücksichtigung der geringen Probenanzahl und der Inhomogenität biologischer Proben korrelieren die mit Hilfe der Clusteranalyse gewonnenen Ergebnisse recht gut mit den Resultaten der klassischen Statistik. So ist für die Elemente P und S der Magenproben, welche als signifikant bzw. tendenziell im Tumorgewebe akkumuliert eingestuft wurden (siehe Tab. 5.1, S. 86), eine gute Übereinstimmung mit den Ergebnissen der Clusteranalyse festzustellen. Für die Kolonproben zeigt sich eine gute Korrelation der Ergebnisse der Clusteranalyse mit denen der klassischen Statistik für das Element Kalium (signifikant im Tumorgewebe angereichert) bzw. für die Elemente S und Cu (Tendenz zur Akkumulation im malignen Gewebe). In den analysierten Rektumproben wurden die Elemente P, S und K als signifikant im Tumorgewebe akkumuliert klassifiziert; die drei genannten Elemente weisen nach der Clusteranalyse eine ähnliche Aufteilung ihrer beiden Cluster von etwa 60% zu 40% auf. Mit einer Aufteilung der Elementgehalte von etwa 50% zu 50% auf die beiden Cluster werden die Elemente Rb und Se (Tendenz zur Akkumulation im Tumorgewebe) sehr gut zugeordnet.

Auffallend bei allen drei Probenarten sind die prozentualen Clusteranteile der Übergangsmetalle Cr, Mn, Fe, Co und Ni. In jeder Probenart sind drei der fünf Elemente mit Werten von 3% zu 97% bzw. 2% zu 98% fast ausschließlich einem Cluster zugehörig.

Die Korrelation zu diesem Phänomen findet sich für die drei letztgenannten Elemente Fe, Co und Ni in ihrer Eigenschaft, im Normalgewebe zu akkumulieren (siehe Tab. 5.1).

Unterzieht man die als signifikant bzw. tendenziell im Tumorgewebe angereicherten Elemente P, S, K und Se als *Gruppe* einer Clusteranalyse (unter den zuvor beschriebenen Bedingungen), so erhält man für die Magenproben einen prozentuale Clusteraufteilung von 67/33. Ein besseres Ergebnis wird für die Kolonproben mit 41/59 erreicht [123]. Für die Rektumproben wird mit einem Anteil von 55/45 die beste Gruppen-Clusterung der vier Elemente erzielt.

Die Resultate der Clusteranalyse zeigen, daß die mit Hilfe der klassischen Statistik als signifikant oder tendenziell in einer der beiden Gewebearten akkumulierten Elemente, nach der Clusteranalyse gute Zuordnungen der Elementgehalte zu den beiden Clustern „Normalgewebe“ und „Tumorgewebe“ ergeben. Sie bestätigen somit die in den vorangegangenen Kapiteln gemachten Aussagen bezüglich der Signifikanz der untersuchten Elemente.

5.6.2 Anwendung der 3-dimensionalen Hauptkomponentenanalyse

Um weitere Beziehungen zwischen den 20 analysierten Elementen aufzuzeigen, werden die Elementgehalte der malignen und normalen Humangewebe mit Hilfe der 3-dimensionalen Hauptkomponentenanalyse (engl. Principal Component Analysis, PCA) untersucht, deren mathematischer Hintergrund sehr ausführlich in [124] beschrieben ist. Eine große Anzahl von Referenzstellen über verschiedene praktische Anwendungen der PCA findet man in [88]. In [120] beschrieben die Autoren die Untersuchung von Spurenelementrelationen in humanen Geweben (Haar, Leber, Niere, Lunge und Gehirn) mit Hilfe ausgewählter Methoden der Korrelationsanalyse. Dabei setzten sie auch die 3-Wege-Hauptkomponentenanalyse ein. Sie ermittelten, daß für die untersuchten inneren Organe bezüglich der 20 analysierten Elemente sehr geringe Unterschiede in den Hauptkomponenten festzustellen waren, während sich zwischen den Komponenten der inneren Organe und der Haarproben deutlichere Differenzen zeigten.

Für die Analyse der Hauptkomponenten der drei Probenarten aus dem menschlichen Verdauungstrakt werden alle Daten der Elementgehalte (nach Ausschluß der Ausreißer) in einer x-y Datenmatrix verwendet. Bei einigen Elementen (z.B: Ti, V, Cr) lagen die Gehalte nach TRF-Analyse unterhalb der Nachweisgrenze der Bestimmungsmethode. Die Behandlung dieser „fehlenden Werte“ (engl. missing values) in der Rohdatenmatrix erfolgte im Statistik-Programm „STATGRAPHICS Plus“ nach dem Prinzip des variablenweisen Ausschlusses (engl. pairwise deletion), d. h. bei Fehlen eines Wertes wurde nicht die gesamte Analyse eliminiert, sondern lediglich die betroffene Variable (d.h. das Element).

Die mit Hilfe der Statistik-Software ermittelten Hauptkomponenten wurden zur Feststellung der Anzahl signifikanter Komponenten dem Eigenwert-Eins-Kriterium nach KAISER [97] unterworfen. In allen drei Probenarten wurden fünf Hauptkomponenten mit Gesamtvarianzen von 84% (Magen), 77,2% (Kolon) und 75,9% (Rektum) gebildet. Dabei mußte das Element Vanadium aufgrund der geringen Gesamtanzahl an Meßwerten über der Nachweisgrenze von der Analyse ausgeschlossen werden, um genügend vollständige „Fälle“ (Analysen) zur Durchführung der Hauptkomponentenanalyse zur Verfügung zu haben.

Die Eigenwerte der fünf Hauptkomponenten sind für die jeweilige Probenart in den Tabellen 5.2 - 5.4 aufgelistet¹⁶. Die Komplexität der humanen Gewebeproben zeigt sich an der Anzahl extrahierter Komponenten. Zur Verdeutlichung sind – exemplarisch für die PCA der Magenproben – alle berechneten Eigenwerte bis zur Gesamtvarianz von 100% in Tabelle 5.2 dargestellt. Für die Kolon- und Rektumproben sind dann nur noch die fünf signifikanten Komponenten aufgelistet.

Tabelle 5.2: Magenproben: Eigenwerte und Varianzen nach Hauptkomponentenanalyse

Komponente	Eigenwert	Varianz (%)	Kumulative Varianz (%)
1	5,53827	32,516	32,516
2	3,67573	21,581	54,097
3	2,09155	12,280	66,377
4	1,67932	9,860	76,236
5	1,32926	7,804	84,041
6	0,869015	5,102	89,143
7	0,529408	3,108	92,251
8	0,497026	2,918	95,169
9	0,343038	2,014	97,183
10	0,178278	1,047	98,230
11	0,142729	0,838	99,068
12	0,0703629	0,413	99,481
13	0,0509493	0,299	99,780
14	0,0241975	0,142	99,922
15	0,0132238	0,078	100,000

¹⁶Die Wichtungen jedes einzelnen Elementes bezogen auf jede der fünf Hauptkomponente sind im Anhang in Tab. A.15 - A.17 aufgeführt.

Tabelle 5.3: Kolonproben: Eigenwerte und Varianzen nach Hauptkomponentenanalyse

Komponente	Eigenwert	Varianz (%)	Kumulative Varianz (%)
1	4,92326	27,284	27,284
2	3,8911	21,564	48,849
3	2,18977	12,136	60,984
4	1,68914	9,361	70,345
5	1,24099	6,877	77,223

Tabelle 5.4: Rektumproben: Eigenwerte und Varianzen nach Hauptkomponentenanalyse

Komponente	Eigenwert	Varianz (%)	Kumulative Varianz (%)
1	5,24517	28,892	28,892
2	3,4118	18,793	47,684
3	2,38661	13,146	60,830
4	1,3969	7,694	68,525
5	1,33958	7,379	75,904

Die Darstellung der Ergebnisse der Hauptkomponentenanalyse erfolgt in 2-dimensionalen Diagrammen. Abgebildet werden die Elementwichtungen der ersten vier Komponenten, wobei die Komponente mit dem höchsten Informationsgehalt (d.h. der größten Varianz) auf der x-Achse, und die übrigen Komponenten jeweils auf der y-Achse dargestellt werden. Dadurch wird eine genaue Betrachtung aller berechneten Hauptkomponenten und der zugehörigen Wichtungen der analysierten Elemente ermöglicht. Als Kriterium für eine Gruppierung von Elementen zu einem Cluster werden die ersten vier Komponenten herangezogen; die Eigenwerte der vier Komponenten haben eine kumulierte Varianz von 70 - 75%, was für eine Klassifizierung als hinreichend betrachtet werden kann [88]. Haben die Wichtungen der Elemente in den Komponenten 1 bis 4

ähnliche Werte, so werden die Elemente zu einem Cluster zusammengefaßt.

Anwendung der PCA auf die Elementgehalte der Magenproben

Abb. 5.33 zeigen die xy-Plots der Komponenten 1/2, 1/3 und 1/4, bzw. 2/3 und 2/4 (Abb. 5.34) der analysierten Magenproben. Die dargestellten Komponentenwichtungen der einzelnen Elemente können bei näherer Betrachtung in einzelne Cluster aufgeteilt werden.

Die in den Magenproben als signifikant klassifizierten Elemente P und K bilden zusammen mit dem Element S (welches als tendenziell akkumuliert eingestuft wurde) ein klar abgegrenztes, erstes Cluster. Ebenso können die Elemente Ca und Zn, die keine Akkumulationstendenz aufwiesen, zu einem Cluster gruppiert werden. Einer dritten Gruppe gehören die Elemente Cr, Fe und Mn an. Sie liegen mit ihren Wichtungen in allen vier Hauptkomponenten eng zusammen. Ein letztes Cluster, wenn auch mit schwächerer Tendenz, bilden die Elemente As und Ti; sie können nur in drei der fünf xy-Plots gruppiert werden.

Die mit Hilfe der PCA bestimmten Elementcluster für die Magenproben bestätigen die in den Kapiteln 5.2 und 5.5 für die Elemente P, K und S getroffene Klassifikation als signifikant (P, K) bzw. tendenziell (S) im Tumorgewebe akkumulierte Elemente. Eine gewisse Sonderstellung nimmt das in den zuvor genannten Kapiteln ebenfalls als signifikant klassifizierte Element Rubidium ein; in den ersten drei xy-Plots (1/2, 1/3 und 1/4) liegt es eng bei der Gruppe von P, K und S. In den beiden anderen Komponenten ist es deutlich außerhalb des P/K/S-Clusters lokalisiert. Zwischen den signifikanten Elementen gibt es folglich qualitative Unterschiede. An dieser Stelle wird der Vorteil der Ergebnisauswertung mit Hilfe multivariater Methoden sichtbar: Zum einen haben Elemente, die als „signifikant akkumuliert“ eingestuft wurden, nicht gleichzeitig ähnliche Komponentenwichtungen. Zum anderen kann durch Anwendung der 3-dimensionalen Hauptkomponentenanalyse und Vergleich der Komponentenwichtungen auch denjenigen Elementen eine „gemeinsame Eigenschaft“ zugeordnet werden, die mit Hilfe der linearen statistischen Auswertung nicht näher charakterisiert werden konnten. Dies gilt für die Elemente des Clusters Cr/Mn/Fe.

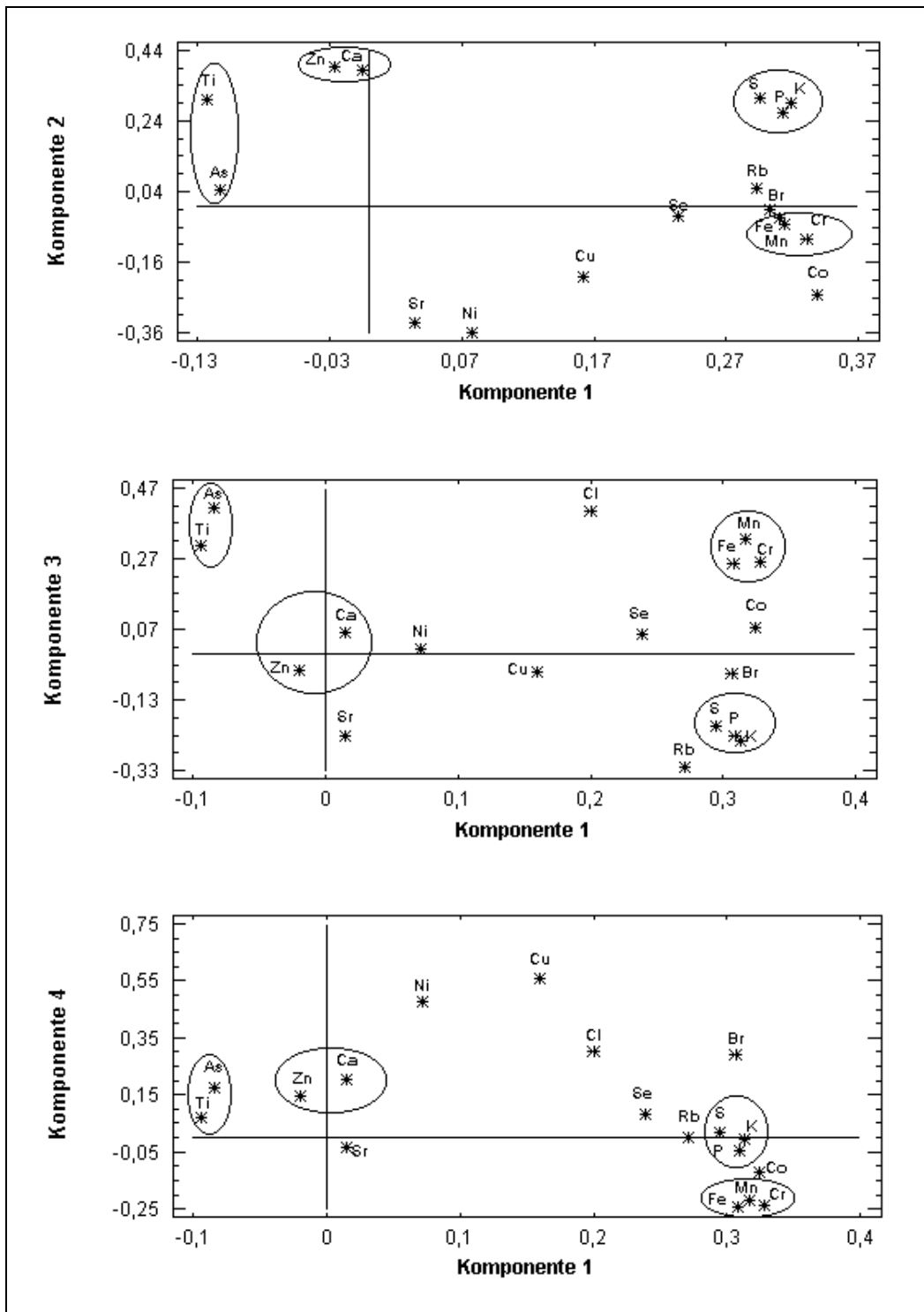


Abbildung 5.33: Magenproben: Hauptkomponentenanalyse – xy-Plot der Komponenten 1/2, 1/3 und 1/4

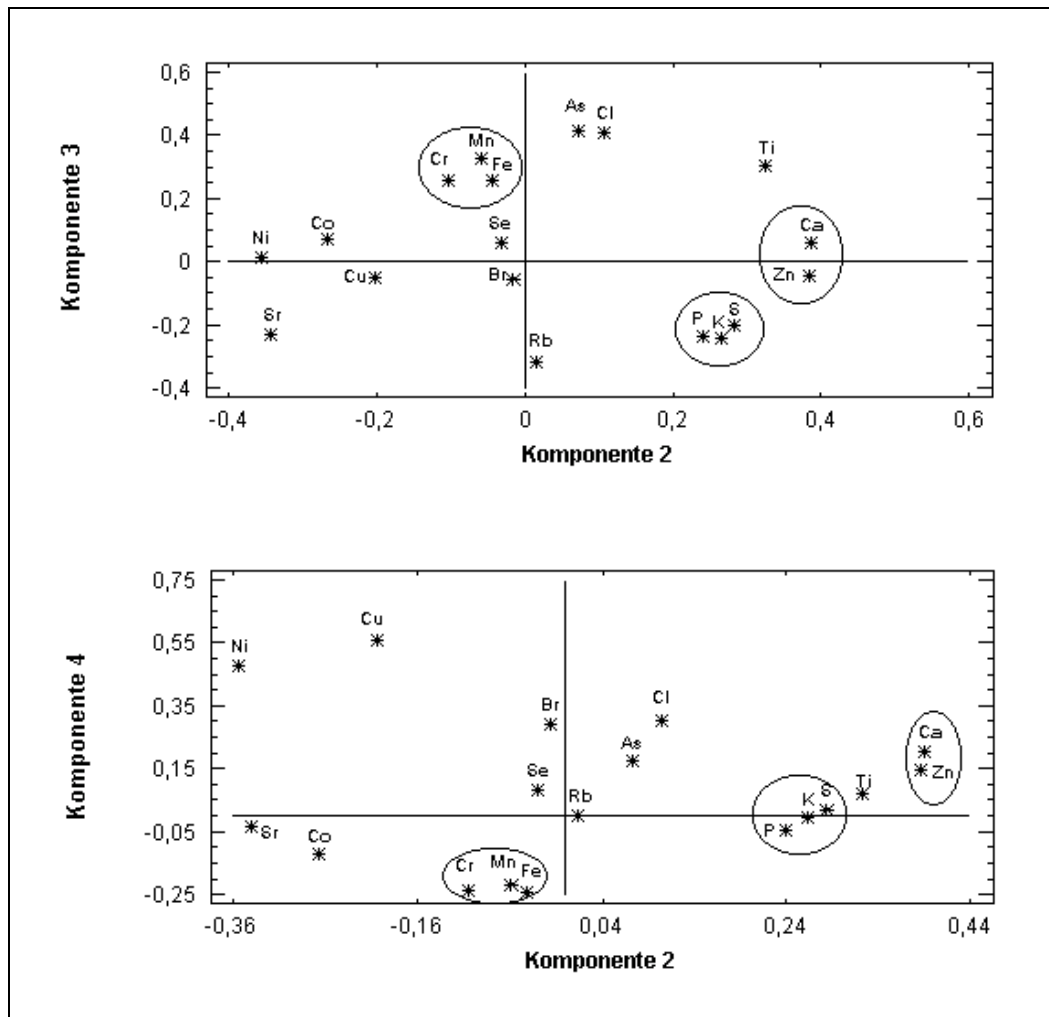


Abbildung 5.34: Magenproben: Hauptkomponentenanalyse – xy-Plot der Komponenten 2/3 und 2/4

Anwendung der PCA auf die Elementgehalte der Kolonproben

Die nach Hauptkomponentenanalyse der Elementgehalte der Kolonproben resultierenden Komponentenwichtungen sind in den Abbildungen 5.35 und 5.36 dargestellt. Im Gegensatz zu den PCA-Ergebnissen der Magenproben liegen die Wichtungen der Komponenten enger beieinander. Auch für die untersuchten Kolongewebeproben können einige Elementcluster gebildet werden.

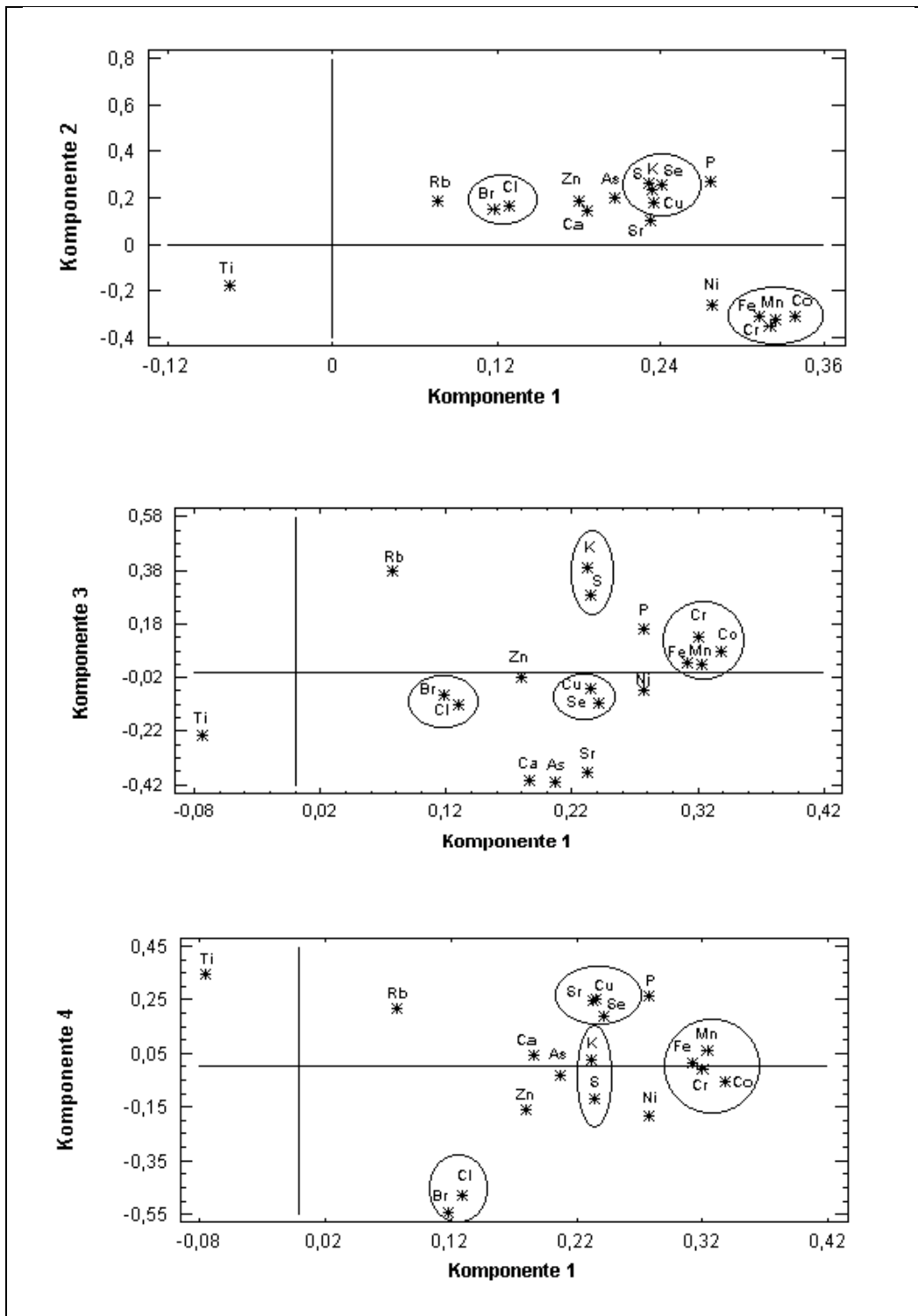


Abbildung 5.35: Kolonproben: Hauptkomponentenanalyse – xy-Plot der Komponenten 1/2, 1/3 und 1/4

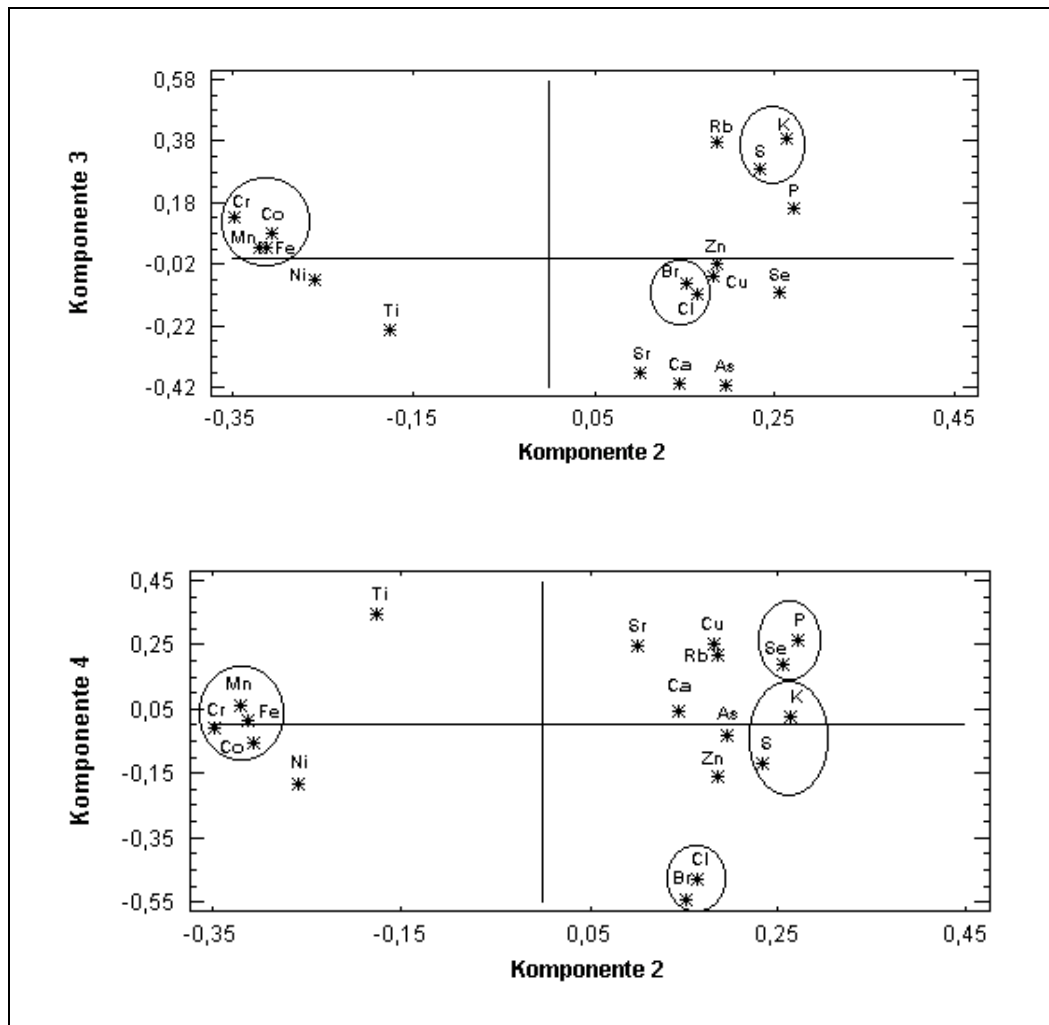


Abbildung 5.36: Kolonproben: Hauptkomponentenanalyse – xy-Plot der Komponenten 2/3 und 2/4

Die Elemente Cr, Mn, Fe und Co können in allen fünf Teilgrafiken als eng begrenztes Cluster identifiziert werden. Im Gegensatz zu den Magenproben ist bei den Kolonproben das Element Co zusätzlich in dem Cluster vertreten. Alle vier Elemente zeigten bei der linearen statistischen Auswertung relativ große Schwankungen in den Elementgehalten und wiesen keine Signifikanz zur Anreicherung in Normal- oder Tumorgewebe auf. Eine weitere Gruppe bilden die Elemente S und K, die tendenziell bzw. signifikant im Tumorgewebe angereichert sind. Das Element Phosphor, welches in den Magenproben noch dem Cluster zugehörig war, nimmt in den Kolonproben eine gewisse

Sonderstellung ein: Es ist in drei der fünf Grafiken in der Nähe des S/K-Clusters positioniert, diesem aber nicht zugehörig und in einem Fall zusammen mit dem Element Se geclustert. Das dritte Elementcluster besteht aus den Elementen Cu und Se, die bei Auftragung der Komponenten 1/2, 1/3 und 1/4 eine gemeinsame Gruppe bilden; sowohl Cu als auch Se wurden als tendenziell im Tumorgewebe akkumuliert eingestuft.

Eine besondere Stellung nimmt das Element Titan in allen fünf xy-Komponentenplots ein. Im Vergleich zu den PCA-Resultaten der Magenproben, wo Ti noch mit As ein Cluster bildete, zeigen die Komponentenwichtungen des Ti keine Gemeinsamkeit mit einem der anderen Elemente. Eine Charakterisierung und Interpretation bezüglich des im Verhältnis zu den Elementen Cr, Mn, Fe und Ni noch relativ wenig untersuchten Elementes Ti wurde bereits in Kap. 5.4.2 vorgenommen.

Anwendung der PCA auf die Elementgehalte der Rektumproben

Auch die Elementgehalte der Rektumproben wurden mit Hilfe der Hauptkomponentenanalyse analysiert. Die Komponentenwichtungen der Rektumproben zeigen eine breitere Verteilung als die zuvor beschriebenen Kolonproben. Eine Clusterung ist auch bei den Rektumproben möglich, wenn die Cluster auch nicht so scharf konturiert sind, wie bei den vorangegangenen beiden Probenarten. Die Abbildungen 5.37 und 5.38 zeigen die Komponentenwichtungen der einzelnen Elemente in Form von xy-Plots.

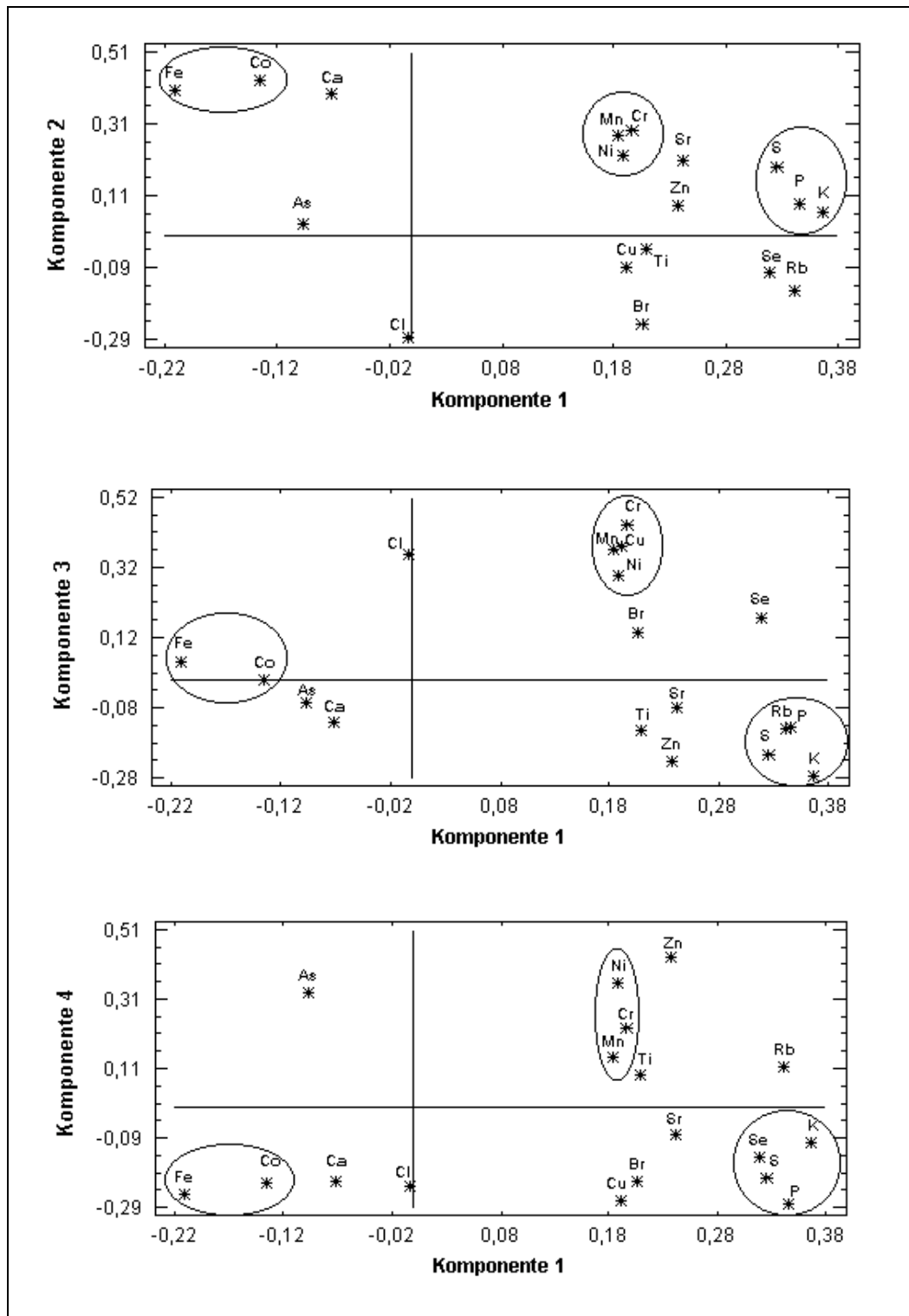


Abbildung 5.37: *Rektumproben: Hauptkomponentenanalyse – xy-Plot der Komponenten 1/2, 1/3 und 1/4*

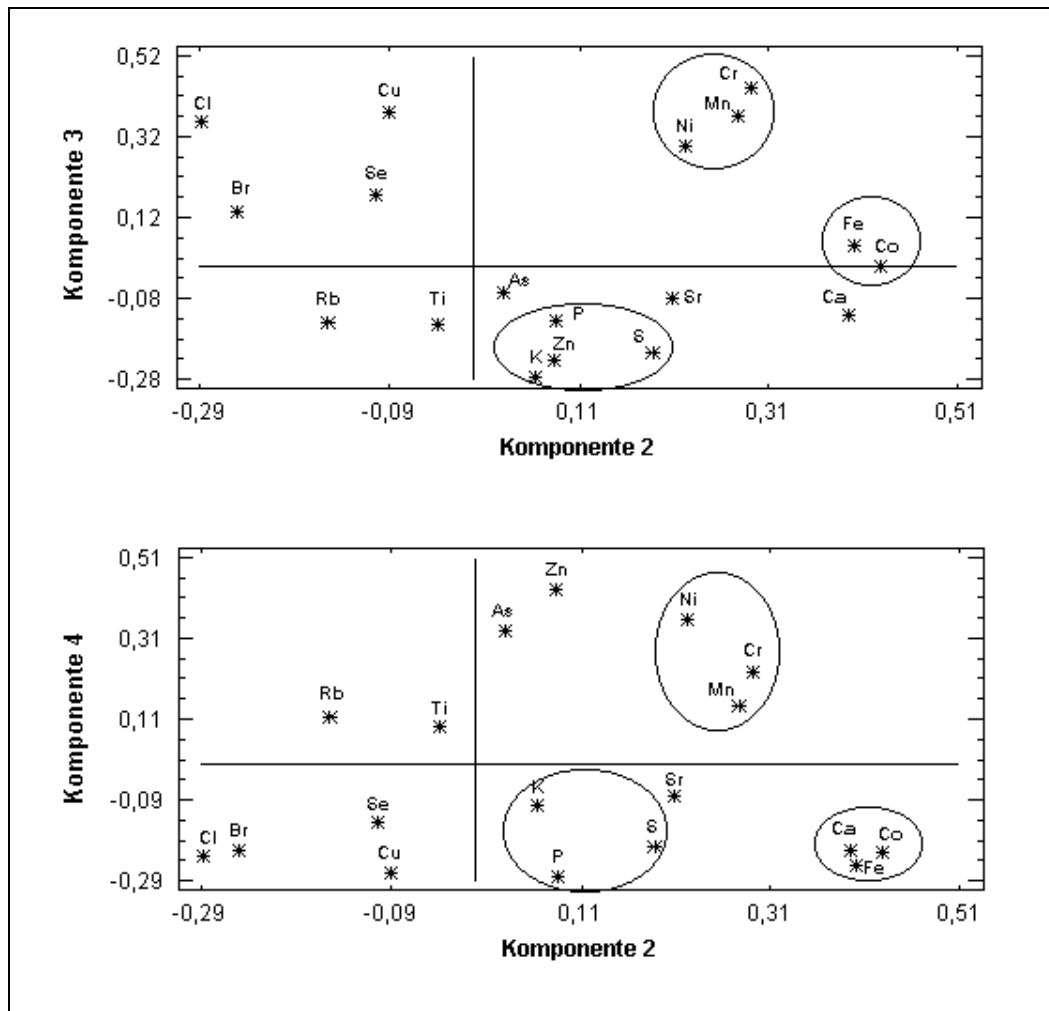


Abbildung 5.38: *Rektumproben: Hauptkomponentenanalyse – xy-Plot der Komponenten 2/3 und 2/4*

Wie schon in den beiden zuvor analysierten Probenarten können die Nebengruppenmetalle Cr und Mn, die keine Akkumulationstendenz aufwiesen, in einem Cluster zusammengefaßt werden. Hinzu kommt im Rektumgewebe noch das Element Ni, welches in Magen- und Kolongewebe eine Tendenz zur Anreicherung im Normalgewebe zeigte, in den Rektumproben aber – wie die beiden erstgenannten Elemente – keine Neigung zur Akkumulation aufweist. Eine zweite, vom Cr/Mn/Ni-Cluster abgesetzte Gruppe, bilden die Elemente Fe und Co. Sie wurden als tendenziell im Normalgewebe akkumuliert eingestuft. Das letzte Cluster, welches in allen fünf xy-Plots identifiziert werden kann, wird von den Elementen P, K und S gebildet; sie akkumulieren in den

Tumorgewebe der Rektumproben. Im Vergleich zu den beiden vorgenannten Clustern ist das P/K-Cluster allerdings in den Komponenten 2, 3 und 4 nur schwach ausgeprägt.

Tabelle 5.5 fasst die Ergebnisse der Hauptkomponentenanalyse für die untersuchten Probenarten Magen, Kolon und Rektum zusammen:

Tabelle 5.5: *Vergleich der Elementcluster der Magen-, Kolon- und Rektumproben nach Hauptkomponentenanalyse*

Probenart	Magen	Kolon	Rektum
Element-Cluster	P / K / S	K / S	P / K / S
	Cr / Mn / Fe	Cr / Mn / Fe / Co	Cr / Mn / Ni
	Ca / Zn	Cu / Se	Fe / Co
	As / Ti		

Die Ergebnisse der Hauptkomponentenanalyse bestätigen zum großen Teil die in den vorangegangenen Kapiteln identifizierten Signifikanzmuster der Magen-, Kolon- und Rektumproben. Der Vergleich der geclusterten Elementgruppen aller drei Probenarten zeigt, daß eine Möglichkeit zur Unterscheidung der humanen Gewebeproben durch Anwendung der 3-dimensionalen PCA besteht, wenn auch die Unterschiede in den Elementclustern noch gering sind. Ein Grund für diese geringen Differenzen liegt sicherlich in der Anzahl der untersuchten Proben. Bei steigender Probenzahl können einerseits die statistische Sicherheit der Ergebnisse erhöht werden und andererseits die Charakteristika der einzelnen Probenarten klarer hervortreten, was sich dann in schärferen Elementclustern und einer deutlicheren Abgrenzung der Cluster resultieren könnte. Darüber hinaus können diejenigen Elemente, die mit Hilfe der klassischen, linearen Statistik (d.h. Bildung von Mittelwerten und Standardabweichungen; Bildung von Elementverhältnissen) nicht näher charakterisiert werden konnten (Cr, Mn, Fe, Co, Ni), durch Anwendung der Hauptkomponentenanalyse näher klassifiziert werden.

Tab. 5.5 zeigt auch, daß sich für einige Elemente mögliche Wechselwirkungen erkennen lassen. Dies gilt z.B. für Cr/Mn oder K/Rb.

PFANNENHAUSER [125] verdeutlichte die möglichen Wechselwirkungen zwischen essentiellen und toxischen Elementen, die ihre Aufnahme durch die Nahrung beeinflussen, mittels der in Abb. 5.39 dargestellten Grafik. Durch die Linien wird eine mögliche gegenseitige Einflußnahme zweier Elemente sichtbar gemacht.

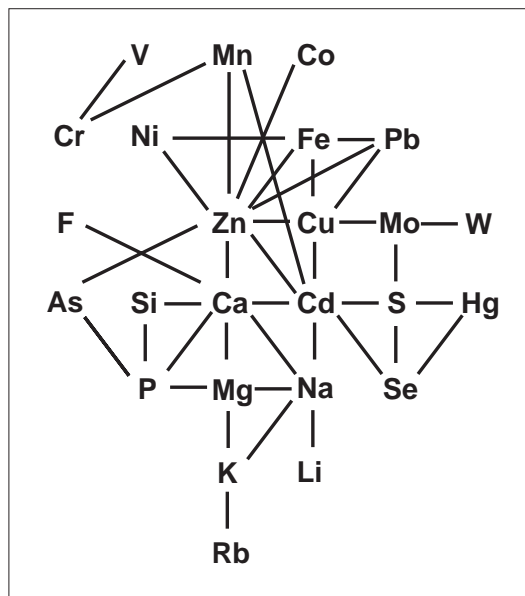


Abbildung 5.39: Mögliche Wechselwirkungen zwischen essentiellen und toxischen Elementen, die durch Nahrungsaufnahme beeinflusst werden [125]

Insgesamt können multivariate Verfahren – wie Cluster- und Hauptkomponentenanalyse – als geeignete Ergänzungen zur klassischen Statistik angesehen werden. Sie ermöglichen es, weitergehende Unterschiede in der Charakteristik der Elementverteilungen sichtbar zu machen, als dies allein mittels linearer statistischer Auswerteverfahren möglich ist.