

5.3.3.5 Anwendung der PCA auf die Barium-Gruppe

Die Spektren der Bariumverbindungen aus Tabelle 5.4 als Vertreter der Verbindungen mit hoher Ordnungszahl, deren K-Fluoreszenzlinien durch die Verpackung über das Bremskontinuum der Röntgenröhre angeregt werden, wurden mit Hilfe der PCA untersucht. Die drei Spektralbereiche (Barium-Fluoreszenzlinien, Compton- und Rayleigh-Streubereich sowie die Kombination aus Fluoreszenzlinien, Compton- und Rayleigh-Streubereich) und die durch die ersten drei HK beschriebene Varianz des entsprechenden Datensatzes sind in Abb. 5.47 dargestellt. Die durch die ersten drei HK beschriebene Varianz ist für alle drei Spektralbereiche sehr gering (siehe auch die Darstellung der Varianzen für Blei).

Die Loading-Plots für die drei Spektralbereiche werden nicht dargestellt. Die für eine Auftrennung der Substanzen in Richtung der ersten zwei HK verantwortlichen Spektralbereiche sind, wie beim Loading-Plot für den Fluoreszenzlinienbereich der Bleiverbindungen, nicht zu erkennen.

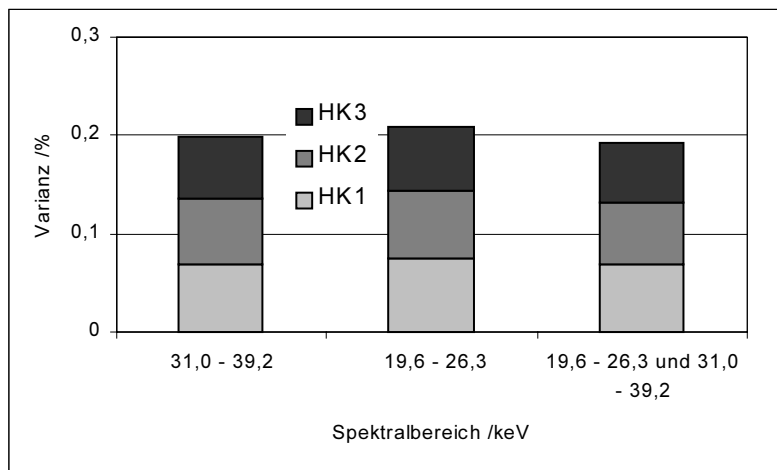


Abb. 5.47 Darstellung der durch die ersten drei HK beschriebene Varianz für die oben aufgeführten Spektralbereiche der vier Bariumverbindungen, berechnet mit der Software SIMCA-P[®], Meßzeit: 30 s

In Abb. 5.48 sind die Spektren der vier Bariumverbindungen gezeigt. Neben den Barium-K-Fluoreszenzlinien (32,19 keV und 36,38 keV) sind bei drei Verbindungen Strontium-Fluoreszenzlinien (14,16 keV und 15,83 keV) zu erkennen. Strontium kommt als Nebenbestandteil (bis 5%) in Bariumverbindungen vor. Aufgrund der Nähe der Strontiumabsorptionskante zur Ag-Primärstrahlungsenergie wird das Strontium besonders gut angeregt.

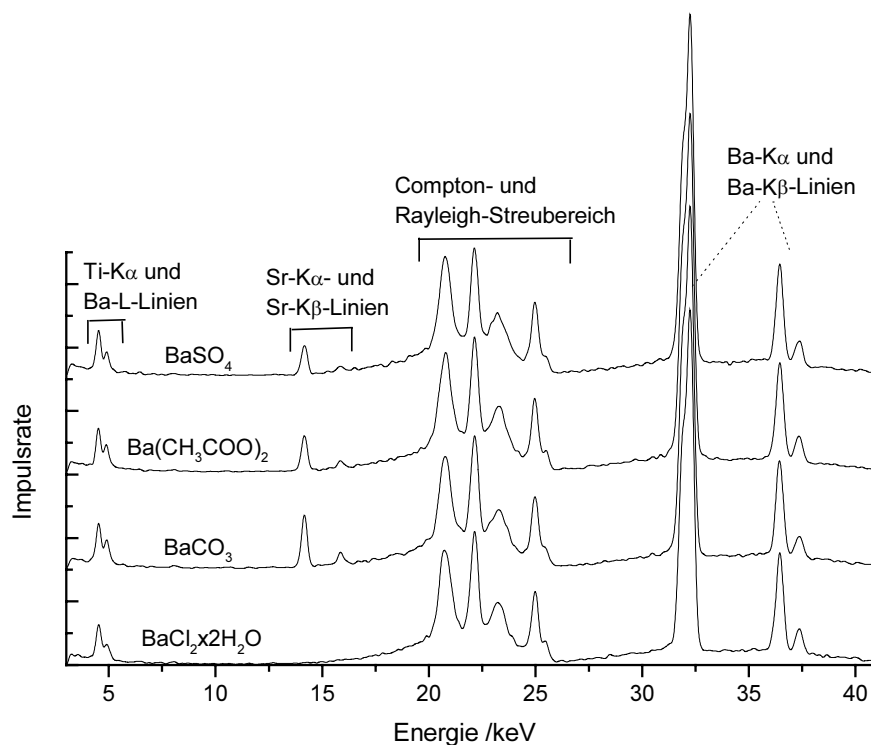


Abb. 5.48 EDRFA-Spektren der vier Bariumverbindungen aus Tabelle 5.4 (PE-Verpackung, weiß, 80 ml), Strom: 2 mA, Spannung: 45 kV, 0,12 mm Ag-Primärstrahlfilter, 3 mm Detektorblende

Die Score-Plots für die vier Bariumverbindungen für die Kombination aus K-Fluoreszenzlinien, Compton- und Rayleigh-Streubereich, für den gesamten Spektralbereich von 3 bis 41 keV und für den Strontiumfluoreszenzlinienbereich (Strontium-K α -Linie) sind in Abb. 5.49 bis Abb. 5.51 dargestellt.

Kombination aus Barium-K-Fluoreszenzlinien, Compton- und Rayleigh-Streubereich	Eine Auftrennung der vier Bariumverbindungen in einzelne Gruppen ist nicht zu erkennen (s. Abb. 5.49).
gesamter Spektralbereich	Die vier Bariumverbindungen sind in drei Gruppen in Richtung der ersten HK aufgetrennt (s. Abb. 5.50).
Strontium-Fluoreszenzlinienbereich	Die vier Bariumverbindungen sind in vier Gruppen in Richtung der HK 1 und teilweise in Richtung der HK 2 aufgetrennt (s. Abb. 5.51).

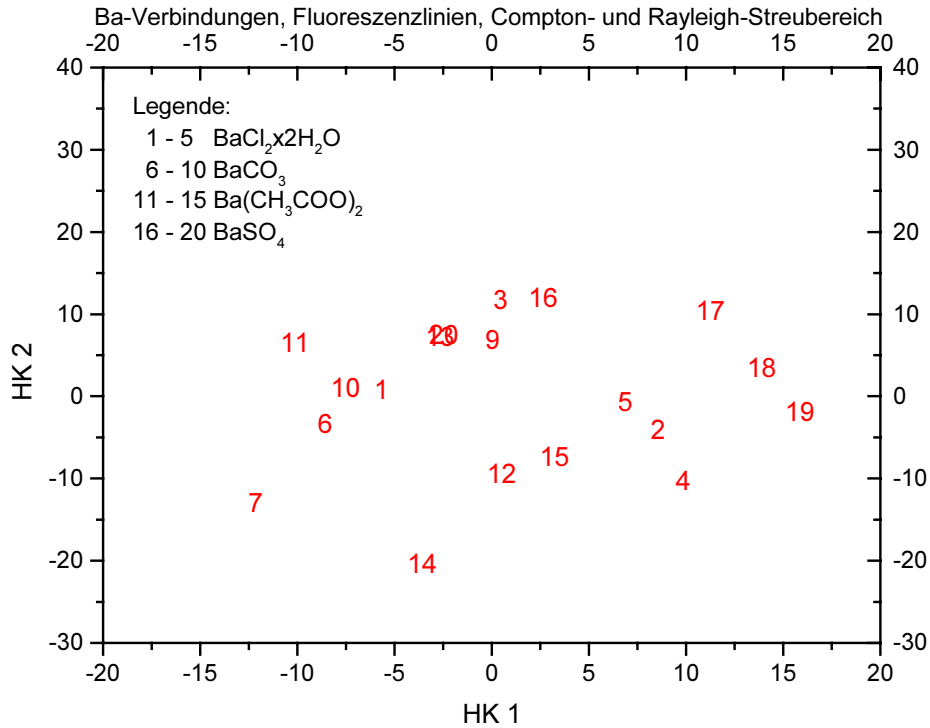


Abb. 5.49 Score-Plot der vier Bariumverbindungen aus Tabelle 5.4, berechnet mit der Software SIMCA-P[®], Meßzeit: 30 s, Spektralbereich: Barium-Fluoreszenzlinien, Compton- und Rayleigh-Streubereiche

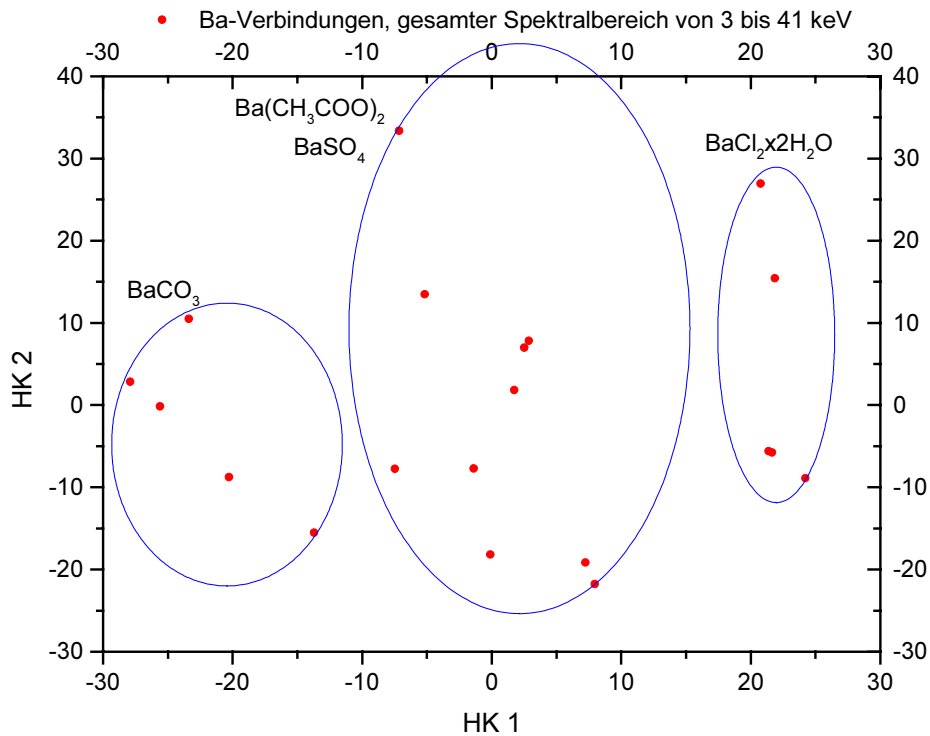


Abb. 5.50 Score-Plot der vier Bariumverbindungen aus Tabelle 5.4, berechnet mit der Software SIMCA-P[®], Meßzeit: 30 s, gesamtes EDRFA-Spektrum

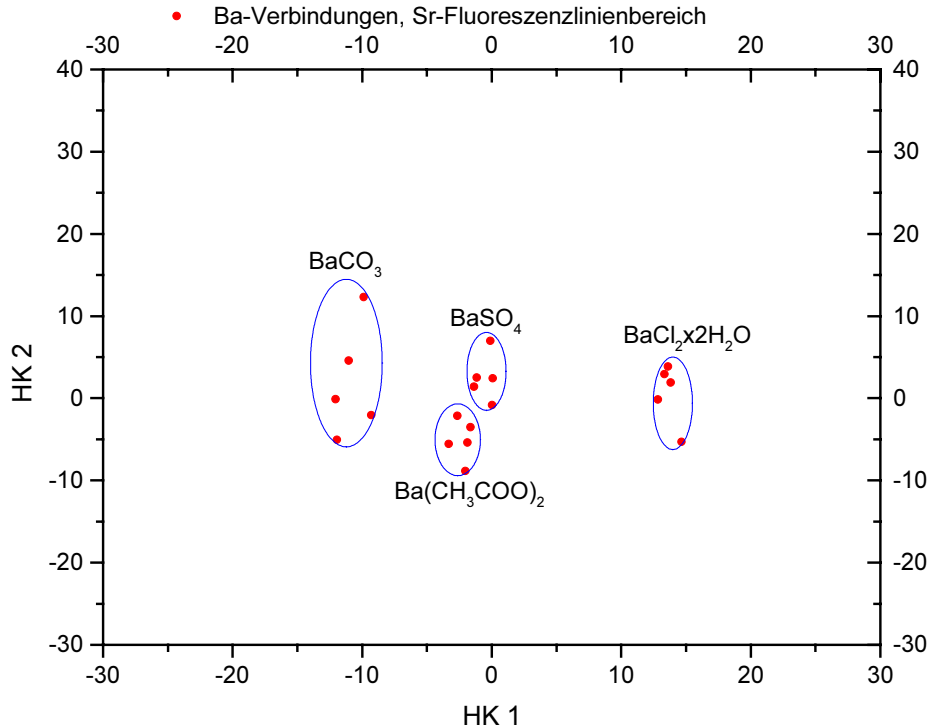


Abb. 5.51 Score-Plot der vier Bariumverbindungen aus Tabelle 5.4, berechnet mit der Software SIMCA-P[®], Meßzeit: 30 s, Spektralbereich: Strontium-K α -Linie

5.3.3.6 Einfluß der physikalischen Parameter auf die Hauptkomponenten (HK) der Substanzen aus der Ein-Element-Gruppe

Für die Chrom- und Eisengruppe ist der Zusammenhang zwischen den physikalischen Parametern der einzelnen Verbindungen (Chromverbindungen: Tabelle 5.6, Eisenverbindungen: Tabelle 5.7) am Beispiel der molaren Masse (s. Abb. 5.52 und Abb. 5.53), der physikalischen Dichte (s. Abb. 5.54 und Abb. 5.55), der \overline{OZ} (s. Abb. 5.56 und Abb. 5.57) und dem Massenschwächungskoeffizienten für die Primärröntgenstrahlungsenergie von 22,162 keV, (s. Abb. 5.58 und Abb. 5.59) auf die ersten zwei HK gezeigt.

Die molare Masse der Chrom- und Eisenverbindungen hat weder auf HK 1 noch auf HK 2 erkennbare Einflüsse (Abb. 5.52 und Abb. 5.53). Einen Einfluß auf HK 1 und HK 2 durch die physikalische Dichte der Chrom- und Eisenverbindungen ist ebenfalls nicht erkennbar (Abb. 5.54 und Abb. 5.55). Ein Einfluß der mittleren Ordnungszahl der Chrom- und Eisenverbindungen ist nur auf HK 1 erkennbar. Für die Chromverbindungen ist bis auf die Substanz (NH₄)₂CrO₄ ein Zusammenhang zu erkennen (Abb. 5.56). Der HK- \overline{OZ} -Verlauf der sieben Eisenverbindungen ist für HK 1 nicht linear (Abb. 5.57). Der Zusammenhang zwischen HK 1 und

dem Massenschwächungskoeffizienten ist im Vergleich zur \overline{OZ} deutlicher (s. Abb. 5.58 und Abb. 5.59). Die Verwendung der HK für Regressionsberechnungen ist durchaus möglich^[184,186]. Eine lineare Anpassung der Zusammenhänge $HK\ 1 = f(\overline{OZ})$ und $HK\ 1 = f(\mu)$ für die fünf Verbindungen der Chromgruppe und eine logarithmische Anpassung für die sieben Eisenverbindungen ergibt für $HK\ 1 = f(\mu)$ ein höheres R^2 (s. Tabelle 5.8). Die in Abschnitt 5.3.1 erwähnte Elektronendichte wird durch den Massenschwächungskoeffizienten^[138,139,140,141] offensichtlich exakter beschrieben als durch die \overline{OZ} .

Tabelle 5.8 Lineare Anpassung (Fit) und deren R^2 für die beiden Funktionen $HK\ 1 = f(\overline{OZ})$ und $HK\ 1 = f(\mu)$ für die fünf Verbindungen der Chromgruppe

Gruppe	Funktion	Fit	R^2
Chrom	$HK\ 1 = f(\overline{OZ})$	$HK\ 1 = -3,11\overline{OZ} + 45,6$	0,733
Chrom	$HK\ 1 = f(\mu)$	$HK\ 1 = -3,89\mu + 25$	0,981
Eisen	$HK\ 1 = f(\overline{OZ})$	$HK\ 1 = 38,5 \ln \overline{OZ} - 106$	0,947
Eisen	$HK\ 1 = f(\mu)$	$HK\ 1 = 19,0 \ln \mu - 36,6$	0,981

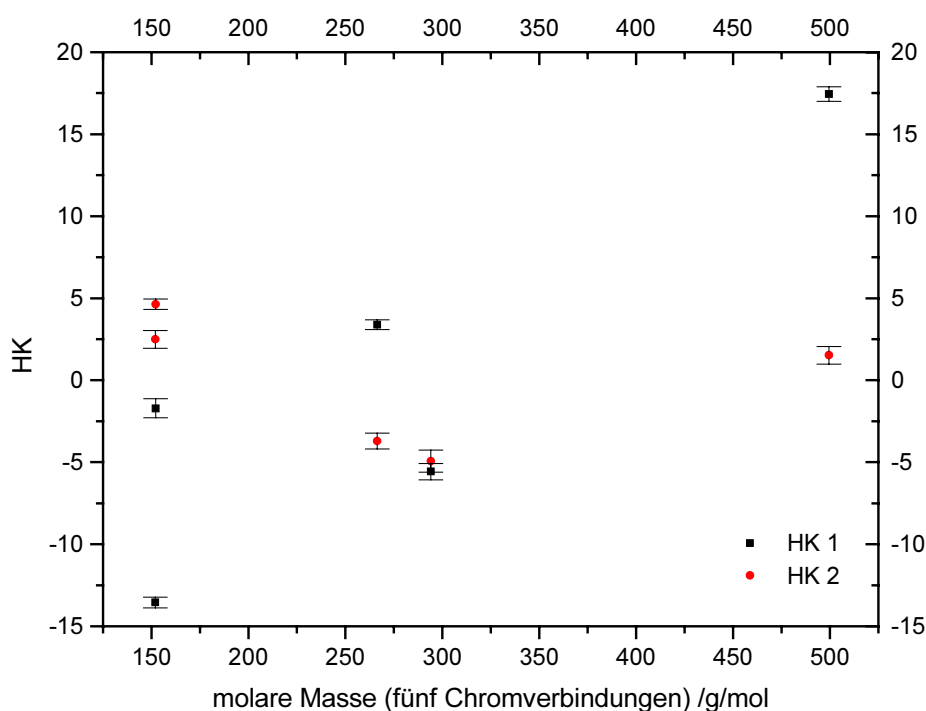


Abb. 5.52 Einfluß der molaren Masse der fünf Chromverbindungen auf die ersten zwei HK

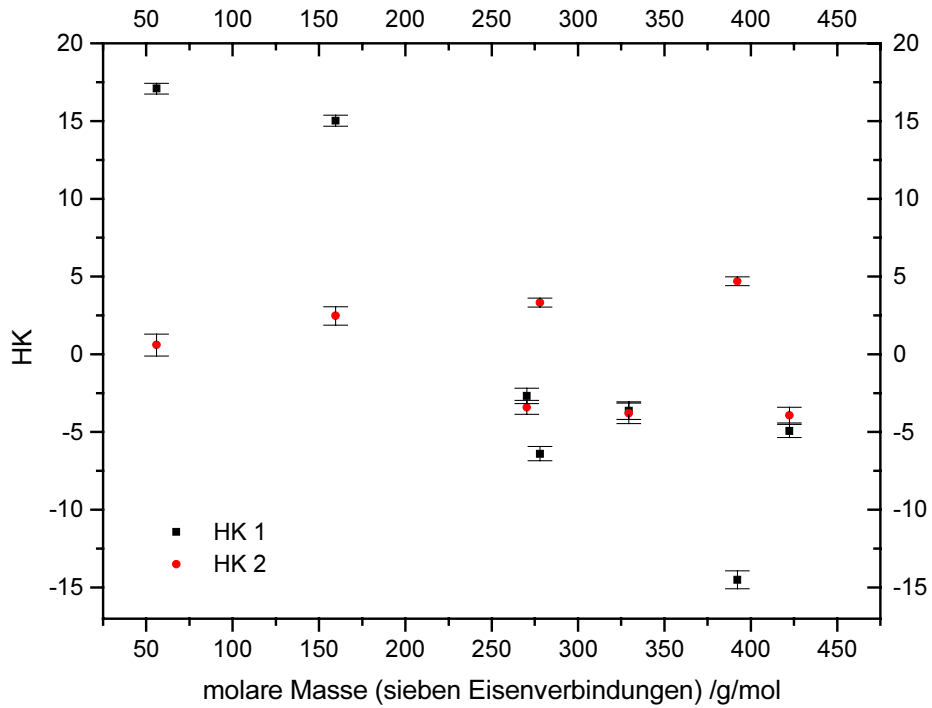


Abb. 5.53 Einfluß der molaren Masse der sieben Eisenverbindungen auf die ersten zwei HK

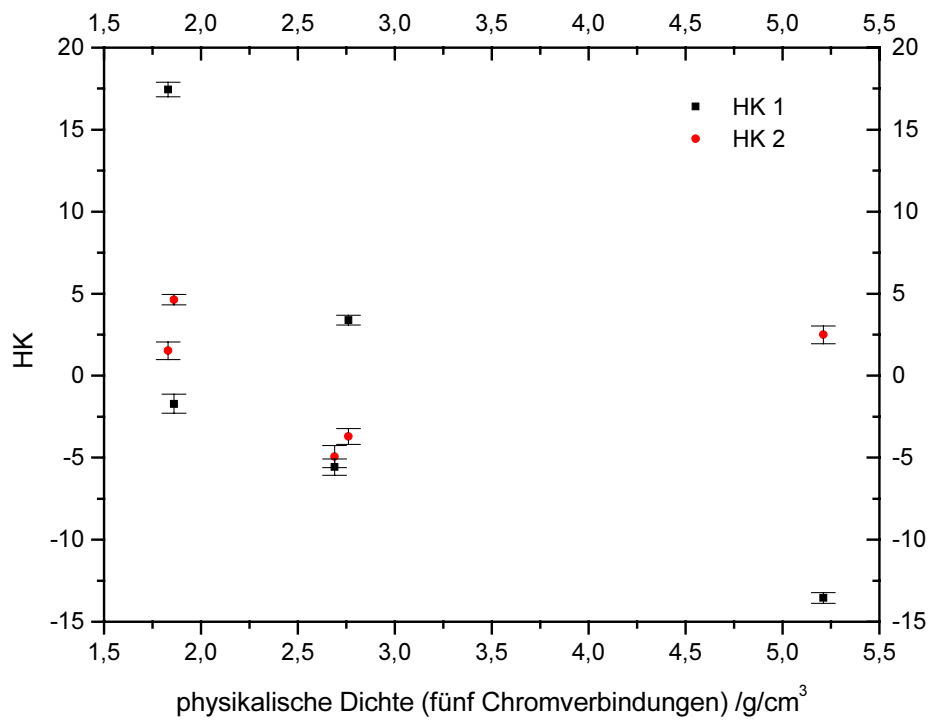


Abb. 5.54 Einfluß der physikalischen Dichte der fünf Chromverbindungen auf die ersten zwei HK

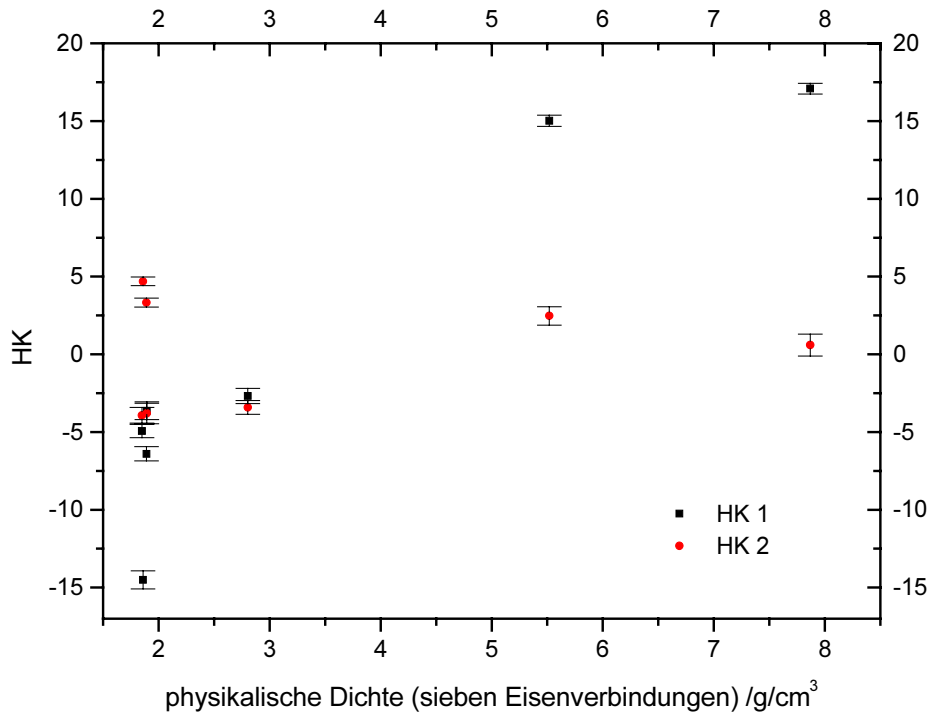


Abb. 5.55 Einfluß der physikalischen Dichte der sieben Eisenverbindungen auf die ersten zwei HK

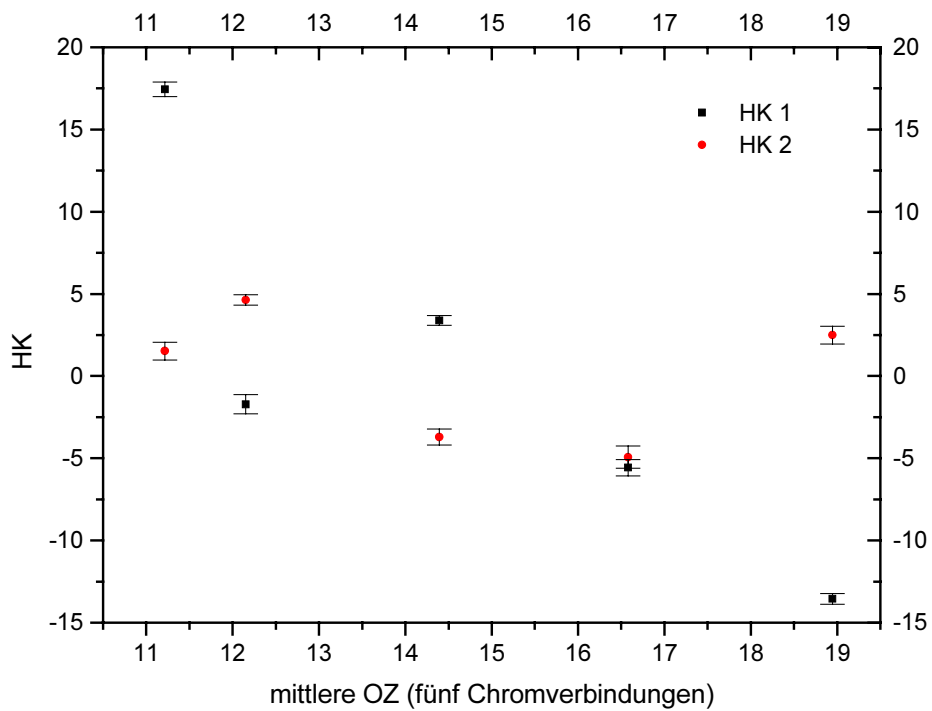


Abb. 5.56 Einfluß der mittleren Ordnungszahl der fünf Chromverbindungen auf die ersten zwei HK

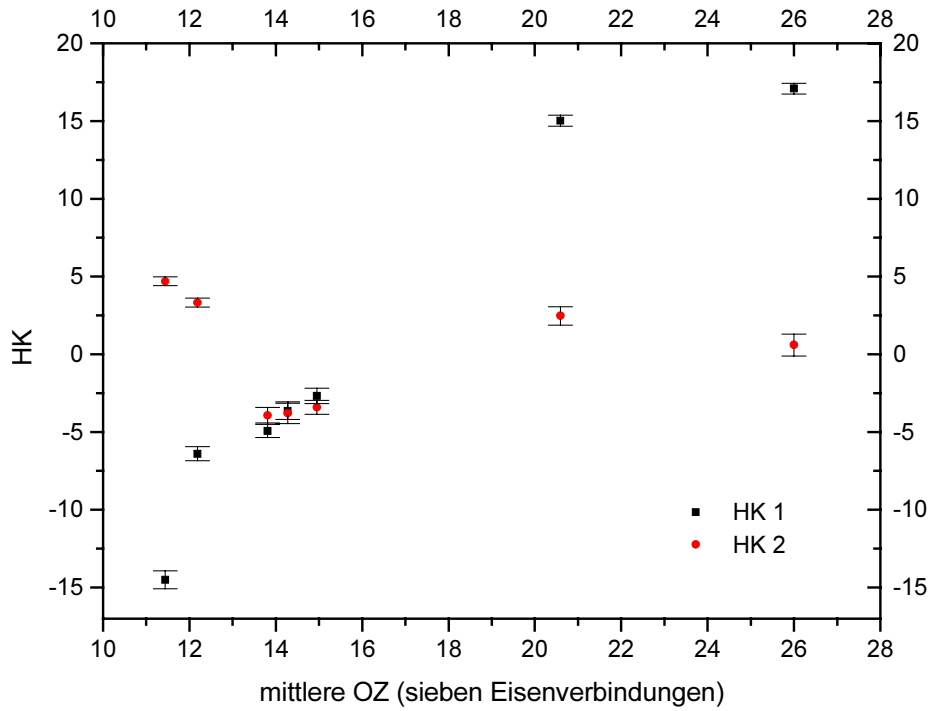


Abb. 5.57 Einfluß der mittleren Ordnungszahl der sieben Eisenverbindungen auf die ersten zwei HK

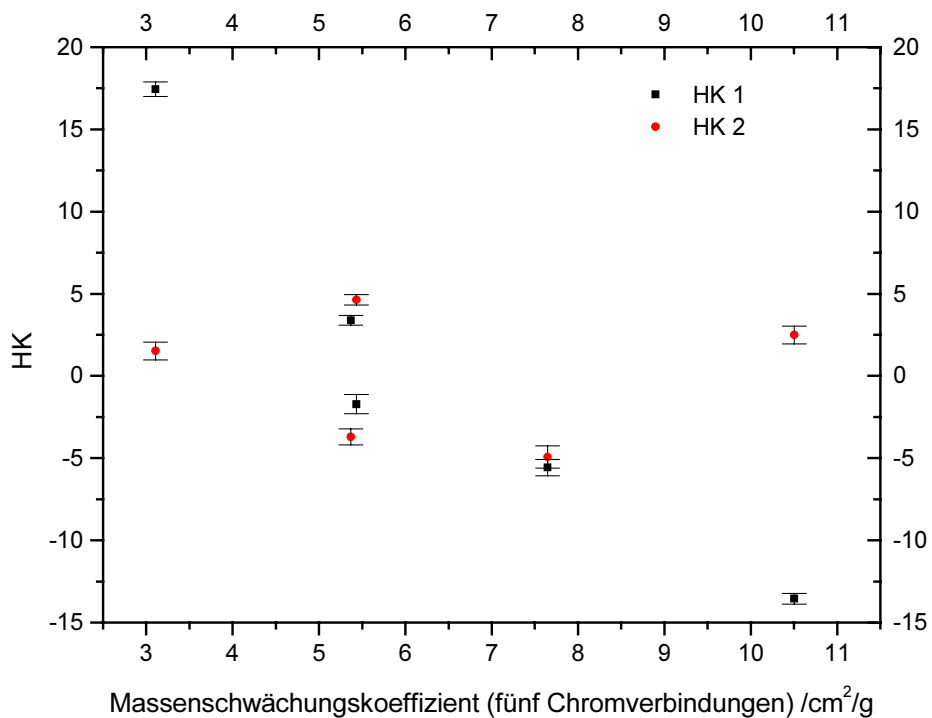


Abb. 5.58 Einfluß des Massenschwächungskoeffizienten der fünf Chromverbindungen auf die ersten zwei HK

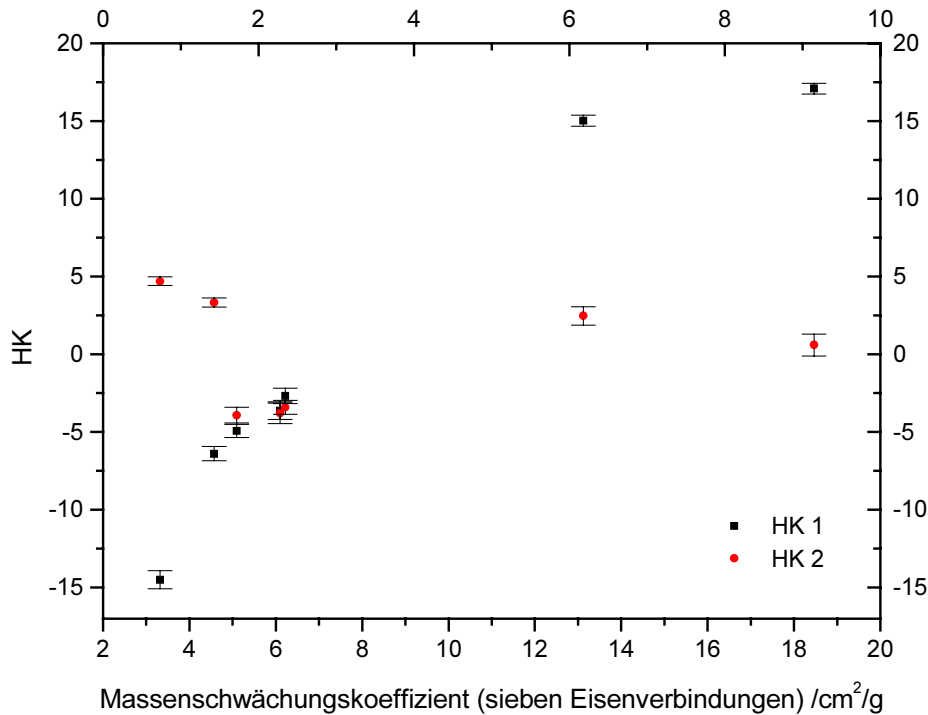


Abb. 5.59 Einfluß des Massenschwächungskoeffizienten der sieben Eisenverbindungen auf die ersten zwei HK

Mit steigender Ordnungszahl nehmen die Eindringtiefe der Röntgenstrahlen sowie die Compton- und Rayleigh-Streuintensitäten ab^[134]. Dadurch verringert sich die substanzspezifische Information bei der Wechselwirkung zwischen Röntgenstrahlung und Probe. Die durch die ersten drei HK beschriebene Varianz (Abb. 5.19, Abb. 5.26, Abb. 5.33 und Abb. 5.40) und der Einfluß der Fluoreszenzlinien, Compton- und Rayleigh-Streubereiche auf HK 2 (Abb. 5.22, Abb. 5.29, Abb. 5.36 und Abb. 5.43) nehmen mit steigender OZ ab. In Abb. 5.60 ist der Einfluß der \overline{OZ} der Substanzen aus den untersuchten Ein-Element-Gruppen auf die Ausdehnung ihrer Gruppen im Score-Plot entlang der Achsen HK 1 und HK 2, die Steigung für diesen Zusammenhang und das dazugehörige Bestimmtheitsmaß (R^2) aus der linearen Regression gezeigt. Ein Zusammenhang zwischen \overline{OZ} und HK 2 ist aufgrund der größeren Steigung und dem höheren R^2 im Vergleich zu dem Zusammenhang zwischen \overline{OZ} und HK 1 zu erkennen. In der PCA ist dieser Zusammenhang zwischen Eindringtiefe, Compton- und Rayleigh-Streuintensität und der \overline{OZ} zu erkennen und wird durch die größer werdende Ausdehnung der zweiten HK bei zunehmender \overline{OZ} in Abb. 5.60 beschrieben. Je größer die Ausdehnung der zweiten HK für eine Gruppierung von Substanzen der gleichen \overline{OZ} im Score-Plot, um so geringer ist die Eindringtiefe sowie die Compton- und Rayleigh-Streuintensität der Röntgenstrahlung in der Probe. Substanzen der Ein-Element-Gruppe mit hohen \overline{OZ} werden dadurch

im Score-Plot schlechter aufgetrennt. Der Grund für die relativ hohe Streuung der Werte in Abb. 5.60 ist im Fehlen der Information über Einflüsse anderer Parameter zu sehen.

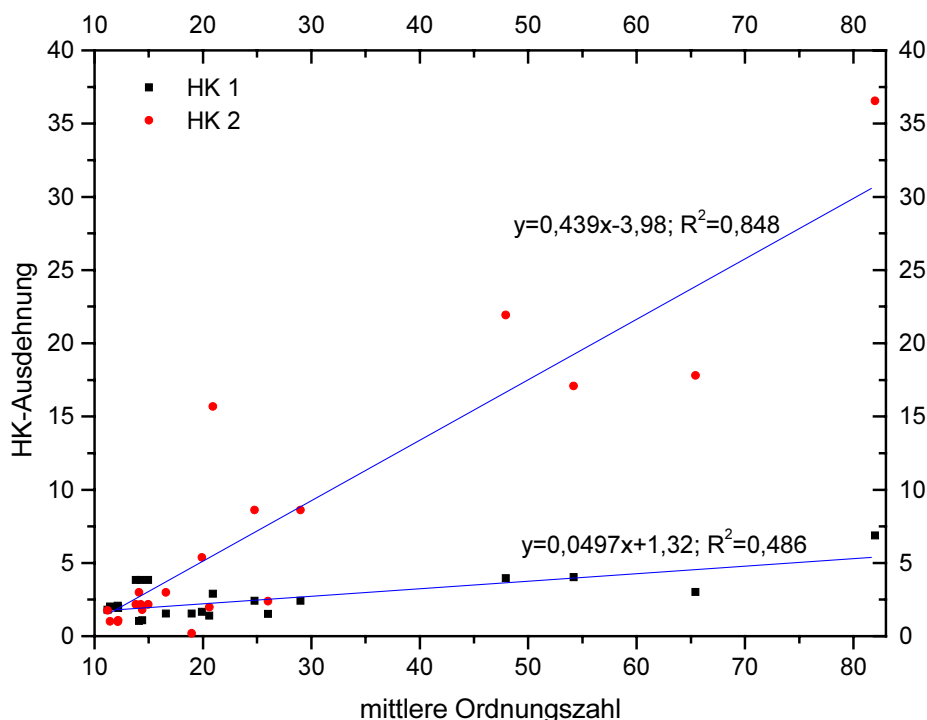


Abb. 5.60 Einfluß der mittleren Ordnungszahl der Substanzen aus den untersuchten Ein-Element-Gruppen auf die Ausdehnung ihrer Gruppen im Score-Plot entlang der Achsen HK 1 und HK 2

5.3.4 Anwendung der PCA auf die EDRFA-Spektren der Kein-Element-Gruppe

Im nächsten Schritt wurde die PCA zur Untersuchung der Kein-Element-Gruppe eingesetzt. Hierzu wurden die in Tabelle 5.9 aufgeführten Substanzen (aufsteigend nach der \overline{OZ} sortiert) in derselben PE-Verpackung (weiß, 80 ml) untersucht und ihre Daten und Spektren für die Berechnungen in den Abschnitten 5.3.4.1 bis 5.3.4.3 und 5.3.5.2 bis 5.3.5.3 verwendet. Die substanzspezifische Information der Kein-Element-Gruppe ist im Compton- und Rayleigh-Streubereich enthalten^[133,137,138,139,140,141], dessen Grenzen mit Hilfe der 1. Ableitung der EDRFA-Spektren bestimmt wurden (s. Abschnitt 5.2.3). Die Meßbedingungen sind in Abschnitt 5.3.2.3 beschrieben. Der untersuchte Spektralbereich liegt zwischen 19,6 und 26,3 keV.

Tabelle 5.9 Unterschiedliche Substanzen und ihre physikalischen Daten aus der Kein-Element-Gruppe

Substanz	M /g/mol	Dichte /g/cm ³	\overline{OZ}	$\mu_{(22,162\text{keV})}$ /cm ² /g
(NH ₄) ₂ CO ₃	96,09	1,6	6,869	0,355
H ₃ BO ₃	61,83	1,435	7,133	0,394
NaCN	49,01	1,546	8,631	0,808
BeSO ₄ x4H ₂ O	177,14	1,713	8,925	1,193
MgSO ₄ x7H ₂ O	246,48	1,68	9,034	1,141
Na ₂ CO ₃	105,99	2,533	9,075	0,865
(NH ₄) ₂ SO ₄	132,14	1,766	9,301	1,462
NaF	41,99	2,79	10,095	1,117
Na ₂ SO ₄	142,04	2,698	10,777	1,732
MgCl ₂ x6H ₂ O	203,31	1,57	11,201	2,419
CaSO ₄ x2H ₂ O	172,17	2,32	12,119	3,345
CaCO ₃	100,09	2,95	12,565	4,056
NH ₄ Cl	53,49	1,531	13,176	3,801
KCN	65,12	1,56	14,020	4,877
K ₂ CO ₃	138,21	2,428	14,049	4,688
K ₂ SO ₄	174,27	2,662	14,407	4,603
NaCl	58,44	2,163	14,641	3,972
CaCl ₂ x2H ₂ O	147,02	1,85	15,420	5,402
KF	58,1	2,49	15,729	5,594
KCl	74,56	1,984	18,047	6,843

5.3.4.1 PCA-Untersuchungen an 20 Substanzen der Kein-Element-Gruppe

In Abb. 5.61 ist der Einfluß der eingesetzten Variablen (Compton- und Rayleigh-Streubereich) auf die erste und zweite HK der 20 Substanzen aus der Kein-Element-Gruppe in einem Loading-Plot gezeigt. Der Einfluß der Spektralbereiche, die eine Aufspaltung in die entsprechende Richtung der HK bewirken, sind durch schraffierte Flächen und Beschriftungen gekennzeichnet.

Die hochenergetischen Flanken des Ag-Compton-K β - und des Ag-Rayleigh-K α -Streupeaks sind für eine Aufspaltung der Substanzen entlang hoher negativer Werte der HK 1 verantwortlich. Eine Auftrennung der Substanzen nach der HK 2 (negative Werte) bewirkt die hochenergetische Flanke des Ag-Rayleigh-K β -Streupeaks. Die niederenergetische Flanke des Ag-

Compton-K α -Streupeaks verteilt die Objekte in Richtung hoher negativer Werte für HK 1 und hohe positive Werte für HK 2 (s. Abb. 5.61).

Basierend auf dem Erkenntnis, daß durch den Compton- und Rayleigh-Streubereich von 19,6 bis 26,3 keV des EDRFA-Spektrums in der PCA die größte Varianz durch die HK beschrieben wird, lassen sich die 20 Substanzen aus Tabelle 5.9 im neuen, durch die ersten zwei HK aufgespannten Datenraum nach der PCA darstellen (s. Abb. 5.62). Weiterhin wurden die Verbindungen der Keim-Element-Gruppe Li₂CO₃, Li₂SO₄·xH₂O, Na₂SO₄, KH₂PO₄ und LiCl aus Tabelle 5.4 mit der PCA berechnet und im neuen Datenraum dargestellt (s. Abb. 5.63).

In dem Score-Plot (Abb. 5.62) sind die 20 Substanzen, die mit der herkömmlichen EDRFA-Auswertung aufgrund fehlender Röntgenfluoreszenzlinien nicht zu unterscheiden sind, entlang der ersten und zweiten HK aufgetrennt. In diesem Score-Plot sind Gruppierungen zu erkennen, die jeweils den 10 Wiederholungsmessungen der Substanzen entsprechen. Die Gruppierungen sind zum größten Teil gut aufgelöst, d.h. die spektralen Unterschiede der Substanzspektren sind für die PCA groß genug. In einigen Fällen ergeben sich allerdings Überlagerungen der Gruppen aufgrund der für die PCA zu geringen spektralen Unterschiede, so daß eine eindeutige Zuordnung nicht möglich ist. Die fünf Verbindungen aus Tabelle 5.4 sind in ihrem Score-Plot gut voneinander getrennt.

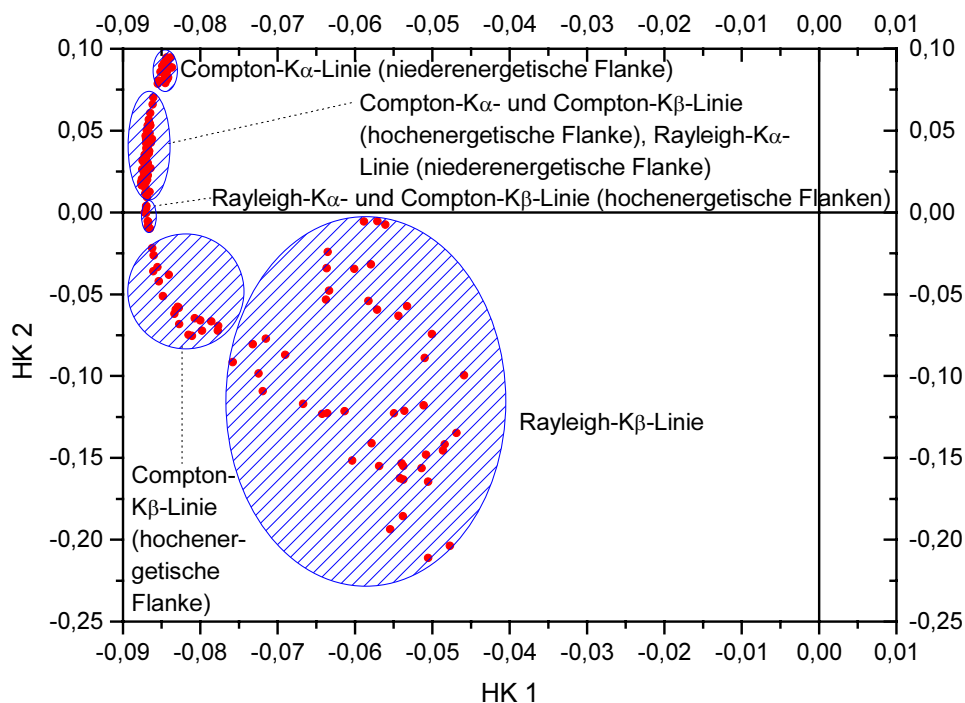


Abb. 5.61 Loading-Plot der Variablen (Compton- und Rayleigh-Streubereich) für die erste und zweite HK der 20 Substanzen aus der Keim-Element-Gruppe, berechnet mit der Software SCAN for Windows[®]

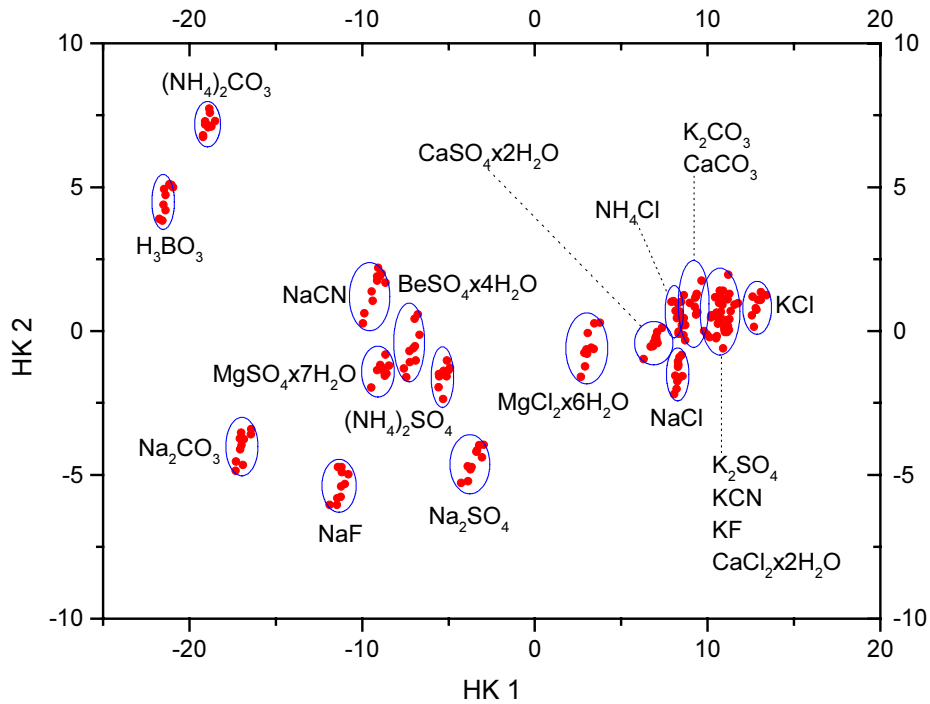


Abb. 5.62 Score-Plot der 20 Substanzen aus Tabelle 5.9, berechnet mit der Software SCAN for Windows[®], Meßzeit: 20 s, Compton- und Rayleigh-Streubereich von 19,6 bis 26,3 keV

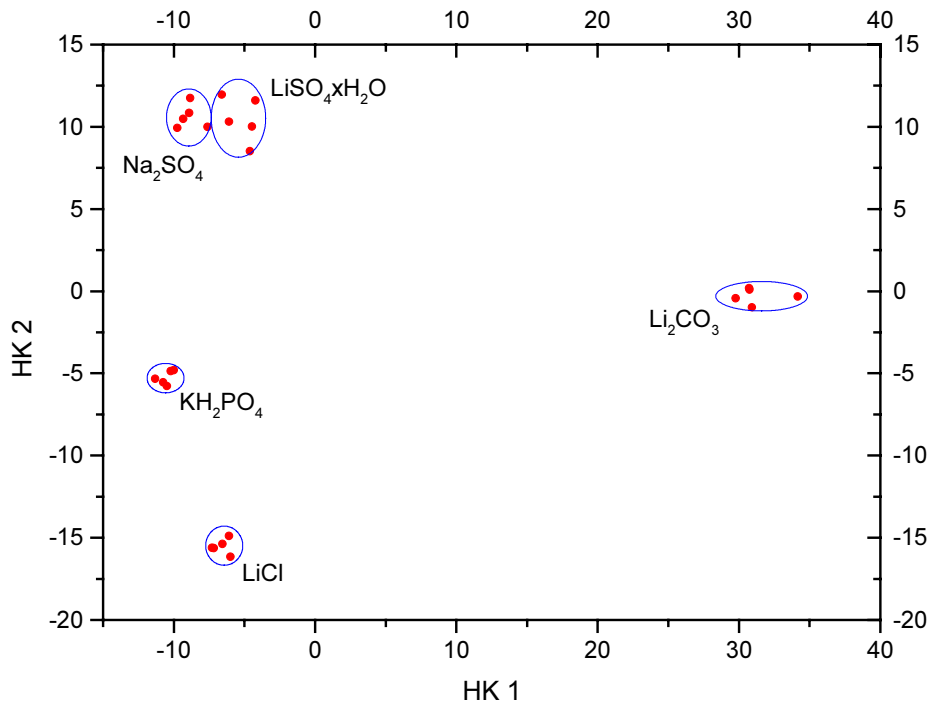


Abb. 5.63 Score-Plot der 5 Substanzen aus Tabelle 5.4, berechnet mit der Software SIMCA-P[®], Meßzeit: 30 s, Compton- und Rayleigh-Streubereich von 19,6 bis 26,3 keV

5.3.4.2 Einfluß der physikalischen Parameter auf die Hauptkomponenten (HK) der Substanzen aus der Kein-Element-Gruppe

In den Abb. 5.64 bis 5.66 sind die Einflüsse der physikalischen Parameter (molare Masse, physikalische Dichte, \overline{OZ} und Massenschwächungskoeffizient) der 20 Substanzen der Kein-Element-Gruppe auf HK 1 und HK 2 graphisch dargestellt.

Ein Zusammenhang zwischen der molaren Masse (s. Abb. 5.64), bzw. der physikalischen Dichte (s. Abb. 5.65) und den ersten zwei HK ist nicht zu erkennen. Der Zusammenhang zwischen der \overline{OZ} und der HK 1 (s. Abb. 5.66) entspricht dem Verlauf zwischen der \overline{OZ} und dem Rayleigh-/Compton-Intensitätsverhältnis (s. Abb. 5.15). Zwischen der \overline{OZ} und der zweiten HK ist kein Zusammenhang zu erkennen (s. Abb. 5.66). Allerdings ist die Streuung für die Auftragung der \overline{OZ} gegen die erste HK geringer als bei der Auftragung der \overline{OZ} gegen das Rayleigh/Compton-Intensitätsverhältnis, denn bei Anwendung der PCA wird die eigentliche Information von den Fremdeinflüssen wie dem Rauschen abgetrennt. Der Zusammenhang zwischen HK 1 und μ ist im Vergleich zur \overline{OZ} eindeutiger (s. Abb. 5.66 und Abb. 5.67). Eine logarithmische Anpassung der Zusammenhänge $HK 1 = f(\overline{OZ})$ und $HK 1 = f(\mu)$ ergibt für $HK 1 = f(\mu)$ ein höheres R^2 (s. Tabelle 5.10).

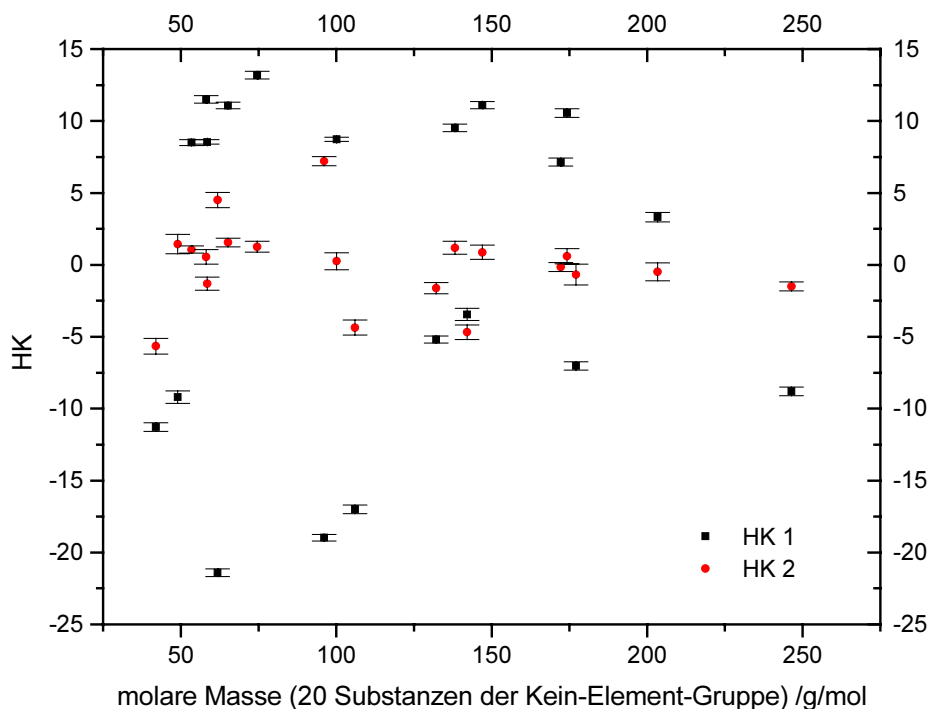


Abb. 5.64 Einfluß der molaren Masse der 20 Substanzen der Kein-Element-Gruppe auf die ersten zwei HK

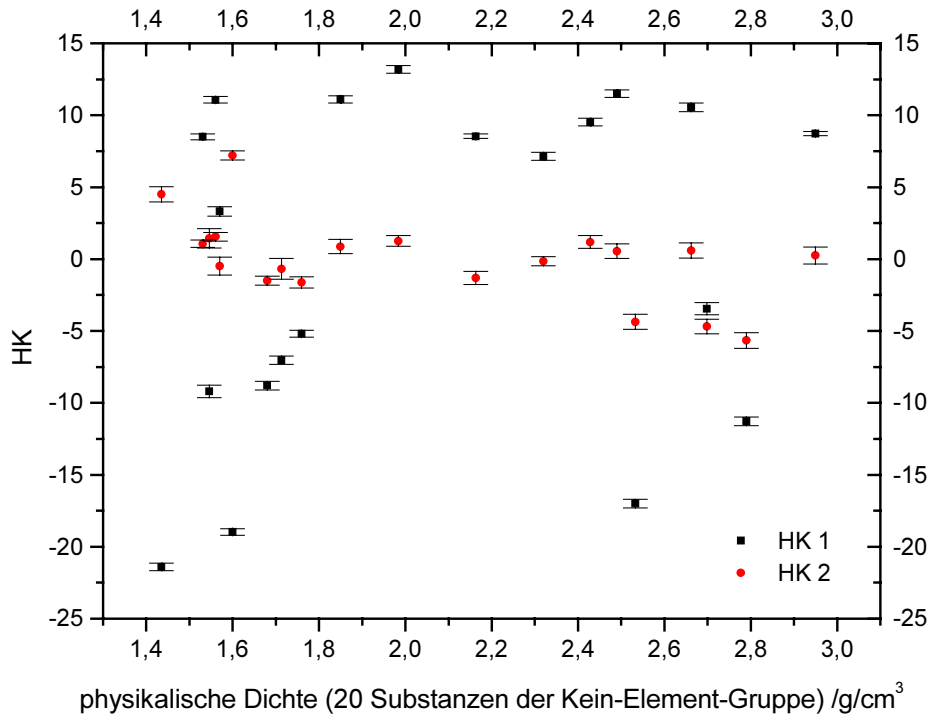


Abb. 5.65 Einfluß der physikalischen Dichte der 20 Substanzen der Kein-Element-Gruppe auf die ersten zwei HK

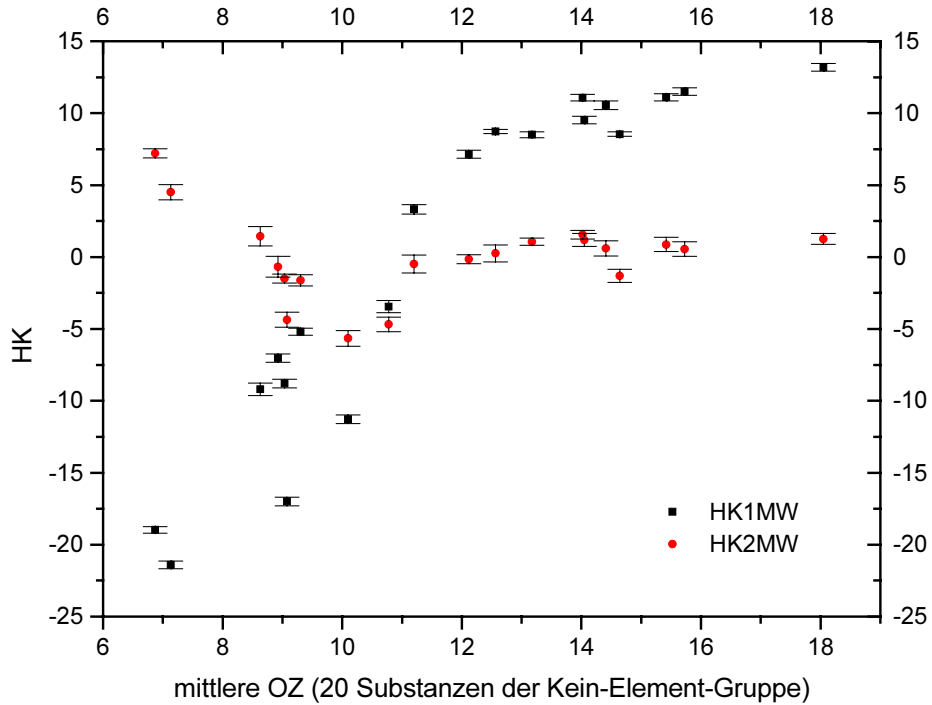


Abb. 5.66 Einfluß der mittleren Ordnungszahl der 20 Substanzen der Kein-Element-Gruppe auf die ersten zwei HK

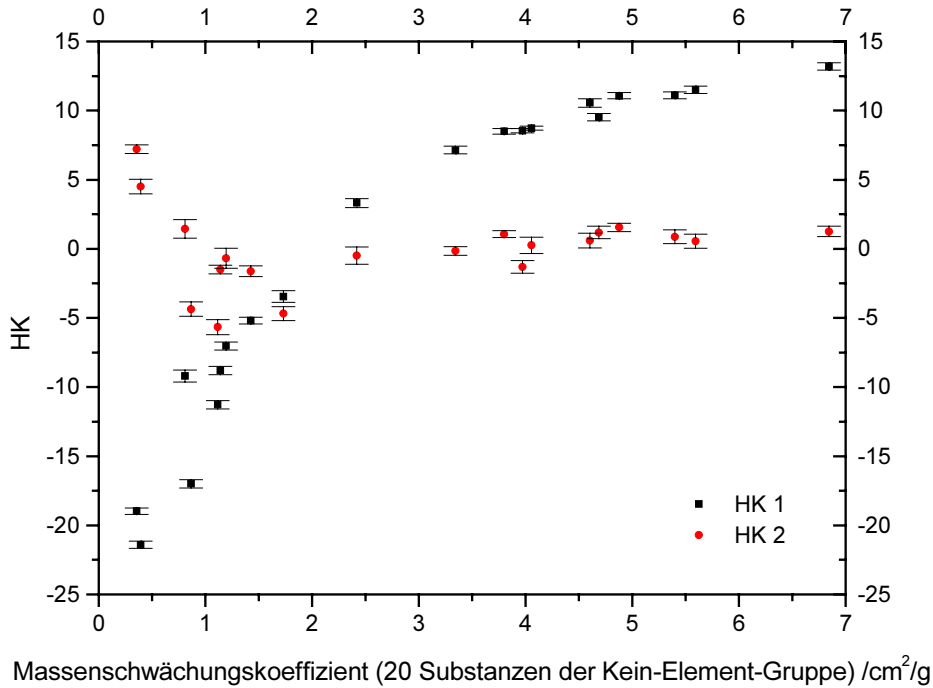


Abb. 5.67 Einfluß des Massenschwächungskoeffizienten der 20 Substanzen der Keim-Element-Gruppe auf die ersten zwei HK

Tabelle 5.10 Logarithmische Anpassung (Fit) und deren R^2 für die beiden Funktionen $HK\ 1 = f(\overline{OZ})$ und $HK\ 1 = f(\mu)$

Funktion	Fit	R^2
$HK\ 1 = f(\overline{OZ})$	$HK\ 1 = 40 \ln \overline{OZ} - 97,2$	0,903
$HK\ 1 = f(\mu)$	$HK\ 1 = 12,5 \ln \mu - 9,47$	0,969

Analog der Korrelationsuntersuchungen an den Ein-Element-Gruppen wird die in Abschnitt 5.3.1 erwähnte Elektronendichte durch den Massenschwächungskoeffizienten exakter als durch die \overline{OZ} beschrieben. Neben den zur Verfügung stehenden physikalischen Eigenschaften beeinflussen noch andere Parameter wie z.B. Verpackungsmaterial, Verpackungswanddicke, Schüttdichte, Korngröße und der Kristallwassergehalt die HK. Da diese Parameter bei den Messungen nicht bekannt waren, ist mit Hilfe der hier vorliegenden Daten kein endgültiger Zusammenhang zwischen verschiedenen physikalischen Eigenschaften und der Identifikation nach HK zu beschreiben.

5.3.4.3 Überlagerungen im Streubereich der Kein-Element-Gruppe durch Substanzen der Ein-Element-Gruppe

Zur Unterteilung innerhalb der Kein-Element-Gruppe wird der Streubereich für die PCA verwendet. Einige Substanzen aus den Ein-Element-Gruppen ergeben EDRFA-Spektren, in denen Fluoreszenzlinien im Compton- und Rayleigh-Streubereich liegen. Das sind Substanzen, die Ag (Überlagerung mit den Ag-Linien aus der Röntgenröhre), Cd (Überlagerung mit der Ag-Compton-K β -Linie) und Sn (Überlagerung mit der Ag-Rayleigh-K β -Linie) beinhalten. Weiterhin werden Signale von Ti-haltige Verbindungen von dem Ti-Fluoreszenzlinien aus dem Weißpigment der PE-Verpackung überlagert. Um eine Falschklassifizierung zu vermeiden, wurden die Substanzen TiO_2 , Ag_2SO_4 , $\text{Cd}(\text{CH}_3\text{COO})_2$ und SnSO_4 (in derselben weißen 80 ml PE-Verpackung) zusammen mit den Substanzen aus Tabelle 5.9 mit der PCA bezüglich ihrer Lage in dem durch die ersten zwei HK aufgespannten neuen Datenraum untersucht. Das Ergebnis ist in Abb. 5.68 gezeigt. Die vier Substanzen aus den Ein-Element-Zuordnungsgruppen Titan, Silber, Cadmium und Zinn können im Score-Plot im Vergleich zu den restlichen 20 Substanzen unterschieden werden.

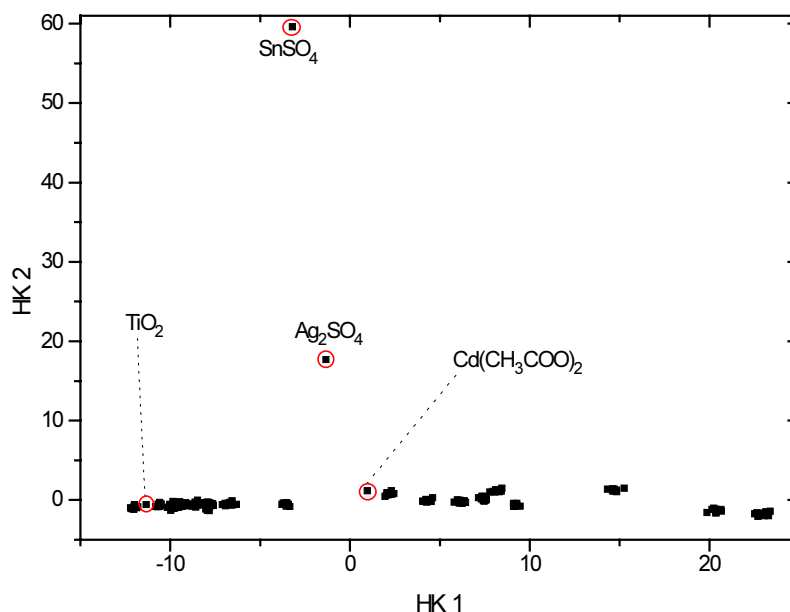


Abb. 5.68 Score-Plot der Substanzen TiO_2 , Ag_2SO_4 , $\text{Cd}(\text{CH}_3\text{COO})_2$ und SnSO_4 und den 20 Substanzen aus Tabelle 5.9, berechnet mit der Software SCAN for Windows[®], Meßzeit: 20 s, Spektralbereiche: 4,0 bis 5,2 und 19,6 bis 26,3 keV

5.3.5 Der Einsatz der RDA auf die EDRFA-Spektren

5.3.5.1 Anwendung der RDA auf die EDRFA-Spektren der Ein-Element-Gruppe

Hauptbestandteil für die abfallspezifische Zuordnung von zurückgenommenen Chemikalien ist die Zuordnung des Verpackungsinhalts zur auf dem Etikett ausgewiesenen Substanz. Diese Zuordnung ist die Klassifizierung und Substanzen werden damit in definierte Zuordnungsgruppen einsortiert.

Wie in Abschnitt 5.3.2.2 erwähnt, ist aufgrund der Datenmatrixstruktur die Klassifizierungsauswahl auf die RDA beschränkt (die Zahl der Variablen ist größer als die Zahl der Objekte).

Die RDA kann mit zwei Arten von Variablen durchgeführt werden:

- Die Variablen sind der ausgewählte Spektralbereich, oder
- die Variablen sind die ersten signifikanten HK aus der PCA.

Die signifikanten HK beschreiben in dem PCA-Modell des ursprünglichen Datensatzes die vom Fehler (z.B. Rauschen) separierte analytische Information. Die Anzahl der ersten signifikanten HK wird von der Software SCAN for Windows[®] über das Eigenwert-Eins-Kriterium berechnet^[186]. Der mittlere Eigenwert der Korrelationsmatrix für die HK (standardisierte Daten) ist 1. Signifikante HK besitzen einen Eigenwert < 1 . Zur Bestimmung der signifikanten Anzahl an HK kann ebenfalls eine heuristische Methode eingesetzt werden. Hierbei sind die HK signifikant, nach denen die Steigung im Diagramm Eigenwert-HK < 1 wird. In Abb. 5.69 sind die ersten zehn HK aus der PCA der Fluoreszenzlinien-, Compton- und Rayleigh-Streubereiche der fünf Chromverbindungen (Tabelle 5.6) und der sieben Eisenverbindungen (Tabelle 5.7) gezeigt.

Alle zehn HK liegen sowohl bei dem Chrom- als auch bei dem Eisendatensatz oberhalb des Eigenwerts von 1. Nach der dritten HK ist die Steigung < 1 , so daß die ersten drei HK als signifikant angesehen werden (s. Abb. 5.69).

Für den Chrom- und Eisendatensatz wurde die RDA mit den ersten drei HK sowie der Kombination aus den Fluoreszenzlinien-, Compton- und Rayleigh-Streubereichen durchgeführt. In Abb. 5.70 ist das Falschklassifizierungsrisiko (mit einer Kreuzvalidierung berechnet) in Abhängigkeit der zwei Arten von Variablen (HK und Spektralbereich) dargestellt^[194]. Das Falschklassifizierungsrisiko gibt für das jeweilige RDA-Modell den Grad der Falschzuordnung im Bereich von 0 bis 1 an. Bei der Kreuzvalidierung wurde nach Definition der Klassen ein bestimmter, freiwählbarer Anteil an Proben aus dem Lerndatensatz entnommen und mit dem RDA-Modell auf

Richtigkeit des Modells überprüft. In diesen Berechnungen wurde die Hälfte der Proben (fünf pro Klasse) für die Kreuzvalidierung verwendet.

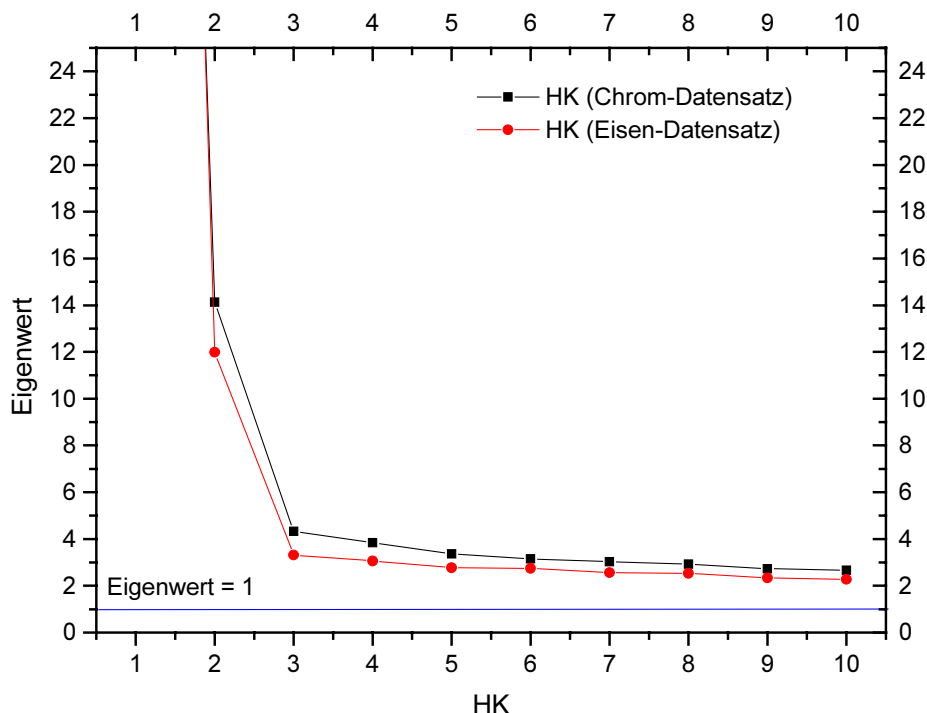


Abb. 5.69 Auftragung der ersten zehn HK als Funktion ihrer Eigenwerte für den Chrom- und Eisendatensatz, berechnet mit der Software SCAN for Windows®

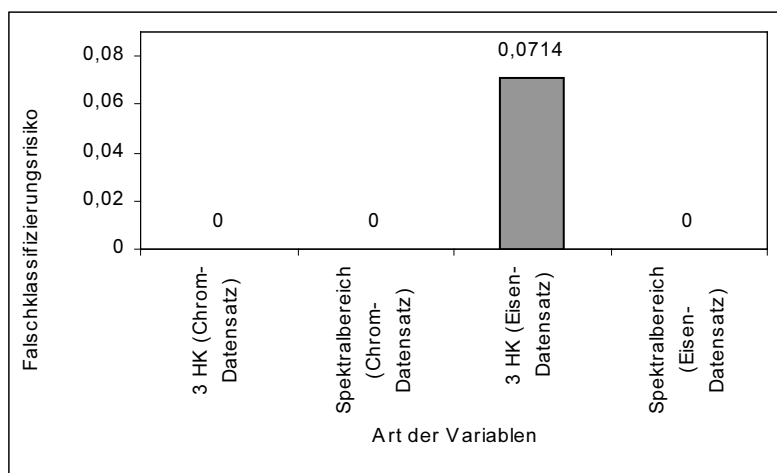


Abb. 5.70 Auftragung des Falschklassifizierungsrisikos (berechnet aus einer Kreuzvalidierung) in Abhängigkeit der Variablenart (HK oder Spektralbereich) für den Chrom- und Eisendatensatz, berechnet mit der Software SCAN for Windows®

Das ermittelte Falschklassifizierungsrisiko ist für den Chromdatensatz unabhängig von der Variablenart jeweils gleich 0. Der Eisendatensatz zeigt bei Verwendung der ersten drei HK als Variablen ein Falschklassifizierungsrisiko von 0,0714 (= 7,14%), während beim Einsatz der Spektralbereiche als Variablen das Falschklassifizierungsrisiko gleich 0 ist (s. Abb. 5.70). So-

mit ist der Einsatz der HK als Variablen für die RDA beim Eisendatensatz aufgrund des höheren Falschklassifizierungsrisikos nicht geeignet.

Das Ergebnis der Klassifizierung für die fünf Chromverbindungen (Variablen: Spektralbereiche) ist in Abb. 5.71 dargestellt. Auf der x-Achse sind die von dem RDA-Algorithmus zugewiesenen Klassen und auf der y-Achse die wahren, zuvor definierten Klassen aufgetragen. Jede Ziffer entspricht einer Substanzklasse. Bei einer richtigen Klassifizierung sind wahre und zugewiesene Klassennummer identisch. Erkennt der RDA-Algorithmus keine spektralen Unterschiede zwischen den einzelnen Klassen, erfolgt eine Falschklassifizierung in spektral ähnliche Klassen. Die eindeutige Klassifizierung des Chromdatensatzes, unabhängig von der Verwendung der ersten drei HK oder der Spektralbereiche als Variablen, war aufgrund der guten Auftrennung der Substanzen im Score-Plot (s. Abb. 5.25) zu erwarten.

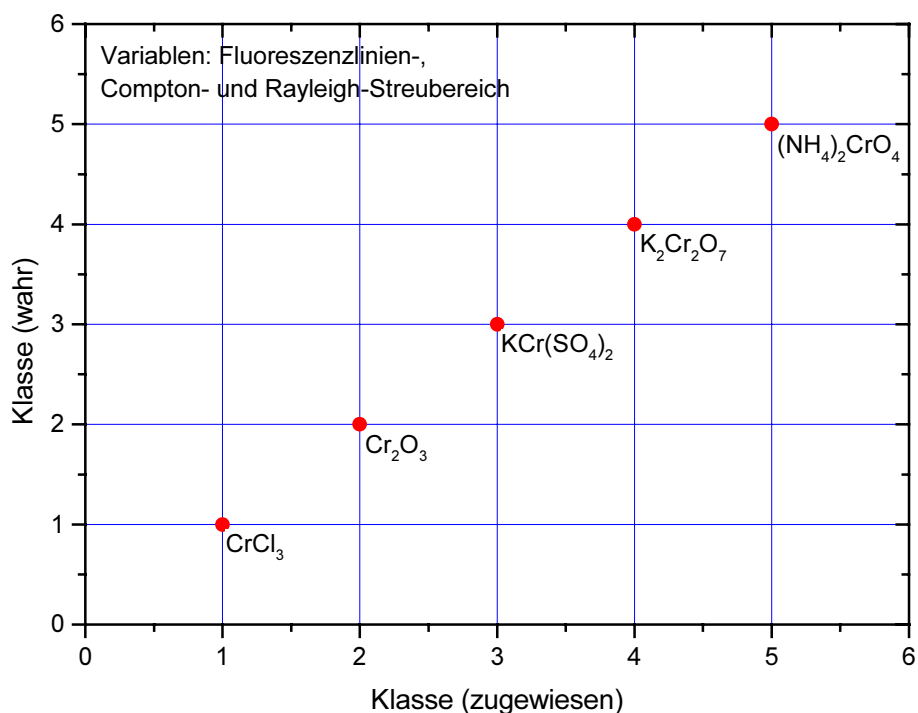


Abb. 5.71 Graphische Darstellung der RDA-Ergebnisse der fünf Chromverbindungen (fünf Klassen) aus Tabelle 5.6 (Klassifizierungsüberprüfung mittels Kreuzvalidierung), berechnet mit der Software SCAN for Windows[®], Meßzeit: 20 s, Variablen: Spektralbereiche

In Abb. 5.72 ist das Ergebnis der RDA-Klassifizierung mit den Spektralbereichen als Variablen für die sieben Eisenverbindungen gezeigt. Im Gegensatz zum Score-Plot (s. Abb. 5.32) ist die Klassifizierung eindeutig. Die spektralen Unterschiede der für die Kreuzvalidierung ausgewählten Spektren der Eisenverbindungen waren entgegen den Erwartungen aus dem Score-Plot groß genug, um die Eisenverbindungen eindeutig zu klassifizieren. Hingegen entspricht das Ergebnis der RDA-Berechnungen mit den HK 1 – 3 als Variablen aus Abb. 5.73 eher den

Erwartungen aus dem Score-Plot der sieben Eisenverbindungen, denn die spektralen Unterschiede der Verbindungen $K_4[Fe(CN)_6] \cdot 3H_2O$, $K_3[Fe(CN)_6]$, und $FeCl_3 \cdot 6H_2O$ sind so ähnlich, daß mit den HK als Variablen keine eindeutige Zuordnung dieser drei Substanzen mit der RDA möglich ist.

5.3.5.2 Anwendung der RDA auf die EDRFA-Spektren der Keim-Element-Gruppe

Der Datensatz für die 20 Substanzen aus Tabelle 5.9 wurde mit der RDA bearbeitet. Analog der Ein-Element-Gruppen standen zwei Arten von Variablen zur Verfügung. Das sind die ersten, signifikanten HK (ermittelt nach dem Eigenwert-Eins-Kriterium, s. Abschnitt 5.3.5.1) bzw. der Compton- und Rayleigh-Streubereich (s. Abschnitt 5.3.4.3). Die Ergebnisse der Eigenwert-Eins-Berechnungen für die Anzahl an signifikanten HK ist in Abb. 5.74 für die ersten fünf HK aus der PCA des Compton- und Rayleigh-Streubereichs des Keim-Element-Gruppe-Datensatzes dargestellt. Aus dieser Abbildung liegen die Eigenwerte der ersten drei HK über 1 und sind somit signifikant.

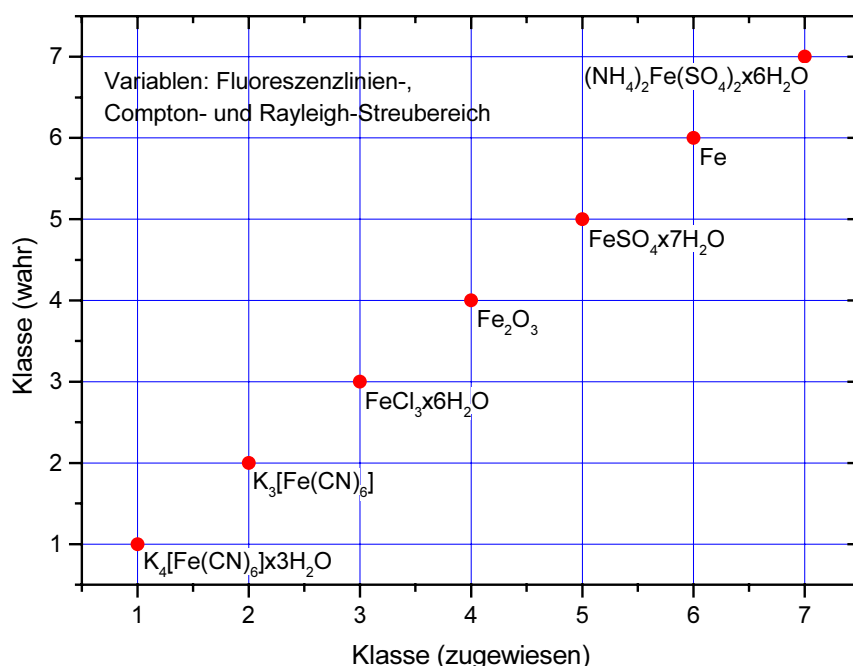


Abb. 5.72 Graphische Darstellung der RDA-Ergebnisse mit den Spektralbereichen als Variablen der sieben Eisenverbindungen (sieben Klassen) aus Tabelle 5.7 (Klassifizierungsüberprüfung mittels Kreuzvalidierung), berechnet mit der Software SCAN for Windows[®], Meßzeit: 20 s

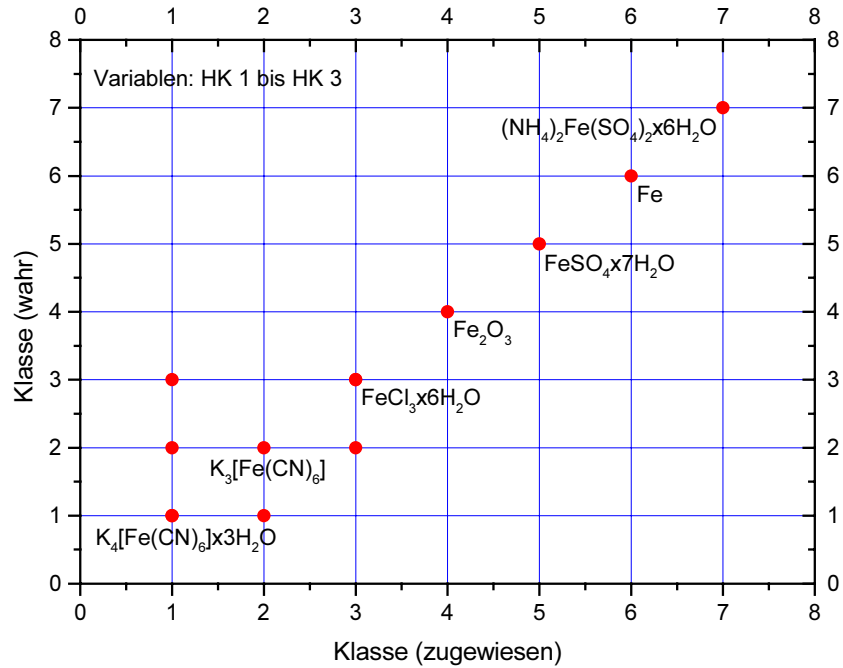


Abb. 5.73 Graphische Darstellung der RDA-Ergebnisse mit den HK 1 bis HK 3 als Variablen der sieben Eisenverbindungen (sieben Klassen) aus Tabelle 5.7 (Klassifizierungsüberprüfung mittels Kreuzvalidierung), berechnet mit der Software SCAN for Windows®, Meßzeit: 20 s

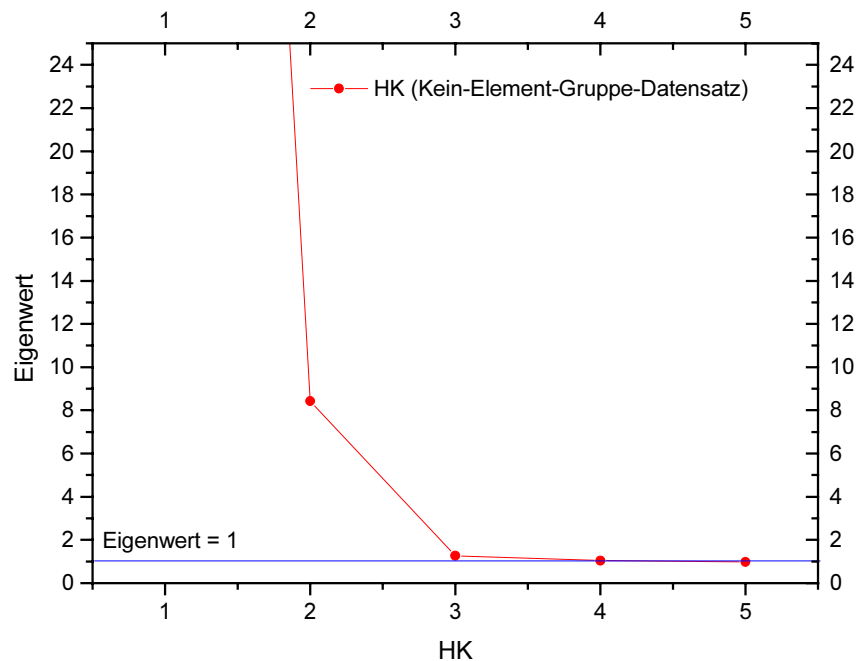


Abb. 5.74 Auftragung der ersten fünf HK als Funktion ihrer Eigenwerte für den Datensatz der Kein-Element-Gruppe, berechnet mit der Software SCAN for Windows®

In Abhängigkeit der zwei Arten von Variablen (HK 1 bis HK 3 bzw. Spektralbereich) wurde mit Hilfe der Kreuzvalidierung das Falschklassifizierungsrisiko des RDA-Modells für den Kein-

Element-Gruppe-Datensatz überprüft. Zusätzlich wurde dieser Datensatz geteilt. Mit dem einen Teil wurden die Klassen definiert (Lerndatensatz), mit dem anderen Teil (Testdatensatz) wurde das RDA-Modell überprüft. In Abb. 5.75 ist in Form eines Balkendiagramms das Falschklassifizierungsrisiko (berechnet mit einer internen Kreuzvalidierung, KV, und für einen Testdatensatz, TD) in Abhängigkeit der Variablenart (HK bzw. Spektralbereich) dargestellt.

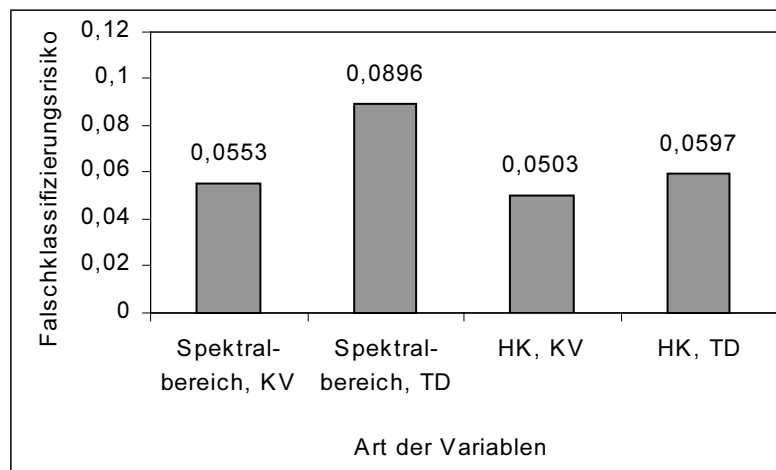


Abb. 5.75 Auftragung des Falschklassifizierungsrisikos (berechnet für eine interne Kreuzvalidierung, KV, und für einen Testdatensatz, TD) in Abhängigkeit der Variablenart (HK oder Spektralbereich) für den Keim-Element-Datensatz, berechnet mit der Software SCAN for Windows®

Aus Abb. 5.75 werden zwei Aussagen deutlich:

- Die Kreuzvalidierung als Überprüfung dieses RDA-Modells weist ein geringeres Falschklassifizierungsrisiko als die Überprüfung mit einem Testdatensatz auf.
- Bei der RDA-Modellüberprüfung mit einem Testdatensatz zeigt der Einsatz der signifikanten HK als Variablen ein geringeres Falschklassifizierungsrisiko als die Verwendung des Compton- und Rayleigh-Streubereichs als Variablen (s. Abb. 5.75).

Die RDA-Ergebnisse sind für den Compton- und Rayleigh-Streubereich als Variablen in Abb. 5.76 und für die HK als Variablen in Abb. 5.77 graphisch dargestellt. In Bezug auf die Anzahl der eindeutig klassifizierten Substanzen aus dem Testdatensatz ist kein signifikanter Unterschied zu erkennen. Bei beiden Arten von Variablen (Spektralbereich und HK) wurden 15 Substanzen eindeutig klassifiziert. Die restlichen fünf Substanzen wurden aufgrund ihrer grossen spektralen Ähnlichkeit falsch klassifiziert.

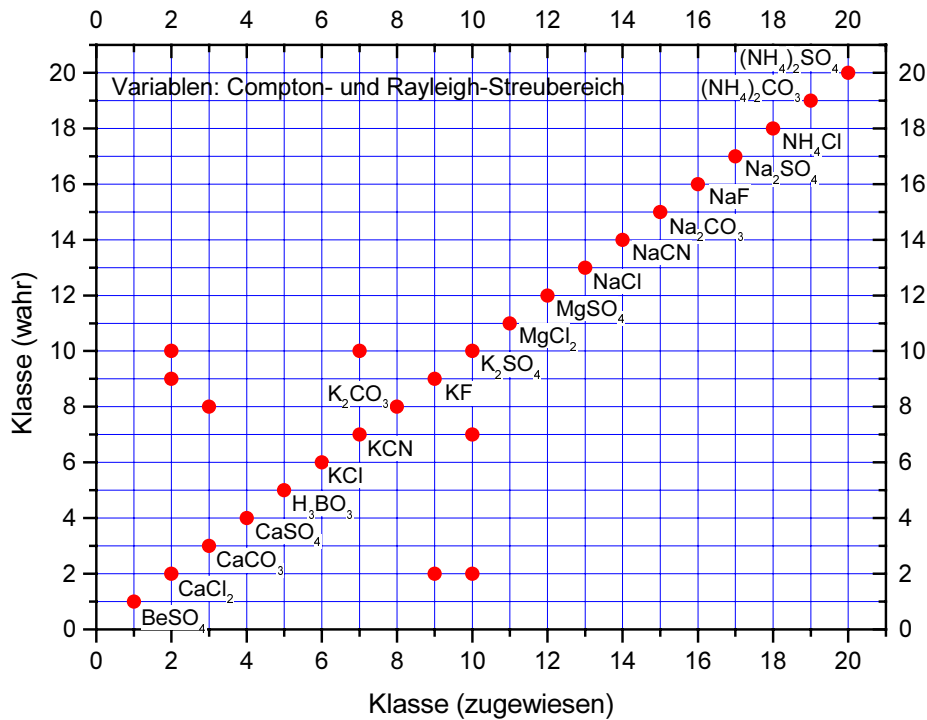


Abb. 5.76 Graphische Darstellung der RDA-Ergebnisse mit dem Compton- und Rayleigh-Streubereich als Variablen von 20 Substanzen (20 Klassen) des Testdatensatzes aus Tabelle 5.9, berechnet mit der Software SCAN for Windows[®], Meßzeit: 20 s,

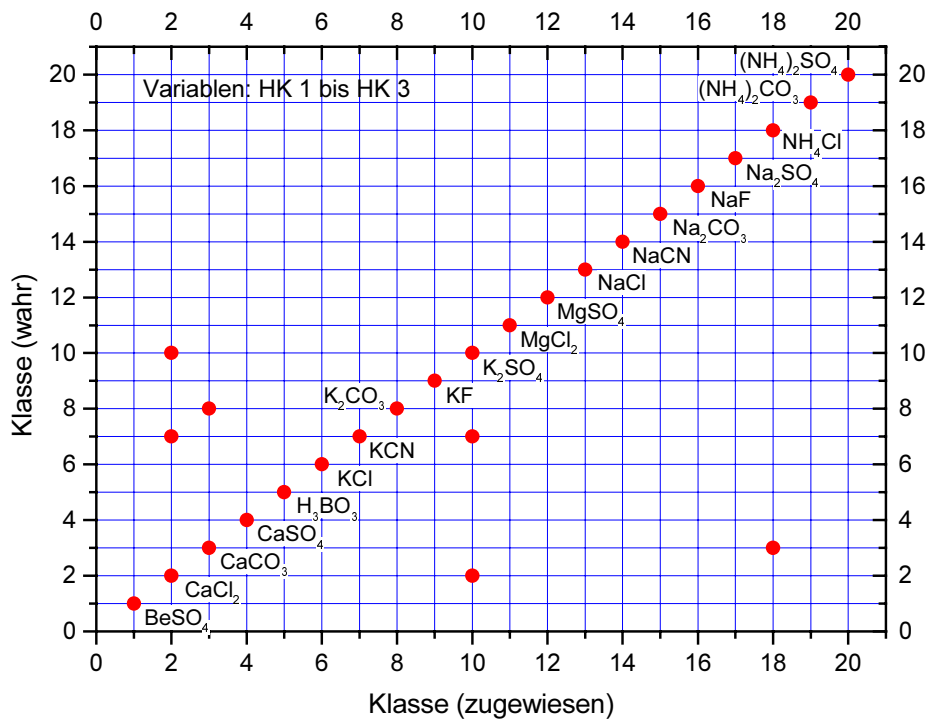


Abb. 5.77 Graphische Darstellung der RDA-Ergebnisse mit den HK 1 bis HK 3 als Variablen von 20 Substanzen (20 Klassen) des Testdatensatzes aus Tabelle 5.9, berechnet mit der Software SCAN for Windows[®], Meßzeit: 20 s

5.3.5.3 Validierung der RDA-Ergebnisse am Beispiel der Kein-Element-Gruppe

Die EDRFA-Spektren der 20 Substanzen aus Tabelle 5.9 wurden in vier Meßreihen bei einer Meßzeit von 30 s aufgenommen. Um zu zeigen, daß die bisherige Klassifizierung einer Meßreihe mit dem dazugehörigen RDA-Modell kein zufälliges Ergebnis war, wurde der Datensatz (jeweils der Compton- und Rayleigh-Streubereich) der ersten Meßreihe als Lerndatensatz für die RDA und eine willkürlich gewählte Anzahl an Substanzen aus den drei anderen Meßreihen als Testdatensatz verwendet.

Diese Berechnungen wurden mit zwei Arten von Variablen (Compton- und Rayleigh-Streubereich bzw. den signifikanten HK) durchgeführt. Die HK für den Lern- und Testdatensatz müssen in einer gemeinsamen Datenmatrix mit der PCA berechnet werden, sonst ist die Richtigkeit des RDA-Modells aufgrund von zwei PCA-Modellen mit unterschiedlichen HK nicht gegeben.

Zunächst wurde die Anzahl der signifikanten HK nach dem Eigenwert-Eins-Kriterium für die RDA-Berechnung mit den HK als Variablen bestimmt. In Abb. 5.78 sind die ersten zehn HK gegen ihren Eigenwert aufgetragen. Der Eigenwert ist ab der vierten HK unterhalb von 1, somit sind die ersten drei HK signifikant und werden als Variablen bei der RDA verwendet. Der Eigenwert für HK 1 ist aufgrund seines hohen Wertes nicht mehr in diesem Maßstab abgebildet.

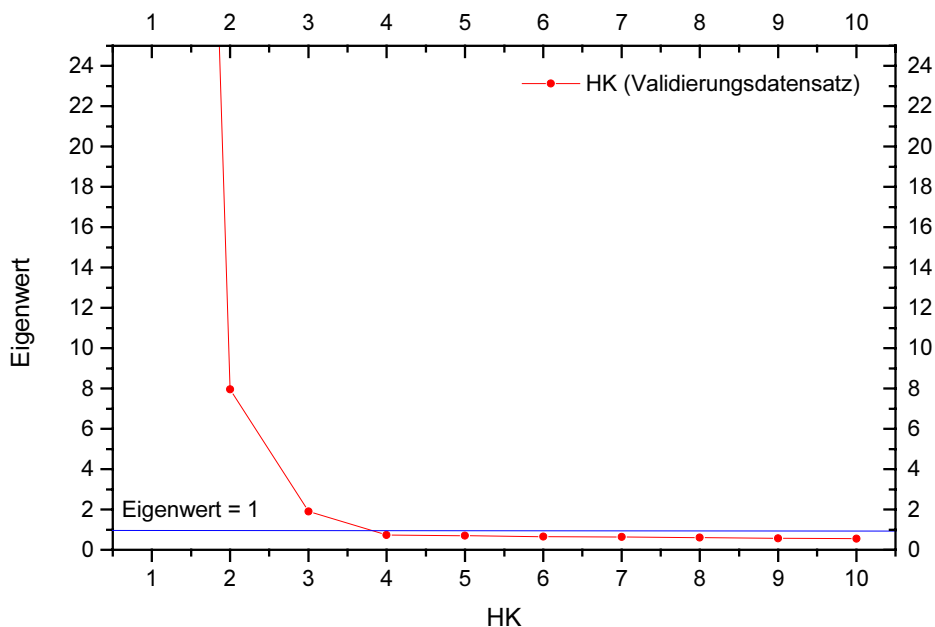


Abb. 5.78 Auftragung der ersten zehn HK als Funktion ihrer Eigenwerte für den Validierungsdatensatz der Kein-Element-Gruppe, berechnet mit der Software SCAN for Windows®

In Abb. 5.79 sind die Falschklassifizierungsrisiken für den Validierungsdatensatz in Abhängigkeit der Variablenart (Spektralbereich und HK 1 bis HK 3) gezeigt. Entgegen den Untersuchungen aus Abschnitt 5.3.5.2 wurde bei dem RDA-Modell des Validierungsdatensatzes unter Verwendung des Compton- und Rayleigh-Streubereichs als Variablen ein geringeres Falschklassifizierungsrisiko als bei dem Einsatz der signifikanten HK als Variablen berechnet.

Die graphische Darstellung der RDA-Ergebnisse des Validierungsdatensatzes für den Spektralbereich und den HK als Variablen sind in Abb. 5.80 und Abb. 5.81 dargestellt. Zwei Aussagen werden aus dem Vergleich der beiden Abbildungen deutlich. Das geringere Falschklassifizierungsrisiko für den Spektralbereich als Variablen (s. Abb. 5.79) zeigt sich in der graphischen Darstellung des RDA-Ergebnisses. Mit dem Compton- und Rayleigh-Streubereich als Variablen lassen sich 17 Substanzen eindeutig mit dem RDA-Modell klassifizieren (s. Abb. 5.80). Bei der Verwendung der signifikanten HK als Variablen sind 15 Substanzen eindeutig mit dem RDA-Modell klassifizierbar (s. Abb. 5.81). Die Ergebnisse aus dem RDA-Modell sind reproduzierbar.

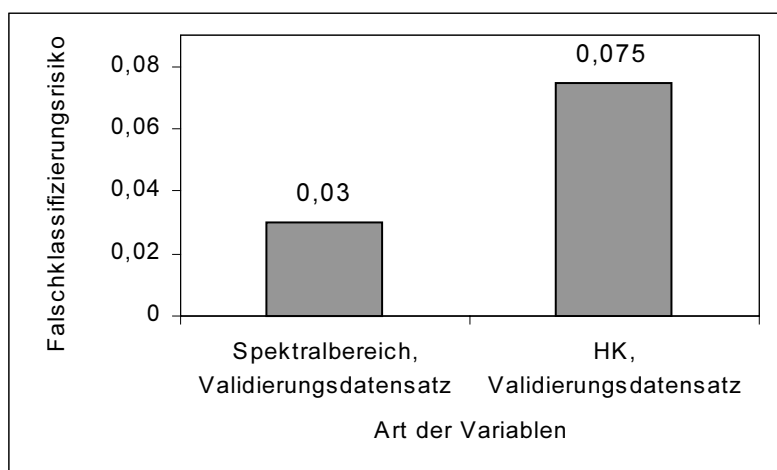


Abb. 5.79 Auftragung des Falschklassifizierungsrisikos (berechnet aus dem Testdatensatz) in Abhängigkeit der Variablenart (HK oder Spektralbereich) für den Validierungsdatensatz, berechnet mit der Software SCAN for Windows®

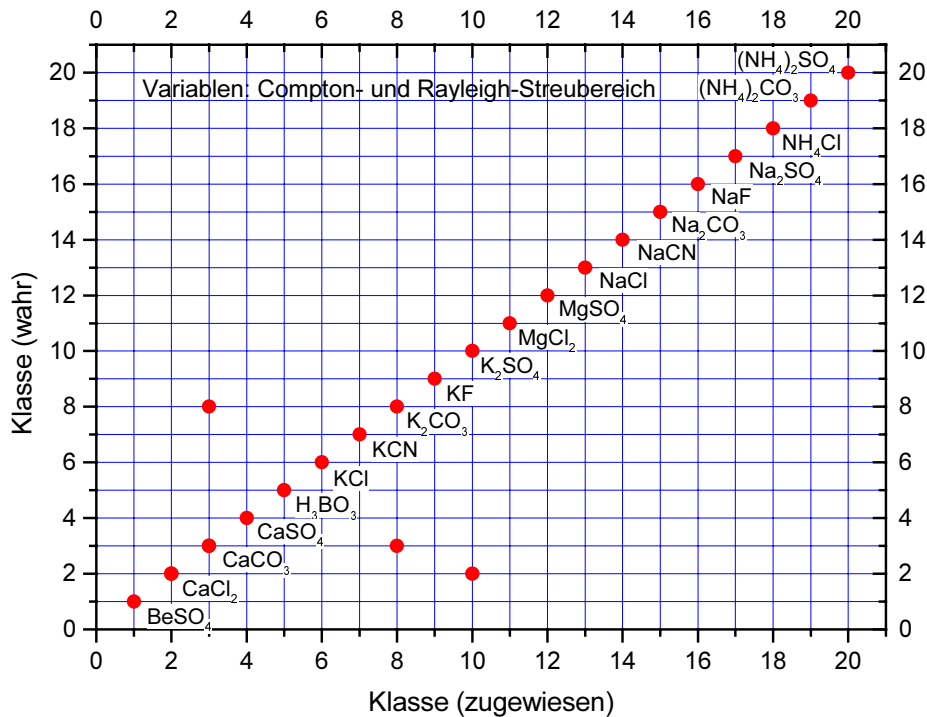


Abb. 5.80 Graphische Darstellung der RDA-Ergebnisse mit dem Compton- und Rayleigh-Streubereich als Variablen von 20 Substanzen (20 Klassen) des Validierungsdatensatzes, berechnet mit der Software SCAN for Windows[®], Meßzeit: 30 s,

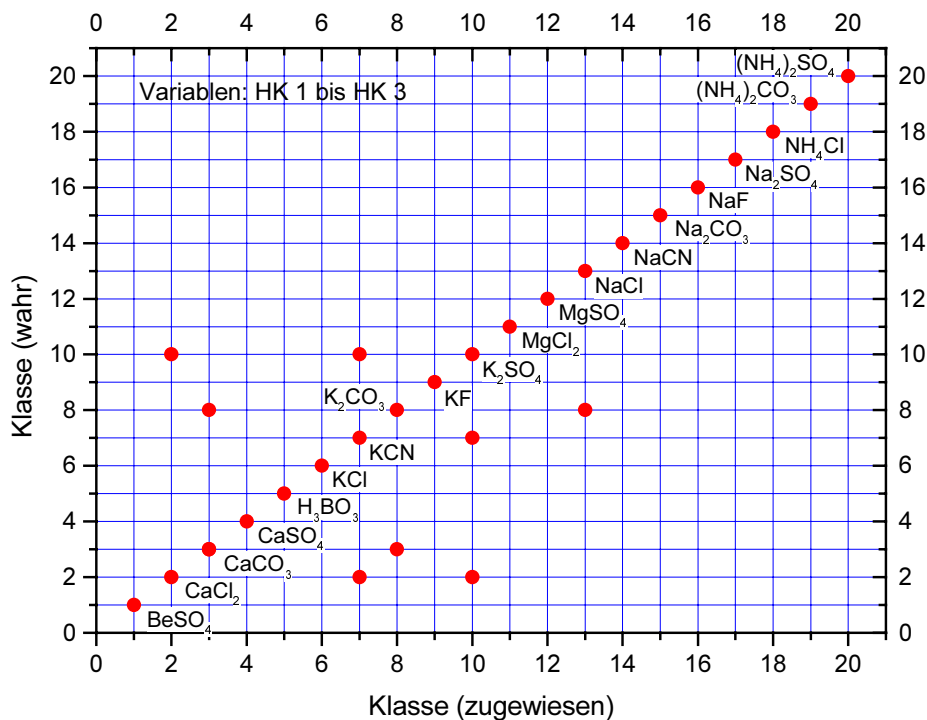


Abb. 5.81 Graphische Darstellung der RDA-Ergebnisse mit den HK 1 bis HK 3 als Variablen von 20 Substanzen (20 Klassen) des Validierungsdatensatzes, berechnet mit der Software SCAN for Windows[®], Meßzeit: 30 s

5.3.5.4 Fazit aus den RDA-Berechnungen

Durch die Verpackung aufgenommene EDRFA-Spektren sind mit Hilfe der RDA voneinander zu unterscheiden, wenn die Spektren für das RDA-Modell erkennbare spektrale Unterschiede aufweisen. Für die Ein-Element-Gruppen Chrom und Eisen ist eine Zuordnung der Substanzen zu einer zuvor definierten Klasse sowohl mit dem Spektralbereich (Fluoreszenzlinien-, Compton- und Rayleigh-Streubereich) als auch mit den signifikanten HK aus der PCA als Variablen möglich. Für die Eisen-Gruppe ist die Klassifizierung mit dem Spektralbereich besser (berechnet durch eine interne Kreuzvalidierung des Eisen-Datensatzes).

Die RDA-Klassifizierung der Substanzen der Kein-Element-Gruppe zu ihren Klassen ist am Beispiel der untersuchten Substanzen möglich. Trotz des geringeren Falschklassifizierungsrisikos durch Verwendung der signifikanten HK als Variablen ist die Anzahl der eindeutig klassifizierten Substanzen für beide Arten von Variablen (Spektralbereich und HK) gleich.

Die Validierung des RDA-Modells für den Lerndatensatz der Kein-Element-Gruppe mit drei voneinander unabhängigen Testdaten zeigt, daß die Klassifizierung mit dem RDA-Modell reproduzierbar ist. Für die RDA-Validierungsberechnungen hat sich der Einsatz des Spektralbereichs als Variablen gegenüber den HK als besser erwiesen. Dies wird durch das geringere Falschklassifizierungsrisiko und die höhere Anzahl an eindeutig klassifizierten Substanzen gezeigt.

5.4 Einfluß der Meßzeit auf die multivariate statistische Auswertung

In Tabelle 5.11 sind sieben Substanzen aufgeführt, mit deren Compton- und Rayleigh-Streubereichen unter Variation der Meßzeit die Varianz und die Auftrennung der Substanzen innerhalb des durch die ersten zwei HK aufgespannten neuen Datenraumes untersucht wurde. Ein Anhaltspunkt ist das Ergebnis der Meßzeitoptimierung aus Abschnitt 4.5. Bei Meßzeiten kleiner 20 s ist eine eindeutige Peakidentifizierung nicht mehr möglich.

In Abb. 5.82 sind die durch die ersten drei HK beschriebenen Varianzen in Abhängigkeit von der Meßzeit in 10 s-Schritten von 10 bis 100 s für den Datensatz der sieben Verbindungen dargestellt. Die durch die ersten drei HK beschriebene Varianz nahm aufgrund der abnehmenden relativen statistischen Schwankung bei größer werdender Meßzeit zu. Die HK bei Meßzeiten kleiner 20 s beschreiben eine sehr geringe Varianz.

Tabelle 5.11 Für die Untersuchungen zur Meßzeit verwendete Substanzen und ihre physikalischen Daten

Substanz	M /g/mol	\bar{OZ}	Dichte /g/cm ³
(NH ₄) ₂ CO ₃	96,09	6,869	1,6
H ₃ BO ₃	61,83	7,133	1,435
Na ₂ CO ₃	105,99	9,075	2,533
Na ₂ SO ₄	142,04	10,777	2,698
KCN	65,12	14,020	1,56
KCl	74,56	18,047	1,984
CaCl ₂ x2H ₂ O	147,02	15,420	1,85

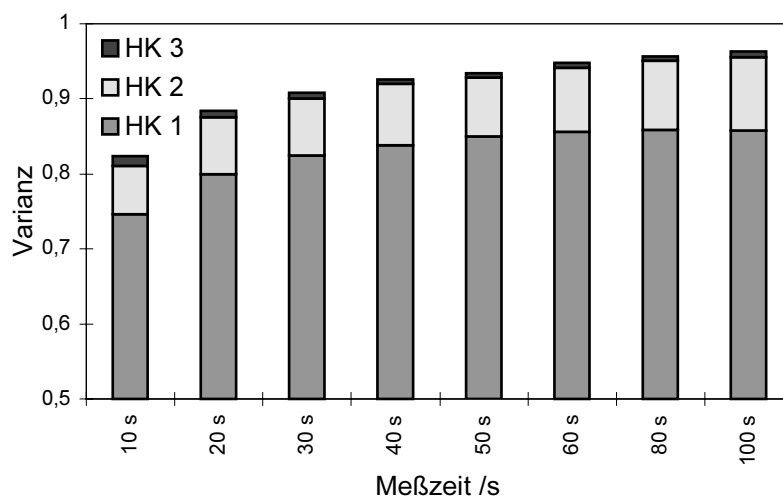


Abb. 5.82 Darstellung der durch die ersten drei HK beschriebene Varianz für die sieben Substanzen aus Tabelle 5.11 in Abhängigkeit von der Meßzeit, berechnet mit der Software SIMCA-P[®]

Die Gruppen in dem neuen Datenraum lassen sich durch Ellipsen beschreiben. Aufgrund der linearen Unabhängigkeit der einzelnen HK zueinander^[184,185,186] liegen die Ellipsenachsen parallel zu den HK-Achsen. Dies gilt streng nur für die Ellipse, die alle Gruppen umschließt. Eine Anwendung dieser Methode auf einzelne Gruppen ist dann erlaubt, wenn die Substanzen in dem neuen Datenraum keine Korrelationen zwischen den HK zeigen. Für die Substanzen KCN und CaCl₂x2H₂O sind die Verteilungen in dem durch die ersten zwei HK aufgespannten neuen Datenraum in Abhängigkeit von den Meßzeiten 10 s (Abb. 5.83), 20 s (Abb. 5.84) und 30 s (Abb. 5.85) in Form von Ellipsenflächen aufgetragen, nachdem die Korrelation der Objekte für die erste und zweite HK bestimmt wurde. Alle Bestimmtheitsmaße (R^2) liegen unterhalb 0,8 und zeigen somit keinen signifikanten Zusammenhang der Lage der Substanzen im Datenraum der HK 1 und HK 2. Der Ellipsenmittelpunkt wurde aus den Mittelwerten der ersten beiden HK-Werte für die entsprechende Gruppe berechnet. Die zweifache Standardabweichung der Werte für die Mittelpunkte wird zur Konstruktion der Ellipsenachsen a und b nach Gl.

5.6^[132] verwendet. Die beiden Ellipsenachsen decken 95% der zur Gruppe gehörenden Substanzen ab.

Gl. 5.6 $F_{\text{Ellipse}} = \pi a b$

10 s	Die Überlagerung der Ellipsenflächen (beinhalten 95% aller Daten) von $\text{CaCl}_2 \cdot x\text{H}_2\text{O}$ und KCN ist im Vergleich zur Meßzeit von 30 s hoch (s. Abb. 5.83).
20 s	Die Ellipsenflächen (beinhalten 95% aller Daten) von $\text{CaCl}_2 \cdot x\text{H}_2\text{O}$ und KCN überlagern sich nicht (s. Abb. 5.84).
30 s	Die Überlagerung der Ellipsenflächen (beinhalten 95% aller Daten) von $\text{CaCl}_2 \cdot x\text{H}_2\text{O}$ und KCN ist im Vergleich zur Meßzeit von 10 s niedrig (s. Abb. 5.85).

Das Ergebnis aus Abschnitt 4.5, daß Meßzeiten kleiner 20 s für eine abfallspezifische Zuordnung mit der EDRFA nicht geeignet sind, wurde hiermit bestätigt.

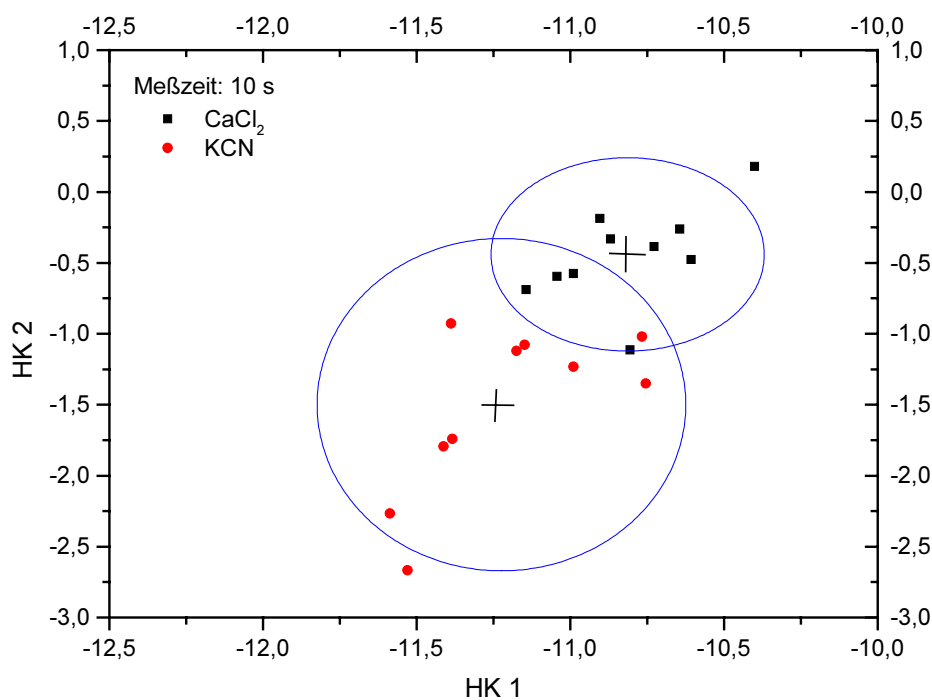


Abb. 5.83 Lage der Substanzen KCN und $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ aus Tabelle 5.11 in dem durch die ersten zwei HK aufgespannten neuen Datenraum (Score-Plot) bei einer Meßzeit von 10 s, berechnet mit Software SCAN for Windows®

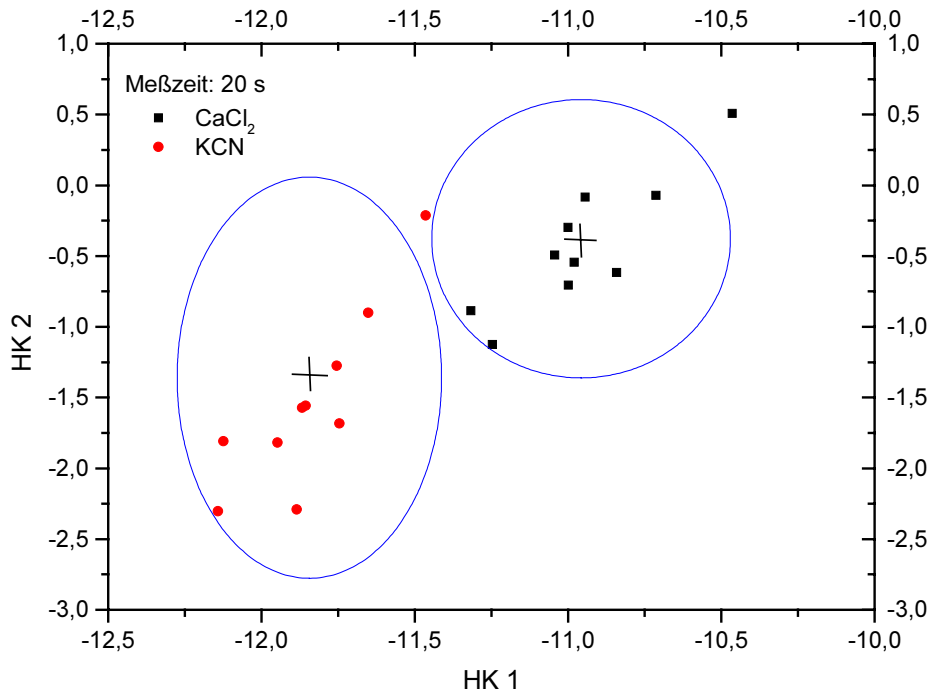


Abb. 5.84 Lage der Substanzen KCN und KCN und $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ aus Tabelle 5.11 in dem durch die ersten zwei HK aufgespannten neuen Datenraum (Score-Plot) bei einer Meßzeit von 20 s, berechnet mit Software SCAN for Windows®

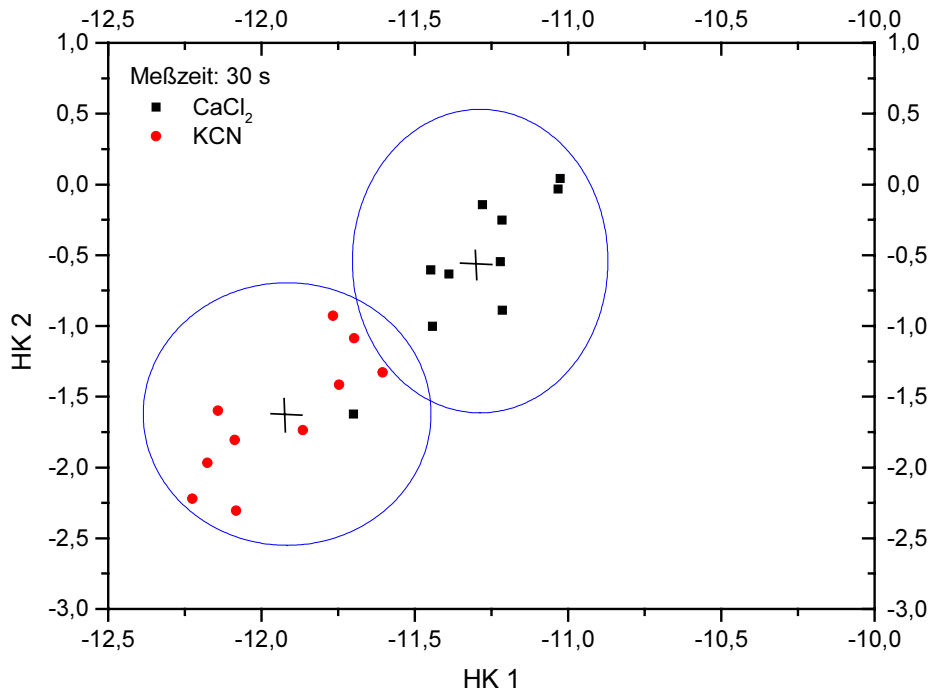


Abb. 5.85 Lage der Substanzen KCN und KCN und $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$ aus Tabelle 5.11 in dem durch die ersten zwei HK aufgespannten neuen Datenraum (Score-Plot) bei einer Meßzeit von 30 s, berechnet mit Software SCAN for Windows®