

3 Bibliothekserstellung

In dieser Arbeit werden Bibliotheken erstellt und getestet, welche der Identitätskontrolle dienen. Das bedeutet, daß jede zu identifizierende Substanz in der Bibliothek enthalten sein muß. Strukturen werden nicht abgespeichert und eine Strukturvorhersage wird im Rahmen der vorliegenden Aufgabenstellung auch nicht gefordert. Vielmehr kommt es darauf an, jede Substanz eindeutig zu identifizieren, wobei zum Teil sehr ähnliche Substanzen in der Bibliothek enthalten sind. Die Unterschiede zwischen Spektren unterschiedlicher Chargen einer Substanz können dabei unter Umständen die gleiche Größenordnung erreichen wie die Unterschiede zwischen Spektren ähnlicher Substanzen. Daher ist es notwendig, daß mehrere Chargen ein und desselben Artikels vermessen werden, um einen möglichst großen Anteil der später zu erwartenden Probenvarianz abzudecken.

Die Vielzahl der so erhaltenen Einzelspektren einer bestimmten Substanz bilden im Datenraum einen Sphäroid, einen sogenannten Cluster (vergleiche Abb. 11, S. 25). Die Grenzen eines solchen Clusters können je nach angewandtem statistischen Kriterium unterschiedlich festgelegt werden. Je größer dieser sogenannte Vertrauensbereich gewählt wird, desto größer ist auch die Wahrscheinlichkeit, daß ein später gemessenes Spektrum der betroffenen Substanz in diesem liegt. Gleichzeitig wächst allerdings die Gefahr, daß es zu Überlappungen mit Clustern anderer Substanzen kommt. Im letzteren Fall ist eine eindeutige Identifizierung nicht mehr möglich.

Zum Aufbau einer Testbibliothek wurden von der Firma Merck ca. 4000 NIR- und Raman-Spektren von ca. 600 Artikeln zur Verfügung gestellt. Von den meisten der ausgewählten Artikeln standen bis zu zehn und mehr Chargen zur Verfügung.

3.1 Aufbau einer Testbibliothek

Schwerpunkt dieser Arbeit ist der Vergleich verschiedener Methoden zur Erstellung und Validierung einer NIR- und einer Raman-Bibliothek. Da auch die Leistungsfähigkeit der beiden Techniken IR und Raman miteinander verglichen werden soll, wird eine Testbibliothek aufgebaut, welche mehrere Forderungen erfüllt:

- Sowohl das NIR- als auch das Raman-Spektrum muß auswertbar sein (z. B. keine Fluoreszenz).
- Es müssen mindestens sechs Spektren von Proben aus verschiedenen Chargen vorhanden sein (10 bis 20 verschiedene Chargen wären wünschenswert).
- Von einer Substanz darf kein Parallelartikel enthalten sein (Substanzen mit identischer Qualität, aber mit verschiedenen Artikelnummern).

Bei dem zur Verfügung gestellten Datensatz handelt es sich um Spektren von Artikeln, welche nach ihrer Verfügbarkeit ausgewählt wurden. Die Artikel können über eine sechsstellige Artikelnummer identifiziert werden. In dem Sortiment enthalten sind verschiedene Artikel einer Substanzen, die unterschiedliche chemische oder physikalische Eigenschaften besitzen und jeweils unter einer eigenen Artikelnummern geführt werden. Hiervon zu unterscheiden sind Parallelartikel, dies sind Artikel mit unterschiedlichen Artikelnummern aber gleicher chemischer Zusammensetzung. Parallelartikel können sich zum Beispiel im Analysenumfang unterscheiden, aber gleichen Ursprungs sein.

Die Spektren liegen alle im OPUS-Format vor, wobei sich der Dateiname aus der Artikelnummer ergibt. Da der Dateiname bei der OS/2-Version der OPUS-Software auf die 8.3 Namenskonvention beschränkt ist, ist der Name aller Einzelspektren einer Substanz gleich. Zur Unterscheidung dient hier die Extension, welche durchnummeriert wird. Aufgrund der großen Anzahl aller Einzelspektren, werden alle Spektren einer Substanz in einem eigenen Verzeichnis abgespeichert. Die NIR- und die Raman-Spektren unterscheiden sich nur in ihrem Dateipfad.

Aus einem Datensatz von ca. 600 Artikeln wurden 360 Substanzen ausgewählt, von denen sowohl geeignete NIR- wie auch Raman-Spektren vorlagen. Hierbei waren jedoch noch Parallelartikel enthalten. Die Spektren dieser Substanzen wurden in die jeweils niedrigere Artikelnummer umbenannt und in *einem* Verzeichnis zusammengefaßt, um für die einzelnen Substanzen eine möglichst große Zahl von Einzelspektren zur Verfügung zu haben. Nicht zusammengefaßt wurden Artikel, die sich physikalisch unterscheiden (z. B. durch unterschiedliche Korngröße).

Die Testbibliothek soll noch nicht völlig robust sein, sondern auch einige Problemsubstanzen enthalten, damit sich Unterschiede zwischen den verschiedenen Identifizierungsmethoden besser herauskristallisieren.

3.2 Thresholdberechnung

Bei der Distanzberechnung (Abschnitt 2.5) muß Bezug auf ein bestimmtes Spektrum genommen werden. Von jedem Artikel liegen jedoch mehrere Einzelspektren vor. Unter Annahme einer Normalverteilung wird von diesen n Einzelspektren ein Mittelwertspektrum berechnet. Die Einzelspektren weisen eine Distanz $D_{i,M}$ zum Mittelwertspektrum auf und streuen um dieses mit folgender Standardabweichung s ^[53]:

$$s = \sqrt{\frac{\sum_{i=1}^n (D_{i,M})^2}{n-1}} \quad \text{Gleichung 22}$$

3.2.1 Konstantes Konfidenzniveau („Constant Confidence Level“)

Aus der Größe dieser Standardabweichung ergibt sich bei der Validierung eine Grenze C_Y , innerhalb der ein zu prüfendes Spektrum liegen muß, damit es als „richtig“ identifiziert wird.

$$C_Y = D_{Y,M} + f * s_Y \quad \text{Gleichung 23}$$

wobei: $D_{Y,M}$ = mittlere Distanz

Für $f=1$ liegen 84,4 % der Einzelspektren innerhalb des Konfidenzbereiches. Für $f=2$ erhöht sich dieser Wahrscheinlichkeitswert auf 97,7 %^[51]. Ein gewisser Anteil der Einzelspektren liegt daher bei dieser Art der Grenzwertsetzung (Thresholding) zwangsläufig außerhalb des Konfidenzbereiches und liefert ein falsch positives Ergebnis.

3.2.2 Fixed Algorithm

Aufgrund der relativ geringen Anzahl der Chargen ist die Unsicherheit der Standardabweichung s relativ groß. Daher ist der „Fixed Algorithm“ zu bevorzugen, denn hierbei geht man von der Distanz des Spektrums mit der größten Abweichung $D_{Y,M}^{max}$ vom Mittelwertspektrum aus und addiert als Sicherheitszuschlag ein Viertel der Standardabweichung:

$$C_Y = D_{Y,M}^{max} + \frac{1}{4} * s_{0,Y} \quad \text{Gleichung 24}$$

Es liegen also auf jeden Fall alle für die Berechnung der Standardabweichung herangezogenen Einzelspektren der Substanz innerhalb der Grenzen, d. h. sie werden bei der Validierung richtig „negativ“ diagnostiziert. Durch den gegenüber dem konstanten Konfidenz-

veau größeren Vertrauensbereich wird die Wahrscheinlichkeit einer falsch positiven Identifizierung gesenkt, gleichzeitig ist die Gefahr einer falsch negativen Identifizierung größer. Die in der vorliegenden Arbeit vorgestellten Validierungsergebnisse beziehen sich, soweit nicht anders angegeben, auf den Fixed Algorithm.

3.3 Vertrauensbereich

Aufgabe einer Bibliothekskalibration ist es, den Vertrauensbereich für die einzelnen Substanzen angemessen festzulegen. In der Regel sollte er so reichlich bemessen sein, daß alle Einzelspektren des Kalibrationsdatensatzes enthalten sind und kleinere zusätzliche Abweichungen abgefangen werden. Hierbei sollte der Kalibrationsdatensatz so umfangreich sein, daß er die zu erwartenden Probenvarianz in realistischem Umfang repräsentiert. Andererseits muß der Abstand zu benachbarten Clustern berücksichtigt werden, die Vertrauensbereiche verschiedener Artikel dürfen sich also nicht überschneiden. Getestet wird dies mit dem im folgenden beschriebenen Kugel- bzw. Doppelkugelmodell.

3.3.1 Einfaches Kugelmodell („Single Sphere Method“)

In der OPUS-Software werden zwei unterschiedlich komplexe Kugelmodelle verwendet, ein Einfachkugelmodell („Single Sphere Method“) und ein Doppelkugelmodell („Double Sphere Method“) ^[51]. Abbildung 15 zeigt beispielhaft die Anwendung des Einfachkugelmodells. X_{mittel} und Y_{mittel} sind die Mittelwertspektren zweier unterschiedlicher Substanzen. Y sei die zu validierende Substanz. C_y ist der Radius der Kugel um Y_{mittel} , die den Konfidenzbereich darstellt. Dieser ist nicht von der Wellenzahl abhängig. Ein Spektrum y gilt dann sowohl bei der Validierung als auch bei der späteren Analyse als eindeutig identifiziert, wenn folgende Bedingungen erfüllt sind:

- 1) Ein benachbartes Mittelwertspektrum X_{mittel} darf nicht näher zu Y_{mittel} als eine dem doppelten Konfidenzbereich ($2C_y$) entsprechende Distanz liegen ($D_{xy} > C_y$).
- 2) Neben der Validierung mit Mittelwertspektren ist auch eine Validierung mit Einzelspektren möglich.
 - a) Ein Einzelspektrum y muß hierbei innerhalb des Konfidenzbereiches C_y des zugehörigen Mittelwertspektrums Y_{mittel} liegen.
 - b) Innerhalb des Konfidenzbereiches des Einzelspektrums y darf kein benachbartes Spektrum X_{mittel} liegen.

Ein Beispiel der Validierung einer Bibliothek mittels Mittelwertspektren („Check Average Spectra“) zeigt Abbildung 15 (links). In diesem Fall liegt kein benachbartes Spektrum innerhalb des doppelten Konfidenzbereiches des getesteten Spektrums Y_{mittel} , es liegt keine Überlappung vor.

Um sich ein besseres Bild über die Lage eines Mittelwertspektrums Y_{mittel} zu einem Mittelwertspektrum X_{mittel} machen zu können, ist es sinnvoll, eine relative Distanz $\delta_{Y,X}$ zu berechnen:

$$\delta_{Y,X} = \frac{D_{X,Y}}{2 * C_Y} = \frac{D_{X,Y}}{2 * (D_{Y,max} + \frac{1}{4} s_{0,Y})} \quad \text{Gleichung 25}$$

Es wird das Verhältnis gebildet aus der Distanz $D_{X,Y}$ zwischen dem zu validierenden Mittelwertspektrum Y_{mittel} und einem benachbarten Mittelwertspektrum X_{mittel} und der Größe des Vertrauensbereiches von Y . Aus $\delta_{Y,X}$ kann man den Grad der Überlappung abschätzen. Ist die relative Distanz von Y_{mittel} zu X_{mittel} kleiner als 1, so liegt eine Überlappung vor. Man beachte, daß bei dieser Art der Validierung immer nur der Vertrauensbereich eines der beiden Mittelwertspektren berücksichtigt wird, d. h. $\delta_{Y,X} \neq \delta_{X,Y}$.

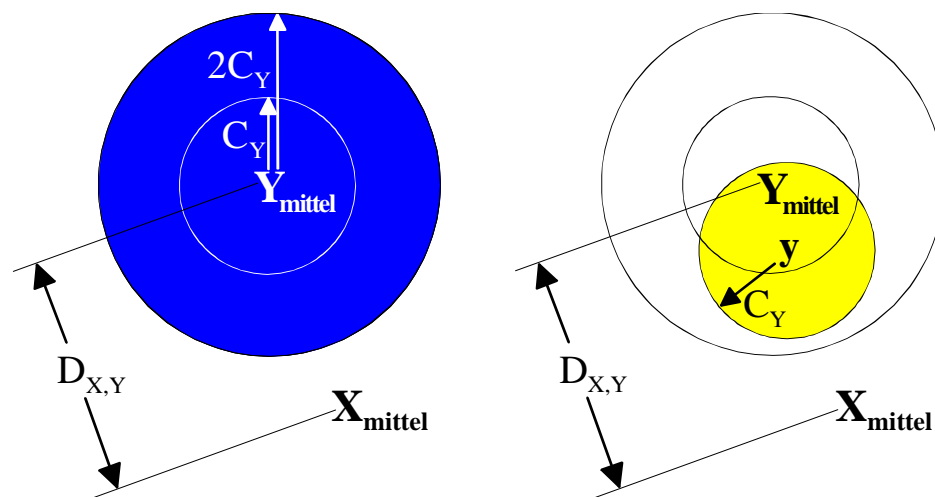


Abb. 15: Einfaches Kugelmodell für die Validierung mit Mittelwertspektren („Check Average Spectra“, links) bzw. Einzelspektren („Check Original Spectra“, rechts).

Darüber hinaus kann die Validierung mit dem Einfachkugelmodell gemäß Punkt 2 in Form der Validierung der Einzelspektren durchgeführt werden („Check Original Spectra“). Hierbei wird geprüft, ob ein benachbartes Mittelwertspektrum in den Konfidenzbereich des getesteten Einzelspektrums* y fällt (Abb. 15, rechts). Das Einzelspektrum selbst muß innerhalb des einfachen Konfidenzbereiches um Y_{mittel} liegen.

Das Beispiel zeigt den Fall, in dem die Distanz des Spektrums y eindeutig kleiner als der Konfidenzbereich des Mittelwertspektrums Y_{mittel} ist. Liegt ein anderes Mittelwertspektrum innerhalb des Konfidenzbereiches um y , so tritt eine Überlappung auf. Es besteht die Gefahr einer Verwechslung. Auch hier muß die Distanz eines benachbarten Mittelwert-

* Für die Einzelspektren einer Substanz wird die Standardabweichung des zugehörigen Mittelwertspektrums angenommen.

3.4 Validierung

Bei der Validierung einer Bibliothek wird deren Konsistenz im Hinblick auf eine eindeutige Identifizierung der enthaltenen Artikel getestet, d. h. ob das Spektrum eines Artikels richtig dem betreffenden Artikel zugeordnet wird und sich eindeutig von allen anderen Artikeln unterscheidet. Ein zu validierendes Spektrum wird hierbei mit der gesamten Bibliothek, also allen Referenzspektren, verglichen. Spektren, die in dem Vertrauensbereich eines Artikels liegen, werden diesem zugeordnet. Dies bedeutet, daß die Distanz zwischen dem getesteten Spektrum und dem Mittelwertspektrum des Artikels kleiner ist als der Grenzwert (Threshold), der für diese Substanz festgelegt wurde. Man erhält eine Hit-Liste, in der alle Bibliothekssubstanzen nach abnehmender spektraler Ähnlichkeit (nach zunehmender Distanz zu dem zu validierenden Spektrum) sortiert werden. Bei einer optimierten Bibliothek sollte an erster Stelle der Hit-Liste die Substanz stehen, von der das zu validierende Spektrum stammt. Ferner sollte die Distanz kleiner als der Threshold sein.

Die Validierung, die vorwiegend nach dem einfachen Kugelmodell durchgeführt wurde, kann mit unterschiedlichen Verfahren erfolgen. Am einfachsten ist es, die Mittelwertspektren gegeneinander zu prüfen („Check Average Spectra“). Fällt ein Mittelwertspektrum in den Vertrauensbereich eines anderen, so wird in einem Validierungsreport eine Meldung ausgegeben, wobei neben der Distanz der beiden Mittelwertspektren auch die Größe der beiden Konfidenzbereiche angegeben wird. Dieser Vergleich der Mittelwertspektren wurde in jedem Fall durchgeführt und gibt einen ersten Eindruck über die Güte einer Bibliothek.

Wesentlich rechenintensiver ist die Methode „Check Original Spectra“, bei der sämtliche Einzelspektren gegen die Mittelwertspektren geprüft werden. Bei Überlappungen kann geprüft werden, ob alle oder lediglich einzelne Spektren einer Substanz nicht eindeutig identifiziert werden können. Solche Spektren sollten jedoch nicht unbegründet aus der Bibliothek entfernt werden, es sei denn, es handelt sich eindeutig um Ausreißer. Denkbar ist nämlich auch, daß einzelne Chargen einer Substanz, abhängig von den Herstellungsparametern, größere Distanzen zum Mittelwert aufweisen. Bei nachfolgenden Chargen würden sich diese Probleme dann wiederholen. Kann durch Optimierung der Bibliothek keine Verbesserung erzielt werden, ist zu prüfen, ob Substanzen mit überlappenden Vertrauensbereichen zu einer *Klasse* zusammengefaßt werden können.

Wird ein konstantes Konfidenzniveau für die Berechnung benutzt, so liegt, abhängig von der gewählten Wahrscheinlichkeit, ein bestimmter Prozentsatz der Einzelspektren außerhalb des jeweiligen Clusters. In dem Validierungsreport werden jeweils die Spektren mit den größten Abweichungen vom Mittelwert und die Größe der Konfidenzbereiche angege-

ben. Auch wenn keine Überlappung mit anderen Clustern vorliegt, erhält man einen Hinweis auf Ausreißer, wenn die Distanz eines Einzelspektrums vom zugehörigen Mittelwertspektrum übermäßig groß ist. Ob tatsächlich ein Ausreißer vorliegt, ist individuell für jedes einzelne Spektrum zu prüfen, da ein Spektrum auch außerhalb des Vertrauensbereiches liegen kann, wenn die Varianz des restlichen Datensatzes zu klein ist.

Zu beachten ist, daß auch bei der Validierung mit Einzelspektren diese bereits zur Berechnung des Mittelwertspektrums herangezogen worden sind. Zusätzlich sollte die Bibliothek daher mit unabhängigen Spektren geprüft werden, welche nicht in die ursprüngliche Berechnung der Mittelwertspektren einbezogen wurden. Eine solche Validierung mit unabhängigen Spektren wird als externe Validierung, die oben beschriebenen Validierungen mit Mittelwert- oder Einzelspektren als interne Validierung bezeichnet. In der Praxis ist es oft zu aufwendig, neben dem Kalibrationsdatensatz für die Bibliothekserstellung einen zweiten ebenso umfangreichen Validationsdatensatz zu generieren. Realistischer ist eine Validation im laufenden Betrieb, denn liegt dann eine Charge außerhalb des Konfidenzbereiches, so erhält man eine Fehlermeldung und kann prüfen, ob die Substanz tatsächlich nicht identisch ist, oder ob der Datensatz erweitert werden muß. Kritisch bleibt der Fall einer falsch negativen Identifizierung durch Verwechslung sehr ähnlicher Spektren. Im Zweifelsfall wird man die ähnlichen Artikel zu einer Klasse zusammenfassen und anschließend in einer zweiten Identifizierungsstufe eine differenziertere Diskriminierung vornehmen.

3.5 Graphische Darstellung der Validierungsergebnisse

3.5.1 Matrixdarstellung

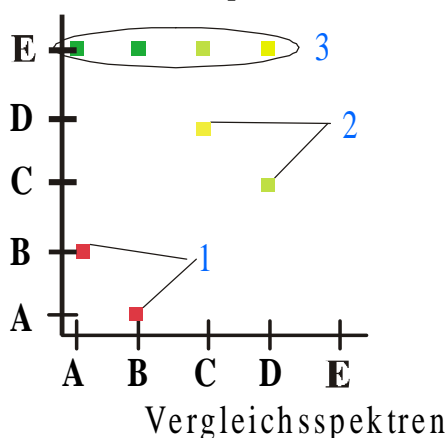
Aufgrund der großen Datenmenge ergibt sich die Problematik, daß im Rahmen der verfügbaren Anwendungssoftware (OPUS) die Ergebnisse einer Bibliotheksvalidierung in umfangreichen Protokolldateien niedergelegt werden. Deren Auswertung ist mühsam und äußerst zeitraubend. Selbst bei der relativ kleinen Testbibliothek können die Reporte mehrere Dutzend Seiten umfassen. Diese Details sind für den Aufbau und die Erweiterung einer Bibliothek zwar notwendig, ein Vergleich verschiedener Methoden ist aber kaum möglich. Es wurde daher eine eigene Schnittstelle zu den Validierungsprotokollen und darauf aufbauend ein Programm für eine übersichtliche graphische Darstellung der Validierungsergebnisse entwickelt, die vor allem einen schnellen Vergleich unterschiedlicher Kalibrationsmodelle ermöglichen soll.

Bei diesem Verfahren erscheint das Validierungsergebnis als Punktmatrix in einem zweidimensionalen Koordinatensystem. Jeder Punkt repräsentiert das Ergebnis des Vergleiches des zu validierenden Spektrums mit einem Referenzspektrum. Die Namen der zu testenden Substanzen (Einzel- bzw. Mittelwertspektren) werden auf der Ordinate, die Namen der Referenzsubstanzen auf der Abszisse aufgetragen. Die Reihenfolge der Notierung der Substanznamen auf den Achsen erfolgt sortiert nach aufsteigenden Artikelnummern.

Möglich wäre auch eine Reihenfolge, welche auf chemischer Ähnlichkeit beruht, damit Clusterbildung besser zu beobachten ist und ähnliche Substanzen leichter zu Klassen zusammengefaßt werden können. Dies wurde allerdings in der Software nicht realisiert.

Der Grad der Eindeutigkeit der Diskriminierung zwischen den zu validierenden Spektren und den unterschiedlichen Referenzspektren wird durch eine farbige Abstufung für die

Zu validierende Spektren



Punkte innerhalb des Koordinatensystems visualisiert. Spektren, die eindeutig gegen das Referenzspektrum diskriminiert werden können, erscheinen als „weiße Punkte“, sind also nicht sichtbar. Als farbige Punkte erscheinen nur solche Substanzen bzw. Spektren, die den Distanzschwellenwert (Threshold) unterschreiten. Hierbei erfolgt eine Abstufung von relativ geringer (grün), mäßiger (grün-gelb) bis starker (orange-rot) Übereinstimmung.

Abb. 17: Punktmatrix von Validierungsergebnissen.

Abbildung 17 zeigt dies schematisch am Beispiel einer Bibliothek mit 5 Substanzen. Die Substanzen A und B sind einander sehr ähnlich

(Fall 1, z. B. gleiche chemische Zusammensetzung mit unterschiedlichen Korngrößen oder unterschiedlicher Reinheit). Die Spektren sind sehr ähnlich und können daher nicht unterschieden werden.

Die Substanzen C und D überlappen mäßig (Fall 2), die relative Distanz ist deutlich größer als bei A und B. Je nach Datenvorbehandlung und Methode kann es möglich sein, solche Substanzen noch zu unterscheiden. Die Substanz E zeigt Überlappungen mit allen anderen Referenzsubstanzen (Fall 3). Da letztere in der Regel deutlich unterschiedliche Spektren besitzen, muß bei der Substanz E die Größe des Konfidenzbereiches überprüft werden.

Die Punktmatrix bei dem in dieser Arbeit angewandten einfachen Kugelmodell ist nicht symmetrisch, denn in die Thresholdberechnung geht nur der Vertrauensbereich des zu validierenden Spektrums ein, der Vertrauensbereich des Referenzspektrums bleibt unberücksichtigt ($\delta_{Y,X} \neq \delta_{X,Y}$, s. S. 40).

Wie bereits beschrieben, wird der Übereinstimmungsgrad zwischen den Spektren durch eine Farbskala visualisiert, die sich auf die relative spektrale Distanz $\delta_{Y,X}$ bezieht. Wird die Validierung mit Mittelwertspektren durchgeführt, so handelt es sich bei letzterem Parameter um das Verhältnis der Distanz zwischen zwei Mittelwertspektren zum Threshold des zu prüfenden Mittelwertspektrums (Gl. 25, s. S. 39).

Auf der Diagonalen der Punktmatrix (Winkelhalbierende der beiden Achsen) treten keine Punkte auf, denn dies würde einem Test eines Mittelwertspektrums gegen sich selbst entsprechen.

Wird eine Validierung mittels Einzelspektren durchgeführt, so wird im alphanumerischen Validierungsreport für jedes einzelne Spektrum eine Überlappung angegeben. Bei der graphischen Darstellung wird für jeden Punkt die relative spektrale Distanz über alle Einzelspektren der jeweiligen Substanz gemittelt.

Bei Verwendung des „Constant Confidence Level“ kann es vorkommen, daß Einzelspektren außerhalb des Konfidenzbereiches liegen. Die Farbskala wird in diesem Fall so erweitert, daß angezeigt wird, wieviel Prozent der Spektren außerhalb des Konfidenzbereiches liegen (schwarz > 0 %, über dunkelblau nach hellblau = 100 %). Diese Ergebnisse der Option „Constant Confidence Level“ wird exemplarisch an einem Beispiel gezeigt, vorwiegend wurde in dieser Arbeit mit dem „Fixed Algorithm“ gearbeitet. Mit dem so entwickelten graphischen Verfahren ist es also mit einem Blick möglich, die Güte einer Validierung zu überprüfen, wodurch sich die Auswirkungen der verschiedenen Datenvorbehandlungsmethoden auf die Validierungsergebnisse wesentlich schneller und effizienter beurteilen läßt als mit einem herkömmlichen Auswerteprotokoll.

3.5.2 Histogramme

Die Zahl der Punkte einer Zeile in der Punktmatrix entspricht der Anzahl der Überlappungen des geprüften Artikels. Die Häufigkeit, mit der eine bestimmte Anzahl von Überlappung auftritt, kann in einem Histogramm dargestellt werden. Auf der Abszisse ist die Anzahl der Überlappungen aufgetragen, die Häufigkeit, mit der diese Anzahl auftritt, wird auf der Ordinate aufgetragen. Üblicherweise erhält man die intensivsten Balken auf der linken Seite des Histogramms, d. h. bei den meisten Artikeln treten nur ein oder zwei Überlappungen auf. Begibt man sich auf der Abszisse weiter nach rechts, so werden die Balken in der Regel kleiner, da bei den meisten Artikeln nur wenige Überlappungen mit anderen Substanzen auftreten.

Abbildung 18 und 19 zeigen zwei Beispiele für solche Histogramme.

Bei der Methode SND19 (Standardmethode mit Vektornormierung und 1. Ableitung mit 9 Glättungspunkten als Datenvorbehandlung, Abb. 18) tritt bei 14 Artikeln nur 1 Überlappung, bei 7 Artikeln 2 Überlappungen usw. auf. Als Maximum treten bei 1 Artikel 15 Überlappungen auf.

Bei der Methode SND29 (Standardmethode mit Vektornormierung und 2. Ableitung mit 9 Glättungspunkten als Datenvorbehandlung, Abb. 19) treten insgesamt mehr Überlappungen (2769) auf. Auffallend ist ferner, daß einige Artikel mit einer großen Anzahl anderer Substanzen überlappt. Bei 7 Substanzen treten sogar 260 Überlappungen auf, diese Substanzen werden praktisch mit allen anderen Bibliotheksspektren verwechselt.

Bei dem zweiten Beispiel können 2/3 der Überlappungen auf 7 Artikel zurückgeführt werden! Dies fällt noch mehr in der Punktmatrixdarstellung auf, denn dort entspricht die Größe der farbigen Fläche der Gesamtzahl der Überlappungen.

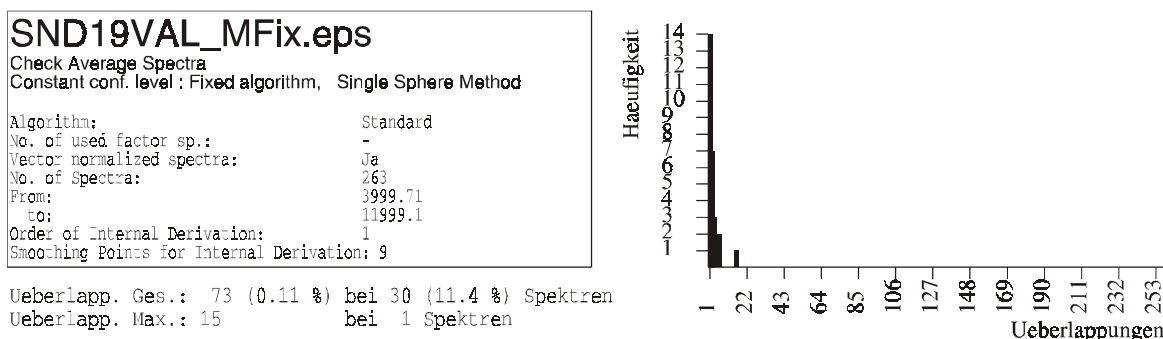
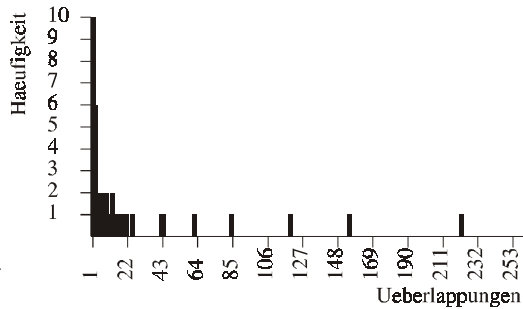


Abb. 18: Histogramm (NIR, Standard-Methode mit Vektornormierung und 1. Ableitung).

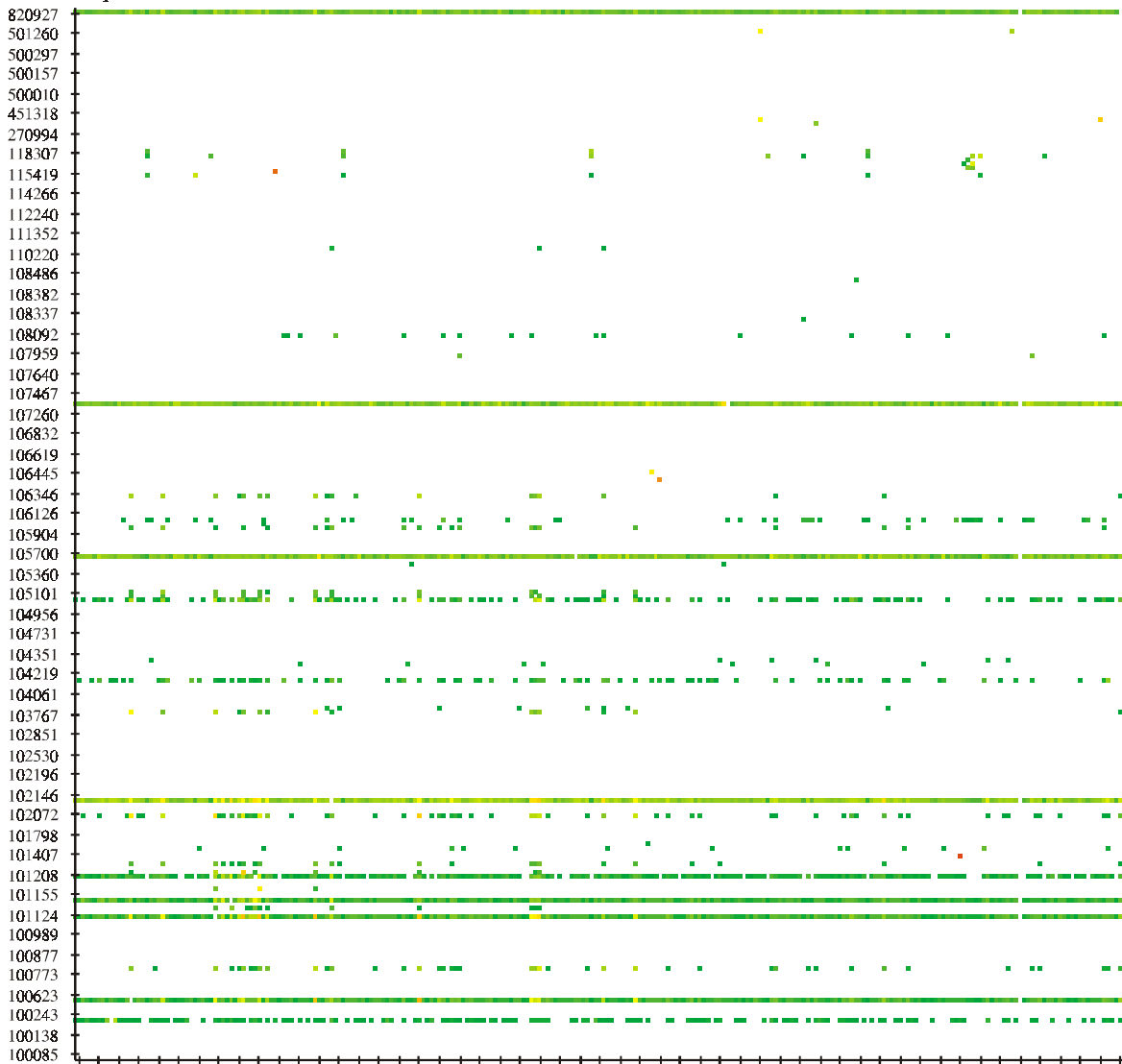
SND29V00_MFix.eps
 Check Average Spectra
 Constant conf. level : Fixed algorithm, Single Sphere Method
 Algorithm: Standard
 No. of used factor sp.: -
 Vector: normalized spectra: Ja
 No. of Spectra: 263
 Prom: 3999.71
 Lo: 11999.1
 Order of Internal Derivation: 2
 Smoothing Points for Internal Derivation: 9



Ueberlapp. Ges.: 2769 (4.02 %) bei 48 (18.2 %) Spektren
 Ueberlapp. Max.: 260 bei 7 Spektren

wahr positiv: 66137 falsch negativ: 2769 Sensitivitaet: 95.981 %

Getestete Spektren



Legende:
 Relative Spektrale Distanz



Abb. 19: Histogramm und Punktmatrix (NIR, Standard-Methode mit Vektornormierung und 2. Ableitung).

3.5.3 Güteparameter zur Beurteilung von Bibliotheken

Neben den in Abschnitt 3.5.1 beschriebenen graphischen Darstellungen, welche die Validierungsergebnisse detailliert wiedergeben ist im Hinblick auf eine automatisierte Optimierung auch eine Bewertung durch pauschale Maßzahlen denkbar.

Hierzu bieten sich die Sensitivität* und Spezifität als statistische Parameter zur Beurteilung von Bibliotheken an. Hierbei wird ein Vergleich zwischen einem Probenspektrum (bzw. einem zu validierenden Spektrum) und einem Referenzspektrum als „Test“ betrachtet, der „positiv“ ausfällt, wenn keine Übereinstimmung vorliegt und „negativ“, wenn Probe und Referenz übereinstimmen. Dann gelten folgende Definitionen^[56]:

$$\text{Sensitivität [\%]} = \frac{n_{TP}}{n_{TP} + n_{FN}} \cdot 100 = \frac{n_{TP}}{n_{\text{nicht o.k.}}} \cdot 100 \quad \text{Gleichung 26}$$

$$\text{Spezifität [\%]} = \frac{n_{TN}}{n_{TN} + n_{FP}} \cdot 100 = \frac{n_{TN}}{n_{\text{o.k.}}} \cdot 100 \quad \text{Gleichung 27}$$

wobei: n_{TP} = Zahl der richtig positiven Tests
 n_{FP} = Zahl der falsch positiven Tests
 n_{TN} = Zahl der richtig negativen Tests
 n_{FN} = Zahl der falsch negativen Tests

Angestrebt wird eine Sensitivität und Spezifität von 100 %. Alle farbigen Punkte in der Punktmatrixdarstellung eines Validierungslaufes, die außerhalb der Diagonalen liegen, kennzeichnen eine Überlappung, d. h. die entsprechenden Substanzen werden unter Umständen bei einer späteren Analyse verwechselt, was einem falsch negativen Testergebnis entspricht.

Bei der Validierung mittels Einzelspektren können auch auf der Diagonale Punkte auftreten. Dies ist der Fall, wenn ein getestetes Einzelspektrum außerhalb des Konfidenzbereiches der zugehörigen Substanz liegt. Ein Punkt auf der Diagonale kennzeichnet also eine falsch positive Identifizierung. Insgesamt sind die Validierungsergebnisse um so besser, je weniger farbige Punkte in der Punktmatrixdarstellung auftreten.

Die Sensitivität ist das Verhältnis der wahr positiven Testergebnisse zur Gesamtzahl der nicht übereinstimmenden Artikel. Wahr positive Testergebnisse entsprechen allen „weißen Punkten“ in der Punktmatrix (die Diagonale ausgenommen).

* Der Begriff Sensitivität (Empfindlichkeit) wird hierbei abweichend von seiner Bedeutung in der quantitativen Analyse benutzt, wo er die Steigung einer Kalibrationsgeraden beschreibt.

⁵⁶ K. Molt, D. Ihlbrock, Principles and applications of quality control by near infrared spectroscopy using the example of polymer additives, Fresenius J. Anal. Chem. (1994) 348, S. 523-529

Bildlich gesprochen ist also die Sensitivität der relative Anteil der weißen Fläche in der Matrixdarstellung.

Wie bereits beschrieben, können nur bei Verwendung des „Constant Confidence Levels“ falsch positive Testergebnisse auftreten. Daher ist nur in diesem Fall (nicht aber bei dem „Fixed Algorithm“) die Angabe der Spezifität sinnvoll. Bei Verwendung des „Constant Confidence Levels“ ist die Spezifität (Gl. 27) das Verhältnis der „weißen Punkte“ auf der Diagonale im Verhältnis zur Gesamtzahl der Punkte auf der Diagonale.

Die Begriffe Sensitivität und Spezifität werden im Rahmen dieser Arbeit pauschal auf die Beurteilung einer Bibliotheksvalidierung angewendet. Bei einer externen Validierung können jedoch die Sensitivität und Spezifität auch für jede einzelne Substanz ermittelt werden. Dies geht jedoch über den Rahmen dieser Arbeit hinaus, da solche Aussagen nur mit einer langfristigen Beobachtung der Performance des Systems möglich sind.