

## **10 Die qualitative Analyse von chromatographischen Signalen**

Um einen Bezug zwischen dem Signal einer mit dem GC-MIP-MED untersuchten Substanz und deren Struktur oder Zusammensetzung herstellen zu können, muss herausgefunden werden, welche Faktoren das GC-MIP-MED-Signal wie beeinflussen können.

### **10.1 Die Normierung von GC-MIP-MED-Signalen**

Die Eigenschaften eines Plasmas hängen sehr stark von der zugeführten Energie und dem Druck ab (siehe Abschnitt 3.1.1). Diese beiden Größen beeinflussen stark die Plasmatemperaturen  $T_e$ ,  $T_{ion}$ ,  $T_{exc}$  und  $T_{gas}$ . Die Temperaturen repräsentieren die durchschnittlichen kinetischen Energien der Elektronen, Ionen und Gasteilchen sowie mit  $T_{exc}$  die Besetzungszahlen der Energiezustände der Neutralteilchen. Gerade die Besetzungszahlen haben einen großen Einfluss auf die Spektraldichten der Einzellinien der Molekülbandensysteme zweiatomiger Spezies (siehe Abschnitt 3.2). Aus diesem Grund wurde bei der Konstruktion der Apparatur größter Wert auf eine möglichst präzise und reproduzierbare Druckmessung (siehe Abschnitt 6.5), sowie eine möglichst reproduzierbare Mikrowellenenergieeinkopplung (siehe Abschnitt 6.7) gelegt. Trotzdem können Veränderungen der äußeren Umweltbedingungen wie die der Außentemperatur, des momentanen Luftdrucks oder der Luftfeuchtigkeit Einfluss auf die Funktion des Generators, sowie auf den Energieabtransport aus dem Plasma in die Umgebung nehmen. Die Energietransportprozesse haben wiederum einen Einfluss auf den Druck direkt im Plasma. Die Druckbestimmung wurde jedoch hinter dem Ansaugstutzen des Entladungsrohres vorgenommen (siehe Abbildung 6.3-2 und Abbildung 6.5-1). Mit einem aus Metallteilen bestehenden sperrigen Druckaufnehmer kann nicht direkt in einem Plasma gearbeitet werden. Diese Probleme haben zur Folge, dass sich die erzeugten Emissionsspektren im Laufe der Zeit etwas verändern. Reproduzierbarkeitsprobleme sind ein allgemeines Phänomen der Plasmaspektroskopie. Ihnen begegnet man im allgemeinen durch Vergleichsmessungen mit bekannten Standardproben.

Abbildung 10.1-1 zeigt bei konstanter Zufuhr von 0,00911 ml/min Propen aufgezeichnete Spektraldichten an verschiedenen Tagen über den Zeitraum von mehreren Monaten. Die Spektraldichten der einzelnen Banden und der Wasserstofflinie schwanken beträchtlich. Auch das Verhältnis der Spektraldichten bleibt nicht konstant, wie in Abbildung 10.1-2 zu sehen ist. Um den Einfluss der äußeren Umweltbedingungen möglichst gering zu halten, sind täglich vor den chromatographischen Messungen mit dem GC-MIP-MED die Spektraldichten der in Abbildung 10.1-1 aufgelisteten Banden und der Wasserstofflinie bei Zufuhr von 0,00911 ml/min Propen bestimmt worden. Mit diesen Messwerten sind alle Chromatogramme normiert worden. Unter Normierung ist die Division der bei der chromatographischen

Messung erhaltenen Spektraldichte durch die bei der Messung des externen Standards erhaltenen Spektraldichte zu verstehen:

$$S_N = \frac{S}{S_{ES}}$$

$S_N$  = normierte Spektraldichte. (10.1-1)

$S$  = während der chromatographischen Trennung gemessene Spektraldichte.

$S_{ES}$  = bei der Messung des externen Standards bestimmte Spektraldichte.

Mit dem GC-MIP-MED können bei Einleitung von Propen die Reaktionsprodukte CC, CH und OH und H nachgewiesen werden. Enthält eine mit dem GC-MIP-MED gemessene Substanz außer Sauerstoff weitere Heteroatome, so entstehen weitere meist zweiatomige Reaktionsprodukte. Bei stickstoffhaltigen organischen Verbindungen können CN und NH nachgewiesen werden. Enthält eine organische Substanz Schwefel, so wird CS gefunden. Bei bromhaltigen organischen Verbindungen können Bromlinien nachgewiesen werden. Die Spektraldichten von Linien und Banden, die Spezies zuzuordnen sind, welche Heteroatome enthalten, können nicht mit Propen als externem Standard normiert werden, da sie bei Einleitung von Propen in den GC-MIP-MED gleich Null sind. In diesen Fällen wird die Spektraldichte durch die einer Propenbande geteilt, die ähnlich auf äußere Umwelteinflüsse reagiert, wie die betreffende Bande. Die Spektraldichte des Rotationskopfes des CH4310Å-Systems ist bei ungünstigen Umweltbedingungen höher als der Messbereich des Spektrometers. Daher wird hier zur Normierung die Spektraldichte CH4310Å 3-Bande herangezogen, welche Teil des gleichen Rotationssystem ist. Tabelle 10.1-1 zeigt welche, Bande bzw. Linie bei der chromatographischen Messung auf welche Bande bzw. Linie des externen Standards normiert wird.

*Tabelle 10.1-1: Mit dem GC-MIP-MED chromatographisch erfasste Banden und Linien und die zur Normierung derselben herangezogenen Banden und Linien*

Im Chromatogramm gemessene Bande/Linie	zur Normierung verwendete Bande/Linie
Swan (2, 1)	Swan (2, 1)
Swan (1, 0)	Swan (1, 0)
Swan (3, 2)	Swan (3, 2)
Swan (0, 0)	Swan (0, 0)
CH 4310Å	CH 4310Å 3
CH 4310Å 2	CH 4310Å 2
CH 4310Å 3	CH 4310Å 3
H	H
OH 3064Å	OH 3064Å
CN-Violett (1, 1)	Swan (2, 1)
CN-Violett (0, 0)	Swan (2, 1)
CN-Violett (0, 1)	Swan (2, 1)
NH 3360Å (0, 0; Q)	H
NH 3360Å (1, 1; Q)	H

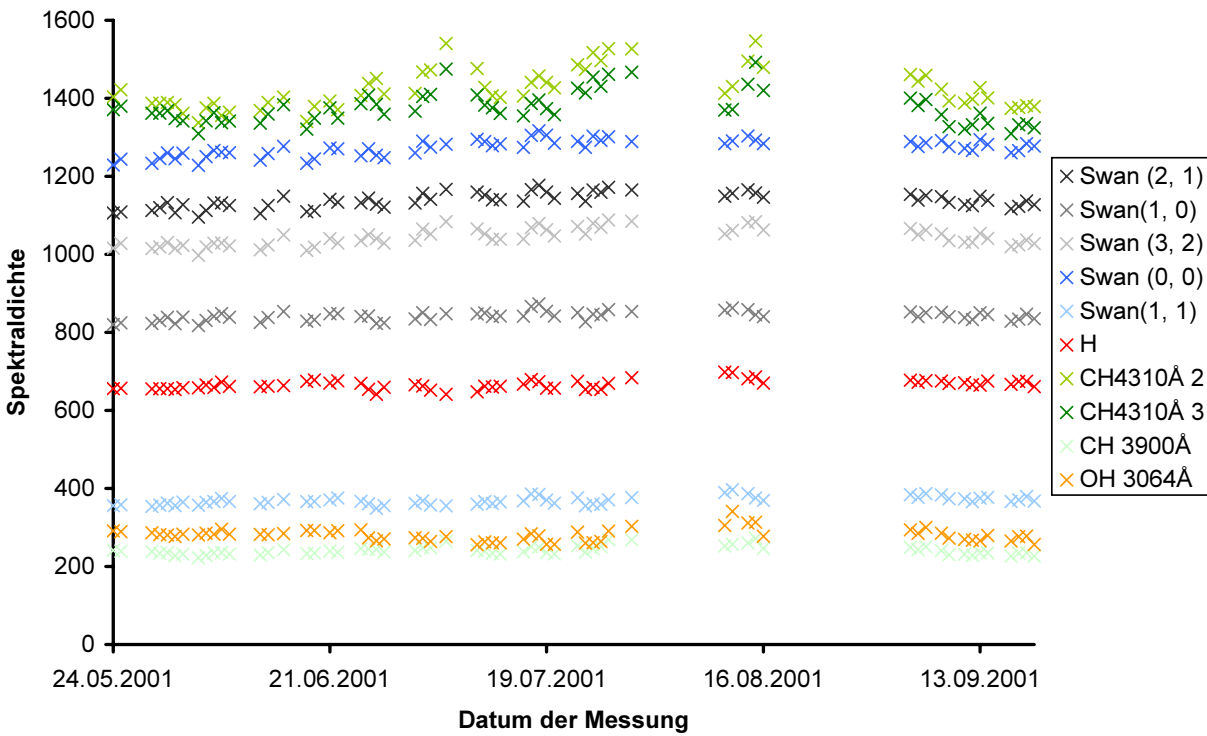


Abbildung 10.1-1: Die Spektraldichten ausgewählter Banden und der Wasserstofflinie bei Einleitung von 0,00911 ml/min des externen Standards Propen gegen das Datum der Messung

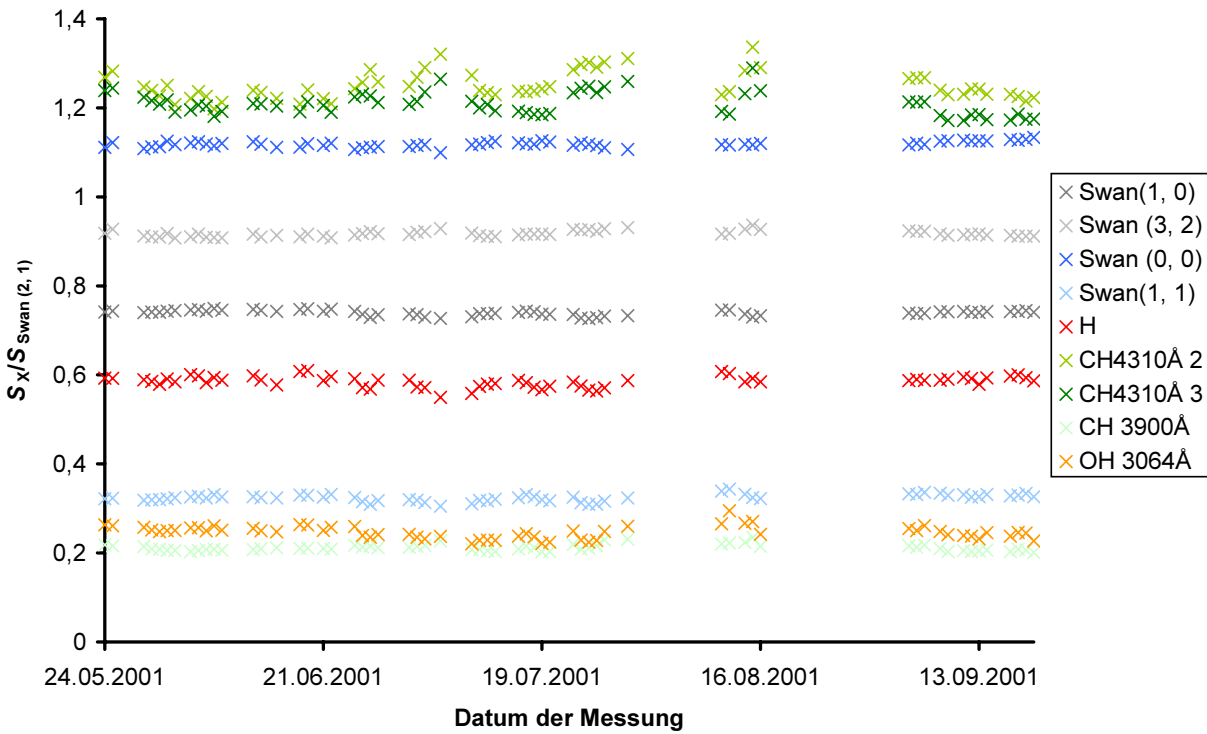


Abbildung 10.1-2: Das Verhältnis der Spektraldichten  $S_x$  ausgewählter Banden und der Wasserstofflinie zu der Spektraldichte der Swan (2, 1)-Bande  $S_{Swan(2,1)}$  bei Einleitung von 0,00911 ml/min des externen Standards Propen gegen das Datum der Messung

## 10.2 Die Klassifizierung chromatographischer GC-MIP-MED Signale

Um eine Beziehung zwischen einer Substanz und deren Detektorsignal analysieren zu können muss das Detektorsignal durch quantifizierbare Größen beschrieben werden können. Ein GC-MIP-MED-Signal besteht aus einer Reihe von Einzelsignalen, die auf dem Spektralbereich der Emissionsbanden und -linien der verschiedenen Reaktionsprodukte der untersuchten Substanz gemessen worden sind. Abbildung 10.2-1 zeigt ein chromatographisches MIP-MED-Signal einer Propenmessung, welches aus vier Einzelsignalen besteht. Das CC- und CH-Signal (Swan (2, 1) und CH 4310Å) besitzen annähernd die für ein chromatographisches Signal optimale Gauß-Glockenkurvenform. Die H- und OH-Signale zeigen Memory-Effekte im Detektor an. Während die CC- und CH-Signale wieder auf die Nulllinie zurückkehren, sobald die untersuchte Substanz den Detektor verlässt, findet man noch viele Sekunden später ein von Null verschiedenes H- und OH-Signal. Dieser Memory-Effekt tritt auf Grund der hohen Affinität des OH-Moleküls oder von eventuell gebildetem Wasser zu den Glasoberflächen im Detektor auf (siehe Abschnitt 9.2).

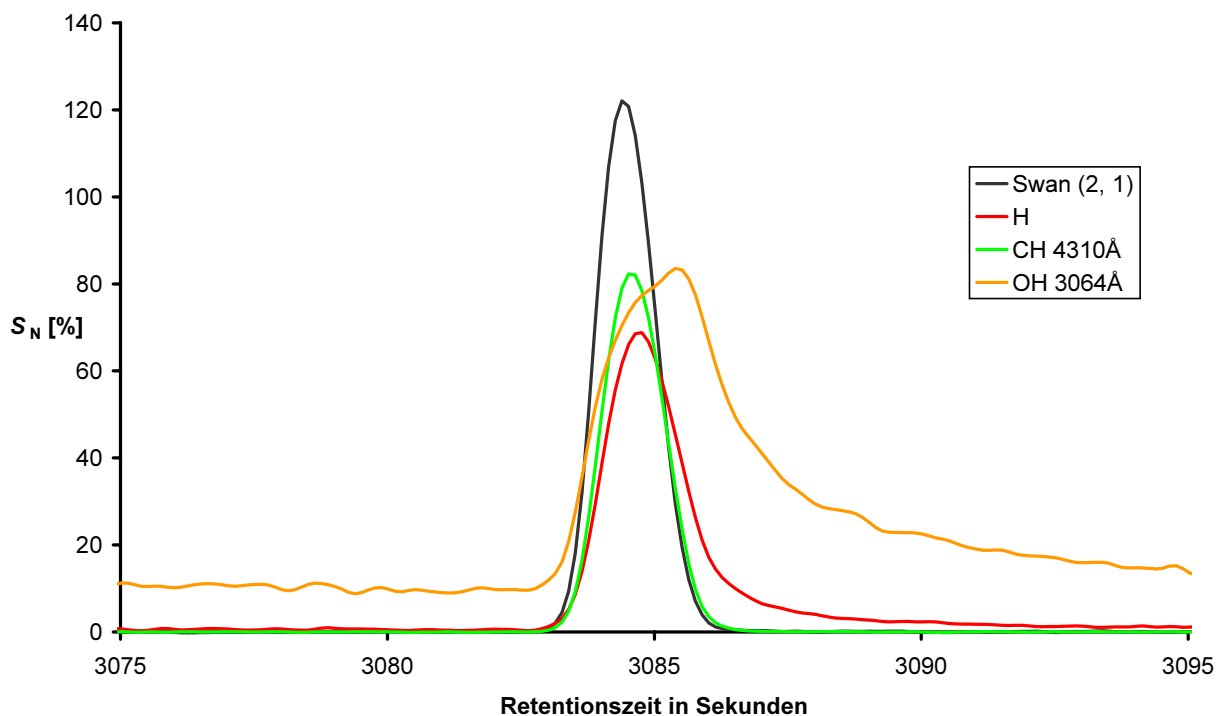


Abbildung 10.2-1: Chromatographisches MIP-MED-Propensignal

Die Signale von CC- und CH korrelieren mit der Menge an Substanz, die sich gerade in dem Detektor befindet. Das OH-Signal zeigt an, wie viele OH- oder Wasser-Moleküle an der Wand des Detektors haften. Daher lässt sich das Signal am besten durch folgende Parameter beschreiben und quantifizieren (siehe Abbildung 10.2-2 sowie Abschnitt 4.8):

1. Die Fläche  $F_{N, \text{Swan } (2, 1)}$  unter dem CC-Signal als Maß für die Substanzmenge, die dem Detektorsignal zu Grunde liegt,
2. die Breite auf halber Höhe  $B$  und der Betrag des Verhältnisses aus maximaler und minimaler Steigung  $D$  des CC-Signals (= Verhältnis der Steigungen an den Wendepunkten) als Maße dafür, wie schnell die Substanz in den Detektor gelangt und ihn auch wieder verlässt,
3. die Intensität  $O$  (= Offset) des OH-Signals kurz vor dem Anstieg des CH-Signals und dessen Fläche  $F_{N, \text{OH } 3064\text{Å}}$  vom ersten Ansteigen des CH-Signals bis zu dessen Wiedererreichen der Nulllinie integriert als Maße für den Memory-Effekt und
4. die Fläche unter dem CH-Signal  $F_{N, \text{CH } 4310\text{Å}}$  als Maß dafür, wie viele CH-Moleküle sich bei der Reaktion von Propen mit dem Plasma bilden. Je größer  $F_{N, \text{CH } 4310\text{Å}}$  ist, desto mehr CH-Moleküle haben sich im Verhältnis zu den CC-Molekülen gebildet.

Alle diese Signalparameter lassen sich mit Hilfe der eigens für den GC-MIP-MED entwickelten Software aus den aufgezeichneten und normierten Chromatogrammen automatisch ermitteln. Sie werden in der Chromatogrammdatei gespeichert und stehen innerhalb der Software in Tabellenform zur Verfügung (siehe Abbildung 8.1-2). Das Vorgehen der Software wird in den chemometrischen Grundlagen (Abschnitt 4.8 und 4.9) erläutert. Als Schwellenwert zur Erkennung von chromatographischen Signalen ist eine Signalstärke von  $S_N$  von 0,7% beim CH 4310Å-Rotationskopf festgelegt worden. Das Einzelchromatogramm, welches auf der Messung des CH 4310Å-Rotationskopfes basiert, weist bei der Messung fast aller Substanzen die größte Signalthöhe auf. Der Wert von 0,7% liegt nur wenig über der Erfassungsgrenze dieses Signals von:

$$S_{N, \text{CH}4310\text{Å}}^{\text{EG}} = 2 \cdot t_{\infty, 99\%} \cdot s_{N, \text{CH}4310\text{Å}} = 0,62\%$$

mit

$$t_{\infty, 99\%} = \text{Studentfaktor für die einseitige Fragestellung, } \infty - \text{viele Freiheitsgrade und ein Wahrscheinlichkeitsniveau von 99\%} = 2,326, \quad (10.2-1)$$

$$s_{N, \text{CH}4310\text{Å}} = \text{Standardabweichung des Rauschens des normierten CH}4310\text{Å-Signales} = 0,1339\% \text{ und}$$

$$S_{N, \text{CH}4310\text{Å}}^{\text{EG}} = \text{Erfassungsgrenze der Spektraldichte des CH}4310\text{Å-Signals.}$$

Punkt 4 beschreibt die eigentliche substanzspezifische Kenngröße. Wie viele CC- und CH-Moleküle sich im Verhältnis zueinander bilden, hängt von der Art der Substanz ab, die mit dem Detektor gemessen wird. Wird statt Propen beispielsweise Methan gemessen, so ist bei

ansonsten gleichen Parametern  $F_{N,CH\ 4310\text{\AA}}$  höher als bei Propen. Um  $F_{N,CH\ 4310\text{\AA}}$  bei verschiedenen Substanzen miteinander vergleichen zu können, dürfen die in den Punkten 1 bis 3 beschriebenen Parameter für die zum Vergleich herangezogenen Messungen allerdings nur wenig variieren. Man unterscheidet demnach zwei Gruppen von Variablen bzw. Parametern:

1. Bei  $F_{N,Swan\ (2,\ 1)}$ ,  $B$ ,  $D$ ,  $O$ , und  $F_{N,OH\ 3064\text{\AA}}$  handelt es sich um die sogenannten **Klassifizierungsparameter**, die das GC-MIP-MED-Signal beschreiben.
2.  $F_{N,CH\ 4310\text{\AA}}$  ist eine Größe, die von der Art der Substanz und der Größe der Klassifizierungsparameter abhängt. Es handelt sich hierbei um die analytische **Zielgröße**.  $F_{N,CH\ 4310\text{\AA}}$  ist nicht die einzige Signalfläche, welche als analytische Zielgröße in Frage kommt (siehe Tabelle 10.2-1). Im Falle von Propen können die Flächen weiterer CC und CH-Signale als analytische Zielgrößen herangezogen werden. Enthält eine organische Verbindung Heteroatome, so können weitere Signalflächen in Betracht gezogen werden. Es existieren also mehrere analytische Zielgrößen.

Die fünf Klassifizierungsparameter eines chromatographischen GC-MIP-MED-Signals werden durch die Betriebsbedingungen des Gaschromatographen, die Probenaufgabe, die Wechselwirkung der Substanz mit der chromatographischen Säule und durch den zeitlichen Abstand, in welchem der GC-MIP-MED zuvor weitere Signale aufgezeichnet hat, beeinflusst. Die Probenaufgabe ist in der Gaschromatographie nicht exakt reproduzierbar und die Betriebsbedingungen des Gaschromatographen werden oft durch äußere Umweltbedingungen mit beeinflusst. Diese Ungenauigkeiten verhindern, dass mit akzeptablem Aufwand Signale mit exakt ein und den selben Klassifizierungsparametern wiederholt gemessen werden können. Durch die unterschiedlichen Wechselwirkungen verschiedener Substanzen mit der chromatographischen Säule ist die Messung von Signalen verschiedener Substanzen mit gleichen Klassifizierungsparametern ein nahezu unmögliches Unterfangen. Bei Substanzgemischen müssen Probenaufgabe und Betriebsparameter des chromatographischen Systems auf eine ausreichende Trennleistung hin optimiert werden. Die Klassifizierungsparameter der chromatographischen MIP-MED-Signale sind hierdurch, abgesehen von den Schwankungen durch die Probenaufgabe, festgelegt. Aus diesen Gründen muss um qualitative Analytik betreiben zu können der Zusammenhang zwischen Klassifizierungsparametern und Zielgrößen quantifiziert werden. Gelingt dies, so können die Zielgrößen bei vorgegebenen Klassifizierungsparametern über eine Berechnung abgeschätzt werden, sobald für eine Substanz eine ausreichend hohe Zahl an Messungen vorliegt. Auf diese Weise können zwei Substanzen verglichen werden, obwohl kein Signal von der einen Substanz gemessen werden konnte, welches die selben Klassifizierungsparameter eines Signals von der anderen Substanz besitzt. Für ein Signal einer nicht bekannten Substanz kann abgeschätzt werden, wie hoch die Zielgrößen wären, wenn es sich um eine bestimmte Substanz handeln würde. Durch Vergleich der berechneten Zielgrößen und der tatsächlich gemessenen Zielgrößen kann eine Aussage

darüber gemacht werden, ob das Signal einer bestimmten Substanz zugeordnet werden kann, für die bereits Messungen vorliegen.

Tabelle 10.2-1: Bezeichnung der analytischen Zielgrößen bzw. der Signalflächen

Im Chromatogramm gemessene Bande/Linie	Bezeichnung der analytischen Zielgröße	
	Gemessene Größe $F_{N,x}$	Schätzwert $F_{S,x}$
Swan (2, 1)	$F_{N, \text{Swan (2, 1)}}$	$F_{S, \text{Swan (2, 1)}}$
Swan (1, 0)	$F_{N, \text{Swan (1, 0)}}$	$F_{S, \text{Swan (1, 0)}}$
Swan (3, 2)	$F_{N, \text{Swan (3, 2)}}$	$F_{S, \text{Swan (3, 2)}}$
Swan (0, 0)	$F_{N, \text{Swan (0, 0)}}$	$F_{S, \text{Swan (0, 0)}}$
CH 4310Å	$F_{N, \text{CH 4310Å}}$	$F_{S, \text{CH 4310Å}}$
CH 4310Å 2	$F_{N, \text{CH 4310Å 2}}$	$F_{S, \text{CH 4310Å 2}}$
CH 4310Å 3	$F_{N, \text{CH 4310Å 3}}$	$F_{S, \text{CH 4310Å 3}}$
OH 3064Å	$F_{N, \text{OH 3064Å}}$	$F_{S, \text{OH 3064Å}}$
CN-Violett (1, 1)	$F_{N, \text{CN-Violett (1, 1)}}$	$F_{S, \text{CN-Violett (1, 1)}}$
CN-Violett (0, 0)	$F_{N, \text{CN-Violett (0, 0)}}$	$F_{S, \text{CN-Violett (0, 0)}}$
CN-Violett (0, 1)	$F_{N, \text{CN-Violett (0, 1)}}$	$F_{S, \text{CN-Violett (0, 1)}}$
NH 3360Å (0, 0; Q)	$F_{N, \text{NH 3360Å (0, 0; Q)}}$	$F_{S, \text{NH 3360Å (0, 0; Q)}}$
NH 3360Å (1, 1; Q)	$F_{N, \text{NH 3360Å (1, 1; Q)}}$	$F_{S, \text{NH 3360Å (1, 1; Q)}}$

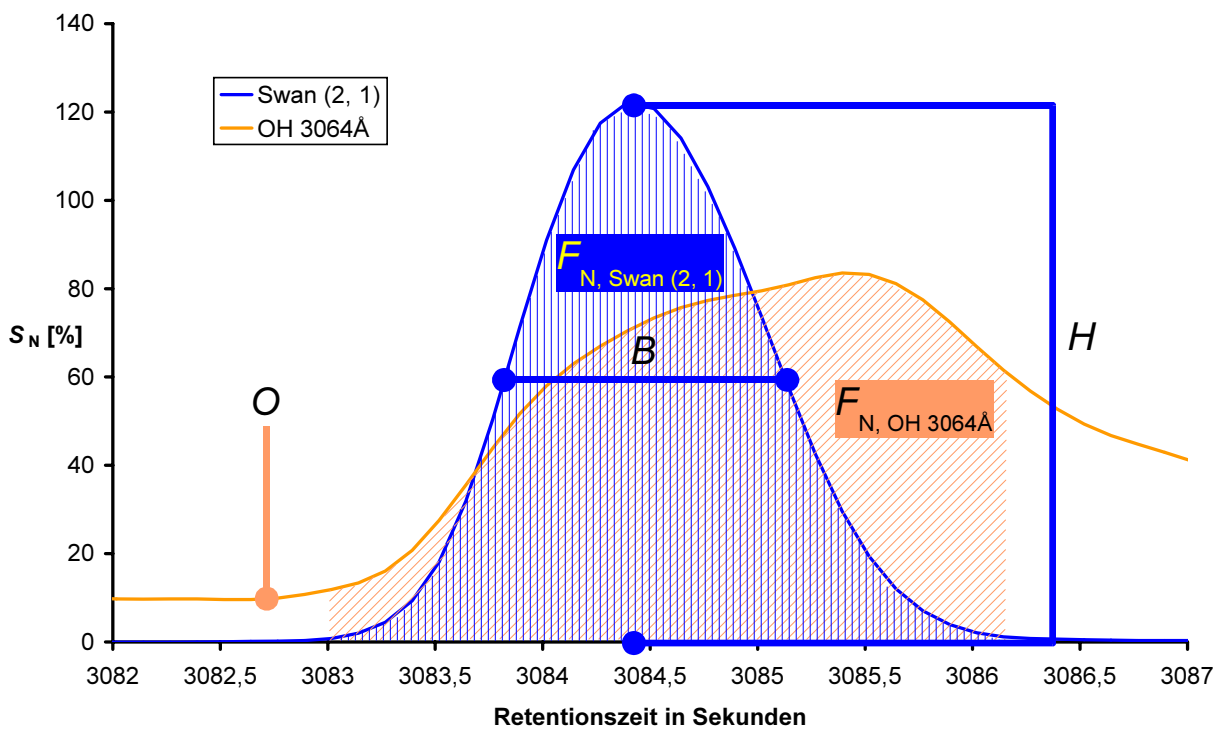


Abbildung 10.2-2: Klassifizierung eines Propensignals

## 10.3 Entwicklung einer Methode zur Auswertung von GC-MIP-MED-Signalen

Um eine Methode zur qualitativen Analyse der chromatographischen Daten zu entwickeln, wurden chromatographische Signale einer Reihe von organischen Verbindungen bei wechselnden Betriebsparametern des Gaschromatographen mit dem GC-MIP-MED aufgezeichnet. Pro Substanz wurden 579 bis 1663 Messungen durchgeführt und in das Datenbankmodul der Auswertesoftware überführt (siehe Tabelle 10.3-1). Tabelle 9.1-1 (Abschnitt 9.1) enthält eine Beschreibung des Verfahrens, mit dem alle Chromatogramme aufgezeichnet und weiterverarbeitet worden sind. Sie wurden stets nach den Vorgaben von Tabelle 10.1-1 in Abschnitt 10.1 normiert. Für stickstoff- und nicht stickstoffhaltige Verbindungen wurden zwei verschiedene Datenbanken angelegt. Diese beiden Datenbanken sind nach den in den Verbindungen enthaltenen Elementen bezeichnet worden. Die CHO-Datenbank enthält die Messwerte aller Verbindungen, die ausschließlich Kohlenstoff, Wasserstoff und Sauerstoff enthalten und CHON-Datenbank die Messwerte aller Verbindungen, die zusätzlich Stickstoff enthalten.

Die Messwerte für die CHO-Datenbank sind zuerst aufgezeichnet worden. Im Verlauf der Messungen wurde festgestellt, dass die Basislinien der Chromatogramme auf Grund von elektronischen Störungen bei der Weiterleitung der Analogsignale von den CCD-Sensoren zum Wandler hin und wieder driften. Bei einer genaueren Untersuchung des Problems wurde festgestellt, dass es sich bei dem Driftproblem um eine Parallelverschiebung der Basislinien aller Einzelchromatogramme handelt. Aus diesem Grund wurde die Spektrometersoftware um die Möglichkeit einer Driftkorrektur erweitert. Das implementierte Driftkorrekturverfahren ist in Abschnitt 4.9.2 beschrieben. Es setzt das Vorhandensein eines Driftkorrekturchromatogramms voraus, welches auf Basis eines Spektralbereiches aufgezeichnet worden ist, in dem kein Signal außer dem Spektrometerrauschen zu messen ist. Für die bereits durchgeführten Messungen ist allerdings kein solches Driftkorrekturchromatogramm aufgezeichnet worden. Zur Kontrolle, ob Stickstoffverunreinigungen im Argon oder Sauerstoff vorhanden sind, sind aber Einzelchromatogramme auf Basis der Spektraldichte für die Banden von CH und NH gespeichert worden. Traten Driftkorrekturprobleme auf, so diente das Einzelchromatogramm für die NH 3360Å (0, 0; Q)-Bande als Driftkorrekturchromatogramm. Dieses Einzelchromatogramm eignet sich deshalb so gut für die Driftkorrektur, weil in dem Bereich wo das Spektraldichtemaximum der NH 3360Å (0, 0; Q)-Bande auftritt keine spektralen Interferenzen zu beobachten sind. Bei der CN Violett (0, 0)- und CN Violett (1, 1)-Bande treten leichte Überlagerungen mit CH-Signalen auf, die angesichts der hohen Spektraldichte dieser Banden für die Untersuchung stickstoffhaltiger Verbindungen bedeutungslos sind. Die spektralen Interferenzen führen jedoch bei der Messung nicht stickstoffhaltiger Verbindungen zu einem minimalen Signal, welches bei der Verwendung des Chromatogramms als Driftkorrekturchromatogramm störend wirken könnte. Bei der anschließenden Messung stickstoffhaltiger Verbindungen ist ein Driftkorrektur-einzelchromatogramm stets mit aufgezeichnet worden (siehe



Tabelle 9.1-3). Die Chromatogramme, deren Signaldaten in der CHON-Datenbank gespeichert sind, wurden immer einer Driftkorrektur unterworfen.

Häufig ist zu beobachten, dass die maximale Spektraldichte eines chromatographischen MIP-MED-Signals den Messbereich des Spektrometers überschreitet. Dies geschieht besonders bei den Einzelchromatogrammen, die auf Basis sehr starker Banden wie dem CH 4310Å-Rotationskopf oder dem CN Violett (0, 0)-Signal aufgezeichnet werden. Auch in solchen Fällen macht es Sinn, diese Signale in die Datenbank aufzunehmen. Im Falle des CH 4310Å-Signals kann für qualitative Untersuchungen auf ein weniger starkes Nebensignal der Rotationsbande wie das CH 4310Å 3-Signal ausgewichen werden. Mit dem CN Violett (1, 1)-Signal steht auch ein schwächeres CN-Signal zur Verfügung. Überschreitet das Signal bei weiteren Einzelchromatogrammen den Messbereich des Spektrometers, so ist es zu verwerfen.

Die Quantifizierung sehr kleiner Signale oder von Signalen mit sehr geringem Offset hat sich als schwierig erwiesen. Bei der Bestimmung des absoluten Minimums muss beachtet werden, dass hierbei die niedrigste Spektraldichte ermittelt wird, die bis zu 10 Punkten vom Signalsockel entfernt liegen darf (vergleiche Abschnitt 4.8). Es handelt sich demnach um den niedrigsten Punkt einer rauschbehafteten Basislinie. Daher ist der so bestimmte Offset systematisch zu niedrig. Wird die Fläche sehr kleiner Signale bestimmt, so werden die Flächenanteile vor und nach den auf Basis eines Schwellenwertes von  $S_N\% = 0,7\%$  beim CH 3410Å-Signal ermittelten Signalgrenzen abgeschnitten. Daher sind die automatisch ermittelten Flächen systematisch zu klein. Aus diesen Gründen sind nur Signale in den Datenbanken gespeichert, deren Offset höher als  $S_N\% = 3\%$  und deren Fläche größer als  $S_N\% = 10\% \cdot s$  sind.

Tabelle 10.3-1: Untersuchte Substanzen und deren Anzahl an Signalen in der Datenbank

Substanz	Anzahl an Signalen in der Datenbank	Substanz	Anzahl an Signalen in der Datenbank
Aceton	1163	Hexan	580
Anilin	763	Isooctan	833
Benzen	1663	Methan	741
Butadien	579	Methanol	895
Butan	738	Octan	863
Butanol	1211	$\alpha$ -Picolin	937
1, 3-Cyclohexadien	600	Propan	678
Cyclohexan	819	Propanal	918
Diethylether	1281	Propanol	1044
Dimethylether	1318	Propen	1474
Ethan	1335	Propylamin	889
Ethanol	1226	Pyridin	771
Ethen	625	Triethylamin	894
Ethylbenzen	816	Trimethylamin	769

### 10.3.1 Voruntersuchungen

Trägt man die Fläche des CH4310Å 3-Signals gegen die des Swan (2, 1)-Signals für verschiedene Substanzen auf, so stellt man fest, dass ein linearer Zusammenhang zwischen diesen beiden Größen besteht (siehe Abbildung 10.3-1). Die lineare Korrelation zwischen den beiden Größen verbessert sich zunehmend, wenn nur die Signale aufgetragen werden, deren Klassifizierungsparameter  $F_{N, OH\ 3064\text{Å}}$  und  $O$  recht ähnlich sind (siehe Abbildung 10.3-2 und Abbildung 10.3-3). Aus diesem Grund kann angenommen werden, dass durch die Klassifizierungsparameter  $F_{N, Swan\ (2, 1)}$ ,  $F_{N, OH\ 3064\text{Å}}$  und  $O$  die Fläche des CH 4310Å 3-Signals  $F_{N, CH4310\text{Å}\ 3}$  für die jeweilige Substanz relativ gut vorhergesagt werden kann. Die Klassifizierungsparameter Halbwertsbreite  $B$  und Deformationsfaktor  $D$  haben demnach wenn überhaupt einen relativ geringen Einfluss auf die Zielgröße. Die Ergebnisse dieses Vortests zeigen, dass die multivariate lineare Regression ein vielversprechender Ansatz zur Bestimmung des Zusammenhangs zwischen Klassifizierungsparametern und Zielgrößen ist.

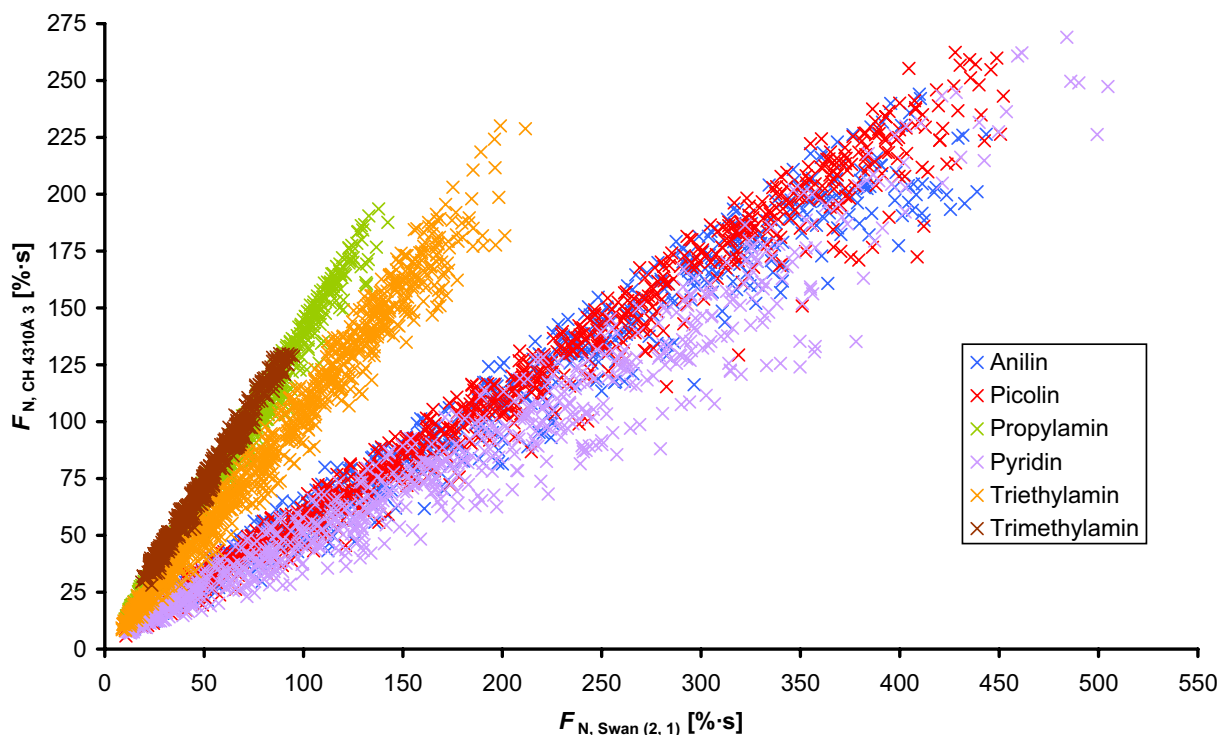


Abbildung 10.3-1: Fläche des Swan (2, 1)-Signals gegen die des CH 4310Å 3-Signals für alle in der Datenbank gespeicherten Signale stickstoffhaltiger Verbindungen

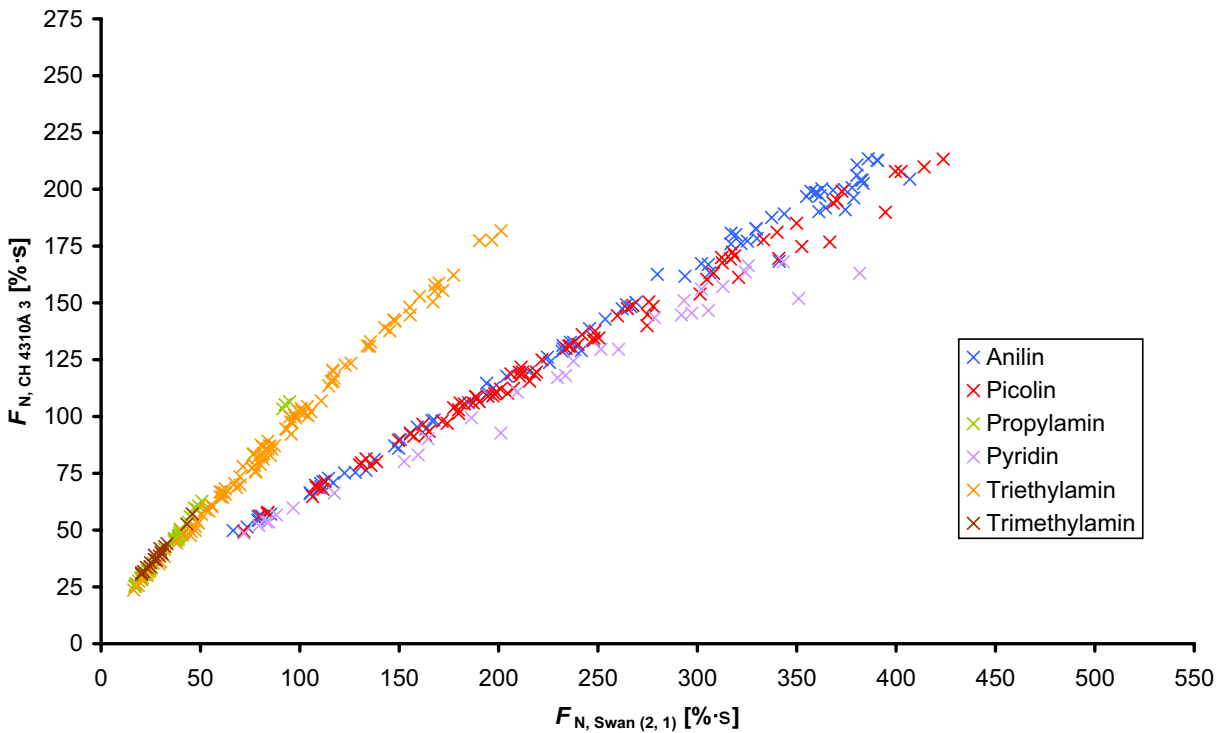


Abbildung 10.3-2: Fläche des Swan (2, 1)-Signals gegen die des CH 4310Å 3-Signals für alle in der Datenbank gespeicherten Signale stickstoffhaltiger Verbindungen, deren Fläche des OH 3064Å-Signals größer als 240 und kleiner als 260%·s ist

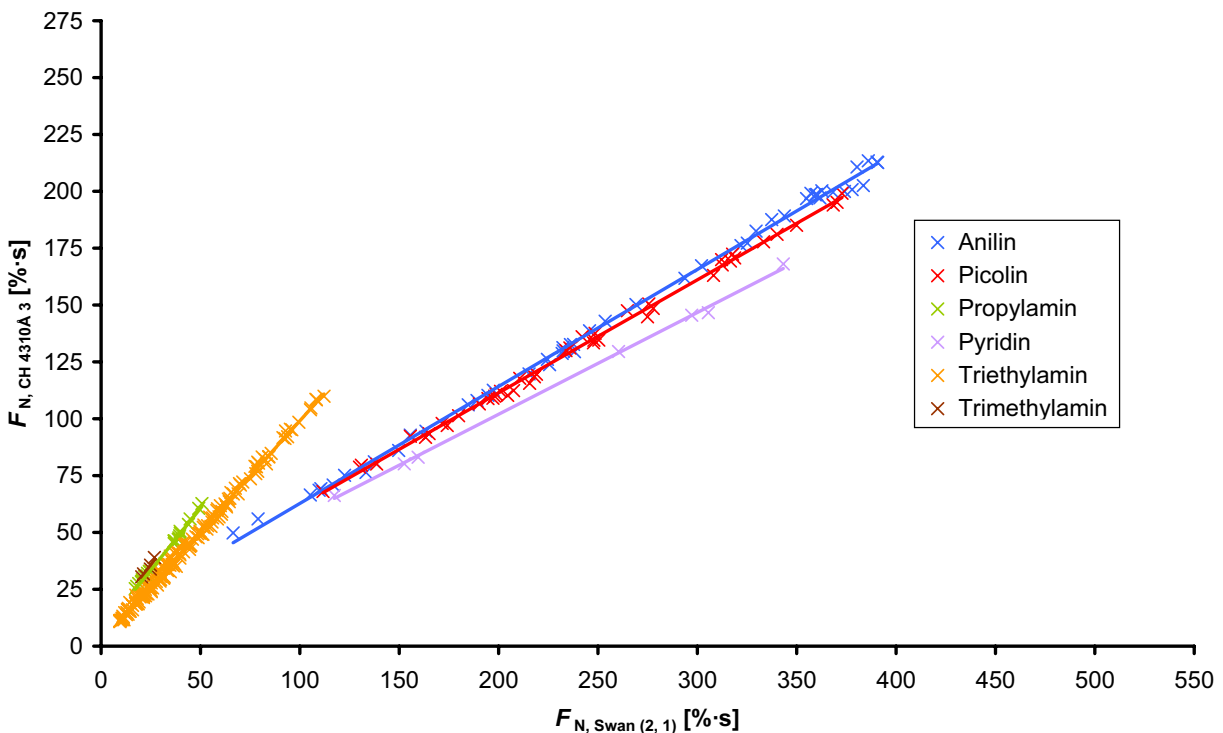


Abbildung 10.3-3: Fläche des Swan (2, 1)-Signals gegen die des CH 4310Å 3-Signals für alle in der Datenbank gespeicherten Signale stickstoffhaltiger Verbindungen, deren Fläche des OH 3064Å-Signals größer als 240 und kleiner als 260%·s ist sowie deren Offset zwischen 15 und 25% liegt

### 10.3.2 Multivariate Regression und Variablenreduktion

Um den Zusammenhang zwischen den Klassifizierungsparametern und der Fläche verschiedener Signale, insbesondere des CH 4310Å 3-Signales, bestimmen zu können, wurden zunächst die kompletten Datensätze in der Datenbank mittels der multivariaten linearen Regression untersucht. Hierbei wurden die Klassifizierungsparameter als Regressoren und die vorherzusagenden Signalflächen als Regressanden betrachtet. Zunächst sind für die Datensätze verschiedener Substanzen Residuenplots erstellt worden. Abbildung 10.3-4 zeigt exemplarisch den für Anilin. Die in Abhängigkeit vom Regressor  $F_{N, \text{Swan}}(2, 1)$  aufgetragenen Residuen streuen nicht normalverteilt um den Nullwert, sondern zeigen einen deutlichen Trend. Für kleine und große  $F_{N, \text{Swan}}(2, 1)$ -Werte sind sie tendenziell höher als Null und für mittlere  $F_{N, \text{Swan}}(2, 1)$ -Werte tendenziell niedriger. Dieser Sachverhalt erschwert weitere statistische Auswertungen. Daher wurde eine Reihe von Regressionsrechnungen durchgeführt, welche nur einen Teil der für eine Substanz vorhandenen Messwerte berücksichtigen. Die Messwerte sollten so ausgewählt werden, dass sie einander möglichst ähnlich sind.

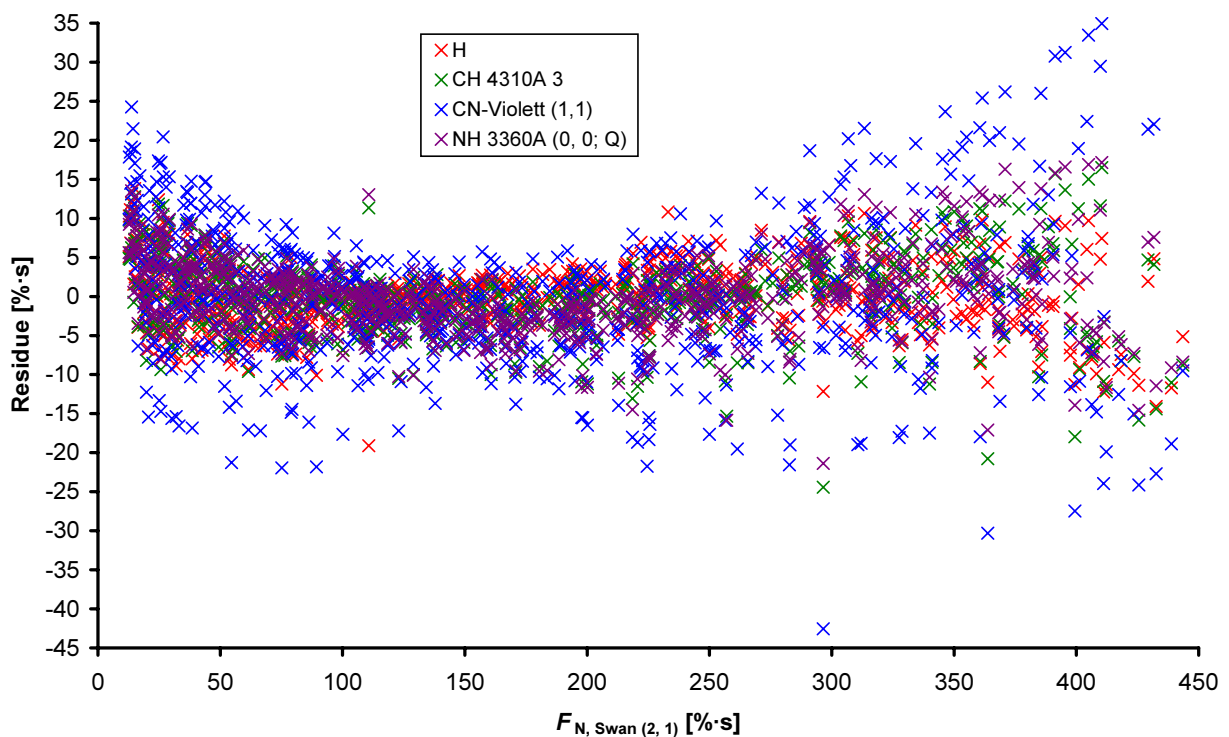


Abbildung 10.3-4: Die aus der Regressionsrechnung mit dem gesamten Anilin-Datensatz ermittelten Residuen in Abhängigkeit von  $F_{N, \text{Swan}}(2, 1)$

Um nur eine Auswahl von Messwerten zur Regression heranziehen zu können, muss zunächst ein Auswahlkriterium geschaffen werden. Zu diesem Zweck wird für jede in der Datenbank gespeicherte Messung folgender Regressorvektor aufgestellt:

$$\vec{r} = \begin{pmatrix} F_{N,Swan(2,1)} \\ F_{N,OH\ 3064\text{\AA}} \\ \mathbf{O} \\ \mathbf{B} \\ \mathbf{D} \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \text{Regressorvektor.} \quad (10.3-1)$$

Anschließend erfolgt eine Distanzbestimmung zwischen jedem einzelnen Regressorvektor und allen übrigen Regressorvektoren nach Gleichung (4.3-7) (Abschnitt 4.3):

Seien  $\vec{r}$  = beliebiger Regressorvektor und  $\{\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n\}$  die Menge aller Regressorvektoren im Datensatz so, wird durch Sortieren und indizieren (siehe Abschnitt 4.3) für alle Vektoren in  $\{\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n\}$  ein Indexvektor erstellt, der die Distanz zwischen  $\vec{r}$  und  $\vec{r}_j \in \{\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n\}$  repräsentiert. Man erhält demnach eine Menge  $\{\vec{i}_1, \vec{i}_2, \dots, \vec{i}_n\}$  von Indexvektoren, deren Länge

$$|\vec{i}_j| = \sqrt{(i_1)^2 + (i_2)^2 + (i_3)^2 + (i_4)^2 + (i_5)^2} \quad \text{mit} \quad |\vec{i}_j| \in \{\vec{i}_1, \vec{i}_2, \dots, \vec{i}_n\}$$

den jeweiligen Abstand zu  $\vec{r}$  repräsentiert.

Sortiert man die Regressorvektoren nach den so erhaltenen Distanzen und verwirft die Vektoren, mit denen zu hohe Distanzen ermittelt worden sind, so erhält man einen Teilansatz aus zu dem ausgewählten Vektor ähnlichen Regressorvektoren. Das Datenbankmodul der Windows-Software ermöglicht es den Prozentsatz an Vektoren auszuwählen, der für den so bestimmten Teilansatz zur Regressionsrechnung herangezogen werden soll. Die Auswirkungen dieser Variablenreduktion ist zunächst mit einer Analyse der erhaltenen Residuenvarianzen untersucht worden. Abbildung 10.3-5 zeigt die Ergebnisse einer solchen Varianzanalyse für die Vorhersage der Flächen der CH 4310Å 3-Signale von Anilin. Hierbei wurden für jeden im Anilindatensatz enthaltenen Regressorvektor nach dem beschriebenen Verfahren Teilansätze in unterschiedlicher Größe gerechnet. In der Abbildung sind die maximale und die minimale Residuenvarianz und der Mittelwert der Varianzen aus allen gerechneten Teilansätzen einer definierten Größe gegen die Größe dieser Teilansätze aufgetragen. Es wird deutlich sichtbar, dass der Varianzmittelwert um so geringer und somit die Güte der Vorhersage im Durchschnitt um so besser ist, je kleiner der Teilansatz ist. Die Abnahme der mittleren Varianz wird zu kleineren Ansätzen hin aber immer geringer. Bei sehr kleinen Teilansätzen können aber unverhältnismäßig hohe Maximalvarianzen auftreten. Analoges gilt auch für die Signalfächen der anderen Banden und Linien. Abbildung 10.3-6 bis Abbildung 10.3-8 zeigen Residuenplots für einen Einzelansatz aus 15% der Signale, die

im Anilin-Datensatz gespeichert worden sind. Es handelt sich hierbei um die Signale, welche zu einem Signal mit folgenden Klassifizierungsparametern bzw. Regressoren am ähnlichsten sind:

$$\bar{\mathbf{r}}_V = \begin{pmatrix} F_{N, \text{Swan (2,1)}} \\ F_{N, \text{OH 3064\AA}} \\ \mathbf{O} \\ \mathbf{B} \\ \mathbf{D} \end{pmatrix} = \begin{pmatrix} 320 \% \cdot \text{s} \\ 230 \% \cdot \text{s} \\ 16 \% \\ 1,75 \text{ s} \\ 1,3 \end{pmatrix} \quad (10.3-3)$$

Die Residuenplots zeigen keine Trends bezüglich eines Regressors, sondern streuen annähernd normalverteilt um Null. Über eine Varianzanalyse kann auch abgeschätzt werden, ob wirklich alle 5 Regressoren benötigt werden (siehe Abschnitt 4.7.5). Zu diesem Zweck sind zusätzlich zu dem Teilansatz mit 15% aller Anilin-Signale mit 5 Regressoren auch die Teilansätze mit allen Regressorensubsets berechnet worden. Tabelle 10.3-5 beinhaltet die Ergebnisse dieser Berechnungen. Bei  $F_H$ ,  $F_{CH\ 4310\AA\ 3}$  und  $F_{NH\ 3360\AA\ (0, 0; Q)}$  ist die Residuenvarianz für die Subsets am niedrigsten, die  $F_{N, \text{Swan (2,1)}}$ ,  $F_{N, \text{OH 3064\AA}}$  und  $\mathbf{O}$  enthalten. Im Falle von  $F_{CN\ \text{Violett (1, 1)}}$  sollte zusätzlich  $\mathbf{B}$  in das Subset aufgenommen werden, um möglichst geringe Varianzen zu erhalten. Die sich aus den Subsets ohne  $\mathbf{B}$  aber mit  $F_{N, \text{Swan (2,1)}}$ ,  $F_{N, \text{OH 3064\AA}}$  und  $\mathbf{O}$  ergebende Residuenvarianzen liegen mit 18,81 bzw. 19,59%·s gegenüber 13,78 bzw. 14,53%·s noch in einer akzeptablen Größenordnung. Die Vermutung liegt nahe, dass die Regressoren  $F_{N, \text{Swan (2,1)}}$ ,  $F_{N, \text{OH 3064\AA}}$  und  $\mathbf{O}$  für eine erfolgreiche Vorhersage der Flächen aller übrigen Signale die wichtigsten sind. Dieser Sachverhalt deutete sich schon bei den Voruntersuchungen an. Untermuert wird diese Theorie durch die inneren Bestimmtheitsmaße (siehe Abschnitt 4.7.3) des Ansatzes mit allen Regressoren und dem Ansatz mit  $F_{N, \text{Swan (2,1)}}$ ,  $F_{N, \text{OH 3064\AA}}$  und  $\mathbf{O}$  als Regressoren (siehe Tabelle 10.3-2). Das starke Absinken der inneren Bestimmtheitsmaße bei der Reduktion der Regressorenzahl lässt den Schluss zu, dass  $\mathbf{B}$  und  $\mathbf{D}$  nur redundante, aber keine zusätzlichen Informationen zu der Vorhersage beisteuern. Betrachtet man die maximale, mittlere und minimale Varianz aus den Teilansätzen, die auf den Regressorvektoren aller in der Datenbank vorhandenen Regressoren basieren, so wird diese These ebenfalls untermuert. Je nach untersuchter Bande bzw. Linie verursacht die Eliminierung von  $\mathbf{B}$  und  $\mathbf{D}$  aus dem Regressorset eine leichten Erhöhung oder Verringerung der Residuenvarianzen (siehe Tabelle 10.3-3 und Tabelle 10.3-4). Lediglich die minimale Residuenvarianz für  $F_{N, \text{CN Violett (1, 1)}}$  als Regressand ist bei Verwendung des kleineren Regressorsets deutlich erhöht. Aus diesem Gründen können  $\mathbf{B}$  und  $\mathbf{D}$  für alle weiteren Regressionsrechnungen aus dem Regressorset eliminiert werden.

Als letzter Schritt zur Optimierung des Regressionsverfahrens wurde die Distanzmessung beim Signalauswahlverfahren durch eine Gewichtung der Regressoren erweitert. Zu diesem Zweck wird Gleichung (10.3-1) für jeden Regressor um einen Gewichtungsfaktor  $g$  erweitert. Da  $\mathbf{B}$  und  $\mathbf{D}$  kaum einen Beitrag zur Verbesserung des durch die lineare Regression ermittelten Schätzwertes leisten, werden die Gewichtungsfaktoren dieser beiden Regressoren

auf Null gesetzt. Die Distanzmessung über das in (10.3-2) beschriebene Verfahren erfolgt also auf Basis des folgenden Regressorvektors:

$$\begin{aligned} \vec{r} &= \begin{pmatrix} F_{N,Swan(2,1)} \cdot g_{Swan(2,1)} \\ F_{N,OH\ 3064\text{\AA}} \cdot g_{OH\ 3064\text{\AA}} \\ \mathbf{O} \quad \cdot \quad g_{\mathbf{O}} \\ \mathbf{B} \quad \cdot \quad g_{\mathbf{B}} \\ \mathbf{D} \quad \cdot \quad g_{\mathbf{D}} \end{pmatrix} = \begin{pmatrix} F_{N,Swan(2,1)} \cdot g_{Swan(2,1)} \\ F_{N,OH\ 3064\text{\AA}} \cdot g_{OH\ 3064\text{\AA}} \\ \mathbf{O} \quad \cdot \quad g_{\mathbf{O}} \\ \mathbf{B} \quad \cdot \quad \mathbf{0} \\ \mathbf{D} \quad \cdot \quad \mathbf{0} \end{pmatrix} \\ &= \begin{pmatrix} F_{N,Swan(2,1)} \cdot g_{Swan(2,1)} \\ F_{N,OH\ 3064\text{\AA}} \cdot g_{OH\ 3064\text{\AA}} \\ \mathbf{O} \quad \cdot \quad g_{\mathbf{O}} \end{pmatrix} \end{aligned} \quad (10.3-4)$$

Wie aus Tabelle 10.3-6 ersichtlich wird, kann durch Variation der Größe der übrigen Gewichtungsfaktoren eine leichte Verringerung der Residuenvarianz bewirkt werden. Die kleinsten Varianzen treten bei der Abschätzung von  $F_{N,H}$ ,  $F_{N,CH4310\text{\AA}3}$  und  $F_{N,NH\ 3360\text{\AA}}(0,0;Q)$  bei dem Gewichtungsfaktor 1 für  $F_{N,Swan(2,1)}$  und 0,5 für die übrigen beiden Gewichtungsfaktoren auf. Wird  $F_{N,CN\ violett(1,1)}$  vorhergesagt, so darf nur der Gewichtungsfaktor für  $\mathbf{O}$  auf 0,5 herabgesetzt werden und alle anderen Gewichtungsfaktoren müssen auf 1 gesetzt bleiben, wenn die mittlere Varianz der Regressionsrechnungen minimal werden soll.

Durch Variablenreduktion lässt sich die Güte der Schätzwerte der Signalflächen erheblich verbessern. Es hat sich gezeigt, dass auf die Regressoren  $\mathbf{B}$  und  $\mathbf{D}$  verzichtet werden kann, weshalb diese Regressoren bei allen folgenden Regressionsrechnungen nicht mehr verwendet werden. Der Regressorvektor wird also zu

$$\vec{r}_V = \begin{pmatrix} F_{N,Swan(2,1)} \cdot g_{Swan(2,1)} \\ F_{N,OH\ 3064\text{\AA}} \cdot g_{OH\ 3064\text{\AA}} \\ \mathbf{O} \quad \cdot \quad g_{\mathbf{O}} \end{pmatrix} = \begin{pmatrix} F_{N,Swan(2,1)} \cdot 1 \\ F_{N,OH\ 3064\text{\AA}} \cdot 0,5 \\ \mathbf{O} \quad \cdot \quad 0,5 \end{pmatrix} = \begin{pmatrix} 1 \cdot F_{N,Swan(2,1)} \\ 0,5 \cdot F_{N,OH\ 3064\text{\AA}} \\ 0,5 \cdot \mathbf{O} \end{pmatrix}. \quad (10.3-5)$$

Die Gewichtung der Regressoren wird stets auf  $g_{Swan(2,1)} = 1$  für  $F_{N,Swan(2,1)}$ ,  $g_{OH\ 3064\text{\AA}} = 0,5$  für  $F_{N,OH\ 3064\text{\AA}}$  und  $g_{\mathbf{O}} = 0,5$  gesetzt. Auch sollten nur die ähnlichsten Signale in die Regressionsrechnung einbezogen werden. Grundsätzlich gilt, dass die Residuenvarianz um so geringer ist, je weniger und damit zum abzuschätzenden Signal ähnlichere Signale in die Regressionsrechnung einbezogen werden. Bei sehr kleinen Ansätzen steigt die Güte der Vorhersage durch eine weitere Verkleinerung der zur Regressionsrechnung herangezogenen Signanzahl nicht mehr so stark, wie bei großen Ansätzen. Die Gefahr, dass einzelne Schätzungen sehr schlecht sind, wächst bei zu starker Verringerung der Signanzahl. Sollen Signalflächen anhand von Regressorkombinationen abgeschätzt werden, welche mit im Datensatz vorhandenen Signalen keine sehr große Ähnlichkeit besitzen, so dürfte die Güte des Schätzwertes bei einem zu sehr verkleinerten Ansatz ebenfalls nicht mehr optimal sein. Aus diesem Grund ist für alle folgenden Regressionsberechnung eine Verringerung des Ansatzes auf 15% der ähnlichsten Signale vorgenommen worden.

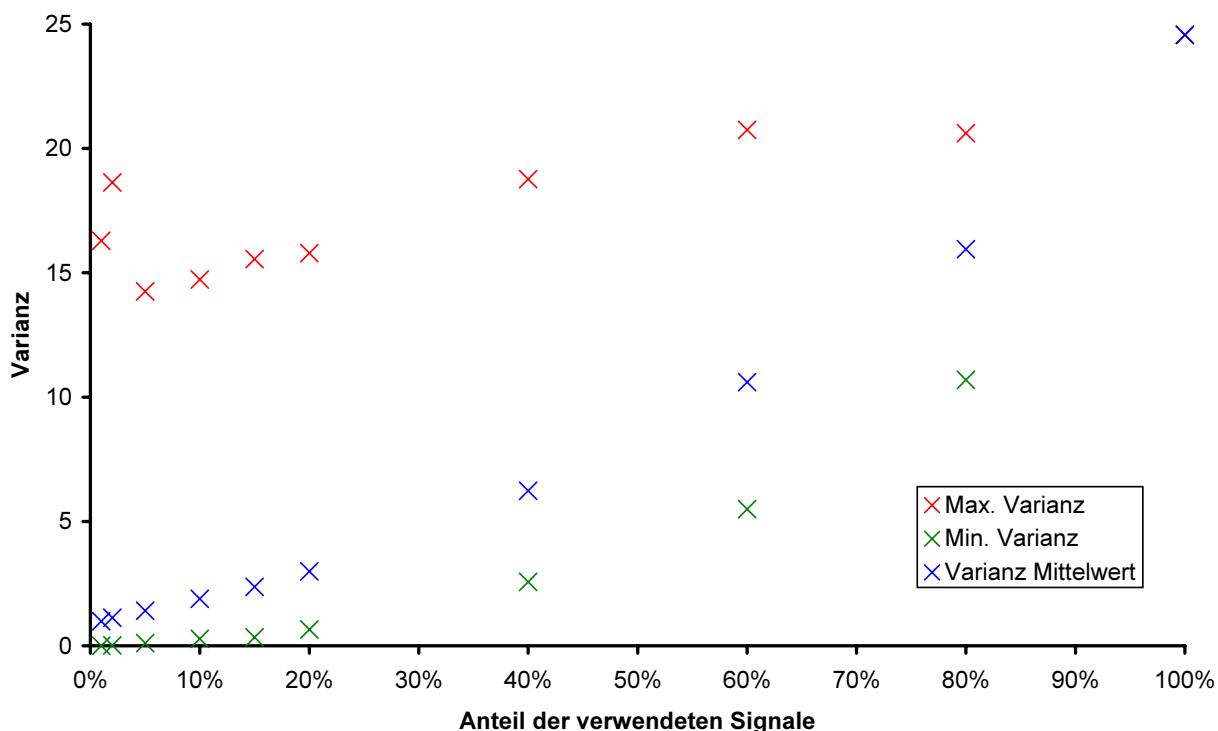


Abbildung 10.3-5: Die Residuenvarianz in Abhängigkeit von der prozentualen Größe des Teilsatzes bezüglich des Totalansatzes bei der Vorhersage der  $F_{CH\ 4310\text{Å}\ 3}$ -Werte für Anilin

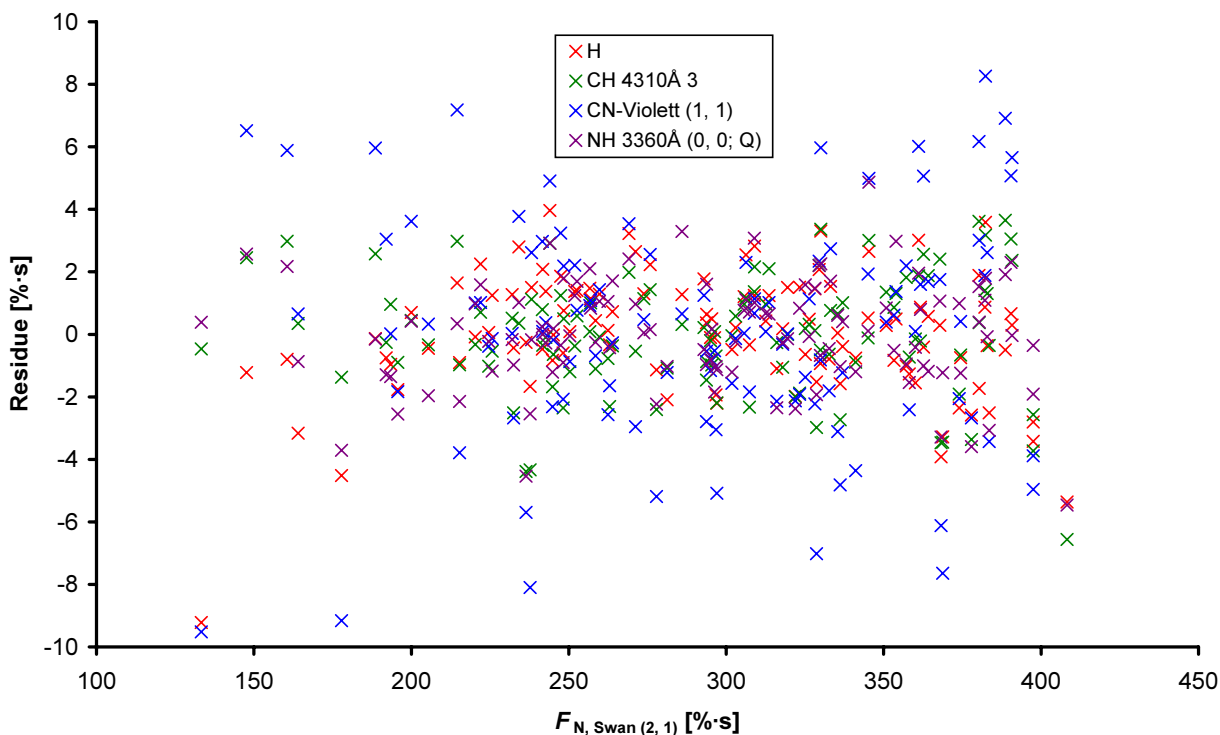


Abbildung 10.3-6: Die Residuen in Abhängigkeit von  $F_{N, Swan(2,1)}$  basierend auf einer Regression mit den 15% der Anilin-Signale, welche einem Signal mit  $F_{N, Swan(2,1)} = 320\% \cdot s$ ,  $F_{N, OH\ 3064\text{Å}} = 230\% \cdot s$ ,  $O = 16\%$ ,  $B = 1,75\ s$  und  $D = 1,3$  am ähnlichsten sind



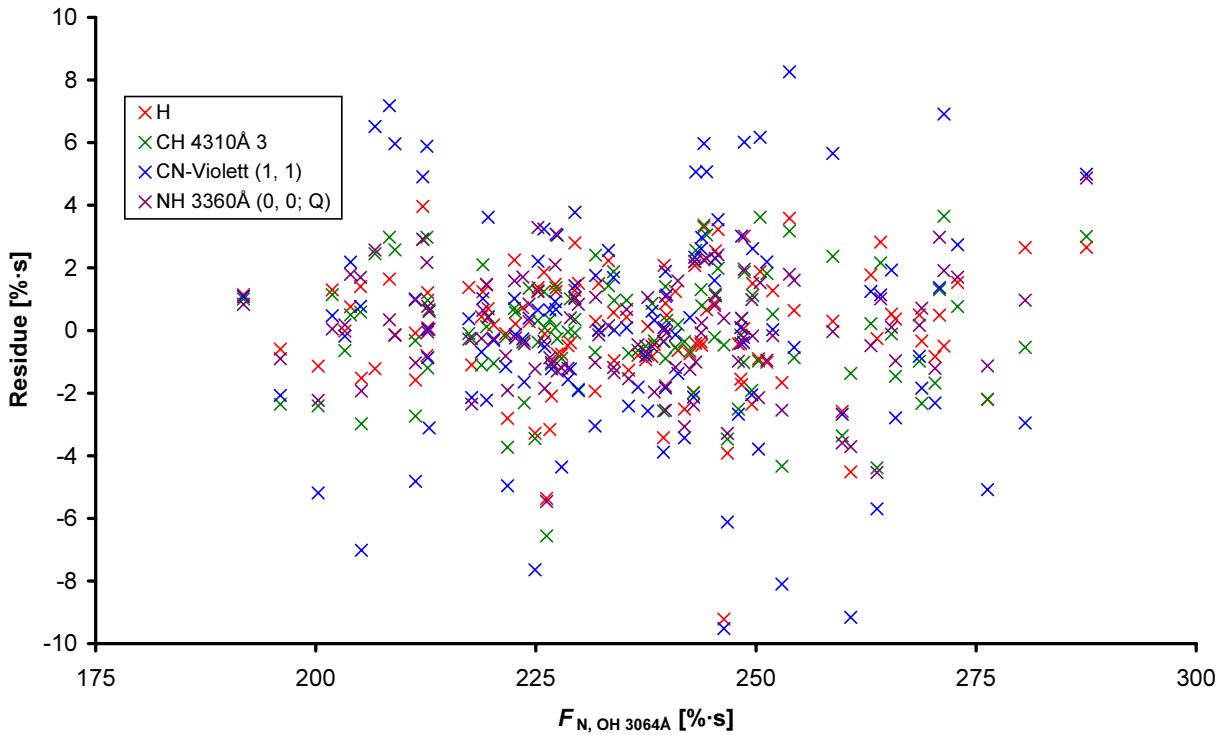


Abbildung 10.3-7: Die Residuen in Abhängigkeit von  $F_{N, OH 3360\text{Å}}$  basierend auf einer Regression mit den 15% der Anilin-Signale, welche einem Signal mit  $F_{N, Swan (2, 1)} = 320\% \cdot s$ ,  $F_{N, OH 3064\text{Å}} = 230\% \cdot s$ ,  $O = 16\%$ ,  $B = 1,75 s$  und  $D = 1,3$  am ähnlichsten sind

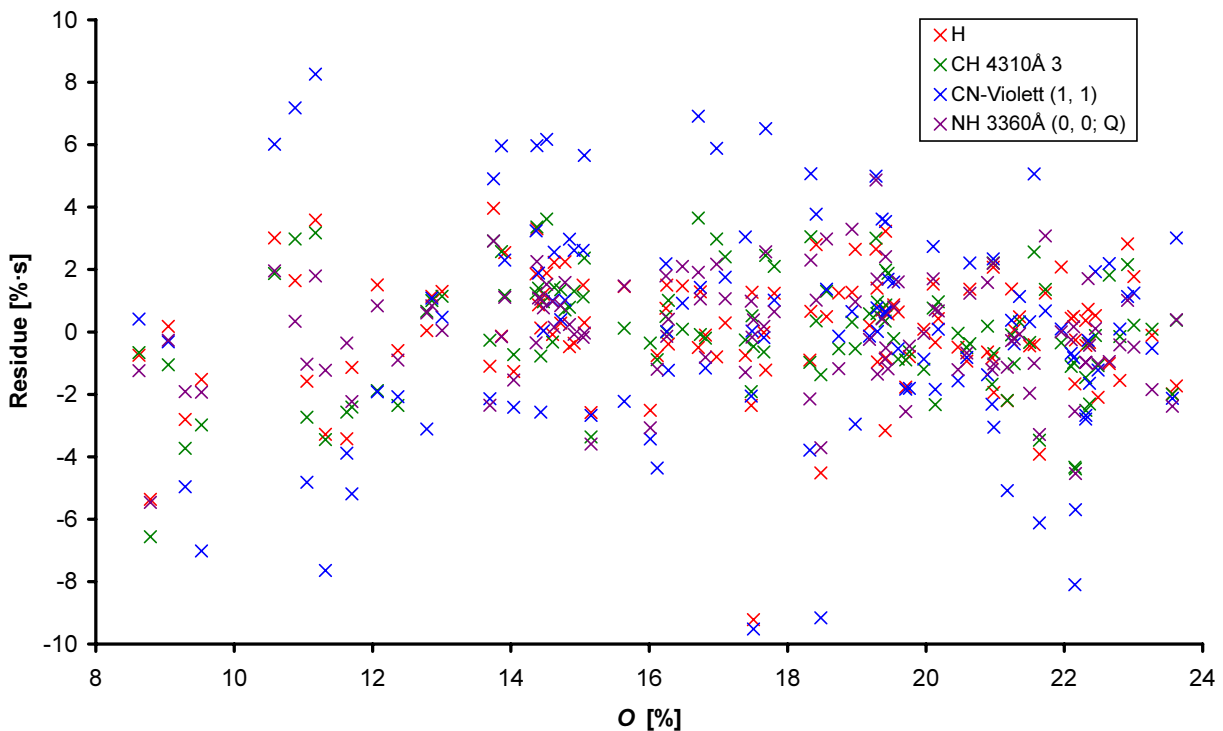


Abbildung 10.3-8: Die Residuen in Abhängigkeit von  $O$  basierend auf einer Regression mit den 15% der Anilin-Signale, welche einem Signal mit  $F_{N, Swan (2, 1)} = 320\% \cdot s$ ,  $F_{N, OH 3064\text{Å}} = 230\% \cdot s$ ,  $O = 16\%$ ,  $B = 1,75 s$  und  $D = 1,3$  am ähnlichsten sind

*Table 10.3-2: Die aus multivariaten Regressionsrechnungen mit dem Anilin-Datensatz resultierenden inneren Bestimmtheitsmaße für den Ansatz mit allen Regressoren und den mit  $F_{N, Swan (2, 1)}$ ,  $F_{N, OH 3064\text{\AA}}$  und  $O$  als Regressoren (Variablenreduktion basierend auf dem Regressorset in Gleichung (10.3-3); 15% aller Anilin-Messwerte)*

$F_{N, Swan (2,1)} [ \% \cdot s ]$	$F_{N, OH 3064\text{\AA}} [ \% \cdot s ]$	$O [ \% ]$	$B [ s ]$	$D$
0,502	0,905	0,815	0,885	0,447
0,181	0,284	0,259		

*Table 10.3-3: Die maximale, mittlere und minimale Residuenvarianz aller Teilansätze aus 15% der im Anilin-Datensatz gespeicherten Messungen in Abhängigkeit von der Bande bzw. Linie, deren Fläche mit allen fünf Regressoren vorhergesagt wird (Variablenreduktion basierend auf dem Regressorset in Gleichung (10.3-3))*

Bande/Linie	Max. Varianz [[%·s) <sup>2</sup> ]	Min. Varianz [[%·s) <sup>2</sup> ]	Varianz Mittelwert [[%·s) <sup>2</sup> ]
H	8,85	0,67	2,52
CH4310\text{\AA} 3	15,55	0,34	2,37
CN Violett (1, 1)	51,87	0,10	7,89
NH 3360\text{\AA} (0, 0; Q)	15,39	0,34	2,73

*Table 10.3-4: Die maximale, mittlere und minimale Residuenvarianz aller Teilansätze aus 15% der im Anilin-Datensatz gespeicherten Messungen in Abhängigkeit von der Bande bzw. Linie, deren Fläche mit Hilfe der Regressoren  $F_{N, Swan (2, 1)}$ ,  $F_{N, OH 3064\text{\AA}}$  und  $O$  vorhergesagt wird (Variablenreduktion basierend auf dem Regressorset in Gleichung (10.3-3) ohne  $B$  und  $D$ )*

Bande/Linie	Max. Varianz [[%·s) <sup>2</sup> ]	Min. Varianz [[%·s) <sup>2</sup> ]	Varianz Mittelwert [[%·s) <sup>2</sup> ]
H	9,50	0,90	2,79
CH4310\text{\AA} 3	9,59	0,35	2,05
CN Violett (1, 1)	49,68	1,92	10,29
NH 3360\text{\AA} (0, 0; Q)	9,14	0,27	2,55

Tabelle 10.3-5: Die aus multivariaten Regressionsrechnungen mit einem Teilansatz aus 15% der im Anilin-Datensatz gespeicherten Signale (Variablenreduktion basierend auf dem Regressorset in Gleichung (10.3-3)) resultierende Residuenvarianz in Abhängigkeit davon, welche Klassifizierungsparameter als Regressoren eingesetzt worden sind

Regressoren (, "+" = verwendet, "-" = nicht verwendet)					Residuenvarianz $[(\% \cdot s)^2]$ bezüglich der Fläche der Signale von:			
$F_{N, \text{Swan}}(2,1)$	$F_{N, \text{OH } 3064\text{Å}}$	<i>O</i>	<i>B</i>	<i>D</i>	H	CH 4310Å 3	CN- Violett (1, 1)	NH 3360Å (0, 0; Q)
+	+	+	+	+	3,92	3,68	13,78	3,01
+	+	+	+	-	3,97	3,68	14,53	3,01
+	+	+	-	+	4,35	3,87	18,81	3,02
+	+	+	-	-	4,40	3,87	19,59	3,01
+	-	+	+	+	8,07	4,13	14,68	3,56
+	-	+	+	-	8,51	4,21	15,90	3,63
+	+	-	+	+	5,55	7,63	19,80	5,02
+	+	-	+	-	5,50	7,73	19,66	5,05
+	-	+	-	+	43,23	10,81	70,45	7,63
+	+	-	-	+	5,87	11,32	19,82	7,72
+	+	-	-	-	5,99	13,32	19,86	8,98
+	-	+	-	-	59,44	13,97	99,26	9,71
+	-	-	+	+	52,89	36,01	67,76	24,06
+	-	-	+	-	53,42	36,94	67,63	24,58
+	-	-	-	-	88,89	41,09	129,84	27,11
+	-	-	-	+	84,88	41,44	119,18	27,35
-	+	+	-	+	71,91	566,04	4473,51	309,04
-	+	+	+	+	72,43	571,07	4514,52	311,66
-	+	-	+	+	74,89	613,62	4993,64	335,31
-	-	-	+	+	112,18	618,40	4949,40	340,14
-	-	+	+	+	92,13	623,80	4891,42	343,23
-	+	-	-	+	87,15	718,12	6301,39	388,40
-	+	+	-	-	109,36	909,14	7135,80	495,43
-	+	+	+	-	110,35	915,56	7192,50	498,56
-	+	-	+	-	110,69	933,05	7440,59	508,46
-	-	-	+	-	155,61	950,04	7408,96	521,08
-	-	+	+	-	127,49	954,48	7465,40	522,34
-	+	-	-	-	114,06	960,46	7975,75	521,24
-	-	+	-	+	251,69	960,68	7568,00	531,30
-	-	-	-	+	253,88	969,56	7928,43	534,14
-	-	+	-	-	258,64	1154,53	9065,99	635,17

Tabelle 10.3-6: Veränderung der Residuenvarianz bei unterschiedlicher Gewichtung der Regressoren für 15%-Teilansätze aus dem Anilin-Datensatz (Variablenreduktion basierend auf dem Regressorset in Gleichung (10.3-3)) wobei die Gewichtungsfaktoren für **B** und **D** auf Null gesetzt worden sind

Gewichtungsfaktor $g$ für			Residuenvarianz bezüglich $F_{N, H} [(\% \cdot s)^2]$		
$F_{N, Swan (2,1)}$	$F_{N, OH 3064\text{Å}}$	$O$	Max. Varianz	Min. Varianz	Mittlere Varianz
1,0	1,0	1,0	9,50	0,90	2,79
0,5	1,0	1,0	13,65	0,93	4,61
1,0	0,5	1,0	7,88	0,99	2,68
1,0	1,0	0,5	5,78	0,67	2,21
0,5	0,5	1,0	13,26	1,26	3,99
<b>1,0</b>	<b>0,5</b>	<b>0,5</b>	<b>5,87</b>	<b>0,61</b>	<b>2,12</b>
0,5	1,0	0,5	8,46	0,76	3,51
Gewichtungsfaktor $g$ für			Residuenvarianz bezüglich $F_{N, CH4310\text{Å} 3} [(\% \cdot s)^2]$		
$F_{N, Swan (2,1)}$	$F_{N, OH 3064\text{Å}}$	$O$	Max. Varianz	Min. Varianz	Mittlere Varianz
1,0	1,0	1,0	9,59	0,35	2,05
0,5	1,0	1,0	10,03	0,50	3,04
1,0	0,5	1,0	9,04	0,39	2,05
1,0	1,0	0,5	7,52	0,30	1,94
0,5	0,5	1,0	9,55	0,76	2,97
<b>1,0</b>	<b>0,5</b>	<b>0,5</b>	<b>9,11</b>	<b>0,27</b>	<b>1,86</b>
0,5	1,0	0,5	7,89	0,36	2,27
Gewichtungsfaktor $g$ für			Residuenvarianz bezüglich $F_{N, CN\text{ Violett} (1, 1)} [(\% \cdot s)^2]$		
$F_{N, Swan (2,1)}$	$F_{N, OH 3064\text{Å}}$	$O$	Max. Varianz	Min. Varianz	Mittlere Varianz
1,0	1,0	1,0	49,68	1,92	10,29
0,5	1,0	1,0	60,33	2,02	12,96
1,0	0,5	1,0	53,14	2,32	12,38
<b>1,0</b>	<b>1,0</b>	<b>0,5</b>	<b>24,59</b>	<b>0,99</b>	<b>7,59</b>
0,5	0,5	1,0	65,38	3,54	15,33
<b>1,0</b>	<b>0,5</b>	<b>0,5</b>	<b>36,52</b>	<b>1,20</b>	<b>8,90</b>
0,5	1,0	0,5	32,19	1,33	8,92
Gewichtungsfaktor $g$ für			Residuenvarianz bezüglich $F_{N, NH 3360\text{Å} (0, 0; Q)} [(\% \cdot s)^2]$		
$F_{N, Swan (2,1)}$	$F_{N, OH 3064\text{Å}}$	$O$	Max. Varianz	Min. Varianz	Mittlere Varianz
1,0	1,0	1,0	9,14	0,27	2,55
0,5	1,0	1,0	10,59	0,30	3,92
1,0	0,5	1,0	9,29	0,36	2,65
1,0	1,0	0,5	8,44	0,23	2,26
0,5	0,5	1,0	11,20	0,49	3,82
<b>1,0</b>	<b>0,5</b>	<b>0,5</b>	<b>8,77</b>	<b>0,22</b>	<b>2,17</b>
0,5	1,0	0,5	8,63	0,26	2,88

### 10.3.3 Fehlerabschätzung

Bei optimaler Kalibration durch den externen Standard kann damit gerechnet werden, dass die durch die lineare Regression berechneten Residuen normalverteilt sind. In diesem Fall ist es mittels t-Statistik möglich die Vertrauensintervalle für die Schätzung des wahren Wertes oder die Wahrscheinlichkeit, bei einer Wiederholungsmessung die Messergebnisse in einem bestimmten Intervall wiederzufinden, zu ermitteln (siehe Abschnitt 4.7.4). Treten jedoch Kalibrationsfehler auf, so kann nicht mehr von normalverteilten Residuen ausgegangen werden. Alle Anilin-Messwerte in der Datenbank wurden beispielsweise am 22.09.2001 aufgezeichnet. Sie wurden mit ein und dem selben externen Standard kalibriert. Die äußeren Umweltbedingungen änderten sich während der Messung kaum. Teilt man die Anilin-Messungen in solche, die zu einem früheren und solche die zu einem späteren Zeitpunkt aufgezeichnet worden sind, so erhält man zwei Messreihen. Diese Messreihen können separat in die Datenbank geladen werden, so dass zwei Datensätze erhalten werden. Auf Basis der Klassifizierungsparameter eines der Datensätze kann nun eine Regressionsanalyse mit beiden Datensätzen durchgeführt werden. Wie aus Abbildung 10.3-9 ersichtlich wird, unterscheiden sich die Schätzwerte für die Fläche des CH 4310Å 3-Signals für beide Datensätze kaum. Angesichts der Tatsache, dass sich die Umweltbedingungen während des Messzeitraum nur unmerklich geändert haben und beide Messreihen auf ein und der selben Kalibration basieren, ist ein solches Ergebnis zu erwarten. Geht man analog zu den Anilin-Messdaten mit den Daten, die bei der Messung von Diethylether am 25. und 27. Juni 2001 gewonnen worden sind, vor, so erhält man die in Abbildung 10.3-10 visualisierten Regressionsergebnisse. Die Messdaten basieren auf zwei verschiedenen Messreihen, die bei anderen Umweltbedingungen aufgezeichnet worden sind, sowie mit einer anderen Messung des externen Standards kalibriert worden sind. Die große Mehrheit der Ergebnisse, welche mit den am 27 Juni aufgezeichneten Messwerten abgeschätzt worden sind, ist höher als die auf Basis der am 25 Juni aufgezeichneten Messwerte ermittelten Schätzergebnisse. Hier liegt ein Kalibrationsfehler vor. Werden die Messwerte von Anilin zu einem Datensatz zusammengefasst, so sind die bei einer Regressionsrechnung erhaltenen Residuen normalverteilt. Fasst man die Diethylethermesswerte zu einem Datensatz zusammen, so sind bei einer Regression die Residuen definitiv nicht normalverteilt, weil sich beide Messreihen unterscheiden lassen. Dieser Sachverhalt zeigt, dass weder von einer optimalen Kalibration noch von normalverteilten Residuen ausgegangen werden darf. Eine Fehlerabschätzung mittels t-Statistik führt definitiv zu falschen Ergebnissen.

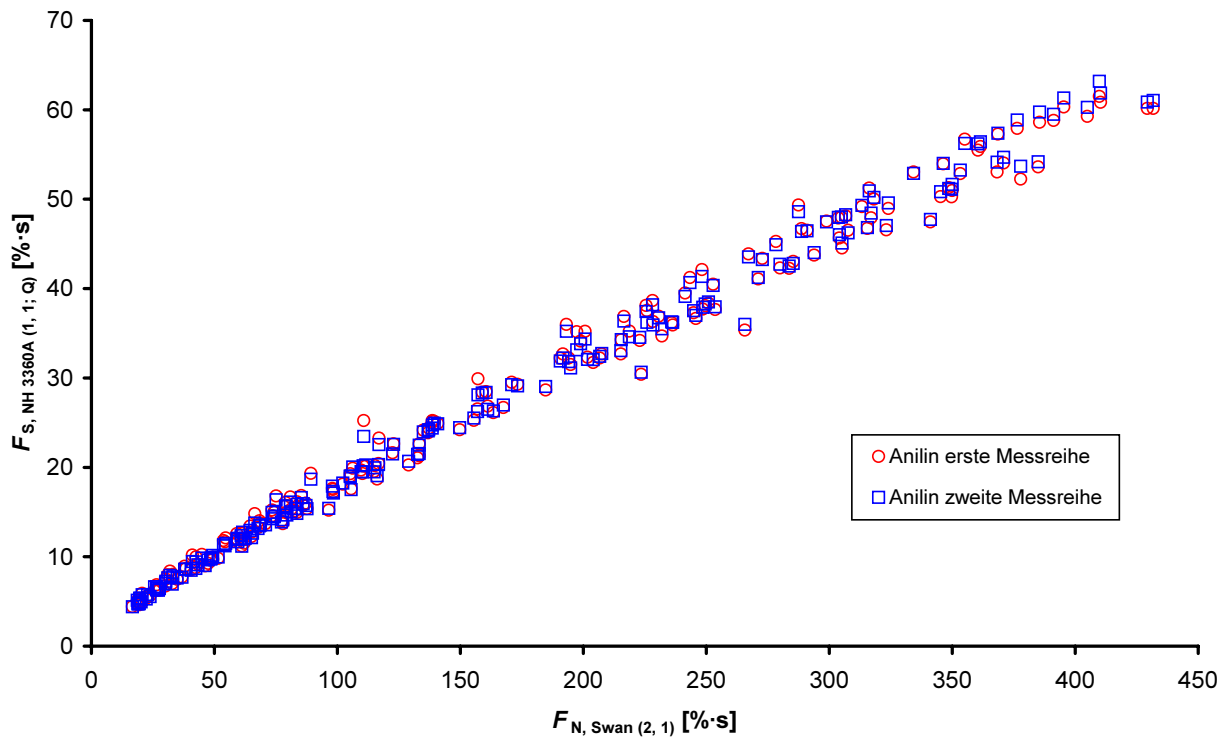


Abbildung 10.3-9:  $F_{N,Swan(2,1)}$  gegen die Schätzwerte von  $F_{N,CH4310A3} = F_{S,CH4310A3}$  auf Basis zweier Anilin-Datensätze, welche zwei verschiedenen Messreihen zu Grunde liegen; die zur Regression herangezogenen Regressoren sind die Klassifizierungsparameter der 2. Messreihe

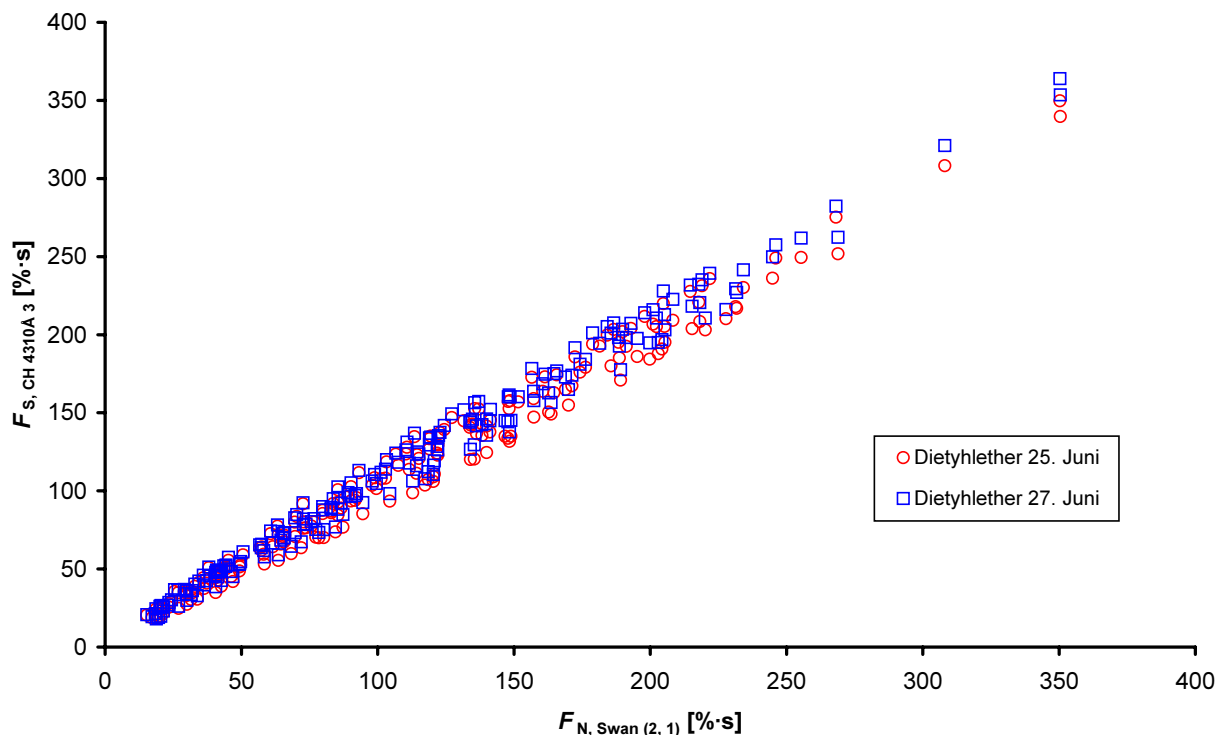


Abbildung 10.3-10:  $F_{N,Swan(2,1)}$  gegen die Schätzwerte von  $F_{N,CH4310A3} = F_{S,CH4310A3}$  auf Basis zweier Diethylether-Datensätze basierend auf Messungen aufgezeichnet am 25. und 27. Juni 2001; die zur Regression herangezogenen Regressoren sind die Klassifizierungsparameter der Messreihe aufgezeichnet am 27. Juni

Die Datensätze in der Datenbank haben demnach folgende Eigenschaften:

1. Varianzhomogenität, insbesondere bezüglich des Zeitpunktes der Aufzeichnung der Daten, kann nicht gewährleistet werden. Eine weitere Quelle von Varianzhomogenität kann die Abweichung des linearen Modells von der Wirklichkeit sein.
2. Für jeden Datensatz existiert eine hohe Anzahl an Messwerten.

Angesichts dieses Sachverhaltes bietet sich eine Fehlerabschätzung mittels einer robusten Methode an. Basis einer solchen Schätzung kann die Differenz zwischen der tatsächlich gemessenen und der mittels Regressionsrechnung abgeschätzten Signalfläche sein:

$$\Delta F_{N,X} = F_{N,X} - F_{S,X}$$

mit

$$\Delta F_{N,X} = \text{Differenz von gemessener und geschätzter Signalfläche,}$$

$$F_{N,X} = \text{gemessene normierte Signalfläche und} \quad (10.3-6)$$

$$F_{S,X} = \text{mittels linearer Regression geschätzte normierte Signalfläche wobei } \mathbf{X}$$

für die Bezeichnung der Bande bzw. Linie steht, auf der das zugehörige Einzelchromatogramm basiert.

Diese Differenz ist für jedes Signal in der Datenbank berechnet worden. Auf diese Weise erhält man für jede abzuschätzende Signalfläche 21400 Werte für alle Messungen nicht stickstoffhaltiger Verbindungen und 5023 für die stickstoffhaltigen Verbindungen. Diese Anzahl an Signalen ist hoch genug für eine repräsentative Abschätzung des Fehlers, welcher bei einer Schätzung mittels multivariater linearer Regression zu erwarten ist. Für die ermittelten  $\Delta F_{N,X}$ -Werte können die Quartil- und Interdecilabstände  $I_{50\%}$  bzw.  $I_{80\%}$  ermittelt werden. Analog können die Intervalle ermittelt werden, in dem sich 95% aller Signale befinden und so die mit  $I_{95\%}$  bezeichneten Abstände berechnet werden (siehe Abschnitt 4.1.3). Die berechneten Intervallabstände sind für die CHO-Datenbank in Tabelle 10.3-7 aufgelistet und die für die CHON-Datenbank in Tabelle 10.3-10. Eine weitere legitime Vorgehensweise ist, nicht die Streuung der  $\Delta F_{N,X}$ -Werte für die gesamte Datenbank, sondern für jede Substanz einzeln zu untersuchen. Es ergeben sich für bestimmte Substanzen größere Intervalle als bei der Betrachtung aller  $\Delta F_{N,X}$ -Werte einer Datenbank. Die größte Streuung ist in der CHO-Datenbank bei Benzen zu finden.

Zum Vergleich sind sowohl für die gesamte CHO-Datenbank (siehe Tabelle 10.3-8) als auch für den Ethanol Datensatz (siehe Tabelle 10.3-9) die Vertrauensintervalle der  $\Delta F_{N,X}$ -Werte nach der t-Statistik ermittelt worden. Läge eine Normalverteilung der  $\Delta F_{N,X}$ -Werte vor, so müssten die Spannweite der Vertrauensintervalle für ein Wahrscheinlichkeitsniveau von  $p$  mit den entsprechenden  $I_p$ -Wert näherungsweise übereinstimmen. Dies gilt aber nur für die  $I_{95\%}$ -Werte, während die Interdecil- und Quartilintervalle  $I_{80\%}$  und  $I_{50\%}$  erheblich kleiner sind als die Spannweite der Vertrauensintervalle. Dies ist ein weiterer Beleg dafür, dass die Streuung der  $\Delta F_{N,X}$ -Werte nicht normalverteilt ist. Die  $I_{95\%}$ -Intervalle stimmen nur zufällig mit der

Spannweite der Vertrauensintervalle überein. Insbesondere die Abschätzung des wahren Wertes für einen Regressanden nach Gleichung (4.7-16) (Abschnitt 4.7.4) wäre grob fehlerhaft.

Die Streuung der  $\Delta F_{N,X}$ -Werte kann folgendermaßen bewertet werden: Wird eine Signalfläche mittels linearer Regression abgeschätzt, so kann mit einer Wahrscheinlichkeit von etwa 95% davon ausgegangen werden, dass die Differenz zwischen wahren Wert und Schätzwert  $\Delta F_{N,X}$  innerhalb der durch  $I_{U,95\%}$  und  $I_{O,95\%}$  definierten Intervallgrenzen liegt. Will man wissen, ob das Ergebnis einer Schätzung und die tatsächlich gemessene Signalfläche sich so weit unterscheiden, dass die Substanz, auf deren Datensatz die Schätzung basiert, mit hoher Wahrscheinlichkeit nicht Urheber des Signals sein kann, so müssen sich Schätzwert und gemessener Wert um den geschätzten maximalen Messfehler  $I_{95\%}$  unterscheiden. Es existieren zwei Methoden  $I_{95\%}$  zu bestimmen. Die eine basiert auf dem gemeinsamen Pool von  $\Delta F_{N,X}$ -Werten aller Datensätze einer Datenbank, die andere auf den  $\Delta F_{N,X}$ -Werten der Substanz, bei der die Diskrepanz zwischen gemessener und geschätzter Signalfläche am größten ist. Es ist nicht auszuschließen, dass das Ausmaß dieser Streuung auch von der Art der Substanz abhängt. So fällt auf, dass vor allem bei aromatischen Verbindungen wie Benzen, Ethylbenzen, Anilin, Picolin und Pyridin ein recht hoher Messfehler zu beobachten ist. Daher ist eine vorsichtiger Fehlerabschätzung auf Basis der bei einem Einzeldatensatz gefundenen maximalen Streuung die bessere (siehe Tabelle 10.3-11 bis Tabelle 10.3-14). Aus diesem Grund werden mit Hilfe dieser Fehlerabschätzung folgende Signifikanzniveaus festgelegt: Unterscheiden sich Schätzwert und tatsächlich gemessene Signalfläche um mehr als  $I_{95\%}$ , so unterscheidet sich die Substanz, deren Signal geschätzt worden ist **signifikant** von der Substanz, auf der der Datensatz basiert. Aufgrund analoger Überlegungen kann auf Basis von  $I_{80\%}$  eine Definition des Signifikanzniveaus **wahrscheinlich** vorgenommen werden. Unterscheiden sich Schätzwert und tatsächlich gemessener Wert um mehr als  $I_{80\%}$ , so ist die Substanz, deren Signalfläche abgeschätzt worden ist, **wahrscheinlich** nicht die Substanz, auf der der zur Schätzung herangezogene Datensatz basiert.

Die Bestimmung  $I_{80\%}$  und  $I_{95\%}$  erfolgte ohne Berücksichtigung der Klassifizierungsparameter der einzelnen Messungen. Diese Fehlerintervalle gelten für ein Signal mit durchschnittlich großen Klassifizierungsparametern. Die Fehlerintervalle sollten daher als durchschnittliche maximale Fehlererwartungswerte für die Analysenmethode interpretiert werden. Soll ein bestimmter Einzelmesswert mit dem Schätzwert verglichen werden, so muss die Abhängigkeit von  $\Delta F_{N,X}$  von den Klassifizierungsparametern berücksichtigt werden. Anhand von Beispielrechnungen mit dem Benzen-Datensatz wird deutlich, dass  $\Delta F_{N,X}$  zwar vom Offset des OH 3064Å-Signals weitgehend unabhängig ist (siehe Abbildung 10.3-13), jedoch mit steigenden  $F_{N,Swan(2,1)}$ -Werten vom Betrag her größer wird (siehe Abbildung 10.3-11). Da  $F_{N,Swan(2,1)}$  und  $F_{N,OH\ 3064\text{Å}}$  interkorrelieren (siehe Abbildung 10.3-14), werden die  $\Delta F_{N,X}$ -Beträge auch mit steigender Fläche  $F_{N,OH\ 3064\text{Å}}$  höher (siehe Abbildung 10.3-12). Teilt man den Benzen-Datensatz in drei Teile, von denen der erste Teil alle Messwerte mit  $F_{N,Swan(2,1)} < 200$ , der zweite alle Messwerte mit  $F_{N,Swan(2,1)}$  zwischen 200 und 400 und der dritte alle Messwerte mit  $F_{N,Swan(2,1)} > 400$  enthält, so erhält man für jede dieser Teildatenbanken unterschiedliche  $I_{95\%}$ -Fehlerintervalle. Diese Intervalle werden um so größer,



je größer  $F_{N, \text{Swan}(2, 1)}$  ist (siehe Abbildung 10.3-15). In Relation zu  $F_{N, \text{Swan}(2, 1)}$  werden die Fehlerintervalle jedoch kleiner (siehe Abbildung 10.3-16).

Die Fehlerintervalle  $I_{80\%}$  und  $I_{95\%}$  lassen sich zur Beantwortung der Frage nutzen, ob sich zwei Substanzen grundsätzlich mit dem GC-MIP-MED voneinander unterscheiden lassen. Zu diesem Zweck müssen zwei Datensätze miteinander verglichen werden. Ein solcher Vergleich kann auf Basis von  $\Delta F_{N, X}$ -Werten durchgeführt werden, wobei die  $\Delta F_{N, X}$ -Werte auf der Abschätzung der Signalflächen des einen Datensatzes mit Hilfe der Klassifizierungsparameter des anderen Datensatzes beruhen. Auf diese Weise können zwei Substanzen miteinander verglichen werden (siehe Gleichung (10.3-6)):

$$\Delta F_{N, X, S1S2} = F_{N, X, S1} - F_{S, X, S1S2} \quad \text{bzw.} \quad \Delta F_{N, X, S2S1} = F_{N, X, S2} - F_{S, X, S2S1}$$

wobei

$$\begin{aligned} F_{N, X, S1} &= \text{mit Substanz 1 (S1) gemessene Signalfläche der Bande X,} \\ F_{N, X, S2} &= \text{mit Substanz 2 (S2) gemessene Signalfläche der Bande X,} \\ F_{S, X, S1S2} &= \text{mit Hilfe des Datensatzes von Substanz 2 per Regressions-} \\ &\quad \text{rechnung geschätzte Signalfläche der Bande X auf Basis der} \\ &\quad \text{zu } F_{N, X, S1} \text{ gehörenden Klassifizierungsparameter.} \\ F_{S, X, S2S1} &= \text{mit Hilfe des Datensatzes von Substanz 1 per Regressions-} \\ &\quad \text{rechnung geschätzte Signalfläche der Bande X auf Basis der} \\ &\quad \text{zu } F_{N, X, S2} \text{ gehörenden Klassifizierungsparameter.} \end{aligned} \quad (10.3-7)$$

Man erhält also pro Klassifizierungsparametersatz für jede Linie bzw. Bande zwei  $\Delta F_{N, X}$ -Werte. D. h. für jede Substanz liegt pro gemessener Bande bzw. Linie ein Satz von  $\Delta F_{N, X, S1S2}$ - bzw.  $\Delta F_{N, X, S2S1}$ -Werten vor. Lassen sich beide Substanzen anhand ihrer Messdatensätze unterscheiden, so ist zu erwarten, dass für mindestens eine Bande oder Linie sowohl mehr als 5% aller  $\Delta F_{N, X, S1S2}$ -Werte als auch mehr als 5% aller  $\Delta F_{N, X, S2S1}$ -Werte größer als  $I_{95\%}$  oder kleiner als  $-I_{95\%}$  sind. Anders formuliert: Sowohl bei mehr als 5% aller  $F_{N, X, S1S2}$ -Werte als auch bei mehr als 5% aller  $F_{N, X, S2S1}$ -Werte sollte ein **signifikanter** Unterschied zwischen Messwert und Schätzwert vorliegen. Um diesen Sachverhalt mathematisch zu beschreiben werden zunächst folgende Größen definiert:

$$\begin{aligned} A_{X, S1-S2} &= \text{Anteil aller } \Delta F_{N, X, S1S2} \text{ - Werte, für die gilt :} \\ &\quad \Delta F_{N, X, S1S2} > I_{95\%} \vee \Delta F_{N, X, S1S2} < -I_{95\%} \cdot \\ A_{X, S2-S1} &= \text{Anteil aller } \Delta F_{N, X, S2S1} \text{ - Werte, für die gilt :} \\ &\quad \Delta F_{N, X, S2S1} > I_{95\%} \vee \Delta F_{N, X, S2S1} < -I_{95\%} \cdot \\ A_{X, S1-S2, (+)} &= \text{Anteil aller } \Delta F_{N, X, S1S2} \text{ - Werte, für die gilt : } \Delta F_{N, X, S1S2} > I_{95\%} \cdot \\ A_{X, S1-S2, (-)} &= \text{Anteil aller } \Delta F_{N, X, S1S2} \text{ - Werte, für die gilt : } \Delta F_{N, X, S1S2} < -I_{95\%} \cdot \\ A_{X, S2-S1, (+)} &= \text{Anteil aller } \Delta F_{N, X, S2S1} \text{ - Werte, für die gilt : } \Delta F_{N, X, S2S1} > I_{95\%} \cdot \\ A_{X, S2-S1, (-)} &= \text{Anteil aller } \Delta F_{N, X, S2S1} \text{ - Werte, für die gilt : } \Delta F_{N, X, S2S1} < -I_{95\%} \cdot \end{aligned} \quad (10.3-8)$$

Auf Basis von  $A_{X, S1-S2}$  und  $A_{X, S2-S1}$  kann nun definiert werden, wann ein **signifikanter** Unterschied zwischen zwei Substanzen S1 und S2 vorliegt:

Gilt für eine Bande oder Linie X :

$$\left(A_{X, S1-S2} > 5\%\right) \wedge \left(A_{X, S2-S1} > 5\%\right) \Rightarrow \text{die Substanzen S1 und S2 unter -} \quad (10.3-9)$$

scheiden sich **signifikant**.

Die Ergebnisse dieses Tests auf Unterscheidbarkeit sollte allerdings bei sehr spektraldichteschwachen Banden, wie der Swan (1, 1)-Bande oder der CH 3900Å-Bande kritisch beurteilt werden. Schon bei der Abschätzung des dynamischen Messbereiches des Detektors hat sich die Spektraldichtebestimmung dieser beiden Banden als sehr fehlerträchtig erwiesen (siehe Abschnitt 9.4.1, insbesondere Abbildung 9.4-3). Die  $\Delta F_{N, X}$ -Werte sind für beide Banden ebenfalls überdurchschnittlich hoch, wenn man sie mit denen anderer CC- bzw. CH-Banden vergleicht. Auch ist bei diesen Banden ein beträchtlicher Kalibrationsfehler nicht auszuschließen, da die Messung des externen Standards fehlerbehaftet sein kann.

Bei der Ermittlung der Messwerte sind für jede Substanz mindestens etwa 600 Messwerte aufgezeichnet worden. Es ist aber nur wenige Male kalibriert worden. Daher ist die Fehlerabschätzung für Signale mit sehr hohen Kalibrationsfehlern auf Grund der geringen Zahl an Wiederholungskalibrationen schlechter als die für Signale mit geringeren Kalibrationsfehlern.

Liegt ein signifikanter Unterschied zwischen zwei Datensätzen der Datenbank vor, so kann es sinnvoll sein, die  $A_{X, S1-S2}$ - bzw.  $A_{X, S2-S1}$ - Werte nach  $A_{X, S1-S2, (+)}$ - und  $A_{X, S1-S2, (-)}$  bzw.  $A_{X, S2-S1, (+)}$  und  $A_{X, S2-S1, (-)}$ -Werten aufzuschlüsseln. Es gilt:

$$\begin{aligned} & \left(A_{X, S1-S2, (+)} > A_{X, S1-S2, (-)}\right) \wedge \left(A_{X, S2-S1, (+)} < A_{X, S2-S1, (-)}\right) \\ \Rightarrow & \text{die Signalfäche der Bande bzw. Linie X ist für die Substanz S1} \\ & \text{tendenziell größer als für die Substanz S2.} \\ & \left(A_{X, S1-S2, (+)} < A_{X, S1-S2, (-)}\right) \wedge \left(A_{X, S2-S1, (+)} > A_{X, S2-S1, (-)}\right) \\ \Rightarrow & \text{die Signalfäche der Bande bzw. Linie X ist für die Substanz S1} \\ & \text{tendenziell kleiner als für die Substanz S2.} \end{aligned} \quad (10.3-10)$$

Es kann also festgestellt werden, welche Substanz bei welcher Bande bzw. Linie höherer bzw. niedrigere Signalfächen aufweist.

Tabelle 10.3-7: Intervalle, in denen sich 80% bzw. 95% aller  $\Delta F_{N,X}$ -Werte der CHO-Datenbank befinden

	$\Delta F_{N, \text{Swan (1, 0)}} [‰\cdot\text{s}]$	$\Delta F_{N, \text{Swan (3, 2)}} [‰\cdot\text{s}]$	$\Delta F_{N, \text{Swan (0, 0)}} [‰\cdot\text{s}]$	$\Delta F_{N, \text{Swan (1, 1)}} [‰\cdot\text{s}]$
Median	0,105	-0,001	0,103	0,114
$I_{U, 80\%}$	-0,63	-0,66	-0,92	-1,50
$I_{O, 80\%}$	0,88	0,64	1,22	1,94
$I_{80\%}$	1,51	1,30	2,14	3,44
$I_{U, 95\%}$	-1,30	-1,14	-2,19	-3,17
$I_{O, 95\%}$	1,46	1,14	2,44	3,80
$I_{95\%}$	2,76	2,28	4,63	6,97
	$\Delta F_{N, H} [‰\cdot\text{s}]$	$\Delta F_{N, CH 4310A 2} [‰\cdot\text{s}]$	$\Delta F_{N, CH 4310A 3} [‰\cdot\text{s}]$	$\Delta F_{N, CH 3900A} [‰\cdot\text{s}]$
Median	-0,524	-0,062	-0,054	0,095
$I_{U, 80\%}$	-4,56	-1,72	-1,68	-2,52
$I_{O, 80\%}$	3,30	1,54	1,52	2,80
$I_{80\%}$	7,86	3,25	3,19	5,32
$I_{U, 95\%}$	-8,15	-3,30	-3,28	-4,64
$I_{O, 95\%}$	7,75	3,44	3,37	5,06
$I_{95\%}$	15,90	6,74	6,65	9,70

Tabelle 10.3-8: Mittelwert und Standardabweichung von  $\Delta F_{N,X}$  sowie Vertrauensintervalle für die Wahrscheinlichkeitsniveaus 50%, 80% und 95% und die zugehörigen  $I_{50\%}$ ,  $I_{80\%}$  sowie  $I_{95\%}$ -Werte aus den Messwerten der CHO-Datenbank

	$\Delta F_{N, \text{Swan (1, 0)}} [‰\cdot\text{s}]$	$\Delta F_{N, \text{Swan (3, 2)}} [‰\cdot\text{s}]$	$\Delta F_{N, \text{Swan (0, 0)}} [‰\cdot\text{s}]$	$\Delta F_{N, \text{Swan (1, 1)}} [‰\cdot\text{s}]$
Mittelwert	0,11	0,00	0,12	0,16
Standardabweichung	0,70	0,56	1,23	1,77
$t_{\infty, 80\%} (t_{\infty, 80\%}\cdot 2)$	$\pm 0,89 (1,78)$	$\pm 0,72 (1,45)$	$\pm 1,58 (3,15)$	$\pm 2,27 (4,54)$
$t_{\infty, 95\%} (t_{\infty, 95\%}\cdot 2)$	$\pm 1,36 (2,73)$	$\pm 1,11 (2,21)$	$\pm 2,41 (4,82)$	$\pm 3,47 (6,94)$
	$\Delta F_{N, H} [‰\cdot\text{s}]$	$\Delta F_{N, CH 4310A 2} [‰\cdot\text{s}]$	$\Delta F_{N, CH 4310A 3} [‰\cdot\text{s}]$	$\Delta F_{N, CH 3900A} [‰\cdot\text{s}]$
Mittelwert	-0,50	-0,05	-0,05	0,11
Standardabweichung	3,91	1,62	1,59	2,45
$t_{\infty, 80\%} (t_{\infty, 80\%}\cdot 2)$	$\pm 5,01 (10,02)$	$\pm 2,07 (4,15)$	$\pm 2,03 (4,07)$	$\pm 3,14 (6,28)$
$t_{\infty, 95\%} (t_{\infty, 95\%}\cdot 2)$	$\pm 7,66 (15,32)$	$\pm 3,17 (6,34)$	$\pm 3,11 (6,22)$	$\pm 4,80 (9,60)$

Tabelle 10.3-9: Mittelwert und Standardabweichung von  $\Delta F_{N,X}$  sowie Vertrauensintervalle für die Wahrscheinlichkeitsniveaus 50%, 80% und 95% und  $I_{50\%}$ ,  $I_{80\%}$  sowie  $I_{95\%}$ -Werte aus den Ethanol-Messwerten des CHO-Datensatzes

	$\Delta F_{N, \text{Swan (1, 0)}}$ [%·s]	$\Delta F_{N, \text{Swan (3, 2)}}$ [%·s]	$\Delta F_{N, \text{Swan (0, 0)}}$ [%·s]	$\Delta F_{N, \text{Swan (1, 1)}}$ [%·s]
Mittelwert	-0,07	-0,02	-0,08	0,03
Standardabweichung	1,01	0,73	2,17	1,99
$t_{\infty, 50\%}$	±0,68	±0,49	±1,47	±1,34
$t_{\infty, 80\%}$	±1,29	±0,94	±2,79	±2,55
$t_{\infty, 95\%}$	±1,98	±1,43	±4,26	±3,90
$t_{\infty, 50\%} \cdot 2$	1,36	0,98	2,93	2,68
$I_{50\%}$	0,90	0,78	1,50	2,09
$t_{\infty, 80\%} \cdot 2$	2,59	1,87	5,57	5,10
$I_{80\%}$	2,11	1,68	3,97	4,67
$t_{\infty, 95\%} \cdot 2$	3,96	2,86	8,53	7,80
$I_{95\%}$	4,54	2,92	8,51	8,54
	$\Delta F_{N, H}$ [%·s]	$\Delta F_{N, CH 4310A 2}$ [%·s]	$\Delta F_{N, CH 4310A 3}$ [%·s]	$\Delta F_{N, CH 3900A}$ [%·s]
Mittelwert	0,62	0,07	0,06	0,06
Standardabweichung	6,81	2,40	2,28	4,00
$t_{\infty, 50\%}$ ( $t_{\infty, 50\%} \cdot 2$ )	±4,60	±1,62	±1,54	±2,70
$t_{\infty, 80\%}$ ( $t_{\infty, 80\%} \cdot 2$ )	±8,74	±3,08	±2,93	±5,13
$t_{\infty, 95\%}$ ( $t_{\infty, 95\%} \cdot 2$ )	±13,37	±4,72	±4,48	±7,84
$t_{\infty, 50\%} \cdot 2$	9,19	3,24	3,08	5,39
$I_{50\%}$	7,55	2,03	1,85	3,84
$t_{\infty, 80\%} \cdot 2$	17,47	6,17	5,86	10,25
$I_{80\%}$	15,19	4,36	4,04	8,37
$t_{\infty, 95\%} \cdot 2$	26,74	9,43	8,96	15,69
$I_{95\%}$	26,30	8,95	8,37	16,44

Tabelle 10.3-10: Intervalle, in denen 80% und 95% aller  $\Delta F_{N,X}$ -Werte bei stickstoffhaltigen Verbindungen gefunden werden

	$\Delta F_{N, \text{Swan (1, 0)}}$ [%·s]	$\Delta F_{N, \text{Swan (3, 2)}}$ [%·s]	$\Delta F_{N, \text{Swan (0, 0)}}$ [%·s]	$\Delta F_{N, \text{Swan (1, 1)}}$ [%·s]
$I_{U, 80\%}$	-0,53	-0,56	-0,88	-1,30
$I_{O, 80\%}$	0,73	0,58	1,18	1,47
$I_{80\%}$	1,26	1,14	2,07	2,77
$I_{U, 95\%}$	-0,99	-0,95	-1,67	-2,25
$I_{O, 95\%}$	1,22	0,94	2,14	2,51
$I_{95\%}$	2,21	1,89	3,80	4,76
	$\Delta F_{N, H}$ [%·s]	$\Delta F_{N, CH 4310A 2}$ [%·s]	$\Delta F_{N, CH 4310A 3}$ [%·s]	$\Delta F_{N, CH 3900A}$ [%·s]
$I_{U, 80\%}$	-2,48	-1,17	-1,10	-1,98
$I_{O, 80\%}$	2,11	1,22	1,19	2,50
$I_{80\%}$	4,59	2,38	2,28	4,48
$I_{U, 95\%}$	-4,65	-2,36	-2,27	-3,65
$I_{O, 95\%}$	4,51	2,34	2,38	3,99
$I_{95\%}$	9,15	4,70	4,65	7,64
	$\Delta F_{N, CN \text{ Violett (1, 1)}}$ [%·s]	$\Delta F_{N, CN \text{ Violett (0, 1)}}$ [%·s]	$\Delta F_{N, NH 3360\text{\AA} (0, 0; Q)}$ [%·s]	$\Delta F_{N, NH 3360\text{\AA} (0, 0; Q)}$ [%·s]
$I_{U, 80\%}$	-3,18	-0,92	-2,17	-1,58
$I_{O, 80\%}$	3,34	1,08	1,98	1,39
$I_{80\%}$	6,51	2,00	4,15	2,97
$I_{U, 95\%}$	-6,76	-2,00	-4,49	-4,88
$I_{O, 95\%}$	6,59	2,00	4,22	4,07
$I_{95\%}$	13,35	4,00	8,71	8,96

Tabelle 10.3-11:  $I_{80\%}$ -Intervalle für die Datensätze der Einzelsubstanzen der CHO-Datenbank

Substanz	$\Delta F_{N, \text{Swan}}(1, 0)$ [%·s]	$\Delta F_{N, \text{Swan}}(3, 2)$ [%·s]	$\Delta F_{N, \text{Swan}}(0, 0)$ [%·s]	$\Delta F_{N, \text{Swan}}(1, 1)$ [%·s]
1, 3-Cyclohexadien	1,45	1,39	2,46	3,65
Aceton	1,44	1,17	1,70	2,81
Benzen	2,00	2,24	2,13	8,54
Butadien	1,42	1,26	1,66	2,20
Butan	1,34	1,32	1,77	2,62
Butanol	1,87	1,26	2,31	3,62
Cyclohexan	1,51	1,34	2,00	3,31
Diethylether	1,72	1,23	2,16	4,20
Dimethylether	1,18	0,98	2,01	2,77
Ethan	1,32	1,31	2,15	2,73
Ethanol	2,11	1,68	3,97	4,67
Ethen	1,17	1,03	1,74	2,65
Ethylbenzen	1,53	1,37	2,27	3,31
Hexan	1,85	1,26	2,34	3,86
Isooctan	1,38	1,27	2,08	2,83
Methan	1,08	0,88	1,35	1,92
Methanol	0,95	0,93	1,48	2,66
Octan	1,64	1,49	2,03	3,18
Propan	1,53	1,17	2,26	3,19
Propanal	1,39	1,20	2,95	2,95
Propanol	1,68	1,32	2,79	3,68
Propen	1,35	1,25	1,86	4,24
<b>Maximalwert</b>	<b>2,11</b>	<b>2,24</b>	<b>3,97</b>	<b>8,54</b>
Substanz	$\Delta F_{N, H}$ [%·s]	$\Delta F_{N, CH\ 4310A\ 2}$ [%·s]	$\Delta F_{N, CH\ 4310A\ 3}$ [%·s]	$\Delta F_{N, CH\ 3900A}$ [%·s]
1, 3-Cyclohexadien	6,77	3,87	4,35	8,71
Aceton	5,12	2,61	2,55	4,05
Benzen	6,93	4,77	4,59	5,54
Butadien	3,18	1,96	1,86	3,30
Butan	5,53	2,29	2,04	4,19
Butanol	13,90	5,76	5,32	8,22
Cyclohexan	6,69	3,59	3,34	6,06
Diethylether	8,63	3,89	4,55	5,79
Dimethylether	5,73	2,65	2,57	4,60
Ethan	10,66	2,87	3,21	5,14
Ethanol	15,19	4,36	4,04	8,37
Ethen	3,42	1,44	1,47	3,89
Ethylbenzen	3,28	3,18	3,09	5,85
Hexan	8,26	4,27	4,24	6,93
Isooctan	6,14	2,94	2,83	4,27
Methan	8,41	1,97	1,73	4,01
Methanol	18,52	2,50	2,50	5,51
Octan	7,40	3,19	3,31	6,56
Propan	8,14	3,33	3,03	3,42
Propanal	5,67	2,57	2,71	4,23
Propanol	10,02	3,98	3,55	5,50
Propen	4,59	2,52	2,39	4,06
<b>Maximalwert</b>	<b>18,52</b>	<b>5,76</b>	<b>5,32</b>	<b>8,71</b>

Tabelle 10.3-12:  $I_{95\%}$ -Intervalle für die Datensätze der Einzelsubstanzen der CHO-Datenbank

Substanz	$\Delta F_{N, \text{Swan (1, 0)}}$ [%·s]	$\Delta F_{N, \text{Swan (3, 2)}}$ [%·s]	$\Delta F_{N, \text{Swan (0, 0)}}$ [%·s]	$\Delta F_{N, \text{Swan (1, 1)}}$ [%·s]
1, 3-Cyclohexadien	3,01	2,25	4,42	6,21
Aceton	2,51	2,02	3,81	4,70
Benzen	3,49	3,89	4,28	15,48
Butadien	2,23	1,95	2,64	3,54
Butan	2,05	2,15	3,12	4,21
Butanol	3,42	2,27	5,68	7,23
Cyclohexan	3,03	2,42	5,13	6,28
Diethylether	3,31	2,07	5,45	7,66
Dimethylether	1,94	1,49	4,16	4,34
Ethan	2,08	2,05	3,76	4,69
Ethanol	4,54	2,92	8,51	8,54
Ethen	1,98	1,69	3,11	3,86
Ethylbenzen	2,60	2,26	4,24	5,62
Hexan	3,27	2,21	5,04	7,07
Isooctan	2,77	2,34	4,59	5,29
Methan	1,70	1,48	2,71	3,11
Methanol	1,59	1,50	2,73	4,79
Octan	2,67	2,43	4,62	4,92
Propan	2,44	1,82	4,02	5,81
Propanal	2,42	2,01	5,73	5,36
Propanol	3,17	2,22	5,31	6,16
Propen	2,29	1,96	3,90	9,11
<b>Maximalwert</b>	<b>4,54</b>	<b>3,89</b>	<b>8,51</b>	<b>15,48</b>
Substanz	$\Delta F_{N, \text{H}}$ [%·s]	$\Delta F_{N, \text{CH 4310A 2}}$ [%·s]	$\Delta F_{N, \text{CH 4310A 3}}$ [%·s]	$\Delta F_{N, \text{CH 3900A}}$ [%·s]
1, 3-Cyclohexadien	12,43	7,94	8,98	14,32
Aceton	9,57	5,81	5,44	7,38
Benzen	11,74	9,73	9,62	10,15
Butadien	5,53	3,46	3,17	5,77
Butan	9,60	4,00	3,79	7,13
Butanol	23,48	9,48	8,62	12,47
Cyclohexan	11,81	6,34	6,39	10,30
Diethylether	16,09	6,78	8,11	10,30
Dimethylether	10,71	5,47	5,46	7,95
Ethan	17,37	5,41	5,52	7,33
Ethanol	26,30	8,95	8,37	16,44
Ethen	5,41	2,56	2,36	5,98
Ethylbenzen	5,71	7,50	7,47	9,46
Hexan	14,32	7,54	7,17	10,77
Isooctan	12,21	5,64	5,38	8,85
Methan	14,13	3,52	2,82	6,34
Methanol	32,35	4,65	4,57	8,77
Octan	12,96	6,61	6,47	10,59
Propan	13,17	6,10	5,24	5,65
Propanal	9,81	5,38	5,35	7,92
Propanol	22,52	8,20	7,28	8,67
Propen	7,39	4,46	4,45	6,77
<b>Maximalwert</b>	<b>32,35</b>	<b>9,73</b>	<b>9,62</b>	<b>16,44</b>

Tabelle 10.3-13: Intervalle  $I_{80\%}$ , in denen 80% aller  $\Delta F_{N,X}$ -Werte bei stickstoffhaltigen Verbindungen gefunden werden für die Datensätze der Einzelsubstanzen der CHON-Datenbank

Substanz	$\Delta F_{N, \text{Swan}} (1, 0)$ [%·s]	$\Delta F_{N, \text{Swan}} (3, 2)$ [%·s]	$\Delta F_{N, \text{Swan}} (0, 0)$ [%·s]	$\Delta F_{N, \text{Swan}} (1, 1)$ [%·s]
Anilin	1,51	1,34	2,36	3,00
Picolin	1,30	1,24	2,16	3,04
Propylamin	1,05	1,02	1,87	2,41
Pyridin	1,78	1,27	3,11	4,18
Triethylamin	1,09	1,05	1,85	2,39
Trimethylamin	0,96	0,91	1,55	2,04
<b>Max</b>	<b>1,78</b>	<b>1,34</b>	<b>3,11</b>	<b>4,18</b>
Substanz	$\Delta F_{N, H}$ [%·s]	$\Delta F_{N, CH 4310A 2}$ [%·s]	$\Delta F_{N, CH 4310A 3}$ [%·s]	$\Delta F_{N, CH 3900A}$ [%·s]
Anilin	3,26	2,45	2,70	5,00
Picolin	3,26	2,34	2,41	4,74
Propylamin	9,53	2,92	2,66	4,44
Pyridin	2,89	3,10	2,49	5,33
Triethylamin	4,44	2,07	1,97	3,98
Trimethylamin	5,40	1,71	1,61	3,49
<b>Max</b>	<b>9,53</b>	<b>3,10</b>	<b>2,70</b>	<b>5,33</b>
Substanz	$\Delta F_{N, CN \text{ Violett}} (1, 1)$ [%·s]	$\Delta F_{N, CN \text{ Violett}} (0, 1)$ [%·s]	$\Delta F_{N, NH 3360\text{\AA}} (0, 0; Q)$ [%·s]	$\Delta F_{N, NH 3360\text{\AA}} (0, 0; Q)$ [%·s]
Anilin	5,99	2,08	2,99	1,35
Picolin	5,10	2,06	2,92	1,75
Propylamin	9,79	2,16	4,98	9,16
Pyridin	9,10	2,94	8,17	6,92
Triethylamin	4,41	1,46	3,70	1,39
Trimethylamin	5,14	1,70	4,23	1,71
<b>Max</b>	<b>9,79</b>	<b>2,94</b>	<b>8,17</b>	<b>9,16</b>



Tabelle 10.3-14: Intervalle  $I_{95\%}$ , in denen 95% aller  $\Delta F_{N,X}$ -Werte bei stickstoffhaltigen Verbindungen gefunden werden für die Datensätze der Einzelsubstanzen der CHON-Datenbank

Substanz	$\Delta F_{N, \text{Swan (1, 0)}}$ [%·s]	$\Delta F_{N, \text{Swan (3, 2)}}$ [%·s]	$\Delta F_{N, \text{Swan (0, 0)}}$ [%·s]	$\Delta F_{N, \text{Swan (1, 1)}}$ [%·s]
Anilin	2,55	2,13	4,02	4,62
Picolin	2,26	2,10	4,09	4,85
Propylamin	1,74	1,65	2,81	4,42
Pyridin	2,99	2,14	5,55	6,73
Triethylamin	1,88	1,66	3,50	3,66
Trimethylamin	1,36	1,46	2,53	3,26
<b>Maximalwert</b>	<b>2,99</b>	<b>2,14</b>	<b>5,55</b>	<b>6,73</b>
Substanz	$\Delta F_{N, \text{H}}$ [%·s]	$\Delta F_{N, \text{CH 4310A 2}}$ [%·s]	$\Delta F_{N, \text{CH 4310A 3}}$ [%·s]	$\Delta F_{N, \text{CH 3900A}}$ [%·s]
Anilin	5,59	5,47	5,63	8,45
Picolin	5,98	4,93	4,88	8,42
Propylamin	18,93	5,59	4,90	7,11
Pyridin	5,28	5,16	4,85	9,97
Triethylamin	7,70	3,86	3,92	7,07
Trimethylamin	8,42	2,63	2,54	5,34
<b>Maximalwert</b>	<b>18,93</b>	<b>5,59</b>	<b>5,63</b>	<b>9,97</b>
Substanz	$\Delta F_{N, \text{CN Violett (1, 1)}}$ [%·s]	$\Delta F_{N, \text{CN Violett (0, 1)}}$ [%·s]	$\Delta F_{N, \text{NH 3360Å (0, 0; Q)}}$ [%·s]	$\Delta F_{N, \text{NH 3360Å (0, 0; Q)}}$ [%·s]
Anilin	12,53	4,44	6,24	2,13
Picolin	10,40	3,64	5,69	3,06
Propylamin	18,69	4,90	9,40	15,50
Pyridin	15,57	5,07	14,19	10,54
Triethylamin	9,05	2,77	5,97	2,16
Trimethylamin	8,20	2,56	7,54	2,72
<b>Maximalwert</b>	<b>18,69</b>	<b>5,07</b>	<b>14,19</b>	<b>15,50</b>

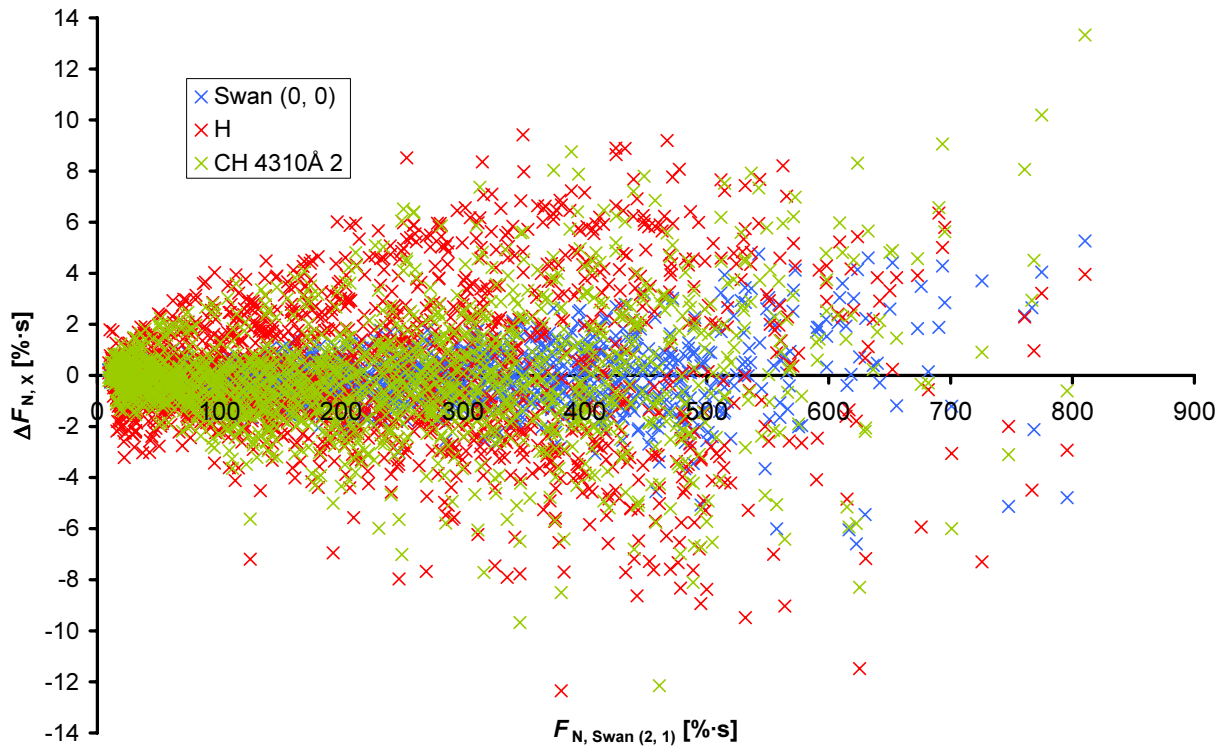


Abbildung 10.3-11: Auftragung der Differenz zwischen Messwert und Schätzwert bei der Swan (0, 0)- und CH 4310Å 2-Bande und der H-Linie gegen die normierte Signalfläche der Swan (2, 1)-Bande für alle im Benzendatensatz gespeicherten Messungen

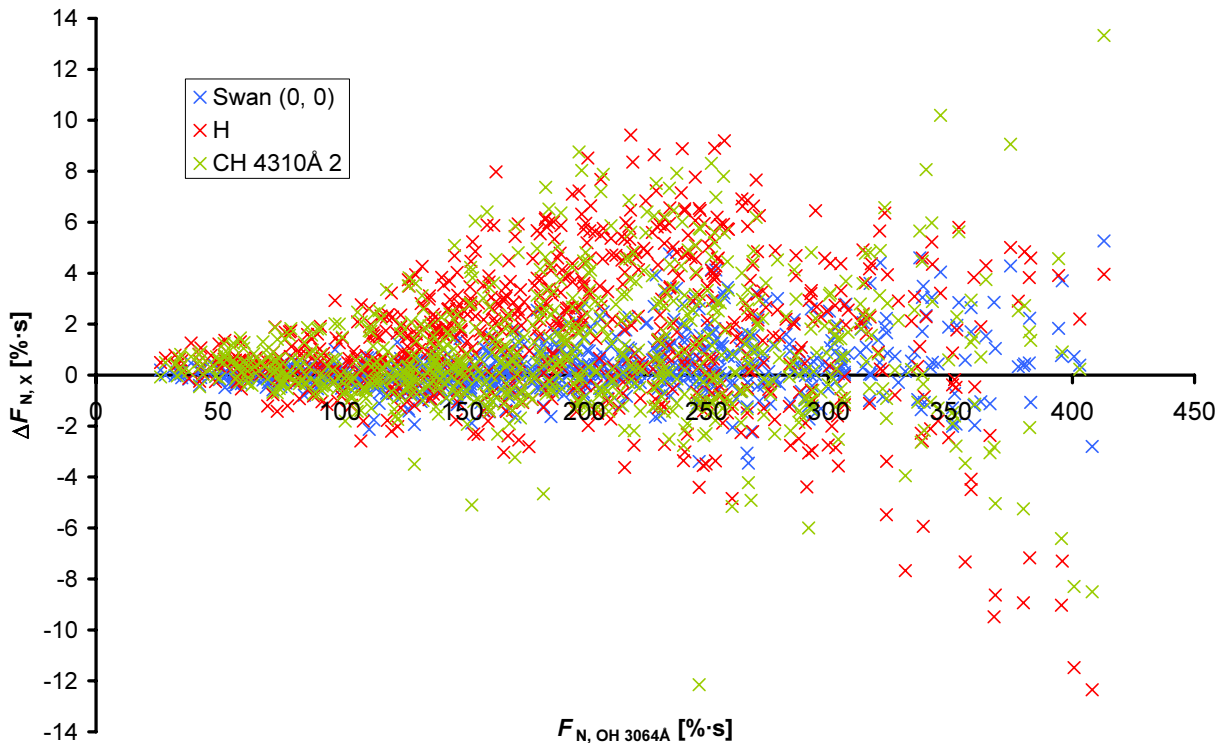


Abbildung 10.3-12: Auftragung der Differenz zwischen Messwert und Schätzwert bei der Swan (0, 0)- und CH 4310Å 2-Bande und der H-Linie gegen die normierte Signalfläche OH3064Å-Bande für alle im Benzendatensatz gespeicherten Messungen

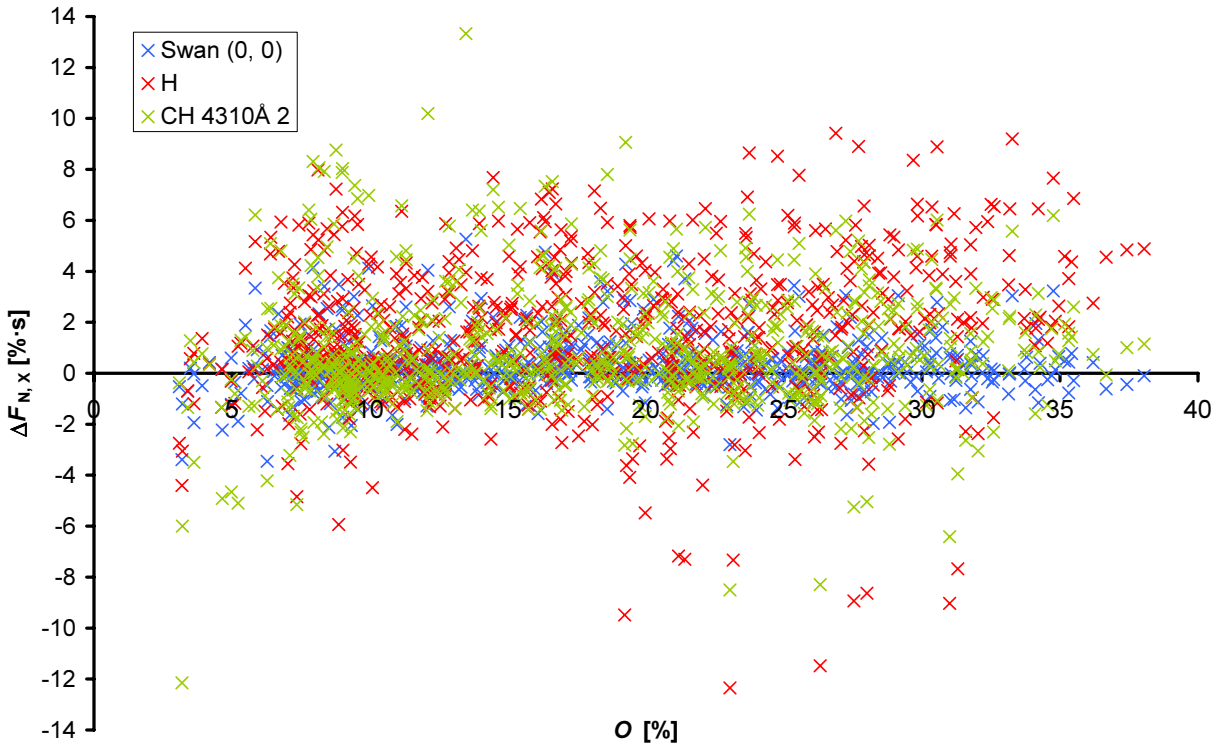


Abbildung 10.3-13: Auftragung der Differenz zwischen Messwert und Schätzwert bei der Swan (0, 0)- und CH 4310Å 2-Bande und der H-Linie gegen den Offset der OH3064Å-Bande für alle im Benzendatensatz gespeicherten Messungen

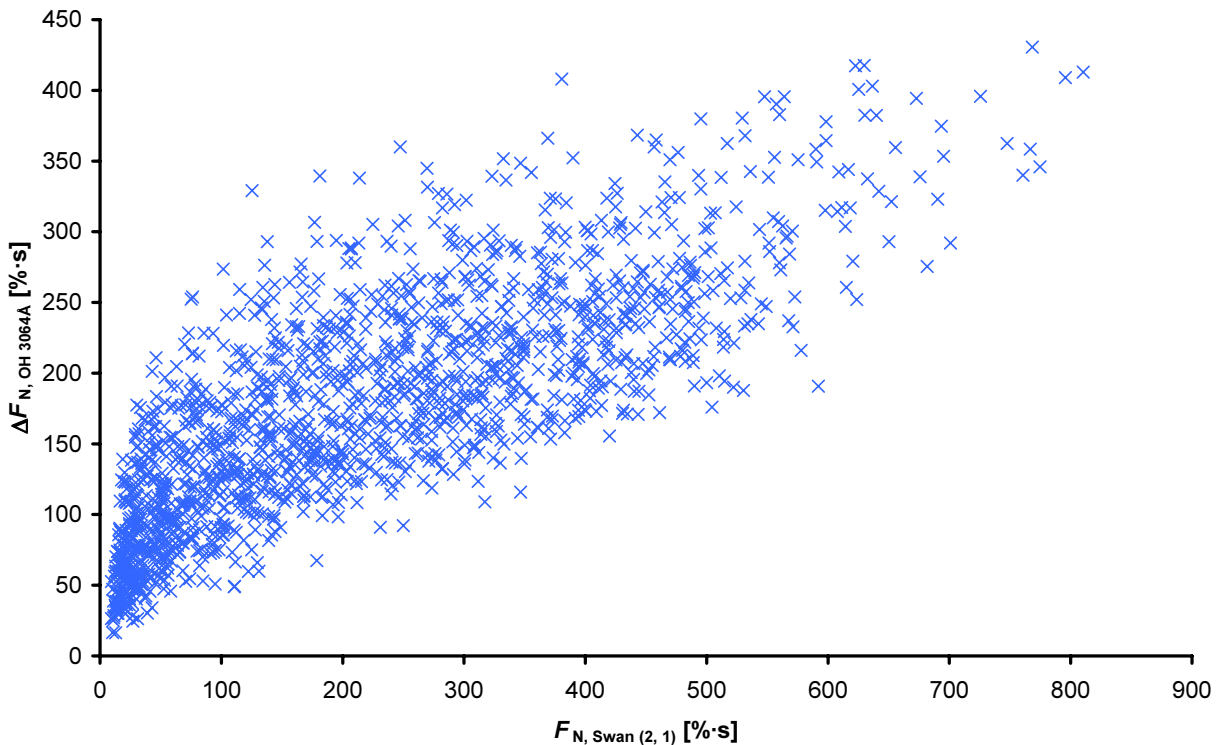


Abbildung 10.3-14: Auftragung der normierten Signalflächen der OH3064Å-Bande gegen die normierte Fläche der Swan (2, 1)-Bande für alle im Benzendatensatz gespeicherten Messungen

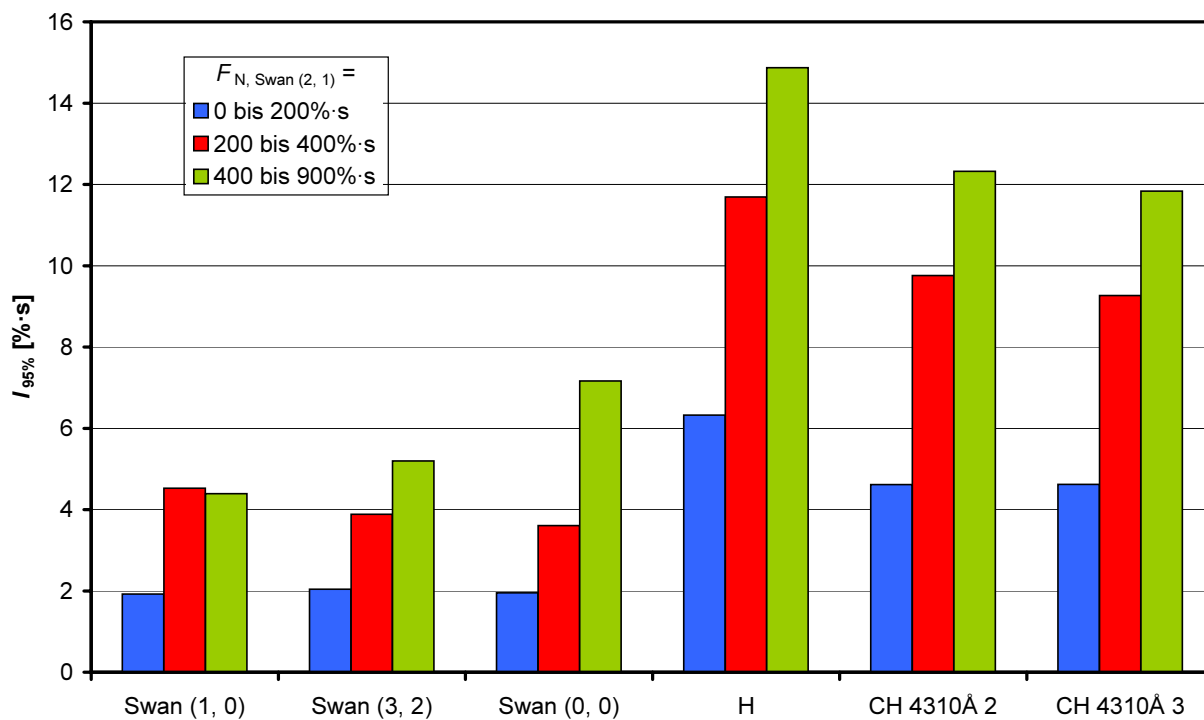


Abbildung 10.3-15:  $I_{95\%}$ -Intervalle für den in drei Teildatensätze nach unterschiedlichen  $F_{N, Swan(2, 1)}$ -Werten aufgespaltenen Benzendatensatz

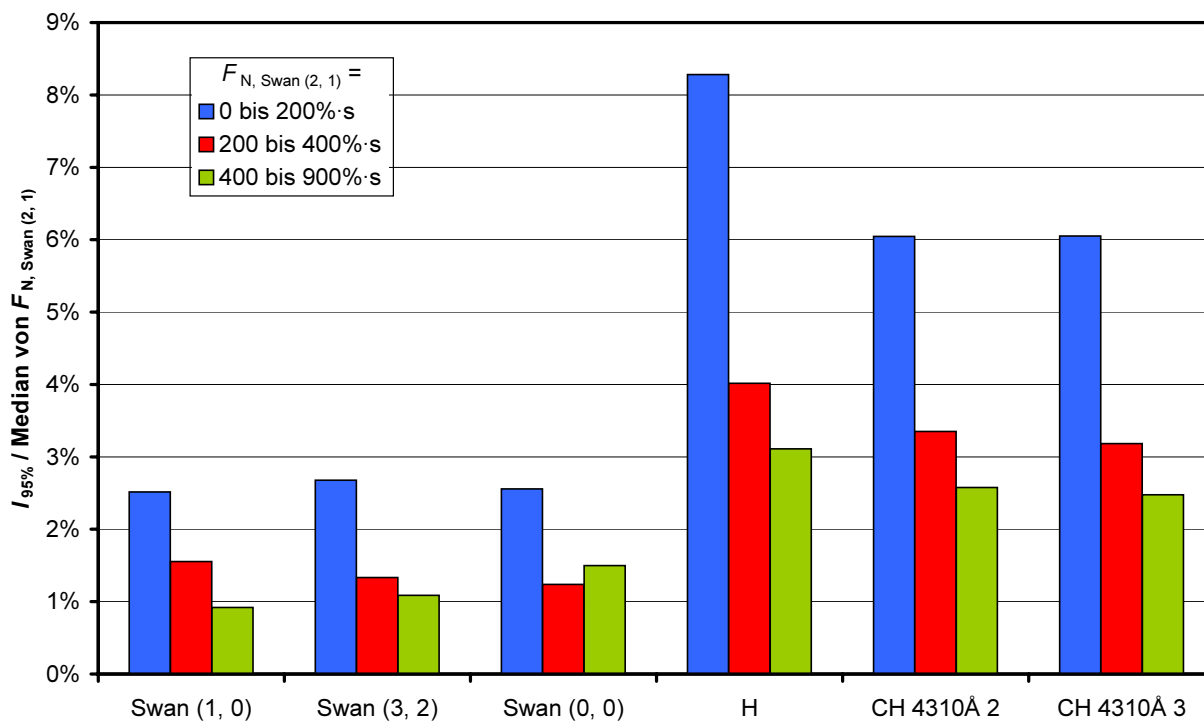


Abbildung 10.3-16:  $I_{95\%}$ -Intervalle in Relation zum Median von  $F_{N, Swan(2, 1)}$  für den in drei Teildatensätze nach unterschiedlichen  $F_{N, Swan(2, 1)}$ -Werten aufgespaltenen Benzendatensatz

### 10.3.4 Substanzvergleich

Um zu untersuchen, ob sich die GC-MIP-MED-Signale verschiedener Substanzen unterscheiden, hat sich der paarweise Vergleich der Datensätze der Datenbanken bewährt. Als Vergleichskriterien kommen die Anteile an signifikant unterscheidbaren Signalen  $A_{X, S1-S2}$  und  $A_{X, S2-S1}$  in Frage. Diese können bei der Abschätzung der Signalflächen aus den Datensätzen von zwei Substanzen S1 und S2 gewonnen werden. Verglichen werden für jeden Klassifizierungsparametersatz im Datensatz von S1 die zugehörigen  $F_{N, X}$ -Werte mit den mittels des Datensatzes von S2 berechneten  $F_{S, X}$ -Werten ( $A_{X, S1-S2}$ ) und umgekehrt ( $A_{X, S2-S1}$ ). Schlüsselst man diese Anteile in  $A_{X, S1-S2, (-)}$  und  $A_{X, S1-S2, (+)}$  bzw.  $A_{X, S2-S1, (-)}$  und  $A_{X, S2-S1, (+)}$  auf, kann nicht nur festgestellt werden, ob für die beiden Datensätze die für eine signifikante Unterscheidbarkeit erforderliche 5%-Marke überschritten ist, sondern auch für welche Substanz höhere oder niedrigere  $F_{S, X}$ -Werte zu erwarten sind.

Um die Ergebnisse eines Substanzvergleiches begutachten zu können ist eine übersichtliche graphische Darstellung von Vorteil. Es hat sich herausgestellt, dass sich hierfür ein Balkendiagramm nach dem Muster von Abbildung 10.3-17 sehr gut eignet. In dieser Abbildung werden die Datensätze von Dimethylether und Ethanol miteinander verglichen. Für jede Bande und die Wasserstofflinie existieren ein blauer und ein roter Balken. Die Höhe des blauen Balkens repräsentiert den Anteil an gemessenen Dimethylethersignalflächen, die sich signifikant von den jeweiligen Schätzwert der Ethanolsignalfläche auf Basis der gleichen Klassifizierungsparametersatzes unterscheiden (S1 = Dimethylether u. S2 = Ethanol). Der rote Balken visualisiert den Anteil an Ethanolsignalflächen, welche sich von der jeweiligen mit dem gleichen Klassifizierungsparametersatz abgeschätzten Dimethylethersignalfläche unterscheiden. Der blaue Balken repräsentiert den Anteil signifikant unterschiedlicher Signale bezüglich des Vergleichs von Dimethylethermesswerten mit Ethanolgeschätzwerten  $A_{X, S1-S2}$  und der rote Balken den Anteil signifikant unterschiedlicher Signale bezüglich des Vergleichs von Ethanolmesswerten mit Dimethyletherschätzwerten  $A_{X, S2-S1}$ . Im Falle der Swan (3, 2)-Bande ist kein roter Balken vorhanden. Hier ist  $A_{X, S2-S1}$  Null. Es existieren für die Swan (3, 2)-Bande keine Ethanolmesswerte, die sich signifikant von den zu den entsprechenden Klassifizierungsparametern gehörenden Dimethyletherschätzwerten unterscheiden. Die Balken sind in einen gefüllten und einen punktierten Bereich aufgeteilt. Der gefüllte Bereich repräsentiert  $A_{X, S1-S2, (+)}$  bzw.  $A_{X, S2-S1, (+)}$  und der punktierte Bereich  $A_{X, S1-S2, (-)}$  bzw.  $A_{X, S2-S1, (-)}$ , weshalb  $A_{X, S1-S2} = A_{X, S1-S2, (+)} + A_{X, S1-S2, (-)}$  bzw.  $A_{X, S2-S1} = A_{X, S2-S1, (+)} + A_{X, S2-S1, (-)}$  die Achsenbezeichnungen bilden. Das Diagramm ist folgendermaßen zu interpretieren: Überschreitet für eine der Molekülbanden oder die Wasserstofflinie sowohl der blaue als auch der rote Balken die 5%-Marke, so unterscheiden sich die Substanzdatensätze signifikant. Da dies bei CH 4310Å 2 und CH 4310Å 3 der Fall ist, liegt ein signifikanter Unterschied zwischen Dimethylether und Ethanol vor. Gilt für ein Balkenpaar, dass die gefüllte blaue Fläche größer ist als die punktierte und die punktierte rote Fläche größer als die gefüllte, so gilt nach Gleichung (10.3-10), dass die normierten Signalflächen der entsprechenden Bande oder Linie der Substanz S1 tendenziell größer sind als die der Substanz S2. Ist hingegen für ein Balkenpaar die gefüllte blaue

Fläche kleiner ist als die punktierte und die punktierte rote Fläche kleiner als die gefüllte, so gilt nach Gleichung (10.3-10), dass die normierten Signalflächen der entsprechenden Bande oder Linie der Substanz S1 tendenziell kleiner sind als die der Substanz S2. Im Falle von Dimethylether und Ethanol tendieren die durch Vermessen von Dimethylether ermittelten Messwerte bei den beiden CH-Banden zu den höheren normierten Signalflächen.

Ein signifikanter Unterschied wird immer dann festgestellt, wenn sowohl  $A_{X, S1-S2}$  als auch  $A_{X, S2-S1}$  größer als 5% sind. Man sollte sich auf keinen Fall auf nur einen dieser beiden Größen als Anhaltspunkt für einen signifikanten Unterschied verlassen. Der Grund für die zwingende Einbeziehung beider Anteile signifikant abweichender Signale liegt in der Häufigkeitsverteilung der Signalflächen, welche durch  $F_{N, \text{Swan}}(2, 1)$  und durch Interkorrelation auch durch  $F_{N, \text{OH } 3064\text{\AA}}$  quantifiziert sind. Die in Abbildung 10.3-18 sowie Abbildung 10.3-19 visualisierten Histogramme zeigen, dass für Ethanol im Vergleich zum Dimethylether eine ganze Anzahl an Messungen vorliegt, die wesentlich höhere  $F_{N, \text{Swan}}(2, 1)$ - und  $F_{N, \text{OH } 3064\text{\AA}}$ -Werte aufweisen. Vergleicht man die Messwerte aus dem Ethanoldatensatz mit den auf Basis ihrer Klassifizierungsparameter berechneten Schätzwerten der analytischen Zielgrößen für Dimethylether, so tritt ein systematischer Fehler auf. Die Klassifizierungsparametersätze der Ethanolmesswerte sind von denen im Datensatz von Dimethylether zu verschieden. Die Abschätzung von Ethanol signalflächen mit Hilfe von Dimethylether-Klassifizierungsparametern unterliegt dieser Problematik nicht; nicht zuletzt deshalb, weil die Häufigkeitsverteilung bei den Offsetwerten für die beiden Datensätze sehr ähnlich ist (siehe Abbildung 10.3-20). Daher kann diesen Schätzwerten wesentlich mehr vertraut werden. Dadurch, dass sowohl  $A_{X, S1-S2}$  als auch  $A_{X, S2-S1}$  zur Feststellung eines signifikanten Unterschiedes herangezogen werden, kann die Gefahr verringert werden, dass ein solcher Unterschied nur auf Grund verschiedener Wertebereiche der Klassifizierungsparameter festgestellt wird.

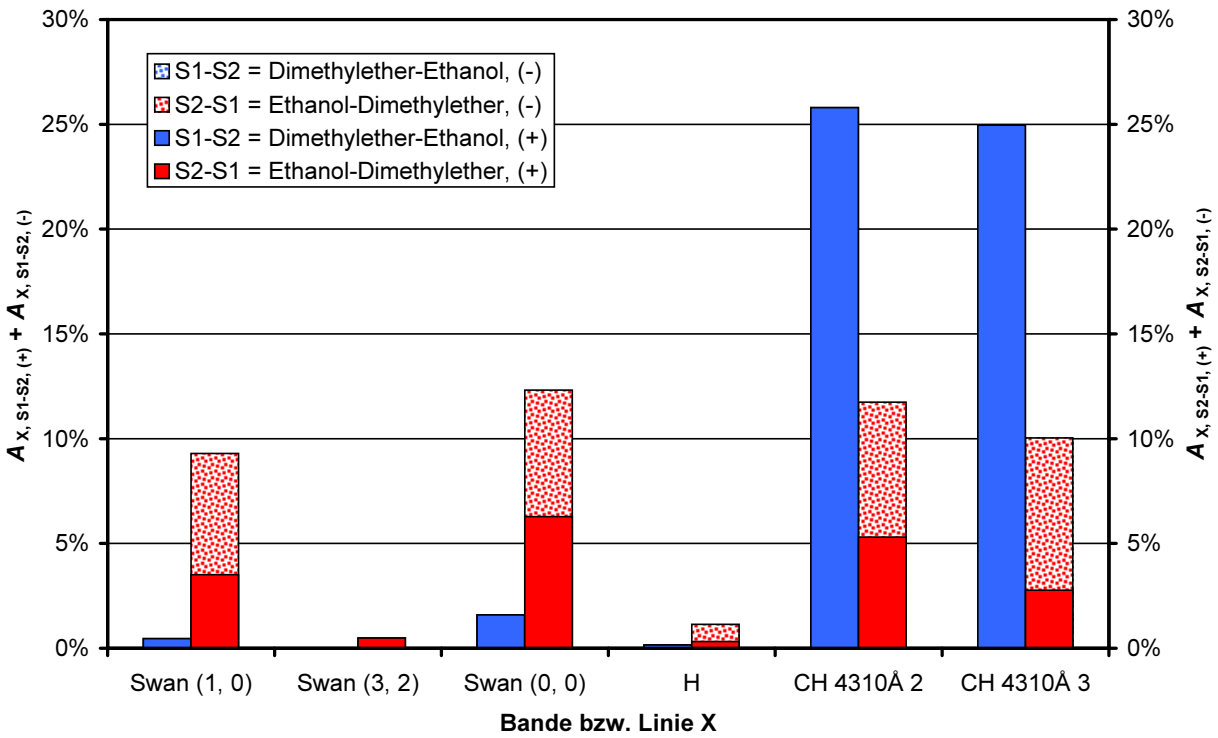


Abbildung 10.3-17: Prozentualer Anteil der bei der Abschätzung von Ethanol-Signalfächen mit Dimethylether-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Signalfächen signifikant zu groß bzw. zu klein sind

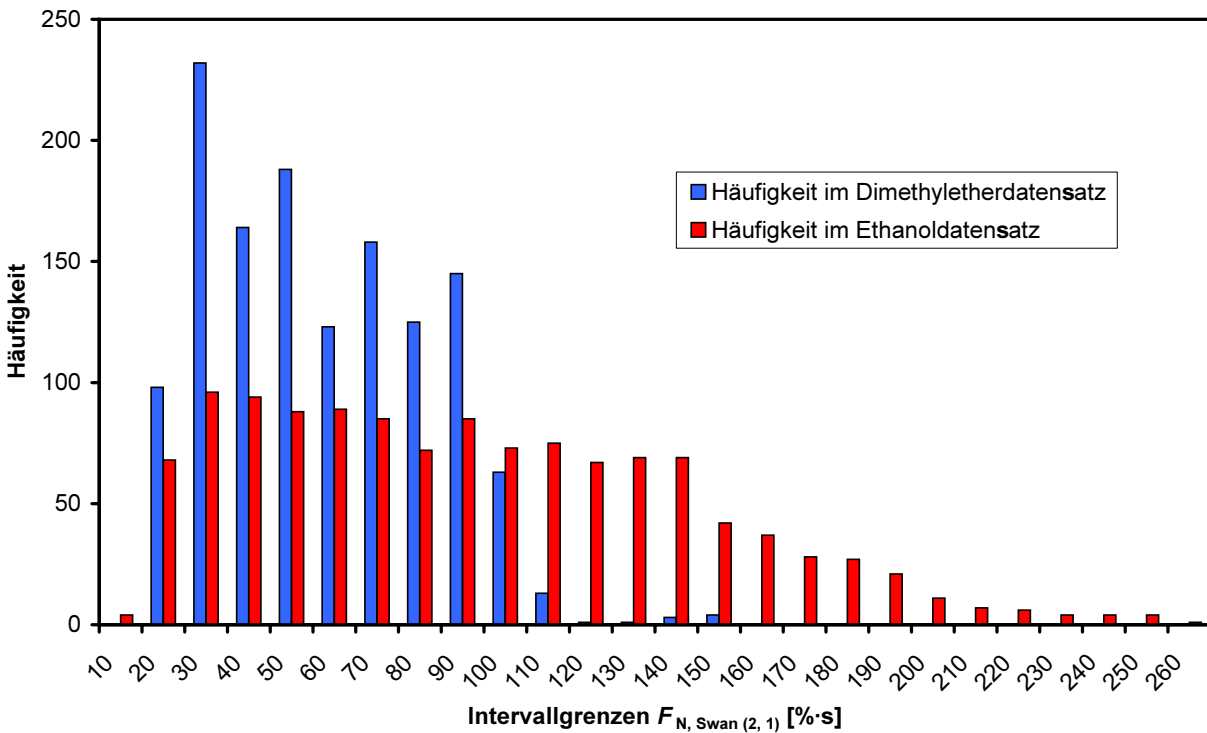


Abbildung 10.3-18: Histogramm aller Signalfächen der Swan (2, 1)-Bande im Dimethylether- und Ethanoldatensatz

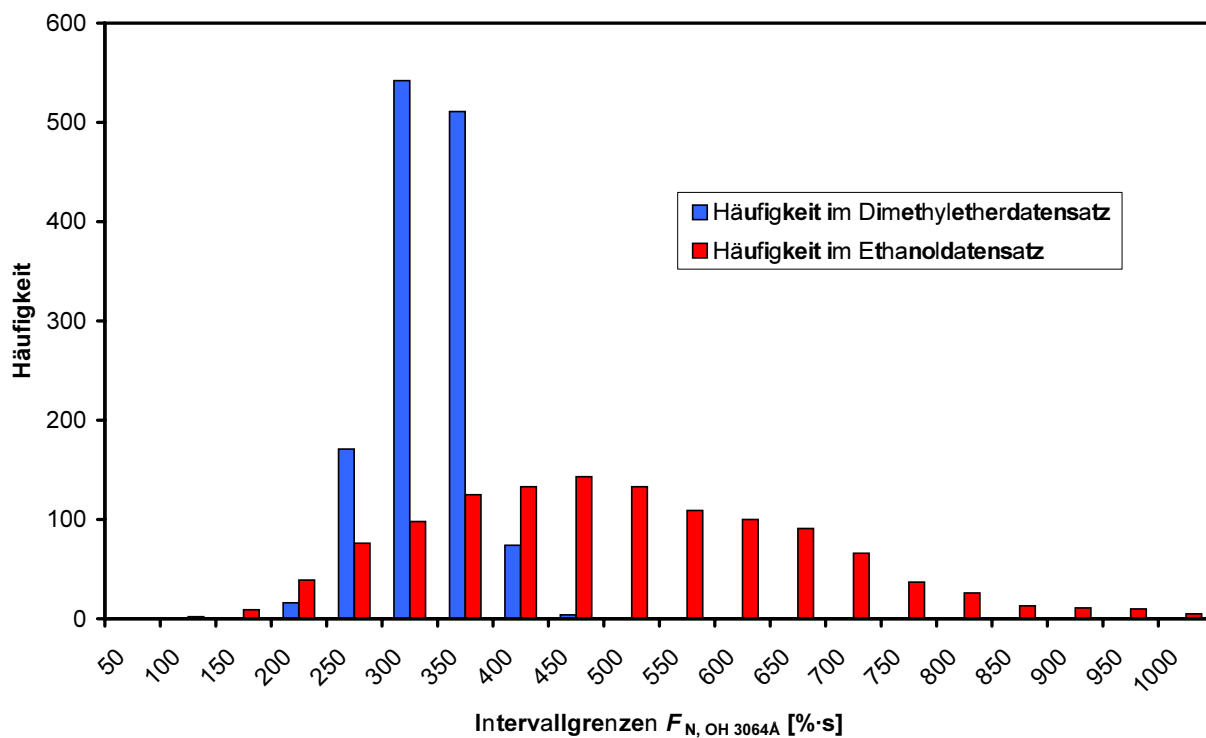


Abbildung 10.3-19: Histogramm aller Signalflächen der OH 3064Å-Bande im Dimethylether- und Ethanoldatensatz

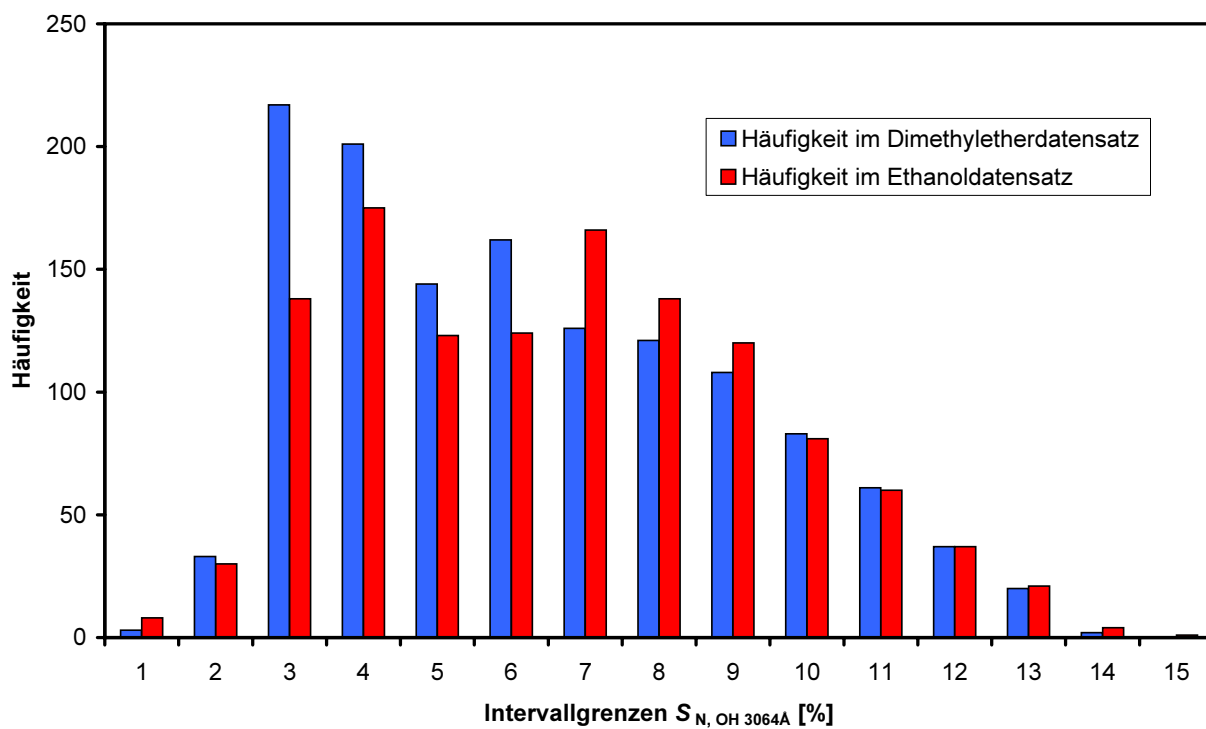


Abbildung 10.3-20: Histogramm aller Offset-Werte im Dimethylether- und Ethanoldatensatz



## 10.4 Unterschiede bei den GC-MIP-MED-Signalen verschiedener Substanzen

### 10.4.1 Verbindungen gleicher Zusammensetzung und unterschiedlicher Konstitution

Um beurteilen zu können, ob nur die Zusammensetzung oder zusätzlich auch die Bindungssequenz einen Einfluss auf GC-MIP-MED-Signale haben, sind eine Reihe von Konstitutionsisomeren vermessen worden. In Tabelle 10.4-1 sind alle Konstitutionsisomere aufgelistet, von denen sich Messungen in den Datenbanken befinden. Die Substanzen sind nach ihrem Kohlenstoffanteil im Molekül sortiert. Abbildung 10.4-1 zeigt die Lewis-Konstitutionsformeln dieser Verbindungen. Anhand der aus den verschiedenen Datensatzpaaren gewonnenen  $A_{X, S1-S2}$ - und der  $A_{X, S2-S1}$ -Werte kann abgeschätzt werden, ob sich Konstitutionsisomere überhaupt anhand von GC-MIP-MED-Signalen unterscheiden lassen (siehe Abbildung 10.4-2 bis Abbildung 10.4-10). Dies ist nur bei zwei Isomerenpaaren, Ethanol und Dimethylether sowie Propylamin und Trimethylamin der Fall (Abbildung 10.4-2 und Abbildung 10.4-3). Nur bei diesen Verbindungspaaren ist für mindestens eine Bande oder Linie X sowohl der  $A_{X, S1-S2}$ - als auch  $A_{X, S2-S1}$ -Wert größer 5%. Beim Vergleich von Ethanol mit Dimethylether zeigen sich signifikante Unterschiede bei den CH4310Å 2- und CH4310Å 3-Signalen, wobei der Ether zu höheren normierten Signalflächen tendiert. Im Falle von Trimethylamin und Propylamin treten bei allen Banden mit Ausnahme der Swan-Banden und der Wasserstofflinie signifikante Unterschiede auf. Trimethylamin tendiert hierbei zu den höheren Signalflächen. Bei den beiden verglichenen Isomerenpaaren besitzt eine der beiden Substanzen keine Kohlenstoffkette. Diese Substanz ist mit Dimethylether bzw. Trimethylamin gleichzeitig diejenige, bei deren GC-MIP-MED-Signalen die Tendenz zu höheren normierten CH4310Å 2- und CH4310Å 3-Signalflächen zu beobachten ist. Es besteht demnach ein Zusammenhang zwischen diesem strukturellen Merkmal und den GC-MIP-MED-Signalen. Bei Isomerenpaaren, welche Kohlenstoffketten mit mindestens zwei aneinander gebundenen C-Atomen besitzen, kann kein Zusammenhang zwischen Struktur bzw. Konstitution der Verbindungen und deren GC-MIP-MED-Signalen nachgewiesen werden.

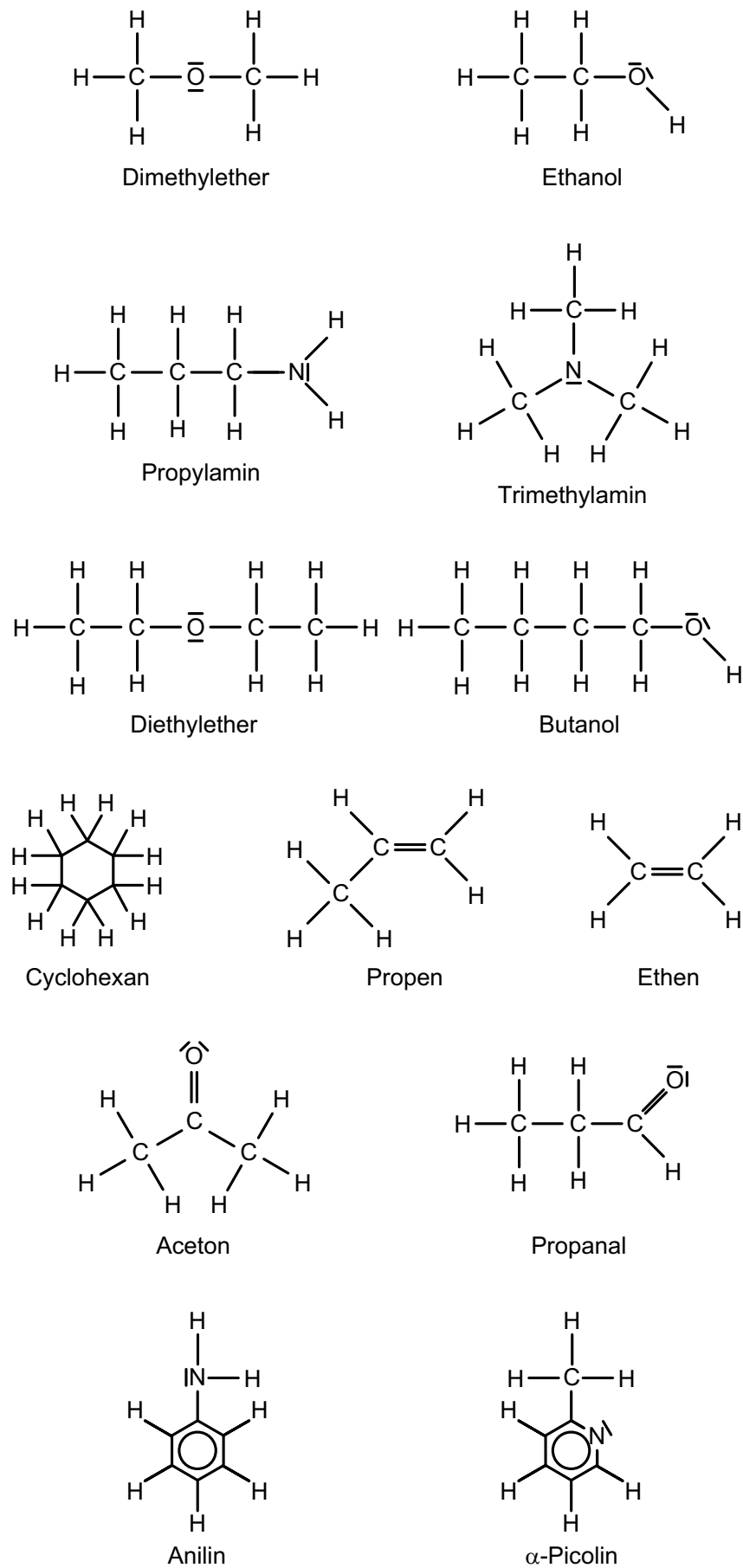


Abbildung 10.4-1: Mit dem GC-MIP-MED vermessene Konstitutionsisomere

Tabelle 10.4-1: Mit dem GC-MIP-MED vermessene Konstitutionsisomere und deren atomare Zusammensetzung

Substanzname	Atomzahl pro Molekül				Atomzahl von C, H, O und N dividiert durch die gesamte Zahl an Atomen im Molekül			
	C	H	O	N	C	H	O	N
Ethanol	2	6	1	0	0,2222	0,6667	0,1111	0
Dimethylether	2	6	1	0	0,2222	0,6667	0,1111	0
Propylamin	3	9	0	1	0,2308	0,6923	0	0,0769
Trimethylamin	3	9	0	1	0,2308	0,6923	0	0,0769
Butanol	4	10	1	0	0,2667	0,6667	0,0667	0
Diethylether	4	10	1	0	0,2667	0,6667	0,0667	0
Octan	8	18	0	0	0,3077	0,6923	0	0
Isooctan	8	18	0	0	0,3077	0,6923	0	0
Cyclohexan	6	12	0	0	0,3333	0,6667	0	0
Ethen	2	4	0	0	0,3333	0,6667	0	0
Propen	3	6	0	0	0,3333	0,6667	0	0
Propanal	3	6	0	0	0,3333	0,6667	0	0
Aceton	3	6	0	0	0,3333	0,6667	0	0
Picolin	6	7	0	1	0,4286	0,5000	0	0,0714
Anilin	6	7	0	1	0,4286	0,5000	0	0,0714

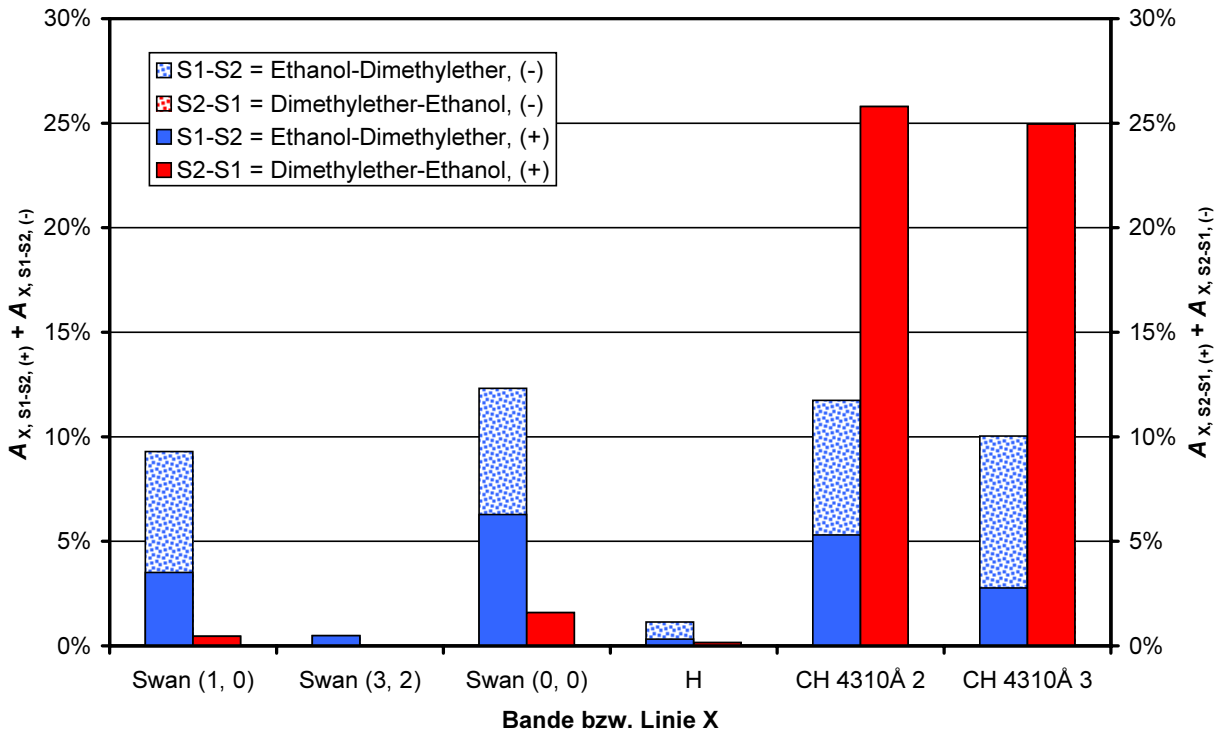


Abbildung 10.4-2: Prozentualer Anteil der bei der Abschätzung von Dimethylether-Signalfächern mit Ethanol-Klassifizierungsparametern und umgekehrt ermittelten Signalfächern, welche gegenüber den gemessenen Signalfächern signifikant zu groß bzw. zu klein sind

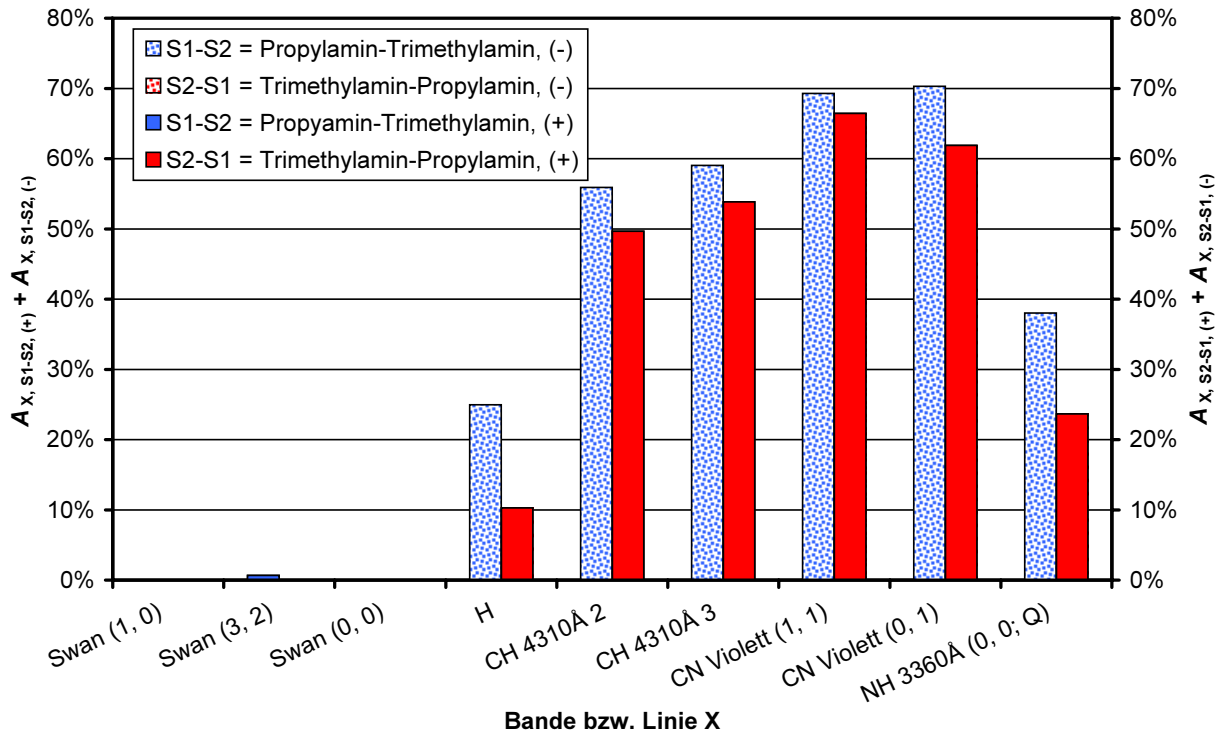


Abbildung 10.4-3: Prozentualer Anteil der bei der Abschätzung von Trimethylamin-Signalfächen mit Propylamin-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

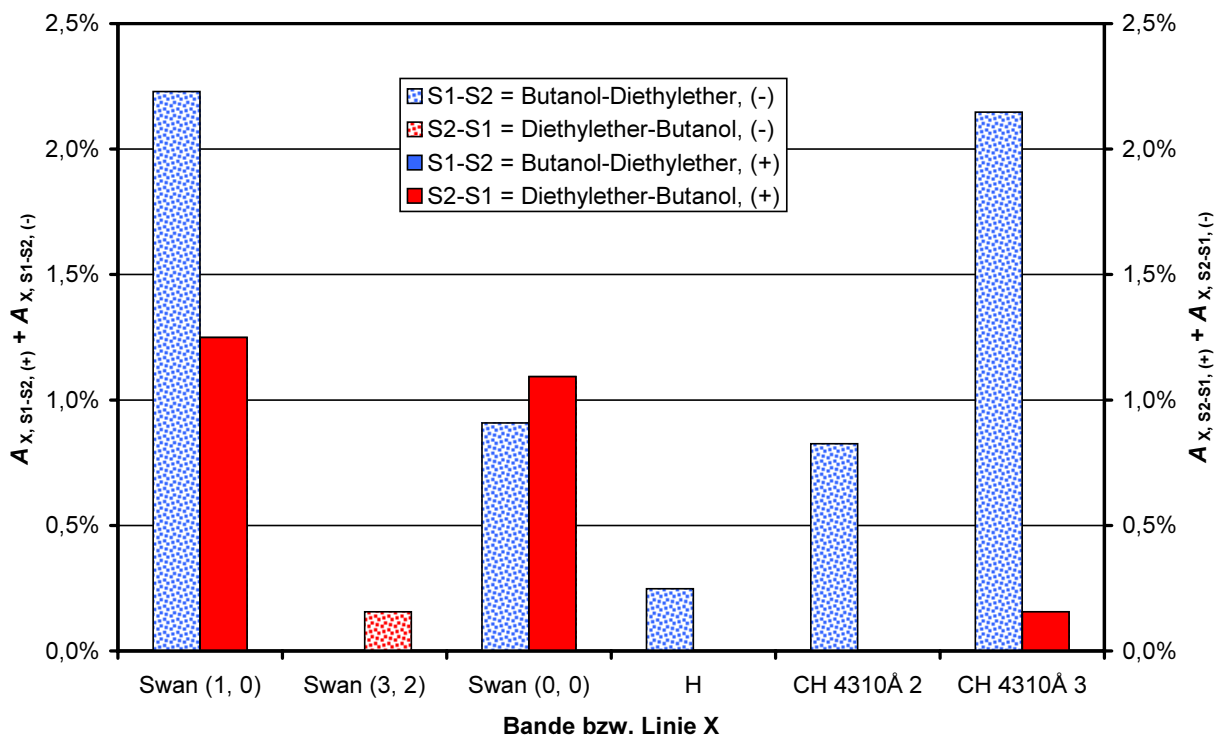


Abbildung 10.4-4: Prozentualer Anteil der bei der Abschätzung von Diethylether-Signalfächen mit Butanol-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

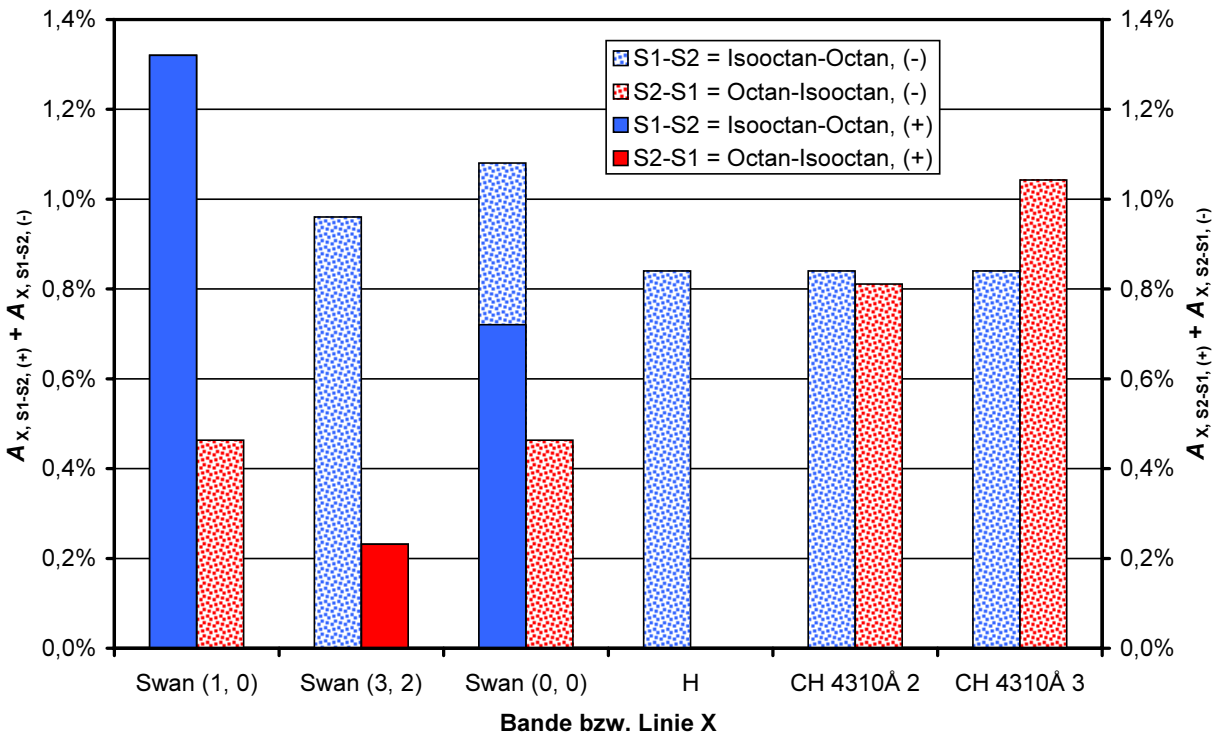


Abbildung 10.4-5: Prozentualer Anteil der bei der Abschätzung von Octan-Signalflächen mit Isooctan-Klassifizierungsparametern und umgekehrt ermittelten Signalflächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

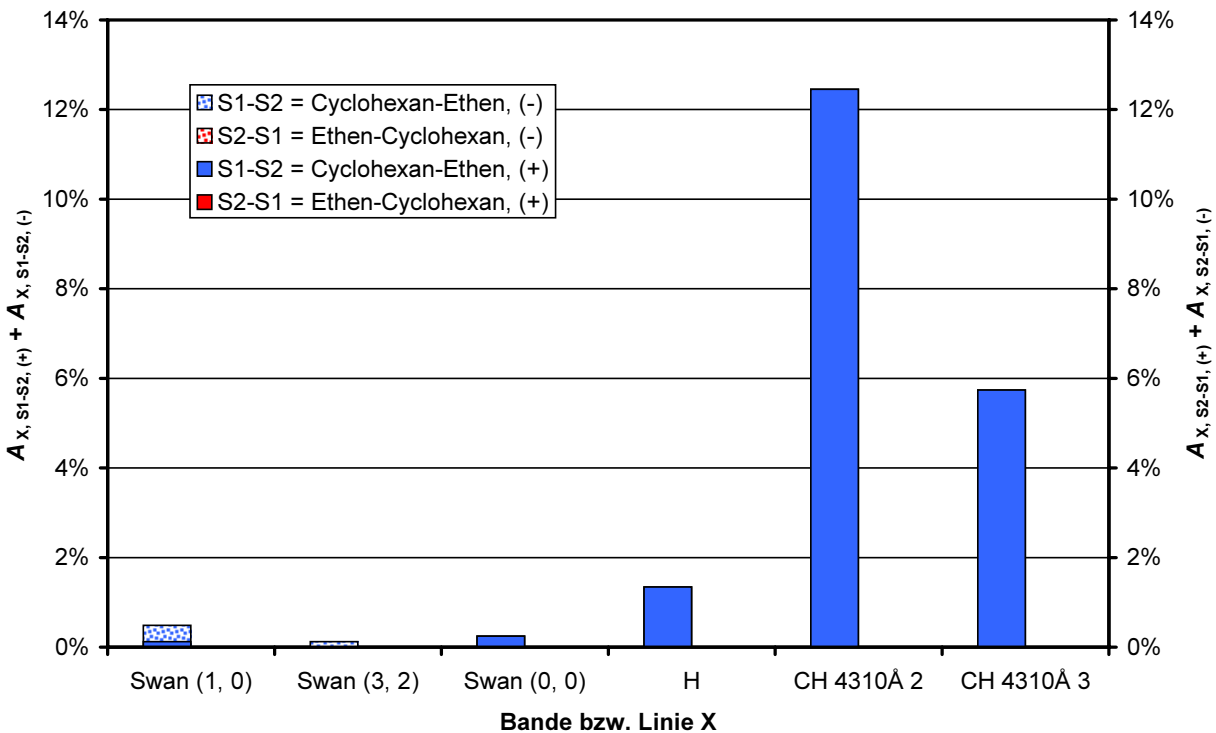


Abbildung 10.4-6: Prozentualer Anteil der bei der Abschätzung von Ethen-Signalflächen mit Cyclohexan-Klassifizierungsparametern und umgekehrt ermittelten Signalflächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

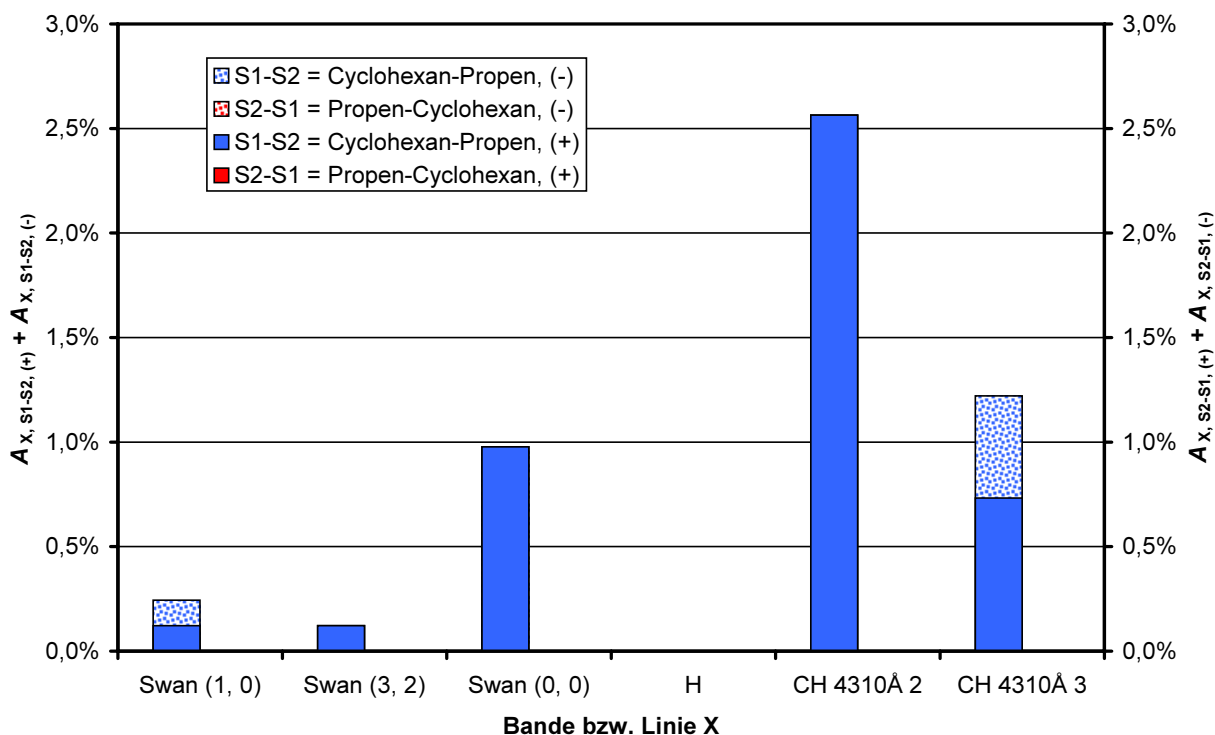


Abbildung 10.4-7: Prozentualer Anteil der bei der Abschätzung von Propen-Signalfächen mit Cyclohexan-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

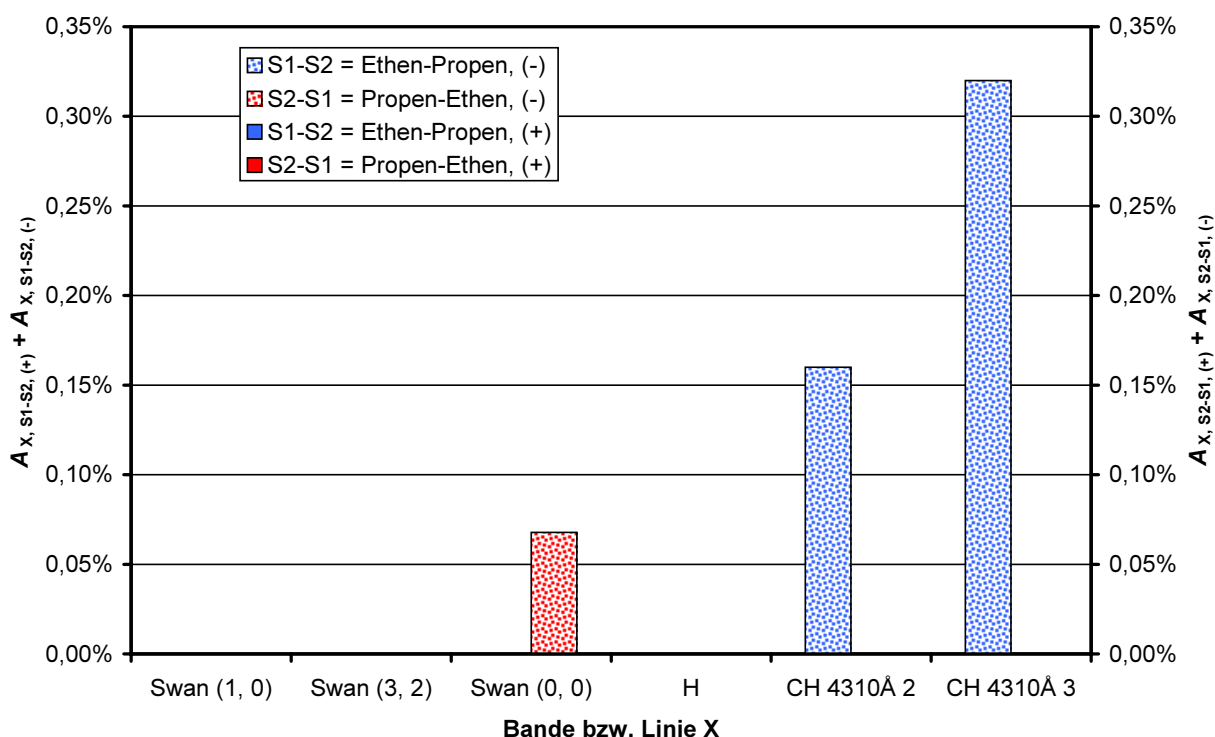


Abbildung 10.4-8: Prozentualer Anteil der bei der Abschätzung von Propen-Signalfächen mit Ethen-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

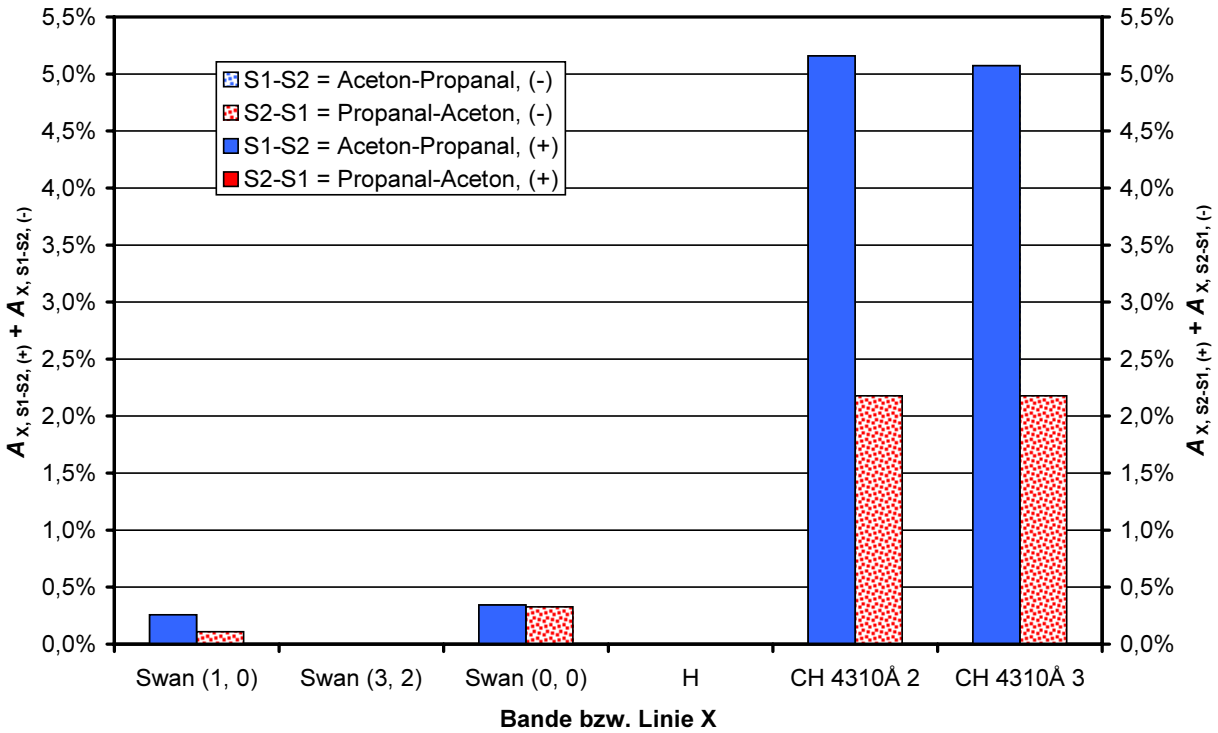


Abbildung 10.4-9: Prozentualer Anteil der bei der Abschätzung von Propanal-Signalfächen mit Aceton-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

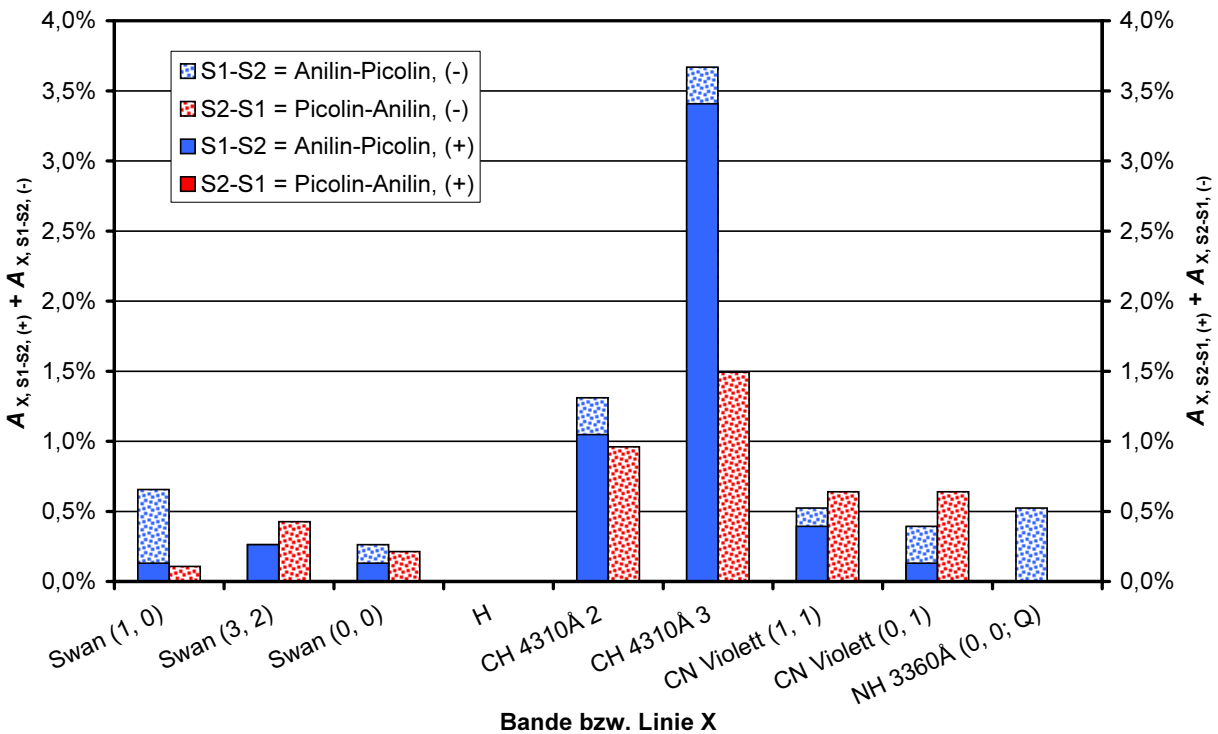


Abbildung 10.4-10: Prozentualer Anteil der bei der Abschätzung von alpha-Picolin-Signalfächen mit Anilin-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

## 10.4.2 Verbindungen ähnlicher Struktur bei leicht unterschiedlicher Zusammensetzung

Konstitutionsisomere lassen sich mit dem GC-MIP-MED nur schlecht unterscheiden. Im Gegensatz dazu ändern sich die GC-MIP-MED-Signalflächen meistens signifikant, sobald sich die Zusammensetzung der gemessenen Verbindung ändert. Die in Tabelle 10.4-2 aufgeführten farblich gekennzeichneten Substanzpaare haben die wichtigsten strukturellen Merkmale gemeinsam, unterscheiden sich allerdings leicht bei der atomaren Zusammensetzung (siehe Abbildung 10.4-11). Propan und Butan sowie Hexan und Octan sind unverzweigte Alkane, die lediglich eine etwas andere Kettenlänge besitzen. Propylamin und Triethylamin sind aliphatische Amine mit mehr oder weniger langen an das N-Atom gebundenen Alkylketten. Cyclohexan weist durch den Ringschluss einen höheren Kohlenstoffgehalt als Octan auf. Beide Moleküle bestehen aus einer Kette von  $\text{CH}_2$ -Gruppen, wobei die C-Atome in beiden Fällen tetraedrisch koordiniert sind. Octan weist als kleinen strukturellen Unterschied zwei endständige H-Atome auf. Bei anderen, vor allem bei kürzerkettigen Cycloalkanen als Cyclohexan ist diese tetraedrische Koordination durch die Ringspannung gestört. Butadien und Cyclohexadien besitzen jeweils zwei konjugierte Doppelbindungen. Pyridin und Picolin besitzen ein heteroaromatisches Ringsystem mit einem N-Atom und Benzen sowie Ethylbenzen sind Aromaten.

Abbildung 10.4-12 bis Abbildung 10.4-18 visualisieren den Grad der Abweichungen zwischen den GC-MIP-MED-Signalen der 7 Verbindungspaare. In fast allen Fällen unterscheiden sich diese signifikant. Nur zwischen Hexan und Octan kann kein signifikanter Unterschied nachgewiesen werden (siehe Abbildung 10.4-14). Bei diesem Substanzpaar liegt auch die geringste Abweichung bei der atomaren Zusammensetzung vor (siehe Tabelle 10.4-2). Sobald  $A_{X, S1-S2}$  und  $A_{X, S2-S1}$  größer Null sind, gilt, dass Verbindungen mit höheren Kohlenstoffanteilen im Molekül höhere normierte Flächen beim Swan (3, 2)-Signal aufweisen, während die normierten Flächen aller anderen Signale niedriger sind. Bei der H-Linie findet man auf Grund der großen Unsicherheit bei der Bestimmung der Signalflächen oft nur einen geringen Anteil an Schätzwerten, welche sich von den Messwerten unterscheiden. Trotz geringer Unsicherheit bei der Signalflächenbestimmung sind auch bei den Swan-Banden eher geringe  $A_{X, S1-S2}$ - bzw.  $A_{X, S2-S1}$ -Werte zu finden. Die normierten Flächen dieser Banden korrelieren demnach nicht sehr stark mit der Substanzzusammensetzung. Die größten  $A_{X, S1-S2}$ - und  $A_{X, S2-S1}$ -Werte sind bei den  $\text{CH}_4 310\text{\AA}$  2- und  $\text{CH}_4 310\text{\AA}$  3-Werten zu beobachten.

Es kann also festgehalten werden, dass die atomare Zusammensetzung einer Substanz einen wesentlich größeren Einfluss auf deren GC-MIP-MED-Signale besitzt als deren Struktur bzw. Konstitution. Aus diesem Sachverhalt folgt, dass die Messungen der Signalflächen der verschiedenen Banden und Linien zum größten Teil redundante Informationen liefern. Die für einen festgelegten Klassifizierungsparametersatz ermittelten Flächenunterschiede korrelieren für jede als analytische Zielgröße betrachtete normierte Signalfläche weitgehend mit der Zusammensetzung der Substanz.



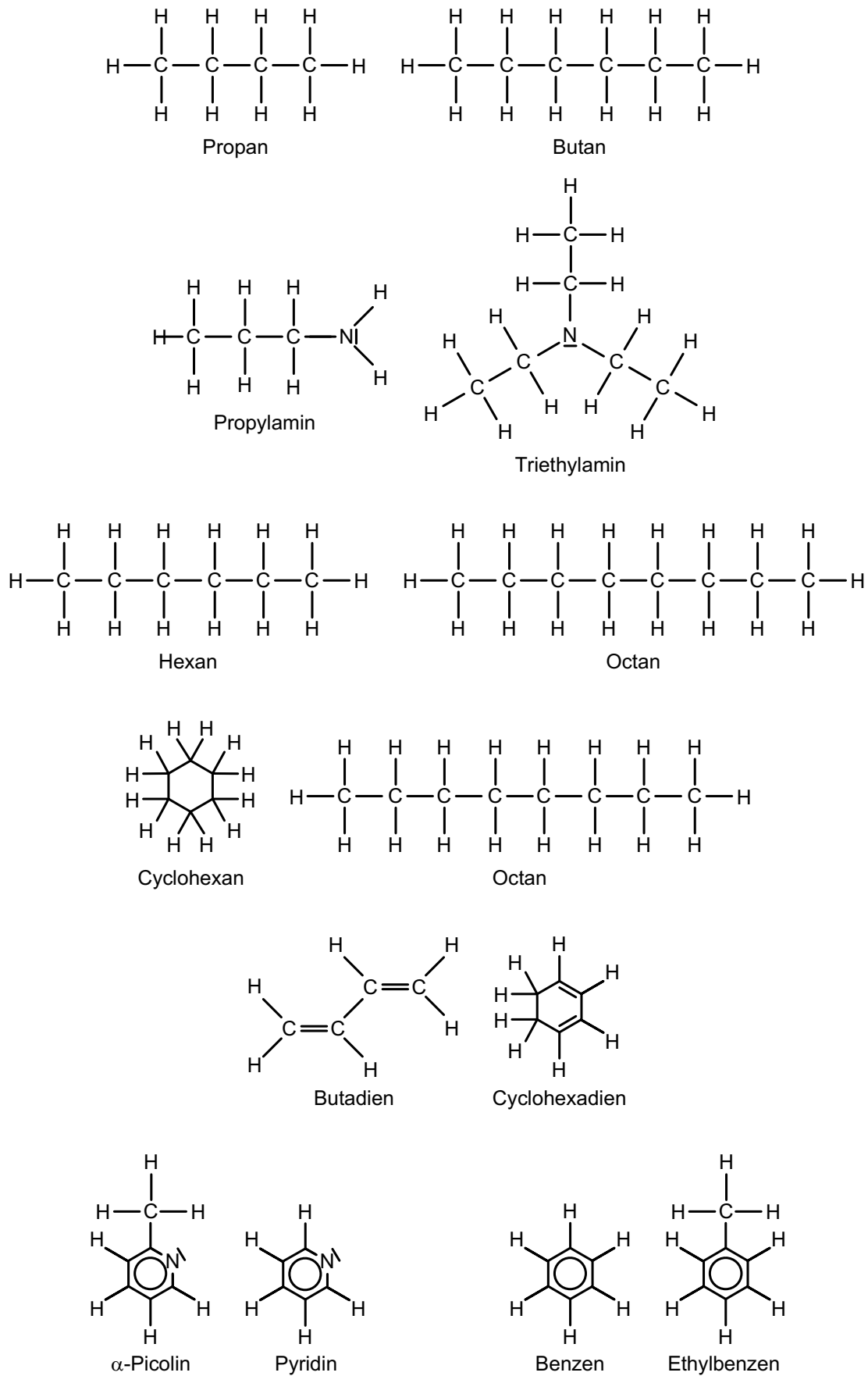


Abbildung 10.4-11: Mit dem GC-MIP-MED vermessene Verbindungen ähnlicher Struktur und etwas unterschiedlicher Zusammensetzung

Tabelle 10.4-2: Mit dem GC-MIP-MED vermessene Substanzen ähnlicher Struktur und etwas unterschiedlicher atomarer Zusammensetzung

Substanzname	Atomzahl pro Molekül			Atomzahl von C, H, O und N dividiert durch die gesamte Zahl an Atomen im Molekül		
	C	H	N	C	H	N
Propan	3	8	0	0,2727	0,7273	0,0000
Butan	4	10	0	0,2857	0,7143	0,0000
Propylamin	3	9	1	0,2308	0,6923	0,0769
Triethylamin	6	15	1	0,2727	0,6818	0,0455
Hexan	6	14	0	0,3000	0,7000	0,0000
Octan	8	18	0	0,3077	0,6923	0,0000
Octan	8	18	0	0,3077	0,6923	0,0000
Cyclohexan	6	12	0	0,3333	0,6667	0,0000
Butadien	4	6	0	0,4000	0,6000	0,0000
1,3-Cyclohexadien	6	8	0	0,4286	0,5714	0,0000
Picolin	6	7	1	0,4286	0,5000	0,0714
Pyridin	5	5	1	0,4545	0,4545	0,0909
Ethylbenzen	8	10	0	0,4444	0,5556	0,0000
Benzen	6	6	0	0,5000	0,5000	0,0000

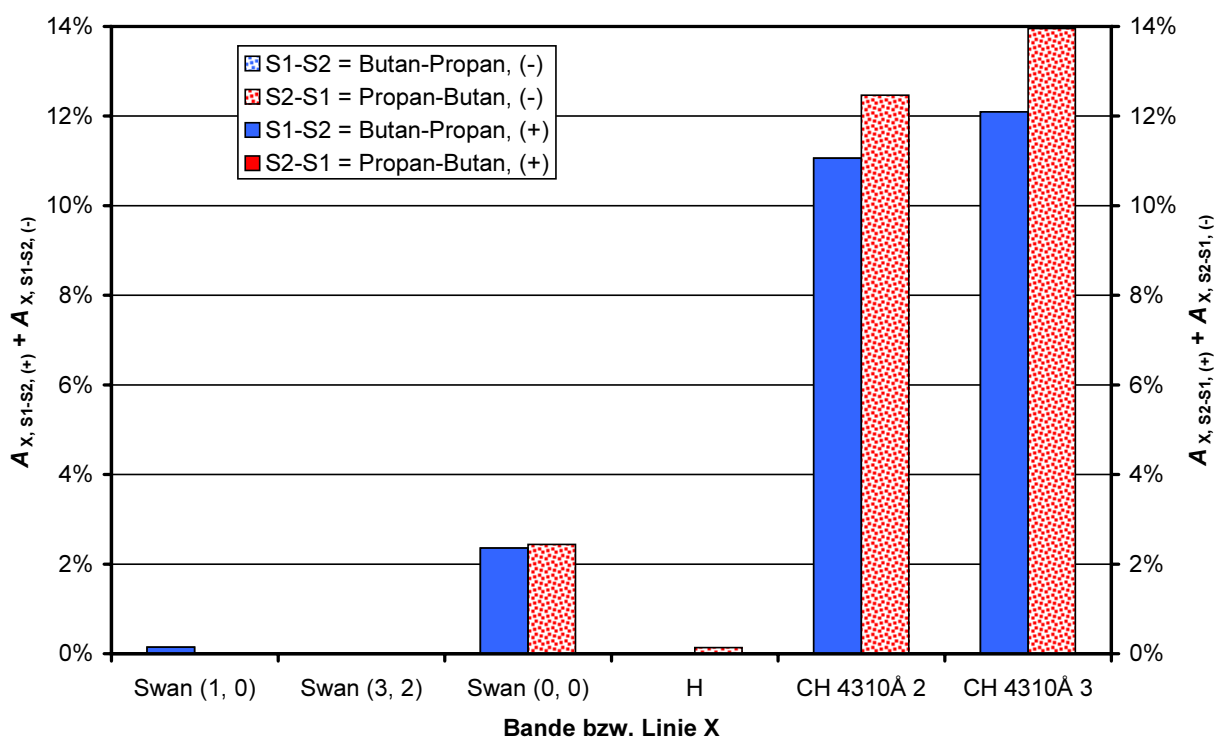


Abbildung 10.4-12: Prozentualer Anteil der bei der Abschätzung von Propan-Signalfächen mit Butan-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

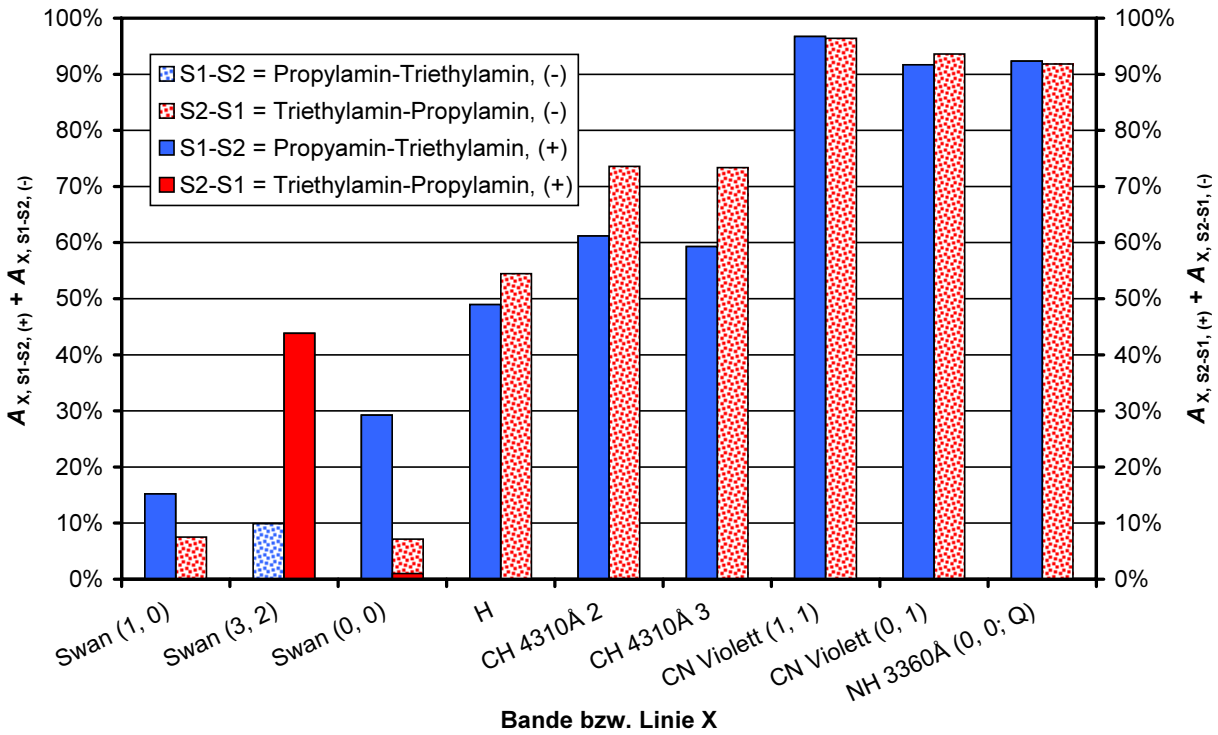


Abbildung 10.4-13: Prozentualer Anteil der bei der Abschätzung von Triethylamin-Signalfächen mit Propylamin-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

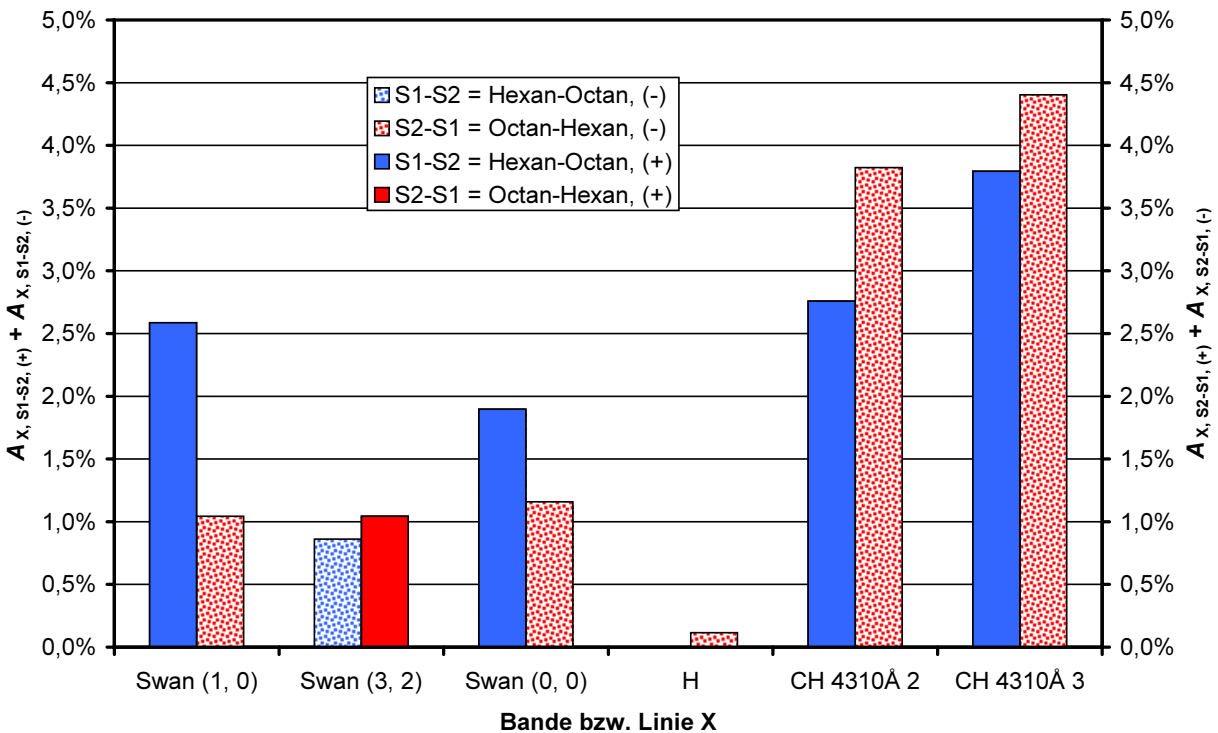


Abbildung 10.4-14: Prozentualer Anteil der bei der Abschätzung von Octan-Signalfächen mit Hexan-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

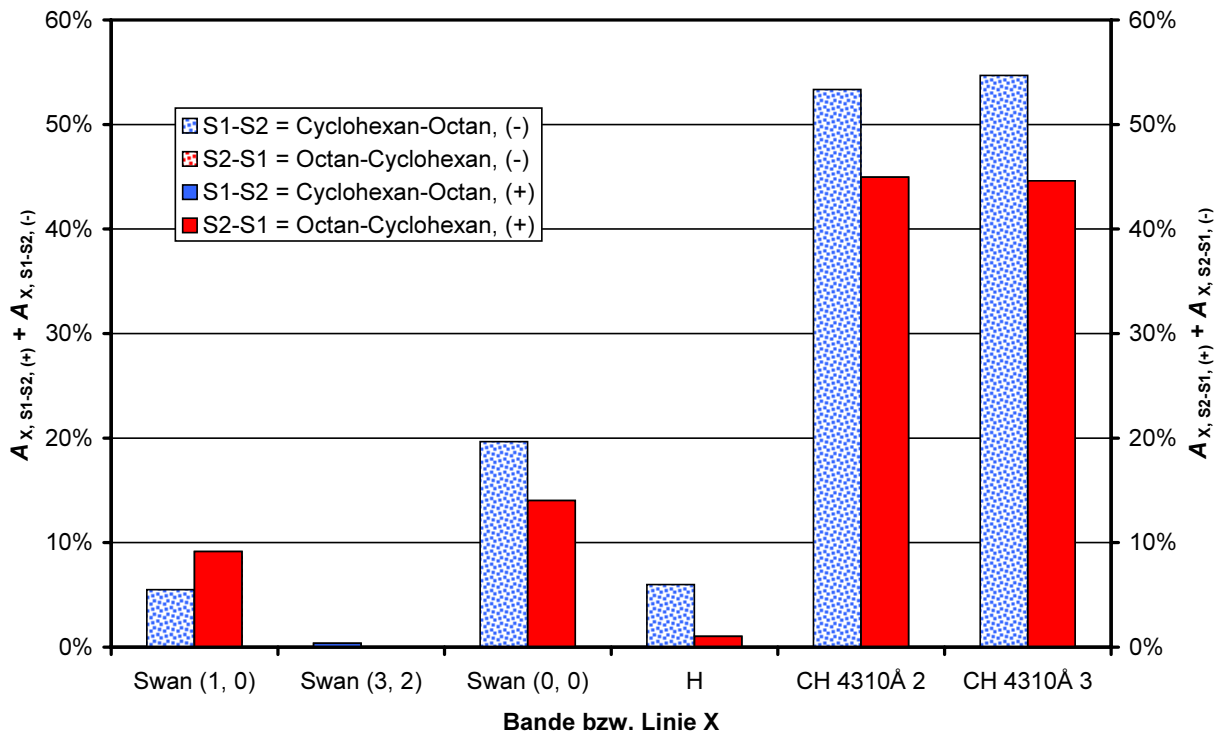


Abbildung 10.4-15: Prozentualer Anteil der bei der Abschätzung von Octan-Signalflächen mit Cyclohexan-Klassifizierungsparametern und umgekehrt ermittelten Signalflächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

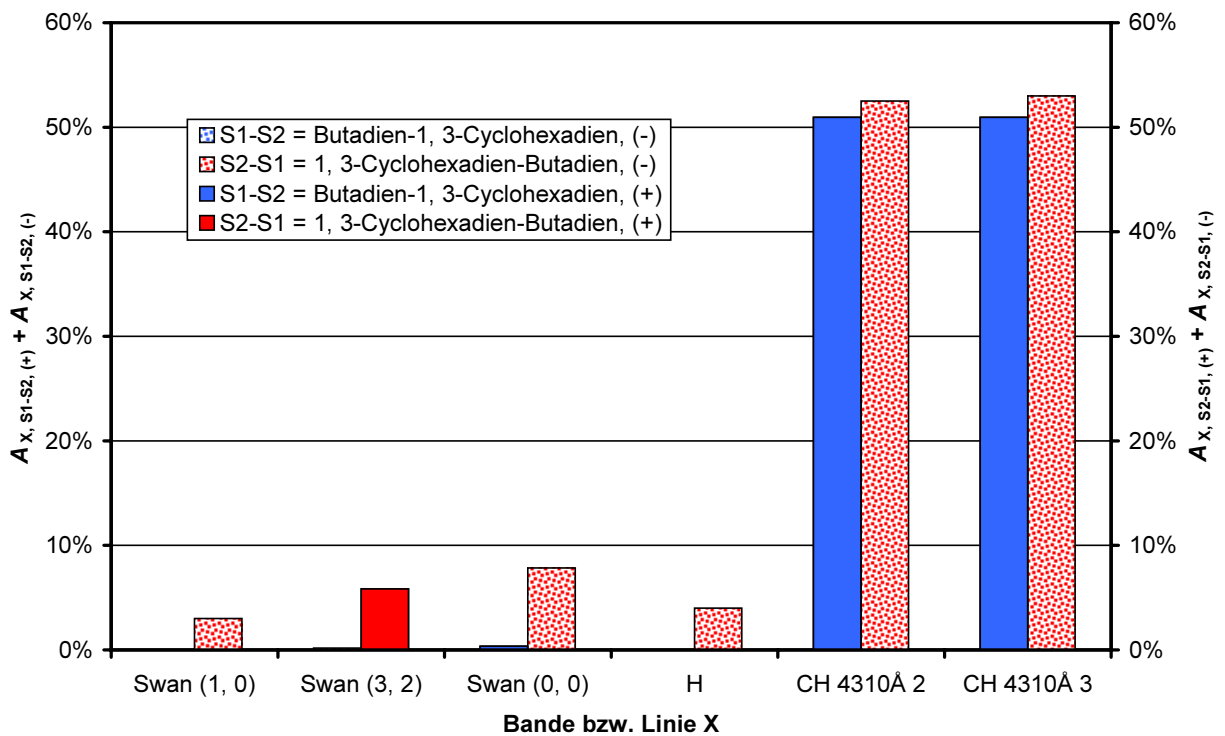


Abbildung 10.4-16: Prozentualer Anteil der bei der Abschätzung von 1,3-Cyclohexadien-Signalflächen mit Butadien-Klassifizierungsparametern und umgekehrt ermittelten Signalflächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

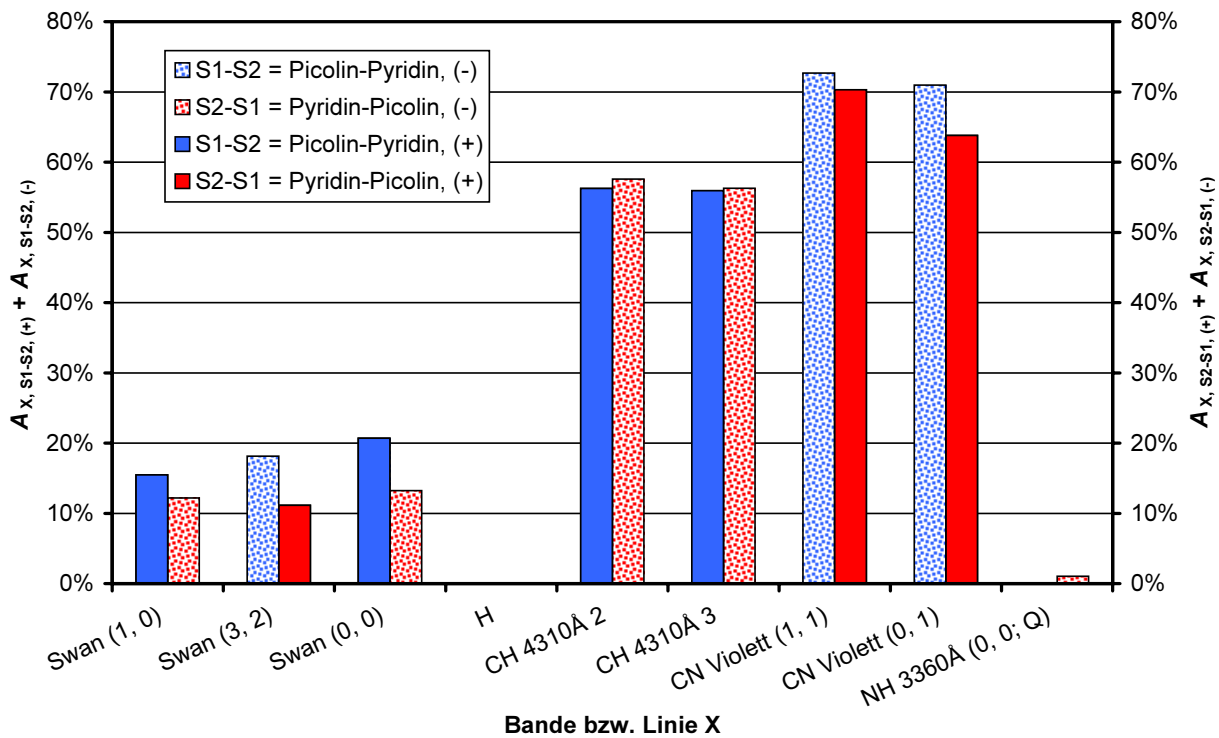


Abbildung 10.4-17: Prozentualer Anteil der bei der Abschätzung von Pyridin-Signalfächen mit  $\alpha$ -Picolin-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

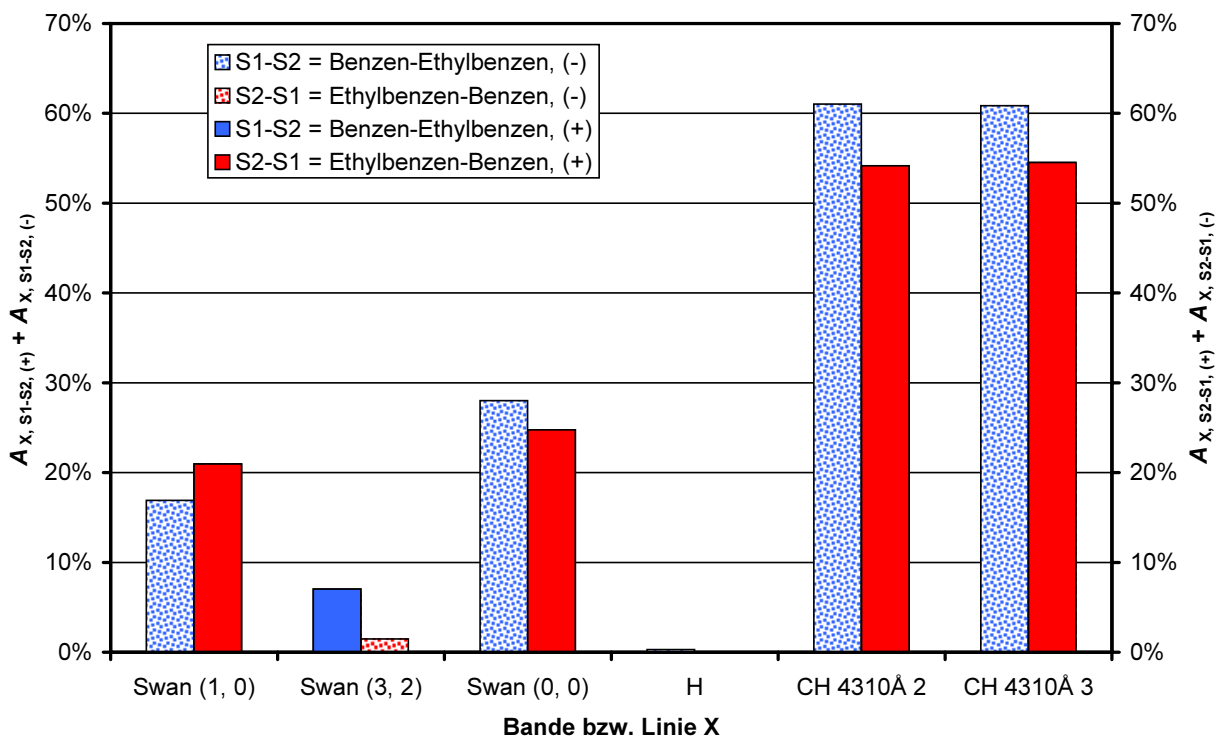


Abbildung 10.4-18: Prozentualer Anteil der bei der Abschätzung von Ethylbenzen-Signalfächen mit Benzen-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

### 10.4.3 Sauerstoffhaltige und nicht sauerstoffhaltige Verbindungen

Da das OH-Signal zu den Klassifizierungsparametern gehört, kann über die Größe von  $F_{N, OH\ 3064\text{\AA}}$  nicht auf den Sauerstoffgehalt eines Moleküls geschlossen werden. Allerdings könnte ein zusätzliches Sauerstoffatom in einem Molekül Einfluss auf die Spektraldichte der Swan-Banden, der CH-Banden sowie der H-Linie haben. Um diesen Einfluss nachzuweisen sind kurzkettige Alkanole mit ihren homologen Alkanen verglichen worden (siehe Tabelle 10.4-3 und Abbildung 10.4-19). Zusätzlich sind die Messdaten von Aceton und Cyclohexan untersucht worden. Beide Verbindungen besitzen 2 Wasserstoffatome pro Kohlenstoffatom (Tabelle 10.4-3, Atomzahlverhältnis). Aceton verfügt zusätzlich über ein Sauerstoffatom.

Anhand der Abweichungen von Messwerten und Schätzwerten bei der Abschätzung von Signalflächen mit den Klassifizierungsparametern der jeweils anderen Substanz können signifikante Unterschiede zwischen Alkanen und homologen Alkanolen sowie zwischen Aceton und Cyclohexan festgestellt werden (siehe Abbildung 10.4-20 bis Abbildung 10.4-24). Bei gleichen Klassifizierungsparametern besitzen die Alkanole tendenziell höhere Signalflächen für die Swan (1, 0)-Bande, die Swan (0, 0)-Bande, die CH 4310Å 2-Bande, die CH 4310Å 3-Bande sowie für die Wasserstofflinie. Für Swan (3, 2) sind für Alkanole hingegen niedrigere Signalflächen als für Alkane zu erwarten.

Auch die Signale der Acetonmessungen unterscheiden sich signifikant von denen der Cyclohexanmessungen (siehe Abbildung 10.4-24). Anhand des Diagramms wird deutlich sichtbar, dass analog zum Vergleich von Alkanolen mit Alkanen für die sauerstoffhaltige Verbindung, Aceton, höhere Signalflächen bei der Swan (1, 0)-, Swan (0, 0)-, CH 4310Å 2- und CH 4310Å 3-Bande sowie der Wasserstofflinie zu erwarten sind. Bei der Swan (3, 2)-Bande sind bei Aceton tendenziell niedrigere Signalflächen gemessen worden, als bei Cyclohexan.

Die Alkane sind zusätzlich mit dem nächst höheren homologen Alkanol verglichen worden. Hierbei ist festgestellt worden, dass sich Methan und Ethanol signifikant unterscheiden (siehe Abbildung 10.4-25), Ethan und Propanol sowie bei Propan und Butanol jedoch nicht (siehe Abbildung 10.4-26 sowie Abbildung 10.4-27). Zwischen Aceton und Octan bzw. Hexan und somit auch zwischen Aceton und Isooctan besteht ebenfalls kein signifikanter Unterschied bezüglich der gespeicherten GC-MIP-MES-Signale (siehe Abbildung 10.4-28 und Abbildung 10.4-29). Die Präzision des Detektors reicht beim bisherigen Entwicklungsstand noch nicht aus, um diese Substanzen anhand der gesammelten Daten zu unterscheiden.

Die allgemeine Erfahrung hat gezeigt, dass sauerstoffhaltige Verbindungen zu höheren OH 3064Å-Signalen neigen. Wie in Abschnitt 10.3.2 bereits erwähnt, korrelieren der Offset wie auch  $F_{N, OH\ 3064\text{\AA}}$  mit hoher Wahrscheinlichkeit mit der Halbwertsbreite und der Deformation des Signals. Der Offset spiegelt zudem die Menge an OH wieder, die kurz vor dem

Signalanstieg auf Grund von Memory-Effekten gebildet wird. Durch Erfahrungswerte, die man durch das Betrachten zahlreicher Chromatogramme sammelt, ist es wahrscheinlich möglich abzuschätzen, ob die dem Signal zu Grunde liegende Verbindung einen hohen Sauerstoffanteil hat oder nicht. Für einen Computer wäre eine solche Vorgehensweise jedoch ein Mustererkennungsproblem, welches nur mit sehr hohem Programmieraufwand gelöst werden kann.

Generell ist zu beobachten, dass die Unterschiede zwischen den sauerstoffhaltigen und nicht sauerstoffhaltigen Verbindungen um so größer werden, je höher der Sauerstoffanteil im Molekül ist. So wächst der Anteil an signifikant unterscheidbaren Signalen an, je kürzerkettige Alkanole man mit dem entsprechenden homologen Alkan vergleicht.

*Tabelle 10.4-3: Mit dem GC-MIP-MED vermessene sauerstoff- und nicht sauerstoffhaltige Verbindungen mit gleicher Zahl an C- und H-Atomen pro Molekül*

Substanzname	Atomzahl pro Molekül			Atomzahl von C, H, O dividiert durch die gesamte Zahl an Atomen im Molekül			Atomanzahlverhältnis H / C
	C	H	O	C	H	O	
Methan	1	4	0	0,2	0,8	0	4
Methanol	1	4	1	0,1667	0,6667	0,1667	4
Ethan	2	6	0	0,25	0,75	0	3
Ethanol	2	6	1	0,2222	0,6667	0,1111	3
Propan	3	8	0	0,2727	0,7273	0	2,667
Propanol	3	8	1	0,25	0,6667	0,0833	2,667
Butan	4	10	0	0,2857	0,7143	0,0714	2,5
Butanol	4	10	1	0,2667	0,6667	0,0667	2,5
Cyclohexan	3	6	0	0,3333	0,6667	0	2
Aceton	3	6	1	0,3	0,6	0,1	2

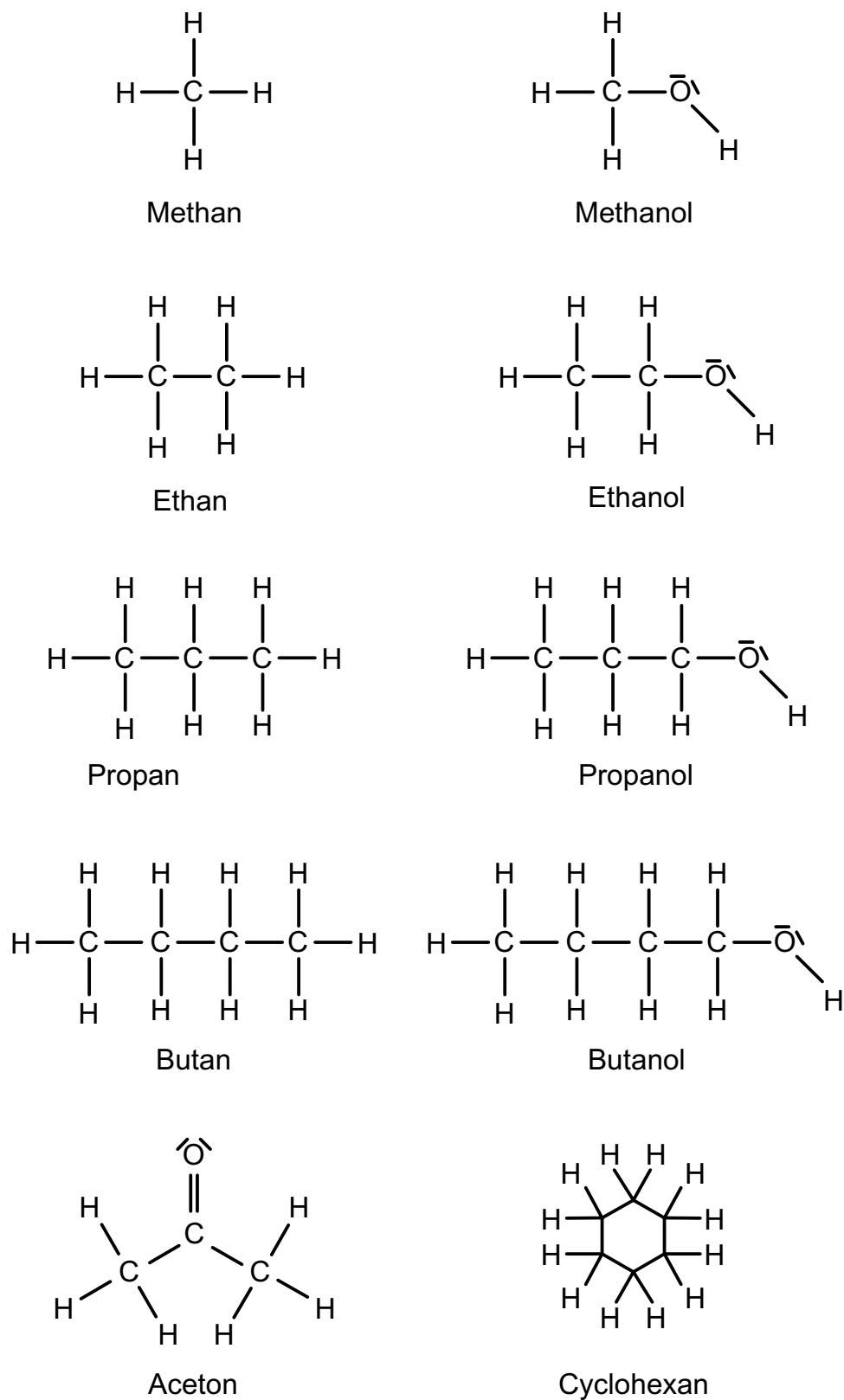


Abbildung 10.4-19: Sauerstoffhaltige und nicht sauerstoffhaltige Verbindungen, deren Moleküle anzahlmäßig das gleiche Verhältnis von H-Atomen zu C-Atomen besitzen



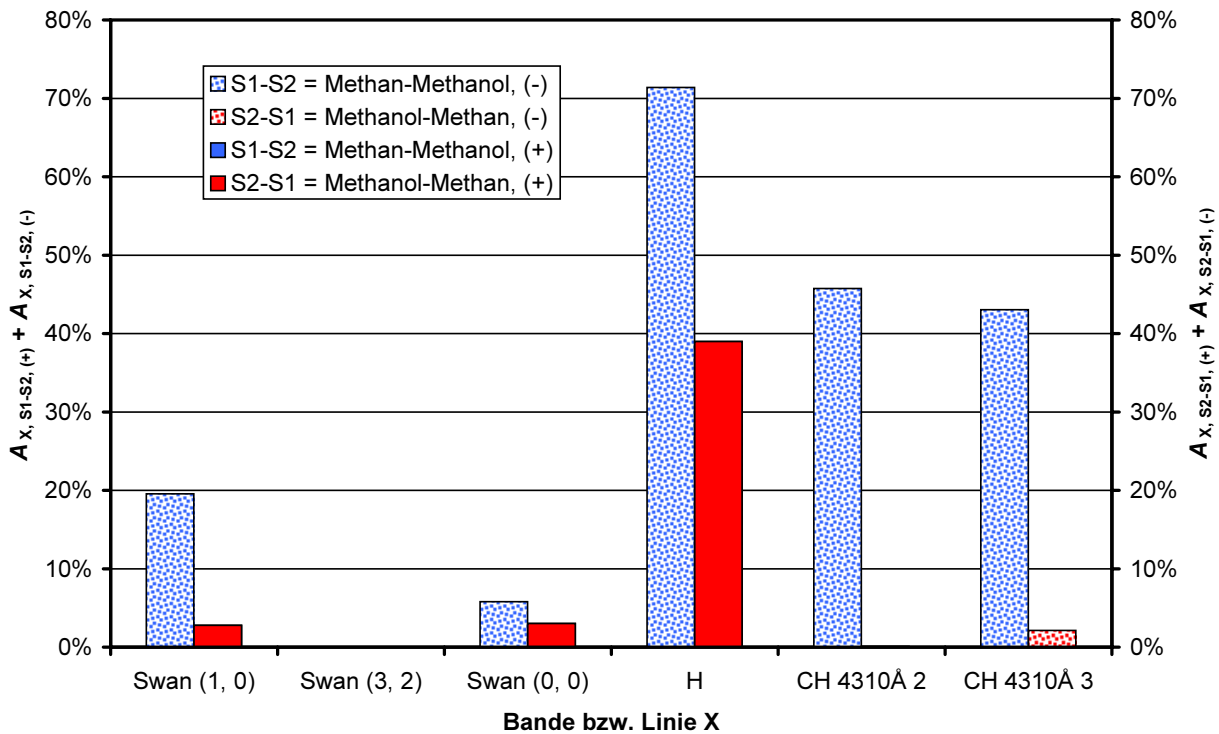


Abbildung 10.4-20: Prozentualer Anteil der bei der Abschätzung von Methanol-Signalfächen mit Methan-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

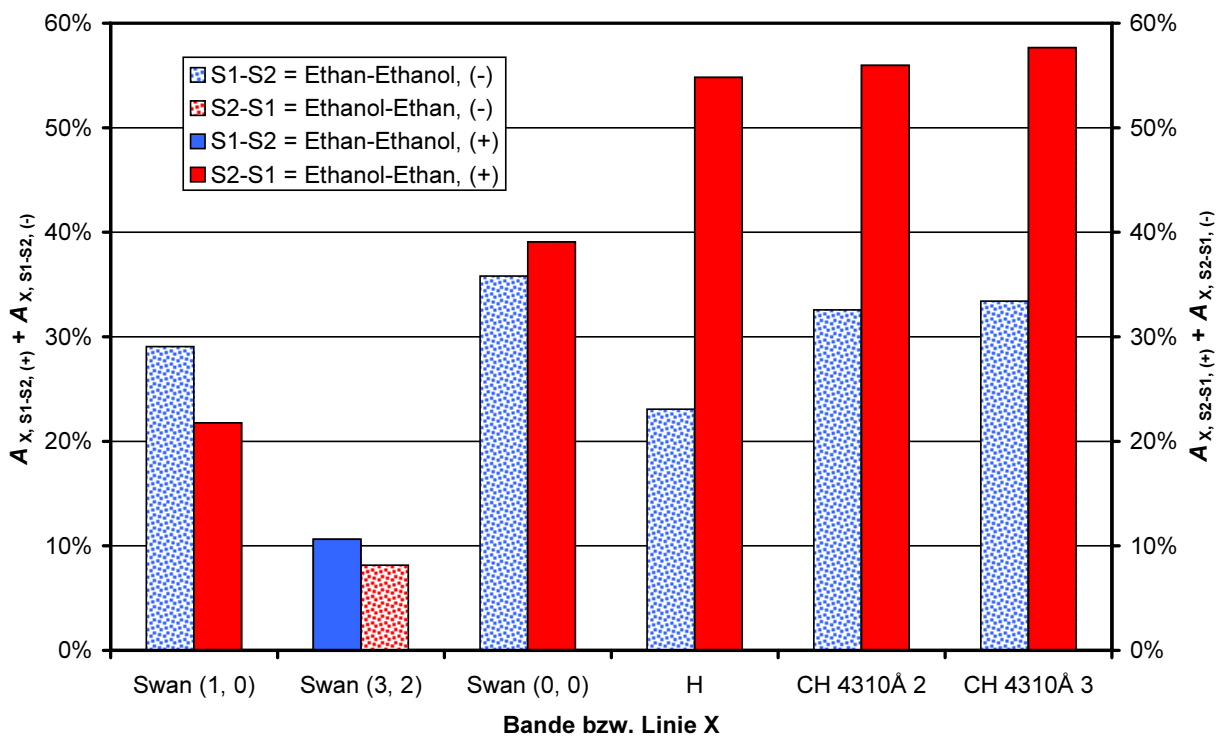


Abbildung 10.4-21: Prozentualer Anteil der bei der Abschätzung von Ethanol-Signalfächen mit Ethan-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

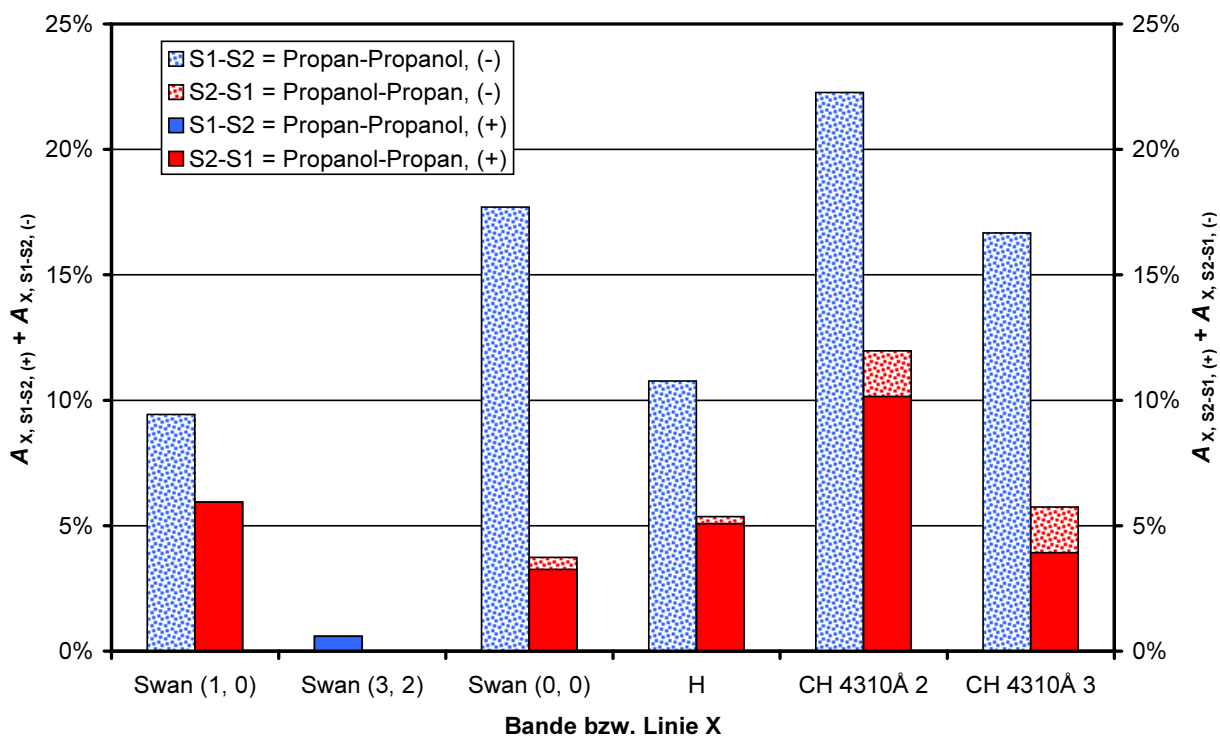


Abbildung 10.4-22: Prozentualer Anteil der bei der Abschätzung von Propanol-Signalflächen mit Propan-Klassifizierungsparametern und umgekehrt ermittelten Signalflächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

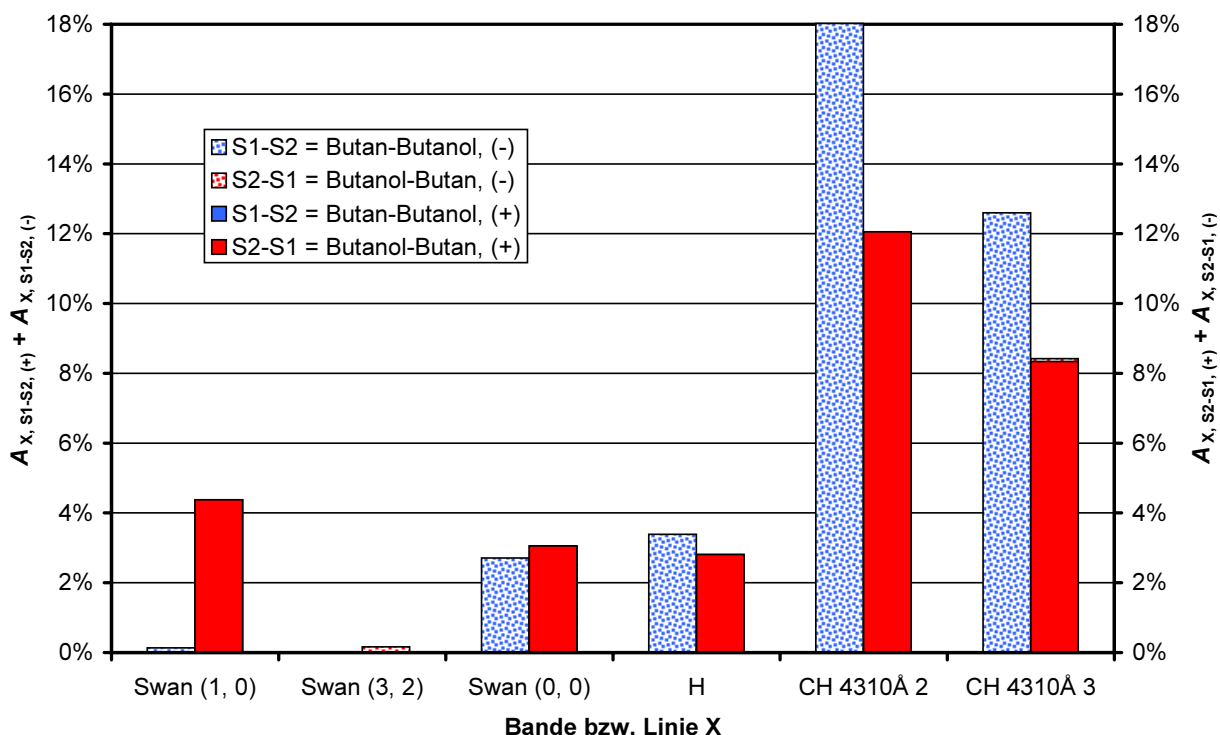


Abbildung 10.4-23: Prozentualer Anteil der bei der Abschätzung von Butanol-Signalflächen mit Butan-Klassifizierungsparametern und umgekehrt ermittelten Signalflächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

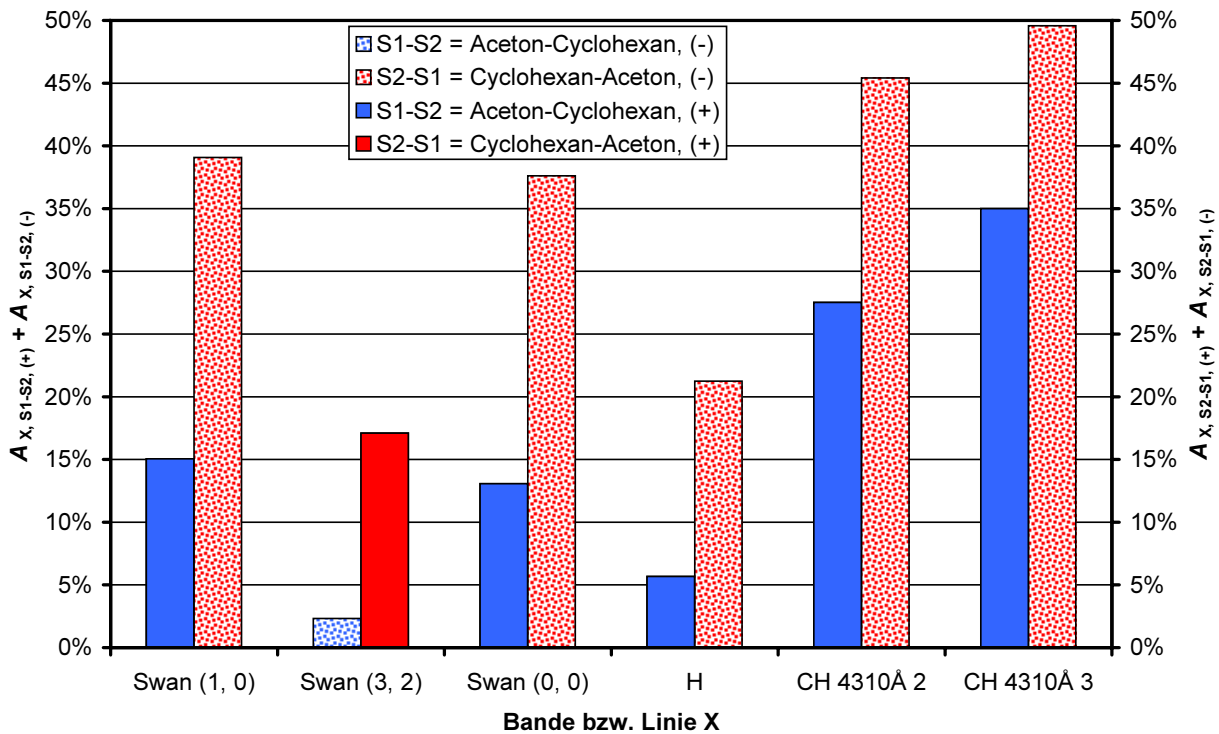


Abbildung 10.4-24: Prozentualer Anteil der bei der Abschätzung von Cyclohexan-Signalfächen mit Aceton-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

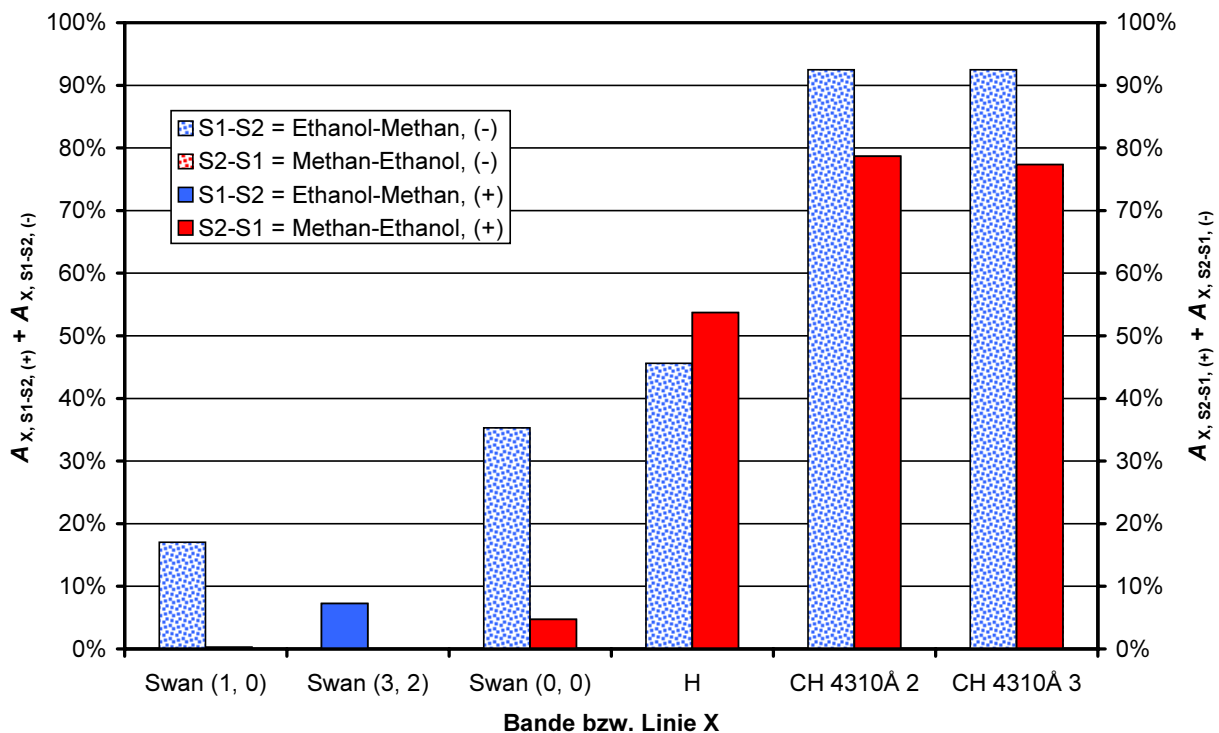


Abbildung 10.4-25: Prozentualer Anteil der bei der Abschätzung von Methan-Signalfächen mit Ethanol-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

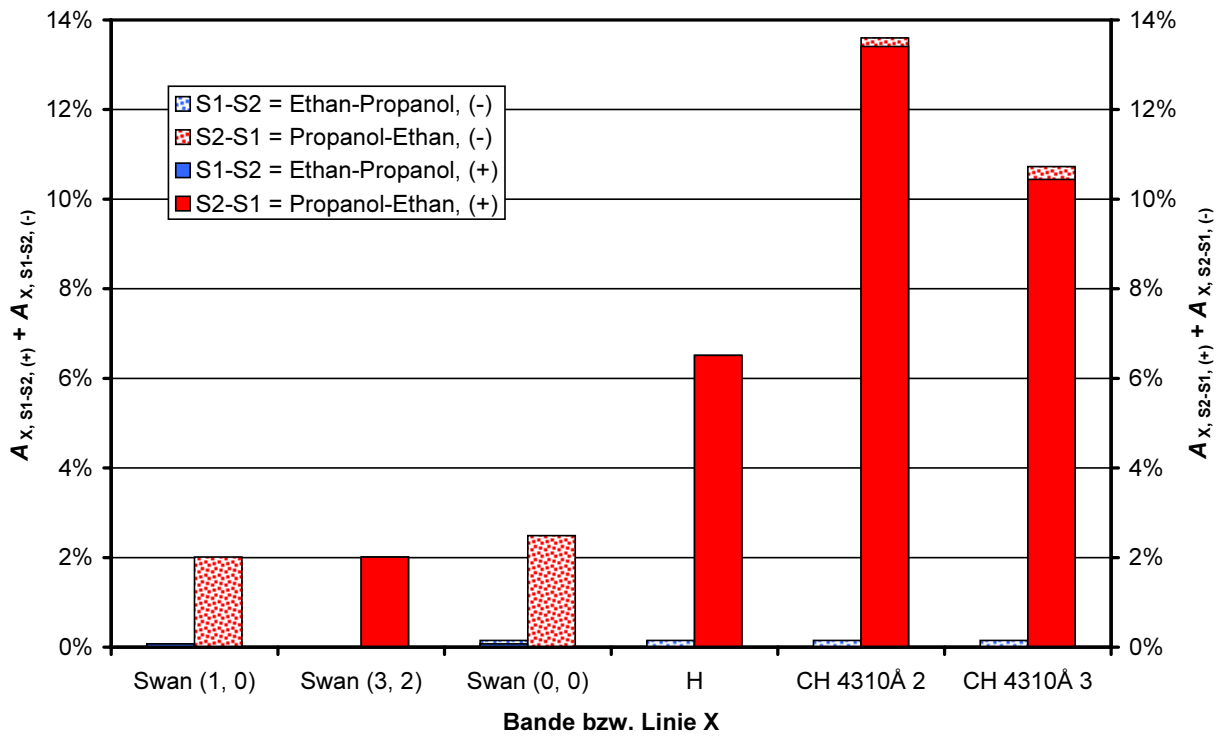


Abbildung 10.4-26: Prozentualer Anteil der bei der Abschätzung von Propanol-Signalfächen mit Ethan-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

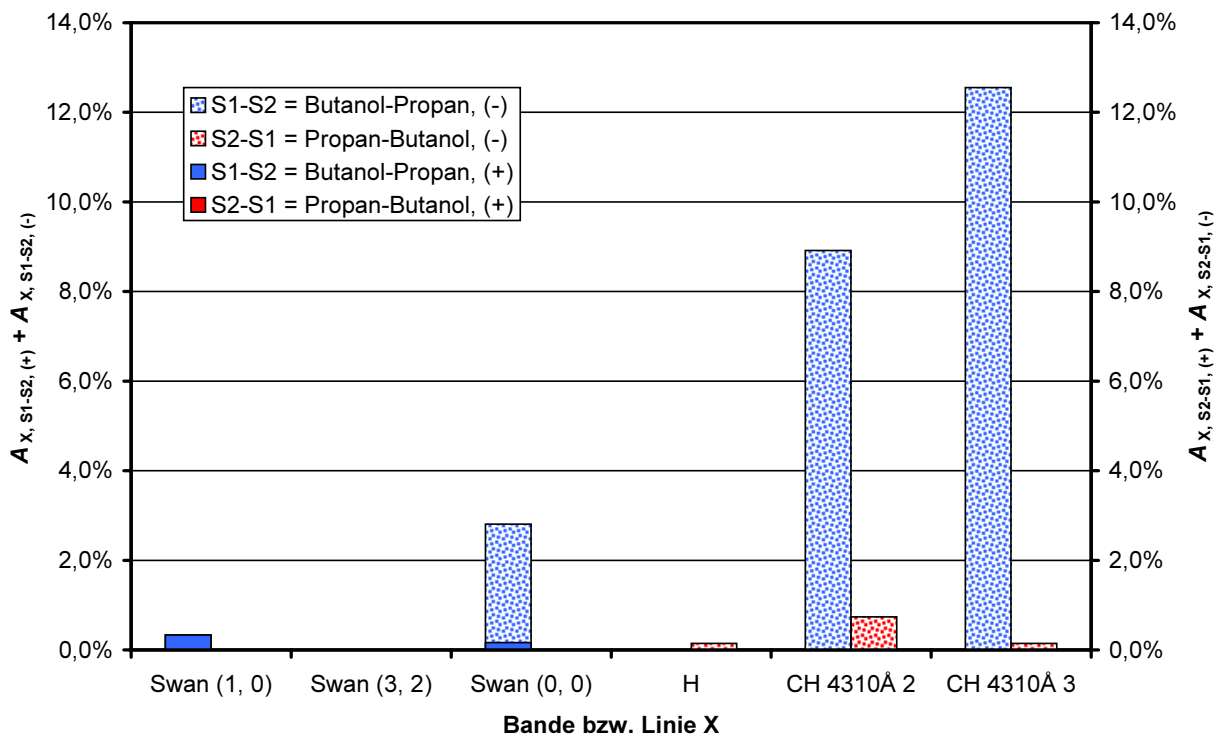


Abbildung 10.4-27: Prozentualer Anteil der bei der Abschätzung von Propan-Signalfächen mit Butanol-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

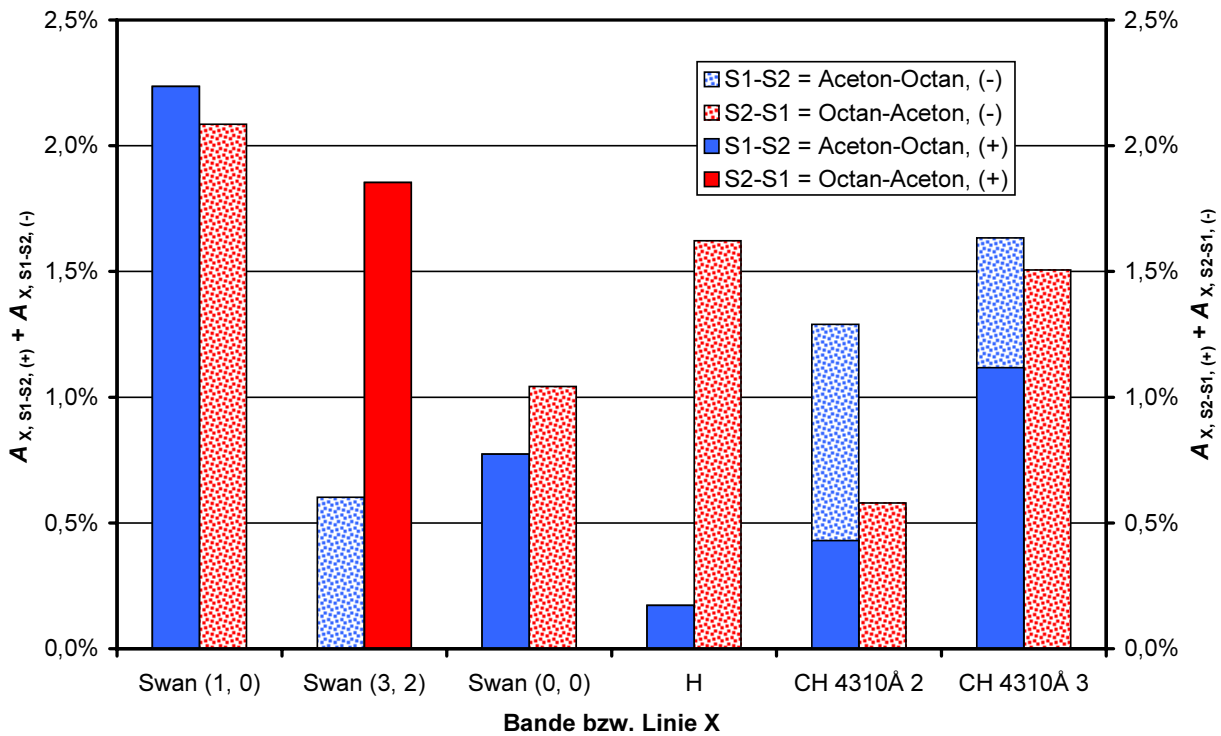


Abbildung 10.4-28: Prozentualer Anteil der bei der Abschätzung von Octan-Signalfächen mit Aceton-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

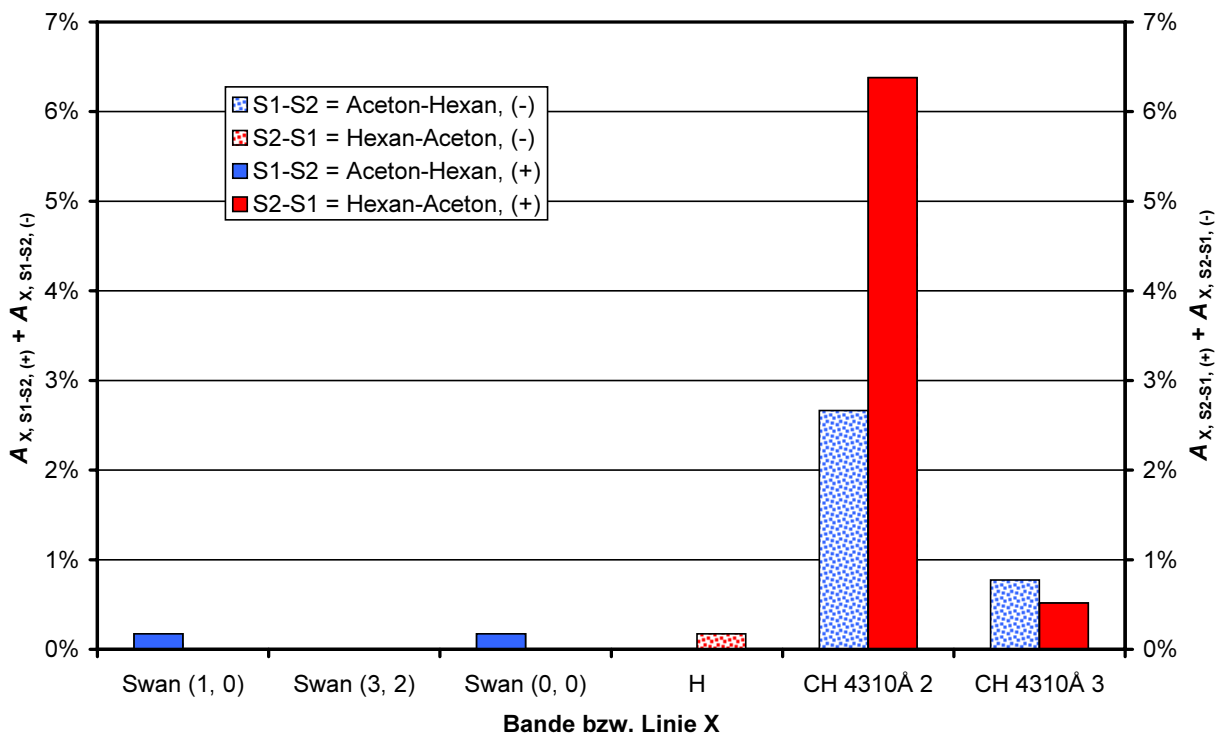


Abbildung 10.4-29: Prozentualer Anteil der bei der Abschätzung von Hexan-Signalfächen mit Aceton-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

### 10.4.4 1, 3-Cyclohexadien und Ethylbenzen

Ethylbenzen und 1, 3-Cyclohexadien unterscheiden sich in Zusammensetzung und Struktur. Im Ethylbenzen ist der Kohlenstoffanteil im Molekül etwas höher und somit der Wasserstoffanteil etwas geringer als im Cyclohexadien (siehe Tabelle 10.4-4). 1, 3-Cyclohexadien enthält nur zwei konjugierte Doppelbindungen, während Ethylbenzen über ein vollständiges aromatisches System verfügt (siehe Abbildung 10.4-11). Trotzdem unterscheiden sich die GC-MIP-MED-Signale dieser beiden Verbindungen nicht signifikant (siehe Abbildung 10.4-30). Der Einfluss der Zusammensetzung auf die GC-MIP-MED-Signale der beiden Verbindungen ist am Größten. Tabelle 10.4-5 zeigt den Unterschied in der Zusammensetzung zwischen Ethylbenzen und 1, 3-Cyclohexadien im Vergleich zum Zusammensetzungsunterschied zwischen Hexan und Octan sowie zwischen Propan und Butan. Hexan und Octan lassen sich wegen zu geringer Unterschiede in der atomaren Zusammensetzung anhand ihrer GC-MIP-MED-Signale nicht unterscheiden. Im Falle von Propan und Butan ist der Unterschied groß genug (siehe Abschnitt 10.4.2). Bei Ethylbenzen und 1, 3-Cyclohexadien ist die Differenz des Kohlenstoffgehaltes mit 0,0158 nicht nur mehr als doppelt so hoch wie bei Hexan und Octan, sondern auch etwas höher als der bei Propan und Butan anzutreffende Unterschied von 0,0130. Die Tatsache, dass sich Ethylbenzen und 1, 3-Cyclohexadien trotz des ausreichend erscheinenden Unterschiedes im Kohlenstoffgehalt anhand ihrer GC-MIP-MED-Signale nicht unterscheiden lassen kann so interpretiert werden, dass sich in diesem Fall die Einflüsse von Struktur und Zusammensetzung auf die GC-MIP-MED-Signale gegenseitig aufheben.

Tabelle 10.4-4: Atomare Zusammensetzung von Propan, Butan, Hexan, Octan, 1, 3-Cyclohexadien und Ethylbenzen

Substanzname	Atomzahl pro Molekül		Atomzahl von C und H dividiert durch die gesamte Zahl an Atomen im Molekül	
	C	H	C	H
Propan	3	8	0,2727	0,7273
Butan	4	10	0,2857	0,7143
Hexan	6	14	0,3000	0,7000
Octan	8	18	0,3077	0,6923
1,3-Cyclohexadien	6	8	0,4286	0,5714
Ethylbenzen	8	10	0,4444	0,5556

Tabelle 10.4-5: Unterschiede im Kohlenstoffgehalt ähnlich zusammengesetzter Verbindungen

Substanzpaar	Differenz im atomaren Gehalt an Kohlenstoff = Unterschied im Stoffmengenanteil von C pro Mol Substanz
1,3-Cyclohexadien und Ethylbenzen	$0,4444 - 0,4286 = 0,0158$
Hexan und Octan	$0,3077 - 0,3000 = 0,0077$
Butan und Propan	$0,2857 - 0,2727 = 0,0130$

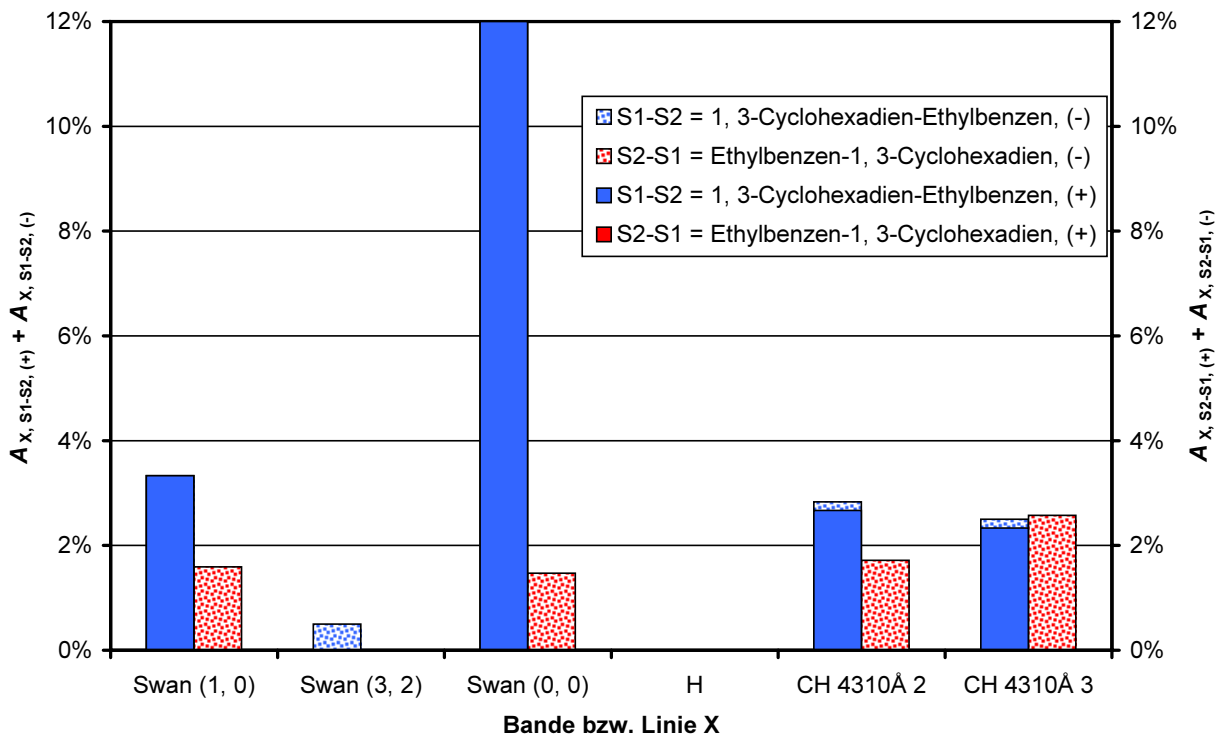


Abbildung 10.4-30: Prozentualer Anteil der bei der Abschätzung von Ethylbenzen-Signalfächen mit 1, 3-Cyclohexadien-Klassifizierungsparametern und umgekehrt ermittelten Signalfächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

## 10.5 Signalidentifikation mit Hilfe der Datenbanken

### 10.5.1 Die Arbeitsweise der Substanzidentifikationsfunktion

Das Datenbankmodul der Windows-Software ist mit einer Funktion zur Identifikation von Substanzen anhand ihrer chromatographischen MIP-MED-Signale ausgestattet worden. Dieses Modul schätzt die normierten Signalfächen der GC-MIP-MED-Signale auf Basis ihrer Klassifizierungsparameter mit allen in einer Datenbank vorhandenen Datensätzen ab. Die so gewonnenen  $F_{S, X}$ -Werte können als Vektor aufgefasst werden. Als Vergleichsvektor wird ein analog aufgestellter Vektor auf Basis der tatsächlich gemessenen Flächen  $F_{N, X}$  herangezogen. Anhand der Länge des Differenzvektors entscheidet die Software, zu welchem Datensatz das

gemessene Signal am besten passt. Der Datensatz, bei dem die geringste Differenzvektorenlänge gefunden wird, basiert am ehesten auf der gleichen Substanz, wie das untersuchte Signal. Es gilt also (siehe auch Gleichung (4.3-8) und (4.3-9)):

$$\left| \vec{F}_N - \vec{F}_S \right|_i = \left( \begin{array}{c} g_{\text{Swan}(1,0)} \cdot F_{N,\text{Swan}(1,0)} \\ g_{\text{Swan}(3,2)} \cdot F_{N,\text{Swan}(3,2)} \\ g_{\text{Swan}(0,0)} \cdot F_{N,\text{Swan}(0,0)} \\ g_H \cdot F_{N,H} \\ g_{\text{CH}431\text{Å}2} \cdot F_{N,\text{CH}4310\text{Å}2} \\ g_{\text{CH}4310\text{Å}3} \cdot F_{N,\text{CH}4310\text{Å}3} \end{array} \right) - \left( \begin{array}{c} g_{\text{Swan}(1,0)} \cdot F_{S,\text{Swan}(1,0)} \\ g_{\text{Swan}(3,2)} \cdot F_{S,\text{Swan}(3,2)} \\ g_{\text{Swan}(0,0)} \cdot F_{S,\text{Swan}(0,0)} \\ g_H \cdot F_{S,H} \\ g_{\text{CH}431\text{Å}2} \cdot F_{S,\text{CH}4310\text{Å}2} \\ g_{\text{CH}4310\text{Å}3} \cdot F_{S,\text{CH}4310\text{Å}3} \end{array} \right) \quad \text{ist minimal}$$

$\Rightarrow$  die Substanz, auf der der Datensatz  $i$  basiert, ist am ehesten auch die Substanz, auf der das zu untersuchende chromatographische MIP - MED - Signal basiert. (10.5-1)

$\vec{F}_N$  = Vektor aller als analytische Zielgröße verwendbaren normierten Signalflächen des zu untersuchenden chromatographischen Signals.

$\vec{F}_S$  = Vektor aller Signalflächen, die mit dem Datensatz  $i$  auf Grundlage der Klassifizierungsparameter des zu untersuchenden Signals abgeschätzt worden sind.

$g_X$  der zur Bande bzw. Linie X gehörende Gewichtungsfaktor.

Wird die CHON-Datenbank statt der CHO-Datenbank eingesetzt, sind die zwei Vektoren um die entsprechenden Signalflächen der CN- bzw. NH-Banden zu erweitern. Die Gewichtungsfaktoren werden nicht über eine Standardisierung ermittelt, sondern können in den Datenbankeinstellungen der Windows-Software frei gewählt werden.

Das in Abschnitt 10.3.2 beschriebene Verfahren zur Variablenreduktion stellt sicher, dass nur die bezüglich der Klassifizierungsparameter ähnlichsten Signale zur Regressionsrechnung herangezogen werden. Trotzdem kann es passieren, dass sich der Klassifizierungsparametersatz des zu bestimmenden Signals stark von den Klassifizierungsparametersätzen der zur Regressionsrechnung benutzten Messdaten unterscheidet. Dies ist immer dann der Fall, wenn trotz Variation der Betriebsparameter des Gaschromatographen mit einer bestimmten Substanz keine Signale erzeugt werden können, deren Klassifizierungsparameter denen des zu bestimmenden Signals ähneln. So können beispielsweise mit Methan keine so hohen  $F_{N,\text{Swan}(2,1)}$ -Werte wie mit Benzen erzeugt werden. Die im Verhältnis zum Benzen viel größeren Wasserstoff- und CH-Signale führen schon bei viel kleineren  $F_{N,\text{Swan}(2,1)}$ -Werten zum Überschreiten des Messbereiches der Apparatur. Aus diesem Grund kommt es zu



systematischen Fehlern. Daher ist die Substanzbestimmungsfunktion der Datenbanksoftware mit einer Überprüfungsroutine für den Wertebereich der Klassifizierungsparameter versehen worden. Gilt

$$\begin{array}{l}
 F_{N, \text{Swan } (2,1), U} \geq F_{N, \text{Swan } (2,1), \text{max}, i} \quad \vee \quad F_{N, \text{Swan } (2,1), U} \leq F_{N, \text{Swan } (2,1), \text{min}, i} \\
 \vee \quad F_{N, \text{OH } 3064\text{\AA}, U} \geq F_{N, \text{OH } 3064\text{\AA}, \text{max}, i} \quad \vee \quad F_{N, \text{OH } 3064\text{\AA}, U} \leq F_{N, \text{OH } 3064\text{\AA}, \text{max}, i} \\
 \vee \quad O_U \geq O_{\text{max}, i} \quad \vee \quad O_U \leq O_{\text{min}, i}
 \end{array} \quad (10.5-2)$$

so stellt die Software eine Verletzung des Wertebereiches der Klassifizierungsparameter fest. Die Klassifizierungsparameter des zu bestimmenden Signals der unbekanntes Substanz sind mit **U** indiziert. Der Maximalwert eines Klassifizierungsparameters in dem Datensatz **i** ist mit **max** gekennzeichnet und der Minimalwert mit **min**.

Die Software ist in der Lage mitzuteilen, bei wie vielen Datensätzen eine Verletzung des Wertebereiches nach Gleichung (10.5-2) bei der Substanzbestimmung aufgetreten ist. Dem Anwender wird ebenfalls mitgeteilt, ob bei dem Datensatz, welcher nach dem Kriterium in (10.5-1) dem zu bestimmenden Signal zuzuordnen ist, eine Wertebereichsverletzung aufgetreten ist. In diesem Fall werden dem Anwender Alternativvorschläge unterbreitet, welcher Substanz das zu bestimmende Signal noch zugeordnet werden könnte. Es wird die Substanz mit den zweit-, dritt-, viert-, ... ähnlichsten Signalen vorgeschlagen bis hin zu der Substanz, bei der bei der Substanzerkennung keine Wertebereichsverletzung mehr auftritt. Werden Einzelsignale bestimmt, so kann abgefragt werden, bei welchen Datensätzen konkret die Wertebereichsverletzungen aufgetreten sind. Zusätzlich werden für jeden Datensatz die Klassifizierungsparametersätze der 5 Signale ausgegeben, welche den Klassifizierungsparametersatz des zu bestimmenden Signals am ähnlichsten sind. Bei der automatischen Bestimmung mehrerer Signale stehen wegen des hohen Datenaufkommens diese zusätzlichen Informationen nicht zur Verfügung.

Um die Güte der Signalidentifikationsmethode abschätzen zu können, ist es sinnvoll, dass die Signalidentifikationsfunktion nicht nur bei Signalen verwendet werden kann, deren Parameter direkt aus dem Chromatogramm bestimmt worden sind (siehe Abbildung 8.1-2 in Abschnitt 8.1.2), sondern auch bei Signalen, deren Parameter in einem der Datensätze in der Datenbank gespeichert sind. Daher kann die Substanzidentifikationsfunktion auch vom Datenbankmodul aus aufgerufen werden (siehe Abbildung 8.1-3).

## 10.5.2 Die Güte der Signalerkennung

Unter der Güte der im vorherigen Abschnitt vorgestellten Substanzen ist die sogenannte Erkennungsrate zu verstehen. Es handelt sich hierbei um den Anteil an chromatographischen GC-MIP-MED-Signalen, welche von der Substanzerkennungsfunktion der Substanz zugeordnet werden, von der auch das Signal stammt. Werden zur Ermittlung der Erkennungsrate

Daten von Signalen verwendet, die bereits in der Datenbank gespeichert sind, so wird die so ermittelte Erkennungsrate im folgenden als Wiedererkennungsrates bezeichnet.

Die Güte der Substanzerkennung und somit die Erkennungsrate hängt zunächst einmal von den Gewichtungsfaktoren für die Regressanden ab (siehe Gleichung (10.5-1)). Daher sind für die CHO-Datenbank fünf Sätze an Gewichtungsparametern aufgestellt worden. Um die Optimierung der Auswahl der Gewichtungsparameter zu vereinfachen sind nur Gewichtungsparameter von der Größe Null und Eins gewählt worden. Tabelle 10.5-1 enthält die verwendeten Gewichtungssätze. In Satz 1 sind die Gewichtungsfaktoren aller für die qualitative Analytik verwendbaren Banden und der Gewichtungsfaktor der Wasserstofflinie Eins. Von diesem Totalansatz sind in den folgenden Sätzen mehr und mehr Regressanden aus dem Regressionsmodell entfernt worden, indem ihre Gewichtung auf Null gesetzt worden ist. In Satz 2 ist die normierte Fläche der Wasserstofflinie aus dem Ansatz entfernt worden, da sie stark fehlerbehaftet ist (siehe Abschnitt 10.3.3). Aus Satz 3 wurde zusätzlich die Fläche der Swan (3, 2)-Bande als Regressor eliminiert, da sie am wenigsten mit Veränderungen in der Substanzzusammensetzung korreliert (siehe Abschnitt 10.4.2). Im Satz 4 sind nur noch die Gewichtungsfaktoren der stark mit der Substanzzusammensetzung korrelierenden CH 4310Å 2 bzw. 3-Banden eins und in Satz 5 ist lediglich der Gewichtungsfaktor von CH 4310Å 3 eins.

Tabelle 10.5-2 enthält die prozentualen Wiedererkennungsrates für alle Substanzen, deren Messwerte sich in der CHO-Datenbank befinden, in Abhängigkeit zum verwendeten Satz an Gewichtungsfaktoren. Diese Wiedererkennungsrates basieren auf allen Messwerten, die in der Datenbank gespeichert sind. Durchschnittlich die höchste Wiedererkennungsrates findet man bei Satz 2. Die von Satz 3 ist nur wenig niedriger. Während die Wiedererkennungsrates bei Satz 1 auch noch recht hoch ist, bricht sie bei Satz 4 und 5 ein. Dieses Ergebnis war zu erwarten, da die Signalflächen aller Banden und auch die der Wasserstofflinie sich als substanzspezifische Größen erwiesen haben. Die hohe Unsicherheit bei der Bestimmung von  $F_{N,H}$  verhindert, dass die Wiedererkennungsrates bei Satz 1 am höchsten ist und die hohe Ähnlichkeit bei den Wiedererkennungsrates von Satz 2 und 3 ist auf die geringe Korrelation von  $F_{N,Swan(3,2)}$  mit der Substanzzusammensetzung zurückzuführen. Die Wiedererkennungsrates erhöhen sich, sobald nicht alle Signale aus der Datenbank zu deren Ermittlung dienen, sondern nur die, bei denen  $F_{N,Swan(2,1)}$  größer 50%-s ist (siehe Tabelle 10.5-3). Dieser Sachverhalt lässt sich dadurch erklären, dass mit zunehmender normierter Signalfläche der Swan (2, 1)-Bande der relative Fehler bei der Bestimmung der Signalflächen abnimmt (siehe Abschnitt 10.3.3, Abbildung 10.3-16). Auch hier ist der Satz 2 der günstigste Gewichtungsfaktorsatz und wird daher im Folgenden für alle Substanzerkennungsberechnungen verwendet.

Die Wiedererkennungsrates liegen im Durchschnitt mit 79,96% bzw. 83,07% (Satz 2 bzw. Satz 2 und Signale größer als 50%-s) hoch genug um von einer funktionstüchtigen Substanzerkennungsroutine sprechen zu können. Für einige Substanzen ist die Wiedererkennungsrates aber auffallend niedrig. Bei Butanol liegt sie je nachdem, ob Signale mit  $F_{N,Swan(2,1)}$  größer 50 %-s oder alle Signale verwendet werden, sogar nur bei 50,84% bzw.

47,65%. Hierbei muss man in Betracht ziehen, dass einige der Substanzen in der CHO-Datenbank sich bezüglich ihrer GC-MIP-MED-Signale nicht signifikant unterscheiden. Sieht man eine Substanz auch dann als wiedererkannt an, wenn die Datenbank eine der von ihr nicht signifikant unterscheidbaren Substanzen vorschlägt, so steigt die Wiedererkennungsrates für alle Substanzen auf durchschnittlich 93,20% (siehe Tabelle 10.5-4). Werden auf Grund von Wertebereichsverletzungen bei der Substanzerkennungsberechnung von der Datenbanksoftware Alternativvorschläge ausgegeben, so ist die richtige oder eine von der richtigen Substanz nicht unterscheidbare Substanz oft der letzte Alternativvorschlag. Bei diesem Vorschlag handelt es sich um die bezüglich der GC-MIP-MED-Signale ähnlichste Substanz, bei der in der Berechnung keine Wertebereichsverletzung aufgetreten ist. Dies gilt besonders, wenn Signale der Alkanole Ethanol, Propanol und Butanol sowie von Diethylether bestimmt werden (siehe Tabelle 10.5-4). Auffällig ist auch, dass die Wiedererkennungsrates bei diesen sauerstoffhaltigen Verbindungen besonders niedrig sind. Die Ursache könnte in den relativ hohen Fehlern, die bei der Signalfächenbestimmung für Signale dieser Verbindungen aufgetreten sind, liegen (siehe Tabelle 10.3-11 und Tabelle 10.3-12). Anders als beim Benzen, bei dem ebenfalls ein recht hoher Signalfächenbestimmungsfehler zu finden ist, ist hier die Verwechslungsgefahr durch ähnlich zusammengesetzte Verbindungen recht groß. So werden 9,58% aller Butanolsignale als Butan identifiziert. Dieser Wert entspricht 116 Verwechslungen. Bei 35 dieser 116 Butanol-Butan-Verwechslungen, also fast einem Drittel der Verwechslungen, lautet der letzte Alternativvorschlag Butanol oder der nicht von Butanol zu unterscheidende Diethylether.

Die Wiedererkennungsrates sind auch für die Messwerte in der CHON-Datenbank bestimmt worden (siehe Tabelle 10.5-5). Zu diesem Zweck ist ein weiterer Satz an Gewichtungsfaktoren zu definieren. Der Gewichtungsfaktorsatz Satz 2 ist um  $g_{\text{CN Violett}}(1, 1) = 1$ ,  $g_{\text{CN Violett}}(0, 1) = 1$  und  $g_{\text{NH } 3360\text{Å}}(0, 0; Q) = 1$  erweitert worden. Der neue Gewichtungsfaktorsatz wird Satz 6 genannt. Die Wiedererkennungsrates sind für die in der CHON-Datenbank gespeicherten Signale höher. Insbesondere gilt dies, wenn Substanzvorschläge der Software auch dann als richtig betrachtet werden, wenn eine von der richtigen Substanz nicht signifikant unterscheidbare Substanz vorgeschlagen wird. Diese Wiedererkennungsrates hängt nicht stark davon ab, wie viele Substanzen, deren Signale in der Datenbank gespeichert sind, sich signifikant unterscheiden lassen. Daher kann sie dazu verwendet werden, die Güte der Substanzerkennung bezüglich der beiden Datenbanken zu vergleichen. Die mit 97,17% für Satz 2 gegenüber 93,20% für die CHO-Datenbank erhöhte Wiedererkennungsrates bei den Signalen in der CHON-Datenbank hat ihre Ursache wohl zum einen in der gegenüber den Signalen in der CHO-Datenbank geringeren Fehlern bei der Messwerterfassung. Zum anderen ist auch die Anzahl an Substanzen geringer, die unterschieden werden müssen.

Tabelle 10.5-1: Verwendete Sätze an Gewichtungsfaktoren

Bezeichnung	Gewichtungsfaktoren					
	$g_{\text{Swan}}(1, 0)$	$g_{\text{Swan}}(3, 2)$	$g_{\text{Swan}}(0, 0)$	$g_{\text{H}}$	$g_{\text{CH 4310Å 2}}$	$g_{\text{CH 4310Å 3}}$
Satz 1	1	1	1	1	1	1
Satz 2	1	1	1	0	1	1
Satz 3	1	0	1	0	1	1
Satz 4	0	0	0	0	1	1
Satz 5	0	0	0	0	0	1

Tabelle 10.5-2: Wiedererkennungsraten bei der Substanzerkennung auf Basis der in der CHO-Datenbank gespeicherten Messwerte in Abhängigkeit vom Gewichtungsfaktorsatz

Substanzname	Wiedererkennungsrate in % für folgende Gewichtungsparametersätze:				
	Satz 1	Satz 2	Satz 3	Satz 4	Satz 5
1,3-Cyclohexadien	78,67	83,17	80,50	66,67	63,17
Aceton	71,11	72,83	71,71	62,17	36,89
Benzen	94,17	94,23	93,63	87,37	86,77
Butadien	98,96	99,48	99,14	98,62	97,93
Butan	84,15	93,36	92,55	88,08	73,71
Butanol	40,38	47,65	47,15	39,64	35,09
Cyclohexan	73,63	80,34	78,75	74,11	61,29
Diethylether	56,67	46,06	45,28	37,31	33,41
Dimethylether	97,04	98,56	98,48	98,03	93,10
Ethan	74,53	79,18	77,23	71,39	68,19
Ethanol	79,69	86,70	85,32	78,79	74,31
Ethen	91,04	96,32	95,36	88,96	55,20
Ethylbenzen	90,93	90,56	90,07	73,28	72,92
Hexan	56,55	60,69	58,79	52,24	47,59
Isooctan	50,90	58,46	55,10	46,22	32,29
Methan	100,00	99,87	99,87	97,57	96,09
Methanol	97,09	98,10	97,65	76,65	64,92
Octan	53,07	60,25	58,05	52,72	41,25
Propan	60,77	72,57	72,42	59,14	47,64
Propanal	70,70	76,58	74,62	70,04	61,55
Propanol	75,86	79,02	77,30	73,47	56,32
Propen	80,19	85,14	84,33	69,27	62,96
<b>Mittelwert</b>	<b>76,19</b>	<b>79,96</b>	<b>78,79</b>	<b>70,99</b>	<b>61,94</b>

*Tabelle 10.5-3: Wiedererkennungsraten bei der Substanzerkennung auf Basis der in der CHO-Datenbank gespeicherten Messwerte in Abhängigkeit vom Gewichtungsfaktorsatz für Signale, deren normierte Swan (2, 1)-Fläche größer als 50 %s ist*

Substanzname	Wiedererkennungsraten in % für folgende Gewichtungssätze:				
	Satz 1	Satz 2	Satz 3	Satz 4	Satz 5
1,3-Cyclohexadien	81,98	86,49	82,88	68,24	63,74
Aceton	78,93	75,88	75,27	66,87	41,90
Benzen	97,36	98,17	97,95	97,21	96,70
Butadien	99,79	100,00	100,00	100,00	100,00
Butan	86,78	93,95	93,15	92,04	83,44
Butanol	42,12	50,84	50,28	43,02	41,34
Cyclohexan	75,64	80,73	79,30	74,68	64,49
Diethylether	59,64	48,21	47,98	39,69	37,67
Dimethylether	98,90	99,21	99,21	99,37	99,21
Ethan	83,31	82,78	82,64	81,44	80,93
Ethanol	83,45	90,75	90,87	85,16	81,05
Ethen	94,99	97,39	96,99	95,19	61,92
Ethylbenzen	95,77	94,56	94,11	76,89	76,89
Hexan	65,34	69,83	68,33	62,59	60,60
Isooctan	57,29	63,73	61,19	48,47	36,78
Methan	100,00	100,00	100,00	100,00	100,00
Methanol	100,00	100,00	100,00	92,75	84,06
Octan	58,04	65,03	63,24	58,63	45,83
Propan	65,61	74,05	73,84	61,18	51,90
Propanal	84,05	89,57	88,52	83,76	76,30
Propanol	78,33	81,28	80,51	76,92	64,49
Propen	85,65	85,07	84,39	75,05	73,22
<b>Mittelwert</b>	<b>80,59</b>	<b>83,07</b>	<b>82,30%</b>	<b>76,33%</b>	<b>69,20%</b>

Tabelle 10.5-4: Wiedererkennungsraten unter Verwendung des Gewichtungsfaktorsatzes Satz 2 für die in der CHO-Datenbank gespeicherten Substanzen

Substanz (von dieser Substanz nicht signifikant unterscheidbare Substanzen)	Wiedererkennungsraten [%]		
	Substanz	Alle Substanzen*	Alternativvorschlag berücksichtigt**
1,3-Cyclohexadien (Ethylbenzen)	83,17	97,33	97,50
Aceton (Hexan, Isooctan, Octan, Propanal)	72,83	94,93	96,22
Benzen	94,23	94,23	94,95
Butadien	99,48	99,48	99,48
Butan	93,36	93,36	93,63
Butanol (Propan, Diethylether)	47,65	74,40	77,46
Cyclohexan (Ethen, Propen)	80,34	95,85	96,21
Diethylether (Butanol, Propan)	46,06	83,29	86,81
Dimethylether	98,56	98,56	98,71
Ethan (Propanol)	79,18	90,19	91,09
Ethanol	86,70	86,70	92,90
Ethen (Cyclohexan, Propen)	96,32	99,52	99,84
Ethylbenzen (Cyclohexadien)	90,56	96,94	96,94
Hexan (Aceton, Isooctan, Octan, Propanal)	60,69	89,48	91,21
Isooctan (Aceton, Hexan, Octan, Propanal)	58,46	97,60	97,72
Methan	99,87	99,87	99,87
Methanol	98,10	98,10	98,88
Octan (Aceton, Hexan, Isooctan, Propanal)	60,25	96,64	97,57
Propan (Butanol, Diethylether)	72,57	93,07	93,36
Propanal (Aceton, Hexan, Isooctan, Octan)	76,58	88,13	89,11
Propanol (Ethan)	79,02	86,02	89,27
Propen (Cyclohexan, Ethen)	85,14	96,81	97,22
<b>Mittelwert</b>	<b>79,96</b>	<b>93,20</b>	<b>94,36</b>

\* Die Substanz gilt auch dann als wiedererkannt, wenn die Datenbank eine der eingeklammerten Substanzen erkennt.

\*\* Die Substanz gilt auch dann als wiedererkannt, wenn die Datenbank eine der eingeklammerten Substanzen erkennt oder der Alternativvorschlag, bei dem keine Verletzung des Klassifizierungsparameterwertebereiches vorliegt, eine der Substanzen in der 1. Spalte ist. (Die Datenbank gibt so lange Alternativvorschläge aus, bis keine Wertebereichverletzung bei der Substanzerkennungsroutine mehr auftritt (siehe vorheriger Abschnitt).)

Tabelle 10.5-5: Wiedererkennungsraten für die in der CHON-Datenbank gespeicherten Substanzen

Substanz (von dieser Substanz nicht signifikant unterscheidbare Substanzen)	Wiedererkennungsraten [%]			
	Ursprungssubstanz		Alle Substanzen*	
	Satz 2	Satz 6	Satz 2	Satz 6
Anilin (Picolin)	76,41	89,25	96,85	100,00
Picolin (Anilin)	80,04	86,55	95,62	99,89
Propylamin	98,43	100,00	98,43	100,00
Pyridin	94,42	100,00	94,42	100,00
Triethylamin	98,21	99,89	98,21	99,89
Trimethylamin	99,48	100,00	99,48	100,00
<b>Mittelwert</b>	<b>91,16</b>	<b>95,95</b>	<b>97,17</b>	<b>99,96</b>

\* Die Substanz gilt auch dann als wiedererkannt, wenn die Datenbank eine der eingeklammerten Substanzen erkennt.

Die Substanzerkennungsfunktion ist auch dahingehend getestet worden, ob auch Signale, die aus einer chromatographischen Trennung gewonnen worden sind, richtig zugeordnet werden. Dieser Test ist mit Hilfe von chromatographischen Signalen durchgeführt worden, die durch Trennung eines Gemisches von Benzen und Propanal erzeugt worden sind. Die GC-MIP-MED-Signale von Benzen und Propanal unterscheiden sich sehr stark. So kann untersucht werden, ob der Memory-Effekt im GC-MIP-MED substanzspezifische Wechselwirkungseffekte verursacht. Diese würden die Brauchbarkeit des GC-MIP-MED für die Gaschromatographie in Frage stellen. In einem Gaschromatogramm muss ein mit ein und den selben Klassifizierungsparametern zu beschreibendes Signal immer den selben analytischen Zielgrößen zuzuordnen sein. Dies muss unabhängig davon sein, welche Substanz vor der Aufzeichnung des klassifizierten Signals im Detektor war. Die Ergebnisse dieses Tests lassen den Schluss zu, dass die Memory-Effekte im GC-MIP-MED keine nachweisbaren substanzspezifischen Wechselwirkungseffekte verursachen. Die Erkennungsrate von Benzen, also der Anteil an Benzensignalen, die dem Benzendatensatz zugeordnet werden, ähnelt mit 94,95% (Tabelle 10.5-6) gegenüber 94,23% (Tabelle 10.5-4) stark der Wiedererkennungsrate für die Benzensignale in der Datenbank. 94 der 99 Signale sind als Benzen wiedererkannt worden. Von den 100 Propanal-Signalen sind aber lediglich 61 als Propanal erkannt worden. Hinzu kommen 4 bzw. 3 Signale, die als Aceton- bzw. Hexan-Signale identifiziert worden sind. Insgesamt sind also 68 von 100 Signalen richtig erkannt oder mit einer mittels GC-MIP-MED nicht zu unterscheidenden Substanz verwechselt worden. Diese Wiedererkennungsquote liegt bei den Propensignalen aus dem Propanaldatensatz mit 88,13% (siehe Tabelle 10.5-4) deutlich höher. Die Ursache liegt darin, dass die Signale im Propanaldatensatz weniger häufig mit Cyclohexan verwechselt worden sind (siehe Tabelle 10.5-7). Bei den Signalen aus der chromatographischen Trennung kam es 23 mal zu einer solchen Verwechslung. Dies entspricht 23% aller Signale. Bei den Propanal-Signalen aus der Datenbank sind es nur 7% (65 von 918 Messungen). Aus diesem Grund sind die Propanalmesswerte aus der chromatographischen Trennung in einen Datensatz geladen worden. Dieser Datensatz ist mit dem Propanaldatensatz aus der CHO-Datenbank verglichen worden. Das Vergleichsergebnis visualisiert Abbildung

10.5-1. Die 918 in der Datenbank gespeicherten Propanalmessungen decken einen viel höheren Wertebereich bezüglich der Klassifizierungsparameter ab, als die 100 Messwerte, die auf den chromatographisch getrennten Propanalsignalen basieren. Daher kommt es zu recht hohen systematischen Fehlern, wenn die 918 Messwerte auf Basis des Datensatzes mit nur 100 Messwerten abgeschätzt werden (blaue Balken). Die Schätzung ist so unzuverlässig, dass sie nicht einmal zur Bestimmung, welche zu dem jeweiligen Datensatz gehörende Menge an Signalen zu höheren oder niedrigeren normierten Flächen neigt, eingesetzt werden sollte. Schätzt man die 100 Messwerte auf Basis der 918 Messwerte ab, so ist das Schätzergebnis nicht mit einem systematischen Fehler behaftet (rote Balken). Aus den roten Balken im Diagramm lässt sich ablesen, dass die Propanal-Peaks aus der chromatographischen Trennung zu niedrigeren CH-, Swan (1, 0)- und Swan (0, 0)- sowie zu höheren Swan (3, 2)-Signalflächen neigen. Die Größenordnung dieses Trends lässt aber auf keinen Fall eindeutig auf einen Unterschied zwischen den verglichenen Datensätzen schließen. Die blauen Balken sind alle viel zu weit von der 5%-Marke entfernt. Allerdings verursacht dieser Trend eine recht hohe Anzahl an Verwechslungen von Propanalsignalen mit Cyclohexansignalen.

*Tabelle 10.5-6: Wiedererkennungsraten unter Verwendung des Gewichtungsfaktorsatzes Satz 2 für Signale aus chromatographischen Trennvorgängen*

Substanz (von dieser Substanz nicht signifikant unterscheidbare Substanzen)	Erkennungsraten	
	Substanz	Alle Substanzen*
Benzen	94,95%	95,96%
Propanal (Hexan, Isooctan, Octan, Aceton)	61,00%	68,00%

\* Die Substanz gilt auch dann als Wiedererkannt, wenn die Datenbank eine der eingeklammerten Substanzen erkennt.

*Tabelle 10.5-7: Anzahl der Verwechslungen bei der Bestimmung von Propensignalen aus der CHO-Datenbank (918 Signale) und von Propensignalen aus einem chromatographischen Trennvorgang (100 Signale)*

Datensatz aus	Verwechselt mit				
	1, 3-Cyclohexadien	Benzen	Butadien	Butan	Butanol
chromatographischer Trennung	3	0	0	0	2
Einzelmessungen	0	1	3	1	1
Datensatz aus	Verwechselt mit				
	Cyclohexan	Diethylether	Ethan	Ethanol	Ethen
chromatographischer Trennung	24	0	0	0	1
Einzelmessungen	65	1	1	1	8
Datensatz aus	Verwechselt mit				
	Ethylbenzen	Propan	Propanol	Propen	
chromatographischer Trennung	1	0	0	1	
Einzelmessungen	0	1	2	24	



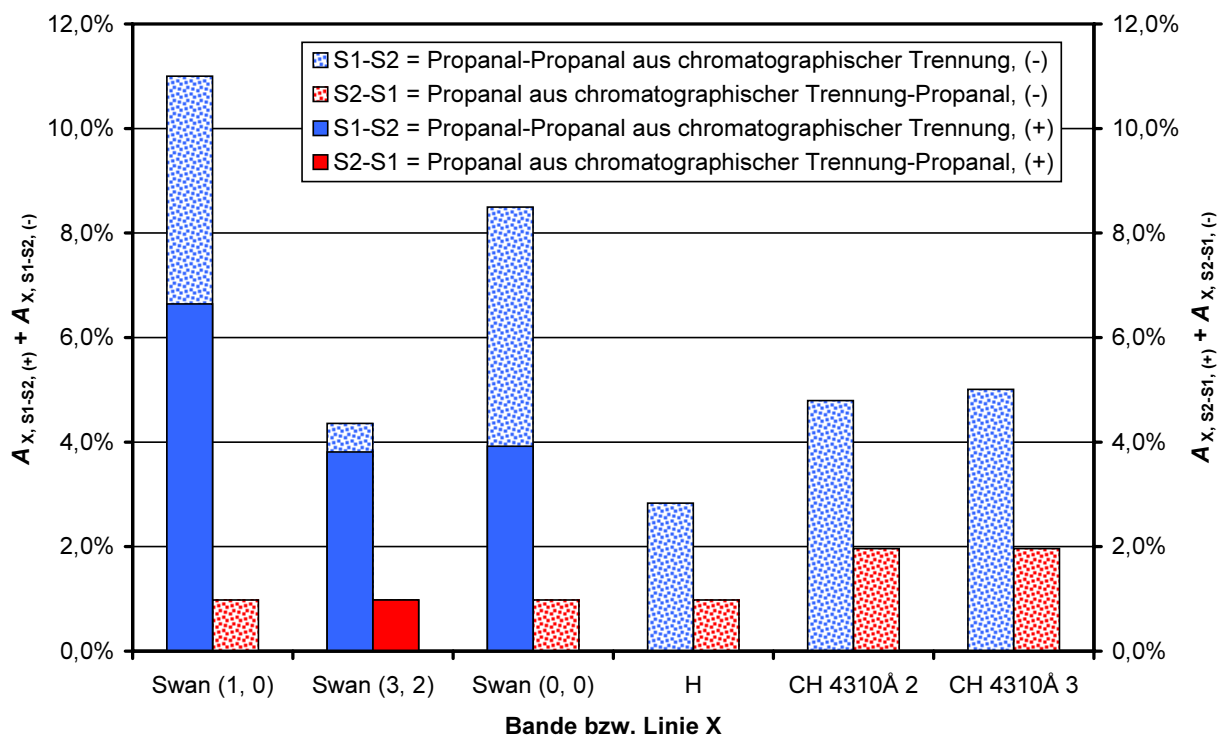


Abbildung 10.5-1: Prozentualer Anteil der bei der Abschätzung von Propanal-Signalflächen aus einer chromatographischen Messung mit Propanal-Klassifizierungsparametern aus dem Propanaldatensatz und umgekehrt ermittelten Signalflächen, welche gegenüber den gemessenen Flächen signifikant zu groß bzw. zu klein sind

## 10.6 Homologe Reihen

Wie bereits festgestellt, basieren die Unterschiede bei den GC-MIP-MED-Signalen verschiedener Substanzen zum größten Teil auf deren unterschiedlicher Zusammensetzung. Strukturelle Merkmale spielen nur bei Isomerenpaaren eine größere Rolle, bei denen die Unterschiede auf nicht an Kohlenstoff gebundene Methylgruppen basieren. Ansonsten ist nur noch bei dem Substanzpaar 1, 3-Cyclohexadien und Ethylbenzen ein Einfluss der Struktur auf die GC-MIP-MED-Signale beobachtet worden. Für Verbindungen, die nur aus Kohlenstoff und Wasserstoff bestehen, lässt sich ein funktioneller Zusammenhang zwischen dem Verhältnis der Zahl der H- zu der Zahl der C-Atome und den als analytische Zielgrößen fungierenden Signalflächen feststellen (siehe Abbildung 10.6-1). Die abgeschätzten Signalflächen für das Beispielsignal mit den Klassifizierungsparametern  $F_{N \text{ Swan } (2, 1)} = 103\% \cdot s$ ,  $F_{OH \text{ 3064Å}} = 270\% \cdot s$  und  $O = 12,3\%$  korrelieren mit einem Polynom 2. Grades. Der Zusammenhang zwischen abgeschätzter Signalfläche und dem Atomverhältnis von C und H ist annähernd linear, d. h. die Krümmung der Trendlinien im Diagramm ist gering. Die Ergebnisse der Berechnungen lassen darauf schließen, dass man kurzkettenige Alkane mit Hilfe des GC-MIP-MED gut unterscheiden kann. Der Unterschied in der atomaren Zusammensetzung zwischen benachbarten Homologen schrumpft, je größer die Kettenlänge ist. Die Unterschiede zwischen Octan bzw. Isooctan (Atomzahlverhältnis H zu C liegt bei 2,25) und Cyclohexan (Atomzahlverhältnis 2), welches

als Alkan mit unendlicher Kettenlänge angesehen werden kann, sind vergleichsweise gering. Daher wird die Bestimmung der Kettenlänge mittels GC-MIP-MED mit wachsender Molekülgröße immer schwieriger. Mit der in dieser Arbeit eingesetzten Apparatur scheitert man bereits daran, Octan und Hexan signifikant zu unterscheiden (siehe Abbildung 10.4-14).

Abbildung 10.6-1 lässt sich dahingehend erweitern, dass man die restlichen Verbindungen aus der CHO-Datenbank, welche sich nur aus C und H zusammensetzen, hinzufügt. Aus der erweiterten Abbildung 10.6-2 wird noch einmal ersichtlich, dass vor allem die Zusammensetzung und weniger stark die Struktur einer Verbindung ihr GC-MIP-MED-Signal bestimmt.

Betrachtet man im Vergleich zu den Alkanen die Abschätzungsergebnisse der homologen Alkanole, so ergibt sich ein ähnliches Bild (siehe Abbildung 10.6-4). Auch hier werden die Unterschiede benachbarter Homologen mit steigender Kettenlänge hinsichtlich der Zusammensetzung und somit auch bezüglich ihrer GC-MIP-MED-Signale immer geringer. Der mit kürzerer Kettenlänge wachsende Einfluss des Sauerstoffatoms (siehe Abschnitt 10.4.3) auf die Messergebnisse führt nur bei den kurzkettingen Alkanolen zu einer besseren Differenzierbarkeit als bei den Alkanen.

Die für die Berechnungen herangezogenen Klassifizierungsparametersätze wurden so ausgewählt, dass sie den Klassifizierungsparametersätzen aller verwendeten Datensätze möglichst ähnlich sind. Zu diesem Zweck sind Klassifizierungsparametersätze untersucht worden, die zu bereits in der Datenbank gespeicherten Messungen gehören. Die Datenbanksoftware ist in der Lage für jeden Datensatz zu überprüfen, ob ein Klassifizierungsparameter eines zur Signalflächenabschätzung herangezogenen Parametersatzes überhaupt innerhalb des Wertebereiches der entsprechenden Klassifizierungsparameter des Datensatzes liegt (siehe Abschnitt 10.5.1, Gleichung (10.5-2)). Mit Hilfe dieser Funktion kann innerhalb der CHO-Datenbank ein Signal gesucht werden, bei dem möglichst wenig Probleme mit dem Wertebereich der Klassifizierungsparameter auftreten. Dessen Klassifizierungsparametersatz kann dann zur Abschätzung von GC-MIP-MED-Signalen verschiedener Substanzen verwendet werden. Der Wertebereich sollte nicht nur für den einzelnen Klassifizierungsparameter eingehalten werden. Auch die Kombination der Klassifizierungsparameter sollte für jeden Datensatz der des abzuschätzenden Signals ähneln. Aus diesem Grund sind für jeden Datensatz die 5 ähnlichsten Klassifizierungsparametersätze ausgegeben und als Auswahlkriterium mit einbezogen worden. Zugleich sollte  $F_{N_{\text{Swan}}(2,1)}$  eine ausreichende Größe besitzen, damit der relative Signalflächenbestimmungsfehler klein ist. Auf Grund der Komplexität des Auswahlverfahrens und der hohen Anzahl an Signalen konnte der am besten für einen Substanzvergleich geeignete Parametersatz nicht über einen Softwarealgorithmus gesucht werden. Statt dessen wurden einige Signale, bei denen nur für wenige Datensätze Wertebereichsverletzungen auftraten, stichprobenartig auf Eignung untersucht. Tabelle 10.6-2 und Tabelle 10.6-3 zeigen, dass bei den zur Berechnung verwendeten Klassifizierungsparametersätzen keine sehr großen Probleme mit dem Wertebereich zu erwarten sind. Wählt man hingegen ein Signal aus, bei dem die Software bei fast allen Datensätzen eine Wertebereichsverletzung feststellt, so unterscheiden sich auch die 5 ähnlichsten Klassifizierungsparametersätze sehr stark von dem Klassifizierungsparametersatz, mit dem die Abschätzung durchgeführt worden ist (siehe Tabelle 10.6-1).

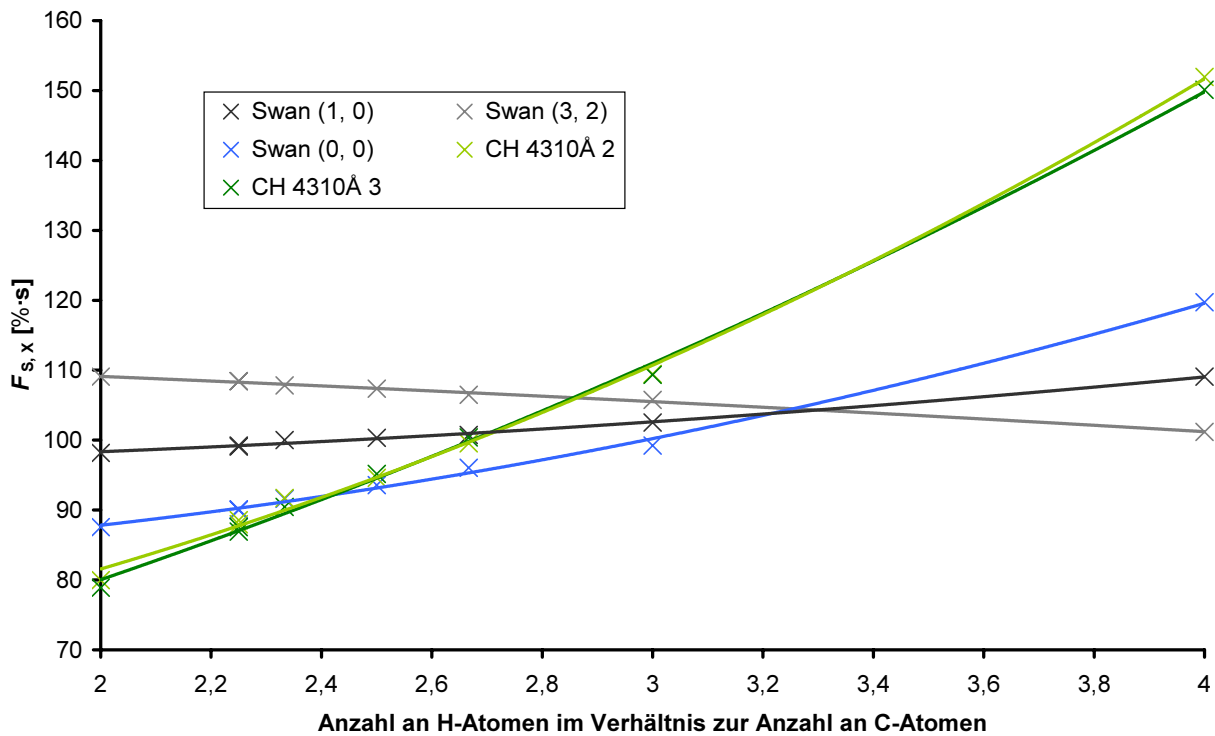


Abbildung 10.6-1: Abgeschätzte Signalflächen verschiedener Banden in Abhängigkeit vom Verhältnis der H- und C-Atomzahlen im Molekül für Alkane und Cyclohexan wobei  $F_{N, \text{Swan } (2, 1)} = 103\% \cdot s$ ,  $F_{N, \text{OH } 3064\text{Å}} = 270\% \cdot s$  und  $O = 12,3\%$

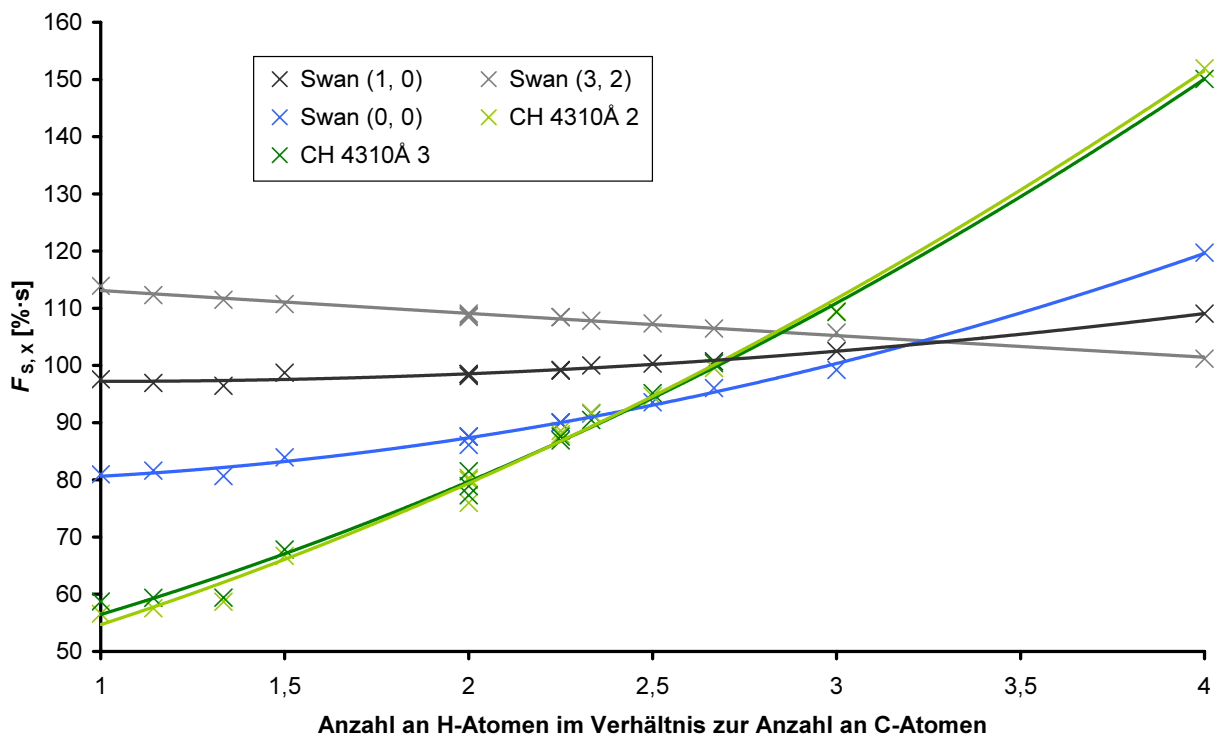


Abbildung 10.6-2: Abgeschätzte Signalflächen verschiedener Banden in Abhängigkeit vom Verhältnis der H- und C-Atomzahlen im Molekül für Verbindungen der Zusammensetzung  $C_aH_b$  ( $a, b \in \mathbb{N}$ ) für  $F_{N, \text{Swan } (2, 1)} = 103\% \cdot s$ ,  $F_{N, \text{OH } 3064\text{Å}} = 270\% \cdot s$  und  $O = 12,3\%$

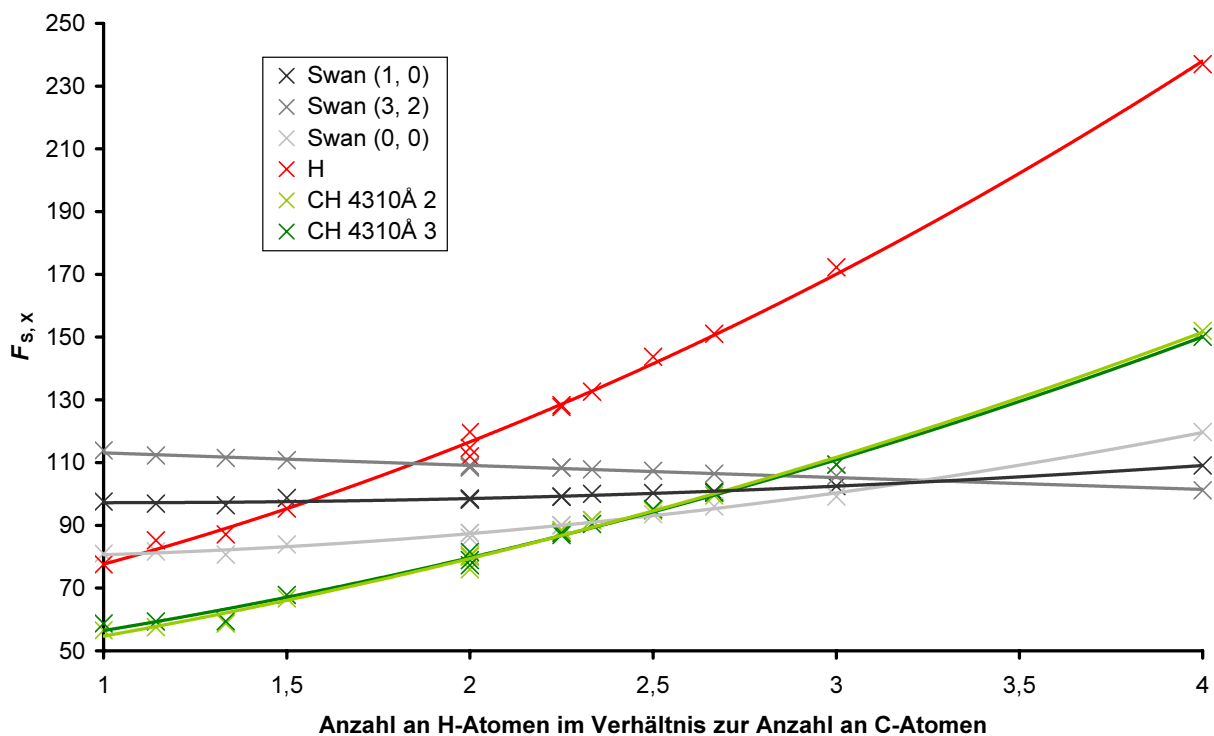


Abbildung 10.6-3: Abgeschätzte Signalflächen verschiedener Banden und der Wasserstofflinie in Abhängigkeit vom Verhältnis der H- und C-Atomzahlen im Molekül für Verbindungen der Zusammensetzung  $C_aH_b$  ( $a, b \in \mathbb{N}$ ) für  $F_{N, \text{Swan}(2,1)} = 103\% \cdot s$ ,  $F_{N, \text{OH } 3064\text{\AA}} = 270\% \cdot s$  und  $O = 12,3\%$

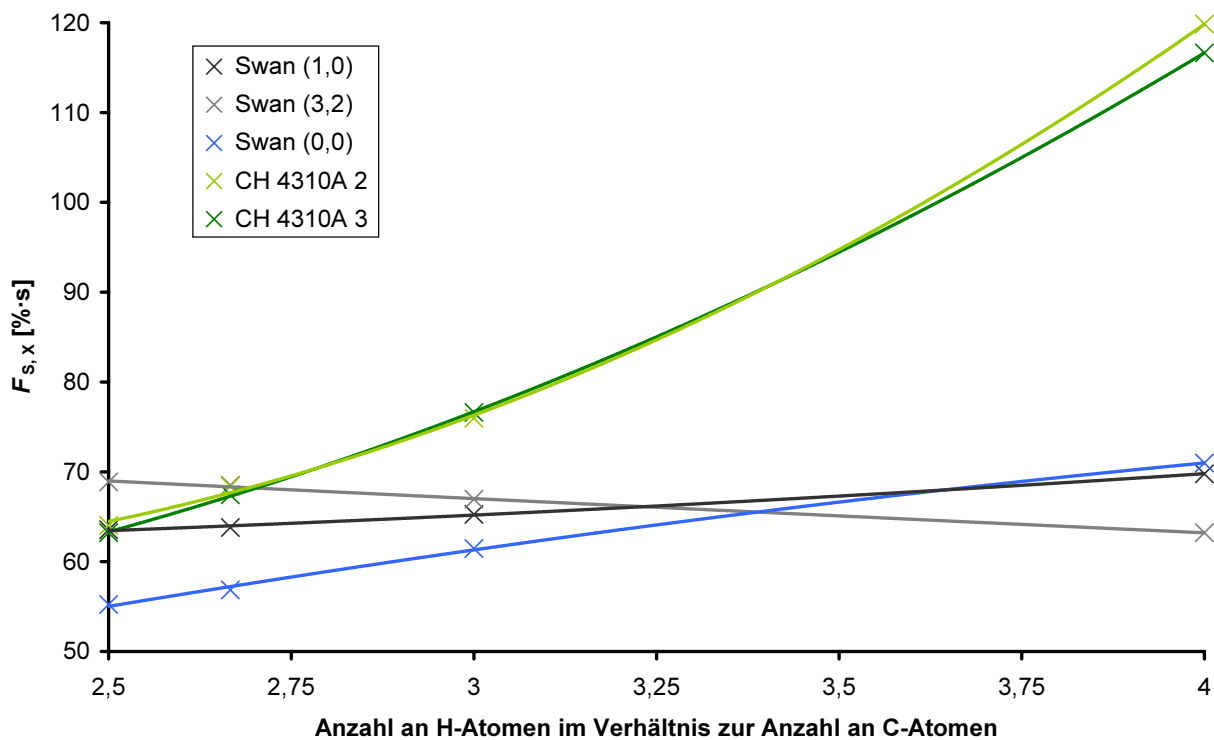


Abbildung 10.6-4: Abgeschätzte Signalflächen verschiedener Banden in Abhängigkeit vom Verhältnis der H- und C-Atomzahlen im Molekül für Alkanole  $F_{N, \text{Swan}(2,1)} = 64,5\% \cdot s$ ,  $F_{N, \text{OH } 3064\text{\AA}} = 293\% \cdot s$  und  $O = 12,1\%$

Tabelle 10.6-1: Die Klassifizierungsparameter der 5 ähnlichsten Signale aus jedem Datensatz der CHO-Datenbank zu dem Klassifizierungsparametersatz

$F_{N, Swan(2,1)} = 810\% \cdot s$ ,  $F_{N, OH\ 3064\text{Å}} = 413\% \cdot s$  und  $O = 13,5\%$  (Parameter aus einer Benzenmessung)

Datensatz	Wertebereiche nach Gleichung (10.5-2) eingehalten	Klassifizierungsparameter	Klassifizierungsparametersätze, zum Vergleich interpoliertes Signal: $F_{N, Swan(2,1)} = 810,3\% \cdot s$ $F_{N, OH\ 3064\text{Å}} = 412,9\% \cdot s$ $O = 13,5\%$				
1, 3-Cyclohexadien	Nein	$F_{N, Swan(2,1)}$ [%·s]	480,4	427,2	449,8	416,0	464,9
		$F_{N, OH\ 3064\text{Å}}$ [%·s]	383,9	325,0	369,4	342,7	370,0
		$O$ [%]	15,7	13,9	16,7	16,3	17,1
Aceton	Nein	$F_{N, Swan(2,1)}$ [%·s]	267,8	300,5	237,3	236,7	303,4
		$F_{N, OH\ 3064\text{Å}}$ [%·s]	436,7	448,7	416,7	412,3	399,4
		$O$ [%]	12,9	12,7	13,4	13,0	11,2
Benzen	Nein	$F_{N, Swan(2,1)}$ [%·s]	810,3	641,9	616,8	774,7	747,5
		$F_{N, OH\ 3064\text{Å}}$ [%·s]	412,9	328,7	344,1	345,9	362,5
		$O$ [%]	13,5	12,7	12,6	12,1	11,8
Butadien	Nein	$F_{N, Swan(2,1)}$ [%·s]	265,9	251,0	262,6	313,1	328,5
		$F_{N, OH\ 3064\text{Å}}$ [%·s]	266,0	261,1	262,7	254,9	253,9
		$O$ [%]	12,7	14,2	15,3	13,4	14,4
Butan	Nein	$F_{N, Swan(2,1)}$ [%·s]	206,8	205,4	235,6	215,4	237,9
		$F_{N, OH\ 3064\text{Å}}$ [%·s]	418,3	410,4	399,4	400,8	407,6
		$O$ [%]	13,6	13,9	14,5	14,2	15,5
Butanol	Nein	$F_{N, Swan(2,1)}$ [%·s]	266,0	203,7	240,9	220,5	256,5
		$F_{N, OH\ 3064\text{Å}}$ [%·s]	379,6	452,8	438,2	475,0	385,0
		$O$ [%]	14,4	12,1	9,6	12,8	17,7
Cyclohexan	Nein	$F_{N, Swan(2,1)}$ [%·s]	374,7	347,6	331,6	335,9	367,5
		$F_{N, OH\ 3064\text{Å}}$ [%·s]	425,6	442,4	435,7	460,3	485,0
		$O$ [%]	12,3	13,9	13,6	12,1	12,2
Diethylether	Nein	$F_{N, Swan(2,1)}$ [%·s]	234,2	255,4	188,7	214,1	239,4
		$F_{N, OH\ 3064\text{Å}}$ [%·s]	436,5	443,3	407,8	402,0	361,6
		$O$ [%]	14,0	12,6	14,2	15,9	11,5
Dimethylether	Nein	$F_{N, Swan(2,1)}$ [%·s]	138,3	111,1	139,1	98,4	91,3
		$F_{N, OH\ 3064\text{Å}}$ [%·s]	400,6	329,0	397,9	329,5	330,9
		$O$ [%]	11,6	12,2	16,2	15,4	15,1
Ethan	Nein	$F_{N, Swan(2,1)}$ [%·s]	179,6	148,0	179,3	170,8	136,0
		$F_{N, OH\ 3064\text{Å}}$ [%·s]	405,2	391,7	373,1	366,1	396,4
		$O$ [%]	14,1	13,1	14,2	13,6	14,4
Ethanol	Nein	$F_{N, Swan(2,1)}$ [%·s]	154,3	139,6	149,7	136,8	144,4
		$F_{N, OH\ 3064\text{Å}}$ [%·s]	439,5	373,2	462,4	446,9	426,5
		$O$ [%]	12,4	12,8	15,8	12,9	9,6
Ethen	Nein	$F_{N, Swan(2,1)}$ [%·s]	259,4	251,7	294,9	247,3	253,1
		$F_{N, OH\ 3064\text{Å}}$ [%·s]	298,2	286,1	282,5	296,8	272,5
		$O$ [%]	13,0	12,9	14,5	15,5	14,9

Datensatz	Wertebereiche nach Gleichung (10.5-2) eingehalten	Klassifizierungsparameter	Klassifizierungsparametersätze, zum Vergleich interpoliertes Signal: $F_{N, \text{Swan } (2, 1)} = 810,3\% \cdot s$ $F_{N, \text{OH } 3064\text{Å}} = 412,9\% \cdot s$ $O = 13,5\%$				
Ethylbenzen	Nein	$F_{N, \text{Swan } (2, 1)}$ [%·s]	497,6	478,5	455,4	680,7	610,8
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	285,5	278,4	289,6	270,5	270,2
		$O$ [%]	11,8	13,8	15,2	11,5	15,6
Hexan	Nein	$F_{N, \text{Swan } (2, 1)}$ [%·s]	289,4	327,7	350,7	227,9	221,3
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	350,5	362,4	363,6	383,8	324,5
		$O$ [%]	13,3	14,2	12,0	14,9	14,5
Isooctan	Ja	$F_{N, \text{Swan } (2, 1)}$ [%·s]	299,5	256,5	242,8	273,4	242,8
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	359,8	420,6	416,5	351,3	407,8
		$O$ [%]	13,4	15,0	12,1	14,7	15,1
Methan	Nein	$F_{N, \text{Swan } (2, 1)}$ [%·s]	104,9	111,9	99,9	89,7	94,4
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	401,6	387,4	438,9	413,7	401,5
		$O$ [%]	14,0	14,6	14,1	12,7	15,5
Methanol	Nein	$F_{N, \text{Swan } (2, 1)}$ [%·s]	38,7	37,5	37,9	38,1	39,9
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	458,0	456,6	465,6	454,2	494,6
		$O$ [%]	10,3	11,7	12,0	10,3	12,6
Octan	Nein	$F_{N, \text{Swan } (2, 1)}$ [%·s]	275,4	347,2	283,5	276,1	279,1
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	395,9	415,3	391,5	385,5	443,1
		$O$ [%]	13,6	15,6	12,4	13,4	13,4
Propan	Nein	$F_{N, \text{Swan } (2, 1)}$ [%·s]	198,4	183,7	198,0	249,4	190,7
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	378,3	381,4	376,4	408,3	373,8
		$O$ [%]	13,3	13,5	15,0	16,7	15,3
Propanal	Nein	$F_{N, \text{Swan } (2, 1)}$ [%·s]	217,9	246,5	233,9	272,6	251,4
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	404,1	424,8	372,5	478,1	412,9
		$O$ [%]	12,6	11,3	12,4	13,3	11,0
Propanol	Nein	$F_{N, \text{Swan } (2, 1)}$ [%·s]	169,9	170,3	175,7	168,9	179,6
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	383,0	456,4	470,0	448,7	428,9
		$O$ [%]	15,0	12,9	14,6	10,9	8,7
Propen	Nein	$F_{N, \text{Swan } (2, 1)}$ [%·s]	244,0	188,9	194,6	212,2	251,4
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	272,4	265,1	258,3	262,1	259,9
		$O$ [%]	14,9	13,4	11,9	15,7	10,8

Tabelle 10.6-2: Die Klassifizierungsparameter der 5 ähnlichsten Signale aus jedem Datensatz der CHO-Datenbank zu dem Klassifizierungsparametersatz

$F_{N, Swan(2,1)} = 103\% \cdot s$ ,  $F_{N, OH\ 3064\text{\AA}} = 270\% \cdot s$  und  $O = 12,3\%$  (Parameter aus einer Octanmessung)

Datensatz	Wertebereiche nach Gleichung (10.5-2) eingehalten	Klassifizierungsparameter	Klassifizierungsparametersätze, zum Vergleich interpoliertes Signal: $F_{N, Swan(2,1)} = 103,0\% \cdot s$ $F_{N, OH\ 3064\text{\AA}} = 269,8\% \cdot s$ $O = 12,3\%$				
1, 3-Cyclohexadien	Ja	$F_{N, Swan(2,1)}$ [%·s]	113,5	135,2	112,6	119,9	146,4
		$F_{N, OH\ 3064\text{\AA}}$ [%·s]	215,7	244,0	222,8	246,2	251,5
		$O$ [%]	13,8	16,6	17,2	18,3	16,5
Aceton	Ja	$F_{N, Swan(2,1)}$ [%·s]	102,4	101,6	110,3	88,3	105,3
		$F_{N, OH\ 3064\text{\AA}}$ [%·s]	263,6	253,7	270,8	253,4	315,1
		$O$ [%]	12,4	11,0	10,8	11,4	11,1
Benzen	Ja	$F_{N, Swan(2,1)}$ [%·s]	134,1	133,7	76,4	101,7	143,7
		$F_{N, OH\ 3064\text{\AA}}$ [%·s]	260,7	244,8	214,2	219,7	231,9
		$O$ [%]	12,2	12,0	11,0	16,5	10,1
Butadien	Ja	$F_{N, Swan(2,1)}$ [%·s]	103,1	123,2	81,2	104,4	105,1
		$F_{N, OH\ 3064\text{\AA}}$ [%·s]	262,8	260,6	257,5	277,4	248,4
		$O$ [%]	13,7	14,4	13,2	16,9	10,7
Butan	Nein	$F_{N, Swan(2,1)}$ [%·s]	106,3	98,6	99,4	108,1	109,8
		$F_{N, OH\ 3064\text{\AA}}$ [%·s]	297,4	318,4	324,6	304,4	315,9
		$O$ [%]	11,6	11,9	11,9	13,5	12,6
Butanol	Ja	$F_{N, Swan(2,1)}$ [%·s]	106,7	94,0	113,3	104,7	110,3
		$F_{N, OH\ 3064\text{\AA}}$ [%·s]	296,2	298,4	248,2	315,7	277,0
		$O$ [%]	9,6	10,6	14,2	10,3	15,4
Cyclohexan	Ja	$F_{N, Swan(2,1)}$ [%·s]	97,7	98,5	86,6	83,6	115,3
		$F_{N, OH\ 3064\text{\AA}}$ [%·s]	281,8	232,9	298,4	294,8	242,3
		$O$ [%]	10,4	10,6	11,5	11,4	9,5
Diethylether	Ja	$F_{N, Swan(2,1)}$ [%·s]	103,8	109,6	106,9	98,0	99,4
		$F_{N, OH\ 3064\text{\AA}}$ [%·s]	259,0	263,6	253,0	252,3	257,1
		$O$ [%]	11,6	13,4	14,1	13,4	15,4
Dimethylether	Ja	$F_{N, Swan(2,1)}$ [%·s]	85,2	88,4	89,7	84,8	82,0
		$F_{N, OH\ 3064\text{\AA}}$ [%·s]	278,4	273,0	282,8	281,1	268,5
		$O$ [%]	11,9	9,7	10,5	13,2	13,2
Ethan	Ja	$F_{N, Swan(2,1)}$ [%·s]	102,7	95,0	100,1	108,9	106,2
		$F_{N, OH\ 3064\text{\AA}}$ [%·s]	285,7	272,2	291,6	285,8	296,0
		$O$ [%]	12,3	12,8	13,0	13,7	12,8
Ethanol	Ja	$F_{N, Swan(2,1)}$ [%·s]	93,9	113,8	102,5	104,5	107,9
		$F_{N, OH\ 3064\text{\AA}}$ [%·s]	266,8	299,5	361,5	374,8	370,4
		$O$ [%]	12,6	13,1	13,7	12,8	13,3
Ethen	Ja	$F_{N, Swan(2,1)}$ [%·s]	101,3	103,7	103,0	109,2	109,9
		$F_{N, OH\ 3064\text{\AA}}$ [%·s]	255,3	240,1	239,6	256,6	243,0
		$O$ [%]	15,8	13,1	13,0	16,5	12,7

Datensatz	Wertebereiche nach Gleichung (10.5-2) eingehalten	Klassifizierungsparameter	Klassifizierungsparametersätze, zum Vergleich interpoliertes Signal: $F_{N, \text{Swan } (2, 1)} = 103,0\% \cdot s$ $F_{N, \text{OH } 3064\text{Å}} = 269,8\% \cdot s$ $O = 12,3\%$				
Ethylbenzen	Ja	$F_{N, \text{Swan } (2, 1)}$ [%·s]	110,1	139,7	98,8	103,2	110,0
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	198,1	224,5	189,2	235,0	219,3
		$O$ [%]	16,6	19,0	17,1	23,9	21,7
Hexan	Ja	$F_{N, \text{Swan } (2, 1)}$ [%·s]	114,8	109,9	110,0	105,1	94,0
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	250,3	230,5	252,3	224,0	222,8
		$O$ [%]	12,4	11,9	14,3	13,7	12,9
Isooctan	Ja	$F_{N, \text{Swan } (2, 1)}$ [%·s]	102,7	102,5	101,3	113,4	98,1
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	262,0	263,4	253,4	280,0	295,8
		$O$ [%]	12,3	10,1	14,3	14,0	10,6
Methan	Ja	$F_{N, \text{Swan } (2, 1)}$ [%·s]	86,9	77,7	73,8	91,3	80,5
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	335,2	322,4	314,4	340,5	330,7
		$O$ [%]	12,0	12,7	11,6	11,9	12,2
Methanol	Nein	$F_{N, \text{Swan } (2, 1)}$ [%·s]	35,4	35,7	36,2	39,1	38,1
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	397,8	403,2	427,0	399,7	454,2
		$O$ [%]	12,9	15,0	10,3	4,6	10,3
Octan	Ja	$F_{N, \text{Swan } (2, 1)}$ [%·s]	103,0	102,6	104,5	102,8	108,6
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	269,8	261,9	255,0	266,2	263,8
		$O$ [%]	12,3	11,9	11,8	11,3	11,3
Propan	Ja	$F_{N, \text{Swan } (2, 1)}$ [%·s]	102,9	99,0	93,8	105,4	96,0
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	305,5	274,5	301,0	315,2	312,9
		$O$ [%]	12,0	14,2	12,6	12,3	11,8
Propanal	Ja	$F_{N, \text{Swan } (2, 1)}$ [%·s]	107,2	93,6	103,1	110,6	92,0
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	277,4	260,0	297,4	284,5	282,0
		$O$ [%]	11,2	12,4	10,8	11,1	11,9
Propanol	Ja	$F_{N, \text{Swan } (2, 1)}$ [%·s]	104,7	98,5	98,8	109,8	100,8
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	306,7	258,3	271,4	313,0	263,2
		$O$ [%]	14,8	15,7	8,6	10,3	8,2
Propen	Ja	$F_{N, \text{Swan } (2, 1)}$ [%·s]	111,0	88,3	90,5	108,6	89,2
		$F_{N, \text{OH } 3064\text{Å}}$ [%·s]	258,9	244,8	247,2	268,2	243,5
		$O$ [%]	10,9	12,0	14,1	15,4	10,3



Tabelle 10.6-3: Die Klassifizierungsparameter der 5 ähnlichsten Signale aus jedem Datensatz der CHO-Datenbank zu dem Klassifizierungsparametersatz

$F_{N, \text{Swan}}(2, 1) = 64,5\% \cdot s$ ,  $F_{N, \text{OH } 3064\text{\AA}} = 293\% \cdot s$  und  $O = 12,1\%$  (Parameter aus einer Probenmessung)

Datensatz	Wertebereiche nach Gleichung (10.5-2) eingehalten	Klassifizierungsparameter	Klassifizierungsparametersätze, zum Vergleich interpoliertes Signal: $F_{N, \text{Swan}}(2, 1) = 64,5\% \cdot s$ $F_{N, \text{OH } 3064\text{\AA}} = 293,0\% \cdot s$ $O = 12,1\%$				
1, 3-Cyclohexadien	Ja	$F_{N, \text{Swan}}(2, 1)$ [%·s]	71,2	67,6	75,2	72,3	68,3
		$F_{N, \text{OH } 3064\text{\AA}}$ [%·s]	221,0	220,4	187,8	228,7	184,7
		$O$ [%]	18,1	18,8	16,1	19,7	16,2
Aceton	Ja	$F_{N, \text{Swan}}(2, 1)$ [%·s]	70,9	60,8	67,9	64,2	73,0
		$F_{N, \text{OH } 3064\text{\AA}}$ [%·s]	276,6	314,2	334,6	309,2	290,2
		$O$ [%]	11,1	10,2	11,0	15,9	9,5
Benzen	Nein	$F_{N, \text{Swan}}(2, 1)$ [%·s]	76,4	78,9	78,2	79,5	79,1
		$F_{N, \text{OH } 3064\text{\AA}}$ [%·s]	214,2	188,8	180,4	170,6	169,3
		$O$ [%]	11,0	13,9	11,3	13,1	12,7
Butadien	Ja	$F_{N, \text{Swan}}(2, 1)$ [%·s]	65,5	81,2	64,7	78,9	66,3
		$F_{N, \text{OH } 3064\text{\AA}}$ [%·s]	251,9	257,5	270,8	277,2	237,5
		$O$ [%]	12,9	13,2	17,8	18,4	11,9
Butan	Ja	$F_{N, \text{Swan}}(2, 1)$ [%·s]	64,7	62,7	64,4	67,2	63,9
		$F_{N, \text{OH } 3064\text{\AA}}$ [%·s]	304,3	279,6	294,1	300,1	309,0
		$O$ [%]	12,4	11,9	10,4	13,6	14,1
Butanol	Ja	$F_{N, \text{Swan}}(2, 1)$ [%·s]	59,9	63,6	67,5	62,0	62,9
		$F_{N, \text{OH } 3064\text{\AA}}$ [%·s]	292,3	314,1	291,9	318,8	316,3
		$O$ [%]	12,0	11,3	10,4	12,2	11,0
Cyclohexan	Ja	$F_{N, \text{Swan}}(2, 1)$ [%·s]	55,8	73,4	70,2	57,3	83,6
		$F_{N, \text{OH } 3064\text{\AA}}$ [%·s]	292,8	303,8	268,9	274,1	294,8
		$O$ [%]	13,5	11,0	10,4	14,7	11,4
Diethylether	Ja	$F_{N, \text{Swan}}(2, 1)$ [%·s]	65,5	61,9	62,8	65,1	69,8
		$F_{N, \text{OH } 3064\text{\AA}}$ [%·s]	279,0	273,1	269,1	301,9	281,9
		$O$ [%]	8,5	7,5	17,0	17,8	18,0
Dimethylether	Ja	$F_{N, \text{Swan}}(2, 1)$ [%·s]	64,2	63,1	65,9	64,7	65,6
		$F_{N, \text{OH } 3064\text{\AA}}$ [%·s]	294,2	291,8	288,1	300,6	287,8
		$O$ [%]	13,2	10,8	10,8	11,6	13,8
Ethan	Ja	$F_{N, \text{Swan}}(2, 1)$ [%·s]	61,4	60,8	67,3	69,1	61,1
		$F_{N, \text{OH } 3064\text{\AA}}$ [%·s]	294,6	290,4	295,8	299,6	295,0
		$O$ [%]	12,1	12,5	13,1	12,4	13,2
Ethanol	Ja	$F_{N, \text{Swan}}(2, 1)$ [%·s]	63,1	61,7	66,8	62,3	59,4
		$F_{N, \text{OH } 3064\text{\AA}}$ [%·s]	278,5	281,4	284,0	342,7	298,6
		$O$ [%]	11,5	12,6	13,7	12,0	13,2
Ethen	Nein	$F_{N, \text{Swan}}(2, 1)$ [%·s]	68,8	68,0	80,2	80,4	81,8
		$F_{N, \text{OH } 3064\text{\AA}}$ [%·s]	244,8	243,3	245,3	250,5	242,5
		$O$ [%]	15,6	18,7	15,8	17,0	14,0

Datensatz	Wertebereiche nach Gleichung (10.5-2) eingehalten	Klassifizierungsparameter	Klassifizierungsparametersätze, zum Vergleich interpoliertes Signal: $F_{N, \text{Swan (2, 1)}} = 64,5\% \cdot s$ $F_{N, \text{OH } 3064\text{Å}} = 293,0\% \cdot s$ $O = 12,1\%$				
Ethylbenzen	Ja	$F_{N, \text{Swan (2, 1)}} [\% \cdot s]$	81,7	63,5	70,4	90,7	90,4
		$F_{N, \text{OH } 3064\text{Å}} [\% \cdot s]$	177,1	172,6	209,4	175,1	219,4
		$O$ [%]	15,8	17,0	22,1	15,5	22,0
Hexan	Nein	$F_{N, \text{Swan (2, 1)}} [\% \cdot s]$	67,8	71,8	59,1	66,8	63,9
		$F_{N, \text{OH } 3064\text{Å}} [\% \cdot s]$	298,2	299,3	242,3	324,0	363,3
		$O$ [%]	12,5	13,0	13,7	15,3	13,1
Isooctan	Ja	$F_{N, \text{Swan (2, 1)}} [\% \cdot s]$	72,3	73,4	63,4	73,9	56,3
		$F_{N, \text{OH } 3064\text{Å}} [\% \cdot s]$	299,0	275,3	286,7	293,5	261,5
		$O$ [%]	14,4	10,7	15,8	15,2	12,9
Methan	Ja	$F_{N, \text{Swan (2, 1)}} [\% \cdot s]$	64,7	67,8	63,9	57,7	61,6
		$F_{N, \text{OH } 3064\text{Å}} [\% \cdot s]$	304,0	306,9	316,0	291,3	310,4
		$O$ [%]	13,1	12,1	11,8	11,2	14,1
Methanol	Ja	$F_{N, \text{Swan (2, 1)}} [\% \cdot s]$	35,4	35,7	39,1	36,2	33,0
		$F_{N, \text{OH } 3064\text{Å}} [\% \cdot s]$	397,8	403,2	399,7	427,0	321,4
		$O$ [%]	12,9	15,0	4,6	10,3	6,9
Octan	Ja	$F_{N, \text{Swan (2, 1)}} [\% \cdot s]$	67,1	73,6	67,4	74,9	68,3
		$F_{N, \text{OH } 3064\text{Å}} [\% \cdot s]$	298,6	289,8	255,5	274,8	252,7
		$O$ [%]	12,2	11,6	12,7	12,8	13,2
Propan	Ja	$F_{N, \text{Swan (2, 1)}} [\% \cdot s]$	64,5	62,2	60,4	61,8	62,6
		$F_{N, \text{OH } 3064\text{Å}} [\% \cdot s]$	293,0	301,7	297,3	304,7	282,3
		$O$ [%]	12,1	11,8	11,7	12,9	14,1
Propanal	Ja	$F_{N, \text{Swan (2, 1)}} [\% \cdot s]$	67,3	59,6	70,8	66,0	71,8
		$F_{N, \text{OH } 3064\text{Å}} [\% \cdot s]$	282,4	294,2	256,5	265,5	256,3
		$O$ [%]	13,2	15,8	12,2	15,6	10,1
Propanol	Ja	$F_{N, \text{Swan (2, 1)}} [\% \cdot s]$	64,8	60,8	66,3	57,5	68,6
		$F_{N, \text{OH } 3064\text{Å}} [\% \cdot s]$	299,1	315,9	324,1	272,9	279,0
		$O$ [%]	12,2	11,6	12,9	12,1	9,3
Propen	Ja	$F_{N, \text{Swan (2, 1)}} [\% \cdot s]$	59,7	59,9	60,0	61,9	61,4
		$F_{N, \text{OH } 3064\text{Å}} [\% \cdot s]$	245,3	259,1	240,6	229,1	235,4
		$O$ [%]	10,7	14,7	9,2	14,0	8,6

## 10.7 Schwefel- und bromhaltige Verbindungen

Um die Eignung des GC-MIP-MED zur Untersuchung von organischen Substanzen mit anderen Heteroatomen als Stickstoff zu untersuchen sind Testmessungen mit Butanthiol und Dibrompropan durchgeführt worden. Zusätzlich zu den in Tabelle 10.1-1 aufgelisteten Banden sowie der H-Linie sind für Butanthiol die Banden CS Hauptsystem (0, 0) und CS Hauptsystem (1, 1) chromatographisch erfasst worden. Im Falle von Dibrompropan wurde auf einer Br-Atomlinie bei 448,0 nm ein zusätzliches Einzelchromatogramm aufgezeichnet. Anhand der in Abbildung 10.7-1 und Abbildung 10.7-2 dargestellten chromatographischen Signale wird deutlich, dass der GC-MIP-MED auch zur chromatographischen Erfassung dieser zusätzlichen Banden und Linien verwendet werden kann. Dieser Sachverhalt zeigt, dass der GC-MIP-MED sich als elementspezifischer Detektor für heteroorganische Verbindungen eignet. H. Jäger<sup>[19]</sup> war in der Lage im Laufe seiner Dissertation auch Chlor und das zweiatomige Molekül CCl emissionspektroskopisch nachzuweisen. Daher ist davon auszugehen, dass auch chlorhaltige Verbindungen sich grundsätzlich mit dem GC-MIP-MED untersuchen lassen. Die Spektraldichten der CS-Banden und der Br-Linie sind gegenüber den Banden von CH, CN und CC recht klein. Daher wäre für qualitative Untersuchungen ein Spektrometer mit einem höheren dynamischen Messbereich erforderlich.

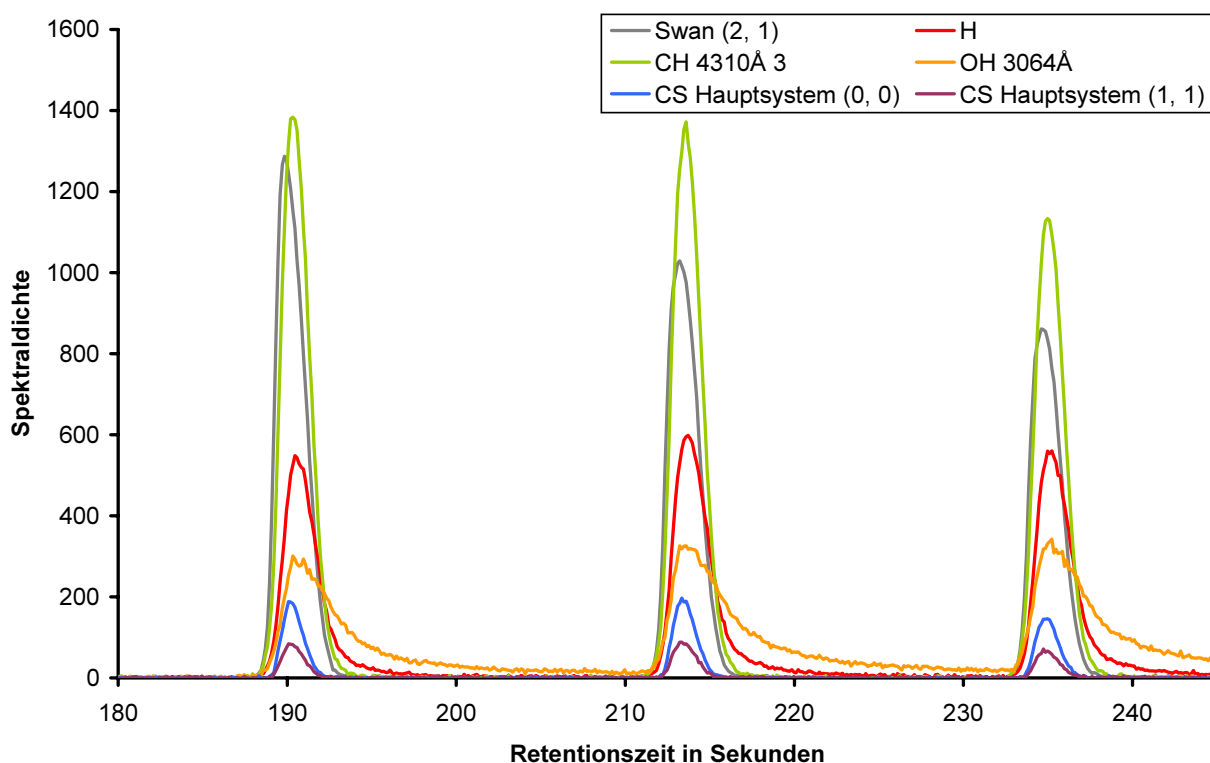


Abbildung 10.7-1: Chromatographische Signale von Butanthiol aufgezeichnet mit dem GC-MIP-MED

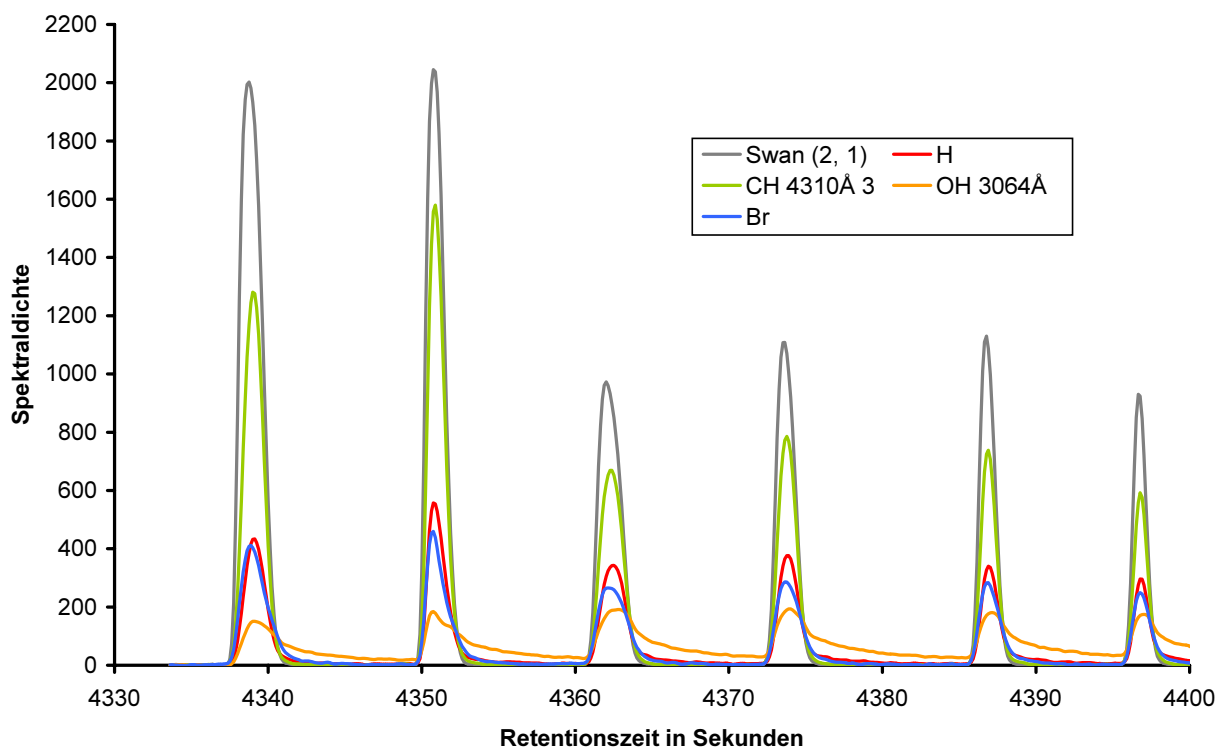


Abbildung 10.7-2: Chromatographische Signale von Dibrompropan aufgezeichnet mit dem GC-MIP-MED

## 10.8 Vergleich der Analyseergebnisse mit Literaturdaten aus anderen Arbeiten

In den Arbeiten von H. Jäger<sup>[19]</sup> und E. Hemstege<sup>[13]</sup> sind vor allem Zusammenhänge zwischen der Struktur und den MIP-MED-Signalen herausgearbeitet worden. H. Jäger konnte beispielsweise signifikante Unterschiede zwischen den Isomerenpaaren Aceton und Propanal bzw. Anilin und  $\alpha$ -Picolin nachweisen. Das von ihm angewandte Messverfahren unterscheidet sich deutlich von der in dieser Arbeit entwickelten Messwerterfassungs- und Auswertemethode. H. Jäger hat Spektrallichte verschiedener Banden und Linien mit einem Monochromator bestimmt. Der zugeführte Massenstrom an Substanz war stets konstant. Es wurde so viel Substanz pro Minute zugeführt, dass die Spektrallichte der Swan (0, 0)-Bande einen definierten Wert erreicht (75% Schreiberausschlag für die Aceton-Propanal-Vergleichsmessung und 25% Schreiberausschlag für die Anilin-Picolin-Vergleichsmessung). Als weiterer Normierungsschritt erfolgt die Division der Spektrallichte der einzelnen Banden und Linien durch die Spektrallichte der Swan (0, 0)-Bande. Neben den Spektrallichteverhältnissen der Linien und Banden steht bei dieser Art der Normierung auch der zugeführte Massenstrom als substanzspezifische Größe zur Verfügung. Bei gleichem Massenstrom ist bei Isomeren auch der Stoffmengenstrom gleich. Tabelle 10.8-1 und Tabelle 10.8-2 enthalten einige von H. Jäger veröffentlichte Spektrallichteverhältnisse, die bei der Messung der erwähnten Isomerenpaare erhalten worden sind. Im Falle von Anilin und  $\alpha$ -Picolin zeigen

sich große Unterschiede bei den CN Violett (0, 1)- und NH 3360Å (0, 0; Q)-Spektraldichteverhältnissen. Im Gegensatz hierzu konnten bei der in dieser Arbeit entwickelten Mess- und Auswertemethode keine signifikanten Unterschiede zwischen Anilin und  $\alpha$ -Picolin festgestellt werden (siehe Abschnitt 10.4.1, Abbildung 10.4-10). Wegen des hohen Fehler bei der Signalflächenbestimmung der NH 3360Å (0, 0; Q)-Bande ist es möglich, dass ein eventueller Unterschied nicht erkennbar ist. Aber auch bei der Messung der Signalflächen der CN Violett (0, 1)-Bande, wo mit einem wesentlich geringeren Fehler zu rechnen ist, konnte mittels GC-MIP-MED kein signifikanter Unterschied zwischen dem Isomerenpaar ermittelt werden. Im Falle von Aceton und Propanal ist der von H. Jäger ermittelte Unterschied zwischen diesen Substanzen bei der relativen Spektraldichte der CH 4310Å-Bande nicht sehr groß. Um so größer ist aber der Unterschied bezüglich des Spektraldichteverhältnisses  $S_H / S_{\text{Swan}(0,0)}$ . Mit dem GC-MIP-MED können Aceton und Propanal nicht unterschieden werden (siehe Abbildung 10.4-9). Ursache könnte die hohe Unsicherheit bei der Bestimmung von  $F_{N,H}$  sein (siehe Tabelle 10.3-12).

Es gibt im wesentlichen zwei große Unterschiede zwischen den von H. Jäger und E. Hemstege verwendeten Detektoren und dem hier vorgestellten GC-MIP-MED. Zum einen kann das auf schnelle hintereinander abfolgende Messungen ausgelegte Spektrometer nicht mit der Präzision eines langsam arbeitenden Monochromators mithalten. Größere Messfehler sind die Folge. Zum anderen scheint es so zu sein, dass bei dem für die Verwendung in der Gaschromatographie optimierten GC-MIP-MED Unterschiede in der Struktur der gemessenen Verbindungen weniger starke Auswirkungen auf die gemessenen Signale haben.

*Tabelle 10.8-1: Spektraldichteverhältnisse zwischen Anilin- und Picolinsignalen bestimmt von H. Jäger sowie die erforderliche Substanzzufuhr zum Erreichen einer Standardintensität für die Swan (0, 0)-Bande*

	Anilin	$\alpha$ -Picolin
$S_H / S_{\text{Swan}(0,0)}^*$	0,100	0,087
$S_{\text{CH}4310\text{Å}} / S_{\text{Swan}(0,0)}$	1,02	0,94
$S_{\text{CN Violett}(0,1)} / S_{\text{Swan}(0,0)}$	0,270	0,205
$S_{\text{NH }3360\text{Å}(0,0;Q)} / S_{\text{Swan}(0,0)}$	2,30	1,40
Substanzverbrauch [mg/min]	0,11	0,035

\*  $S_X$  = Spektraldichte der Bande bzw. Linie X

*Tabelle 10.8-2: Spektraldichteverhältnisse zwischen Aceton- und Propanalsignalen bestimmt von H. Jäger sowie die erforderliche Substanzzufuhr zum Erreichen einer Standardintensität für die Swan (0, 0)-Bande*

	Aceton	Propanal
$S_H / S_{\text{Swan}(0,0)}^*$	0,125	0,195
$S_{\text{CH}4310\text{Å}} / S_{\text{Swan}(0,0)}$	1,90	1,85
Substanzverbrauch [mg/min]	0,192	0,122

\*  $S_X$  = Spektraldichte der Bande bzw. Linie X