

5 Ergebnisse

Während des Untersuchungszeitraumes von vierzehn Monaten wurden 57 Messungen mit 190 bakteriellen Kulturen (siehe Tab. 1) mit dem Prototyp Olfaktograph® durchgeführt.

5.1 Vergleich der Normierung/Referenznormierung

In der Datenauswertung wurden drei unterschiedliche Normierungen getestet. Die Referenznormierung auf sterilen Agar, bzw. Luft (leere Kulturschale) ergab keine Verbesserungen für die Unterscheidbarkeit der bakteriellen Geruchsklassen. Das LDA-Modell aus Median normierten Daten erreichte im Vergleich zu beiden LDA-Modellen aus Referenz normierten Daten geringfügig bessere Erkennungsraten. Alle folgenden Datenauswertungen beziehen sich deshalb auf Median normierte Messwerte.

In den folgenden Punkten 5.2 bis 5.9 werden Auswertungen der Daten des Wolframtrioxid-Sensorarrays vorgestellt. Unter Punkt 5.10 werden die Auswertungen für beide Sensorarrays und ihre gemeinsame Nutzung miteinander verglichen.

5.2 LDA-Modellbildung für fünf Laborspezies

Die Abb. 15 zeigt zwei Ebenen des vierdimensionalen LDA-Modells, das aus 150 Datensätzen der fünf bakteriellen Spezies, jeweils 30 Datensätze von Messungen einer Spezies, berechnet wurde. 99,9 % der Messpunkte liegen innerhalb der Ellipsen. Durch den relativ großen Streudurchmesser der Messpunkte ergeben sich sehr große Ellipsen, die sich zu durchdringen scheinen. Es sind fünf Ellipsen für die Spezies *A. actinomycetemcomitans* (AA), *E. corrodens* (EC), *F. nucleatum* (FN), *P. gingivalis* (PG) und *P. intermedia* (PI) dargestellt. Man erkennt, dass die LDA 1 vor allem die Gruppen *P. gingivalis*, *E. corrodens* und *F. nucleatum* unterscheidet. Die Gruppen *A. actinomycetemcomitans* und *P. intermedia* werden zwar von den erstgenannten zum Teil separiert, können aber voneinander nicht unterschieden werden. Die LDA 2 unterscheidet vor allem die Gruppen *P. gingivalis* und *E. corrodens*. *F. nucleatum* wird nur geringfügig von *A. actinomycetemcomitans* und *P. intermedia* unterschieden. Die LDA 3 und LDA 4 zeigen so gut wie keine Separierung der fünf Klassen. Der mittlere relative Abstand der Klassen vom Zentrum hat einen sehr geringen Wert von 2,518. Die fünf bakteriellen Klassen sind geringfügig unterscheidbar.

Die Tab. 3 zeigt die Werte für die Separationsquotienten. Der Separationsquotient Lambda der ersten Dimension/LDA 1 hat einen Wert von 0,83141. Das heißt mit 83,14 % wird für das erste Diskriminanzmerkmal, die LDA 1, ein relativ hoher Separationsquotient erreicht. Λ_2 mit einem Wert von 42,29 % ist dagegen schon sehr gering. Als Gesamtseparationsmaß kann die Summe 1,66465 der vier Separationsquotienten angesehen werden. Ein höherer Wert ist mit gegebenen Daten mit Hilfe von Linearkombinationen der Ausgangsvariablen nicht erreichbar. Die Anteile der ersten beiden Diskriminanzmerkmale (LDA 1 und LDA 2) am Gesamtseparationsmaß betragen 75,35 %. Die Hauptseparierung der fünf Klassen ist damit durch diese beiden Komponenten gegeben. Λ_3 und Λ_4 sind mit 0,25877, bzw. 0,15159 sehr klein. Mit 25,9 %, bzw. 15,16 % haben die LDA 3, bzw. die LDA 4 kaum eine Bedeutung für die Unterscheidbarkeit der fünf bakteriellen Geruchsklassen.

Die schwarzen Linien in der Abbildung stellen die Schnittebenen zwischen zwei Ellipsen dar. Sie grenzen den Raum ein, in dem Messwerte liegen, die nicht voneinander zu unterscheiden sind.

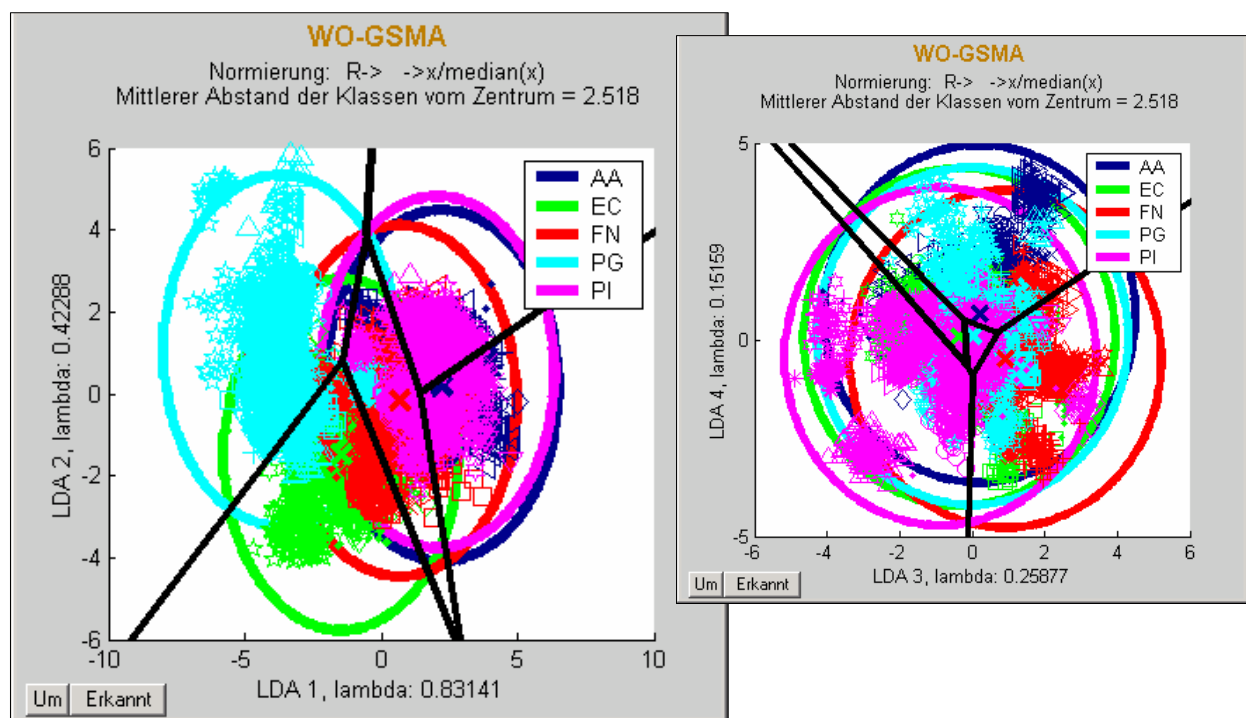


Abb. 15 Zwei Ebenen des Fünf-Klassen-Modells für die fünf bakteriellen Spezies

Jeweils 30 Datensätze für *A. actinomycetemcomitans* (AA), *E. corrodens* (EC), *F. nucleatum* (FN), *P. gingivalis* (PG), *P. intermedia* (PI) bilden die Datenbasis des LDA-Modells.

| $\Lambda 1$ | $\Lambda 2$ | $\Lambda 3$ | $\Lambda 4$ |
|-------------|-------------|-------------|-------------|
| 0,83141 | 0,42288 | 0,25877 | 0,15159 |

Tab. 3 die Separationsquotienten $\Lambda 1$ bis $\Lambda 4$ für das Fünf-Klassen-Modell der fünf Laborstämme *A. actinomycetemcomitans*, *E. corrodens*, *F. nucleatum*, *P. gingivalis*, *P. intermedia*

5.3 LDA-Modelltest für fünf Laborspezies

Unter Nutzung der „leave-one-out“-Methode wurde für das LDA-Modell der 150 Datensätze der Laborspezies ein Modelltest durchgeführt. Ein unbekannter Testdatensatz wird einer Geruchsklasse/Ellipse korrekt zugeordnet, wenn er in der richtigen Ellipse liegt, außerhalb der Überschneidungsebene zwischen verschiedenen Ellipsen.

Die Ergebnisse zeigen, dass nur 54,9 % aller Bakterienmessungen den Ellipsen im Fünf-Klassen-Modell der Laborstämme korrekt zugeordnet werden (Tab. 4). Betrachtet man die Erkennungsraten für die einzelnen Spezies, wird die Spezies *P. gingivalis* mit 72,4 % am besten erkannt. Eine relativ hohe Erkennungsrate wird mit 61,3 % für die Messungen der Spezies *E. corrodens* erreicht. Jedoch werden die restlichen drei Spezies sehr schlecht identifiziert mit Erkennungsraten von weniger als 50 %.

| korrekt erkannte | Messungen | Messpunkte | außerhalb der Ellipse |
|---------------------------------|-----------|------------|-----------------------|
| <i>gesamt</i> | 54,9 % | 54,7 % | |
| <i>A. actinomycetemcomitans</i> | 40,0 % | 40,3 % | 3,4 % |
| <i>E. corrodens</i> | 61,3 % | 64,5 % | 0,6 % |
| <i>F. nucleatum</i> | 37,9 % | 40,4 % | <0,1 % |
| <i>P. intermedia</i> | 32,0 % | 33,4 % | 0,2 % |
| <i>P. gingivalis</i> | 72,4 % | 71,7 % | 6,1 % |

Tab. 4 Erkennungsraten für die Laborspezies von *A. actinomycetemcomitans*, *E. corrodens*, *F. nucleatum*, *P. gingivalis*, *P. intermedia* im Fünf-Klassen-Modell berechnet für die Erkennung nach Ellipsen

5.4 LDA-Modell der klinischen Isolate von *P. gingivalis* und *P. intermedia*

In einem Zwei-Klassen-Modell der LDA sind die beiden Spezies der klinischen Isolate von *P. gingivalis* und *P. intermedia* deutlich voneinander zu differenzieren (Abb. 16). Rein morphologisch sind Kolonien von *P. gingivalis* und *P. intermedia* durch ihre schwarze Pigmentierung nicht eindeutig zu unterscheiden.

Das Geruchsunterscheidungsvermögen dieses Modells, beschrieben durch den Wert des mittleren relativen Abstands zwischen den Klassen vom Zentrum, ist mit einem Wert von 28,21 sehr groß. Die beiden maximalen Separationsquotienten Lambda 1 und 2 gehen nahezu gegen 1 mit einem Wert von 0,99855. 99,86 % aller Messwerte liegen voneinander getrennt.

Der Modelltest des LDA-Modells ergibt eine sehr hohe Wahrscheinlichkeit der Erkennung von Messungen und Messpunkten mit nahezu 100 %, bzw. 99,1 %.

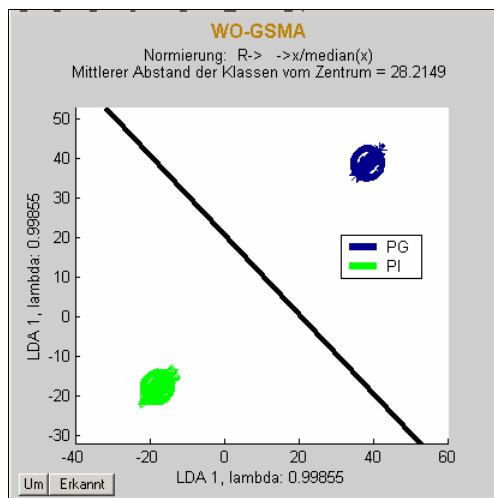


Abb. 16 Zwei-Klassen-Modell für die klinischen Isolate der Spezies *P. gingivalis* und *P. intermedia*. Acht Datensätze der Spezies *P. intermedia* (PI) und vier der Spezies *P. gingivalis* (PG) bilden die Datenbasis des LDA-Modells.

| korrekt zugeordnete | Messungen | Messpunkte |
|-----------------------------------------------------------|-----------|------------|
| <i>P. intermedia</i> und <i>P. gingivalis</i> zusammen | 100% | 99,1% |

Tab. 5 Erkennungsraten für die klinischen Isolate von *P. gingivalis* und *P. intermedia* im Zwei-Klassen-Modelltest

5.5 Erkennung der klinischen Isolate von *P. intermedia* und *P. gingivalis* im Zwei-Klassen-Modell der Laborstämme

In der Tab. 6 sind die Erkennungsraten aufgeführt, mit denen die Messungen der klinischen Isolate von *P. intermedia* und *P. gingivalis* in einem Zwei-Klassen-Modell der Laborstämme von *P. intermedia* und *P. gingivalis* (DSM 20706, bzw. DSM 20709) erkannt werden. Bis auf eine Messung des *P. intermedia* werden alle Patientenproben mit 100 % sicher zugeordnet.

| Spezies | Nr. der Messung | Erkennung der Messpunkte der Messung in % |
|----------------------|-----------------|-------------------------------------------|
| <i>P. gingivalis</i> | 13 | 100 |
| | 19 | 100* |
| | 20 | 100 |
| | 21 | 100 |
| <i>P. intermedia</i> | 15 | 94,4 |
| | 16 | 100 |
| | 18 | 100 |
| | 22 | 100 |
| | 28 | 100* |
| | 29 | 100 |
| | 43 | 100 |
| | 44 | 100 |

Tab. 6 Erkennungsraten für die Messungen der klinischen Isolate von *P. intermedia* und *P. gingivalis* im Zwei-Klassen-Modell aus Laborkeimen
* Messung wird als Beispiel in Abb. 17 gezeigt

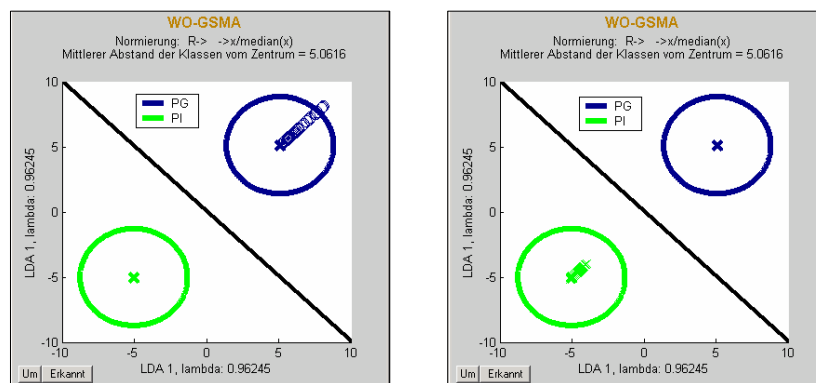


Abb. 17 Testmessungen zwei klinischer Isolate von *P. gingivalis* und *P. intermedia*

links: Test der Messung 19 (eine klinische Kultur von *P. gingivalis*) im Zwei-Klassen-Modell der LDA aus den Laborspezies von *P. gingivalis* und *P. intermedia*

rechts: Test der Messung 28 (eine klinische Kultur von *P. intermedia*) im Zwei-Klassen-Modell der LDA aus den Laborspezies von *P. gingivalis* und *P. intermedia*

5.5.1 Erkennung der klinischen Isolate der Spezies *P. intermedia* und *P. gingivalis* im Fünf-Klassen-Modell der Laborspezies

In der Tab. 7 sind die Erkennungsraten aufgeführt, mit denen die klinischen Isolate in dem Fünf-Klassen-Modell der Laborstämme erkannt werden. Es wird die Anwendung des LDA-Modells der Laborstämme für die Klassifikation von klinischen Isolaten untersucht.

Drei Datensätze der klinischen Isolate der Spezies *P. gingivalis* werden mit einer Erkennung von 100 % sehr gut zugeordnet. Ein Datensatz wird nicht erkannt, nur 42,9% der Messpunkte werden korrekt zugeordnet. Im Schnitt werden 85,7 % der Messpunkte der Datensätze korrekt erkannt.

Eine korrekte Zuordnung der klinischen Isolate der Spezies *P. intermedia* ist nicht möglich. Nur ein Datensatz wird korrekt klassifiziert im Fünf-Klassen-Modell der Laborstämme mit korrekter Zuordnung für 75,5 % der Messpunkte. Die Wahrscheinlichkeit für eine Erkennung der klinischen Isolate von *P. intermedia* beträgt im Schnitt nur 13,1 %.

| Spezies | Nr. der Messung | Erkennung der Messpunkte der Messung in % |
|----------------------------------|-----------------|-------------------------------------------|
| <i>P. gingivalis</i> | 13 | 100 |
| | 19 | 100 |
| | 20 | 100 |
| | 21 | 42,9 |
| Schnitt für <i>P. gingivalis</i> | | 85,7 |
| <i>P. intermedia</i> | 15 | 0 |
| | 16 | 25,6 |
| | 18 | 3,7 |
| | 22 | 0 |
| | 28 | 0 |
| | 29 | 75,5 |
| | 43 | 0 |
| | 44 | 1,7 |
| Schnitt für <i>P. intermedia</i> | | 13,1 |

Tab. 7 Erkennung der klinischen Isolate der Spezies *P. intermedia* und *P. gingivalis* in einem Fünf-Klassen-Modell der Laborkeime

5.5.2 Vergleich der Erkennung von *P. gingivalis* und *P. intermedia* im Modell aus Laborstämmen und im Modell aus Laborstämmen und klinischen Isolaten zusammen

Der folgende Vergleich soll zeigen, ob ein Fünf-Klassen-Modell, in dem die Datensätze der klinischen Isolate mit in die Datenbasis eingingen, vergleichbar ist mit dem Fünf-Klassen-Modell der reinen Laborstämme. Dazu wurden beide Modelle überprüft mit Testdatensätzen von *P. gingivalis* und *P. intermedia* der Laborstämme und der klinischen Isolate. Die ermittelten Erkennungsraten sind in der Tab. 8 aufgeführt.

Die Erkennungsraten für die Spezies *P. gingivalis* sind in dem Modell aus Laborstämmen und klinischen Isolaten zusammen nur geringfügig höher als in dem Fünf-Klassen-Modell der Laborstämme allein. Laborproben der Spezies *P. gingivalis* werden zu 80,2 % korrekt erkannt im Vergleich zu 77,6 % im Laborkeim-Modell. Klinische Proben werden im zusammengefassten Modell mit 87,25 % erkannt im Vergleich zu 85,7 % im Laborkeim-Modell. Die Unterschiede sind so gering, dass man für die Spezies *P. gingivalis* beide Modelle gleichwertig verwenden kann.

Laborproben der Spezies *P. intermedia* werden in beiden Modellen zu 33 % korrekt erkannt. Die Erkennung ist sehr schlecht, nicht einmal die Hälfte aller unbekannt Proben dieser Spezies würde identifiziert werden. Testdatensätze von klinischen Isolaten werden mit 13,1 % im Modell aus Laborstämmen kaum erkannt. Die Erkennung mit 42 % ist für das Modell der Laborstämmen und klinischen Isolate zusammen zwar besser, liegt aber immer noch unter der 50 %-Grenze. Die Erkennung ist also auch hier sehr schlecht und kann mehr als zufällig angesehen werden.

| | Fünf-Klassen-Modell aus Laborstämmen | Fünf-Klassen-Modell aus Laborstämmen und klinischen Isolaten |
|--------------------------|---------------------------------------------|---------------------------------------------------------------------|
| Laborstämmen | Erkennung Messpunkte | Erkennung Messpunkte |
| <i>P. gingivalis</i> | 77,6 % | 80,2 % |
| <i>P. intermedia</i> | 33,9 % | 33,3 % |
| Klinische Isolate | | |
| <i>P. gingivalis</i> | 85,7 % | 87,25 % |
| <i>P. intermedia</i> | 13,1 % | 42,0 % |

Tab. 8 Vergleich der Erkennung für die Spezies *P. gingivalis* und *P. intermedia* im Fünf-Klassen-Modell der Laborspezies mit der Erkennung im Fünf-Klassen-Modell der fünf Laborspezies zusammen mit den klinischen Isolaten

5.6 Untersuchung der Stabilität des LDA-Modells

In der Abb. 18 ist die erste Ebene der vierdimensionalen Darstellung des Fünf-Klassen-Modells zu sehen, das aus 150 Datensätzen (jeweils 30 pro Spezies) der fünf Laborstämmen gebildet wurde. In die Gruppen von *P. gingivalis* und *P. intermedia* sind die Datensätze der klinischen Isolate mit eingegangen.

Die Ellipsen der fünf bakteriellen Geruchsklassen überschneiden sich. Die Unterscheidung der fünf Klassen ist gering. Der mittlere relative Abstand der Klassen vom Zentrum beträgt 2,6986. Er ist geringfügig höher als in dem Fünf-Klassen-Modell der Laborstämmen unter Punkt 5.2 mit einem Wert von 2,518.

Die beiden maximalen Separationsquotienten Lambda 1 und Lambda 2 haben einen Wert von 0,86975, bzw. 0,33229. Das erste Diskriminanzmerkmal, die LDA 1, erreicht einen relativ hohen Separationsquotienten. Lambda 2 ist dagegen schon sehr gering. Die Hauptseparierung der fünf Klassen ist durch diese beiden Komponenten gegeben.

Im Modelltest werden unbekannte Messungen mit einer Wahrscheinlichkeit von 60,3 % korrekt klassifiziert (Tab. 9). Die Erkennungsrate in diesem Modell ist etwas höher als die Erkennungsrate von 54,9 % für das Fünf-Klassen-Modell der Laborspezies (siehe Punkt 5.3).

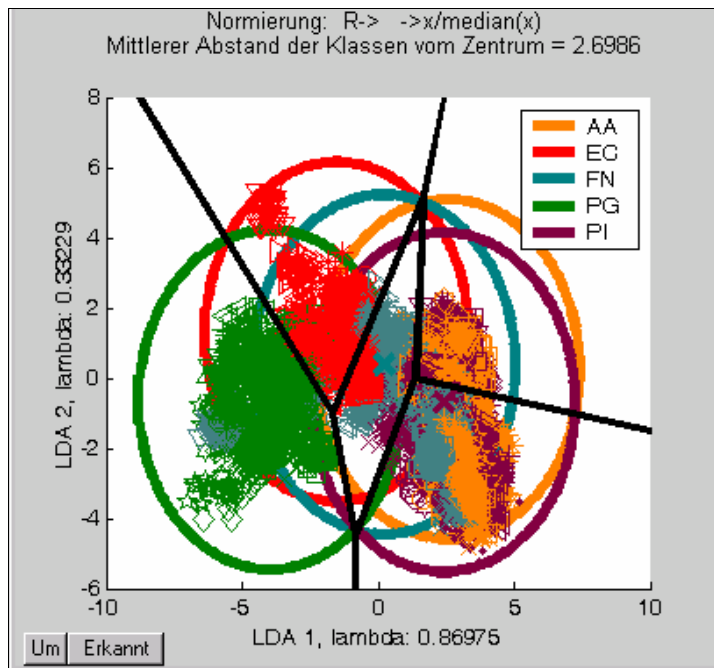


Abb. 18 LDA-Modell der fünf Bakterienklassen der Laborspezies und der klinischen Isolate
Die Datenbasis des LDA-Modells umfasst 150 Datensätze, jeweils 30 pro Spezies.

| Erkennungsrates Messungen in % | Erkennungsrates Messpunkte in % |
|--------------------------------|---------------------------------|
| 60,3 | 55,5 |

Tab. 9 Erkennungsrates für das Fünf-Klassen-Modell aus Datensätzen der Laborspezies und klinischen Isolate

5.6.1 Verringerung der Datenbasis des Fünf-Klassen-Modells unter Weglassen von Datensätzen aus ersten Messungen

Die Ellipsen des Fünf-Klassen-Modells in Abb. 18 überschneiden sich stark, da die Streuung innerhalb der Messpunkte sehr groß ist. Ob die Streuung auf eine Veränderung der Sensorantworten im Laufe der Zeit zurückzuführen ist, soll mit den folgenden Punkten überprüft werden.

Die Datenbasis des Fünf-Klassen-Modells von Punkt 5.6 (Abb. 18) wird um 20 % der Datensätze verringert, die zu Beginn des Untersuchungszeitraumes von vierzehn Monaten aufgenommen wurden. Die Abb. 19 zeigt die Darstellung des Fünf-Klassen-Modells, dessen Datenbasis 120 Datensätze umfasst. Der mittlere relative Abstand der Klassen zum Zentrum der Ellipsen ist 2,81. Das Geruchsunterscheidungsvermögen ist vernachlässigbar höher als mit 2,6986 im Fünf-Klassen-Modell mit einer Datenbasis von 150 Datensätzen (Punkt 5.6). Der Separationsquotient Lambda 1 hat einen Wert von 0,86631. Mit 86,63 % wird für das erste Diskriminanzmerkmal, die LDA 1, ein relativ hoher Separationsquotient erreicht. Im Vergleich zum Fünf-Klassen-Modell, dessen Datenbasis 150 Datensätze umfasst (Punkt 5.6), erreicht der

LDA-Modell-Test mit 68,1 % eine Verbesserung der Erkennungsrate für unbekannte Messungen um 8 %. Die Erkennungsrate für korrekt erkannte Messpunkte steigt nur gering um 3 % auf 58,4 % (Tab. 10).

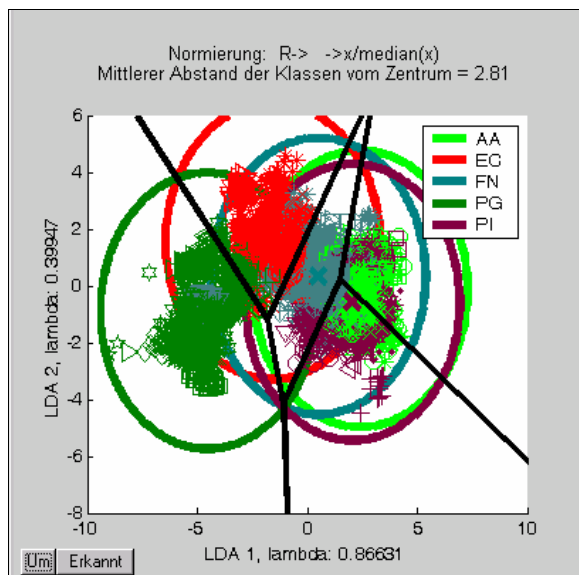


Abb. 19 Fünf-Klassen-Modell für die fünf Bakterienklassen der Laborspezies und klinischen Isolate

Die Datenbasis des LDA-Modells umfasst 120 Datensätze. 30 Datensätze der ersten Messungen wurden nicht in der Modellbildung verwendet.

| Erkennungsrate Messungen in % | Erkennungsrate Messpunkte in % |
|-------------------------------|--------------------------------|
| 68,1 | 58,4 |

Tab. 10 Erkennungsraten für das Fünf-Klassen-Modell der fünf Bakterienklassen der Laborspezies und klinischen Isolate mit einer Datenbasis von 120 Datensätzen, verringert um Datensätze der ersten aufgenommenen Messungen.

5.6.2 Verringerung der Datenbasis des Fünf-Klassen-Modells unter Weglassen von Datensätzen aus letzten Messungen

Um den Anstieg der Erkennungsrate von Messungen um 8 % (siehe Punkt 5.6.1) mit einer Veränderung der Sensorantworten im Laufe der Anwendungszeit zu überprüfen, wurde die Datenbasis von 150 Datensätzen (Punkt 5.6) um 20 % der Datensätze, die in den letzten Messungen der Untersuchung aufgenommen wurden, verringert. Tab. 11 zeigt die Erkennungsraten des Modelltestes. Unbekannte Messungen werden mit einer Wahrscheinlichkeit von 68,1 % korrekt klassifiziert. Die Erkennungsrate unterscheidet sich nicht von der des Fünf-Klassen-Modells, in dem Datensätze der ersten Messungen nicht in die Modellbildung eingingen (Tab. 10). Die Erkennungsrate für unbekannte Messpunkte steigt leicht auf 60,2 % um 4,7 %. Eine kontinuierliche Veränderung der Sensorantworten im Laufe der Untersuchungszeit von vierzehn Monaten ist nicht vorhanden.

| Erkennungsrate Messungen in % | Erkennungsrate Messpunkte in % |
|-------------------------------|--------------------------------|
| 68,1 | 60,2 |

Tab. 11 Erkennungsraten für das Fünf-Klassen-Modell der fünf Bakterienklassen der Laborspezies und klinischen Isolate mit einer Datenbasis von 120 Datensätzen, verringert um Datensätze der letzten aufgenommenen Messungen.

5.6.3 Überprüfung der Zufälligkeit des Anstiegs der Erkennungsraten bei verringerter Datenbasis

Da die Erkennungsraten der Fünf-Klassen-Modelle mit verringerter Datenbasis keinen eindeutigen Bezug zu einer Veränderung der Sensorantworten mit der Zeit zeigen (siehe Punkt 5.6.1 und 5.6.2), ist zu überprüfen, ob der Anstieg der Erkennungsraten zufällig ist. Die Datenbasis von 150 Datensätzen (Punkt 5.6) wird dafür um 20 % der Datensätze verringert, die in Messungen in der Mitte des gesamten Untersuchungszeitraumes aufgenommen wurden. In diesem Fünf-Klassen-Modell werden unbekannte Messungen mit 63,8 % und unbekannte Messpunkte mit 55,1 % korrekt erkannt (siehe Tab. 12).

| Erkennungsrate Messungen in % | Erkennungsrate Messpunkte in % |
|-------------------------------|--------------------------------|
| 63,8 | 55,1 |

Tab. 12 Erkennungsraten für das Fünf-Klassen-Modell der fünf Bakterienklassen der Laborspezies und klinischen Isolate mit einer Datenbasis von 120 Datensätzen, verringert um Datensätze von Messungen, die chronologisch in der Mitte des Untersuchungszeitraumes aufgenommen wurden.

5.7 Überprüfung der Entwicklung der Modellerkennung

Die Datenbasis des Modells der fünf Laborspezies umfasst 150 Datensätze. Die Überprüfung der Modellstabilität unter den Punkten von 5.6 zeigt, dass eine Verringerung der Datenbasis um 20 % der Datensätze die Modellstabilität nicht beeinflusst. Die Erkennungsraten des Fünf-Klassen-Modells verändern sich nur geringfügig.

Abb. 20 bis Abb. 27 ist die Veränderung der Erkennungsraten und der Separationsquotienten λ_1 bis λ_4 für Fünf-Klassen-Modelle in Abhängigkeit zur Größe der Datenbasis, der Anzahl von Messungen in der Modellbildung, dargestellt. Berechnet wurden Fünf-Klassen-Modelle, die mehr als 20 % der Datensätze, das heißt mehr als elf Messungen in der Datenbasis enthalten. Mit dieser Menge von Datensätzen ist eine statistisch sinnvolle LDA-Modellbildung möglich. Das Fünf-Klassen-Modell mit der höchsten Anzahl von Datensätzen enthält alle 57 aufgenommenen Messungen (siehe Tab. 1). Für die LDA-Modellbildung wurde die Datenbasis jeweils so verändert, dass enthaltene Messungen unterschiedlich chronologisch geordnet sind,

um den Einfluss zeitlicher Veränderungen der Sensorantworten auf die Modellerkennung zu überprüfen.

5.7.1 Entwicklung der Modellerkennung für die chronologische Abfolge der Messungen

Die Abb. 20 stellt die Entwicklung der Erkennungsraten des Fünf-Klassen-Modells in Abhängigkeit von der chronologischen Abfolge der zunehmenden Anzahl von Messungen in der Datenbasis über den Untersuchungszeitraum von vierzehn Monaten dar. Mit den ersten zwölf aufgenommenen Messungen wurde das erste Fünf-Klassen-Modell in dieser Darstellung berechnet.

Die Datensätze der zu Beginn der Untersuchung durchgeführten 17 Messungen ergeben Modelle mit sehr guten Erkennungsraten für unbekannte Bakterienproben von 75 % bis über 90 %. Gehen folgende Messungen mit in die Datenbasis ein, fallen die Erkennungsraten ab unter 70 %. Werden die ersten 36 Messungen für die Datenbasis verwendet, erreicht das Fünf-Klassen-Modell nur eine Erkennungsrate von weniger als 60 %. Bei 47 Messungen, deren Datensätze in die Modellbildung eingehen, gehen die Erkennungsraten der Fünf-Klassen-Modelle gegen 60 %. Mit der Größe der Datenbasis von 47 Messungen wird ein Stabilitäts-Maximum für das Fünf-Klassen-Modell erreicht.

Die Entwicklung der maximalen Separationsquotienten Lambda 1 bis 4 zeigt einen ähnlichen Verlauf wie die Entwicklung der Erkennungsraten (Abb. 21). Alle Lambda-Werte gehen schon bei einer Anzahl von 42 Messungen für die Modellbildung gegen einen konstanten Grenzwert. Der Separationsquotient Λ_1 erreicht die höchsten Werte mit 85 %. Mit stabilem Modell liegt damit die Mehrheit der voneinander differenzierbaren fünf Geruchsklassen in der ersten Ebene der vierdimensionalen Darstellung. Für die Bildung eines stabilen Fünf-Klassen-Modells sind 42 bis 47 Messungen in der Datenbasis ausreichend.

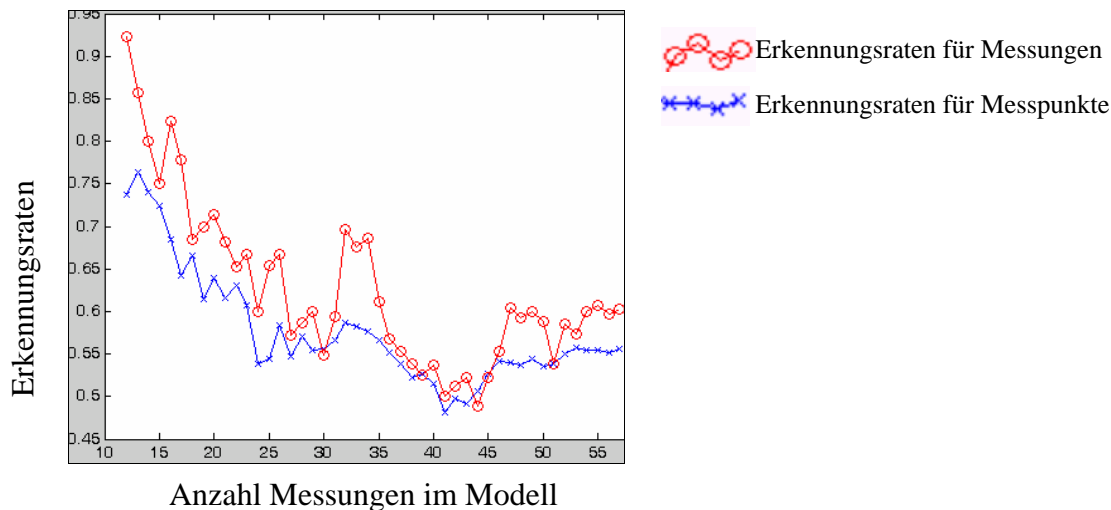


Abb. 20 Entwicklung der Modellerkennung in Abhängigkeit von der Größe der Datenbasis des Fünf-Klassen-Modells

Die Größe der Datenbasis wird verändert durch die zunehmende Anzahl von Messungen, die in chronologischer Abfolge in die Datenbasis aufgenommen werden.

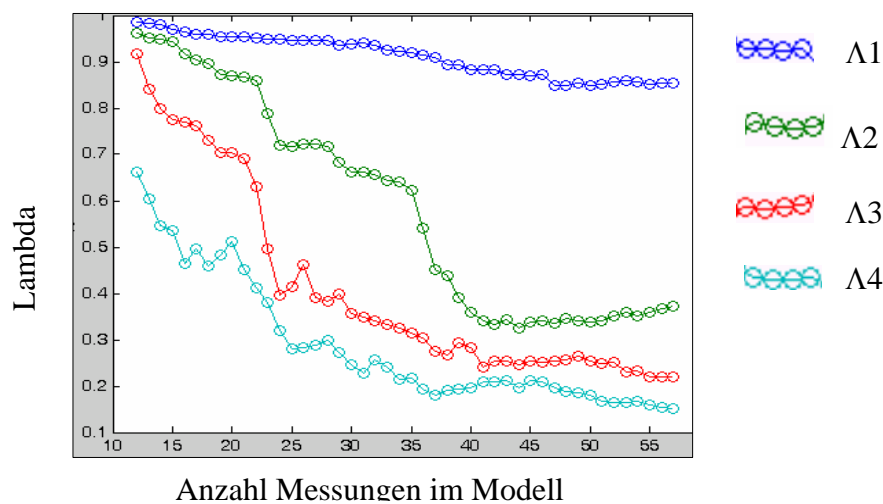


Abb. 21 Entwicklung der Separationsquotienten Lambda 1 bis 4 in Abhängigkeit von der Größe der Datenbasis des Fünf-Klassen-Modells

Die Größe der Datenbasis wird verändert durch die zunehmende Anzahl von Messungen, die in chronologischer Abfolge in die Datenbasis aufgenommen werden.

5.7.2 Datumsbezogene Entwicklung der Modellerkennung

Die Abb. 22 stellt die Entwicklung der Erkennungsraten des Fünf-Klassen-Modells in Abhängigkeit einer weiteren chronologischen Abfolge der zunehmenden Anzahl von Messungen in der Datenbasis dar. Die LDA-Modellberechnung beginnt in dieser Darstellung mit den ersten zwölf Messungen der zweiten Hälfte des Untersuchungszeitraumes. Die Datensätze der ersten zwölf bis 20 Messungen der zweiten Hälfte des Untersuchungszeitraumes bilden Fünf-Klassen-

Modelle mit Erkennungsraten von mehr als 68 % bis zu 85 % für unbekannte Bakterienproben (Abb. 22). Mit einer Anzahl von 21 bis 34 Messungen in der LDA-Modellberechnung erreichen die Erkennungsraten nur Werte zwischen 58 % und 75 %. Ab der Anzahl von 35 Messungen in der Datenbasis des Fünf-Klassen-Modells werden Messungen der ersten Hälfte des Untersuchungszeitraumes zur LDA-Modellbildung hinzugenommen. Die Erkennungsraten fallen nun mit steigender Anzahl der Messungen auf Werte unter 60 % ab. Erst mit einer Anzahl von 47 Messungen erreichen die Fünf-Klassen-Modelle wieder Erkennungsraten um 60 % und ein stabiles Maximum. Eine Ausnahme bildet das Fünf-Klassen-Modell mit 51 Messungen in der Datenbasis, dessen Erkennung unter 55 % liegt. Der Verlauf der maximalen Separationsquotienten Lambda 1 bis 4 ähnelt dem der in Abb. 21 gewählten Abfolge der Messungen. Alle Lambda-Werte gehen gegen einen konstanten Wert schon bei einer Anzahl von 42 Datensätzen. Lambda 1 erreicht die höchsten Werte mit 85 %.

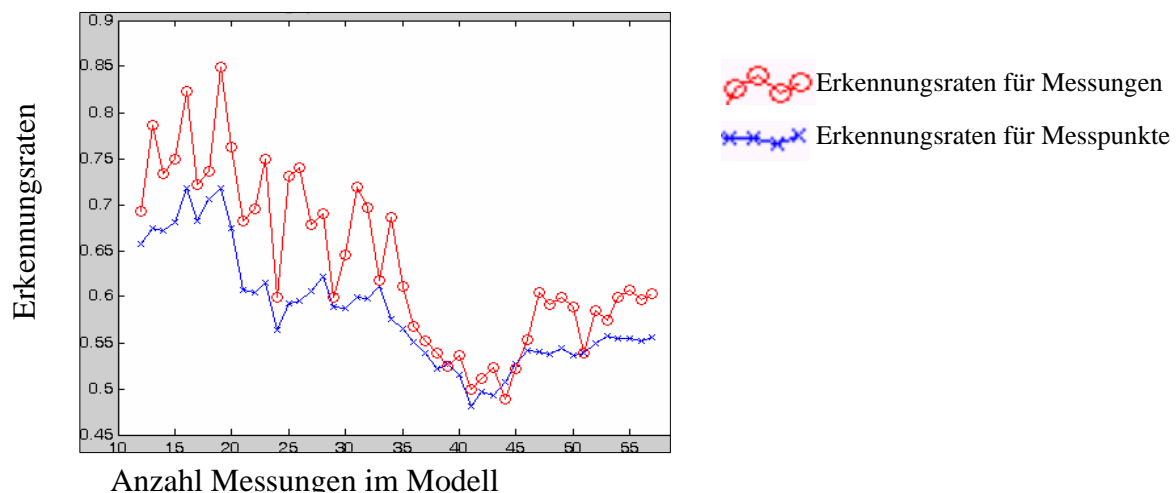


Abb. 22 Entwicklung der Modellerkennung in Abhängigkeit von der Größe der Datenbasis des Fünf-Klassen-Modells

Die Größe der Datenbasis wird verändert durch die zunehmende Anzahl von Messungen. Die LDA-Modell-Bildung beginnt mit den Messungen der zweiten Hälfte des Untersuchungszeitraumes von vierzehn Monaten.

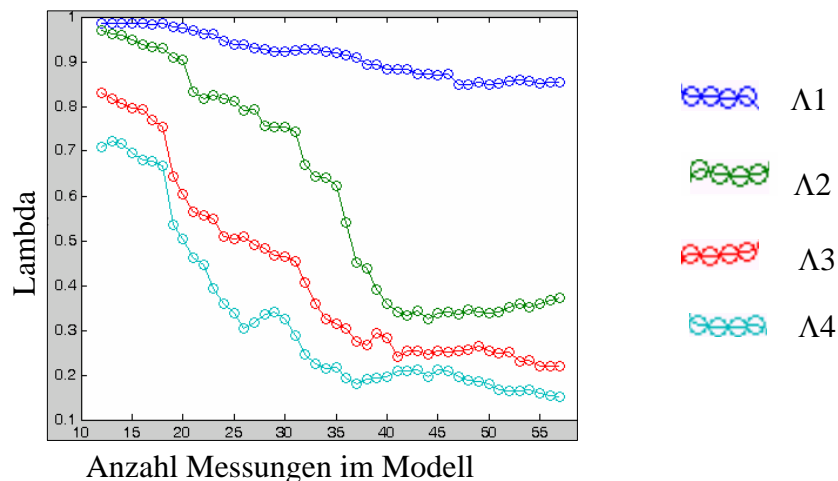


Abb. 23 Entwicklung der Separationsquotienten Lambda 1 bis 4 in Abhängigkeit von der Größe der Datenbasis des Fünf-Klassen-Modells

Die Größe der Datenbasis wird verändert durch die zunehmende Anzahl von Messungen. Die LDA-Modell-Bildung beginnt mit den Messungen der zweiten Hälfte des Untersuchungszeitraumes von vierzehn Monaten.

5.7.3 Entwicklung der Modellerkennung für zwei zufällige Abfolgen der Messungen

Die Abb. 24 und Abb. 26 stellen die Entwicklung der Erkennungsraten des Fünf-Klassen-Modells in Abhängigkeit von zwei chronologisch zufälligen Abfolgen der zunehmenden Anzahl von Messungen in der Datenbasis dar. Die Entwicklung der Erkennungsraten für die LDA-Modelle zeigt einen mehr oder weniger monotonen Anstieg mit steigender Anzahl von Messungen in der Datenbasis. Die Erkennungsraten erreichen in Abb. 24 ein Maximum bei 70 % für ein LDA-Modell, gebildet aus 45 Messungen. Für eine Anzahl von mehr als 45 Messungen in der LDA-Modellberechnung sinken die Erkennungsraten grenzwertig zwischen 60 % und 65 %. In der Abb. 26 erreichen die Erkennungsraten maximal 63 %. Ab einer Anzahl von 46 Messungen in der Datenbasis gehen die Erkennungsraten gegen Werte zwischen 58 % und 60 %.

Der Verlauf des Separationsquotienten Lambda 1 ähnelt für beide zufälligen Reihenfolgen der Messungen in der Datenbasis (siehe Abb. 25 und Abb. 27) dem Verlauf der Separationsquotienten für die chronologische Abfolge von Messungen in der Datenbasis (Abb. 21 und Abb. 23). Lambda 2 bis 3 erreichen dagegen für die zweite Berechnung der Erkennungsraten (siehe Abb. 27) schon bei einer Größe der Datenbasis von 30 Messungen einen

Grenzwert im Vergleich zu 42 Messungen für die chronologischen Abfolgen der Messungen in der Datenbasis (siehe Abb. 21 und Abb. 23).

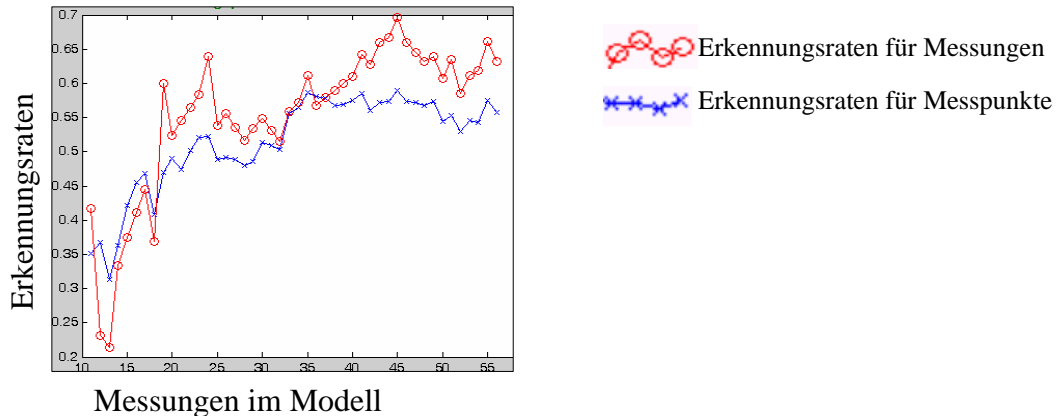


Abb. 24 Entwicklung der Modellerkennung in Abhängigkeit von der Größe der Datenbasis des Fünf-Klassen-Modells

Die Größe der Datenbasis wird verändert durch die zunehmende Anzahl von Messungen. Die Abfolge der Messungen in der Datenbasis des LDA-Modells ist zufällig gewählt.

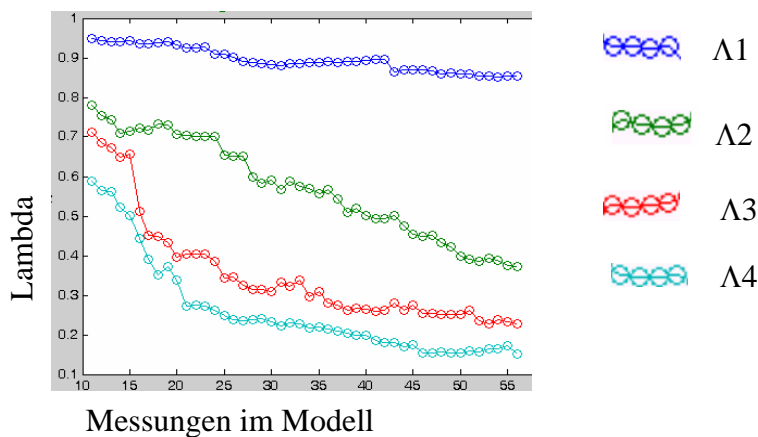


Abb. 25 Entwicklung der Separationsquotienten Lambda 1 bis 4 in Abhängigkeit von der Größe der Datenbasis des Fünf-Klassen-Modells

Die Größe der Datenbasis wird verändert durch die zunehmende Anzahl von Messungen. Die Abfolge der Messungen in der Datenbasis des LDA-Modells ist zufällig gewählt.

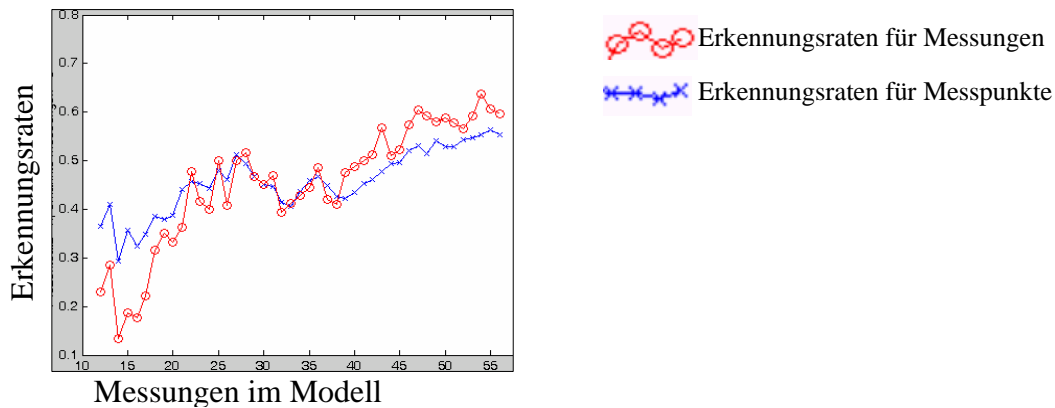


Abb. 26 Entwicklung der Modellerkennung in Abhängigkeit von der Größe der Datenbasis des Fünf-Klassen-Modells

Die Größe der Datenbasis wird verändert durch die zunehmende Anzahl von Messungen. Die Abfolge der Messungen in der Datenbasis des LDA-Modells ist zufällig gewählt.

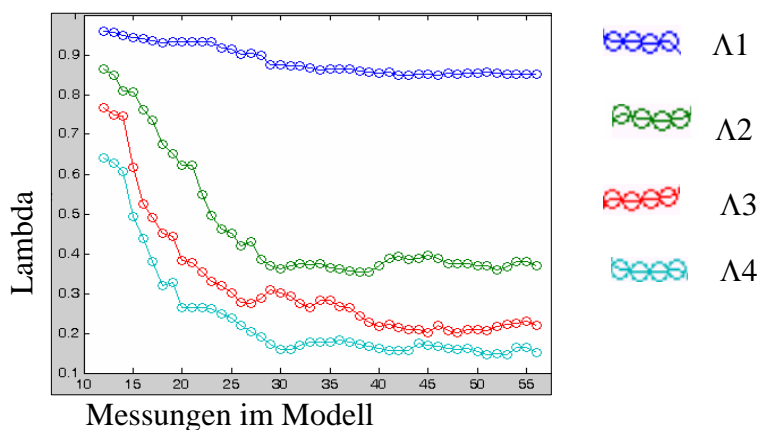


Abb. 27 Entwicklung der Separationsquotienten Lambda 1 bis 4 in Abhängigkeit von der Größe der Datenbasis des Fünf-Klassen-Modells

Die Größe der Datenbasis wird verändert durch die zunehmende Anzahl von Messungen. Die Abfolge der Messungen in der Datenbasis des LDA-Modells ist zufällig gewählt.

5.8 Vergleich der Erkennung in Abhängigkeit von der Kultivierungsdauer der bakteriellen Spezies

Die Tab. 13 zeigt die Erkennungsraten in Abhängigkeit von der Kultivierungsdauer der fünf bakteriellen Spezies im Vergleich. Es wurden für jede Spezies pro Altersstufe je vier Datensätze ausgewählt, um eine gleichmäßige Datenbasis zu erhalten. Die Erkennungsraten wurden im Fünf-Klassen-Modelltest für eine Kultivierungsdauer von drei, vier, sechs, acht und zehn Tagen berechnet. Wenn es mehr als vier verfügbare Datensätze für eine Spezies gab, wurden alle Kombinationsmöglichkeiten getestet. Die erhaltenen Ergebnisse wurden dann gemittelt.

| Kultivierungsdauer in Tagen | Erkennung von Messungen in % | | | | |
|-----------------------------|---------------------------------|---------------------|---------------------|----------------------|----------------------|
| | <i>A. actinomycetemcomitans</i> | <i>E. corrodens</i> | <i>F. nucleatum</i> | <i>P. gingivalis</i> | <i>P. intermedia</i> |
| 3 | 35 | 100 | 70 | 67 | 10 |
| 4 | 25 | 45 | 50 | 90 | 75 |
| 6 | 55 | 40 | 20 | 55 | 40 |
| 8 | 33 | 100 | 25 | 25 | 0 |
| 10 | 50 | 100 | 30 | 50 | 40 |

Tab. 13 Vergleich der Erkennung der einzelnen Spezies für verschiedene Gruppen der Kultivierungsdauer der bakteriellen Proben

Die Erkennungsdaten wurden im Fünf-Klassen-Modelltest mit jeweils vier Datensätzen pro Spezies in der Datenbasis ermittelt. Bei mehr als vier verfügbaren Datensätzen für eine Spezies wurden alle Kombinationsmöglichkeiten getestet und der Durchschnitt für die Erkennung angegeben.

Die Spezies *A. actinomycetemcomitans* wird mit 25 % bis 55 % in allen Gruppen der Kultivierungsdauer sehr schlecht erkannt. *E. corrodens* erreicht für drei Tage, acht und zehn Tage alte Proben sehr gute Erkennungsdaten von 100%. Vier, bzw. sechs Tage alte Proben werden sehr schlecht mit nur 45 % erkannt. *F. nucleatum* erreicht für die Kultivierung von drei Tagen eine gute Erkennung von 70 %. Für die Gruppen der vier bis zehn Tage alten Proben ist die Erkennung mit 25 % bis 50 % sehr schlecht. *P. gingivalis* wird mäßig gut mit 67 % in der Gruppe der drei Tage und sehr gut in der Gruppe der vier Tage kultivierten Proben mit 90 % erkannt. Sechs bis zehn Tage alte Keime erreichen nur sehr schlechte Erkennungsdaten von 25 % bis 55 %. *P. intermedia* zeigt für alle Kultivierungsgruppen eine schlechte Erkennung mit Ausnahme der vier Tage alten Proben.

Die Erkennungsdaten der jeweiligen Spezies schwanken stark zwischen den einzelnen Kultivierungsgruppen. Der Vergleich lässt keine Relation zwischen der Erkennung und der Kultivierungsdauer der Proben schlussfolgern. Die Schwankungen sehen eher statistisch als systematisch aus. Eine Ausnahme könnte für die Spezies *F. nucleatum* vermutet werden. Für diese Spezies ist die Erkennung der drei, bzw. vier Tage alten Keime deutlich besser als die der älteren Keime.

Die Tab. 14 zeigt die Erkennungsdaten der Spezies in Fünf-Klassen-Modellen der zwei bis vier Tage und sechs bis zehn Tage alten Keime im Vergleich zur Erkennung im Fünf-Klassen-Modell für zwei bis zehn Tage alte Keime. Für die jeweilige Datenbasis wurden gleiche Anzahlen von Datensätzen pro Spezies für jede Gruppe verwendet. Die Erkennung der Spezies in einem LDA-Modell aus allen Altersstufen erreicht mit Werten von nahezu 100 % für *P. gingivalis* bis 36,4 % für *P. intermedia* ähnliche Werte wie in den LDA-Modellen der Gruppe der zwei bis vier Tage

alten Keime. Je nach Auswahl der Messungen schwanken die Erkennungsraten in den Modellen für alle Gruppen der Kultivierungsdauer (Hier ist nur eine Kombination aufgeführt.) Der Vergleich der Erkennungsraten für die vier Spezies *A. actinomycetemcomitans*, *E. corrodens*, *P. gingivalis* und *P. intermedia* zeigt keine Relation zum Alter der Probe. Für die Spezies *F. nucleatum* lässt sich eine Relation zwischen Alter der Keime und Erkennung im LDA-Modell vermuten. Das LDA-Modell der zwei bis vier Tage alten Keime erreicht die höchsten Werte für die Erkennung mit 62,5 %. Vier bis sechs Tage alte Keime werden fast nicht erkannt mit nur 29 %. Diese Aussage ist jedoch nicht statistisch abgesichert.

| Dauer der Kultivierung in Tagen | Erkennungsraten Messungen in % | | | | |
|---------------------------------|---------------------------------|---------------------|---------------------|----------------------|----------------------|
| | <i>A. actinomycetemcomitans</i> | <i>E. corrodens</i> | <i>F. nucleatum</i> | <i>P. gingivalis</i> | <i>P. intermedia</i> |
| 2 bis 4 | 50,0 | 77,8 | 62,5 | 100,0 | 36,4 |
| 6 bis 10 | 66,7 | 77,8 | 28,6 | 69,2 | 15,4 |
| 2 bis 10 | 53,8 | 77,8 | 28,6 | 100,0 | 36,4 |

Tab. 14 Vergleich der Erkennung der einzelnen Spezies für zusammengefasste Gruppen der Kultivierungsdauer der bakteriellen Proben
Die Erkennungsraten wurden im Fünf-Klassen-Modelltest ermittelt.

5.9 Anwendbarkeit von Modelltests

5.9.1 Zwei-Klassen-Modell

Um das erstellte Modell für die Identifizierung von unbekanntem bakteriellen Proben der fünf Spezies im Labor anwenden zu können, sind verschiedene Tests möglich. Die einfachste Anwendung ist ein Zwei-Klassen-Modelltest der LDA. Jeweils eine bakterielle Spezies wird als eine Klasse definiert. Die verbleibenden vier werden in der zweiten Klasse zusammengefasst. Auf diese Weise kann überprüft werden, ob es sich bei der unbekanntem Laborprobe um die eine Spezies handelt oder nicht. Man bildet fünf LDA-Modelle, eines für jede Spezies. Es wird die Erkennung der einzelnen Spezies gegenüber den anderen vier, zusammengefasst in einer Klasse, ermittelt.

5.9.2 Schrittweise Erkennung

Eine weitere Möglichkeit der Überprüfung besteht darin, die Bakterien Schritt für Schritt voneinander zu unterscheiden. Man testet zuerst, ob die unbekanntem Probe zu *P. gingivalis* gehört. Wenn nicht, dann testet man weiter:

Es werden Teilmodelle für den jeweiligen Zwei-Klassen-Fall gebildet. Im ersten Modelltest wird die Erkennung von *P. gingivalis* gegenüber den anderen vier in einer Klasse zusammengefassten Spezies berechnet. *P. gingivalis* wird mit 91,4 % sehr gut erkannt (Tab. 15). Im zweiten Teilmodell wurde die Erkennung für *E. corrodens* berechnet gegen eine Klasse der verbleibenden drei Spezies (*A. actinomycetemcomitans*, *P. intermedia*, *F. nucleatum*). *E. corrodens* ist die Spezies, die als zweitbeste erkannt werden kann mit 83,9 %. Das dritte Teilmodell unterscheidet *A. actinomycetemcomitans* gegen die Klasse von *F. nucleatum* zusammengefasst mit *P. intermedia*. Hier wird *A. actinomycetemcomitans* mit einer Wahrscheinlichkeit von 69,2 % korrekt klassifiziert. Das vierte und letzte Modell unterscheidet dann *P. intermedia* von *F. nucleatum*. In diesem Modell wird *F. nucleatum* mit einer Wahrscheinlichkeit von 78,1 %, bzw. *P. intermedia* mit einer Wahrscheinlichkeit von 80,0 % korrekt erkannt.

Bei der schrittweisen Unterscheidung errechnen sich die Gesamt-Erkennungsraten für jede Spezies durch die Multiplikation der Erkennungsraten für jedes der Teilmodelle. Die korrekte Klassifizierung im ersten Modell ist Voraussetzung für korrekte Identifikation im zweiten Modelltest usw. Dadurch sinken die Erkennungsraten für jede einzelne Bakterienklasse.

| | Erkennungsrate bei schrittweisem Vorgehen für Messungen in % | Gesamt-Erkennungsrate für Messungen in % |
|---------------------------------|---------------------------------------------------------------------|-------------------------------------------------|
| <i>P. gingivalis</i> | 91,4 | 91 % |
| <i>E. corrodens</i> | 83,9 | 77 % |
| <i>A. actinomycetemcomitans</i> | 69,2 | 53 % |
| <i>P. intermedia</i> | 80,0 | 43 % |
| <i>F. nucleatum</i> | 78,1 | 43 % |

Tab. 15 Erkennungsraten für die fünf bakteriellen Geruchsklassen beim schrittweisen Vorgehen im LDA-Modelltest

Die Gesamt-Erkennung berechnet sich durch Multiplikation der Erkennungsraten für jedes der Teilmodelle.

5.9.3 Vergleich der Erkennung durch die drei Methoden

Um unbekannte Proben der fünf bakteriellen Spezies im Labor mit dem Multigassensensorsystem identifizieren zu können, wäre das angestrebte Ziel ein Fünf-Klassen-Modell, mit dem die Probe sofort klassifiziert werden kann. Das Fünf-Klassen-Modell erreicht jedoch die geringsten Erkennungsraten. Nur die Spezies *P. gingivalis* und *E. corrodens* werden mit 85,7%, bzw. 67,7 % sehr gut erkannt (Tab. 16). In der Tab. 16 sind die Erkennungsraten der drei Methoden des Modelltests für die fünf Spezies vergleichend gegenüber gestellt. Die höchsten

Erkennungsraten werden mit der Methode des Zwei-Klassen-Tests erreicht. Mit dieser Möglichkeit werden bis auf die Spezies *F. nucleatum* alle anderen vier mit 60 % bis 91 % relativ gut klassifiziert. Darauf folgt die Methode der schrittweisen Überprüfung der zu identifizierenden Keime.

| | Erkennung im Zwei-Klassen-Modell in % | | Erkennung im Fünf-Klassen-Modell in % | | Erkennung im schrittweisen Mod.test in % |
|---------------------------------------|---------------------------------------------|-----------|---------------------------------------------|-----------|------------------------------------------------|
| | Messungen | Messpkt.e | Messungen | Messpkt.e | Messungen |
| <i>P. gingivalis</i> | 91,4 | 87,8 | 85,7 | 83,4 | 91 |
| <i>A. actinomycetem- comitans</i> | 79,5 | 79,0 | 56,4 | 55,4 | 53 |
| <i>E. corrodens</i> | 71,0 | 69,7 | 67,7 | 65,9 | 77 |
| <i>P. intermedia</i> | 60,0 | 67,1 | 37,1 | 40,2 | 43 |
| <i>F. nucleatum</i> | 43,8 | 49,1 | 46,9 | 45,2 | 43 |

Tab. 16 Vergleich der Erkennungsraten des Zwei-Klassen-Modells, Fünf-Klassen-Modells und der schrittweisen Erkennung für die fünf bakteriellen Gruppen

5.10 Vergleich der Erkennung durch die zwei Sensorarrays und ihrer gemeinsamen Nutzung im Zwei-Klassen-Modell

Der untersuchte Prototyp des Multigassensensorsystems enthält zwei Gassensormikroarrays. Das Zinndioxid-Gassensormikroarray (SnO₂-GSMA) spricht hauptsächlich auf organische Verbindungen an. Das Wolframtrioxid-Gassensormikroarray (WO₃-GSMA) reagiert dagegen hauptsächlich auf anorganische Substanzen. Durch die Kombination beider Sensorarrays erhoffte man sich eine Steigerung des Informationsgehaltes. Da die Bakterien sowohl organische als auch anorganische Stoffwechselendprodukte bilden, wurde vermutet, dass die gemeinsame Auswertung der zwei Datenbasen zu einer Verbesserung der Erkennung führt.

Im Vergleich unterscheidet sich die Erkennung durch die zwei Sensorarrays nur minimal (Tab. 17). Wenn die Datensätze der beiden Gassensormikroarrays in einer Datenbasis verwendet werden, verändern sich für die Spezies *P. gingivalis*, *E. corrodens* und *A. actinomycetemcomitans* die Erkennungsraten für unbekannte Proben nur geringfügig. Jedoch verbessern sich die Erkennungsraten für die zwei Spezies *F. nucleatum* und *P. intermedia* um fast 10 %, wenn alle Datensätze der zwei Sensorarrays in einem Zwei-Klassen-Modell zusammengefasst werden.

| | Zwei-Klassen-Modell / Erkennung Messungen in % | | |
|---------------------------------|-------------------------------------------------------|----------------------------|----------------------------------------------------------|
| | SnO₂-GSMA | WO₃-GSMA | SnO₂-/WO₃-GSMA zusammen |
| <i>P. gingivalis</i> | 93 | 91 | 94 |
| <i>A. actinomycetemcomitans</i> | 71 | 80 | 81 |
| <i>P. intermedia</i> | 64 | 60 | 73 |
| <i>E. corrodens</i> | 69 | 71 | 71 |
| <i>F. nucleatum</i> | 41 | 44 | 54 |

Tab. 17 Übersicht der Erkennungsraten im Zwei-Klassen-Modell für Datensätze des Zinndioxid-Gassensormikroarrays (SnO₂-GSMA), für Datensätze des Wolframtrioxid-Gassensormikroarrays (WO₃-GSMA) und für die Nutzung der Datensätze beider Gassensormikroarrays in einer Datenbasis

5.11 Zusammenfassung der Ergebnisse

Die Ergebnisse der vorliegenden Studie lassen sich in folgender Weise zusammenfassen:

Prinzipiell ist es möglich, mit dem Prototyp des Multigassensensorsystems Olfactograph bakterielle Gerüche zu erkennen. Die Erkennung ist jedoch nicht für alle Spezies ausreichend.

Die Laborstämme der fünf Spezies *A. actinomycetemcomitans*, *E. corrodens*, *P. intermedia*, *F. nucleatum* und *P. gingivalis* sind in einem Fünf-Klassen-Modell der linearen Diskriminanzanalyse nur unzureichend durch ihre „Gerüche“ voneinander zu unterscheiden. Für die Spezies *E. corrodens* und *P. gingivalis* können zufrieden stellende Erkennungen ermittelt werden. In Abhängigkeit von gewählter Methode des Modelltests lassen sich die Erkennungen für die einzelnen Spezies auf 60 bis 90 % für vier der fünf Spezies (im Zwei-Klassen-Modell) verbessern.

In Bezug auf eine Abweichung der Sensorantwort im Laufe der Zeit können keine eindeutigen Einflüsse auf die Streuung der Messwerte festgestellt werden. Ähnliches lässt sich für den Einfluss der Kultivierungsdauer der Proben auf die Signalbildung feststellen. Es ist keine Relation zwischen dem Alter der bakteriellen Proben und der Erkennung im LDA-Modell nachweisbar.

Das Fünf-Klassen-Modell mit einer Datenbasis von 150 Datensätzen ist ausreichend stabil genug. Mit den Datensätzen von 42 bis 47 Messungen der bakteriellen Proben ist die Bildung eines stabilen Modells durch die lineare Diskriminanzanalyse möglich.

Kulturen klinischer Isolate der Spezies *P. gingivalis* und *P. intermedia* lassen sich sehr gut mit der vorgestellten Methode voneinander unterscheiden. Testmessungen von beiden werden sehr gut im Zwei-Klassen-Modell der Laborstämme identifiziert. Jedoch ist ihre Erkennung im Fünf-Klassen-Modell der Laborspezies nur für die Spezies *P. gingivalis* zufrieden stellend.

Werden jedoch wenige Datensätze von klinischen Isolaten bei der Modellbildung zusammen mit den Datensätzen der Laborspezies verwendet, verbessert sich die Erkennung für die klinischen Isolate von *P. gingivalis* und *P. intermedia* noch etwas im Fünf-Klassen-Modell.

Obwohl sich die Erkennung der Spezies durch das Wolframtrioxid- bzw. das Zinndioxid-Gassensormikroarray im Vergleich nicht unterscheiden, verbessert sich die Erkennung der Spezies *P. intermedia* und *F. nucleatum*, wenn die Datensätze der zwei Sensorarrays zusammen für die LDA-Modellbildung genutzt werden.