

Aus dem Institut für Agrar- und Ernährungswissenschaften
(Geschäftsführender Direktor: Prof. Dr. Klaus Eder)

der Naturwissenschaftlichen Fakultät III
(Dekan: Prof. Dr. Peter Wycisk)

der Martin-Luther-Universität Halle-Wittenberg

Fachgebiet: Agrarinformatik

Mathematisch-statistische Modellierung und Evaluierung von Vorhersagefunktionen zur sicheren Identifizierung tierindividueller Leistungsabweichungen beim Milchrind

Dissertation

zur Erlangung des akademischen Grades
doctor agriculturarum (Dr. agr.)

von

Diplomagraringenieur Christian Ammon

Halle/Saale 2008

urn:nbn:de:gbv:3-000013823

[<http://nbn-resolving.de/urn/resolver.pl?urn=nbn%3Ade%3Agbv%3A3-000013823>]

Aus dem Institut für Agrar- und Ernährungswissenschaften
(Geschäftsführender Direktor: Prof. Dr. Klaus Eder)

Der Naturwissenschaftlichen Fakultät III
(Dekan: Prof. Dr. Peter Wycisk)

der Martin-Luther-Universität Halle-Wittenberg

Mathematisch-statistische Modellierung und Evaluierung von Vorhersagefunktionen zur sicheren Identifizierung tierindividueller Leistungsabweichungen beim Milchrind

Dissertation

zur Erlangung des akademischen Grades
doctor agriculturarum (Dr. agr.)

vorgelegt von

Diplomagraringenieur Christian Ammon
geb. am 23.03.1975 in Dillenburg

Gutachter: Prof. Dr. J. Spilke
Prof. Dr. J. Krieter
Prof. Dr. H. H. Swalve

Verteidigung am: 07.04.2008

Halle/Saale 2008

urn:nbn:de:gbv:3-000013823

[<http://nbn-resolving.de/urn/resolver.pl?urn=nbn%3Ade%3Agbv%3A3-000013823>]

Inhaltsverzeichnis

Inhaltsverzeichnis.....	V
Abkürzungsverzeichnis	IX
Tabellenverzeichnis.....	X
Abbildungsverzeichnis	XIII
1 Einleitung und Zielstellung	1
2 Material	3
2.1 Betrieb 1	3
2.2 Betrieb 2	6
2.3 Betrieb 3	7
3 Methoden.....	10
3.1 Gemischtes Lineares Modell.....	10
3.1.1 Grundlagen	10
3.1.2 Modellierung von Laktationskurven	11
3.1.3 Modellwahl.....	12
3.1.3.1 Likelihood-Ratio-Test zur Wahl der geeigneten Variante des gemischten linearen Modells.....	14
3.1.3.2 Informationskriterien zur Wahl des Ansatzes zur Laktationskurvenmodellierung.....	15
3.1.4 Verwendungsmöglichkeiten gemischter linearer Modelle zur Qualifizierung des Herdenmanagements	17
3.2 Exponentially-weighted moving-average control charts.....	18
3.2.1 Grundlagen	18
3.2.2 Berechnung.....	18
3.2.3 Verwendung eines EWMA-Charts im Milchviehbetrieb.....	19
3.3 Fuzzy Logic.....	20
3.3.1 Grundlagen	20
3.3.2 Fuzzifizierung.....	21
3.3.2.1 Linguistische Variablen.....	21
3.3.2.2 Zugehörigkeitsfunktionen	22
3.3.2.3 Vorgehensweise bei der Fuzzifizierung	24
3.3.3 Inferenz.....	24
3.3.4 Defuzzifizierung.....	28
3.3.5 Bedeutung.....	29

3.4	Künstliche neuronale Netze	31
3.4.1	Grundlagen	31
3.4.2	Biologische Grundlagen	32
3.4.3	Funktionsweise und Architektur	33
3.4.4	Training und Lernen	35
3.4.5	Planung und Entwurf	37
3.4.6	Auswertung	38
3.4.7	Verwendungsmöglichkeiten für künstliche neuronale Netze im Milchviehbereich	39
3.4.7.1	Mehrschichtige Feedforward-Netzwerke	39
3.4.7.2	Selbstorganisierende Karten	41
3.5	Vergleich der Methoden und Modelle	44
4	Ergebnisse	46
4.1	Nutzung der Milchmenge	46
4.1.1	Definition von Milchminderungen	46
4.1.2	Anwendung des gemischten linearen Modells	48
4.1.2.1	Vergleich der Modellierungsansätze für Laktationskurven und der Modellvarianten Fixed und Random Regression	48
4.1.2.2	Erkennung von Milchminderungen mit dem gemischten linearen Modell	53
4.1.2.3	Berechnung des Abweichungsfaktors zwischen Trends von Vorhersage- und beobachteten Milchmengen	55
4.1.3	Erkennung von Milchminderungen mit EWMA-Charts	57
4.1.4	Anwendung von Fuzzy Logic	60
4.1.4.1	Modelle	60
4.1.4.2	Erkennung von Milchminderungen mittels Fuzzy Logic	67
4.1.5	Anwendung künstlicher neuronaler Netze	71
4.1.5.1	Modelle	71
4.1.5.2	Modellauswahl	71
4.1.5.3	Erkennung von Milchminderungen mittels künstlicher neuronaler Netze	73
4.2	Nutzung der elektrischen Leitfähigkeit der Milch	76
4.2.1	Erkennung von Eutererkrankungen mittels Fuzzy Logic	76
4.2.1.1	Modelle	77
4.2.1.2	Ergebnisse zur Erkennung von Eutererkrankungen	85
4.2.2	Fallbeispiele	86

4.2.3	Mastitiserkennung unter Nutzung der Leitfähigkeit	91
4.3	Nutzung des gemischten linearen Modells für das Herdenmanagement	91
4.3.1	Schätzung von herdenweiten Milchmengen	91
4.3.2	Nutzung von Testtageffekten.....	93
5	Diskussion und Schlussfolgerungen.....	97
5.1	Datenqualität und betriebliche Besonderheiten der Datenbasis.....	97
5.2	Kriterien für den Vergleich der untersuchten Methoden	100
5.3	Vergleich der untersuchten Methoden	102
5.4	Anforderungen an Merkmale zur Entscheidungsunterstützung.....	109
6	Zusammenfassung.....	111
7	Summary	113
8	Literaturverzeichnis.....	115

Abkürzungsverzeichnis

AF	-	Abweichungsfaktor
AIC	-	Akaike-Information-Criterion
AMS	-	Automatisches Melksystem
CoA	-	center-of-area
CoG	-	center-of-gravity
CoM	-	center-of-maximum
EL	-	elektrische Leitfähigkeit
EWMA	-	exponentially-weighted moving-average
FG	-	Freiheitsgrad
FRM	-	Fixed-Regression-Model
KNN	-	künstliches neuronales Netz
LCL	-	lower control limit
LL	-	log L, Log-Likelihood
LoM	-	left-of-maximum
LR	-	Likelihood-Ratio
LRT	-	Likelihood-Ratio-Test
ML	-	Maximum-Likelihood
MLP	-	multilayer Perceptron
MoM	-	mean-of-maximum
REML	-	Restricted Maximum-Likelihood
RoM	-	right-of-maximum
RNN	-	recurrent neural network
RRM	-	Random-Regression-Model
SOM	-	self-organizing map
UCL	-	upper control limit

Tabellenverzeichnis

Tabelle 1: Übersicht über die verfügbaren Betriebe	3
Tabelle 2: Herdenstruktur in Betrieb 1	4
Tabelle 3: Herdenstruktur in Betrieb 2	6
Tabelle 4: Herdenstruktur in Betrieb 3	8
Tabelle 5: Feste und zufällige Effekte der untersuchten Modellvarianten.....	14
Tabelle 6: Vergleich verschiedener Methoden zur Defuzzifizierung, nach Altrock (1997)	28
Tabelle 7: Anzahl Datensätze mit negativer Milchabweichung bei dokumentierten Euter- und Stoffwechselerkrankungen	48
Tabelle 8: Anteil dokumentierter Euter- und Stoffwechselerkrankungen und negativer Milchmengenabweichungen gegenüber den Tagesmilchmengen je Betrieb	48
Tabelle 9: Anpassungsgüte der Modelle nach Ali und Schaeffer (1987) im Vergleich	49
Tabelle 10: Anpassungsgüte der Modelle nach Wilmink (1987) im Vergleich.....	49
Tabelle 11: Anpassungsgüte der Modelle nach Guo (1998) im Vergleich	49
Tabelle 12: Ergebnisse des gemischten linearen Modells für Betrieb 1	54
Tabelle 13: Ergebnisse des gemischten linearen Modells für Betrieb 2	55
Tabelle 14: Ergebnisse des gemischten linearen Modells für Betrieb 3	55
Tabelle 15: Klassifikationsergebnisse in % für EWMA-Charts, Betrieb 1.....	58
Tabelle 16: Klassifikationsergebnisse in % für EWMA-Charts, Betrieb 2.....	58
Tabelle 17: Klassifikationsergebnisse in % für EWMA-Charts, Betrieb 3.....	59
Tabelle 18: Linguistische Variable und die entsprechenden Zugehörigkeitsfunktionen für Modell MM1	60
Tabelle 19: Regelgerüst für Modelle MM1 bis MM4.....	63
Tabelle 20: Standardabweichungen der beiden Variablen AbwAbs und AbwRel in den Betrieben	64
Tabelle 21: Linguistische Variable und die entsprechenden Zugehörigkeitsfunktionen für Modelle MM2-MM4	65
Tabelle 22: Regelgerüst für Modelle MM5 bis MM7.....	66
Tabelle 23: Auswertung für Betrieb 1 bei Grenzwert von 0.5	67
Tabelle 24: Auswertung für Betrieb 1 bei Grenzwert von 0.6	67
Tabelle 25: Auswertung für Betrieb 1 bei Grenzwert von 0.7	68
Tabelle 26: Auswertung für Betrieb 2 bei Grenzwert von 0.5	68

Tabelle 27: Auswertung für Betrieb 2 bei Grenzwert von 0.6	68
Tabelle 28: Auswertung für Betrieb 2 bei Grenzwert von 0.7	69
Tabelle 29: Auswertung für Betrieb 3 bei Grenzwert von 0.5	69
Tabelle 30: Auswertung für Betrieb 3 bei Grenzwert von 0.6	69
Tabelle 31: Auswertung für Betrieb 3 bei Grenzwert von 0.7	70
Tabelle 32: Klassifikationsergebnisse für ein KNN mit 3 Neuronen in der versteckten Schicht	72
Tabelle 33: Klassifikationsergebnisse für ein KNN mit 4 Neuronen in der versteckten Schicht	72
Tabelle 34: Klassifikationsergebnisse für ein KNN mit 5 Neuronen in der versteckten Schicht	73
Tabelle 35: Klassifikationsergebnisse für ein KNN mit 6 Neuronen in der versteckten Schicht	73
Tabelle 36: Klassifikationsergebnisse für ein KNN mit 60, 90, 120 oder 281 Tagen Vorlaufzeit zur Berechnung des Abweichungsfaktors, Betrieb 1	74
Tabelle 37: Klassifikationsergebnisse für ein KNN mit 60, 90, 120 oder 240 Tagen Vorlaufzeit zur Berechnung des Abweichungsfaktors, Betrieb 2	75
Tabelle 38: Klassifikationsergebnisse für ein KNN mit 60, 90, 120 oder 245 Tagen Vorlaufzeit zur Berechnung des Abweichungsfaktors, Betrieb 3	75
Tabelle 39: Linguistische Variable und die entsprechenden Zugehörigkeitsfunktionen für Modell LF1	78
Tabelle 40: Regelgerüst für Modell LF1	78
Tabelle 41: Vereinfachtes Regelgerüst für Modell LF1	79
Tabelle 42: Linguistische Variablen und die entsprechenden Zugehörigkeitsfunktionen für Modell LF2	79
Tabelle 43: Regelgerüst für Modell LF2	80
Tabelle 44: Linguistische Variablen und die entsprechenden Zugehörigkeitsfunktionen für Modell LF3	80
Tabelle 45: Regelgerüst für Modell LF3	81
Tabelle 46: Regelgerüst für Modell LF4	82
Tabelle 47: Linguistische Variablen und die entsprechenden Zugehörigkeitsfunktionen für Modell LF5	83
Tabelle 48: Regelgerüst für Modell LF5	84
Tabelle 49: Übersicht der Modellvarianten	84

Tabelle 50: Alarme in Abhängigkeit der Modellvariante	85
Tabelle 51: Beobachtete und vorhergesagte Herdenmilchmengen im Vergleich	92
Tabelle 52: Vergleich der untersuchten Methoden	108

Abbildungsverzeichnis

Abbildung 1: Datensätze je Tag für Betrieb 1.....	5
Abbildung 2: Datensätze je Tag für Betrieb 2.....	7
Abbildung 3: Datensätze je Tag für Betrieb 3.....	8
Abbildung 4: Standard-Zugehörigkeitsfunktionen, nach Altrock (1997)	22
Abbildung 5: Entsprechungen von Operatoren der booleschen Logik (oben) und der Fuzzy-Logic (unten), erweitert nach Borgelt et al. (2003), S. 157.....	25
Abbildung 6: Grafische Darstellung boolescher Logik im Vergleich zu Fuzzy Logic.....	26
Abbildung 7: Aggregationsschema (verändert nach MathWorks, 2000).....	26
Abbildung 8: Graphische Darstellung der Defuzzifizierung mit der MoM-, LoM- und RoM-Methode, nach Altrock (1997).....	29
Abbildung 9: Schematische Darstellung eines einfachen künstlichen neuronalen Netzes (verändert nach Patterson, 1997).....	31
Abbildung 10: Schematische Darstellung zweier Neuronen (Nauck et al., 1996).....	32
Abbildung 11: Anordnung von Neuronen in Schichten.....	34
Abbildung 12: Typische Aktivierungsfunktionen neuronaler Netze (Patterson, 1997).....	35
Abbildung 13: Probleme bei ungünstiger Lernrate: lokale Minima (links) und oszillieren um ein Minimum herum (rechts) (nach Nauck et al., 1996).....	35
Abbildung 14: Die Probleme des Underfitting und Overfitting (Zhou, 2003).....	38
Abbildung 15: Schematische Darstellung eines mehrschichtigen Feedforward- Perceptrons (verändert nach Patterson, 1997).....	39
Abbildung 16: Nachbarschaftsbeziehungen in einer quadratischen selbstorganisierenden Karte (Demuth und Beale, 2004)	41
Abbildung 17: Beispiele für eine Anordnung von Neuronen in der Ausgangsschicht einer selbstorganisierenden Karte (Borgelt et al., 2003).....	42
Abbildung 18: Entfaltung einer selbstorganisierenden Karte während des Trainings (Borgelt et al., 2003).....	43
Abbildung 19: Selbstorganisierende Karte nach Initialisierung mit ungünstigen Parametern (Borgelt et al., 2003)	43
Abbildung 20: Herdenweite Laktationskurven für Betrieb 1 nach Ali und Schaeffer (1987) bei Nutzung des Fixed-Regression-Modells.....	51
Abbildung 21: Herdenweite Laktationskurven für Betrieb 1 nach Ali und Schaeffer (1987) bei Nutzung des tierspezifischen Random-Regression-Modells	51

Abbildung 22: Herdenweite Laktationskurven für Betrieb 1 nach Ali und Schaeffer (1987) bei Nutzung des tier- und laktationsspezifischen Random- Regression-Modells	52
Abbildung 23: Herdenweite Laktationskurven für Betrieb 1 nach Wilmink (1987) bei Nutzung des Fixed-Regression-Modells	52
Abbildung 24: Herdenweite Laktationskurven für Betrieb 1 nach Guo (1998) bei Nutzung des Fixed-Regression-Modells	53
Abbildung 25: Zugehörigkeitsfunktionen der Eingabevariable absolute Abweichung der Milchmenge gegenüber dem Vortag	61
Abbildung 26: Zugehörigkeitsfunktionen der Eingabevariable prozentuale Abweichung der Milchmenge gegenüber dem Vortag	61
Abbildung 27: Zugehörigkeitsfunktionen der Eingabevariable Abweichungsfaktor	62
Abbildung 28: Zugehörigkeitsfunktionen der Ausgabevariable negative Abweichung	62
Abbildung 29: Verlauf von geschätzter und beobachteter Milchmenge sowie Leitfähigkeit in Fallbeispiel 1	86
Abbildung 30: Ergebnisse der Fuzzy-Logic-Modelle für Fallbeispiel 1	87
Abbildung 31: Verlauf von geschätzter und beobachteter Milchmenge sowie Leitfähigkeit in Fallbeispiel 2	88
Abbildung 32: Ergebnisse der Fuzzy-Logic-Modelle für Fallbeispiel 2	88
Abbildung 33: Verlauf von geschätzter und beobachteter Milchmenge sowie Leitfähigkeit in Fallbeispiel 3	90
Abbildung 34: Ergebnisse der Fuzzy-Logic-Modelle für Fallbeispiel 3	90
Abbildung 35: Laktationsverlauf mit guter Persistenz nach dem Laktationshöhepunkt, beispielhaft für einige Tiere in Betrieb 3	93
Abbildung 36: LSMEANS der Testtage in Betrieb 1 im untersuchten Zeitraum	94
Abbildung 37: LSMEANS der Testtage in Betrieb 2 im untersuchten Zeitraum	94
Abbildung 38: LSMEANS der Testtage in Betrieb 3 im untersuchten Zeitraum	95

1 Einleitung und Zielstellung

Die derzeitige Situation in der Milcherzeugung ist geprägt von hohem Kostendruck durch niedrige Milchpreise. Damit kommt kostenbeeinflussenden Faktoren wie Fruchtbarkeit, Nutzungsdauer und Krankheiten eine besondere Bedeutung zu. Fruchtbarkeitsprobleme machen sich durch erst spät oder gar nicht erfolgende Trächtigkeit nach Besamung bemerkbar. Damit einhergehend sind kurze Nutzungsdauern, derzeit von zwei oder weniger Laktationen. Krankheiten, hier vor allem Eutererkrankungen, sorgen ebenfalls häufig für ein verfrühtes Ausscheiden von Tieren aus der Herde. Auf der anderen Seite stehen in Milchviehbetrieben durch moderne Melkanlagen mit Messeinrichtungen für verschiedene Merkmale zunehmend einzeltierbezogene Daten zur Verfügung. Dies gilt sowohl für herkömmliche Melkstände, aber vor allem auch für automatische Melksysteme (AMS), bei denen es sich wegen der geringeren Zahl von Melkplätzen eher lohnt, Sensoren zur Messung von Merkmalswerten zu verwenden.

Der Nutzen dieser Daten hängt vor allem davon ab wie es gelingt, diesen Datenbestand kontinuierlich zur Entscheidungsunterstützung auf Einzeltierebene heranzuziehen. Herdenmanagementprogramme greifen dazu auf die vorhandene Datenbasis zu und stellen die Merkmalswerte der einzelnen Tiere in Tabellen- oder Diagrammform dar. Der entscheidende zusätzliche Schritt ist allerdings die Aufwertung dieser Daten zu Informationen, sei es auf Ebene des Einzeltiers, einer Tiergruppe oder der gesamten Herde. Während die nur visualisierten Daten schon die Aufmerksamkeit des Personals auf bestimmte Tiere lenken können, lassen weiterverarbeitete, zu Informationen aufbereitete Daten präzise Aussagen zu und ermöglichen somit möglicherweise schnelleres Handeln seitens des Personals.

Während die Entwicklungen in der Prozesstechnik für Melkanlagen insbesondere zu vermehrter und verbesserter Erfassung von Merkmalswerten geführt haben, sind bei der Auswertung und Aufbereitung dieser Daten zu Informationen in der Praxis wenige Neuerungen zu beobachten. Die derzeit verwendeten Methoden zur Entscheidungsunterstützung in verfügbaren Herdenmanagementsystemen beschränken sich vorwiegend auf Nutzung gleitender Mittelwerte zur Identifikation von Merkmalsabweichungen. Dabei wird meist nur ein einzelnes Merkmal überwacht, beispielsweise die Leitfähigkeit der Milch zur Identifikation von Eutererkrankungen. In den letzten Jahren wurden jedoch weitere Methoden hinsichtlich ihrer Nutzungsmöglichkeiten bei verschiedenen Fragestellungen im Rahmen der Milcherzeugung untersucht. Gegenstand der Untersuchungen waren vor allem eine bessere Kontrolle von Eutergesundheit und Brunst durch

Nutzung von *exponentially-weighted moving-average control charts* (EWMA-Charts), Fuzzy Logic und künstlichen neuronalen Netzen sowie die Zuchtwertschätzung und Überwachung von Milchmengen mittels gemischter linearer Modelle. Diese Methoden bieten gegenüber einfachen gleitenden Mittelwerten verschiedene neue Ansätze, wie unterschiedliche zeitliche Gewichtung von Daten, tierindividuelle Anpassung, rechnergerechte Umsetzung von Expertenwissen und Möglichkeit der gleichzeitigen Überwachung von Merkmalskombinationen. Insbesondere die Aussicht, qualifizierte Aussagen über mögliche Gründe einer verminderten Milchmenge treffen zu können, kann die Verwendung von Methoden, die gleichzeitig mehrere Merkmale berücksichtigen, für das Herdenmanagement interessant machen. Beispielsweise ließe bei einer stark negativen Abweichung der Milchmenge eine gleichzeitig erhöhte elektrische Leitfähigkeit der Milch den Schluss auf eine Eutererkrankung zu, während bei normalen Leitfähigkeitswerten eher Stoffwechselerkrankungen in Frage kommen. Ein umfassender Methodenvergleich hinsichtlich ihrer Möglichkeiten und Leistungsfähigkeit liegt jedoch noch nicht vor.

Vor diesem Hintergrund ergeben sich für die vorliegende Arbeit als Zielstellungen

- (1) die Bewertung der verfügbaren Daten hinsichtlich der Datenqualität und der sich daraus ergebenden Nutzungsmöglichkeiten oder -einschränkungen,
- (2) die Modellbildung und -anpassung für ein gemischtes lineares Modell zur tierspezifischen Vorhersage von Milchmengen unter Berücksichtigung verschiedener Ansätze für die Beschreibung des Laktationskurvenverlaufs,
- (3) der Vergleich der Methoden EWMA-Chart, gemischtes lineares Modell, Fuzzy Logic und künstliche neuronale Netze, basierend auf Erkennung von Abweichungen beim Merkmal Milchmenge, sowie
- (4) die Darstellung der Möglichkeiten der organisatorischen Einordnung der untersuchten Methoden innerhalb des betrieblichen Informationsmanagements.

2 Material

Für die Untersuchung standen Daten aus drei Betrieben zur Verfügung. Zwei dieser Betriebe sind normale Produktionsbetriebe, deren Daten aus der laufenden Produktion stammen und aus den jeweiligen Herdenmanagementprogrammen oder Programmen der Prozessrechner entnommen wurden. Damit wird der Stand der Datenerfassung unter Produktionsbedingungen charakterisiert. Die Produktionsbetriebe unterscheiden sich im Umfang des Tierbestandes und in der verwendeten Melktechnik. Der dritte Betrieb ist ein Versuchsbetrieb. Dies bietet eine vergleichsweise zuverlässigere, nahezu lückenlose Datenbasis mit einer Anzahl von zusätzlich erfassten Merkmalen, die unter Produktionsbedingungen in der Regel nicht zur Verfügung stehen. Andererseits ist aber vorstellbar, dass eine solche Datenbasis wegen der höheren Genauigkeit bessere Ergebnisse erzielt, als das Daten aus den Produktionsbetrieben zulassen. Eine Vergleichbarkeit der Produktionsbetriebe und des Versuchsbetriebs ist damit nicht gegeben.

Im Weiteren werden die Betriebe mit Betrieb 1, Betrieb 2 und Betrieb 3 bezeichnet. Nähere Informationen sind in Tabelle 1 zusammengestellt. Für alle drei Betriebe gilt, dass nur Datensätze mit Tagesmilchmengen von mindestens 5 kg je Kuh berücksichtigt wurden. Ebenso wurden nur Datensätze bis zum 400. Laktationstag verwendet.

Tabelle 1: Übersicht über die verfügbaren Betriebe

Betrieb	Art	Melktechnik	Herden-durchschnittsbestand	Zeitraum der Datenverfügbarkeit
Betrieb 1	Produktionsbetrieb	AMS ^a	~130	806 Tage
Betrieb 2	Produktionsbetrieb	Melkstand	~330	1007 Tage
Betrieb 3	Versuchsbetrieb	Melkstand	~50	1475 Tage

^a Automatisches Melksystem

2.1 Betrieb 1

Von Betrieb 1 sind Daten über einen Zeitraum von 806 Tagen vorhanden. Die verwendete Melktechnik ist ein automatisches Melksystem (AMS). In diesem Zeitraum wurden von 394 Kühen Merkmalswerte erfasst. Der Bestand lag zu Anfang bei etwa 120 Tieren, eingeteilt in drei Gruppen (je eine Melkbox pro Gruppe), gegen Ende des verfügbaren Zeitraums nach Aufstockung bei etwa 160 Tieren. Mit der Aufstockung des Bestandes einher geht die Inbetriebnahme einer vierten Melkbox und somit Unterteilung der Tiere in vier Melkgruppen. Die Tiere werden durchschnittlich 2.97mal am Tag gemolken. Die mittlere Laktationszahl beträgt 1.92. Insgesamt liegen 235657 Datensätze einzelner Melkungen vor. Die Milchmenge wird nur auf das Gesamteuter bezogen erfasst,

während Leitfähigkeit, Melkdauer, mittlerer und maximaler Milchfluss als viertelbezogene Merkmale vorliegen. Tabelle 2 verschafft einen Überblick über die Herdenstruktur im untersuchten Zeitraum.

Tabelle 2: Herdenstruktur in Betrieb 1

Laktationsnummer	Anzahl Kühe	Anzahl Laktationen	Anzahl Einzelmelke	Anzahl Leitfähigkeitswerte	Anzahl Tagesmilchleistungen	Milchmenge je Tag (kg)			
						Mittelwert	Standardabweichung	Minimalwert	Maximalwert
total	394	608	235657	198104	79254	31.98	10.02	0.88	77.13
1		295	105449	86400	37029	28.42	7.42	1.31	59.42
2		162	75780	64443	23962	35.0	10.23	0.94	76.33
3		93	33466	29230	11131	34.85	11.91	0.88	77.13
4+		58	20962	18031	7132	35.89	11.45	1.8	73.59

Bedeutsam im vorliegenden Sachzusammenhang ist die tatsächliche Datenabdeckung des beobachteten Zeitraums. Innerhalb dieser 806 Tage liegen nur an 607 Tagen Messwerte vor, also nur für 75.3% des Beobachtungszeitraums (vgl. Abbildung 1). Die Verfügbarkeit der viertelspezifischen Leitfähigkeitswerte beziffert sich auf etwa 84.1% der Einzelgemelksdatensätze. Dies sind keine Ausfälle einzelner Tiere, sondern technische Probleme bei der Datenerfassung oder -übertragung. Einzelne fehlende Werte treten durchschnittlich alle vier Tage auf. Aus Abbildung 1 ist zu ersehen, dass gerade mit der Installation der vierten Melkbox in der zweiten Hälfte des Datenerfassungszeitraums technische Probleme auftraten, die dazu führten, dass an mehreren aufeinander folgenden Tagen keine Daten zur Verfügung standen. Auch während des ersten halben Jahres sind Fehler bei der Datenübertragung zu vermerken, die jedoch nicht zu Totalausfällen führten.

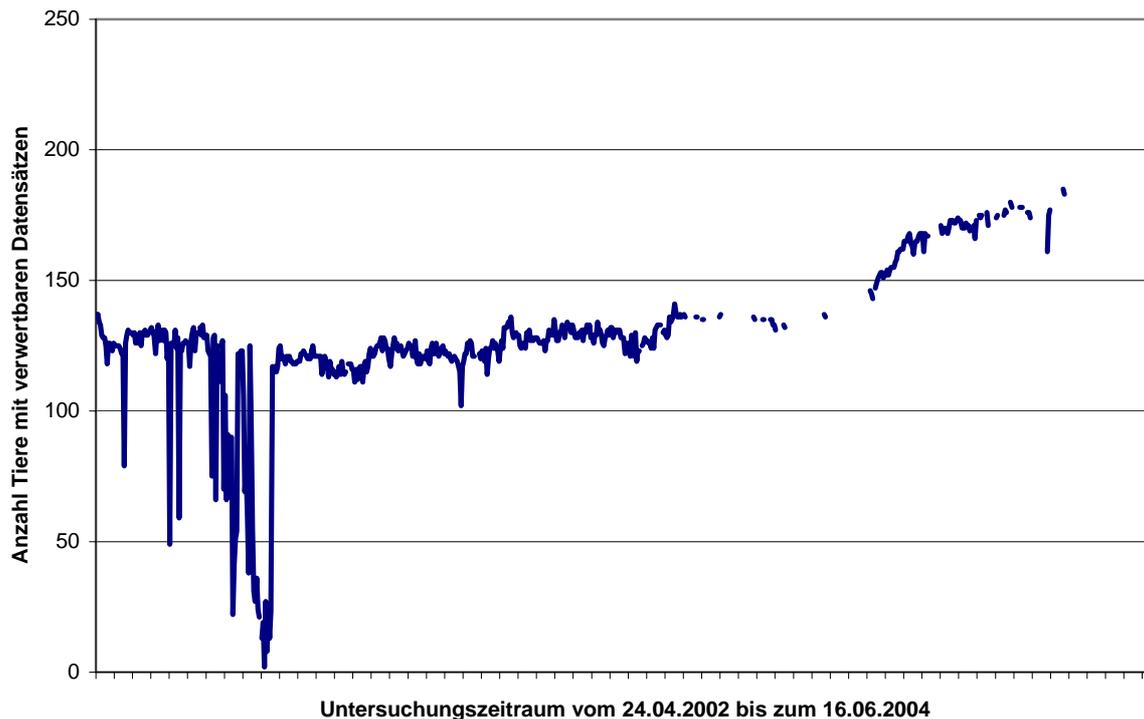


Abbildung 1: Datensätze je Tag für Betrieb 1

Für die Anwendung der in Abschnitt 3 beschriebenen Methoden zur Erkennung von Milchmengenabweichungen wurde daher nur der Zeitraum ohne Fehltage bzw. stärkere Schwankungen verwendet. Dieser Zeitraum umfasst 281 Tage.

Ein selten auftretendes Problem ist die Fehlerkennung einer Kuh in der Melkbox. Dies kann auftreten, wenn eine nicht in der Melkbox befindliche Kuh mit ihrem Transponder in den Antennenbereich der Melkbox gerät. Dadurch wird die ermolkenen Milchmenge dieser Kuh anstelle der eigentlich gemolkenen Kuh zugewiesen. Erkennbar sind diese Fehlerkennungen im Datensatz häufig durch sehr kurze Melkzeiten bei sehr hohen Milchflusswerten.

Fehltage einzelner Tiere weisen häufig auf Krankheiten hin, die einen Wechsel der betroffenen Kuh in die Krankengruppe nötig machen. Dort werden die kranken Tiere in einem Melkstand gemolken, von dem keine Daten erfasst werden. Entsprechend der Zielstellung der vorliegenden Arbeit werden auch gesicherte Daten zu Krankheitsfällen während des Beobachtungszeitraums benötigt. Die Erfassung von Krankheiten und sonstigen Vorfällen für das Einzeltier entsprach in ihrer Genauigkeit den derzeit üblichen Produktionsbedingungen (Dokumentation in Herdbuch bzw. Herdenmanagementsoftware). Entsprechend ist nicht von einer lückenlosen Erkennung und Dokumentation der Krankheiten auszugehen.

2.2 Betrieb 2

Der zweite Betrieb ist ebenfalls ein Produktionsbetrieb, von dem die Daten aus der laufenden Milchproduktion vorlagen. Über einen Zeitraum von 1007 Tagen liegen Datensätze von insgesamt 420 Kühen vor. Die Datenverfügbarkeit innerhalb des Zeitraums beträgt dabei 95.5%, graphisch dargestellt in Abbildung 2. Der Durchschnittsbestand liegt bei ca. 330 Tieren, die mittlere Nutzungsdauer liegt bei etwa 2.3 Laktationen. Die Herdenstruktur im untersuchten Zeitraum kann Tabelle 3 entnommen werden. Neben der Milchleistung steht auch die Leitfähigkeit als Merkmal zur Verfügung. Allerdings werden hier keine viertelspezifischen Messwerte erfasst, sondern ein Gesamtwert für das Euter. Die Kühe werden in unterschiedlichen Leistungsgruppen gemolken. Zu Beginn der Laktation wird in den Hochleistungsgruppen dreimal am Tag gemolken, gegen Ende der Laktation nur noch zweimal. Probleme bei der Datenübertragung zwischen dem Prozessrechner des Melkstandes und der Herdenmanagementsoftware führten zu fehlenden Datensätzen im Herdenmanagementprogramm, so dass die Daten direkt aus dem Programm des Prozessrechners ausgelesen wurden.

Tabelle 3: Herdenstruktur in Betrieb 2

Laktationsnummer	Anzahl Kühe	Anzahl Laktationen	Anzahl Tagesmilchleistungen	Milchmenge je Tag (kg)			
				Mittelwert	Standardabweichung	Minimalwert	Maximalwert
total	420	778	198562	28.91	10.06	0.2	83.2
1		249	65145	25.83	8.33	0.2	79
2		213	58841	29.61	10.33	0.2	69.6
3		186	42441	31.31	10.61	0.2	83.2
4+		130	32135	30.79	10.53	0.2	72.4



Abbildung 2: Datensätze je Tag für Betrieb 2

Die Kühe werden in einem 2x16 Side-by-side-Melkstand gemolken. Zur Identifikation sind je Melkplatz Antennen zur Einzeltiererkennung verfügbar. Dennoch kann es auch hier dazu kommen, dass Tiere falsch erkannt werden. Dies tritt auf, wenn Transponder in den Antennenbereich eines benachbarten Tierplatzes kommen. Zu Fehlerfassungen führen auch Zwischenfälle beim Melken die es erforderlich machen, die Tiere umzutreiben und erneut in den Melkstand zu lassen, obwohl sie für diesen Melkdurchgang schon erkannt und verbucht wurden. Dadurch kommt es auch in diesem Betrieb in einigen Fällen dazu, dass Milchmengen einem Tier falsch zugeordnet sind.

2.3 Betrieb 3

Dieser Versuchsbetrieb lieferte eine zeitlich, aber auch die Merkmale betreffend umfangreiche Datenbasis. Der gesamte verfügbare Beobachtungszeitraum (dargestellt in Abbildung 3) umfasst 1475 Tage. Es nehmen insgesamt 77 Tiere an Versuchen teil. Die Daten für jedes Tier sind täglich dokumentiert worden. Die Datenabdeckung des untersuchten Zeitraums beträgt demnach fast 100%. Alle Tiere treten wenige Tage nach der Kalbung in einen Versuch ein und scheiden zu Ende der Laktation mit dem Trockenstellen wieder aus. Diese Vorgehensweise führt dazu, dass über mehrere Wochen gegen Anfang und Ende der Versuche nur von wenigen Tieren Daten vorliegen. Um eine ausreichende Mindestanzahl von Tieren je Tag und somit eine adäquate Datenbasis mit einer

Tabelle 4: Herdenstruktur in Betrieb 3

Laktationsnummer	Anzahl Kühe	Anzahl Laktationen	Anzahl Tagesmilchleistungen	Milchmenge je Tag (kg)			
				Mittelwert	Standardabweichung	Minimalwert	Maximalwert
total	54	58	12863	27.37	7.86	5.5	62.9
1		22	4984	25.41	5.2	9.1	43.6
2		12	2224	28.25	6.07	6.9	44.3
3		11	2866	25.3	9.03	6.1	51.7
4+		13	2789	31.79	9.29	5.5	62.9

möglichst gut den üblichen Produktionsbedingungen entsprechenden Verteilung der Laktationsstadien für die Berechnungen zur Verfügung zu haben, wurde der verwendete Zeitraum auf einen Bereich von 245 Tagen eingeschränkt. In diesem Zeitraum lagen täglich Daten von mindestens 35 Tieren vor. Dadurch verringerte sich die Anzahl der untersuchten Tiere auf 54. Die Zusammensetzung der Herde in diesem verwendeten Zeitraum ist in Tabelle 4 dargestellt.

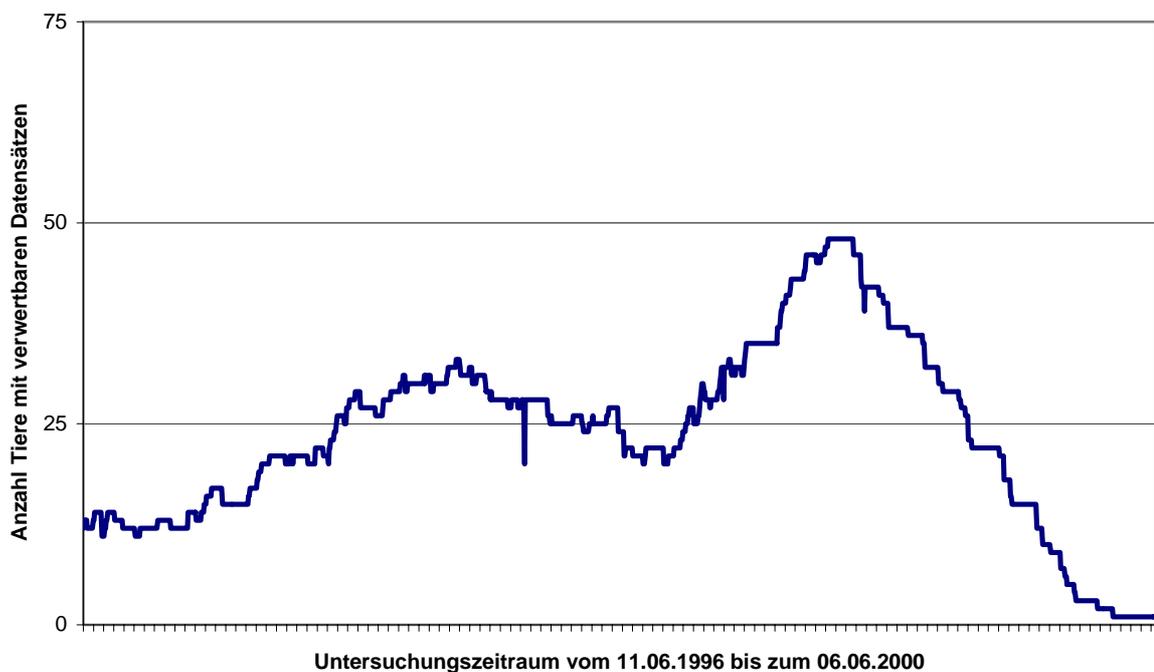


Abbildung 3: Datensätze je Tag für Betrieb 3

Gemolken werden die Kühe zweimal täglich in einem Melkstand. Außer der Milchmenge und dem Milchfluss werden während des Melkens jedoch keine weiteren Merkmale erfasst. Die Merkmale liegen tierspezifisch vor und sind pro Tag in einem Datensatz zusammengefasst. Weitere Kennzahlen zur Milch (Fettgehalt, Eiweißgehalt, Laktosegehalt, Harnstoffgehalt und Zellzahl) liegen als Laborwerte vor. Diese Merkmale werden als Laborwerte allerdings entsprechend weniger häufig erfasst und sind somit nicht „tagesaktuell“. Die übliche Frequenz der Erfassung liegt bei ein oder zwei Wochen. Somit scheiden diese Merkmale zur kurzfristigen Entscheidungsunterstützung weitgehend aus und sind eher für langfristige Managemententscheidungen nutzbar.

3 Methoden

In diesem Abschnitt sollen zunächst die verwendeten Methoden beschrieben werden.

3.1 Gemischtes lineares Modell

3.1.1 Grundlagen

Die tägliche Milchleistung eines Tieres ist Ergebnis genetischer und umweltbedingter Faktoren. So spielen neben dem genetischen Merkmalswert z.B. auch die Laktationsnummer, der aktuelle Laktationstag und die aktuellen Umweltbedingungen am Untersuchungstag eine Rolle. Wenn über diese Einflussfaktoren Strukturdaten vorliegen, lassen sie sich als erklärende Variable in einem Vorhersagemodell für eine tierindividuelle Leistung verwenden. Die Grundidee für die Nutzung gemischter linearer Modelle besteht darin, mit vorhandenen Daten bis zum Zeitpunkt t einen Vorhersagewert für die Milchleistung zum Zeitpunkt $t+1$ zu schätzen und anschließend mit der zum Zeitpunkt $t+1$ tatsächlich beobachteten Milchleistung zu vergleichen. Je nach Abweichung der Beobachtungswerte vom Schätzwert können dann Rückschlüsse auf den Zustand des Tieres gezogen werden, so z.B. auf Krankheiten bei stark negativer Leistungsabweichung.

Zur Vorhersage soll hier ein gemischtes lineares Modell der Form

$$\underline{y} = X\underline{\beta} + Z\underline{u} + \underline{e} \quad (1)$$

mit

$\underline{y} = n \times 1$ – Vektor der Beobachtungswerte,

$\underline{\beta} = p \times 1$ – Vektor der p unbekanntesten festen Effekte,

$\underline{u} = q \times 1$ – Vektor der q unbekanntesten zufälligen Effekte,

$\underline{e} = n \times 1$ – Vektor der n unbekanntesten zufälligen Resteffekte,

$X =$ bekannte $n \times p$ – Versuchsplanmatrix der festen Effekte,

$Z =$ bekannte $n \times q$ – Versuchsplanmatrix der zufälligen Effekte

verwendet werden.

Angenommen wird des Weiteren:

$$\underline{u} \sim N(0; G), \quad \underline{e} \sim N(0; R),$$

$$E(\underline{y}) = X \beta, \quad \text{Var}(\underline{y}) = ZGZ' + R = V,$$

$$\text{Var} \begin{pmatrix} \underline{y} \\ \underline{u} \\ \underline{e} \end{pmatrix} = \begin{bmatrix} V & ZG & R \\ GZ' & G & 0 \\ R & 0 & R \end{bmatrix}$$

(HENDERSON, 1990, S. 1ff).

Unter Berücksichtigung der vorliegenden Datenstruktur bilden der Testtag, die Laktationsnummer und die verwendeten Regressionskoeffizienten zur Modellierung einer Laktationskurve die festen Effekte in Vektor β . Vektor \underline{u} beinhaltet den zufälligen Tiereffekt. Alle weiteren zufälligen Effekte beinhaltet der Vektor \underline{e} als Resteffekte. Ein Modell dieses Typs bezeichnet man als Fixed-Regression-Testtagsmodell. Es wird beispielsweise in der Zuchtwertschätzung beim Milchrind verwendet, wobei aber mehrere Merkmale geschätzt werden und auch die Verwandtschaftsmatrix berücksichtigt wird (SWALVE, 2000). Eine Erweiterung stellt das Random-Regression-Modell dar, bei dem neben dem zufälligen Tiereffekt noch zufällige Effekte für den Verlauf von tierspezifischen Laktationskurven geschätzt werden. Da diese Modelle sowohl feste als auch zufällige Effekte beinhaltet, werden sie als gemischte lineare Modelle bezeichnet. Spezialformen davon sind Modelle, die nur feste Effekte und den zufälligen Resteffekt enthalten (allgemein als Modell I bezeichnet), oder Modelle, die neben dem allgemeinen Mittel nur zufällige Effekte enthalten (Modell II).

3.1.2 Modellierung von Laktationskurven

Den eigentlich nichtlinearen Verlauf der Laktationskurve kann man im linearen Modell mit geeigneten Covariablen abbilden, ohne den linearen Ansatz verlassen zu müssen. In der vorliegenden Arbeit werden die Ansätze zur Modellierung von Laktationskurven von ALI und SCHAEFFER (1987), WILMINK (1987) und von GUO (1998) verwendet. Die Laktationskurve nach ALI und SCHAEFFER (1987) sieht vier Covariablen zur Beschreibung vor:

$$X_1 = \text{Laktationstag}/400$$

$$X_2 = (X_1)^2$$

$$X_3 = \ln(400/\text{Laktationstag})$$

$$X_4 = (X_3)^2$$

Der Wert von 400 bezieht sich dabei auf die Länge der beobachteten Laktation in Tagen. Falls die durchschnittliche Dauer einer Laktation in einem Betrieb stark von dieser Zahl abweicht, ist eine Anpassung an die tatsächliche durchschnittliche Laktationsdauer in Tagen sinnvoll.

WILMINK (1987) verwendet demgegenüber nur zwei Parameter zur Modellierung der Laktationskurve:

$$X_1 = \text{Laktationstag}$$

$$X_2 = e^{-0.05 \text{ Laktationstag}}$$

Dieser Ansatz wird beispielsweise im kanadischen Testtagsmodell verwendet.

Der Ansatz von GUO (1998) sieht ebenfalls zwei Parameter vor:

$$X_1 = \text{Laktationstag}^{0.5}$$

$$X_2 = \ln(\text{Laktationstag})$$

Je nach verwendetem Modell wird der Vektor der festen Effekte β demnach neben Testtag und Laktationsnummer vier oder zwei weitere Regressionskoeffizienten beinhalten. Weitere Ansätze sind bei GUO (1998) beschrieben.

3.1.3 Modellwahl

Der Vergleich der Modelle erfolgt mit Hilfe von testbasierten und analytischen Kriterien. Dafür stehen in der verwendeten Software zur Berechnung der gemischten linearen Modelle verschiedene Kriterien zur Modellselektion zur Verfügung. Es werden davon die -2 Log Likelihood und das Akaike-Information-Criterion (AIC; AKAIKE, 1969) betrachtet (WADA und KASHIWAGI, 1990; MIELENZ et al., 2005; SPILKE und MIELENZ, 2006). Für beide Größen gilt, dass die Anpassung eines Modells umso besser ist, je niedriger der Wert des Kriteriums ausfällt. Um eine Modellbewertung bei Beachtung fester und zufälliger Effekte vornehmen zu können, wurde für die Aufstellung der Likelihood-Funktion die Methode ML (*Maximum Likelihood*) zur Optimierung der Erwartungs-

wertstruktur gegenüber der Methode REML (*Restricted Maximum Likelihood*) zur Optimierung der Kovarianzstruktur verwendet. Der berechnete Wert der Likelihood-Funktion kann zur Bewertung der Modellanpassung zwischen den Funktionen nach ALI und SCHAEFFER (1987) sowie nach WILMINK (1987) und nach GUO (1998) verwendet werden. Hierauf wird detailliert im Rahmen des analytischen Kriteriums in Abschnitt 3.1.3.2 eingegangen.

Neben der Wahl des Modells zur Beschreibung der Laktationskurve ist weiterhin die übrige Modellstruktur festzulegen. Im *Fixed-Regression*-Modell (im Weiteren als FRM bezeichnet) ist im Allgemeinen der Tiereffekt der einzige zufällige Effekt. Für alle Tiere innerhalb einer Laktation werden dieselben Regressionskoeffizienten für den Laktationskurvenverlauf geschätzt, d.h. die Laktationskurven innerhalb einer Laktation unterscheiden sich nur hinsichtlich des tierspezifischen Leistungsniveaus, jedoch nicht im generellen Kurvenverlauf. Dieser Zusammenhang ist aber aufgrund der biologischen Varianz als unrealistisch anzusehen, weshalb es sinnvoll erscheint, weitere zufällige Regressionskoeffizienten einzuführen und mit einem *Random-Regression*-Modell (im Weiteren als RRM bezeichnet) zu arbeiten (vgl. MIELENZ et al., 2006). Zu diesem Zweck wurden zwei Varianten als RRM formuliert. Die erste Variante beinhaltet ebenfalls tierspezifische zufällige Effekte und zusätzlich tierspezifische zufällige Regressionskoeffizienten zur Laktationskurvenschätzung.

Diese Modellierung hat eine Schätzung einer eigenen Laktationskurve für jedes Tier zum Ergebnis, da neben den festen Regressionskoeffizienten zur Beschreibung der Laktationskurve auf Herdenniveau zusätzlich tierspezifische Regressionskoeffizienten berücksichtigt werden. Die Laktationskurven zweier Tiere in derselben Laktation unterscheiden sich demnach nicht durch eine über den Laktationsverlauf konstante Differenz der Tiereffekte wie im FRM, sondern können auch unterschiedliche tierspezifische Steigungen und Gefälle aufweisen. Bei dieser Variante wird jedoch weiterhin davon ausgegangen, dass sich die tierspezifischen Regressionskoeffizienten für den Laktationskurvenverlauf in aufeinander folgenden Laktationen nicht ändern. Diese Möglichkeit wird in der zweiten Variante berücksichtigt, da auch die Laktationen eines Tieres durchaus genetisch und umweltbedingte unterschiedliche Steigerungsraten haben können (vgl. LIU et al., 2003).

In Variante 2 des RRM wird daher auch für jede Laktation tierspezifisch zusätzlich ein separater zufälliger Effekt geschätzt. Je Tier sind also für jede betrachtete Laktation unterschiedliche Steigungen und Leistungsniveaus in den Laktationskurven zu erwarten. Eine Vergleichsübersicht über die verschiedenen festen und zufälligen Effekte aller drei Varianten zeigt Tabelle 1. Da diese Modellvarianten hierarchisch aufeinander aufbauen, kann zum Vergleich der Varianten innerhalb einer Laktationskurvenmodellierung der Likelihood-Ratio-Test verwendet werden.

Tabelle 5: Feste und zufällige Effekte der untersuchten Modellvarianten

Modellvariante		Fixed Regression	Random Regression I	Random Regression II
Feste Effekte	Laktation	x	x	x
	Testtag	x	x	x
	Regressionskoeffizienten innerhalb der Laktationsnummer	x	x	x
Zufällige Effekte	Tier	x	x	
	Tier innerhalb Laktation			x
	Regressionskoeffizienten zur Modellierung der Laktationskurve innerhalb Tier		x	
	Regressionskoeffizienten zur Modellierung der Laktationskurve innerhalb Tier und Laktation			x

3.1.3.1 Likelihood-Ratio-Test zur Wahl der geeigneten Variante des gemischten linearen Modells

Mit zunehmender Anzahl zufälliger Effekte steigt die Komplexität der Modelle. Dabei sind die Modellvarianten hierarchisch aufgebaut, denn durch Festlegen eines oder mehrerer Parameter (in diesem Fall werden die entsprechenden Varianzkomponenten auf 0 gesetzt) können die restriktiven Varianten als Spezialfälle der komplexeren Modelle angesehen werden (vgl. auch MIELENZ et al., 2006). Um zu überprüfen, ob das komplexere und weniger eingeschränkte Modell tatsächlich eine signifikant bessere Anpassung liefert als das einfachere und damit restringierte Modell, wird der *Likelihood-Ratio-Test* (LRT, auch Likelihood-Verhältnis-Test oder Likelihood-Quotienten-Test) verwendet. Wenn kein signifikanter Unterschied zu finden ist, ist das restringierte Modell zu bevorzugen, da es bei vergleichbarer Erklärungskraft das einfachere Modell ist. Die LRT-Statistik ist wie folgt definiert (vgl. LEHMANN, 1997, S. 16f):

$$LR = -2 \ln \left(\frac{L_r}{L_u} \right) = (-2) \cdot [\ln(L_r) - \ln(L_u)] = 2 \cdot \ln(L_u) - 2 \cdot \ln(L_r) \quad (2)$$

L_u ist dabei die Likelihood der uneingeschränkteren, komplexeren Variante, L_r die des einfacheren, restriktiveren Modells, \ln ist der natürliche Logarithmus. Anstelle der Likelihood, die eine sehr kleine Größe ist, arbeiten Statistikprogramme meist mit dem besser handhabbaren Logarithmus der Likelihood (*Log Likelihood*, abgekürzt $\log L$), oder wie SAS mit der $-2 LL$. Die $\log L$ ist eine negative Zahl, da die Likelihood nur Werte zwischen 0 und 1 annehmen kann und ihr Logarithmus somit kleiner als 0 ist. Daher erhält man bei der Verwendung der $-2 \log L$ eine positive Zahl.

Die LR folgt annähernd der Chi-Quadrat-Verteilung. Die Anzahl der Freiheitsgrade ergibt sich aus der Anzahl der Restriktionen im eingeschränkten Modell, die notwendig sind, um es als Spezialfall des komplexeren Modells aufzufassen (MIELENZ et al., 2006). Für ein zu wählendes Signifikanzniveau α kann nun der Tabellenwert der Chi-Quadrat-Verteilung (χ^2_{Tab}) abgelesen werden. Es gilt:

$$LR \leq \chi^2_{Tab} \Rightarrow H_0$$

$$LR > \chi^2_{Tab} \Rightarrow H_1$$

Annahme der Nullhypothese H_0 bedeutet, dass keine Abweichung zwischen komplexem und einfacherem Modell vorliegt. Wird H_1 angenommen, ist das komplexere Modell signifikant besser als das restringierte Modell.

3.1.3.2 Informationskriterien zur Wahl des Ansatzes zur Laktationskurvenmodellierung

Zur Wahl des geeigneten Ansatzes zur Modellierung des Laktationskurvenverlaufs steht das *Akaike-Information-Criterion* (AIC) zur Verfügung. Bei der nachfolgend genutzten Form gilt, dass die Anpassungsgüte besser ist, je kleiner die entsprechenden Werte der Kriterien für die Anpassungsgüte sind.

Für den Vergleich der Ansätze von ALI und SCHAEFFER (1987) einerseits sowie WILMINK (1987) und GUO (1998) andererseits ist, wie bereits angesprochen, die Verwendung der Maximum-Likelihood-Methode (ML) notwendig. Die übliche und bei der späteren Berechnung von Vorhersagewerten zu nutzende REML-Methode berücksichtigt bei der Optimierung der Kovarianzstruktur nur die zufälligen Effekte. Da sich bei den untersuchten Ansätzen zur Modellierung der Laktationskurven jedoch die Anzahl der festen Effekte, für die Regressionskoeffizienten geschätzt werden, bei den drei Ansätzen unterscheidet, ist die REML-Methode hier nicht für den Zweck des Modellvergleichs geeignet.

Das AIC ist definiert als (AKAIKE, 1969):

$$AIC = -2 \log L + 2d \quad (3)$$

Dabei ist d die Summe der frei wählbaren festen Effekte und der Varianzkomponenten im betrachteten Modell. AIC berücksichtigt also zusätzlich zur Likelihood die Komplexität des Modells, indem je zusätzlichem Parameter ein Strafterm von 2 Einheiten vergeben wird, der den AIC erhöht (vgl. auch NGO und BRAND, 1997). Für kleine Stichproben mit $n/d < 40$, wobei n die Anzahl der Beobachtungen ist, wird in der Literatur die Verwendung des korrigierten AIC empfohlen, das in SAS als AICC vorliegt (vgl. MIELENZ et al., 2006; HURVICH und TSAI, 1989).

$$AICC = -2 \log L + \frac{n(n+d)}{(n-d-2)} \quad (4)$$

Da das günstigste Modell dasjenige mit dem niedrigsten AIC bzw. AICC ist, werden somit einfachere Modelle gegenüber komplexeren bevorzugt, wenn das komplexere Modell keine bessere Anpassung an das unbekanntes „wahre“ Modell bietet. Der komplexere Ansatz von ALI und SCHAEFFER (1987) muss also hinsichtlich der Anpassung der Laktationskurve an den tatsächlichen Verlauf besser geeignet sein als die mit weniger Parametern auskommenden Ansätze von WILMINK (1987) oder GUO (1998), um diesen Strafterm für Komplexität wieder wettzumachen.

3.1.4 Verwendungsmöglichkeiten gemischter linearer Modelle zur Qualifizierung des Herdenmanagements

Zunächst ist die bereits angedeutete Nutzungsmöglichkeit zur Milchmengenvorhersage zum anschließenden Vergleich mit tatsächlich beobachtbarer Milchmenge zur Identifikation von negativen Milchmengenabweichungen, beispielsweise durch Krankheiten bedingt, zu nennen (vgl. VAN BEBBER et al., 1999). Dies geschieht zunächst auf Einzeltierbasis. Mit Hilfe der geschätzten Regressionskoeffizienten für die festen und zufälligen Effekte aus den zum Zeitpunkt t vorliegenden Beobachtungen kann für $t+1$ eine Vorhersage angegeben werden. Zum Zeitpunkt $t+1$ wird der beobachtete Wert mit dem für $t+1$ vorhergesagten Wert verglichen. Liegt der beobachtete Wert innerhalb einer vorgegebenen Konfidenzschranke, liegt keine bedeutsame Abweichung vor.

Eine weitere Möglichkeit besteht in der Nutzung der geschätzten Parameter zur Erzeugung von laktationsspezifischen Milchleistungskurven auf höherer Organisationsebene (Herdengruppen, ganze Herde). Diese Kurven können beispielsweise zur Überprüfung der Leistung in verschiedenen Phasen der Laktation verwendet werden. Ebenso sind Schätzungen von kumulierten Milchleistungen über die gesamte Herde möglich, die zur Unterstützung des betrieblichen Quotenmanagements herangezogen werden können.

Des Weiteren können die geschätzten Testtagseffekte, womit die täglichen betrieblichen Umweltschwankungen abgebildet werden, als Hinweise auf das betriebliche Management genutzt werden. Dabei ist vor allem von Interesse, wie hoch die Testtagseffekte sind und wie sehr sie im Zeitverlauf schwanken. Gleichmäßige geringe Testtagseffekte bedeuten dabei eine gleich bleibende betriebliche Umwelt (beispielsweise Futterqualität, Stallklima), und ein stark abweichender Wert kann Störungen aufzeigen, deren Ursache daraufhin gesucht und auf die entsprechend reagiert werden kann.

Auf die Nutzung gemischter linearer Modelle im Rahmen der Zuchtwertschätzung soll entsprechend des Ziels der vorliegenden Arbeit nicht eingegangen werden.

3.2 Exponentially-weighted moving-average control charts

3.2.1 Grundlagen

Ein *Exponentially-weighted moving-average* (EWMA) *control chart* bzw. eine EWMA-Karte ist ein in der Industrie verwendetes Kontrollinstrument zur Überwachung von Produktionsabläufen in Form einer Zeitreihe mit einem „Gedächtnis“. Dabei wird ein gewogener Mittelwert aus einer Reihe von Beobachtungen berechnet, wobei je nach gewünschter Gedächtnisleistung mehr Gewicht auf weit zurückliegende oder aktuelle Beobachtungen gelegt werden kann. Mehrere zeitlich nacheinander liegende Berechnungen kann man dadurch zu einer Mittelwertkurve verbinden. Gleichzeitig werden in der Regel obere und untere Grenzwerte vorgesehen, mit denen der ermittelte Wert verglichen werden muss. Diese Grenzwerte werden üblicherweise durch die Varianz der Beobachtungen bestimmt. Da während des Produktionsprozesses immer neue Beobachtungen dazukommen, können sich auch die Varianz und damit der jeweilige Grenzwert verändern. Solange der EWMA innerhalb der Grenzwerte liegt, verläuft der Prozess wie gewünscht. Werden die Grenzwerte über- oder unterschritten, muss ein Alarm erfolgen. Die Varianz des Prozesses hat das übliche Niveau überschritten, und ein Eingreifen seitens des Personals ist nötig. Im Fall der Nutzung in einem Milchviehbetrieb kann dies z.B. eine Änderung der Milchmenge sein. Denkbar sind ein Rückgang der Milchleistung durch Euter- oder Stoffwechselerkrankungen, aber auch Fehlbuchungen von einzelnen Gemelken können zu Alarmen führen.

3.2.2 Berechnung

Der EWMA ist definiert als

$$EWMA_{Y,t} = \lambda Y_t + (1 - \lambda)EWMA_{t-1} \quad (5)$$

für $t = 1, 2, \dots, n$ (vgl. ROBERTS, 1959; MONTGOMERY, 1997, S. 299ff). Für $t = 1$ wird $EWMA_0$ als Startwert verwendet. Man benötigt dazu einen „historischen“ Mittelwert oder Zielwert μ . Y_t ist die Beobachtung zum Zeitpunkt t , n ist entsprechend die gesamte Anzahl vorliegender Beobachtungen. Der Parameter λ bestimmt die Güte des „Gedächtnisses“, es gilt $0 < \lambda \leq 1$. Ist λ nahe 0 werden aktuelle Beobachtungen nur schwach gewichtet und weiter zurückliegende stark. Ist umgekehrt λ nahe 1, wird großes Gewicht auf aktuelle Beobachtungen gelegt und zurückliegende Daten vernachlässigt. Da ein niedriger Wert für λ zu geringeren Schwankungen bedingt durch aktuelle Beobachtungen führt, aber gerade diese Schwankungen interessant zur Identifikation von Milchleistungsabweichungen sind, empfiehlt sich hier ein hoher Wert für λ .

Die Varianz für $EWMA_t$ ist gegeben durch

$$\sigma_{EWMA_{y,t}}^2 = \sigma_y^2 \left(\frac{\lambda}{2-\lambda} \right) \left[1 - (1-\lambda)^{2t} \right] \quad (6)$$

wodurch sich für den oberen Grenzwert (*upper control limit*, UCL) und den unteren Grenzwert (*lower control limit*, LCL) folgende Gleichung zur Berechnung ergibt (MONTGOMERY, 1997, S. 301, KRIETER et al. 2005):

$$UCL_t = \mu + L\sigma \sqrt{\left(\frac{\lambda}{2-\lambda} \right) \left[1 - (1-\lambda)^{2t} \right]} \quad (7)$$

$$LCL_t = \mu - L\sigma \sqrt{\left(\frac{\lambda}{2-\lambda} \right) \left[1 - (1-\lambda)^{2t} \right]} \quad (8)$$

L ist dabei ein weiterer vorzugebender konstanter Parameter mit $L > 0$ der bestimmt, wie weit die oberen und unteren Grenzwerte vom Zielwert μ entfernt sind. Ein klein gewähltes L sorgt für einen engen „Korridor“ in dem sich der EWMA bewegen darf ohne einen Alarm auszulösen, wodurch aber auch eine Anfälligkeit gegenüber Fehlalarmen entsteht. Fehlalarme in diesem Sinne sind Alarme, die auf normale Schwankungen und nicht auf Erkrankungen zurückzuführen sind. Andererseits kann ein zu groß gewählter Wert für L dazu führen, dass geringere krankheitsbedingte Schwankungen keinen Alarm hervorrufen und übersehen werden können.

3.2.3 Verwendung eines EWMA-Charts im Milchviehbetrieb

Es ist je nach betrachtetem Merkmal denkbar, auf den oberen oder unteren Grenzwert zu verzichten. Bei der Milchleistung sind in erster Linie die Abweichungen außerhalb des LCL von Interesse. Abweichungen nach oben, die durch Fehlzuordnungen eines Gemelks entstehen, sollten bei einem anderen Tier der Herde durch ein fehlendes Gemelk zu einer negativen Leistungsabweichung und einem entsprechenden Alarm führen.

Zur Überwachung der Leitfähigkeit ist vor allem das UCL wichtig, um z.B. Mastitiden anhand erhöhter Werte zu erkennen. Eine Verwendung zur Östrusüberwachung mit Hilfe des Merkmals Aktivität wird von KRIETER et al. (2005) beschrieben.

3.3 Fuzzy Logic

3.3.1 Grundlagen

Das ursprüngliche Anliegen von Fuzzy-Systemen (engl. *fuzzy*: unscharf) bestand darin, von der klassischen, mathematischen Vorgehensweise in der Regeltechnik abweichend eine eher der menschlichen Regeltechnik nachempfundene Strategie zu verfolgen (BORGELT et al., 2003, S. 151). Menschliche Regeltechnik bedeutet dabei die häufige Benutzung qualitativer Angaben anstelle von exakten quantitativen Angaben. Wenn man sich beispielsweise als Autofahrer mit *hoher* Geschwindigkeit auf der Landstrasse einer *scharfen* Linkskurve nähert, sollte man *etwas fester* auf die Bremse treten und das Lenkrad bei Erreichen der Kurve *stark nach links* einschlagen. Man benutzt also im allgemeinen Sprachgebrauch Ausdrücke wie „hoch“ um die Geschwindigkeit zu charakterisieren, anstatt die exakten Werte zu benutzen. Man könnte auch sagen, man erfasst die Situation intuitiv und handelt dementsprechend angemessen. Ebenso kann man an diesem einfachen Beispiel erkennen, dass man oft mit Merkmalsausprägungen konfrontiert ist, die fließend ineinander übergehen. Wo hört z.B. „normale“ Geschwindigkeit auf und fängt „hohe“ Geschwindigkeit an? Hier kann der Fuzzy Logic-Ansatz helfen. Das hierbei verfolgte Grundprinzip besteht darin, einen Vektor von Eingabevariablen mit Hilfe intuitiver Regeln zu klassifizieren und in einem entsprechenden Ausgabektor abzubilden. Das Ergebnis kann dabei zwar wieder exakt sein, wichtiger ist aber die Aussagekraft.

Solche unscharfen Messwerte und Klassifizierungen sind auch typisch für biologische Daten, die immer genetisch- und umweltbedingten Schwankungen unterliegen. Unterschiedliche Leistungsniveaus in einer Milchviehherde lassen eine scharfe Abgrenzung bei festgelegten täglichen Milchmengen zur Klassifizierung wenig sinnvoll erscheinen. Der Fuzzy Logic-Ansatz erlaubt, die biologische Varianz angemessen zu berücksichtigen.

Mittels Fuzzy Logic kann man solche qualitativen oder unscharfen Beschreibungen von Merkmalszuständen in rechnergerechte Form übersetzen und verarbeiten (YANG, 1998, S. 27). Im Gegensatz zur klassischen binären Mengenlehre, bei der ein Element entweder zu einer Menge gehört oder nicht, wird in der unscharfen Fuzzy-Theorie die gleitende Zuordnung eines Elements zu einer Menge durch den Grad der Zugehörigkeit bestimmt. Der im Intervall von 0 bis 1 liegende Zugehörigkeitsgrad ist durch Zugehörigkeitsfunktionen gegeben. Der binäre Ansatz von entweder 0 oder 1 stellt einen Extremfall dar, der ebenfalls durch entsprechende Zugehörigkeitsfunktionen abgebildet werden kann. Jede relevante Merkmalsausprägung wird durch eine separate Zugehörigkeitsfunktion dargestellt. Diese Formalisierung umgangssprachlicher Daten versetzt uns in die Lage, unscharfe

Mengen (Fuzzy-Mengen) zu beschreiben (BORGELT et al., 2003, S. 151f). Im Unterschied zur Wahrscheinlichkeitsrechnung, bei der die Summe aller Einzelwahrscheinlichkeiten 1 ergeben muss, ist man bei Zugehörigkeiten zu Merkmalszuständen nicht an diese Beschränkung gebunden.

Der Aufbau eines Fuzzy-Logic-Systems lässt sich in drei Bereiche unterteilen:

Fuzzifizierung, Fuzzy-Inferenz und Defuzzifizierung. Diese sollen nun im Einzelnen betrachtet werden.

3.3.2 Fuzzifizierung

3.3.2.1 Linguistische Variablen

Bei der Fuzzifizierung werden die als Messwerte vorliegenden Merkmalsausprägungen in Zugehörigkeiten linguistischer Variable (auch: Terme) umgewandelt. Für jedes Merkmal müssen für die entsprechenden linguistischen Variablen Zugehörigkeitsfunktionen angelegt werden. Dabei ergibt sich häufig eine ungerade Anzahl von Termen, um das Konzept der „Mitte“ und entsprechende nach oben und unten abweichende Zustände abzubilden (YANG, 1998, S. 29). Daraus kann man ableiten, dass es in der Regel mindestens drei Abstufungen (Mitte und zwei Extreme) und selten mehr als sieben Abstufungen geben wird – sieben Abstufungen ergeben sich aus der Beschränkung des menschlichen Kurzzeitgedächtnisses auf sieben Symbole, die den sprachlichen Termen entsprechen (ALTROCK, 1997, S. 325). Je nach Fragestellung ist es aber auch möglich und sinnvoll, von diesem Konzept abzuweichen. Dies ist zum Beispiel der Fall, wenn nur ein Extrem von Interesse ist und alle linguistischen Variablen von der Mitte (die häufig den Normalzustand wiedergibt) zum anderen Extrem gleich behandelt werden. In diesem Fall wäre hier keine weitere Abstufung notwendig, und eine einzige Zugehörigkeitsfunktion, die diese Merkmalsausprägungen abdeckt, reicht aus. Eine gute Möglichkeit dies herauszufinden ist das Vorziehen der Regelsektion (vgl. Fuzzy-Inferenz). Durch Formulierung von typischen Fuzzy Logic-Regeln (vgl. Abschnitt 3.3.3) kann man einen Überblick bekommen, welche und wie viele Terme man benötigt, um die Regelbasis abzudecken (ALTROCK, 1997, S. 326).

3.3.2.2 Zugehörigkeitsfunktionen

Der nächste Schritt ist die Festlegung der Art und der Lage der Zugehörigkeitsfunktionen. Üblicherweise werden alle Zugehörigkeitsfunktionen eines Merkmals in einem gemeinsamen Koordinatensystem aufgetragen. Das Koordinatensystem ist dabei auf der x-Achse durch die Minimal- und Maximalwerte begrenzt, die das Merkmal annehmen kann. Auf der y-Achse wird die Zugehörigkeit abgetragen. Die Skala reicht also von 0 bis 1. Gerade bei extremen hohen oder niedrigen Merkmalswerten, die in Bereichen liegen, die entweder Zugehörigkeit 0 oder 1 ergeben, ist es möglich und erscheint es sinnvoll, einen Maximal- und Minimalwert auf der x-Achse festzulegen und den außerhalb liegenden Beobachtungen diese Extremwerte zuzuweisen. Nachdem der Wertebereich festgelegt ist, können die Funktionsformen ausgewählt werden. Die einzige Einschränkung hierbei ist, dass die Funktion vereinfachend Werte von 0 bis 1 annehmen können muss.

Die einfachsten Funktionen (Standard-Zugehörigkeitsfunktionen) benutzen dabei gerade Linien. Als einfaches Beispiel ist hier die Dreiecksfunktion (auch: Lambda-Typ) zu nennen. Diese lässt sich durch drei Punkte definieren: der Punkt, bis zu dem sie eine Zugehörigkeit von 0 aufweist, der Punkt an dem die Zugehörigkeit 1 ist und der Punkt, an dem die Zugehörigkeit wieder 0 ist. Diese Punkte (und die Nullpunkte an beiden Enden des Wertebereichs) werden durch Geraden verbunden, so dass man den Gesamtverlauf der Zugehörigkeitsfunktion erhält. Weitere einfache Funktionsformen sind die Trapezfunktion (auch: Pi-Typ) oder die Funktionen vom Z- oder S-Typ, die jeweils nach einer Seite hin „offen“ sind und deren Zugehörigkeiten nur von 1 bis 0 oder 0 bis 1 verlaufen (vgl. Abbildung 4). Sie können als Verallgemeinerung der Dreiecksfunktionen angesehen werden (BORGELT et al., 2003, S. 155).

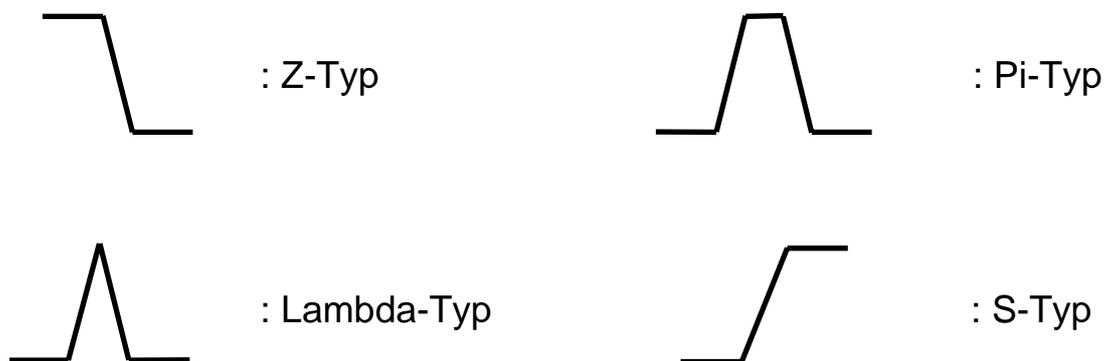


Abbildung 4: Standard-Zugehörigkeitsfunktionen (ALTROCK, 1997, S. 327)

Weitere Möglichkeiten, die Zugehörigkeitsfunktionen zu gestalten, führen zu nichtlinearen Funktionen mit „weichen“ Übergängen. Diese werden weiter unten diskutiert. Exakt an eine Fragestellung angepasste polynomiale Funktionsverläufe könnten eine Genauigkeit suggerieren, die bei einem

prinzipiell zunächst einmal unscharfen Ansatz am Ende zu vergleichbaren Ergebnissen wie ein linearer Funktionsverlauf führt. Andererseits können fortgeschrittene Funktionen eine genauere Abbildung der linguistischen Terme liefern und daher gerade für Modelle zur Entscheidungsunterstützung geeigneter sein (ALTROCK, 1997, S. 328).

Zur Festlegung der Definitionspunkte der Zugehörigkeitsfunktionen gibt es verschiedene Möglichkeiten. Man kann sich an (Grenz)werten aus entsprechender Literatur oder eigener Erfahrung orientieren (Experten-Ansatz). Es kann aber auch sinnvoll sein, sich an Lage- und Streumaße des konkreten Datensatzes zu halten. Da im vorliegenden Fall in den untersuchten Betrieben unterschiedliche Leistungsniveaus vorhanden sind und auch die Varianzen je nach Herde verschieden sind, ist eine betriebspezifische Ausrichtung der Zugehörigkeitsfunktionen ratsam. Allgemein kann man zur Festlegung der Definitionspunkte in vier Schritten vorgehen (ALTROCK, 1997, S. 327f):

1. Für jeden Term wird der Merkmalswert definiert, der der linguistischen Bedeutung am nächsten kommt. Diesen typischen Werten wird jeweils der Zugehörigkeitsgrad $\mu=1$ zugewiesen.
2. Dort, wo die jeweils benachbarten Terme ihren typischsten Wert annehmen, wird der Zugehörigkeitsgrad des betrachteten Terms auf $\mu=0$ gesetzt.
3. Die so definierten Punkte mit Zugehörigkeitsgraden $\mu=0$ und $\mu=1$ werden mit geraden Linien verbunden, woraus für innen liegende Terme Zugehörigkeitsfunktionen vom Lambda-Typ resultieren.
4. Für die an den Enden des Wertebereichs liegenden Terme existieren keine weiter außen liegenden benachbarten Terme, so dass alle weiter außen liegenden Werte ebenfalls eine Zugehörigkeit von 1 zu diesen Extremtermen aufweisen. Daraus ergibt sich für den links liegenden Term eine Zugehörigkeitsfunktion vom Z-Typ und für den rechts liegenden Term eine vom S-Typ.

Zu Zugehörigkeitsfunktionen vom Pi-Typ kommt man dann, wenn man anstelle von typischen Punkten ein Intervall annimmt.

Alle Formen von Zugehörigkeitsfunktionen haben gemein, dass sie die vordefinierten Punkte mit den Zugehörigkeitsgraden $\mu=0$ bzw. $\mu=1$ miteinander verbinden. Weitere Möglichkeiten für Formen von Zugehörigkeitsfunktionen sind z.B. Polynome, analytische Funktionen, stochastische Funktionen wie die Gaußsche Glockenkurve, stückweise lineare Funktionen oder empirisch abgeleitete Funktionen für die verschiedenen Terme. Die empirischen Ansätze stehen allerdings hinsichtlich

des Nutzens nach ALTROCK (1997, S. 331f) in keinem Verhältnis zum notwendigen Aufwand. Abgesehen von den polynomischen Funktionen spielen die anderen Ansätze in der praktischen Anwendung keine Rolle.

Weiterhin gilt das genannte fast ausschließlich für Zugehörigkeitsfunktionen für Inputvariablen. Für die Outputvariablen werden in den meisten Anwendungen nur die Standard-Zugehörigkeitsfunktionen vom Lambda-Typ verwendet (ALTROCK, 1997, S. 332).

3.3.2.3 Vorgehensweise bei der Fuzzifizierung

Die eigentliche Fuzzifizierung besteht darin, den Eingangsvariablen je nach Merkmalswert entsprechende Zugehörigkeiten zu der jeweiligen Fuzzy-Menge zuzuweisen. Dies geschieht, indem man in die vorher festgelegten Zugehörigkeitsfunktionen für den Teilabschnitt, in dem der Merkmalswert liegt, den Merkmalswert einsetzt. Dabei wird für jede Zugehörigkeitsfunktion eine Zugehörigkeit von 0 bis 1 berechnet. Der Unterschied zu Wahrscheinlichkeiten zeigt sich hier besonders, da die Summe der Zugehörigkeiten zu den einzelnen Zugehörigkeitsfunktionen für einen Merkmalswert nicht gleich 1 sein muss. Beispielsweise könnten einem Wert von 5.7 mS/cm für die elektrische Leitfähigkeit der Milch Zugehörigkeiten von 0.95 zur Funktion „normal“, von 0.25 zur Funktion „leicht erhöht“ und von 0.01 zur Funktion „stark erhöht“ zugeordnet werden. Gleichzeitig besteht noch die Zugehörigkeit von 0 zu der Funktion „niedrig“.

3.3.3 Inferenz

Die Fuzzy-Inferenz bezeichnet den Abschnitt des regelbasierten Schließens auf Basis unscharfer Aussagen (KOCH, 1996, S. 106). Um Nutzen aus der Darstellung von Fuzzy-Mengen zu ziehen ist es erforderlich, diese mit Regeln zu verknüpfen. Diese Regeln können ebenfalls umgangssprachlich formuliert werden und ergeben sich entweder intuitiv aus Erfahrung oder Expertenwissen. Die Aufgabe von Fuzzy-Inferenz besteht darin, eine Regelmenge zu bilden (YANG, 1998, S. 36). Mit Hilfe dieser Regelmenge werden die Fuzzy-Mengen der Eingangsvariablen zu Fuzzy-Mengen für die Ausgabevariable(n) weiterverarbeitet. Allgemein lassen sich die Regeln folgendermaßen modellieren:

WENN <Bedingung(en)>, **DANN** <Schlussfolgerung(en)>,

wobei gilt:

<Bedingung(en)> = <Eingangsvariable> **IST** <Linguistischer Term> und
 <Schlussfolgerung(en)> = <Ausgangsvariable> **IST** <Linguistischer Term>.

Die Bedingungen bzw. der Wenn-Teil wird auch als Antezedenz bezeichnet, die Schlussfolgerung bzw. der Dann-Teil entsprechend als Konsequenz (BORGELT et al., 2003, S. 156f).

Entscheidend hierbei ist die Möglichkeit zur Kombination mehrerer Bedingungen durch mit der Booleschen Logik vergleichbare Operatoren wie UND, ODER oder NICHT. Entsprechende Fuzzy Logic-Operatoren sind Minimum, Maximum und $1-\mu$ (ALTROCK 1997, S. 41). Nachdem die Zugehörigkeitsfunktionen also den unscharfen Teil von Fuzzy Logic abdecken, kommt bei den Regelmechanismen der Inferenz der Logik-Teil zum Tragen. Dabei ist der Logik-Teil als eine Verallgemeinerung der Booleschen Logik anzusehen.

A	B	A und B	A	B	A oder B	A	nicht A
1	1	1	1	1	1	1	0
1	0	0	1	0	1	0	1
0	1	0	0	1	1		
0	0	0	0	0	0		
AND			OR			NOT	

A	B	min(A, B)	A	B	max(A, B)	A	1-A
1	1	1	1	1	1	1	0
1	0	0	1	0	1	0	1
0	1	0	0	1	1		
0	0	0	0	0	0		
AND			OR			NOT	

Abbildung 5: Entsprechungen von Operatoren der booleschen Logik (oben) und der Fuzzy Logic (unten), erweitert nach BORGELT et al. (2003, S. 157)

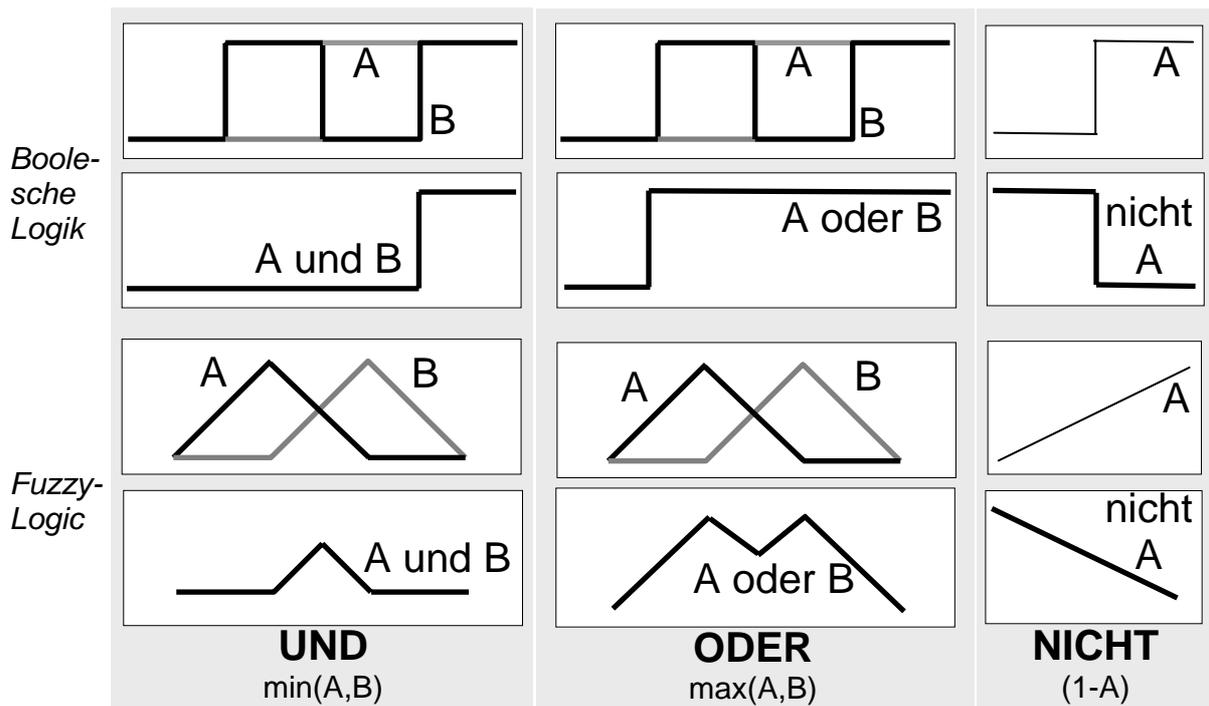


Abbildung 6: Grafische Darstellung boolescher Logik im Vergleich zu Fuzzy Logic

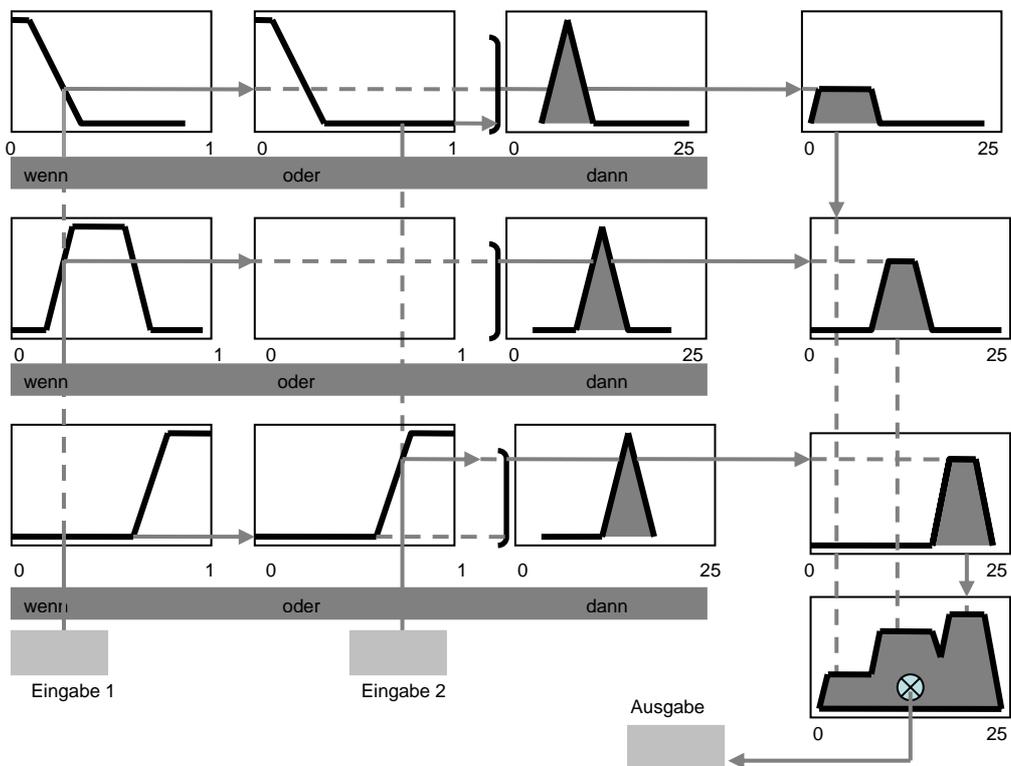


Abbildung 7: Aggregationsschema (verändert nach MATHWORKS, 2000, S. 2-27)

Wenn man als Zugehörigkeiten nur die Extremwerte 0 und 1 zulässt, gelten die bekannten Booleschen Regeln wie in Abbildung 5 dargestellt. Ein einfaches Beispiel für eine Regel, die bei der Mastitiserkennung Verwendung finden könnte und bei der mehrere Merkmale kombiniert abgefragt werden, lautet in umgangssprachlicher Form wie folgt:

Wenn die elektrische Leitfähigkeit der Milch *stark erhöht*
und die Milchmenge gegenüber dem Vortag *verringert* ist,
dann ist die Wahrscheinlichkeit für eine Mastitiserkrankung der Kuh *hoch*.

Die UND-Verknüpfung bedeutet hier, dass der Minimum-Operator bei der Aggregation verwendet wird. Von den Fuzzy-Mengen je Merkmalswert werden also jeweils diejenigen ausgewählt, die das Minimum an Zugehörigkeiten aufweisen (siehe Abbildung 6). Übertragen auf die Mengenlehre entsprechen der Minimum-Operator der Schnittmenge und der Maximum-Operator der Vereinigungsmenge. Diese Mengen geben an, zu welchem Grad eine Regel eine Output-Fuzzy-Menge unterstützt. Die Output-Sets für alle Regeln werden dann zu einem einzelnen Output-Fuzzy-Set aggregiert. Regeln können dabei auch unterschiedlich stark gewichtet werden. Das Resultat der Aggregation ist schließlich eine Fläche, die von den durch die Regeln eingeschränkten Zugehörigkeitsfunktionen der Terme der Outputvariable(n) begrenzt wird (Abbildung 7).

Die maximale Anzahl der nötigen Regeln ergibt sich, wenn man die Anzahl der Terme für die Eingangsvariablen miteinander multipliziert. Daraus lässt sich ersehen, dass bei einer größeren Anzahl von gleichzeitig betrachteten Merkmalen und Termen das Regelgerüst schnell unübersichtlich werden kann. Die Regeln für zwei Merkmale lassen sich noch problemlos in einer Tabelle darstellen, bei der in den Tabellenfeldern die entsprechende Output-Fuzzy-Menge angegeben wird, die bei Kombination der Spalten- und Zeilterme das Ergebnis ist. Bei drei oder mehr Merkmalen müsste man aber schon mehrdimensionale Darstellungen erstellen. Dies kann man vermeiden, indem man sinnvoll zusammenfassbare Merkmale in separaten Fuzzy-Systemen berechnet und dann die dort erhaltenen Outputs in einem nachgelagerten Fuzzy-System als Inputvariablen verwendet. Man erhält dadurch ein kaskadiertes System, das zur Vereinfachung und Übersichtlichkeit beiträgt. Als Beispiel für eine Anwendung kann hier ein Fuzzy-System genannt werden, dass verschiedene Merkmale zur elektrischen Leitfähigkeit der Milch wie generelles Niveau, kurzfristige Tendenz und längerfristige Tendenz bereits vorverarbeitet und als Output einen Wert für den Komplex „Leitfähigkeit“ liefert.

3.3.4 Defuzzifizierung

Das aggregierte Output-Set wird bei der Defuzzifizierung wieder auf eine einzelne, scharfe Zahl reduziert. Mathematisch entspricht das der Abbildung eines Vektors (dem Wert der unscharfen linguistischen Variable) auf eine reelle Zahl (ein „harter“ bzw. „scharfer“ Wert) (ALTROCK, 1997). Die übliche Methode dabei ist der *center-of-area*-Ansatz (CoA, auch: *center-of-gravity*, CoG), bei der der Schwerpunkt der Aggregationsfläche berechnet wird (vgl. Abbildung 7). Die Koordinate dieses Punktes auf der x-Achse ist dabei der zurückgegebene Wert, das gewünschte Ergebnis der entsprechenden Kombination von Inputmerkmalen. Andere Möglichkeiten zur Defuzzifizierung sind die *center-of-maximum*-Methode (CoM), bei der die jeweils typischen Merkmalsausprägungen der einzelnen Terme gewichtet werden, oder die *mean-of-maximum*-Methode (MoM, mit den Varianten *left-of-maximum* LoM und *right-of-maximum* RoM), die sich nur an dem Term mit der höchsten Zugehörigkeit orientiert (vgl. Abb. 10). Tabelle 6 bietet eine Vergleichsübersicht der drei Methoden.

Bei Entscheidungsunterstützungssystemen hängt die Wahl der Defuzzifizierungsmethode auch vom Zusammenhang der Entscheidung ab. Bei quantitativen Entscheidungen ist die CoM-Methode besser geeignet, bei qualitativen Entscheidungen bietet sich die MoM-Methode an. Diese gilt auch für Mustererkennung als am Besten geeignet. Die CoA- bzw. CoG-Methode ist historisch bedingt weit verbreitet. Die in Tabelle 6 angeführten Probleme können durch Änderungen bei Zugehörigkeitsfunktionen oder Regeln ausgeglichen werden.

Tabelle 6: Vergleich verschiedener Methoden zur Defuzzifizierung (ALTROCK, 1997, S. 363)

	CoA, CoG	CoM	MoM
Linguistische Charakteristik	„Bester Kompromiss“	„Bester Kompromiss“	„Plausibelste Lösung“
Übereinstimmung mit intuitiver Herangehensweise	Unplausibel bei verschiedenen Formen und starker Überlappung von Zugehörigkeitsfunktionen	Gut	Gut
Kontinuität	Ja	Ja	Nein
Rechnerische Effizienz	Sehr niedrig	Hoch	Sehr Hoch
Anwendungen	Steuerung, Entscheidungsunterstützung, Datenanalyse	Steuerung, Entscheidungsunterstützung, Datenanalyse	Mustererkennung, Entscheidungsunterstützung, Datenanalyse

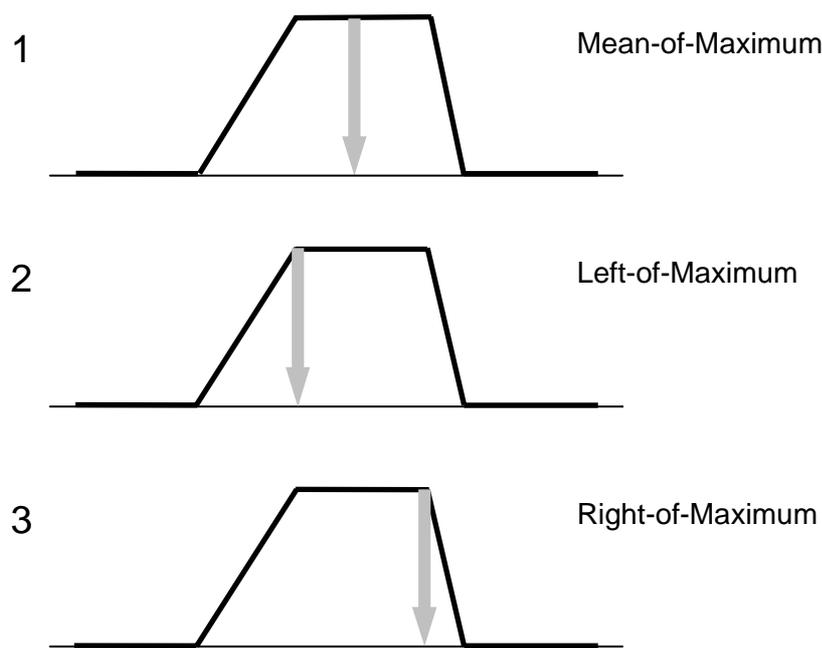


Abbildung 8: Graphische Darstellung der Defuzzifizierung mit der MoM-, LoM- und RoM-Methode (ALTROCK, 1997, S. 362)

3.3.5 Bedeutung

Der Fuzzy Logic-Ansatz zeichnet sich dadurch aus, dass er leicht und intuitiv zu interpretieren ist, und zudem recht einfach durch Änderungen bei den Zugehörigkeitsfunktionen und im Regelgerüst an spezielle Bedürfnisse und Vorgaben eines Betriebs angepasst werden kann. Dazu ist auch kein Experte für statistische Modellierung nötig, sondern Experten auf dem jeweiligen Fachgebiet, z.B. Mastitiserkennung (Betriebsleiter, Betreuungspersonal, Veterinär etc.), können solche Modifikationen direkt vor Ort vornehmen (DE MOL, 2000, S. 142). Bei der Anwendung wird in erster Linie darauf hinzuwirken sein, Fehlerraten durch die Kombination mehrerer Merkmale zu verringern, während gleichzeitig Erkennungsrate und Spezifität möglichst hohe Werte erreichen.

Die Bedeutung von Fuzzy-Systemen ergibt sich aus der Betrachtung von Merkmalskombination und somit einerseits besserer Klassifizierung von Daten, was in einem Entscheidungsunterstützungssystem zu einer geringeren Zahl von Fehlalarmen führen sollte, andererseits aber auch Erkennung von mit einzelnen Merkmalen nicht erfassbaren Fällen.

Zur Anwendung in der Landwirtschaft, speziell im Milchviehbereich, wurden bereits Fuzzy-Logic-Modelle zur Östruserkennung (YANG, 1998, S. 25ff; DE MOL und WOLDT, 2001; FIRK et al., 2003a; FIRK et al., 2003b; PACHE et al., 2005) und zur Mastitiserkennung (DE MOL und WOLDT, 2001; KÖHLER, 2003, S. 24ff; CAVERO et al., 2006a) untersucht. Bei all diesen Anwendungen werden je-

weils verschiedene Merkmale zur Entscheidungsfindung herangezogen. Zur Östruserkennung ist das Hauptmerkmal jeweils die Bewegungsaktivität, während als weitere Informationskriterien Liegeverhalten (*PACHE* et al., 2005), vergangene Tage seit der letzten Brunst (*FIRK* et al., 2003b; *DE MOL* und *WOLDT*, 2001) bzw. allgemeiner Status der Kuh (abgekalbt, Brunst, besamt, tragend), Anteil der Tiere mit erhöhter Aktivität in der Herde (zur Verringerung von Fehlalarmen bei allgemein erhöhtem Aktivitätsniveau der Herde durch andere Ereignisse), Abweichungen der Temperatur und Milchmenge (jeweils *DE MOL* und *WOLDT*, 2001) sowie Milchleistung und Kraftfutteraufnahme (*YANG*, 1998, S. 56ff) untersucht wurden. Für die Erkennung von Mastitis wurde als Hauptmerkmal die elektrische Leitfähigkeit (EL) verwendet. *DE MOL* und *WOLDT* (2001) nutzten sowohl die standardisierten viertelweisen Abweichungen der EL als auch die viertelweisen absoluten Messwerte für die EL als Zusatzinformation. *KÖHLER* (2003, S. 47ff) nutzte neben der EL noch Milchbildungsrate, Zwischenmelkzeit und Milchfluss, wobei hier alle Merkmale viertelweise ausgewertet wurden. *CAVERO* et al. (2006a) untersuchten ebenfalls ein Fuzzy-Modell mit den Merkmalen EL (sowohl Abweichung als auch Maximalwert), Abweichung der Milchflussrate und Abweichung der Milchbildungsrate.

3.4 Künstliche neuronale Netze

3.4.1 Grundlagen

Ein künstliches neuronales Netz (KNN) ist ein Daten verarbeitendes System, dessen Struktur und Funktionsweise dem Nervensystem und speziell dem Gehirn von Tieren und Menschen nachempfunden ist (BORGELT et al., 2003, S. 3). Dabei folgt ein KNN dem aus der Informatik bekannten EVA-Prinzip (Eingabe \rightarrow Verarbeitung \rightarrow Ausgabe). Die Eingabe gestaltet sich als Präsentation eines Musters, üblicherweise ein Vektor von Merkmalswerten, aus denen Informationen gewonnen werden sollen (HEUER, 1997, S. 14). Diese werden je nach Anzahl der Zielwerte als entsprechende Ausgabemuster in Skalar- oder Vektorform erzeugt. Abbildung 9 zeigt diesen grundlegenden Aufbau, wobei als Ausgabe ein Skalar vorgesehen ist. Neuronen sind für die Verarbeitung von Daten dabei die wichtigsten Bestandteile des biologischen Vorbilds, des Nervensystems. Daher bietet es sich an, kurz den Aufbau und die Funktionsweise eines solchen Neurons darzustellen.

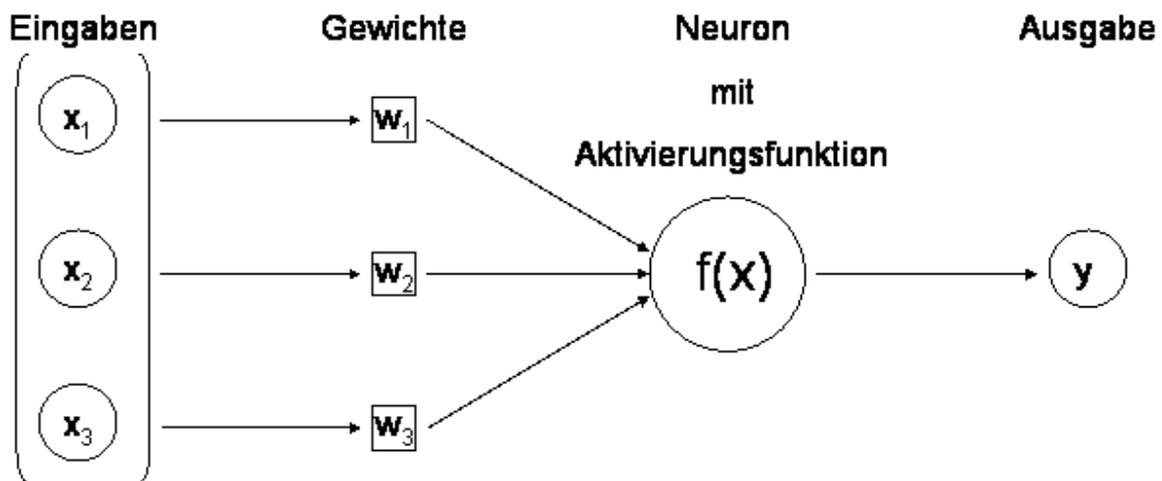


Abbildung 9: Schematische Darstellung eines einfachen künstlichen neuronalen Netzes (verändert nach PATTERSON, 1997, S. 15)

3.4.2 Biologische Grundlagen

Ein Neuron ist durch drei typische Strukturen gekennzeichnet: Dendritenbaum, Zellkörper und Axon (RITTER et al., 1994, S. 17ff). Der Dendritenbaum ist eine verästelte Struktur von Zellfortsätzen um das Neuron herum, die Reize von anderen umgebenden Neuronen in Form eines elektrischen Potentials aufnimmt und an den Zellkörper (Soma) weiterleitet. Wird dabei ein Schwellenwert überschritten, sendet der Zellkörper seinerseits einen elektrischen Impuls (Spike) das Axon entlang. Das Axon ist eine wenige Millimeterbruchteile bis zu mehreren Metern lange Nervenfasern, die ebenfalls verästelt ist und so den Impuls mehreren Zielneuronen zuführen kann. An den Endpunkten des Axons befinden sich Synapsen, d.h. Kontaktstellen zu Dendriten oder direkt zum Zellkörper der Zielneuronen (vgl. Abbildung 10). Die Synapsen schütten bei Aktivierung durch den elektrischen Impuls chemische Botenstoffe (Neurotransmitter) aus, die ihrerseits an den Zielstrukturen eine Änderung des elektrischen Potentials auslösen. Je nach Synapse wird das Potential erhöht oder erniedrigt, man spricht hierbei von erregenden bzw. exzitatorischen und hemmenden bzw. inhibitorischen Synapsen.

Trotz dieser elektrochemischen Übertragung von Reizen, die langsamer als die elektrische Übertragung vonstatten geht, werden durch die über das Netz der Neuronen verteilten parallelen Berechnungen hohe Verarbeitungsgeschwindigkeiten erreicht.

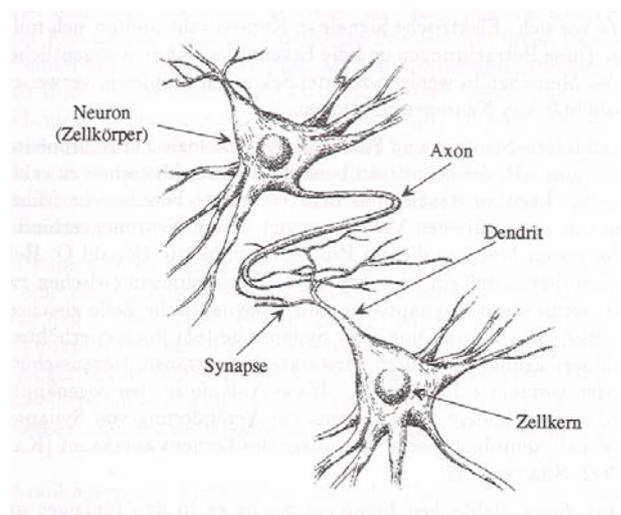


Abbildung 10: Schematische Darstellung zweier Neuronen (NAUCK et al., 1996, S. 13)

Der Lernprozess innerhalb des Gehirns beruht wahrscheinlich auf metabolischem Wachstum durch erhöhte Zellaktivität, wodurch die von einer Synapse erzeugte Ladung variieren kann (PATTERSON, 1997, S. 20).

Auf ein künstliches neuronales Netz übertragen entspricht der Dendritenbaum den Eingaben, der Zellkörper dem Berechnungsteil und das Axon der Ausgabe. Die von den Synapsen erzeugten

variieren Ladungen entsprechen den in KNNs vorkommenden Gewichtungen. Dabei stellen negative Gewichtungen inhibitorische Synapsen und positive Gewichte exzitatorische Synapsen dar. Das Überschreiten des Schwellenwertes zur Aktivierung des Neurons entspricht der Aktivierungs- bzw. Transferfunktion, die einem Neuron zugeordnet ist.

3.4.3 Funktionsweise und Architektur

Neuronen sind die Verarbeitungseinheiten eines KNNs. Die Struktur des KNNs kann grafisch dargestellt werden mit Knoten (als Verarbeitungseinheiten/Neuronen) und Kanten, die die gewichteten Kommunikationsverbindungen repräsentieren (NAUCK et al., 1996, S. 20). Im Prinzip kann jedes Neuron mit jedem anderen und sich selbst durch eine Kante verbunden sein. Wie oben schon erwähnt, führen positive Gewichte zu erregend wirkenden Verbindungen und negative Gewichte zu hemmend wirkenden Verbindungen. Ist das Gewicht für eine Kante auf Null gesetzt, gelten die beiden Neuronen als nicht verbunden. Die Gewichte sind also ausschlaggebend für die Struktur des Netzes. Üblicherweise fasst man die Neuronen zu Schichten zusammen. Diese Schichten umfassen Eingabe- und Ausgabeschicht sowie zumindest eine, je nach Komplexität des Netzes aber auch mehrere verdeckte Verarbeitungsschichten. Diese haben keine Verbindung zur Umwelt (Abbildung 11, vgl. auch KRIETER et al., 2005).

Wenn man die Gewichte so bestimmt, dass sich das Signal nur von Eingabeknoten zu Ausgabeknoten vorwärts fortpflanzen kann, spricht man von einem *Feedforward*-Netzwerk (PATTERSON, 1997, S. 33). In diesem Fall sind keine Verbindungen zu Neuronen in derselben Schicht oder zu vorangegangenen verarbeitenden Schichten vorhanden. Die Gewichte stellen das gespeicherte, verteilte Wissen des Netzwerks dar. Nach der Netzwerkarchitektur nennt man diese KNNs *multilayer feedforward networks*, also mehrschichtige vorwärts betriebene Netzwerke oder auch *multilayer perceptrons* (MLPs) (ZHOU, 2003). Ein Perceptron ist ein einfaches neuronales Netz, das nur aus einer Verarbeitungseinheit besteht (NAUCK et al., 1996, S. 39ff).

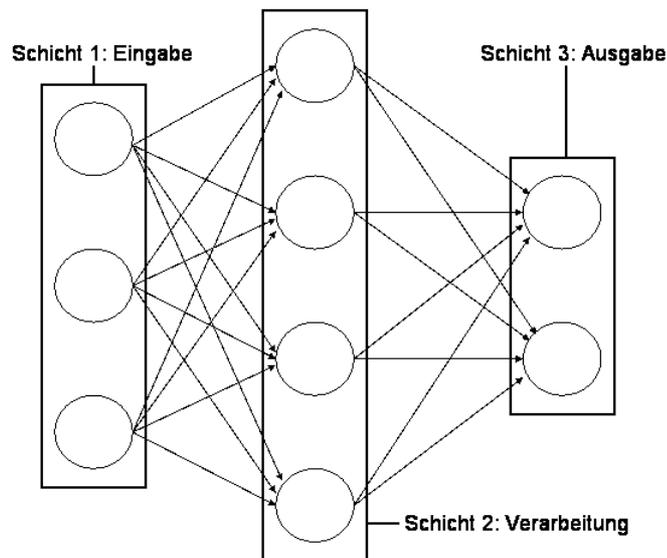


Abbildung 11: Anordnung von Neuronen in Schichten

Die nach der Netzwerkarchitektur unterschiedene zweite wichtige Kategorie von KNNs sind rekursive neuronale Netzwerke (auch *recurrent neural networks*, RNNs), bei denen Rückkopplungsschleifen vorhanden sind. Es existieren also Verbindungen zu anderen Neuronen in derselben oder vorangegangenen Schichten – weshalb man eigentlich auch nicht mehr von streng getrennten Schichten sprechen kann, sondern eher von miteinander verknüpften Verarbeitungseinheiten. RNNs sind durch das mit Hilfe der Feedback-Schleifen erreichte dynamische Verhalten auch für komplexe Berechnungen geeignet, wie z.B. das Erlernen von temporären oder kontextabhängigen Mustern (PATTERSON, 1997, S. 265ff).

Jedem Neuron ist eine Aktivierungsfunktion zugeordnet, die in Abhängigkeit einer Eingabe den Aktivierungszustand des Neurons bestimmt. Durch diese Funktion wird die Eingabe in den Aktivierungswert transformiert. Die Funktion kann unterschiedliche Formen annehmen (z.B. linear, Stufenfunktionen, sigmoid, exponentiell, vgl. auch Abbildung 12). Ebenso kann der Aktivierungswert positiv, negativ, reell, binär oder bipolar sein (PATTERSON, 1997, S. 14f). Bei mehrschichtigen Modellen werden in der Regel sigmoide Aktivierungsfunktionen verwendet, die kontinuierliche Aktivierungszustände aus $[0, 1]$ oder $[-1, 1]$ ergeben (NAUCK et al., 1996, S. 22). Diese nichtlinearen Transformationsfunktionen versetzen ein KNN bei entsprechendem Design und Training des Netzwerks in die Lage, beliebige nichtlineare Funktionen approximieren zu können (ZHOU, 2003).

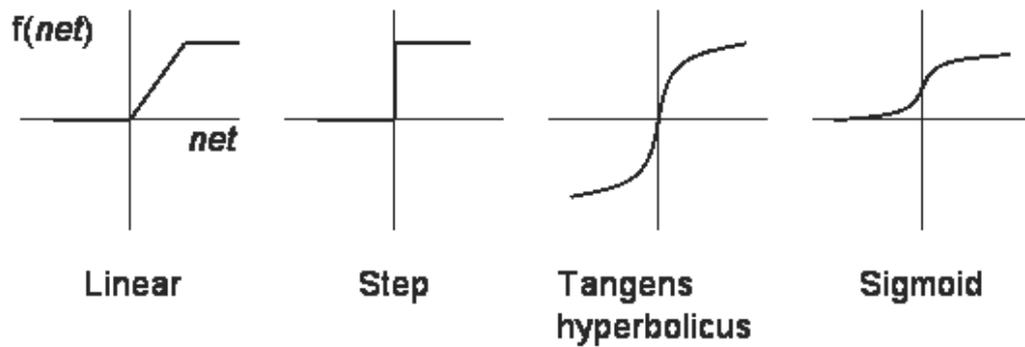


Abbildung 12: Typische Aktivierungsfunktionen neuronaler Netze (PATTERSON, 1997, S. 16)

3.4.4 Training und Lernen

Das Training des Netzwerkes beruht hauptsächlich auf zwei Lernstrategien: überwachtem und unüberwachtem Lernen. Als Lernen bezeichnet man den Anpassungsprozess, der innerhalb des KNNs als Reaktion auf die Eingabesignale abläuft. Begrifflich dazu abgegrenzt bezeichnet Training die wiederholte Durchführung des Lernens durch mehrmalige Präsentation von Eingabemustern (HEUER, 1997, S. 16f). Beim überwachten Lernen existieren vorgegebene Zielwerte, mit denen die im KNN erzeugten Ausgabewerte verglichen werden.

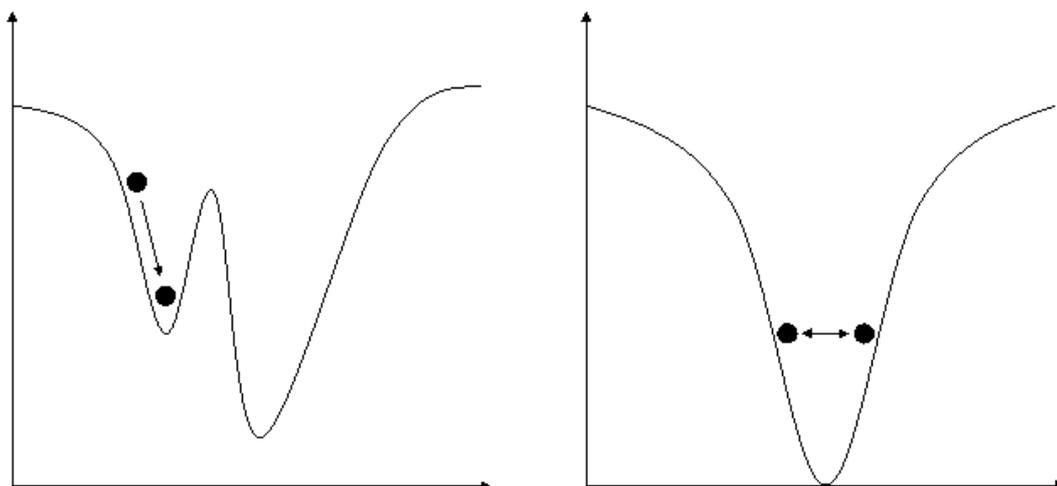


Abbildung 13: Probleme bei ungünstiger Lernrate: lokale Minima (links) und oszillieren um ein Minimum herum (rechts) (nach NAUCK et al., 1996, S. 79)

Nach dem Vergleich von Ausgabe- und Zielwert kann ein Annäherungsfehler berechnet werden, und die Netzwerkparameter (also die Gewichte für die Verbindungen zwischen den Neuronen) werden angepasst, um diesen Fehler zu minimieren. Dies wird in Iterationen weitergeführt, bis ein vor-

gegebenes akzeptables Fehlerniveau erreicht wird. Dabei wird eine differenzierbare Aktivierungsfunktion benötigt, um die Gewichtungen der Neuronen basierend auf der Steigung der Fehlerfunktion in Abhängigkeit der Gewichtung und des Bias schrittweise anzupassen. Dieses Verfahren der Parameterermittlung in Richtung eines Fehlerminimums durch Feststellung der Steigung bzw. des Gradienten an einem Punkt der Fehlerfunktion bezeichnet man als Gradientenabstieg (BORGELT et al., 2003, S. 55ff). Das Ausmaß der Gewichtsänderung je Durchgang wird als Lernrate bezeichnet (z.B. ZHOU, 2003). Dieser Parameter ist bei Erstellung eines KNNs vorzugeben. Idealerweise endet man bei einem absolut minimierten Fehler. Es ist aber auch möglich, bei zu groß gewählter Schrittweite nie ein Minimum zu erreichen, sondern darum herum zu oszillieren und somit keine Lösung zu finden. Ein weiteres Problem kann auftreten, wenn die Gewichte ungünstig initialisiert wurden und man bei der Fehlerminimierung (und möglicherweise zu klein gewählter Schrittweite) auf ein lokales Minimum ungleich dem globalen Minimum stößt, woraufhin der Lernalgorithmus stoppt (Abbildung 13). In beiden beschriebenen Fällen ist der Lernvorgang nicht erfolgreich gewesen (vgl. PATTERSON, 1997, S. 45f; NAUCK et al., 1996, S. 78ff).

Durch den Lernalgorithmus wird eine Generalisierung des Modells angestrebt, um auf unbekannte Eingaben, die ähnlich einer der Eingaben im Lernprozess sind, ähnlich zu reagieren. Es sollen Muster erkannt werden und Eingaben klassifiziert werden können. Die Fähigkeit der Generalisierung ist daher vor allem bei Eingabe gestörter oder unvollständiger Muster von Vorteil. Neuronale Netze bieten in diesem Sinne eine gewisse Fehlertoleranz (NAUCK et al., 1996, S. 30).

Eine weitere Möglichkeit ein KNN zu trainieren ist das konkurrierende Lernen. Dabei haben alle Neuronen einer Schicht zu Beginn kleine, aber nicht gleiche Gewichtungen (PATTERSON, 1997, S. 44). Wenn dem Netzwerk nun ein Eingabevektor präsentiert wird, reagiert eines der Neuronen am stärksten und wird dementsprechend stark in Richtung des Eingabemusters verändert. Je nach Modell können benachbarte Neuronen geringfügiger in dieser Richtung verstärkt werden oder sogar Gewichtungen von nicht durch die Eingabe aktivierten Neuronen reduziert werden. Auf dieses Prinzip des unüberwachten, konkurrierenden bzw. Wettbewerbslernens wird im Abschnitt über *Self-Organizing Maps* (SOMs, selbstorganisierende Karten) näher eingegangen. Entscheidend ist, dass beim unüberwachten Lernen nur Eingabemuster vorliegen, während bei überwachtem Lernen für jedes Eingabemuster auch entsprechend gewünschte Ausgabemuster vorhanden sind. Dabei ist es dem Netz nur möglich, abhängig von den im Training präsentierten Mustern zufällige Klassifikationen (im Sinne von nicht vorgegebenen Klassen) vornehmen zu können (HEUER, 1997, S. 18).

Zusammenfassend sind neuronale Netze bei überwachtem Lernen verwendbar, wenn Trainingsdaten (sowohl Eingabe- als auch Ausgabevektoren) vorliegen. Mathematische Modelle des Problems oder Vorwissen über die Lösung sind nicht notwendig. Allerdings ist die Lösung durch ein neuro-

nales Netz auch nicht interpretierbar. Vielmehr handelt es sich um ein *"black box"*-System, dessen Struktur sich nicht in Form von verständlichen Regeln ausdrücken lässt (BORGELT et al., 2003, S. 152, S. 185). Wenn man keine vorliegenden Trainingsdaten hat, sind neuronale Netze allerdings in der Lage, die Eingabedaten zu klassifizieren, d.h. auf ähnliche Eingaben ähnlich zu reagieren (vgl. unüberwachtes Lernen).

Ein weiterer nennenswerter Aspekt ist die Wiederholbarkeit des Trainings. Wenn man dem zu trainierenden Netzwerk die Trainingsdatensätze präsentiert, wird die Reihenfolge üblicherweise randomisiert. Dadurch reagieren die Neuronen bei unterschiedlichen Reihenfolgen konsequenterweise ebenfalls unterschiedlich. Gewichtungen werden anders angepasst, und der Lernprozess endet möglicherweise bei anderen Fehlerminima. Ein identisches Ergebnis bei Verwendung eines Trainingsdatensatzes kann demnach auch bei Verwendung desselben KNNs nur erreicht werden, wenn sowohl die Ausgangsgewichtungen als auch die Reihenfolge der Datensätze identisch sind. Umgekehrt bedeutet dies aber auch, dass man einen erneuten Trainingsdurchgang mit neu initialisierten Gewichten vornehmen kann, wenn ein Training nicht bei einem minimalen Fehler endete. Dadurch ist es mitunter möglich, bessere Ergebnisse zu erzielen.

3.4.5 Planung und Entwurf

Die Anzahl der Knoten (bzw. Neuronen) in der Ein- und Ausgabeschicht ist prinzipiell durch die vorhandenen Eingaben und die benötigten Ausgaben vorgegeben. In den inneren, versteckten Schichten kann die Anzahl der Knoten allerdings zunächst beliebig gewählt werden – aber es muss der Einfluss der Anzahl von Knoten auf das Gesamtmodell berücksichtigt werden. Je mehr Knoten man vorsieht, desto besser kann die Fähigkeit zur Anpassung des KNNs an die Eingabemuster werden. Andererseits wird dadurch aber auch die Anzahl der zu trainierenden Gewichtungen und damit die generelle Komplexität des Modells erhöht. Zu viele Knoten führen zum Problem des *„overfitting“*: das Netzwerk wird unzuverlässig und unberechenbar, die angestrebte Generalisierungsfähigkeit geht verloren. Das Netzwerk wird dabei zu speziell auf den Trainingsdatensatz abgestimmt. Als Resultat sind zwar bei der Evaluierung für den Trainingsdatensatz gute bis sehr gute Ergebnisse zu erwarten, die sich aber bei Anwendung auf einen unbekanntem Testdatensatz nicht erzielen lassen. Andererseits führt eine zu geringe Anzahl von Knoten zu verminderter Lernfähigkeit und entsprechend schlechter Annäherung der Gewichtungen an die zu modellierenden Sachverhalte, was zum Problem des *„underfitting“* führt. Diese Probleme bei der Modellierung sind in Abbildung 14 dargestellt (ZHOU, 2003). Für die genaue Bestimmung der notwendigen Anzahl von Knoten gibt es demnach keine klare Richtlinie. Ein Ansatz besteht darin, bei einem gegebenen Netzwerk mit ge-

wünschter Anpassungsfähigkeit nicht benötigte Neuronen zu identifizieren und zu entfernen. Wenn ein Neuron trotz Veränderung seiner Gewichte keinen Einfluss auf die Ausgabe des Gesamtnetzes hat, kann es ohne Beeinträchtigung des Netzwerks entfernt werden. Dazu führt man in der Regel eine Sensitivitätsanalyse durch.

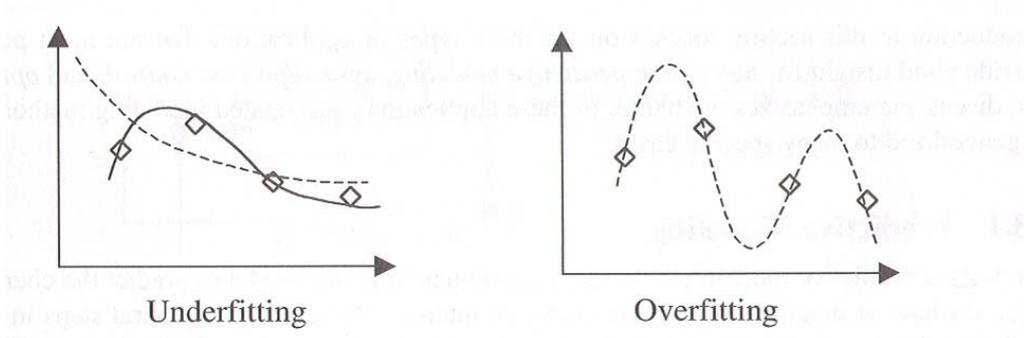


Abbildung 14: Die Probleme des Underfitting und Overfitting (ZHOU, 2003)

Ein anderer Ansatz ist das Erstellen verschiedener Netzwerke mit unterschiedlicher Anzahl von Schichten und Neuronen, die alle mit demselben Datensatz trainiert werden. Dabei werden die jeweiligen Anpassungsfehlerwerte festgehalten. Zur Bestimmung des Generalisierungsfehlers werden die Netzwerke anschließend mit anderen Datensätzen getestet und eine Kreuzvalidierung vorgenommen. Je nach wiederum subjektiver Gewichtung von Anpassungs- und Generalisierungsfehler kann dann das bestgeeignete Netzwerkmodell ausgewählt werden.

3.4.6 Auswertung

Für das Training der künstlichen neuronalen Netze muss der Datensatz in Trainings- und Validierungsdaten unterteilt werden. In der vorliegenden Arbeit wurde das Verfahren der einfachen Kreuzvalidierung verwendet. Dazu werden die jeweils N verfügbaren Datensätze zufällig in $k = 8$ Teildatensätze aufgeteilt. Jeweils $k - 1$ Teildatensätze werden dann zum Training eines KNNes verwendet und mit dem entsprechend trainierten KNN ein Durchlauf mit dem verbliebenen k -ten Teildatensatz simuliert. Um die Gesamtwerte für Erkennungs- und Fehlerrate sowie Spezifität jeder Variante (vgl. Abschnitt 3.5) zu erhalten, werden die entsprechenden Kriterien über alle acht Einzeldurchläufe gemittelt.

3.4.7 Verwendungsmöglichkeiten für künstliche neuronale Netze im Milchviehbereich

3.4.7.1 Mehrschichtige Feedforward-Netzwerke

In erster Linie soll in der vorliegenden Arbeit die Eignung mehrschichtiger Feedforward-Netzwerke (wie z.B. in Abbildung 15 schematisch dargestellt) zur Erkennung von Krankheiten überprüft werden. Hauptaugenmerk liegt dabei auf Euter- und Stoffwechselerkrankungen und der Annahme, dass diese Erkrankungen mit einer erkennbaren Milchminderleistung einhergehen (vgl. WENDT et al., 1999, S. 13f). Eine vergleichbare landwirtschaftsbezogene Anwendung zur Mastitiserkennung auf Basis von Varianten für Zellzahlen und Behandlungen als Zielwerte beschreiben CAVERO et al. (2006b). Als Eingabevariablen wurden dazu elektrische Leitfähigkeit der Milch, Milchbildungsrate, Milchflussrate und Laktationstag verwendet. HEALD et al. (2000) nutzten ein vierschichtiges Feedforward-Perceptron für ein Mastitisiagnosesystem zur Klassifizierung von Mastitisfällen je nach bakterieller Krankheitsursache. Die Einflussfaktoren umfassten einerseits tierspezifische Daten wie Alter, Laktation und Laktationstag, prozentuale Änderung der Milchmenge gegenüber dem Vormonat, geschätzte 305-Tage-Produktion, diverse Zellzahlparameter (beispielsweise aktuell, Vormonate, Laktationsdurchschnitt, Anzahl Tage mit erhöhter Zellzahl) und andererseits betriebspezifische Parameter wie Aufstellungsform, Liegeboxengestaltung und Melkpraktiken (Dippen, Voruntersuchungen des Gemelks, Anrüstzeiten, Melkdauer). Zielwerte wurden aus der bakteriologischen Untersuchung von Vorgemelken abgeleitet. Unterteilt wurden dabei vier Klassen

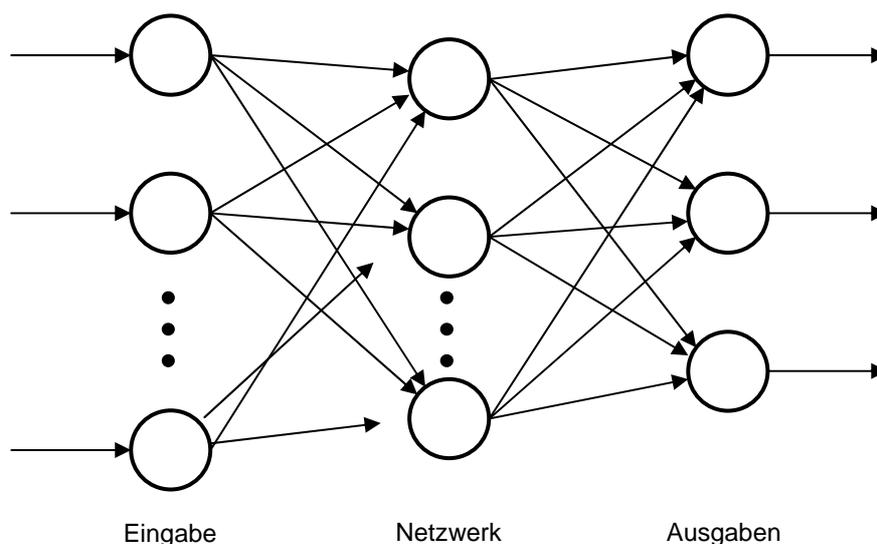


Abbildung 15: Schematische Darstellung eines mehrschichtigen Feedforward-Perceptrons (verändert nach PATTERSON, 1997, S. 161)

(dementsprechend bestand die Ausgabeschicht aus vier Neuronen): kein signifikantes Bakterienwachstum; ansteckende, von Kuh zu Kuh übertragbare Bakterien wie beispielsweise *Streptococcus agalactiae*, *Staphylococcus aureus* und *Staphylococcus chromogenes/hyicus*; umweltassoziierte Erreger wie beispielsweise *Streptococcus nonagalactiae*, *Streptococcus dysgalactiae* oder *Escherichia coli*; andere mastitiserregende Bakterien. SALEHI et al. (1998) entwickelten ein zusammenhängendes System mehrerer KNN zur Klassifizierung von Datensätzen mit hoher und niedriger Milchleistung im ersten Schritt und zur anschließenden speziellen Vorhersage von 305-Tage-Milchleistungen (Milch-, Fett- und Eiweißmenge; dies sind auch die Zielvariablen). Als Eingabevariablen werden hier neben Werten für aktuelle und kumulative Mengen von Milch, Fett und Eiweiß auch Laktation und Laktationstag, Alter, Kalbesaison, Gewicht der Kuh, der Logarithmus der Zellzahl, Fütterungsparameter wie Energie-, Protein- und Trockenmassegehalt der Ration sowie die durchschnittliche 305-Tage-Milchmenge der Herde verwendet. Die beiden Anwendungsbeispiele von HEALD et al. (2000) und SALEHI et al. (1998) greifen dabei allerdings auf Merkmale zurück, die im normalen Produktionsbetrieb derzeit nicht täglich verfügbar sind.

Eine weitere Anwendungsmöglichkeit besteht, wie auch bei der Fuzzy Logic, in der Nutzung zur Östruserkennung. KRIETER et al. (2005) untersuchten ein KNN mit zwei Eingabemerkmalen (Aktivität, vergangene Tage seit letzter Brunst) und einer Zielvariable (Brunststatus: ja/nein).

Weiterhin kann ein Feedforward-Netzwerk auch zur Vorhersage von Milchleistungen, vergleichbar zum gemischten linearen Modell, trainiert werden. Dabei werden allerdings keine kompletten Laktationskurvenparameter geschätzt, sondern durch Eingaben wie Laktationstag/-stadium und Milchmengen vorangegangener Tage wird das Netzwerk trainiert, um die aktuelle Tagesmilchleistung wiederzugeben. Das trainierte Netzwerk muss daraufhin in der Lage sein, bei vorgelegten neuen Eingabevektoren Ähnlichkeiten zu erkennen und ein entsprechendes Ergebnis zu liefern.

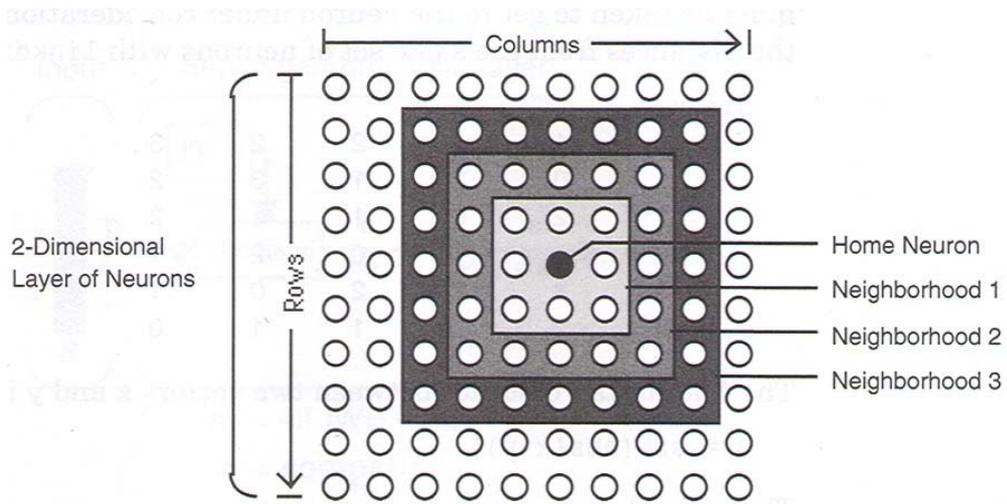


Abbildung 16: Nachbarschaftsbeziehungen in einer quadratischen selbstorganisierenden Karte (DEMUTH und BEALE, 2004, S. 8-15)

3.4.7.2 Selbstorganisierende Karten

Mittels Wettbewerbslernens kann durch nicht-überwachtes Lernen die Klassifikation eines Datensatzes erzielt werden (NAUCK et al., 1996, S. 117ff). Wettbewerbslernen beruht auf dem Prinzip, dass dem neuronalen Netz ein Eingabevektor vorgelegt wird, aber nur ein Neuron, das dem Eingabevektor am ähnlichsten ist „gewinnt“ und erfährt entsprechend der gewählten Lernregeln eine Veränderung der Gewichtungen zum Eingabevektor hin. Dazu müssen die Neuronen zu Beginn des Lernens alle kleine, aber leicht unterschiedliche Gewichtungen aufweisen (PATTERSON, 1997). Alle anderen Neuronen gehen in der Regel leer aus, d.h. sie erfahren keine Gewichtungsanpassung und lernen nicht. Dies führt dazu, dass das gewinnende Neuron bei der nächsten Vorlage eines ähnlichen Eingabevektors mit erhöhter Wahrscheinlichkeit erneut als Gewinner hervorgeht und sich weiter auf „seine“ Kategorie von Eingabevektoren spezialisieren kann (DEMUTH und BEALE, 2004, S. 8-9). Zur Bestimmung des Abstands der Neuronen zum Eingabevektor werden beispielsweise der euklidische Abstand (9) oder die Manhattan- bzw. City-Block-Distanz (10) verwendet (n entspricht dabei der Dimension des Vektors). Mit p und q werden hier die Koordinaten des Eingabevektors und eines Neurons bezeichnet.

$$dist(p, q) = \sqrt{(p - q)^2} = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (9)$$

$$dist(p, q) = |p - q| = \sum_{i=1}^n |p_i - q_i| \quad (10)$$

Bei selbstorganisierenden Karten wird dieses Prinzip um eine Nachbarschaftsbeziehung zwischen den Neuronen erweitert (BORGELT et al., 2003, S. 98ff, vgl. Abbildung 16). Bei zweidimensionalen selbstorganisierenden Karten werden die Ausgabeneuronen gewöhnlich in quadratischen oder hexagonalen Gittern angeordnet (vgl. Abbildung 17). Die verbindenden Kanten stellen die Nachbarschaftsbeziehungen dar. Dadurch ist es möglich, dass benachbarte Neuronen ebenfalls lernen können, d.h. ihre Gewichtungen im Lernprozess auch verändert werden, wenn sie nicht das „Gewinnerneuron“ sind, wenn auch weniger stark. Je nach dem wie die „Nachbarschaft“ um ein Neuron herum definiert wurde, wird eine unterschiedliche Anzahl von Neuronen angepasst. Im Beispiel in Abbildung 16 würden bei einer Nachbarschaftsdistanz von 1 nur die acht unmittelbar benachbarten Neuronen des Gewinnerneurons angepasst, bei einer Nachbarschaftsdistanz von 3 hingegen müssten 48 benachbarte Neuronen Gewichtsänderungen erfahren. Während des Trainings „entfalten“ sich die anfangs nahe um einen Punkt konzentrierten Neuronen zu einer Karte, je nach vorgegebener Struktur der Ausgangsschicht (Abbildung 18). Allerdings kann es bei ungünstiger Initialisierung der Gewichtungen, zu geringer Lernrate oder zu kleiner Nachbarschaft zu Verdrehungen beim Entfalten der Karte kommen (BORGELT et al., 2003, S. 109, vgl. Abbildung 19). Wenn man durch die Nachbarschaftsbeziehungen zwischen den Ausgabeneuronen die relative Lage der verschiedenen Eingabevektoren zueinander wiedergeben kann, erhält man durch die selbstorganisierenden Karten eine topologieerhaltende Abbildung (vgl. z.B. auch Robinson-Projektion der Oberfläche einer Kugel in eine Ebene bei Weltkarten). Dadurch ist man in der Lage, hochdimensionale Strukturen in niedrigdimensionale Räume abbilden zu können (BORGELT et al., 2003, S. 101). Dies kann man sich zur Visualisierung höherdimensionaler Eingabevektoren zunutze

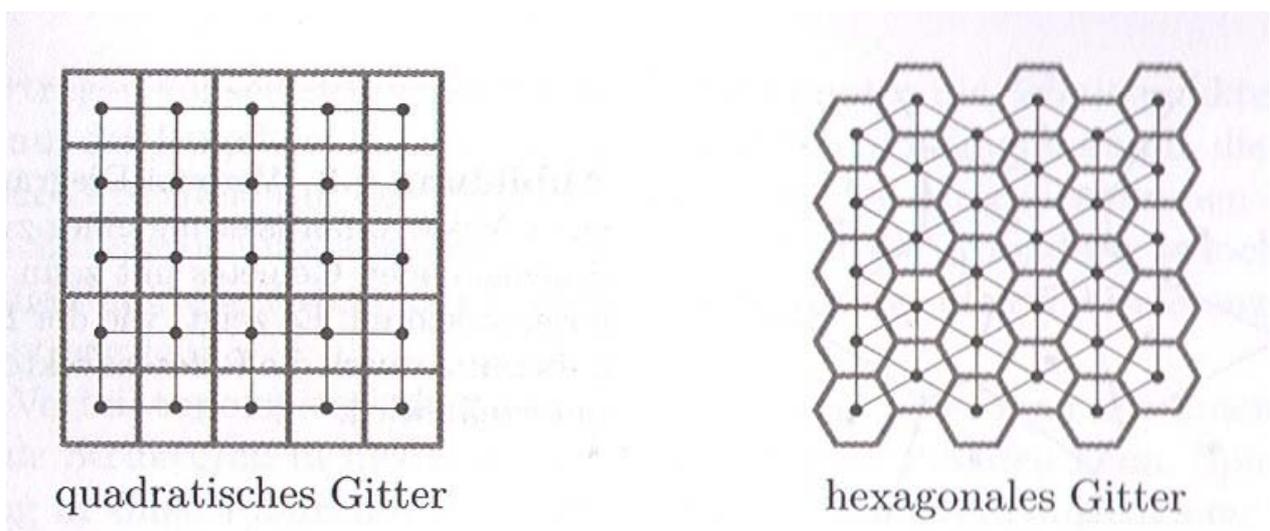


Abbildung 17: Beispiele für eine Anordnung von Neuronen in der Ausgangsschicht einer selbstorganisierenden Karte (BORGELT et al., 2003, S. 99)

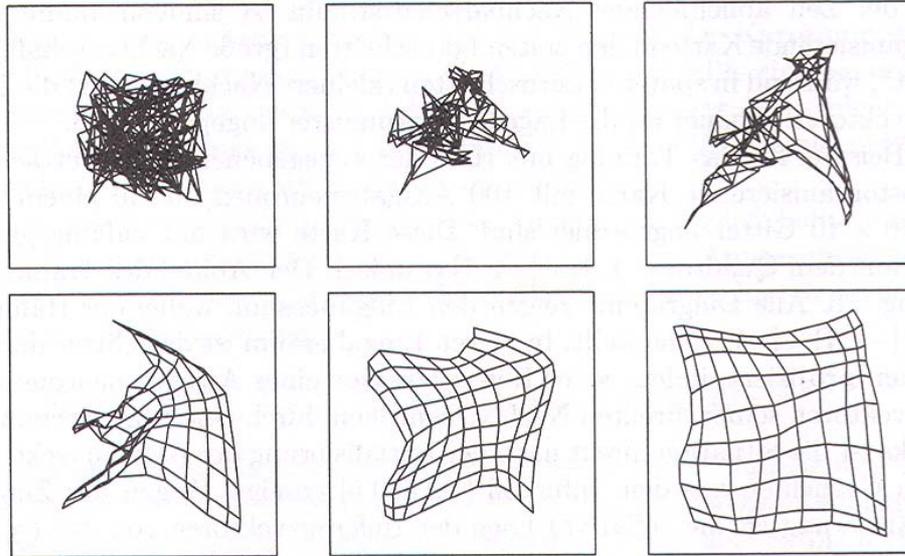


Abbildung 18: Entfaltung einer selbstorganisierenden Karte während des Trainings (BORGELT et al., 2003, S. 108)

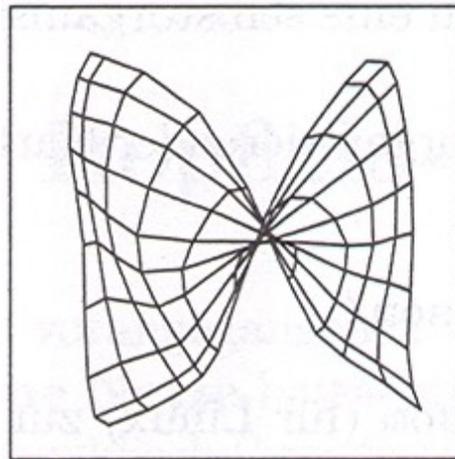


Abbildung 19: Selbstorganisierende Karte nach Initialisierung mit ungünstigen Parametern (BORGELT et al., 2003, S. 109)

machen. Die in der vorliegenden Arbeit genannten Anwendungen kombinieren in der Regel mehr als drei Merkmale miteinander, so dass eine verständliche grafische Darstellung Schwierigkeiten bereitet. Ziel soll es auch hierbei wieder sein, dass ähnliche Eingabevektoren gleiche oder benachbarte Neuronen der Ausgabeschicht aktivieren. Somit können Alarme ausgelöst werden, wenn sich bestimmte Ausgabeneuronen als typische Krankheiten klassifizierende Neuronen erweisen. Selbstorganisierende Karten eignen sich aber nur für unüberwachtes Lernen, d.h. man kann den Ausgabeneuronen erst nach dem Training Klassen zuordnen, indem man überprüft, welche Neuronen häufig auf Eingabevektoren von kranken Tieren reagiert haben. Sobald man diese

Neuronen in der Ausgangsschicht identifiziert hat ist es denkbar, die selbstorganisierende Karte als Kontrollmechanismus benutzen, um die täglich anfallenden Datensätze je Tier auf Ähnlichkeiten zu bekannten Erkrankungsmustern zu untersuchen und so möglicherweise eine frühzeitigere Erkennung von Krankheitsfällen ermöglichen. Dieser Ansatz der selbstorganisierenden Karten wurde in der vorliegenden Arbeit nicht weiter verfolgt, bietet aber Ansätze für weitere Untersuchungen, vor allem bei schlechter Datenqualität für die Zielwerte.

3.5 Vergleich der Methoden und Modelle

Zum Vergleich der in den Abschnitten 3.1 bis 3.4 verwendeten Methoden und Modelle werden Erkennungs- und Fehlerraten sowie die Spezifität herangezogen. Dazu müssen die Ausgabewerte der einzelnen Methoden unter verschiedenen vorgegebenen Grenzwerten mit Zielwerten verglichen werden, die für den jeweiligen Datensatz ein Ereignis bzw. eine Anomalie (beispielsweise eine Erkrankung oder negative Leistungsabweichung) festlegen oder ausschließen. Durch den Vergleich kann man die einzelnen Datensätze in vier Gruppen klassifizieren (vgl. DE MOL et al., 1997; KÖHLER, 2003, S. 58; KRIETER et al., 2005):

- (1) Wahr positiv (*true positive*, TP): ein Ausgabewert, der einen vorgegebenen Grenzwert überschreitet, stimmt mit einer vorliegenden Anomalie überein.
- (2) Falsch positiv (*false positive*, FP): ein Ausgabewert, der einen vorgegebenen Grenzwert überschreitet, tritt ohne eine vorliegende Anomalie auf.
- (3) Wahr negativ (*true negative*, TN): ein Ausgabewert überschreitet einen vorgegebenen Grenzwert nicht, und es liegt keine Anomalie vor.
- (4) Falsch negativ (*false negative*, FN): ein Ausgabewert überschreitet einen vorgegebenen Grenzwert nicht, obwohl eine Anomalie vorliegt.

Die Erkennungsrate in Prozent ist definiert als der Anteil der Datensätze die einen richtigen Alarm auslösen an den insgesamt vorhandenen Anomalien:

$$\text{Erkennungsrate} = \frac{TP}{(TP + FN)} \cdot 100 \quad (11)$$

Die Fehlerrate in Prozent ist gegeben als Anteil der falschen Alarme an den insgesamt auftretenden Alarmen:

$$\text{Fehlerrate} = \frac{FP}{(FP + TP)} \cdot 100 \quad (12)$$

Die Spezifität schließlich gibt die Relation der wahr negativen Datensätze zur Gesamtzahl der Datensätze ohne Anomalie wieder:

$$\text{Spezifität} = \frac{TN}{(FP + TN)} \cdot 100 \quad (13)$$

Erstrebenswert sind eine möglichst hohe Erkennungsrate bei gleichzeitig möglichst niedriger Fehlerrate und hoher Spezifität. Gerade Erkennungs- und Fehlerrate hängen dabei jedoch neben der verwendeten Methode auch vom jeweils gesetzten Grenzwert ab, so dass ein hoher Grenzwert dazu führen wird, dass zwar die Fehlerrate sehr gering ist, aber gleichzeitig auch wenige Anomalien tatsächlich erkannt werden. Umgekehrt führt ein niedrig gewählter Grenzwert dazu, dass auch schwieriger zu erkennende (weil schwächer ausgeprägte) Anomalien erkannt werden, aber gleichzeitig erhöht sich dabei auch die Anzahl der Falschalarme und dadurch die Fehlerrate. Die Spezifität hängt wiederum stark von der Anzahl der insgesamt untersuchten Anzahl von Datensätzen und der Häufigkeit des Auftretens von Anomalien ab. Bei einer wie in der vorliegenden Arbeit hohen Anzahl von Datensätzen gegenüber einer vergleichsweise geringen Anzahl von Ereignissen wird die Spezifität gemäß der Definition immer recht hoch sein, da die Anzahl der wahr negativen Datensätze sehr hoch ist.

4 Ergebnisse

4.1 Nutzung der Milchmenge

Gemäß der Zielstellung sollten Erkrankungen, insbesondere Euter- und Stoffwechselerkrankungen, anhand von negativen Milchmengenabweichungen identifiziert werden. Nach den ersten Auswertungen der verfügbaren Daten zeigte sich jedoch, dass die generelle Eignung der Milchmenge als Indikator für eine Erkrankung zu hinterfragen ist. Einerseits traten sehr häufig negative Milchmengenabweichungen auf, an denen keine Erkrankung des Tieres dokumentiert war, andererseits waren Krankheiten an Tagen dokumentiert, an denen sogar ansteigende Milchmengen zu verzeichnen waren. Es war unter den vorliegenden Produktionsbedingungen nicht zu klären, ob es sich um Probleme in der Dokumentation oder der Erkennung von Krankheiten im Bestand handelte. Demnach erschien eine Bewertung der im vorangegangenen Abschnitt 3 vorgestellten Methoden anhand der dokumentierten Krankheiten nicht sinnvoll.

Es wurde daher der grundlegende Gedanke weiterverfolgt, dass eine negative Abweichung der Milchmenge über die normalen physiologischen Schwankungen hinaus der Aufmerksamkeit des Betreuungspersonals bedarf. Entsprechend sollten die untersuchten Methoden in der Lage sein, normalen Schwankungen von alarmwürdigen Abweichungen zu unterscheiden, d.h. normale Gemelke von Milchminderungen abzugrenzen. Bei Beachtung der oben beschriebenen Schwierigkeiten sollten daher zur Bewertung also nicht die erkannten Krankheiten, sondern die als Milchminderungen klassifizierten Tagesgemelke herangezogen werden. Um die drei untersuchten Betriebe einheitlich bewerten zu können, ist zunächst die Definition einer Milchmindermenge notwendig.

4.1.1 Definition von Milchminderungen

Bei Eutererkrankungen ergeben sich durch die verminderte Milchbildungsfähigkeit des Drüsengewebes bei klinischen Mastitiden und durch den erhöhten Zellgehalt der Milch bei subklinischen Mastitiden Milchminderungen. Diese Milchminderungen können je nach Ausprägung und Verlauf der Mastitis unterschiedlich stark ausgeprägt sein. WENDT et al. (1998, S. 13) nennen für Zellgehalte von über 400.000 Zellen pro ml Milch Verluste von 7.5 bis 12%. ZSCHÖCK et al. (1998, S. 508) beziffern die Leistungsminderung bei von subklinischen Mastitiden erkrankten Vierteln gegenüber den Parallelvierteln mit durchschnittlich 20%. Demnach müsste bei Betrachtung des gesamten Euters schon eine Verminderung der Milchmenge von etwa 5% als

auffällig gelten. Besonders bei Berücksichtigung der Häufigkeit des Auftretens subklinischer Mastitiden im Verhältnis zu klinischen Mastitiden (nach ZSCHÖCK et al. (1998, S. 508) treten subklinische Mastitiden um den Faktor 20 bis 50 häufiger auf als klinische Mastitiden) wird deutlich, dass hier viertelweise vorliegende Milchmengen von Vorteil für eine krankheitsspezifische Erkennung wären. Leider war das Merkmal Milchmenge für die vorliegende Arbeit nur auf das Gesamteuter bezogen verfügbar.

Ebenso ist bei Stoffwechselerkrankungen mit sinkenden Milchmengen zu rechnen. Zunächst wurde daher ein Kriterium für eine vorliegende Milchmindermenge definiert, um die Nutzbarkeit der in Abschnitt 3 beschriebenen Methoden zur Erkennung dieser Milchmindermengen zu untersuchen. Unter Berücksichtigung der genannten Literaturwerte und der in den Betrieben dokumentierten Erkrankungen erschien eine negative Abweichung von 7% gegenüber dem Vortag als Definition sinnvoll. Um vorkommende kurzfristige Schwankungen in der Milchmenge (beispielsweise durch die in Abschnitt 2 angesprochenen Fehlzuordnungen von Gemelken) zu relativieren wurde als Zusatzkriterium gefordert, dass auch gegenüber dem gleitenden Durchschnitt der letzten drei Tage ein Rückgang der Milchmenge um mindestens 7% vorliegen musste, damit ein Datensatz als negative Milchmengenabweichung klassifiziert wurde. Durch diese Vorgehensweise ergaben sich die in Tabelle 7 dargestellte Anzahl negativer Abweichungen bei dokumentierten Euter- (EU) und Stoffwechselerkrankungen (SW) in den untersuchten Betrieben. Tabelle 8 beinhaltet die insgesamt vorliegende Anzahl der entsprechenden Erkrankungen aus der Dokumentation der Betriebe und die sich nach obiger Definition ergebende Anzahl von Datensätzen mit negativen Milchmengenabweichungen, bezogen auf die Anzahl der von den jeweiligen Betrieben genutzten Tagesmilchmengen. In den Betrieben 1 und 3 wurde hierbei nur der erste Tag einer Erkrankung festgehalten, während für Betrieb 2 auch nachfolgende Tage dokumentiert und enthalten sind. Aus den beiden Tabellen ist einerseits ersichtlich, dass Erkrankungen sich bereits vor der Dokumentation durch negative Abweichungen der Milchmenge bemerkbar machen können, dass aber andererseits der Großteil der dokumentierten Euter- und Stoffwechselerkrankungen mit geringeren Milchminderleistungen als 7% einhergehen.

Diese Datensätze mit negativen Milchmengenabweichungen sollten jeweils mit den im Folgenden dargestellten Modellen erkannt werden. Das besondere Erkennungsproblem besteht dabei vor allem darin, dass zwischen physiologisch normalen Schwankungen und Abweichungen, die einer genaueren Betrachtung des Tieres bedürfen, unterschieden werden muss. Diese Vorgehensweise ist besonders günstig im Hinblick auf den in dieser Arbeit im Mittelpunkt stehenden Methodenvergleich.

Tabelle 7: Anzahl Datensätze mit negativer Milchabweichung bei dokumentierten Euter- und Stoffwechselerkrankungen

	Betrieb 1			Betrieb 2			Betrieb 3		
	Vortag	2. Vortag	3. Vortag	Vortag	2. Vortag	3. Vortag	Vortag	2. Vortag	3. Vortag
EU ^a	6	11	5	620	628	512	0	1	2
SW ^b	4	3	1	24	20	18	0	0	0

^a Eutererkrankungen

^b Stoffwechselerkrankungen

Tabelle 8: Anteil dokumentierter Euter- und Stoffwechselerkrankungen und negativer Milchmengenabweichungen gegenüber den Tagesmilchmengen je Betrieb

	Betrieb 1	Betrieb 2	Betrieb 3
Anteil EU ^a gesamt	0.07%	0.98%	0.19%
Anteil SW ^b gesamt	0.02%	0.01%	0.00%
Anteil negative Abweichungen	2.22%	8.26%	4.07%

^a Eutererkrankungen

^b Stoffwechselerkrankungen

4.1.2 Anwendung des gemischten linearen Modells

4.1.2.1 Vergleich der Modellierungsansätze für Laktationskurven und der Modellvarianten Fixed und Random Regression

Zum Vergleich der drei Modellierungsansätze für Laktationskurven wurde der Datensatz aus Betrieb 1 als Datengrundlage verwendet, da von diesem zu Beginn des Projektes bereits Daten verfügbar waren. Für die Berechnungen wurde PROC MIXED aus dem Statistikpaket SAS (Version 9.1.3, Service Pack 2) genutzt. Als Einflussfaktoren auf die tierspezifische Milchleistung wurden Laktationsnummer und Testtag verwendet. Der Laktationstag wurde in Form der Covariablen der jeweils zu vergleichenden Ansätze zur Laktationskurvenmodellierung berücksichtigt (vgl. Abschnitt 3.1.2).

Nicht berücksichtigt wurden Datensätze mit einer Tagesmilchmenge von weniger als 5 kg oder von höheren Laktationstagen als 400. Wegen der geringeren Anzahl vorhandener Tiere wurden die

Tabelle 9: Anpassungsgüte der Modelle nach Ali und Schaeffer (1987) im Vergleich

Kriterium	-2 Log Likelihood	AIC	Likelihood-Quotient
Modell 1: Fixed Regression	392197.7	393395.7	Vergleich mit ...
Modell 2: Random Regression I ^a	367076.2	368282.2	Modell 1: 25121.5 ^c (4 FG)
Modell 3: Random Regression II ^b	347744.1	348972.1	Modell 2: 19332.1 ^c (15 FG)

^a tierspezifische Laktationskurvenmodellierung

^b tier- und laktationsspezifische Laktationskurvenmodellierung

^c signifikant bei $\alpha=5\%$

Tabelle 10: Anpassungsgüte der Modelle nach Wilmlink (1987) im Vergleich

Kriterium	-2 Log Likelihood	AIC	Likelihood-Quotient
Modell 1: Fixed Regression	395306.3	396482.3	Vergleich mit ...
Modell 2: Random Regression I ^a	374553.5	375739.5	Modell 1: 20752.8 ^c (3 FG)
Modell 3: Random Regression II ^b	352203.7	353407.7	Modell 2: 22349.8 ^c (9 FG)

^a tierspezifische Laktationskurvenmodellierung

^b tier- und laktationsspezifische Laktationskurvenmodellierung

^c signifikant bei $\alpha=5\%$

Tabelle 11: Anpassungsgüte der Modelle nach Guo (1998) im Vergleich

Kriterium	-2 Log Likelihood	AIC	Likelihood-Quotient
Modell 1: Fixed Regression	398431.9	399613.6	Vergleich mit ...
Modell 2: Random Regression I ^a	379713.3	380899.3	Modell 1: 18718.6 ^c (3 FG)
Modell 3: Random Regression II ^b	359035.6	360239.6	Modell 2: 20677.7 ^c (9 FG)

^a tierspezifische Laktationskurvenmodellierung

^b tier- und laktationsspezifische Laktationskurvenmodellierung

^c signifikant bei $\alpha=5\%$

Tiere in der vierten und fünften Laktation zu einer Gruppe zusammengefasst, während die erste bis dritte Laktation jeweils einzelne Laktationsgruppen bildeten. Daraus ergab sich für den Modellvergleich eine Anzahl von über 69500 Datensätzen. Die Resultate des Modellvergleichs sind in den Tabellen 9 bis 11 dargestellt.

Über alle Modellvarianten hinweg erzielt der Ansatz von ALI und SCHAEFFER (1987) im Vergleich zu den Ansätzen von WILMINK (1987) und GUO (1998) damit, gemessen an dem analytischen

Kriterium AIC, die jeweils beste Anpassungsgüte der Laktationskurven an den unbekanntem tatsächlichen Laktationskurvenverlauf.

Gegenüber der alleinigen Betrachtung der tierindividuellen Abweichungen von einer Laktationskurve im Fixed-Regression-Modell, die den Herdendurchschnitt für die jeweils betrachtete Laktation darstellt, sind die Random-Regression-Modelle mit der Möglichkeit der Einbeziehung laktations- und einzeltierspezifischer Effekte vorzuziehen. Der zu erwartende unterschiedliche Verlauf der Laktationskurven eines Tieres in unterschiedlichen Laktationen wird durch das Random-Regression-Modell besser abgebildet, als das mit dem Fixed-Regression-Modell möglich ist (vgl. JAMROZIK et al., 1997), was sich in den durchweg besseren Kriterien zur Anpassungsgüte zeigt.

Daraus ergibt sich für die weiterführenden Untersuchungen die Benutzung des Modellierungsansatzes für Laktationskurven nach ALI und SCHAEFFER (1987) mit dem einzeltier- und laktationsspezifischen Random-Regression-Modell.

Die Abbildungen 20 bis 24 zeigen die auf den festen Effekten basierenden herdenweiten Laktationskurven für die untersuchten Regressionsmodelle nach den Ansätzen zur Laktationskurvenmodellierung von ALI und SCHAEFFER (1987) (Abbildungen 20-22) sowie die Fixed-Regression-Modelle nach den Ansätzen von WILMINK (1987; Abbildung 23) und GUO (1998; Abbildung 24). Für die eigentlichen Laktationskurven je Tier müssen in den Fixed-Regression-Modellen noch der zufällige Tiereffekt und in den Random-Regression-Modellen noch die zufälligen Effekte für die Kovariablen der Ansätze zur Modellierung der Laktationskurven addiert werden.

Während die Ansätze von ALI und SCHAEFFER (1987) sowie GUO (1998) zu Beginn der Laktation einen etwas verzögerten Milchmengenanstieg aufweisen, sieht der Ansatz von WILMINK (1987) einen steilen Anstieg vor. Nach Überschreiten der Höchstmenge erfolgt der weitere Laktationsverlauf nach GUO (1998) und WILMINK (1987) quasi als linear sinkende Funktion, während der Ansatz von ALI und SCHAEFFER (1987) die Laktationskurve gegen Ende der Laktation flacher fallen lässt.

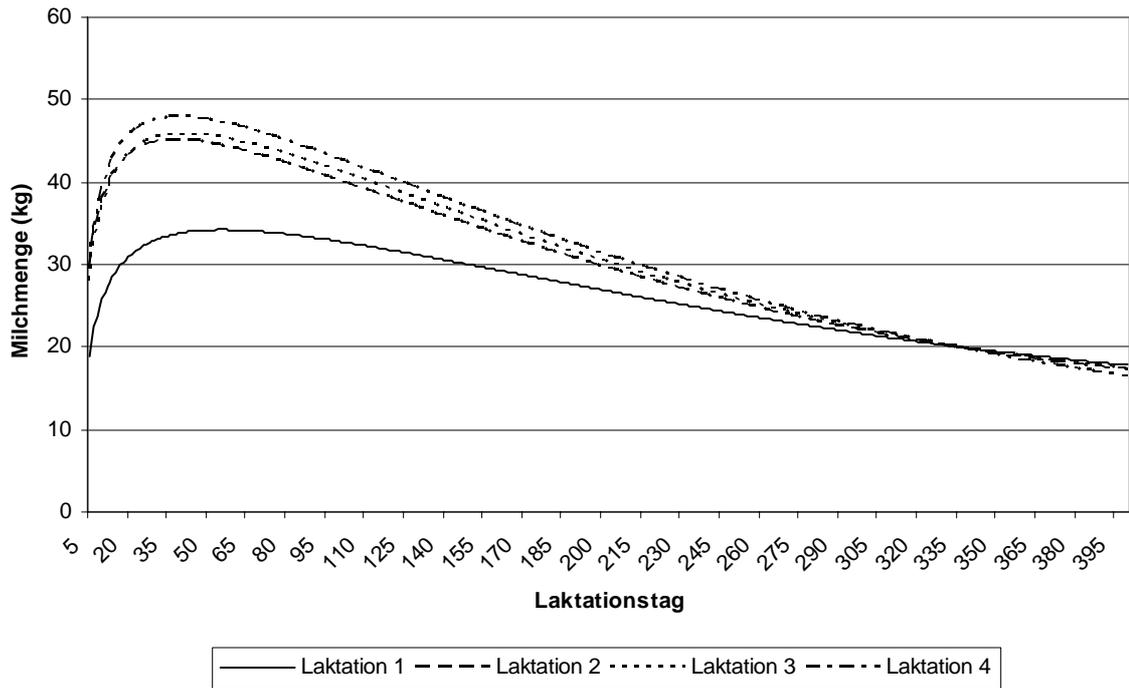


Abbildung 20: Herdenweite Laktationskurven für Betrieb 1 nach ALI und SCHAEFFER (1987) bei Nutzung des Fixed-Regression-Modells

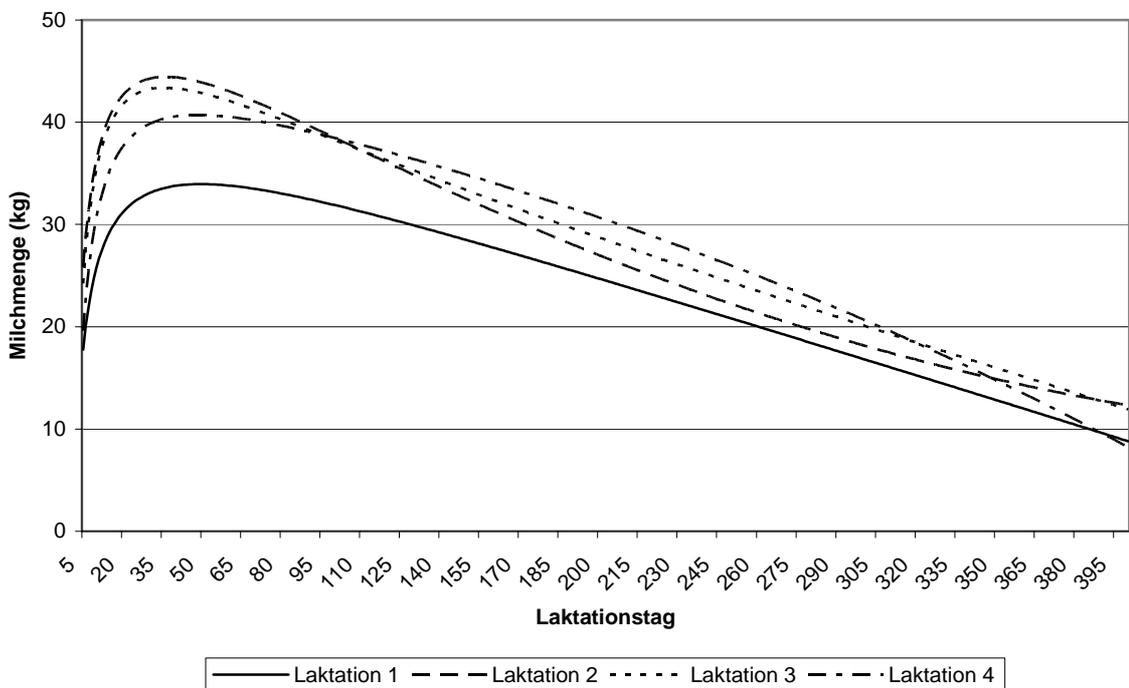


Abbildung 21: Herdenweite Laktationskurven für Betrieb 1 nach ALI und SCHAEFFER (1987) bei Nutzung des tierspezifischen Random-Regression-Modells

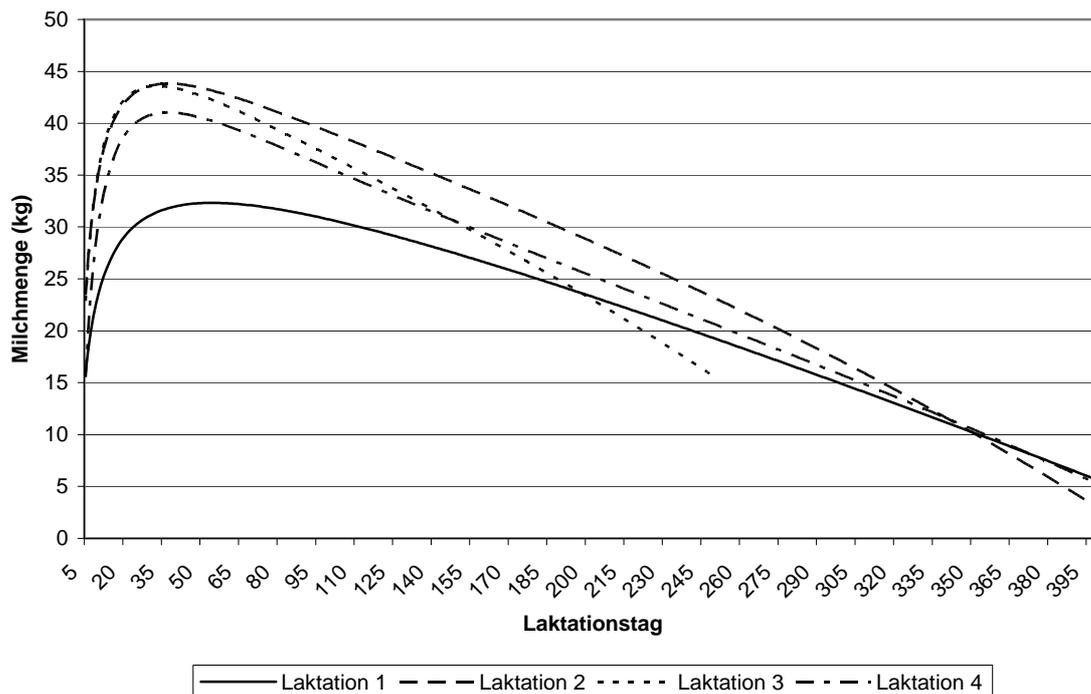


Abbildung 22: Herdenweite Laktationskurven für Betrieb 1 nach ALI und SCHAEFFER (1987) bei Nutzung des tier- und laktationsspezifischen Random-Regression-Modells

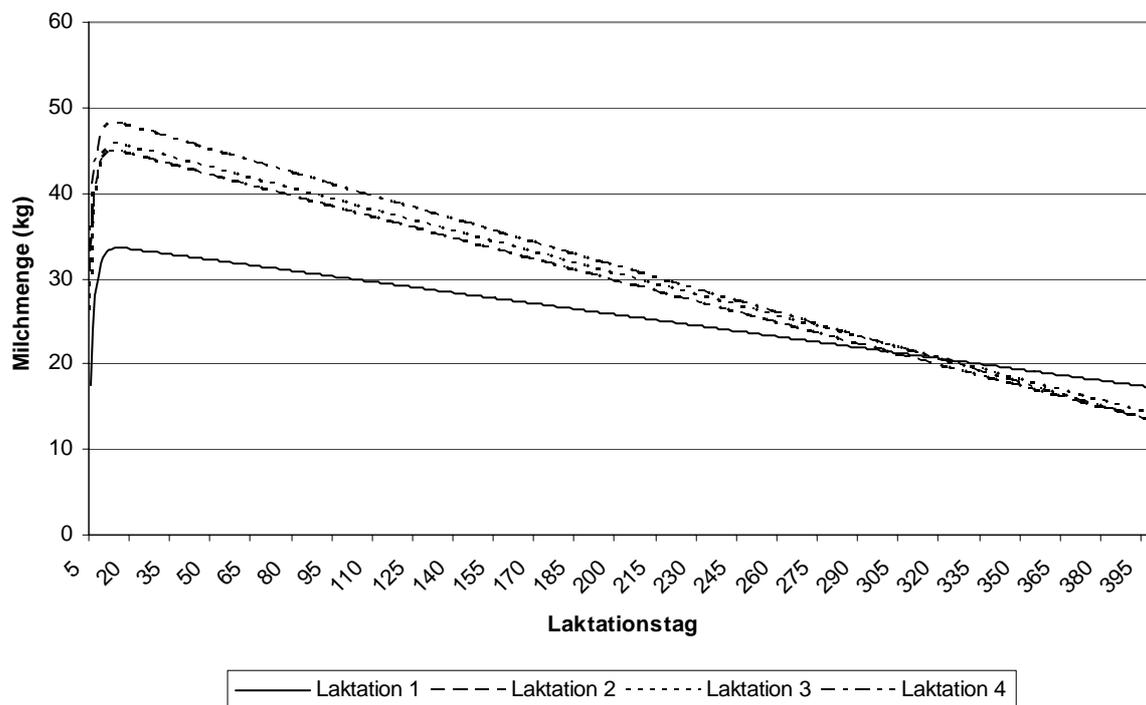


Abbildung 23: Herdenweite Laktationskurven für Betrieb 1 nach WILMINK (1987) bei Nutzung des Fixed-Regression-Modells

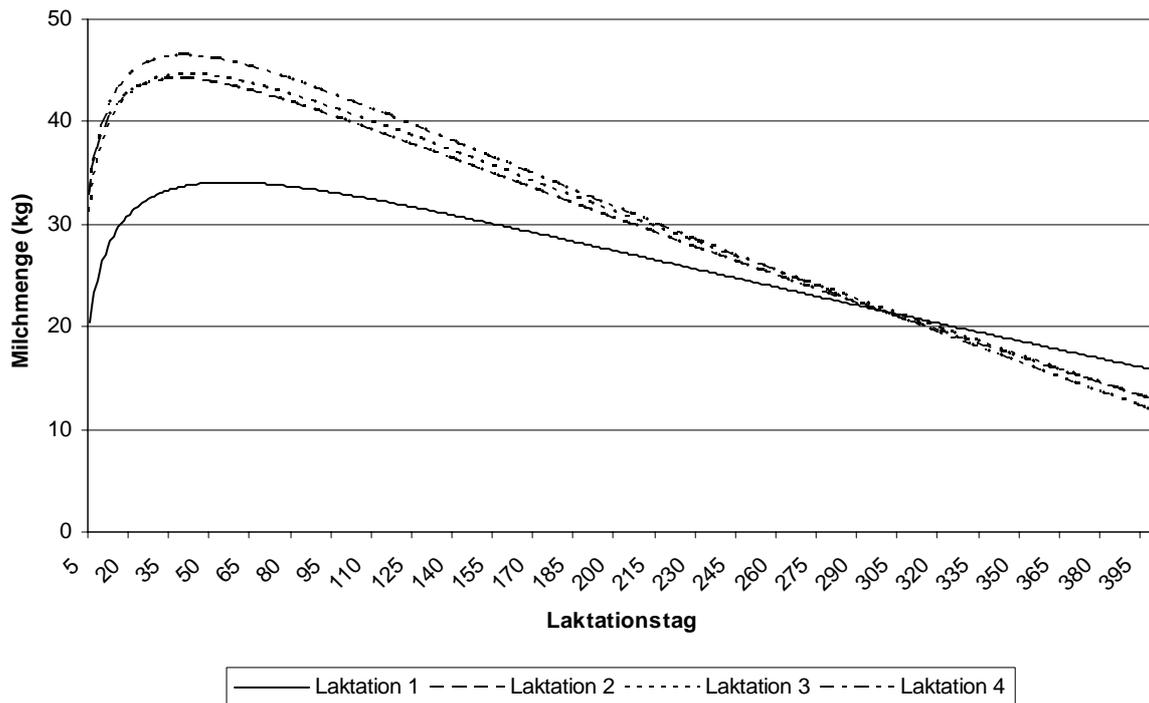


Abbildung 24: Herdenweite Laktationskurven für Betrieb 1 nach GUO (1998) bei Nutzung des Fixed-Regression-Modells

Bei dem tier- und laktationsspezifischen Random-Regression-Modell standen für die dritte Laktation nach dem 250. Laktationstag nicht mehr genügend Daten von Tieren zur Verfügung, um plausible herdenweite Effekte zu schätzen (Abbildung 22). Über zwei Drittel der Tiere der dritten Laktation waren bis zu diesem Laktationstag aus der Herde ausgeschieden. Die Daten der Tiere aus der vierten Laktation beinhalten dagegen wieder eine gleichmäßige Anzahl von Tieren, da diese Daten vorwiegend dem Beginn der Erfassungsperiode entstammen.

4.1.2.2 Erkennung von Milchminderungen mit dem gemischten linearen Modell

Um den Einfluss der Menge und der Aktualität der zum Zeitpunkt der Schätzung verfügbaren Daten auf die Erkennung der Milchminderungen zu untersuchen, wurden für jeden der drei Betriebe vier Modelle mit unterschiedlicher Menge verfügbarer Datensätze gerechnet. Minimum waren dabei 60 Tage, die weiteren Varianten umfassten 90, 120 für alle drei Betriebe und 281 Tage in Betrieb 1, 240 Tage für Betrieb 2 und 245 Tage in Betrieb 3. In den Betrieben 1 und 3 ergeben sich die längsten Zeiträume durch die maximale Anzahl zusammenhängend verfügbarer Tage.

Als Grenzwert, bei dessen Unterschreiten eine negative Milchmengenabweichung als erkannt gelten sollte, wurde der Standardfehler der mit PROC MIXED berechneten Vorhersagewerte verwendet.

Auch hier wurden mehrere Varianten ausgewertet, die sich durch einen Faktor unterscheiden, mit denen der Standardfehler multipliziert wurde. Dadurch wurden unterschiedliche Grenzwertniveaus erreicht. Als Faktoren wurden 0.5 sowie 1 bis 6 in ganzzahligen Schritten gewählt. Die Ergebnisse sind den Tabellen 12 bis 14 zu entnehmen. ER bezeichnet die Erkennungsrate, FR die Fehlerrate und Spez die Spezifität; die Angaben sind jeweils Prozentwerte.

Die durchweg hohen Werte für die Spezifität sind mit der vergleichsweise geringen Anzahl von negativen Abweichungen im Bezug zur Gesamtzahl der untersuchten Datensätze zu erklären. Auch die hohen Erkennungsraten bei den engeren Konfidenzintervall mit niedrigen Faktoren waren zu erwarten. Allerdings fallen die Fehlerraten ebenfalls sehr hoch aus, so dass eine Eignung des gemischten linearen Modells zur Erkennung von Milchminderungen nur sehr bedingt gegeben ist. Selbst bei Wahl des Faktors 2, der etwa dem üblichen Konfidenzniveau von $P = 0.95$ entspricht und bei dem für die untersuchten Betriebe Erkennungsraten nahe 80% erreicht werden, bleiben die Fehlerraten in vergleichbaren Größenordnungen von 60 bis 80%. Dies ist dadurch erklärbar, dass sich abschnittsweise deutliche Über- oder Unterschätzungen der tatsächlich beobachtbaren Milchmenge ergaben. Vor allem trat dies nach Zeiträumen auf, in denen Daten fehlten oder in denen sich das Milchleistungsniveau in kurzer Zeit stark änderte. Die Anpassung der Vorhersagewerte an das neue Milchmengenniveau gelang danach zeitlich nur verzögert. Im Falle von Überschätzungen führte dies vermehrt zu Fehlalarmen, im (selteneren) Falle von Unterschätzungen zu fehlenden Alarmen. Daher erschien die alleinige Verwendung des gemischten linearen Modells zur Erkennung von Milchminderungen nicht ausreichend.

Tabelle 12: Ergebnisse des gemischten linearen Modells für Betrieb 1 (%)

Faktor		0.5	1	2	3	4	5	6
60 Tage	ER ^a	95.68	93.92	86.70	76.35	63.90	52.59	42.87
	FR ^b	85.44	82.88	76.14	68.80	61.53	53.49	46.56
	Spez ^c	99.63	99.52	99.09	98.51	97.83	97.23	96.73
90 Tage	ER ^a	95.45	93.12	86.24	76.58	64.41	53.21	43.09
	FR ^b	85.70	83.28	76.93	69.32	61.53	53.89	46.69
	Spez ^c	99.61	99.46	99.05	98.52	97.86	97.27	96.74
120 Tage	ER ^a	91.19	89.71	82.38	72.60	60.43	49.46	39.91
	FR ^b	86.05	83.46	77.03	68.77	60.86	51.75	44.64
	Spez ^c	99.25	99.20	98.80	98.28	97.63	97.07	96.58
281 Tage	ER ^a	93.69	91.53	83.91	73.62	61.46	49.97	39.34
	FR ^b	86.31	83.97	78.23	70.94	64.01	57.14	50.50
	Spez ^c	99.45	99.33	98.88	98.32	97.67	97.08	96.53

^a Erkennungsrate; ^b Fehlerrate; ^c Spezifität

Tabelle 13: Ergebnisse des gemischten linearen Modells für Betrieb 2 (%)

Faktor		0.5	1	2	3	4	5	6
60 Tage	ER ^a	90.85	86.99	77.79	67.42	58.28	51.01	44.96
	FR ^b	75.4	70.33	58.08	45.61	34.86	27.17	22.36
	Spez ^c	98.59	98.2	97.33	96.35	95.5	94.82	94.26
90 Tage	ER ^a	90.15	86.59	77.71	67.72	59.28	51.89	46.1
	FR ^b	76	71.48	60.5	48.59	38.12	30.27	24.66
	Spez ^c	98.46	98.12	97.28	96.35	95.58	94.89	94.36
120 Tage	ER ^a	89.81	86.22	77.33	67.74	58.8	51.69	45.96
	FR ^b	76.22	71.9	61.2	49.85	39.21	31.42	25.21
	Spez ^c	98.4	98.06	97.23	96.34	95.52	94.87	94.34
240 Tage	ER ^a	88.47	85.22	76.77	67.39	58.86	51.79	46.03
	FR ^b	77.05	73.05	63.67	53.21	43.78	35.67	29.24
	Spez ^c	98.17	97.89	97.11	96.26	95.49	94.85	94.33

^a Erkennungsrate

^b Fehlerrate

^c Spezifität

Tabelle 14: Ergebnisse des gemischten linearen Modells für Betrieb 3 (%)

Faktor		0.5	1	2	3	4	5	6
60 Tage	ER ^a	90.82	89.29	83.37	73.80	61.76	50.48	43.21
	FR ^b	88.00	85.89	80.15	71.93	64.54	55.63	45.28
	Spez ^c	99.22	99.18	98.90	98.44	97.83	97.28	96.94
90 Tage	ER ^a	88.72	87.76	82.22	72.47	60.23	49.14	41.87
	FR ^b	87.96	85.81	80.32	73.40	64.80	57.02	46.06
	Spez ^c	99.06	99.07	98.83	98.35	97.75	97.21	96.87
120 Tage	ER ^a	87.38	85.85	81.07	69.60	59.66	49.33	39.77
	FR ^b	88.15	86.20	80.80	74.26	65.18	58.12	51.40
	Spez ^c	98.95	98.92	98.75	98.18	97.71	97.21	96.75
245 Tage	ER ^a	91.97	91.01	84.32	76.48	65.97	54.49	45.12
	FR ^b	88.14	86.15	81.52	74.55	66.04	58.33	49.89
	Spez ^c	99.31	99.30	98.94	98.56	98.05	97.48	97.03

^a Erkennungsrate

^b Fehlerrate

^c Spezifität

4.1.2.3 Berechnung des Abweichungsfaktors zwischen Trends von Vorhersage- und beobachteten Milchmengen

Wegen der den Laktationstagen angepassten Modellierung der Laktationskurve ist davon auszugehen, dass der Trend der mit dem gemischten linearen Modell vorhergesagten Werte für die Milchmenge dem in der jeweiligen Laktationsphase zu erwartenden Trend entspricht. Vor diesem Hintergrund wurde versucht, die berechneten Vorhersagewerte für die weiterhin untersuchten

Methoden (vgl. Abschnitte 3.3, 3.4) nutzbar zu machen. Zu diesem Zweck wurden zunächst für beobachtete (y) und vorhergesagte Milchmengen (z) zum Zeitpunkt t die relativen Abweichungen a_y und a_z gegenüber dem Mittel der Tage $t-3$ bis $t-1$ berechnet:

$$a_{y_t} = \left(\frac{y_t}{\frac{1}{3} \sum_{i=3}^{t-1} y_{t_i}} \right) - 1 \quad (14)$$

bzw.

$$a_{z_t} = \left(\frac{z_t}{\frac{1}{3} \sum_{i=3}^{t-1} z_{t_i}} \right) - 1 \quad (15)$$

Aus den in (14) und (15) berechneten relativen Abweichungen a_y und a_z gegenüber dem Mittel der vorangegangenen drei Tage lässt sich nun errechnen, um welchen Faktor der beobachtete kurzfristige Trend der Milchmenge für ein Tier vom für diesen Zeitraum vorhergesagten Trend für dieses Tier abweicht. Dieser Abweichungsfaktor wird im Folgenden mit AF bezeichnet.

$$AF = \begin{cases} a_y & | \quad a_z = 0 \\ -\frac{a_y}{a_z} & | \quad (a_y < a_z) \wedge (((a_y > 0) \wedge (a_z > 0)) \vee ((a_y < 0) \wedge (a_z < 0))) \\ \frac{a_y}{a_z} & | \quad (a_y > a_z) \wedge (((a_y > 0) \wedge (a_z > 0)) \vee ((a_y < 0) \wedge (a_z < 0))) \\ -\frac{a_y}{a_z} + 1 & | \quad (a_y > a_z) \wedge (a_y > 0) \wedge (a_z < 0) \\ \frac{a_y}{a_z} - 1 & | \quad (a_y < a_z) \wedge (a_y < 0) \wedge (a_z > 0) \\ -a_z & | \quad a_y = 0 \end{cases} \quad (16)$$

Beträgt beispielsweise die vorhergesagte Abweichung der Milchmenge $a_z = 0.05$ und die tatsächlich beobachtete Abweichung der Milchmenge $a_y = -0.1$, gelten $(a_y < a_z)$, $(a_y < 0)$ und $(a_z > 0)$, sodass der AF nach $(a_y/a_z)-1$ berechnet wird. Der AF beträgt in diesem Fall -3 .

Den Faktor AF kann man in dieser Form als Zusatzinformation in weiteren entscheidungsunterstützenden Methoden einsetzen, da ein hoher negativer AF für Datensätze mit

Milchminderungen zu erwarten ist. Als alleiniges Merkmal ist die Verwendung eingeschränkt, da auch sehr kleine Abweichungen hohe Faktoren erzeugen können. Beispielsweise würde bei einer Abweichung der beobachteten Milchmenge um -0.1% bei gleichzeitiger Abweichung der geschätzten Milchmenge von -0.01% im selben Zeitraum ein Abweichungsfaktor von -10 resultieren, während sich bei einer Abweichung der beobachteten Milchmenge um -7% bei gleichzeitiger Abweichung der geschätzten Milchmenge von -1% nur ein AF von -7 ergibt. In Verbindung mit absoluten und relativen Abweichungen ist die Verwendung des AF als Zusatzinformation jedoch denkbar.

4.1.3 Erkennung von Milchminderungen mit EWMA-Charts

Um negative Milchmengenabweichungen mit einem EWMA-Chart zu erfassen, wurden für die drei Betriebe gemäß Abschnitt 3.2 der EWMA-Wert und die oberen und unteren Grenzwerte berechnet. Der Gewichtungsfaktor λ zur stärkeren Gewichtung zurückliegender ($\lambda \rightarrow 0$) oder aktueller ($\lambda \rightarrow 1$) Beobachtungen wurde von 0.1 bis 0.9 in Schritten von 0.2 variiert. Der L-Wert zur Vorgabe eines Grenzwertes nahe ($L \rightarrow 0$) oder weiter entfernt ($L \rightarrow \infty$) vom Mittelwert wurde von 0.5 bis 2 in Schritten von 0.5 variiert. Dadurch ergeben sich je Betrieb 20 Varianten. Die benötigten Statistiken wurden mit PROC MEANS aus dem Statistikpaket SAS (Version 9.1.3, Service Pack 2) berechnet. Als Startwert $EWMA_0$ wurde jeweils der erste verfügbare Beobachtungswert jeden Tieres verwendet, um das Milchmengenniveau der aktuellen Laktationsphase als Bezugspunkt zu nutzen. Ein „historischer“ Mittelwert beispielsweise aus einer vorangegangenen Laktation würde dazu führen, dass das EWMA-Chart je nach gewähltem λ längere Zeit benötigt, um sich mit Hilfe der dazukommenden aktuellen Beobachtungen auf dem derzeitigen Niveau einzupendeln. Zudem liegen bei erstlaktierenden Kühen noch keine „historischen“ Mittelwerte vor, was die Nutzung aktueller Werte ohnehin notwendig macht.

Um Milchminderleistungen zu erkennen, ist nur die Berücksichtigung von Unterschreitungen des unteren Grenzwertes durch den EWMA-Wert notwendig. Bei Unterschreitung des unteren Grenzwertes durch den EWMA wurde der Tag demnach als Alarmtag gewertet. Die entsprechend klassifizierten Tage wurden dann mit den in Abschnitt 4.1.1 definierten negativen Milchmengenabweichungen verglichen und entsprechende Erkennungs- und Fehlerraten sowie Spezifitäten für die einzelnen Modelle berechnet. Die Ergebnisse für die Betriebe 1 bis 3 sind den Tabellen 15 bis 17 zu entnehmen. ER bezeichnet die Erkennungsrate, FR die Fehlerrate und Spez die Spezifität.

Allgemein ist festzustellen, dass die Erkennungsraten wie auch die Fehlerraten und die Spezifität in allen Varianten bei hohem λ , also stärkerer Gewichtung der aktuellen Werte, besser werden. Erwartungsgemäß führt ein niedriger L-Wert zu besseren Erkennungsraten, da auch schwächere Abweichungen als Alarme gemeldet werden, dies geht jedoch gerade bei hohem λ mit hohen Fehlerraten einher, da auch Schwankungen im normalen Bereich häufiger als Abweichungen gemeldet werden. Umgekehrt führt ein hoher L-Wert zu geringen Erkennungsraten, da nur offensichtliche Abweichungen einen Alarm auslösen und entsprechend weniger Falschalarme

Tabelle 15: Klassifikationsergebnisse in % für EWMA-Charts, Betrieb 1

L	λ	0,1	0,3	0,5	0,7	0,9
0,5	ER ^a	24,53	37,09	48,49	56,45	62,15
	FR ^b	95,89	92,77	88,20	83,21	77,76
	Spez ^c	94,03	95,34	96,45	97,15	97,61
1	ER ^a	15,23	16,28	22,85	26,92	29,07
	FR ^b	96,45	92,87	84,38	74,74	63,66
	Spez ^c	94,08	94,88	95,50	95,84	96,02
1,5	ER ^a	9,77	6,10	8,95	11,86	12,73
	FR ^b	97,05	94,70	84,62	70,39	57,80
	Spez ^c	94,11	94,62	94,95	95,16	95,24
2	ER ^a	7,03	2,56	3,31	5,00	5,93
	FR ^b	97,38	95,77	87,55	72,96	60,77
	Spez ^c	94,17	94,59	94,73	94,85	94,91

^a Erkennungsrate

^b Fehlerrate

^c Spezifität

Tabelle 16: Klassifikationsergebnisse in % für EWMA-Charts, Betrieb 2

L	λ	0,1	0,3	0,5	0,7	0,9
0,5	ER ^a	37,63	51,46	57,13	61,55	63,87
	FR ^b	87,50	79,75	72,84	65,16	57,17
	Spez ^c	91,28	93,63	94,68	95,42	95,83
1	ER ^a	26,66	34,26	38,79	41,67	43,18
	FR ^b	86,89	72,47	57,61	43,92	32,67
	Spez ^c	90,98	92,63	93,38	93,81	94,03
1,5	ER ^a	18,67	21,22	25,07	27,33	28,31
	FR ^b	86,57	65,24	45,38	31,58	21,67
	Spez ^c	90,73	91,75	92,28	92,56	92,68
2	ER ^a	12,74	12,14	14,33	16,37	17,16
	FR ^b	86,90	60,74	40,15	26,63	16,99
	Spez ^c	90,50	91,08	91,37	91,58	91,68

^a Erkennungsrate

^b Fehlerrate

^c Spezifität

Tabelle 17: Klassifikationsergebnisse in % für EWMA-Charts, Betrieb 3

L	λ	0,1	0,3	0,5	0,7	0,9
0,5	ER ^a	11,76	32,90	43,36	51,63	57,73
	FR ^b	97,29	92,09	88,62	84,68	80,95
	Spez ^c	94,94	96,19	96,85	97,37	97,75
1	ER ^a	6,97	14,81	20,48	25,27	27,89
	FR ^b	97,26	92,95	80,21	66,57	54,29
	Spez ^c	95,16	95,67	96,17	96,45	96,59
1,5	ER ^a	3,70	6,75	8,93	10,46	11,76
	FR ^b	97,72	88,85	74,69	54,72	32,50
	Spez ^c	95,22	95,60	95,75	95,85	95,92
2	ER ^a	2,61	2,18	3,92	5,23	5,45
	FR ^b	97,71	92,19	67,86	42,86	16,67
	Spez ^c	95,28	95,45	95,56	95,63	95,65

^a Erkennungsrate

^b Fehlerrate

^c Spezifität

auftreten. Die besten Varianten aller Betriebe sind dementsprechend diejenigen mit $\lambda = 0.9$ und $L = 0.5$, jedoch sind die erzielten Erkennungsraten vergleichsweise niedrig und die Fehlerraten für eine praktische Anwendung zu hoch. Es zeigt sich, dass in Betrieb 3 noch wesentlich schlechtere Erkennungsraten als in den Betrieben 1 und 2 erzielt werden. Dies lässt sich durch die in diesem Betrieb geringeren durchschnittlichen Abweichungen der Milchmenge gegenüber dem Vortag als in den Betrieben 1 und 2 erklären (vgl. dazu Tabelle 20 in Abschnitt 4.1.4.1).

Die sich daraus insgesamt ergebende schlechte Verwendbarkeit des EWMA-Charts zur Erkennung von Milchminderungen ist mit dem typischen Verlauf der Laktationskurve mit einem sich über die Zeit verändernden Zielwert zu begründen, während das EWMA-Chart üblicherweise in der industriellen Fertigung bei Produkten mit gleich bleibenden Mittelwerten für das überwachte Merkmal eingesetzt wird. Zusätzlich führen auch die gelegentlich kurzfristig stark schwankenden Milchmengen bei Fehlzuordnungen zu Fehlalarmen. Datenausfälle über mehrere Tage oder krankheitsbedingte Datenlücken beeinträchtigen die Aussagekraft des EWMA-Charts in den nachfolgenden Tagen ebenfalls, da eine Anpassung des Grenzwertes an das neue Niveau erfolgen muss. Der EWMA bewegt sich durch den hohen λ -Wert sofort wieder nahe an der beobachteten Milchmenge, während die Grenzwertkurve stärker geglättet ist und sich erst verzögert auf die aktuelle Laktationsphase einstellt.

4.1.4 Anwendung von Fuzzy Logic

4.1.4.1 Modelle

Zunächst wurde ein allgemeines Modell (im Weiteren mit MM1 bezeichnet) aufgestellt, das bei allen drei Betrieben verwendet wurde. Die Eingabevariablen und die Ausgabevariable mit den jeweiligen Termen und Zugehörigkeitsfunktionen (wie in Abschnitt 3.3.2.2 beschrieben, vgl. Abbildung 5) sind Tabelle 18 zu entnehmen. Es wurden drei Eingangsvariablen (absolute und prozentuale Abweichung der Milchmenge gegenüber dem Vortag, Abweichungsfaktor) und eine Ausgangsvariable (negative Abweichung) vorgesehen. Die Unterteilung der Terme erfolgte bei den Eingabevariablen zur Abweichung der Milchmenge in sieben Stufen, um eine möglichst feine Regelabstimmung zu ermöglichen. Für den Abweichungsfaktor erschienen fünf Abstufungen ausreichend, die Ausgabevariable kommt mit drei Ausprägungen aus. Die entsprechenden Zugehörigkeitsfunktionen sind in den Abbildungen 25 bis 29 dargestellt.

Tabelle 18: Linguistische Variable und die entsprechenden Zugehörigkeitsfunktionen für Modell MM1

Variable	Term	Zugehörigkeitsfunktionstyp	Definitionspunkte
Absolute Abweichung Milchmenge gegenüber Vortag	stark negativ	Z	-15; -5
	negativ	Lambda	-15; -7; -2
	gering negativ	Lambda	-7; -3; 0
	normal	Lambda	-3; 0; 3
	gering positiv	Lambda	0; 3; 7
	positiv	Lambda	2; 7; 15
	stark positiv	S	5; 15
Prozentuale Abweichung Milchmenge gegenüber Vortag	stark negativ	Z	-0.35; -0.1
	negativ	Lambda	-0.25; -0.08; -0.03
	gering negativ	Lambda	-0.1; -0.05; 0
	normal	Lambda	-0.03; 0; 0.3
	gering positiv	Lambda	0; 0.05; 0.1
	positiv	Lambda	0.03; 0.08; 0.25
	stark positiv	S	0.1; 0.35
Abweichungsfaktor	stark negativ	Z	-35; -5
	negativ	Lambda	-25; -5; 0
	normal	Lambda	-5; 0; 5
	positiv	Lambda	0; 5; 25
	stark positiv	S	5; 35
Negative Abweichung	keine	Z	0.2; 0.4
	unsicher	Pi	0.2; 0.4; 0.6; 0.8
	stark	S	0.6; 0.8

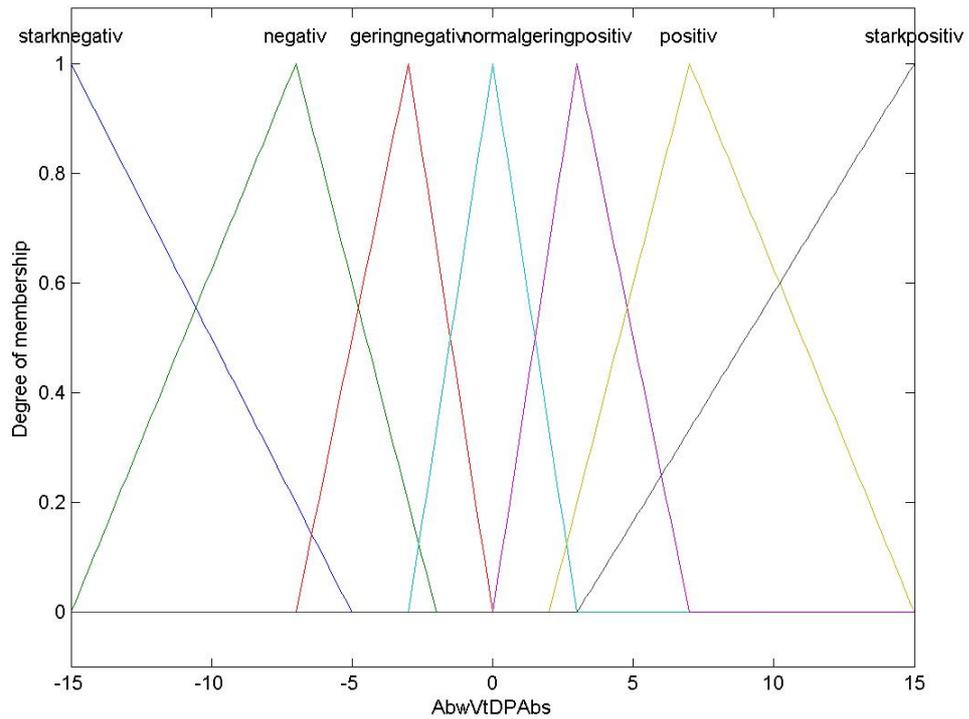


Abbildung 25: Zugehörigkeitsfunktionen der Eingabevariable absolute Abweichung der Milchmenge gegenüber dem Vortag

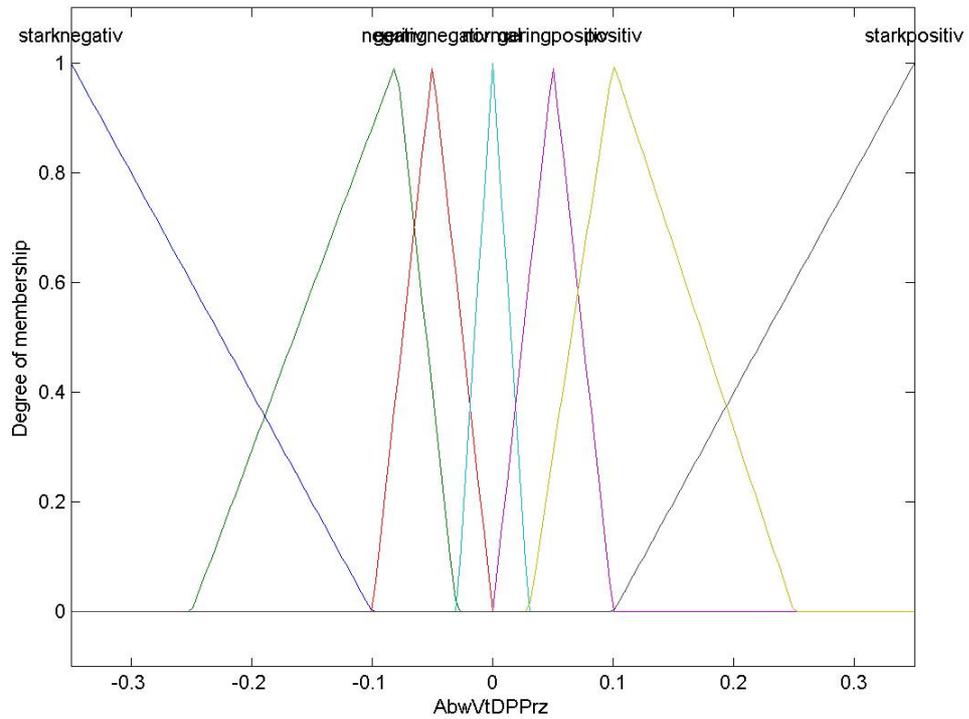


Abbildung 26: Zugehörigkeitsfunktionen der Eingabevariable prozentuale Abweichung der Milchmenge gegenüber dem Vortag

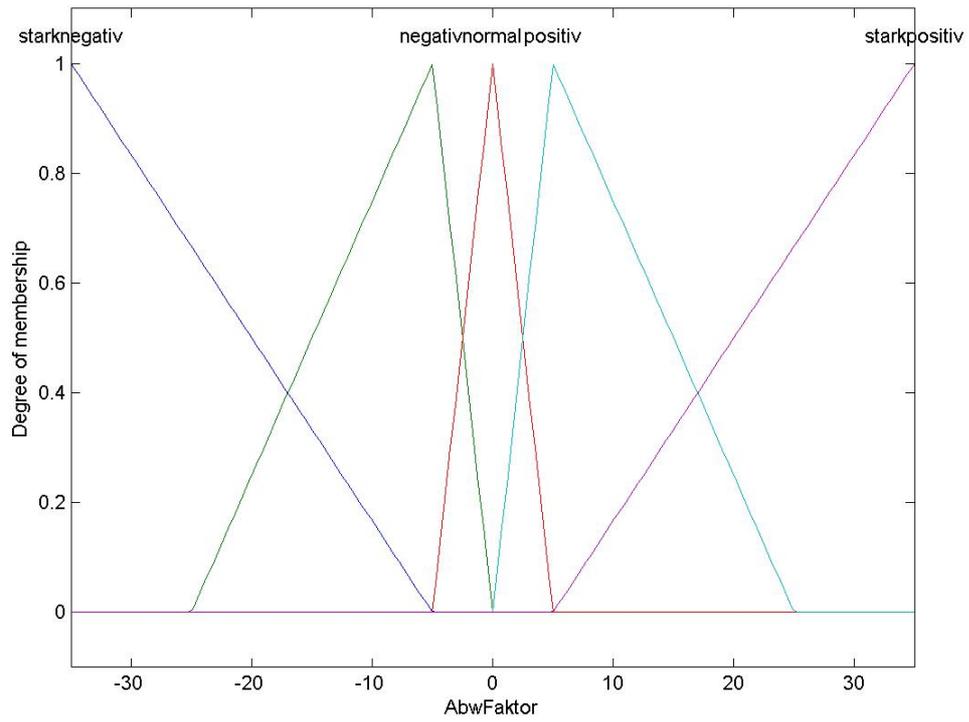


Abbildung 27: Zugehörigkeitsfunktionen der Eingabevariable Abweichungsfaktor

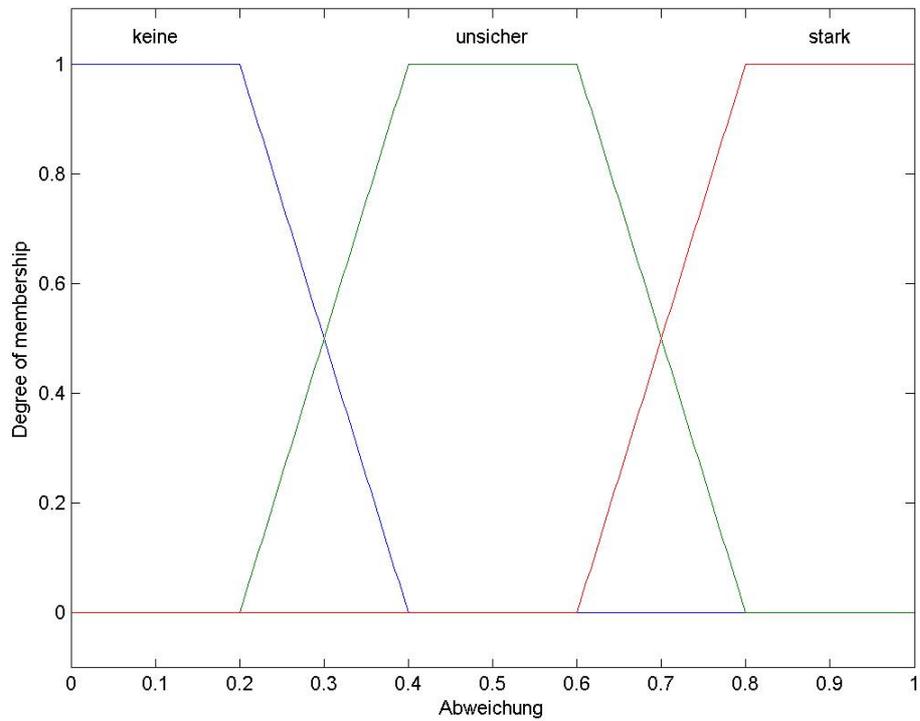


Abbildung 28: Zugehörigkeitsfunktionen der Ausgabevariable negative Abweichung

Im nächsten Schritt wurden die relevanten Entscheidungsregeln aufgestellt (vgl. Tabelle 19). Dabei bezeichnet *AbwAbs* die absolute Abweichung der Milchmenge, *AbwRel* die prozentuale Abweichung, *AF* den Abweichungsfaktor und *NegAbw* die Ausgabevariable für die negative Abweichung.

Tabelle 19: Regelgerüst für Modelle MM1 bis MM4

		AbwAbs^a		AbwRel^b		AF^c	NegAbw^d		
1	Wenn	gering positiv						Dann	keine
2	Wenn	positiv						Dann	keine
3	Wenn	stark positiv						Dann	keine
4	Wenn			gering positiv				Dann	keine
5	Wenn			positiv				Dann	keine
6	Wenn			stark positiv				Dann	keine
7	Wenn					normal		Dann	keine
8	Wenn					positiv		Dann	keine
9	Wenn					stark positiv		Dann	keine
10	Wenn	normal	Und	normal	Und	negativ		Dann	keine
11	Wenn	normal	Und	normal	Und	stark negativ		Dann	keine
12	Wenn	normal	Und	gering negativ	Und	negativ		Dann	keine
13	Wenn	normal	Und	gering negativ	Und	stark negativ		Dann	keine
14	Wenn	normal	Und	negativ	Und	negativ		Dann	unsicher
15	Wenn	normal	Und	negativ	Und	stark negativ		Dann	unsicher
16	Wenn	normal	Und	stark negativ	Und	negativ		Dann	unsicher
17	Wenn	normal	Und	stark negativ	Und	stark negativ		Dann	unsicher
18	Wenn	gering negativ	Und	normal	Und	negativ		Dann	keine
19	Wenn	gering negativ	Und	normal	Und	stark negativ		Dann	keine
20	Wenn	gering negativ	Und	gering negativ	Und	negativ		Dann	unsicher
21	Wenn	gering negativ	Und	gering negativ	Und	stark negativ		Dann	unsicher
22	Wenn	gering negativ	Und	negativ	Und	negativ		Dann	unsicher
23	Wenn	gering negativ	Und	negativ	Und	stark negativ		Dann	unsicher
24	Wenn	gering negativ	Und	stark negativ	Und	negativ		Dann	stark
25	Wenn	gering negativ	Und	stark negativ	Und	stark negativ		Dann	stark
26	Wenn	negativ	Und	normal	Und	negativ		Dann	unsicher
27	Wenn	negativ	Und	normal	Und	stark negativ		Dann	unsicher
28	Wenn	negativ	Und	gering negativ	Und	negativ		Dann	unsicher
29	Wenn	negativ	Und	gering negativ	Und	stark negativ		Dann	unsicher
30	Wenn	negativ	Und	negativ	Und	negativ		Dann	stark
31	Wenn	negativ	Und	negativ	Und	stark negativ		Dann	stark
32	Wenn	negativ	Und	stark negativ	Und	negativ		Dann	stark

Tabelle 19: Regelgerüst für Modelle MM1 bis MM4 (Fortsetzung)

		AbwAbs^a		AbwRel^b		AF^c	NegAbw^d	
33	<i>Wenn</i>	negativ	<i>Und</i>	stark negativ	<i>Und</i>	stark negativ	<i>Dann</i>	stark
34	<i>Wenn</i>	stark negativ	<i>Und</i>	normal	<i>Und</i>	negativ	<i>Dann</i>	unsicher
35	<i>Wenn</i>	stark negativ	<i>Und</i>	normal	<i>Und</i>	stark negativ	<i>Dann</i>	unsicher
36	<i>Wenn</i>	stark negativ	<i>Und</i>	gering negativ	<i>Und</i>	negativ	<i>Dann</i>	stark
37	<i>Wenn</i>	stark negativ	<i>Und</i>	gering negativ	<i>Und</i>	stark negativ	<i>Dann</i>	stark
38	<i>Wenn</i>	stark negativ	<i>Und</i>	negativ	<i>Und</i>	negativ	<i>Dann</i>	stark
39	<i>Wenn</i>	stark negativ	<i>Und</i>	negativ	<i>Und</i>	stark negativ	<i>Dann</i>	stark
40	<i>Wenn</i>	stark negativ	<i>Und</i>	stark negativ	<i>Und</i>	negativ	<i>Dann</i>	stark
41	<i>Wenn</i>	stark negativ	<i>Und</i>	stark negativ	<i>Und</i>	stark negativ	<i>Dann</i>	stark

^a absolute Abweichung der Milchmenge

^b prozentuale Abweichung der Milchmenge

^c Abweichungsfaktor

^d Ausgabevariable für die negative Abweichung

Als Variante wurden dann bei unveränderter Inferenz die Zugehörigkeitsfunktionen an die einzelnen Betriebe angepasst. Als Definitionspunkte für die Zugehörigkeitsfunktionen wurden dabei die Standardabweichungen für die beiden Variablen *AbwAbs* und *AbwRel*, die aus den Datensätzen berechnet wurden, oder entsprechende Vielfache gewählt (vgl. Tabelle 20).

Tabelle 20: Standardabweichungen der beiden Variablen AbwAbs und AbwRel in den Betrieben

	Betrieb 1	Betrieb 2	Betrieb 3
<i>AbwAbs</i> ^a	2.8451	4.3941	1.9762
<i>AbwRel</i> ^b	0.2533	0.2366	0.0852

^a absolute Abweichung der Milchmenge

^b prozentuale Abweichung der Milchmenge

Die sich daraus ergebenden Zugehörigkeitsfunktionen sind in Tabelle 21 dargestellt. Der Wert *s* kennzeichnet dabei, wo jeweils die betriebspezifische Standardabweichung der entsprechenden Variablen verwendet wird. In der Reihenfolge der Betriebe 1 bis 3 werden die Modelle im weiteren MM2 bis MM4 bezeichnet.

Tabelle 21: Linguistische Variable und die entsprechenden Zugehörigkeitsfunktionen für Modelle MM2-MM4

Variable	Term	Zugehörigkeitsfunktionstyp	Definitionspunkte
Absolute Abweichung Milchmenge gegenüber Vortag	stark negativ	Z	-15; -s
	negativ	Lambda	-15; -2s; -0.5 s
	gering negativ	Lambda	-2s; -s; 0
	normal	Lambda	-s; 0; s
	gering positiv	Lambda	0; s; 2s
	positiv	Lambda	0.5s; 2s; 15
	stark positiv	S	s; 15
Prozentuale Abweichung Milchmenge gegenüber Vortag	stark negativ	Z	-0.35; -0.5s
	negativ	Lambda	-s; -0.5s; -0.03
	gering negativ	Lambda	-0.5s; -0.03; 0
	normal	Lambda	-0.03; 0; 0.3
	gering positiv	Lambda	0; 0.03; 0.5s
	positiv	Lambda	0.03; 0.5s; s
	stark positiv	S	0.5s; 0.35
Abweichungsfaktor	stark negativ	Z	-35; -5
	negativ	Lambda	-25; -5; 0
	normal	Lambda	-5; 0; 5
	positiv	Lambda	0; 5; 25
	stark positiv	S	5; 35
Negative Abweichung	keine	Z	0.2; 0.4
	unsicher	Pi	0.2; 0.4; 0.6; 0.8
	stark	S	0.6; 0.8

Aufbauend auf den betriebsspezifischen Modellen wurde in weiteren Varianten der Abweichungsfaktor als Eingabevariable weggelassen um zu untersuchen, wie sich die Verwendung dieser Variable auf die Erkennung von Milchminderungen auswirkt. Die entsprechenden Modelle MM5 bis MM7 verwenden die in Tab. 21 dargestellten Zugehörigkeitsfunktionen, ausgenommen derer für den Abweichungsfaktor. Das entsprechend angepasste Regelgerüst findet sich in Tabelle 22.

Für die Berechnung des Abweichungsfaktors werden die Schätzungen aus dem gemischten linearen Modell verwendet, wobei zunächst alle verfügbaren Daten bis zum jeweilig untersuchten Tag verwendet werden. Je mehr Daten vorhanden sind, umso besser sollte die Anpassung der geschätzten Milchmengen an die tatsächlich beobachteten Milchmengen sein und somit die erhaltenen Vorhersagewerte. Allerdings nimmt mit zunehmender Anzahl von Datensätzen je Tier auch die Berechnungszeit zu, so dass ebenfalls untersucht werden sollte, wie sich die Beschränkung auf eine geringere, gleich bleibende Anzahl von Datensätzen zur Schätzung der Vorhersagewerte auf die Aussagekraft des Abweichungsfaktors auswirkt. Zu diesem Zweck wurden von jedem Modell vier Varianten untersucht, bei denen die Vorhersagewerte mit 60, 90, 120 und 240 Tagen

Tabelle 22: Regelgerüst für Modelle MM5 bis MM7

		AbwAbs^a		AbwRel^b		NegAbw^c
1	<i>Wenn</i>	gering positiv			<i>Dann</i>	keine
2	<i>Wenn</i>	positiv			<i>Dann</i>	keine
3	<i>Wenn</i>	stark positiv			<i>Dann</i>	keine
4	<i>Wenn</i>			gering positiv	<i>Dann</i>	keine
5	<i>Wenn</i>			positiv	<i>Dann</i>	keine
6	<i>Wenn</i>			stark positiv	<i>Dann</i>	keine
7	<i>Wenn</i>	normal	<i>Und</i>	normal	<i>Dann</i>	keine
8	<i>Wenn</i>	normal	<i>Und</i>	gering negativ	<i>Dann</i>	keine
9	<i>Wenn</i>	normal	<i>Und</i>	negativ	<i>Dann</i>	unsicher
10	<i>Wenn</i>	normal	<i>Und</i>	stark negativ	<i>Dann</i>	unsicher
11	<i>Wenn</i>	gering negativ	<i>Und</i>	normal	<i>Dann</i>	keine
12	<i>Wenn</i>	gering negativ	<i>Und</i>	gering negativ	<i>Dann</i>	unsicher
13	<i>Wenn</i>	gering negativ	<i>Und</i>	negativ	<i>Dann</i>	unsicher
14	<i>Wenn</i>	gering negativ	<i>Und</i>	stark negativ	<i>Dann</i>	stark
15	<i>Wenn</i>	negativ	<i>Und</i>	normal	<i>Dann</i>	unsicher
16	<i>Wenn</i>	negativ	<i>Und</i>	gering negativ	<i>Dann</i>	unsicher
17	<i>Wenn</i>	negativ	<i>Und</i>	negativ	<i>Dann</i>	stark
18	<i>Wenn</i>	negativ	<i>Und</i>	stark negativ	<i>Dann</i>	stark
19	<i>Wenn</i>	stark negativ	<i>Und</i>	normal	<i>Dann</i>	unsicher
20	<i>Wenn</i>	stark negativ	<i>Und</i>	gering negativ	<i>Dann</i>	stark
21	<i>Wenn</i>	stark negativ	<i>Und</i>	negativ	<i>Dann</i>	stark
22	<i>Wenn</i>	stark negativ	<i>Und</i>	stark negativ	<i>Dann</i>	stark

^a absolute Abweichung der Milchmenge

^b prozentuale Abweichung der Milchmenge

^c Ausgabevariable für die negative Abweichung

(bzw. 245 Tagen für Betrieb 1 und 281 Tagen für Betrieb 3) vorher verfügbarer Daten berechnet wurden.

Schließlich wurde noch der Grenzwert, ab dem eine Abweichung als erkannt gelten sollte, variiert. Abweichend zum Konfidenzintervall des gemischten linearen Modells oder zu den Kontrollgrenzen des EWMA-Charts liefert das Fuzzy Logic-Modell Werte von 0 bis 1. Man kann daher einen Wert festlegen, bei dessen Überschreiten ein Datensatz als Abweichung gewertet werden soll. Als Mindestwert wurde ein Grenzwert von 0.5 angenommen, außerdem wurden die Erkennungs- und Fehlerraten für Grenzwerte von 0.6 und 0.7 errechnet. Die höheren Grenzwerte dienen dabei auch der Untersuchung des Einflusses der als „unsicher“ klassifizierten Datensätze, die zu Ausgabewerten in diesem Bereich führen können, auf die Bewertungskriterien.

4.1.4.2 Erkennung von Milchminderungen mittels Fuzzy Logic

Die Ergebnisse für Betrieb 1 finden sich in den Tabellen 23 bis 25. ER bezeichnet die Erkennungsraten, FR die Fehlerraten und Spez die Spezifität. Der Wert oben links in den Tabellen gibt den Grenzwert an, die Zeile AF die zugrunde liegende Datenbasis in Tagen. Alle Angaben für Ergebnisse sind Prozentwerte.

Tabelle 23: Auswertung für Betrieb 1 bei Grenzwert von 0.5 (%)

0,5	AF ^a	240	120	90	60
MM1	ER ^b	79.51	79.47	79.09	77.52
	FR ^c	27.59	27.88	27.93	29.55
	Spez ^d	98.49	98.49	98.46	98.35
MM2	ER ^b	84.31	84.54	84.74	82.81
	FR ^c	35.52	35.85	35.41	37.28
	Spez ^d	98.83	98.85	98.86	98.72
MM5	ER ^b	87.86	87.86	87.86	85.82
	FR ^c	46.74	46.82	46.74	47.98
	Spez ^d	99.07	99.07	99.07	98.91

^a für Berechnung des Abweichungsfaktors zugrunde liegende Datenbasis

^b Erkennungsrate

^c Fehlerrate

^d Spezifität

Tabelle 24: Auswertung für Betrieb 1 bei Grenzwert von 0.6 (%)

0,6	AF ^a	240	120	90	60
MM1	ER ^b	58.9	44.74	45.07	44.35
	FR ^c	7.29	11.95	11.87	13.68
	Spez ^d	96.09	96.11	96.13	96.08
MM2	ER ^b	60.22	60.26	60.64	59.98
	FR ^c	17.87	18.04	18.5	19.39
	Spez ^d	97.15	97.16	97.18	97.13
MM5	ER ^b	63.34	63.39	63.34	61.72
	FR ^c	28.25	28.16	28.25	30.09
	Spez ^d	97.35	97.35	97.35	97.23

^a für Berechnung des Abweichungsfaktors zugrunde liegende Datenbasis

^b Erkennungsrate

^c Fehlerrate

^d Spezifität

Die Ergebnisse für Betrieb 2 sind in den Tabellen 26 bis 28 zusammengestellt, die Ergebnisse für Betrieb 3 finden sich in den Tabellen 29 bis 31.

Tabelle 25: Auswertung für Betrieb 1 bei Grenzwert von 0.7 (%)

0,7	AF ^a	240	120	90	60
MM1	ER ^b	51.26	24.28	24.7	23.92
	FR ^c	1.93	6.28	5.95	9.34
	Spez ^d	94.77	94.76	94.79	94.73
	ER ^b	32.09	32.29	32.45	31.55
MM2	FR ^c	6.97	7.36	7.06	9.64
	Spez ^d	95.27	95.28	95.29	95.23
	ER ^b	36.06	36.11	36.06	35.1
MM5	FR ^c	15.25	15.4	15.25	17.51
	Spez ^d	95.52	95.53	95.52	95.45

^a für Berechnung des Abweichungsfaktors zugrunde liegende Datenbasis

^b Erkennungsrate

^c Fehlerrate

^d Spezifität

Tabelle 26: Auswertung für Betrieb 2 bei Grenzwert von 0.5 (%)

0,5	AF ^a	240	120	90	60
MM1	ER ^b	76.35	77.23	77.14	77.25
	FR ^c	15.18	13.93	14.1	13.99
	Spez ^d	97.33	97.43	97.42	97.43
	ER ^b	68.51	69.33	69.23	69.32
MM3	FR ^c	8.27	6.75	6.95	6.88
	Spez ^d	96.51	96.6	96.59	96.6
	ER ^b	75.67	76.64	76.64	76.64
MM6	FR ^c	16.1	15.03	15.03	15.03
	Spez ^d	97.25	97.36	97.36	97.36

^a für Berechnung des Abweichungsfaktors zugrunde liegende Datenbasis

^b Erkennungsrate

^c Fehlerrate

^d Spezifität

Tabelle 27: Auswertung für Betrieb 2 bei Grenzwert von 0.6 (%)

0,6	AF ^a	240	120	90	60
MM1	ER ^b	55.81	56.47	56.58	56.52
	FR ^c	4.71	3.35	3.47	3.38
	Spez ^d	95.18	95.25	95.27	95.26
	ER ^b	50.55	51.16	51.06	51.09
MM3	FR ^c	3.1	1.8	1.94	1.79
	Spez ^d	94.65	94.71	94.7	94.7
	ER ^b	55.31	56	56	56
MM6	FR ^c	7.7	6.55	6.55	6.55
	Spez ^d	95.12	95.2	95.2	95.2

^a für Berechnung des Abweichungsfaktors zugrunde liegende Datenbasis

^b Erkennungsrate

^c Fehlerrate

^d Spezifität

Tabelle 28: Auswertung für Betrieb 2 bei Grenzwert von 0.7 (%)

0,7	AF ^a	240	120	90	60
MM1	ER ^b	44.72	45.33	45.3	45.18
	FR ^c	2.53	1.26	1.37	1.25
	Spez ^d	94.06	94.12	94.12	94.11
	ER ^b	42.11	42.7	42.68	42.56
MM3	FR ^c	2.02	0.73	0.8	0.75
	Spez ^d	93.79	93.86	93.85	93.84
	ER ^b	47.62	48.23	48.23	48.23
MM6	FR ^c	5.32	4.12	4.12	4.12
	Spez ^d	94.34	94.41	94.41	94.41

^a für Berechnung des Abweichungsfaktors zugrunde liegende Datenbasis

^b Erkennungsrate

^c Fehlerrate

^d Spezifität

Tabelle 29: Auswertung für Betrieb 3 bei Grenzwert von 0.5 (%)

0,5	AF ^a	240	120	90	60
MM1	ER ^b	59.34	7.99	58.22	55.71
	FR ^c	20.11	94.8	21.78	21.54
	Spez ^d	97.74	94.37	97.56	97.42
	ER ^b	78.22	79.45	79.22	77.4
MM4	FR ^c	50.33	47.19	47.02	48.32
	Spez ^d	98.73	98.74	98.73	98.62
	ER ^b	90.87	92.47	92.47	92.47
MM7	FR ^c	60.89	58.72	58.72	58.72
	Spez ^d	99.44	99.52	99.52	99.52

^a für Berechnung des Abweichungsfaktors zugrunde liegende Datenbasis

^b Erkennungsrate

^c Fehlerrate

^d Spezifität

Tabelle 30: Auswertung für Betrieb 3 bei Grenzwert von 0.6 (%)

0,6	AF ^a	240	120	90	60
MM1	ER ^b	25.93	3.2	23.52	23.97
	FR ^c	3.85	95.54	3.74	3.67
	Spez ^d	95.99	94.36	95.66	95.68
	ER ^b	59.75	57.99	59.82	58.68
MM4	FR ^c	22.58	23.49	22.02	21.17
	Spez ^d	97.76	97.54	97.65	97.59
	ER ^b	71.78	73.29	73.29	73.29
MM7	FR ^c	33.59	32.7	32.7	32.7
	Spez ^d	98.4	98.41	98.41	98.41

^a für Berechnung des Abweichungsfaktors zugrunde liegende Datenbasis

^b Erkennungsrate

^c Fehlerrate

^d Spezifität

Tabelle 31: Auswertung für Betrieb 3 bei Grenzwert von 0.7 (%)

0,7	AF ^a	240	120	90	60
MM1	ER ^b	11.83	1.14	9.59	10.05
	FR ^c	1.72	97.08	0	0
	Spez ^d	95.27	94.34	94.91	94.93
	ER ^b	28.22	29	28.08	29
MM4	FR ^c	8.72	7.97	6.82	5.93
	Spez ^d	96.11	95.95	95.9	95.95
	ER ^b	37.55	36.76	36.76	36.76
MM7	FR ^c	16.97	15.71	15.71	15.71
	Spez ^d	96.59	96.37	96.37	96.37

^a für Berechnung des Abweichungsfaktors zugrunde liegende Datenbasis

^b Erkennungsrate

^c Fehlerrate

^d Spezifität

Die höchsten Erkennungsraten erzielen hierbei jeweils die Modelle MM5, MM6, und MM7, die speziell an die Betriebe angepasst sind und den Abweichungsfaktor nicht berücksichtigen. Die Modelle MM2, MM3 und MM4 weisen diesen Modellen gegenüber geringere Erkennungsraten, aber auch geringere Fehlerraten auf. Das allgemeine Modell MM1 erzielt bei Betrieb 1 und Betrieb 2 vergleichsweise gute Ergebnisse, während bei Betrieb 3 schlechte Erkennungsraten zu beobachten sind. Das gilt vor allem bei MM1 in der Variante mit 120 Tagen Datenvorlauf. Die geringeren Fehlerraten bei höheren Grenzwerten werden durch die wesentlich geringeren Erkennungsraten nicht gerechtfertigt. Das allgemeine Modell MM1 erzielt sogar bessere Erkennungsraten als die speziell auf den Betrieb abgestimmten Modelle, im Fall von MM3 jedoch auf Kosten von höheren Fehlerraten. Hinsichtlich der Frage des günstigsten Datenvorlaufs zur Bestimmung des Abweichungsfaktors lässt sich keine eindeutige Tendenz erkennen: in Betrieb 1 erzielen die Modelle mit den meisten verfügbaren Daten die besten Ergebnisse, in Betrieb 2 die schlechtesten. Allerdings ist die Datenabdeckung des untersuchten Zeitraums in Betrieb 1 auch schlechter als in Betrieb 2. Die häufiger auftretenden Datenausfälle müssen daher dort durch einen größeren Zeitraum kompensiert werden, während bei einer weniger lückenhaften Datenbasis ein geringerer Datenvorlauf ausreichend erscheint, um ähnlich gute Ergebnisse zu erzielen. Die Spezifität erweist sich bei allen Modellen durchgängig als hoch, was aber auch durch die große Anzahl an Datensätzen im Verhältnis zu den Datensätzen mit negativen Milchmengenabweichungen zu erwarten ist. Generell ist zu erkennen, dass bei Betrieb 1 die Erkennungsraten im wünschenswerten Bereich von über 80% liegen, das Ergebnis jedoch unter den hohen Fehlerraten leidet. Bei Betrieb 2 dagegen finden sich hervorragende Fehlerraten von unter 15%, aber entsprechend schlechtere Erkennungsraten. Bei Betrieb 3 machen sich die im Vergleich zu den beiden anderen Betrieben

geringeren Standardabweichungen für die beiden Milchmengenabweichungsmerkmale in höheren Fehlerraten bei den betriebspezifischen Modellen bemerkbar.

4.1.5 Anwendung künstlicher neuronaler Netze

4.1.5.1 Modelle

Zur Erkennung von negativen Milchmengenabweichungen wurde ein Netz vom Typ Feedforward Backpropagation erstellt. Die Eingabeschicht des neuronalen Netzes bestand aus drei Neuronen, jeweils ein Neuron für die Variablen absolute Abweichung der Milchmenge gegenüber dem Vortag, prozentuale Abweichung der Milchmenge gegenüber dem Vortag und Abweichungsfaktor von beobachteten und vorhergesagten Milchmengen. Die Ausgabeschicht beinhaltet ein einzelnes Neuron, das nach dem Training in der Lage sein soll, den Status für jeden Eingabevektor anzugeben. Anhand des Ausgabewertes soll entschieden werden, ob eine negative Milchmengenabweichung vorliegt oder nicht. In einer versteckten Schicht wurden weitere Neuronen vorgesehen. Es wurden Netze mit drei, vier, fünf und sechs Neuronen in der versteckten Schicht untersucht, um die bestgeeignete Anzahl von Neuronen zu ermitteln.

4.1.5.2 Modellauswahl

Zum Zweck der Auswahl des bestgeeigneten Modells wurde der Datensatz von Betrieb 1 mit 60 Tagen Vorlaufzeit zur Berechnung der für den Abweichungsfaktor notwendigen Vorhersagewerte verwendet. Der Gesamtdatensatz wurde in acht Teildatensätze aufgeteilt. Zu diesem Zweck wurde mittels des in der Access-Datenbank nutzbaren Visual Basic eine gleichverteilte Zufallszahl von 1 bis 8 für jeden einzelnen Datensatz (je Tier und Tag) erzeugt und diese Datensätze entsprechend der zugewiesenen Zahl in entsprechende Tabellen sortiert. Jeweils sieben Teildatensätze wurden benutzt, um das Netzwerk zu trainieren. Mit dem trainierten Netzwerk wurde ein Testdurchlauf mit dem verbleibenden Datensatz durchgeführt. Die Ausgabewerte des Testdurchlaufs wurden dann mit den vorgegebenen Zielwerten für negative Milchmengenabweichungen verglichen und bei verschiedenen Grenzwerten den Vorgaben in Abschnitt 4.1.1 entsprechend klassifiziert. Die Zielwerte waren dabei 0 und 1 für Datensätze ohne und mit vorliegenden Milchminderungen gemäß der Definition in Abschnitt 4.3. Entsprechend wurden für die Neuronen log-sigmoide Aktivierungsfunktionen gewählt, um Ausgabewerte im Wertebereich von 0 bis 1 zu erhalten.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (17)$$

Für das Training des Netzwerkes wurde der Levenberg-Marquardt-Algorithmus genutzt, der in der verwendeten MatLab-Software die Standardeinstellung ist und zu einer schnellen Optimierung des Gradientenabstiegs führt (DEMUTH und BEALE, 2004, S. 5-28ff). Zur Überprüfung der Anpassung des künstlichen neuronalen Netzes wurden die Trainingsdatensätze klassifiziert, der jeweils verbleibende Validierungsdatensatz dient zur Bewertung der Generalisierungsfähigkeit des neuronalen Netzes. Als mögliche Grenzwerte zur Abgrenzung normaler Gemelke von einer Milchmindermenge wurden Werte von 0.1 bis 0.9 in Schritten von 0.1 überprüft, um Aufschlüsse über die Trennschärfe des Netzwerkes zu bekommen. Die Ergebnisse sind den Tabellen 32 bis 35 zu entnehmen. Angegeben sind jeweils die über alle acht Trainings- bzw. Validierungsdatensätze gemittelten Werte.

Tabelle 32: Klassifikationsergebnisse für ein KNN mit 3 Neuronen in der versteckten Schicht (%)

3 Neuronen	Grenzwert	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Trainings- datensätze	ER ^a	99.2	97.6	94.5	90.5	85.9	80.4	72.2	60.0	31.6
	FR ^b	35.0	28.2	24.3	21.4	18.7	15.7	13.1	9.6	5.4
	Spez ^c	99.9	99.8	99.6	99.3	99.0	98.6	98.0	97.1	95.3
Validierungs- datensatz	ER ^a	99.3	97.8	94.7	89.8	85.8	79.6	71	59.7	30.4
	FR ^b	35.1	27.8	24	21.8	18.3	16.4	13.3	9.65	6.17
	Spez ^c	99.9	99.8	99.6	99.2	99	98.5	97.9	97.1	95.1

^a Erkennungsrate

^b Fehlerrate

^c Spezifität

Tabelle 33: Klassifikationsergebnisse für ein KNN mit 4 Neuronen in der versteckten Schicht (%)

4 Neuronen	Grenzwert	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Trainings- datensätze	ER ^a	98.5	97.0	93.8	89.8	85.2	80.0	71.9	59.5	29.1
	FR ^b	34.2	27.8	24.3	21.4	18.8	15.8	12.9	9.5	3.8
	Spez ^c	99.9	99.8	99.5	99.2	98.9	98.5	98.0	97.1	95.1
Validierungs- datensatz	ER ^a	99.6	98	93.8	90	85.6	80.2	72.1	59.5	28.5
	FR ^b	33.7	27	23.7	21.2	17.8	15.7	12.4	8.5	3.2
	Spez ^c	100	99.8	99.5	99.3	98.9	98.6	98	97.1	95

^a Erkennungsrate

^b Fehlerrate

^c Spezifität

Tabelle 34: Klassifikationsergebnisse für ein KNN mit 5 Neuronen in der versteckten Schicht (%)

5 Neuronen	Grenzwert	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Trainings- datensätze	ER ^a	99.5	97.8	94.3	89.8	84.1	78.0	69.6	57.6	35.2
	FR ^b	36.5	27.8	23.0	19.7	16.7	13.7	10.6	7.6	4.0
	Spez ^c	100.0	99.8	99.6	99.2	98.8	98.4	97.8	97.0	95.5
Validierungs- datensatz	ER ^a	99.3	97.7	93.7	89	84.5	78.8	69.6	57.5	35.2
	FR ^b	36.3	28	23.2	20.1	16.7	13.8	11.4	9.02	4.57
	Spez ^c	99.9	99.8	99.5	99.2	98.9	98.5	97.8	97	95.5

^a Erkennungsrate

^b Fehlerrate

^c Spezifität

Tabelle 35: Klassifikationsergebnisse für ein KNN mit 6 Neuronen in der versteckten Schicht (%)

6 Neuronen	Grenzwert	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Trainings- datensätze	ER ^a	99.6	97.3	93.1	87.0	80.3	72.7	64.6	53.4	33.7
	FR ^b	42.9	31.5	24.3	19.8	16.2	13.0	9.9	6.8	3.5
	Spez ^c	100.0	99.8	99.5	99.1	98.6	98.1	97.5	96.7	95.3
Validierungs- datensatz	ER ^a	99.4	96.8	93.4	87.5	79.3	72.6	64.7	54.5	34.4
	FR ^b	43.9	32.7	24.3	19.7	16.6	13.6	10.2	6.52	3.57
	Spez ^c	100	99.8	99.5	99.1	98.5	98	97.5	96.8	95.4

^a Erkennungsrate

^b Fehlerrate

^c Spezifität

Die Variante mit fünf Neuronen in der versteckten Schicht erzielte gute Erkennungsraten bei gleichzeitig möglichst geringen Fehlerraten und hoher Spezifität und wurde daher für die weiteren Modelle verwendet. Ausschlaggebend für diese Entscheidung waren vor allem die Ergebnisse für die Grenzwerte 0.4 bis 0.6. Generell sind die Unterschiede der Varianten jedoch sehr gering, und auch die Trainings- und Validierungsdatensätze führen zu vergleichbaren Ergebnissen hinsichtlich der Anpassungsfähigkeit (d.h. Durchlaufen eines trainierten neuronalen Netzes mit den Trainingsdaten und Vergleich des Ergebnisses mit den Zielwerten) und der Generalisierungsfähigkeit (d.h. Durchlaufen eines trainierten neuronalen Netzes mit unbekanntem Daten und Vergleich des Ergebnisses mit den Zielwerten) der Netzwerke.

4.1.5.3 Erkennung von Milchminderungen mittels künstlicher neuronaler Netze

Analog zur Erkennung von Milchmengenabweichungen mittels Fuzzy Logic wurden nun Varianten des künstlichen neuronalen Netzes für die drei Betriebe und die vier unterschiedlichen Zeiträume zur Berechnung des Abweichungsfaktors berechnet. Verwendet wurde das Feedforward

Backpropagation-Netzwerk mit drei Neuronen in der Eingangsschicht, fünf Neuronen in der verdeckten Schicht und einem Neuron in der Ausgabeschicht. Das Training wurde mittels Levenberg-Marquardt-Backpropagation durchgeführt, Anpassung der Gewichte erfolgte mit Gradientenabstieg mit Impulsterm zur Verstärkung von Gewichtsänderungen mit gleichem Trend (bzw. Abschwächung bei widersprüchlichem Trend). Ziel war die Minimierung des mittleren quadratischen Fehlers der Differenz zwischen Ausgabeneuronen und Zielwerten. Diese Einstellungen entsprechen den Standardeinstellungen der Neural Network Toolbox (Version 4) aus dem Mathematiksoftwarepaket MatLab (Version 7.1.0.246 (R14) Service Pack 3). Als Aktivierungsfunktionen wurden abweichend von den Standardeinstellungen log-sigmoide Funktionen benutzt. Die Ergebnisse für die drei Betriebe sind in den Tabellen 36 bis 38 dargestellt. Auch hier bezeichnet ER die Erkennungsraten, FR die Fehlerraten und Spez die Spezifität. Alle Angaben für Ergebnisse sind Prozentwerte.

Allen Modellen gemein ist die hohe Spezifität. Auch hier ist dies wieder durch die geringe Anzahl von Ereignissen im Verhältnis zur Gesamtzahl betrachteter Datensätze zu erklären. Die durchgängig niedrigen Fehlerraten schon bei geringen Grenzwerten deuten auf eine gute Trennschärfe hin. Gegenüber Betrieb 2 sind die Erkennungsraten bei hohen Grenzwerten in den Betrieben 1 und 3 vergleichsweise schlecht. Daraus lässt sich schließen, dass in diesen Betrieben gegenüber Betrieb 2 weniger eindeutige bzw. auffällige negative Milchmengenabweichungen auftreten. Dies bestätigt die in Abschnitt 4.1.4.1 dargestellten Unterschiede in den mittleren Milchmengenabweichungen

Tabelle 36: Klassifikationsergebnisse für ein KNN mit 60, 90, 120 oder 281 Tagen Vorlaufzeit zur Berechnung des Abweichungsfaktors, Betrieb 1 (%)

AF ^a	Grenzwert	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
60 Tage Datenvorlauf	Er ^b	99.3	97.7	93.7	89	84.5	78.8	69.6	57.5	35.2
	FR ^c	36.3	28	23.2	20.1	16.7	13.8	11.4	9.02	4.57
	Spez ^d	99.9	99.8	99.5	99.2	98.9	98.5	97.8	97	95.5
90 Tage Datenvorlauf	Er ^b	95.6	92.7	87.9	82.5	76.4	67.2	54.1	29.5	5.9
	FR ^c	38.1	29.8	24.6	21.2	17.5	14.3	11.2	6.3	11.9
	Spez ^d	99.6	99.4	99.1	98.7	98.3	97.6	96.7	95.1	93.6
120 Tage Datenvorlauf	Er ^b	99.3	96.8	92.2	86.1	79.4	67.3	37.6	0.6	0.0
	FR ^c	34.3	26.5	22.6	19.3	15.5	11.6	3.8	0.0	0.0
	Spez ^d	99.9	99.8	99.4	99.0	98.5	97.6	95.6	93.2	93.2
281 Tage Datenvorlauf	Er ^b	99.8	97.7	93.6	89.0	83.4	75.8	64.8	42.5	3.0
	FR ^c	34.1	26.9	23.0	19.9	17.0	14.2	10.1	3.9	0.0
	Spez ^d	100.0	99.8	99.5	99.2	98.8	98.3	97.5	96.0	93.4

^a für Berechnung des Abweichungsfaktors zugrunde liegende Datenbasis

^b Erkennungsrate

^c Fehlerrate

^d Spezifität

Tabelle 37: Klassifikationsergebnisse für ein KNN mit 60, 90, 120 oder 240 Tagen Vorlaufzeit zur Berechnung des Abweichungsfaktors, Betrieb 2 (%)

AF ^a	Grenzwert	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
60 Tage Datenvorlauf	Er ^b	98.6	97.5	95.4	91.6	88.7	84.9	80.0	73.2	61.6
	FR ^c	27.9	21.3	18.4	15.4	12.9	10.8	8.2	5.9	3.6
	Spez ^d	99.8	99.7	99.5	99.0	98.7	98.3	97.8	97.0	95.8
90 Tage Datenvorlauf	Er ^b	98.2	97.3	94.2	90.7	85.9	79.5	71.7	62.9	49.6
	FR ^c	31.8	28.2	24.6	21.6	18.2	14.6	11.2	8.4	5.2
	Spez ^d	99.8	99.7	99.3	98.9	98.4	97.7	96.9	95.9	94.6
120 Tage Datenvorlauf	Er ^b	98.3	97.3	95.2	91.6	88.8	85.5	80.4	73.6	62.3
	FR ^c	27.8	21.5	19.0	15.8	13.3	11.2	8.7	6.3	3.9
	Spez ^d	99.8	99.7	99.4	99.0	98.7	98.4	97.8	97.1	95.9
240 Tage Datenvorlauf	Er ^b	99.5	98.3	96.3	93.1	89.9	86.6	81.5	74.2	61.8
	FR ^c	26.3	21.0	17.9	14.9	12.1	10.1	7.7	5.0	2.3
	Spez ^d	95.9	97.0	97.6	98.1	98.6	98.9	99.2	99.6	99.8

^a für Berechnung des Abweichungsfaktors zugrunde liegende Datenbasis

^b Erkennungsrate

^c Fehlerrate

^d Spezifität

Tabelle 38: Klassifikationsergebnisse für ein KNN mit 60, 90, 120 oder 245 Tagen Vorlaufzeit zur Berechnung des Abweichungsfaktors, Betrieb 3 (%)

AF ^a	Grenzwert	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
60 Tage Datenvorlauf	Er ^b	91.1	83.8	74.3	65.7	51.8	35.8	13.9	5.4	3.0
	FR ^c	37.2	26.9	20.5	16.4	14.2	14.2	22.6	26.3	13.9
	Spez ^d	99.5	99.0	98.5	98.0	97.2	96.3	95.2	94.7	94.6
90 Tage Datenvorlauf	Er ^b	98.2	88.7	78.4	66.9	49.4	26.5	6.7	2.4	0.5
	FR ^c	38.0	26.4	21.5	17.0	11.3	9.2	11.9	5.2	0.0
	Spez ^d	99.9	99.3	98.7	98.1	97.1	95.8	94.8	94.6	94.5
120 Tage Datenvorlauf	Er ^b	97.5	91.5	81.4	73.0	59.9	45.4	8.3	0.0	0.0
	FR ^c	38.4	27.1	23.5	17.6	13.2	8.6	1.1	0.0	0.0
	Spez ^d	99.8	99.5	98.9	98.4	97.7	96.9	94.9	94.5	94.5
245 Tage Datenvorlauf	Er ^b	97.6	91.0	82.0	71.1	57.3	41.4	7.9	0.9	0.4
	FR ^c	38.9	27.2	22.4	18.3	12.2	7.0	3.1	5.0	0.0
	Spez ^d	99.9	99.5	99.0	98.4	97.7	96.8	95.1	94.8	94.7

^a für Berechnung des Abweichungsfaktors zugrunde liegende Datenbasis

^b Erkennungsrate

^c Fehlerrate

^d Spezifität

gegenüber dem Vortag in den Betrieben. Insgesamt erweist sich bei den Betrieben 1 und 2 ein Grenzwert von 0.4 oder 0.5 als günstig, während er bei Betrieb 3 im Bereich von 0.2 oder 0.3

gewählt werden sollte, um ein bei vergleichbares Ergebnis bei den Erkennungsraten zu erzielen. Dabei muss in Betrieb 3 jedoch eine höhere Fehlerrate in Kauf genommen werden.

Vergleicht man die Modelle hinsichtlich des verfügbaren Datenvorlaufs zur Berechnung des Abweichungsfaktors, so ist keine eindeutige Tendenz zu einem bestimmten optimal benötigten Zeitraum erkennbar. Der erwarteten höheren Genauigkeit der Schätzungen bei größerer Zahl verfügbarer Datensätze steht die höhere Aktualität und Relevanz der Datensätze bei Beschränkung auf einen kürzeren Zeitraum gegenüber. In diesem Fall ist die Entscheidung zugunsten des weniger aufwändigen Modells mit weniger Datensätzen und dadurch geringerer Berechnungszeit zu treffen. Speziell bei einem Grenzwert von 0.5 zur Unterscheidung der normalen Gemelke von Gemelken mit Abweichungen erzielt dieses Modell mit 60 Tagen Datenvorlauf zur Berechnung des Abweichungsfaktors in den Betrieben 1 und 2 auch die besten Erkennungsraten bei gleichzeitig möglichst geringen Fehlerraten, während in Betrieb 3 bei einem Grenzwert von 0.3 die Wahl auf das Kompletmodell fallen würde.

4.2 Nutzung der elektrischen Leitfähigkeit der Milch

Die elektrische Leitfähigkeit ist ein Indikator für die Konzentration von Elektrolyten in der Milch (WENDT et al., 1998, S. 79f). Neben Schwankungen in Abhängigkeit von beispielsweise Laktation, Laktationsstadium oder Fettgehalt ist vor allem auch bei Eutererkrankungen eine Veränderung zu erwarten, da es im Zuge einer Euterentzündung zu einem Anstieg der Konzentration von Natrium- und Chloridionen und einer Verringerung der Laktosekonzentration kommt (KÖHLER, 2003, S. 18f). In Betrieb 1 stehen Werte für die Leitfähigkeit je Euterviertel zur Verfügung. Die Leitfähigkeit wird im AMS zu Beginn im Hauptmelk und zum Ende des Melkvorgangs erfasst und danach gemittelt. Die höchste Aussagekraft ist für Leitfähigkeitswerte aus dem Vorgemelk gegeben (WOOLFORD et al., 1998), die jedoch nicht verfügbar sind. Die Durchschnittswerte je Euterviertel sollen insbesondere hinsichtlich der Möglichkeit zur frühzeitigen Erkennung von Euterkrankheiten sowohl einzeln als auch in Kombination mit dem Merkmal Milchmenge untersucht werden.

4.2.1 Erkennung von Eutererkrankungen mittels Fuzzy Logic

Die Kombination mit der Tagesmilchmenge erfordert eine Anpassung der Leitfähigkeitswerte aus den Einzelmelken. Daher wurde für jedes Viertel ein Tagesmittelwert berechnet. Diese viertelweisen Leitfähigkeitswerte wurden weiter zu einem Gesamtleitfähigkeitswert über alle Euterviertel gemittelt. Neben diesem mittleren Leitfähigkeitswert des Gesamteuters ist auch die

Differenz der Leitfähigkeitswerte der einzelnen Euterviertel gegenüber dem Gesamteuterdurchschnitt ein wichtiges Merkmal für die Erkennung einer Eutererkrankung. In der Regel sind zunächst nicht alle Euterviertel gleichmäßig betroffen, so dass das kranke Viertel durch einen höheren Leitfähigkeitswert identifiziert werden kann. Zu diesem Zweck wurde außerdem die Differenz des Viertels mit der höchsten Leitfähigkeit zur Durchschnittsleitfähigkeit des Gesamteuters berechnet.

Während des untersuchten Zeitraums wurden in Betrieb 1 59 Eutererkrankungen festgestellt. Darunter waren vier Fälle von Strichverletzungen, die spontan auftreten und bei der Erkennung daher nicht berücksichtigt werden, während es sich bei den restlichen 55 Fällen um Mastitiden handelte. Für den ersten dokumentierten Erkrankungstag liegen dabei allerdings in allen Fällen schon keine Daten mehr vor. Wenn man Krankheitsmeldungen innerhalb weniger Tage pro Tier als nur einen Krankheitsfall auffasst, verbleiben 23 Mastitisfälle. Von diesen liegen in 20 Fällen auch Leitfähigkeitswerte für die Tage vor der Erkrankung vor, weshalb diese als Fallbeispiele für die Krankheitserkennung mit Hilfe von Leitfähigkeitswerten in Frage kommen. Davon sollen im Folgenden drei Beispiele exemplarisch eingehend untersucht werden. Zu diesem Zweck werden die verfügbaren Leitfähigkeitswerte und Milchmengen in der Woche vor Feststellung der Erkrankung verwendet.

Als Methode wurde nur ein Fuzzy Logic-System gewählt, da keine brauchbaren Zielwerte für das Trainieren eines neuronalen Netzes (Krankheitstage mit verfügbaren Werten für Milchmenge und Leitfähigkeit) vorliegen.

4.2.1.1 Modelle

Für ein gesundes Euterviertel bewegen sich die Leitfähigkeitswerte zwischen 4.5 und 5.9 mS/cm bei 20°C. Die Differenz der Viertel sollte dabei weniger als 0.5 mS/cm betragen (WENDT et al., 1998, S. 81), größere Werte sind als eine abnorme Abweichung anzusehen. In einem ersten Modell sollte zunächst nur die Leitfähigkeit betrachtet werden. Entsprechend wurden für die Merkmale durchschnittliche Leitfähigkeit (*AvgLF*) und maximale Differenz der Leitfähigkeit eines Viertels zum Euterdurchschnitt (*MaxDiff*) sowie für die Ausgabevariable Erkrankungswahrscheinlichkeit (*EWSK*) Terme erstellt (Tabelle 39). Dabei ist *EWSK* keine Wahrscheinlichkeit im statistischen Sinn, sondern nur als Alarmwert für eine mögliche Erkrankung zu interpretieren. Das Modell wird im Weiteren als LF1 bezeichnet.

Tabelle 39: Linguistische Variable und die entsprechenden Zugehörigkeitsfunktionen für Modell LF1

Variable	Term	Zugehörigkeitsfunktionstyp	Definitionspunkte
Durchschnittliche Leitfähigkeit	nicht erhöht	Z	5.4; 5.9
	erhöht	Lambda	5.4; 5.9; 6.4
	hoch	S	5.9; 6.4
Maximale Differenz zwischen Vierteln und Gesamteuter	gering	Z	0; 0.25
	erhöht	Lambda	0; 0.25; 0.5
	hoch	Pi	0.25; 0.5; 1; 1.25
	sehr hoch	S	1; 1.25
Erkrankungswahrscheinlichkeit	minimal	Z	0; 0.2
	niedrig	Pi	0; 0.2; 0.25; 0.45
	mittel	Pi	0.25; 0.45; 0.55; 0.75
	hoch	S	0.55; 0.75

Aus der Kombination der einzelnen Terme ergibt sich die in Tabelle 40 dargestellte entsprechende Fuzzy-Inferenz.

Tabelle 40: Regelgerüst für Modell LF1

		AvgLF ^a		MaxDiff ^b		EWSK ^c
1	Wenn	nicht erhöht	Und	gering	Dann	minimal
2	Wenn	nicht erhöht	Und	erhöht	Dann	niedrig
3	Wenn	nicht erhöht	Und	hoch	Dann	mittel
4	Wenn	nicht erhöht	Und	sehr hoch	Dann	hoch
5	Wenn	erhöht	Und	gering	Dann	minimal
6	Wenn	erhöht	Und	erhöht	Dann	niedrig
7	Wenn	erhöht	Und	hoch	Dann	mittel
8	Wenn	erhöht	Und	sehr hoch	Dann	hoch
9	Wenn	hoch	Und	gering	Dann	mittel
10	Wenn	hoch	Und	erhöht	Dann	hoch
11	Wenn	hoch	Und	hoch	Dann	hoch
12	Wenn	hoch	Und	sehr hoch	Dann	hoch

^a durchschnittliche Leitfähigkeit

^b Differenz des Euterviertels mit der höchsten Leitfähigkeit zum Euterdurchschnitt

^c Erkrankungswahrscheinlichkeit

Allerdings lässt sich die Anzahl der Regeln durch Verwendung anderer Logik-Operatoren verringern und damit die Übersichtlichkeit durch entsprechende Anpassungen verbessern, indem mehrere Terme einer Variablen zusammengefasst werden, wenn sie zum selben Ergebnis führen

(Tabelle 41). Dies kann vor allem bei einer höheren Anzahl von Termen je Variable und Nutzung weiterer zusätzlicher Eingabevariablen vorteilhaft sein.

Tabelle 41: Vereinfachtes Regelgerüst für Modell LF1

		AvgLF^a		MaxDiff^b		EWSK^c
1	<i>Wenn Nicht</i>	hoch	<i>Und</i>	gering	<i>Dann</i>	minimal
2	<i>Wenn Nicht</i>	hoch	<i>Und</i>	erhöht	<i>Dann</i>	niedrig
3	<i>Wenn Nicht</i>	hoch	<i>Und</i>	hoch	<i>Dann</i>	mittel
4	<i>Wenn Nicht</i>	hoch	<i>Und</i>	sehr hoch	<i>Dann</i>	hoch
5	<i>Wenn</i>	hoch	<i>Und</i>	gering	<i>Dann</i>	mittel
6	<i>Wenn</i>	hoch	<i>Und Nicht</i>	gering	<i>Dann</i>	hoch

^a durchschnittliche Leitfähigkeit

^b Differenz des Eutervierts mit der höchsten Leitfähigkeit zum Euterdurchschnitt

^c Erkrankungswahrscheinlichkeit

Tabelle 42: Linguistische Variablen und die entsprechenden Zugehörigkeitsfunktionen für Modell LF2

Variable	Term	Zugehörigkeitsfunktionstyp	Definitionspunkte
Durchschnittliche Leitfähigkeit	nicht erhöht	Z	5.4; 5.9
	erhöht	Lambda	5.4; 5.9; 6.4
	hoch	S	5.9; 6.4
Maximale Differenz zwischen Vierteln und Gesamteuter	gering	Z	0; 0.25
	erhöht	Lambda	0; 0.25; 0.5
	hoch	Pi	0.25; 0.5; 1; 1.25
	sehr hoch	S	1; 1.25
Prozentuale Änderung zum Vortagemittel	sinkend	Z	-0.1; 0
	normal	Lambda	-0.1; 0; 0.1
	steigend	S	0; 0.1
Erkrankungswahrscheinlichkeit	minimal	Z	0; 0.2
	niedrig	Pi	0; 0.2; 0.25; 0.45
	mittel	Pi	0.25; 0.45; 0.55; 0.75
	hoch	S	0.55; 0.75

Hier zeigt sich, dass die Unterteilung der Eingabevariable *AvgLF* in nur zwei Terme, beispielsweise „normal“ und „hoch“, zum selben Regelgerüst geführt hätte, da effektiv nur zwischen zwei Zuständen unterschieden wird. Man kann nun entweder die Terme neu formulieren oder die Regeln daraufhin untersuchen, ob die Zustände „nicht erhöht“ und „erhöht“ in Kombination mit Termen der Variable *MaxDiff* nicht doch zu unterschiedlichen Ausgaben führen sollen.

Tabelle 43: Regelgerüst für Modell LF2

		AvgLF^a		MaxDiff^b		RelAbwLF^c		EWSK^d
1	Wenn Nicht	hoch	Und	gering	Und	steigend	Dann	minimal
2	Wenn Nicht	hoch	Und	erhöht	Und	steigend	Dann	niedrig
3	Wenn Nicht	hoch	Und	hoch	Und	steigend	Dann	hoch
4	Wenn Nicht	hoch	Und	sehr hoch	Und	steigend	Dann	hoch
5	Wenn	hoch					Dann	hoch
6	Wenn					sinkend	Dann	minimal
7	Wenn			gering	Und	normal	Dann	minimal
8	Wenn			erhöht	Und	normal	Dann	niedrig
9	Wenn			hoch	Und	normal	Dann	hoch
10	Wenn			sehr hoch	Und	normal	Dann	hoch

^a durchschnittliche Leitfähigkeit

^b Differenz des Euterviertels mit der höchsten Leitfähigkeit zum Euterdurchschnitt

^c relative Abweichung der Leitfähigkeit gegenüber dem Mittel der drei Vortage

^d Erkrankungswahrscheinlichkeit

Eine weitere Variante (Modell LF2) sah die Nutzung der prozentualen Änderung der Leitfähigkeit im Vergleich zum Mittel der drei vorherigen Tage (*RelAbwLF*) als weitere Eingabevariable vor, woraus sich die in Tabelle 42 dargestellten Terme und Zugehörigkeitsfunktionen ergeben.

Im nächsten Schritt wurde das Regelgerüst entsprechend verändert und an die zusätzliche Variable angepasst (Tabelle 43).

Tabelle 44: Linguistische Variablen und die entsprechenden Zugehörigkeitsfunktionen für Modell LF3

Variable	Term	Zugehörigkeitsfunktionstyp	Definitionspunkte
Durchschnittliche Leitfähigkeit	nicht erhöht	Z	5.4; 5.9
	erhöht	Lambda	5.4; 5.9; 6.4
	hoch	S	5.9; 6.4
Prozentuale Änderung zum Vortagemittel	sinkend	Z	-0.1; 0
	normal	Lambda	-0.1; 0; 0.1
	steigend	S	0; 0.1
Erkrankungswahrscheinlichkeit	minimal	Z	0; 0.2
	niedrig	Pi	0; 0.2; 0.25; 0.45
	mittel	Pi	0.25; 0.45; 0.55; 0.75
	hoch	S	0.55; 0.75

Tabelle 45: Regelgerüst für Modell LF3

		AvgLF^a		RelAbwLF^b		EWSK^c
1	<i>Wenn</i>	nicht erhöht	<i>Und</i>	sinkend	<i>Dann</i>	minimal
2	<i>Wenn</i>	nicht erhöht	<i>Und</i>	normal	<i>Dann</i>	minimal
3	<i>Wenn</i>	nicht erhöht	<i>Und</i>	steigend	<i>Dann</i>	mittel
4	<i>Wenn</i>	erhöht	<i>Und</i>	sinkend	<i>Dann</i>	niedrig
5	<i>Wenn</i>	erhöht	<i>Und</i>	normal	<i>Dann</i>	mittel
6	<i>Wenn</i>	erhöht	<i>Und</i>	steigend	<i>Dann</i>	hoch
7	<i>Wenn</i>	hoch	<i>Und</i>	sinkend	<i>Dann</i>	mittel
8	<i>Wenn</i>	hoch	<i>Und</i>	normal	<i>Dann</i>	hoch
9	<i>Wenn</i>	hoch	<i>Und</i>	steigend	<i>Dann</i>	hoch

^a durchschnittliche Leitfähigkeit

^b relative Abweichung der Leitfähigkeit gegenüber dem Mittel der drei Vortage

^c Erkrankungswahrscheinlichkeit

Mit einem dritten Modell (LF3) sollte untersucht werden, wie gut die Erkennungsmöglichkeiten von Eutererkrankungen ohne Berücksichtigung der Differenz der Leitfähigkeit einzelner Viertel im Vergleich zu den vorher betrachteten Modellen LF1 und LF2 sind. Dies wäre eine Alternativmöglichkeit für Melkanlagen, bei denen keine Messwerte für die einzelnen Euterviertel vorliegen. Es wurden also nur durchschnittliche Leitfähigkeit und relative Änderung der Leitfähigkeit als Inputvariablen verwendet (Tabelle 44). Die entsprechenden Regeln für Modell LF3 sind in Tabelle 45 dargestellt.

Ein viertes Modell (LF4) basiert auf den Eingabevariablen aus LF2 (vgl. Tabelle 42). Dieses Modell verwendet jedoch ein detaillierteres Regelgerüst, um zu untersuchen, inwiefern die Verwendung einer feineren Abstufung unter Ausnutzung aller vorhandenen Terme sinnvoll ist oder ob man auf einige Terme ohne eine Verschlechterung der Erkennung von Euterkrankheiten verzichten kann (Tabelle 46). Durch die hier vorgenommene Abstufung wird aber beispielsweise auch die Eingabevariable *MaxDiff* nicht immer zu einer Einstufung als hoch für die Erkrankungswahrscheinlichkeit führen, sondern auch verstärkt Augenmerk auf das durchschnittliche Leitfähigkeitsniveau über alle Euterviertel und Veränderung des Niveaus gegenüber den Vortagen gelegt.

Tabelle 46: Regelgerüst für Modell LF4

		AvgLF^a		MaxDiff^b		RelAbwLF^c		EWSK^d
1	<i>Wenn</i>	nicht erhöht	<i>Und</i>	gering	<i>Und Nicht</i>	steigend	<i>Dann</i>	minimal
2	<i>Wenn</i>	nicht erhöht	<i>Und</i>	gering	<i>Und</i>	steigend	<i>Dann</i>	niedrig
3	<i>Wenn</i>	nicht erhöht	<i>Und</i>	erhöht	<i>Und Nicht</i>	steigend	<i>Dann</i>	minimal
4	<i>Wenn</i>	nicht erhöht	<i>Und</i>	erhöht	<i>Und</i>	steigend	<i>Dann</i>	niedrig
5	<i>Wenn</i>	nicht erhöht	<i>Und</i>	hoch	<i>Und</i>	sinkend	<i>Dann</i>	niedrig
6	<i>Wenn</i>	nicht erhöht	<i>Und</i>	hoch	<i>Und</i>	normal	<i>Dann</i>	mittel
7	<i>Wenn</i>	nicht erhöht	<i>Und</i>	hoch	<i>Und</i>	steigend	<i>Dann</i>	hoch
8	<i>Wenn</i>	nicht erhöht	<i>Und</i>	sehr hoch	<i>Und</i>	sinkend	<i>Dann</i>	niedrig
9	<i>Wenn</i>	nicht erhöht	<i>Und</i>	sehr hoch	<i>Und</i>	normal	<i>Dann</i>	mittel
10	<i>Wenn</i>	nicht erhöht	<i>Und</i>	sehr hoch	<i>Und</i>	steigend	<i>Dann</i>	hoch
11	<i>Wenn</i>	erhöht	<i>Und</i>	gering	<i>Und</i>	sinkend	<i>Dann</i>	minimal
12	<i>Wenn</i>	erhöht	<i>Und</i>	gering	<i>Und</i>	normal	<i>Dann</i>	niedrig
13	<i>Wenn</i>	erhöht	<i>Und</i>	gering	<i>Und</i>	steigend	<i>Dann</i>	mittel
14	<i>Wenn</i>	erhöht	<i>Und</i>	erhöht	<i>Und</i>	sinkend	<i>Dann</i>	niedrig
15	<i>Wenn</i>	erhöht	<i>Und</i>	erhöht	<i>Und Nicht</i>	sinkend	<i>Dann</i>	mittel
16	<i>Wenn</i>	erhöht	<i>Und</i>	hoch	<i>Und Nicht</i>	steigend	<i>Dann</i>	mittel
17	<i>Wenn</i>	erhöht	<i>Und</i>	hoch	<i>Und</i>	steigend	<i>Dann</i>	hoch
18	<i>Wenn</i>	erhöht	<i>Und</i>	sehr hoch	<i>Und Nicht</i>	steigend	<i>Dann</i>	mittel
19	<i>Wenn</i>	erhöht	<i>Und</i>	sehr hoch	<i>Und</i>	steigend	<i>Dann</i>	hoch
20	<i>Wenn</i>	hoch	<i>Und</i>	gering	<i>Und</i>	sinkend	<i>Dann</i>	niedrig
21	<i>Wenn</i>	hoch	<i>Und</i>	gering	<i>Und</i>	normal	<i>Dann</i>	mittel
22	<i>Wenn</i>	hoch	<i>Und</i>	gering	<i>Und</i>	steigend	<i>Dann</i>	hoch
23	<i>Wenn</i>	hoch	<i>Und</i>	erhöht	<i>Und</i>	sinkend	<i>Dann</i>	mittel
24	<i>Wenn</i>	hoch	<i>Und</i>	erhöht	<i>Und Nicht</i>	sinkend	<i>Dann</i>	hoch
25	<i>Wenn</i>	hoch	<i>Und</i>	hoch	<i>Und</i>	sinkend	<i>Dann</i>	mittel
26	<i>Wenn</i>	hoch	<i>Und</i>	hoch	<i>Und Nicht</i>	sinkend	<i>Dann</i>	hoch
27	<i>Wenn</i>	hoch	<i>Und</i>	sehr hoch			<i>Dann</i>	hoch

^a durchschnittliche Leitfähigkeit

^b Differenz des Euterviertels mit der höchsten Leitfähigkeit zum Euterdurchschnitt

^c relative Abweichung der Leitfähigkeit gegenüber dem Mittel der drei Vortage

^d Erkrankungs Wahrscheinlichkeit

Tabelle 47: Linguistische Variablen und die entsprechenden Zugehörigkeitsfunktionen für Modell LF5

Variable	Term	Zugehörigkeitsfunktionstyp	Definitionspunkte
EWSK(LF) ^a	kein Alarm	Z	0.45; 0.5
	Alarm	S	0.5; 0.55
Absolute Abweichung Milchmenge	negativ	Z	-3; -1
	normal	Lambda	-3; 0; 3
Prozentuale Abweichung Milchmenge	positiv	S	1; 3
	negativ	Z	-0.08; -0.03
gering negativ	Lambda	-0.1; -0.05; 0	
	normal	Lambda	-0.05; 0; 0.05
positiv	S	0; 0.05	
	negativ	Z	-5; 0
Abweichungsfaktor	normal	Lambda	-5; 0; 5
	positiv	S	0; 5
Erkrankungswahr- scheinlichkeit	niedrig	Z	0.2; 0.4
	mittel	Pi	0.2; 0.4; 0.6; 0.8
	hoch	S	0.6; 0.8

^a Erkrankungswahrscheinlichkeit, berechnet anhand der Leitfähigkeit

Die vier bisher vorgestellten Modelle wurden zusätzlich als Vorstufe eines Modells (LF5) verwendet, das die Tagesmilchmenge zur Bestimmung des Gesundheitsstatus eines Tieres mit einbezieht. Neben dem in den bisher behandelten Modellen erhaltenen Ausgabewert ($EWSK(LF)$) werden drei weitere Eingabevariablen einbezogen: die absolute Differenz der aktuellen Milchmenge zur Milchmenge des Vortags ($AbsAbw$), die prozentuale Abweichung der aktuellen Milchmenge von der Milchmenge des Vortags ($RelAbw$) und der Abweichungsfaktor (AF), der den tendenziellen Verlauf der beobachteten Milchleistungskurve der vergangenen drei Tage mit dem im gemischten Modell geschätzten tendenziellen Verlauf der Milchleistung im selben Zeitraum vergleicht. Für die Variablen ergeben sich die in Tabelle 47 dargestellten Terme.

Das dazugehörige Regelgerüst ist in Tabelle 48 dargestellt. Es ergeben sich daraus insgesamt acht Modellvarianten, die in Tabelle 49 zusammengestellt sind.

Tabelle 48: Regelgerüst für Modell LF5

	EWSK (LF) ^a		AbsAbw ^b		RelAbw ^c		AF ^d		EWSK ^e	
1	Wenn	Alarm						Dann	hoch	
2	Wenn	Kein Alarm	Und	positiv				Dann	niedrig	
3	Wenn	Kein Alarm			Und	positiv		Dann	niedrig	
4	Wenn	Kein Alarm					Und	positiv	Dann	niedrig
5	Wenn	Kein Alarm			Und	normal		Dann	niedrig	
6	Wenn	Kein Alarm	Und	normal	Und	gering negativ	Und	normal	Dann	niedrig
7	Wenn	Kein Alarm	Und	negativ	Und	gering negativ	Und	normal	Dann	niedrig
8	Wenn	Kein Alarm	Und	negativ	Und	gering negativ	Und	negativ	Dann	hoch
9	Wenn	Kein Alarm	Und	normal	Und	gering negativ	Und	negativ	Dann	mittel
10	Wenn	Kein Alarm	Und	normal	Und	negativ	Und	normal	Dann	niedrig
11	Wenn	Kein Alarm	Und	negativ	Und	negativ	Und	normal	Dann	mittel
12	Wenn	Kein Alarm	Und	negativ	Und	negativ	Und	negativ	Dann	hoch
13	Wenn	Kein Alarm	Und	normal	Und	negativ	Und	negativ	Dann	mittel

^a Erkrankungswahrscheinlichkeit, berechnet anhand der Leitfähigkeit

^b absolute Abweichung der Milchmenge gegenüber dem Vortag

^c relative Abweichung der Milchmenge gegenüber dem Vortag

^d Abweichungsfaktor

^e Erkrankungswahrscheinlichkeit

Tabelle 49: Übersicht der Modellvarianten

Modell	Verwendete Eingabevariablen						
	AvgLF ^a	RelAbwLF ^b	MaxDiff ^c	EWSK(LF) ^d	AbsAbw ^e	RelAbw ^f	AF ^g
LF1	X		X				
LF2	X	X	X				
LF3	X	X					
LF4	X	X	X				
LF1 & LF5	(X)	(X)	(X)	X	X	X	X
LF2 & LF5	(X)	(X)	(X)	X	X	X	X
LF3 & LF5	(X)	(X)	(X)	X	X	X	X
LF4 & LF5	(X)	(X)	(X)	X	X	X	X

^a durchschnittliche Leitfähigkeit

^b relative Abweichung der Leitfähigkeit gegenüber dem Mittel der drei Vortage

^c Differenz des Euterviertels mit der höchsten Leitfähigkeit zum Euterdurchschnitt

^d Erkrankungswahrscheinlichkeit, berechnet anhand der Leitfähigkeit

^e absolute Abweichung der Milchmenge gegenüber dem Vortag

^f relative Abweichung der Milchmenge gegenüber dem Vortag

^g Abweichungsfaktor

4.2.1.2 Ergebnisse zur Erkennung von Eutererkrankungen

Die Anzahl der Alarme an den drei Tagen vor Dokumentation einer Eutererkrankung sind Tabelle 50 zu entnehmen. Es ist ersichtlich, dass die Verwendung der Differenz der Leitfähigkeit zwischen dem Euterviertel mit der höchsten Leitfähigkeit und dem Gesamteuterdurchschnitt maßgeblich für die Alarme verantwortlich ist. Die Modelle LF3 und (LF5 & LF3) liefern an den beiden Tagen vor Feststellung einer Erkrankung die wenigsten Alarme. Die relative Abweichung der Leitfähigkeit gegenüber dem Vortag führt gegenüber der alleinigen Berücksichtigung der durchschnittlichen Leitfähigkeit ebenfalls zu einer geringeren Anzahl Alarme. Die Hinzunahme der Merkmalsgruppe "Tagesmilchmenge" in den kombinierten Modellen führt in einigen Fällen zu einer Verringerung der Alarme. In diesen Fällen handelt es sich um Alarme, die in den Basismodellen LF1 bis LF4 direkt am gewählten Grenzwert lagen und nun ohne gleichzeitiges Vorliegen einer negativen Milchmengenabweichung nicht mehr als Alarm gewertet werden. In anderen Fällen kommen Alarme hinzu, wenn zwar vorher keine auffällige Leitfähigkeitsdifferenz vorlag, gleichzeitig aber eine starke negative Milchmengenabweichung vorliegt. Es wird außerdem deutlich, dass die Anzahl der Alarme und somit der möglichen Symptome für eine Erkrankung generell zunimmt, je näher der erste dokumentierte Krankheitstag rückt.

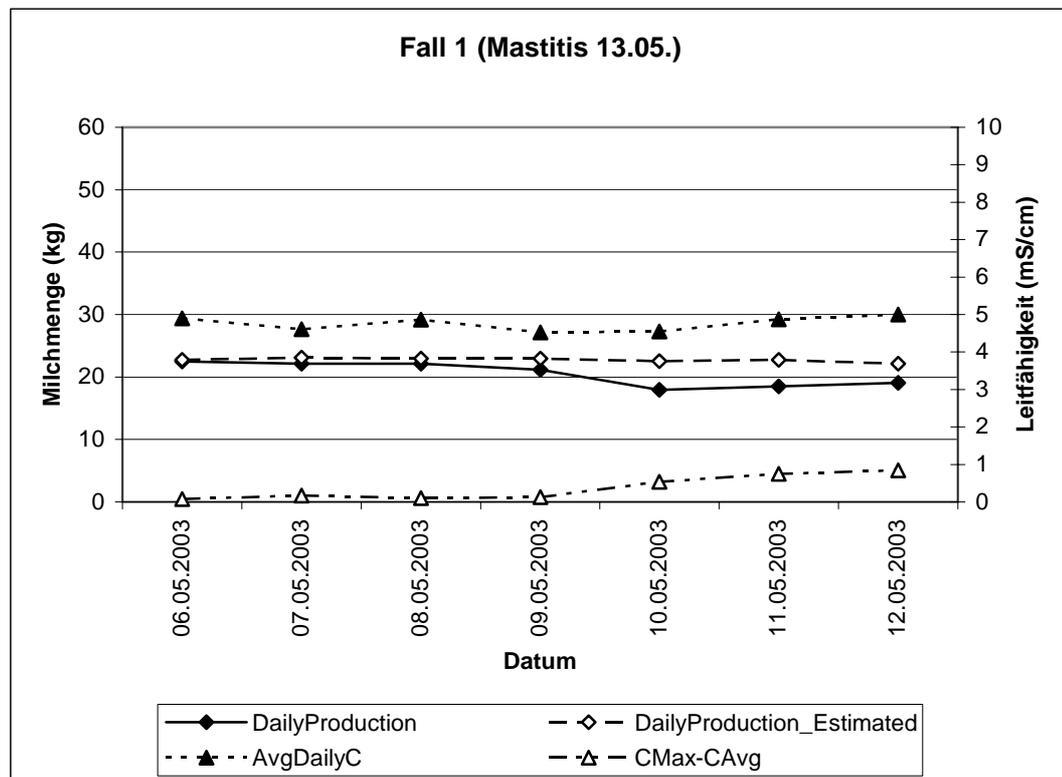
Tabelle 50: Alarme in Abhängigkeit der Modellvariante

Modell	Fälle	Erster Vortag (20)	Zweiter Vortag (20)	Dritter Vortag (22)	Weitere Vorwoche
LF1		16	7	7	20
LF2		15	5	6	22
LF3		5	1	3	8
LF4		11	4	2	11
LF1 & LF5		12	8	5	20
LF2 & LF5		12	8	5	26
LF3 & LF5		9	5	4	14
LF4 & LF5		11	7	4	16

Anmerkung: Im Gegensatz zur Erkennung von Milchminderungen wurde hier mit der Woche vor den dokumentierten Eutererkrankungen nur ein kleiner Teil des Gesamtdatensatzes untersucht. Eine nachträgliche Klassifizierung der einzelnen untersuchten Tage in "krank" und "gesund" ist im Nachhinein objektiv nicht möglich, weshalb hier von der Berechnung von Fehler- und Erkennungsraten sowie Spezifität abgesehen wurde.

4.2.2 Fallbeispiele

Es werden nun in den folgenden Abbildungen 29 bis 34 drei Fälle von Mastitiserkrankungen in der Woche vor Feststellung der Erkrankung betrachtet. Damit soll grafisch dargestellt werden, wie die einzelnen Modellvarianten auf die untersuchten Merkmale reagieren und Alarmer erzeugen. Auch die Schwierigkeit der nachträglichen Klassifikation der untersuchten Tage in gesunde und Krankentage wird in Fallbeispiel 2 nochmals verdeutlicht. *DailyProduction* bezeichnet jeweils die Tagesmilchmenge, *DailyProduction_Estimated* die mit dem gemischten linearen Modell vorhergesagte Tagesmilchmenge, *AvgDailyC* die über den Tag gemittelte Leitfähigkeit sowie *CMax-CAvg* die Differenz der Leitfähigkeit zwischen dem Euterviertel mit dem höchsten Wert und dem Gesamteuter. Die Modellvarianten LF1 bis LF5 sind in Tabelle 49 dargestellt.



Dailyproduction:	beobachtete Tagesmilchmenge
Dailyproduction_Estimated:	vorhergesagte Tagesmilchmenge
AvgDailyC:	durchschnittliche Leitfähigkeit über alle Euterviertel am Tag
CMax-CAvg:	Differenz der Leitfähigkeit zwischen dem Euterviertel mit dem Tageshöchstwert und dem Gesamteuterdurchschnitt

Abbildung 29: Verlauf von geschätzter und beobachteter Milchmenge sowie Leitfähigkeit in Fallbeispiel 1

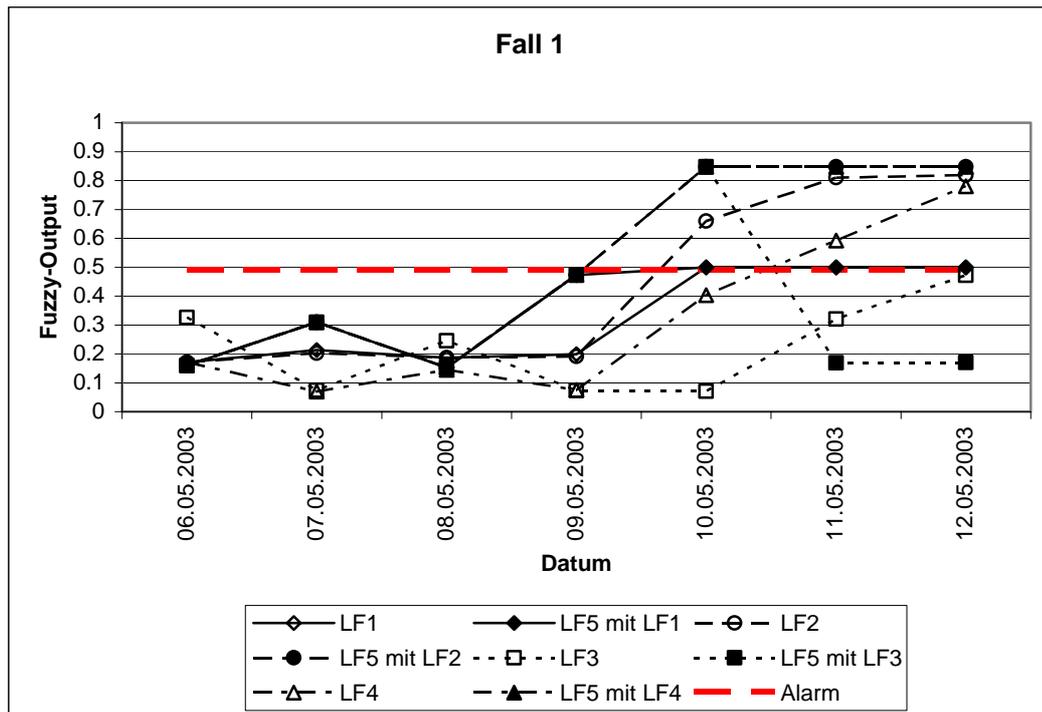
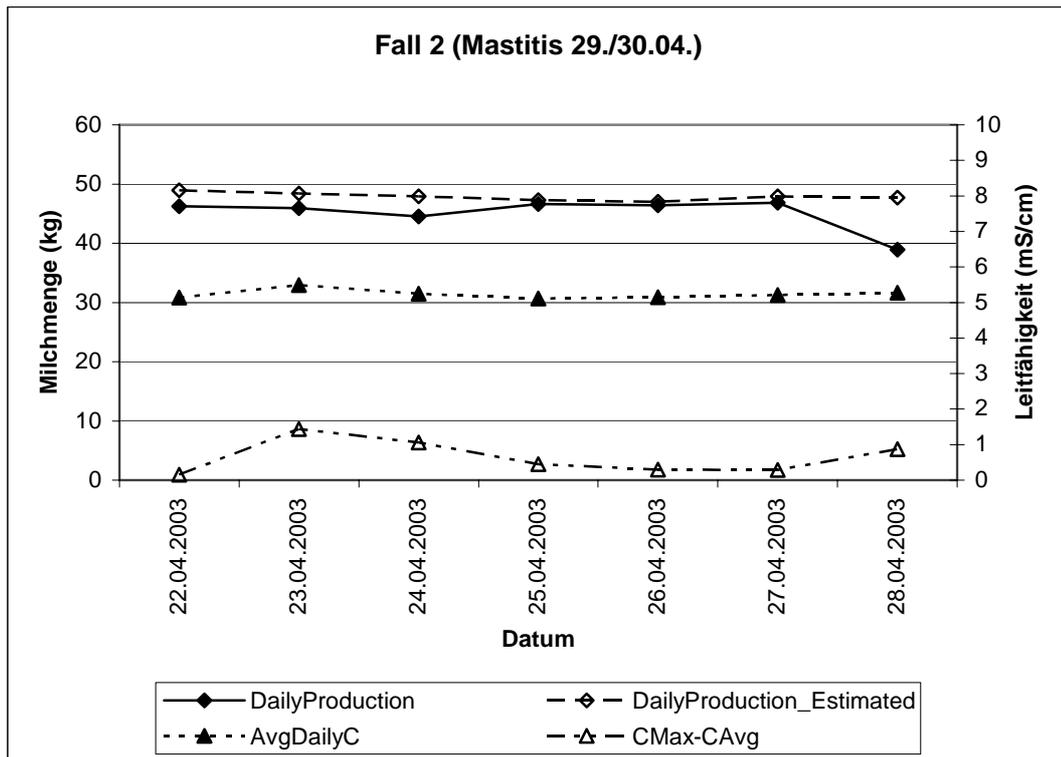


Abbildung 30: Ergebnisse der Fuzzy-Logic-Modelle für Fallbeispiel 1

Im ersten Beispiel ist der Anstieg der Leitfähigkeitsdifferenz ab dem fünften Tag bei gleichzeitiger Milchmindermenge gut erkennbar (Abbildungen 29 und 30). Die durchschnittliche Leitfähigkeit bewegt sich allerdings innerhalb des unauffälligen Bereiches. Entsprechend erzielt Modell LF3 auch nur mit Einbeziehung der Milchmenge in Variante LF5 einen Alarm an Tag 5. Modell LF4 mit dem feiner abgestimmten Regelgerüst überschreitet die Alarmschwelle erst an Tag 6, für alle anderen Varianten ist die Leitfähigkeitsdifferenz entscheidendes Merkmal, das an allen drei Tagen vor Krankheitserkennung zu Alarmen führt. Auch in diesen Fällen sorgt die gleichzeitige Betrachtung der Milchmenge für höhere Alarmwerte. Modell LF1 erreicht auch hier in beiden Varianten nur den minimalen Alarmwert von 0.5, da die sich im physiologischen Normalbereich befindliche durchschnittliche Leitfähigkeit im vorgesehenen Regelwerk bei maximalen Leitfähigkeitsdifferenzen kleiner als 1mS/cm nur eine mittlere Erkrankungswahrscheinlichkeit als Ausgabewert zulässt.



Dailyproduction: beobachtete Tagesmilchmenge
 Dailyproduction_Estimated: vorhergesagte Tagesmilchmenge
 AvgDailyC: durchschnittliche Leitfähigkeit über alle Euterviertel am Tag
 CMax-CAvg: Differenz der Leitfähigkeit zwischen dem Euterviertel mit dem Tageshöchstwert und dem Gesamteuterdurchschnitt

Abbildung 31: Verlauf von geschätzter und beobachteter Milchmenge sowie Leitfähigkeit in Fallbeispiel 2

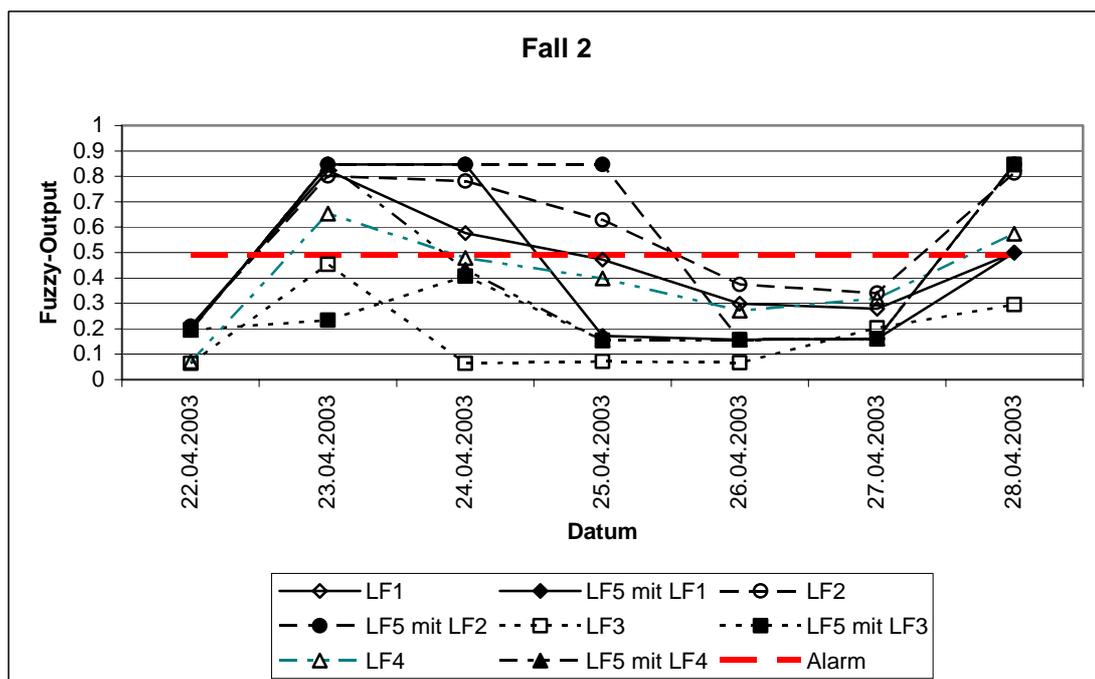
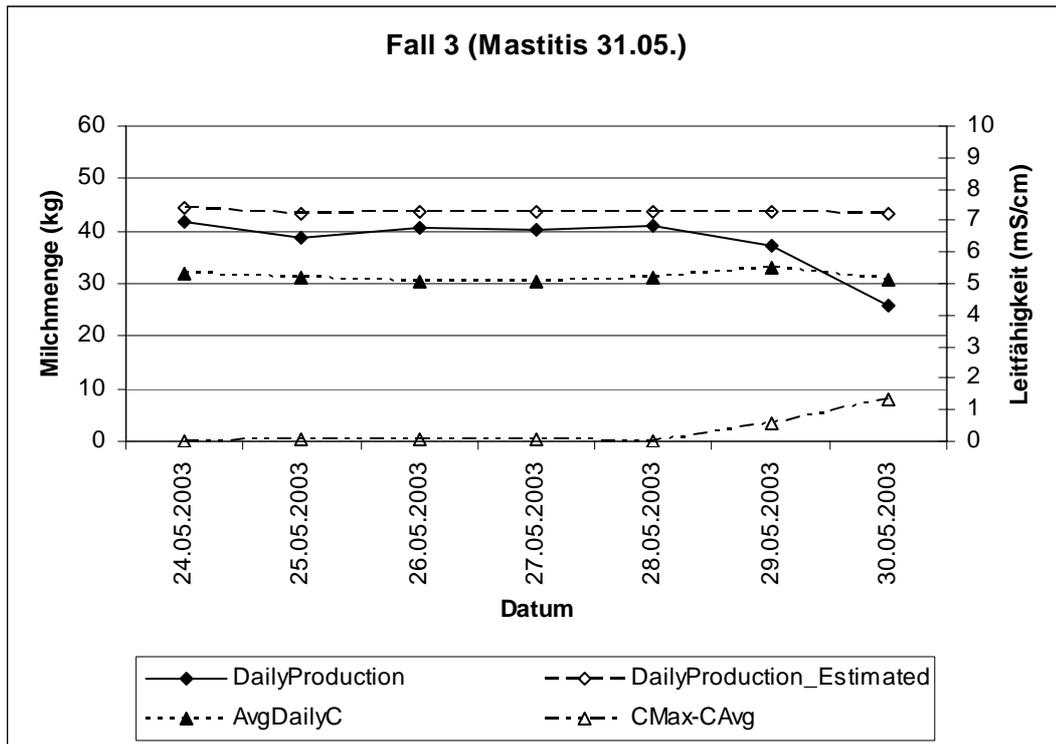


Abbildung 32: Ergebnisse der Fuzzy-Logic-Modelle für Fallbeispiel 2

Auch in diesem Fall (Abbildungen 31 und 32) erkennen alle auf Leitfähigkeitsdifferenz basierenden Modelle den starken Anstieg bereits am zweiten Wochentag. Am letzten Tag der Woche ist ein weiteres Mal eine erhöhte Leitfähigkeitsdifferenz zu verzeichnen, gleichzeitig tritt auch eine Milchmindermenge auf, die bei allen Modellen mit Ausnahme von LF3 ebenfalls zu einem Alarm führt. Die beiden Tage 26. und 27.04. weisen in den beobachteten Merkmalen keine auffälligen Abweichungen auf. Damit werden dort auch berechtigt keine Alarme erzeugt. Andererseits ist es nicht gerechtfertigt, diese kurze Übergangszeit zwischen zwei Perioden mit deutlich erhöhter Leitfähigkeit in einem Euterviertel als "gesund" zu klassifizieren.

Der dritte Fall (Abbildungen 33 und 34) stellt ein Paradebeispiel für den erwarteten Verlauf einer Eutererkrankung dar. Die Modellvarianten von LF5 reagieren am zweiten Tag gleichermaßen mit einem Alarm auf eine Milchmindermenge von etwa 7%. Am Tag sechs und sieben ist dann der bei einer Eutererkrankung zu erwartende Anstieg der Leitfähigkeit in einem Euterviertel sowie gleichzeitig eine Mindermenge erkennbar. Mit Ausnahme des Modells LF3 erreichen alle Varianten hier den Alarmgrenzwert von 0.5 oder mehr.



Dailyproduction: beobachtete Tagesmilchmenge
 Dailyproduction_Estimated: vorhergesagte Tagesmilchmenge
 AvgDailyC: durchschnittliche Leitfähigkeit über alle Euterviertel am Tag
 CMax-CAvg: Differenz der Leitfähigkeit zwischen dem Euterviertel mit dem Tageshöchstwert und dem Gesamteuterdurchschnitt

Abbildung 33: Verlauf von geschätzter und beobachteter Milchmenge sowie Leitfähigkeit in Fallbeispiel 3

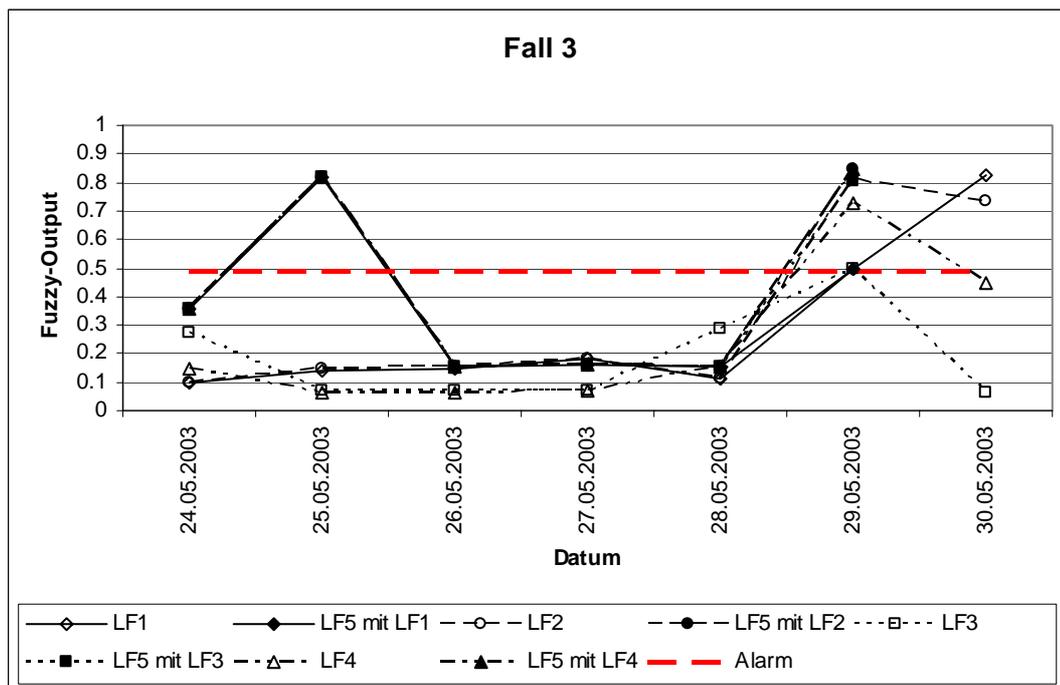


Abbildung 34: Ergebnisse der Fuzzy-Logic-Modelle für Fallbeispiel 3

4.2.3 Mastitiserkennung unter Nutzung der Leitfähigkeit

Es zeigt sich, dass die Betrachtung der Leitfähigkeit auf Viertelebene in den in dieser Arbeit vorliegenden Fällen ein entscheidendes Merkmal für die frühzeitige Erkennung von Mastitiserkrankungen darstellt. Dies wird insbesondere klar, wenn man wie im Beispielbetrieb 1 feststellt, dass bezogen auf die in der Literatur zu findenden Werte auf das Gesamteuter gesehen keine außergewöhnlich hohen Leitfähigkeitswerte auftreten. Insbesondere ist auffällig, dass die erkrankten Viertel mit Hilfe der Differenz der Leitfähigkeit zwischen dem Euterviertel mit dem Tageshöchstwert und dem Gesamteuterdurchschnitt oft schon wenige Tage, in einzelnen Fällen aber auch bis zu einer Woche vor Dokumentation der Erkrankung im Betrieb identifiziert werden können.

Ein geeignet modelliertes Fuzzy-System ist in der Lage, diese Abweichungen zu erkennen. Die Milchmenge ist ein verwendbares Merkmal, problematisch ist allerdings, dass auch Fälle von Eutererkrankungen ohne nennenswerte Mindermengen bzw. im Gegenteil sogar Anstiegen der Milchmenge auftreten. Daher ist die Milchmenge als zusätzliches unterstützendes Merkmal denkbar, sollte jedoch mit der idealerweise je Viertel vorliegenden Leitfähigkeit kombiniert werden.

4.3 Nutzung des gemischten linearen Modells für das Herdenmanagement

4.3.1 Schätzung von herdenweiten Milchmengen

Eine weitere Nutzungsmöglichkeit zur Unterstützung des Herdenmanagements mit dem gemischten linearen Modell besteht in der Verwendung der geschätzten Parameter für den Verlauf der Laktationskurven der Einzeltiere, um daraus längerfristige Milchmengenvorhersagen, beispielsweise zur Unterstützung des Quotenmanagements oder, bei Wegfall der Milchquote, zur Vorausplanung von Milchmengen für Lieferverträge mit Molkereien, zu machen. Diese können dann über alle Tiere zu Herdenmilchmengen aufsummiert werden.

Mit zunehmendem Abstand zwischen der letzten verfügbaren Leistung und dem Vorhersagetag ist mit zunehmender Ungenauigkeit der Schätzung zu rechnen, da keine aktuellen Daten mehr in die Schätzung einfließen. Daher wurde der Vorhersagezeitraum in der vorliegenden Arbeit auf einen Monat begrenzt. Als Datenbasis wurden die Daten aus Betrieb 3 verwendet, da dort keine Fehltag auftraten und so die beobachteten herdenweiten Tagesmilchmengen komplett dokumentiert waren. Die im gemischten linearen Modell für diesen Betrieb verwendeten 245 Tage für die Vorhersage

wurden in sieben Abschnitte zu je 30 Tagen eingeteilt. Die ersten zehn Tage wurden zusammen mit 90 Tagen weiterhin verfügbarer Daten der schon in Versuchen befindlichen Tiere als Datenvorlauf verwendet. Es wurden jeweils die mit dem gemischten linearen Modell für den letzten Tag vor einem 30-Tage-Abschnitt geschätzten Effekte verwendet, um für die folgenden 30 Tage Milchmengen je Tier und Tag vorherzusagen. Dabei wurden nur die Tiere berücksichtigt, für die am Tag vor Beginn des Vorhersagezeitraums Effektschätzungen vorlagen. Diese Tagesmilchmengen der Einzeltiere wurden dann zu einer Gesamtmilchmenge für die Herde zunächst je Tag und dann für die untersuchten 30 Tage aufsummiert. Diesen Vorhersagewerten wurde die Summe der im selben Zeitraum tatsächlich beobachteten Milchmengen derselben Tiere gegenübergestellt. Milchmengen von Tieren, die während des gerade untersuchten 30-Tage-Zeitraums abkalbten und zur Herde hinzukamen, wurden dabei nicht berücksichtigt, da für diese auch keine Effektschätzungen zu Beginn des Vorhersagezeitraums vorlagen und daher keine Tagesmilchmengen vorhergesagt werden konnten. Als Vergleichsmaß zwischen den vorhergesagten und den beobachteten Monatsmengen wurde die prozentuale Abweichung gewählt. Die Ergebnisse sind in Tabelle 51 dargestellt.

Tabelle 51: Beobachtete und vorhergesagte Herdenmilchmengen im Vergleich

Monat ^a	1	2	3	4	5	6	7
Anzahl Tiere zu Monatsbeginn ^a	34	34	40	42	44	38	36
Anzahl Tagesmilchmengen	1020	1020	1197	1241	1289	1129	1008
beobachtete Gesamtmenge (kg)	30540.5	28845.1	32868.5	33992.8	33710.1	29361.1	25100
vorhergesagte Gesamtmenge (kg)	29921.7	27043.1	33312	34963.5	34066.6	29907	25975
Abweichung (kg) ^b	618.8	1802	-443.5	-970.7	-356.5	-545.9	-875
Abweichung (%) ^b	2.07	6.66	-1.33	-2.78	-1.05	-1.83	-3.37

^a ein Monat entspricht hier jeweils einem 30-Tage-Zeitraum

^b jeweils Abweichung der beobachteten Menge von der vorhergesagten Menge

Die prozentualen Abweichungen der Schätzung von den tatsächlich ermolkenen Milchmengen bewegen sich in vertretbarem Rahmen und scheinen für die Unterstützung des Quotenmanagements geeignet. Es ist eine etwas stärkere Unterschätzung der realisierten Milchmenge im zweiten Monat des untersuchten Zeitraums zu erkennen. Eine Erklärungsmöglichkeit besteht in der Zusammensetzung der Herde. Zu Beginn stehen vor allem Daten von Tieren, die sich im ersten

Laktationsdrittel befinden, als Datengrundlage für eine Schätzung zur Verfügung. Die Unterschätzung kann daher dadurch zustande kommen, dass modellseitig bereits ein Absinken der Milchmenge erwartet wird, die Tiere jedoch eine gute Persistenz zeigen (vgl. Abbildung 35) oder neue Tiere mit zu Beginn der Laktation höheren Milchmengen als die zuvor im Versuch befindlichen dazukommen.

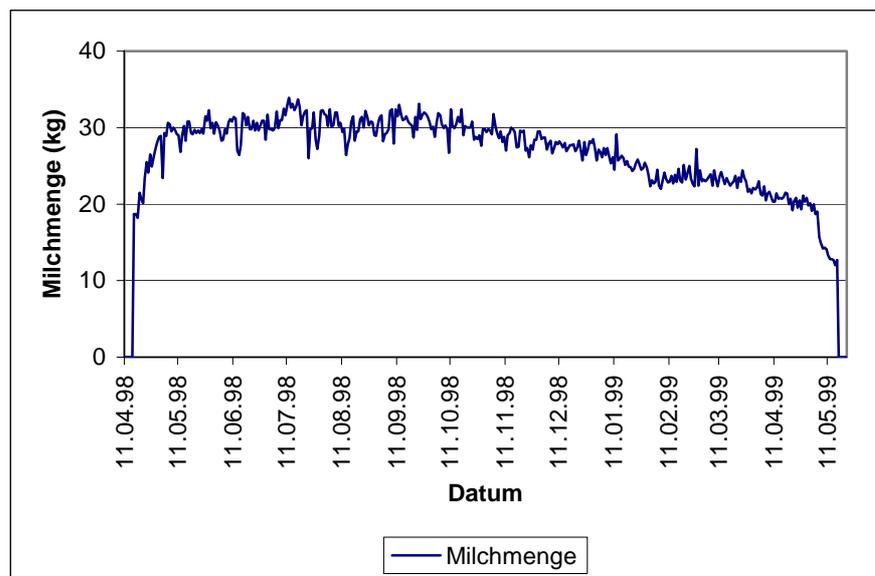


Abbildung 35: Laktationsverlauf mit guter Persistenz nach dem Laktationshöhepunkt, beispielhaft für einige Tiere in Betrieb 3

4.3.2 Nutzung von Testtagseffekten

In den genutzten gemischten linearen Modellen sind zur genaueren Schätzung der individuellen Laktationskurven der Einzeltiere die Testtagseffekte einbezogen. Wie in 3.1.4 kurz dargestellt, sind diese aber auch zur Kontrolle des betrieblichen Managements nutzbar. Die Abbildungen 36 bis 38 zeigen die LSMEANS in den drei untersuchten Betrieben, berechnet mit dem tier- und laktationsspezifischen Random-Regression-Modell und Modellierung des Laktationskurvenverlaufs nach Ali und Schaeffer (1987).

Erstrebenswert für einen Betrieb sind hierbei möglichst gering schwankende Werte der LSMEANS auf hohem Leistungsniveau. Dies wäre als gleich bleibende Umweltbedingungen und entsprechend gutes Management des Betriebes zu interpretieren.

Abbildung 36 zeigt für Betrieb 1 eine weitestgehend gleichmäßige Kurve ohne starke Schwankungen. Der auffällige negative Peak am 127. Testtag wird durch einen Datenausfall von

einer einzelnen Melkbox an diesem Tag verursacht. Gründe für den negativen Peak am zweiten Testtag konnten im Nachhinein nicht rekonstruiert werden.

Betrieb 2 weist stärkere Schwankungen auf (Abbildung 37). Es ist eine generelle Abwärtstendenz zu erkennen, unterbrochen von unregelmäßigen Aufschwüngen. Die aggregierte Darstellung über einen längeren Zeitraum, wie in diesem Fall, sollte generell dazu geeignet sein, etwaige saisonale

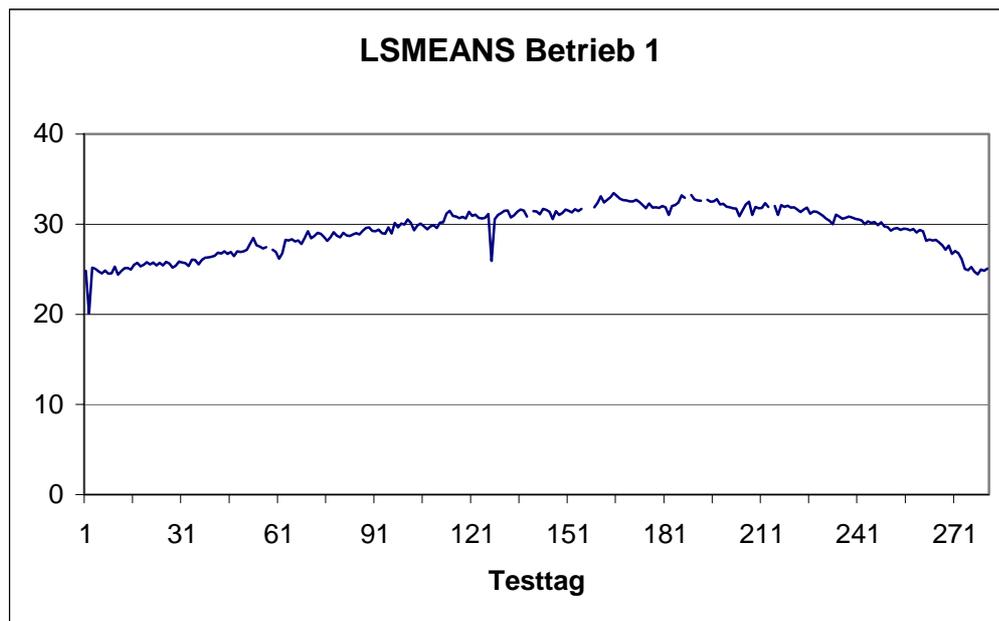


Abbildung 36: LSMEANS der Testtage in Betrieb 1 im untersuchten Zeitraum

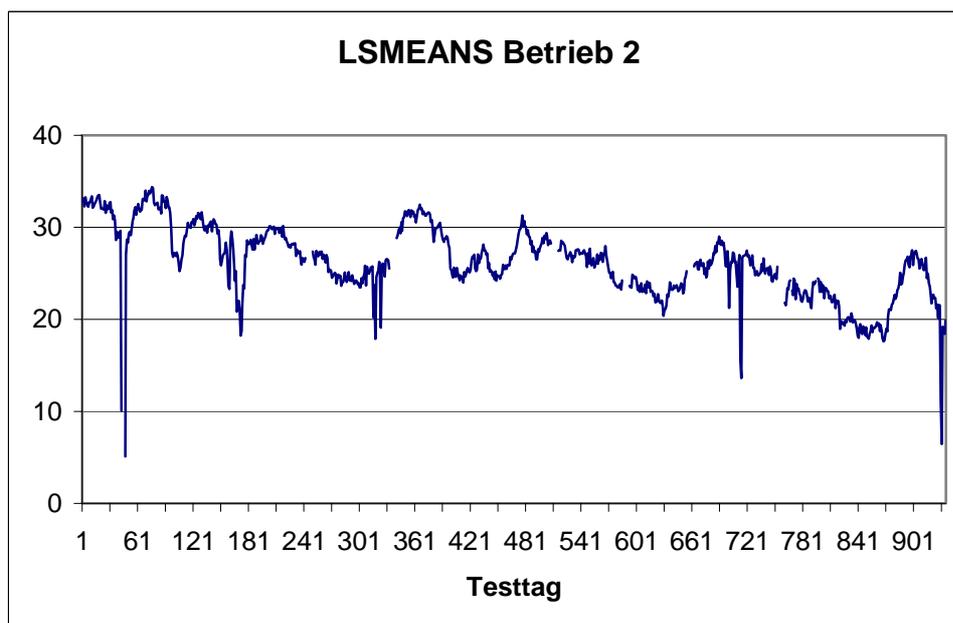


Abbildung 37: LSMEANS der Testtage in Betrieb 2 im untersuchten Zeitraum

Einflüsse der betrieblichen Umwelt abzubilden. Zwar deuten sich Ansätze von zyklischem Verhalten im Abstand von ungefähr 300 Tagen an, dies ist aber nicht begründbar. Eine Ursache für die generellen Schwankungen kann in der ungleichmäßigen Qualität der Futterrationen, hier vor allem der wirtschaftseigenen Grundfuttermittel, gegeben sein. Genauere Angaben über Ereignisse in der Betriebsumwelt, die zu solchen Schwankungen führen könnten, können durch den zeitlichen Abstand zwischen Datenerhebung und Auswertung jedoch nicht mehr gemacht werden, da keine Dokumentation über mögliche Ursachen für betriebliche Umweltschwankungen vorliegt.

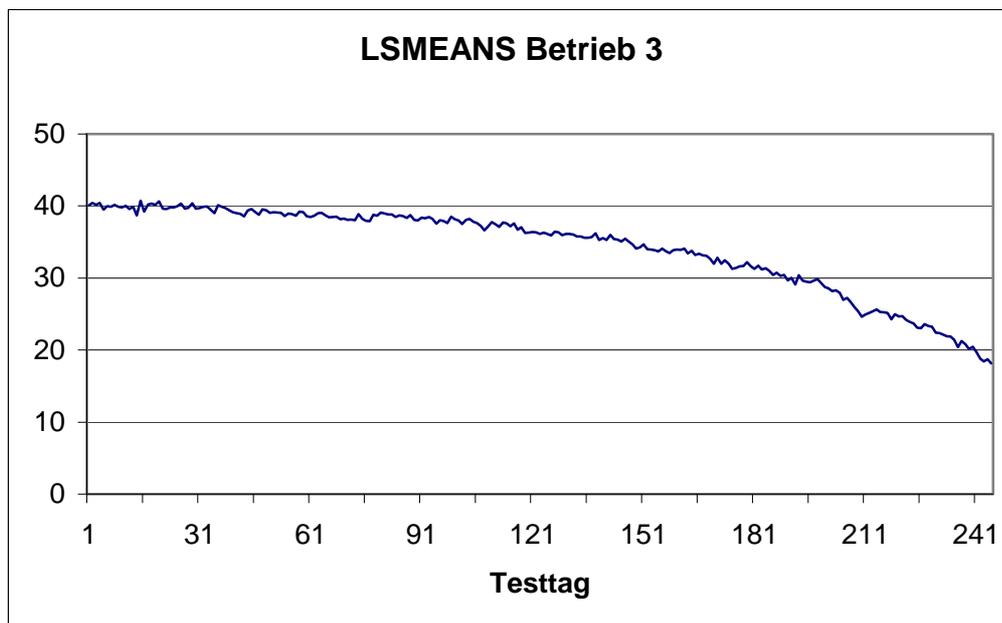


Abbildung 38: LSMEANS der Testtage in Betrieb 3 im untersuchten Zeitraum

Die in Abbildung 38 dargestellte Kurve für Betrieb 3 zeigt demgegenüber insgesamt gering schwankende Werte, die jedoch zum Ende des Untersuchungszeitraums abnehmen. Als Begründung kommt hier vor allem wieder die ungleichmäßige Verteilung der Laktationsstadien in diesem Betrieb in Frage. Gegen Ende des untersuchten Zeitraums befinden sich alle Tiere in späteren Stadien der Laktation, in denen die Milchmenge stetig abnimmt. Dadurch sinken auch die LSMEANS der Testtage, da der Laktationseinfluss und Testtag teilweise vermischt und daher schwer zu trennen sind.

Der Vorteil einer zeitnahen Berechnung von Testtageeffekten zur Überwachung der betrieblichen Umweltbedingungen liegt darin, dass beim Auftreten von (vor allem negativen) Schwankungen schnellstmöglich nach Ursachen gesucht und angemessen reagiert werden kann. Damit stellen die Testtageeffekte eine bedeutsame Managementhilfe dar.

5 Diskussion und Schlussfolgerungen

5.1 Datenqualität und betriebliche Besonderheiten der Datenbasis

Im Rahmen der vorliegenden Arbeit traten verschiedene Datenqualitätsprobleme auf, worauf noch nachfolgend detailliert eingegangen werden soll. Generell lassen sich die aufgetretenen Probleme bei der Datenerfassung in drei Gruppen klassifizieren:

- a) Fehlzusordnungen,
- b) Datenausfälle,
- c) Softwarekompatibilitätsprobleme.

Zu a)

Fehlzusordnungen ergeben sich aus der Erkennung von Tieren, die sich während des Melkvorgangs von außerhalb der Melkbox (Betrieb mit AMS) oder von benachbarten Melkplätzen (Betrieb mit Side-by-Side-Melkstand) in den Antennenbereich bewegen. Auch wenn bereits zum Melken registrierte Tiere wegen Problemen während des Melkvorgangs umgetrieben und erneut in den Melkstand gebracht werden müssen, kann es zu Fehlmengen oder nur teilweise erfassten Gemelken kommen. Als Konsequenz führen Fehlzusordnungen von einem einzelnen Gemelk zu einer Fehlmenge von je nach Melkrhythmus einem Drittel bis zur Hälfte der Tagesmilchmenge. Daraus ergibt sich in einem Entscheidungsunterstützungssystem, das Milchminderungen als Merkmal zur Gesundheitsüberwachung verwendet, folgerichtig ein Alarm an diesem Tag – in diesem Fall allerdings ein Fehlalarm. Gleichzeitig wird das Gemelk einem anderen Tier zugerechnet, wodurch dessen Vergleichsmilchmenge über mehrere Tage hinweg erhöht wird. Durch diese Erhöhung ist es ebenfalls möglich, dass zufällige normale Schwankungen in den folgenden Tagen ebenfalls Grenzwerte unterschreiten und zu Fehlalarmen führen, weil die Bezugsmilchmenge durch die Fehlzusordnung verfälscht wurde. Bei einem AMS wie in Betrieb 1 ist nach technischen Lösungen zu suchen, um die Fehlzusordnungen der Datensätze zu minimieren. In diesem Fall erlaubt beispielsweise die günstigere Platzierung der Antennen zur Tiererkennung oder alternativ bzw. ergänzend das Fernhalten anderer Tiere vom Ausgangsbereich der Melkbox und somit eine verbesserte Organisation des Kuhverkehrs, dass tatsächlich nur das in der Melkbox befindliche Tier identifiziert werden kann. Dadurch wird die Zusordnung von Gemelken zum falschen Tier praktisch ausgeschlossen.

Zu b)

Datenausfälle werden durch technische Probleme verursacht. Defekte Sensoren können zum Ausfall einzelner Merkmale führen. Aber auch Ausfälle der Daten kompletter Melkboxen traten während der Datenerfassung auf. Von den in der vorliegenden Arbeit untersuchten Betrieben war dies vor allem in Betrieb 1 ein Problem. Gegenüber den Betrieben 2 (ca. 95%) und 3 (ca. 100%) betrug die Datenabdeckung des Untersuchungszeitraums dort nur etwa 75%, d.h. durchschnittlich entstand jeden vierten Tag eine betriebsweite Datenlücke. Diese Datenausfälle führen dazu, dass ein Merkmal nicht zur Berechnung zur Verfügung steht. Dies betrifft einerseits die Bildung von gleitenden Mittelwerten über mehrere Gemelke in zeitlicher Abfolge zur Berechnung von Vergleichswerten bzw. Berechnung von Abweichungen, andererseits können für diese Fehltage mit dem gemischten linearen Modell keine aktuellen Effekte berechnet werden. Methoden, die mehrere Merkmale verwenden, können den Ausfall von unterstützenden Merkmalen bei entsprechender Modellierung verkraften, fehlt jedoch das entscheidende Merkmal, ist eine Klassifizierung des Fehltages nicht möglich. Durch Datenlücken ergeben sich also generell Tage, an denen eine Klassifizierung von Datensätzen in unbedenkliche und Alarme erfordernde Datensätze entweder gar nicht oder nur mit ungenauen bzw. veralteten Daten erfolgen kann. Diesen Problemen kann nur durch zuverlässige Technik bei der Datenerfassung, Speicherung und Weitergabe an auswertende Software entgegengewirkt werden. Nicht nur die Melkanlage mit dem Prozessrechner, sondern auch die verbundenen Rechner mit Datenbanken und Auswertungssoftware sollten beispielsweise durch Notstromaggregate gegen Stromausfälle abgesichert sein.

Zu c)

Probleme bei der Datenkommunikation zwischen Prozessrechner und Herdenmanagementsystem traten nur bei Betrieb 2 auf und führten jeweils entweder zu falsch zugeordneten Daten (beispielsweise Aufteilung der Milchmenge aus zwei Gemelken auf drei Gemelke) oder fehlenden Daten (keine Datenübertragung vom Prozessrechner möglich, wenn zum Zeitpunkt der Übertragung das Herdenmanagementprogramm verwendet wurde). Daraus ergaben sich in der Konsequenz wieder die beiden eingangs genannten Probleme. In Betrieb 2 führt die gleichzeitige Verwendung von Melkstandsoftware zur Datenerfassung und -auswertung sowie eines Herdenmanagementprogramms zu Problemen. So erfolgt z.B. die Datenübertragung der Milchleistungen zum Herdenmanagementprogramm nur, wenn dieses nicht aktiv ist. Wird es während der Übertragung benutzt oder nach vorheriger Benutzung nicht beendet, werden keine Daten übertragen. Diese werden auch nachträglich nicht mehr abgerufen, sodass Fehltage im Herdenmanagementprogramm entstehen. Zur weiteren Verwertung sind daher als einfachste Lösung die

Daten aus der Melkstandsoftware zu verwenden. Aber auch dabei können bei Betrachtung der einzelnen Gemelke Ungenauigkeiten auftreten. Neu in die Herde kommende oder nach Krankheit wieder zur Milcherzeugung genutzte Tiere müssen für die Melkberechtigung aktiviert werden. Ist dies nicht der Fall, werden zwar Tiernummer und Milchmenge erfasst, aber für diesen Tag die Gesamtmilchmenge gleichmäßig auf alle Melkvorgänge verteilt. Dabei werden allerdings nicht die tatsächliche Anzahl Melkungen an diesem Tag berücksichtigt. Ein Tier, das nur zweimal gemolken wurde, findet sich also trotzdem mit drei Einträgen zu jeweils einem Drittel der gesamten Tagesmilchmenge wieder. Dadurch verlieren diese Datensätze an Aussagekraft, wenn die Zwischenmelkzeiten und damit die durchschnittliche Milchbildung als Merkmale herangezogen werden sollen. Um diese Probleme zu umgehen und durchgängig konsistente und plausible Daten nutzen zu können, ist eine Datenbank unverzichtbar, die Daten regelmäßig und automatisch aus dem Prozessrechner übernimmt und für die Anwendungen des Herdenmanagementprogramms dauerhaft speichert. Das Herdenmanagementprogramm muss für betriebliche Besonderheiten wie unterschiedliche Melkrhythmen konfigurierbar sein. Lösungsmöglichkeiten für die Probleme bei der Datenerfassung sind auf technischer bzw. organisatorischer Seite und bei Verbesserung von Software- und Schnittstellenkompatibilität zu finden. Als Anforderung für die Verwendung von einzeltierspezifischen Daten ist es unumgänglich, dass konsistente Daten im Sinne einer eindeutigen Zuordenbarkeit kontinuierlich vorliegen. Anderenfalls wird die Verwendung von Methoden, die auf aktuelle oder zurückliegende Daten zugreifen müssen, entweder erschwert oder unmöglich gemacht.

Generell sind von einem reibungslos funktionierenden AMS die qualitativ besten Daten zu erwarten, da hier die viertelweise Erfassung von Merkmalen durch geeignete Sensoren möglich und durch die geringere Zahl von Melkplätzen im Vergleich zu Melkständen auch ökonomisch sinnvoll ist. Wie in Abschnitt 4.2 am Beispiel der elektrischen Leitfähigkeit dargestellt, bietet gerade für die Erkennung von Eutererkrankungen die Überwachung der Merkmalsdifferenzen zwischen einzelnen Vierteln große Vorteile. Allerdings muss hier auch erwähnt werden, dass die im vorliegenden Fall erfassten Werte für die Leitfähigkeit Durchschnittswerte über das Gesamtgemelk hinweg sind. Nach WOOLFORD et al. (1998) ist allerdings die elektrische Leitfähigkeit des Vorgemelks für Aussagen über den Eutergesundheitszustand besser geeignet. Dieser Wert wird jedoch in der Praxis noch nicht separat erfasst. Eine Entwicklung in diese Richtung wäre daher wünschenswert, solange besser geeignete Merkmale wie die somatische Zellzahl noch nicht direkt während des Melkvorgangs erfasst werden können.

Die Herkunft der Daten aus Versuchen in Betrieb 3 ist insofern von Vorteil, dass die Datengüte (im Sinne von Fehltagen, falschen Messwerten, Eindeutigkeit und Zuordenbarkeit von Datensätzen) besser ist als bei den Produktionsbetrieben. Bei Nutzung des gemischten linearen Modells führt die Vorgehensweise, die Tiere jeweils über eine komplette Laktation in einen Versuch zu übernehmen, zum Problem der diskontinuierlich verteilten Laktationsstadien auf Herdenebene. Die Auswirkungen wurden in Abschnitt 4.3 dargestellt. Es zeigt sich, dass für aussagekräftige Schätzungen mit dem gemischten linearen Modell im Sinne von Unverzerrtheit eine annähernd gleichmäßige Verteilung der Laktationsstadien innerhalb der Herde über den betrachteten Zeitraum unverzichtbar ist, um eine Trennung von Testtageeffekten und Laktationsverlauf zu ermöglichen. Letztendlich stellt die Bereitstellung aussagekräftiger Daten für ein Entscheidungsunterstützungssystem in Milchviehbetrieben eine ebenso große Herausforderung dar, als diese Daten anschließend mit geeigneten Methoden auszuwerten. Der Erfolg der Datenauswertung hängt wesentlich von der verfügbaren Datenqualität ab. Bei Beachtung der eigenen Untersuchung folgt, dass trotz der potentiellen Möglichkeit einer Erfassung der Milchmengen für eine kontinuierliche Datennutzung erhebliche Verbesserungen der Datenqualität hinsichtlich Fehlmengen, Softwarekompatibilität und Stabilität der technischen Systeme erforderlich sind.

5.2 Kriterien für den Vergleich der untersuchten Methoden

Ein wichtiger Aspekt für eine vergleichende Betrachtung der in dieser Arbeit untersuchten Methoden ist die Bewertung hinsichtlich der Zielstellung, im vorliegenden Fall also vor allem der Vergleich von Erkennungs- und Fehlerraten bei der Erkennung von Milchminderungen. Allerdings hängen diese beiden Kriterien sowie die Spezifität maßgeblich von der Qualität der Datenbasis ab, besonders hinsichtlich der Definition und Identifikation von Milchminderungen. In der vorliegenden Arbeit wurde durch die Vorgabe einer Definition für Milchminderungen ein optimaler Fall geschaffen, da auf diese Weise alle zu erkennenden Ereignisse bekannt sind. Dies wäre bei Verwendung der in den Betrieben dokumentierten Erkrankungen nicht möglich gewesen, da dort auch immer von einer Anzahl nicht erkannter oder nicht dokumentierter Erkrankungen auszugehen ist. Zwar wird der Vergleich der einzelnen untersuchten Methoden untereinander dadurch nicht beeinflusst, da alle Auswertungen auf derselben Datenbasis und demselben Informationsstand beruhen. Die mit den Methoden erreichten Ergebnisse würden jedoch insgesamt sehr viel schlechter ausfallen, falls undokumentierte Erkrankungen erkannt werden, die nachträglich nicht bestätigt werden können und somit als Falschalarm gewertet werden müssen. Damit sind vor allem unannehmbar hohe Fehlerraten verbunden.

Erkennungs- und Fehlerrate verhalten sich bei den untersuchten Methoden (mit Ausnahme des EWMA-Charts bei niedrigen λ -Werten) bei Verschiebung von Grenzwerten oder Konfidenzintervallen gleich. Ein verringerter Grenzwert oder entsprechend ein engeres Konfidenzintervall machen eine Erkennung von Abweichungen leichter, gleichzeitig werden aber auch häufiger zufällige Schwankungen einen Fehllarm auslösen. Umgekehrt sorgen hohe Grenzwerte und weite Konfidenzintervalle für weniger falsche Alarmer, jedoch entgehen dabei weniger ausgeprägte Abweichungen der Erkennung. Der optimale Fall wäre die eindeutige Klassifizierung der Datensätze in normale Milchmengen (inklusive zufälliger Abweichungen) und begründete Abweichungen. Dies setzt jedoch eine sehr hohe Trennschärfe der verwendeten Methode und die Verwendung geeigneter, hinsichtlich der Fragestellung aussagekräftiger Merkmale voraus. Wenn man gleichzeitig hohe Erkennungsraten, hohe Spezifitäten und niedrige Fehlerraten anstrebt, muss man daher einem dieser Kriterien Priorität einräumen. Der Grenzwert ist dann zunächst auf das Erreichen einer Mindestvorgabe für dieses Kriterium einzustellen.

Bei selten auftretenden Ereignissen (zu denen Erkrankungen bzw. Milchminderungen hier zu zählen sind) wird die Anzahl der wahr negativen Datensätze generell sehr hoch sein. Das führt zu hohen Werten für die Spezifität. Dadurch ist deren Aussagekraft vergleichsweise gering.

Die Erkennungsrate stellt ein wichtiges Kriterium dar, um das Hauptanliegen einer Methode zu bewerten: die Identifizierung von Abweichungen, die über die normalen Schwankungen hinausgehen. Ein entsprechend niedrig gewählter Grenzwert wird jedoch immer dazu führen, dass eine Erkennungsrate von 100% erreicht werden kann. Dies bringt allerdings bei wenig ausgeprägter Trennschärfe der verwendeten Methode eine entsprechend hohe Anzahl falsch positiver Alarmer mit sich, was zu einer hohen Fehlerrate führt.

Die Fehlerrate gibt an, welcher Anteil der auftretenden Alarmer falsch ist. Eine hohe Fehlerrate ist für die Entscheidungsunterstützung in einem Herdenmanagementprogramm nicht hilfreich, da eine Vielzahl von Fehllarmen gleichbedeutend mit einer hohen Zahl zu untersuchender Tiere ist. Damit wird demnach eine Arbeiterschwerung anstelle der angestrebten Unterstützung erreicht. Es ist daher sinnvoll, neben einer hohen Erkennungsrate auch eine möglichst niedrige Fehlerrate zu erzielen, um etwaige Alarmlisten kurz und übersichtlich zu halten.

Eine Beschränkung nur auf die Erzielung einer hohen Erkennungsrate von 80% und mehr ist folglich nicht empfehlenswert. Gleichzeitig sollten Fehlerraten von weniger als 20% angestrebt werden. Durch die Definition der Fehlerrate als Anteil der Falschalarmer an den insgesamt auftretenden Alarmen nimmt ihre Wichtigkeit mit der Herdengröße zu, da entsprechend mehr Tiere kontrolliert werden müssen.

5.3 Vergleich der untersuchten Methoden

Der Vergleich der untersuchten Methoden basiert auf den Kriterien Erkennungsrate, Fehlerrate und Spezifität. Ziel war die Erkennung der in Abschnitt 4.1.1 definierten Milchmengenabweichungen. Dabei ist zunächst eine Unterscheidung zwischen Einmerkmalsmethoden und Mehrmerkmalsmethoden zu treffen.

Die in der vorliegenden Arbeit untersuchten Einmerkmalsmethoden umfassen EWMA-Charts und gemischte lineare Modelle. Verwendetes Merkmal bei beiden Methoden war die Milchmenge. Beide Methoden beruhen auf dem Prinzip der Vorgabe oberer und unterer Grenzwerte bzw. Konfidenzintervalle, innerhalb derer sich die Milchmenge im Normalfall bewegen soll. Mit dem EWMA-Chart wird untersucht, ob der gewichtete gleitende Mittelwert der Milchmenge innerhalb dieser Grenzen liegt. Beim gemischten linearen Modell wird überprüft, ob die beobachtete Milchmenge innerhalb des für den Vorhersagewert gegebenen Konfidenzintervalls liegt. Für die Entscheidung, ob es sich um eine Milchmengenabweichung handelt, wird jeweils nur eine Merkmalsausprägung (EWMA bzw. beobachtete Milchmenge) herangezogen. Zwar werden im gemischten linearen Modell mehrere Einflussgrößen zur bestmöglichen Anpassung der Vorhersagewerte an den tatsächlichen Laktationskurvenverlauf berücksichtigt, doch dies dient nur der Bereitstellung des Vorhersagewertes und somit der Lage der oberen und unteren Grenzwerte für das einzelne Merkmal beobachtete Milchmenge.

Mit den EWMA-Charts wurden die günstigsten Ergebnisse bei hohen λ -Werten und niedrigem L-Wert erreicht. Dies entspricht einer stärkeren Gewichtung der aktuellen Daten und einem schmalen Konfidenzintervall. Die höchsten Erkennungsraten wurden in allen drei Betrieben bei einem λ -Wert von 0.9 und einem L-Wert von 0.5 erzielt. Die Erkennungsraten betragen 62.15% bei Betrieb 1, 63.87% bei Betrieb 2 und 57.73% bei Betrieb 3. Die Spezifitäten für diese Variante liegen zwischen 95.83% und 97.75%, sind also wie erwartet hoch. Das Problem neben den noch zu niedrigen Erkennungsraten sind vor allem die Fehlerraten von 57.17% bis 80.95%. Dies ist vor allem mit der Anpassung der Grenzwerte an den Verlauf der Laktationskurve zu begründen. Während der EWMA sich durch den hohen λ -Wert gut dem Laktationsverlauf anpasst, erfolgt die Anpassung der Grenzwerte nur verzögert. Dadurch liegt der EWMA beispielsweise zu Beginn der Laktation mit dem starken Anstieg der Milchmenge oft außerhalb des Intervalls für „normale“ Werte. Aber auch nach Wechsel von Tieren in eine andere Fütterungsgruppe wurden stärkere Abweichungen festgestellt, auf die der EWMA schnell reagierte, während sich die Grenzwerte nur langsam an das neue Milchmengenniveau anpassten. Dies führt zu vermehrten Fehlalarmen und folglich zu hohen

Fehlerraten. Eine Möglichkeit zur Lösung dieses Problems könnte die Einführung eines Gewichtungsfaktors ähnlich des λ -Wertes sein, mit dem die Grenzwerte sich schneller an das sich während des Laktationsverlaufes verändernde Milchmengenniveau anpassen können. Da die Grenzwerte aber immer noch ein Intervall für den normalen Verlauf der Milchmenge darstellen müssen, sollte der Gewichtungsfaktor kleiner als λ sein, d.h. stärkeres Gewicht auf zurückliegende Daten legen. So würde die Erkennung von akuten Abweichungen durch den Vergleich der hochaktuellen Daten des EWMA bei hohem λ mit normalen Intervallen für den gegenwärtigen Laktationsabschnitt (im Sinne des Milchmengenniveaus der letzten Wochen) ermöglicht. Die Einführung zusätzlicher anzupassender Parameter würde allerdings auch den wichtigen Vorteil der einfachen Erstellung und Nutzung des EWMA-Charts verringern. Ohne diese Anpassung ist ein EWMA-Chart jedoch nicht für die Überwachung von sich über den Zeitverlauf im Niveau ändernden Merkmalen zu empfehlen. Bei Berücksichtigung des Trends einer Laktationskurve ist es dagegen vorstellbar, dass bessere Ergebnisse erzielt werden können. Ebenso könnten multivariate Varianten des EWMA-Charts zusätzliche Merkmale berücksichtigen und so ebenfalls zu besseren Resultaten führen. Eine Verwendung des univariaten EWMA-Charts ist vor allem bei vorwiegend stetigen Merkmalen ohne längerfristige Trends wie beispielsweise der Leitfähigkeit denkbar.

Eine ähnliche Vorgehensweise wie beim EWMA-Chart wird auch mit dem gemischten linearen Modell verfolgt. Auch hier werden Grenzwerte bestimmt, bei deren Über- oder Unterschreiten durch einen Kontrollwert eine nicht zufällige Abweichung der Milchmenge vorhanden ist. Der Kontrollwert hierbei ist die tatsächlich beobachtete Milchmenge je Tag. Die Grenzwerte werden gebildet, indem man von dem mittels des gemischten linearen Modells bestimmten Vorhersagewertes je Tag Vielfache des Standardfehlers der Schätzung subtrahiert bzw. addiert. Ein Faktor von 2, etwa entsprechend des üblichen Konfidenzniveaus von $P = 0.95$, erreichte bei den Betrieben 1 und 3 mindestens die angestrebte Erkennungsrate von 80%, bei Betrieb 2 wurden damit Erkennungsraten zwischen 76% und 78% erzielt. Für die Verbesserung der Kriterien erscheint vor allem eine möglichst gute Anpassung der Vorhersagewerte an den aktuellen tatsächlichen Laktationskurvenverlauf wichtig zu sein. Mit Ausnahme von Betrieb 3 erzielte jeweils das Modell mit der kürzeren Datenvorlaufzeit von 60 Tagen das beste Ergebnis. In Betrieb 3 führte die Verwendung des gesamten verfügbaren Datensatzes zur besten Erkennungsrate. Insgesamt unterscheiden sich die Varianten mit unterschiedlichem Datenvorlauf innerhalb der jeweiligen Betriebe jedoch nur um zwei bis vier Prozentpunkte. Als Referenz erzielten die Modelle mit 60 Tagen Datenvorlauf und bei Verwendung des zweifachen Standardfehlers als Grundlage zur Berechnung der oberen und unteren Grenzwerte Erkennungsraten von 86.7% in Betrieb 1, 77.79%

in Betrieb 2 und 83.37% in Betrieb 3. Die Spezifität lag jeweils im erwarteten hohen Bereich von 97.33% bis 99.09%. Aber während im Vergleich zu den EWMA-Charts mit dem gemischten linearen Modell deutlich bessere Erkennungsraten erreicht wurden, zeigte sich auch hier das Problem der hohen Fehlerraten. Diese reichten von 50.08% im günstigsten Fall in Betrieb 2 über 76.14% in Betrieb 1 bis zu 80.15% in Betrieb 3. Diese hohen Fehlerraten sind vor allem mit ungenügender Anpassung der Vorhersagewerte an die tatsächlichen Milchmengen in der Zeit nach einer beispielsweise krankheitsbedingten stärkeren Milchmengenabweichung zu erklären. Oft erreichen die Tiere nach einer Erkrankung nicht wieder das für die Laktationsphase übliche Milchmengenniveau. Die notwendige Anpassung der Vorhersagewerte an das neue, niedrigere Niveau erfolgt, ähnlich wie die Anpassung der Grenzwerte an den Laktationsverlauf beim EWMA-Chart, nur verzögert. Während sich der Vorhersagewert noch auf dem alten Milchmengenniveau befindet, kann es daher durch die ständig außerhalb des unteren Grenzwertes liegenden tatsächlichen Milchmengen zu gehäuften Fehlalarmen kommen. Andererseits führt eine Unterschätzung der tatsächlichen Milchmenge über mehrere Tage hinweg zu reduzierten Erkennungsraten, wenn die Unterschätzung groß genug ist und Milchmengen trotz vorliegender Abweichung den unteren Grenzwert nicht unterschreiten. Eine Möglichkeit zur Verbesserung der Anpassung der Vorhersagewerte könnte die Einbeziehung von Zusatzinformationen über Erkrankungen des Tieres während der aktuellen Laktation sein, die für eine Niveauekorrektur sorgen könnten. Dies setzt allerdings eine möglichst fehlerfreie und lückenlose Dokumentation der Erkrankungen voraus, die im vorliegenden Fall nicht gegeben war. Ebenfalls wird dafür die tägliche aktuelle Führung einer Krankheitsdatenbank benötigt, während die in der vorliegenden Arbeit untersuchte Modellvariante auf vorhandene Bestandsdaten und Daten aus dem Prozessrechner der Melkanlage zugreift und daher die Berechnung automatisiert stattfinden könnte.

Die in der vorliegenden Arbeit untersuchten Mehrmerkmalsmethoden umfassen Fuzzy Logic-Systeme und künstliche neuronale Netze. Ihre Funktionsweise unterscheidet sich von den vorgestellten Einmerkmalsmethoden dahingehend, dass keine oberen und unteren Grenzwerte mehr für das untersuchte Merkmal vorgegeben werden. Stattdessen wird ein Datensatz, bestehend aus mehreren Merkmalswerten, mit Hilfe der in der jeweiligen Methode vorgesehenen Mechanismen klassifiziert - beispielsweise entweder als „normal“ oder „abweichend“, aber es sind auch weiter unterteilte Klassifizierungen denkbar. Der Vorteil eines Fuzzy Logic-Systems beruht dabei darauf, dass basierend auf Expertenwissen Regeln aufgestellt werden können, die daraufhin rechnerisch umgesetzt werden können und eine Klassifizierung von Datensätzen ermöglichen. Aber auch wenn, wie im vorliegenden Fall, wenig Expertenwissen verfügbar ist, lässt sich intuitiv ein entsprechendes

Regelgerüst erstellen. Allerdings kann dann die Anpassung und Verbesserung des Systems sehr zeitaufwendig sein, wenn die Lage einzelner Zugehörigkeitsfunktionen systematisch verändert und die Auswirkung auf das Gesamtergebnis überprüft werden muss.

Im vorliegenden Fall wurden Modelle erstellt, bei denen als Ergebnis ein Wert zwischen 0 und 1 ausgegeben wurde. Hohe Werte deuten auf eine Milchmengenabweichung hin, niedrige Werte kennzeichnen normale Datensätze. Um das Verhalten von Erkennungs- und Fehlerraten zu untersuchen, wurden drei unterschiedliche Grenzwerte zur Klassifizierung genutzt: 0.5, 0.6 und 0.7. Es wurden außerdem Modellvarianten berechnet, die einerseits auf allgemeinen und andererseits speziell angepassten Zugehörigkeitsfunktionen für die drei untersuchten Betriebe beruhten. Zur Verdeutlichung der Wirkung des Abweichungsfaktors wurden zusätzlich betriebsspezifische Varianten ohne dessen Berücksichtigung berechnet, wobei die der Berechnung zugrunde liegende Datenbasis ebenfalls variiert wurde.

Um die angestrebte Erkennungsrate von mindestens 80% zu erzielen, musste in allen drei Betrieben ein Grenzwert von 0.5 gewählt werden. Höhere Grenzwerte wiesen zwar geringere Fehlerraten auf, allerdings sanken die Erkennungsraten in starkem Maße.

Das günstigste Modell in Betrieb 1 war das Modell mit allgemeinen Zugehörigkeitsfunktionen. Es erreichte eine Erkennungsrate von 79.5% bei einer gleichzeitigen Fehlerrate von 27.6%. Das entsprechende betriebsspezifische Modell erreichte zwar eine um 4.8 Prozentpunkte bessere Erkennungsrate, jedoch auch eine um 7.9 Prozentpunkte schlechtere Fehlerrate. Das Modell ohne Abweichungsfaktor erreichte demgegenüber eine weiter um 3.5 Prozentpunkte verbesserte Erkennungsrate, führte aber ebenfalls zu einer weiteren Verschlechterung der Fehlerrate um 11.2 Prozentpunkte. Dabei wurde die Datenbasis genutzt, in der auch die weiter zurückliegenden Daten verfügbar waren.

In Betrieb 2 führten dagegen die auf den aktuellen Daten der vergangenen 60 Tage basierenden Modelle zu den besten Ergebnissen. Das allgemeine Modell erwies sich mit einer Erkennungsrate von 77.3% und einer Fehlerrate von 14% in allen Kriterien besser als das vergleichbare betriebsspezifische Modell ohne Berücksichtigung des Abweichungsfaktors. Die Nutzung des betriebsspezifischen Modells mit Abweichungsfaktor führte zu einer Verbesserung der Fehlerrate um 7.1 Prozentpunkte, gleichzeitig aber auch zu einer Verschlechterung der Erkennungsrate um 7.9 Prozentpunkte.

Betrieb 3 unterschied sich von den beiden anderen Betrieben durch geringere Standardabweichungen sowohl bei absoluten als auch relativen Abweichungen der Milchmenge, wodurch die Unterscheidung normaler und abweichender Datensätze den anderen Betrieben gegenüber erschwert war. Daher weisen hier die betriebsspezifischen Modelle mit den höchsten

Erkennungsraten (92.5% ohne AF; 79.5% mit AF) vergleichsweise hohe Fehlerraten von 58.8% bzw. 47.2% auf. Das allgemeine Modell erreicht nur eine Erkennungsrate von 59.3%, allerdings ist hier die Fehlerrate mit 20.1% am geringsten.

Insgesamt führt die Verwendung von Fuzzy Logic gegenüber den Einmerkmalsmethoden vor allem zu deutlich verbesserten Fehlerraten, auch durch Nutzung des aus dem gemischten linearen Modell abgeleiteten Abweichungsfaktors. Die Erkennungsraten bewegen sich in den besten Fällen auf Höhe derjenigen des gemischten linearen Modells und liegen über denen der EWMA-Charts. Auffällig waren hier auch die unterschiedlichen Auswirkungen des Datenvorlaufs zur Berechnung des AF. In Betrieb 1 erzielten die Modelle mit den meisten verfügbaren Daten die besten Ergebnisse, in Betrieb 2 die schlechtesten, auch wenn sich die Unterschiede je Betrieb nur im Bereich von einem bis zwei Prozentpunkten bewegen. Die verfügbare Datenmenge war in Betrieb 2 zwar grundsätzlich größer, zu Vergleichszwecken jedoch auf annähernd gleiche Zeiträume begrenzt. Dies lässt den Schluss zu, dass die aktuellen Daten zur Erkennung von Abweichungen von größerer Bedeutung sind, an dieser Stelle aber der lückenhaftere Datenbestand in Betrieb 1 dazu führt, dass die Ergebnisse unter Einbeziehung aller verfügbaren Daten bessere Ergebnisse erzielen. Bei vollständiger Datenbasis sollte ein Datenvorlauf von einem bis zwei Monaten daher geeignete Ergebnisse erzielen. Fuzzy Logic-Systeme sind dabei vor allem vorteilhaft, wenn auf vorhandenes Expertenwissen zurückgegriffen werden kann und sich Regelsystem sowie Zugehörigkeitsfunktionen leicht aufstellen lassen. Sind diese Bedingungen nicht gegeben, kann durch systematisches Probieren ebenfalls eine Lösung gefunden werden, jedoch ist dies mit hohem Aufwand verbunden, da es keine formalen Methoden zur Anpassung eines Fuzzy Logic-Systems gibt. Bei zu starker Ausrichtung der Zugehörigkeitsfunktionen an betriebsspezifischen Daten besteht zudem die Gefahr des overfitting. Damit können zwar anhand der verfügbaren Daten gute Ergebnisse erzielt werden, diese Resultate können jedoch nicht bestätigt werden, wenn das zu spezielle Modell auf andere bzw. neue Tiere in der Herde angewandt wird.

Künstliche neuronale Netze sind nicht wie Fuzzy Logic-Systeme auf Expertenwissen oder zumindest ein intuitives Verständnis der zu klassifizierenden Vorgänge angewiesen, sondern sind in der Lage, durch Training die nötigen Klassifikationsregeln selbst zu „erlernen“. Hier zeigt sich bereits ein wichtiges notwendiges Kriterium für die Nutzung eines KNN: der für das Training benötigte Datensatz von hervorragender Qualität, vor allem hinsichtlich der Klassifizierung. Dies war im vorliegenden Fall durch die notwendige Definition der Mindermengen gegeben. Dadurch war es möglich, in allen drei Betrieben gute Ergebnisse zu erzielen. In Betrieb 1 wurden bei einem Grenzwert von 0.5 und 60 Tagen Datenvorlauf zur Berechnung des AF eine Erkennungsrate von

84.5% und eine Fehlerrate von 16.7% erzielt. Betrieb 2 erzielte bei diesen Vorgaben eine Erkennungsrate von 88.7% und eine Fehlerrate von 12.9%, bei 240 Tagen Datenvorlauf zu Berechnung des AF und einem Grenzwert von 0.3 wurde sogar eine Erkennungsrate von 86.3% bei einer Fehlerrate von 17.9% erreicht. Letztere Vorgaben führten auch in Betrieb 3 zum besten Ergebnis, einer Erkennungsrate von 82% bei einer Fehlerrate von 22.4%. Damit wurden mit dem KNN in den Betrieben 1 und 2 die angestrebten Kriterien von mindestens 80% Erkennungsrate und gleichzeitig höchstens 20% Fehlerrate erreicht. In Betrieb 3 ist das Ergebnis zumindest sehr nahe an diesen Kriterien. Das verwendete Modell mit nur fünf Knoten in der verdeckten Schicht sollte gegenüber overfitting unempfindlich sein. Die Resultate im vorangegangenen Schritt der Modellwahl deuten auf eine gute Generalisierungsfähigkeit hin. Die Fuzzy Logic-Modelle schnitten demgegenüber schlechter ab, was den Schluss zulässt, dass dort noch nicht die ideale Lage und Form der einzelnen Zugehörigkeitsfunktionen gefunden wurde. In diesem Fall hat das KNN den Vorteil, auch ohne Hintergrundinformation über einen Sachverhalt (wie im Fuzzy Logic-System durch das Regelgerüst und die Zugehörigkeitsfunktionen gegeben) sehr gute Ergebnisse erzielen können. Allerdings handelt es sich um ein Black-Box-System, das nicht einfach verständlich ist und Vertrauen seitens des Anwenders erfordert. Eine interessante Weiterführung könnte daher in der Kombination beider Methoden bestehen. Diese sogenannten Neuro-Fuzzy-Systeme benutzen einerseits das vorgegebene Regelgerüst aus dem Fuzzy Logic-Modell, andererseits können die Definitionspunkte der Zugehörigkeitsfunktionen mittels eines KNN „gelernt“ werden und entlasten den Anwender so im möglicherweise langwierigen Anpassungsprozess.

Während die Mehrmerkmalsmethoden im vorliegenden Fall die besseren Ergebnisse geliefert haben, hängen beide Methoden stark von der Güte der Datenbasis ab. KNNs benötigen vor allem eindeutig klassifizierte Datensätze, um trainiert werden zu können. Fuzzy Logic-Systeme benötigen bei fehlenden Daten ein eigens darauf abgestimmtes Regelsystem, um dennoch ein Ergebnis liefern zu können. Die Notwendigkeit von Trainingsdaten in entsprechend guter Qualität für Fuzzy Logic-Modelle und KNNs kann daher ein Grund sein, einfacher zu handhabende Methoden wie EWMA-Charts einzusetzen. Die Möglichkeit zur Reduzierung von Fehlerraten bei Verwendung von je nach Fragestellung geeigneten Merkmalskombinationen spricht allerdings generell dafür, Methoden anzuwenden, die mehrere Merkmale verarbeiten können.

In Form des Abweichungsfaktors zwischen vorhergesagter und beobachteter Milchmenge sind die Berechnungen des gemischten linearen Modells als Eingabevariablen für die untersuchten Mehrmerkmalsmethoden geeignet und führen in den besten Modellvarianten zu Verbesserungen in

Tabelle 52: Vergleich der untersuchten Methoden

	univariates EWMA-Chart	gemischtes lineares Modell	Fuzzy Logic	künstliches neuronales Netz
Anforderungen an die Datenbasis	Datenvorlauf nötig, um effektive obere und untere Grenzwerte zu bilden	Datenvorlauf nötig, um eine gute Regression zu erhalten (~60d)	Kein Datenvorlauf notwendig, jedoch hilfreich zur Anpassung der Zugehörigkeitsfunktionen; kann mit lückenhaften Daten umgehen, wenn in der Regelbasis vorgesehen	Hervorragende Datenbasis für das Training notwendig; kann mit lückenhaften Daten umgehen
Berechnung	Im Betrieb, eingebunden in ein Herdenmanagementsystem	Dienstleister, täglich bei Erkennung von Abweichungen nötig, anderenfalls je nach Bedarf	Im Betrieb, eingebunden in ein Herdenmanagementsystem	Im Betrieb, eingebunden in ein Herdenmanagementsystem
Anpassungsmöglichkeiten durch den Nutzer	Justierung von λ - und L- Werten	Modellwahl, funktioneller Ansatz zur Modellierung der Laktationskurve, Konfidenzintervall	Regelgerüst, Zugehörigkeitsfunktionen, Grenzwert für Outputs	Wahl des Netzwerkes, Grenzwert für Outputs

Form von geringeren Fehlerraten. Dadurch werden zwar auch die Erkennungsraten reduziert, allerdings ist der Rückgang der Fehlerraten in Prozentpunkten meist größer als die Einbußen bei den Erkennungsraten. Werden gemischte lineare Modelle ohnehin zur Unterstützung des Herdenmanagements durch Schätzung von herdenweiten Milchmengen oder Laktationskurven verwendet, bietet sich daher auch die Nutzung zur tierindividuellen Entscheidungsunterstützung an. Gegenüber den anderen untersuchten Methoden ist dazu allerdings die Auslagerung der vergleichsweise aufwendigen Berechnungen an einen außerbetrieblichen Dienstleister notwendig, während die anderen Methoden in ein übliches Herdenmanagementprogramm eingebunden denkbar sind. Tabelle 52 stellt die untersuchten Methoden noch einmal in einer Übersicht dar.

5.4 Anforderungen an Merkmale zur Entscheidungsunterstützung

Nach der Erfahrung, wonach mit Vorliegen einer Erkrankung (hier vor allem Stoffwechsel- und Eutererkrankungen) nicht zwingend eine negative Abweichung der Milchmenge verbunden ist und im Gegenteil bei einigen Tieren sogar steigende Milchmengen zu verzeichnen waren, muss auf die Eignung von Merkmalen für verschiedene Fragestellungen eingegangen werden.

Wenn Tiere unterschiedlich auf Erkrankungen reagieren, muss eine Methode zur Entscheidungsunterstützung diese unterschiedlichen Reaktionen (in Form von Änderungen der Werte geeigneter Merkmale) erfassen und klassifizieren können. Probleme entstehen, wenn kein offensichtlicher Unterschied zwischen der Merkmalsreaktion eines kranken Tieres und eines gesunden Tieres besteht. In diesem Fall ist eine Eignung des Merkmals zur automatisierten Erkennung nicht gegeben. Zusätzliche Informationen in Form anderer Merkmale werden benötigt, um zu aussagefähigen Ergebnissen zu kommen. Dies ist in Abschnitt 4.2.1 beispielhaft an den Merkmalen Leitfähigkeit und Milchmenge zur Erkennung von Eutererkrankungen dargestellt worden. Einerseits ist nicht bei jeder dokumentierten Eutererkrankung eine negative Abweichung der Milchmenge zu verzeichnen, andererseits sind auch die auf das gesamte Euter bezogenen Werte für die Leitfähigkeit deutlich weniger aussagekräftig als die Werte der einzelnen Viertel. Die Vierteldifferenzen müssen hier trotz der oben angeführten nur bedingten Eignung von Durchschnittswerten des Gesamtgemelks gegenüber dem Vorgemelk als entscheidendes Merkmal gesehen werden. Die elektrische Leitfähigkeit des Gesamtgemelks trägt nur in einem Viertel der Fälle zur Erkennung einer Eutererkrankung bei (vgl. Abschnitt 4.2.1.2), und die Berücksichtigung der Milchmenge führt im Vergleich zum besten nur die Leitfähigkeit betrachtenden Modell nur am zweiten und dritten Tag vor der Dokumentation der Erkrankung zu einem bzw. zwei zusätzlichen Alarmen. Im Vergleich zur Vierteldifferenz dienen diese Merkmale daher nur als unterstützende

Information. Abgeleitet davon ergeben sich Anforderungen für Merkmale, die zur Entscheidungsunterstützung genutzt werden sollen. Sie müssen auf einer geeigneten Aggregationsebene vorliegen, sie sollten bei Abweichungen möglichst eindeutig von zufälligen normalen Schwankungen abgrenzbar sein und zudem möglichst zeitnah nach der Erfassung bereitstehen. Unter diesen Aspekten ist die Milchmenge zwar als unterstützendes Merkmal zur Entscheidungsunterstützung in Milchviehbetrieben denkbar, andere Merkmale werden aber als entscheidendes Kriterium benötigt. Die Vorzüglichkeit viertelweise verfügbarer Merkmale zur Überwachung der Eutergesundheit spricht für die Verwendung mit entsprechenden Sensoren ausgestatteter Melkzeuge. Dies ist derzeit jedoch nur im Rahmen eines AMS sinnvoll zu erreichen. Zusätzlich ist auch das in der vorliegenden Arbeit verwendete Zielmerkmal zu hinterfragen. Die verminderte Milchmenge ist nur eines von verschiedenen Merkmalen, die zur Unterstützung der Überwachung des Gesundheitszustandes von Kühen denkbar sind. Neben der elektrischen Leitfähigkeit der Milch und der Anzahl somatischer Zellen sind des Weiteren beispielsweise auch Aktivität und Liegeverhalten als Merkmale denkbar. Wie in Abschnitt 4.1 dargestellt, ergab sich die Konzentration auf die Erkennung von Milchminderungen aus der Schwierigkeit, die Krankentage in den drei Betrieben im Nachhinein eindeutig festzustellen. Ebenso stellt auch der Wert von 7% als Definition einer Milchmengenabweichung keinen allgemein gültigen Wert dar. Die Unterschiede zwischen den drei untersuchten Betrieben (vgl. Tabelle 8 in Abschnitt 4.1.1) deuten darauf hin, dass betriebspezifisch gewählte Werte zur Definition von Milchminderungen zu anderen Ergebnissen führen können. Ziel muss allerdings langfristig die möglichst frühzeitige Erkennung von Erkrankungen sein. Wichtigste Voraussetzung zur Bereitstellung einer geeigneten Datenbasis dafür ist allerdings ständige tierärztliche Überwachung und Dokumentation der auftretenden Erkrankungen sowie die sichere und zeitnahe Erfassung der anderen genannten Merkmale. Diese Voraussetzungen sind jedoch derzeit unter Praxisbedingungen wohl nicht zu erfüllen und auch unter Versuchsbedingungen nur mit enormem Aufwand zu leisten.

6 Zusammenfassung

Zielstellungen der vorliegenden Arbeit waren die Bewertung der verfügbaren Daten hinsichtlich der Datenqualität und der sich daraus ergebenden Nutzungsmöglichkeiten oder -einschränkungen, die Modellbildung und -anpassung für ein gemischtes lineares Modell zur tierspezifischen Vorhersage von Milchmengen unter Berücksichtigung verschiedener Ansätze für die Beschreibung des Laktationskurvenverlaufs, der Vergleich der Methoden EWMA-Chart, gemischtes lineares Modell, Fuzzy Logic und künstliche neuronale Netze, basierend auf Erkennung von Abweichungen beim Merkmal Milchmenge, und die Darstellung der Möglichkeiten der organisatorischen Einordnung der untersuchten Methoden innerhalb des betrieblichen Informationsmanagements.

Die Datenqualität unterschied sich in den untersuchten Betrieben hinsichtlich der Zahl der erfassten Merkmale, der Zuverlässigkeit der Datenerfassung und der Plausibilität der Daten. Fehlzuordnungen, Datenlücken und Softwarekompatibilitätsprobleme würden die Nutzbarkeit der untersuchten Methoden im Praxisbetrieb einschränken.

Bei der Wahl des besten linearen gemischten Modells erwies sich ein Random-Regression-Modell mit tier- und laktationsspezifischen Effekten zur Schätzung von täglichen Milchmengen als am Besten geeignet.

Die Modellierung der Laktationskurve mit dem Ansatz von ALI and SCHAEFFER (1987) lieferte die beste Anpassung der geschätzten Laktationskurve an den tatsächlichen Kurvenverlauf.

Die errechneten Werte für feste und zufällige Effekte aus dem gemischten linearen Modell eignen sich zur Schätzung von Herdenmilchmengen für kurzfristige Entscheidungen im Milchkontingentmanagement, zur Erstellung von herdenweiten laktationsspezifischen Laktationskurven im Hinblick auf Herdenmanagement und Überwachung von Zuchtzielen und zur Berechnung kurzfristiger Trends für die Milchmenge, die zu Vergleichszwecken mit der tatsächlichen Milchmenge in weiteren Methoden genutzt werden können. Dieser Vergleich wird in Form eines Abweichungsfaktors der beiden Trends voneinander quantifiziert.

Das gemischte lineare Modell allein sowie univariate EWMA-Charts sind wegen der hohen Fehlerraten nur bedingt als Methoden zur Erkennung von Milchmengenabweichungen zu empfehlen.

Bei der Untersuchung der Leitfähigkeit als Merkmal zur Erkennung von Eutererkrankungen erwiesen sich verfügbare Werte je Euterviertel einem Durchschnittswert des Gesamteuters gegenüber als deutlich überlegen und nutzbar. Die viertelweise Erfassung ist sinnvoll, aber derzeit auf ein AMS beschränkt.

Fuzzy Logic und künstliche neuronale Netze sind als Methoden zur Erkennung von Milchmengenabweichungen aus methodischer Sicht sehr gut geeignet. Bei entsprechender Modellierung können hohe Erkennungsraten bei gleichzeitig möglichst niedrigen Fehlerraten erzielt werden. Dabei ist vor allem die Möglichkeit der Kombination mehrerer Merkmale bzw. verschiedener Aspekte eines Merkmals als entscheidendes Kriterium für eine bessere Bewertung gegenüber den untersuchten Einmerkmalsmethoden zu nennen.

Der vermutete Zusammenhang zwischen Milchminderungen und Erkrankungen konnte anhand der vorliegenden Daten nicht durchgehend beobachtet werden.

Wird der Abweichungsfaktor als Zusatzmerkmal bei einer der Mehrmerkmalsmethoden verwendet, ist mit einer verbesserten Fehlerrate zu rechnen.

Für alle untersuchten Methoden gilt, dass eine kontinuierlich geführte, konsistente und möglichst fehlerfreie Datenbasis und Dokumentation zur Verfügung stehen muss, um Nutzen für das Herdenmanagement zu ziehen. Sind für ein Tier nicht täglich Istwerte von verwendeten Merkmalen vorhanden, ist der Vergleich mit Sollwerten nicht möglich, und es können keine aktuellen Alarmlisten erstellt werden.

Organisatorisch lassen sich die Methoden EWMA-Charts, Fuzzy Logic und künstliche neuronale Netze innerhalb eines betrieblichen Herdenmanagementsystems realisieren, während für Berechnungen eines gemischten linearen Modells außerbetriebliche Dienstleister sinnvoller erscheinen.

Aims of the paper at hand were the evaluation of the available data regarding data quality and the resulting possibilities or restrictions of use, modelling and fitting of a mixed linear model for animal-specific predictions of milk yields taking into account different approaches for the modelling of lactation curves, the comparison of the methods EWMA-chart, mixed linear model, fuzzy logic and artificial neural networks based on the detection of deviations of the trait milk yield and to submit a statement of the options for integrating the methods under examination into the on-farm information management organisationally.

Data quality differed between the farms under examination regarding number of documented traits, reliability of data recording and plausibility of the data. False mapping of records, missing data or problems with software compatibility would restrict the investigated methods' usability in practice. When choosing the suited mixed linear model, a random-regression model with animal- and lactation-specific effects for predicting daily milk yields was found best.

Modelling the lactation curves following the approach of ALI and SCHAEFFER (1987) provided the best fit of the estimated lactation curve to the real curve progression.

The values for fixed and random effects calculated with the mixed linear model are suited for estimating herd-wide milk yields for short-term decisions within management of milk contingents, for providing herd-wide lactation-specific lactation curves for herd management and monitoring of goals in breeding and for calculation of short-term trends for milk yields that can be used for comparison with actual milk yields in other methods. This comparison is quantified in form of a deviation factor between those trends.

As a result of the high error rates the mixed linear model as well as the base form of the univariate EWMA-chart that was used can be recommended as methods for use in detection of milk yield deviations only to limited extent.

Upon investigation of electrical conductivity of milk for use as a trait for detecting udder diseases the available values per quarter proved to be noticeable better and usable compared to an average value throughout the udder. It makes sense to record electrical conductivity for each quarter separately, but this is currently restricted to automatic milking systems.

Fuzzy logic and artificial neural networks are well-suited methods for detection of deviations in milk yield from a methodical point of view. Provided appropriate modelling, high detection rates with at the same time as low error rates as possible can be achieved. Especially the option to easily

combine different traits or different aspects of a trait is to be stated as the essential criterion for an improved validation compared to the methods under investigation that used only single traits.

The assumed correlation between lowered milk yields and diseases could not be observed continuously based on the available data.

When the deviation factor is used as an additional trait in one of the multi-trait methods, an improved error rate can be expected.

All methods under investigation require a continuous, consistent and as error-free as possible data base and documentation to be useful for herd management. If daily actual values are not available for an animal, comparison with reference values is impossible and up-to-date alarm lists cannot be provided.

EWMA-charts, fuzzy logic and artificial neural networks can be implemented within on-farm herd management organisationally, while external service providers seem to be preferable for calculations with the mixed linear model.

8 Literaturverzeichnis

- AKAIKE, H. (1969): Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, 21, 243-247.
- ALI, T.E.; SCHAEFFER, L.R. (1987): Accounting for covariance among test day milk yields in dairy cows. *Canadian Journal of Animal Science* 67: 637-644.
- ALTROCK, C. v. (1997): *Fuzzy Logic and NeuroFuzzy Applications in Business and Finance*. Prentice Hall PTR, Upper Saddle River, New Jersey.
- BORGELT, C.; KLAWONN, F.; KRUSE, R.; NAUCK, D. (2003): *Neuro-Fuzzy-Systeme*. Vieweg, Wiesbaden.
- CAVERO, D.; TÖLLE, K.-H.; BUXADÉ, C; KRIETER, J. (2006a): Mastitis detection in dairy cows by application of fuzzy logic. *Livestock Science* 105: 207-215.
- CAVERO, D.; TÖLLE, K.-H.; HENZE, C.; BUXADÉ, C.; KRIETER, J. (2006b): Mastitis detection in dairy cows by application of neuronal networks. *Livestock Science*, in Druck.
- DE MOL, R.M.; KROEZE, G.H.; ACHTEN, J.M.F.H; MAATJE, K.; ROSSING, W. (1997): Results of a multivariate approach to automated oestrus and mastitis detection. *Livestock production Science* 48, 219-227.
- DE MOL, R.M. (2000): Automated detection of oestrus and mastitis in dairy cows. Thesis Wageningen University.
- DE MOL, R.M.; WOLDT, W.E. (2001): Application of Fuzzy Logic in Automated Cow Status Monitoring. *Journal of Dairy Science* 84: 400-410.
- DEMUTH, H.; BEALE, M. (2004): *Neural Network Toolbox User's Guide*. The MathWorks, Inc.
- FIRK, R.; STAMER, E.; JUNGE, W.; KRIETER, J. (2003a): Oestrus detection in dairy cows based on serial measurements using univariate and multivariate analysis. *Archiv für Tierzucht, Dummerstorf* 46: 127-142.
- FIRK, R.; STAMER, E.; JUNGE, W.; KRIETER, J. (2003b): Improving oestrus detection by combination of activity measurements with information about previous oestrus cases. *Livestock Production Science* 82, 97-103.
- GUO, Z. (1998): Modelle zur Beschreibung der Laktationskurve des Milchrindes und ihre Verwendung in Modellen zur Zuchtwertschätzung. Dissertation, Universität Göttingen.
- HEALD, C.W.; KIM, T.; SISCHO, W.M.; COOPER, J.B.; WOLFGANG, D.R. (2000): A Computerized Mastitis Decision Aid Using Farm-Based Records: An Artificial Neural Network Approach. *J. Dairy Sci.* 83, 711-722.
- HENDERSON, C.R. (1990): Statistical methods in animal improvement: historical overview, In: *Advances in Statistical Methods for Genetic Improvement of Livestock*. Springer, New York.

- HEUER, J. (1997): Neuronale Netze in der Industrie: Einführung - Analyse – Einsatzmöglichkeiten. Deutscher Universitäts-Verlag, Wiesbaden.
- HURVICH, C.M.; TSAI, C.L. (1989): Regression and time series model selection in small samples. *Biometrika* 76, 297-397.
- JAMROZIK, J.; SCHAEFFER, L.R.; DEKKERS, J.C.M. (1997): Genetic Evaluation of Dairy Cattle Using Test Day Yields and Random Regression Model. *J. Dairy Sci.* 80, 1217-1226.
- KOCH, M.; KUHN, TH.; WERNSTEDT, J. (1996): Fuzzy Control, Optimale Nachbildung und Entwurf optimaler Entscheidungen. R. Oldenbourg Verlag, München.
- KÖHLER, S. (2003): Nutzung von Prozessparametern automatischer Melksysteme für die Erkennung von Eutererkrankungen unter Verwendung der Fuzzy Logic. *Berichte aus der Agrarwissenschaft*, Shaker Verlag GmbH, Aachen.
- KRIETER, J.; STAMER, E.; JUNGE, W. (2005): Oestrus detection in dairy cows using control charts and neural networks. *Papers of 56th Annual Meeting of the EAAP 2005*, Uppsala.
- LEHMANN, E.L. (1997): *Testing Statistical Hypotheses*. John Wiley & Sons, New York.
- LIU, Z; REINHARDT, F.; BÜNGER, A.; REENTS, R.; RENSING, S. (2003): Das neue Random-Regression-Testtag-Modell. *Milchrind* 12: 12-15.
- MATHWORKS (2000): *Fuzzy Logic Toolbox User's Guide*. The MathWorks, Inc.
- MIELENZ, N.; SPILKE, J.; KREJCOVA, H.; SCHÜLER, L. (2005): Nutzung des Informationskriteriums von Akaike innerhalb von Testtagsmodellen – eine Simulationsstudie. In: Haerting, J.; Becker, C.; Spilke, J.; Kuß, O.; Röhmel, J. (Hrsg.): 51. Biometrisches Kolloquium der Deutschen Region der Internationalen Biometrischen Gesellschaft (DR-IBS) in Halle (Saale), März, Programm und Abstracts, 90.
- MIELENZ, N.; SPILKE, J.; KREJCOVA, H.; SCHÜLER, L. (2006): Statistical analysis of test-day milk yields using random regression models for the comparison of feeding groups during the lactation period. *Archives of Animal Nutrition* 60 (5), 341-357.
- MONTGOMERY, D.C. (1997): *Introduction to statistical quality control*. John Wiley & Sons, New York [u.a.].
- NAUCK, D.; KLAWONN, F.; KRUSE, R. (1996): *Neuronale Netze und Fuzzy-Systeme*. Vieweg, Braunschweig/Wiesbaden.
- NGO, L.; BRAND, R. (1997): Model Selection in Linear Mixed Effects Models using SAS[®] PROC MIXED. Paper 284 in *Proceedings of SAS Users Group International Conference 22*, San Diego CA.
- PACHE, S.; AMMON, C.; BREHME, U.; RUDOVSKY, H.J.; SPILKE, J.; BERGFELD, U. (2005): Oestrus detection in dairy cattle using ALT-Pedometer and Fuzzy-Logic. In: *Book of abstracts of the 56th annual meeting of the European Association for Animal Production (EAAP)*, Bd. 56, 383.

- PATTERSON, D.W. (1997): Künstliche neuronale Netze: das Lehrbuch. Prentice Hall Verlag GmbH, Haar bei München [u.a.].
- RITTER, H.; MARTINETZ, T.; SCHULTEN, K. (1994): neuronale Netze: Eine Einführung in die Neuroinformatik selbstorganisierender Netzwerke. Addison-Wesley, Bonn [u.a.].
- ROBERTS, S.W. (1959): Control chart tests based on geometric moving averages. *Technometrics* 1, 239-250.
- SALEHI, F.; LACROIX, R.; WADE, K.M. (1998): Improving dairy yield predictions through combined record classifiers and specialized artificial neural networks. *Comput. Electron. Agric.* 20, 199-213.
- SPIELKE, J.; MIELENZ, N. (2006): Vergleich von Kriterien zur Modellwahl bei der multiplen linearen Regression. In: Kaiser, K.; Bödeker, R.-H. (Hrsg.): *Statistik und Datenanalyse mit SAS, Proceedings der 10. Konferenz der SAS-Anwender in Forschung und Entwicklung (KSFE)*. Shaker, 215-223.
- SWALVE, H.H. (2000): Theoretical Basis and Computational Methods for Different Test-Day Genetic Evaluation Methods. *Journal of Dairy Science* 83: 1115-1124.
- VAN BEBBER, J.; REINSCH, N.; JUNGE, W.; KALM, E. (1999): Monitoring Daily Milk Yields with a Recursive Test day Repeatability Model (Kalman Filter). *Journal of Dairy Science* 82: 2421-2429.
- WADA, Y.; KASHIWAGI, N. (1990): Selecting statistical models with information statistics. *Journal of Dairy Science* 73: 3575-3582.
- WENDT, K.; LOTTHAMMER, K.-H.; FEHLINGS, K.; SPOHR, M. (1999): *Handbuch Mastitis*. Kamlage, Osnabrück.
- WILMINK, J.B.M. (1987): Comparison of different methods of predicting 305-day milk yield using means calculated from within-herd lactation curves. *Livestock Production Science* 17: 1-8.
- WOOLFORD, M.W.; WILLIAMSON, J.H.; HENDERSON, H.V. (1998): Changes in electrical conductivity and somatic cell count between milk fractions from quarters subclinically infected with particular mastitis pathogens. *Journal of Dairy Research* 65, 187-198.
- YANG, Y. (1998): *Rechnergestützte Östrusüberwachung bei Milchkühen unter Anwendung der Fuzzy-Logic-Methode*. Herbert Utz Verlag, Wissenschaft, München.
- ZHOU, M. (2003): *Neural Network Techniques and Their Engineering Applications*. In Leondes, C.T. (Hrsg.): *Intelligent Systems: Technology and Applications, Vol. II – Fuzzy Systems, Neural Networks and Expert Systems*, CRC Press LLC, Boca Raton [u.a.], 1-19.
- ZSCHÖCK, M.; KLOPPERT, B.; WOLTER, W.; SEUFERT, H.; SCHWARZ, H.-P.; KÖTTING, C. (1998): *Zellzahlen der Milch in großen Beständen*. Rationalisierungs-Kuratorium für Landwirtschaft (RKL), Schrift 4.2.1.1, Rendsburg.

Erklärung

Hiermit versichere ich, dass ich die eingereichte Dissertation selbständig angefertigt habe und keine anderen als die angegebenen Hilfsmittel verwendet habe.

Halle/Saale, Dezember 2007

Wissenschaftlicher Werdegang

- 04/1996 – 09/2002 Studium der Agrarwissenschaften an der Justus-Liebig-Universität
Gießen
- Spezialisierungsfachrichtung: Wirtschafts- und Sozialwissenschaften
des Landbaus
- Diplomarbeit: „Qualitätssicherung im Lebensmittelsektor – Risiken
und Kontrollverfahren“
- 11/2002 Abschluss: Diplom-Agraringenieur
Abschlussnote: 1,9
- 12/2003 – 01/2007 Wissenschaftlicher Mitarbeiter in der AG Biometrie und
Agrarinformatik der Martin-Luther-Universität Halle-Wittenberg

Danksagung

Mein Dank gilt Prof. Spilke nicht nur für die Überlassung des Themas, sondern vor allem die Betreuung und nicht zuletzt die Geduld mit mir. Zu bedanken habe ich mich ebenfalls bei Prof. Krieter und Prof. Swalve für die Begutachtung der Arbeit. Die Mitarbeiter der Arbeitsgruppe Biometrie und Agrarinformatik, Frau Dr. Warnstorff, Frau Häßler und Herr Giebler sowie die studentischen Hilfskräfte Frau Gehweiler und Herr Hildebrand haben ebenso dazu beigetragen, dass ich die Arbeit in der vorliegenden Form fertig stellen konnte. Auch ihnen gebührt mein Dank.

Ich widme diese Arbeit meinen Eltern, mit Dank für die Unterstützung.

Kleinigkeiten können manchmal große Wirkung zeigen.