

Professur für Unternehmensforschung und Informationsmanagement
der Technischen Universität München

**Data Mining und Wissensentdeckung im Precision Farming -
Entwicklung von ökonomisch optimierten Entscheidungsregeln zur
kleinräumigen Stickstoff-Ausbringung**

Georg Weigert

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Agrarwissenschaften (Dr. agr.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. Dr. h. c. A. Heißenhuber

Prüfer der Dissertation:

1. Univ.-Prof. Dr. P. Wagner,
Martin-Luther-Universität Halle-Wittenberg
2. Univ.-Prof. Dr. H. Auernhammer
3. Univ.-Prof. Dr. L. Reiner, i. R.

Die Dissertation wurde am 22.02.2006 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 23.06.2006 angenommen.

Vorwort

Nach Abschluss meiner Dissertation möchte ich allen herzlich danken, die zum Gelingen dieser Arbeit beigetragen haben:

Herrn Prof. Dr. Wagner für die Überlassung des Themas und die Betreuung, für seine Hinweise und Anregungen, aber vor allem auch für seine Förderung und Unterstützung und den gewährten Freiraum.

Herrn Prof. Dr. Auernhammer, dem Sprecher der DFG Forschergruppe IKB-Dürnast, für seinen unermüdlichen und engagierten Einsatz für das IKB-Projekt sowie für die Übernahme des Koreferats.

Herrn Prof. Dr. Dr. h. c. Heißenhuber für die Übernahme des Prüfungsvorsitzes.

Herrn Prof. Dr. Schmidhalter und Herrn Dr. Maidl für die umfangreiche Unterstützung bei der Versuchsdurchführung.

Danken möchte ich auch allen Kolleginnen und Kollegen an der Professur für Unternehmensforschung und Informationsmanagement, am Lehrstuhl für Volkswirtschaftslehre sowie allen Mitgliedern der DFG Forschergruppe IKB-Dürnast für eine erlebnisreiche und schöne Zeit.

Nicht zuletzt gehört mein Dank meiner Familie und meinen Freunden.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Problemstellung und Zielsetzung	4
1.2	Aufbau dieser Arbeit	5
2	Wissensentdeckung in Datenbanken	8
2.1	Data Mining und Wissensentdeckung in Datenbanken	8
2.2	Die Aufgabe des überwachten Lernens	10
2.3	Der Prozess der Wissensentdeckung in Datenbanken	14
2.4	Datenvorbereitung und Exploration	15
2.4.1	Regeln zur Datenvorbereitung	16
2.4.2	Merkmalsraum und Qualität von Daten	17
2.4.3	Veränderungen des Merkmalsraumes	20
2.5	Data Mining mit künstlichen neuronalen Netzen	23
2.5.1	Aufbau und Technik	24
2.5.2	Lernverfahren	27
2.5.3	Erweiterungen für ein optimiertes Lernen	29
2.6	Data Mining mit Entscheidungsbaumverfahren	34
2.6.1	Aufbau und Technik	35
2.6.2	Konstruktion eines Entscheidungsbaumes	37
2.6.3	Pruning-Techniken und spezialisierte Algorithmen	39
2.7	Verifikation, Evaluation und Erkenntnisgewinn	40
2.7.1	Leistungsvorhersage bei begrenztem Datenumfang	40
2.7.2	Erkenntnisgewinn	43
3	Forschungsstand der Bereiche Stickstoffdüngung, Precision Farming und Data Mining	45
3.1	Schnittmenge I: Stickstoffdüngung und Precision Farming	45
3.1.1	Stickstoffdüngung bei Winterweizen	46

3.1.2	Teilflächenspezifische Stickstoff-Applikation	50
3.1.3	Kleinräumige Sensordaten	55
3.2	Schnittmenge II: Precision Farming und Data Mining	61
4	Entwicklung von Entscheidungsregeln zur kleinräumigen Stickstoff-Ausbringung	68
4.1	Ziel und Überblick eines modifizierten WED-Prozesses	68
4.2	Material: Versuchsanlagen und Versuchsdurchführung	70
4.2.1	Versuchsanlagen und Versuchsdurchführung TU München-Weihenstephan.....	72
4.2.2	Versuchsanlage und Versuchsdurchführung Görzig	78
4.2.3	Witterung und Pflanzenentwicklung im Untersuchungszeitraum.....	83
4.2.4	Standortbeschreibung	85
4.3	Datenquellen, Datenattribute und Datenvorverarbeitung	85
4.3.1	Datenvorverarbeitung IKB-Versuche	87
4.3.2	Datenvorverarbeitung der Versuche in Görzig.....	91
4.3.3	Fazit Datenvorverarbeitung	94
4.4	Umsetzung: Entwicklung von Entscheidungsregeln und Extraktion von Mustern	95
4.4.1	Teilflächenspezifische Ertragsprognose	96
4.4.2	Ökonomische Optimierung von N-Gaben mittels trainierter neuronaler Netze zur Ertragsprognose.....	101
4.4.3	Extraktion von ökonomisch optimierten Entscheidungsregeln.....	104
4.4.4	(Vor-)Überlegungen zur Extraktion von Mustern.....	105
4.5	Data Mining I: Versuch Schafhof 02	111
4.5.1	Datenexploration	111
4.5.2	Ertragsprognose für die zweite N-Applikation	114
4.5.3	Ertragsprognose zur dritten Stickstoff-Applikation	116
4.5.4	Zwischenfazit	119
4.6	Data Mining II: Versuch D4 03.....	119
4.6.1	Datenexploration	120

4.6.2	Data Mining bei begrenztem Datenumfang	122
4.6.3	Ertragsprognose und Entscheidungsregeln zur ersten N-Applikation.....	123
4.6.4	Ertragsprognose und Entscheidungsregeln zur zweiten N-Applikation.....	128
4.6.5	Ertragsprognose und Entscheidungsregeln zur dritten N-Applikation.....	131
4.6.6	Zwischenfazit	135
4.7	Data Mining III: Versuch Görzig 04	136
4.7.1	Datenexploration	136
4.7.2	Erste Stickstoff-Applikation.....	141
4.7.3	Zweite Stickstoff-Applikation.....	148
4.7.4	Dritte Stickstoff-Applikation.....	153
4.7.5	Zwischenfazit	157
4.8	Verifikation und Erkenntnisgewinn des WED-Prozesses	159
4.8.1	Ergebnisse der Feldversuche	159
4.8.2	Extrahierte Muster	163
5	Diskussion und Schlussfolgerungen.....	168
5.1	Zusammenfassung und Interpretation der Ergebnisse	168
5.2	Wissensentdeckung in Datenbanken und Precision Farming.....	173
5.3	Verbesserungspotenzial und Ausblick	177
6	Zusammenfassung.....	182
7	Literatur	186
8	Anhang	199

Tabellenverzeichnis

Tabelle 1: Leistungsmaße für die numerische Vorhersage	42
Tabelle 2: Anforderungen an einen Precision-Farming-Ansatz aus ökonomischer Sicht....	55
Tabelle 3: Übersicht von Arbeiten zur Ertragsprognose mit neuronalen Netzen.....	64
Tabelle 4: Übersicht der durchgeführten Feldversuche.....	71
Tabelle 5: Düngesysteme und Faktorstufen des Versuches <i>Schafhof 02</i>	74
Tabelle 6: Düngesysteme und Faktorstufen des Versuches <i>D4 03</i>	76
Tabelle 7: Düngesysteme und Faktorstufen des Versuches <i>Schafhof 04</i>	77
Tabelle 8: Faktorstufen der Variante <i>Merkmalsraum</i> (Versuch <i>Görzig 04</i>)	80
Tabelle 9: Düngesysteme und Faktorstufen des Versuches <i>Görzig 04</i>	81
Tabelle 10: N-Applikationen der untersuchten Varianten im Versuch <i>Görzig 05</i>	83
Tabelle 11: Das Spektrum der verwendeten teilflächenspezifischen Attribute.....	86
Tabelle 12: Beispielhafter Datensatz zur Erzeugung von Entscheidungsregeln	104
Tabelle 13: Analyse eines trainierten Netzes – vorgestellt an einem Beispiel.....	108
Tabelle 14: Neuronales Netzwerk zur zweiten N-Applikation (Versuch <i>Schafhof 02</i>)	115
Tabelle 15: Neuronales Netzwerk zur dritten N-Applikation (Versuch <i>Schafhof 02</i>)	118
Tabelle 16: Neuronales Netzwerk zur ersten N-Applikation (Versuch <i>D4 03</i>)	124
Tabelle 17: Neuronales Netzwerk zur zweiten N-Applikation (Versuch <i>D4 03</i>)	128
Tabelle 18: Neuronales Netzwerk zur dritten N-Applikation (Versuch <i>D4 03</i>)	132
Tabelle 19: Korrelationsmatrix des Datensatzes <i>Görzig 04</i>	139
Tabelle 20: Neuronales Netzwerk zur ersten N-Applikation (Versuch <i>Görzig 04</i>)	141
Tabelle 21: Balance-Regeln zur Optimierung des Trainings	143
Tabelle 22: Neuronales Netzwerk zur ersten N-Applikation (Versuch <i>Görzig 04</i> , mit Datenbalance)	143
Tabelle 23: Ergebnisse der Kreuzvalidierung für zwei Datenaufbereitungsmethoden	144
Tabelle 24: Neuronales Netzwerk zur zweiten N-Applikation (Versuch <i>Görzig 04</i>).....	148
Tabelle 25: Ergebnisse der Kreuzvalidierung zur zweiten N-Applikation.....	149
Tabelle 26: Neuronales Netz zur dritten N-Applikation (Versuch <i>Görzig 04</i>)	153
Tabelle 27: Ergebnisse des Versuches <i>Görzig 05</i>	161
Tabelle A-28: Algorithmus zur dritten N-Gabe (<i>Schafhof 02</i> , Variante Sensor Overlay) .	200
Tabelle A-29: Bodenwerte der Domäne <i>Görzig</i>	200
Tabelle A-30: Korrektur der Ertragsdaten (Quelle: Agri Con®)	203
Tabelle A-31: Neuronale Netze zur ersten N-Applikation (Versuch <i>D4 03</i>).....	204
Tabelle A-32: Neuronale Netze zur zweiten N-Applikation (Versuch <i>D4 03</i>)	205
Tabelle A-33: Neuronale Netze zur dritten N-Applikation (Versuch <i>D4 03</i>)	207

Abbildungsverzeichnis

Abbildung 1: Prozentualer Anteil der Flächen in den USA, auf denen Teilflächenbewirtschaftung angewendet wird (verändert nach GRIFFIN ET AL., 2004, S. 5)	2
Abbildung 2: Aufbau der Arbeit.....	6
Abbildung 3: Die Datenmatrix des überwachten Lernens	11
Abbildung 4: Veranschaulichung des „Überlernens“	13
Abbildung 5: WED-Prozessmodell (verändert nach NAKHAEIZADEH, 1998, S. 4).....	15
Abbildung 6: Ausschnitt eines Merkmalsraumes (verändert nach PYLE, 1999, S. 80).....	19
Abbildung 7: Mehrlagiges, vorwärtsgekoppeltes neuronales Netzwerk (JAFAR- SHAGHAGI, 1996, S. 185)	25
Abbildung 8: Aufbau eines Neurons	26
Abbildung 9: Darstellung eines Fehlergebirges	29
Abbildung 10: Vermeidung von Überlernen bei Clementine®.....	31
Abbildung 11: Funktionsweise der Trainingsmethode <i>Exhaustive Prune</i> (Software <i>Clementine</i> ®)	33
Abbildung 12: Komponenten eines Entscheidungsbaumes	36
Abbildung 13: Unterteilung des Merkmalsraumes mit einem Entscheidungsbaum (verändert nach PYLE, 1999, S. 105).....	37
Abbildung 14: Vorgehensweise einer fünffachen Kreuzvalidierung	41
Abbildung 15: Messung der scheinbaren elektrischen Leitfähigkeit	57
Abbildung 16: Zugkraftmessung bei der Stoppelbearbeitung.....	58
Abbildung 17: Modifizierter WED-Prozess zur Entwicklung von teilflächenspezifischen Entscheidungsregeln.....	69
Abbildung 18: Teilflächen in der IKB-Versuchsanlage.....	72
Abbildung 19: Versuchsanlage <i>Schafhof 02</i>	73
Abbildung 20: Versuchsanlage <i>D4 03</i>	75
Abbildung 21: Versuchsanlage <i>Schafhof 04</i>	77
Abbildung 22: Versuchsanlage <i>Görzig 04</i>	79
Abbildung 23: Versuchsanlage <i>Görzig 05</i>	82
Abbildung 24: Datenvorverarbeitung IKB-Versuche bezüglich Ertragkarten.....	89
Abbildung 25: Technische Umsetzung der N-Applikation (Versuch <i>Görzig 04</i> und <i>Görzig 05</i>)	92
Abbildung 26: Instanzen des Versuches <i>Görzig 04</i>	93
Abbildung 27: Verwendete Attribute zur Ertragsprognose in den Teilapplikationen.....	98
Abbildung 28: Ökonomische Optimierung von N-Gaben mit Ertragsprognosemodellen .	103
Abbildung 29: Konstruktion eines Testvektors zur Erstellung von <i>Response Surfaces</i>	109

Abbildung 30: <i>Response Surfaces</i> zur Analyse neuronaler Netze	110
Abbildung 31: <i>Data Audit</i> des Datensatzes <i>Schafhof 02</i>	112
Abbildung 32: Korrelationsmatrix der Attribute <i>Schafhof 02</i>	113
Abbildung 33: <i>Scatterplot</i> aus Ertragsdaten und Applikationsmengen (<i>N_2</i>)	114
Abbildung 34: <i>Response Surface</i> zur zweiten N-Gabe (<i>Ertrag_98</i> , Versuch <i>Schafhof 02</i>)	116
Abbildung 35: Scatterplot der Instanzen <i>Schafhof 02</i> (<i>REIP_3</i> , <i>Ertrag_02</i> und <i>N_3</i>)	117
Abbildung 36: <i>Response Surface</i> zur dritten N-Applikation (<i>REIP_3</i> , Versuch <i>Schafhof 02</i>)	118
Abbildung 37: <i>Data Audit</i> des Datensatzes <i>D4 03</i>	120
Abbildung 38: Korrelationsmatrix Versuch <i>D4 03</i>	122
Abbildung 39: <i>Response Surface</i> zur ersten N-Applikation (<i>Ertrag_98</i> , Versuch <i>D4 03</i>)	125
Abbildung 40: <i>NKL-Response</i> zur ersten N-Applikation (<i>Ertrag_98</i> , Versuch <i>D4 03</i>) ...	126
Abbildung 41: Entscheidungsbaum zur ersten N-Gabe (Versuch <i>D4 03</i>)	127
Abbildung 42: <i>Response Surface</i> zur zweiten N-Applikation (<i>REIP_2</i> , Versuch <i>D4 03</i>) .	129
Abbildung 43: <i>Response Surface</i> zur zweiten N-Applikation (<i>Zugkraft</i> , Versuch <i>D4 03</i>)	130
Abbildung 44: Entscheidungsbaum zur zweiten N-Applikation (Versuch <i>D4 03</i>)	131
Abbildung 45: <i>Response Surface</i> zur dritten N-Applikation (<i>REIP_3</i> , Versuch <i>D4 03</i>) ...	133
Abbildung 46: <i>Response Surface</i> zur dritten N-Applikation (<i>Ertrag_98</i> , Versuch <i>D4 03</i>)	134
Abbildung 47: Entscheidungsbaum zur dritten N-Applikation (Versuch <i>D4 03</i>)	134
Abbildung 48: <i>Data Audit</i> des Datensatzes <i>Görzig 04</i>	137
Abbildung 49: <i>Scatterplot</i> des Datensatzes <i>Görzig 04</i> (<i>EM_38</i> , <i>Ertrag_03</i> , <i>Ertrag_04</i>) ..	140
Abbildung 50: Verteilung des Attributes <i>N_2</i> im Versuch <i>Görzig 04</i>	140
Abbildung 51: Das Ergebnis der Option „Reduce“ in Bezug auf die Datenbalance	142
Abbildung 52: <i>Response Surface</i> zur ersten N-Applikation (<i>Ertrag_03</i> , Versuch <i>Görzig 04</i>)	145
Abbildung 53: <i>Response Surface</i> zur ersten N-Applikation (<i>EM_38</i> , Versuch <i>Görzig 04</i>)	146
Abbildung 54: Entscheidungsbaum zur ersten N-Applikation	147
Abbildung 55: <i>Response Surface</i> zur zweiten N-Applikation (<i>REIP_2</i> , Versuch <i>Görzig 04</i>)	149
Abbildung 56: <i>Response Surface</i> zur zweiten N-Applikation (<i>Ertrag_03</i> , Versuch <i>Görzig 04</i>)	150
Abbildung 57: <i>Response Surface</i> zur zweiten N-Applikation (<i>EM_38</i> , Versuch <i>Görzig 04</i>)	151
Abbildung 58: Entscheidungsbaum zur zweiten N-Applikation (Versuch <i>Görzig 04</i>)	152

Abbildung 59: Response Surface zur dritten N-Applikation (Versuch <i>Görzig 04</i>)	154
Abbildung 60: <i>Response Surface</i> zur dritten N-Gabe (<i>Ertrag_03</i> , Versuch <i>Görzig 04</i>)....	155
Abbildung 61: <i>Response Surface</i> zur dritten N-Gabe (<i>EM_38</i> , Versuch <i>Görzig 04</i>).....	156
Abbildung 62: Entscheidungsbaum zur dritten N-Applikation (Versuch <i>Görzig 04</i>).....	157
Abbildung 63: <i>Data Audit</i> der Instanzen des Versuches <i>Görzig 05</i>	160
Abbildung 64: Abstimmung der Arbeitsbreiten und angepasste Rasterung	179
Abbildung 65: Zweidimensionale Sinus-Wellen (verändert nach PRINGLE ET AL. 2004a, S. 621)	181
Abbildung A-66: Konzeption des Forschungsprojektes IKB-Dürnast.....	199
Abbildung A-67: Monatliche Temperaturmittelwerte in den Jahren 2002, 2003 und 2004 der Wetterstation Weihenstephan (Datengrundlage: Deutscher Wetterdienst).....	201
Abbildung A-68: Niederschlagsmenge der Versuchsjahre 2002, 2003 und 2004 (IKB-Versuche, Datengrundlage: Deutscher Wetterdienst, Wetterstation Weihenstephan)	201
Abbildung A-69: Niederschlagsmenge der Versuchsjahre 2004 und 2005 (Versuch <i>Görzig</i>).....	202
Abbildung A-70: Temperaturverlauf der Versuchsjahre 2004 und 2005 (Versuch <i>Görzig 04</i>).....	202
Abbildung A-71: NKL-Response zur ersten N-Applikation (<i>Zugkraft</i> , Versuch <i>D4 03</i>)	205
Abbildung A-72: NKL-Response zur zweiten N-Applikation (<i>REIP_2</i> , Versuch <i>D4 03</i>)	206
Abbildung A-73: NKL-Response zur zweiten N-Applikation (<i>Zugkraft</i> , Versuch <i>D4 03</i>)	206

Abkürzungsverzeichnis

€	Euro
bzw.	beziehungsweise
ca.	circa
DGPS	Differentielles Globales Positions System
d. h.	das heißt
dt.	deutsch
EDV	Elektronische Datenverarbeitung
EC	Dezimaler Code für das Entwicklungsstadium von Getreide (TOTTMANN, 1987)
<i>EM₃₈</i>	Scheinbare Elektrische Bodenleitfähigkeit (Messwerte eines EM38-Sensor von Geonics®)
engl.	englisch
evtl.	eventuell
f	folgende Seite
ff	folgende Seiten
GPS	Globales Positions System
ha	Hektar
IKB	Informationssystem Kleinräumige Bestandesführung
inkl.	Inklusive
KDD	Knowledge Discovery in Databases
KNN	künstliche neuronale Netze
kg	Kilogramm
l	Liter
m ²	Quadratmeter
max.	maximal
Mio.	Millionen
N	Stickstoff
NKL	stickstoffkostenfreie Leistung
<i>N_{opt}</i>	optimale spezielle Intensität der Stickstoffdüngung
org.	organisch
PF	Precision Farming
R ²	Bestimmtheitsmaß
<i>REIP</i>	Red Edge Inflection Point
S.	Seite
<i>s. p.</i>	sine pagina (Quellen ohne Seitenangabe)
TS:	Trockensubstanz
SEP	Standard Error of Prediction (vgl. Tabelle 1)
t	Metrische Tonne
SSCM	Site-Specific Crop Management
tsd.	tausend
u. U.	unter Umständen
vgl.	vergleiche
VB	Vegetationsbeginn (Entwicklungsstadium nach TOTTMANN, 1987)
vs.	versus
WED	Wissensentdeckung in Datenbanken
z. T.	zum Teil

1 Einleitung

Während der Begriff *Precision Farming* (PF) noch bis Anfang der 90er Jahre für eine Zukunftsvision stand, stellt sich die aktuelle Situation anders dar. Die mit diesem Begriff verbundenen Technologien und Produkte haben mittlerweile einen Standard erreicht, der weit über das Versuchsstadium hinausgeht. Auf dem Markt für Landtechnik haben sich diese Produkte bereits einen festen Platz erobert. Nach AUERNHAMMER (2004, S. 33) umfasst der Überbegriff *Precision Farming* die Teilbereiche Dokumentation, Flottenmanagement, Feldrobotik und die Teilflächenbewirtschaftung¹. Die aktuellen Trends zeigen, dass die einzelnen Teiltechnologien sehr differenziert vom Markt angenommen werden (WHIPKER UND AKRIDGE, 2005, S. 16).

Die Forschungsanstrengungen im Bereich *Precision Farming* konzentrierten sich seit Mitte des letzten Jahrzehntes insbesondere auf die standortangepasste Ausbringung von Betriebsmitteln (Teilflächenbewirtschaftung). Die Idee besteht darin, kleinräumig Daten zu erfassen und damit den Betriebsmitteleinsatz in diesem Maßstab zu optimieren. Ausgehend von der bisherigen Praxis des schlageinheitlichen Faktoreinsatzes, eröffnet die Teilflächenbewirtschaftung durch die Berücksichtigung von Heterogenitäten innerhalb eines Schlages neue Optimierungspotenziale. Dieser leicht nachvollziehbare Gedankengang führte zu einer relativ schnellen Akzeptanz in der Praxis (vgl. Abbildung 1). Bereits im Jahre 1998 wurden in den USA 12% der Mais-Fläche nach dieser neuen Methode bewirtschaftet. Die beiden Kulturen Soja und Weizen haben in den USA produktionsbedingt schlechtere Voraussetzungen für den Einsatz dieser Technologie. Dennoch werden Anteile von rund 3 bzw. 7% erreicht. Die Technologien zur Erfassung und Aufbereitung von kleinräumigen Daten (insbesondere Ertragskartierung) waren bereits zu dieser Zeit in einem fortgeschrittenen Stadium und auch am Markt entsprechend verfügbar.

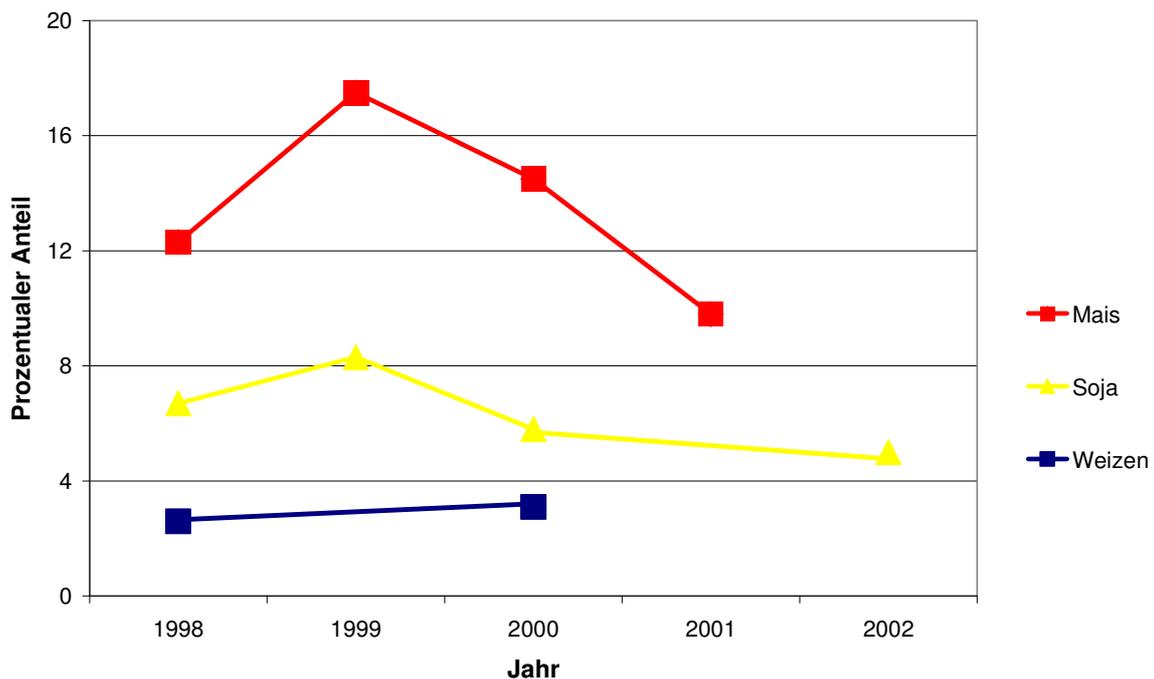
Überraschend ist allerdings die Trendwende in der Akzeptanz dieser Technologie ab dem Jahre 1999. Zu diesem Zeitpunkt wurden bereits über 17% der Mais-Fläche, rund 8% der Soja-Fläche und rund 3% der Weizen-Fläche teilflächenspezifisch bewirtschaftet. Von diesem Zeitpunkt an sinken die Anteile in allen Kulturen beständig². Die letzten Jahre aus den Daten dieser Studie zeigen, dass die Anteile in allen Kulturen unter den Ausgangswert im Jahre 1998 sinken.

Dies ist umso überraschender, als sich in dieser Zeit die technischen Voraussetzungen enorm verbessert haben. Beobachten lässt sich dies an der Marktdurchdringung von Systemen zur Ertragsmessung – einer der möglichen Datengrundlagen zur Teilflächenbewirtschaftung. Obwohl sich die Zahl der Mähdrescher, die mit einer Ertragserfassung ausgestat-

¹ Engl. Site-Specific Crop Management (SSCM) bzw. Variable Rate Technology (VRT).

² Für die Furchtart Weizen existieren keine Angaben für das Jahr 1999. Es ist zu vermuten, dass dieser Wert ebenfalls höher liegt, als der Ausgangswert im Jahre 1998.

tet sind, in diesem Zeitraum stetig erhöht hat, wurden diese Daten nicht in dem entsprechenden Maße dazu eingesetzt, um Ertragskarten zu erstellen (GRIFFIN ET AL., 2004, S. 6). Eine entsprechende Weiterentwicklung kann auch für die Systeme zur Verarbeitung von kleinräumigen Daten angenommen werden. Insbesondere durch die Weiterentwicklung von Softwarepaketen bzw. durch die Verfügbarkeit von Precision-Farming-Dienstleistungen, die die aufwändige Datenaufbereitung erleichtern, wird der Einstieg in diese Technologien von Jahr zu Jahr erleichtert. Nicht zuletzt sollte sich in diesem Zeitraum auch das Wissen bzw. die Erfahrungen des Anwenders im Umgang mit dieser Technologie verbessert haben.



Quelle: USDA ARMS, folgende Daten nicht verfügbar: Mais 2002, Soja 2001, Weizen 1999, 2001, 2002

Abbildung 1: Prozentualer Anteil der Flächen in den USA, auf denen Teilflächenbewirtschaftung angewendet wird (verändert nach GRIFFIN ET AL., 2004, S. 5)

Damit stellt sich die Frage nach den Gründen für diese auffällige Trendwende, die in der Anwendung dieser Technologie liegen müssen. Der Schluss liegt nahe, dass sich in der Praxis nicht die erhofften ökonomischen Vorteile eingestellt haben. Auch wissenschaftliche Untersuchungen bestätigen diesen Verdacht. In Studien, die den ökonomischen Erfolg einer Teilflächenbewirtschaftung untersucht haben, zeigt sich ein sehr heterogenes Bild mit negativen und positiven Effekten auf die Wirtschaftlichkeit (vgl. LAMBERT UND LOWENBERG-DEBOER, 2000, S. 14; DABBERT UND KILIAN, 2002, S. 426). Aus ökonomischer Sicht konnten bisher keine „durchschlagenden“ Ergebnisse aufgezeigt werden. In der Wissenschaft herrscht weitgehende Einigkeit darüber, dass dies ein wesentlicher Grund für die sinkende Akzeptanz dieser Technik ist (DABERKOW UND MCBRIDE, 2000, s.p.; LOWENBERG-DEBOER, 1996, S. 1283).

Allerdings vermitteln die dazu durchgeführten Versuche ein uneinheitliches Bild und ein näherer Blick auf die betrachteten Untersuchungen zeigt, dass es aus mehreren Gesichtspunkten heraus noch zu früh für eine abschließende ökonomische Beurteilung ist (vgl. Kap. 3.1.2). Insbesondere bestehen immer noch große Lücken in der Ableitung von teilflächenspezifischen Düngestrategien (vgl. WENKEL ET AL., 2002, s.p.). In den Versuchen werden daher sehr unterschiedliche Algorithmen verwendet. Die anfangs verwendete Übertragung von schlageinheitlichen Düngerstrategien auf ein teilflächenspezifisches Regime erweist sich vielfach als nicht zielführend (vgl. Kap. 3.1.2). In vielen Studien werden zur Optimierung der teilflächenspezifischen Varianten Bodenbeprobungen verwendet, die sehr kostenintensiv durchzuführen bzw. räumlich zu niedrig aufgelöst sind. Dies beeinträchtigt sowohl die Kosteneffizienz als auch das Potenzial zur Optimierung. Weiterhin basiert die Optimierung des variablen Input-Faktors in vielen Studien ausschließlich auf einer einzelnen kleinräumigen Datenlage. In vielen Fällen wird beispielsweise auf Basis von historischen Ertragskarten optimiert, in anderen Studien nur nach reflexionsoptischen Messungen. Doch genau in der Nutzung aller verfügbaren kleinräumigen Informationen, die sich typischerweise gegenseitig beeinflussen, sollte ökonomisches Potenzial zu finden sein.

Zusammenfassend kann festgehalten werden, dass die Akzeptanzhemmnisse bezüglich des Einsatzes einer teilflächenspezifischen Bewirtschaftung in der Praxis aus den unsicheren ökonomischen Vorteilen resultieren (vgl. GANDORFER, 2005, S. 138). Die Unsicherheit wiederum ergibt sich dadurch, dass bis dato keine allgemein akzeptierte Regeln zur Umsetzung einer teilflächenspezifischen Bewirtschaftung entwickelt wurden, die zu positiven ökonomischen Effekten führen (vgl. PETERS ET AL., 1999, S. 719). Die Aussage von LOWENBERG-DEBOER (1996, S. 1283) beweist nach wie vor Aktualität:

"One of the key factors limiting adoption of precision farming technology is the lack of decision support. There is too much data to sort and analyze manually or mentally, and little software to automate the process. Someone needs to estimate the surface-generating production function ... someone needs to develop the optimization algorithm that will apply that information to generating next season's cropping strategy."

Was also die Akzeptanz dieser Technologie in der Praxis bremst, sind Wissensdefizite in der Nutzung von teilflächenspezifischen Daten zur kleinräumigen Optimierung. Es ist wenig Wissen über die vielfältigen Interaktionen der teilflächenspezifischen Attribute untereinander und insbesondere im Zusammenwirken mit Stickstoff auf den kleinräumigen Ertrag vorhanden (vgl. Kap. 3). An diesem Punkt knüpft die vorliegende Arbeit an.

1.1 Problemstellung und Zielsetzung

Wie aus der Einleitung deutlich wurde, besteht ein konkreter Forschungsbedarf zur Entwicklung von Algorithmen, mit Hilfe derer kleinräumige Daten zur Optimierung des teilflächenspezifischen Stickstoffeinsatzes verwendet werden können. Es stellt sich die Frage, wie diese Wissensdefizite beseitigt werden können. Die Lösungsmöglichkeiten für dieses Problem fassen DRUMMOND ET AL. (2000, s.p.) folgendermaßen zusammen:

„Several means exist to investigate these topics [site specific crop response, Anm. d. Verfassers]. Agronomic methods, involving countless small plot trials over multiple site-years would seem to be the most traditional, and possibly the best means to compile the necessary data. However, this is extremely time and labor intensive and seems unrealistic for the near future. Another approach which may hold promise for understanding these response relationships involves the use of crop growth models [...]. A third method consists of empirical analysis of spatial, multivariate data sets – just the type of data sets produced by the implementation of precision agriculture technologies.”

Der traditionelle pflanzenbauliche Exaktversuch kann für eine systematische Entwicklung von komplexen Algorithmen kurz- bis mittelfristig nicht weiterhelfen: Da diese Versuchsanstellung auf homogene Bodenverhältnisse angewiesen ist, bedarf es schon alleine deswegen einer enormen Anzahl von Parzellenversuchen, um das Spektrum möglicher Bodenverhältnisse abzudecken. Dies wiederum stellt eine extrem zeit- und kostenintensive Aufgabe dar. Diese Versuchsform hat daher kurz- bis mittelfristig wenig Potenzial, die Informationsdefizite zu beseitigen. Ähnliches gilt für Simulations- bzw. Wachstumsmodelle, die auf dieser Versuchsform basieren. Die dritte Alternative besteht in der empirischen Analyse multivariater räumlicher Daten, also genau der Daten, die bei der Implementierung von Precision-Farming-Technologien in der Praxis entstehen. Insbesondere dieser empirische Ansatz erscheint in diesem Zusammenhang aussichtsreich, da die Anwendung von Precision Farming dazu führt, dass empirische Daten in bisher unbekannter Quantität erfasst werden.

Dieses Phänomen taucht nicht nur im Bereich Precision Farming, sondern in sehr verschiedenen Disziplinen und Branchen auf. Aus diesem Antrieb heraus hat sich eine neue Disziplin herausgebildet, deren Kerngebiet die Analyse von großen multidimensionalen Datenbanken ist. Die Methode der Wissensentdeckung in Datenbanken (WED) hat das Ziel, Muster, d. h. spezifisches Wissen, über die Zusammenhänge im Datenmaterial zu finden (vgl. NAKHAEIZADEH, 1998, S. 2ff). Dazu werden in dieser Disziplin Techniken des maschinellen Lernens angewendet.

Diese Methode bietet sich demnach insbesondere auch dazu an, die aufgezeigten Informationsdefizite für eine kleinräumige Optimierung der Betriebsmittel-Ausbringung zu schließen. Dieser Weg soll im Rahmen dieser Arbeit am Beispiel der teilflächenspezifischen Stickstoff-Applikation in Winterweizen eingeschlagen werden.

Das Ziel dieser Arbeit besteht zunächst darin, die Methode WED näher vorzustellen und

ihre Einsatzmöglichkeiten für Precision Farming aufzuzeigen. Dazu sollen die derzeitigen Ansätze zur kleinräumigen Optimierung näher untersucht werden, um daraus konkrete Wissensdefizite abzuleiten. Weiterhin sollen dazu Arbeiten untersucht werden, die Elemente dieser Methode im Bereich Precision Farming verwenden.

Basierend auf diesen Ergebnissen soll die Methode WED speziell darauf ausgerichtet werden, um ökonomisch optimierte Entscheidungsregeln für eine kleinräumige Stickstoff-Ausbringung zu entwickeln. Dazu soll ein Konzept aufgebaut werden, das den Einsatz der Techniken des maschinellen Lernens für diese spezielle Anwendung ermöglicht. Die konkrete Entwicklung von Entscheidungsregeln soll anhand mehrerer empirischer Versuche aufgezeigt werden. Das Ziel des empirischen Teils besteht zum einen darin, die praktische Umsetzung des Verfahrens vorzustellen und daraus Erkenntnisse für die Weiterentwicklung zu gewinnen. Da es sich um eine neuartige Herangehensweise handelt, gilt es also zunächst, Erfahrungen mit der Methode zu sammeln und daraus grundlegende Erkenntnisse für die Versuchsgestaltung und die Datenanalyse zu gewinnen. Zum anderen besteht das Ziel darin, soweit als möglich spezifisches Wissen zur kleinräumigen Optimierung der Stickstoff-Düngung zu generieren.

Diese Arbeit entstand im Rahmen des Precision-Farming-Projektes „IKB-Dürnast“ (vgl. Abbildung A-66, im Anhang). Die Idee besteht darin, Entscheidungsregeln, die mit dieser Methode entwickelt werden, in einen In-field-Controller zu implementieren (vgl. OSTERMEIER UND AUERNHAMMER, 2004, s.p.) und damit zu einer automatisierten teilflächenspezifischen Stickstoff-Ausbringung zu gelangen.

1.2 Aufbau dieser Arbeit

Die Arbeit gliedert sich in zwei Hauptteile: Im ersten Teil (Kap. 1-3) wird das Theoriefundament anhand des aktuellen Forschungsstands aufgebaut. Im zweiten Teil (ab Kap. 4) wird daraus ein Ansatz zur Entwicklung von Entscheidungsregeln im Bereich Precision Farming abgeleitet und an verschiedenen empirischen (Versuchs-)Daten angewendet. Eine Übersicht der Struktur dieser Arbeit ist in Abbildung 2 dargestellt.

Die vorliegende Arbeit basiert auf Elementen aus drei wissenschaftlichen Bereichen. Im ersten Bereich handelt es sich um die gewählte Methode „*Wissensentdeckung in Datenbanken*“. Die beiden anderen Bereiche stellen das Anwendungsgebiet „*Precision Farming*“ und „*Stickstoff-Düngung*“ dar.

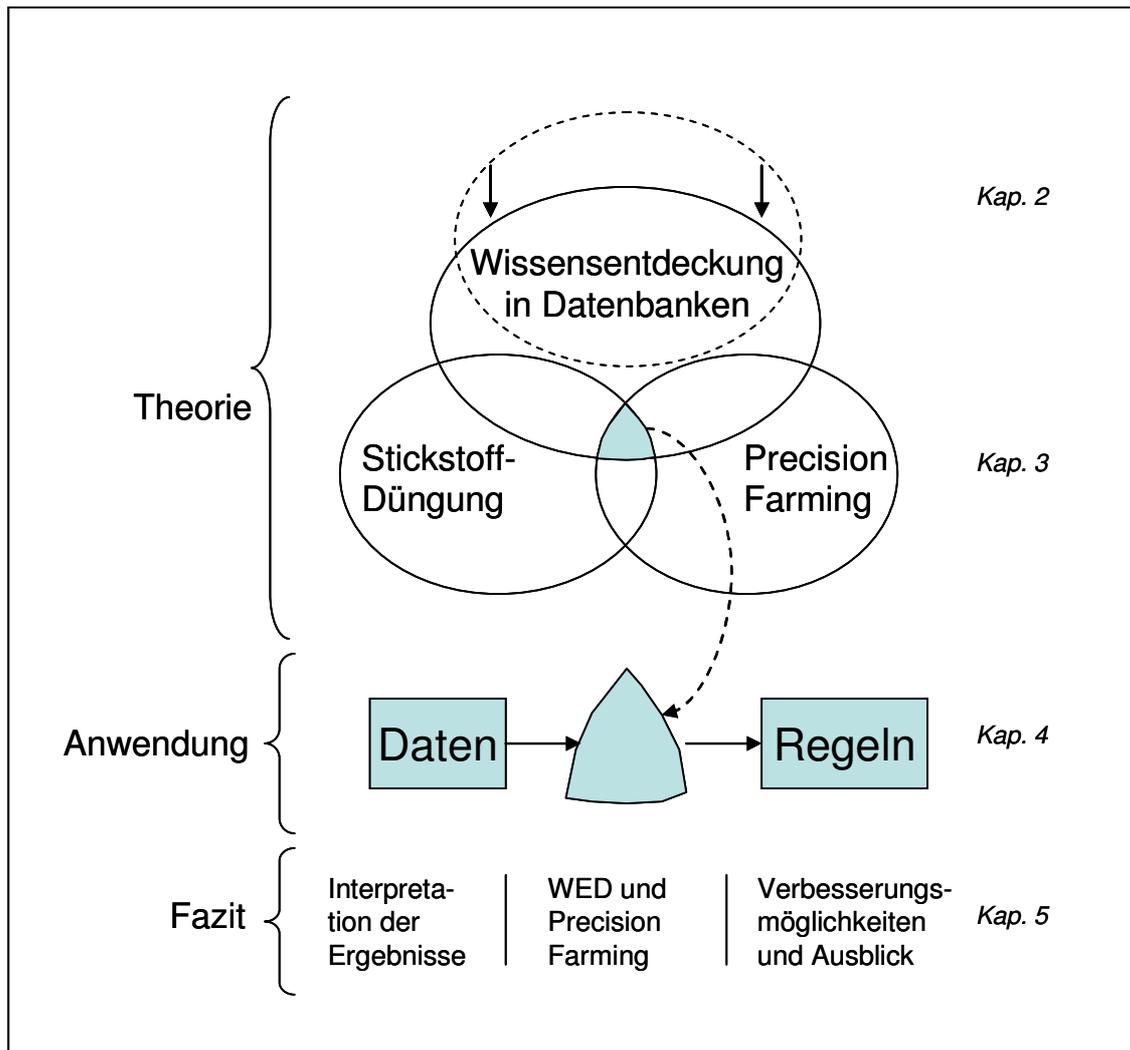


Abbildung 2: Aufbau der Arbeit

Das zweite Kapitel behandelt die für diese Arbeit relevanten Ausschnitte der Forschungsdisziplin Wissensentdeckung in Datenbanken. Hier sollen insbesondere die Data-Mining-Algorithmen des überwachten Lernens dargestellt werden, die im zweiten (Haupt-) Teil dieser Arbeit verwendet werden³. Weiterhin wird auf die Datenaufbereitung und den Methoden zu Validierung, die für diese Arbeit ebenfalls von hoher Bedeutung sind, näher eingegangen.

Im dritten Abschnitt wird der Forschungsstand der Bereiche Stickstoffdüngung und Precision Farming hinzugezogen. Die Ausführungen beschränken sich dabei weitgehend auf die Schnittmengen zwischen diesen drei Bereichen⁴. Dazu gehören insbesondere die Praxis der Stickstoffdüngung und die daraus abgeleiteten Methoden zur teilflächenspezifischen Stickstoff-Ausbringung. Aus dieser Analyse wird der konkrete Forschungsbedarf abgeleitet, der

³ Die Ergebnisse der Modellierung hängen z. T. stark von den eingesetzten Algorithmen (Software) ab. Daher ist eine genaue Darstellung unabdingbar.

⁴ Die Schnittmenge Stickstoff-Düngung und Data Mining ist dabei ausgenommen. Hier existieren nur vereinzelt Überschneidungen.

im empirischen Teil aufgegriffen wird. Weiterhin wird in diesem Kapitel kurz auf Sensoren eingegangen, die im Rahmen einer teilflächenspezifischen Stickstoff-Düngung eingesetzt werden und im weiteren Verlauf dieser Arbeit verwendet werden. Zum Stand der Forschung gehört weiterhin eine Betrachtung der Arbeiten, die im Rahmen von Precision Farming bzw. Stickstoff-Düngung bereits Elemente aus dem Bereich Data Mining verwenden.

Der zentrale Schnittpunkt definiert dabei den Forschungsstand, an den diese Arbeit anknüpft. Ausgehend von dieser Basis wird im zweiten Hauptteil (ab Kap. 4) ein Modell zur Entwicklung von Entscheidungsregeln entworfen. Anhand von verschiedenen Versuchsdaten wird gezeigt, wie Entscheidungsregeln aus empirischen Daten erzeugt und getestet werden können. Dieser Teil beginnt mit einem Überblick über einen modifizierten Prozess der Wissensentdeckung in Datenbanken für Precision Farming (vgl. Kap. 4.1) und einer detaillierten Beschreibung der verwendeten Versuchsdaten. Es folgt für die jeweiligen Versuche eine Darstellung der Modellierungsphase und eine Präsentation der Ergebnisse. Es wird versucht, aus diesen Ergebnissen stabile Muster zu extrahieren, die über die einzelnen Versuchsjahre hinweg auftauchen.

Das Schlusskapitel ist dreigeteilt. Im ersten Teil werden die Ergebnisse des empirischen Teiles zusammengefasst und die Schlussfolgerungen daraus gezogen. Im zweiten Teil wird nochmals auf das Thema Wissensentdeckung in Datenbanken eingegangen. Aus der Erfahrung dieser Arbeit erfolgt eine kritische Reflektion zur Eignung dieser Methode im Rahmen der Forschung zu Precision Farming. Im letzten dieser drei Teile werden, ausgehend von den beiden ersten Teilen des Schlusskapitels, Vorschläge für ein verbessertes Versuchsdesign und für weiteren Forschungs- und Entwicklungsbedarf in diesem Bereich gegeben.

2 Wissensentdeckung in Datenbanken

Moderne Datenerfassungsgeräte und leistungsfähige Datenbanken machen es möglich, große Datenmengen kostengünstig und automatisiert zu erfassen und zu speichern. In wissenschaftlichen Testserien werden beispielsweise Daten automatisch erzeugt und gespeichert. In Unternehmen dient die massive Sammlung und Speicherung von Daten aller Art zunächst oft für operative Zwecke, wie z. B. für die Kundenverwaltung oder das betriebliche Rechnungswesen. Gleichzeitig bieten sich diese Datensammlungen aber für Analysen und Auswertungen an. Bereiche, die exemplarisch für große Datensammlungen stehen, sind die Biomedizin, die Kapitalmarktforschung, die Telekommunikationsbranche und seit der Installation der Scanner-Kassen der Einzelhandel (vgl. HAN UND KAMBER, 2001, S. 3ff.; WESTPHAL UND BLAXTON, 1998, S. 7ff.). In verschiedensten Bereichen der Forschung und in Unternehmen ist es interessant, diese Daten auszuwerten, um beispielsweise Prozesse zu verbessern (vgl. BIGUS, 1996, S. 16ff).

Die Abfrage und Analyse dieser Datenmassen gestaltet sich jedoch schwierig und komplex. Aus dieser Anforderung heraus entwickelte sich ein Forschungsbereich, der von Autoren aus sehr verschiedenen Fachrichtungen getragen wird und sich nur mit der Analyse von multidimensionalen und umfangreichen Datenmengen beschäftigt. Die Forschungsanstrengungen konzentrieren sich insbesondere auf die Definition von Prozessen und auf die Entwicklung von Instrumenten, die eine effiziente Auswertung ermöglichen.

In diesem Kapitel werden die Grundlagen dieses Forschungsbereiches vorgestellt. Im ersten Teil werden dazu die Hauptelemente eines Prozessmodells eingeführt und beschrieben. Als Data-Mining-Werkzeuge werden zwei Techniken des überwachten Lernens, neuronale Netze und Entscheidungsbaumalgorithmen vorgestellt, die in dieser Arbeit angewendet werden.

2.1 Data Mining und Wissensentdeckung in Datenbanken

Im Wesentlichen existieren zwei konkurrierende Begriffe, die dieses Forschungsgebiet kennzeichnen: *Data Mining* (DM) und *Knowledge Discovery in Databases* (KDD). Eine deutsche Übersetzung des Begriffes Data Mining hat sich nicht etabliert, während Knowledge Discovery in Databases oft mit *Wissensentdeckung in Datenbanken* (WED) übersetzt wird. Oft werden diese beiden Begriffe synonym verwendet. In jüngerer Zeit hat sich jedoch die Auffassung durchgesetzt, Data Mining als einen – wenn auch sehr wichtigen – Teil des Prozesses der Wissensentdeckung anzusehen. Der Begriff Data Mining wird auch in dieser Arbeit dann verwendet, wenn der Fokus auf der Modellierungsphase liegt.

Nach FAYYAD ET AL. (1996, S. 6) ist Knowledge Discovery in Databases der nicht triviale Prozess der (semi-)automatischen Extraktion von Wissen aus Datenbanken, das gültig, bis-

her unbekannt, potenziell nützlich und schlussendlich verständlich ist.

In dieser oft zitierten Definition sind die wichtigsten Anforderungen dargestellt. Im Einzelnen bezieht sich diese Definition auf Wissen, das gültig im statistischen Sinne ist. Mit „bisher unbekannt“ soll gewährleistet werden, dass es sich bei dem gefundenen Wissen um kein Allgemeinwissen handeln sollte und weiterhin sollte das neue Wissen in einer Form verwertbar, also gewinnbringend, einzusetzen sein. Gleichzeitig wird der Prozesscharakter dieses Ansatzes herausgestellt.

Ein weiteres Kennzeichen eines WED-Prozesses ist das so genannte *Hybride Agieren* bzw. die *Interaktion*. Damit wird herausgestellt, dass der Prozess erst unter Zusammenwirken von „menschlicher Ratio, Kreativität, etc. und leistungsfähigen Maschinen“ (SCHOMMER 2003, S. 9) zu Erkenntnissen führt. Die Data-Mining-Werkzeuge werden lediglich als Hilfsmittel angesehen, die der Benutzer anwendet, um Informationen zu gewinnen und schlussendlich Wissen zu generieren.

Eine Übersicht einiger Definitionen für dieses komplexe Forschungsgebiet, die sich sehr ähnlich sind, aber den Fokus jeweils auf unterschiedliche Bereichen setzen, wird beispielsweise in SÄUBERLICH (2000, S. 7ff) gegeben.

Weiterhin gibt es aber auch Autoren, die sich von den Problemen mit der Terminologie distanzieren und den Fokus auf die Kombination der Techniken legen, die in diesem Bereich verwendet werden (vgl. KAMATH, 2001, S. 2f). Damit kommt man zu einem Ansatz, der hochgradig multidisziplinär ist. In diesem Bereich treffen Elemente von Statistik und Informatik (insbes. Maschinelles Lernen und Datenbanksysteme) aufeinander und charakterisieren so ein neues Forschungsgebiet. Eine tragende Rolle kommt beispielsweise den Elementen aus der „klassischen“ Statistik zu. Statistische Methoden werden oftmals direkt oder in abgewandelter Form als Data-Mining-Werkzeug eingesetzt. Die Überschneidungen zwischen diesen beiden Forschungsrichtungen sind groß. Es gibt allerdings auch konzeptuelle Unterschiede. Generell eignen sich Werkzeuge des Data Mining insbesondere zur Analyse von großen Datenmengen und weiterhin von Daten mit vielen Attributen, d. h. multidimensionalen Daten (WEGMAN UND SOLKA, 2005, S. 2ff).

In der klassischen Statistik steht das Testen von Hypothesen im Vordergrund, während Data Mining von einem absoluten Nicht-Wissen über Zusammenhänge in den Daten ausgeht. Eine typische Frage in diesem Kontext lautet: „*What are all the hypotheses that this data supports?*“ (PYLE, 1999, S. 31). Eine scharfe Trennung zwischen hypothesenfreien und hypothesengetriebenen Problemen wird dabei nur als idealtypische Systematik zur Verdeutlichung der unterschiedlichen Herangehensweisen gesehen (vgl. GEHRA, 2005, S. 40). Zu einer kritischen Reflektion des Einsatzes von Data Mining in der Wissenschaft sei auf das Schlusskapitel verwiesen. Anhand der Ergebnisse und Erfahrungen des empirischen Teils werden darin die damit verbundenen Chancen und Probleme nochmals aufgegriffen.

Innerhalb dieser Disziplin wird grundlegend zwischen *unüberwachten* (deskriptiven) Verfahren und *überwachten* (prädiktiven) Verfahren unterschieden. Diese Unterscheidung wird anhand der Autonomie des Verfahrens vorgenommen. Vereinfacht lernen überwachte Verfahren aus Beispielen, die bereits klassifiziert sind bzw. einen empirischen Zielwert haben. Unüberwachte Verfahren suchen selbständig nach Mustern, Regeln oder Gruppierungen innerhalb der Daten und betreiben damit eine Form der Selbstorganisation (SCHOMMER, 2003, S. 33ff). Beispiele für derartige Aufgaben sind das Clustering bzw. Assoziationsregeln (ESTER UND SANDER, 2000, S. 5ff).

In dieser Arbeit werden ausschließlich Techniken des überwachten Lernens eingesetzt und behandelt. Dabei nimmt das Lernverfahren eine Menge bereits „klassifizierter“ Beispiele entgegen, aus denen es lernen soll, unbekannte Beispiele zu klassifizieren. Die numerische Vorhersage ist eine Variante des klassifizierenden Lernens, wobei das Ergebnis statt einer Kategorie ein numerischer Wert ist.

2.2 Die Aufgabe des überwachten Lernens

Zur formalen Darstellung dieser Aufgabe müssen dazu die Grundelemente des WED-Prozesses aufgeschlüsselt werden. Ein zentrales Element sind *Daten*. Im Allgemeinen wird unter *Daten* eine Ansammlung von „Datenpunkten bzw. Datensätzen“ verstanden“ (PYLE, 1999, S. 67). Daten stellen den Input für einen WED-Prozess dar.

Im Gegensatz dazu können *Informationen* in diesem Kontext als (vor-)verarbeitete Daten verstanden werden und somit als Resultat einer Datenvorbereitung. Am Ende dieser Prozesskette steht idealtypisch das *Wissen*, als „Resultat einer fachspezifischen Diskussion und Interpretation“ (SCHOMMER, 2003, S. 8).

Daten, die für Data Mining verwendet werden, liegen in der Regel auf einem Computersystem vor. Dabei ist weitgehend unerheblich, in welcher Form sie auf dem Computer gespeichert wurden. Entscheidend ist, dass sich der zu analysierende Ausschnitt der Daten in Form einer Matrix darstellen lässt.

Eine Datenmatrix für überwachtetes Lernen lässt sich wie in Abbildung 3 darstellen (vgl. HASTIE ET AL. 2001). In dieser Tabelle (auch: *spreadsheet data format*, WEISS UND INDURKHYA 1998, S. 53) ist die formale Beschreibung dieser Problemstellung dargestellt. Die Terminologie lehnt sich dabei an WITTEN UND FRANK (2001, S. 41ff) an. Die einzelnen Datensätze d_i , die im Folgenden als *Instanzen* bezeichnet werden⁵, sind in den Zeilen dargestellt. Eine Instanz kann beispielsweise für einen Kunden stehen. In der vorliegenden Arbeit handelt es sich dabei meist um Teilflächen.

⁵ In der Literatur trifft man auch oft auf die Begriffe „Fälle“, „Beispiele“, oder „records“.

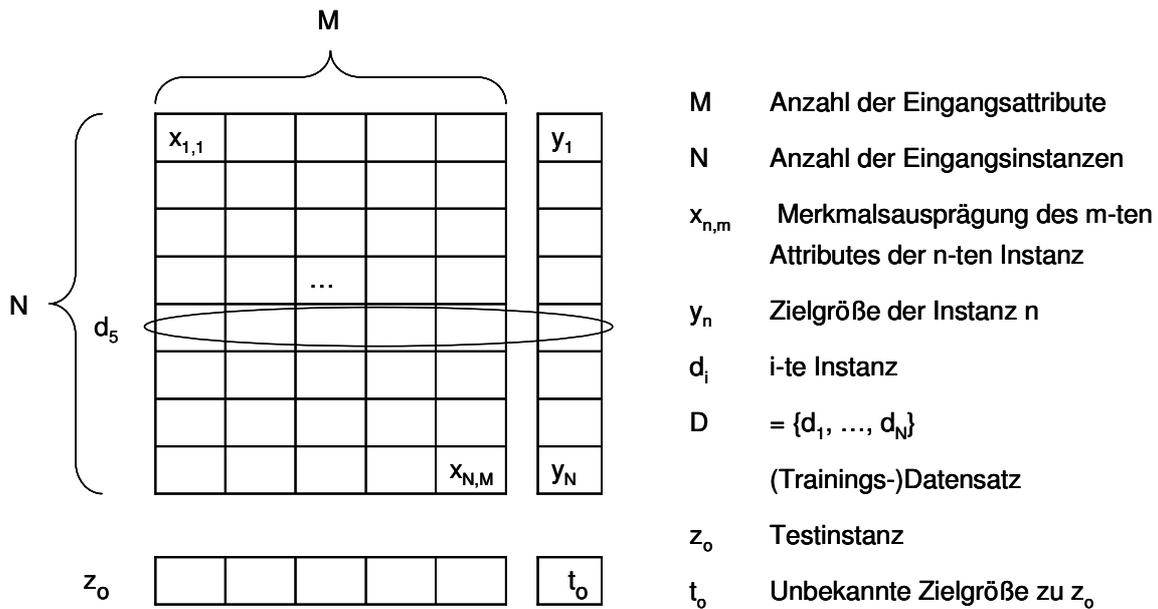


Abbildung 3: Die Datenmatrix des überwachten Lernens

In den Spalten sind die *Attribute*⁶ sowie die *Zielgröße* Y für jede Instanz abgetragen. In dieser Arbeit sind dies beispielsweise Boden-Attribute, die eine Teilfläche beschreiben. Eine Instanz d_n zeichnet sich durch konkrete Merkmalsausprägungen für die Attribute $1, \dots, M$ und für die Zielgröße y_n aus. Über diese Attribute wird im Idealfall eine Instanz vollständig charakterisiert und die jeweilige Zielgröße erklärt. Es wird vereinfacht von einer einzigen Zielvariablen ausgegangen, obwohl einige Data-Mining-Werkzeuge auch mehrere Zielgrößen verwenden können. Die gesamte *Instanzenmenge* stellt den Trainings-Datensatz D dar.

Mittels Lernverfahren des überwachten Lernens wird aus dem Trainingsdatensatz eine Funktion approximiert, die dann für Testinstanzen, für die keine Ausprägung der Zielgröße vorliegt, diese entsprechend klassifiziert bzw. vorhersagt. Ein klassisches Beispiel für ein solches Lernverfahren ist die multiple lineare Regression.

Das formale Ziel besteht darin, aus der Trainingsmenge eine Funktion zu approximieren, die den Zusammenhang zwischen der Matrix der Eingabewerte

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1M} \\ \vdots & \ddots & \vdots \\ x_{N1} & \cdots & x_{NM} \end{pmatrix}$$

und dem Vektor der Zielvariablen

$$y = (y_1, \dots, y_N)$$

möglichst gut beschreibt. Gesucht wird also eine Funktion, die von den Input-Werten auf

⁶ Synonym dazu stehen die Begriffe Variable, Feature, Merkmal, etc.

einen Output-Wert schließen lässt.

Die Attribute und Zielgrößen können dabei in unterschiedlichen Ausprägungsformen vorliegen. Eine der Hauptunterscheidungen sind die Kategorien „nominal“ (bzw. kategorisch oder symbolisch) oder „numerisch“ (vgl. WITTEN UND FRANK, 2001, S. 50f). Nominale Attribute können Werte von vordefinierten Kategorien annehmen, während numerische Attribute ganzzahlige bzw. reelle Werte annehmen können. In dieser Arbeit werden fast ausschließlich kontinuierliche numerische Attribute verwendet. Diese Attribute können also theoretisch jeden Wert in beliebiger Genauigkeit (mit Einschränkungen in der Messung) annehmen. Lediglich für die applizierte N-Menge werden „Stufen“ verwendet. In den praktischen Implementierungen wird dieses Merkmal daher entweder als nominal und sortierbar oder näherungsweise numerisch aufgefasst⁷.

Die Trainingsinstanzen stammen in der Regel von automatisierten Datenerfassungsgeräten. Während die Daten z. B. von Scannerkassen eindeutig vorliegen, ist der Umgang mit physischen Messungen schwieriger. Das Grundproblem einer physischen Messung ist, dass sie einem Bias, z. dt. einer Verzerrung erliegen können. Die Messung erzeugt dadurch einen systematischen Fehler, der beispielsweise aus einer fehlerhaften Kalibrierung resultiert. Im Gegensatz dazu steht das Messrauschen (engl. noise), durch das ein unsystematischer, also zufälliger Fehler entsteht.

Dieses Messrauschen sollte aber nicht als Fehler angesehen werden, es ist integraler Bestandteil der Messung (PYLE, 1999, S. 53) und kann demnach auch nicht korrigiert werden. Dies gilt auch für alle Precision-Farming-Sensordaten in dieser Arbeit (vgl. Kap. 3.1.3). Beim überwachten Lernen kann das Messrauschen durch eine Überanpassung des Modells auch als Information angesehen und gelernt werden.

Überlernen

Eines der großen Probleme besteht demnach darin, das richtige Maß zu finden, mit dem die Funktion angenähert werden soll. Es stellt sich immer wieder heraus, dass scheinbar perfekt angepasste Modelle bei der Anwendung auf neue Daten versagen. Die Ursache hierfür ist das so genannte *Überlernen* bzw. engl. *Overfitting* (ESTER UND SANDER, 2000, S. 131f).

Mit Überlernen bezeichnet man das Problem, dass die Modelle zu stark an die von Störeinflüssen geprägten Daten angepasst werden und nicht die ihnen zugrunde liegende Struktur erfasst wird. Somit verlieren diese Modelle die Eigenschaft der Generalisierungsfähigkeit, also der Fähigkeit, Daten so zu modellieren, dass auch neue Eingabemuster sinnvoll verarbeitet werden können.

Es existieren zwei Ansätze zur Begründung dieses Verhaltens (vgl. WIDMANN, 2001): Die erste Begründung ist das Einsetzen von einer Art Auswendiglernen. Das bedeutet mathema-

⁷ In der verwendeten Software existiert für nominale, aber sortierbare Attribute der Begriff „ordered set“.

tisch-statistisch gesehen, dass die Modelle im Regelfall zu groß bzw. komplex für die den Daten zugrunde liegenden Strukturen sind (vgl. für neuronale Netze Kap. 2.5.3., für Entscheidungsbäume Kap. 2.6.3). Die zweite Begründung dafür ist, dass die Modelle das Messrauschen lernen. Dieser Effekt rührt daher, dass die einzelnen Instanzen in der Regel sehr oft für den wiederholten Lern- bzw. Anpassungsprozess verwendet werden. Insbesondere bei einer geringen Quantität an Lerninstanzen führt dies zu dem Problem des Überlernens.

Veranschaulicht ist dieses Phänomen in Abbildung 4. Hier sind verschiedene xy -Messpunkte eingetragen. Das zu erfassende Muster bzw. die zu erfassende Struktur kann im Wesentlichen mit einer ansteigenden Gerade beschrieben werden. Durch das Messrauschen liegen diese Punkte jedoch nie genau auf dieser Gerade. Durch eine entsprechende Überanpassung wie sie beispielsweise bei komplexen neuronalen Netzen vorkommen kann (vgl. Kap. 2.5), verläuft die angepasste Funktion durch alle Punkte. Damit wurde das Messrauschen auswendig gelernt. Bei neuen Daten führt diese Funktion in der Regel zu schlechteren Ergebnissen.

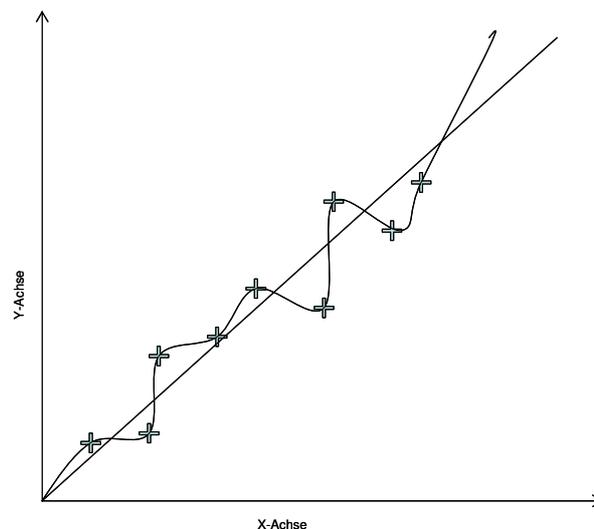


Abbildung 4: Veranschaulichung des „Überlernens“

Andererseits dürfen die Modelle die Struktur in den Daten auch nicht zu sehr vereinfachen. Es gilt, einen Kompromiss aus diesen Anforderungen in der Modellierung zu finden. Dazu werden bei der Modellbildung unterschiedliche Verfahren angewendet (vgl. Kap. 2.5.3 und 2.6.3). Gleichzeitig gilt es, durch die entsprechende Konzeption eines (Gesamt-)Prozesses diese Aufgabe zu unterstützen. Dieser Prozess soll im folgenden Kapitel vorgestellt werden.

2.3 Der Prozess der Wissensentdeckung in Datenbanken

Es existieren verschiedene Modelle, die den Prozess der Wissensentdeckung in Datenbanken beschreiben. Typischerweise basieren diese Modelle aber auf gemeinsamen (Haupt-) Abschnitten (SÄUBERLICH, 2000, S. 22ff). Für eine genauere Behandlung der einzelnen Phasen eines WED-Prozesses wird auf WITTEN UND FRANK (2001, S. 41ff) verwiesen. Beispielhaft ist ein Konzept von NAKHAEIZADEH (1998, S. 5) in Abbildung 5 dargestellt.

Ein solcher Prozess beginnt typischerweise mit einem Abschnitt, der *Fokussierung/Planung/Task Analysis* oder *Business Understanding* benannt wird. In diesem Abschnitt geht es darum, das Anwendungsgebiet zu analysieren. Gleichzeitig sollen in diesem Abschnitt die Anforderungen, Ziele und die Machbarkeit überprüft werden. Wichtig ist hierbei auch die Zieldefinition, was mit der Data-Mining-Aktivität erreicht werden soll. Dazu gehört eine umfassende Analyse des Anwendungsgebietes, die bereits bestehendes Wissen mit einschließt. In der hier vorliegenden Arbeit wird dieser Anforderung mit dem zweiten und dritten Kapitel Rechnung getragen, in denen der „Stand der Forschung“ vorgestellt wird. Ebenso großer Bedeutung wird in dieser Phase einer gründlichen Inspektion der vorhandenen Datenquellen zugemessen.

Der zweite Hauptabschnitt wird oft als *Datenvorbereitung* bzw. *Datenvorverarbeitung* bezeichnet. In diesem Abschnitt stehen bereits die Rohdaten zur Verfügung. Diese stammen oftmals aus unterschiedlichen Datenquellen, so dass diese in ein Format integriert werden müssen. Gleichzeitig findet in diesem Abschnitt eine Überprüfung bzw. eine genaue *Exploration* der Daten statt.

Der dritte Abschnitt kann dann als eigentlicher *Data-Mining-Abschnitt* bezeichnet werden. Hier werden die Daten genau untersucht und dann mit Hilfe von Data-Mining-Werkzeugen modelliert. Insbesondere zwischen diesen beiden Phasen wird oftmals vor- und zurückgesprungen, so dass Erkenntnisse aus der Modellierung zu einer entsprechenden Exploration führen und oftmals auch Daten dann entsprechend aufbereitet werden. Im Prozessmodell der Abbildung 5 wird diese Phase als „Anwendung von Modellierungs- und Entdeckungstechniken“ bezeichnet.

Im anschließenden Abschnitt werden die Ergebnisse der Modellierung bewertet und interpretiert. Dazu kann eine Überprüfung anhand von Expertenwissen gehören bzw. die Anwendung von Validierungstechniken.

Die einzelnen Phasen werden in Interaktion mit dem Anwender meist iterativ in Rückkopplungsschleifen durchlaufen (KLÖSGEN UND ZYTKOW, 2002, S. 10). Es handelt sich also keinesfalls um einen linearen Prozess.

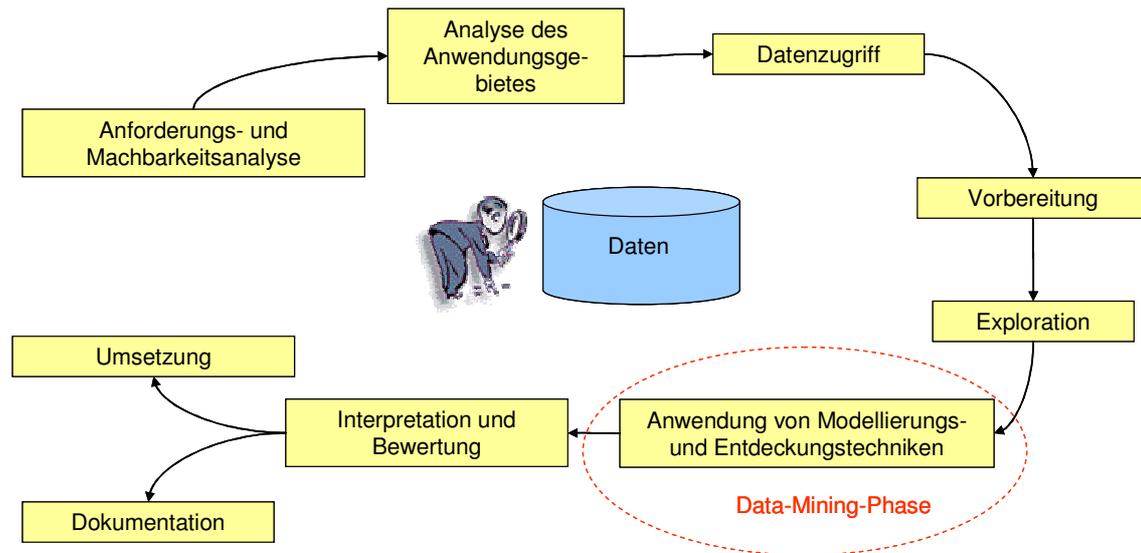


Abbildung 5: WED-Prozessmodell (verändert nach NAKHAEIZADEH, 1998, S. 4)

Der WED-Prozess wird in der Regel durch Softwarepakete unterstützt, die für alle Phasen dieses Prozesses geeignete Werkzeuge liefern. Gleichzeitig hängen wesentliche Funktionalitäten sowie die Ausgestaltung der Algorithmen von diesen Softwarepaketen ab. Ein Vergleich dieser Softwarepakete ist beispielsweise in GENTSCH ET AL. (2000, S. 64ff.) zu finden. Für die vorliegende Arbeit wurde das Softwarepaket Clementine© von SPSS verwendet. Eine Beschreibung dieser Software findet sich in WESTPHAL UND BLAXTON (1998, S. 320ff.). In den folgenden Abschnitten werden die einzelnen Phasen insbesondere im Hinblick auf die im empirischen Teil verwendeten Methoden und Algorithmen genauer vorgestellt.

2.4 Datenvorbereitung und Exploration

Die Datenvorbereitung wird grundsätzlich als eine sehr zeitintensive und wichtige Aufgabe angesehen (vgl. ESTER UND SANDER, 2000, S. 3). Der Unterschied zur Datenaufbereitung der „traditionellen Statistik“ ist in folgender Eigenheit zu finden (WIEDERHOLD, 1996, S. VI):

„The major barrier in obtaining high-quality knowledge from data is due to the limitations of the data itself. The data are rarely collected for the mining of knowledge, but usually as a byproduct of other tasks. [...] The Data may have limited depth, and essential variables may be missing.”

Während in der traditionellen Statistik überwiegend ausgewählte Daten aus speziell entwickelten Versuchen verwendet werden, gelangen in den Data-Mining-Prozess oftmals Daten von wesentlich schlechterer Qualität. Beispielsweise kann es den Eingangsdaten eines WED-Prozesses an einer beschränkten „Tiefe“ mangeln. In diesem Fall decken die Daten nur einen Ausschnitt der in der Realität möglichen Fälle ab. Insbesondere wirkt sich in diesem Zusammenhang das Fehlen von wichtigen erklärenden Variablen negativ aus. Sofern

es sich um wissenschaftliche Versuche handelt, kann unter Umständen in die Datenerzeugung eingegriffen werden, um die Qualität der Auswertungen zu verbessern (vgl. Kap. 4.2.1 und Kap. 5.3).

Die Anforderungen an die Datenvorbereitung sind deshalb in diesem Forschungsbereich sehr hoch. Es geht nicht nur darum, eine bessere und schnellere Modellierung zu erreichen, sondern gleichzeitig soll der Anwender, der diesen Prozess steuert, ein besseres Verständnis für die Daten gewinnen, um Fehlinterpretationen zu vermeiden (WESTPHAL UND BLAXTON, 1998, S. 75). Die Datenvorbereitung kann demzufolge nicht durch Algorithmen automatisiert werden und hängt sehr stark vom Anwendungsgebiet ab.

2.4.1 Regeln zur Datenvorbereitung

Die Datenvorbereitung wird als eine anwendungsspezifische Aufgabe angesehen (WEISS UND INDURKHYA, 1998, S. 16f). Für die Datenvorbereitung gibt es deshalb keine Standards, die bestimmte Tests zur Überprüfung des Datenmaterials nahe legen. Ersatzweise liefert beispielsweise PYLE (2003, S. 368ff) eine Zusammenstellung von Regeln, die dem Anwender helfen können, Fehler in der Datenvorbereitung zu vermeiden.

1. "Select clearly defined problems that will yield tangible benefits.
2. Specify the required solution.
3. Define how the solution delivered is going to be used.
4. Understand as much as possible about the problem and the data set (the domain).
5. Let the problem drive the modeling (i.e., tool selection, data preparation).
6. Stipulate assumptions.
7. Refine the model iteratively.
8. Make the model as simple as possible – but not simpler.
9. Define instability in the model (critical areas where change in output is drastically different for small changes in inputs).
10. Define uncertainty in the model (critical areas and ranges in the data set where the model produces low confidence predictions/insights)."

Diese Regeln reichen auch in die anderen Abschnitte des WED-Prozesses hinein. Insbesondere die Regeln 4, 5, 6, 7 und 10 besitzen besondere Relevanz für Datenvorbereitung. Diese Regeln sind allerdings sehr allgemein verfasst und lassen Spielraum für vielfältige Interpretationen. Dennoch beinhalten sie die wichtigsten Anhaltspunkte.

Von Bedeutung ist zunächst der Bezug zur Problemstellung, aus der sich relevante Ziele ableiten lassen. Die Ausgestaltung der Datenvorbereitung basiert auf einer genauen Analyse des Anwendungsgebietes und der verwendeten Daten. Zur Modellierung sollten Annahmen postuliert und zusammen mit den Modellen schrittweise verfeinert werden. Eine exakte

Analyse der Schwachstellen im Datenmaterial und in den erzeugten Modellen verringert das Risiko von Fehlinterpretationen. Die grundsätzliche Frage, wie Daten beschaffen sein müssen, damit Data-Mining-Modelle daraus lernen können, wird mit diesen Regeln allerdings nicht beantwortet. Dazu muss weiter ausgeholt werden.

2.4.2 Merkmalsraum und Qualität von Daten

Für die grundsätzliche Anwendbarkeit von Modellen des maschinellen Lernens werden an die Daten keine besonderen Ansprüche gestellt. Die Qualität von Daten im Hinblick auf ein Data-Mining-Ziel ist demnach schwer zu definieren. Im Folgenden werden einige Anforderungen dargestellt, die auf der Vorstellung (bzw. auf der Veranschaulichung) eines *Merkmalsraumes*⁸ aufbauen (vgl. PYLE, 1999, S. 84ff; KANTARDZIC, 2003, S. 21f).

Ein Merkmalsraum ist in diesem Kontext ein n-dimensionaler Raum, in dem die Merkmalsausprägungen der Attribute und Zielgrößen aller Instanzen der Trainingsmenge abgetragen sind. Die Anzahl der Dimensionen entspricht der Anzahl der Attribute und Zielgrößen. Die Instanzen werden dabei als Punkte in diesem Raum dargestellt. Aus diesen Punkten wird mit Hilfe von Data-Mining-Werkzeugen eine Oberfläche angepasst, die bestmöglich zu diesen Punkten passt, die sog. *Manifold*.

Die Aufgabe der Datenvorbereitung und Exploration ist es nun, diesen Merkmalsraum in Verbindung mit der Mannigfaltigkeit zu untersuchen. Es können dabei nach PYLE (1999, S. 85) mehrere Probleme auftauchen:

- “If, for instance the “fuzz” is such that the manifold hardly represents the data at all over some portion of its surface, modeling in that area is not likely to produce good results
- In another area there may be very few data points around in state space to define the shape of the manifold. Here, explore the shape as we might, the results will be poor too, but for a different reason than described above.
- Elsewhere the shape of the manifold may be well defined by the data, but have problematic shapes. For instance, it might be folded over on itself rather like a breaking wave. Many modeling tools simply cannot deal with such a shape.
- It is possible that the manifold has donutlike holes in it, or higher-dimensional forms of the anyway.
- There could be tunnels through the manifold.”

Es müssen sich also in jedem Teilbereich des Merkmalsraumes, in dem sich der Anwendungsfall bewegt, genügend Trainingsinstanzen befinden. Je weiter entfernt bzw. schwächer vertreten die Datenpunkte sind, umso wahrscheinlicher sind schlechtere Modellierungsergebnisse. Eine exakte Analyse des Merkmalsraumes erzeugt ein Verständnis dafür,

⁸ Engl. *State Space*. Oftmals werden dazu synonym die Begriffe *Zustandsraum* und *Ereignisraum* verwendet. Ein systematischer Unterschied besteht zu dem Lösungsraum, der durch die Data-Mining-Werkzeuge bestimmt wird.

wie Ergebnisse des Data-Mining-Abschnittes zustande kommen und verbessert somit die Transparenz des gesamten WED-Prozesses.

Es stellt sich die Frage, wie der Merkmalsraum am besten erkundet werden kann. Eine umfangreiche Darstellung dieser Werkzeuge ist beispielsweise in SCOTT UND SAIN (2005, S. 213ff) oder in WESTPHAL UND BLAXTON (1998, S. 123ff) zu finden. Einfache Instrumente der Visualisierung sind in diesem Zusammenhang Histogramme bzw. Scatterplots. Gleichzeitig können dazu Elemente aus der deskriptiven Statistik eingesetzt werden. In einem Histogramm kann die Verteilung eines Attributes erkundet werden. Wesentlich bedeutsamer sind bei multidimensionalen Daten jedoch mehrdimensionale Verteilungen⁹. Mittels Farb- bzw. Formcodierungen können beispielsweise mit Scatterplots auch mehr als zwei Dimensionen dargestellt werden, wobei mit Einbußen in der Übersichtlichkeit zu rechnen ist. 3D-Darstellungen von Punktwolken im Raum leiden an denselben Problemen.

In Abbildung 6 ist ein Scatterplot bezüglich zweier Attribute dargestellt, die miteinander korrelieren. Die beiden Ellipsen spannen beispielhaft eine Fläche auf, innerhalb derer sich die Instanzen lokalisieren lassen. Dabei wird zwischen einer großen Ellipse unterschieden, innerhalb derer sich reale Werte befinden können und zwischen einer kleineren Ellipse, innerhalb derer sich die tatsächlich gemessenen Werte befinden. Diese Abbildung soll einerseits zeigen, dass bei gemessenen Werten jeweils nur ein Ausschnitt der Realität erfasst wird. Wie im Rahmen dieser Arbeit untersucht wird, interagieren beispielsweise Bodenattribute, die einen Schlag beschreiben, untereinander. Es existiert also u. U. ein Zusammenhang zwischen der scheinbaren elektrischen Leitfähigkeit und den historischen Ertragsdaten. Gleichzeitig gibt es in der realen Welt (also beispielsweise in einer anderen Region) Kombinationen dieser beiden Attribute, die innerhalb der weißen, größeren Ellipse liegen (oder innerhalb eines völlig anders geformten Bereiches). Auf einem betrachteten Schlag findet sich jedoch nur ein kleiner Ausschnitt dieser möglichen Kombinationen, die beispielsweise in einer kleineren Ellipse liegen. Es muss also untersucht werden, inwieweit der Ausschnitt mit den Gegebenheiten in der Realität, bzw. im Anwendungsfall vergleichbar ist.

Weiterhin soll anhand dieser Abbildung die Problematik von multidimensionalen Daten dargestellt werden. Falls noch weitere (erklärende) Attribute existieren, die einen Einfluss auf die Zielgröße haben, reicht ein Scatterplot nicht mehr aus, um Verteilungen und Zusammenhänge in den Daten zu erkennen.

Je nach Art des Zusammenhangs der beiden Attribute tauchen sehr verschiedene Formen der Bereiche auf, in denen die realen und gemessenen Datenpunkte liegen. Je mehr zusätzliche Attribute beteiligt sind, desto schwerer wird dieser Zusammenhang (visuell) erkennbar. Den Formen dieser Bereiche sind in der Realität keine Grenzen gesetzt. Es kann also

⁹ In diesem Zusammenhang wird oftmals der Begriff *Joint Occurrence* verwendet.

durchaus der Fall eintreten, dass das x-Attribut den gesamten möglichen Bereich (Range) abdeckt, dass aber gleichzeitig das y-Attribut nur einen Bruchteil des möglichen Bereiches abdeckt, weil noch andere Attribute ihre Wirkung zeigen. Damit ist der betrachtete Ausschnitt nicht repräsentativ für die Grundgesamtheit der möglichen realen Werte. Es ist daher ratsam, die *Joint-Occurrence*, d. h. die Zusammenhänge der Verteilungen der einzelnen Attribute untereinander zu betrachten. Dies kann jedoch bei multidimensionalen Daten schnell zum Problem werden. Bei sieben Attributen lassen sich beispielsweise schon 42 mögliche zweidimensionale Scatterplots erstellen, aus denen diese Effekte abgeschätzt werden können.

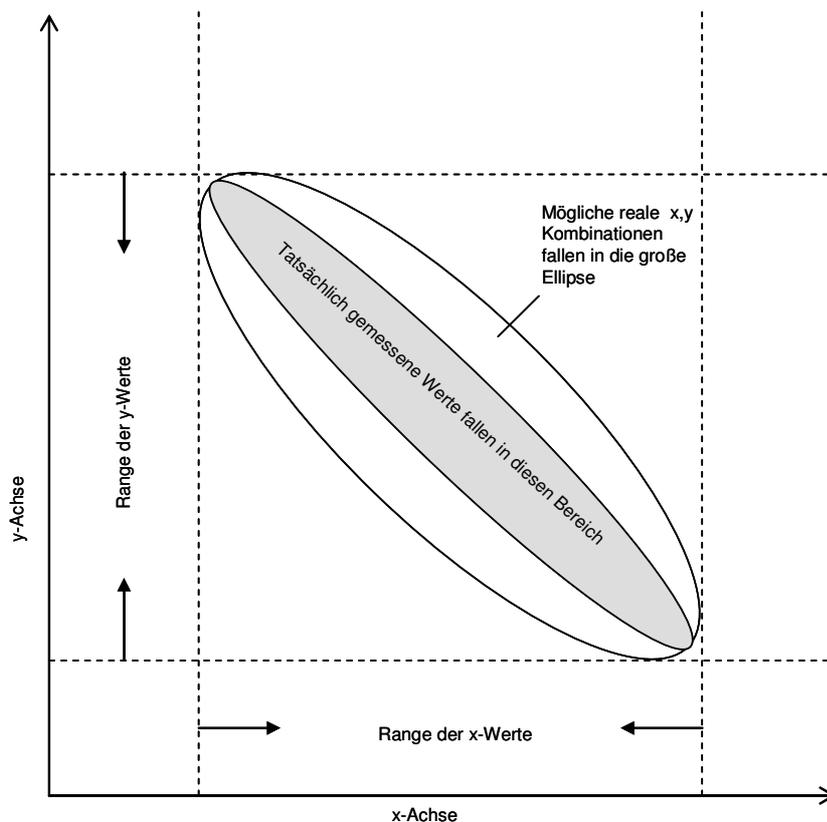


Abbildung 6: Ausschnitt eines Merkmalsraumes (verändert nach PYLE, 1999, S. 80)

Zusammenfassend ist für eine erfolgreiche Data-Mining-Phase nicht in erster Linie der Datenumfang entscheidend. Vielmehr sollte die Datenbasis an Begriffen wie **Datenrepräsentativität**, **Datenhomogenität** und **Datenbalance** gemessen werden. Diese Begriffe kennzeichnen alle den Merkmalsraum der zur Verfügung stehenden Daten. Eine Bewertung der Qualität von Daten kann allerdings immer nur in Verbindung mit der angestrebten Anwendung von Modellen gesehen werden. Es kommt, anders gesagt, auf das Ziel des WED-Projektes an, wie die Qualität von Daten zu beurteilen ist.

Einen ersten Anhaltspunkt für die Qualität von Daten liefert die **Datenrepräsentativität**. Im Allgemeinen sollte jedes Attribut mit den Ausprägungen in den Trainingsdaten vertreten sein, die später bei der Anwendung vorkommen. Entsprechendes gilt für die Häufigkeit der

Merkmalsausprägungen. Der Merkmalsraum der Trainingsdaten sollte dem der Daten in der Anwendung sehr ähnlich sein.

Ein ähnlicher Begriff ist die **Datenbalance**. Diese Umschreibung bezieht sich auf die Verteilung der Ausprägungen von Attributen. Falls numerische Werte verwendet werden, sollten die Daten idealerweise in dem Bereich des Merkmalsraumes, aus dem Wissen generiert werden soll, „gleichmäßig“ vertreten sein¹⁰. Speziell in der Anwendung von neuronalen Netzen kann ein Training mit nicht ausbalancierten Daten zu schlechten Ergebnissen führen.

In eine andere Richtung zielt der Begriff der **Homogenität**. Das bedeutet insbesondere, dass die Trainingsdatenbasis in sich nicht widersprüchlich sein sollte. Die in ihr dargestellte Abbildung sollte funktional sein. Wie auch bei der Datenbalance bzw. der Datenrepräsentativität sollte diese Homogenität auch für Teilmengen der Datenmenge erhalten werden. Das bedeutet beispielsweise, dass Histogramme der Variablen in Test-, Trainings-, und Validationsmengen (vgl. Kap. 2.7.1) eine ähnliche Form haben sollten.

Dennoch spielt auch die schiere Menge an Instanzen eine Rolle: Data-Mining-Algorithmen haben Probleme mit einem beschränkten Datenumfang. Probleme wie Überlernen entstehen unter anderem durch ein wiederholtes Durchlaufen derselben Datensätze. Durch eine große Datenmenge werden die zufälligen Einflüsse des Messrauschens für diese Modelle erst erkennbar.

2.4.3 Veränderungen des Merkmalsraumes

Es existieren einige Möglichkeiten, um einen Merkmalsraum entsprechend des Data-Mining-Zieles und der Data-Mining-Werkzeuge anzupassen. Dieser Vorgang stellt sich allerdings nicht trivial dar. Ein Beispiel für die Komplexität dieser Aufgabe ist in folgendem (vereinfachten) Beispiel zu erkennen:

“Suppose you have two outcomes: low or high. Ninety percent of cases are low, and only 10% are high. Neural networks will respond badly to such biased data. They will learn only the low outcomes and tend to ignore the high ones. Their chance of learning to make accurate predictions is greatly increased if there are roughly equal numbers of each output value. One way of balancing the data in this example would be to use only one-ninth of the low cases for training” (SPSS, 2003, S. 129).

In diesem Beispiel ist eine sehr schlechte Datenbalance offenkundig. Gesetzt den Fall, dass in der Anwendung wiederum dieselbe Verteilung vorherrscht, also wiederum 90 Prozent der Fälle „low“ und 10 Prozent der Fälle „high“, so wird ein mit dieser Verteilung trainiertes Modell eine Funktion annähern, die für den gesamten Datensatz die beste Approximation liefert. Um den Gesamtfehler zu minimieren, passt sich das Modell entsprechend der

¹⁰ Geeignete Verteilungen wären beispielsweise Normalverteilungen bzw. Gleichverteilungen. Problematisch hingegen sind unstetige Verteilungen mit der Konzentration auf wenige Werte.

Häufung an. Die wenigen „Ausreißer“ könnten für das Ergebnis eine untergeordnete Rolle spielen. Im Sinne der Approximationaufgabe wäre mit einem guten Ergebnis zu rechnen.

Das eigentliche Problem stellt eine wechselnde Datenbalance im Trainings- und Anwendungsfall dar. Es kommt also anders gesagt auf das WED-Ziel an. Besteht das Ziel darin, für genau diesen kleinen Bereich des Attributes („high“) eine Funktion zu liefern, bzw. ist dieser Fall besonders interessant, so müsste die Datenbalance für die Modellerstellung angepasst werden.

Die Datenbalance kann verändert werden, indem zum einen die Anzahl der Instanzen mit bestimmten Merkmalen verringert wird. Das bedeutet, dass beispielsweise nur die Hälfte der Instanzen für das Training verwendet wird. Eine andere Möglichkeit besteht darin, Instanzen aus unterrepräsentierten Bereichen zu vervielfältigen oder manuell zu erstellen (vgl. BIGUS, 1996, S. 57).

Insbesondere bei neuronalen Netzen kann es sich auszahlen, gezielt die Verteilung von Merkmalsausprägungen eines Attributes zu steuern. Diese Funktion wird von einigen Data-Mining-Softwarepaketen unterstützt¹¹. In der Regel werden dazu numerische Attribute in kategorische Werte umgewandelt, indem der Bereich, innerhalb dessen die Merkmalsausprägungen liegen, in kleinere Abschnitte unterteilt wird¹². Zwischen diesen Bereichen wird dann eine Gleichverteilung erzeugt. PYLE (2003, S. 369) knüpft die Bedeutung der Datenaufbereitung bei der Anwendung von neuronalen Netzwerken an bestimmte Voraussetzungen:

„Neural Networks, on the other hand, are affected, sometimes severely, by the magnitude of the difference in values. This difference immediately impacts data preparation. When the algorithm is sensitive to value magnitude, two preparation issues become important: (a) when the data contains univariate value clusters in a variable, it is often beneficial to normalize its distribution, or remap the values to be more evenly spread across the variable's range; and (b) single values that are wildly different from the bulk of the values –“outliers”- can create tremendous distortion in the final model. Redistribution also alleviates genuine outlier problems.”

In den genannten Fällen zahlt es sich aus, die Verteilung der Attribute neu zu gestalten. Wenn allerdings in der Anwendung wieder dieselbe Verteilung vorherrscht, sind die Fehler für sehr selten vertretene Klassen bzw. Bereiche u. U. für das Gesamtergebnis uninteressant. Kann, wie beispielsweise in dieser Arbeit, der Effekt der Stickstoff-Applikation mangels Trainingsinstanzen für bestimmte Schlagverhältnisse nicht gut modelliert werden, ist das lediglich ein Problem, sofern der Düngealgorithmus genau auf diese Verhältnisse angewendet werden soll.

Gleichzeitig ist mit der Anpassung der Verteilung eines Attributes auch eine Veränderung des Merkmalsraumes in Bezug auf die restlichen Attribute verbunden. Wird beispielsweise

¹¹ Bei Clementine© wird beispielsweise eine Gleichverteilung eines Attributes erreicht, indem die Funktionen *Boost* oder *Reduce* eingesetzt werden.

¹² Diese Funktion wird als *Binning* bezeichnet.

eine Instanz vervielfältigt, um eine Gleichverteilung eines Attributes zu erreichen, wird der Einfluss dieser Instanz auf die Modellierung verstärkt. Unterliegen die restlichen Attribute dieser Instanzen einer Verzerrung, beeinflusst diese Veränderung die Ergebnisse. Eine Anpassung der Verteilung sollte demnach nur bei sehr wenigen und sehr wichtigen Attributen erfolgen.

Wenn der Datensatz zu Validierungszwecken aufgespaltet werden muss, sollte die Datenbalance und Datenrepräsentativität in allen Teilmengen erhalten bleiben (Datenkonsistenz). Dies erfordert einen gewissen Datenumfang.

Damit zeigt sich, dass das Ausbalancieren von Datensätzen eine Herausforderung darstellt, die aufgrund ihrer Komplexität in der Regel nicht mit Standard-Algorithmen bewältigt werden kann. Es existieren zahlreiche Forschungsanstrengungen zu dem so genannten „class imbalance problem“ (vgl. JAPKOWICZ UND STEPHEN, 2002, S. 429). Allerdings zeichnen sich noch keine konkreten Lösungsmöglichkeiten ab, sodass in der Praxis die Verteilung gegebenenfalls manuell angepasst wird. Welche Datenbalance optimal im Sinne der Modellierungsaufgabe ist, kann objektiv kaum ermittelt werden.

Beseitigung von Ausreißern

Ein Ausreißer ist eine einzelne, oder sehr niedrig frequentierte Beobachtung, die sehr weit von der Masse der anderen Beobachtungen entfernt liegt (PYLE, 1999, S. 73f). Wenn sichergestellt werden kann, dass es sich dabei um einen Fehler handelt, kann er korrigiert werden. Eine Methode zur Korrektur besteht darin, diesen Wert als „fehlenden Wert“ aufzufassen. Einige Modelle reagieren sehr sensibel auf Ausreißer. Ein Problem entsteht, wenn größere Klumpen von Ausreißern existieren, die zum Beispiel darauf hindeuten können, dass ein Messgerät fehlerhaft kalibriert wurde, und sich dadurch eine Verzerrung der Daten ergibt. Während es bei der Betrachtung eines Attributes relativ einfach ist, Ausreißer über ein Histogramm zu erkennen, ist dies bei mehrdimensionalen Aufgaben durchaus problematisch (vgl. HAND ET AL., 2001, S. 62ff; KANTARDZIC, 2003, S. 33f). Eine ausführliche Darstellung von Methoden zur Beseitigung von multivariaten Ausreißern findet sich in HUBERT ET AL. (2005, S. 264ff). Sofern diese Funktionen nicht von den eingesetzten Data Mining Softwarepaketen unterstützt werden, gestaltet sich die Durchführung aufwändig.

Fehlende Werte

Weiterhin wird die Qualität von Daten durch fehlende Werte in Datensätzen beeinträchtigt. Der Unterschied zwischen fehlenden (missing value) und leeren (empty value) Werten besteht darin, dass der fehlende Wert einen korrespondierenden Wert hätte, der einfach nicht erfasst wurde, während es für den leeren Wert keinen solchen Wert gibt. Für die Datenvorbereitung ist es wichtig zu wissen, um welchen dieser beiden Fälle es sich handelt.

Leere Werte können auf ein Muster in Daten hindeuten und sollten auch als solche in den

Prozess mit einfließen. Beispielsweise kann es einen wichtigen Grund geben, warum bestimmte Attribute leer sind. Demgegenüber bergen fehlende Werte die Gefahr in sich, dass ein Muster vorhanden wäre, das aber nicht erfasst wird, weil entsprechende Messungen fehlen (vgl. WITTEN UND FRANK, 2001, S. 57f).

Das Problem besteht darin, dass einige Modelle sehr empfindlich auf fehlende bzw. leere Werte reagieren, wie zum Beispiel neuronale Netze. Es ist aber nicht unbedingt zielführend, eine ganze Instanz wegzulassen, wenn ein Attribut nicht zu ermitteln ist. Die Schwierigkeit besteht darin, Ersatzwerte zu finden, die weder Information hinzufügen noch Information abziehen. Insbesondere sollte dadurch keine Verzerrung entstehen, wie beispielsweise durch den häufigen Einsatz des Mittelwertes als Ersatz zu erwarten ist (PYLE, 2003, S. 370). Eine Aufstellung an alternativen Methoden liefern beispielsweise HAN UND KAMBER (2003).

Damit wurden die wichtigsten Punkte angesprochen, die als Vorbereitung für die Modellierungsphase Beachtung finden sollten. Dabei existieren große Unterschiede, die zum einen von der Anwendung und zum anderen von den gewählten Modellen abhängen. Beispielsweise sind Entscheidungsbaumverfahren robust gegenüber fehlenden Werten und schlechter Datenbalance. Die spezifisch (vor-)verarbeiteten Instanzen werden in der anschließenden Phase zum Erstellen von Modellen benutzt.

Im Folgenden werden zwei Data-Mining-Methoden des überwachten Lernens vorgestellt, nämlich neuronale Netze und Entscheidungsbäume. Diese beiden Werkzeuge werden im empirischen Teil der Arbeit angewendet. Die Auswahl dieser beiden Werkzeuge ergibt sich durch die Vorarbeiten anderer Autoren im Bereich Precision Farming (vgl. Kap. 3.2) und den Zielen dieser Arbeit (vgl. Kap. 1.1). In der vorliegenden Arbeit wurde die Software Clementine® von SPSS® verwendet, daher wird auf die Lernalgorithmen dieser Software genauer eingegangen.

2.5 Data Mining mit künstlichen neuronalen Netzen

Die ersten Publikationen über *künstliche neuronale Netze* (KNN) liegen weit zurück: Bereits in den 60er Jahren erlebten die Arbeiten aus diesem Themenbereich eine frühe Hochzeit, mittlerweile haben sich neuronale Netze einen festen Platz in der Datenanalyse geschaffen (vgl. u. a. SCHERER, 1997, S. 7f; PATTERSON, 1996, S. 22ff). Das Wort „künstlich“ (oder engl. „artificial“) wird verwendet, um sicherzustellen, dass es sich um ein künstliches Modell und nicht um die wirklichen biologischen neuronalen Netze handelt, wie sie beispielsweise beim Menschen anzutreffen sind (CALLAN, 2002, S. 15).

Künstliche neuronale Netze werden in verschiedensten Disziplinen als Instrumente zur Klassifikation und Prognose angewendet. Sie sind im weitesten Sinne mathematische Modelle, die auf den Prinzipien biologischer Systeme aufbauen, obwohl das biologische Vor-

bild des Gehirns eine weitaus komplexere Struktur hat¹³. Dennoch besteht der Grundgedanke darin, durch die Nachbildung biologischer Systeme deren Eigenschaften reproduzieren zu können (WIDMANN, 2000).

Wenn KNN zum überwachten Lernen eingesetzt werden, sollen sie durch Adaption des Netzwerkes in der Lage sein, „zum einen die Trainingsbeispiele weitestgehend zu reproduzieren und zum anderen bei nicht trainierten Fällen eine angemessene Reaktion durch Assoziation bzw. Interpolation zu erreichen“ (KRATZER, 1993, S. 18).

Aus einer Vielzahl von existierenden Modellen künstlicher neuronaler Netze wird in dieser Arbeit ausschließlich das Konzept der *Feedforward-Backpropagation-Netzwerke* behandelt. Dieses Konzept wurde von RUMELHART ET AL. (1986) entwickelt und stellt das in der Praxis am häufigsten verwendete Modell dar. Die universelle Einsatzfähigkeit ergibt sich dadurch, dass mit diesen Netzen jede beliebige Funktion approximiert werden kann¹⁴. Insbesondere eignen sie sich zur Modellierung komplexer nichtlinearer Zusammenhänge, die von den klassischen statistischen Modellen nicht erfasst werden. Gleichzeitig ist der Einsatz von KNN dann interessant, wenn keine Theorien zur Verfügung stehen, die die konkrete Form des anzuwendenden Modells angeben (vgl. ZIMMERMANN, 1994, S. 13ff), wie es auch in vorliegender Arbeit der Fall ist.

2.5.1 Aufbau und Technik

Neuronale Netze verwenden Bausteine, die dem biologischen Vorbild, dem Nervensystem, angelehnt sind. Ein neuronales Netz aus der Klasse der *Multi-Layer-Perceptrons* besteht zunächst aus mehreren *Neuronen*¹⁵, die in verschiedenen *Schichten* angeordnet und nach bestimmten Regeln miteinander verbunden sind. Im Falle von vorwärtsgekoppelten¹⁶ neuronalen Netzen erfolgt der Signalfluss ausschließlich in eine Richtung, und zwar von allen Neuronen der vorgelagerten Schicht an alle Neuronen der nachgelagerten Schicht. Innerhalb einer Schicht werden keine Signale ausgetauscht. Ein Beispiel eines vorwärtsgekoppelten Netzwerkes mit einer versteckten Schicht ist in Abbildung 7 dargestellt.

Die Neuronen, die Input-Daten aus der externen Welt aufnehmen, so genannte *Eingabeneuronen*, liegen auf der Eingabeschicht. Die *Ausgabeneuronen* haben die Aufgabe, an der Ausgangsschicht Informationen an die externe Welt abzuleiten. Die Anzahl der Input- bzw. Outputneuronen ist dabei flexibel und hängt davon ab, wie viele Input- bzw. Output-Attribute in der Datenmatrix (vgl. Abbildung 3) vorliegen. Die Dimensionierung der Eingabe- und Ausgabeschicht ist in der Regel durch die Problemstellung vorgegeben: Die An-

¹³ Ein neuronales Netz besteht im Allgemeinen aus nicht mehr als 10^2 Neuronen, ein menschliches Gehirn hat ungefähr 10^{11} Neuronen.

¹⁴ Eine Zusammenstellung der mathematischen Beweise für die Approximationsfähigkeit von sigmoiden Netzen ist beispielsweise bei Widmann, 2001, S.21 dargestellt.

¹⁵ Alternativ: Zelle, Element, Prozessoren, (Verarbeitungs-)Einheiten, units, etc.

¹⁶ Engl. feedforward.

zahl der Neuronen der Eingabeschicht wird determiniert durch die Anzahl der Input-Attribute. Es wird also für jedes Attribut eine eigene Eingabeeinheit verwendet. Für die Zielgröße (Output-Attribute) gilt selbiges. Jede Zielgröße wird durch ein Neuron in der Ausgabeschicht repräsentiert.

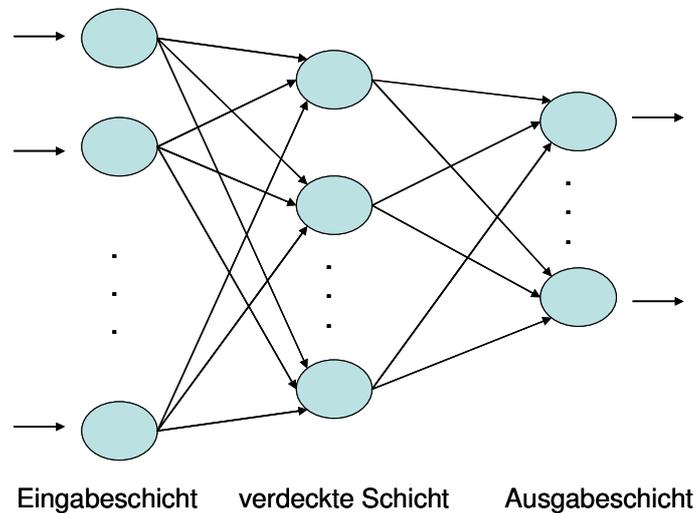


Abbildung 7: Mehrlagiges, vorwärtsgekoppeltes neuronales Netzwerk (JAFAR-SHAGHAGI, 1996, S. 185)

Dazwischen liegen die verdeckten Neuronen, die in einer oder mehreren verdeckten Schichten angeordnet sind. Der Aufbau bzw. die Architektur des Netzes und insbesondere die Dimensionierung der verdeckten Schicht(en) werden unter dem Begriff *Netztopologie* zusammengefasst. Die Netztopologie wird üblicherweise in der Form

$$N_E - N_{V1} - (N_{V2} - \dots) - N_A$$

angegeben, wobei N_E die Anzahl der Neuronen in der Eingabeschicht, N_{V1} , N_{V2} usw. die Anzahl der Neuronen in den verdeckten Schichten sowie N_A die Anzahl der Neuronen in der Ausgabeschicht angibt. Ein 5-7-1 Netz besteht also aus fünf Neuronen in der Eingabeschicht, sieben Neuronen in einer verdeckten Schicht sowie einem Neuron in der Ausgabeschicht.

Die Neuronen sind über Verbindungen, die als Gewichte (w) bezeichnet werden, miteinander verbunden. Die Gewichte erregen oder hemmen das zu kommunizierende Signal. Über das Vorzeichen und den Absolutbetrag des Verbindungsgewichtes w kann somit formal im Modell ein anregender bzw. ein hemmender Einfluss mit variabler Intensität¹⁷ abgebildet werden (vgl. OBERHOFER UND ZIMMERER, 1996, S. 8f).

In Abbildung 8 ist ein Neuron in seiner Funktion dargestellt. Das Neuron (j -Element) empfängt Signale von den Neuronen der vorgelagerten Schicht i (Output, o). Diese eingehenden

¹⁷ Nimmt w_{ij} den Wert Null an, so besteht beispielsweise keine Verbindung.

Signale werden durch Aufsummieren der gewichteten Signale kombiniert (net_j).

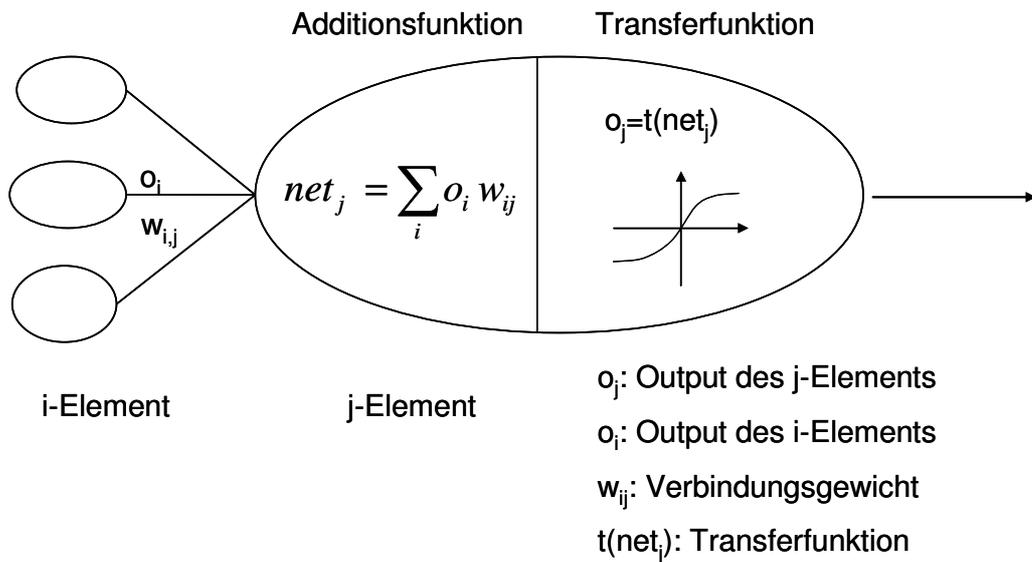


Abbildung 8: Aufbau eines Neurons

Aus dieser Summe wird dann mit einer Transferfunktion¹⁸ das Ausgangssignal (Output) definiert. In der Regel werden sigmoide Transferfunktionen angewendet, da diese sich insbesondere für die Modellierung nicht-linearer Beziehungen zwischen Ein- und Ausgabedaten eignen. Zudem erfüllen sie die Anforderungen nach Stetigkeit, Differenzierbarkeit und Monotonie, was insbesondere bei dem hier angewendeten Backpropagation-Lernalgorithmus eine große Rolle spielt. Die sigmoide Transferfunktion ist definiert als:

$$t(net_j) = \frac{1}{1 + e^{-net_j}}$$

Ihr Wertebereich ist auf das offene Intervall $]0,1[$ beschränkt. Durch das asymptotische Verhalten wird sichergestellt, dass sich ein extremes Verhalten einzelner Neuronen nicht über das ganze Netzwerk fortpflanzen kann. Transferfunktionen werden allerdings nur in den versteckten Schichten angewendet. Die Transferfunktion für Eingabe- und Ausgabe-Neuronen ist die Identitätsfunktion; das Signal wird somit von diesen Schichten untransferriert weiter- bzw. ausgegeben.

Die Eigenschaft neuronaler Netze, Funktionen von hoher Komplexität approximieren zu können, resultiert daraus, dass aus der Verbindung von Einzelelementen, die funktionell miteinander verbunden sind, äußerst komplexe Gesamtstrukturen entstehen können.

Die Muster oder Zusammenhänge, die über eine Trainingsmenge gelernt werden, werden über die Stärke der Verbindungen (Verbindungsgewichte) zwischen den einzelnen Einheiten (Neuronen) gespeichert. Es erfolgt also eine Art Adaption, wie sie in biologischen Systemen charakteristisch ist.

¹⁸ Auch: Aktivierungsfunktion (engl. Activation Function).

2.5.2 Lernverfahren

Das Ziel der Funktionsapproximation wird bei neuronalen Netzen dadurch erreicht, dass die Verbindungsgewichte angepasst werden. Bei komplexeren Lernverfahren, die Überlernen verhindern sollen, wird gleichzeitig die Netztopologie angepasst. Die Gewichte werden durch die Lernbeispiele aus der Trainingsmenge adaptiert. Viele Jahre lang existierte keine Regel für die Anpassung der Gewichte eines mehrschichtigen Netzes. Mit dem *Backpropagation-Algorithmus* entwickelten RUMELHART ET AL. (1986, S. 319ff) eine Methode, die neuronale Netze zu einem universell einsetzbaren Werkzeug machte.

Das Grundprinzip des überwachten Lernens besteht darin, die Ausgabe des KNN mit dem Zielwert aus den einzelnen Instanzen zu vergleichen. Die Abweichung zwischen dem Output und dem Zielwert dient als Steuer- bzw. Fehlersignal, um die Verbindungsgewichte zu verändern.

Die Basis für den Backpropagation-Algorithmus bildet die **Delta-Regel** (WIDROW UND HOFF, 1960), die im Folgenden kurz dargestellt werden soll. Die Notation ist identisch zu Abbildung 3 und Abbildung 8. Allerdings gilt die Delta-Regel nur für ein sehr einfaches Netz mit lediglich einer Eingabe- und einer Ausgabeschicht. Die Gewichtsänderung nach der einfachen Delta-Regel für ein Netzwerk ohne Zwischenschichten gestaltet sich folgendermaßen:

$$\Delta w_{ij} = \eta * (t_j - o_j) * x_i \quad \text{bzw.} \quad \Delta w_{ij} = \eta * \delta_j * x_i$$

mit

- w_{ij} : Gewichte zwischen den Neuronen i und j
- η : Lernrate (Anpassungsgeschwindigkeit für das Lernen zwischen 0 und 1)
- t_j : Erwarteter Ausgabewert des Neurons j in der Ausgabeschicht (Zielwert)
- o_j : Ausgabewert des Neurons j in der Ausgabeschicht, der vom Netz berechnet wurde
- x_i : Eingabe des Neurons j aus dem Neuron i (Netzeingabe)
- δ_j : Fehlerfaktor bzw. Fehlersignal für das Neuron j

Die Gewichte des Netzes werden dabei zu Beginn des Trainings auf zufällige Werte festgelegt. Für jede Trainingsinstanz wird ein Vorgang der Netzanpassung (Gewichtsänderung) nach der Delta-Regel durchgeführt. Mit jeder Trainingsinstanz wird der jeweilige Zielausgabewert der Ausgabeschicht j (t_j) und der tatsächliche Ausgabewert (o_j) bestimmt. Die Differenz aus den beiden Werten wird als Faktor zur Ermittlung der Gewichtsänderung verwendet. Über die Lernrate η wird bestimmt, wie schnell die Anpassung erfolgt. Der dritte Faktor, der zur Gewichtsänderung verwendet wird, ist der Eingabewert für das jeweilige Attribut, d. h. die jeweilige Attributausprägung einer Instanz.

Aus diesen drei Faktoren kann die nötige Gewichtsänderung für jedes der Verbindungsge-

wichte zwischen Ein- und Ausgabeneuronen ermittelt werden. Das neue Gewicht ist dann das alte Gewicht zuzüglich der Anpassung ($w = \Delta w + w$). Durch die wiederholte Anwendung der Delta-Regel konvergiert die Ist-Ausgabe des Netzes in Richtung der gewünschten Soll-Ausgabe. Der Durchlauf aller Trainingsdatensätze wird als Epoche bezeichnet. Dieser Trainingsvorgang wird so oft wiederholt, bis **alle** Ausgaben während einer Epoche innerhalb eines Toleranzbereiches liegen.

Die einfache Delta-Regel funktioniert allerdings nur bei einem zweischichtigen Netz, bei dem jeweils der Output des Netzes direkt mit dem Zielwert verglichen werden kann. Die Abweichungsfeststellung für versteckte Neuronen kann nicht erreicht werden, da für die mittlere(n) Schicht(en) kein Zielwert vorliegt.

Dieses Problem wurde mit dem Backpropagation-Algorithmus gelöst. Der Backpropagation-Algorithmus basiert auf einer verallgemeinerten Delta-Regel, die die Anwendung für mehrschichtige Netze erlaubt. Die Herleitung und genauen Rahmenbedingungen für den Backpropagation-Algorithmus können in RUMELHART (1986, S. 319ff) bzw. ZHANG (2000, S. 117) gefunden werden. Die Anwendung dieses Lernverfahrens wird u. a. von KOLLIG (1993, S. 64ff) an einem Beispiel dargestellt.

Die Idee besteht darin, die Fehlersignale der Neuronen der übergeordneten Schicht anteilmäßig auf die Neuronen der Zwischenschicht aufzuteilen. Dabei wird ein Neuron mit einem größeren Fehlersignal eine stärkere Auswirkung auf die Fehlererzeugung der nachfolgenden Neuronen haben (JAFAR-SHAGHAGI, 1996).

Gleichzeitig muss dazu die Transferfunktion beachtet werden, die bisher ausgeblendet wurde. Um zu erfahren, in welche Richtung und wie schnell sich das Fehlersignal verändert, wenn die Aktivität eines Ausgabeneurons verändert wird, wird die Ableitung dieser Transferfunktion verwendet.

Mit dem Backpropagation-Algorithmus kann eine Gewichtsangpassung in den versteckten Schichten erreicht werden. Die Grundidee besteht darin, das Fehlersignal in einem Rückwärtsdurchlauf durch das Netz zu erzeugen. Dazu werden zwei Netzdurchläufe definiert: zuerst einen Vorwärtsdurchlauf von der Eingabeschicht zur Ausgabeschicht und anschließend einen Rückwärtsdurchlauf von der Ausgabeschicht zur Eingabeschicht. Die Gewichtsangpassung im Backpropagation-Algorithmus erfolgt proportional zur partiellen Ableitung der Fehlerfunktion nach den Verbindungsgewichten (vgl. OBERHOFER UND ZIMMERER, 1996, S. 17ff). Für die Berechnung der aktuellen Gewichtsmodifikationen gehen die Fehlersignale sämtlicher nachgelagerter Verbindungsgewichte mit ein.

Ausgehend von diesem Grundprinzip wird das allgemeine Ziel des Lernverfahrens deutlich. Es müssen alle Gewichte schrittweise angepasst werden, um den Gesamtfehler zu minimieren. Dieses Minimierungsprinzip wird als Gradientenabstiegsverfahren bezeichnet.

2.5.3 Erweiterungen für ein optimiertes Lernen

Es existieren zahlreiche Möglichkeiten, den Backpropagation-Algorithmus zu erweitern, bzw. ihn für spezielle Anwendungsfälle zu optimieren. Diese Aufgabe wird insbesondere auch von den Entwicklern der Data-Mining-Softwarelösungen erledigt. Die eingesetzten Erweiterungen hängen damit stark von der verwendeten Software ab. In der vorliegenden Arbeit wird das Software-Paket Clementine© von SPSS® verwendet. In diesem Kapitel werden daher die Optionen dargestellt, die in diesem Programm realisiert sind.

Allerdings sind die detaillierten Algorithmen nicht veröffentlicht. Anstelle der Algorithmen müssen daher überwiegend Beschreibungen der Handbücher verwendet werden¹⁹. Dies gilt insbesondere für die Lernmethode *Exhaustive-Prune* die zur Vermeidung von Überlernen standardmäßig im empirischen Teil dieser Arbeit verwendet wird. Verschiedene Varianten ähnlicher Lernmethoden sind beispielsweise in BISHOP (1995, S. 353ff) zu finden.

Das Ziel aller Gradientenabstiegsverfahren stellt die Erreichung eines globalen Minimums im so genannten Fehlergebirge dar, das durch die Anpassung der Gewichte (bzw. der Topologie) erreicht werden kann. Ein Beispiel dieses Fehlergebirges ist in Abbildung 9 dargestellt. Auch bei neuronalen Netzen besteht die Gefahr, in lokalen Minima steckenzubleiben und das globale Minimum, d. h. die beste Approximation einer Funktion, somit niemals zu erreichen.

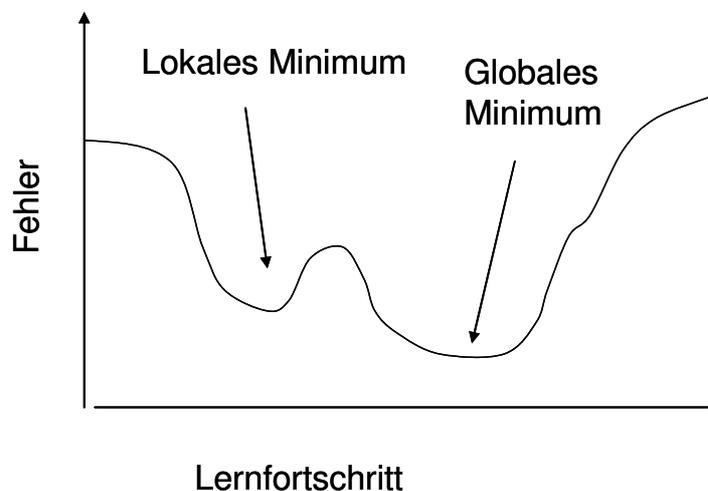


Abbildung 9: Darstellung eines Fehlergebirges

Eta-Anpassung und Momentum-Term

Ein Problem in diesem Zusammenhang ist die ex-ante Bestimmung der Lernrate (PATTERSON, 1996, S. 180ff). Wenn der Wert für den Anpassungsschritt zu groß gewählt wird, besteht die Gefahr, dass keine Fehlerminimierung erfolgen kann, weil ein Minimum mit ei-

¹⁹ Die Beschreibung der folgenden Algorithmen stützt sich u. a. auch auf Antworten des „technical supports“ von SPSS®.

nem Modifizierungszyklus übersprungen wird. Kleinere Lernschritte bedingen langsamere Fortschritte und damit einen höheren Zeitbedarf. In der Regel führen sie aber zu besseren Ergebnissen (GUERNEY, 2002, S. 70f). Eine gängige Strategie ist die Variation der Lernrate. Beginnend mit relativ großer Lernrate, wird diese dann im Verlauf des Trainings sukzessive verkleinert.

Der Algorithmus der verwendeten Software startet bei einem frei einstellbaren Wert (*initial eta*). Dieser Wert wird dann logarithmisch verringert, bis er eine untere Grenze (*low eta*) erreicht. Von diesem Punkt springt er auf eine obere Grenze (*high eta*) und wird dann wieder logarithmisch verringert²⁰.

Eine weitere Möglichkeit, ein konvergentes Lernverhalten zu erzielen wird durch die Einführung einer Konstante in der Delta-Regel erreicht. Diese wird in Anlehnung an die physikalische Mechanik oft als Moment oder Momentum-Term bezeichnet. Die Gewichtsänderungen berechnen sich in dieser Variation aus der bekannten Delta-Regel und dem Produkt einer Konstanten (μ) mit der Gewichtsänderung der vorangegangenen Periode.

$$\Delta w_{ij}(t+1) = \eta * \delta * x_i + \mu * \Delta w_{ij}(t)$$

Der Momentum-Term verringert ebenfalls die Gefahr von Oszillationen (SCHERER, 1997, S. 79f). In der Vorstellung des Fehlergebirges werden die Effekte der zwei „Stellschrauben“ oftmals mit einem Ball veranschaulicht, der das Fehlergebirge hinunterrollt. Wenn das Fehlergebirge zu steil ist (hohes eta) bzw. der Ball zu klein ist (Momentum-Term), bleibt er unter Umständen in einem lokalen Minimum hängen. Ist der Ball zu groß bzw. das Fehlergebirge zu flach, kann er unter Umständen ein globales Minimum überspringen.

Vermeidung von Überlernen und selbstanpassende Netzwerke

Eines der zentralen Probleme bei der Anwendung von neuronalen Netzen ist das Überlernen (vgl. Kap. 2.2). Redundante Eingabeneuronen erhöhen die Anzahl der Parameter des Netzes, ohne einen nennenswerten Erklärungsbeitrag zu leisten. Sie verstärken damit die Gefahr des Überlernens. Dasselbe gilt für die versteckten Neuronen.

Die Aufgabe des Lernens besteht also nicht nur darin, das Netz den Trainingsdaten beliebig genau anzupassen, sondern parallel Komplexitätsbeschränkungen durchzuführen, die das Überlernen begrenzen. Für eine Zusammenstellung dieser Techniken sei auf ZIMMERMANN (1994, S. 40ff) verwiesen. Nach KRATZER (1993, S. 143ff) existiert jedoch bisher kein „Königspfad“, wie die Topologie eines Netzes bestimmt werden kann.

In der verwendeten Software Clementine® wird zur Vermeidung von Überlernen die zur Verfügung stehende Datenmenge in einem frei wählbaren Verhältnis in zwei Teilmengen aufgeteilt. Die Instanzen des Trainingssamples werden dazu verwendet, die Netzgewichte

²⁰ Im empirischen Teil wird folgende Einstellung verwendet: Alpha (Momentum Term) = 0.9, Initial Eta: 0.4, High Eta: 0.15, Low Eta: 0.01.

nach dem Backpropagation-Verfahren anzupassen.

Das Test-Sample wird zunächst dazu verwendet, die Trainingsphase zu beenden. Dazu wird versucht, den „optimalen“ Zeitpunkt zu ermitteln, bei dem das Netz die Zusammenhänge erfasst hat und gleichzeitig die Gefahr des Überlernens noch nicht zu groß ist.

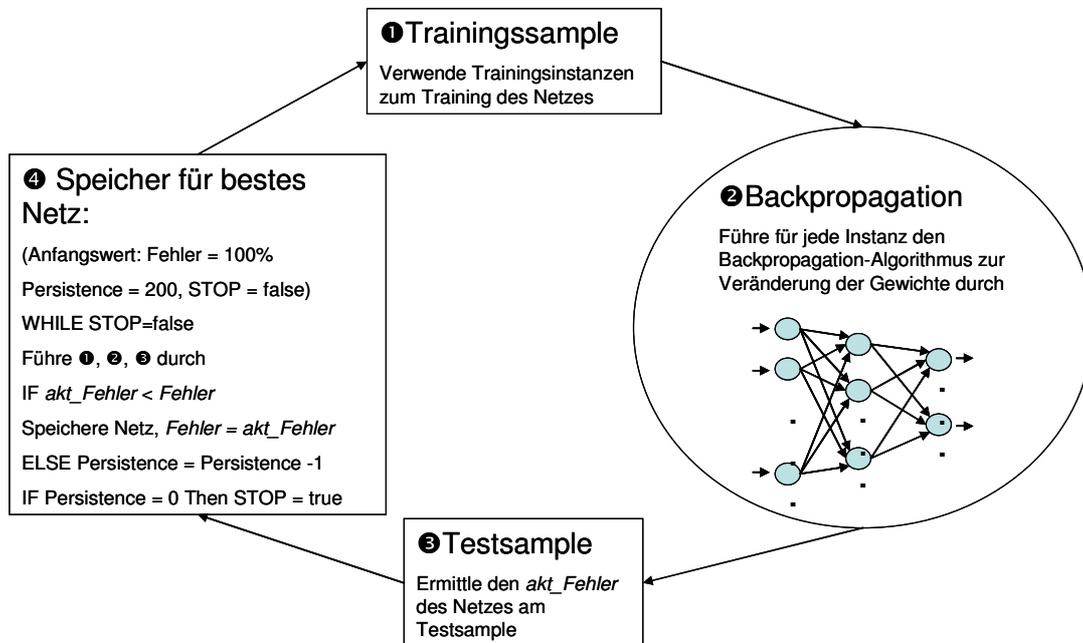


Abbildung 10: Vermeidung von Überlernen bei Clementine®

Der Trainingsalgorithmus basiert auf dem Begriff „Persistence“. Nach jedem Trainingszyklus wird der Fehler in Bezug auf die Test-Stichprobe gemessen und ergibt sich aus der Summe der quadrierten Abweichungen. Wenn dieser Fehler der bis dato geringste ist, wird der Zustand des Netzwerkes gespeichert. Das Training läuft dann weiter. Wenn sich ein Netzwerk nach einer frei wählbaren Anzahl von Zyklen nicht verbessert, gilt das Netzwerk als „trainiert“. Diese Anzahl wird *Persistence* bezeichnet. Das Training wird dann abgebrochen, oder geht weiter in die nächste Phase. Das Ausgabenetzwerk ist dasjenige, bei dem der Fehler am geringsten ist. Diese Phase ist mit dem entsprechenden Pseudocode in Abbildung 10 dargestellt.

In Anlehnung an den *Exhaustive-Prune-Algorithmus*, der in dieser Arbeit verwendet wurde, wird eine *Persistence* von 200 gewählt. Das bedeutet, jedes Mal, wenn ein bis dato „bestes Netz“ erzeugt wurde, werden noch 200 Zyklen durchgeführt und getestet, ob sich das Netz noch weiter verbessern lässt. Sind diese Zyklen durchgeführt, ohne dass sich das Netz verbessert hat, wird das Training abgebrochen.

Eine weitere Möglichkeit, das Lernverhalten von neuronalen Netzen zu optimieren, liegt in der Optimierung der Topologie. Bei einer großen Dimensionierung neigt das Netz zum Überlernen. Die Anzahl der Netzgewichte nimmt mit einer großen Topologie entsprechend

zu und erhöht damit die Anzahl der benötigten Trainingsinstanzen (BIGUS, 1996, S. 57). Wird die Zahl der versteckten Neuronen verringert, neigt das Netz zu einem interpolativen Verhalten, das die Qualität des überwachten Lernens verschlechtert.

Zur Gestaltung der Topologie existieren keine Regeln, die den Anwender in der Entscheidung unterstützen könnten. Aus diesem Grund wurden selbstanpassende Netzwerke entwickelt. Diese testen im Wesentlichen verschiedene Topologien an einem Testsample und messen jeweils den Fehler. Es wird mit einem sehr großen Netzwerk mit allen Input-Einheiten und sehr vielen Knoten in der versteckten Schicht begonnen und dann schrittweise Neuronen entfernt²¹. Die Reihenfolge ergibt sich aus der *Sensitivität*. Die Sensitivität eines Input-Attributes wird im verwendeten Softwarepaket ermittelt, indem der Input-Wert für jede Instanz im Testsample über fünf verschiedene Werte [0, 0.25, 0.5, 0.75, 1] variiert wird und dann jeweils die maximale Differenz zwischen den Output-Werten eines trainierten neuronalen Netzes gemessen wird. Dieser Wert wird dann über alle Instanzen aufsummiert und normalisiert.

Als Pseudocode stellt sich dieser Prozess folgendermaßen dar (WATKINS, 1997, S. 6):

For each input attribute

For each record

Vary the value and record the max difference in the output

Calculate the mean of the max differences

Diese Sensitivitätsanalyse ergibt Hinweise über die Bedeutung eines Attributes für die Prognose (*Relative Importance*). Die Sensitivität hat nichts mit der Genauigkeit bzw. Fehlerhöhe zu tun. Es wird jeweils lediglich ermittelt, wie sich die Variation eines Attributes, d. h. eines Input-Neurons auf den Output auswirkt.

In ähnlicher Form lässt sich eine Sensitivitätsanalyse für die Neuronen der versteckten Schicht(en) durchführen. Hier werden zuerst die Output-Werte eines trainierten neuronalen Netzes an einem Testsample erzeugt und gespeichert. Dann wird jeweils ein Neuron der versteckten Schicht(en) entfernt (indem das Gewicht auf Null gesetzt wird) und die Differenzen aus diesen Werten und den gespeicherten Werten bestimmt. Daraus kann eine Art Standardabweichung ermittelt werden. Je größer dieser Wert ist, desto nützlicher ist dieses Neuron für das Netzwerk.

Die Funktionsweise dieser Methode ist in Abbildung 11 dargestellt. Als Kernelement dient die in Abbildung 10 dargestellte Phase 1 des Trainings. Nach dem Training dieses Netzes erfolgt eine Sensitivitäts-Analyse auf die Einheiten der versteckten Schichten und ein defi-

²¹ 30 Einheiten in der versteckten Schicht 1 und 20 in der versteckten Schicht 2.

nierter Anteil²² der schwächsten Neuronen wird entfernt. Dieser (Teil-)Prozess aus Trainieren und Entfernen von Verbindungen wird wiederum über ein Maß der *Persistence* gesteuert. Er wird so lange wiederholt, bis sich innerhalb der *Hidden Persistence*, in diesem Fall über 100 Zyklen, keine Verbesserung ergibt.

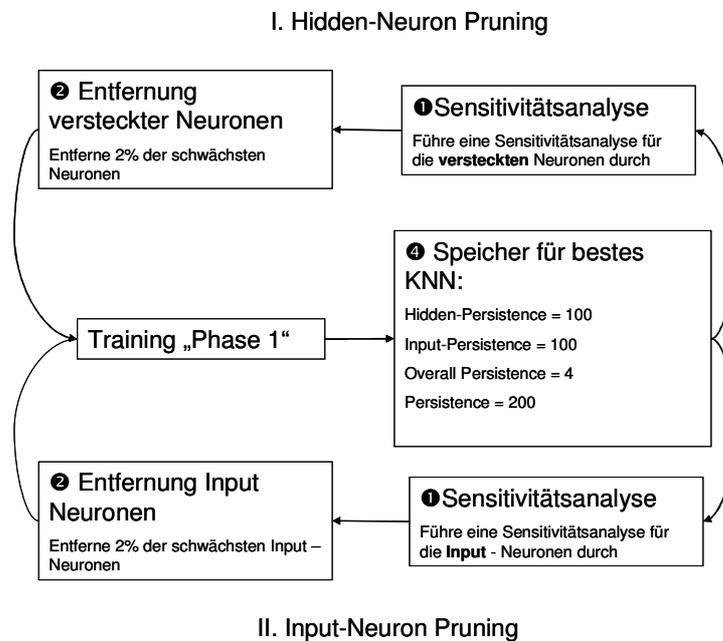


Abbildung 11: Funktionsweise der Trainingsmethode *Exhaustive Prune* (Software *Clementine*®)

Nachdem diese Schleife beendet ist, wird ein ähnliches Verfahren mit den Input-Einheiten wiederholt. Das heißt, es wird ein Prozess aus Trainieren und Wegschneiden wiederholt und das beste Netz gespeichert. Tritt nach 200 Zyklen keine Verbesserung des Fehlers auf, wird das Input-Neuron-Pruning beendet und das Hidden-Neuron-Pruning beginnt erneut.

Der gesamte Pruning-Prozess, bestehend aus den beiden Teilprozessen, wird maximal viermal wiederholt, ohne dass sich eine Verbesserung ergibt. Dieses Abbruchkriterium wird als *Overall-Persistence* bezeichnet. Wenn alle *Persistence-Zyklen* durchgeführt sind und damit am bis dato besten Netz auch in den 200 *Persistence-Zyklen* keine Verbesserung erzielt wurden, wird der gesamte Prozess abgebrochen.

Insbesondere das Abschneiden von Input-Neuronen ist für den Lernfortschritt von entscheidender Bedeutung (WIDMANN, 2000, S. 100ff). Dieses Abschneiden kann natürlich auch vom Benutzer selbst erfolgen, wenn sich herausstellt, dass bestimmte Variablen für die Prognose von geringer Bedeutung sind. Als Ergebnis dieser Algorithmen wird ein Netz ausgegeben, in dem unwichtige Einheiten automatisch entfernt wurden. Damit wird Überlernen verhindert.

²² Dieser Anteil wird als *hidden rate factor* bezeichnet und ist bei Verwendung des *Exhaustive-Prune*-Algorithmus auf 0.02 eingestellt.

Datenvorbereitung

Zur Datenvorbereitung wird insbesondere auf Kapitel 2.4 verwiesen. Zusätzlich existieren für neuronale Netze weitere Anforderungen. Besonderer Fokus sollte bei neuronalen Netzen auf der Datenbalance liegen (OLSON, 2004, S. 71ff).

Für neuronale Netze mit einer sigmoiden Transferfunktion ist es weiterhin von Bedeutung, die Eingabewerte für ein neuronales Netz in der Form vorzubereiten, dass sie bezüglich ihres Mittelwertes und ihrer Streuung standardisiert werden (vgl. BISHOP, 1995, S. 298f).

Bei neuronalen Netzen mit einer sigmoiden Transferfunktion muss bedacht werden, dass die Aktivierungsfunktion sowohl am linken als auch am rechten Rand einen Sättigungsbereich besitzt, in dem Änderungen der Gewichte kaum Auswirkungen auf die Netzausgabe haben. Sind die Eingabewerte nicht standardisiert, so ist die Gefahr groß, dass man sich durch ungünstige Anfangswerte für die Netzwerkgewichte schon zu Beginn des Netzwerktrainings in diesem Sättigungsbereich befindet. Dies behindert den Lernprozess, da in den einzelnen Anpassungsschritten nur geringfügige Änderungen vorgenommen werden. Es ist sogar möglich, dass der Lernprozess abgebrochen wird, weil durch die geringfügigen Änderungen der Netzwerkgewichte beim Training das Abbruchkriterium erfüllt ist. Die numerischen Daten (Eingabedaten und Ausgabedaten) werden in der vorliegenden Arbeit deshalb auf das Intervall $[0, 1]$ transformiert, indem folgende Formel angewendet wird:

$$(\text{Wert} - \text{Untere Grenze}) / (\text{Obere Grenze} - \text{Untere Grenze})$$

Zusammenfassend kann festgestellt werden, dass einige Möglichkeiten existieren, die Leistung des Backpropagation-Ansatzes zu erhöhen. Die Anwendung der vorgestellten Algorithmen verbessert die Leistung der Funktionsapproximation, da die Ergebnisse ständig anhand einer unabhängigen Stichprobe überprüft werden. Gleichzeitig verringern sich dadurch die Aufgaben des Anwenders, da aufwändige Untersuchungen zur Dimensionierung des Netzwerkes und zur Vermeidung von Überlernen wegfallen. Allerdings wird mit den Erweiterungen die Komplexität der Modellierung weiter erhöht und die Ergebnisse für den Benutzer dadurch schwerer zu verstehen.

2.6 Data Mining mit Entscheidungsbaumverfahren

Ein alternatives Verfahren des überwachten Lernens stellen so genannte *Entscheidungsbaumverfahren* dar. Die Einsatzgebiete von Entscheidungsbäumen sind zum Teil ähnlich wie die von neuronalen Netzen, aber die Ergebnisse unterscheiden sich grundsätzlich gegenüber neuronalen Netzen. Entscheidungsbaumverfahren leisten **gleichzeitig** eine Approximation einer Funktion und eine strukturierte Beschreibung (WITTEN UND FRANK, 2001, S. 228). Ein Entscheidungsbaumverfahren liefert also explizites Wissen (ESTER UND SANDER, 2000, S. 107). In dieser Beschreibung liegt ein wesentlicher Unterschied zu neuronalen

Netzen: Diese liefern als Ergebnis eine „Black Box“. Das Muster bzw. das Wissen ist in der Netztopologie und in den Verbindungsgewichten enthalten, kann aber nicht interpretiert bzw. verstanden werden. Der Vorteil von Entscheidungsbaum-Klassifikatoren ist also in der einfachen Interpretation des gefundenen Baumes zu sehen und wird deshalb in der Praxis häufig angewendet.

Entscheidungsbaumverfahren können den nicht-parametrischen Regressions-Methoden zugeordnet werden. Die hier vorgestellten Algorithmen sind so genannte Top-Down- oder divisive Verfahren. Die gesamte Menge der Trainingsdaten wird zunächst der Wurzel eines zu generierenden Baumes zugeordnet. Dann wird die augenblicklich interessierende Trainingsdaten(teil)menge mit Hilfe bestimmter Tests in weitere Teilmengen unterteilt, die neuen Knoten des Baumes zugewiesen werden. Ebenso wird mit jedem dieser neuen Knoten verfahren, bis solche Knoten ein bestimmtes Abbruchkriterium erfüllen und damit zu Endknoten werden. Eine formale Darstellung dieser Vorgehensweise ist beispielsweise in SÄUBERLICH (2000, S. 80ff) dargestellt.

In dieser Arbeit werden so genannte CART (Classification and Regression Trees)-Verfahren angewendet. Dieses Verfahren wurde erstmals von BREIMAN ET AL. (1984) beschrieben und kann sowohl für numerische Attribute (Regressionsbaum) als auch für kategoriale Attribute (Klassifikation) eingesetzt werden.

2.6.1 Aufbau und Technik

Ein Entscheidungsbaum ist nach ESTER UND SANDER (2000, S. 126) ein Baum mit folgenden Eigenschaften:

- ein Knoten (Astgabel) repräsentiert ein Attribut,
- eine Kante repräsentiert einen Test auf dem Attribut des Vaterknotens,
- ein Blatt repräsentiert eine der Klassen.

Entscheidungsbäume werden von oben nach unten gelesen und beginnen mit dem Ausgangsknoten. An jedem Knoten wird eine Entweder-Oder-Entscheidung auf Grund eines Attributwertes getroffen²³ (vgl. Abbildung 12). Am darauf folgenden Knoten wird die nächste Variable abgefragt. Dies wird solange fortgesetzt, bis keine weiteren Verzweigungen mehr vorhanden sind. Am letzten Blatt jedes Astes lässt sich das entsprechende Ergebnis der Klassifikation ablesen. Bei der Anwendung des Entscheidungsbaumes gelangt man auf einem eindeutigen Pfad von der Wurzel zu einem der Blätter. Das Objekt wird zur Klasse des erreichten Blattes zugeordnet. Ist das Input-Attribut numerisch, entscheidet die Überprüfung in einem Knoten darüber, ob der Attributwert größer oder kleiner als eine be-

²³ Die überwiegende Anzahl der Entscheidungsbaum-Klassifikatoren führen binäre Splits durch, manche bilden aber auch mehr Partitionen.

stimmte Konstante ist, wodurch sich zwei Verzweigungen ergeben. Ein Beispiel für einen Entscheidungsbaum aus dem empirischen Teil der vorliegenden Arbeit ist in Abbildung 12 dargestellt.

Ein Entscheidungsbaum wird anhand der Trainingsmenge mit einem Top-Down-Ansatz konstruiert. Die Menge der Objekte wird in disjunkte Teilmengen zerlegt. Das heißt, dass jedes Objekt in nur genau einer Teilmenge enthalten sein darf. Das Kriterium, nach dem die Aufteilung erfolgt, wird aus den Attributen bestimmt. Dabei wird das Attribut ausgewählt, das die Objekte untereinander möglichst am stärksten separiert, aber gleichzeitig einen möglichst einfachen Baum mit wenigen Knoten erzeugt.

Als alternative Darstellungsform zu Entscheidungsbäumen existieren Klassifikations- bzw. Entscheidungsregeln. Ein Entscheidungsbaum kann in der Regel einfach in Klassifikationsregeln der Form „If ..., then ...“ übersetzt werden. Eine exakte Übersetzung kann zu umfangreichen Regeln führen, die aber gekürzt werden können (WITTEN UND FRANK, 2001, S. 105ff). Einer bereits existierenden Regelmenge können neue Regeln hinzugefügt werden, ohne die bereits vorhandenen Regeln zu stören, während die Erweiterung einer Baumstruktur den gesamten Baum kippen kann.

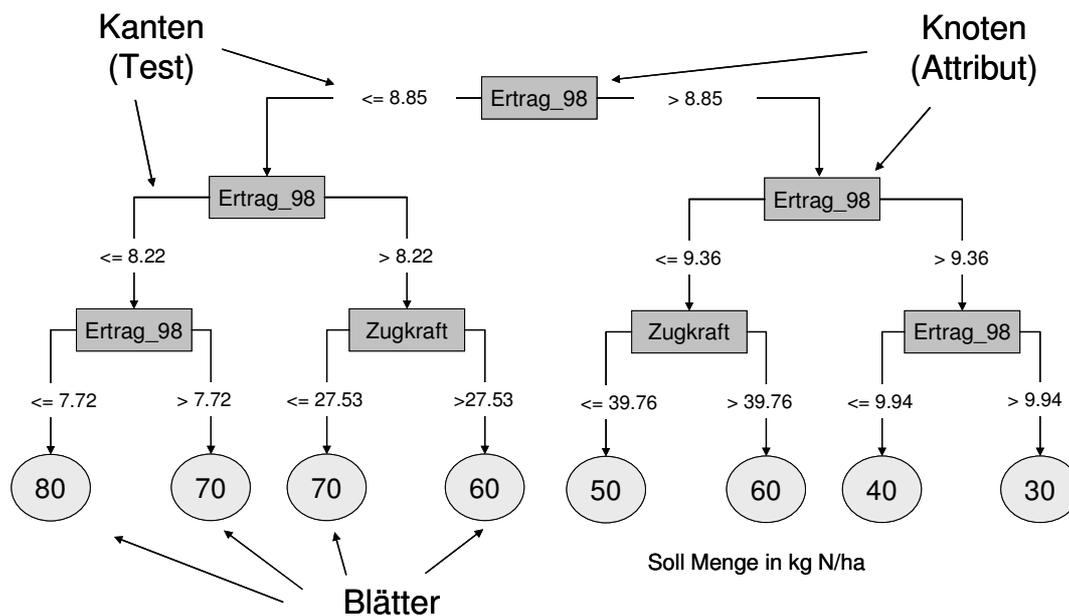


Abbildung 12: Komponenten eines Entscheidungsbaumes

Ein weiterer Spezialfall sind Bäume für numerische Vorhersagen. Hier liegt das Output-Attribut in numerischer Form vor. Dabei enthalten die Blattknoten des Baumes einen numerischen Wert, der den Durchschnitt aller Trainingsmengenwerte darstellt, auf die das Blatt oder die Regel anzuwenden sind. Diese Bäume werden auch als *Regressionsbäume* bezeichnet (WITTEN UND FRANK, 2001, S. 74ff).

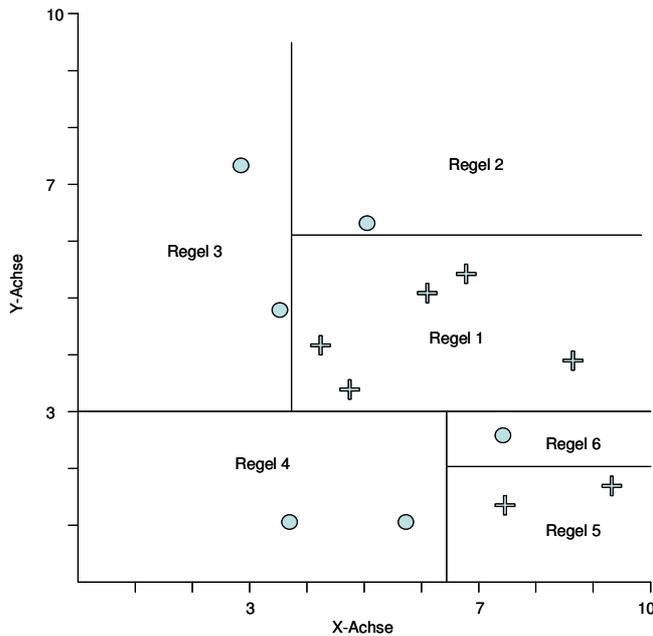


Abbildung 13: Unterteilung des Merkmalsraumes mit einem Entscheidungsbaum (verändert nach PYLE, 1999, S. 105)

In der Anwendung eines Entscheidungsbaumverfahrens wird der Merkmalsraum, der von den Trainingsinstanzen aufgespannt wird, „zerschnitten“. Jede Komponente einer Regel zerschneidet damit den Merkmalsraum in zwei Teile. Dieser Effekt ist für einen zweidimensionalen Merkmalsraum mit zwei (Ziel-) Klassen in Abbildung 13 dargestellt.

In dieser Abbildung stehen die Kreuze für die Zugehörigkeit zu einer Klasse, die kleinen Kreise für eine andere Klasse. Die Kreuze bzw. Kreise stehen jeweils für eine Instanz. Auf der x- bzw. y-Achse sind die Merkmalsausprägungen bezogen auf zwei Attribute abgetragen, die annahmegemäß die Zielgröße erklären. Ein Entscheidungsbaumverfahren unterteilt den Merkmalsraum mit geeigneten Regeln so, dass alle Instanzen richtig klassifiziert (vorhergesagt) werden. Innerhalb des Bereiches einer Regel liegen im Idealfall nur noch die Instanzen einer Klasse. Die Regel 5 prognostiziert beispielsweise die Klasse „Kreuz“, falls der y-Wert größer 6.5 und der y-Wert kleiner 3 ist. Dabei verlaufen die Trennlinien in den meisten Verfahren linear²⁴.

2.6.2 Konstruktion eines Entscheidungsbaumes

Ein Entscheidungsbaum wird mit einem Basis-Algorithmus konstruiert, der anfangs alle Trainingsdatensätze der Wurzel zuordnet. Aus den Datensätzen wird ein Attribut mit Hilfe einer Splitstrategie ausgewählt. Die Trainingsdatensätze werden dann unter Nutzung des Splitattributs partitioniert. Dieses Verfahren wird nun rekursiv für die Partitionen fortgesetzt. Wenn keine weiteren Splitattribute vorhanden sind, oder alle Trainingsdatensätze

²⁴ Es existieren Verfahren, die auch komplexere Zusammenhänge darstellen können. Dies führt jedoch wiederum zu komplexen Entscheidungsbäumen, die schwer interpretierbar sind.

eines Knotens zur selben Klasse gehören, wird dieser Algorithmus abgebrochen. Ein Pseudocode dieses Algorithmus' (in Anlehnung an ESTER UND SANDER, 2000, S. 127) lautet:

```

Entscheidungsbaum_Konstruktion (Trainingsmenge T, Float min conf)
If mindestens min-conf der Objekte aus T in Klasse c Then Return
Else
  For each Attribut A do
    For each möglicher Split von A do
      Bewerte die Qualität der Partitionierung, die durch den Split entstehen würde;
    Führe den besten aller dieser Splits durch;

```

Der entscheidende Punkt in diesem Algorithmus ist die Bewertung der Qualität der Partitionierung (Splits). Dazu gibt es verschiedene Auswahlmaße, mit Hilfe derer die Aufteilungen bewertet werden können (vgl. SÄUBERLICH, 2000, S. 88ff). Im Wesentlichen basieren die Kennzahlen darauf, die Änderung der Entropie²⁵ beim Übergang zur bedingten Verteilung bezüglich eines Splits zu messen. Je größer diese Entropieminderung ausfällt, umso reiner sind die neu gebildeten Teilmengen in Bezug auf die Klassenstruktur. Es wird diejenige Unterteilung präferiert, die dieses Maß maximiert. Für eine ausführliche Darstellung dazu siehe HASTIE ET AL. (2001), WITTEN UND FRANK (2000, S. 95ff) oder ESTER UND SANDER (2000, S. 107ff).

Die Basis der Algorithmen ist daher ein so genanntes *Impurity-Auswahlmaß*. Diese Maßzahl umfasst den Grad, wie stark die Ausprägungen in einem Ast auf eine Ausprägung konzentriert sind. Ein „reiner“ Ast ist ein Ast, bei dem alle Instanzen in eine einzelne Output-Klasse fallen. Ein Ast mit einer maximalen Unreinheit würde dieselbe Anzahl von Ausprägungen in jeder der Output-Klassen besitzen. Ein enger Zusammenhang besteht zu dem Begriff des Klassifikationsfehlers, denn ein „unreiner“ Ast bedeutet immer, dass noch Klassifikationsfehler vorliegen. *Impurity* kann allerdings sehr verschiedenartig definiert werden. In der vorliegenden Arbeit wird die Gini-Maßzahl verwendet:

$$Gini = 1 - \sum_i P(t)_i^2$$

$P(t)_i$ entspricht dabei dem Anteil der Instanzen in der Abzweigung t mit der Klasse i . Wenn zwei Abzweigungen in zwei unterschiedliche Verteilungen in Bezug auf die Output-Klassen resultieren, z. B. $(1,0,0)$ und $(1/3, 1/3, 1/3)$, wird also dasjenige mit der größeren Konzentration auf eine einzelne Klasse (die erste) verwendet²⁶. Die Gini-Maßzahl liegt zwischen 0 und 1, obwohl der maximale Wert eine Funktion aus der Anzahl von Output-Kategorien ist.

²⁵ Als Entropie wird in diesem Fall die minimale Anzahl von Bits zum Codieren der Nachricht definiert, mit der man die Klasse eines zufälligen Trainingsobjekts mitteilen möchte (vgl. WITTEN UND EIBE, 2000, S. 99f).

²⁶ Im vorliegenden Beispiel hat die Verteilung $(1, 0, 0)$ eine Gini-Maßzahl von 0, die Verteilung $(1/3, 1/3, 1/3)$ dagegen einen Gini-Index von 0.66.

2.6.3 Pruning-Techniken und spezialisierte Algorithmen

Zur Bestimmung der Endknoten werden deshalb Abbruchkriterien benutzt, die dafür sorgen, dass der zu generierende Baum nicht zu groß wird und damit kein Überlernen auftritt. Allgemein werden solche Verfahren als Pruning-Techniken bezeichnet. Es werden zwei Klassen von Pruning-Techniken unterschieden. So genannte Pre-Pruning-Techniken, die beim Aufbau des Baumes durch Anwendung bestimmter Abbruchkriterien das Anwachsen des Baumes verhindern, sowie so genannte Post-Pruning-Techniken, die erst nach dem Aufbau des Baumes durch „Zurückschneiden“ von Ästen den Baum verkleinern (ESTER UND SANDER, 2000, S. 133f; BERRY UND LINOFF, 2001, S. 117f). Die folgenden Abbruchkriterien werden in den verwendeten Algorithmen dieser Arbeit im Pre-Pruning verwendet:

- Der Ast ist rein (alle Instanzen in diesem Ast haben denselben Zielwert).
- Alle Datensätze eines Knotens besitzen die gleichen Input-Merkmalausprägungen.
- Der Baum hat eine einstellbare Anzahl von Ebenen erreicht (*maximum tree depth*).
- Die Zahl der Datensätze in einem Knoten ist kleiner als ein einstellbarer Schrankenwert (*minimum node size*).
- Der beste Split für den Ast erreicht eine Verringerung der Unreinheit um weniger als ein einstellbarer Wert (*minimum change in impurity*).

Als Post-Pruning-Technik wird Cost-Complexity-Pruning (ECP) eingesetzt (vgl. SUTTON, 2005, S. 313f). Durch die Pruning-Maßnahmen wird damit gleichzeitig die Transparenz der Entscheidungsbäume erhöht: Große Entscheidungsbäume sind in der Regel schlecht darstellbar und dadurch wird auch der Zugriff zum dargestellten Wissen erschwert (WEISS UND INDURKHYA, 1998, S. 136ff). Um übersichtliche Entscheidungsbäume zu erreichen, werden in dieser Arbeit zusätzlich die Anzahl der Verzweigungen (*maximum tree depth*) auf drei Ebenen beschränkt.

Allerdings wird durch diese Einschränkung ein inhärentes Problem von Entscheidungsbaumverfahren verstärkt. Dadurch, dass diese Verfahren nur unstetige Ergebnisse produzieren können, arbeiten diese Verfahren nicht optimal, wenn die zugrunde liegende Funktion stetig bzw. glatt ist. Ein Entscheidungsbaum bildet diesen Sachverhalt nur mit entsprechenden Einbußen in der Performance ab (HASTIE ET AL., 2001, „Lack of Smoothness“). Bei kleineren Entscheidungsbäumen sind die einzelnen Splits umso gröber und liefern daher auch schlechtere Abbildungen dieser Funktionen.

In den letzten beiden Abschnitten wurden damit zwei sehr unterschiedliche Werkzeuge des Data Mining vorgestellt. Diese Unterschiede äußern sich auch in der folgenden Phase des WED-Prozesses, in dem die Erkenntnisse der Modellierungsphase herausgearbeitet werden.

2.7 Verifikation, Evaluation und Erkenntnisgewinn

Das Ergebnis eines Prozesses der Wissensentdeckung in Datenbanken sind allgemein definiert *Muster*. Dabei können diese *Muster* in sehr unterschiedlicher Form vorliegen bzw. beschrieben werden (BRACHMAN UND ANAND, 1996, S. 47f). Dies gilt auch für die beiden vorgestellten Werkzeuge. Während neuronale Netze eine „Black Box“ liefern, deren Interna nicht interpretierbar sind, erlauben Entscheidungsbäume nachvollziehbare Beschreibungen. Neuronale Netze liefern somit keine Darstellung des produzierten Wissens. Das Wissen von (übersichtlichen) Entscheidungsbäumen dagegen ist direkt verfügbar. Durch diese Unterschiede müssen unterschiedliche Werkzeuge herangezogen werden, um die jeweiligen Ergebnisse bewerten zu können.

Generell sind die erzeugten Modelle in der Theorie nicht verifizierbar (KRATZER, 1993, S. 89). Zur Verifikation der Modelle müssen diese in der Realität überprüft werden, was mit Schwierigkeiten verbunden ist. Insbesondere können Veränderungen in der Anwendungsumgebung dazu führen, dass die Modelle in der Praxis nicht die Erwartungen erfüllen.

Zur Evaluation bzw. Validierung existieren Verfahren, mit denen die Güte der Modelle anhand der Trainingsdatenmenge beurteilt werden kann.²⁷ Zentrale Aufgabe der Evaluation ist die „Schätzung der Vorhersagekraft“ (ESTER UND SANDER, 2000, S. 6), d. h. es muss **geschätzt** werden, wie gut die in den Trainingsdaten gefundenen Muster sich auf zukünftige Daten verallgemeinern lassen.

2.7.1 Leistungsvorhersage bei begrenztem Datenumfang

Falls es eine ausreichend große Datenmenge gibt, ist die Vorgehensweise relativ problemlos: Es muss lediglich eine genügend große Datenmenge ausgewählt werden, mit der das Modell trainiert wird. An der restlichen unabhängigen Datenmenge wird dieses Modell überprüft (WEISS UND INDURKHYA, 1998, S. 36f).

Bei begrenzten, wenn nicht sogar knappen Datenmengen ist die Leistungsvorhersage schwierig. Wegen der Gefahr des Überlernens verbietet es sich, dieselben Instanzen für das Training und für die Validierung zu benutzen. Für die Leistungsvorhersage bei begrenztem Datenumfang wird der Datenbestand deshalb in drei Gruppen eingeteilt: Ein Teil der Datenmenge wird dazu verwendet, die Netzgewichte zu bestimmen. Dieser Teil der Daten wird im Allgemeinen als *Trainingsample* bezeichnet. Das *Testsample* wird einerseits dazu verwendet, die Trainingsphase zu dem Zeitpunkt zu terminieren, an dem die Gefahr von Überlernen zu groß wird. Gleichzeitig wird es beispielsweise bei neuronalen Netzwerken

²⁷ Neben den hier vorgestellten Methoden können auch „klassische“ statistische Methoden, wie beispielsweise die Residuenanalyse eingesetzt werden. Damit kann überprüft werden, ob das Modell korrekt spezifiziert, d. h., ob die richtigen exogenen Variablen in der korrekten Form in das Modell miteinbezogen wurden (WIDMAN, 2000, S.110ff).

zur Optimierung der Netzstruktur verwendet (vgl. Kap. 2.5.3). Die Fehlerrate an diesem *Testsample* liefert nicht notwendigerweise Aufschlüsse für die Fehlerrate bei neuen Daten. Zur Optimierung des Data-Mining-Prozesses ist es jedoch nützlich, sie zu kennen (WITTEN UND FRANK, 2001, S. 127ff).

Der dritte Teil der Daten (*Validierungssample*) wird dazu verwendet, die Maße für die Leistungsvorhersage zu bestimmen. Datensätze aus dieser Menge dürfen also weder für das Training noch für das Testsample benutzt werden.

Zur Validierung der erzeugten Modelle wird in der Regel eine Kreuzvalidierung empfohlen (HASTIE ET AL., 2001). Bei einer fünffachen Kreuzvalidierung dient ein Fünftel der Instanzen als Validierungssample, der Rest der Instanzen wird für das Trainings- und Testsample verwendet und daraus ein Modell erzeugt. Dieser Vorgang wird fünfmalig wiederholt durchgeführt, so dass sich jede Instanz einmal in einem Validierungssample befunden hat. In jedem dieser fünf Durchgänge werden Maßzahlen errechnet, die die Leistungsfähigkeit der Prognose an dem Validierungssample beurteilen lassen. Der Durchschnittswert dieser Maßzahlen leistet einen guten Schätzer für die Güte der Prognosen. Das Verfahren der Kreuzvalidierung ist in Abbildung 14 dargestellt.

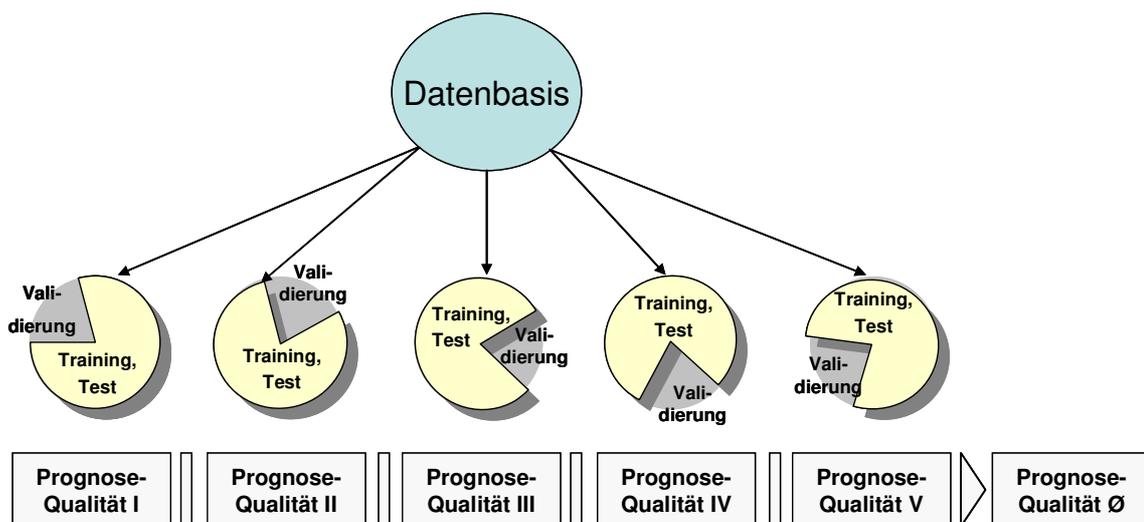


Abbildung 14: Vorgehensweise einer fünffachen Kreuzvalidierung

Innerhalb der einzelnen Wiederholungen der Validierung werden also mit den Validierungsinstanzen die vorhergesagten Werte (p_1, p_2, \dots, p_n) mit den tatsächlichen Werten (a_1, a_2, \dots, a_n) verglichen. Aus dem Vergleich dieser Werte lassen sich Leistungsmaße für die numerische Vorhersage definieren. In Tabelle 1 sind einige dieser Maßzahlen abgetragen. Insbesondere wurden hier die Maßzahlen aufgeführt, die in betrachteten Publikationen (vgl. Kap. 3.2) verwendet werden. Eine Übersicht ist beispielsweise in WITTEN UND FRANK (2001, S. 158) zu finden. Eine gewisse Ausnahme stellt die Maßzahl *Estimated Accuracy* dar. Diese Maßzahl wird in der verwendeten Software zur Bestimmung des Fehlers am

(nicht unabhängigen) Test-Sample verwendet und wird dementsprechend auch in dieser Arbeit (vgl. Kap. 4.5-4.7) eingesetzt. Weiterhin werden in dieser Arbeit selbsterklärende Maßzahlen, wie beispielsweise der maximale Fehler bzw. die Standardabweichung des Fehlers, verwendet (vgl. Kap. 4.7.2).

Tabelle 1: Leistungsmaße für die numerische Vorhersage

Mittlerer absoluter Fehler (Mean Absolute Error, MAE)	$\frac{ p_1 - a_1 + \dots + p_n - a_n }{n}$
Korrelationskoeffizient (Coefficient of Determination, R^2)	$\frac{S_{PA}}{S_P S_A}, \text{ wobei}$ $S_{PA} = \frac{\sum_i (p_i - \bar{p})(a_i - \bar{a})}{n-1}$ $S_P = \frac{\sum_i (p_i - \bar{p})^2}{n-1}, S_A = \frac{\sum_i (a_i - \bar{a})^2}{n-1}$
Mittlerer quadratischer Fehler (Mean Square Error, MSE)	$\frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{n}$
Standard-Fehler der Vorhersage (Standard Error of Prediction, SEP)	$\sqrt{\frac{(p_1 - a_1)^2 + \dots + (p_n - a_n)^2}{n}}$
Mittlere prozentuale Abweichung (MPA)	$\sqrt{\frac{\sum_{i=1}^n \left(\frac{(p_i * 100)}{a_i} - 100 \right)^2}{n}}$
Relativer quadratischer Fehler (RMS Error)	$\sqrt{\frac{\sum_{i=1}^n \left(\frac{(p_i - a_i)}{a_i} \right)^2}{n}}$
Estimated Accuracy (interne Maßzahl von Clementine®)	$\sum_{i=1}^n \frac{1 - a_i - p_i }{\max(a) - \min(a)} / n$
Mit:	
p_1, p_2, \dots, p_n : vorhergesagte Werte:	
a_1, a_2, \dots, a_n : tatsächliche Werte	
n : Anzahl der Instanzen	

Das gebräuchlichste Maß ist der mittlere quadratische Fehler, während mit dem mittleren absoluten Fehler eine bessere Vorstellung verbunden ist. Der mittlere absolute Fehler weist unabhängig von der Richtung, auf die durchschnittliche Fehlergröße hin. Mit dem mittleren quadratischen Fehler wird die Wirkung von Ausreißern verstärkt. Es existiert eine gewisse Schwierigkeit, ein geeignetes Leistungsmaß für die numerische Vorhersage zu auswäh-

len, jedoch stellen WITTEN UND FRANK (2001, S. 160) fest: „*Glücklicherweise stellt es sich heraus, dass in den meisten praktischen Situationen die beste numerische Vorhersagemethode immer die beste ist, egal welches Fehlermaß verwendet wird.*“

Die Kreuzvalidierungstechnik ermöglicht es, verschiedene Modellierungstechniken oder verschiedene Formen der Datenaufbereitung zu vergleichen. Die Interpretation der absoluten Zahlen erschließt sich allerdings erst im Kontext, also beispielsweise über vergleichbare Studien.

Die Aufteilung der Input-Instanzen in die drei Gruppen sollte im Allgemeinen zufällig erfolgen, gleichzeitig sollte aber beachtet werden, dass die Datenrepräsentation in allen drei Teilmengen gleich ist. Selbst mit Unterstützung aufwändiger Algorithmen ist diese Aufgabe allerdings bei multidimensionalen und inhomogenen Daten praktisch nicht durchführbar (vgl. Kap. 2.4.2).

Nachdem die Güte der Prognose ermittelt wurde, können die Validierungsinstanzen wieder in die Trainings- bzw. Testdaten aufgenommen werden, um ein besseres Modell zur Prognose zu schaffen. Damit wird eine größere Datenmenge geschaffen, und so das Risiko vermindert, dass wichtige Instanzen des Merkmalsraumes in der Erstellung des Modells fehlen. Insbesondere beim Einsatz in der Praxis ist diese Vorgehensweise zu empfehlen. Dasselbe gilt für die Testdaten. Falls sich eine „optimale“ Netzform bestimmen lässt und die Effekte des Überlernens bekannt sind, kann unter Umständen mit vordefinierten Parametern gearbeitet werden. Gleichzeitig bedeutet dies, dass bei begrenztem Datenumfang die Prognoseleistung innerhalb der Kreuzvalidierung in der Regel unterschätzt wird.

Mit der Kreuzvalidierung und den vorgestellten Kennzahlen ist es also möglich, die Prognoseleistung zu beurteilen. Dennoch ist es für einen erfolgreichen WED-Prozess im Sinne der Definition nicht ausreichend, lediglich eine gute Prognose zu erstellen. Dazu müssen weitere Punkte erfüllt sein, die im folgenden Abschnitt beschrieben werden.

2.7.2 Erkenntnisgewinn

Das Ziel eines WED-Prozesses ist es allgemein, Zusammenhänge bzw. Muster in Daten zu finden. Das Ergebnis sollte dabei zentralen Anforderungen genügen (vgl. Kap. 2.1). Diese Anforderungen lassen sich nach NAKHAEIZADEH (1998, S. 22) mit den Begriffen *Neuigkeit*, *Nützlichkeit* und *Verständlichkeit* charakterisieren.

Die *Neuigkeit* eines Ergebnisses ist nur bedingt automatisch messbar. In der Regel muss dazu Expertenwissen herangezogen werden. In wissenschaftlichen Anwendungen bietet sich dazu der Vergleich zu (Fach-)Publikationen an.

Die *Nützlichkeit* eines Ergebnisses ergibt sich durch die Umsetzungsmöglichkeiten. Falls die Anwendung des Ergebnisses in der Praxis zu einer Verbesserung beiträgt, ist diese An-

forderung erfüllt.

Die *Verständlichkeit* hängt hingegen von der Anwendung und den beteiligten Personen ab. Daher sollte im Rahmen des WED-Prozesses darauf Wert gelegt werden, die Ergebnisse mit geeigneten Werkzeugen graphisch zu veranschaulichen. Als Muster bzw. Output eines WED-Prozesses wird daher selbst ein Diagramm angesehen, aus dem Zusammenhänge eines Modells hervorgehen (vgl. BRACHMAN UND ANAND, 1996, S. 47f). Insbesondere gilt dies für die numerische Vorhersage, die in der vorliegenden Arbeit eine zentrale Rolle spielt:

„Die Erfahrung hat gezeigt, dass in vielen Anwendungen des maschinellen Lernens für das Data Mining die expliziten Wissensstrukturen, die strukturierten Beschreibungen, mindestens ebenso wichtig und nicht selten wichtiger sind als die Fähigkeit, eine gute Leistung für neue Beispiele zu erbringen“ (WITTEN UND FRANK, 2001, S. 8).

Eine gute Beschreibung liefert auch gleichzeitig Erklärungen und somit Wissen aus den Daten. Ein Entscheidungsbaum kann beispielsweise direkt von Experten überprüft werden und damit direkt zu Wissen werden. Es existieren jedoch auch einige Interpretationshilfen, mit denen sich die trainierten Netzwerk-Modelle erschließen lassen. Hierzu zählen insbesondere

- Sensitivitätsanalysen (Überprüfung der Bedeutung einzelner Attribute für die Vorhersage, vgl. Kap. 2.5.3, *Relative Importance*),
- Analysen der Topologie (Einschätzung der Gefahr für Überlernen, vgl. Kap. 2.5.3), sowie
- Visualisierungen (z. B. über *Response Surfaces*, vgl. Kap. 4.4.4).

Eine entscheidende Rolle bei der Interpretation des Datenmaterials und der Entdeckung von Mustern spielt in allen Abschnitten des WED-Prozesses die Visualisierung (vgl. SCHOMMER, 2003). Die Aufgabe der Visualisierung ist, geeignete Repräsentationen von Daten und Informationen zu finden, die zu einer explorativen oder konfirmativen Auswertung führen. Die Visualisierung kann die Analyse, sowie das Verständnis und die Kommunikation von Modellen, Konzepten und Daten erleichtern.

Damit wurden die wichtigsten Konzepte eines WED-Prozesses vorgestellt. In den folgenden Abschnitten werden diese Konzepte an einem praktischen Problem angewendet. Wie in einem typischen WED-Prozess, beginnt dieser Teil mit einer Analyse des Anwendungsgebietes, aus dem heraus Ziele definiert werden.

3 Forschungsstand der Bereiche Stickstoffdüngung, Precision Farming und Data Mining

Zur Darstellung des Forschungsstandes müssen drei wissenschaftliche Bereiche herangezogen werden, die das Fundament für diese Arbeit bilden: Stickstoffdüngung, Precision Farming, Data Mining und vor allem die Überschneidungen dieser drei Gebiete²⁸. Mit dem vorausgegangenen Kapitel wurde der Bereich „Data Mining“ bereits vorgestellt. In diesem Kapitel werden nun die anderen Bereiche herangezogen.

Das Kapitel 3 besitzt eine „Doppelfunktion“. Einerseits beschreibt es den Stand der Forschung. Gleichzeitig stellt dieses Kapitel den ersten Teil eines WED-Prozesses dar. Diese Phase umfasst eine Analyse des Anwendungsgebietes. Ausgehend von dieser Basis wird das Ziel der empirischen Analyse spezifiziert und die Rahmenbedingungen erläutert.

3.1 Schnittmenge I: Stickstoffdüngung und Precision Farming

Zuerst wird daher ein Blick auf den Bereich Stickstoffdüngung bei Winterweizen (Kap. 3.1.1) geworfen und darauf aufbauend die Schnittmenge der Bereiche Precision Farming und Stickstoffdüngung (Kap. 3.1.2) betrachtet. In diesem Abschnitt werden insbesondere verschiedene Ansätze zur teilflächenspezifischen Optimierung der Stickstoffdüngung vorgestellt. Unter anderem wird gezeigt, wo die Schwachstellen der bisherigen Ansätze liegen und wie diese möglicherweise beseitigt werden können. Die kleinräumige Stickstoffdüngung basiert zu einem großen Teil auf Daten, die mit Sensoren erfasst werden. Im folgenden Abschnitt (Kap. 3.1.3) werden deshalb Charakteristiken dieser Sensoren beschrieben. Gleichzeitig stellt dieser Abschnitt einen Vorgriff auf den empirischen Teil dar, in dem diese Sensor-Daten zur Modellierung verwendet werden.

Das Gros der Ansätze (bzw. Algorithmen) zur teilflächenspezifischen Stickstoff-Ausbringung basiert auf dem Transfer von Wissen aus der „traditionellen“ pflanzenbaulichen Forschung zur Optimierung der mineralischen Stickstoffdüngung. Vor diesem Hintergrund werden die Grundprinzipien und Methoden zur Optimierung der schlageinheitlichen Stickstoffdüngung kurz vorgestellt.

Gleichzeitig werden die schlageinheitlichen Düngestrategien oftmals als „Referenzszenarien“ gewählt: Zur Untersuchung der ökonomischen Vorteilhaftigkeit von Precision-Farming-Ansätzen werden diese Düngestrategien als Vergleichsvarianten eingesetzt.

²⁸ Die Überschneidung aus den Teilgebieten Data Mining und N-Düngung wird davon ausgenommen, da hier praktisch keine Überschneidungen existieren.

3.1.1 Stickstoffdüngung bei Winterweizen

Mit der Bestimmung der optimalen speziellen Intensität eines Betriebsmittels, also beispielsweise der Stickstoffmenge, die zum ökonomisch optimalen Naturalertrag führt, beschäftigt sich die landwirtschaftliche Produktionstheorie (STEINHAUSER ET AL., 1982, S. 73ff). Zur Ermittlung dieser Intensitäten werden so genannte *Produktionsfunktionen* bzw. *Ertragsfunktionen* (vgl. WAGNER, 1999, S. 39ff) geschätzt. Im Rahmen dieser Arbeit sind insbesondere die Produktionsbeziehungen zwischen dem Input „Stickstoff“ und dem Output „Korn- oder Naturalertrag“ von Interesse. In der englischsprachigen Literatur wird diese funktionelle Beziehung als *N-Response* bezeichnet.

Zur Schätzung der Produktionsfunktionen werden üblicherweise pflanzenbauliche Exaktversuche (hier: Stickstoffsteigerungsversuche) durchgeführt. Diese Methode ist in vielen Ländern verbreitet und wird von diversen Institutionen angewendet (vgl. SWINTON UND LIU, 2002, S. 1). Die Ermittlung der optimalen speziellen Intensität (N_{opt}) erfolgt in der Regel ex post aus geschätzten Produktionsfunktionen.

Um aus Produktionsfunktionen Düngestrategien abzuleiten, muss berücksichtigt werden, dass es sehr viele Faktoren gibt, die die Ertragsbildung von Winterweizen beeinflussen. Eine ausführliche Darstellung der Einflussfaktoren auf den Ertrag ist beispielsweise bei REINER ET AL. (1992) zu finden. Dabei sind nicht alle Faktoren kontrollier- bzw. steuerbar. Insbesondere Witterungs- und Standortverhältnisse gehören zu dieser Gruppe. Ein Beispiel für die Komplexität der Zusammenhänge innerhalb dieser Faktoren stellen bodenbürtige Stickstoffquellen dar. Hier ist ein großer Vorrat an organischem Stickstoff gespeichert, der bei entsprechenden Klimaverhältnissen pflanzenverfügbar wird (vgl. GUTSER, 2005, S. 5ff). Der ausgebrachte mineralische Stickstoff ist damit nicht die einzige Versorgungsmöglichkeit dieses Nährstoffes für die Pflanzen. Eine Prognose der Stickstoffnachlieferung aus dem Boden ist an das Klima gebunden und gestaltet sich daher schwierig. Dennoch ist die Nutzung dieses N-Angebotes des Bodens ökonomisch von großer Bedeutung (vgl. AUFHAMMER, 1998, S. 283f).

Die einzelnen Faktoren stehen also in vielfältigen dynamischen Interaktionen zueinander, die schwer zu quantifizieren sind. BULLOCK UND BULLOCK (2000, S. 96) fassen die Situation wie folgt zusammen: *„Despite decades of study, major disagreements and ignorance still exist about how crop yield responds to managed inputs and non-managed factors of production.“*

Alle Einflussfaktoren führen potenziell zu unterschiedlichen Verläufen der Produktionsfunktionen und bedingen variierende optimale spezielle Intensitäten (vgl. OLFS ET AL. 2005, S. 415f). Ein Exaktversuch ermöglicht damit (isoliert betrachtet) lediglich ex post Schätzungen der speziellen Intensität für eine spezifische Situation. Diese Situation wird über die Dokumentation von Bodenverhältnissen, Witterungsverhältnissen und produktionstechni-

schen Rahmenbedingungen erfasst und dokumentiert. Zur Ableitung von Düngestrategien werden diese Versuche mehrere Jahre hintereinander wiederholt und somit die spezifischen Jahreseinflüsse ausgeglichen. Der Vergleich von Studien aus verschiedenen Jahren wird durch ein sich dynamisch veränderndes Umfeld in Bezug auf züchterischen Fortschritt, verbesserten Produktionsmethoden und sonstigen Umwelteinflüssen erschwert.

Zur Modellierung dieser komplexen Zusammenhänge werden in der Regel die Ergebnisse von mehreren Exaktversuchen an unterschiedlichen Standorten und in unterschiedlichen Jahren kombiniert und in Modelle eingebaut (vgl. MAKOWSKI UND WALLACH, 2001, S. 197f). Dadurch können auch die Einflüsse von erklärenden Variablen auf N_{opt} untersucht werden.

Um auf spezifische Jahresverhältnisse besser reagieren zu können, wird die Stickstoffdüngung in Deutschland üblicherweise in drei Teilgaben aufgespaltet (vgl. REINER ET AL., 1992, S. 111; WENKEL ET AL., 2002, s.p.)²⁹. Die erste Applikation erfolgt als „Startgabe“ bei Vegetationsbeginn, die zweite als Nachdüngung“ bei EC 30-32 (nach TOTTMANN, 1987, Beginn des Schossens) sowie die dritte als „Ährengabe“ ab EC 49 (dto., Ährenschieben).

Es muss daher nicht nur die Gesamtmenge optimiert werden, sondern auch die Verteilung dieser Menge auf die einzelnen Teilgaben. Die Verteilungseffekte, werden in der Produktionstheorie oftmals nicht beachtet, obwohl die Bedeutung dieser Effekte relativ hoch einzuschätzen ist (STURM ET AL., 1994, S. 271). Anstelle dessen wird in der Regel lediglich die Gesamtdüngermenge optimiert, ohne auf die Verteilung einzugehen.

Ausgehend von dieser Komplexität und der hohen Anzahl von durchgeführten Versuchen, gestaltet es sich schwierig, einen stichhaltigen Überblick über die Ergebnisse der einzelnen Studien und die Ableitung von Düngestrategien für Winterweizen zu erstellen³⁰. Beispielsweise führen STURM ET AL. (1994, S. 273ff) und DIEPOLDER (1994) zahlreiche Düngestrategien an, die sich in ihrem Ansatz und in den Ableitungen deutlich unterscheiden und partiell widersprechen. Weiterhin muss angemerkt werden, dass die ökonomischen Effekte einer Variation in der Düngermenge wegen flach verlaufender Produktionsfunktionen teilweise schwach ausgeprägt sind (vgl. PANNELL, 2004, s.p.).

Praxisumsetzung

Aus den oben angeführten Gründen ergibt sich, dass für eine spezifische Praxissituation zur Bestimmung von N_{opt} keine eindeutigen Regeln, sondern sehr viele Methoden, Empfehlungen und Anhaltspunkte zur Verfügung stehen. Da unmöglich für das breite Spektrum an Standorten und Versuchszielen Exaktversuche durchgeführt werden können, muss für die

²⁹ Darin unterscheidet sich das Management beispielsweise von den nordamerikanischen Verhältnissen, hier wird Stickstoff oft in einer oder in maximal zwei Gaben für Winterweizen ausgebracht. Die Applikation erfolgt oftmals auch schon im Herbst.

³⁰ Ein Blick in die Literatur kann leicht den Eindruck erwecken, dass über die Stickstoffdüngung bereits alles geschrieben steht, nur vielleicht noch nicht von jedem.

Umsetzung in der Praxis „interpoliert“ werden. Die Umsetzungen in der Praxis beruhen daher im Allgemeinen zuerst auf dem Erfahrungswissen und auf der damit verbundenen saisonalen Einschätzung des Betriebsleiters in Verbindung mit Düngestrategien. Selbst moderne teilflächenspezifische Düngesysteme verzichten häufig nicht auf Bestandesbeobachtungen (vgl. WENKEL ET AL. 2002, s.p.).

Hinzu kommt, dass die Ansätze der Praxis nicht ausschließlich auf ökonomischen Überlegungen beruhen. Mit vielen Düngeempfehlungen wird eine Maximierung des Ertrages angestrebt. Gleichzeitig stellen viele Ansätze einen Kompromiss zwischen ökologischen und ökonomischen Gesichtspunkten dar (DIEPOLDER, 1994, S. 2f). Ein exaktes Zielsystem wird jedoch in der Regel nicht angegeben³¹.

Die Nutzung der Verfahren zur Abschätzung des Düngerbedarfs in der Praxis und in der Beratung ist dabei stark differenziert (vgl. STURM ET AL. 1994, S. 270ff). Eine umfangreiche Aufstellung von eingesetzten Düngestrategien liefert beispielsweise DIEPOLDER (1994, S. 12ff).

In der Regel liefert die Methode der *N-Bilanzierung* und die damit verbundene *Düngung nach Entzug* einen Rahmen für die Höhe der benötigten Gesamtdüngermenge (vgl. REINER ET AL. 1992, S. 106ff; AUFHAMMER, 1998, S. 272ff). Die Idee besteht darin, die Menge an Nährstoffen, die dem Boden mit dem Erntegut „entzogen“ wird, über die Düngung wieder zuzuführen. Dazu wird anhand der Schlaggeschichte und evtl. aktuellen Bestandsbegutachtungen ein erwarteter Kornertrag geschätzt und daraus ein proportional abhängiger Stickstoffbedarf abgeschätzt. Das bedeutet, es wird generell ein höherer Stickstoffbedarf für Standorte mit einem höheren Ertragspotenzial empfohlen. In diese Kalkulation finden weiterhin die N-Nachlieferung des Bodens, zu erwartende N-Verluste bei der Ausbringung sowie der erwartete N-Gehalt des Korns Eingang. Aufgrund der Ungenauigkeit der Schätzungen der Eingangsparameter liefert diese Methode allerdings nur sehr grobe Bereiche, innerhalb derer die optimale Düngermenge liegen sollte. Durch die Anwendung der Methode soll eine langfristig konstante Bodenfruchtbarkeit und Ertragsleistung sichergestellt werden (HEYLAND, 1996, S. 171). Es erfolgt daher mit dieser Methode keine Optimierung nach ökonomischen Gesichtspunkten (!). Diese Methode kann teilflächenspezifisch umgesetzt werden, wie im folgenden Kapitel dargestellt wird.

Weiterhin existieren Ansätze, mit denen die „Feinsteuerung“ der Düngung erfolgen kann. AUFHAMMER (1998, S. 284) unterscheidet dazu zwischen *analytischen Methoden*, *indirekten Methoden* und *Kombinationsmethoden*.

Die *analytischen Methoden* basieren auf physikalischen Messungen. Eine weit verbreitete

³¹ Das Chaos der Zielsysteme in der Praxis wird durch eine aktuelle Düngeempfehlung der BASF (<http://www.basf.de/file/1226762.pdf1>) sehr treffend wiedergespiegelt: Die Autoren empfehlen hierin eine Düngung speziell „für schwache Nerven“. Damit werden zu allem Überfluss noch psychologische Elemente in eine Düngeempfehlung gebracht.

Methode ist die N_{\min} -Methode (nach SCHARPF U. WEHRMANN, 1975) für die erste Stickstoffgabe. Dazu wird die pflanzenverfügbare Stickstoffmenge über Bodenbeprobungen ermittelt. Die Differenz aus dem Teil-Stickstoffbedarf und dem verfügbaren Stickstoff im Boden ergibt die Sollmenge für die Startgabe.

Weitere Beispiele sind der Nitratschnelltest (WOLLRING UND WEHRMANN, 1981) und Chlorophyllmessungen (WOLLRING, 1995). Diese Methoden erfassen indirekt den N-Gehalt in der Pflanze. Über entsprechende Versuchsserien kann der optimale N-Gehalt, bzw. die optimale N-Aufnahme bestimmt werden. Dadurch können aus Düngermengen für die zweite und dritte Stickstoffgabe abgeleitet werden. Wiederum erfolgt mit den analytischen Methoden keine Optimierung nach ökonomischen Gesichtspunkten. Diese Methoden streben mit einer optimalen Ernährung der Pflanzen einen Maximalertrag an.

Die *indirekten Methoden* basieren auf einer Beurteilung der (saisonalen) Bedingungen. In diese Kategorie fallen die bereits erwähnten Bestandesbeobachtungen bzw. Bestandesbonituren. Zur Unterstützung des Landwirts werden so genannte *Hilfstabellen* erstellt³². Zur Nachdüngung wird der Bedarf an Stickstoff überwiegend durch die Einschätzung des Betriebsleiters bestimmt. Durch die Anwendung analytischer Methoden kann er bei dieser Aufgabe unterstützt werden.

Die *Kombinationsmethoden* basieren auf Elementen aus den analytischen und indirekten Methoden. Dabei finden einerseits physikalische Messungen und andererseits Bestandesbonituren Eingang. In diesen Bereich fallen vor allem Prognose- bzw. Simulationsmodelle und Düngeberatungssysteme (vgl. z. B. Expert-N, ENGEL, 1991). Es wird davon ausgegangen, dass der Landwirt (Betriebsleiter) die Ergebnisse der analytischen Methoden auf seine Standorte übertragen kann, wie folgendes Zitat zeigt: „*Im allgemeinen weiß der Betriebsleiter – vor allem wenn er schon mehrere Jahre mit der N_{\min} Methode gearbeitet hat – auf welchem Schlag bzw. auch Teilstück eines Schlages etwas mehr oder weniger Düngung notwendig ist*“ (STURM ET AL., 1994, S. 274).

Bewertung der Düngesysteme

Der Vergleich unterschiedlicher Düngesysteme stellt eine schwierige Aufgabe dar. Das Problem einer objektiven ökonomischen Bewertung von Düngestrategien besteht im Wesentlichen darin, den Einfluss des Betriebsleiters, der über die Bestandesbonituren stark in die Düngesysteme einwirkt, zu erfassen und zu bewerten. Es kann nicht angenommen werden, dass diese Aufgabe von unterschiedlichen Betriebsleitern in einheitlicher Qualität durchgeführt wird. Es muss sehr kritisch hinterfragt werden, ob der Betriebsleiter in der

³² In diesen Tabellen sind Zu- bzw. Abschläge zu einer Düngermenge für schlagspezifische Verhältnisse (z. B. Auswinterung, Vorfrucht, etc.) abgetragen. Die Standort-Empfehlungen beschränken sich dabei auf Grobkategorien (beispielsweise auf „Leichte Böden“, „Lößboden“, „Tonböden“). Weiterhin werden für diese Grobkategorien große Spannbreiten angegeben, in denen sich die Düngung bewegen sollte. Damit ist die Übertragungsmöglichkeit dieses Wissens für eine teilflächenspezifische Düngung begrenzt.

Lage ist, die ökonomischen Effekte in der notwendigen Präzision abzuschätzen. Gleichzeitig stellt dieses angewendete Wissen zur Bonitur von Beständen ein idiosynkratisches Wissen dar, das also beispielsweise nicht in Regeln transformiert werden kann³³. Damit kann es nur begrenzt an andere Betriebsleiter weitergegeben werden und steht keinesfalls als Basis für eine automatisierbare Düngung zur Verfügung.

Der Vergleich von Düngesystemen wird dadurch erschwert, dass für viele Systeme kein eindeutiges Zielsystem zu ermitteln ist. Häufig wird unterstellt, dass die Systeme auf ökonomische Zielsysteme ausgerichtet wären. In der Regel beschränken sich Studien, die verschiedene Düngesysteme vergleichen, auf den Vergleich der erzielten Erträge (vgl. HEGE ET AL., 2002, S. 28ff), und auch in der Praxis ist die Orientierung am Maximalertrag weit verbreitet. Die Orientierung der Stickstoff-Düngerberatung am Ertragsmaximum war aus ökonomischen Gesichtspunkten so lange richtig, so lange der Quotient zwischen Faktorpreis (Stickstoff) und Produktpreis (Winterweizenpreis) sehr klein war. Ist dies der Fall, so besteht kaum ein Unterschied zwischen der Stickstoffmenge, die zum Ertragsmaximum führt, und derjenigen, die zur Erreichung des maximalen ökonomischen Ergebnisses notwendig ist. In der jüngeren Vergangenheit hat jedoch vor allem eine Veränderung der agrarpolitischen Rahmenbedingungen zu einer Absenkung der Produktpreise geführt. Zeitgleich hat sich der Faktorpreis für Stickstoff seit 1999/2000 nahezu verdoppelt. Beide Entwicklungen haben dazu geführt, dass sich die optimale spezielle Intensität der Stickstoffdüngung, also die ökonomische optimale Stickstoffdüngergabe, immer weiter vom Ertragsmaximum entfernt; der oben beschriebene Quotient aus Faktorpreis und Produktpreis wird immer größer.

Aufgrund dieser Entwicklung sollten Düngestrategien diesen Sachverhalt berücksichtigen und entsprechend angepasst werden, wenn sich Änderungen in diesem Quotienten ergeben. Diese Forderung muss demnach auch in der vorliegenden Arbeit berücksichtigt werden (vgl. Kap. 3.1.2 und Kap. 4.4.2).

3.1.2 Teilflächenspezifische Stickstoff-Applikation

Die Ansätze für eine teilflächenspezifische Düngung beruhen überwiegend auf den in Abschnitt 3.1.1 vorgestellten Methoden, obwohl die Umsetzung mit Problemen behaftet ist. Beispielsweise können die analytischen Methoden teilflächenspezifisch umgesetzt werden, indem die Messungen innerhalb eines Schlages in einem bestimmten Raster erfolgen. Darauf aufbauend kann ein teilflächenspezifisches Management implementiert werden. Die notwendige hohe Auflösung ist dabei nur mit hohem Aufwand möglich (vgl. EARL ET AL.,

³³ Vgl. ECKERT, 1998, S.26f: Es existiert ein so genanntes „Wissen neben der Sprache“, d. h. ein Wissen, das sich nicht ausdrücken lässt. In diesem Sinne fertigt jedes Individuum idiosynkratische Konstruktionen der Wirklichkeit an. Diese Konstruktionen sind von anderen Individuen ohne einer genauen Kenntnis des Kontextes nicht mehr nachzuvollziehen.

2003, S. 438f; HEGE ET AL., 2002, S. 34). Als Alternative zu den destruktiven Methoden gewinnen daher Sensoren an Bedeutung (vgl. Kap. 3.1.3). Mit zerstörungsfreien bzw. berührungslosen Sensoren kann eine hohe räumliche Auflösung kostengünstig realisiert werden (SWINTON UND JONES, 1998, S. 10).

In der derzeitigen Forschung und in der Praxis für teilflächenspezifische Stickstoff-Ausbringung werden überwiegend zwei Ansätze angewendet. Diese Ansätze können unmittelbar aus den schlageinheitlichen Ansätzen abgeleitet werden.

Der erste Ansatz besteht darin, anhand von Ertrags- oder Bodenkarten Managementzonen zu definieren. Innerhalb der einzelnen Zonen wird dann ein einheitliches Management angewendet. Der Ansatz wird in der Regel als *Mapping-Ansatz* bezeichnet. Die Basis für diese Ansätze ist in den Hilfstabellen (für betrachtete Bodenverhältnisse) sowie in der Düngung nach Entzug zu finden.

Der zweite Ansatz besteht darin, mit Hilfe von Sensoren während der Vegetationszeit³⁴ den aktuellen Stickstoffbedarf der Pflanzen abzuschätzen und darauf aufbauend die Düngermenge anzupassen. In einem in der Praxis bereits angewendeten Ansatz werden dazu reflexionsoptische Messungen durchgeführt, mit denen der aktuelle Ernährungszustand von Pflanzen erfasst werden kann (vgl. 3.1.3)³⁵. Dieser Ansatz basiert damit auf den analytischen Methoden und wird wie selbige nur zur zweiten und dritten N-Gabe eingesetzt. Dazu erfolgt eine Kalibrierung mit Hilfe des Nitratschnelltests. Die kleinräumigen Anpassungen erfolgen mit einem berührungslosen Sensor. Der Ansatz wird allgemein als *Sensor-Ansatz* bezeichnet. Er zeichnet sich dadurch aus, dass Online-Sensoren verwendet werden, die beispielsweise am Schlepper montiert sind und während der Applikation Signale liefern, die direkt in eine Düngermenge verrechnet werden. Diese Menge wird dann unmittelbar ausgebracht. Es existieren alternative Ansätze, die sich beispielsweise an Luftbildern bzw. Satellitenaufnahmen orientieren. Diese Verfahren haben im europäischen Raum jedoch bisher keine große praktische Bedeutung erlangt.

Die Kombination der beiden Ansätze, ein *Sensor-Ansatz mit Kartenüberlagerung*, ist bisher nur vereinzelt untersucht worden. Eine Möglichkeit besteht darin, die Sensorsteuerung an verschiedene Ertragspotenziale anzupassen (vgl. LINSEISEN, 2003, S. 117ff; MAIDL ET AL., 2004, s.p.). Die Ertragspotenziale werden dabei über Bestandsbegutachtungen und/oder Ertragskarten geschätzt.

In zahlreichen Studien wurde untersucht, ob sich mit teilflächenspezifischen Ansätzen ein höherer wirtschaftlicher Erfolg als mit schlageinheitlichen Verfahren erzielen lässt. Einen umfassenden Überblick über die Studien, in denen Precision-Farming-Ansätze mit einheit-

³⁴ In der englischsprachigen Literatur wird hierbei von *in-season* gesprochen.

³⁵ Es existieren auch andere Methoden um den aktuellen Aufwuchs teilflächenspezifisch zu beurteilen, wie beispielsweise das Crop-Meter (vgl. EHLERT ET AL., 2004, s.p.). Die Problematik der Kalibrierung ist jedoch ähnlich, bzw. verstärkt sich sogar.

lichen Ansätzen verglichen werden, geben beispielsweise GANDORFER (2005, S. 8ff) sowie SCHNEIDER (2005, S. 26f). In den Versuchen wird mehrheitlich entweder eine Sensor-Ansatz oder einen Mapping-Ansatz nach Ertragskarten verwendet.

Trotz der hohen Anzahl an Versuchen zur teilflächenspezifischen Stickstoffausbringung existieren wenige ansprechend publizierte Studien, die erstens die Fruchtart Winterweizen betreffen und die zweitens *ex-ante* definierte und reproduzierbare Entscheidungswege für die Bestimmung der N-Mengen angeben. Dies betrifft sowohl die Precision-Farming-Varianten als auch das zu vergleichende Referenzszenario der einheitlichen Düngung. Die Ursache liegt darin, dass lediglich wenige Versuchsansteller bereit sind, die menschlichen Eingriffsmöglichkeiten abzuschaffen.

Weitere Probleme sollen nur am Rande angesprochen werden: Es existieren bisher keine Standards für das Design von Versuchen zum Vergleich teilflächenspezifischer und schlag-einheitlicher Düngestrategien. Insbesondere die Interpretation der Versuchsergebnisse gestaltet sich schwierig (vgl. WEIGERT ET AL. 2004, S. 320ff). Die Ursache liegt zum einen darin, dass das ökonomische Ergebnis von Standortverhältnissen beeinflusst wird. So führt beispielsweise eine geringe räumliche Variabilität dazu, dass auch der mögliche Vorteil einer teilflächenspezifischen Düngung stark eingeschränkt ist. Weiterhin bestimmt sich die ökonomische Vorteilhaftigkeit aus dem Vergleich mit der schlageinheitlichen Variante. Diese Variante wird oft als „betriebsübliche“ Variante bezeichnet und kann sehr unterschiedlich bestimmt werden (vgl. Kap. 3.1.1). Ist diese Variante besonders gut, wird die ökonomische Differenz zwischen beiden Varianten in der Regel kleiner. Andererseits ist es ein leichtes, die Differenz besonders groß erscheinen zu lassen, wenn nämlich die Vergleichsvariante besonders schlecht gewählt wird. Da die Düngermengen der Teilgaben in der Regel nicht *ex ante* festgesetzt werden, sondern über subjektive Bestandesbonituren angepasst werden, wird die Intransparenz erhöht und eine ökonomische Interpretation dieser Vergleiche erschwert.

Insgesamt gesehen ergibt sich bei den Studien kein eindeutiges Bild. Die ökonomische Vorteilhaftigkeit konnte in vielen Fällen nicht belegt werden. Allerdings stellen die Ergebnisse dieser Studien lediglich eine Momentaufnahme dar. Aus den Ergebnissen sollten demnach auch keine Rückschlüsse auf das zugrunde liegende ökonomische Potenzial einer teilflächenspezifischen Düngung gezogen werden. Dass dieses Potenzial mit den verwendeten Ansätzen nicht ausgeschöpft wird, ergibt sich aus folgenden Punkten:

Es existieren in erster Linie spezifische Probleme sowohl für den Mapping-Ansatz als auch für den Sensor-Ansatz. Vereinfacht dargestellt wird mit dem Mapping-Ansatz basierend auf Ertragskarten in denjenigen Bereichen, in denen in vergangenen Jahren bei einheitlicher Düngung ein relativ hoher Ertrag erzielt wurde, eine hohe Gesamtmenge an Stickstoff appliziert. Damit werden in diesem Ansatz wesentliche Elemente aus der Methode Düngung

nach Entzug verwirklicht. Im Regelfall werden dazu drei Zonen eingeteilt, die sich am Kartenmaterial orientieren: eine Hohertragszone, eine Mittlerertragszone und eine Niedrigertragszone. Obwohl der Ansatz in der Praxis erfolgreich angewendet wird (vgl. LISSO, 2003, S. 1142), tauchen auch diametral abweichende Ergebnisse auf: GODWIN ET AL. (2003, S. 535) verglichen unterschiedliche Mapping-Ansätze in verschiedenen Jahren. Unter anderem stellten sie einen Mapping-Ansatz, der auf den eben beschriebenen Prinzipien basiert, einem Ansatz, bei dem genau umgekehrt verfahren wird, gegenüber. In diesem Ansatz wird auf Hohertragsbereichen relativ wenig Stickstoff appliziert und vice versa. Dieser „inverse“ Ansatz schnitt ökonomisch gesehen im Durchschnitt besser ab als der normale Mapping-Ansatz. LINK UND JASPER (2003, S. 359) schließen aus der Auswertung von 31 Stickstoff-Steigerungsversuchen, dass eine teilflächenspezifische Düngung, die nur auf Ertragsdaten basiert, ökonomisch nicht erfolgreich sein kann.

Weiterhin existiert das Problem der zeitlichen Instabilität von Ertragskarten (vgl. BLACKMORE ET AL. (2003, S. 458ff). Eine Umsetzung mit wenigen Managementzonen ist daher in der Regel zu grob, um die teilflächenspezifischen Optimierungsmöglichkeiten auszuschöpfen. BULLOCK UND BULLOCK (2000, S. 96ff) kritisieren in diesem Zusammenhang die gängige Praxis, sich im Mittelwert (Mittlerertragszonen) an die regionalen Düngeempfehlungen zu halten und zwischen der Hohertragszone und der Niedrigertragszone umzuverteilen. Durch dieses Vorgehen werden die Möglichkeiten einer teilflächenspezifischen Optimierung stark eingeschränkt. Außerdem wirkt sich bei diesem Ansatz negativ aus, dass die Feinsteuerung während der Saison, die bei einer schlageinheitlichen Implementierung der *Düngung nach Entzug* über Bestandesbonituren (Expertenwissen) erfolgt, nur unter unverhältnismäßig großem Aufwand übertragen werden kann. Es müsste für jede einzelne Management-Zone eine Bonitur erfolgen, was in der Praxis sowohl ein zeitliches als auch ein kostentechnisches Problem darstellt.

Ähnliche Kritikpunkte lassen sich für den Sensor-Ansatz finden, obwohl die technische Umsetzung des Sensor-Ansatzes ausgereift ist und auch in der Praxis eingesetzt wird³⁶. Da der Sensor-Ansatz viele Gemeinsamkeiten mit dem Nitratschnelltest besitzt, konnte Erfahrung bzw. Wissen in diesen Ansatz transferiert werden. Die ökonomische Vorteilhaftigkeit konnte jedoch bisher nicht eindeutig belegt werden. Erneut muss auf einen Zielkonflikt hingewiesen werden: Die analytischen Methoden, auf denen der Sensor-Ansatz basiert, wurden zur optimalen Ernährung der Pflanzen und somit zur Erreichung des Maximalertrages entwickelt.

Mit dem Sensor-Ansatz werden der Standorteinfluss und das N-Nachlieferungsvermögen des Bodens vollkommen ausgeblendet (vgl. LIEBLER, 2003, S. 150ff). Kritisiert wird insbe-

³⁶ Die Argumentation bezieht sich vor allem auf die reflexionsoptischen Messungen, kann aber ohne weiteres auf analoge Sensorsysteme (z. B. Crop-Meter, vgl. EHLERT ET AL., 2004, s.p.) übertragen werden.

sondere die sich daraus ergebende Überdüngung der Niedrigertragsbereiche (vgl. MAIDL ET AL., 2004, s.p.). Darüber hinaus kann mit diesem Ansatz keine teilflächenspezifische Optimierung der ersten N-Gabe zu Vegetationsbeginn erfolgen, da der Biomasse-Aufwuchs nicht für eine reflexionsoptische Bonitur ausreicht.

Wie bereits dargestellt, wurden beide Ansätze (Sensor bzw. Mapping) **nicht** zur ökonomischen Optimierung der Stickstoff-Ausbringung entwickelt. Es kann allerdings nicht vorausgesetzt werden, dass ein teilflächenspezifischer Ansatz im Vergleich zu einem schlageinheitlichen Ansatz **ökonomisch** erfolgreicher sein sollte, wenn beide auf ein- und demselben Konzept basieren, welches aber **nicht** zur ökonomischen Optimierung entwickelt wurde. Vor diesem Hintergrund müssen auch die Beurteilungen der Sensor- und Mapping-Ansätze nach ökologischen Aspekten (EBERTSEDER ET AL., 2003, S. 195ff, Maßstab: Stickstoff-Effizienz) gesehen werden. Die teilflächenspezifischen Ansätze sollten demnach auf ein konkretes Zielsystem ausgerichtet werden, das auch bei der Beurteilung herangezogen wird. Entsprechendes gilt für die schlageinheitlichen Ansätze.

Notwendigkeit einer neuen Konzeption

Es kann somit festgehalten werden, dass mit den bis dato verwendeten Ansätzen die ökonomischen Potenziale nicht ausgeschöpft werden können. Die angeführten Punkte verdeutlichen, dass mit einem separaten Karten- bzw. Sensor-Ansatz jeweils wichtige Informationen nicht beachtet werden. In der Umsetzung der teilflächenspezifischen Strategien werden die Probleme verstärkt, die bereits für schlageinheitliche Düngesysteme offenkundig waren: Das zentrale Element, „Expertenwissen“, ist im Hinblick auf die Komplexität der teilflächenspezifischen Optimierungsaufgabe kritisch zu sehen. Aus beiden Gründen können falsche Ableitungen gezogen werden, die zu suboptimalen Düngergaben führen. BULLOCK UND BULLOCK (2000, S. 89ff) stellen fest, dass die Fehler, die mit einer teilflächenspezifischen Düngung begangen werden, ebenso groß sein können, wie die Fehler, die gemacht werden, wenn uniform gedüngt wird. Die Kosten von falschen Ableitungen bei Precision Farming steigen aber dadurch, dass mehr Informationen verwendet werden und dadurch höhere Kosten der Informationsgewinnung entstehen.

All diese Faktoren bedingen die Forderung nach einer neuen Konzeption des Sensor-Ansatzes mit Kartenüberlagerung. Die Herausforderung besteht darin, einen Ansatz zu entwickeln, der sowohl die Informationen des Sensor-Ansatzes verwendet und gleichzeitig die Informationen des Bodens berücksichtigt. Nur dadurch können Fehler bei einer teilflächenspezifischen Düngung vermieden werden. Um das ökonomische Potenzial auszuschöpfen, sollte ein solches Konzept für die einzelnen Teilgaben alle erhältlichen Informationen berücksichtigen und dadurch auch die Verteilung der Gesamtdüngermenge auf die Teilgaben optimieren. Als Informationen können also alle verfügbaren kleinräumigen Sensordaten (vgl. Kap. 3.1.3) dienen. Dies setzt ein Wissen voraus, wie sich diese Attribute im Zu-

sammenspiel auf die kleinräumigen Produktionsfunktionen auswirken. Aus den Anforderungen an die Präzision und an die Möglichkeit zur Automatisierung sollte eine Optimierung ohne „menschliche“ Bonituren auskommen.

Dieser Ansatz muss darüber hinaus auf ein Zielsystem ausgerichtet werden, mit dem er sich auch bewerten lässt. Hier bietet sich ein streng **ökonomisches** Zielsystem an (vgl. BULLOCK ET AL., 2002, S. 238ff). Dieses Zielsystem ist auch aus ökologischer Sicht vorteilhafter als die Maximierung des Ertrages, da die optimale spezielle Intensität niedriger ist als die Stickstoffmenge, die zum Ertragsmaximum führt.

Der Status Quo sowie die erforderlichen Anpassungen sind in Tabelle 2 nochmalig zusammengetragen. Bevor in Kapitel 4 ein solches Konzept aufgebaut wird, muss jedoch näher auf die möglichen Informationsquellen eingegangen werden.

Tabelle 2: Anforderungen an einen Precision-Farming-Ansatz aus ökonomischer Sicht

	Bisher:	Gewünscht:
Zielgröße	Ertrag, Stickstoffeffizienz, langfristiges N-Gleichgewicht, ...	Stickstoffkostenfreie Leistung
Verwendete teilflächenspezifische Informationslagen (Sensoren)	1 -2	Alle verfügbaren
Optimierung von Teilapplikationen?	N-Gesamt bzw. N_2 u. N_3	Jede Teilapplikation
Expertenwissen (Landwirt)	Ertragsschätzung, Ableitung des Düngenniveaus, etc.	Automatisierte Düngung mit nachvollziehbaren Algorithmen

3.1.3 Kleinräumige Sensordaten

Wie im vorigen Kapitel dargestellt wurde, werden wegen der offensichtlichen ökonomischen Vorteile überwiegend berührungslose bzw. zerstörungsarme Sensoren verwendet. Aktuelle Studien kommen zu dem Schluss, dass die Kombination aus kleinräumigen Informationen wertvoller als konventionelle Bodenbeprobungen sind (HURLEY ET AL. 2001, S. 10). Der Wert zusätzlicher Informationen, seien es höher aufgelöste Daten (also mehr Daten pro Flächeneinheit) oder zusätzliche Attribute, ist bei der Teilflächenbewirtschaftung wesentlich höher als bei der uniformen Bewirtschaftung (BULLOCK ET AL. 2002, S. 243f). In diesem Kapitel sollen die verwendeten Sensoren, die in dieser Arbeit verwendet werden, kurz vorgestellt werden. Die Messung dieser Sensoren erfolgt in Kombination mit der Erfassung des exakten räumlichen Standorts. Dadurch können räumliche Sensor-Daten erfasst werden. Weitere Sensoren, mit denen sich eine Schlag-Variabilität erfassen lässt, sind beispielsweise in GODWIN UND MILLER (2003, S. 393ff) beschrieben.

Lokale Ertragserfassung

Besonderer Bedeutung wird der automatischen Ertragserfassung zugemessen. Diese dient einerseits zur Erfassung der Zielgröße, des Naturalertrages. Andererseits kann der teilflächenspezifische Ertrag auch als ein Attribut verwendet werden, das die Standortverhältnisse beschreibt: Eine Ertragsvariabilität auf einem Schlag wird bei einheitlichem Management durch den Einfluss der Bodenheterogenität hervorgerufen.

Ertragserfassungssysteme in Mähdreschern ermitteln die Durchflussmenge [t/h] des gedroschenen Gutes am Elevator (vgl. AUERNHAMMER ET AL., 1993). Diese Menge wird dann über die Fläche integriert, die mit Hilfe der eingestellten Arbeitsbreite und des zurückgelegten Weges berechnet werden kann. Die entstehenden Erträge [t/ha] werden einem Punkt zugeordnet, dessen geographische Position mit der Navigationstechnik bestimmt wird (vgl. AUERNHAMMER ET AL., 1994). Insbesondere dadurch, dass verschiedene Vorgänge im Mähdrescher ablaufen, bevor das Gut am Elevator (Messpunkt) ankommt, ist dieses Messverfahren verschiedenen Fehlerquellen ausgesetzt (vgl. DRUMMOND UND SUDDUTH, 2004, s.p.). Eine ausführliche Aufstellung dieser Fehlerquellen liefert beispielsweise STEINMAYR (2002, S. 8ff). Diese Fehler werden zum einen bereits während der Aufzeichnung durch entsprechende Sensoren und Algorithmen minimiert. Zum anderen kann eine Korrektur auf Basis der Ertragskarten erfolgen. In mehreren Arbeiten wurde untersucht, wie sich dieses Kartenmaterial verbessern lässt. Dazu werden von diversen Autoren Filter vorgeschlagen, mit denen sich kritische Datenpunkte entfernen lassen (vgl. THYLEN ET AL. 2000, s.p.). Eine Alternative besteht darin, die Umgebung der einzelnen Ertragspunkte zu untersuchen, um mit verschiedenen Verfahren der räumlichen Statistik Korrekturmöglichkeiten abzuleiten (vgl. NOACK ET AL. 2003, S. 446ff; BACHMEIER UND AUERNHAMMER, 2004, S. 742ff).

Bodenleitfähigkeit

Eine weit verbreitete nicht-destruktive Methode zur Erfassung der Bodenheterogenität ist die Messung der scheinbaren elektrischen Bodenleitfähigkeit (EC_a). Eine ausführliche Beschreibung der verwendeten physikalischen Prinzipien und der praktischen Anwendung ist beispielsweise in CORVIN UND LESCH (2003) zu finden.

Ursprünglich wurde diese Methode entwickelt, um den Salzgehalt des Bodens bei bewässerten Schlägen zu erfassen. Während in Salzböden die elektrische Leitfähigkeit (EC_a) im Allgemeinen durch die Leitfähigkeit der Bodenlösung bestimmt wird, verspricht man sich für Böden humiden Klimas mit geringen Salzgehalten Aufschlüsse insbesondere auf den Ton- und Wassergehalt des Bodens.

Zur Messung dieser Größe werden der elektrische Widerstand bzw. die elektromagnetische Induktion verwendet. Während sich die Messung des elektrischen Widerstandes vor allem auf trockenen und steinigten Böden aufwändig gestaltet, da dazu Elektroden in den Boden geführt werden müssen, kann die elektromagnetische Induktion mit einem einfachen Gerät

aus zwei Spulen, das über den Boden gezogen wird, kleinräumig und schnell erfasst werden.

Dazu wird oft der kommerziell erhältliche EM38 Sensor von Geonics® verwendet. Dieser Sensor verwendet einen Spulenabstand von 1 m und bestimmt dadurch die Werte bis zu einer Tiefe von ungefähr 1.5 m, was in etwa der durchwurzelbaren Tiefe entspricht. Die Messwerte im empirischen Teil dieser Arbeit wurden mit diesem Gerät erfasst. Eine Darstellung der Messung ist in Abbildung 15, der genaue technische Aufbau dieses Gerätes in CORVIN UND LESCH (2003, S. 462) abgebildet.



Abbildung 15: Messung der scheinbaren elektrischen Leitfähigkeit

DURLESSER (1999, S. 109f) stellte für den Standort „Scheyern“ (45 km nordwestlich von München) fest, dass die Variation von EC-Werten im Wesentlichen auf den Tongehalt, auf den Wassergehalt und auf die Bodentemperatur zurückzuführen ist. Während die Temperatur relativ einfach auf einen Normwert von 25° C zu korrigieren ist (EC_{25}), führt die zeitliche Variation des Wassergehaltes im Oberboden zu Problemen (BREVIK ET AL., 2004, S. 148ff): Mit der vegetations- und reliefbedingten Variation des Wassergehaltes ändert sich nicht nur das Niveau, sondern auch die Variation der scheinbaren elektrischen Leitfähigkeit (SUDDUTH ET AL. 2001, S. 248ff). Der Messwert setzt sich daher aus mehreren Bodeneigenschaften und –zuständen zusammen.

Eine „absolute Skalierung“, d. h. eine jahreszeitenunabhängige Bestimmung einzelner Bodeneigenschaften kann nur in Verbindung mit einer Bodenbeprobung erreicht werden.

Zugkraft

Das Beispiel der *Zugkraftmessung* zeigt, dass nicht immer aufwändige Apparaturen notwendig sind, um Informationen über Bodenvariabilitäten zu erhalten. Zur Messung der Zugkraft [\sim kN] werden lediglich die Messprotokolle der elektronischen Hubkraftregelung

während der Bodenbearbeitung abgegriffen. In der Funktionsweise entspricht die Messung einem horizontalen Penetrometer. Unter Einbindung der GPS-Positionsdaten können räumliche Daten erzeugt werden, die einen Einblick in die (Ober-)Bodenverhältnisse ermöglichen. Für die vorliegende Arbeit wurden Zugkraftmessungen während der Stoppelbearbeitung an einem Lemken Smaragd-Grubber verwendet (vgl. Abbildung 16). Das Protokoll der Prozessdaten erfolgt über das landwirtschaftliche BUS-System LBS nach DIN 9684 / ISO 11783 (vgl. ROTHMUND ET AL. 2003, S. 306).

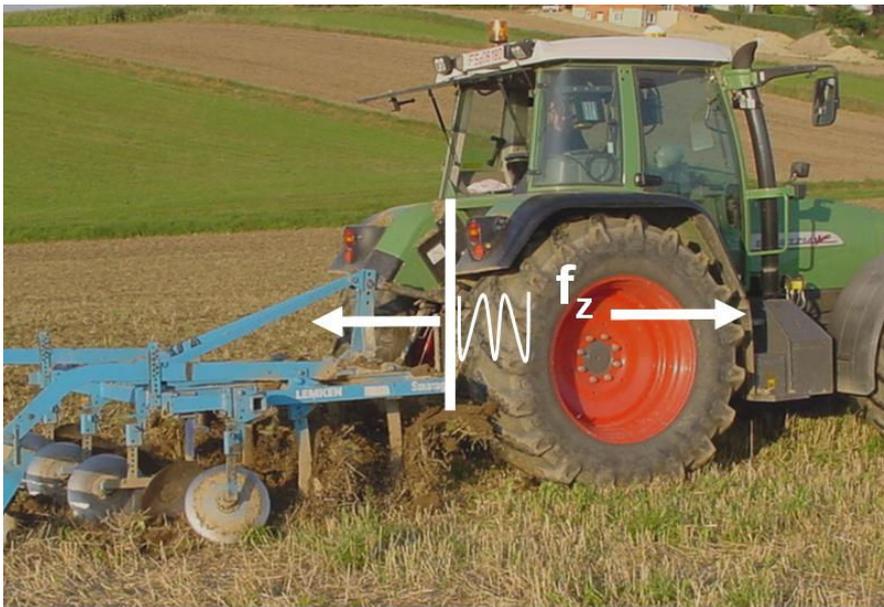


Abbildung 16: Zugkraftmessung bei der Stoppelbearbeitung

Bisher haben sich nur vereinzelt Autoren mit einer Analyse dieser Bodeninformation beschäftigt. LAPEN ET AL. (2002, S. 241ff) untersuchten die Zugkraft zur wendenden Bodenbearbeitung und verglichen diese Daten mit teilflächenspezifischen Ertragsdaten. Sie fanden negative Korrelationen zum Ertrag bei Mais und wechselnde Korrelationen bei Weizen (in Abhängigkeit vom Ertragsniveau). ROTHMUND ET AL. (2003, S. 308f) konnten eine Korrelation zwischen der Zugkraft und dem Tongehalt im Oberboden nachweisen (Bestimmtheitsmaß: 0.6).

Der Vorteil dieses Sensors ist darin zu sehen, dass die räumliche Auflösung, bedingt durch die geringe Arbeitsbreite des Bodenbearbeitungsgerätes, sehr hoch ist. Weiterhin sind die geringen Kosten hervorzuheben, da wesentliche Komponenten des Messaufbaus zur Standardausstattung moderner Traktoren gehören.

Nachteilig wirkt sich aus, dass diese Bodeninformation noch nicht umfassend untersucht wurde. Beispielsweise darf angenommen werden, dass der Wassergehalt des Bodens wie auch bei der Messung der elektrischen Leitfähigkeit einen Einfluss auf das Niveau und auf die Variation dieses Attributes hat. Hinzu kommen die Einflüsse des Managements beispielsweise über lokale Boden-Verdichtungen durch den Einsatz schwerer Maschinen.

Ebenso gilt dies für die Bedingungen während der Aufzeichnung (Geschwindigkeit, Arbeitstiefe, Hangneigung, etc.). Die gemessenen Werten dürfen demnach nicht als absolute Bodenparameter interpretiert werden, sondern erlauben maximal die relative Beurteilung des Bodens im Hinblick auf den Zugwiderstand, sofern die Aufzeichnung in einem Schlag unter konstanten Bedingungen erfolgt ist.

Reflexionsoptische Messungen

Eine Technologie mit stetig wachsender Bedeutung im Bereich Precision Farming ist die reflexionsoptische Bonitur des Bestandes. Dazu wird mit einer Sensoreinheit die Reflexion des Pflanzenbestandes im roten und nah-infraroten Spektralbereich erfasst. Es stellte sich heraus, dass anhand der Reflexionseigenschaften des Pflanzenbestandes auf den Ernährungszustand geschlossen werden kann. Ein Effekt, der dabei verwendet wird, ist die Verschiebung des *Hauptwendepunktes* (Red Edge, bzw. Red Edge Inflection Point, REIP) bei unterschiedlichen Chlorophyllgehalten. Da der Chlorophyllgehalt eng mit der Stickstoffversorgung korreliert, kann über die Lage des Hauptwendepunktes der Ernährungszustand der Pflanzen erfasst werden (vgl. REUSCH, 1997, S. 131ff, LIEBLER, 2003, S. 19ff und SCHÄCHTL, 2004, S. 9ff). Die Berechnung der Wellenlänge des Hauptwendepunktes stützt sich auf die Näherungsformel von GUYOUT ET AL. (1988):

$$REIP = 700 + 40 \frac{(R_{670} + R_{780}) / 2 - R_{700}}{R_{740} - R_{700}}$$

Verschiedene Autoren konnten gute Korrelationen dieses Vegetationsindex' mit der N-Aufnahme und dem N-Gehalt feststellen (vgl. LIEBLER ET AL. 2001, S. 84ff, SCHMIDHALTER ET AL. 2003, S. 617f)

Die Kalibrierung dieses Sensors erfolgt über Stickstoffsteigerungsversuche. Zur Ableitung einer Düngermenge werden mit dem N-Tester Messungen durchgeführt und daraus Soll-Werte für vordefinierte Entwicklungsstadien abgeleitet, oder aber der Landwirt legt aufgrund von Erfahrungswerten einen Soll-Ernährungszustand für das jeweilige Stadium fest. Die Differenz zwischen diesem Soll-Ernährungs-Zustand und dem Ist-Zustand wird in eine proportional abhängige Düngermenge verrechnet.

Zur korrekten Ableitung von Düngergaben aus diesen Kennzahlen bestehen noch einige ungeklärte Fragen. Insbesondere wird von mehreren Autoren angezweifelt, dass eine Düngung, die nur auf diesen Ernährungszuständen (ohne Berücksichtigung der Bodenverhältnisse, des Wasservorrats und der bodenbürtigen Stickstoffnachlieferung) basiert, erfolgreich sein kann (PIEKIELEK ET AL. 1995, S. 407f, LIEBLER, 2003, S. 150ff). MAIDL ET AL. (2004, s.p.) berücksichtigen diese Effekte mit einer Anpassung an (geschätzte) Ertragswartungen. Sie verwenden dazu die Schätzung der optimalen N-Aufnahme von DIEPOLDER (1994, S. 152).

Es existieren verschiedene technische Umsetzungen für eine reflexionsoptische Sensoreinheit. In vielen Versuchen wurden Hand-Sensoren verwendet (vgl. LIEBLER, 2003, S. 39ff). Für größere Versuchspartzen und in der Praxis werden am Traktor montierte Sensoren verwendet (z. B. Yara-N-Sensor). Als vorteilhaft erwiesen sich Sensoren, die mehrere Einkoppelungsoptiken verwenden, die in unterschiedliche Richtungen weisen (oligo-view). Die eingehenden Signale werden dabei zu einem Wert verrechnet. Dadurch wird die Messung unabhängiger vom Blickwinkel und dem Sonnenstand (vgl. MISTELE, 2004, s.p.). In der vorliegenden Arbeit wurde ein Handsensor³⁷ sowie ein Sensor der Firma Yara®³⁸ (oligo-view) verwendet.

Die Technologie der reflexionsoptischen Messungen bezieht ihren Reiz daraus, dass während der Vegetation (also im Allgemeinen während oder kurz vor den Düngezeitpunkten) aktuelle Werte zur Stickstoffversorgung der Pflanze (insbes. N-Aufnahme) abgebildet werden können. Vereinfacht werden von der Pflanze „verarbeitete“ Boden- und Klimaverhältnisse während der Vegetationsperiode („in-season“) erfasst. Die Pflanze wird somit als Indikator benutzt. Diese Informationen ermöglichen eine Optimierung der Ertragsprognose (vgl. Kap. 3.2): Anstelle von Witterungskennzahlen, die lediglich Rahmenbedingungen beschreiben, können nun direkt Informationen von der Pflanze erfasst werden, die ein Integral der Wachstumsbedingungen bis zu diesem Zeitpunkt abbilden (vgl. Kap. 3.2).

Sonstige Sensoren

Neben den Sensoren, die in dieser Arbeit verwendet werden, existieren noch eine Reihe von Sensoren, die in der Praxis eine gewisse Bedeutung erlangt haben. Erwähnt sei der Pendelsensor (Crop-Meter), der über den Grad der Auslenkung eines Pendels Rückschlüsse über die Bestandesdichte zulässt, die dann für die Bemessung der Düngergabe herangezogen wird. Gleichzeitig kann die GPS-Navigation selbst als „Sensor“ aufgefasst werden. Mit Hilfe der Positionsbestimmung bei Precision Farming wird gleichzeitig die Höhe (über Meeresspiegel) erfasst. Dadurch ist es relativ einfach, Geländemodelle zu erstellen (vgl. GODWIN UND MILLER, 2003, S. 398ff). Mögliche Attribute, die die Standortverhältnisse beschreiben, sind davon ausgehend beispielsweise die Erhebung und Geländeneigung. Bei unebenen Schlägen kann die Information, die in diesen Attributen enthalten ist, von hoher Bedeutung für eine teilflächenspezifische Ertragsprognose sein.

Zusammenfassend kann festgehalten werden, dass mittlerweile sehr viele Instrumente zur Verfügung stehen, um eine Variabilität der Standortverhältnisse zu erfassen. Neben den „konventionellen“ Bodenbeprobungen stehen teils aufwändige, teils sehr einfache Apparaturen zur Verfügung, mit denen sich die Variabilität erfassen lässt. Ungünstig wirkt sich

³⁷ Es handelt sich dabei um ein portables zweikanaliges Spektralsensorsystem der Firma tec5 mit 2 MMM-VIS Spektralsensoren der Firma Carl Zeiss. Zu einer genaueren Beschreibung vgl. LIEBLER (2003, S.39f.)

³⁸ Für den genauen Aufbau vgl. MISTELE (2004, s.p.).

aus, dass mit den beiden Bodensensoren (Zugkraft und scheinbare Leitfähigkeit) zwar Variabilitäten eines Schlages erfasst werden können, aber die Messgrößen sehr vom Kontext abhängig sind und dadurch keiner absoluten Skala zugeordnet werden können. Beispielsweise liefert eine gemessene Zugkraft von 25 (kN) für sich genommen keine Information, da die externen Effekte bei der Messung nicht quantifiziert werden können. Der Wert kann nur interpretiert werden, wenn die Werte mit anderen Messwerten unter den gleichen Verhältnissen verglichen werden. Eine Wiederholung der Zugkraft-Messung bzw. der scheinbaren elektrischen Leitfähigkeit unter anderen Witterungsverhältnissen kann zu unterschiedlichen Ergebnissen bezüglich des Niveaus und der Variation führen³⁹.

Neben einer isolierten Betrachtung dieser Attribute bietet sich jedoch die Kombination dieser Sensoren an. Ein erster Schritt zu einem besseren Verständnis des Zusammenwirkens dieser Attribute kann durch eine empirische Analyse erreicht werden, wie im folgenden Abschnitt gezeigt wird.

3.2 Schnittmenge II: Precision Farming und Data Mining

Im diesem Abschnitt werden die Schnittpunkte zwischen Data Mining und Precision Farming untersucht. In diesen Bereich fallen die Arbeiten, in denen empirische Daten aus dem Bereich Precision Farming mit Methoden des Data Mining analysiert werden. Dieses Kapitel soll Anhaltspunkte für die Entwicklung der Methodik für den empirischen Teil dieser Arbeit liefern.

Ein entscheidender Schritt hin zur Erzeugung eines teilflächenspezifischen Ansatzes zur Stickstoffausbringung ist zunächst eine intensive Analyse der Ertragsvariabilität (SHATAR UND MCBRATNEY 1999, S. 249f). Dazu werden in der Regel Modelle zur kleinräumigen bzw. teilflächenspezifischen Ertragsprognose entwickelt (DOBERMANN, 2004, s.p.). Insbesondere dafür können Werkzeuge des Data Mining eingesetzt werden. Dabei hilft, dass diese Zusammenhänge empirisch relativ kostengünstig untersucht werden können. Die Durchführung von Versuchen im Schlagmaßstab wird durch den Einsatz von Precision-Farming-Technologien einfach und kostengünstig (GANDORFER ET AL., 2004, S. 227).

Verschiedene Autoren untersuchten empirische Datensätze und fanden lineare Korrelationen zwischen Ertrag und verschiedenen Bodenattributen, die aber innerhalb eines Schlages und zwischen mehreren Schlägen stark variierten (vgl. KHAKURAL ET AL. 1999, 577ff.; KRAVCHENKO UND BULLOCK, 2000, S. 78ff, SPAULDING ET AL. 2005, s.p.). Gleichzeitig erwiesen sich diese Korrelationen zeitlich gesehen instabil (LAMB ET AL., 1997, S. 412ff). Auch die Anwendung von komplexeren linearen Methoden wie beispielsweise einer mul-

³⁹ Bei den Spektralindices deutet sich an, dass diese Größen absolute Aussagen ermöglichen, dass also mit einem konkreten Wert auch entsprechend absolute Aussagen zum Ernährungszustand getroffen werden können (vgl. MISTELE, 2005, S. 56).

tiplen linearen Regression zeigte bei verschiedenen Autoren nicht zufrieden stellende Resultate (KRAVCHENKO UND BULLOCK, 2000, S. 80, DRUMMOND ET AL., 2003, S. 12ff).

Als überlegen erwiesen sich nicht-lineare Methoden. DRUMMOND ET AL. (2003) verglichen verschiedene Modelle zur Ertragsprognose und konnten die Überlegenheit von neuronalen Netzen im Vergleich zu verschiedenen linearen Methoden zeigen. In Tabelle 3 ist eine Auswahl von Studien zusammengestellt, in denen neuronale Netze zur Ertragsprognose verwendet werden. Die Autoren verwenden dazu „Multilayer-Feedforward-Netze“ mit einem Backpropagation-Lernalgorithmus. Die Topologie ist jeweils in der Übersichtstabelle angegeben. Die Notation erfolgt in Anlehnung an das Kapitel 2.5.1. Darüber hinaus wird in dieser Übersichtstabelle jeweils die Methode der Validierung sowie die verwendete Maßzahl der numerischen Vorhersage angegeben (vgl. hierzu Kap. 2.7.1 und Tabelle 1).

In den Studien werden einerseits Daten aus Exaktversuchen (Parzellendaten) und andererseits Daten aus Precision-Farming-Versuchen im Schlag-Maßstab verwendet. Die Daten eines Schlages aus einem Jahr werden in der Regel als Schlag-Jahr (engl. site-year) bezeichnet. Die Unterschiede zwischen den einzelnen Studien bestehen in erster Linie in den verwendeten Attributen zur Ertragsprognose. Aufgrund der Vielzahl der verwendeten Attribute werden diese in Gruppen zusammengefasst. Grob lassen sich die Attribute gliedern in

- Bodeneigenschaften über Bodenbeprobung (Nährstoffversorgung, pH-Wert, Kationen-Austausch-Kapazität, Humusanteil, Lehmanteil, Tonanteil, etc.),
- scheinbare elektrische Leitfähigkeit (EC),
- Managementattribute (Sorte, applizierter Stickstoff, Bewässerungsintensität, etc.),
- topographische Attribute (Erhebung, Hangneigung, etc.) und
- Klimadaten (Witterungsdaten einzelner Monate, Vegetationszeit, etc.).

Der Aufwand zur Erfassung der Attribute ist damit entsprechend unterschiedlich. Ebenso lassen sich nicht alle Attribute kleinräumig erheben, bzw. sind auch nicht kleinräumig variabel (vgl. z. B. Witterung).

Eine weitere Untergliederung liegt in der Quantität der verwendeten Daten, also in der Anzahl der Instanzen. In diesem Rahmen muss weiterhin unterschieden werden, ob die trainierten neuronalen Netze nur zu einer ex post Ertragsprognose für die Erträge eines bestimmten Schlages in einem Jahr verwendet werden, oder ob mit den Modellen zukünftige Erträge geschätzt werden. In diesem Fall wird von einer „echten“ Prognose gesprochen. Zur Durchführung einer „echten“ Prognose werden in der Regel Witterungsdaten in die Modelle einbezogen. Weiterhin unterscheiden sich die Studien in Bezug auf

- die Datenaufbereitung,

- die verwendeten Algorithmen und Topographien zum Training der neuronalen Netze (vgl. Kap. 2.5.3) und
- die verwendeten Methoden zu Validierung (vgl. Kap. 2.7.).

In der Datenaufbereitung für Precision-Farming-Datensätze ist unter anderem bedeutsam, mit welcher Methode die Datenpunkte interpoliert wurden bzw. mit welcher Rastergröße (Teilflächen) gearbeitet wurde.

Eine der ersten deutschsprachigen Arbeiten in diesem Gebiet wurde von KOLLIG (1993) angefertigt. Er versuchte, mit neuronalen Netzen aus Parzellendaten von Bundessortenversuchen Erträge von Zuckerrüben in Abhängigkeit von diversen Attributen zu schätzen. Das Datenmaterial umfasste 70 verschiedene Standorte aus fünf Jahren. Dabei zeigte sich ein großer Einfluss klimatischer Daten auf den Ertrag. Wenn die jahreszeitlichen Einflüsse als Input-Attribute weggelassen werden, versiebenfachte sich der Fehler der Prognose. Für die Validierung wurde eine Mischung aus Trainings- und Validierungsdaten verwendet, so dass kein unabhängiger Fehler ermittelt werden konnte. Deshalb kann Überlernen nicht ausgeschlossen werden.

SUDDUTH ET AL. (1998) testeten erstmals neuronale Netze für eine teilflächenspezifische Ertragsprognose, indem sie räumliche Daten aus der Anwendung von Precision-Farming-Technologien benutzten. Dazu verwendeten sie teilflächenspezifische Daten eines Schlates und versuchten ex post Erträge dieses einzelnen Schlag-Jahres zu schätzen. Es wurde eine 5-fache Kreuzvalidierung eingesetzt, um Überlernen zu vermeiden. Das beste neuronale Netz erzielte einen Standard Error of Prediction (SEP) von 0.20 t/ha und übertraf damit die Ergebnisse anderer Methoden zur Ertragsprognose.

Diese Untersuchung wurde erweitert und DRUMMOND ET AL. (2000) veröffentlichten die Ergebnisse. Das Datenmaterial wurde in diesem Fall von drei Schlägen in einer Soja-Mais-Fruchtfolge in Missouri, USA, erhoben. Insgesamt standen ihnen 10 Schlag-Jahre⁴⁰ zur Verfügung. Zusätzlich wurden in dieser Studie Klimadaten miteinbezogen und versucht, Erträge in unterschiedlichen Jahren zu schätzen. Dies führte jedoch zu einem beträchtlichen Überlernen. Den Grund sehen die Autoren in der eingeschränkten Anzahl von klimatisch verschiedenen Jahren. Je nach Schlag standen dem Modell nur maximal drei verschiedene Ausprägungen von Klimadaten zur Verfügung, so dass ein Lernerfolg nur eingeschränkt möglich war⁴¹. Die „Interpolation“ aus diesen Einzelwerten ist erwartungsgemäß schlecht. Die Autoren folgern, dass eine beträchtlich höhere Anzahl von klimatisch verschiedenen Jahren nötig wäre, um eine Überanpassung zu vermeiden.

⁴⁰ engl. site-years. Die Daten eines Schlates in einem Jahr ergeben ein Schlag-Jahr.

⁴¹ Klimadaten ändern sich teilflächenspezifisch kaum, deshalb existiert beispielsweise für das Attribut „Niederschlagsmenge Mai“ nur eine Ausprägung pro Jahr. Bei einer Trainingsmenge aus drei Jahren müsste quasi eine Funktion aus drei Punkten geschätzt werden.

Tabelle 3: Übersicht von Arbeiten zur Ertragsprognose mit neuronalen Netzen

Autor	Fruchtart	Daten	Attribute	Daten-aufbereitung	Topologie*	Validierung**
Kollig, 1993	Zucker- rüben	5 Jahre 1171 Instanzen	31 Attribute: Klima, Boden, Management	Parzellendaten	v.a. 31-17-1	Hold-Out MPA (4.21 %)
Sudduth et al., 1998	Soja	1 Schlag- Jahr 340 Instanzen	7 Attribute: Boden, Topo- graphie, EC	30 m-Raster Kriging	7-10-1	Kreuzvalidierung SEP (0.20 t/ha)
Schearer et al., 1999	Mais	1 Schlag- Jahr 200 Instanzen	4 Attribute: EC, Topogra- phie, Satelliten- daten	63 m-Raster Inverse-Distance	-variierend-	Hold-Out SEP
Drummond et al., 2000	Mais, Soja	10 Schlag- Jahre à ca. 300 Instanzen	10 Attribute: Boden	25 m-Raster	-variierend-	Kreuzvalidierung SEP (0.2–1.1 t/ha)
Liu et al., 2001	Mais, Soja	360 Instanzen	15 Attribute: Boden, Klima, Management	Parzellendaten	15-20-1	Hold-Out RMS-Error (20%)
Drummond et al., 2002	Mais, Soja	44 Schlag- Jahre à ca. 3000 Instanzen	4 Attribute: EC, Topogra- phie	10 m-Raster Kriging	4-10-1	Hold-Out R ² (0.09-0.67) SEP (0.1-1.19 t/ha)
Diker et al., 2004	Mais	6 Schlag- Jahre à ca 100 Instanzen	12 Attribute: Boden, Klima, Management, Topographie	76 m-Raster	12-9-2-1	Hold-Out R ² (0.94) SEP (0.05 t/ha)

* zur Darstellung der Topologie sei auf Abschnitt 2.5.1 verwiesen

** verwendete Methoden und Validierungskennzahlen werden in Abschnitt 2.7 sowie Tabelle 1 erklärt

SHEARER ET AL. (1999) verwendeten zur Ertragsprognose zusätzlich zu Bodeneigenschaften und Topographiedaten erstmals Satellitenbilder der aktuellen Vegetation. Als Datensatz stand allerdings nur ein Schlag-Jahr zur Verfügung. Der Datensatz beschränkte sich auf 63 Instanzen zum Trainieren und 137 Instanzen zum Testen. Methoden zur Erkennung und Vermeidung von Überanpassung wurden trotz der geringen Anzahl von Instanzen nicht eingesetzt.

LIU ET AL. (2001) griffen die Ergebnisse von DRUMMOND ET AL. (2000) auf und entwickelten eine Ertragsprognose für Mais. Dazu verwendeten sie Daten aus einem langjährigen Exaktversuch. Damit gelangten sie zwar zu Klimadaten und damit verbundenen Ertragsdaten, die bis zu 30 Jahre zurück reichten, pro Jahr standen allerdings nur 12 (!) Trainingsda-

tensätze zur Verfügung. Zusammen mit einer sehr hohen Anzahl von Prognose-Attributen (15 Input-Neuronen) ist das Auftreten von Überlernen höchst wahrscheinlich. Diesen Verdacht entkräften die Autoren nicht, da sie zur Validierung lediglich die Hold-Out-Methode verwenden.

Als eine der umfangreichsten Arbeiten in diesem Gebiet untersuchten DRUMMOND ET AL. (2002) in einem Projekt über mehrere Jahre und mehrere US-Staaten (Illinois, Iowa, Michigan, Missouri, South Dakota und Wisconsin) hindurch den Einsatz von neuronalen Netzen zur teilflächenspezifischen Ertragsprognose. In dieser Studie verzichteten die Autoren auf Bodenbeprobungen und erzielten dennoch gute Resultate. In einer Ex-post-Prognose konnten sie für die verschiedenen Schlag-Jahre 9-67 % der Ertragsvariabilität mit Sensordaten (Topographie und EC) erklären, im Durchschnitt über alle Schlag-Jahre waren dies 38 %. Die Ergebnisse legen nahe, dass die Prognoseleistung eng mit der Variabilität der Schläge zusammenhängt. Aus der Analyse von Fehlerkarten schließen sie, dass Verbesserung durch die Hinzunahme weiterer Input-Attribute erreicht werden könnten.

DIKER ET AL. (2004) verwendeten sechs Schlag-Jahre von zwei bewässerten (!) Schlägen. Als Besonderheit klassifizierten sie die Instanzen nach dem Jahr, in dem sie erhoben wurden und verwenden dieses Attribut als Input-Attribut. Das neuronale Netz wurde mit den Instanzen aller Schlag-Jahre trainiert. Im trainierten Modell besitzt genau dieses Attribut die größte Beutung für die Ertragsprognose, ohne dass das Modell dann für eine „echte Prognose“ verwendet werden kann, da für neue Jahre kein Jahres-Index zur Verfügung steht. Weiterhin wird die Menge an zugeführtem Wasser für die Prognose verwendet. Damit sind die guten Ergebnisse der Ertragsprognose zu erklären.

Es wird spätestens mit diesem Beispiel offensichtlich, dass sich die bis dato erstellten Studien in ihrer Prognoseleistung nur schwer miteinander vergleichen lassen. Es existieren große Unterschiede wie in Tabelle 3 dargestellt ist. Weiterhin existieren viele kleinere Unterschiede, in den Rahmenbedingungen, deren Auswirkungen schwer abzuschätzen sind. Ein großes Problem ist, dass deskriptive statistische Untersuchungen der Input-Instanzen meist nicht angegeben werden, so dass die Beurteilung der Ergebnisse schwierig ist.

Dennoch liefern die Ergebnisse einen Rahmen, um neue Resultate einordnen zu können⁴². Gleichzeitig sollen in vorliegender Arbeit Methoden aus diesen Studien eingesetzt werden, die sich bei den betrachteten Arbeiten bewährt haben. Dazu gehören insbesondere die *Sensitivitätsanalyse* (LIU ET AL., 2001, S. 711) sowie *Response Curves* und *Response Surfaces* (POCRAJAC UND OBRADOVIC, 2001, S. 31, DRUMMOND ET AL., 2002, s.p.; LIU ET AL., 2001, S. 711). Nicht zuletzt können einige Punkte zur Ausgestaltung von Ertragsprognosemodellen abgeleitet werden:

⁴² Es wird davon ausgegangen, dass mit der Fruchtart Winterweizen die Ertragsprognose nach vergleichbaren Prinzipien erfolgen kann.

Generell zeigt sich, dass die Ergebnisse von Bodenbeprobungen nicht notwendiger Bestandteil für eine Ertragsprognose sind. Dadurch, dass diese Labordaten nicht auf einen Wert komprimierbar sind, erhöhen sie die Anzahl der Input-Attribute und damit die Gefahr von Überanpassung (DRUMMOND ET AL., 2003, S. 10f). Die Bedeutung einzelner Bodeneigenschaften und speziell der Nährstoffversorgung ist für die Prognose eher gering.

Ein weiterer Punkt betrifft die Anzahl der Instanzen. Eine geringe Anzahl von Instanzen beeinträchtigt das Lernergebnis enorm. Zusammen mit einer hohen Anzahl von Input-Attributen ist das Auftreten von Überlernen unvermeidbar. Deshalb eignen sich insbesondere die im vorigen Kapitel beschriebenen kleinräumigen Sensoren, mit denen große Mengen an Instanzen zu sehr geringen Kosten erhoben werden können. Die stabilsten Ergebnisse werden mit diesen Precision-Farming-Daten und wenigen Input-Attributen erreicht, da die Gefahr von Überlernen stark eingeschränkt ist.

Weiterhin zeigte sich, dass, falls Klimadaten verwendet werden und damit eine „echte Prognose“ erreicht werden soll, eine sehr hohe Anzahl von verschiedenen Jahresdaten benötigt wird, um eine Überanpassung zu vermeiden. Einen Ausweg liefern unter Umständen reflexionsoptische Messungen. Das Grundproblem der „echten“ Prognose liegt in der zeitlichen Instabilität der Ertragsmuster (WENDROTH ET AL., 2003, 258ff). Über reflexionsoptische Daten könnten die jahreszeitlichen Einflüsse erfasst werden⁴³. Die Gefahr von Überlernen und damit der Bedarf an multiplen Jahresdaten könnte gesenkt werden, da durch die teilflächenspezifische Erhebung auf heterogenen Schlägen pro Jahr sehr viele Ausprägungen dieses Attributes erhoben werden. Diese Idee wurde in Ansätzen von SHEARER ET AL. (1999) durch die Hinzunahme von Satellitendaten der aktuellen Vegetation verwirklicht und soll auch in vorliegender Arbeit getestet werden.

Mit den vorgestellten neuronalen Netzwerk-Modellen wurde also die Interaktion von Boden- und Pflanzencharakteristiken auf den Ertrag erfolgreich abgebildet. Für eine Ableitung von **teilflächenspezifischen** Düngestrategien müssen jedoch die Interaktionen von Boden, Pflanze und Stickstoff-Management erfasst werden. Die betrachteten teilflächenspezifischen Ertragsprognosen wurden stets ohne Management-Attribute erstellt.

Um diese Interaktionen zu modellieren, müssen verschiedene Voraussetzungen erfüllt werden: Mit diesem Modell muss eine Ertragsprognose für variierende Stickstoff-Mengen ermöglicht werden. Deshalb muss in den Trainingsinstanzen das Attribut „Stickstoffmenge“ in verschiedenen Ausprägungen vorliegen und gleichzeitig mit teilflächenspezifisch wechselnden Boden- bzw. Pflanzenattributen in Verbindung gebracht werden. Dies erfordert für die Versuchsanstellung den Einsatz von Randomisierungsstrategien auf Großschlägen

⁴³ Dahinter steckt die Annahme, dass ein bestimmter (Vegetations-)Indexwert in jedem Jahr dieselbe Bedeutung hat. Ein Indexwert von X zum Zeitpunkt T deutet demnach in jedem Jahr auf den gleichen Ertrag hin. Eng damit verbunden ist die absolute Skalierbarkeit dieser Indices.

(POCRAJAC UND OBRADOVIC, 2001) und wurde im Rahmen dieser Arbeit erstmalig durchgeführt. Mit Daten aus Parzellenversuchen konnte die Modellierung dieser Effekte bereits gezeigt werden (LIU ET AL., 2001). Weiterhin wurde von CHAUDAHARY ET AL. (2005, s.p.) die Anwendung von verschiedenen Werkzeugen des Data Mining an Daten aus einem Stickstoff-Steigerungsversuch im Schlagmaßstab (field-scale) vorgestellt. POCRAJAC UND OBRADOVIC (2001, S. 5ff) gelang die Ableitung von teilflächenspezifischen Düngestrategien mit simulierten Daten mit Hilfe neuronaler Netze. In dieser Arbeit soll dies an empirischen Daten versucht werden.

Mit diesem Kapitel endet die Analyse des Anwendungsgebietes. In dieser Analyse wurde der Status Quo sowie die Problemfelder in den Bereichen Stickstoffdüngung und Precision Farming aufgezeigt. In diesem Abschnitt wurden erste Berührungspunkte zwischen diesen beiden Gebieten vorgestellt. Es zeigt sich, dass neuronale Netze ein geeignetes Data-Mining-Werkzeug zur Analyse und zur Prognose von teilflächenspezifischen Erträgen darstellen. Gleichzeitig wurden in den vorherigen Kapiteln die notwendigen Entwicklungen zur angestrebten Neuentwicklung eines Sensor-Ansatzes mit Kartenüberlagerung angedeutet. Die Zusammenführung dieser Ansätze erfolgt im folgenden Abschnitt.

4 Entwicklung von Entscheidungsregeln zur kleinräumigen Stickstoff-Ausbringung

Nachdem die Anforderungen in den vorausgegangenen Kapiteln deutlich gemacht wurden, wird in diesem Kapitel eine Methode zur Erzeugung von Entscheidungsregeln vorgestellt, das auf einem modifizierten Konzept der Wissensentdeckung in Datenbanken basiert.

Der Aufbau des Kapitels selbst orientiert sich an diesem Prozess (vgl. Kap.1). Dadurch, dass es sich um einen neuen Ansatz handelt, konnte insbesondere zum Versuchsdesign und zur Durchführung des Prozesses der Wissensentdeckung in Datenbanken (WED) auf keine vergleichbaren Arbeiten und Erfahrungen zurückgegriffen werden. Für den empirischen Teil dieser Arbeit wurden Feldversuche angelegt, deren Design deshalb mehrfach angepasst wurde. Gleiches gilt für die Methode zur Analyse. Zu Beginn dieses Kapitels werden die Ziele sowie die wichtigsten Schritte eines idealtypischen Prozesses dargestellt, der sich gleichzeitig zur Gliederung der einzelnen Schritte eignet. Im Anschluss daran werden Elemente dieses Prozesses ausführlich behandelt.

4.1 Ziel und Überblick eines modifizierten WED-Prozesses

Das Ziel des empirischen Teiles dieser Arbeit besteht zunächst darin, eine Methode zu entwickeln, mit Hilfe derer Entscheidungsregeln zur kleinräumigen Optimierung generiert werden können. Dazu müssen Feldversuche angelegt werden, aus denen die Instanzen für einen WED-Prozess gewonnen werden können. Weiterhin sollen die Ergebnisse auch in der Realität überprüft werden. Aus dieser Anforderung heraus wurde ein Konzept entwickelt, bei dem ein einzelner Feldversuch gleichzeitig als Datenquelle sowie zur Evaluierung der erzeugten Regeln dient⁴⁴.

Abbildung 17 stellt einen modifizierten WED-Ansatz für Precision Farming dar und zeigt vereinfacht ein idealtypisches Vorgehen auf. Dieses Modell basiert auf dem Ansatz des kontinuierlichen Lernens (PYLE, 1999, S. 39ff.). Nach diesem Ansatz werden die Ergebnisse kontinuierlich in der Realität evaluiert und verifiziert und können dadurch stetig verbessert werden.

Das zentrale Element stellen dabei Feldversuche dar. Das genaue Design dieser Versuchsanlagen sowie die Anpassungen, die im Laufe der Versuchsjahre vorgenommen wurden, werden im Abschnitt 4.2 ausführlich beschrieben. In diesen Versuchen werden im ersten Zugriff Daten generiert, die in Form von Instanzen als Basis für das überwachte Lernen verwendet werden.

⁴⁴ Natürlich können dazu auch zwei unterschiedliche Versuche durchgeführt werden, die Idee bestand darin, das Data Mining Konzept anzuwenden, bei dem der versuchstechnische Aufwand gering bleibt.

Die Besonderheit dieses WED-Prozesses ist darin zu sehen, dass die Rohdaten als räumliche Daten erhoben werden. Die Speicherung und Bearbeitung der Daten erfolgt in Geographischen Informationssystemen (GIS) (vgl. LINSEISEN, 2003, S. 35ff, ROTHMUND UND AUERNHAMMER 2004, S. 3ff). Die Aufgabe der Datenvorverarbeitung und Datenaufbereitung ist es, diese Rohdaten zu analysieren und daraus Instanzen für die Data-Mining-Phase zu bilden (vgl. Abschnitt 4.3.). Die Datenaufbereitung ist dabei wiederum abhängig von der Versuchsanlage.

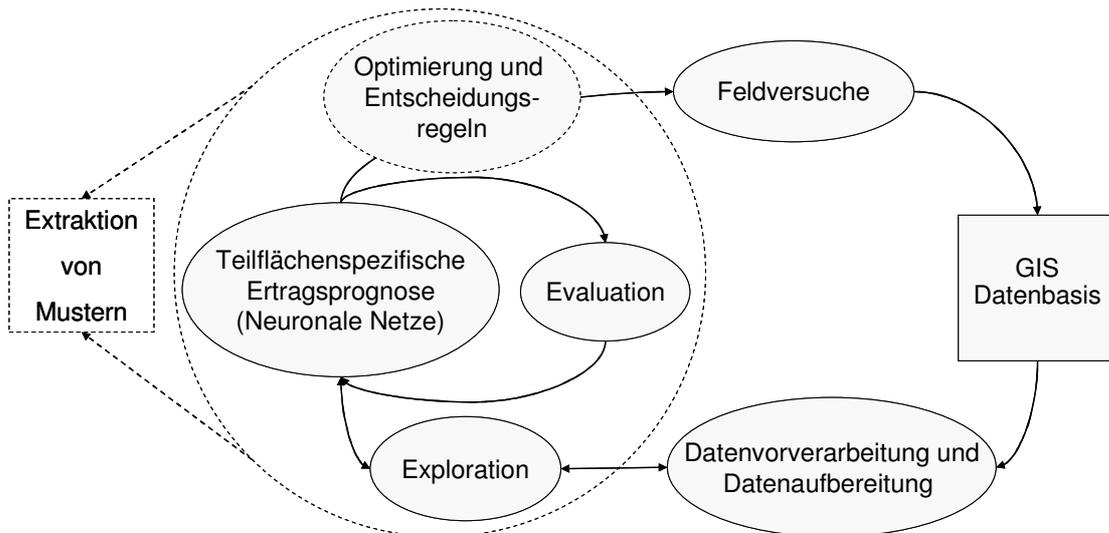


Abbildung 17: Modifizierter WED-Prozess zur Entwicklung von teilflächenspezifischen Entscheidungsregeln

Die Data-Mining-Phase orientiert sich an einem Konzept, das in Abschnitt 4.4 ausführlich dargestellt wird. Das Hauptelement dieser Phase ist die Modellierung der teilflächenspezifischen Ertragsprognose mittels neuronaler Netze. Mit den trainierten neuronalen Netzen kann daraufhin eine ökonomische Optimierung erfolgen. In einem optionalen Schritt können Teilflächen mit ex-post-optimierten N-Applikationen als Input für Entscheidungsbaumverfahren verwendet werden, um Entscheidungsregeln zu generieren. Indem diese Regeln in einen In-Field-Controller eingebaut werden, kann eine automatisierte Umsetzung einer teilflächenspezifischen N-Applikation ermöglicht werden (vgl. OSTERMEIER UND AUERNHAMMER, 2004, s.p.).

Die Überprüfung dieser Regeln bzw. der trainierten neuronalen Netze erfolgt wiederum in Feldversuchen, wodurch eine Verifikation angestrebt wird (vgl. Kap. 4.8). Gleichzeitig dient dieser Versuch wiederum der Datengenese. Das bedeutet, dass der abgebildete Prozess für jedes Versuchsjahr wiederholt wird und die Modelle mit den aggregierten Daten aus den bisherigen Versuchsjahren trainiert werden können.

Obwohl dieses idealtypische Vorgehen durch notwendige Lernprozesse und den damit verbundenen Anpassungen der Versuchsdurchführung nicht vollständig umgesetzt werden

konnte, dient es dennoch als Leitbild für den WED-Prozess, der in den folgenden Abschnitten beschrieben wird. Um einzelne Ergebnisse vorwegzunehmen: Es zeigte sich, dass die Aggregation von Daten aus verschiedenen Versuchsjahren bei vorliegendem Datenmaterial nicht sinnvoll ist. Deshalb wurden die einzelnen Versuchsdaten separat betrachtet und jeweils in einem abgeschlossenen WED-Prozess analysiert.

Weiterhin zeigte es sich, dass die Anwendung von Entscheidungsbaumverfahren in der hier vorgestellten Methode zu sehr komplexen Wissensdarstellungen führt. Die produzierten Entscheidungsregeln in der optimalen Komplexität eignen sich somit nicht, um daraus über die Versuchsjahre hinweg stabile Muster zu identifizieren. Deshalb wurden in der Data-Mining-Phase neben „gekürzten“ Entscheidungsregeln alternative Verfahren zur Extraktion von Mustern benutzt (vgl. Kap. 4.4). Die einzelnen Data-Mining-Phasen werden dazu ausführlich in den Abschnitten 4.5 bis 4.7 dargestellt.

Das Ziel der folgenden Abschnitte ist es in erster Linie, Erfahrungen bzw. Erkenntnisse in der Anwendung der Methode WED in allen Elementen dieses Prozesses zu gewinnen. Darüber hinaus wird versucht, spezifisches Wissen für eine teilflächenspezifische Stickstoff-Applikation nach ökonomischen Gesichtspunkten zu finden.

4.2 Material: Versuchsanlagen und Versuchsdurchführung

Mit der Anwendung des Modells des kontinuierlichen Lernens müssen Feldversuche, wie beschrieben, zweierlei Anforderungen genügen: Sie müssen einerseits die Instanzen zur Modellierung liefern, andererseits müssen mit diesen Versuchen die gefundenen Regeln verifiziert werden können.

In diesem Abschnitt werden Versuche dargestellt, die zu diesem Zweck angelegt wurden. Sie bilden die Datenbasis für die späteren Modellierungen. Deshalb ist ein genauerer Blick auf die Versuche notwendig. Dazu gehören die klimatischen Bedingungen der einzelnen Versuchsjahre, die bodenkundliche Beschreibung der Versuchsstandorte sowie die Beschreibung der Versuchsanlage und –durchführung. Damit soll ein Kontext aufgebaut werden, der die Interpretation der Ergebnisse erleichtert.

Das Datenmaterial besteht aus fünf Einzelversuchen, deren Versuchsanlage den jeweiligen Gegebenheiten angepasst wurde. Drei der fünf Versuche wurden auf den Versuchsgütern der TU München in Freising im Rahmen des IKB-Projektes durchgeführt. Zwei Versuche wurden in Sachsen-Anhalt auf einer Versuchsfläche der Martin-Luther-Universität Halle durchgeführt. In Tabelle 4 ist ein Überblick der Versuche dargestellt.

Die betrachtete Kultur ist jeweils Winterweizen. Es wurden Stickstoff-Düngesysteme mit drei Einzelgaben verglichen, die zu den gleichen Entwicklungsstadien des Pflanzenbestandes appliziert wurden. Mit den Versuchen wurde jeweils die ökonomische Vorteilhaftigkeit

teilflächenspezifischer Düngesysteme im Vergleich zu schlageinheitlichen Düngesystemen geprüft (vgl. GANDORFER, 2005, S. 84ff). Dazu wurden ein Streifenversuchsdesign angelegt (*Hard-Edge-Design*, vgl. PRINGLE ET AL., 2004a, S. 619f).

Tabelle 4: Übersicht der durchgeführten Feldversuche

Bezeichnung	Schlag	Jahr	Versuchsstation	Universität	Versuchsansteller
<i>Schafhof 02</i>	S 04	2002	Dürnast	TU München	IKB-Projekt
<i>D4 03</i>	D 4	2003	Dürnast	TU München	IKB-Projekt
<i>Schafhof 04</i>	S 04	2004	Dürnast	TU München	IKB-Projekt
<i>Görzig 04</i>	550	2004	Domäne Görzig	MLU Halle	Professur für landwirtschaftliche Betriebslehre, Prof. Dr. P. Wagner
<i>Görzig 05</i>	350	2005	Domäne Görzig	MLU Halle	Professur für landwirtschaftliche Betriebslehre, Prof. Dr. P. Wagner

Die Versuche an der TU München-Weihenstephan unterscheiden sich von den Versuchen an der MLU Halle in Bezug auf Versuchsanlage und -durchführung wesentlich. Bei den Versuchen in Freising wurden die Schläge in Parzellen aufgeteilt, die entlang der Fahrgassen ausgerichtet und damit exakt ansteuerbar waren. Für die Versuche in Görzig wurden Standard-Algorithmen zur Erzeugung von Rasterparzellen verwendet (vgl. DRUMMOND ET AL., 2002, s.p. bzw. Kap. 4.2.2). Weiterhin konnte zur Versuchsdurchführung am Standort Weihenstephan auf Geräte der Versuchsstation Dürnast zurückgegriffen werden, während für den Versuch *Görzig 04* ausschließlich ein praxisüblicher Maschinenpark zur Verfügung stand.

Der erste Versuch fand im Jahr 2002 auf dem Schlag *Schafhof* statt. Hier wurden verschiedene Algorithmen aus der ersten Phase des IKB Projektes getestet. Im Rahmen dieser Arbeit dient dieser Versuch ausschließlich zur Erzeugung von Trainings-Instanzen. Der IKB-Versuch des Jahres 2003 fand auf dem Schlag D4 statt. In diesem Versuch wurden wiederum Trainings-Instanzen erzeugt, aber zugleich bereits die Anwendung erster Entscheidungsregeln aus den Daten des Vorjahres überprüft. Der letzte Versuch des IKB-Projektes wurde im Jahr 2004 wiederum auf den Schlag *Schafhof* durchgeführt. Für diesen Versuch stand die Datenbasis aus den Jahren 2002 und 2003 zur Verfügung, anhand derer die Entscheidungsregeln entwickelt wurden. Im selben Jahr wurde der Versuch *Görzig 04* durchgeführt, der im Rahmen dieser Arbeit nur zur Datenerfassung dient. Die daraus entwickelten Entscheidungsregeln wurden im Versuch *Görzig 05* im darauf folgenden Jahr getestet.

Im Laufe der Versuchsserie wurden einige konzeptionelle Details in der Versuchsanlage und Versuchsdurchführung verändert. Insbesondere zeigte sich nach den ersten Data-Mining-Ergebnissen des Versuches *Schafhof 02* (vgl. Kap. 4.5), dass in die Versuchsanlage Elemente eines Stickstoff-Steigerungsversuches eingebaut werden müssen. Die Umsetzung dieser Anforderung wird in den folgenden Abschnitten dargestellt.

4.2.1 Versuchsanlagen und Versuchsdurchführung TU München-Weihenstephan

Im Versuchsjahr 2002 wurde die Winterweizensorte Pegassos (A-Weizen) verwendet. Für die Versuchsjahre 2003 und 2004 wurde im Zuge des Saatgutwechsels des Versuchsbetriebes auf die Sorte Ludwig (A-Weizen) umgestellt. Die Vorfrucht des Versuches *Schafhof 02* und *Schafhof 04* war jeweils Silomais, die Vorfrucht des Versuches *D4 03* war Sommeraps. Die Bodenbearbeitung, Saat sowie die Pflanzenschutzanwendungen erfolgten in allen Versuchen standardmäßig und betriebsüblich.

In den Versuchen an der TU München-Weihenstephan wurde als Applikationsgerät ein Pneumatikdüngerstreuer der Firma Rauch mit einer Arbeitsbreite von 15 m verwendet. Dazu wurden bei der Saat Fahrgassen im Abstand von 15 m angelegt. Mit dem Streuer können Teilbreiten à 7.5 m (d. h. links und rechts von der Fahrgasse) separat angesteuert werden. Deshalb können pro Fahrgasse zwei Streifen angelegt werden. Die *Streifen* verlaufen also entlang der Bearbeitungsrichtung und können von den Fahrgassen aus exakt angesteuert, d. h. mit einer individuellen Stickstoffmenge versorgt werden. Als Streifenbreite ergeben sich damit 7.5 m (vgl. Abbildung 18).

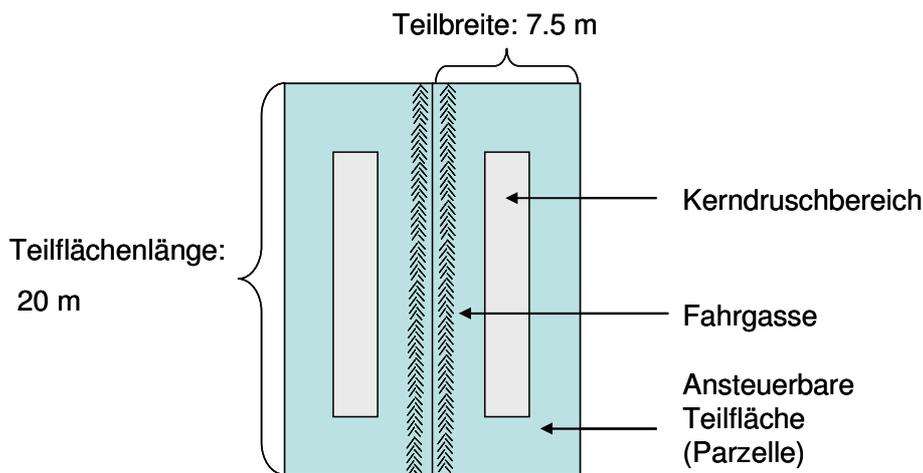


Abbildung 18: Teilflächen in der IKB-Versuchsanlage

Zur Erzeugung von Teilflächen wurden alle Streifen entlang der Fahrgasse in Teilflächen (bzw. Großparzellen) aufgeteilt. Es wurde versucht, diese Flächen sehr klein zu halten, um kleinräumige Heterogenitäten erfassen zu können. Gleichzeitig wird eine gewisse Länge benötigt um Fehler in der Ausbringungsmenge und in der Ertragerfassung zu vermeiden (vgl. COLVIN ET AL., 2000, s.p.). Die korrekte Ausbringungsmenge ist nur für einen Kernbereich der Teilfläche gesichert. In den IKB-Versuchen wurde als Kompromiss dieser Anforderungen die Länge von 20 m gewählt, wobei der Kernbereich (ca. 12 m) mit einem Parzellenmähdrescher abgeerntet wurde. Eine einzelne Teilfläche hat damit eine Fläche von 150 m². Jede dieser Flächen wird im weiteren Verlauf dieser Arbeit als Instanz für ein überwachtetes Lernen verwendet.

Die einzelnen Versuche unterscheiden sich in den getesteten Varianten, in der Größe, sowie in der Versuchsanlage. Im Hinblick auf die spätere Modellierungsphase sind insbesondere die getesteten Düngesysteme von Bedeutung. Durch diese wird eine Variation der Applikationsmengen erzeugt, die für den Lernerfolg in der Data-Mining-Phase von entscheidender Bedeutung ist.

Schafhof 02

Der Versuchsaufbau des ersten Versuches *Schafhof 02* entspricht einem einfaktoriellen Streifenversuch. Dabei wurden 4 Varianten bzw. Faktorstufen getestet. Diese vier verschiedenen Varianten wurden zufällig in einem Block aus vier Streifen gelegt. Der Versuch besteht aus 9 Blöcken mit je 4 Streifen. Die Versuchsanlage ist in Abbildung 19 dargestellt. Es wurden insgesamt 36 Streifen mit zusammen 412 Teilflächen angelegt.

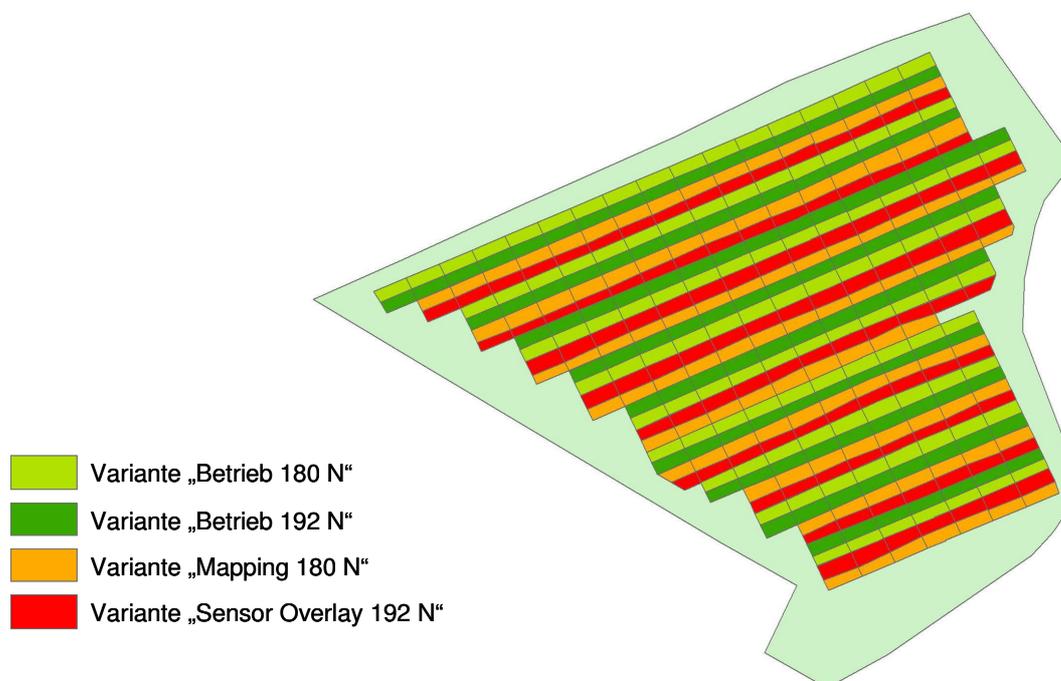


Abbildung 19: Versuchsanlage *Schafhof 02*

Die vier Varianten setzten sich aus zwei schlageinheitlichen Varianten (*Betrieb 180 N*, *Betrieb 192 N*) und zwei Precision-Farming-Varianten (*Mapping 180 N*, *Sensor-Overlay 192 N*) zusammen. In den beiden Precision-Farming-Varianten wurde den Teilflächen eine Stickstoff-Menge appliziert, die sich aus den teilflächenspezifischen Verhältnissen ergibt.

Zur ersten Teilapplikation wurde in allen Varianten einheitlich mit 70 kg N/ha ange düngt. Zur zweiten N-Gabe wurde in den schlageinheitlichen Varianten jeweils 50 kg N/ha ausgebracht. In den beiden Precision-Farming-Varianten wurde ein Mapping-Ansatz angewendet. Dazu wurde der Schlag in drei Ertragszonen aufgeteilt (Niedrig, Mittel, Hoch). Die Einteilung der Zonen erfolgte in Orientierung an die Durchschnittserträge und Maximalerträge von Winterweizen der Jahre 1996 und 1998 und der Einschätzung von Pflanzenbau-

Experten. Die in diesen Varianten durchschnittlich applizierte Menge entspricht den schlageinheitlichen Varianten.

Zur dritten Gabe erfolgte in den schlageinheitlichen Varianten eine Applikation von 60 kg N/ha bzw. 72 kg N/ha. Die Gaben der Variante *Mapping 180 N* richteten sich wiederum nach den eingeteilten Ertragszonen. Zur Bestimmung der Applikationen der Variante *Sensor-Overlay 192 N* wurde ein Algorithmus verwendet, der neben den historischen Ertragsdaten auch den aktuellen Versorgungsstand aus Reflektionsmessungen berücksichtigt (vgl. Tabelle A-28, im Anhang). Insgesamt wurde in dieser Variante durchschnittlich 180 kg N/ha appliziert, während in der Variante *Sensor-Overlay 192 N* entsprechend der schlageinheitlichen Variante *Betrieb 192 N* durchschnittlich 192 kg N/ha appliziert wurden. Eine Übersicht der Einzelgaben findet sich in Tabelle 5.

Tabelle 5: Düngesysteme und Faktorstufen des Versuches Schafhof 02

Varianten Ertragszone	N-Düngung [kg/ha]							
	insgesamt	VB Gesamt	EC32			EC49		
			Niedrig	Mittel	Hoch	Niedrig	Mittel	Hoch
Betrieb 180	180	70	50			60		
Betrieb 192	192	70	50			72		
Mapping 180	180	70	40	50	60	50	60	70
Sensor-Overlay 192	192	70	40	50	60	*	*	*

* N-Gabe wird nach einem Düngealgorithmus durchgeführt; sie variiert je nach Standort und Reflexionsmessung zwischen 60 und 90 N [kg/ha]

Es zeigt sich, dass sich die einzelnen Varianten in Bezug auf die Stickstoffapplikationen nur marginal unterscheiden: Zur ersten Teilgabe ergibt sich keine Differenzierung, zur zweiten Teilgabe entsteht eine maximale Differenz von 20 kg N/ha. Dies führt zu einer stark eingeschränkten Abdeckung des Merkmalsraumes der N-Gaben (vgl. Kap. 2.4.2). Gleichzeitig wird durch die Wahl der Varianten eine Verzerrung des Merkmalraumes erzeugt (vgl. Kap. 4.5.2). Beispielsweise wird auf Teilflächen mit relativ niedrigen historischen Ertragsdaten überwiegend relativ wenig Stickstoff appliziert. Es mangelt daher an Instanzen, die den umgekehrten Fall abbilden. Damit können die kleinräumigen Interaktionen der N-Applikationen mit diesen Daten nur unzureichend abgebildet werden. Aus diesen Überlegungen heraus wurden die folgenden Versuche modifiziert und die Variation der N-Applikationen in einer separaten Variante künstlich erzeugt.

Feldversuch D4 2003

Diese Anforderung wurde erstmals im Versuch *D4 03* umgesetzt. In diesem Jahr konnte ein vollständig randomisierter Streifenversuch realisiert werden. Im Gegensatz zu dem Versuch *Schafhof 02* wurde darauf verzichtet, gleiche durchschnittliche N-Mengen zu erzeugen. Die Varianten unterscheiden sich also in der Gesamtmenge und deren Verteilung. Die zweige-

teilte Versuchsanlage ist in Abbildung 20 dargestellt. Der eine Teil wurde am nördlichen Ende des Schlages platziert, der andere am südlichen Teil hangaufwärts. Für den Versuch stand eine relativ kleine Versuchsfläche zur Verfügung. Es ergaben sich 19 Streifen mit insgesamt lediglich 208 Teilflächen.

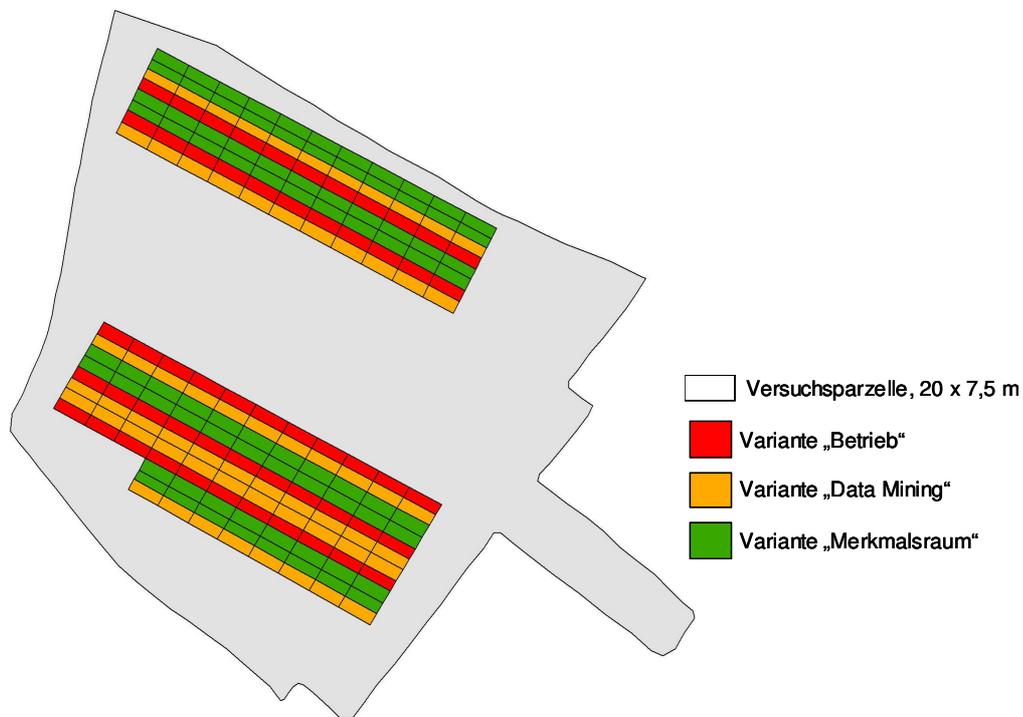


Abbildung 20: Versuchsanlage D4 03

Im Jahr 2003 wurden drei Varianten angelegt. Die Variante *Betrieb*, die als Referenzszenario dient, wurde vom Betriebsleiter des Staatsgutes Dürnast gestaltet. Die entsprechenden Applikationsmengen wurden auf den Streifen der Variante *Betrieb* schlageinheitlich ausgebracht. Mit der Variante *Data Mining* erfolgt eine erste Verifikation von Entscheidungsregeln. Diese Entscheidungsregeln wurden mit dem in dieser Arbeit vorgestellten WED-Prozess erzeugt. Dazu wurden die Instanzen des Versuches *Schafhof 02* in Kombination mit denen eines N-Steigerungsversuches verwendet. Aufgrund der bereits angesprochenen Probleme in der Modellierung dieser Instanzen sowie der Kompatibilitätsprobleme mit dem Datensatz des N-Steigerungsversuches sind diese Regeln kritisch zu sehen (vgl. Kap. 4.5). Insbesondere konnten für die erste N-Applikation keine Regeln erzeugt werden, so dass ersatzweise ein Mapping-Ansatz verwendet wurde. Eine genaue Beschreibung dieses Ansatzes findet sich in WEIGERT ET AL. (2004, S. 321). Aus diesen Gründen muss der erste Versuch zur Verifikation als Testlauf angesehen werden; das ökonomische Ergebnis (vgl. GANDORFER, 2005, S. 84) besitzt dadurch lediglich eine geringe Relevanz.

Schließlich wurde mit einer Variante *Merkmalsraum* das Spektrum möglicher Düngergaben abgedeckt. Dazu wurde für jede Teilflächen innerhalb dieser Variante zufällig eine Appli-

kationsmenge für jede der drei Teilgaben ausgewählt. Die Applikationsmenge konnte zur ersten und zweiten Teilapplikation Werte aus der Menge [0, 10, ..., 120 kg N/ha] sowie zur dritten Teilapplikation Werte aus der Menge [0, 10, ..., 140 kg N/ha] annehmen. Durch die eingeschränkte Anzahl von Teilflächen wurden nur die zufällig erzeugten Verteilungen verwendet, die nicht schon mit den anderen Varianten abgedeckt wurden. Im Wesentlichen entspricht diese Variante also einem N-Steigerungsversuch, bei dem zusätzlich die Verteilung der N-Applikationen verändert wird.

Tabelle 6: Düngesysteme und Faktorstufen des Versuches D4 03

	VB	EC32	EC49	Gesamt (Durchschnitt)
	[kg/ha]	[kg/ha]	[kg/ha]	[kg/ha]
Betrieb*	70*	40*	40*	150
Data Mining**	60-80**	70-90**	40-60**	203
Merkmalsraum	0-120	0-120	0-140	157

* basierend auf der Entscheidung des Betriebsleiters des Staatsgutes Dürnast

**basierend auf einer teilflächenspezifischen Entscheidungsregel nach WEIGERT ET AL. (2004, S. 321)

Eine Übersicht der Teilapplikationen der einzelnen Varianten ist in Tabelle 6 zusammengestellt. Die Variation der Teilgaben erhöht sich in diesem Versuch beträchtlich. Dadurch wird die Ausgangslage zur Modellierung verbessert.

Schafhof 04

Zum letzten der drei IKB-Versuche wurde das Design der Versuchsanlage nochmals verändert. Der Unterschied besteht darin, dass der Schlag nicht mehr streifenweise nach einer Düngevariante bewirtschaftet wird, sondern jeder Teilfläche zufällig einer Variante zugelost wurde. Es ergibt sich somit um eine vollständig randomisierte Versuchsanlage (vgl. Abbildung 21). Der Vorteil dieser Vorgehensweise ist darin zu sehen, dass jede Parzelle als Wiederholung in die varianzanalytische Auswertung eingehen kann. Im Vergleich zum Streifenversuch *D4 03*, bei dem die Streifen als Wiederholung angesehen werden, erhöht sich die Anzahl der Wiederholungen beträchtlich. Dadurch können auch kleinere Mittelwertsdifferenzen gegebenenfalls statistisch abgesichert werden können. Ein weiterer Vorteil ist darin zu sehen, dass der versuchsbedingte Fehler über eine Kovarianzanalyse kontrolliert werden kann (vgl. GOMEZ UND GOMEZ, 1984, S. 424ff).

In diesem Jahr wurden vier Varianten angelegt. Die Variante *Merkmalsraum* wurde wie im Versuch *D4 03* zufällig erzeugt. Die obere Grenze der dritten Teilgabe wurde allerdings bei 120 kg N/ha gesetzt. Weiterhin wurde ein Mapping-Ansatz mit fünf Ertragsklassen getestet (Variante *Mapping*).

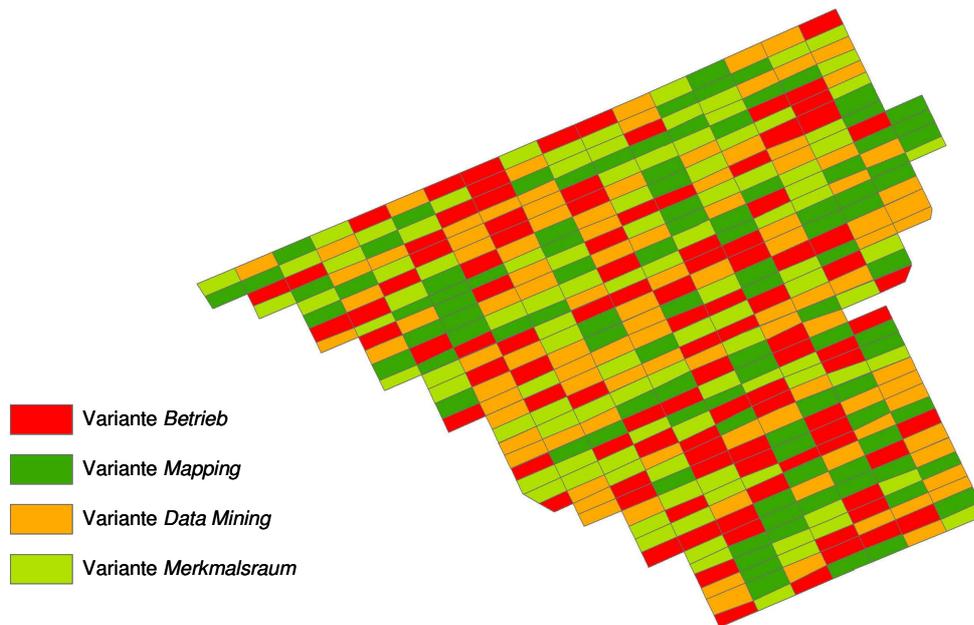


Abbildung 21: Versuchsanlage Schafhof 04

In der Variante *Data Mining* wurden die Entscheidungsregeln eines WED-Prozesses verwendet, in dem die Kombination der Daten des Versuches *Schafhof 02* und *D4 03* als Input dienen. Eine ausführliche Darstellung der Entwicklung dieser Regeln findet sich in WEIGERT UND WAGNER (2004, s.p.). Wiederum müssen diese Ergebnisse als suboptimal bezeichnet werden, da der eingeschränkte Merkmalsraum des Versuches *Schafhof 02* zu Überlernen führt.

In der schlageinheitlichen Variante *Betrieb* sowie in der Variante *Mapping* wurden die durchschnittlichen Düngermengen aus der Variante „Data Mining“ abgeleitet. Die sich dadurch ergebenden Teilapplikationen sind in Tabelle 7 abgetragen.

Tabelle 7: Düngesysteme und Faktorstufen des Versuches Schafhof 04

Düngesystem	Stickstoff-Düngergaben in [kg/ha]			
	VB ¹	EC32	EC49	Schlagdurchschnitt
Betrieb	40	50	50	140
Mapping	20-60	30-70	30-70	140
Data Mining	20-70	0-70	20-100	140
Merkmalsraum	0-120	0-120	0-120	170

Sowohl in Bezug auf die Verwendung der Teilflächendaten zu einer Ertragsprognose mit Hilfe neuronaler Netze als im Hinblick auf den Variantenvergleich kann diese Versuchsanstellung als optimal angesehen werden. Einerseits führt die vollständige Randomisierung zu umfassenden Möglichkeiten im Sinne der Verifikation, andererseits führt die Randomisierung auch zu einem reichhaltigeren und verzerrungsfreien Merkmalsraum im Sinne des Data-Mining-Ziels. Durch die Anlage von Streifen (wie in den vorausgegangenen Versu-

chen) befinden sich die Teilflächen eines Düngesystems immer in einer systematischen Anordnung. Dadurch werden immer systematisch zusammenhängende Ausschnitte der Bodenheterogenität betrachtet. Falls sich das Düngesystem an Bodenattributen orientiert, ergeben sich Gruppierungen von ähnlichen Merkmalsausprägungen der Attribute „Boden“ und „N-Input“. Dies wurde in diesem Versuch vermieden. Gleichzeitig konnten in diesem Versuch ausreichend viele Teilflächen (103) dieser Variante angelegt werden.

Jedoch zeigte sich in der Auswertung, dass ein Fehler in der Versuchsdurchführung des Vorjahres die Ergebnisse stark beeinflusst⁴⁵. Stark negative Stickstoffbilanzsalden deuten darauf hin, dass durch das Erntegut wesentlich mehr Stickstoff entzogen als durch mineralische Stickstoffdüngung zugeführt wurde. Dies bedeutet, dass entweder sehr viel Stickstoff im Boden durch Mineralisation freigesetzt wurde, oder, was in diesem Fall wahrscheinlicher ist, dass sehr hohe, versuchsbedingte Stickstoffgaben (eines stabilisierten Stickstoffdüngers) im Vorjahr, die aufgrund sehr geringer Niederschläge in der vegetationsfreien Zeit zudem nicht ausgewaschen wurden, zu einem Stickstoffüberangebot geführt haben. Der Versuch kann demnach **nicht** zur Entwicklung von Entscheidungsregeln verwendet werden; auf die Darstellung der Data-Mining-Phase wird daher in dieser Arbeit verzichtet.

4.2.2 Versuchsanlage und Versuchsdurchführung Görzig

Zusätzlich zu den vorher betrachteten Versuchen wurden zwei Versuche im Raum Sachsen-Anhalt durchgeführt. Ausgewählt wurde der Schlag 550 und 350 des Lehr- und Versuchsgutes „Domäne Görzig“ der Landwirtschaftlichen Fakultät der Martin-Luther-Universität Halle. Die Domäne liegt ca. 10 km südlich von Köthen in Sachsen-Anhalt. Die Versuche wurden von der Professur für Landwirtschaftliche Betriebslehre der Martin-Luther-Universität Halle-Wittenberg (Prof. Dr. P. Wagner) angelegt. Der Versuch des Jahres 2004 wurde im Rahmen dieser Arbeit ausschließlich zur Entwicklung von Entscheidungsregeln verwendet (vgl. Kap. 4.7), während der Versuch des Jahres 2005 ausschließlich zur Verifikation, d. h. zum Test dieser Entscheidungsregeln genutzt wurde (vgl. Kap. 4.8.1).

Görzig 04

Der Versuchsschlag 550 hat eine Größe von ca. 65 ha und übersteigt somit die Größe der Versuchsanlagen im Rahmen des IKB-Projekts um ein Vielfaches. Die betrachtete Fruchtart ist Winterweizen der Sorte Aron (Elitequalität, E-Weizen). Im vorausgegangenen Jahr wurde Körnermais geerntet. Die Vorvorfrucht war wiederum Winterweizen.

Die teilflächenspezifische Düngung sowie die Sensormessungen wurden von der Firma Agri Con® erledigt, während die restlichen Bewirtschaftungsmaßnahmen betriebsüblich

⁴⁵ Es wurde ein stabilisierter Stickstoffdünger in der falschen Dosierung ausgebracht, was zu sehr hohen N-Applikationen führte (bis zu 300 N [kg/ha]). Da in diesem Versuch nach Ertragszonen gedüngt wurde, ergeben sich zudem unterschiedliche Stickstoffüberhänge.

vom Betriebsverbund „Hof Pfaffendorf“ durchgeführt wurden. Die eingesetzten Maschinen unterscheiden sich von den IKB-Versuchen in mehreren Punkten: Zum einen wurde zur Stickstoff-Applikation ein Schleuderstreuer verwendet und mit einer Fahrgassenbreite von 24 m gearbeitet. Weiterhin erfolgte die Ertragsmessung im Versuchsjahr sowie im Vorjahr mit einer automatischen Ertragsaufzeichnung (Mähdrescher: *Claas-Lexion*, Ertragsmesssystem: *Quantimeter*).

Bei diesem Versuch wurden drei verschiedene Bewirtschaftungsvarianten (*Betrieb*, *Sensor*, *Mapping*) in jeweils drei Streifen getestet. Ein Streifen besteht aus drei Fahrgassen mit einer Breite von jeweils 24 m. Auf zwei Streifen wurde entsprechend zu den Varianten Merkmalsraum bei den IKB-Versuchen eine Variante *N-Steigerung* angelegt. Die Versuchsanlage ist in Abbildung 22 dargestellt. Mit der Versuchsanlage selbst wurden keine Teilflächen festgelegt; die Teilflächen ergeben sich erst durch die verwendete Methode der Datenaufbereitung (vgl. Kap. 4.3.2).

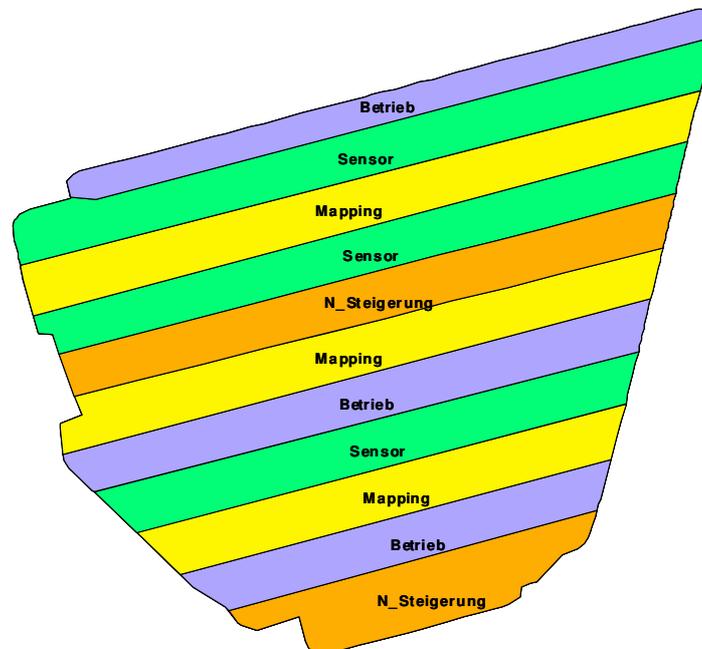


Abbildung 22: Versuchsanlage Görzig 04

Zur ersten N-Gabe bei Vegetationsbeginn wurde in allen Varianten einheitlich 60 kg N/ha appliziert. Die Bemessung der Teilapplikationen zu EC 32 und EC 49 erfolgte entsprechend der Ansätze. In der Variante *Mapping* wurde eine Düngung nach Entzug umgesetzt. Das Ertragspotenzial leitet sich aus dem teilflächenspezifischen Vorjahresertrag (Feuchtertrag in dt/ha) ab. Die errechnete N-Gesamtmenge wurde abzüglich der Menge der ersten N-Gabe zu gleichen Teilen auf die folgenden zwei Applikationen aufgeteilt⁴⁶.

In der Variante *Sensor* wurden die Teilflächen entsprechend des Algorithmus bestimmt,

⁴⁶ Damit bestimmt sich die Höhe der N-Applikation in N [kg/ha] für die 2. oder 3. N-Gabe nach folgender Formel: $N_2 = N_3 = [(Ertrag_{03} * 2) - 60] / 2$.

wie er von der Firma Agri Con® standardmäßig vorgenommen wird: als Bemessungsgrundlage dienen jeweils Sensormessungen mit dem Yara® N-Sensor® und der Einsatz eines N-Testers. Zur genaueren Beschreibung dieses Verfahrens sei auf LINK ET AL. (2002, s.p.) verwiesen.

Das Referenzszenario wurde mit einer schlageinheitlichen betriebsüblichen Düngung (Variante *Betrieb*) definiert, in der zur ersten N-Applikation 60 kg N/ha, sowie zur zweiten und dritten N-Gabe jeweils 40 kg N/ha verabreicht wurden.

Gleichzeitig wurde auf diesem Schlag in zwei Streifen eine Variante zur Abdeckung des Merkmalsraumes in Form eines N-Steigerungsversuches angelegt. Durch den Einsatz von Schleuderstreuern und automatischer Ertrags erfassung mussten die „Parzellen“ relativ groß gewählt werden, damit zumindest im Kernbereich der Input- bzw. Output exakt erfasst werden konnte. Dazu wurden die Streifen in 14 bzw. 7 Teilflächen⁴⁷ aufgeteilt. Die einzelnen Teilflächen waren damit ca. 65 m lang und 72 m breit. Jeder dieser 21 Teilflächen wurde eine der in Tabelle 8 dargestellten Düng-Varianten zugewiesen.

Tabelle 8: Faktorstufen der Variante *Merkmalsraum* (Versuch *Görzig 04*)

Teilfläche	N_1	N_2	N_3	N-Gesamtmenge
	kg N/ha	kg N/ha	kg N/ha	kg N/ha
1	0	0	0	0
2	0	0	0	0
3	0	0	0	0
4	0	70	0	70
5	0	70	0	70
6	70	0	0	70
7	70	0	0	70
8	0	0	70	70
9	0	0	70	70
10	30	30	30	90
11	30	30	30	90
12	40	40	40	120
13	40	40	40	120
14	50	50	50	150
15	50	50	50	150
16	80	80	80	240
17	80	80	80	240
18	90	90	90	270
19	90	90	90	270
20	100	100	100	300
21	100	100	100	300

⁴⁷ Der Begriff Teilfläche ist hier etwas missverständlich: In jedem dieser Bereich liegen mehrere Teilflächen, aus denen die Instanzen erzeugt werden.

Die Ausgestaltung dieser Variante orientiert sich an N-Steigerungsversuchen, wobei die N-Gesamt mengen, die von den anderen Varianten abgedeckt werden, also der Bereich zwischen 150 und 240 kg N/ha, ausgelassen wurden. Mit sechs Teilflächen (Teilflächen 4-9) wurde versucht, die Verteilung der Gesamtmenge auf die Teil-Applikationen zu variieren. Dies konnte allerdings nur ansatzweise erreicht werden (vgl. Kap. 4.7). Durch die eingeschränkte Anzahl von ansteuerbaren Teilflächen wurde auf eine zufällige Auswahl der Teil-Applikationen (vgl. IKB-Versuch) verzichtet. Dies führt dazu, dass die Korrelation der Teil-Applikationen hoch und die Abstufung der Mengen sehr grob ist.

Wie sich aus den Teil-Applikationen der einzelnen Varianten in Tabelle 9 ergibt, wurde allerdings auch mit den Precision-Farming-Düngesystemen zur zweiten und dritten N-Applikation eine beträchtliche Varianz erzeugt. Die Anwendung des Yara® N-Sensors® führt beispielsweise zu Applikationen von 23-85 kg N/ha zur zweiten und dritten N-Gabe.

Tabelle 9: Düngesysteme und Faktorstufen des Versuches Görzig 04

Düngesystem	Stickstoff-Düngergaben in [kg/ha]			
	VB	EC32	EC49	Schlagdurchschnitt
Betriebsüblich	60	40	40	140
Mapping	60	0-78	0-78	125
Sensor	60	23-85	23-63	147
Merkmalsraum	0-100	0-100	0-100	170

Mit diesem Versuch ergeben sich neue Voraussetzungen für die Datenanalyse. Der Vorteil dieses Versuches ist sicherlich in der Größe des Schlags und der damit verbundenen Variation der Standortverhältnisse zu sehen. Andererseits ist - ausgehend von den IKB-Versuchen - durch den Einsatz von Maschinen aus der Praxis mit Einbußen in der Präzision der Applikation und Datenerfassung zu rechnen.

Görzig 05

Der Versuch *Görzig 05* wird im Rahmen dieser Arbeit ausschließlich zur Untersuchung der ökonomischen Vorteilhaftigkeit der Entscheidungsregeln aus dem Versuch *Görzig 04* (vgl. Kap. 4.7) verwendet. Der Feldversuch wurde im Anbaujahr 2004/05 auf dem Schlag 350 (63 Hektar) der Domäne Görzig, angelegt. Im Versuchszeitraum wurde die Winterweizensorte *Compliment* (A-Weizensorte) angebaut (Vorfrucht: Winterraps). In einem Streifendesign wurden die vier Versuchsglieder auf dem Schlag angelegt. Jedes Glied wurde in zwei Streifen wiederholt. Die Streifen hatten jeweils eine Breite von zwei bzw. drei Fahrspuren bei einer Arbeitsbreite der Düngetechnik von 24 Metern (vgl. *Görzig 04*). Die Versuchsanlage ist in Abbildung 23 illustriert.

Auf eine zufällige Verteilung der Wiederholungen im Feld wurde verzichtet. Dieses Vor-

gehen findet seine Begründung im begrenzten Platzumfang des Schlages und der damit verbundenen niedrigen Anzahl an möglichen Wiederholungen. Für jeden Streifen wurde das arithmetische Mittel der räumlich interpolierten Bodenleitfähigkeitsmessung und des Ertrages 2003 bestimmt. Die Aufteilung der Wiederholungen auf die Streifen erfolgte so, dass jedem Versuchsglied annähernd der gleiche durchschnittliche Wert dieser Heterogenitätsindikatoren zukam. Durch diese Vorgehensweise kann über einen Vergleich der Ertragsergebnisse (Mittelwerte) eine näherungsweise Bewertung der einzelnen Varianten ermöglicht werden, ohne dass die Standortvoraussetzungen berücksichtigt werden müssen (vgl. Abschnitt 4.8.1).

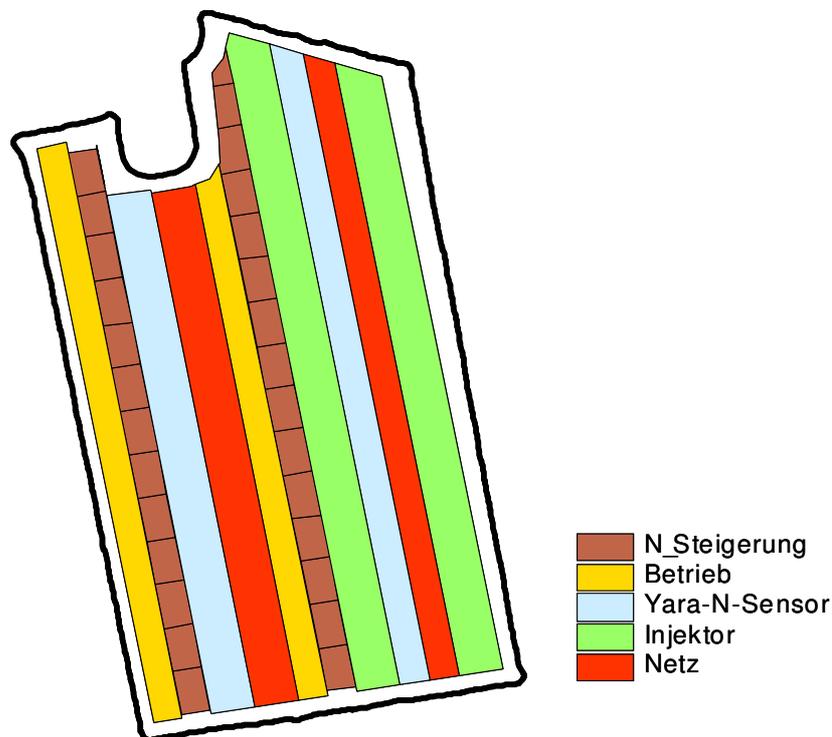


Abbildung 23: Versuchsanlage Görzig 05

Es wurden vier Düngestrategien (*Betrieb*, *Sensor*, *Injektor* und *Netz*) verglichen. Eine ausführliche Beschreibung dieses Versuches findet sich in WAGNER ET AL. (2005). Der Einfachheit halber beschränkt sich die folgende Darstellung auf die drei Varianten, in denen ein Düngesystem mit drei Teilapplikationen verwendet wurde. Die Variante *Merkmalsraum* sowie die Variante *Injektor* werden im Rahmen dieser Arbeit nicht betrachtet. In Tabelle 10 sind die Applikationsmengen der einzelnen Varianten zusammengefasst.

Die Teilmengen der Variante *Betrieb* orientieren sich an den Vorgaben des Pflanzenbauleiter des Betriebsverbundes „Hof Pfaffendorf“ und liegen bei 55 kg N/ha für die erste N-Applikation und jeweils 60 kg N/ha für die zweite und dritte Applikation.

In der Variante *Sensor* wurde zur Bemessung der zweiten und dritten N-Applikation der Yara® N-Sensor® verwendet (vgl. Versuch Görzig 04). Die erste Stickstoffgabe wurde

schlageinheitlich, entsprechend der Variante *Betrieb*, durchgeführt. Die beiden weiteren Ausbringungsmengen liegen im Mittelwert bei ca. 60 kg N/ha. Die Streuung ergibt sich durch den verwendeten teilflächenspezifischen Sensor-Ansatz.

Tabelle 10: N-Applikationen der untersuchten Varianten im Versuch Görzig 05

Versuchsglied	1. Gabe		2. Gabe		3. Gabe	
	Mittelwert (kg/ha)	Streuung (kg/ha)	Mittelwert (kg/ha)	Streuung (kg/ha)	Mittelwert (kg/ha)	Streuung (kg/ha)
<i>Betrieb</i>	55	0	60	0	60	0
<i>Sensor</i>	55	0	61	44 – 80	66	53 – 80
<i>Netz</i>	44	30 – 60	47	10 – 60	57	30 – 70

Die Variante *Netz* wurde entsprechend der in dieser Arbeit vorgestellten Methode optimiert. Dazu wurden die trainierten neuronalen Netze aus Kap. 4.7.2, 4.7.3 und 4.7.4 verwendet. Es ergaben sich zur ersten N-Gabe Applikationsmengen innerhalb des Bereiches von 30-60 kg N/ha, mit einer Dominanz der Werte 40 und 50 kg N/ha (vgl. Kap. 4.7.2). Zur zweiten N-Applikation wurden Mengen innerhalb des Bereiches 10-60 kg N/ha und zur dritten Applikation Werte von 30-70 kg N/ha ausgebracht.

4.2.3 Witterung und Pflanzenentwicklung im Untersuchungszeitraum

Eine wichtige Rolle in der Interpretation der Ergebnisse spielen die klimatischen Gegebenheiten der Versuchsjahre (vgl. Kap. 3.1.1). Der Witterungsverlauf für die IKB-Versuche ist anhand der Wetterstation Weihenstephan des Deutschen Wetterdienstes erfasst (vgl. Abbildung A-67 und Abbildung A-68 im Anhang). Im langjährigen Mittel fallen dort 800 mm Niederschlag. Die mittlere Jahresdurchschnittstemperatur beträgt 7,5°C.

Im Vergleich dazu liegen die Jahresniederschläge der Versuchstation Görzig bei 480 mm, also weit unter den Freisinger Verhältnissen. Charakteristisch für das östliche Harzvorland ist weiterhin, dass ein Großteil der Niederschläge als Starkregen während der Vegetationsperiode auftritt. Die mittlere Jahresdurchschnittstemperatur beträgt 9,1°C. Zur Dokumentation des Witterungsverlaufes wurden die Daten der Wetterstation Halle-Seeben verwendet, die ca. 15 km südlich der Versuchstation Görzig liegt (vgl. Abbildung A-70 sowie Abbildung A-70 im Anhang).

Versuchsjahr 2002 Weihenstephan

Am 10. bzw. 11. Oktober 2001 wurde Winterweizen für den Versuch 2002 gesät. Aufgrund starker Niederschläge im März wurde ein Großteil des vorhandenen mineralischen Stickstoffs im Boden vertikal und horizontal verlagert. Im April führten geringe Niederschläge

bei zugleich kalten Temperaturen zu geringen Mineralisationsprozessen und zu einer entsprechend schlechten Entwicklung der Weizenbestände

Die Pflanzen konnten den Entwicklungsrückstand aufgrund guter Vegetationsbedingungen im Mai wieder ausgleichen. Niedrige Niederschlagsmengen im Juni und extrem hohe Tagestemperaturen von bis zu 38°C führten zu einer vorzeitigen und beschleunigten Abreife der Weizenbestände (vgl. SCHÄCHTL, 2004, S. 47).

Versuchsjahr 2003 Weihenstephan

Der nasse Herbst führte in diesem Jahr zu späteren Saatterminen beim Winterweizen und zog eine schwache Jugendentwicklung nach sich. Nach dem Winter entwickelten sich die Weizenbestände gut. Die Stickstoffversorgung aus Düngung und mineralisiertem Bodennstickstoff war meist gut. Extremer Wassermangel in der Kornfüllungsphase ab Mitte Juni verhinderte dann allerdings gute Erträge. Insgesamt sorgten im Jahr 2003 extreme Witterungsverhältnisse bei Winterweizen für niedrige Erträge.

Versuchsjahr 2004 Weihenstephan

Die Aussaat des Winterweizens konnte unter optimalen Bedingungen durchgeführt werden. Der milde November begünstigte eine rasche Jugendentwicklung mit guter Wurzelentwicklung. Aufgrund des sehr trockenen, ertragsarmen Vorjahres waren die Bodennstickstoffgehalte meist relativ hoch. Die ausgebildeten Bestandesdichten lagen auf dem Niveau normaler Jahre. Im Juni waren die Niederschläge durchschnittlich, und der Weizenbestand war sehr gut mit Wasser versorgt. Während der Kornfüllungsphase herrschte eine gemäßigte Witterung ohne große Temperaturspitzen mit ausreichender Bodenfeuchtigkeit. Insgesamt wurden zur Ernte 2004 extrem hohe Kornerträge erzielt.

Versuchsjahr 2004 Görzig

Die Witterungsbedingungen des Jahres 2004 am Standort Görzig waren ähnlich denen zu Weihenstephan. Die Aussaat erfolgte am 15.10.2003. Die Temperaturen verliefen im Einklang zum langjährigen Mittel, sodass gute Bestände aufgebaut werden konnten. In den für den Kornertrag entscheidenden Monaten Mai und Juni stand eine überdurchschnittliche Niederschlagsmenge zur Verfügung.

Versuchsjahr 2005 Görzig

Das Versuchsjahr 2005 zeigt große Ähnlichkeiten zu dem vorausgegangenen Jahr. Die Aussaat konnte aufgrund der Vorfrucht Winterraps bereits am 10.09.2004 durchgeführt werden. Die durchschnittlichen Temperaturen während der Kornbildungsphase erreichten ein etwas höheres Niveau als im Vorjahr. Lediglich die Niederschlagsmenge im Monat Juni zeigte starke Einbußen im Vergleich zum Jahr 2004. Direkt vor der Ernte waren starke Niederschläge zu verzeichnen. Dadurch kam es zu Ertrags- und Qualitätseinbußen.

4.2.4 Standortbeschreibung

Der Versuchsstandort **S04** („Schafhof“, 8,78 ha) liegt im tertiären Hügelland in der Nähe von Freising in Oberbayern. Der Schlag ist gekennzeichnet von einer mäßigen Variabilität. In den Hohertragsbereichen handelt es sich um Braunerde aus Lößlehm, die leicht pseudovergleyt ist. Der Kiesanteil beträgt in diesen Bereichen 0%. Der Bodentyp in den Niedrigertragsbereichen kann als pseudovergleyte Braunerde aus Lößlehm klassifiziert werden. Hier beträgt der Kiesanteil ca. 5%. Die Bodenart erstreckt sich von schwach sandigem Lehm im Hohertragbereich bis hin zu stark sandigem Lehm im Niedrigertragbereich. Die einzelnen Teilflächen weisen unterschiedliche Hangneigungen auf. Die Erhebung liegt im Bereich von 460-483 m über NN.

Beim Versuchsschlag **D4** können aus bodenkundlicher Sicht tendenziell zwei verschiedene Bereiche abgegrenzt werden. In den Tallagen und in einer ausgeprägten Erosionsrinne, in die Erosions-Massen abgetragen wurden, handelt es sich um eine basenreiche, kolluviale Braunerde, die durch den Grundwassereinfluss (Tallagen) zum Teil schwach vergleyt ist. In den Hanglagen findet sich Braunerde mit geringer Basensättigung aus umlagertem sandig-lehmigen Tertiärmaterial. Die Bodenart kann überwiegend als sandiger Lehm charakterisiert werden (Quelle: Bodenkundliche Karte Freising Nord Nr. 7536).

Der beiden benachbarten Standorte **Schlag 550** und **Schlag 350** der Domäne Görzig befinden sich in einer Randlage des Lößgebietes und gehören zum Köthener Ackerland, einem Teil des Magdeburg-Halleschen Schwarzerdegebietes. Dieser Bereich zeichnet sich durch eine vergleichsweise geringe Mächtigkeit der Lößauflage (bis max. 100 cm) und Kies im Untergrund aus. Bodentypologisch handelt es sich um eine Schwarzerde aus Löß. Begleitend treten Übergangssubtypen (u. a. Braunerde und Tschernoseme) auf. Beide Schläge weisen eine sehr geringe Hangneigung auf. Die durchschnittliche Bodenwertzahl der Domäne liegt bei 78. Die Bodencharakteristiken der Domäne Görzig sind in Tabelle A-29 im Anhang zu finden. Für beide Schläge zeigen sich sehr ähnliche Standortvoraussetzungen.

Zusammenfassend weisen die beiden Standorte Freising und Görzig sehr unterschiedliche Voraussetzungen hinsichtlich der natürlichen Voraussetzungen auf. Dies gilt vor allem für die Bodencharakteristiken in Verbindung mit der Witterung.

4.3 Datenquellen, Datenattribute und Datenvorverarbeitung

Dieser Abschnitt stellt im WED-Prozessmodell die Phase der Datenvorverarbeitung und Datenaufbereitung dar. Die Aufgabe der Datenvorverarbeitung ist es, mit den räumlichen Rohdaten der verschiedenen Sensormessungen Instanzen aus den Teilflächen zu erstellen, die in dem Abschnitt „Modellierung“ verwendet werden. Die Aufgabe der Datenaufbereitung dagegen besteht in der Analyse und in der entsprechenden Anpassung der Instanzen

und erfolgt in dieser Arbeit erst vor der entsprechenden Modellierung.

Bei den Datenquellen handelt es sich um die in Kap. 3.1.3 vorgestellten Sensoren. Zusammen mit dem Messwert wird dabei jeweils die exakte räumliche Position erfasst. Die so erhaltenen räumlichen Punktdaten können in einem Geographischen Informationssystem (z. B. ESRI® ArcGIS®) bearbeitet werden.

In Tabelle 11 sind die Attribute (bzw. Messgrößen) und ihre Kurzbezeichnungen zusammengestellt, die im Laufe der Arbeit verwendet werden. Eine Instanz sollte nach der Datenvorverarbeitung jeweils Merkmalsausprägungen zu diesen Attributen besitzen. Bei den (Korn-)Erträgen kennzeichnet eine Jahresangabe im zweistelligen Format das Jahr der Aufzeichnung. Die restlichen Daten sind entweder Bodenattribute, die als zeitlich stabil angenommen werden, oder beziehen sich auf das Versuchsjahr. Dies gilt für die Stickstoffapplikationen und die Vegetationskennzahlen. Die einzelnen Stickstoff-Applikationen finden zu festgelegten Entwicklungsstadien des Winterweizen-Bestandes statt. N_1 wurde jeweils bei Vegetationsbeginn (VB), N_2 bei EC 32 und N_3 bei EC 49 ausgebracht.

Tabelle 11: Das Spektrum der verwendeten teilflächenspezifischen Attribute

<i>Ertrag_jj</i>	Korn-Ertrag (Fruchtart: Winterweizen) des Jahres jj (t/ha)	Bodenattribute
<i>EM_38</i>	Scheinbare elektrische Leitfähigkeit (mS/m) normiert auf 25°C (EC ₂₅)	
<i>Zugkraft</i>	Zugkraft (gemessen bei der Stoppelbearbeitung, (~kN)	
<i>REIP_2</i>	Red Edge Inflection Point, gemessen vor der 2. N Applikation (nm)	In-season-
<i>REIP_3</i>	Red Edge Inflection Point, gemessen vor der 3. N Applikation (nm)	Attribute
<i>N_1</i>	Erste N-Gabe zu Vegetationsbeginn (kg N/ha)	Stickstoff-
<i>N_2</i>	Zweite N-Gabe zu EC 32 (kg N/ha)	Applikationen
<i>N_3</i>	Dritte N-Gabe zu EC 49 (kg N/ha)	
<i>Ertrag_jj</i>	Korn-Ertrag im Versuchsjahr jj (t/ha bei 14% Feuchtigkeit)	Output

Eine erste Aufgabe besteht darin, die erhaltenen Kartierungen zu überprüfen und gegebenenfalls zu korrigieren. Darauf aufbauend, müssen die Daten der verschiedenen Sensoren miteinander kombiniert werden. Dadurch, dass die Punktmessungen für die jeweiligen Sensoren (zufällig) an unterschiedlichen Positionen erfolgen, können die Punktdaten nicht einfach zu Instanzen kombiniert werden. Dies wird in der Regel so gelöst, dass die Schläge in Flächeneinheiten aufgeteilt und für diese Einheiten Messwerte ermittelt werden. Die Teilflächen werden als Instanzen verwendet. Aus den Sensormessungen werden Attributsausprägungen, die diese Instanz beschreiben.

In den IKB-Versuchen sind diese Teilflächen bereits mit der Versuchsanlage definiert. In

den Görziger Versuchen werden diese Teilflächen erst in der Datenaufbereitung (virtuell) in einem GIS-System erzeugt.

In der Data-Mining-Phase werden räumlichen Informationen nicht verwendet. Dies bedeutet, dass die Information „Ort“ und damit alle räumlichen Abhängigkeiten in der Modellierung nicht mehr berücksichtigt werden. Alle Informationen, die sich aus räumlichen Zusammenhängen ergeben, müssen in der Phase der Datenvorverarbeitung benutzt werden. Beispielsweise orientieren sich einige Methoden zur Korrektur von Ertragsdaten an Ertragspunkten in der räumlichen Umgebung. Diese Methoden müssen deshalb vor der Bildung von Instanzen eingesetzt werden.

Im Verlaufe dieser Arbeit wird in zwei verschiedene Datenaufbereitungskonzepte unterschieden, die eng mit der Versuchsanstellung zusammenhängen. In den folgenden Abschnitten werden diese beiden Methoden vorgestellt. Das Ziel ist jeweils die Erzeugung von validen Instanzen zur späteren Modellierung.

4.3.1 Datenvorverarbeitung IKB-Versuche

Die erste Methode der Datenvorverarbeitung wurde entsprechend der Versuchsbedingungen der IKB-Versuche entworfen. Die geringen Schlaggrößen erlauben den Einsatz von (Versuchs-)Geräten mit geringer Arbeitsbreite. Dies ermöglicht eine hohe räumliche Auflösung, die durch eine entsprechende Versuchsdurchführung und Datenvorverarbeitung ausgenutzt werden kann.

Um die Chancen dieser hohen Auflösung zu nutzen, wurde die Anlage der Teilflächen dem Schlag angepasst. Dazu wurden mit Hilfe Geographischer Informationssysteme entlang der Fahrgassen manuell Teilflächen mit einer vordefinierten Größe angelegt. Dadurch, dass die Bearbeitungsrichtung nicht streng geradlinig ist, wird damit auch kein streng rechtwinkliges Raster-Design erreicht, und es ergeben sich kleinere Unterschiede in der tatsächlichen Größe der Parzellen. Die Größe der Parzellen wurde in der Breite so gewählt, dass sie mit den Ausbringungsgeräten exakt angesteuert werden können. Um die Präzision zu erhöhen, wurden Parzellenmähdrescher eingesetzt und nur der Kernbereich einer Teilfläche erfasst. Für diesen Bereich kann auch eine exakte N-Applikation angenommen werden, da die Applikation abgesetzt erfolgte, d. h. jeweils an der Grenze zu einer neuen Teilfläche wird beim Stillstand des Traktors eine neue N-Menge eingestellt. Die Input-Output-Beziehung kann demnach für den Kernbereich der Teilflächen exakt ermittelt werden.

Zur Erzeugung der restlichen teilflächenspezifischen Attribute wurden Messungen durchgeführt bzw. auf Daten der IKB-Datenbank (vgl. LINSEISEN, 2003, S. 35ff, ROTHMUND UND AUERNHAMMER 2004, S. 3ff) zurückgegriffen. Auf dem Staatsgut Dürnast sind beispielsweise alle Traktoren mit einer automatischen mobilen Prozessdatenerfassung ausgestattet. Diese Daten werden in regelmäßigen Abständen auf diese Datenbank übertragen. Dadurch

konnte insbesondere auf gespeicherte Zugkraft- und Ertragsdaten zurückgegriffen werden. Die Aufgabe besteht jeweils darin, aus den räumlichen Messpunkten, die Merkmalsausprägung eines Attributes für eine Teilfläche zu ermitteln.

Zu dieser Aufgabe wird eine Funktion von Geographischen Informationssystemen verwendet, die untersucht, welche Punkte in einer betrachteten Teilfläche liegen (Spatial Join). Von diesen Messwerten wird der Mittelwert gebildet und dieser dann der Teilfläche als Merkmalsausprägung des Attributes zugewiesen. In der Regel liegen von den einzelnen Messungen genügend valide Datenpunkte in einer Teilfläche, dennoch traten grenzwertige Fälle auf, wie in den nachfolgenden Abschnitten beschrieben wird. Bei sehr präzisen Messungen mit hoher Auflösung (insbes. Yara N-Sensor und EM_38) wurden die Messpunkte außerhalb des Kernbereiches manuell entfernt, so dass der Mittelwert die Input-Output-Beziehung genauer beschreibt. Die Anwendung von Korrekturalgorithmen beschränkte sich auf die Ertragsdaten.

Ertragsdaten

Für die betrachtete Kulturart Winterweizen konnte für beide Schläge auf zwei Ertragskarten der Jahre 1996 und 1998 zurückgegriffen werden. Das Ertragsmesssystem war jeweils ALPF 3000. In diesen Jahren erfolgte das Management schlageinheitlich. Die einzelnen Schritte der Datenvorverarbeitung sind beispielhaft für die Ertragsdaten des Jahres 1996 in Abbildung 24 illustriert.

Zunächst wird der Feuchtegehalt jeweils auf den marktüblichen Wert von 14% normiert. Für jeden Datensatz wird daraufhin aus den Histogrammen ein individueller Filter abgeleitet, mit dem Ausreißer entfernt werden (vgl. DRUMMOND UND SUDDUTH, 2004, s.p.). Beispielsweise werden von den Ertragsdaten *D4* des Jahres 1998 die Ertragspunkte entfernt, bei denen Erträge mit weniger als 4 t/ha oder mehr als 11 t/ha gemessen wurden.

Ein kritischer Bereich bei der automatischen Ertragserfassung ist die Qualität der Daten, die nahe am Vorgewende liegen (vgl. STEINMAYER, 2002, S. 25). Die Versuchsanlagen der IKB-Versuche reichen jedoch relativ nahe an das Vorgewende heran. Aus der visuellen Sichtung der Ertragskarten zeigt sich, dass sich die Ertragsmessungen erst innerhalb der ersten Teilfläche, ausgehend von der Schlaggrenze, stabilisieren. Deshalb wurden diese Messpunkte (bis in die Hälfte der Randparzellen hinein) entfernt, um so Ausreißer zu vermeiden. In Fällen, in denen nur eine sehr geringe Anzahl von Messpunkten vorliegt, wurde der entsprechende Wert als fehlender Wert aufgefasst. Aufgrund des begrenzten Datenumfanges konnten die Teilflächen am Rand der Versuchsanlagen jedoch nicht gänzlich von der Modellierung ausgeschlossen werden (vgl. Abbildung 24, Darstellung 1 und 2)

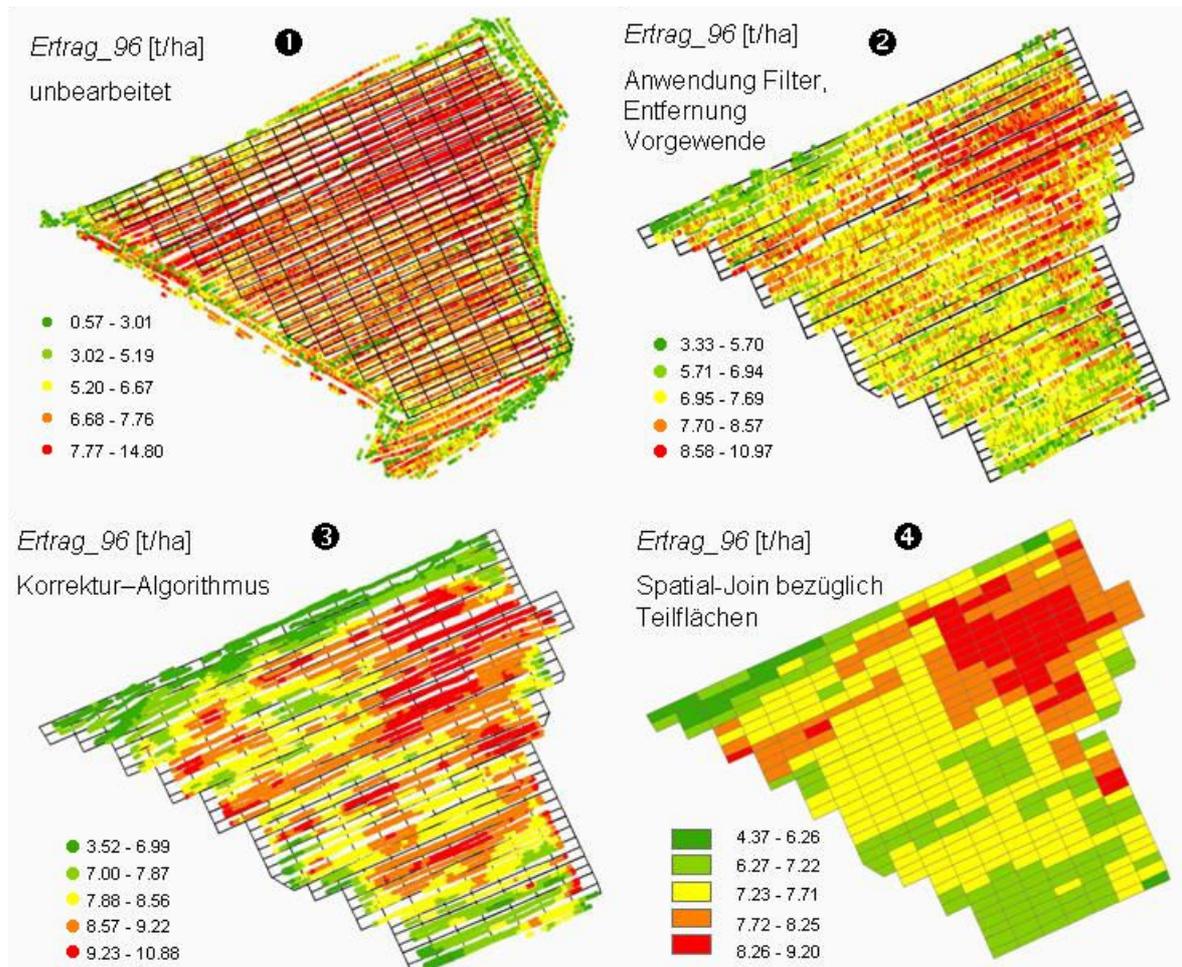


Abbildung 24: Datenvorverarbeitung IKB-Versuche bezüglich Ertragskarten

Nach diesen manuellen Eingriffen wurden die Rohdaten zusätzlich nach einer Methode von BACHMEIER UND AUERNHAMMER (2004, S. 375ff) korrigiert. Dazu wurde die Umgebung der jeweiligen Messpunkte betrachtet. Als Umgebung wird ein kreisförmiger bzw. schmetterlingsförmiger Bereich definiert (BACHMEIER UND AUERNHAMMER, 2005, S. 743f). Aus den darin liegenden Ertragsmesspunkten wird eine paraboloidale Struktur der Ertragsoberfläche erzeugt. Der betrachtete Ertragsmesspunkt im Zentrum des schmetterlingsförmigen Bereiches wird dann jeweils mit dieser Struktur verglichen und gegebenenfalls korrigiert (vgl. Abbildung 24, Darstellung 3).

Aus den korrigierten Punktdaten wurde dann mit Hilfe des Spatial Joins die durchschnittliche Merkmalsausprägung bezüglich der Teilflächen gebildet. Ergeben sich dabei Ausreißer, die am Feldrand lokalisiert sind, werden diese Teilflächendaten als fehlende Werte aufgefasst. Ein Beispiel für die dadurch entstehenden Teilflächen-Karten ist in Abbildung 24, Darstellung 4, illustriert.

Zugkraft

Die Zugkraftdaten werden in der Modellierungsphase als „Zugkraft“ bezeichnet. Für den

Schlag **Schafhof** wurden die Zugkraftdaten der Stoppelbearbeitung vom 14.08.2000 verwendet. Dabei erfolgte die Bodenbearbeitung mit einem Schlepper *Fendt 714 Vario* und einem Grubber *Lemken Smaragd*. Die Qualität der Messung kann aus der visuellen Sichtung als sehr gut bezeichnet werden. Ein Einfluss der Fahrtrichtung ist auf der Karte nicht zu erkennen. Da die Auflösung der Daten durch die geringe Arbeitsbreite (3m) sehr hoch ist, wurde der Mittelwert aller in einer Parzelle liegenden Punktdaten verwendet. Im Durchschnitt lagen über zwanzig Messpunkte innerhalb einer Teilfläche.

Die zeitliche Stabilität der Zugkraft-Muster konnte in diesem Rahmen nicht untersucht werden, da für diesen Schlag nur eine einzelne Karte vorliegt. Es muss daher eingeräumt werden, dass sich diese Werte vor allem in der Höhe aber auch in der Relation untereinander u. U. stark verändern können. Für den Versuch 2004, der ja wiederum am Schafhof stattfand, wurden ebenfalls diese Zugkraft-Daten verwendet.

Für den Schlag **D4** wurden die Zugkraftdaten der Stoppelbearbeitung am 04.09.2002 sowie am 06.09.2002 verwendet. An diesen beiden Tagen wurde der Schlag jeweils zur Hälfte bearbeitet. Die Daten dieses Jahres erscheinen ebenfalls sehr gut, es kann beispielsweise genau die Erosionsrinne und der damit verbundene Wechsel der Oberbodenart erkannt werden. In diesem Fall wurde ein Schlepper der Marke *Deutz* mit einem Grubber *Lemken Smaragd* verwendet. Für diesen Schlag existiert eine weitere Zugkraft-Erhebung des Jahres 2000. Der Vergleich der beiden Karten zeigt lediglich geringe Übereinstimmungen⁴⁸.

Reflexionsmessungen

Die spektralen Messungen *REIP_2* fanden kurz vor der zweiten und *REIP_3* kurz vor der dritten Stickstoff-Applikation statt und damit wiederum zu den festgelegten EC-Stadien statt.

Für den Versuch *Schafhof 02* wurden zu EC 49 Reflexionsmessungen mit einem Handsensor durchgeführt. Es wurden dabei ausschließlich die Teilflächen der Precision-Farming Düngesysteme erfasst. Innerhalb einer Teilfläche wurden fünf Messungen durchgeführt und jeweils der Hauptwendepunkt berechnet. Der Mittelwert aus den Einzelmessungen wurde der Teilfläche als Attribut *REIP_3* zugewiesen.⁴⁹ In den Folgeversuchen wurde zu EC 30 und EC 49 der Yara-N-Sensor im Scannermodus verwendet, der zum einen eine wesentlich höhere Anzahl von Messungen und zum anderen eine Positionsbestimmung der einzelnen Messung innerhalb der Parzelle erlaubt. Dadurch konnten die Messpunkte im Kerndruschbereich ausgewählt werden und daraus der Mittelwert gebildet werden.

⁴⁸ Die Korrelation zwischen diesen beiden Datensätzen, bezogen auf ein 10x10m-Raster (Räumliche Interpolation: Kriging, ohne Vorgewende), ergibt ein R^2 von lediglich 0.2.

⁴⁹ Der Varianz der Messwerte innerhalb einer Parzelle ist dabei sehr hoch. Die hohen Standardabweichungen könnten aus kleinräumigen Variabilitäten sowie aus Messungenauigkeiten entstanden sein. Eine Erhöhung der Anzahl der Messungen aber würde zeitliche Probleme ergeben, da die Messung von den PF-Parzellen (ohnehin nur die Hälfte aller Parzellen) mit einem Handsensor ca. zwei Tage in Anspruch nahm.

Scheinbare elektrische Leitfähigkeit

Die Messungen der scheinbaren elektrischen Leitfähigkeit wurden für beide Versuchstandorte vom Lehrstuhl für Pflanzenernährung (Prof. Schmidhalter) der TU München-Weihenstephan durchgeführt. Es wurde dabei ein EM38-Messgerät der Firma Geonics® verwendet. Die Messwerte wurden auf die Bodentemperatur von 25° C normiert. Wiederum wurden nur die Messpunkte im Kernbereich der Parzellen berücksichtigt.

Fazit

Grundsätzlich stellt diese Methode der Versuchsanlage und Datenvorverarbeitung einen Spezialfall dar. Die Durchführung ist nur bei kleineren Schlägen mit den entsprechenden Versuchsgeräten aus den Exaktversuchen möglich.

Die Verwendung von historischen Ertragsdaten gestaltet sich für die IKB-Versuche schwierig, da die Versuchsanlage nahe an das Vorgewende heranreicht. Für diese Bereiche ist die Qualität der automatischen Ertragserfassung oft nicht ausreichend. Allerdings darf für alle Bereiche davon ausgegangen werden, dass die Auswirkungen von Messfehlern auf die Kartierung durch die Anwendung des Korrekturverfahrens von BACHMEIER UND AUERNHAMMER (2004, S. 375ff) stark eingeschränkt werden. Dadurch, dass mit dieser Methode insbesondere benachbarte Fahrspuren zur Glättung herangezogen werden, erfolgt die Bewertung der Validität stärker an unabhängigen Beobachtungen.

Als vorteilhaft erwies sich eine reflexionsoptische Messung mit dem Yara-N-Sensor gegenüber der Messung von Hand. Durch eine spezielle Aufhängung konnten die Kernbereiche der Teilflächen mit hoher Präzision erfasst werden. Die Ergebnisse der Zugkraftmessungen müssen mit einer gewissen Vorsicht interpretiert werden da hierzu keine langfristigen Ergebnisse vorliegen. Es deutet sich an, dass die Messungen jahresspezifischen Einflüssen unterliegen.

Für diese Versuchsanlage spricht eindeutig die hohe Präzision, die an Exaktversuche heranreicht. Der Nachteil dieser Versuchsanstellung ist, dass Standard-Software-Pakete zur räumlichen Interpolation dafür nicht standardmäßig ausgelegt sind, da die Parzellen keine feste Größe und Form haben. Weiterhin sind zahlreiche manuelle Tätigkeiten zu erledigen, die sehr zeitintensiv und fehleranfällig sind.

4.3.2 Datenvorverarbeitung der Versuche in Görzig

Alternativ dazu steht die Datenvorverarbeitungsmethode, die für die Versuche *Görzig 04* und *Görzig 05* angewendet wurde. Diese Methode ist vor allem für den On-Farm-Einsatz, d. h. für den Einsatz auf Praxisschlägen ohne der Gerätschaft von Versuchsstationen geeignet. Hier werden typischerweise Maschinen mit entsprechend großer Arbeitsbreite für die Bewirtschaftung eingesetzt.

Diese Methode wird in vielen Precision-Farming-Versuchen verwendet und ausführlich beschrieben (vgl. z. B. DIKER ET AL. 2004, s.p.; DRUMMOND ET AL. 2004, s.p.). Eine Standard-Software dazu ist *SS-Toolbox* mit der Anwendung *Surfer* zur räumlichen Interpolation. In diesem Software-Konzept wird ein Schlag automatisiert in Nord-Süd-Richtung in ein quadratisches Raster eingeteilt und alle Sensordaten auf diese Rasterflächen räumlich interpoliert. Zur räumlichen Interpolation wird bevorzugt die Methode „Kriging“ verwendet. Diese Methode ist von LINSEISEN (2003, S. 55ff) ausführlich beschrieben.

In beiden Versuchen wurden alle räumlichen Sensor-Daten dieses Versuches in der Software *SS-Toolbox*⁵⁰ verarbeitet und mit dieser Methode auf einheitliche Rasterflächen (Grid) von 10 x 10 m interpoliert. Die Einteilung der Parzellen erfolgt also unabhängig von der Bewirtschaftungsrichtung, was die Präzision der Applikation verringert. Diese Konsequenz wird mit Abbildung 25 verdeutlicht. Im Hintergrund dieser Graphik ist ein Raster abgetragen, in der die Soll-Mengen zur Stickstoff-Ausbringung abgetragen sind, die sich aus Teilflächendaten errechnen. Die Applikationsmenge wird von der Parzelle bestimmt, in der sich der GPS-Empfänger befindet. Das Streubild eines Schleuderstreuers ist weit größer als die Arbeitsbreite (Fahrgassenabstand) (vgl. WELTZIEN ET AL., 2002, S. 162). Weiterhin ist die Arbeitsbreite nicht kongruent mit der Rastergröße und die Ausrichtung des Rasters nicht kongruent mit der Fahrgassenausrichtung.

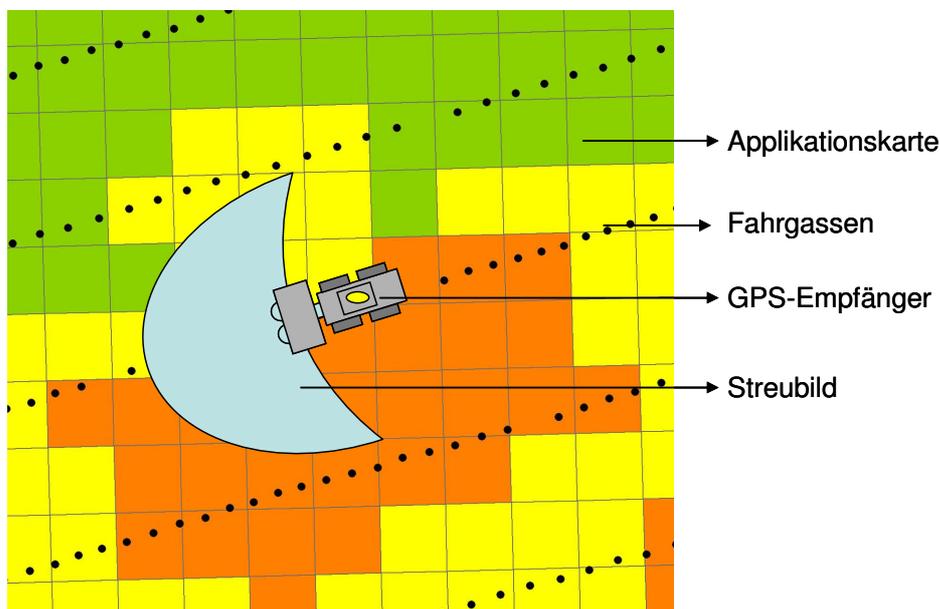


Abbildung 25: Technische Umsetzung der N-Applikation (Versuch Görzig 04 und Görzig 05)

Aus diesen Gründen weicht die Ist-Ausbringungsmenge mehr oder weniger stark von der Soll-Ausbringungsmenge ab. Im vorliegenden Fall (Rastergröße von 10x10 m) erhalten die Nachbar-Rasterflächen dieselbe Menge an Stickstoff wie die Parzellen, die direkt auf der Fahrgasse liegen, obwohl die Soll-Menge dieser Parzellen abweichen können. Es kann da-

⁵⁰ Hierbei handelt es sich um ein Produkt der Firma SST Development Group. In dieser Software ist das Produkt „Surfer“ integriert, mit dem die räumliche Interpolation erfolgt.

Görzig 04

Im Rahmen dieses Versuches wurden, ähnlich wie in den IKB-Versuchen, folgende räumliche Daten erhoben: Ertrag im Versuchsjahr (*Ertrag_04*), Vorjahresertrag (*Ertrag_03*), scheinbare Bodenleitfähigkeit (*EM_38*) sowie die spektralen Vegetationsindizes zur zweiten und dritten N-Gabe (*REIP_2* bzw. *REIP_3*).

Die Ertragsmessung erfolgte mit einer automatischen Ertragsaufzeichnung mit Claas-Mähreschern des Typs Lexion und den Ertragsmesssystemen Quantimeter. Die Rohdaten wurden von der Firma Agri Con® nach verschiedenen Methoden korrigiert (vgl. Tabelle A-30, im Anhang). Die Erträge des Vorjahres (Körnermais) wurden dabei nicht auf einen Feuchtegehalt normiert. Die Messung der scheinbaren elektrischen Leitfähigkeit und die Durchführung der reflexionsoptischen Messungen wurden ebenfalls von der Fa. Agri Con® mit einem EM38-Gerät der Firma Geonics und einem Yara® N-Sensor® im Scannermodus durchgeführt.

Görzig 05

Für den Versuch Görzig 05 wurden dieselben Daten mit den entsprechenden Systemen erhoben und korrigiert. Zusätzlich standen in diesem Versuch die Erträge des Vorjahres (*Ertrag_03*) sowie Proteinwerte der einzelnen Streifen (stichprobenartige Ährenschnitte) zur Verfügung.

4.3.3 Fazit Datenvorverarbeitung

In dieser Arbeit wurden zwei sehr verschiedene Konzepte zur Versuchsanlage und zur entsprechenden Datenvorverarbeitung vorgestellt. Das IKB-Konzept besticht durch eine höhere Präzision in der Applikation und in der Ertrags Erfassung. Nachteilig wirkt sich aus, dass damit ein enormer Arbeitsaufwand in der Versuchsanlage, Versuchsdurchführung und Versuchsauswertung verbunden ist. Insbesondere zur Datenaufbereitung existieren keine entsprechenden Softwarepakete, die diese Arbeit unterstützen. Deshalb ist dieses Konzept nur für kleinere Schläge und für wissenschaftliche Anwendungen geeignet.

Das Konzept, das für den Versuch in Halle angewendet wurde, ist hingegen auch für Großschläge geeignet. Dadurch können wesentlich mehr Instanzen erzeugt werden und größere räumliche Bereiche betrachtet werden. Nachteilig wirkt sich hier insbesondere die geringere Präzision in der Applikation aus. Dies ist vor allem für die Verifikation von Bedeutung: Selbst wenn die Soll-Mengen exakt bestimmt werden können, können die Ist-Mengen stark von diesen Werten abweichen, da die Teilflächen nicht exakt ansteuerbar sind. Für die Datenerfassung können diese Effekte durch die Anwendung der räumlichen Interpolation jedoch bis zu einem gewissen Maß berücksichtigt werden. Grundsätzlich ist diese Methode auch in der Praxis einsetzbar, da sich der versuchsbedingte Mehraufwand in Grenzen hält (vgl. Kap. 5.1).

Ein Vergleich der beiden Methoden zur Datenvorverarbeitung, bzw. eine Bewertung der Qualität der daraus entstehenden Instanzen, ist nach dem derzeitigen Stand der Forschung schwierig und kann im Rahmen dieser Arbeit nicht näher untersucht werden. Es existiert nach wie vor kein Standard zur Beurteilung der Qualität von Rohdaten bzw. vorverarbeiteten Daten. Die gerne verwendeten Validierungsverfahren sind nicht unumstritten, da nach diesem Konzept die korrigierten Daten mit Rohdaten verglichen werden, die vielfältigen Fehlerquellen ausgeliefert sind (vgl. Abschnitt 3.1.3). Die Methode Kriging zur Interpolation von Sensorkarten ist weitgehend akzeptiert und wird dementsprechend häufig zur Korrektur eingesetzt. Jedoch deutet sich an, dass mit Methoden, die die benachbarten Fahrspuren stärker berücksichtigen, bessere Ergebnisse erzielt werden (vgl. NOACK ET AL., 2005, S. 705; BACHMEIER UND AUERNHAMMER, 2004, S. 375ff). Inwiefern dieser Vorteil sich für eine Data-Mining-Anwendung auswirkt, kann nicht beurteilt werden. Beide hier vorgestellten Methoden führen dazu, dass das Risiko von fehlerhaften Instanzen stark eingeschränkt wird.

Im Ergebnis führen beide Konzepte zu Instanzen, die im ersten Schritt zur Modellierung einer kleinräumigen Ertragsprognose verwendet werden. Dazu werden die Instanzen in einer Tabellenform abgespeichert (vgl. Kap. 2.4.). In diesem Format werden die Instanzenmengen aus den einzelnen Jahren vor der Modellierung nochmalig in der Exploration analysiert und gegebenenfalls korrigiert, wie in den Abschnitten 4.5 bis 4.7 dargestellt wird.

4.4 Umsetzung: Entwicklung von Entscheidungsregeln und Extraktion von Mustern

Den Anforderungen aus Kap. 3.1.2 entsprechend muss die Vorgehensweise in der Modellierung definiert werden. Ein Ziel dieser Arbeit besteht in der Entwicklung einer Methode, mit der aus Precision-Farming-Daten ökonomische Entscheidungsregeln abgeleitet werden können. Die Ausgangselemente liegen - wie in Kap. 3.2 dargestellt wurde - bereits vor. Diese Bausteine müssen also aufgegriffen, angepasst und zusammengefügt werden, um sie für den Zweck der Entwicklung von Entscheidungsregeln für die kleinräumige Stickstoff-Applikation nutzen zu können.

Die Aufgabe wurde in drei Teilbereiche unterteilt, die aufeinander aufbauen. Als erstes wird versucht, ein neuronales Netz zur teilflächenspezifischen Ertragsprognose für variable N-Mengen zu erzeugen, um die kleinräumigen Zusammenhänge abbilden zu können.

Dieses Modell wird unmittelbar dazu verwendet, um für eine bestimmte teilflächenspezifische Situation die ökonomisch optimale Stickstoff-Applikation herauszufinden. Dies wird durch die Simulation mit einer eingegrenzten Auswahl von technisch möglichen Düngergaben erreicht. Die Auswahl einer Applikationsmenge N_{opt} soll nach streng ökonomischen Gesichtspunkten erfolgen.

Um Entscheidungsregeln zu erzeugen, wird dieser Vorgang für eine Anzahl von spezifischen Teilflächen (und entsprechend verschiedenen Merkmalsausprägungen) wiederholt und jeweils eine optimale Stickstoffmenge berechnet. Der daraus entstehende Datensatz wird dazu verwendet um mit Entscheidungsbaumverfahren Regeln zu erzeugen. Dieser Schritt ist dazu notwendig, um zu explizitem Wissen zu gelangen, das in einem In-Field-Controller transferiert werden kann und somit eine automatisierte Umsetzung des Ansatzes ermöglicht.

Gleichzeitig ermöglicht diese Aufteilung, überprüfbare Zwischenschritte zu finden. Beispielsweise können die Ergebnisse der teilflächenspezifischen Ertragsprognose mit den im Abschnitt 3.2 beschriebenen Studien verglichen werden.

Weiterhin ist es Ziel dieser Arbeit, konkrete Ansätze (Muster) für eine ökonomische Optimierung zu liefern. Wie sich in der Data Mining Phase herausstellt, helfen Entscheidungsregeln nur bedingt bei der Suche nach stabilen Mustern über die einzelnen Versuchsjahre hinweg (vgl. 4.4.3). Deshalb werden im Abschnitt 4.4.4 alternative Techniken vorgestellt und ein Konzept zur Extraktion von Mustern innerhalb der Data Mining Phase aufgebaut.

4.4.1 Teilflächenspezifische Ertragsprognose

Die hier verwendete Methode der teilflächenspezifischen Ertragsprognose basiert auf dem Ansatz von DRUMMOND ET AL. (2002, s.p.), in dem ausschließlich räumliche Sensordaten verwendet wurden. Der zentrale Unterschied zu diesem Ansatz besteht darin, dass zur Ertragsprognose auch die Menge an appliziertem Stickstoff verwendet wird. Die Stickstoffgaben werden als kontrollierbares Attribut aufgefasst (vgl. POKRAJAC UND OBRADOVIC, 2001b, S. 2 ff).

Eine weitere Anpassung muss erfolgen, um die Verteilung der Stickstoffmenge auf drei Teilapplikationen zu berücksichtigen (vgl. Kap. 3.1.). Dazu müssen drei Applikationsmengen zu unterschiedlichen Zeitpunkten optimiert werden. Die Ertragsprognose muss damit zu den **drei Zeitpunkten der N-Ausbringung** (an festgelegten EC-Stadien) erfolgen. Zu diesen Ausbringungszeitpunkten liegen aber jeweils unterschiedliche Informationen vor. Die Modelle müssen somit der zeitlichen Abfolge des Informationsgewinnes Rechnung tragen. Die Entscheidung über die Höhe einer N-Gabe soll auf der Grundlage aller bis zu diesem Zeitpunkt erhältlichen Informationen getroffen werden.

Modellbildung

Im Grundmodell soll der teilflächenspezifische Ertrag (Y_T) annahmegemäß von kontrollierbaren und nicht-kontrollierbaren Faktoren abhängig sein. Als kontrollierbar wird die Stickstoff-Applikation (N) angesehen. Als nicht kontrollierbar werden die teilflächenspezifischen Sensor-Attribute (s_1 bis s_N) aufgefasst. Diese Attribute werden damit als *Prediktoren* bzw. als Input-Variablen für die Modellierung verwendet. Es können hier per se alle mess-

technisch erfassbaren Größen verwendet werden, die in den Feldversuchen erhoben wurden und denen unterstellt wird, dass sie einen Einfluss auf die Ertragsvariabilität besitzen⁵¹. Mit diesen Annahmen sieht die Ertragsfunktion im Grundmodell folgendermaßen aus:

$$Y_T = f(N; s_1, s_2, \dots, s_M)$$

mit M = Anzahl der Sensor-Attribute

Dieses Modell muss entsprechend angepasst werden, um die drei Düngezeitpunkte zu berücksichtigen. Es soll zu den drei Düngezeitpunkten eine ökonomische Entscheidung zur Höhe der Ausbringung durchgeführt werden. Zu jedem dieser Zeitpunkte muss also ein Modell zur Verfügung stehen, das aus den bis dato vorhandenen Informationen den Ertrag möglichst genau schätzt.

Das Modell zur ersten Stickstoff-Applikation darf beispielsweise ausschließlich Bodenattribute und die applizierte Stickstoffmenge (N_I) zur Ertragsprognose verwenden. In der Anwendung dieses Modells wird dabei die N_I -Menge als variabel aufgefasst, da diese Menge erst zu diesem Zeitpunkt ausgebracht wird. Mit dem trainierten Modell soll der Ertrag für verschiedene Ausprägungen des Attributes N_I (unter sonst konstanten Bedingungen) prognostiziert werden und daraus Rückschlüsse für eine optimale Intensität gezogen werden.

Zur zweiten N-Gabe ist das Attribut N_I nicht mehr variabel, da die Ausbringung bereits erfolgt ist. N_I wird daher zu einem nicht-kontrollierbaren Attribut. Zu diesem Zeitpunkt liegen zusätzlich die Informationen der ersten reflexionsoptischen Messungen vor. Entsprechendes gilt für die dritte N-Applikation.

Diese Ertragsfunktion soll somit zu drei Zeitpunkten geschätzt werden: Zum Zeitpunkt der ersten N-Gabe (N_I) bei Vegetationsbeginn, zum Zeitpunkt der zweiten N-Gabe (N_2) bei EC 32 und zum Zeitpunkt der dritten N-Gabe (N_3) bei EC 49 (vgl. Kap. 3.1.1). Dabei ist die Anzahl der Attribute zur dritten Düngergabe maximal. Nach diesem Zeitpunkt werden keine Attribute mehr erfasst. Die Abbildung der Ertragsfunktion wird demnach zu diesem Zeitpunkt die höchste Präzision erreichen.

In Abbildung 27 sind die Attribute dargestellt, die in dieser Arbeit zum jeweiligen Zeitpunkt maximal zur Verfügung stehen. Im Einzelfall muss jedoch entschieden werden, ob wirklich alle Attribute verwendet werden, ob Attribute weggelassen oder ob mehrere Attribute zu einem verdichtet werden. Beispielsweise kann u. U. in der Modellierung der Ertragsfunktion zum Zeitpunkt EC 32 auf das Attribut N_I verzichtet werden, da der Einfluss dieser Teilapplikation auf den Ertrag im Ernährungszustand der Pflanzen über das Attribut $REIP_2$ abgebildet sein könnte.

⁵¹ Es sei darauf hingewiesen, dass nur Messgrößen verwendet werden können, die sich teilflächenspezifisch verändern. Eine einheitliche Ausbringung von Pflanzenschutz liefert beispielsweise nur eine Ausprägung über alle Instanzen und hat damit keinen Einfluss auf die teilflächenspezifische Ertragsfunktion.

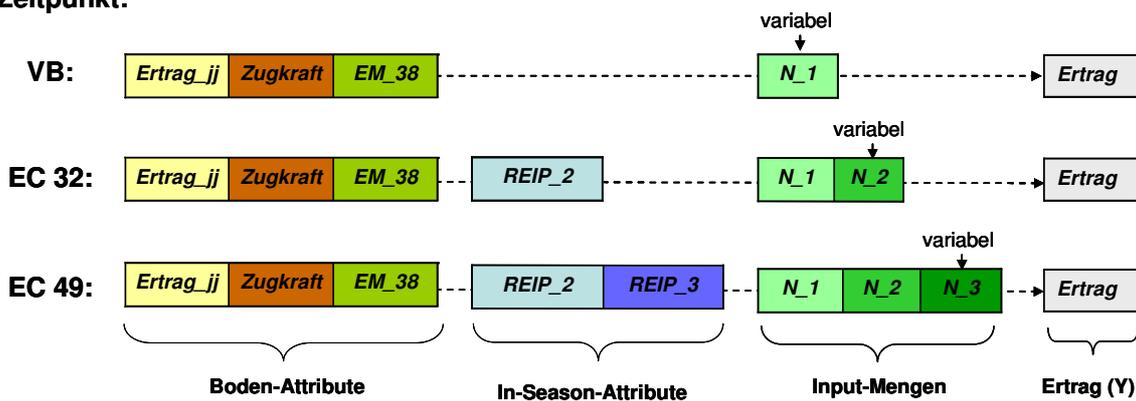
Zeitpunkt:

Abbildung 27: Verwendete Attribute zur Ertragsprognose in den Teilapplikationen

Zur Modellierung der Ertragsfunktion werden künstliche neuronale Netzwerke verwendet (vgl. Kap. 2.5). Als Lernalgorithmus wird die Methode Backpropagation eingesetzt (vgl. Kap. 2.5.2). Aufgrund der eingeschränkten Anzahl von Instanzen kann der rechenintensive *Exhaustive-Prune-Algorithmus* verwendet werden, um eine angepasste Topologie zu erzeugen (vgl. Kap. 2.5.3). Das Training der neuronalen Netzwerke erfolgt über die Instanzen (Teilflächen) aus den Feldversuchen. Beispielsweise liegen vom Versuch *D4 03* ca. 200 vorverarbeitete Instanzen vor. Zu jeder dieser Instanzen liegen Merkmalsausprägungen der einzelnen Attribute und der Zielgröße (Ertrag) vor. Die jeweiligen Merkmalsausprägungen werden als absolute Werte in das Netz eingegeben. Dadurch können in der Ex-post-Betrachtung die Einflussfaktoren für den Ertrag wiederum in absoluten Größen analysiert werden.

Die Instanzen werden vor dem Training zufällig in ein Trainings- und ein Testsample aufgeteilt⁵². Wie in Kap. 2.5.3 beschrieben werden mit Hilfe des Trainingsamples die Gewichte angepasst. Die Anpassung der Topologie wird über das Testsample erreicht. Dazu wird die Präzision der Vorhersage über die Kennzahl *Estimated Accuracy* (vgl. Tabelle 1) geschätzt. Das Training wird automatisch abgebrochen, wenn sich keine weiteren Verbesserungen durch weiteres Training bzw. durch Anpassung der Topologie erreichen lassen (vgl. *Persistence*, 2.5.3).

Zusammenfassend besteht das Ziel des Netzwerktrainings darin, eine Ertragsfunktion zu schätzen. Diese Funktion lässt sich allerdings nicht beschreiben, sondern ist in der Netztopologie mit den entsprechenden Verbindungsgewichten abgebildet. Erst in der Anwendung des trainierten neuronalen Netzwerkes lässt sich dann feststellen, wie sich eine Veränderung eines oder mehrerer kleinräumiger Attribute bei konstanten Werten der anderen Attribute auf den Ertrag auswirkt. Eine Frage, die beispielsweise damit beantwortet werden

⁵² In der Regel wird eine hälftige Aufteilung verwendet. Bei geringen Datenumfängen kann die Größe des Testsamples verringert werden.

kann, lautet: Welche Korntrträge sind für eine Teilfläche mit festgelegten Bodencharakteristiken zu erwarten, wenn die Teilapplikation N_1 von 30 kg N/ha auf 50 kg N/ha erhöht wird? Bei einem optimal trainierten Netz sind (theoretisch) alle funktionalen Zusammenhänge der Attribute entsprechend abgespeichert und es können beliebig komplexe Fragestellungen (z. B. die Variation mehrerer Attribute) abgerufen werden⁵³.

Umsetzung

In der Praxis des Netzwerktrainings ergeben sich vielfach Probleme, die eine „optimales“ Training und somit die Abbildung der funktionellen Zusammenhänge verhindern. Überwiegend resultieren diese Probleme aus der eingeschränkten Qualität der Trainingsdatenbasis. Vereinfacht dargestellt kann das neuronale Netz keine Zusammenhänge lernen, für die es keine Trainingsbeispiele gibt. Gleichzeitig lernt das neuronale Netz auch aus fehlerhaften bzw. verzerrten Daten. Dem neuronalen Netz sollten demnach Trainingsbeispiele präsentiert werden, die nicht durch versuchsbedingte Rahmenbedingungen oder sonstige Einflüsse verfälscht sind. Beispielsweise kann das Lernergebnis schon durch eine unglückliche Auswahl von Teilflächen, denen hohe Stickstoffmengen appliziert werden, beeinträchtigt werden. Wenn die ausgewählten Teilflächen gleichzeitig jeweils ein sehr hohes Ertragspotenzial besitzen, könnte die Wirkung einer hohen N-Applikation auf den Ertrag überschätzt werden. Hinweise auf mögliche Verzerrungen liefert beispielsweise eine Korrelationsmatrix bzw. ein Data Audit (vgl. 4.4.4). Mögliche Verzerrungen bzw. Fehler können dabei allerdings nur (subjektiv) vom Anwender beurteilt werden.

Weitere Risiken liegen in der Modellbildung. Aus den einzelnen Datensätzen (*Schafhof 02, D4 03, Görzig04*) wird für jede der drei Stickstoff-Applikationen jeweils ein neuronales Netzwerk trainiert. Das heißt, es wird jeweils dieselbe Datenbasis verwendet und lediglich eine unterschiedliche Anzahl von Attributen zur Prognose ausgewählt. Dadurch genügen die Instanzen eines Feldversuches zur Optimierung der drei Teilapplikationen.

Diese Vorgehensweise birgt allerdings Risiken in sich: Die Zielgröße wird unter Umständen von (kontrollierbaren) Attributen, die in den Teil-Modellierungen nicht verwendet werden, entscheidend beeinflusst. Insbesondere kann der Ertrag durch die späteren Stickstoff-Applikationen beeinflusst werden. Korrelieren die Teil-Applikationen (*N_1, N_2* u. *N_3*) miteinander (beispielsweise durch Anwendung eines Mapping-Ansatzes), kann nicht gewährleistet werden, dass der Einfluss **einer** Teilgabe adäquat abgebildet wird. In einem entsprechend trainierten Modell würde der Einfluss einer Teilgabe u. U. zu stark dargestellt werden.

⁵³ Zur Beantwortung der Beispielsfrage müssen dazu zwei Instanzen in ein trainiertes Netz eingegeben werden. Beide Instanzen besitzen dieselben (vorher definierten) Ausprägungen für Bodenattribute. Im einen Fall wird der Wert „30“ im anderen Fall der Wert „60“ für die N_1-Applikation eingesetzt. Dann können die prognostizierten Erträge verglichen werden.

Anwendung

In der Anwendung der trainierten Netze ergeben sich ähnliche Risiken wie beim Training selbiger. In der vorliegenden Arbeit werden die trainierten neuronalen Netzwerke überwiegend zur Ex-post-Analyse verwendet (vgl. Abschnitt 4.5-4.7). Es erfolgt damit eine Analyse der kleinräumigen N-Applikation für ein vergangenes Jahr. Dennoch ist es ein Ziel dieser Arbeit, die trainierten neuronalen Netze, bzw. die daraus abgeleiteten Entscheidungsregeln, in neuen Versuchen zu testen. Für die Anwendung zur Optimierung der kleinräumigen N-Düngung ergeben sich einige diskussionswürdige Punkte.

Zunächst müssen für eine „echte“ Prognose, d. h. für den Einsatz der trainierten neuronalen Netze in zukünftigen Jahren, wieder die Voraussetzungen geschaffen werden, auf die das neuronale Netz „geeicht“ wurde. Dies bedeutet unter anderem, dass die gleichen Attribute zur Verfügung stehen sollten. Bei der Verwendung von Ertragsattributen aus bestimmten Jahren zum Training sollten demnach die entsprechenden Ertragsdaten auch zur Anwendung vorliegen. Bei unterschiedlichen Fruchtfolgen müssen demnach Anpassungen vorgenommen werden, die schwer quantifizierbar sind (vgl. 4.8.1). Es stellt sich für alle Attribute die Frage, ob die Datenerfassung unter vergleichbaren Voraussetzungen erfolgt ist, oder ob Anpassungen vorgenommen werden können.

Ein weiterer kritischer Punkt ist der Einfluss der Witterung. Wie in Abschnitt 3.2 dargestellt ist der Einsatz von Witterungskennzahlen zum Training der neuronalen Netzwerke unter den Voraussetzungen dieser Arbeit nicht zielführend. Es stellt sich die Frage, wie das Netzwerk auf unterschiedliche jahresspezifische Einflüsse angepasst werden kann. Für diese Arbeit soll –mangels Alternativen– ein pragmatischer Weg eingeschlagen werden, zu dem es bisher keine Erfahrungswerte gibt. Es wird die Hypothese aufgestellt, dass die jahresspezifischen Einflüsse über die reflexionsoptische Kennzahl REIP kalibriert werden können. Diese Hypothese basiert auf den Untersuchungen von MISTELE (2005, S. 56) zur absoluten Skalierbarkeit von Reflexionsindices. Es deutet sich an, dass mit diesen Kennzahlen absolute (jahresunabhängige) Aussagen zum Ernährungszustand eines Bestandes getroffen werden könnten⁵⁴. Übertragen auf diese Arbeit wird angenommen, dass die ökonomisch optimale N-Applikation unter gleichen Bedingungen (gleiche Ausprägungen der erfassten teilflächenspezifischen Attribute) unabhängig vom Jahr konstant ist. Schlussendlich muss die Annahme getroffen werden, dass sich die relevanten kleinräumigen Zusammenhänge (einschließlich der reflexionsoptischen Kennzahlen) zur Ertragsbildung aus den Versuchsdaten eines Jahres (!) bestimmen lassen.

Unter diesen Annahmen wurden die Ertragsprognosemodelle des Versuches *Görzig 04* zur

⁵⁴ In den Untersuchungen von Mistle (2005) tauchen auch Fälle auf, in denen diese Annahme nicht zutrifft. Dafür werden insbesondere die Rahmenbedingungen der Versuche (unterschiedliches Sensorsystem, Sorte, Vorfrucht, Bodenfarbe) verantwortlich gemacht. Im Rahmen dieser Arbeit kann dies nicht näher untersucht werden.

Optimierung von N-Applikationen im darauf folgenden Jahr verwendet (Görzig 05, vgl. Kap. 4.8). In diesem Rahmen wurde eine „echte“ Prognose erstellt, obwohl die (Trainings-) Daten nur aus einem Jahr stammen.

Diskussionsbedarf besteht auch darüber, ob eine Datenbasis aus verschiedenen Jahren verwendet werden kann, um aus den kombinierten Instanzen ein Modell zur „echten“ Ertragsprognose zu entwickeln. Die Idee einer solchen Versuchsserie besteht darin, ein neuronales Modell zu entwickeln und es über die Jahre hinweg jeweils mit der sich vergrößernden Datenbasis zu trainieren. Nach DRUMMOND ET AL. (2003, s.p.) werden für die Trainingsdatenbasis eine möglichst große Anzahl an Schlag-Jahren benötigt, um dann eine echte Prognose für ein Jahr erstellen zu können, das nicht in der Trainingsmenge verwendet wurde. Dieser Ansatz wurde mit den Versuchsdaten *Schafhof 02* und *D4 03* getestet. Es stellte sich heraus, dass die Gefahr von Überlernen alleine durch die Kompatibilitätsprobleme der beiden Datensätze ansteigt (WEIGERT UND WAGNER, 2004, s.p.). In dieser Arbeit wird auf diesen Ansatz nicht näher eingegangen.

Dieser Abschnitt zeigt, dass einige kritische Punkte in der Umsetzung und Anwendung der neuronalen Prognosemodelle existieren. Für diese Arbeit wurde ein pragmatischer Weg eingeschlagen, der Elemente aus verschiedenen Ansätzen verwendet. Die getroffenen Annahmen können in diesem Rahmen nicht näher untersucht werden.

4.4.2 Ökonomische Optimierung von N-Gaben mittels trainierter neuronaler Netze zur Ertragsprognose

Der folgende Schritt basiert auf dem Ansatz von POKRAJAC UND OBRADOVIC (2001a, S. 2111ff). In ihrer Arbeit verwenden sie Ertragsprognosemodelle zur ökonomischen Optimierung von Düngermengen (Methode: Direct Modelling). Die Ausgangslage stellt sich übertragen auf die vorliegende Arbeit beispielsweise folgendermaßen dar: zum Zeitpunkt „Vegetationsbeginn“, also zur ersten N-Gabe, soll für eine bestimmte Teilfläche in einem Schlag die Entscheidung zur ökonomisch-optimalen Höhe der Stickstoff-Applikation getroffen werden. Von dieser Teilfläche sind zu diesem Zeitpunkt die Ausprägungen der Bodenattribute bekannt. Zur Verfügung steht weiterhin das Ertragsprognosemodell, das im vorausgegangenen Schritt erzeugt wurde. Der Ansatz der beiden Autoren besteht darin, eine Auswahl an *möglichen Ausbringungsmengen* der Stickstoff-Applikation in Kombination mit den Bodenattributen in dieses Modell einzugeben (Input) und daraus jeweils den Ertrag (Output) zu prognostizieren.

In der vorliegenden Arbeit müssen entsprechend die Teilflächendaten zu den jeweiligen Düngezeitpunkten in Kombination mit der Teilapplikation in den möglichen Ausbringungsmengen als Input in das trainierte neuronale Netz eingegeben werden. Aus diesen Input-Output-Kombinationen kann eine Auswahl der ökonomisch optimalen N-Menge er-

folgen.

Dazu muss allerdings das Auswahlkriterium festgelegt werden. Das bedeutet, es muss eine Zielgröße definiert werden, anhand derer die Auswahl erfolgen soll. Dazu die stickstoffkostenfreie Leistung (NKL) verwendet (vgl. Kap. 3.1). Eine teilflächenspezifische stickstoffkostenfreie Leistung NKL_T für eine Stickstoffmenge N errechnet sich für einen teilflächenspezifischen Ertrag Y_T (der wiederum von N abhängig ist) wie folgt:

$$NKL_T(N) = p_y Y_T(N) - p_n N$$

Dabei steht p_y für den Faktorpreis pro Einheit des Ertrages Y . Der Faktorpreis p_n bezieht sich auf die Einheit des Inputs N . In der hier vorliegenden Arbeit wird durchgängig ein Winterweizenpreis p_y von 120 € pro t und ein Stickstoffpreis p_n von 0,65 € pro kg N angenommen. Zur Auswahl von N-Gaben muss jeweils der prognostizierte Wert $\hat{Y}(N_i)$ des neuronalen Netzwerkes für einen spezifischen Wert N_i eingesetzt werden. Damit kann für jeden Input N_i ein teilflächenspezifischer Wert NKL_T errechnet werden:

$$NKL_T(N_i) = q \hat{Y}_T(N_i) - c N_i$$

Zur Ermittlung der optimalen N-Gabe in Abhängigkeit von den gegebenen Attributen muss die Funktion von $NKL_T(N_i)$ maximiert werden (vgl. ROBERTS ET AL, 2000, S. 134ff). Formal also⁵⁵:

$$N_{opt} = \arg \max_N \{NKL_T(N_i)\}$$

Wegen der begrenzten Präzision der Düngerstreuer können die applizierbaren Düngermengen N_i als diskret aufgefasst werden (POKRAJAK UND OBRADOVIC, 2001, S. 2111). In der vorliegenden Arbeit werden die zur Auswahl stehenden Teilapplikationsmengen N_i aus der Menge

$$\{0, 10, 20, \dots, Max\}$$
 ausgewählt.

Die maximale Höhe der N-Gabe (Max) ergibt sich aus der maximal applizierten Menge in den jeweiligen Trainingsinstanzen. Für jede dieser N-Mengen wird daher die jeweilige stickstoffkostenfreie Leistung berechnet und die maximale Leistung ausgewählt. Die entsprechende N-Gabe wird als die ökonomisch optimale Menge N_{opt} bezeichnet werden. Das Optimierungsproblem wird dadurch in ein Auswahlproblem umgewandelt. Die Auswahl kann mit Hilfe einfacher Algorithmen (linear bzw. sequential search) durchgeführt werden⁵⁶.

⁵⁵ Arg max ist eine Funktion der Komparativen Statik, die ein Argument auswählt, für das eine Funktion ihr Maximum erreicht (vgl. SYDSAETER ET AL., 2000, S. 151).

⁵⁶ Bei der Optimierung einer Teilapplikation werden die restlichen Teilapplikationen als konstant angesehen; diese fließen aber in vorliegender Arbeit nicht in die Berechnung der NKL mit ein. Damit wird die NKL in der Regel überschätzt. Für die Optimierung ist dies aber von keiner Bedeutung.

In Abbildung 28 ist der gesamte Vorgang vereinfacht dargestellt. In diesem Fall wird die Düngung N_1 für eine Teilfläche X optimiert. Die möglichen Düngermengen liegen dabei zwischen 0 und 30 kg N/ha. Als diejenige N_1 -Menge, die zu der höchsten stickstoffkostenfreien Leistung führt, wird in diesem Beispiel die Menge 10 kg N/ha ausgewählt. Mit dieser Methode können analog die speziellen Intensitäten der weiteren Teilgaben ($N_{2,opt}$, $N_{3,opt}$) bestimmt werden.

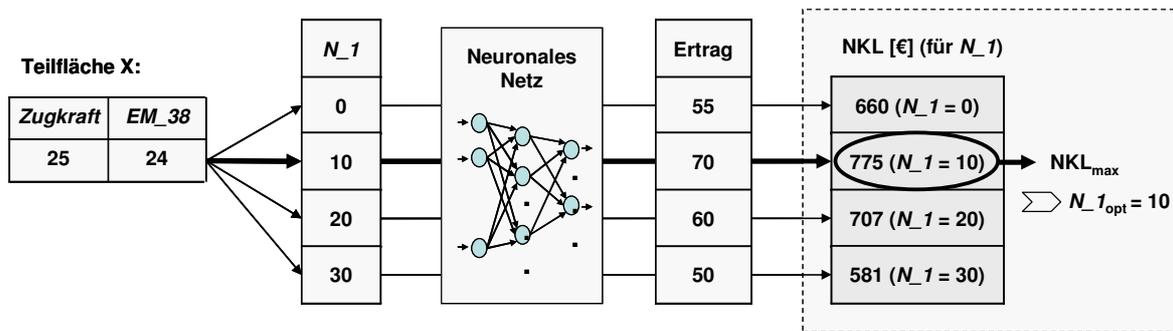


Abbildung 28: Ökonomische Optimierung von N-Gaben mit Ertragsprognosemodellen

Diese beiden Schritte (Ertragsprognose und ökonomische Auswahl) stellen also ein Konzept dar, wie aus einer Menge an Trainingsinstanzen Modelle erzeugt werden können, mit denen die Stickstoff-Ausbringung ökonomisch optimiert werden kann. Die Trennung der beiden Schritte ermöglicht es, zum einen an agronomischem Wissen zur Ertragsprognose anzuknüpfen. Zum anderen können einmal erstellte teilflächenspezifische Ertragsprognosemodelle auch dann verwendet werden, wenn sich Faktorpreise verändern. Die Anpassung erfolgt lediglich in der NKL-Berechnung. Gleichzeitig könnten aber auch andere Zielsysteme implementiert werden. Es wäre beispielsweise denkbar, zusätzlich Qualitätsmerkmale des Ernteguts (z. B. Rohproteingehalt) zu bewerten, sofern die entsprechenden Attribute auch in den Ertragsprognosemodellen eingebaut sind.

Die Aufgabe der ökonomischen Optimierung von teilflächenspezifischen N-Gaben könnte damit bereits mit diesen beiden Schritten erledigt werden. In den Ansätzen zur Verifikation wurde dies auch so durchgeführt. Allerdings ergeben sich dadurch zwei gewichtige Nachteile: Zum einen ist der Optimierungsvorgang relativ intransparent. Insbesondere die trainierten neuronalen Netze zur Ertragsprognose können nicht „verstanden“ werden und kein Wissen extrahiert werden. Es existieren zwar einige Methoden, mit denen die Transparenz erhöht werden kann (vgl. Abschnitt 4.4.4), jedoch können damit nur Teilaspekte des Ertragsprognosemodells beleuchtet werden. Zum anderen soll die Optimierung automatisiert in einem In-Field-Controller ablaufen, was mit dieser Methode nur umständlich möglich wäre. Um diese beiden Anforderungen zu erfüllen, ist ein weiterer Schritt notwendig, nämlich die Extraktion von ökonomisch optimierten Entscheidungsregeln.

4.4.3 Extraktion von ökonomisch optimierten Entscheidungsregeln

Die Aufgabe des folgenden Schrittes besteht darin, Beschreibungen bzw. Muster des Optimierungsvorganges zu erhalten, die dann interpretiert, bewertet und verglichen werden können. Die Beschreibungen sollen in Form von Regeln bzw. Entscheidungsbaum dargestellt werden. Mit dieser Darstellung soll die Transparenz erhöht werden. Es soll für Experten möglich sein, die Regeln zu verstehen und sie gegebenenfalls anzupassen.

Die erzeugten Regeln können gleichzeitig auf einem In-Field-Controller eingesetzt werden, um einen Online-Ansatz zur Düngung (Sensor-Ansatz mit Kartenüberlagerung) zu verwirklichen. Zur Erzeugung von Regeln bieten sich Entscheidungsbaumverfahren (vgl. Kap. 2.6) an.

Dazu werden wiederum die Ergebnisse der zwei vorausgegangenen Schritte verwendet. Mit diesen beiden Schritten ist es möglich, die optimalen speziellen Intensitäten für teilflächenspezifische Verhältnisse zu bestimmen. Werden diese beiden Schritte für alle Instanzen (Teilflächen) eines Feldversuches wiederholt, lässt sich mit den beiden Methoden ein Datensatz erzeugen, in dem für kleinräumige Attributs-Ausprägungen die jeweilige optimale N-Gabe abgetragen ist (vgl. Tabelle 12). Dieser Datensatz bildet die Trainingsmenge für den folgenden Schritt des überwachten Lernens.

Tabelle 12: Beispielhafter Datensatz zur Erzeugung von Entscheidungsregeln

Ertrag_96	Zugkraft	EM_38	N_1	REIP_2	N_2 _{opt}
8.1	33	21	40	721.1	30
7.5	27	23	10	723.4	40
8.1	40	22	50	724,3	60
...

Der Entscheidungsbaumalgorithmus wird dazu verwendet, um aus diesem Datensatz Regeln (bzw. einen Entscheidungsbaum) zu erzeugen. Die Zielgröße ist in diesem Fall die optimierte N-Menge N_{opt} . Die erklärenden Attribute sind die jeweiligen teilflächenspezifischen Daten der Teilapplikationen. Im Vergleich zur Ertragsprognose wird also nur für jede Instanz die Zielgröße Ertrag durch die optimierte N-Menge ersetzt. Zur Erstellung von Entscheidungsbaum wird der CART-Algorithmus (Software-Clementine®) in den Standard-Einstellungen verwendet. Mit diesem Algorithmus lassen sich sowohl Klassifikationsbäume (N_{opt} als kategorisches Attribut) als auch Regressionsbäume (N_{opt} als numerisches Attribut) produzieren (vgl. Kap. 2.6)⁵⁷.

Im Ergebnis wird das Wissen, das in den neuronalen Netzen zur Ertragsprognose abgebildet

⁵⁷ Die Anwendung des C 5.0-Algorithmus führt zu sehr ähnlichen Ergebnissen in der Klassifikation. Die Unterschiede im Lernverfahren fallen im Rahmen der hier verwendeten Klassifikationsaufgaben kaum ins Gewicht.

ist, in ökonomische Regeln überführt. Dies erfolgt unter Einbußen von Qualität, da eine stetige Funktion in unstetige Regeln abgebildet wird (vgl. Kap. 2.6). Die Anpassung an die Stetigkeit führt in der Anwendung von Entscheidungsbaumverfahren zu zahlreichen Unterteilungen des Lösungsraumes. Im Ergebnis bedeutet dies vielfach, dass die produzierten Regeln sehr unübersichtlich werden und sich damit nicht mehr zur Darstellung und Interpretation eignen. Dies führt zu der Auftrennung der Data-Mining-Ziele, die im Eingang dieses Kapitels angesprochen wurde. Die exakten Regeln können nicht als verständliche Muster angesehen werden, es müssen weitere Werkzeuge zur Interpretation eingesetzt werden.

Zur Darstellung in dieser Arbeit wird die *Maximum-Tree-Depth* auf drei Ebenen beschränkt, um eine übersichtliche Baumstruktur zu erzeugen. Es werden damit nur die wichtigsten Regeln dargestellt. Für die Anwendung in einem In-Field-Controller kann jedoch die optimale Komplexität verwendet werden, um die Präzision hoch zu halten.

Mit der Anwendung dieser drei beschriebenen Schritte können ökonomisch optimierte Entscheidungsregeln zur kleinräumigen Stickstoffdüngung erzeugt werden. Zur Beurteilung müssen allerdings die möglichen Fehlerquellen der einzelnen Schritte betrachtet werden. Funktioniert beispielsweise die Ertragsprognose nicht, können auch keine validen Regeln gefunden werden. Gleichzeitig hängen die erzeugten Entscheidungsregeln wiederum stark von den jeweiligen Versuchsanlagen und den damit verbundenen Merkmalsräumen der Trainingsinstanzen ab. Die erzeugten Entscheidungsregeln decken zwar per Definition alle möglichen Fälle der Realität ab, allerdings darf nicht angenommen werden, dass die Klassifikationen auch für Attributskombinationen valide sind, die im Merkmalsraum der Trainingsinstanzen nicht aufgetreten sind (vgl. Kap. 2.4.). Das bedeutet, dass zur Interpretation wiederum der Merkmalsraum der Trainingsinstanzen herangezogen werden muss.

4.4.4 (Vor-)Überlegungen zur Extraktion von Mustern

Um zu verstehen, welche Herausforderungen sich bei der Umsetzung ergeben, muss etwas ausgeholt werden. Ein erstes Problem ist die Darstellung des WED-Prozesses. Da es sich dabei um einen interaktiven Prozess mit möglichen Rückkopplungsschleifen handelt, ist eine lineare Darstellung schwierig. Um die Übersichtlichkeit zu erhalten, beschränkt sich die Argumentation auf Ausschnitte dieses Prozesses.

Ein weiteres Problem stellt die Auswahl von überlegenen Varianten der Datenaufbereitungen und den damit verbundenen Modellen dar. Zur Bewertung von Datenaufbereitungen und den damit erzeugten Modellen wird eine Zielgröße bzw. ein Zielsystem benötigt. Die Maßzahlen für die Prognosegüte (vgl. Kap. 2.7) beziehen sich allerdings nur auf die jeweilige Datenaufbereitung. Das bedeutet, dass die Methode der Datenaufbereitung zum Training eines Modells auch für die Validierungsdaten dieser Datenaufbereitung angewendet

wird. Damit können streng genommen nur Modelle mit derselben Datenaufbereitung verglichen werden.

Für den Vergleich unterschiedlicher Datenaufbereitungen müssen weitere Punkte beachtet werden, die sich nur aus dem Anwendungsfall ergeben: Beispielsweise können in der Datenaufbereitung bestimmte Regionen des Merkmalsraumes ausgeschlossen werden, falls die Datendichte nicht ausreichend erscheint. Dadurch wird die Präzision der Vorhersage u. U. erhöht. Falls in einer späteren Anwendung eine Prognose für diesen Bereich erfolgen sollte, schneidet allerdings in der Regel das Modell besser ab, zu dessen Training auch Instanzen aus diesem Bereich verwendet wurden. Ähnliches gilt für die verwendeten Input-Attribute.

Die Zielgröße „Prognosegüte“ ist nicht immer kongruent mit dem Ziel des WED-Prozesses. Das Ziel dieser Arbeit ist es beispielsweise, Entscheidungsregeln für die Stickstoff-Ausbringung zu finden. Damit müssen gegebenenfalls Datenaufbereitungsmethoden bevorzugt werden, die den Einfluss der Stickstoff-Applikation in den Vordergrund rücken. Es kann es durchaus vorkommen dass die Vorhersagekraft der N-Gaben auf den Ertrag gering ist und unter Umständen von Pruning-Algorithmen abgeschnitten wird. In diesem Fall muss auf Präzision verzichtet werden, um den Zusammenhang von N-Applikation auf den Ertrag modellieren zu können.

Ein weiteres Problem ist, dass nicht alle Möglichkeiten in der Datenaufbereitung und Modellierung einer Systematik folgend „durchgetestet“ werden können. Die Erstellung eines neuronalen Netzes nimmt bei vorliegenden Datenumfängen unter Verwendung des *Exhaustive-Prune-Algorithmus* bis zu 2 Stunden Rechenkapazität in Anspruch⁵⁸. Um die Ergebnisse einer Datenaufbereitung in Verbindung mit einer Attributauswahl und verschiedenen Trainingsparametern zu testen, werden in einer fünffachen Kreuzvalidierung fünf neuronale Netze trainiert. Die Auswahl möglicher Input-Attribute und Datenaufbereitungen lassen die Zahl der Varianten explodieren. Weiterhin empfiehlt es sich insbesondere bei beschränktem Datenumfang, eine Methode der Datenauswahl wiederholt zu testen, um zu vermeiden, dass die Approximation an lokalen Minima abgebrochen wird. All diese Möglichkeiten führen dazu, dass die zur Verfügung stehende Zeit das begrenzende Element für einen WED-Prozess ist (vgl. Kap. 5.2).

Zusammenfassend bleibt festzuhalten, dass ex ante keine Struktur eines WED-Prozesses vorgegeben werden kann und kein objektives Zielsystem für diesen Prozess existiert. Der Wert der Prognose erschließt sich oftmals nur im Kontext, bzw. im Anwendungsfall. Dadurch besteht die Gefahr, dass subjektive Entscheidungen des „Data Miners“ in den Vordergrund rücken.⁵⁹ In dieser Arbeit werden aus diesen Gründen verschiedene Indizien bzw. Orientierungspunkte zur Beurteilung des Prozesses verwendet, die im Folgenden erläutert

⁵⁸ Verwendet wurde ein 1800 MHz Prozessor, Intel© Pentium© M.

⁵⁹ Vgl. hierzu auch Kap. 5.1.

werden.

Datenbeschreibung und -visualisierung

Zum einen wird für jeden Versuchsdatensatz eine Übersicht der einzelnen Attribute erstellt, in der Kennzahlen der deskriptiven Statistik sowie ein Histogramm der Merkmalsausprägungen abgetragen sind. Diese Funktion wird als *Data Audit (Software Clementine ®)* bezeichnet. Weiterhin werden die linearen Zusammenhänge der Attribute untereinander in einer Korrelationsmatrix dargestellt, um erste Zusammenhänge in den Daten zu erkennen. DRUMMOND ET AL. (2002, s.p.) weisen insbesondere auf die Analyse der Datendichte hin. Erst durch eine gewisse Anzahl von Datenpunkte in einem Bereich kann gewährleistet werden, dass für diesen Bereich auch eine gute Prognose mit einem neuronalen Netzwerk erstellt wird. Erst dann wird zur eigentlichen Modellierungsphase übergegangen.

Analyse der neuronalen Netze

In einem ersten Schritt werden mehrere Datenaufbereitungsverfahren getestet, die möglichst alle Attribute einschließen. Dazu werden Neuronale Netze mit dem *Exhaustive Prune-Algorithmus* (vgl. Kap.2.5.3) und einer hälftigen Aufteilung des Datensatzes in Trainings- und Testsample trainiert.

Die Werte der *Estimated Accuracy* (vgl. Tabelle 1) geben einen Anhaltspunkt für die Vorhersagekraft des Netzes. Die Prognosegüte wird anhand des Testsamples gemessen. Da jedoch dieses Sample nicht unabhängig ist, ist der absolute Wert nicht notwendigerweise objektiv (vgl. Kap. 2.7).

Die Anwendung des *Exhaustive-Prune-Algorithmus'* liefert jeweils eine optimierte Netztopologie. Daraus ergeben sich Interpretationsmöglichkeiten für die gelernten Zusammenhänge. Von Bedeutung sind dabei in erster Linie die Input-Attribute, die in dem ausgegebenen Netzwerk verwendet werden. Damit können „unwichtige“ Input-Attribute erkannt werden. Generell gilt, dass mit zu vielen „unwichtigen“ Input-Attributen Überlernen riskiert wird. Ähnliches gilt für die Komplexität der Topologie: Komplexe Netzformen neigen zum Überlernen, sehr kompakte Netztypen neigen zu einer Vereinfachung der Zusammenhänge.

Ein wichtiger Anhaltspunkt ist weiterhin die *Relative Importance*, das Ergebnis der Sensitivitätsanalyse (vgl. Kap. 2.5.3). Dadurch können unwichtige Attribute von wichtigen unterschieden werden, und somit eine Auswahl für weitere Datenaufbereitungsmethoden erreicht werden. Nacheinander können dann verschiedene Attribute zusammengefasst bzw. ausgeschlossen werden. In dieser Arbeit werden die Eckpunkte der trainierten neuronalen Netze jeweils angegeben. Ein Beispiel dafür ist in Tabelle 13 an einem Beispiel dargestellt. Die Einstellungen beziehen sich dabei auf die Optionen des Trainingsalgorithmus'(vgl. Kap. 2.5).

Die Optimierung des Merkmalsraumes in Bezug auf die *Datenbalance* wird in dieser Arbeit

nur eingeschränkt verwendet. Der Merkmalsraum wird nur in Bezug auf die N-Mengen optimiert, da die Abdeckung der Merkmalsräume für dieses Attribut in der Regel nicht optimal war. Eine *Datenbalance* kann unter Umständen die Präzision der Vorhersage senken.

Zum Vergleich verschiedener neuronaler Netze wird eine fünffache *Kreuzvalidierung* angewendet. Dazu wird ein wechselndes Validierungssample erzeugt, das nur zur Ermittlung der Prognosegüte verwendet wird. Dieses Sample wird vor der Modellierung von den anderen Datensätzen getrennt und erst zur Ermittlung der Leistungsvorhersage benutzt (vgl. Kap. 2.7.1).

Tabelle 13: Analyse eines trainierten Netzes – vorgestellt an einem Beispiel

Estimated accuracy: 90.71	Relative Importance of Inputs:	Einstellungen:
Input Layer: 4 neurons	<i>N_2</i> 0.30	Methode: Exhaustive prune
Hidden Layer 1: 3 neurons	<i>EM_38</i> 0.28	Stop on: Default
Hidden Layer 2: 1 neuron	<i>Zugkraft:</i> 0.18	Prevent overtraining: true
Output Layer: 1 neuron	<i>Ertrag_98:</i> 0.17	Testsample: 50 %
Analyse des Netzwerkes und der Vorhersagekraft	Sensitivitätsanalyse	Parameter des Netzwerk-Trainings

Response Surfaces

Eine Möglichkeit zur Analyse und zur Extraktion von Wissen aus der „Black Box“ eines trainierten neuronalen Netzwerkes stellen *Response Surfaces* dar (POCRAJAC UND OBRADOVIC, 2001, S. 31, DRUMMOND ET AL., 2002, s.p.; LIU ET AL., 2001, S. 711). Mit dieser Methode können die funktionalen Zusammenhänge von maximal zwei Input-Attributen in Bezug auf das Output-Attribut visualisiert werden. Für die vorliegende Arbeit ist dabei die Wirkung der Stickstoffmenge auf den Ertrag unter Variation eines teilflächenspezifischen Attributes besonders interessant. Dazu wird ein Test-Datensatz konstruiert, mit dem systematisch Input-Kombinationen simuliert werden. Vom Merkmalsraum des Trainingsdatensatzes werden die Bereiche ausgewählt, die von der Datenbasis angemessen vertreten sind.⁶⁰

Die betrachteten Attribute werden dazu in diskrete Werte unterteilt und in ihren Linear-kombinationen in Tabellenform abgetragen (vgl. Abbildung 29). Die daraus entstehenden Input-Instanzen werden dann zusammen mit den restlichen Attributen, die konstant am Mittelwert gehalten werden, in das trainierte Modell eingegeben und der Output (prognostizier-

⁶⁰ Es existieren hierzu keine Regeln, aber für Bereiche, die außerhalb des Merkmalsraumes oder in schlecht vertretenen Bereichen liegen, können keine guten Ergebnisse erwartet werden.

ter Ertrag) gespeichert. Die Ergebnisse lassen sich in einer dreidimensionalen Grafik darstellen.

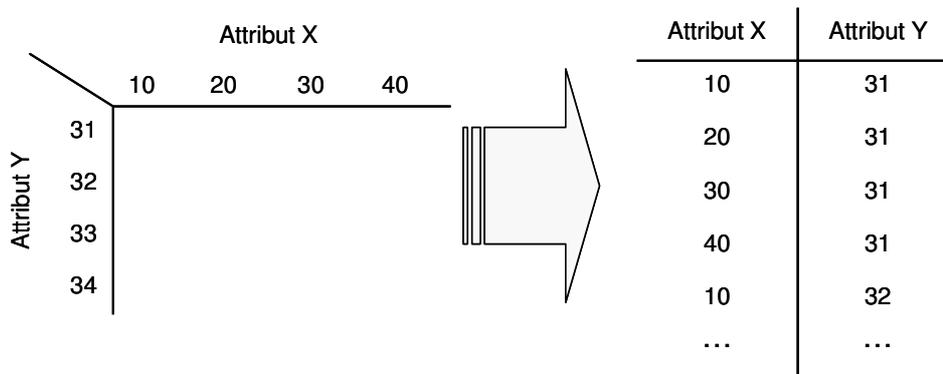


Abbildung 29: Konstruktion eines Testvektors zur Erstellung von *Response Surfaces*

Der Nachteil dieser Methode ist darin zu sehen, dass sich über *Response Surfaces* ausschließlich die Zusammenhänge zweier Input-Variablen auf den Output des neuronalen Netzes untersuchen lassen können – die restlichen Input-Werte werden im Test-Datensatz konstant am Mittelwert gehalten. Hierbei besteht die Gefahr für Fehlinterpretationen: Die Oberfläche könnte bei Variation der nicht betrachteten Input-Attribute ganz andere Formen annehmen. Diese Gefahr steigt insbesondere dann an, wenn die „nicht beachteten“ Attribute einen großen Einfluss auf die Prognose haben (vgl. *Relative Importance*) bzw. wenn komplexe Netzstrukturen verwendet werden.

Als Erweiterung dieser Methode können die Linearkombinationen dreier Input-Attribute in das trainierte neuronale Netzwerk eingegeben werden. Mit Hilfe einer Pivot-Tabelle lassen sich damit die entsprechenden Prognosen für zwei Input-Attribute bei ausgewählten Werten des dritten Input-Attributes bestimmen. Alternativ kann auch der durchschnittliche Einfluss des dritten Attributes ausgegeben werden.

In dieser Arbeit werden *Response-Surfaces* überwiegend zur Visualisierung der Abhängigkeit von Düngermengen auf den Ertrag bei den verschiedenen teilflächenspezifischen Attributen verwendet. Die Abbildung stellt also ein Bündel an Produktionsfunktionen, also eine Art Produktionsoberfläche, dar. Ein Beispiel dafür ist in Abbildung 30 dargestellt. An der x-Achse ist ein teilflächenspezifisches Attribut abgetragen, an der y-Achse die Düngermenge einer Teilapplikation. Der Output des neuronalen Netzes, d. h. die Ertragsprognose ist auf der z-Achse abgetragen. Aus diesem Diagramm kann der Ertragsverlauf bei einer Veränderung der Eingangs-Attribute abgelesen werden. Zur besseren Orientierung sei angemerkt, dass die in x-Achsen-Richtung verlaufenden Linien in der Ertragsoberfläche jeweils eine konstante Düngermenge darstellen, die in 10er-Schritten erhöht wird. Gleichzeitig wird der prognostizierte Ertrag in Klassen aufgeteilt und den einzelnen Klassen Farben zugewiesen (rechte Legende). In der Oberfläche stellen die Trennlinien zwischen den Klassen

somit Höhenlinien dar. Aus dieser Beispiels-Abbildung kann beispielsweise gefolgert werden, dass eine Erhöhung der N-Applikation bei niedrigen *REIP_2*-Werten zu einem Ertragsanstieg führt, während höhere N-Mengen bei hohen *REIP_2*-Werten praktisch keine Ertragswirkung zeigen. Es können also Muster bzw. Ansätze für eine teilflächenspezifische Applikation abgeleitet werden. Allerdings muss untersucht werden, ob sich diese Muster auch zeigen, falls die restlichen Input-Attribute nicht am Mittelwert eingegeben werden (Pivot-Tabelle). Mit dieser Methode können somit Zusammenhänge (bzw. Muster) interpretiert werden, die in einem trainierten neuronalen Netz abgebildet sind.

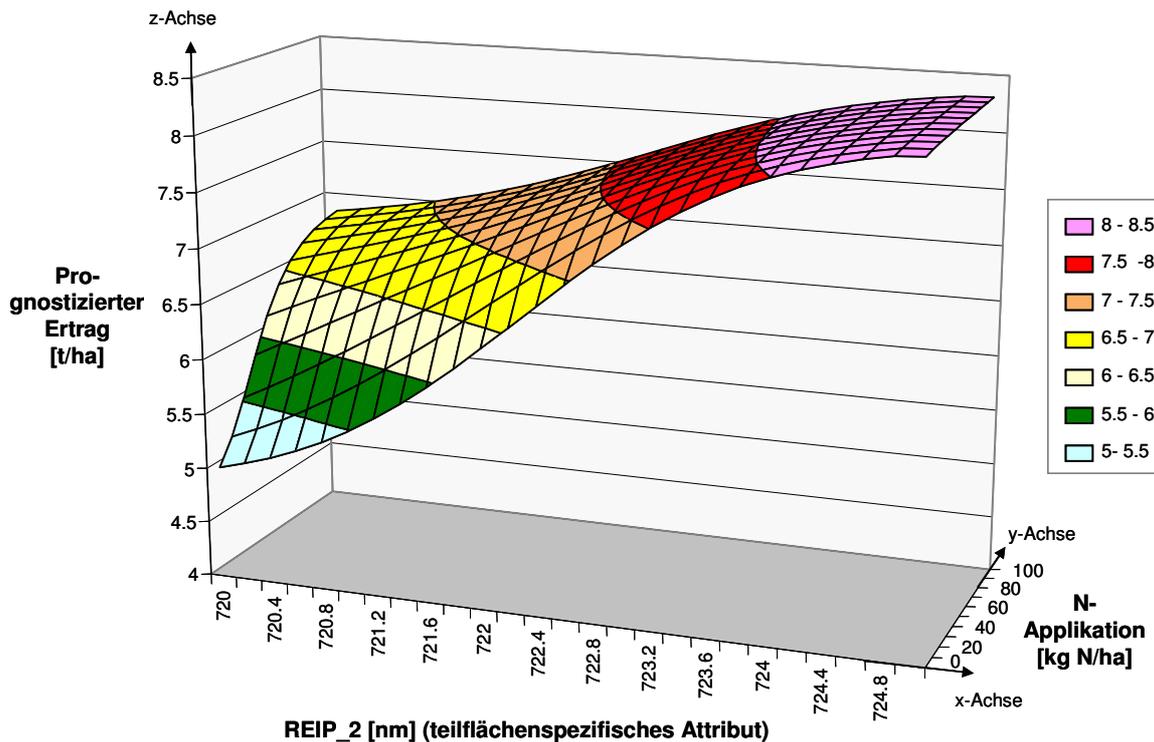


Abbildung 30: *Response Surfaces* zur Analyse neuronaler Netze

Responses Surfaces sind insbesondere deshalb interessant, weil hier vorhandenes Wissen über Produktionsfunktionen genutzt werden kann, um die Modelle zu beurteilen oder zumindest vergleichen zu können. Ein Problem besteht jedoch darin, dass sich die in der Literatur dargestellten Produktionsfunktionen überwiegend auf die Stickstoff-Gesamtmenge beziehen während in der vorliegenden Arbeit der Effekt von Teilmengen abgebildet wird. Mit der Berechnung der NKL (vgl. Kap. 4.4.3) kann auf der Z-Achse auch diese Zielgröße abgetragen werden. Damit lassen sich aus dem Diagramm aus den jeweiligen Maxima ablesen, welche Düngermenge zu einem maximalen ökonomischen Ergebnis führt.

Entscheidungsregeln

Als weiteres Indiz dienen die Entscheidungsregeln selbst, die wie beschrieben erstellt werden. Durch die Beschränkung auf drei Ebenen soll erreicht werden, dass nur die wichtigsten Regeln abgebildet werden. Zur leichteren Interpretation wird an jeder Abzweigung angege-

ben, ob die Anwendung der Regel zu einem höheren, niedrigeren oder gleich bleibenden (prognostizierten) N_{opt} führt. In dieser Arbeit werden zur Erzeugung der Entscheidungsregeln die Teilflächen des Schlages verwendet, der zur Prognose des neuronalen Netzes verwendet wurde⁶¹.

Mit diesen „Werkzeugen“ kann die Data-Mining-Phase bis zu einem gewissen Grad beurteilt und gesteuert werden. Der genaue Ablauf orientiert sich jedoch an den jeweiligen Daten. In den folgenden Abschnitten werden für jeden dieser Versuche die zentralen Schritte der Exploration und der Data-Mining-Ergebnisse aufgezeigt.

4.5 Data Mining I: Versuch *Schafhof 02*

Der erste Versuch zur Erzeugung von Entscheidungsregeln erfolgt mit den Daten des Versuches *Schafhof 02*. Dieser Versuch wurde ursprünglich nur zum Testen von Düngesystemen im Streifenversuchsdesign angelegt. In ungefähr der Hälfte aller Instanzen wurde die N-Gabe schlageinheitlich konstant gehalten, während in der anderen Hälfte zwei Algorithmen angewendet wurden, die zur zweiten Applikation identisch und zur dritten Applikation ähnlich waren. Dies wirkt sich negativ darauf aus, geeignete Instanzen für ein überwachtes Lernen zu liefern. Dennoch soll die Modellierung und insbesondere die Exploration auch an diesem Datensatz vorgestellt werden, um zu einem besseren Verständnis für die Anwendung der Methode zu gelangen.

4.5.1 Datenexploration

Aus der Versuchsanlage *Schafhof 02* konnten insgesamt 412 Instanzen erzeugt werden. In Abbildung 31 ist der *Data Audit* dieser Datensätze abgebildet. Die durchschnittlichen Erträge des Versuchsjahres (*Ertrag_02*) dieses Schlages bewegen sich im Bereich von etwa 8-9 t/ha. Die Standardabweichung der Erträge bei einheitlicher Düngung in den Jahren 1996 und 1998 liegt bei 0.59 bzw. 0.84. Die Heterogenität des Schlages ist damit als gering einzuschätzen. Sichtbar werden auch fehlende Werte für *Ertrag_96* und *Ertrag_98* und für den Ertrag des Versuchsjahres 2002, die in der Datenvorverarbeitung entfernt wurden. Es handelt sich hierbei um insgesamt 8 Instanzen, die von der Modellierung ausgenommen wurden. Weiterhin konnte ein auffälliger Ausreißer im *Ertrag_02* identifiziert werden⁶². Die entsprechende Instanz wurde ebenfalls entfernt. Damit gelangt man zu 403 Instanzen mit vollständigen Bodenattributen. In diesem Fall wurde darauf verzichtet, leere Werte zu ersetzen, da es sich um relativ wenige Instanzen handelte, die ausgeschlossen werden mussten. Gravierender stellt sich das Problem in Bezug auf die Spektralmessungen dar, da zur

⁶¹ Dies führt dazu, dass sich die Entscheidungsregeln auch den natürlichen Gegebenheiten anpassen. Wahlweise könnten nur bestimmte Bereiche eines Schlages ausgewählt werden oder künstliche Teilflächen erzeugt werden, was jedoch i. d. R. zu sehr verschiedenen Regeln führt.

⁶² Es handelt sich hierbei um eine Teilfläche am Rand des Schlages.

zweiten N-Gabe keine Spektralmessung durchgeführt wurde. Das Attribut *REIP_2* ist damit in allen Instanzen fehlend und es können prinzipiell keine Regeln für einen Sensor-Overlay-Ansatz zur zweiten N-Gabe generiert werden. Das Attribut *REIP_3* wurde nur für ungefähr die Hälfte aller Instanzen erfasst.

In Bezug auf die Modellierung des Einflusses der N-Düngung hat dieser Datensatz mehrere gravierende Unzulänglichkeiten: Da der Schlag in diesem Versuch einheitlich angedüngt wurde, existiert nur eine Ausprägung des Attributes *N_1*. Weiterhin ist die Variation der Stickstoff-Gaben zur zweiten und dritten N-Gabe sehr gering. Sie erstreckt sich jeweils nur über einen Bereich von 30 kg N/ha mit Ab- und Zuschlägen von 10 kg N/ha ausgehend von der Mitteltragszone. Der mittlere Wert ist dabei unverhältnismäßig oft vertreten.

Damit wird deutlich, dass dieser Versuch **nicht** für eine Ertragsprognose für variierende N-Mengen in der **ersten N-Gabe** verwendet werden kann, da diese Ereignisse außerhalb des Merkmalsraumes der Lerninstanzen liegen (vgl. Kap. 2.4).

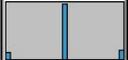
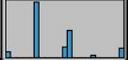
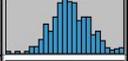
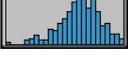
Attribut	Histogramm	Minimum	Maximum	Mittelwert	Std. Abw.	Anzahl
N_1		70	70	-	-	412
N_2		40	60	50.29	4.51	412
N_3		50	90	65.97	9.28	412
REIP_3		720.2	731.26	727.22	1.52	201
ZUGKRAFT		15.67	37.12	27.46	3.35	412
EM_38		19.25	50.75	36.22	5.14	412
ERTRAG_96		5.26	9.2	7.51	0.59	411
ERTRAG_98		5.55	10.01	8.32	0.84	409
ERTRAG_02		5.2	9.31	7.81	0.66	408

Abbildung 31: *Data Audit* des Datensatzes *Schafhof 02*

Eine Ertragsprognose zur zweiten N-Gabe kann nur auf drei Ausprägungen (40, 50 und 60 kg N/ha) der zweiten N-Applikation aufbauen. Die Auswirkungen von Applikationen außerhalb dieser Werte kann nur extrapoliert werden. Die Gefahren, die mit einer Extrapolation verbunden sind, zeigen sich in der Analyse des trainierten neuronalen Netzwerkes (vgl. Kap. 4.5.2). Die Verteilung der Merkmalsausprägungen von *N_3* ist ebenfalls unzureichend. Es liegen nur Ausprägungen im Bereich von 50 – 90 kg N/ha vor, wobei der Wert der schlageinheitlichen Düngermenge unverhältnismäßig oft vorkommt.

Das Hauptproblem des Datensatzes wird jedoch aus der Korrelationsmatrix in Abbildung

32 deutlich. In den Instanzen, in denen die N-Menge variiert, wurden überwiegend ähnliche Algorithmen angewendet (vgl. Kap. 4.2.1). Die Applikationsmengen resultieren größtenteils aus dem verwendeten Mapping-Ansatz. Damit wird der Ereignisraum stark eingeschränkt. Es gibt sehr wenige Instanzen, die hohe historische Erträge und gleichzeitig niedrige Applikationen zur zweiten N-Gabe aufweisen. Gleichzeitig existieren wenige Instanzen für den umgekehrten Sachverhalt, d. h. niedrige historische Erträge und hohe N-Applikation. Dadurch zeigt sich ein relativ starker Zusammenhang zwischen (historischen) Ertragsattributen und den Teilapplikationen (N_2 und N_3). Es handelt sich hier um einen Zusammenhang, der nur durch die Art der Versuchsdurchführung zustande kommt.

Die Abbildung einer verzerrungsfreien Ertragsprognose für Teilapplikationen wird dadurch erschwert, bzw. unmöglich gemacht. Vereinfacht dargestellt kann das neuronale Netz beispielsweise nicht unterscheiden, ob ein höherer Ertrag durch die N-Menge oder durch die Bodenattribute hervorgerufen wird, wenn zwischen diesen beiden Attributen ein starker Zusammenhang besteht. Insbesondere wird eine Ertragsprognose für variierende N-Gaben kritisch.

	N_2	N_3	REIP_3	ZUGKRAFT	EM_38	ERTRAG_96	ERTRAG_98	ERTRAG_02
N_2		0.3	0.52	-0.15	0.01	0.38	0.3	0.46
N_3	0.3		0.04	0.04	-0.02	0.12	0.13	0.17
REIP_3	0.52	0.04		-0.3	0.28	0.16	0.31	0.73
ZUGKRAFT	0.15	0.04	-0.3		0.06	-0.15	-0.16	-0.29
EM_38	0.01	0.02	0.28	0.06		0.05	0.15	0.3
ERTRAG_96	0.38	0.12	0.16	-0.15	0.05		0.61	0.35
ERTRAG_98	0.3	0.13	0.31	-0.16	0.15	0.61		0.41
ERTRAG_02	0.46	0.17	0.73	-0.29	0.3	0.35	0.41	

Abbildung 32: Korrelationsmatrix der Attribute Schafhof 02.

Das Problem wird durch die „unscharfe“ Einteilung von Management-Zonen (vgl. Kap. 4.2.1), wie in Abbildung 33 zu erkennen ist, etwas abgeschwächt. Hier sind in einem Scatterplot die Instanzen des Versuches in ihren Attributen *Ertrag_96* sowie *Ertrag_98* abgetragen. Farb- und Formcodiert ist in dieser Abbildung außerdem die Applikation N_2 dargestellt. Es zeigen sich die angesprochenen Gruppierungen der N_2 -Ausprägungen, die das Lernverhalten behindern. Die blauen Dreiecke (Applikation von 40 kg N/ha) liegen überwiegend näher am Ursprung als die magentafarbenen Kreise (Applikation von 60 kg N/ha). Bei einer Düngung nach einem Mapping-Ansatz auf Basis der einzelnen Teilflächendaten müsste dies allerdings durchgängig der Fall sein. Die Verwendung von Management-Zonen

erzeugt somit eine unscharfe Trennung dieser beiden Gruppen. Zur dritten N-Gabe wird die Verzerrung durch die Anwendung eines Sensor-Overlay-Algorithmus' weiter abgeschwächt.

Prinzipiell könnte eine derartige Verzerrung durch eine entsprechende Selektion von Instanzen weiter verringert werden, indem die Anzahl von Instanzen mit hohen Düngermengen in Hohertragsbereichen beschränkt werden würde. Allerdings würde dann die Zahl der Instanzen sehr niedrig werden. Zum Training der neuronalen Netze werden deshalb alle Instanzen verwendet.

Von den beiden Düngergaben N_2 und N_3 hat die frühere Applikation eine stärkere Korrelation mit dem Ertrag im Versuchsjahr. Der Vegetationsindex $REIP_3$ korreliert sehr stark mit dem Ertrag im Versuchsjahr und ist damit für eine Ertragsprognose prinzipiell gut geeignet. Die Korrelation der Bodenattribute mit $Ertrag_{02}$ ist insgesamt schwächer und liegt im Bereich zwischen 0.3 und 0.4. Auffallend ist, dass das Attribut *Zugkraft* negativ mit dem Ertrag des Versuchsjahres korreliert.

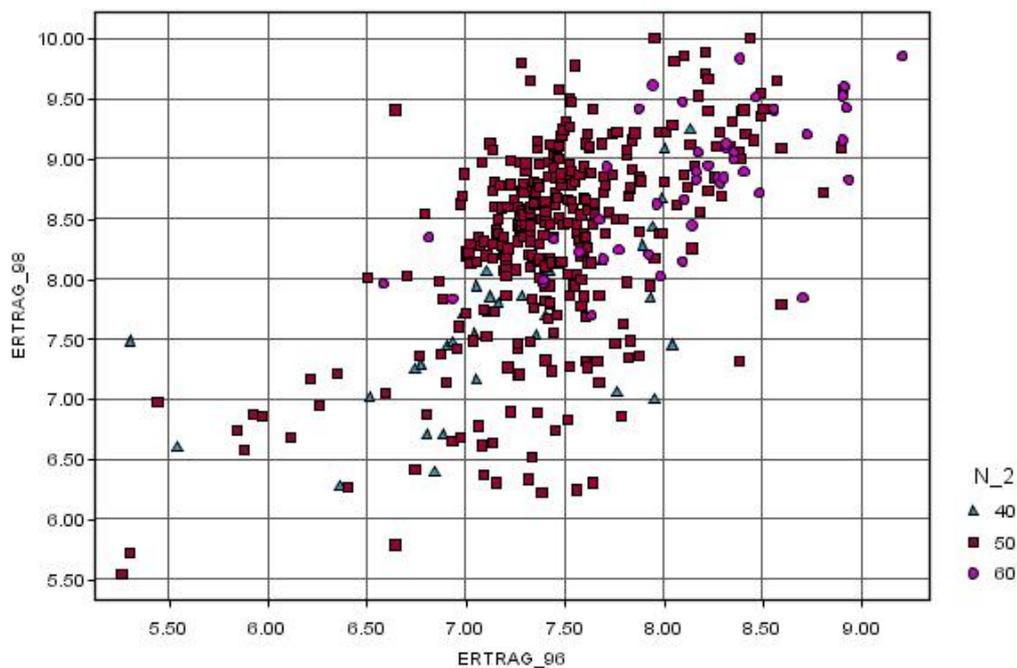


Abbildung 33: Scatterplot aus Ertragsdaten und Applikationsmengen (N_2)

4.5.2 Ertragsprognose für die zweite N-Applikation

In einem ersten Schritt werden alle Instanzen mit allen zum Zeitpunkt N_2 vorhandenen Attribute zum Training eines neuronalen Netzes verwendet. Das Ergebnis der Trainingsphase ist in Tabelle 14 dargestellt. Die *Estimated Accuracy* dieses Netzwerkes liegt bei 90.71. Das erstellte Netz hat eine relativ kompakte Topologie mit vier Neuronen in der Eingabeschicht, drei Neuronen in der ersten versteckten Schicht, einem Neuron in der zwei-

ten versteckten Schicht sowie einem Neuron in der Ausgabeschicht. Das Attribut *Ertrag_96* wurde durch den Pruning-Algorithmus abgeschnitten. Bei der Sensitivitätsanalyse fällt auf, dass das Attribut *N_2* einen sehr hohen Einfluss auf die Prognose hat, während das Attribut *Ertrag_98* von geringer Bedeutung für die Prognose ist. Zur Interpretation kann wiederum die Verzerrung im Datensatz herangezogen werden. Durch die (Schein-)Korrelation von *N_2* mit den historischen Ertragsattributen könnten die Auswirkungen auf den Ertrag über das Attribut *N_2* im neuronalen Netz abgebildet worden sein. Die scheinbare Bodenleitfähigkeit *EM_38* besitzt ebenfalls eine hohe Bedeutung für die Prognose.

Im nächsten Schritt wird die Funktionalität des neuronalen Netzwerkes in einem *Response Surface* untersucht. Dazu werden die Bereiche des Merkmalsraumes verwendet die ausreichend abgedeckt sind⁶³. Dieser Ausschnitt wird durch die Grundfläche des Diagramms repräsentiert.

Tabelle 14: Neuronales Netzwerk zur zweiten N-Applikation (Versuch Schafhof 02)

Estimated accuracy: 90.71	Relative Importance of Inputs:	Einstellungen:
Input Layer: 4 neurons	<i>N_2</i> 0.30	Method: Exhaustive prune
Hidden Layer 1: 3 neurons	<i>EM_38</i> 0.28	Stop on: Default
Hidden Layer 2: 1 neuron	<i>Zugkraft</i> : 0.18	Prevent overtraining: true
Output Layer: 1 neuron	<i>Ertrag_98</i> : 0.17	Testsample: 50 %

In Abbildung 34 sind die entsprechenden Ertragsprognosen für diesen Ausschnitt graphisch dargestellt. In diesem Diagramm werden in diesem Fall „Produktionsfunktionen“ abgetragen, die sich auf die Teilgabe *N_2* beziehen. Es kann abgelesen werden, wie sich die Produktionsfunktionen verändern, wenn das teilflächenspezifische Attribut *Ertrag_98* verändert wird (vgl. Kap. 4.4.4). Dabei zeigt sich eine stufenförmige Oberfläche. In den Bereichen, der nicht von Trainingsdaten abgedeckt wird, liegt die N-Response auf einem Plateau. Dies ist sowohl im Bereich $N_2 < 40$ kg N/ha als auch $N_2 > 60$ kg N/ha der Fall. Im Bereich zwischen 40 und 60 kg N/ha steigt der prognostizierte Ertrag für alle Werte von *Ertrag_98* enorm steil an. Aus den zugrunde liegenden Daten ergibt sich ein durchschnittlicher Ertragszuwachs von 0,68 t/ha bei einer Erhöhung von *N_2* von 40 auf 50 kg N/ha. Diese Werte können als realitätsfern betrachtet werden.

Der gesamte Verlauf der hier abgebildeten Produktionsfunktionen ist extrem untypisch und kann somit verworfen werden⁶⁴. Dieses Ergebnis zeigt deutlich, dass die Rahmenbedingungen, die auch schon in der Datenexploration angesprochen wurden zu Fehlinterpretationen

⁶³ In diesem Fall werden *Ertrag_98*-Werte von 7-9 verwendet. *N_2* wird von 0 bis 100 variiert, um die Interpolationsmöglichkeiten eines neuronalen Netzwerkes darzustellen.

⁶⁴ Vgl. Ergebnisse von N-Steigerungsversuchen im selben Jahr in der Region in GANDORFER ET AL. (2004a, S.31).

führen können. In diesem Fall scheinen die Zusammenhänge der Teil-Applikationen mit den Bodenattributen zu einer Überschätzung der Ertragswirkung von Stickstoff zu führen. Gleichzeitig zeigt dieses Beispiel auch sehr deutlich, dass die Fähigkeit neuronaler Netze auf Bereiche außerhalb der Trainingsdaten zu extrapolieren, als kritisch zu bewerten sind. Der plateauförmige Verlauf zeigt praktisch keine Reaktion auf eine variierende N-Gabe.

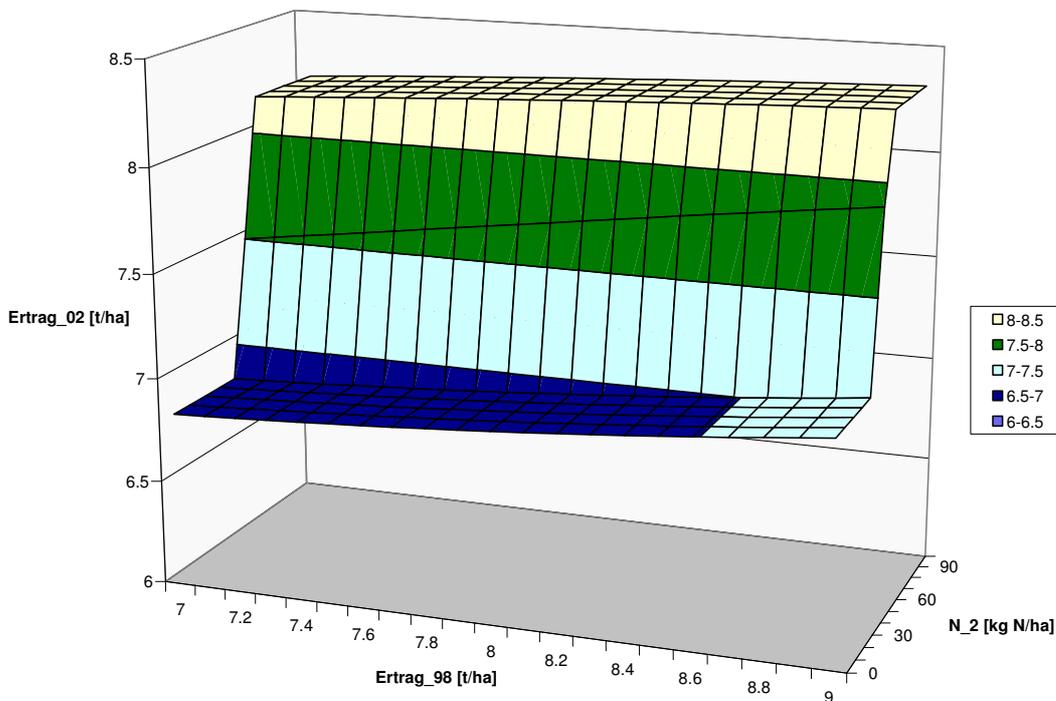


Abbildung 34: Response Surface zur zweiten N-Gabe (Ertrag_98, Versuch Schafhof 02)

Dies wurde auch von DRUMMOND ET AL. (2002, s.p.) bei ähnlicher Datenlage festgestellt:

“This example demonstrates the danger of extrapolation from such models. Neural Network models generalize far better when the data used to train them contains many observations. Where training data was relatively sparse (or non-existent) the models produced unreasonable results.”

Damit wird die Bedeutung einer exakten Exploration als Ausgangsbasis für eine Modellierung und für eine Interpretation der Ergebnisse offenkundig. Weiterhin macht dieses Beispiel die Notwendigkeit einer Variante deutlich, die den Merkmalsraum aus Sicht der N-Gabe verzerrungsfrei abdeckt.

4.5.3 Ertragsprognose zur dritten Stickstoff-Applikation

Die Vorzeichen für eine Modellierung der Zusammenhänge von N_3 mit dem Ertrag des Versuchsjahres sehen eindeutig besser aus. Zum einen gibt es einen größeren Bereich an Merkmalsausprägungen (50-90 kg N/ha), zum anderen wird die Verzerrung des Datensatzes durch den Mapping-Ansatz durch den hinzukommenden Sensor-Overlay-Algorithmus

in den Versuchsdaten weiter „verwischt“.

Bezogen das Attribute *REIP_3* und der *N_3*-Gabe ist der Merkmalsraum verhältnismäßig gut abgedeckt, wie aus Abbildung 35 zu erkennen ist. Das Attribut *N_3* wurde hierzu zweifach codiert (Farbe und Größe des Symbols)⁶⁵. Hier liegen die Kombinationen dieser beiden Attribute in allen Varianten verhältnismäßig gleichmäßig vor. Es tauchen keine größeren Anhäufungen ähnlicher Kombinationen auf. Wiederum zeigt sich in dieser Abbildung der starke Zusammenhang von *REIP_3* und dem Ertrag im Versuchsjahr.

Für das Training des neuronalen Netzes werden nur vollständige Instanzen verwendet. Es stellt sich heraus, dass wiederholte Trainingsvorgänge zu unterschiedlichen Netztopologien führen. Das liegt im Wesentlichen an der geringen Anzahl von Datensätzen und an möglichen Gruppierungen von ähnlichen Datensätzen in den zufällig ausgewählten Trainings- und Testmengen. Auffällig ist der hohe Einfluss des Attributs *REIP_3* für die Prognose.

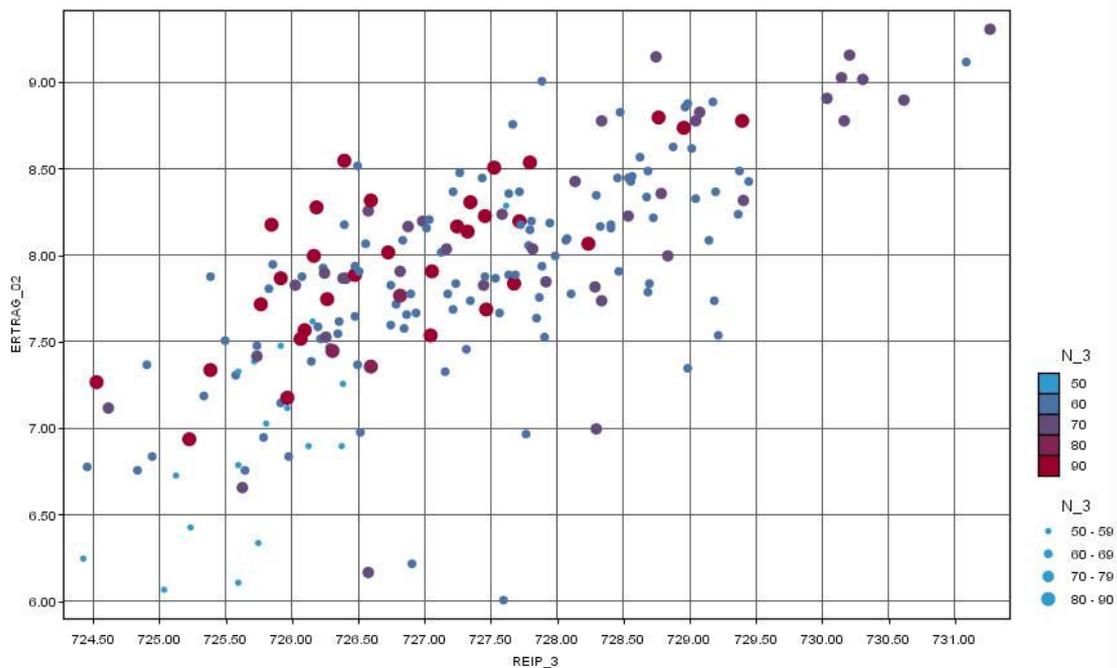


Abbildung 35: Scatterplot der Instanzen Schafhof 02 (*REIP_3*, *Ertrag_02* und *N_3*)

Die *Estimated Accuracy* verbessert sich, wie zu erwarten, wenn mehr Input-Attribute verwendet werden. Allerdings wird damit die Gefahr von Überlernen erhöht. Werden alle Input-Attribute zum Zeitpunkt *N_3* verwendet, so zeigt sich in der Validierung eine hohe Überanpassung auf die Trainings- und Testdaten. Daher werden schrittweise die Eingabe Neuronen entfernt, die die geringsten Werte in der Sensitivitäts-Analyse besitzen, um der geringen Datenquantität Rechnung zu tragen.

Beispielhaft soll ein Netzwerk vorgestellt werden, bei dem die Attribute *N_2* und *Er-*

⁶⁵ Ein niedriger Wert von *N_3* ist beispielsweise mit einem kleinen Kreis **und** einer blauen Farbe dargestellt.

trag_96 nicht als Input-Faktoren verwendet werden. Die Analyse dieses Netzes ist in Tabelle 15 abgetragen. Es zeigt sich eine *Estimated Accuracy*, die schlechter ist als im neuronalen Netz zu *N_2*. Die Topologie des Netzes ist verhältnismäßig komplex. Das Attribut *N_3* besitzt keinen besonders ausgeprägten Einfluss auf die Ertragsprognose. Dies gilt allerdings auch für die Bodenattribute. Die entscheidende Bedeutung besitzt, wie zu erwarten war, das Attribut *REIP_3*.

Tabelle 15: Neuronales Netzwerk zur dritten N-Applikation (Versuch *Schafhof 02*)

Estimated accuracy: 90.19	Relative Importance of Inputs	Einstellungen:
Input Layer: 5 neurons	<i>REIP_3</i> 0.45	Method: Exhaustive prune
Hidden Layer 1: 6 neurons	<i>EM_38</i> 0.25	Stop on: Default
Hidden Layer 2: 3 neurons	<i>Zugkraft</i> : 0.19	Prevent overtraining: true
Output Layer: 1 neuron	<i>N_3</i> 0.16	Sample: 50%
	<i>Ertrag_98</i> : 0.15	

In Abbildung 36 ist die Ertragsprognose für die Attribute *N_3* und *REIP_3* abgetragen. Wiederum können die Unzulänglichkeiten des Trainings-Datensatzes erkannt werden. Außerhalb der N-Mengen im Trainingsdatensatz zeigt sich wiederum ein plateauförmiger Verlauf. Der Anstieg des Ertrages bei höheren N-Mengen fällt mit diesem Modell gemäßiger und damit realitätsnäher aus.

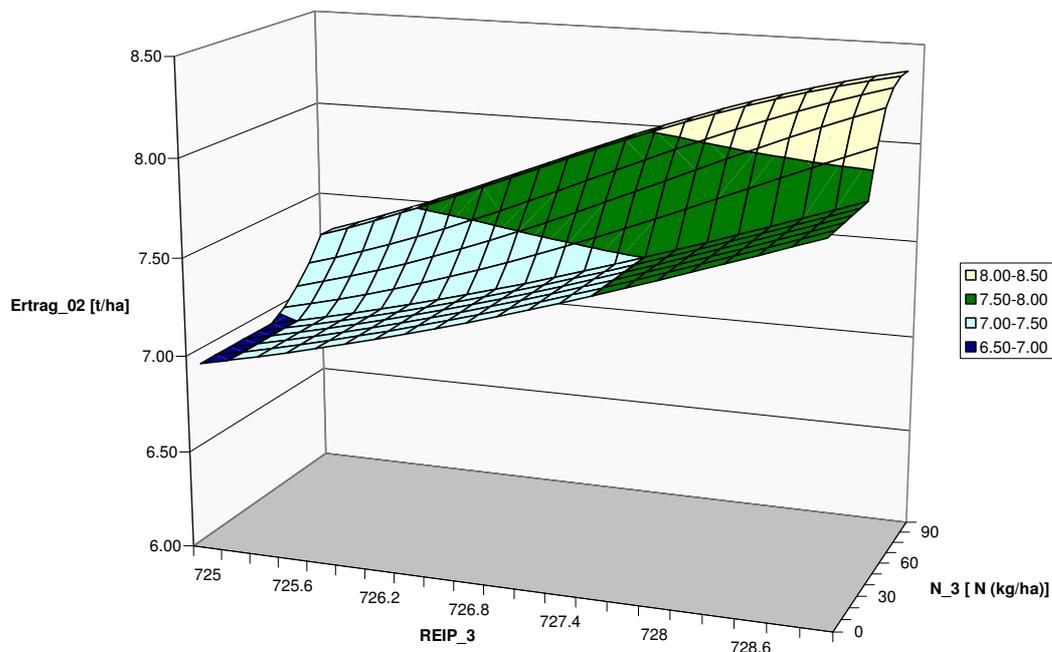


Abbildung 36: Response Surface zur dritten N-Applikation (*REIP_3*, Versuch *Schafhof 02*)

Der Anstieg des prognostizierten Ertrages für höhere *REIP_3*-Werte verläuft annähernd

linear. Auffällig ist der steilere Anstieg der prognostizierten Erträge bei niedrigen *REIP_3*-Werten im Vergleich zu hohen *REIP_3*-Werten. Dies deutet auf eine relativ höhere optimale N-Gabe bei niedrigen *REIP_3*-Werten hin. Diese Ableitung muss allerdings vor dem Hintergrund der beschriebenen Unzulänglichkeiten im Datensatz gesehen werden. Aufgrund dieser Voraussetzungen wird von weiteren Analysen und von der Erzeugung von Entscheidungsregeln abgesehen.

4.5.4 Zwischenfazit

Die wichtigste Lektion aus diesem Versuch ist darin zu sehen, dass Verzerrungen innerhalb des Datensatzes zu gewichtigen Fehlinterpretationen führen können. Korrelationsmatrizen können in der Analyse des Datensatzes dazu wertvolle Aufschlüsse liefern. Aus diesen Ergebnissen heraus wurde für die folgenden Versuche versucht, den Merkmalsraum in Bezug auf die N-Gabe zu „bereichern“. Wie auch schon von anderen Autoren festgestellt, konnte wiederum gezeigt werden, dass neuronale Netze praktisch keine Fähigkeit zur Extrapolation auf Werte außerhalb des Merkmalsraumes haben.

Es konnte auch gezeigt werden, dass *Response Surfaces* sehr nützlich sind, die abgebildeten Muster in den neuronalen Netzen in verständlicher Weise darstellen. Damit kann agronomisches Wissen zur Einordnung bzw. Verwerfung der Modelle herangezogen werden.

Um aus diesem Versuch dennoch Entscheidungsregeln für den Versuch des folgenden Jahres (*D4 03*) abzuleiten, wurden die Instanzen notgedrungen mit denen eines Stickstoff-Steigerungsversuches kombiniert. Ohne näher auf diese Daten einzugehen, erwies sich diese Idee nach heutigem Kenntnisstand als nicht zielführend: Zum einen standen für den N-Steigerungsversuch nicht dieselben Attribute zur Verfügung. Es fehlen beispielsweise Zugkraft- und Bodenleitfähigkeits-Messungen sowie historische Ertragsdaten. Dies führt zu zahlreichen Instanzen mit fehlenden Werten, die die Modellierung beeinträchtigen. Weiterhin wurde der Parzellenversuch wie üblich auf annähernd homogenen (Boden-) Verhältnissen durchgeführt, sodass Versuche, diese fehlenden Werte zu ersetzen, zu zahlreichen Instanzen mit denselben Ausprägungen der Bodenattribute führte. Dadurch wird der Merkmalsraum stark eingeschränkt und das Lernergebnis wiederum fragwürdig. Weitere Probleme ergaben sich dadurch, dass in diesem Versuch die N-Gesamtmenge auf vier Teilgaben verteilt wurde und das Ertragsniveau vom Versuch *Schafhof 02* stark abweicht.

4.6 Data Mining II: Versuch D4 03

Mit dem Versuch *D4 03* wird eine verbesserte Abdeckung des Merkmalsraumes erreicht als im Vorjahresversuch. Insbesondere die Variante *Merkmalsraum* gewährleistet eine gleichmäßigere Datendichte bezüglich der Düngergaben. Allerdings stand in diesem Versuchsjahr eine kleinere Versuchsfläche zur Verfügung. Der daraus entstehende geringe Datenumfang

ist kritisch zu sehen. Insbesondere bei Verwendung der Data-Mining-Algorithmen und dem sequentiellen Eingang der Trainingsinstanzen ist es hier schwierig, stabile Modelle zu entwickeln bzw. eine Auswahl für Modelle zu treffen.

Mit diesem Datensatz kann damit insbesondere untersucht werden, welchen Einfluss die Quantität der Daten auf die Modellierungsergebnisse hat, und ob sich Lösungsmöglichkeiten für dieses Problem ergeben.

4.6.1 Datenexploration

Aus der Versuchsanlage *D4 03* wurden 208 Instanzen erzeugt. Die Ergebnisse des *Data Audit* sind in Abbildung 37 abgetragen. Die drei Stickstoff-Applikationen decken den Merkmalsraum zwar in einer großen Spanne von 0-120 kg N/ha ab, die Verteilung ist jedoch für die Modellierungsaufgabe nicht optimal. Insbesondere die Düngergaben der schlageinheitlichen Variante sind weiterhin dominant vorherrschend.

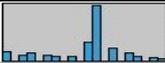
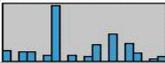
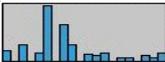
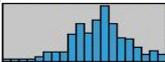
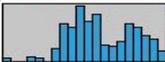
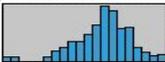
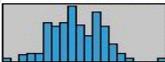
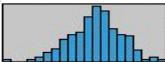
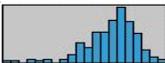
Attribut	Histogramm	Minimum	Maximum	Mittelwert	Std. Abw.	Anzahl
N_1		0	120	60.94	25.75	208
N_2		0	120	55.48	29.51	208
N_3		0	140	52.86	31.79	208
REIP_2		719.16	724.86	722.54	0.99	208
REIP_3		721.88	729.22	727.23	1.21	208
Zugkraft		19.46	42.95	33.75	4.6	208
EM_38		11.71	30.53	23.34	3.17	208
Ertrag_96		5.21	8.71	6.8	0.52	190
Ertrag_98		6.67	10.53	8.86	0.58	200
Ertrag_03		4.07	9.14	7.45	0.78	207

Abbildung 37: Data Audit des Datensatzes D4 03

Die geringe Anzahl von Instanzen verhindert ein Ausbalancieren, d. h. eine Reduktion der betriebsüblich gedüngten Instanzen. In Relation zum Versuchsdatensatz *Schafhof 02* sind die Standard-Abweichungen der applizierten Teilmengen dennoch beträchtlich höher.

Die spektralen Messungen konnten in diesem Versuch vollständig durchgeführt werden, es ergeben sich demnach keine fehlenden Werte für *REIP_2* und *REIP_3*. Der Mittelwert der ersten Messung liegt dabei deutlich unter der zweiten (722,54 nm und 727,23 nm). Die Standardabweichung erhöht sich zur zweiten Messung, was auf die Variante „Merkmals-

raum“ zurückzuführen ist. Insbesondere extreme Applikationen führen hier zu wachsenden Heterogenitäten im Bestand. Der Mittelwert von *REIP_3* liegt im Vergleich zum Vorjahresversuch auf ungefähr demselben Niveau.

Die Bodenattribute unterschieden sich vom Versuch *Schafhof 02* in einigen Punkten. Das Niveau der Zugkraftdaten liegt höher, das der Leitfähigkeitsdaten wesentlich niedriger, als im Schlag Schafhof, obwohl die beiden Schläge aus bodenkundlicher Sicht Ähnlichkeiten aufweisen (vgl. Kap. 4.2.4). Diese Unterschiede dürfen jedoch möglicher Verzerrungen der Messwerte nicht überbewertet werden, sondern liefern lediglich ein Indiz für die Interpretation (vgl. Kap. 3.1.3).

Die Ertragsdaten führen in den einzelnen Jahren zu unterschiedlichen Kennzahlen. Der Mittelерtrag des Jahres 1998 ist beispielsweise deutlich höher als im Jahr 1996. Die Werte des Versuchsjahres liegen in der Mitte dieser beiden Jahre. Bei diesen Ertragsdaten treten als Ergebnis der Datenvorverarbeitung fehlende Werte auf. Hier wurde den Teilflächen, für die keine validen Ertragsmesspunkte vorliegen (Vorgewende), keine Ausprägung zugewiesen. Die Standardabweichung der historischen Ertragsdaten ist, wie auch im Versuch *Schafhof 02*, gering. Die Standardabweichung der Ertragsdaten des Versuchsjahres ist deutlich höher als die der historischen Ertragsdaten, was wiederum auf die Variante *Merkmalsraum* und auf die teilflächenspezifischen Applikationen zurückzuführen ist.

In der Abbildung 38 sind die Korrelationen der Attribute untereinander abgetragen. Dabei fallen mehrere Eigenheiten des Datensatzes auf: Die erste N-Gabe besitzt den stärksten linearen Zusammenhang mit dem Ertrag des Versuchsjahres. Die beiden folgenden Applikationen haben tendenziell einen eher schwachen Einfluss auf den Ertrag. Ein auffallend starker Zusammenhang besteht zwischen den N-Applikationen und den darauf folgenden Vegetationsindices.

Die Korrelationen der Düngegaben mit den Bodenattributen sind in diesem Versuch stark eingeschränkt. Lediglich zur ersten N-Gabe, zu der ein Mapping-Ansatz angewendet wurde, existieren noch sehr schwache lineare Zusammenhänge mit den Ertragsdaten. Die beiden *REIP*-Kennzahlen korrelieren stark positiv mit den Erträgen des Versuchsjahres (0,69 und 0,70).

Der Zusammenhang zwischen den beiden historischen Ertragsdaten der Jahre 1996 und 1998 ist mittelstark ausgeprägt. Beide Attribute korrelieren mit den Ertragsdaten des Versuchsjahres ebenfalls mittelstark. Dabei ist verwunderlich dass, obwohl die Ertragskarte des Jahres 1998 augenscheinlich sehr schlecht war, die Korrelation zum Ertrag des Versuchsjahres deutlich höher ist als die des Jahres 1996.

Das Attribut *EM_38* verhält sich zu den anderen Attributen nicht konsistent: Zu den historischen Ertragsdaten korreliert es einmal schwach positiv, einmal schwach negativ, zu allen anderen Attributen sind praktisch keine Korrelationen festzustellen. Im Vergleich zum Ver-

such *Schafhof 02* ist der Zusammenhang mit den Ertragsdaten sehr viel schwächer. Damit erhärtet sich der Verdacht, dass mit diesem Attribut keine standortunabhängigen Aussagen ermöglicht werden (vgl. Kap. 3.1.3 und 4.7.1).

	N_1	N_2	N_3	REIP_2	REIP_3	ZUGKRAFT	EM_38	ERTRAG_96	ERTRAG_98	ERTRAG_03
N_1		0.28	0.24	0.53	0.7	-0.12	0.01	0.16	0.08	0.48
N_2	0.28		0.25	0.09	0.67	0.01	0.01	0.09	0.06	0.3
N_3	0.24	0.25		0.16	0.29	0.08	0.02	-0.02	0.06	0.17
REIP_2	0.53	0.09	0.16		0.62	-0.35	0.03	0.19	0.31	0.70
REIP_3	0.7	0.67	0.29	0.62		-0.16	0.04	0.19	0.18	0.69
ZUGKRAFT	0.12	0.01	0.08	-0.35	-0.16		0.21	-0.52	-0.34	-0.43
EM_38	0.01	0.01	0.02	0.03	0.04	0.21		-0.17	0.32	0.07
ERTRAG_96	0.16	0.09	0.02	0.19	0.19	-0.52	-0.17		0.24	0.28
ERTRAG_98	0.08	0.06	0.06	0.31	0.18	-0.34	0.32	0.24		0.36
ERTRAG_03	0.48	0.3	0.17	0.70	0.69	-0.43	0.07	0.28	0.36	

Abbildung 38: Korrelationsmatrix Versuch D4 03

Das Attribut *Zugkraft* ist mit den Ertragsattributen, wie aus dem Schafhof-Versuch zu erwarten war, durchgängig negativ korreliert. Die Stärke der linearen Zusammenhänge ist im Vergleich zu den anderen Bodenattributen sehr hoch.

Im Vergleich zu den von SPAULDING ET AL. (2004, s.p.) festgestellten Zusammenhänge erscheinen die des vorliegenden Datensatzes sehr stark. Der Grund kann zum einen in der höheren Präzision der Versuchsanstellung, aber auch an den naturräumlichen Gegebenheiten liegen.

4.6.2 Data Mining bei begrenztem Datenumfang

Für die Data-Mining-Phase stehen 208 unvollständige, bzw. 182 vollständige (ohne fehlende Werte) Instanzen zur Verfügung. Diese Menge ist kritisch zu sehen. Ein ernst zu nehmendes Problem stellt die Anpassung der Topologie bzw. die Terminierung des Trainings dar (vgl. Kap. 2.5.3). Im vorliegenden Fall stehen bei einer hälftigen Aufteilung in Trainings- und Testsample jeweils 91 Instanzen zur Verfügung. Dadurch, dass im Feldversuch *D4 03* mindestens ein Drittel der Instanzen dieselben N-Applikation erhalten hat, muss eine entsprechend schlechte Datenhomogenität bzw. Datenrepräsentativität in den Teilmengen angenommen werden (vgl. Kap. 2.4.2). Bei der zufälligen Aufteilung in Trainings und Testmenge können kleine Unterschiede einen entscheidenden Einfluss auf das Ergebnis

ausüben.

Dieses Problem verschärft sich, da es bei diesem geringem Datenumfang nicht mehr sinnvoll ist, eine Kreuzvalidierung durchzuführen und die Modellierung an deren Ergebnissen auszurichten. Im vorliegenden Fall werden bei einer fünffachen Kreuzvalidierung an jeweils einem Fünftel der Instanzen, also an lediglich ca. 45 Instanzen, die Maßzahlen zur Prognosegüte ermittelt. Dies führt zu einer enormen Variation der Maßzahlen in den aufeinander folgenden Phasen. Gleichzeitig bleiben dann nur noch jeweils vier Fünftel der Instanzen übrig, um die jeweiligen neuronalen Netze zu trainieren, wobei auch in diesem Fall nicht auf ein Test-Sample verzichtet werden kann, da durch die Heterogenität des Datensatzes der optimale Trainingszeitraum (Anzahl der Zyklen) wechselt.

Es kann nicht gewährleistet werden, dass die Abdeckung des Merkmalsraumes für die benötigten Trainings-, Test- und Validierungsmengen gewährleistet ist, wenn der Datensatz zufällig aufgespaltet wird. Eine manuelle Unterstützung dieser Aufspaltung ist zwar prinzipiell möglich, kann aber bei der verwendeten hohen Anzahl von Attributen nicht durchgeführt werden. Innerhalb der Teilphasen der Validierung entstehen dadurch sehr unterschiedliche Netze, die dann auch noch an sehr unterschiedlichen Validierungsstichproben getestet werden. Damit können keine verlässlichen Konfidenz-Kennziffern ermittelt werden.

Ersatzweise muss auf „rustikale“ Methoden zurückgegriffen werden. Die Auswahl der Modelle kann demnach nur durch wiederholtes Training, durch die Auswahl nach der *Estimated Accuracy* und durch die Analyse von trainierten Netzen und deren *Response Surfaces* und Entscheidungsregeln geschehen. Die Gefahr, nur lokale Minima in der Approximation zu erreichen ist entsprechend hoch.

Um eine Überanpassung zu vermeiden muss in diesem Versuch darauf geachtet werden, dass nur wenige Attribute verwendet werden bzw. dass die Attribute verdichtet werden, d. h. z. B. die beiden Ertragsdaten zu einem durchschnittlichen historischen Ertrag zusammenzufassen.

4.6.3 Ertragsprognose und Entscheidungsregeln zur ersten N-Applikation

Im Umgang mit den fehlenden Werten wurden drei verschiedene Datenaufbereitungsmethoden getestet. Als erste Alternative wurde versucht, die fehlenden Attribute mit dem Mittelwert der Trainingsmenge zu ersetzen. Dadurch wurde dieses Attribut jedoch weniger wertvoll für die Prognose und überwiegend durch den Pruning-Algorithmus (vgl. Kap. 2.5.3) abgeschnitten. Bei dem Versuch, die fehlenden Werte mit einem Entscheidungsbaumverfahren zu ersetzen, ergaben sich Netze mit sehr komplexen Topologien, die auf starkes Überlernen hinweisen (vgl. Kap. 2.6.3). Es gelang nicht, die fehlenden Werte mit dem entsprechenden Messrauschen zu erzeugen. Daher wurden alle Instanzen mit leeren Werten ausgeschlossen und mit 182 vollständigen Instanzen gearbeitet.

In der Anwendung des *Exhaustive-Prune-Modus* ergeben sich selbst unter identischen Voraussetzungen bei wiederholten Trainingsdurchläufen sehr verschiedene Topologien. Zum einen werden sehr große Netzwerke ausgegeben (oftmals jeweils 20 Neuronen in den zwei versteckten Schichten), was auf starkes Überlernen hindeutet. Im anderen Extremfall werden sehr triviale Netze gebildet, die nur ein Neuron in einer versteckten Schicht besitzen. Dadurch wird ein stark interpolatives Verhalten wahrscheinlich (vgl. Kap. 2.6.3).

In allen erzeugten neuronalen Netzen dominieren die beiden Attribute *Zugkraft* und *N_1* - wie nach der Korrelationsmatrix zu erwarten war. Diese beiden Attribute erreichen in der Sensitivitätsanalyse die höchsten Werte. Im Allgemeinen hat das Attribut *Ertrag_98* einen höheren Einfluss auf die Ertragsprognose als das Attribut *Ertrag_96*, das in der Regel vom *Exhaustive-Prune-Algorithmus* abgeschnitten wird. Dasselbe gilt für das Attribut *EM_38*. Es zeigen sich allerdings keine Vorteile, wenn diese Attribute ex ante nicht als Input-Attribute verwendet werden. Die *Estimated Accuracy* schwankt im Bereich zwischen 91 und 93. Die Zusammenfassung der beiden Attribute *Ertrag_96* und *Ertrag_98* zu einem durchschnittlichen Ertrag (*AVG_Ertrag*) führt zu einer sehr geringen Bedeutung dieses Attributes für die Prognose. Die Ergebnisse sind im Anhang (Tabelle A-31, im Anhang) zusammengefasst.

Neuronales Netzwerk

Für die weitere Analyse wird deshalb ein repräsentatives Netz mit der höchsten *Estimated Accuracy* verwendet. Dieses Netzwerk zeigt eine „durchschnittliche“ Topologie und verwendet nur die drei Attribute *Zugkraft*, *N_1* und *Ertrag_98* für die Prognose. Die anderen Input-Attribute wurden im Netzwerktraining abgeschnitten. Die Ergebnisse des Trainings sind in Tabelle 16 dargestellt. Die verwendeten Attribute besitzen in diesem Netzwerk sehr ähnliche Bedeutungen für die Prognose.

Tabelle 16: Neuronales Netzwerk zur ersten N-Applikation (Versuch D4 03)

Estimated accuracy: 93.21	Relative Importance of Inputs	Einstellungen:
Input Layer: 3 neurons	<i>Zugkraft</i> : 0.22	Method: Exhaustive prune
Hidden Layer 1: 4 neurons	<i>N_1</i> : 0.21	Stop on: Default
Hidden Layer 2: 2 neurons	<i>Ertrag_98</i> : 0.19	Prevent overtraining: true
Output Layer: 1 neuron		Sample: 50 %

Eine hohe Bedeutung bekommt unter den dargestellten Voraussetzungen die Analyse mittels *Response Surfaces*. Zu diesem Zweck wurde ein Testvektor mit den Attributen *N_1*, *Zugkraft* und *Ertrag_98* konstruiert und als Input in das trainierte neuronale Netzwerk eingegeben.

Response Surfaces

Bezüglich des *Ertrags_98* ergibt sich unter Verwendung des dargestellten Netzwerkes die in Abbildung 39 abgebildete *Response*. Die daraus ersichtlichen Produktionsfunktionen weisen eine Form auf, wie sie bei der Schätzung von Produktionsfunktionen aus N-Steigerungsversuchen an einem vergleichbaren Standort beschrieben werden (vgl. GANDORFER, 2005, S. 50ff). Über alle Ertragszonen hinweg führt eine Erhöhung von Stickstoff erst zu einem steilen Anstieg, der sich allmählich abflacht, d. h. es zeigt sich eine Funktion mit abnehmendem Grenzertrag.

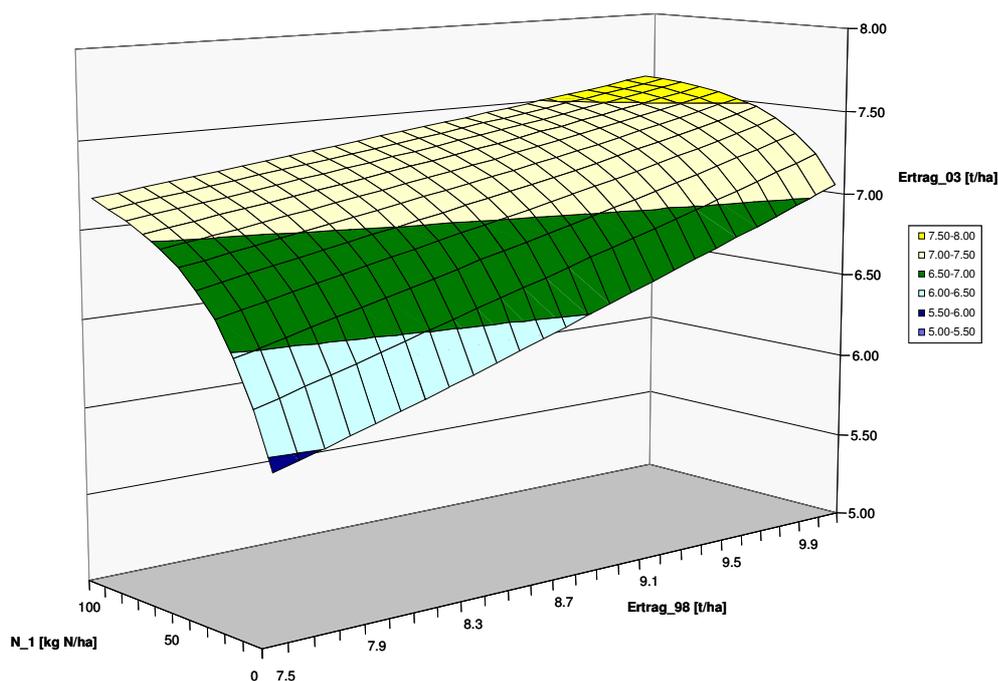


Abbildung 39: Response Surface zur ersten N-Applikation (*Ertrag_98*, Versuch D4 03)

Hierbei ergeben sich zwei Trends. Ein höherer Ertrag des Jahres 1998 führt in diesem Diagramm auch zu einem höheren prognostiziertem Ertrag (*Ertrag_03*). Allerdings ist der Ertragsanstieg bei hohen N_1 -Mengen nicht so hoch, wie bei niedrigen N-Mengen. Bezüglich der ertragsmaximierenden Stickstoffmenge gilt, dass die maximale Düngermenge stets auch zum maximalen Ertrag führt. Dennoch scheint der Ertragsanstieg bei der maximalen Applikation von 100 kg N/ha noch nicht zu Ende zu sein. Auffällig ist, dass der Ertragsanstieg in Folge einer erhöhten N-Applikation in niedrigen Ertragsbereichen höher als in hohen Ertragsbereichen ist, was in diesem Fall ökonomische Effekte verursacht. Dies zeigt die Analyse der stickstoffkostenfreien Leistung.

Dazu werden die Faktoren N_1 und der prognostizierte Ertrag mit marktüblichen Preisen bewertet (vgl. Kap. 4.4.3). Die dadurch entstehende Oberfläche ist in Abbildung 40 dargestellt. Auf der z-Achse ist die (prognostizierte) stickstoffkostenfreie Leistung (in €/ha) abgetragen. Hierbei wurde die Perspektive der 3-D-Abbildung so gewählt, dass die Oberfläche frontal aus Sicht der X-Achse betrachtet wird. Damit ist der Teil der gekrümmten Ober-

fläche, der hinter dem Maximum liegt, nicht mehr zu sehen. Aus diesem Diagramm kann abgelesen werden, dass das ökonomische Optimum bei niedrigen *Ertrag_98*-Werten höher liegt, als bei hohen. Die ökonomisch optimale *N_1*-Menge liegt beispielsweise in Teilflächen mit einem *Ertrag_98* von 7.5 t/ha bei ca. 70 kg N/ha. Bei einem *Ertrag_98* von 10 t/ha sind dies ca. 30 kg N/ha.

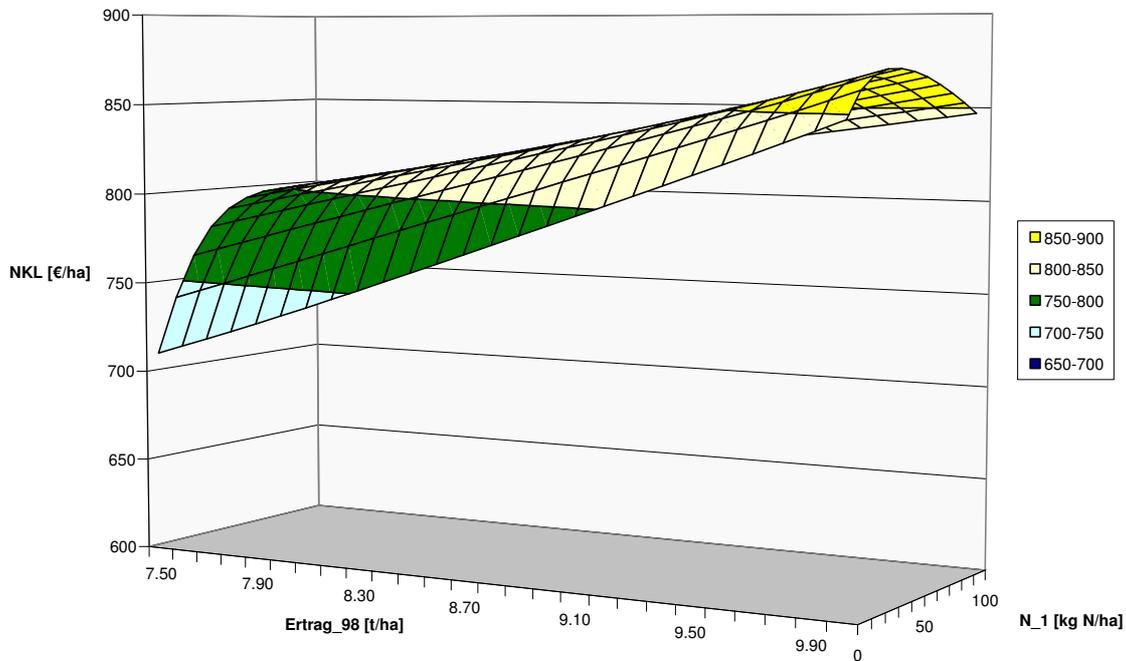


Abbildung 40: *NKL-Response* zur ersten N-Applikation (*Ertrag_98*, Versuch D4 03)

Hieraus kann ein Muster bzw. ein Ansatz zur ökonomischen Optimierung der ersten N-Gabe abgeleitet werden. Entgegen dem „klassischen“ Mapping-Ansatz (vgl. Kap. 3.1.2) liegt nach diesen Ergebnissen N_{opt} in Niedrigertragsbereichen höher als in Hohertragsbereichen. Die resultierende Düngestrategie kann als „inverser Mapping-Ansatz“ umschrieben werden. Der ökonomische Effekt, der mit einer Düngung an N_{opt} erreicht werden kann, ist bei niedrigen Ertragsbereichen wesentlich höher als bei hohen Ertragsbereichen. So führt beispielsweise nach dieser Schätzung eine Null-Düngung auf einer Teilfläche mit einem *Ertrag_98* von 7.5 t/ha zu einer stickstoffkostenfreien Leistung von ca. 700 €/ha. Mit einer Erhöhung der Düngung auf 60 kg/ha kann dieser Betrag auf ca. 800 €/ha gesteigert werden. In höheren Ertragsbereichen fällt dieser Effekt wesentlich kleiner aus, obwohl das durchschnittliche Niveau höher liegt.

In ähnlicher Weise werden *Responses Surfaces* für das Attribut *Zugkraft* untersucht. Die Basis für eine teilflächenspezifische Optimierung ist hier jedoch stark eingeschränkt, da sich die optimalen N-Gaben nur minimal unterscheiden. Der Verlauf der Response-Diagramme unterscheidet sich jedoch bei verschiedenen *Ertrag_98* Werten. Der Verlauf der *NKL-Response* ist im Anhang (Abbildung A-71) dargestellt.

Entscheidungsregeln

Zur weiteren Analyse werden Entscheidungsbaumverfahren verwendet. Mit dem vorgestellten neuronalen Netzwerk wird der Datensatz *D4 03* ex post ökonomisch optimiert (vgl. Kap. 4.4.2) und diese Daten mit einem Entscheidungsbaumalgorithmus modelliert (vgl. Kap. 4.4.3). Mit diesem Data-Mining-Werkzeug sollen die wichtigsten Regeln für eine teilflächenspezifische Optimierung basierend auf den Attributen *Zugkraft* und *Ertrag_98* ermittelt werden. In dieser Modellierung stellen diese beiden Attribute den Input dar. Das Output-Attribut (Zielgröße) ist die ex post optimierte N-Menge (N_{opt}). Aus diesem Datensatz lässt sich der Entscheidungsbaum aus Abbildung 41 erstellen⁶⁶.

Es wird deutlich, dass das entscheidende Attribut für die Bestimmung von N_{opt} der *Ertrag_98* ist und in allen Abzweigungen ein höherer Wert dieses Attributes zu einem niedrigeren N_{opt} führt. Damit zeigt sich wiederum der „inverse Mapping-Ansatz“, der bereits aus den *Response Surfaces* abgeleitet wurde. Das Attribut *Zugkraft* verhält sich nicht konsistent: in niedrigen *Ertrag_98*-Bereichen führt eine niedrigere Zugkraft zu einem höheren N_{opt} . In vergleichsweise hohen *Ertrag_98*-Bereichen führt eine niedrigere Zugkraft zu niedrigeren N_{opt} und vice versa. Insgesamt werden N_{opt} -Gaben in einer „praxisüblichen“ Spanne von 30-80 kg N/ha abgeleitet.

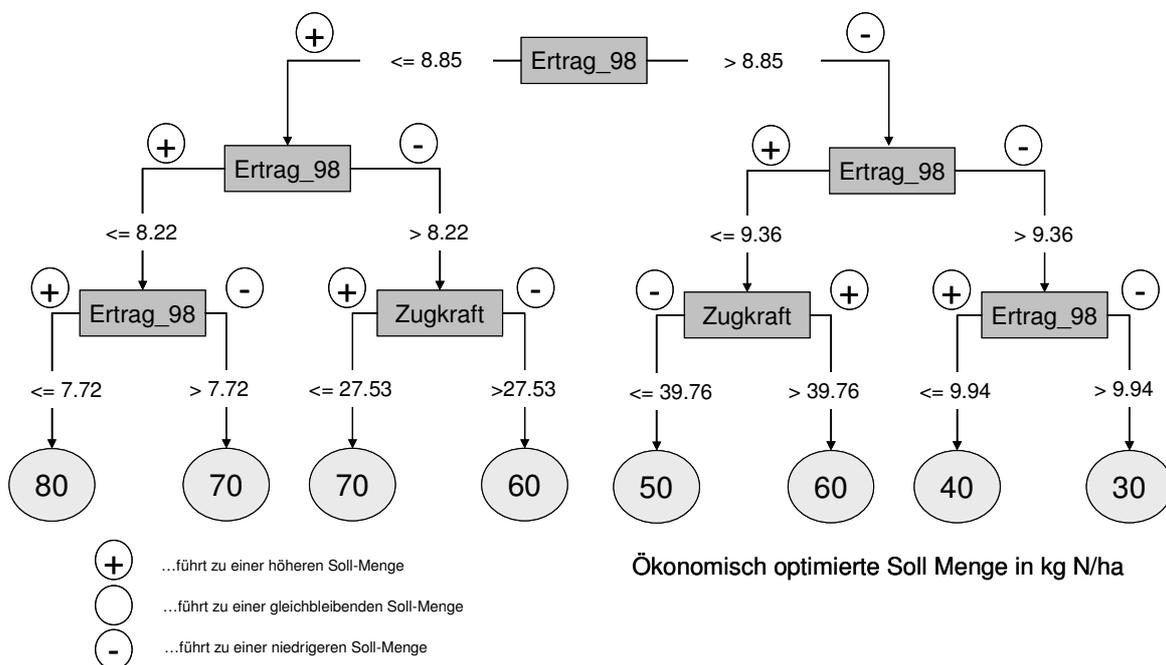


Abbildung 41: Entscheidungsbaum zur ersten N-Gabe (Versuch D4 03)

In weiteren Testreihen mit gleichen bzw. veränderten Trainingsparametern und Datenaufbereitungen tauchen die beschriebenen (Grob-)Muster regelmäßig auf. Allerdings unter-

⁶⁶ Wie beschrieben, wird die Darstellung auf drei Ebenen beschränkt. Der vollständig ausgebildete Entscheidungsbaum besteht aus sieben Ebenen.

scheiden sich die Ergebnisse, wenn einzelne *Response Surfaces* bzw. Entscheidungsregeln miteinander verglichen werden, beträchtlich. Die Oberfläche der *Response Surfaces* ändert sich in ihrer Struktur unter Umständen sehr auffällig. Insbesondere in Bereichen mit geringer Datendichte produzieren neuronale Netze teils sehr unregelmäßig geformte Oberflächen. Gleiches gilt für die *Splits* der Entscheidungsbäume (Kanten). Hier können keine konstanten Werte beobachtet werden.

Als Ergebnis kann demnach festgehalten werden, dass die geringe Anzahl von Instanzen in Verbindung mit einer mäßigen Abdeckung des Merkmalsraumes die Data-Mining-Phase stark erschwert. Eine Änderung in der Datenaufbereitung resultiert immer in mehr oder weniger gravierenden Abweichungen der approximierten Ertragsfunktion. Gleichzeitig steigt das Risiko, ein nicht repräsentatives Netz auszuwählen, sofern dies unter diesen Voraussetzungen überhaupt beurteilt werden kann. Die Extraktion von stabilen Mustern gelingt nur dann, wenn ausschließlich grobe Zusammenhänge betrachtet werden.

4.6.4 Ertragsprognose und Entscheidungsregeln zur zweiten N-Applikation

Zur zweiten Stickstoff-Applikation wurden wiederum verschiedene Datenaufbereitungen und Trainingsdurchläufe getestet. Alle trainierten Netze zeigen dieselbe Rangfolge der Attribute bezogen auf die *Relative Importance*. Als stärkstes Prediktor-Attribut erweist sich *REIP_2*, gefolgt von der *Zugkraft* und *N_2*. Die Ertragsattribute werden zumeist abgeschnitten. Eine Übersicht der trainierten Netze ist im Anhang (Tabelle A-32) angeführt.

Neuronales Netzwerk

Aus diesen Netzen wurde als repräsentatives Beispiel ein neuronales Netz ausgewählt, dessen Eigenschaften in Tabelle 17 zusammengefasst sind. Dieses Netz zeichnet sich durch die höchste *Estimated Accuracy* und eine kompakte Netztopologie aus. Die Ergebnisse der Sensitivitätsanalyse liegen im Einklang mit den alternativen neuronalen Netzen. Durch den *Exhaustive-Prune-Algorithmus* wurden lediglich die drei Attribute *REIP_2*, *Zugkraft* und *N_2* verwendet; alle anderen wurden abgeschnitten. Im Vergleich zur Ertragsprognose zu *N_1* hat sich die *Estimated Accuracy* nicht verbessert, obwohl weitere Attribute hinzugezogen wurden. Dies bestätigt die Schwierigkeit, globale Minima zu erreichen.

Tabelle 17: Neuronales Netzwerk zur zweiten N-Applikation (Versuch D4 03)

Estimated accuracy: 93.21	Relative Importance of Inputs	Einstellungen:
Input Layer: 3 neurons	<i>REIP_2</i> : 0.45	Method: Exhaustive prune
Hidden Layer 1: 2 neurons	<i>Zugkraft</i> : 0.15	Stop on: Default
Hidden Layer 2: 1 neuron	<i>N_2</i> : 0.13	Prevent overtraining: true
Output Layer: 1 neuron		Sample: 50 %

Das ausgewählte neuronale Netzwerk wird im folgenden Schritt dazu verwendet, um Erträge für einen Testvektor zu prognostizieren. Zunächst wird die Wirkung der Teilapplikation bei unterschiedlichen *REIP_2*-Werten untersucht.

Das entsprechende Response Surface ist in Abbildung 42 dargestellt. Deutlich erkennbar ist der stark positive Zusammenhang zwischen den *REIP_2*-Werten und dem prognostizierten Ertrag. Für die höchsten Werte des Vegetationsindex (725 nm) werden mehr als acht Tonnen Kornenertrag pro Hektar prognostiziert, während diese Menge bei einem sehr niedrigen Versorgungszustand bei lediglich ca. sechs Tonnen pro Hektar liegt. Der Einfluss der zweiten Teilapplikation ist ebenfalls durchgängig positiv, allerdings fällt der Anstieg des prognostizierten Ertrages für niedrig versorgte Teilflächen sehr viel höher aus, als bei gut versorgten Beständen. Bei *REIP*-Werten von über 724 nm zeigt sich eine sehr geringe Ertragswirkung der N-Düngung. Der Anstieg erfolgt zunächst vergleichsweise steil und flacht dann allmählich ab.

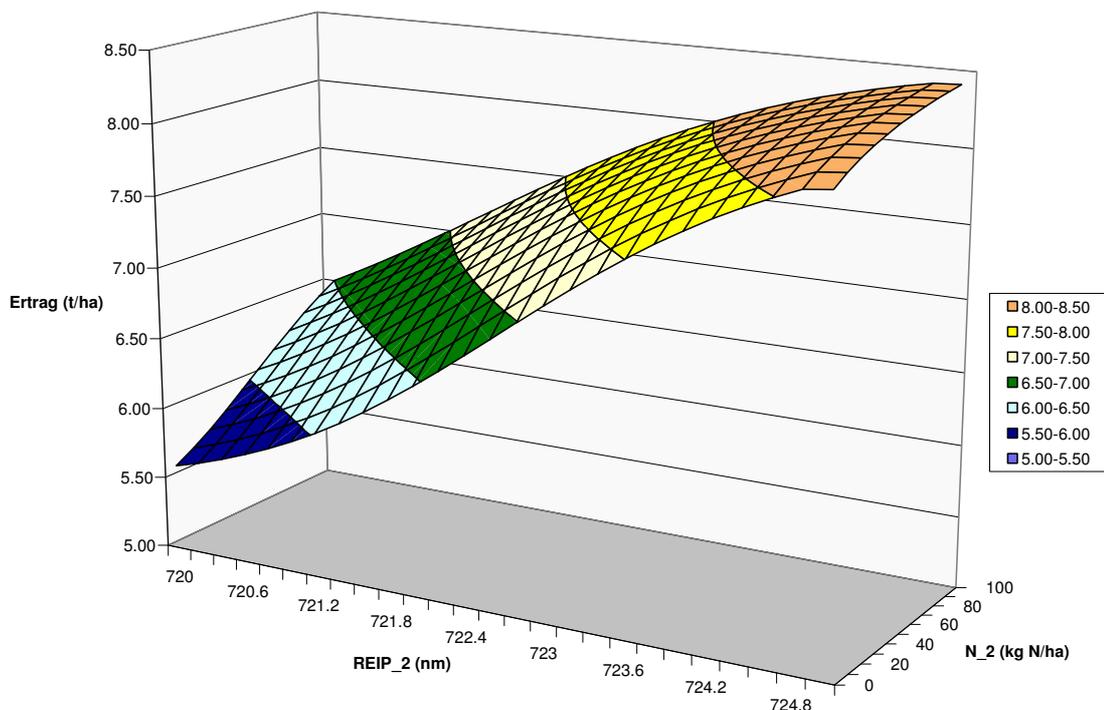


Abbildung 42: Response Surface zur zweiten N-Applikation (*REIP_2*, Versuch D4 03)

Die Berechnung der entsprechenden prognostizierten stickstoffkostenfreien Leistung zeigt, dass das ökonomische Optimum bei sehr niedrigen *REIP*-Werten bei 100 kg N/ha liegt und dann schrittweise abnimmt, bis es zu einer Höhe von 0 kg N/ha bei sehr hohen *REIP*-Werten gelangt. Das entsprechende Diagramm ist im Anhang (Abbildung A-72) dieser Arbeit zu finden.

Als nächsten Schritt wurde die Response dieses neuronalen Netzwerkes im Hinblick auf das Attribut *Zugkraft* untersucht. Das *Response Surface* bezüglich der *Zugkraft* und der zweiten Stickstoff-Applikation ist in Abbildung 43 dargestellt. Es ergibt sich in diesem Fall

ein ausgeprägter negativer Zusammenhang von Zugkraft und dem prognostizierten Ertrag. Bezüglich der zweiten N-Applikation zeigt sich ein durchgängig gleichgerichteter Zusammenhang mit dem Ertrag des Versuchsjahres. Im Vergleich zu dem Attribut *REIP_2* ist die Sensitivität hinsichtlich der Zugkraft schwächer⁶⁷.

Aus ökonomischer Sicht ist der Anstieg des Ertrages bei zunehmender N-Gabe interessant. Zunächst zeigt sich in allen Bereichen ein abnehmender Grenzertrag. Die Abflachung der Steigung erfolgt in hohen Zugkraft-Bereichen später als in niedrigen Zugkraft-Bereichen. Die durchschnittliche Steigung ist ebenfalls in hohen Zugkraft-Bereichen höher. Daraus ergeben sich ökonomische Konsequenzen, wie aus der Berechnung der stickstoffkostenfreien Leistung hervorgeht. Die optimale N-Gabe sinkt mit zunehmender Zugkraft von ca. 80 auf 20 kg N/ha.

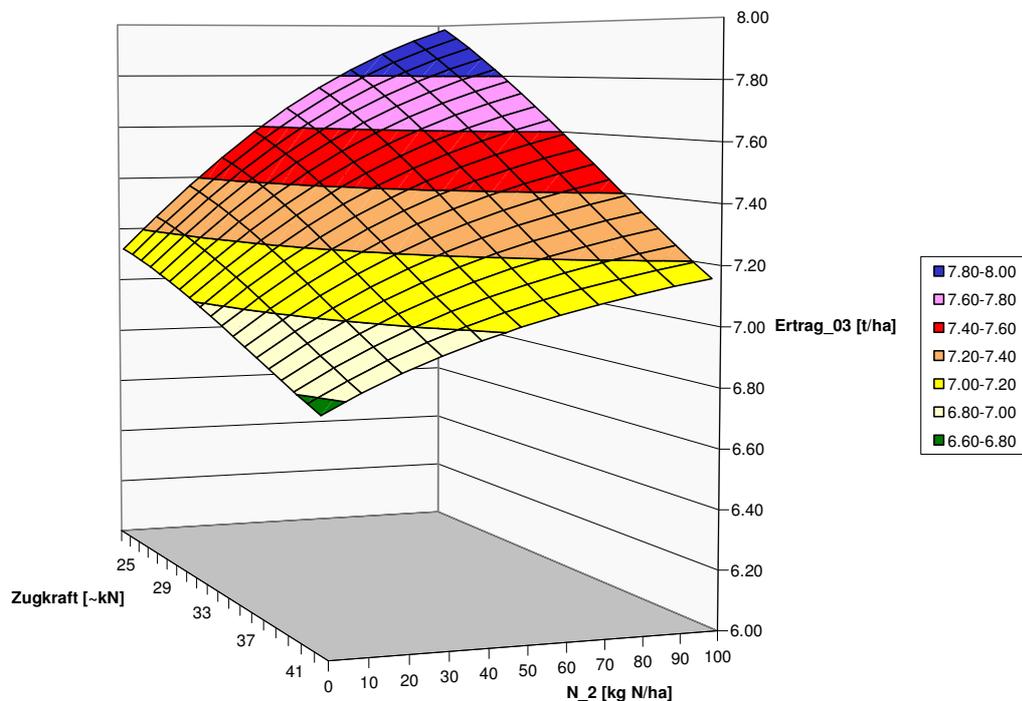


Abbildung 43: Response Surface zur zweiten N-Applikation (Zugkraft, Versuch D4 03)

Aus der Analyse der *Response-Surfaces* ergeben sich damit eindeutige Hinweise für die ökonomische Optimierung. Wiederum wird versucht, mit den ex post optimierten Teilflächen des Versuchsdatensatzes Entscheidungsregeln zu generieren um diese Indizien näher zu untersuchen. Der daraus entstehende Entscheidungsbaum ist in Abbildung 44 dargestellt.

Wie zu erwarten war, nimmt das Attribut *REIP_2* eine dominierende Stellung in diesem Entscheidungsbaum ein. Erst in der dritten Ebene tritt das Attribut Zugkraft in Erscheinung. Die damit verbundenen Regeln führen allerdings zu einer großen Änderung von N_{opt} (je-

⁶⁷ Dies kann an den maximalen Änderungen des prognostizierten Ertrages bei einer maximalen Variation des Input-Attributes in den zwei Diagrammen abgelesen werden.

weils 30 kg N/ha). Die Variation der optimierten N-Mengen erscheint sehr groß. Dabei ist aber zu beachten, dass die Teilflächen, für die die Anwendung des Entscheidungsbaumes zu extremen Applikationen (z. B. 100 kg N/ha) führt, auch sehr selten auftreten. Dies kann annähernd über die Histogramme des *Data Audit* nachvollzogen werden⁶⁸. Das bedeutet, es würden wenige Instanzen des Schlages *D4 03* ex post mit 0 bzw. 100 kg N/ha gedüngt werden. Die Wirkungszusammenhänge verlaufen so, wie es aus den *Response Surfaces* zu erwarten ist. Ein niedriger Versorgungszustand führt durchgängig zu einem höheren N_{opt} . Für niedrige *REIP*-Werte wird zur Entscheidung das Attribut *Zugkraft* hinzugezogen. In diesen Bereichen führt eine niedrigere *Zugkraft* jeweils zu einem höheren N_{opt} .

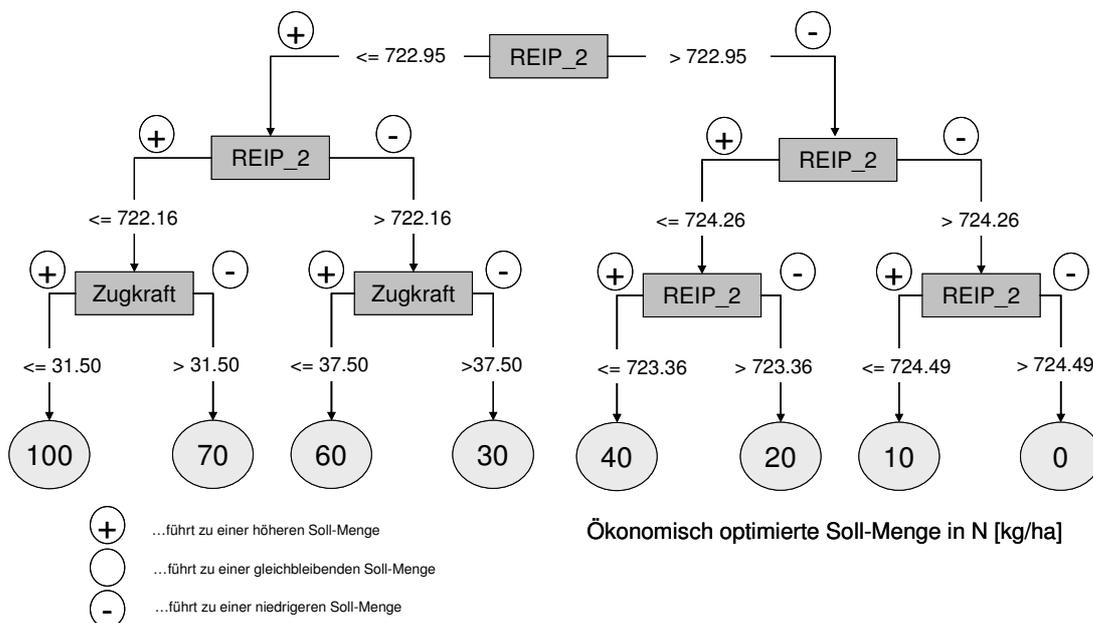


Abbildung 44: Entscheidungsbaum zur zweiten N-Applikation (Versuch D4 03)

Damit können zur zweiten N-Applikation wiederum (Grob-)Muster identifiziert werden. Zu diesem Zeitpunkt entspricht die Optimierung einem „klassischen“ Sensor-Ansatz. Ein schlechterer Versorgungszustand führt nach diesem Modell zu einem höheren N_{opt} . Das Attribut *Zugkraft* scheint ebenfalls einen Einfluss auf die Bestimmung von N_{opt} zu haben. Nach diesem Modell sollte in Bereichen mit hohen *Zugkraft*-Werten aus ökonomischer Sicht die zweite Teilapplikation verringert werden – allerdings nur in schlecht versorgten Beständen.

4.6.5 Ertragsprognose und Entscheidungsregeln zur dritten N-Applikation

Für das Training der neuronalen Netze zur dritten N-Gabe stehen zwei weitere Input-Attribute, *REIP_3* und *N_3*, zur Verfügung. Während der Data-Mining-Phase werden mehrere Varianten der Input-Attribute sowie Veränderungen in den Trainingsparametern (v.a.

⁶⁸ Wie dem *Data Audit* zu entnehmen ist, treten beispielsweise *REIP_2*-Werte von unter 722.16 nm vergleichsweise selten auf.

Größe des Test-Samples) getestet. Ein Ausschnitt an trainierten Netzen ist im Anhang (Tabelle A-33) zu finden. Wie aus der Datenexploration zu erwarten war, fällt bei allen getesteten Netzen der Einfluss der dritten Teilgabe auf den Ertrag relativ gering aus (*Relative Importance* < 0.1). In einigen Fällen wird dieses Attribut vom Trainingsalgorithmus abgeschnitten. Daraus kann der Hinweis abgeleitet werden, dass die Teilgabe keine ausgeprägte Ertragswirkung hat. Im Allgemeinen wird diese Teilgabe auch als „Qualitätsgabe“ bezeichnet, d. h. die Stickstoffgabe erhöht zwar die Konzentration von Stickstoff im Korn, aber nicht unbedingt die Masse (vgl. Kap. 3.1.1). Vor diesem Hintergrund lassen sich die Ergebnisse interpretieren.

Als stärkste Prediktoren erweisen sich *REIP_3*, *Zugkraft* und in einigen Fällen *Ertrag_98*. Wiederum zeigen sich unterschiedliche Ergebnisse bei Wiederholung des Trainingsvorgangs – selbst unter identischen Voraussetzungen. Die *Estimated Accuracy* schwankt entsprechend stark.

Um die Zahl der Input-Attribute gering zu halten, werden die vorausgegangenen N-Applikationen, *REIP_2*, *EM_38* sowie *Ertrag_96* nicht zur Prognose verwendet. Dadurch kann die Gefahr von Überlernen etwas verringert werden. Als Beispiel zur Analyse wird ein Netzwerk verwendet, das in Tabelle 18 dargestellt ist. Dieses Netz hat eine relativ kompakte Struktur, die vergleichbar ist mit der zur zweiten N-Applikation. Die Präzision steigt durch den Einbezug der beiden neuen Attribute etwas an. Die Bedeutung des Attributes *REIP_3* für die Prognose ist sehr hoch, wohingegen die dritte N-Gabe eine *Relative Importance* von lediglich 0.09 erreicht. Wiederum zeigt sich ein hoher Einfluss der *Zugkraft* auf die Ertragsprognose.

Tabelle 18: Neuronales Netzwerk zur dritten N-Applikation (Versuch D4 03)

Estimated accuracy: 93.31	Relative Importance of Inputs	Einstellungen:
Input Layer: 4 neurons	<i>REIP_3</i> : 0.45	Method: Exhaustive prune
Hidden Layer 1: 2 neurons	<i>Zugkraft</i> : 0.21	Stop on: Default
Hidden Layer 2 : 2 neurons	<i>Ertrag_98</i> : 0.15	Prevent overtraining: true
Output Layer: 1 neuron	<i>N_3</i> : 0.09	Sample: 40 %

Aus diesem Netz können wiederum die abgebildeten Zusammenhänge analysiert werden. Zunächst ist dabei der Zusammenhang zwischen den Attributen *REIP_3* und *N_3* auf den prognostizierten Ertrag interessant (vgl. Abbildung 45). Es zeigt sich ein ähnlicher Verlauf wie zur zweiten N-Applikation (entsprechend *REIP_2*). Die Ertragswirkung aus einer Variation dieser beiden Attribute heraus ist allerdings schwächer ausgeprägt. In Bereichen mit sehr niedrigen *REIP_3*-Werten führt eine Erhöhung der Düngung zu einer Erhöhung des prognostizierten Ertrages. In sehr gut versorgten Bereichen dagegen tritt der umgekehrte

Fall ein: Eine Erhöhung der Düngung führt zu einem leichten Absinken des Ertrages. Aus ökonomischer Sicht können daraus Ansätze für eine teilflächenspezifische Düngung abgeleitet werden. Das ökonomische Optimum für sehr niedrige $REIP_3$ -Werte (725 nm) liegt bei 50 kg N/ha. Bei einer Erhöhung des Versorgungszustandes verringert sich das ökonomische Optimum schrittweise bis zur Nulldüngergabe. Ab einem Wert ca. 727 nm liegt N_{opt} durchgängig bei diesem Wert.

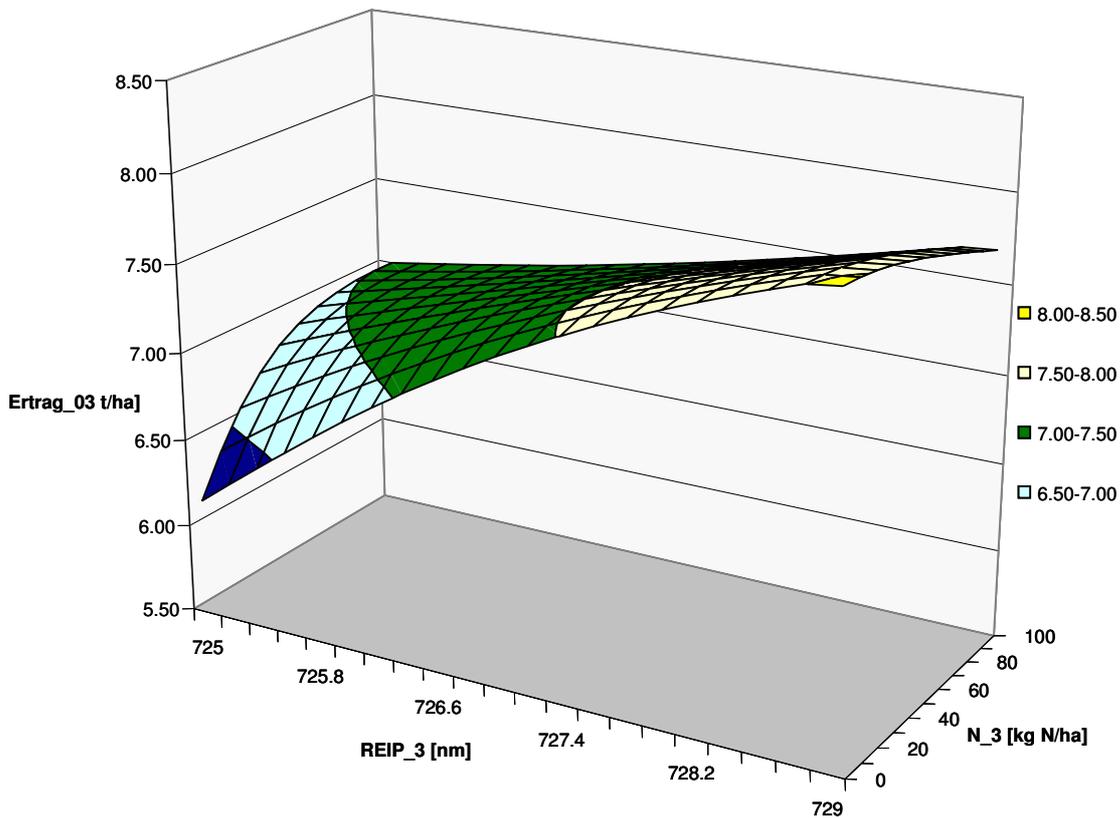


Abbildung 45: Response Surface zur dritten N-Applikation ($REIP_3$, Versuch D4 03)

Aus der Analyse des *Response Surfaces* der Zugkraft im Zusammenhang mit der N_3 -Düngung lassen sich hingegen keine Ansätze zu einer teilflächenspezifischen Düngung ableiten. Die Produktionsfunktionen verlaufen über alle Zugkraft-Bereiche sehr ähnlich.

Dagegen lassen sich für die historischen Ertragsdaten durchaus Ableitungen finden. Mit Hilfe eines Response-Surface (vgl. Abbildung 46) wird dieser Zusammenhang illustriert. Dabei zeigt sich zum einen, dass relativ niedrige N-Gaben zur Erreichung des Ertragsmaximums ausreichen. Diese Intensität reicht von ca. 20 kg N/ha bei sehr niedrigen $Ertrag_98$ -Werten bis hin zu 50 kg N/ha bei sehr hohen $Ertrag_98$ -Werten. Bei $Ertrag_98$ -Werten ab 10.8 t/ha lässt sich wiederum ein plateauförmiger Verlauf erkennen. Dieser Bereich ist im Merkmalsraum der Trainingsdaten nur sehr schwach abgedeckt. Die entsprechende ökonomische Ableitung wäre also - wie zur zweiten N-Applikation - in einem „klassischen“ Mapping-Ansatz zu sehen. Zu beachten ist allerdings, dass das Attribut einen relativ geringen Einfluss auf die Prognose hat. Es ist beispielsweise zu erwarten, dass

die Zusammenhänge von *REIP_3* mit der dritten N-Applikation diese Ergebnisse dominieren.

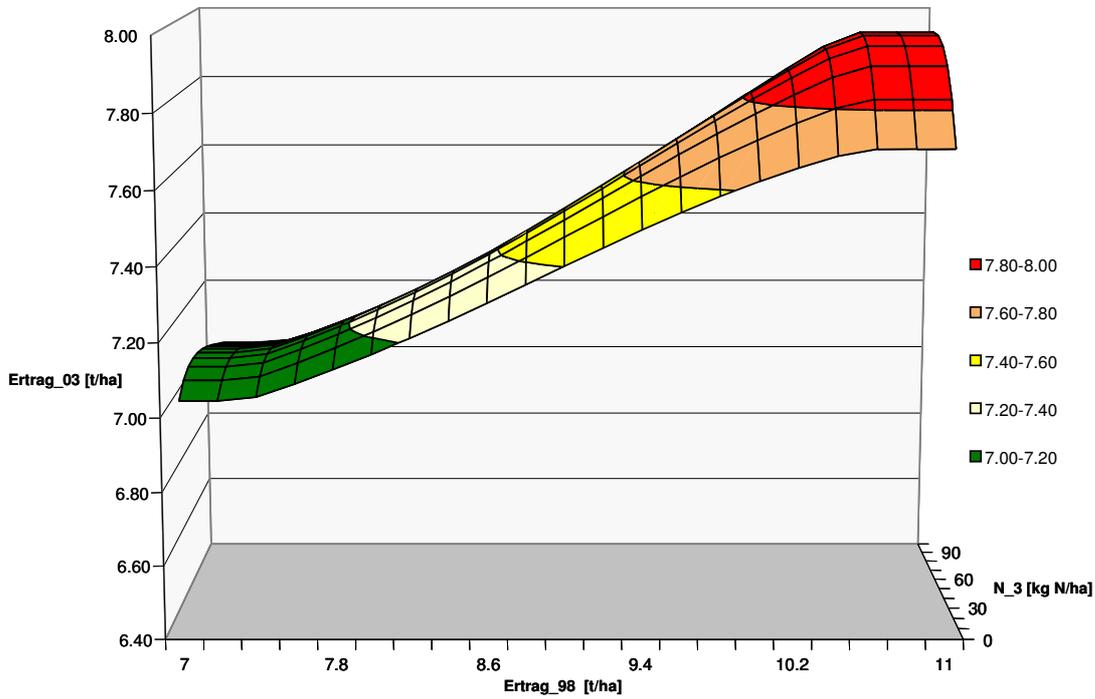


Abbildung 46: Response Surface zur dritten N-Applikation (Ertrag_98, Versuch D4 03)

Dies läßt sich mit Hilfe von Entscheidungsbaumverfahren nachvollziehen. Werden die Düngermengen der Datensätze des *D4 03* ex post mit Hilfe dieses neuronalen Netzes optimiert und daraus Entscheidungsregeln abgeleitet, ergibt sich ein Entscheidungsbaum, der in Abbildung 47 dargestellt ist.

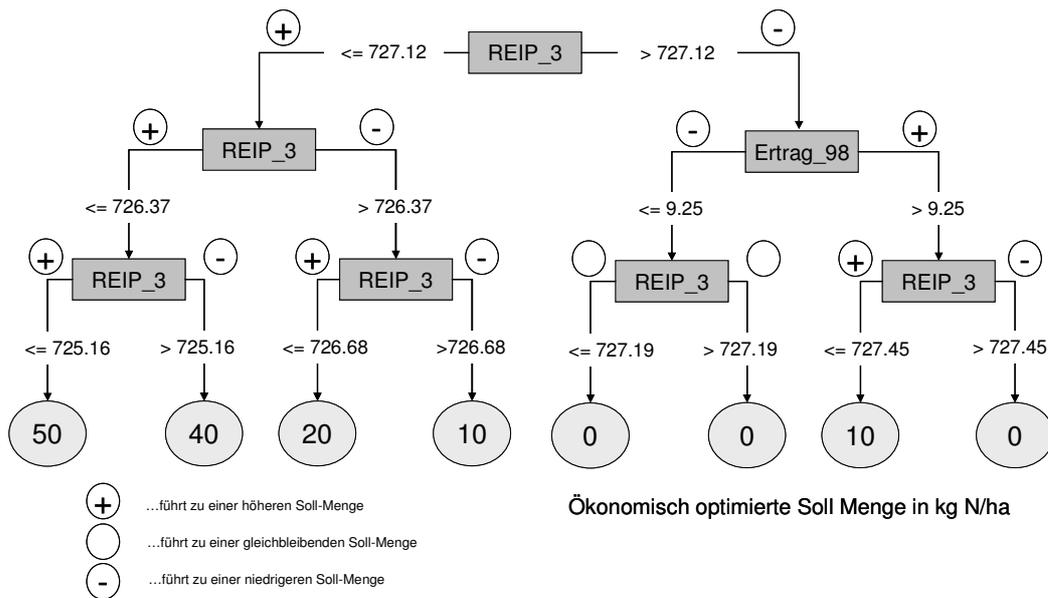


Abbildung 47: Entscheidungsbaum zur dritten N-Applikation (Versuch D4 03)

Wie zu erwarten dominiert das Attribut *REIP_3* diese Regeln. Es ergeben sich optimale spezielle Intensitäten von 0-50 kg N/ha. Es zeigt sich eindeutig, dass die optimale spezielle Intensität mit abnehmenden *REIP_3*-Werten zunimmt. In einem Knoten wird das Attribut *Ertrag_98* angewendet. Ein höherer historischer Ertrag führt dabei auch zu einer höheren empfohlenen N₃-Applikation und vice versa. Dies ist insofern überraschend, als dass bei der ersten N-Applikation der Einfluss dieses Attributes auf N_{opt} in die entgegen gesetzte Richtung zeigt.

4.6.6 Zwischenfazit

Die Auswertung des Versuches *D4 03* brachte interessante Ergebnisse. Es zeigen sich konkrete Ansätze zur teilflächenspezifischen Optimierung der Stickstoff-Applikationen. Dabei spielen die Attribute in den jeweiligen Teilgaben sehr unterschiedliche Rollen.

Aus der Datenexploration wurde deutlich, dass die Teilapplikationen einen unterschiedlichen Einfluss auf den Ertrag haben. Während die erste N-Gabe einen relativ starken Zusammenhang mit dem Ertrag hat, sinkt die Stärke dieses Zusammenhangs für die folgenden Teilapplikationen.

Zur Bemessung der ersten N-Gabe dominieren historische Ertragsdaten. Es kristallisiert sich ein Ansatz heraus, der als „inverser Mapping-Ansatz“ bezeichnet werden kann. Die Andüngung sollte demnach bei Niedrigertragsbereichen höher erfolgen als in Hohertragsbereichen. Das Attribut *Zugkraft* verhält sich in den verschiedenen Ertragsbereichen nicht konsistent.

Zur zweiten N-Gabe wandelt sich das Bild. Historische Ertragsdaten spielen für die Düngung zu diesem Zeitpunkt keine Rolle. Die optimale Düngung wird überwiegend über den Vegetationsindex *REIP_2* bemessen. Eine untergeordnete Rolle kommt dem Attribut *Zugkraft* zu. Die optimale Düngung wird in einigen Fällen erhöht, falls die Zugkraftwerte unter bestimmten Grenzen liegen und vice versa.

Zur dritten N-Gabe spielt wiederum der aktuelle Vegetationsindex (*REIP_3*) die dominierende Rolle. Auffällig ist, dass sich die Wirkungsweise des Attributes Ertrag umkehrt: Die dritte Teilgabe sollte mit steigenden historischen Erträgen ebenfalls erhöht werden.

Die Ergebnisse dieses Versuches stehen somit im Widerspruch zur teilflächenspezifischen Umsetzung der Düngung nach Entzug (vgl. Kap. 3.1.2). Eine Übereinstimmungen mit dem zugrunde liegende Konzept kann lediglich für die dritte Teilapplikation beobachtet werden, wobei diese Applikation eine sehr geringe Bedeutung für den Ertrag aufweist. Nach den Ergebnissen dieser Data-Mining-Phase kann mit Hilfe von Ertragskarten ein sehr hohes Potenzial zur teilflächenspezifischen Optimierung realisiert werden – allerdings größtenteils zur ersten N-Applikation und in der gegenteiligen Richtung.

Auffallend ist, dass das Attribut Zugkraft generell eine sehr große Bedeutung für die Ertragsprognose besitzt, aber sich diese Bedeutung nicht entsprechend in der Bemessung der Düngergaben widerspiegelt.

Die durchschnittliche Verteilung der Düngergaben kann aus den ex post optimierten Versuchsdatensätzen bestimmt werden. Es ergibt für dieses Jahr eine durchschnittliche optimale N-Menge von ca. 50 kg N/ha für die erste Gabe, von 54 kg N/ha für die zweite N-Gabe und von lediglich 10 kg N/ha für die dritte Gabe. Damit würde das durchschnittliche Niveau im Vergleich zu den tatsächlich applizierten Mengen (vgl. Tabelle 6) insgesamt stark absinken. Auffällig ist das niedrige Niveau zur dritten N-Gabe.

Einschränkend muss auf die Probleme hingewiesen werden, die sich aus dem begrenzten Datenumfang ergeben. Wie bereits dargestellt, musste dadurch auf eine Validierung verzichtet werden. Die Auswahl von Modellen erschwert sich, da die Entscheidung für ein „besseres“ neuronales Netzwerk gegebenenfalls anhand von subjektiven Einschätzungen erfolgen muss.

4.7 Data Mining III: Versuch Görzig 04

Die Datenquantität stellt sich für den Versuch *Görzig 04* wesentlich besser dar. Nach der Datenaufbereitung stehen 4329 vollständige Instanzen zur Verfügung. Dies entspricht in etwa der zehnfachen Menge wie in den IKB-Versuchen. Die Abdeckung des Merkmalsraumes in Bezug auf die Stickstoff-Applikationen ist allerdings durch die geringe Anzahl von Teilflächen der Variante „Merkmalsraum“ eingeschränkt (vgl. Tabelle 8).

4.7.1 Datenexploration

Zunächst wurde eine Korrektur vorgenommen, die sich aus der visuellen Sichtung der Histogramme ergab. In diesen lassen sich vereinzelte Ausreisser in Bezug auf das Attribut *EM_38* erkennen. Deshalb wurden alle Instanzen, deren Wert größer als 50 mS/m betrug, von der Auswertung ausgeschlossen⁶⁹.

Der damit entstandene Datensatz ist in Abbildung 48 mit einem *Data Audit* dargestellt. Die Attribute werden mit Histogrammen sowie mit Kennzahlen der deskriptiven Statistik (Minimum, Maximum, arithmetischer Mittelwert, Standardabweichung und Schiefe der Verteilung) beschrieben.

Bei den Stickstoff-Applikationen lässt sich der Einfluss der Versuchsvarianten erkennen. Zur ersten N-Gabe wurde in allen Varianten schlageinheitlich 60 kg N/ha ausgebracht. Lediglich über die Versuchs-Variante *Merkmalsraum* wurde ein gewisser Ausgleich erreicht. Das Resultat ist ein Histogramm mit einer Dominanz der Ausprägung 60 kg N/ha. Es exis-

⁶⁹ Diese Instanzen befinden sich in einem zusammenhängenden Bereich des Schlages.

tieren ca. 3500 Instanzen mit einer Applikation von ca. 60 kg N/ha zur ersten N-Gabe, aber lediglich ca. 40 Instanzen mit einer Applikation von 90 kg N/ha. In einigen Stichproben zeigte sich allerdings, dass selbst von diesen wenigen Instanzen das Spektrum der damit verbundenen Bodendaten vergleichsweise gut abgedeckt wurde.

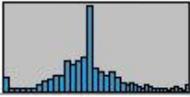
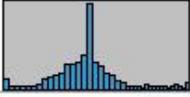
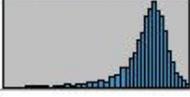
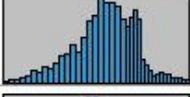
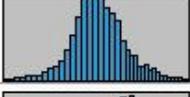
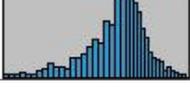
Attribut	Histogramm	Minimum	Maximum	Mittelwert	Std. Abw.	Schiefe
N_1		0	100	57.64	14.13	-2.13
N_2		0	100	39.69	17.13	0.6
N_3		0	100	38.49	16.03	0.92
REIP_2		721.71	727.18	725.69	0.62	-1.01
REIP_3		722.92	729.53	728.17	0.64	-1.84
EM_38		17.97	49.95	33.68	4.93	0.01
ERTRAG_03		1.19	12.38	6.27	1.48	0.59
ERTRAG_04		6.42	11.37	9.15	0.74	-0.52

Abbildung 48: Data Audit des Datensatzes Görzig 04

In den darauf folgenden Applikationen wird eine gleichmäßigere Datendichte des entsprechenden Attributes erreicht. Dennoch konzentriert sich die Verteilung immer noch auf einen Bereich. Die Mittelwerte der Teil-Applikationen liegen nahe bei den Werten der schlageinheitlichen Variante (vgl. Tabelle 7, Kap. 4.2.1).

In den Histogrammen der beiden Vegetationsindices zeigen sich jeweils ähnliche, links-schiefe Verteilungen. Die Werte im unteren Bereich stammen überwiegend aus den Teilflächen der Variante Merkmalsraum mit sehr geringen Applikationsmengen⁷⁰. Ansonsten liegen die Werte am Mittelwert konzentriert vor. Das Maximum liegt nur ca. 1nm vom Mittelwert entfernt. Das Niveau und damit die N-Versorgung ist im Vergleich zu den beiden anderen Versuchen sehr hoch (vgl. Abbildung 31 sowie Abbildung 37). Die Standardabweichung liegt unter der der anderen Versuche.

⁷⁰ Die Instanzen wurden dazu in der Graphik interaktiv ausgewählt und genauer analysiert.

Das Attribut *EM_38* zeigt sich linksseitig annähernd normalverteilt, wobei sich die Werte im Bereich zwischen 30 und 40 mS/m häufen. Der Mittelwert liegt bei 33.68 mS/m und damit zwischen den Mittelwerten der Schläge *Schafhof* und *D4*. Die Standardabweichung ist verhältnismäßig hoch.

Der Vorjahresertrag (*Ertrag_03*, Körnermais, feucht) erstreckt sich in einer sehr breiten Spanne zwischen 1 und 12 t/ha mit einer mittleren Standardabweichung. Das Niveau liegt sehr niedrig, was auf die trockenen Verhältnisse in diesem Jahr hindeutet. Dagegen weist der Ertrag im Versuchsjahr (*Ertrag_04*) einen sehr hohen Mittelwert mit einer geringen Standardabweichung (in etwa halb so groß wie im Vorjahr) auf.

In Tabelle 19 sind die Pearson-Korrelationen zwischen den einzelnen Attributen dargestellt. Die Zusammenhänge sind aus mehreren Gesichtspunkten aufschlussreich. Auffallend hoch sind die Korrelationen der Stickstoff-Applikationen untereinander. Dies ist damit zu erklären, dass innerhalb der einzelnen Varianten derselbe Algorithmus in der zweiten und dritten Applikation verwendet wurde. Weiterhin wurde die überwiegende Anzahl der Teilflächen zur ersten N-Gabe schlageinheitlich gedüngt. Die zweite und dritte Stickstoff-Applikation der Mapping-Variante korrelieren beispielsweise vollständig. Die Variante Merkmalsraum kann diesem Effekt nur bedingt entgegenwirken: Zum einen ist die Anzahl der Teilflächen dieser Variante zu gering. Zum anderen konnten die Teilgaben nicht zufällig bestimmt werden (vgl. Kap. 4.2.2). Daher korrelieren selbst in dieser Variante die Teilmengen miteinander. Insbesondere der starke Zusammenhang zwischen *N_2* und *N_3* ist kritisch für die Modellierung der Ertragswirkung einer einzelnen Teilgabe. Es ist zu erwarten, dass der Einfluss einer einzelnen Teilapplikation auf den Ertrag überschätzt wird. Zum Zeitpunkt der ersten N-Applikation werden beispielsweise die weiteren Applikationen *N_2* und *N_3* nicht als Input-Attribut verwendet. Unter den genannten Voraussetzungen, ist die Wahrscheinlichkeit hoch, dass einer Instanz mit einer niedrigen N-Applikation zur ersten Teilgabe auch in den folgenden Applikationen eine niedrige N-Menge appliziert wird und umgekehrt. Damit kann „das Netz“ schwer lernen, welche Bedeutung die einzelnen Teilapplikationen auf den Ertrag haben. Die genauen Auswirkungen dieser Verzerrung sind jedoch schwer abzuschätzen.

Auffallend hoch sind die Zusammenhänge der Bodenattribute (*EM_38* und *Ertrag_03*) untereinander und mit dem Ertrag im Versuchsjahr. Von diesen beiden Attributen ist daher eine hohe Bedeutung für die Ertragsprognose zu erwarten. Werden nur Instanzen mit gleichen N-Applikationen ausgewählt, steigt die Stärke der linearen Zusammenhänge nochmals bedeutend an.

Tabelle 19: Korrelationsmatrix des Datensatzes Görzig 04

	<i>N_1</i>	<i>N_2</i>	<i>N_3</i>	<i>REIP_2</i>	<i>REIP_3</i>	<i>EM_38</i>	<i>Ertrag_03</i>	<i>Ertrag_04</i>
<i>N_1</i>		0.54	0.44	0.36	0.44	0.16	-0.03	0.19
<i>N_2</i>	0.54		0.77	0.03	0.34	0.12	0.13	0.22
<i>N_3</i>	0.44	0.77		0.12	0.3	0.17	0.23	0.22
<i>REIP_2</i>	0.36	0.03	0.12		0.54	0.27	0.12	0.25
<i>REIP_3</i>	0.44	0.34	0.3	0.54		0.42	0.22	0.49
<i>EM_38</i>	0.16	0.12	0.17	0.27	0.42		0.39	0.50
<i>Ertrag_03</i>	-0.03	0.13	0.23	0.12	0.22	0.39		0.54
<i>Ertrag_04</i>	0.19	0.22	0.22	0.25	0.49	0.50	0.54	

Demgegenüber sind die Korrelationen der spektralen Kennzahlen mit dem Ertrag im Versuchsjahr bei weitem nicht so hoch, wie aus den IKB-Versuchen zu erwarten war. Eine Erklärung dafür könnte in der geringeren Präzision der Messung (durch den großen Fahrgassenabstand bzw. durch die räumliche Interpolation) liegen. Der Zusammenhang von Stickstoff und Ertrag, d. h. die Ertragswirksamkeit der Stickstoffapplikation, im Versuchsjahr zeigt sich schwach ausgeprägt.

Im nächsten Schritt der Exploration wird der Merkmalsraum grafisch untersucht. Hierzu wird zunächst ein *Scatterplot* der Bodenattribute *EM_38* und *Ertrag_03* erstellt, in dem gleichzeitig der Zusammenhang mit dem *Ertrag_04* farblich illustriert ist (vgl. Abbildung 49). Wie dargestellt wurde, korrelieren diese beiden Bodenattribute untereinander mittelstark (0,39). Beide besitzen eine Verteilung, bei denen sich die Häufigkeiten zentral konzentrieren. Aus Abbildung 49 ist ersichtlich, dass sich die Punkte ebenfalls auf ein Zentrum konzentrieren (erkennbar durch die Größe der Punkte).

Das Zentrum der kombinierten Verteilung reicht von *EM_38*-Werten von 30 bis 40 mS/m und von *Ertrag_03*-Werten von 5-7 t/ha. Um diesen Bereich herum liegt ein Bereich, der verhältnismäßig gut abgedeckt ist. Dieser Bereich, an dem sich auch der Testvektor für die *Responses Surfaces* orientiert, umfasst *EM_38* Werte von 24 – 44 mS/m und gleichzeitig Ertragswerte von 3-9 t/ha, wobei Einschränkungen in den Ecken (hoher Ertrag und niedriger *EM_38*-Wert sowie umgekehrt) hingenommen werden müssen. Erkennbar ist in dieser Abbildung 49 weiterhin der hohe Einfluss dieser beiden Attribute auf den *Ertrag_04*, der farblich codiert dargestellt ist. Der Anstieg des Attributes *Ertrag_04* verläuft diagonal. Niedrige *EM_38* und *Ertrag_03* Werte bedingen in der Regel einen niedrigen *Ertrag_04* und vice versa.

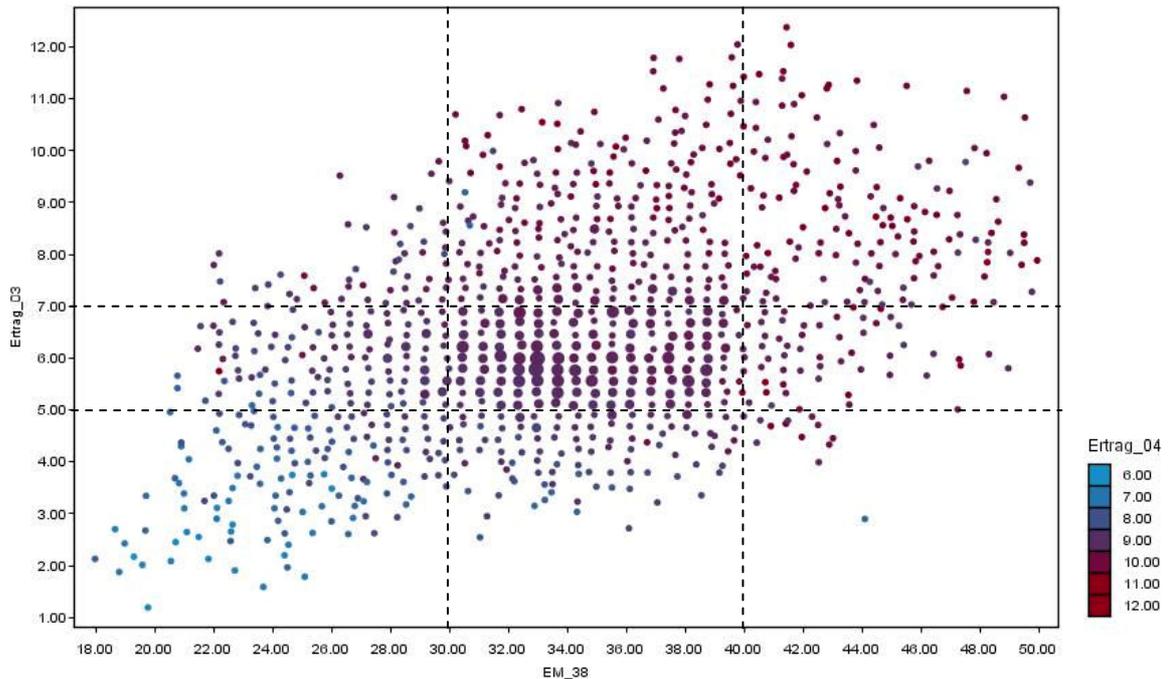


Abbildung 49: Scatterplot des Datensatzes Görzig 04 (EM_{38} , $Ertrag_{03}$, $Ertrag_{04}$)

Die Verteilung der N-Gaben innerhalb dieses Ausschnittes kann als ausreichend, aber nicht optimal angesehen werden, wie in Abbildung 50 illustriert ist. In diesem Fall wurden jeweils das Attribut N_2 kombiniert mit einem der beiden Bodenattribute ein einem Scatterplot dargestellt. Der Ausschnitt entspricht dem eben vorgestellten zentralen Bereich innerhalb der Trainingsdaten. Zur besseren Darstellung wurde (zufällig) nur die Hälfte aller Instanzen ausgewählt, sodass eine ähnliche Situation wie bei der Aufspaltung der Trainingsdaten in Training- und Testsample erzeugt wird.

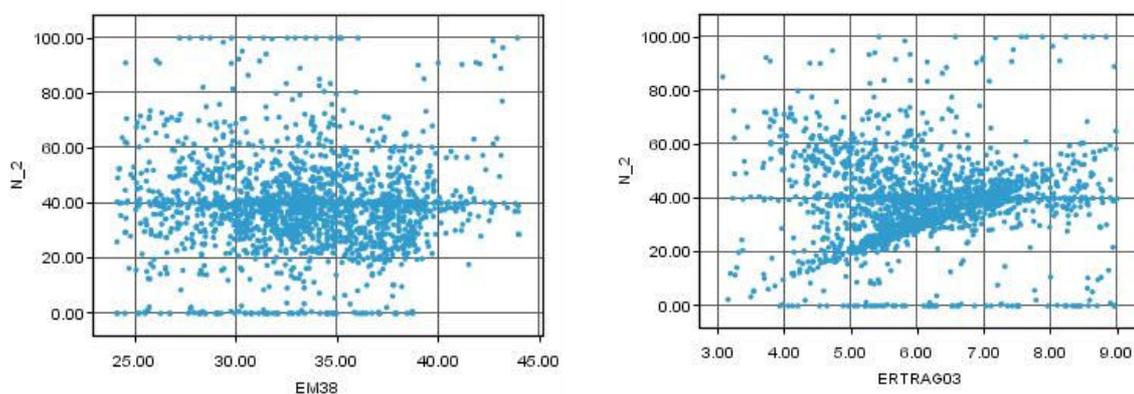


Abbildung 50: Verteilung des Attributes N_2 im Versuch Görzig 04

Dabei wird deutlich, dass einige Lücken existieren. Beispielsweise existieren sehr wenige Null-Applikationen in hohen EM_{38} -Bereichen, sowie sehr wenige hohe Applikationsmengen für niedrige Vorjahresbereiche.

Deutlich sind in der rechten Grafik in Abbildung 50 die Einflüsse des Mapping-Ansatzes zu erkennen, der zu den Instanzen führt, die auf der Diagonalen, dicht besetzten Linie führt.

Damit wurden die Ausgangsbedingung für die Data-Mining-Phase vorgestellt. Die Ergebnisse müssen vor diesem Hintergrund interpretiert werden. Beispielsweise können für Regionen mit geringer Datendichte keine validen Ergebnisse erwartet werden.

4.7.2 Erste Stickstoff-Applikation

Die Ertragsprognose zur ersten N-Gabe gestaltet sich aufgrund der eingeschränkten Variation bezüglich N_1 schwierig. Zunächst soll deshalb untersucht werden, ob durch eine Datenbalance, d. h. durch eine entsprechende Selektion der Instanzen, bessere Modellierungsergebnisse erreicht werden. Dazu werden in der Modellierung zwei Verfahren getestet.

Vergleich von Datenaufbereitungsmethoden

In der ersten Methode der Datenaufbereitung werden alle Instanzen verwendet und ein neuronales Netzwerk mit dem *Exhaustive-Prune-Algorithmus* trainiert. Das Ergebnis ist in Tabelle 20 abgetragen. Die *Estimated Accuracy* (91,76) erscheint im Vergleich zum Versuch *D4 03* niedrig. Allerdings muss dazu die geringere Präzision in der Applikation und Datenerfassung des Versuches *Görzig 04* gesehen werden. Die Topologie des trainierten neuronalen Netzwerkes weist große versteckte Schichten auf. Die relative Bedeutung der einzelnen Attribute für die Prognose entspricht ungefähr den Korrelationen dieser Attribute zum Ertrag des Versuchsjahres. Der Vorjahresertrag stellt das bedeutendste Attribut für die Ertragsprognose dar, während der ersten N-Gabe lediglich eine geringe Bedeutung (0.14) zugemessen wird. Die scheinbare elektrische Leitfähigkeit erweist sich von mittlerer Bedeutung.

Tabelle 20: Neuronales Netzwerk zur ersten N-Applikation (Versuch Görzig 04)

Estimated accuracy: 91,76	Relative Importance of Inputs	Einstellungen:
Input Layer: 3 neurons	<i>ERTRAG_03</i> : 0.55	Method: Exhaustive prune
Hidden Layer 1: 7 neurons	<i>EM_38</i> : 0.35	Stop on: Default
Hidden Layer 2: 3 neurons	<i>N_1</i> : 0.14	Prevent overtraining: true
Output Layer: 1 neuron		Sample: 50 %

In Wiederholungen des Trainingsprozesses zeigen sich sehr ähnliche Ergebnisse. Die Topologie bleibt stabil, die geschätzte Genauigkeit verändert sich kaum. Dies zeigt, dass mit diesem Datensatz die erforderliche Datenmenge erreicht wurde.

Das Ziel der Datenbalance besteht in diesem Fall darin, eine ausgeglichene Verteilung bezüglich des Attributes N_1 zu erreichen. Standardmäßig wird dies mit der Option *Reduce*

(Clementine®) erreicht (vgl. 2.4.3). Die Ergebnisse dieser Methode sind für die vorliegende Aufgabe allerdings nicht zufrieden stellend. Die Option führt dazu, dass lediglich 200 Instanzen übrig bleiben, da mit dieser Option eine Gleichverteilung erzeugt wird. Damit wird die geringste Häufigkeit in der Bandbreite als Niveau verwendet. Dadurch verschlechtert sich die Abdeckung des Merkmalsraumes auch bezüglich der restlichen Attribute. Abbildung 51 illustriert diesen Zusammenhang. In einem *Scatterplot* sind die Instanzen nach der Datenselektion in den Dimensionen *Ertrag_03*, *EM_38* und (farblich codiert) *N_1* abgetragen.

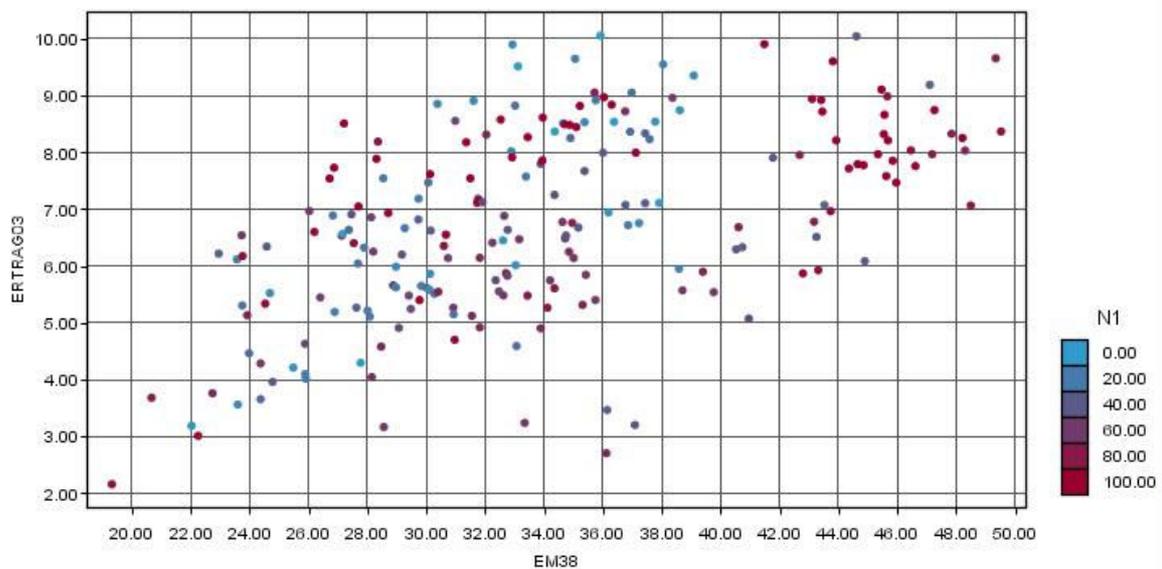


Abbildung 51: Das Ergebnis der Option „Reduce“ in Bezug auf die Datenbalance

Es wird deutlich, dass große Lücken im Merkmalsraum entstehen. Dies gilt insbesondere im Vergleich zum Ausgangsdatensatz (vgl. Abbildung 49). Beispielsweise befinden sich im *EM_38*-Bereich von 38-44 mS/m sehr wenige Punkte. Ein weiterer Effekt zeigt sich in der Variation: Dadurch, dass hier zufällig Instanzen ausgeschlossen werden, reduziert sich auch die Anzahl der Datenpunkte in den interessanten Randbereichen. Dadurch müsste der Bereich, für den valide Prognosen erstellt werden können, weiter eingengt werden. Die Erzeugung einer Gleichverteilung bezüglich der ersten Stickstoff-Applikation führt damit zu einer schlechten Ausgangslage für die Modellierung der kleinräumigen Zusammenhänge. In einigen Versuchen zeigte sich, dass mit dieser Instanzenmenge keine zufrieden stellenden Modellierungsergebnisse erreicht werden können. Die ausgegebenen trainierten Netzwerke neigen zu Überlernen oder vereinfachen die Zusammenhänge, indem sie extrem kompakte Topologien produzieren.

Aus diesem Grund wurde versucht, manuell einen Filter zur Datenbalance zu erstellen, der als Kompromiss zwischen den beiden extremen Varianten gesehen werden kann. Wie in Abschnitt 2.4.3 dargestellt wurde, kann aufgrund des „class imbalance problems“ (vgl. JAP-

KOWICZ UND STEPHEN, 2002, S. 429) keine optimale Balance ermittelt werden. Es kann nur versucht werden, den Einfluß größerer Ungleichgewichte zu verringern. Die Balance-Faktoren orientieren sich deshalb an den Histogrammen. Die Instanzen werden nach den Regeln in Tabelle 21 ausbalanciert.

Beispielsweise werden dazu auch die Instanzen, die eine Nulldüngung zur ersten N-Gabe erhielten, um die Hälfte verringert. Für die stark vertretene Gruppe der Applikationen mit ca. 60 kg N/ha wurde die Anzahl um 70% reduziert, allerdings wurden dazu die Instanzen ausgewählt, die in den restlichen Attributen nahe am Mittelwert liegen, um deren Variation nicht zu verringern. Die Reduktion der Teilflächen mit einer N-1-Applikation von über 90 kg N/ha wurde auf 50 Prozent festgesetzt.

Tabelle 21: Balance-Regeln zur Optimierung des Trainings

Balance-Faktor	Bedingung
0.5	$N_1 < 10$
0.3	$N_1 \geq 56$ UND $N_1 < 64$ UND $ERTRAG_{03} < 7$ UND $ERTRAG_{03} > 5$ UND $EM_{38} < 40$ UND $EM_{38} > 30$
0.5	$N_1 > 90$

Als Resultat zeigt sich ein relativ gleichmäßig abgedeckter Merkmalsraum mit geringeren Häufungen. Mit den selektierten Instanzen wurde wiederum ein neuronales Netzwerk trainiert (vgl. Tabelle 22). Das Lernergebnis, bezogen auf die *Estimated Accuracy*, ist geringfügig besser als ohne Datenbalance. Die Topologie bleibt dagegen gleich. Die Bedeutung des Attributes N_1 für die Prognose steigt an. Dasselbe gilt für den Vorjahresertrag.

Tabelle 22: Neuronales Netzwerk zur ersten N-Applikation (Versuch Görzig 04, mit Datenbalance)

Estimated accuracy: 91,81	Relative Importance of Inputs	Einstellungen:
Input Layer: 3 neurons	$ERTRAG_{03}$: 0.66	Method: Exhaustive prune
Hidden Layer 1: 7 neurons	EM_{38} : 0.32	Stop on: Default
Hidden Layer 2 : 3 neurons	N_1 : 0.17	Prevent overtraining: true
Output Layer: 1 neuron		Sample: 50 %

Die Analysen der neuronalen Netzwerke zeigen, dass sich die beiden Varianten nur leicht unterscheiden. Daher wird eine Kreuzvalidierung durchgeführt, um die beiden Varianten zu testen. Daraus ergeben sich ebenfalls ähnliche Ergebnisse, wie Tabelle 23 zu entnehmen ist. Der maximale bzw. minimale Fehler liegt bei beiden Varianten auf demselben Niveau. Selbiges gilt für die weiteren Maßzahlen. In der Maßzahl R^2 verbessert sich das Ergebnis al-

lerdings merklich von 0.47 auf 0.52.

Gleichzeitig bietet sich diese Analyse an, um die Ergebnisse mit denen anderer Autoren einzuordnen. Als Vergleich bietet sich die Studie DRUMMOND ET AL. (2002, s.p.) an, die ähnliche Attribute (d. h. Bodenattribute) sowie eine ähnliche Datenaufbereitung verwendet (vgl. Kap. 3.2). Der Median dieser Ergebnisse liegt bei einem R^2 von 0.38. Die Ergebnisse liegen somit im oberen Bereich der Spanne der ausgewerteten Versuche.

Tabelle 23: Ergebnisse der Kreuzvalidierung für zwei Datenaufbereitungsmethoden

	Minimaler Fehler	Maximaler Fehler	Mittelwert des Fehlers	Mittlerer absoluter Fehler	SEP	Lineare Korrelation (R^2)
Mit Datenbalance	-1.81	2.18	0.00	0.42	0.54	0.52
Ohne Datenbalance	-1.99	2.20	-0.01	0.41	0.54	0.47

Als Zwischenfazit kann somit festgehalten werden, dass sich zum Training eines neuronalen Netzes unter den vorgestellten Voraussetzungen im Wesentlichen zwei Datenaufbereitungsmethoden eignen. Entweder werden alle Instanzen in der Modellierungsphase verwendet, oder es werden komplexere (aber manuelle) Verfahren zur Ausbalancierung angewendet. Wie die Sensitivität zeigt, kann mit einer Ausbalancierung der Einfluss von N_1 auf die Ertragsprognose verstärkt werden und damit die Zusammenhänge dieses Attributes besser abgebildet werden. Aus diesem Grund werden zur Ableitung von Entscheidungsregeln diese Datenaufbereitungsmethode und das entsprechende Modell verwendet.

Extraktion von Mustern

Die Reaktion des trainierten Netzes aus den balancierten Daten wird wiederum in einem *Response Surface* untersucht. In einem ersten Test wird das Attribut EM_{38} konstant am Mittelwert gehalten und die Attribute N_1 und $Ertrag_{03}$ variiert. Aus Abbildung 52 können dadurch Produktionskurven für verschiedene Vorjahreserträge abgelesen werden.

Hier ist zu erkennen, dass der Vorjahreserträge einen sehr ausgeprägten Einfluss auf den prognostizierten Ertrag besitzt. Nach dieser Prognose kann bei Vorjahreserträgen von 3 t/ha ein Maximalertrag von ca. 8 t/ha erreicht werden, während für einen Vorjahresertrag von 9 t/ha ein Ertrag von ca. 10 t/ha vorhergesagt wird.

Gleichzeitig wird der prognostizierte Ertrag von der ersten N-Gabe beeinflusst. Die Verläufe sind in den jeweiligen Ertragsregionen unterschiedlich. In den niedrigen Ertragsbereichen steigt der Ertrag bis zu eine N-Gabe von 70 kg N/ha vergleichsweise steil an und fällt dann wieder gemäßigt steil ab. In mittleren Ertragsbereichen ist der Anstieg ebenfalls steil. Das Ertragsmaximum wird mit 60 kg N/ha erreicht. Danach bleibt der Ertrag ungefähr auf

diesem Niveau. In hohen Ertragsbereichen des Vorjahres erfolgt der Anstieg erst später, bei ungefähr 20 kg N/ha. Das Maximum liegt unter diesen Voraussetzungen bei 50 kg N/ha. Bei einer weiteren Erhöhung der N-Applikation fällt der prognostizierte Ertrag vergleichsweise steil ab.

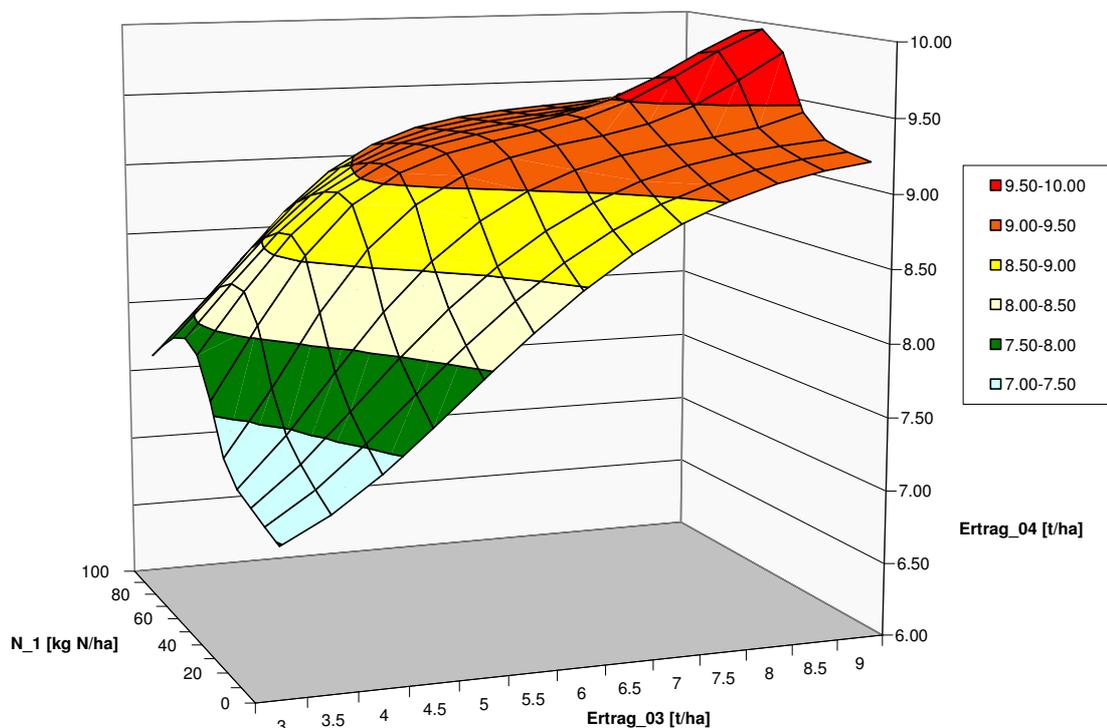


Abbildung 52: Response Surface zur ersten N-Applikation (Ertrag_03, Versuch Görzig 04)

Aus den Produktionsflächen wird deutlich, dass die N-Menge, die zu einem Ertragsmaximum führt, sich innerhalb der verschiedenen Ertragszonen nur wenig unterscheidet. Allerdings werden hier auch mögliche Risiken deutlich: Eine zu hohe oder zu niedrige N_1 -Applikation in hohen Ertragsbereichen führt nach diesem Modell zu Mindererträgen. In mittleren und niedrigen Ertragsbereichen führt vor allem eine zu geringe N_1 -Menge zu beträchtlichen Ertragsverlusten (bis zu einer Tonne pro Hektar).

Zu beachten ist allerdings, dass der Merkmalsraum wie oben erwähnt nicht verzerrungsfrei ist. Insbesondere durch die Datenbalance wird der Einfluss der Variante Merkmalsraum verstärkt. In dieser Variante wurde den Teilflächen mit einer hohen N_1 Gabe auch oftmals in den späteren N-Gaben hohe Mengen appliziert und umgekehrt. Dadurch könnten mit dem Netzwerktraining die Effekte aller Teilapplikationen partiell dieser einen Teilapplikation (N_1) zugerechnet worden sein. Es kann also nicht unterschieden werden, ob eine Ertragswirkung von N_1 ausgeht, oder von den beiden folgenden N-Applikationen.

Aus ökonomischer Sicht ergeben sich mit dieser Prognose optimale N_1 -Applikationen von 50-60 kg N/ha, wobei die optimale N-Gabe bei niedrigen Vorjahreserträgen höher ist als bei hohen Vorjahreserträgen. Es ist somit ansatzweise das Muster „inverser Mapping-

Ansatz“ (vgl. Kap. 4.6.3) zu erkennen. Gleichzeitig kann festgehalten werden, dass eine Düngung außerhalb von N_{opt} zu stark negativen ökonomischen Konsequenzen führen kann. Diese Schlussfolgerung kann allerdings aus den dargestellten Gründen nicht ausschließlich auf die erste Teilapplikation bezogen werden.

Bodenleitfähigkeit und N-Applikation

Als nächsten Schritt wurde der Einfluss der Bodenleitfähigkeit auf die optimale Düngung untersucht. Hierzu wurde wiederum ein Response Surface (Abbildung 53) erstellt, bei dem diesmal das Attribut EM_{38} und N_1 variiert werden und der Vorjahresertrag konstant gehalten wird. Dabei ergibt sich ein ähnliches Bild. Auch hier wird das Ertragsmaximum für alle EM_{38} -Werte bei einer N-Gabe 50-60 kg N/ha erreicht. In diesem Fall führt aber eine Applikation über diesen Bereich hinaus insbesondere in niedrigen EM_{38} -Bereichen zu einem starken Ertragsabfall. Der Verlauf der Ertragsfunktion in diesem Bereich erscheint extrem: Der Ertrag steigt bis zu einer Applikation von 50 kg/ha sehr steil an und fällt dann ebenso stark wieder ab. Dieser Bereich des Merkmalsraumes zeichnet sich allerdings auch nicht durch eine hohe Datendichte aus (insbesondere in den „Ecken“), sodass die Validität dieser extremen Steigungen angezweifelt werden darf.

Aus ökonomischer Sicht ergibt sich durchgängig eine optimale N-Gabe von 50 kg N/ha. Der Ertragsanstieg, der mit einer Erhöhung der Applikation auf 60 kg N/ha in den Bereichen niedriger Bodenleitfähigkeit prognostiziert wird (Ertragsmaximum), reicht nicht aus, um die Grenzkosten zu decken.

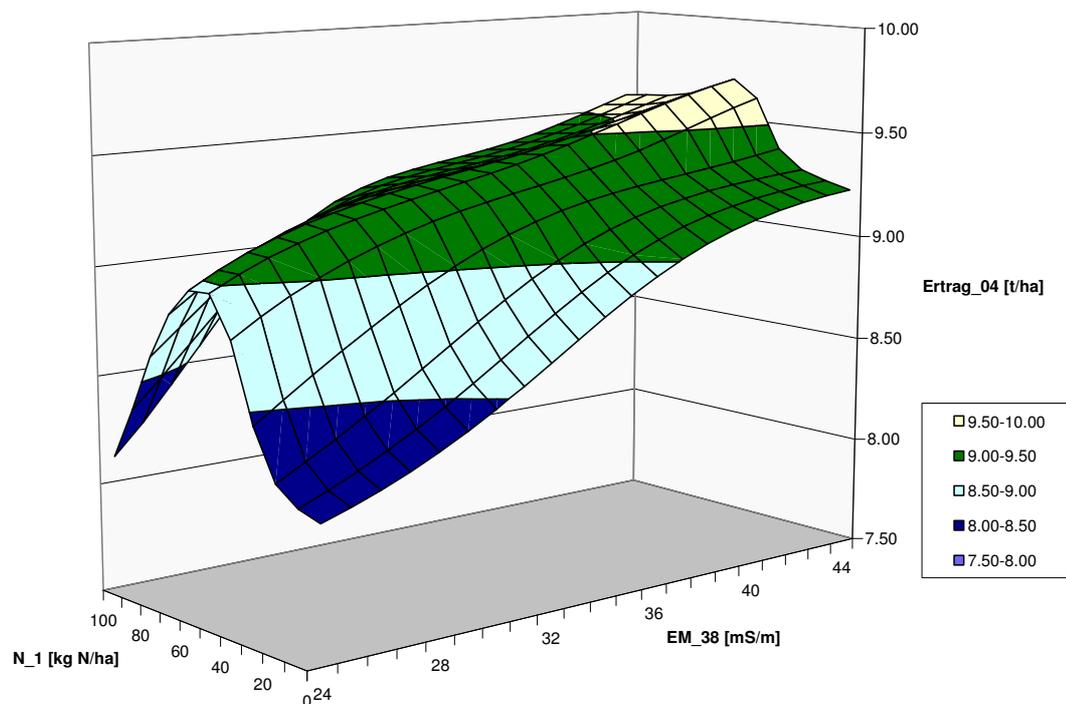


Abbildung 53: Response Surface zur ersten N-Applikation (EM_{38} , Versuch Görzig 04)

Die Interpretation der Response Surfaces aus ökonomischer Sicht führt also zu Regeln, die eine optimale Stickstoff-Gabe im Bereich von 50-60 kg N/ha erwarten lassen. Tendenziell führen niedrigere Werte in den beiden Attributen *EM_38* und *Ertrag_03* zu geringfügig höheren empfohlenen N-Gaben. Eine teilflächenspezifische Applikation verspricht nach diesem Modell kein großes Potenzial, da sich die optimalen N-Gaben nur geringfügig unterscheiden. Allerdings reagieren die prognostizierten Erträge in einigen Fällen stark (negativ), falls eine Applikation außerhalb dieses Bereiches gewählt wird.

Die optimierten N-Gaben weichen also aus Sicht der *Response Surfaces* nicht weit von den betriebsüblichen Gaben ab. Durch die Anwendung des Entscheidungsbaum-Modells kann hier weiter überprüft werden, welche Regeln zur optimierten Düngung generiert werden. Dazu wird das neuronale Netzwerk (mit Datenbalance) ex post auf die Schlagdaten des Versuches *Görzig 04* angewendet und die ökonomisch optimale N-Menge für alle Teilflächen des Schlag ermittelt. Die optimierten N-Gaben werden zusammen mit den Teilflächendaten als Input in einen Entscheidungsbaum-Algorithmus eingegeben. Der daraus entstehende Entscheidungsbaum ist in Abbildung 54 dargestellt.

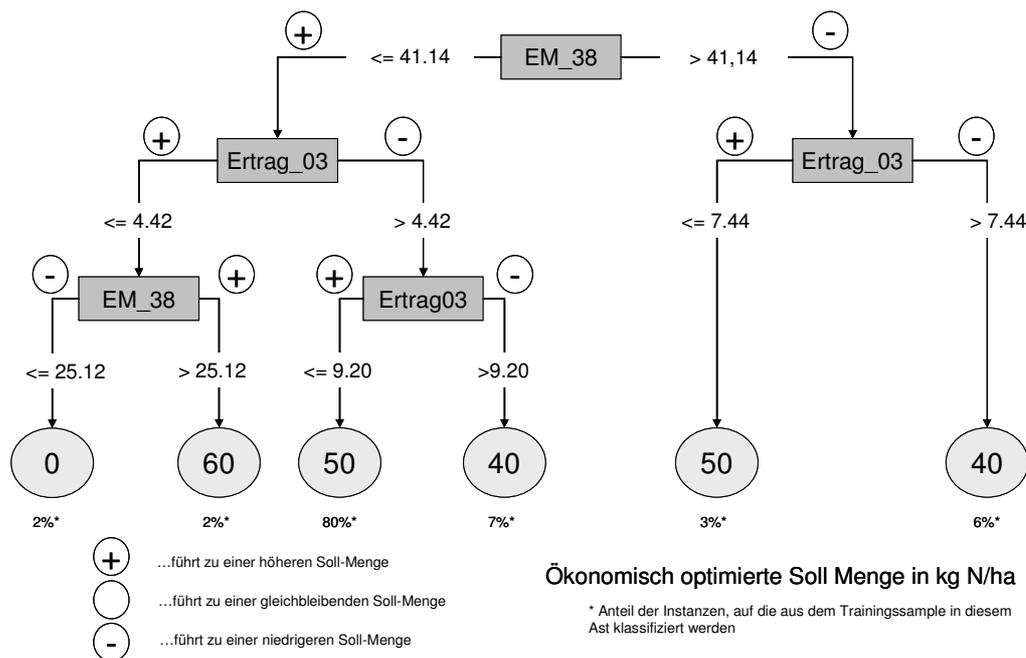


Abbildung 54: Entscheidungsbaum zur ersten N-Applikation

Aus dem Entscheidungsbaum wird deutlich, dass sich die optimalen N-Mengen nur wenig unterscheiden. Auffällig ist die linke Abzweigung, die zu einer empfohlenen Null-Düngergabe führt. Dabei handelt es sich um Teilflächen, die in den Vorjahren einen sehr geringen Ertrag hatten und gleichzeitig eine sehr niedrige Bodenleitfähigkeit aufweisen. Diese Ableitung ist aus den *Response Surfaces* nicht zu beobachten. Es fallen allerdings nur sehr wenige Instanzen des Schlags in diesen Bereich (ca. 2%) und die Abdeckung des entsprechenden Ausschnittes im Merkmalsraum ist entsprechend schlecht.

Der Großteil des Schlages würde somit ex post mit 50 kg N/ha gedüngt werden. Das entspricht einer Absenkung um 10 kg N/ha vom betrieblichen Niveau. Wie auch im Versuch *D4 03* sind die Muster eines „inversen Mapping-Ansatzes“ zu erkennen, allerdings in diesem Fall nicht stark ausgeprägt.

4.7.3 Zweite Stickstoff-Applikation

Die Ausgangsbedingungen zur Modellierung der zweiten N-Gabe gestalten sich besser. Die zweite N-Gabe wurde nach verschiedenen Varianten ausgebracht, was zu einer wesentlich besseren Abdeckung des Entscheidungsraumes führen sollte. Es wird zwar durch jede Versuchs-Variante eine Verzerrung bezüglich der N-Applikationen erzeugt, aber da die N-Applikation nach unterschiedlichen Entscheidungsgrundlagen bemessen wird, „entschärft“ sich diese Verzerrung bezogen auf den gesamten Datensatz. Der am häufigsten vertretene Wert der zweiten N-Applikation ist 40 kg N/ha. Die Verteilung konzentriert sich nicht mehr so intensiv auf diesen Wert wie zur ersten N-Applikation.

Datenaufbereitung

Zur Modellierung der zweiten N-Applikation werden wiederum verschiedene Varianten der Datenaufbereitung getestet. Insbesondere die Attributauswahl sowie die Datenbalance können variiert werden. Dazu wurden verschiedene Varianten getestet. Als Ergebnis zeigen sich sehr ähnliche *Response Surfaces* und Entscheidungsbäume. Die Grobstrukturen tauchen beständig auf, während sich die detaillierten Strukturen selbst unter gleichen Trainingsbedingungen der neuronalen Netze verändern. Im Folgenden wird deshalb auf eine Analyse der Datenaufbereitung verzichtet.

Neuronales Netzwerk

Unter Verwendung aller Attribute zu diesem Zeitpunkt wird ein neuronales Netz trainiert, dessen Eigenschaften in Tabelle 24 zusammengefasst sind. Die komplexe Topologie deutet auf leichtes Überlernen hin.

Tabelle 24: Neuronales Netzwerkes zur zweiten N-Applikation (Versuch *Görzig 04*)

Estimated accuracy: 92,58	Relative Importance of Inputs:	Einstellungen:
Input Layer: 5 neurons	<i>ERTRAG_03</i> : 0.41	Method: Exhaustive prune
Hidden Layer 1: 14 neurons	<i>EM_38</i> : 0.28	Stop on: Default
Hidden Layer 2: 9 neurons	<i>N_1</i> : 0.23	Prevent overtraining: true
Output Layer: 1 neuron	<i>N_2</i> : 0.19	Sample: 50 %
	<i>REIP_2</i> : 0.18	

Es kann beobachtet werden, dass sich durch die zusätzlichen Attribute (*N_2*, *REIP_2*) die geschätzte Genauigkeit der Vorhersage merklich erhöht (von 91.81 auf 92.58). Vor diesem

Hintergrund ist auffallend, dass die zusätzliche Attribute (N_2 und $REIP_2$) eine geringe Bedeutung für die Ertragsprognose haben, obwohl die Erhöhung der Genauigkeit nur über diese beiden Attribute zustande kommen kann.

Die Bodenattribute haben wie schon zur ersten N-Gabe eine hohe Bedeutung für die Ertragsprognose. Die Bedeutung der ersten N-Gabe erhöht sich gegenüber der Ertragsprognose zur ersten N-Applikation. Die Prognoseleistung, die mit Hilfe einer Kreuzvalidierung geschätzt wird, verbessert sich ebenfalls, wie Tabelle 25 zeigt. Im Vergleich zur ersten N-Applikation verbessern sich alle Gütemaße der Prognose merklich.

Tabelle 25: Ergebnisse der Kreuzvalidierung zur zweiten N-Applikation

Minimaler Fehler	Maximaler Fehler	Mittelwert des Fehlers	Mittlerer absoluter Fehler	SEP	Bestimmtheitsmaß (R^2)
-1.99	1.99	-0.01	0.39	0.51	0.53

In der Analyse dieses Modells ist vor allem das neue Attribut $REIP_2$ interessant. Dazu wird wiederum ein Response Surface erstellt. Das Attribut $REIP_2$ sowie die N_2 -Menge werden variiert und die restlichen Werte am Mittelwert gehalten. Abbildung 55 illustriert die ausgegebenen Werte des neuronalen Netzwerkes.

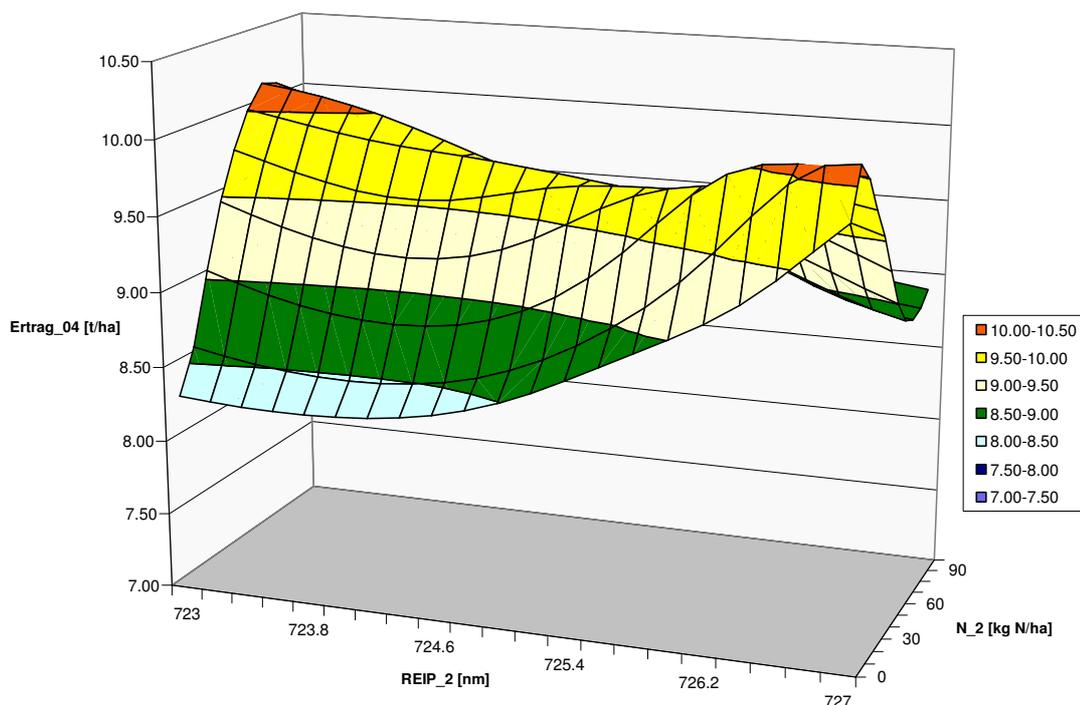


Abbildung 55: Response Surface zur zweiten N-Applikation ($REIP_2$, Versuch Görzig 04)

In Teilflächen mit niedrigen $REIP_2$ -Werten (schlechte N-Versorgung) führt eine Erhöhung der N_2 -Gabe zu einem steilen Anstieg des prognostizierten Ertrages. Das Ertragsmaximum wird bei ca. 60 kg N/ha erreicht. Im Vergleich zu einer Null-Applikation kann der

Ertrag demnach um ca. 1.5 t/ha (!) erhöht werden. Dieser Anstieg erscheint übertrieben. Wiederum muss dieses Ergebnis vor dem Hintergrund der schlechten Abdeckung des Merkmalsraumes in diesen Bereichen gesehen werden. Nach Erreichung des Ertragmaximums führt eine Erhöhung der Applikation zu einem ebenso steilen Ertragsabfall.

Je höher das *REIP_2*-Attribut wird, desto schneller wird mit einer Erhöhung der N-Applikation das Ertragsmaximum erreicht. Bei 725 nm wird das Ertragsmaximum bereits mit 50 kg N/ha erreicht. Bei 727 nm reichen 10 kg N/ha aus, um den maximalen Ertrag zu erreichen. Auffällig ist weiterhin, dass das prognostizierte Ertragsmaximum (ca. 10 t/ha) eines sehr schlecht versorgten Bestandes ähnlich hoch ist, wie das eines sehr gut versorgten Bestandes, nur dass dazu eine geringere N-Applikation nötig ist.

Aus diesen Werten lassen sich Muster zur optimalen ökonomischen N-Gabe ableiten. Die optimale N-Gabe für sehr schlecht versorgte Bestände liegt bei 60 kg N/ha und verringert sich schrittweise bis hin zu 10 kg N/ha bei sehr gut versorgten Beständen. Dies entspricht einer „klassischen“ Sensorsteuerung (vgl. Kap. 3.1.2).

Zu untersuchen ist weiterhin, wie sich eine Variation in den Bodenattributen auf die optimale N-Menge auswirkt. Dies wird wiederum anhand von *Response Surfaces* untersucht. Das Ergebnis bezüglich des Vorjahresestrages ist in Abbildung 56 dargestellt.

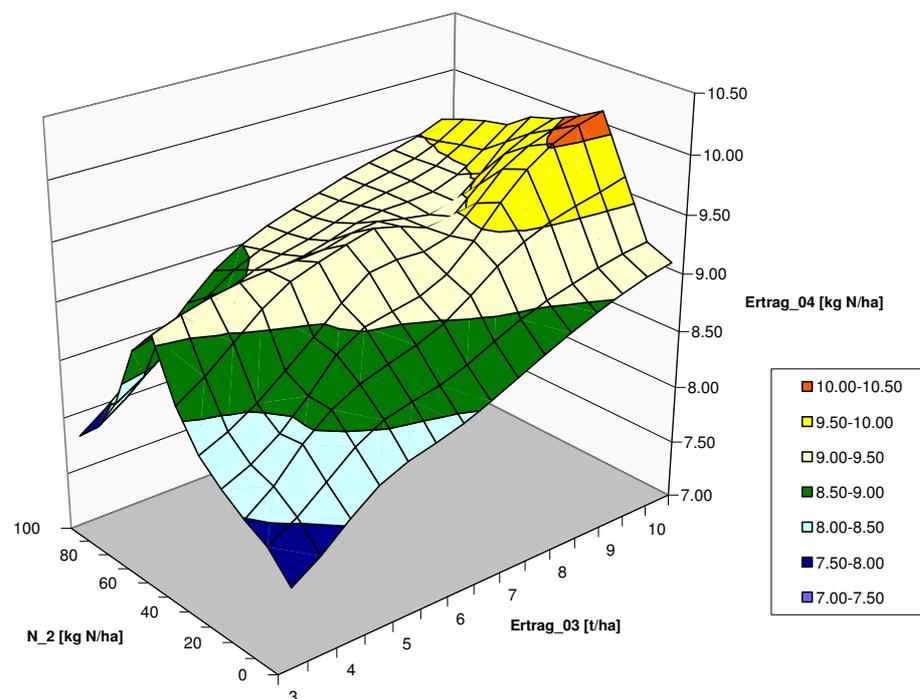


Abbildung 56: Response Surface zur zweiten N-Applikation (Ertrag_03, Versuch Görzig 04)

Es zeigt sich eine sehr unregelmäßig geformte Oberfläche, was auf Überlernen hindeutet. Generell wird für einen hohen Vorjahresetrag auch ein hoher Ertrag prognostiziert. In niedrigen Ertragsbereichen des Vorjahres führt eine Erhöhung der N-Applikation zu einem

steilen Anstieg des prognostizierten Ertrages, der wiederum nach Erreichung des Ertragsmaximums steil absinkt. Das Ertragsmaximum für einen Vorjahresertrag von 3 t/ha wird bei 60 kg N/ha erreicht. Bei hohen Ertragbereichen wird selbiges bereits mit einer N-Gabe von 30 kg N/ha erreicht. Der Einbruch des Ertrages bei einer Applikation über das Ertragsmaximum hinaus fällt in diesem Modell bei hohen Ertragbereichen schwach aus.

Die ökonomische Konsequenz ist wie zur ersten Applikation in einem inversen Mapping-Ansatz zu sehen. Die optimale N-Menge sinkt im Übergang von niedrigen auf hohen Ertragbereichen von 60 kg N/ha auf 30 kg N/ha. Als letztes soll der Einfluss der scheinbaren Bodenleitfähigkeit im Ertragsprognosemodell untersucht werden. Das entsprechende *Response Surface* ist in Abbildung 57 dargestellt.

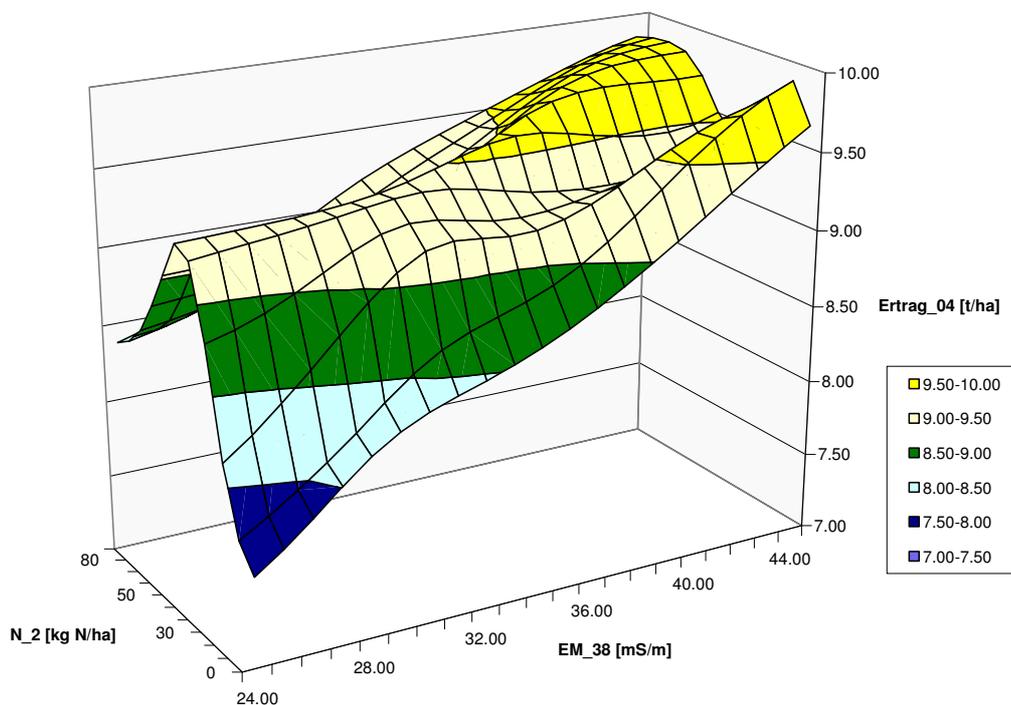


Abbildung 57: Response Surface zur zweiten N-Applikation (EM_38, Versuch Görzig 04)

Wiederum zeigt sich eine unregelmäßig geformte Oberfläche. Das Ertragsmaximum wird bei Teilflächen mit niedrigen EM_{38} -Werten bereits bei ca. 40 kg N/ha erreicht. Der An- und Abstieg bezogen auf den maximal prognostizierten Ertrag verläuft in diesem Bereich sehr steil. In Bereichen mit höherer Bodenleitfähigkeit entsteht eine Oberfläche mit zwei Gipfeln, die ungefähr gleich hoch liegen. Aufgrund der eingeschränkten Anzahl von Trainingsinstanzen für den Bereich mit sehr hoher Bodenleitfähigkeit und sehr niedriger N-Applikation muss der „vordere“ Gipfel jedoch kritisch gesehen werden. Ohne diesen ungewöhnlichen Gipfel ergibt sich ein relativ konsistenter Verlauf der N-Menge, die zur Erreichung des Ertragsmaximums führt. Diese Menge wird umso größer, je höher die Bodenleitfähigkeit wird. Bei sehr hohen EM_{38} -Werten liegt die ertragsmaximierende N-Menge bei 50-70 kg N/ha. Auffällig ist weiterhin, dass der Ertragsabfall bei sehr hohen EM_{38} -

Werten sehr flach ausfällt.

Die optimalen N-Mengen aus ökonomischer Sicht verlaufen ähnlich zu denen zur Erreichung des Ertragsmaximums. Damit wird ein Trend erkennbar, der sich bei der ersten N-Applikation nicht gezeigt hat: ein „klassischer“ Mapping-Ansatz bezüglich der scheinbaren elektrischen Leitfähigkeit.

Mit diesen drei Ausschnitten konnten einige stabile Zusammenhänge dargestellt werden. Wiederum reagieren die prognostizierten Erträge in einigen Fällen stark (negativ), falls eine Applikation neben der jeweiligen optimalen N-Applikation gewählt wird.

Die Frage ist allerdings, wie diese Zusammenhänge in Regeln abgebildet werden. Dazu wird wiederum das Entscheidungsbaumverfahren angewendet. Das Ergebnis ist in Abbildung 58 abgetragen. Es zeigt sich ein dominanter Einfluss des Attributes *REIP_2*. Wie aus den *Response Surfaces* zu erwarten war führt ein niedriger *REIP_2*-Wert zu einer Erhöhung der optimalen *N_2*-Menge. Damit zeigt sich, dass ein niedriger Wert in der Sensitivitätsanalyse des neuronalen Netzwerkes nicht notwendigerweise auf die Bedeutung bezüglich der ökonomischen Optimierung schließen lässt.

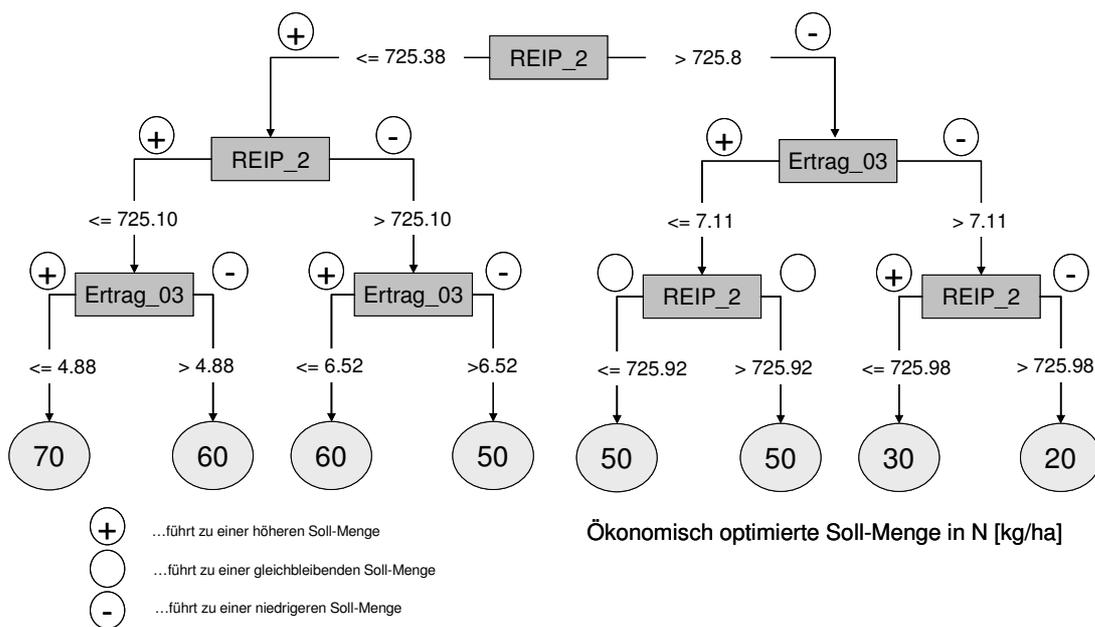


Abbildung 58: Entscheidungsbaum zur zweiten N-Applikation (Versuch Görzig 04)

Es wird lediglich ein weiteres Attribut, nämlich der Ertrag des Vorjahres, zur Entscheidung verwendet. Allerdings muss angemerkt werden, dass falls die Beschränkung der *Maximum Tree Depth* auf drei Ebenen aufgehoben wird, die restlichen Attribute durchaus zur Entscheidung herangezogen werden. Die Entscheidungsrichtung über die scheinbare elektrische Bodenleitfähigkeit entspricht dem Wirkungszusammenhang, der sich über die *Response Surfaces* gezeigt hat.

4.7.4 Dritte Stickstoff-Applikation

Auch zur dritten Stickstoff-Applikation wurden verschiedene Verfahren der Datenaufbereitung getestet. Wiederum zeigen sich nur Abweichungen in den „Feinstrukturen“. Schwerpunktmäßig wurde in den trainierten Modellen die Beziehung *REIP_3* und *N_3* untersucht.

In einem ersten Modell wurden sämtliche zur Verfügung stehenden Attribute als Input-Attribute verwendet. Das Ergebnis ist in Tabelle 26 abgebildet. Das Netz hat im Vergleich zu den Netzen der vorausgegangenen Applikationen eine höhere *Estimated Accuracy* (92.86). Durch die Hinzunahme der beiden Attribute *REIP_3* und *N_3* verbesserte sich somit die Prognose. Es entsteht ein relativ komplexes Netz mit zwei großen versteckten Schichten. Mit der Anwendung des *Exhaustive-Prune-Algorithmus* werden keine Input-Einheiten abgeschnitten. Somit sind alle Attribute für die Prognose bedeutend.

Tabelle 26: Neuronales Netz zur dritten N-Applikation (Versuch Görizg 04)

Estimated accuracy: 92,86	Relative Importance of Inputs:		Einstellungen:
Input Layer: 7 neurons	<i>ERTRAG_03</i>	0.47	Method: Exhaustive Prune
Hidden Layer 1: 11 neurons	<i>EM_38</i>	0.39	Stop on: Default
Hidden Layer 2: 7 neurons	<i>REIP_3</i>	0.34	Prevent overtraining: true
Output Layer: 1 neuron	<i>N_1</i>	0.24	Sample: 50 %
	<i>N_3</i>	0.21	
	<i>N_2</i>	0.21	
	<i>REIP_2</i>	0.16	

Bei der *Relative Importance* der Input-Einheiten rückt das Attribut *REIP_3* mit 0.34 auf den dritten Rang vor, während die Stickstoff-Applikationen relativ einheitlich an den hinteren Plätzen landen. *REIP_2* wird durch die spätere Messung *REIP_3* verdrängt und befindet sich an letzter Position. Zur Beurteilung der trainierten Netze müssen nochmals die in der Exploration angeführten hohen Korrelationen zwischen den beiden Teilgaben *N_2* und *N_3* sowie den entsprechenden Vegetationsindices herangezogen werden. Durch die starken Zusammenhänge können die Einzeleffekte nicht isoliert betrachtet werden. In der Prognose kann unter Umständen der Einfluss eines dieser Attribute durch das entsprechende korrelierende Attribut abgebildet werden.

In einem ersten Schritt werden die Zusammenhänge der neuen Attribute (*N_3* und *REIP_3*) auf den prognostizierten Ertrag untersucht. In umfangreichen Analysen mit Hilfe von Pivot-Tabellen zeigt sich, dass der Zusammenhang stark von den Einflüssen der restlichen Attribute, insbesondere von *Ertrag_03* und *EM_38*, geprägt ist. Die Verläufe der *Response-Surfaces* unterscheiden sich stark. Es wird dazu übergegangen, den durchschnittlichen prognostizierten Ertrag darzustellen. Der in den Modellen abgebildete Zusammenhang zwi-

schen N_3 , $REIP_{49}$ und dem Ertrag im Versuchsjahr ist in Abbildung 59 dargestellt. Es zeigt sich eine unregelmäßig geformte Oberfläche. Aus dieser Darstellung kann abgeleitet werden, dass eine höhere N-Aufnahme zu höheren prognostizierten Erträgen führt.

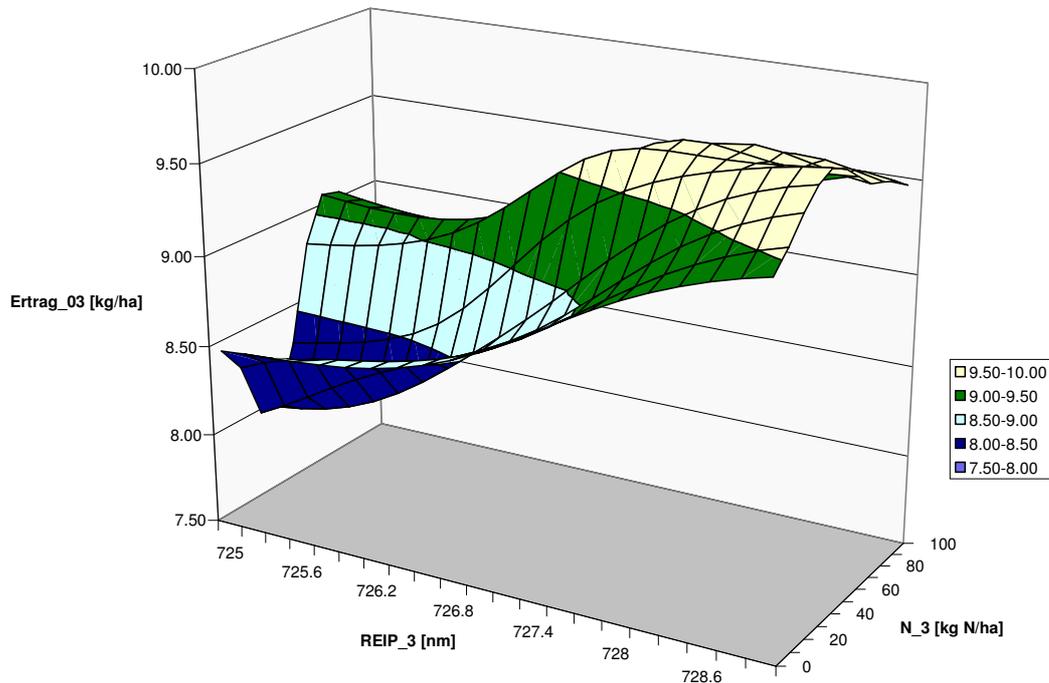


Abbildung 59: Response Surface zur dritten N-Applikation (Versuch Görzig 04)

Auffällig ist der Verlauf bei niedrigen $REIP_3$ -Werten. Hier fällt der prognostizierte Ertrag mit zunehmender N_3 -Applikation zuerst ab, dann steigt er bis zu einem Gipfel bei ca. 60 kg N/ha. Nach dieser N-Stufe fällt der Ertrag wieder ab. Dieser wellenförmige Verlauf macht aus pflanzenbaulicher Sicht keinen Sinn. Die Erklärung dafür kann in der geringen und wechselnden Datendichte in diesem Bereich gesehen werden. Aus Sicht der $REIP_3$ -Werte liegt der Schwerpunkt der Verteilung bei ungefähr 728 nm. Das bedeutet, dass die überwiegende Anzahl der Teilflächen der Trainingsmenge im rechten Teil dieses Diagramms zu finden wäre. Während es relativ viele Teilflächen mit einer Nullapplikation gibt, existieren für den Bereich von 10-30 kg N/ha nur sehr wenige Teilflächen mit niedrigen $REIP_3$ -Ausprägungen. Dieser Verlauf kann somit das Resultat einer Überanpassung sein.

Bei mittleren $REIP$ -Werten wird das Ertragsmaximum bei 50 kg N/ha erreicht. Bei hohen $REIP$ -Werten genügen 40 kg N/ha zur Erreichung des Ertragsmaximums. Der Ertragsabfall bei Applikationen über diesen Düngermengen hinaus verläuft unterschiedlich steil: Während der Abstieg bei niedrig versorgten Teilflächen sehr steil ausfällt, ist bei besser versorgten Teilflächen nur ein mäßiger Abstieg zu beobachten.

Aus ökonomischer Sicht ergeben sich optimale N-Gaben von 30 kg N/ha bei sehr hohen $REIP_3$ -Werten bis hin zu 60 kg N/ha bei sehr niedrigen $REIP_3$ -Werten. Es ist zu beobachten, dass die ökonomischen Verluste bei einer Applikation neben den N_{opt} in Relation

zur zweiten N-Applikation geringer ausfallen. Schlimmstenfalls ist mit einer Differenz zur maximalen NKL von ca. 50 €/ha zu rechnen.

Als nächsten Schritt soll die Reaktion des neuronalen Netzes auf eine Variation der dritten Teilapplikation und dem Attribut *Ertrag_03* untersucht werden (vgl. Abbildung 60). Auffällig ist zunächst ein extrem steiler Abfall des prognostizierten Ertrages, falls bei niedrigen Ertragsbereichen über das Ertragsmaximum hinaus gedüngt wird. In niedrigen Ertragsbereichen reichen 30 kg N/ha zur Erreichung des Ertragsmaximums. In mittleren Ertragsbereichen verschiebt sich das Ertragsmaximum zu einem Gipfel der erst mit einer N-Gabe von ca. 70 kg N/ha erreicht wird. Allerdings ist in diesem Bereich der Einfluss der N-Applikation auf den Ertrag relativ gering. In hohen Vorjahresertragsbereichen wird das Ertragsmaximum bei ca. 50 kg N/ha erreicht.

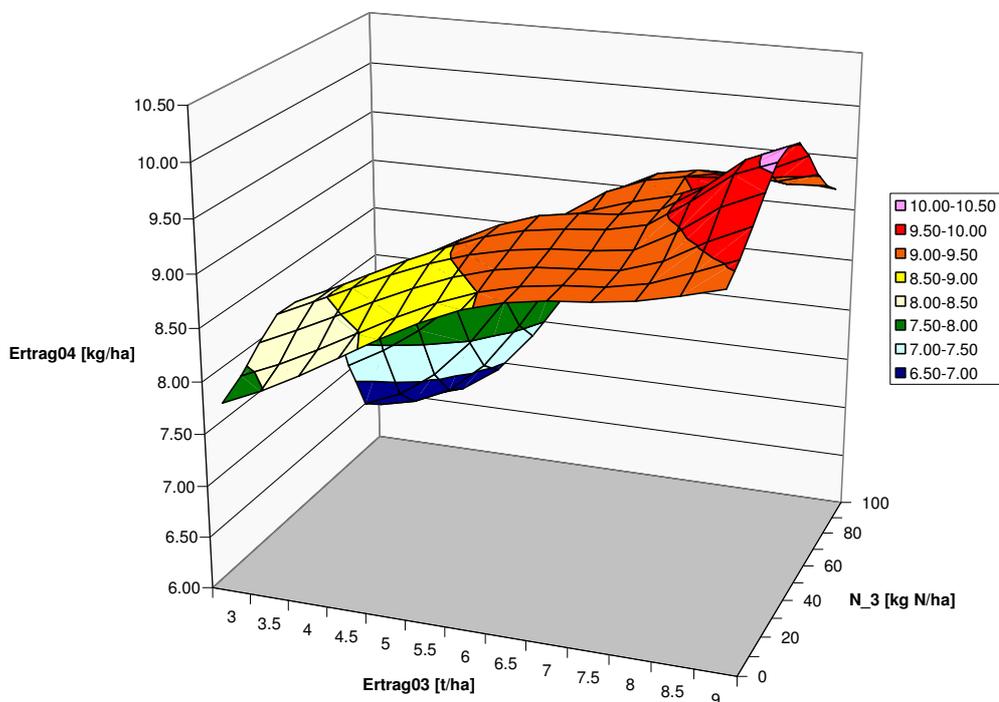


Abbildung 60: Response Surface zur dritten N-Gabe (*Ertrag_03*, Versuch *Görzig 04*)

Aus ökonomischer Sicht liegt die optimale spezielle Intensität in niedrigen Vorjahresbereichen bei 30 kg N/ha und steigt auf 40 kg N/ha für sehr hohe Ertragsbereiche an. Damit werden hier wieder Anzeichen für einen konventionellen Mapping-Ansatz deutlich.

Als letzten Schritt in der Untersuchung des trainierten neuronalen Netzwerkes wird auf das Attribut *EM_38* eingegangen. Die Zusammenhänge des prognostizierten Ertrages und *N_3* bei verschiedenen Leitfähigkeitsbereichen ist in Abbildung 61 dargestellt. Es zeigt sich eine relativ gleichmäßig geformte Oberfläche. Ausgenommen davon ist der Bereich mit hohen *N_3*-Applikationen und niedrigen *EM_38*-Werten. Hier fällt der prognostizierte Ertrag steil ab. Erneut könnte für diese Prognose die schlechte Abdeckung dieses Bereiches in der Trainingsdatenmenge verantwortlich gemacht werden. In niedrigen *EM_38* Bereichen wird

das Ertragsmaximum mit niedrigen N-Gaben erreicht, in Bereichen mit hoher Bodenleitfähigkeit wird dazu eine größere Menge benötigt. Im Übergang befindet sich ein wellenförmiger Bereich.

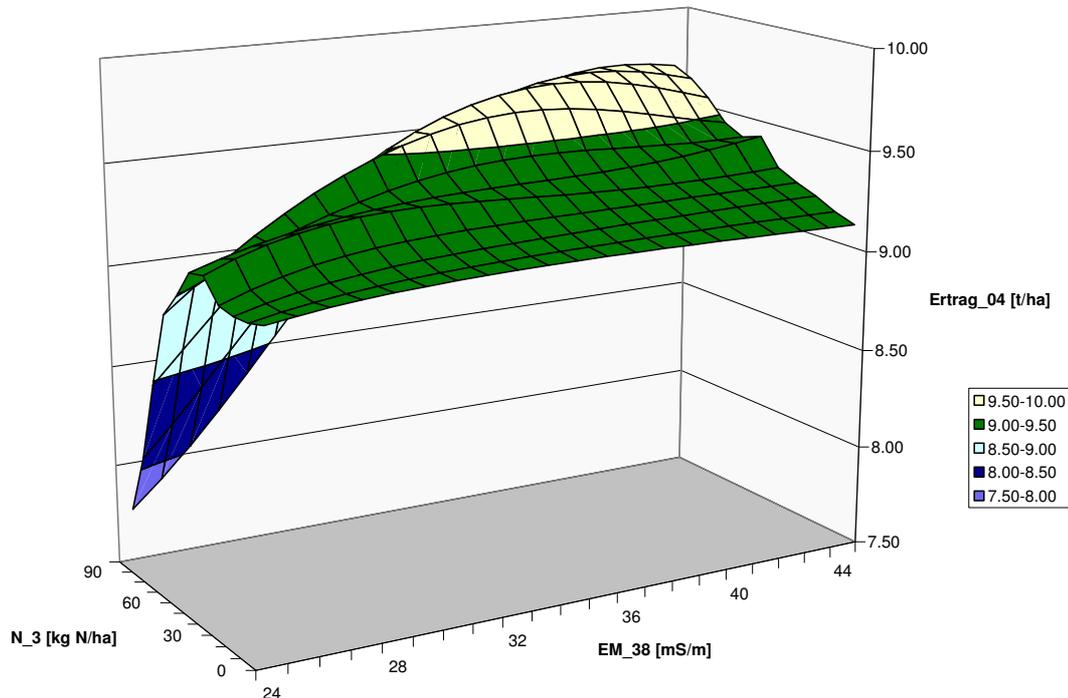


Abbildung 61: *Response Surface* zur dritten N-Gabe (EM_{38} , Versuch Görzig 04)

Aus ökonomischer Sicht liegt die optimale spezielle Intensität in niedrigen EM_{38} -Bereichen bei 40 kg N/ha. Bei mittleren EM_{38} -Bereichen erfolgt ein Sprung hin zu 60-70 kg N/ha. Bei extrem hohen EM_{38} -Bereichen fällt sie allerdings wieder zurück auf 50 kg N/ha. Hier reicht der Anstieg des Ertrages nicht mehr aus, um die Grenzkosten zu decken.

Entscheidungsregeln

Mit diesen Visualisierungen wurden die Zusammenhänge der wichtigsten Attribute im trainierten neuronalen Netz analysiert. Es kann festgehalten werden, dass sich einige Ansätze für eine teilflächenspezifische Applikation aus den Ergebnissen ableiten lassen. Allerdings beschränken sich die Visualisierungen auf einzelne Attribute. In den Entscheidungsbaumalgorithmus fließen dagegen alle Attribute, die im trainierten neuronalen Netzwerk verwendet werden. Genau dieser Vorteil des Entscheidungsbaumverfahrens zeigt sich im vorliegenden Fall.

In Abbildung 62 ist der Entscheidungsbaum für diese Applikation abgebildet. Es zeigt sich, dass die Bestimmung der optimalen speziellen Intensität in erster Linie über das Attribut $REIP_2$, d. h. über die reflexionsoptische Messung zu EC 32 bestimmt wird. Das ist insofern verwunderlich, da dies nicht dem aktuellen Versorgungszustand zu N_3 entspricht. Wiederum müssen dazu die möglichen Verzerrungen innerhalb des Datensatzes herangezogen

gen werden. Da die beiden reflexionsoptischen Messungen untereinander einen starken Zusammenhang aufweisen und zwischen den beiden Teilapplikationen N_2 und N_3 ebenfalls ein starker Zusammenhang besteht, können hier mit dem neuronalen Netz Approximationen erzeugt werden, die speziell für diesen Datensatz passen.

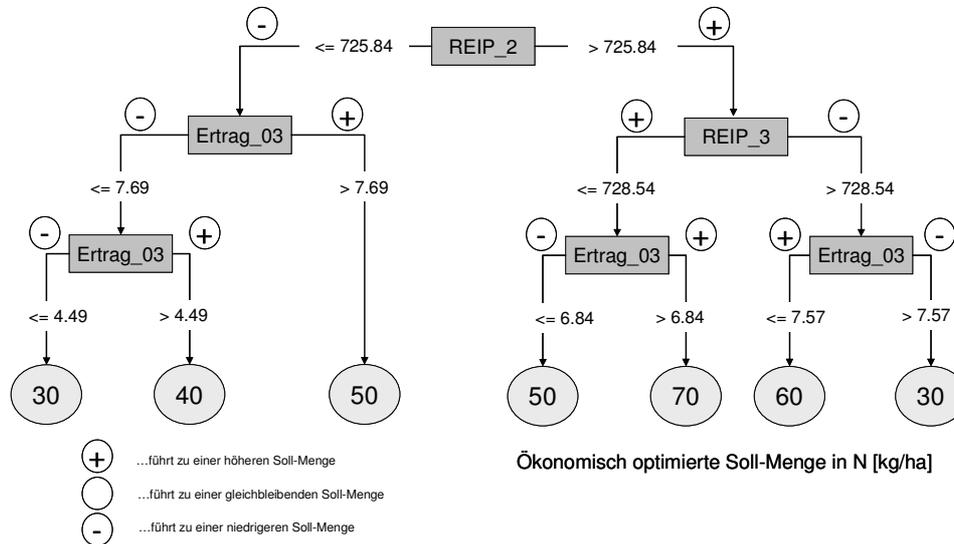


Abbildung 62: Entscheidungsbaum zur dritten N-Applikation (Versuch Görzig 04)

Auffällig ist, dass ein hoher $REIP_2$ -Wert im Durchschnitt zu einer niedrigeren N_3 -Applikation führt. Bei vergleichsweise niedrigen $REIP_2$ -Werten (linker Teil des Baumes) wird zur Entscheidung das Attribut $Ertrag_{03}$ herangezogen. Ein niedriger $Ertrag_{03}$ führt nach diesem Modell zu einer verringerten speziellen Intensität und umgekehrt. Bei hohen $REIP_2$ -Werten (rechter Teil der Baumstruktur) folgt zunächst eine Aufspaltung nach dem aktuellen Vegetationsindex $REIP_3$. Falls dieser Wert unter einer Grenze (728.54 nm) liegt, erhöht sich die durchschnittliche spezielle Intensität und umgekehrt. Dies entspricht dem Zusammenhang aus den *Response Surfaces*. Das Attribut $Ertrag_{03}$ zeigt in der Wirkungsweise nach diesem Split keine konstante Richtung: Bei vergleichsweise gut versorgten Beständen (hoher $REIP_3$) wird N_{opt} mit zunehmenden Vorjahresertrag verringert und vice versa. Bei vergleichsweise schlecht versorgten Beständen tritt genau der umgekehrte Fall ein.

Dies zeigt dass die hohe Anzahl von verwendeten Attributen zu einer entsprechend hohen Komplexität in der Interpretation führt. Insbesondere gilt dies, wenn innerhalb des Datensatzes (versuchsbedingte) Verzerrungen auftauchen.

4.7.5 Zwischenfazit

Mit diesem Versuchsdatensatz kann gezeigt werden, dass der notwendige Datenumfang für eine konsistente Durchführung des Netzwerk-Trainings erreicht wurde. Allerdings können in der Datenexploration einige Verzerrungen beobachtet werden, die die Modellierung und

Interpretation von Teilapplikationen erschweren. Insbesondere sei dazu die starke Korrelation der N-Applikationen untereinander erwähnt. Mit der Feldversuchs-Variante *Merkmalsraum* werden diese Verzerrungen unter Umständen verstärkt, da keine zufällige Auswahl der Teil-Applikationsmengen vorgenommen wurde.

In der Modellierung der neuronalen Netzwerke bilden sich im Allgemeinen relativ komplexe Netzwerke heraus, die ein Überlernen wahrscheinlicher machen. Allerdings werden mit diesen Strukturen bessere Ergebnisse in der Validierung erreicht.

Für alle Teilapplikationen gilt, dass auch bei unterschiedlichen Datenaufbereitungen dieselben (Grob-)Muster zu beobachten sind. Falls detaillierte Ergebnisse, wie beispielsweise einzelne Äste eines Entscheidungsbaumes oder einzelne Produktionsfunktionen innerhalb eines Response Surfaces betrachtet werden, können nur noch Ähnlichkeiten festgestellt werden. Die exakten Ergebnisse unterscheiden sich unter Umständen beträchtlich.

Zur ersten N-Applikation werden bessere Resultate in der Validierung erreicht, wenn die Instanzen bezüglich der Stickstoff-Gaben ausbalanciert werden. In der Analyse zeigt sich ansatzweise die Optimierung nach einem inversen Mapping-Ansatz (vgl. Kap. 4.6.3). Bezüglich des Attributes *EM_38* konnten keine Ansätze für eine teilflächenspezifische Optimierung beobachtet werden. Allerdings deutet sich an, dass, falls die Applikation nicht bei N_{opt} erfolgt, die prognostizierte NKL stark einbricht. Für diese Teilapplikation können damit insbesondere Erkenntnisse zur schlageinheitlichen Intensitätsanpassung gewonnen werden. Die durchschnittliche ex post optimierte Intensität liegt ca. 15 kg N/ha niedriger als die betriebsübliche Andüngung.

Zur zweiten N-Applikation stellt die aktuelle N-Versorgung (*REIP_2*) das entscheidende Attribut zur Bemessung der N-Applikation dar. Es kann deutlich beobachtet werden, dass die optimale N-Applikation bei niedrig versorgten Beständen höher liegt als bei gut versorgten Beständen. Gleichzeitig liefert dieses Attribut sehr enge Bereiche, innerhalb derer die N-Applikation liegen muss, ohne dass mit hohen Einbußen hinsichtlich der NKL zu rechnen ist. Damit können wiederum Anhaltspunkte für die Intensitätsanpassung festgestellt werden. Die durchschnittliche ex post optimierte Intensität liegt ca. 5 kg N/ha höher als die betriebsübliche zweite N-Applikation – allerdings können große Differenzen je nach teilflächenspezifischen Verhältnissen festgestellt werden. Ein nicht zu vernachlässigender Einfluss auf die Bemessung der N-Gabe geht vom Vorjahresertrag aus. Hier zeigt sich wiederum das Muster eines inversen Mapping-Ansatzes. Bezüglich des Attributes *EM_38* kann das Muster eines klassischen Mapping-Ansatzes beobachtet werden, wenn die Überanpassung an die Trainingsdaten nicht beachtet wird.

Der Einfluss der dritten N-Applikation auf den Ertrag ist vergleichsweise schwach ausgeprägt. Entsprechend niedrig sind die Einbußen hinsichtlich der NKL, wenn das ökonomische Optimum nicht getroffen wird. Bezüglich des Attributes *REIP_3* zeigt sich dasselbe

Muster wie zur zweiten N-Applikation mit dem entsprechenden Attribut *REIP_2*. Das N_{opt} liegt bei schlecht versorgten Teilflächen höher als bei gut versorgten Teilflächen. Auffällig ist weiterhin, dass eine höhere NKL erreicht werden kann, wenn Teilflächen mit einem hohen Vorjahresertrag auch eine vergleichsweise hohe N-Menge appliziert wird. Damit kehrt sich die Richtung der Intensitätsanpassung um – hin zu einem klassischen Mapping-Ansatz.

4.8 Verifikation und Erkenntnisgewinn des WED-Prozesses

In diesem Kapitel gilt es nun, die Ergebnisse des WED-Prozesses zusammenzustellen. Zunächst sollen dazu die Resultate der Verifikation, d. h. der Feldversuche dargestellt werden. Daraufhin wird entsprechend dem Ziel dieser Arbeit versucht, gemeinsame Muster aus der Analyse der einzelnen Versuche abzuleiten.

4.8.1 Ergebnisse der Feldversuche

Die trainierten neuronalen Netze zur Ertragsprognose sollen im Sinne einer Verifikation innerhalb des WED-Prozess jeweils in den darauf folgenden Jahren getestet werden (vgl. Kap. 4.1). Konkret wurde die vorgestellte Methode im Rahmen des IKB-Projektes in den Feldversuchen *D4 03* und *Schafhof 04* getestet. Die neuronalen Netzwerke des Versuches *Görzig 04* konnten im Versuch *Görzig 05* getestet werden.

IKB-Versuche

Für den Versuch *D4 03* wurden dazu Entscheidungsregeln aus einer Kombination der Instanzen des Versuches *Schafhof 02* und eines N-Steigerungsversuches erzeugt. Wie bereits in Kap. 4 dargestellt wurde, erwies sich dieser Ansatz als nicht zielführend. Insbesondere konnte kein neuronales Netzwerk zur ersten N-Applikation erstellt werden. Die Kombination der beiden Datensätze „krankte“ an der schlechten Kompatibilität.

Zur Verifikation im Versuch *Schafhof 04* wurden die Instanzen aus dem Versuch *Schafhof 02* und *D4 03* verschnitten. Auch hier zeigten sich Probleme, die in WEIGERT UND WAGNER (2004, s.p.) ausführlich beschrieben sind. Darüber hinaus existierten versuchstechnische Probleme (lokale N-Überhänge aus dem Vorjahr, vgl. Kap. 4.2.1), die eine Interpretation der Ergebnisse behindern.

Die ökonomischen Ergebnisse dieser Feldversuche sind ausführlich von GANDORFER (2005) beschrieben. Zusammenfassend konnten keine signifikanten ökonomischen Vorteile gegenüber den Vergleichsvarianten aufgezeigt werden. Im Sinne einer Verifikation des hier vorgestellten Verfahrens eignen sich diese Feldversuche aus den eben vorgestellten Gründen nur bedingt und sind daher als Zwischenevaluierung zu betrachten.

Versuch *Görzig 05*

Die Entscheidungsregeln aus dem Versuch *Görzig 04* konnten im darauf folgenden Jahr im Versuch *Görzig 05* getestet werden (vgl. WAGNER ET AL., 2005).

Die Variante *Netz* wurde entsprechend der in dieser Arbeit vorgestellten Methode optimiert. Dazu wurden die trainierten neuronalen Netze aus Kap. 4.7.2, 4.7.3 und 4.7.4 verwendet. Zur Optimierung konnte in der Anwendung auf dieselben Attribute zurückgegriffen werden, die auch im Training verwendet wurden (*EM_38*, *REIP_2*, *REIP_3*, *Ertrag_03*). Das Trainings-Attribut *Ertrag_03* (Körnermais) des Versuches am Schlag 550 wurde entsprechend ersetzt durch den *Ertrag_03* auf dem Schlag 350 (Braugerste)⁷¹.

Das *Data Audit* der Instanzen des Versuches *Görzig 05* ist in Abbildung 63 abgetragen. Es ergaben sich 1692 vollständige Instanzen, die mit derselben Methode vorverarbeitet wurden, wie im Versuch *Görzig 04* in dieser Arbeit (vgl. Kap. 4.3.2). In diesem Versuch wurde ein 20 x 20 m-Raster verwendet. Jede Teilfläche ist somit 400 m² groß.

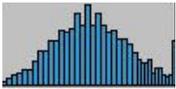
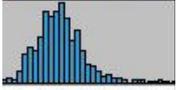
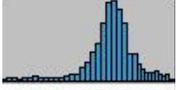
Attribut	Histogramm	Minimum	Maximum	Mittelwert	Std. Abw.	Anzahl
<i>ERTRAG_03</i>		4.29	9.67	6.94	1.14	1692
<i>EM_38</i>		23.32	34.46	26.68	1.33	1692
<i>REIP_2</i>		721.7	728.21	725.49	0.83	1692
<i>REIP_3</i>		720.45	727.3	725.74	0.84	1692

Abbildung 63: Data Audit der Instanzen des Versuches *Görzig 05*

In Relation zum Versuch *Görzig 04* ergeben sich dadurch vergleichbare statistische Kennzahlen. Insgesamt erscheint die Heterogenität dieses Schlages schwächer ausgeprägt zu sein als im Schlag 550. Das Attribut *Ertrag_03* zeigt einen leicht höheren Mittelwert. Die Standardabweichung dagegen liegt niedriger. Das Attribut *EM_38* ist im Durchschnitt in diesem Schlagausschnitt niedriger als in dem vergleichbaren Versuch *Görzig 03* (26.68 zu 33.68). Die Variabilität bezüglich dieser beiden Bodenattribute ist sehr niedrig einzuschätzen.

Die erste reflexionsoptische Messung zu *EC 32* zeigte sehr ähnliche Ergebnisse wie im Versuch des vorausgegangenen Jahres. Die Kennzahlen zeigen eine große Übereinstimmung. Dagegen liegt das Niveau der zweiten reflexionsoptischen Messung auffällig niedrig

⁷¹ Dadurch, dass die Attribute in absoluten Größen in das Netz Eingang finden, ergibt sich in diesem Versuch ein Problem in der Umrechnung: Es stellt sich die Frage, wie eine Ertragseinheit Braugerste in eine Ertragseinheit Körnermais umzurechnen ist. Dies ist insofern elementar, als dass das neuronale Netz auf die Ertragseinheit Körnermais „geeicht“ ist. Nach Absprache mit der Firma AgriCon© wurden dazu die Braugerstenerträge um 6% nach unten korrigiert.

(728,17 zu 725,74). In diesen Werten zeigen sich allerdings bereits die Auswirkungen der Teilapplikationen der einzelnen Systeme.

In der Anlage des Versuches wurde darauf geachtet, dass die einzelnen Varianten bezüglich des Mittelwertes der Attribute *Ertrag_03* und *EM_38* nur sehr leicht differieren (vgl. Kap. 4.2.2). Dadurch konnten vergleichbare Voraussetzungen für die betrachteten Systeme erreicht werden. Die Auswertung dieses Versuches kann somit über einen Mittelwertsvergleich der Instanzen erreicht werden, ohne dass die Kovariaten *EM_38* und *Ertrag_03* berücksichtigt werden müssen. Die Ergebnisse der einzelnen Varianten sind in Tabelle 27 zusammengefasst.

Tabelle 27: Ergebnisse des Versuches Görzig 05

	Betrieb	Sensor	Netz
Stickstoff (kg/ha)	175	182	148
Ertrag (dt/ha)	7.04	7.34	7.68
NKL	731	763	825.4
Proteingehalt (in TS, %)	14,34	14,88	12,97
NKL _{Prot, Akt} * (€/ha)	606	633	647

*NKL-Berechnung unter Bewertung des Proteingehaltes und unter Verwendung von aktuellen Preisen

Die schlageinheitliche Variante *Betrieb* erzielte einen Ertrag von 7.04 t/ha unter Einsatz von 175 kg N/ha. Die teilflächenspezifische Variante *Sensor* konnte einen höheren Ertrag erreichen, wobei dazu geringfügig mehr Stickstoff eingesetzt wurde. Die Variante *Netz* erzielte den höchsten Ertrag (7.68 t/ha), wobei beträchtlich weniger Stickstoff gedüngt wurde. Bezogen auf die Zielgröße NKL mit den Faktorpreisen, die in dieser Arbeit verwendet und zur Optimierung der Variante herangezogen wurden, ergibt sich ein Vorteil der Variante *Netz* von knapp 100 € pro Hektar gegenüber der schlageinheitlichen Variante. Im Vergleich zur Variante *Sensor* sind dies ca. 60 € pro Hektar. Werden aktuelle Faktorpreise verwendet und gleichzeitig der Proteingehalt bewertet, der im Versuch stichprobenartig erfasst wurde, verringert sich der ökonomische Vorteil auf 40 bzw. 15 € pro Hektar (Auszahlungspreise für die Region Halle/Saale, August 2005).

Bezüglich beider Bewertungsmaßstäbe zeigen sich damit beträchtliche ökonomische Vorteile der Variante *Netz*, die in dieser Höhe nur selten beobachtet werden konnten (vgl. Kap. 3.1.2). Dieses Ergebnis erscheint vor allem dadurch überraschend hoch als dass zur Anwendung dieser Methode nur ein Trainingsjahr verwendet wurde und dass der Schlag als homogen klassifiziert werden kann. Nochmalig sei erwähnt, dass die Applikationsmengen in dieser Variante autonom unter dem Einsatz der trainierten neuronalen Netze bestimmt

wurden und dass keine manuellen Eingriffe hinsichtlich einer Anpassung der N-Mengen erfolgten.⁷²

In den Schlussfolgerungen muss beachtet werden, dass der Versuch nur eine Einzelbeobachtung darstellt. Der Versuch muss mehrfach wiederholt werden, um zufällige Einflüsse auf das Ergebnis auszuschließen. Im Sinne einer Verifikation liefert dieser Versuch damit lediglich ein (wenngleich starkes) Indiz für die ökonomische Vorteilhaftigkeit des in dieser Arbeit vorgestellten Systems zur Optimierung - unter den dargestellten Rahmenbedingungen.

Die in Abschnitt 4.4.1 aufgestellte Hypothese, dass die trainierten neuronalen Netze über die reflexionsoptischen Kennzahlen kalibriert werden und damit jahresunabhängig eingesetzt werden können, scheint sich zu bestätigen. Wie bereits dargestellt, wurde zum Training der neuronalen Netze nur ein Versuchsjahr verwendet. Obwohl die Witterungsverhältnisse des Trainingsjahres und des Anwendungsjahres voneinander (leicht) abweichen, liefern die neuronalen Netze sehr gute Ergebnisse in der Optimierung. Die Konsequenzen für einen möglichen Praxiseinsatz wären beträchtlich: Es wäre mit minimalem Aufwand (vgl. Versuchsanlage *Görzig 04*) möglich, Entscheidungsregeln für einen spezifischen Standort zu erzeugen.

Es ist davon auszugehen, dass für das gute Abschneiden der neuronalen Netzwerke mitunter die Ähnlichkeit der Merkmalsräume im Training und in der Anwendung verantwortlich ist (vgl. Abbildung 48 und Abbildung 63). Alle Attributsausprägungen der Anwendungsbeispiele liegen durchgängig innerhalb des Bereiches, der von den Trainingsbeispielen aufgespannt wird. Eine Extrapolation wie beispielsweise im Versuch *Schafhof 02* (vgl. Kap. 4.5.2) konnte vermieden werden. Wie in diesem Versuch gezeigt wurde, haben neuronale Netze praktisch keine Fähigkeit zur Extrapolation. Dies kann auch als Hinweis für die Übertragbarkeit auf andere Standorte verstanden werden: Die Einsatzbedingungen trainierter neuronaler Netze sollten demnach sehr ähnlich sein zu den Trainingsbedingungen. Um dies beurteilen zu können, sollten die Instanzen mit Hilfe eines *Data Audits* untersucht werden.

Weiterhin liefert dieser Versuch ein Indiz für die These, dass eine ökonomische Optimierung gleichzeitig zu positiven ökologischen Effekten führt (vgl. Kap. 3.1.2). In diesem Fall konnte der Einsatz von Stickstoff um ca. 30 kg N/ha im Vergleich zu den anderen Systemen verringert werden und damit auch das Risiko für N-Auswaschung.

Über die Gründe für die Überlegenheit dieser Variante in diesem Versuch kann jedoch nur spekuliert werden. Es können keine Einzelteile der immanenten Entscheidungsregeln he-

⁷² Um diesen Punkt zu verdeutlichen: Die Applikationsmengen wurden „am Computer“ erstellt, ohne den Schlag oder Bestand vorher zu begutachten. Es wurden lediglich die Daten der Teilflächen in die jeweiligen trainierten neuronalen Netze eingegeben und die N-Mengen nach dem vorgegebenen Zielsystem ausgewählt (vgl. Kap. 4.4.2).

rausgegriffen werden, denen die ökonomischen Effekte zugeschrieben werden können. Dazu sind nähere Untersuchungen bzw. alternative Versuchsanlagen nötig⁷³.

4.8.2 Extrahierte Muster

Dadurch, dass mit dieser Arbeit in einigen Punkten wissenschaftliches Neuland betreten wurde und dementsprechend heterogene Durchläufe des WED-Prozesses produziert wurden, gestaltet sich die Ableitung von Mustern bzw. Erkenntnissen schwierig.

Erkenntnisse aus einem WED-Prozess lassen sich allgemein in den Kategorien *Neuigkeit*, *Verständlichkeit* und *Nützlichkeit* beurteilen (vgl. Kap. 2.7.) Für diese Arbeit ergeben sich sehr differenzierte Beurteilungen der Ergebnisse nach diesen Kategorien.

Zum einen können die trainierten neuronalen Netze als Muster angesehen werden. Wie sich mit dem Versuch *Görzig 05* andeutet, können mit Hilfe der trainierten Netze – selbst wenn dazu nur ein Versuchsjahr zum Training angewendet wird – überraschend hohe ökonomische Effekte realisiert werden. Allerdings ist es mit der vorgestellten Versuchsanlage nicht möglich, die Ursachen für diese Effekte zu analysieren. Es kann also nicht ergründet werden, welche der (Teil-)Muster für den ökonomischen Erfolg ausschlaggebend sind. Der vorgestellte Ansatz zur Verifikation bezieht sich auf den gesamten Optimierungsvorgang. In der Verifikation deutet sich streng genommen lediglich an, dass mit der Anwendung der drei vorgestellten neuronalen Netzwerke stark positive ökonomische Effekte erzielen lassen. Bezüglich der Kategorien *Neuigkeit* und *Nützlichkeit* entspricht dies sehr guten Ergebnissen. Im Sinne der *Verständlichkeit* bzw. der Gewinnung von beschreibbaren Erkenntnissen ist dieses (Teil-) Ergebnis allerdings nicht zielführend. Die positiven Effekte lassen sich nicht einzelnen verständlichen Mustern zuweisen.

Unter Zuhilfenahme von Entscheidungsbaumalgorithmen können verständliche Regeln abgeleitet werden. Allerdings ergibt sich ein *Trade off* zwischen der *Verständlichkeit* und der *Nützlichkeit* einer Entscheidungsregel. Wenn neuronale Netzwerke exakt in Regeln überführt werden, ergeben sich unüberschaubar viele Regeln, von denen eine einzelne Regel eine vernachlässigbar große Bedeutung in einer möglichen Anwendung hat. Wenn beispielsweise aus einem *Response Surface* abgelesen werden kann, dass eine Erhöhung der Düngung um 10 kg N/ha bei Teilflächen mit einem *EM_38*-Wert von 44 mS/m zu einem Ertragszuwachs von 0.1 t/ha führt, ist dieses „Muster“ nur nützlich, wenn es in der Anwendung auch Teilflächen mit diesem *EM_38*-Wert gibt. Eine *Nützlichkeit* einer einzelnen Regel kann somit schwer eingeschätzt werden, da sie stark vom Anwendungsfall abhängt.

Gleichzeitig führt eine hohe Anzahl von detaillierten Regeln, wenn sie als Ganzes betrach-

⁷³ Beispielsweise könnte der inverse Mapping-Ansatz zur ersten N-Gabe isoliert angewendet werden und mit einer betriebsüblichen Düngung verglichen werden. Daraus kann abgeleitet werden, welchen ökonomischen (Teil-)Beitrag dieser Ansatz liefert.

tet werden, zu einer Unübersichtlichkeit und damit zu einer unzureichenden *Verständlichkeit*. Selbst die in der Arbeit dargestellten und auf drei Ebenen beschränkten Entscheidungsbäume mit numerischen Splits sind schwer verständlich, da keine Vorstellung mit einer konkreten Ausprägung eines Attributes (z. B. falls $EM_{38} > 35.4$) verbunden ist⁷⁴.

Es kommt also darauf an, ein geeignetes Maß zur Darstellung von Erkenntnissen zu finden. Innerhalb der einzelnen Data-Mining-Phasen wurde diese Maß relativ fein gewählt. Die Muster (bzw. Regeln) gelten streng genommen nur unter den gezeigten Voraussetzungen und sind dementsprechend in den Zwischenfazits festgehalten.

In diesem Abschnitt wird dagegen ein gröberes Raster an die Muster angelegt. Es gilt, die Muster, die sich in den einzelnen Data-Mining-Phasen gezeigt haben, auf den kleinsten gemeinsamen Nenner zu bringen. Diese Muster müssen als Indizien verstanden werden, die innerhalb des vorgestellten WED-Prozesses nicht verifiziert werden können und deren Nützlichkeit nicht gut abgeschätzt werden kann. Die Auswahl kann damit nicht nach objektiven Gesichtspunkten erfolgen. Umso wichtiger ist deshalb die Interpretation vor dem Hintergrund des derzeitigen fachspezifischen Stands des Wissens (vgl. Kap. 3 bzw. SCHOMMER, 2003, S. 8).

Muster 1: Bedeutung der räumlichen Informationen

Es deutet sich an, dass über die **Kombination** mehrerer **teilflächenspezifischer Attribute** ein entscheidender Vorteil zur ökonomischen Optimierung erreicht werden kann. Zur Optimierung werden in der Regel mehrere Informationen über eine Teilfläche verwendet. In dieser Arbeit stellte sich kein Attribut heraus, dass per se unwichtig für die Bemessung der N-Gabe ist. Es existieren allerdings bezüglich des Standortes und der Teilapplikation große Unterschiede bezüglich der Bedeutung der jeweiligen Attribute. Wie auch von anderen Autoren festgestellt, ist der Zusammenhang der Bodenattribute (EM_{38} und Ertragskarten) mit dem teilflächenspezifischen Ertrag von standortspezifischen Einflüssen geprägt (vgl. KHAKURAL ET AL. 1999, 577ff.; KRAVCHENKO UND BULLOCK, 2000, S. 78ff; SPAULDING ET AL. 2005, s.p.).

Muster 2: Optimierung der Teilapplikationen

Als weiteres Muster deutet sich an, dass die **Optimierung der Teilapplikationen** auf Basis der maximal verfügbaren Informationen einen entscheidenden ökonomischen Vorteil bringt. Dies leitet sich aus folgenden Indizien ab:

- Die Ertragswirkung und die damit verbundenen ökonomischen Effekte der Stickstoff-Applikation sind in den einzelnen Teilgaben nicht konstant. Dies ist ersichtlich

⁷⁴ Ein Ausweg, der in einigen WED-Anwendungen verwendet wird, ist die Unterteilung der numerischen Trainingsattribute in kategorische Werte, also beispielsweise EM_{38} ist „hoch“ oder „niedrig“. Dadurch ist allerdings Vorwissen zur Einteilung nötig und die Präzision wird in der Regel stark eingeschränkt.

aus den Korrelationsmatrizen, aus der Sensitivitätsanalyse der neuronalen Netzwerke und aus den produzierten Entscheidungsbäumen. Die beiden ersten N-Gaben haben in der Regel eine vergleichsweise hohe Bedeutung für den Ertrag, die dritte N-Applikation besitzt demgegenüber lediglich eine geringe Bedeutung.

- In den einzelnen Teilapplikationen ändert sich die Bedeutung teilflächenspezifischer Attribute zur Bemessung der N-Gabe. Gleichzeitig ändern sich in einigen Fällen die Wirkungszusammenhänge, wie dies beispielsweise bei historischen Ertragsdaten der Fall ist.
- Die aus der Ex-post-Optimierung erzeugten durchschnittlichen Stickstoffmengen der Teilapplikationen unterscheiden sich stark.

Muster 3: Ertragskarten

Ertragskarten spielen für die Optimierung der N-Gabe eine große Rolle, allerdings wirken sich die daraus abgeleiteten Attribute nicht konsistent auf die Bemessung der N-Gabe aus. Ein stabiles Muster (*D4 03* und *Görzig 04*) kann im Ergebnis als **inverser Mapping-Ansatz zur ersten N-Applikation** umschrieben werden. Nach diesem Muster sollte in Teilschlägen, auf denen ein vergleichsweise niedriger Ertrag produziert wurde, die erste N-Applikation erhöht werden und umgekehrt.

Es existieren Arbeiten, die ebenfalls auf dieses Muster hinweisen (vgl. DIEPOLDER, 1994, S. S. 159, HEGE UND OFFENBERGER, 2005, S. 35). Aus N-Steigerungsversuchen leiten HEGE UND OFFENBERGER (2005) ab, dass sich für gute Standorte (im Kontext dieser Arbeit: vergleichsweise hohe historische Erträge) eine vergleichsweise niedrige Andüngung empfiehlt. Für ungünstige Standorte (im Kontext dieser Arbeit: vergleichsweise niedrige historische Erträge) empfehlen diese Autoren eine relativ hohe erste N-Applikation. DIEPOLDER konnte beobachten, dass eine Erhöhung der Andüngung auf leichteren sandigen Böden zu deutlich positiven Ertragseffekten führt. Die Kategorie „leichtere und sandige Böden“ deutet wiederum auf Böden mit historisch niedrigen Erträgen hin.

Weitere Indizien dafür lassen sich auch bei GODWIN (2003, S. 535) erkennen. In dem beschriebenen Versuch führt ein eingeschränkter inverser Mapping Ansatz zu den besten Ergebnissen. Diese Ergebnisse werden erzielt, wenn nur die Niedrigertragsbereiche besonders hoch gedüngt werden und in den restlichen Ertragszonen Standard-Applikationen verwendet werden.

Die Inversion beschränkt sich größtenteils auf die erste N-Gabe. Zur zweiten N-Applikation zielen die Regeln noch mehrheitlich in diese Richtung. Spätestens zur dritten N-Gabe kehrt sich der Effekt auf N_{opt} um. Zu dieser Applikation wird überwiegend in Hohertragsbereichen eine höhere N-Gabe appliziert. Allerdings richtet sich die Bemessung der N-Gabe zur zweiten und dritten N-Applikation überwiegend nach reflexionsoptischen Kennzahlen.

Es finden sich in dieser Arbeit lediglich zur dritten N-Applikation Indizien, die auf eine Vorteilhaftigkeit der Methode „Düngung nach Entzug“ hindeuten. Diese N-Applikation besitzt allerdings nach den hier vorgestellten Ergebnissen (Muster 1) die geringste Bedeutung für den Ertrag.

Muster 4: Reflexionsoptische Messungen (REIP)

In allen Versuchen dieser Arbeit lassen sich zur zweiten und dritten N-Applikation ein Ansatz zur teilflächenspezifischen N-Ausbringung beobachten, der auf reflexionsoptischen Messungen und den daraus bestimmten Vegetationsindices (REIP) basiert. Dieser Index liefert *entscheidende* Anhaltspunkte für die Bemessung einer ökonomisch optimierten N-Gabe in einem konstanten Muster: In Teilflächen mit schlechter N-Versorgung (vergleichsweise niedriger REIP-Wert) erhöht sich die ökonomisch optimale N-Menge und umgekehrt. Es ergeben sich dadurch große Ähnlichkeiten zu der Sensorsteuerung, die beispielsweise mit dem Yara® N-Sensor® verfolgt wird (vgl. LINK, 2002, s.p.). Insbesondere zur zweiten N-Applikation eröffnen sich mit diesem Attribut große Chancen, die maximale stickstoffkostenfreie Leistung zu erreichen. Allerdings zeigt es sich, dass dieses Attribut in den generierten Entscheidungsregeln im Allgemeinen nicht isoliert verwendet wird. Zur Optimierung werden häufig Bodenattribute hinzugezogen.

Muster 5: Bodenleitfähigkeit (EM_38)

Bezüglich des Attributes *EM_38* kann kein starkes Muster ausgemacht werden. Der Einfluss auf die Bemessung der Düngergabe im Versuch *D4 03* ist vernachlässigbar. Im Versuch *Görzig 04* steht dieses Attribut in engem Zusammenhang mit den Ertragskarten. Bezüglich der ersten N-Applikation zeigen sich überwiegend inverse Einflüsse zur Bemessung der N-Applikation. Das bedeutet, ähnlich wie bei den Ertragskarten, führen niedrige *EM_38*-Werte zu einer Erhöhung der optimierten ersten N-Applikation und umgekehrt. In den folgenden Applikationen kehrt sich dieser Einfluss tendenziell um. In hohen *EM_38*-Bereichen sollte demnach die N-Applikation erhöht werden und umgekehrt. Damit zeigen sich auch in dieser Arbeit die von mehreren Autoren festgestellten standortspezifischen Zusammenhänge dieses Attributes mit dem teilflächenspezifischen Ertrag.

Muster 6: Zugkraft

Für dieses Attribut kann ebenfalls kein starkes bzw. verständliches Muster beobachtet werden. Zum einen liegt dies daran, dass im Versuch *Görzig 03* keine Zugkraft-Daten verwendet wurden. Zum anderen liegt das an dem undurchsichtigen Einfluss auf die Bemessung der Düngergabe, die sich im Versuch *D4 03* gezeigt hat. Es fehlt für dieses Attribut das nötige Hintergrundwissen, um dieses Verhalten interpretieren zu können. Das liegt im Wesentlichen daran, dass sich bisher sehr wenige Autoren mit diesen Sensordaten beschäftigt haben (vgl. Kap. 3.1.3). Dennoch deutet sich an, dass dieses Attribut einen nicht geringen Einfluss auf die Bemessung der teilflächenspezifischen N-Applikation, vor allem zur ersten

und zweiten Teilapplikation, besitzt.

Muster 7: Steile Produktionsfunktionen

Auffällig ist ein Muster, das allerdings größtenteils nur für den Versuch *Görzig 03* zu beobachten ist: Es zeigen sich teilweise extrem steil verlaufende Produktionsfunktionen für die Teilapplikationen. Die prognostizierten Erträge steigen in einigen Fällen mit zunehmender N-Applikation extrem stark an. Gleichzeitig ist ein starker Ertragseinbruch zu beobachten, falls über das Ertragsmaximum hinaus die N-Applikation erhöht wird. Dies ist insbesondere deshalb interessant, als dass im Versuchsjahr kein Lagergetreide zu beobachten war. Dieser Ertragsabfall resultiert nicht wie bei der Produktionstheorie aus der Anwendung quadratischer Produktionsfunktion, sondern zeigt sich hier durch parameterfreie, selbstanpassende Modelle des überwachten Lernens. Nach diesem Muster ergeben sich hohe Verluste hinsichtlich der stickstoffkostenfreien Leistung, falls die Düngung nicht am ökonomischen Optimum erfolgt (vgl. im Gegensatz dazu PANELL, 2004, s.p.). Es deutet sich damit an, dass hohe ökonomische Effekte durch eine (nicht unbedingt teilflächenspezifische) Intensitätsanpassung erreicht werden können.

Es zeigt sich damit, dass der Anwendung der Methode WED interessante Erkenntnisse (im Sinne einer hohen *Neuigkeit*) aus empirischen Daten abgeleitet werden können. Gleichzeitig können Muster entdeckt werden, die zwar aus pflanzenbaulicher Sicht keine Neuigkeit darstellen, aber bestätigen, dass mit dieser Methode valide Erkenntnisse aufgedeckt werden können. Beispielsweise gilt dies für den Einfluss der Teilapplikationen auf den Kornertrag oder die Sensorsteuerung.

Mit diesem Abschnitt schließt sich die Beschreibung des Prozesses der Wissensentdeckung in Datenbanken, der anhand empirischer Daten durchgeführt wurde. Es wurden alle Phasen durchlaufen und detailliert beschrieben. Im folgenden Schlusskapitel sollen der gewählte Ansatz sowie die einzelnen Phasen zusammenfassend betrachtet werden und daraus Schlussfolgerungen gezogen werden.

5 Diskussion und Schlussfolgerungen

Die vorliegende Arbeit stellt den ersten Versuch für den Einsatz der Methode Wissensentdeckung in Datenbanken (WED) im Bereich Precision Farming dar. Deshalb musste in einigen Bereichen, insbesondere in der Versuchsanlage, in der Datenvorverarbeitung und in der Durchführung der Data-Mining-Phase „Pionierarbeit“ geleistet werden.

Im folgenden Abschnitt sollen die in dieser Arbeit aufgebauten Konzepte zur Entwicklung von Entscheidungsregeln zur kleinräumigen Stickstoffausbringung und zur Extraktion von Mustern in allen Bereichen kritisch reflektiert werden. Es erfolgt weiterhin eine Beurteilung der Methode WED im Bereich Precision Farming sowie Ideen zur weiteren Verbesserung des aufgestellten Konzeptes. Schließlich wird noch bestehender Forschungsbedarf für die Anwendung dieser Methode im agrarwissenschaftlichen Bereich identifiziert.

5.1 Zusammenfassung und Interpretation der Ergebnisse

Aufgrund der heterogenen Rahmenbedingungen der einzelnen Versuche ergibt sich eine gewisse Schwierigkeit, die Ergebnisse dieser Arbeit zusammenzufassen und vergleichend zu interpretieren. Das verwendete Datenmaterial der verschiedenen Feldversuche repräsentiert eine große Breite an möglichen Versuchen für eine Data-Mining-Auswertung.

Konzept Wissensentdeckung in Datenbanken

Die Konzeption der Data-Mining-Phase wurde entsprechend den Ergebnissen aus der Literaturanalyse und den Anforderungen für diese Arbeit aufgebaut. Aus den Ergebnissen mehrerer Autoren (vgl. Kap. 3.2) konnte abgeleitet werden, dass mit neuronalen Netzwerken im Vergleich zu alternativen statistischen Methoden (z. B. multivariate Regression) hinsichtlich einer Ertragsprognose die besten Ergebnisse zu erwarten sind. Weiterhin werden die Aufspaltung der Stickstoff-Düngung in drei Teil-Applikationen und der zunehmende Informationsstatus in späteren N-Applikationen (v. a. durch reflexionsoptische Messungen) berücksichtigt. Daher werden für jede der drei Teilapplikationen neuronale Netze mit dem jeweiligen Informationsstatus trainiert.

Zur Entwicklung von ökonomischen Entscheidungsregeln werden die neuronalen Netze zur Ertragsprognose verwendet und über ein ökonomisches Zielsystem die optimale spezielle Intensität für eine Teilapplikation ermittelt. Aus den ex post optimierten Teilflächen können unter Anwendung von Entscheidungsbaumverfahren Regeln generiert werden.

Es stellte sich heraus, dass mit dieser Methode zwar Entscheidungsregeln generiert werden können, dass aber die detaillierten Regeln sehr umfangreich werden und daher die Verständlichkeit leidet. Zur Extraktion von Mustern werden daher zwei Verfahren angewendet: Zum einen wurden die Entscheidungsregeln in ihrer Komplexität beschränkt und zum ande-

ren wurden zur Analyse der neuronalen Netzwerke *Response Surfaces* verwendet.

Feldversuche

Die Feldversuche wurden an unterschiedlichen Standorten und in unterschiedlichen Jahren durchgeführt. Gleichzeitig variieren die Rahmenbedingungen beispielsweise hinsichtlich der Fruchtfolge, der Sorte und des Managements.

Die Versuchsdaten der einzelnen Jahre zeigen individuelle Vor- und Nachteile hinsichtlich der Anwendung von Data-Mining-Werkzeugen. Die IKB-Versuche zeichnen sich beispielsweise durch eine sehr hohe Präzision in der Versuchsdurchführung und Datenerfassung aus. Allerdings kann damit nur ein geringer Datenumfang erzeugt werden. Dadurch wird insbesondere die Validierung erschwert und die Gefahr für Überanpassung steigt. Beispielsweise konnten im Versuchsjahr nur ca. 200 Teilflächen angelegt werden. Der umgekehrte Fall gilt für den Versuch *Görzig 04*. Hier konnte eine hohe Datenquantität erreicht werden – allerdings mit Einbußen in der Qualität. In den einzelnen Versuchen wird also mit unterschiedlicher **Datenquantität** und **Datenqualität** gearbeitet.

Über alle Versuche hinweg lassen sich Schwachstellen bezüglich der Abdeckung des Merkmalsraumes (*Data Coverage*) insbesondere für die Stickstoff-Applikationen erkennen. Diese Schwachstellen konnte teilweise mit der Einführung einer Variante *Merkmalsraum*, behoben werden. Mit dieser Variante wurden Applikationsmengen für Bereiche mit geringer Datendichte erzeugt.

Datenvorverarbeitung

Die Datenvorverarbeitung gestaltet sich insbesondere für die IKB-Versuche aufwändig, da in diesen Versuchen ansteuerbare Teilflächen erzeugt werden, die in Fahrgassenrichtung verlaufen (vgl. Kap. 4.2.1). Dadurch können keine Standard-Software-Anwendungen zur Datenvorverarbeitung verwendet werden. Die Kartierungen müssen daher individuell korrigiert - und daraus Ausprägungen für die Teilflächen bezüglich der Attribute ermittelt werden. Die Qualität des Kartenmaterials kann schwer beurteilt werden. Augenscheinlich traten diesbezüglich große Unterschiede auf. Aufgrund der neuartigen Versuchsanlage existieren keine Anhaltspunkte zur konkreten Ausgestaltung der Datenvorverarbeitung. Insbesondere gilt dies für vergleichsweise neue Sensoren, wie die Zugkraftmessung. Im Versuch *Görzig 03* wird eine Methode zur Datenvorverarbeitung angewendet, die für den Bereich Precision Farming häufig in der Literatur zu finden ist. Damit verbunden sind allerdings Einbußen in der Präzision. Die Instanzen für das überwachte Lernen, die das Resultat der Datenvorverarbeitung darstellen, haben somit unterschiedliche Charakteristiken.

Data-Mining-Phase

Die Modellierung in der Data-Mining-Phase wird im Wesentlichen durch die versuchsbedingten Voraussetzungen bestimmt. Dadurch, dass nicht speziell für die Anwendung er-

zeugte Versuchsdaten verwendet werden, treten - wie bei WED-Anwendungen üblich – Verzerrungen innerhalb des Datenmaterials auf. Die Auswirkungen der Verzerrungen können schwer beurteilt werden und auch nicht durch eine entsprechende Datenaufbereitung beseitigt werden.

Eine intensive Exploration der Datensätze ermöglicht es allerdings, mögliche Schwachstellen der produzierten neuronalen Netzwerke zu ermitteln. Diese Schwachstellen werden bei der jeweiligen Interpretation der Ergebnisse berücksichtigt.

Im Training der neuronalen Netzwerke zeigten sich mitunter Schwierigkeiten hinsichtlich der Auswahl von „besseren“ Netzwerken. Selbst unter identischen Voraussetzungen (gleiche Einstellungen, gleiche Attribute) ergeben sich unterschiedliche Ergebnisse als Resultat des Netzwerktrainings. Dies liegt im Wesentlichen an der Verwendung von selbstanpassenden Netzwerken und an der zufälligen Aufspaltung von Trainings- und Testmengen. Dennoch zeigt sich, dass mit diesen Verfahren das Risiko von Überlernen beschränkt werden kann. Aufgrund der heterogenen Instanzenmengen aus den Feldversuchen können hier zufällige Einflüsse auf das Ergebnis wirken.

Es kann allerdings mit Hilfe von *Response Surfaces* beobachtet werden, dass die unterschiedlichen Netze zwar im Detail unterschiedliche Reaktionen zeigen, aber dass die „groben“ Muster aber in einer hohen Beständigkeit auftreten.

Im Relation zu vergleichbaren Arbeiten (vgl. Kap. 3.2) können gute Ergebnisse hinsichtlich der Prognosegüte erzielt werden. Dazu wurde das Verfahren der Kreuzvalidierung angewendet.

Generell deutet es sich an, dass größere Datenmengen den Modellierungsprozess vereinfachen, da sich hier stabile Netzformen herausbilden. Problematisch ist die Modellierung bei einer geringen Datenmenge, wie beispielsweise im Versuch *D4 03*.

Ergebnisse

Mit den Daten der einzelnen Versuchsjahre kann gezeigt werden, dass sich Entscheidungsregeln zum Einsatz auf einem In-Field-Controller erzeugen lassen. Eine automatisierte Düngung ist somit prinzipiell umsetzbar. Das Ziel dieser Arbeit, einen durchgängigen Weg aufzuzeigen, wie Entscheidungsregeln generiert werden, kann anhand von empirischen Datensätzen vorgeführt werden.

Zur Erreichung des weiteren Ziels, Erkenntnisse hinsichtlich einer optimalen kleinräumigen N-Applikation zu gewinnen, werden weitere Verfahren angewendet. *Response Surfaces* und „beschränkte“ Entscheidungsbäume ermöglichen eine Interpretation der trainierten neuronalen Netzwerke im pflanzenbaulichen Kontext. Mit den vorgestellten Data-Mining-Werkzeugen wird damit eine Zugänglichkeit zu den empirischen Modellen erreicht und die Verständlichkeit verbessert.

Erkenntnisse, die aus dieser Arbeit gezogen werden können, bestehen aus (Grob-)Mustern. Trotz heterogener Datenlage zeigen sich diesbezüglich stabile (Grob-)Muster, die in Abschnitt 4.8.2 zusammengefasst sind. Die Muster bestätigen die Bedeutung einer hohen Anzahl von unterschiedlichen teilflächenspezifischen Informationen. Es kann kein Sensor identifiziert werden, der bedeutungslos für die Optimierung wäre. Gleichzeitig deuten sich entscheidende ökonomische Vorteile an, wenn auf die unterschiedlichen Charakteristiken der Teilapplikationen eingegangen wird. Dies kann in vorliegender Arbeit beispielsweise an Ertragskarten beobachtet werden. Zur ersten Teilapplikation geht von ihnen ein „inverser“ Einfluss für die Bemessung der N-Applikation nach ökonomischen Gesichtspunkten aus. Demnach liegt die optimale spezielle Intensität in Niedrigertragsbereichen höher als in Hohertragsbereichen. In den weiteren Teilapplikationen kehrt sich diese Richtung um.

Reflexionsoptische Messungen erweisen sich von hoher Bedeutung zur Optimierung der zweiten und dritten N-Applikation. In den optimierten Entscheidungsregeln sind Ähnlichkeiten zur Steuerung, wie sie beispielsweise mit einem N-Sensor verfolgt wird, nicht zu übersehen. Allerdings sollten auch zur Optimierung der zweiten und dritten N-Applikation Bodeninformationen verwendet werden, um das Optimierungspotenzial auszuschöpfen. Hinsichtlich der Bodenleitfähigkeit und der Zugkraft können keine konsistenten Ergebnisse erzielt werden, obwohl deren Einfluss zur Bemessung der Düngergabe nicht gering erscheint. Damit können die derzeitigen Ansätze z. T. bestätigt aber auch Widersprüche aufgedeckt werden.

Eine Gesamt-Interpretation der Ergebnisse ist in dieser Arbeit aufgrund der spezifischen Rahmenbedingungen der einzelnen Versuche schwierig, da die Modelle sehr komplex sind und das Herauspicken einzelner Teile zu Missverständnissen führen kann.

Die „besten“ Ergebnisse sind streng genommen die jeweiligen trainierten neuronalen Netze, die einen teilflächenspezifischen Ertrag für variierende N-Mengen vorhersagen können. In einem ersten Ansatz zur Verifikation zeigten sich überraschend gute Ergebnisse, die allerdings noch in Folgeversuchen bestätigt werden müssen. Weiterhin sind bis dato die Erfolgsfaktoren, d. h. die Elemente, die das bessere ökonomische Ergebnis bewirkt haben, noch nicht identifiziert.

Unter vorliegenden Bedingungen dürfen die Teilergebnisse nicht überinterpretiert werden: Beispielsweise dürfen die erzeugten Regeln zur N-Applikation nicht als allgemeingültig betrachtet werden (vgl. Kap. 5.2.). Es darf nicht übersehen werden, dass die Voraussetzungen bezüglich der Feldversuche modifiziert und weiter verbessert werden können. Dadurch können stabilere Ergebnisse erreicht werden. Das Potenzial dieser Methode ist weitaus höher einzuschätzen, als mit den vorliegenden Ergebnissen angedeutet wird.

Durch die Analyse der vorliegenden vier Versuche können tendenzielle Aussagen zu einer modifizierten Versuchsgestaltung für die Entwicklung von Entscheidungsregeln gemacht

werden: Es kann gezeigt werden, dass eine Grundvoraussetzung für ein Lernen eine Variante zur Bereicherung des Merkmalsraumes ist.

Die Qualität der Variante *Merkmalsraum* determiniert die Qualität der Modellierung in hohem Maße. Dabei konnte gelernt werden, dass eine zufällige Auswahl von Varianten um das geschätzte Optimum die beste Voraussetzung dafür ist. Dabei darf die Zahl dieser notwendigen Variationen nicht unterschätzt werden. Die kostengünstige Implementierung einer solchen Variante kann in den On-Farm-Experimenten in Halle gezeigt werden. Grundlegende Verbesserungen könnten durch eine Ausweitung der Variante Merkmalsraum erreicht werden, die im letzten Abschnitt dieser Arbeit vorgestellt wird.

Gleichzeitig ist diese Variante die Basis für Interpretationen des Versuchsergebnisses. Diese Variante erlaubt es, Produktionsfunktionen zu schätzen, die beispielsweise für die Bewertung des Referenzszenarios (einheitliche Bewirtschaftung) und zur Abschätzung des Optimierungspotenzials elementar sind (vgl. WEIGERT ET AL., 2004, S. 320ff; GANDORFER, 2006).

Weiterhin kann gezeigt werden, dass für die Data-Mining-Anwendung Teilflächen in einer hohen Anzahl benötigt werden. Die eingeschränkte Größe der Versuchsflächen (IKB-Versuche) führt zu einem sehr geringen Datenumfang, für den die Data-Mining-Algorithmen nicht ausgelegt sind. Insbesondere gilt dies dann, wenn die Datenrepräsentativität nicht angemessen ist. Ein größerer Datenumfang eröffnet erst die Möglichkeit zu einer Daten-Selektion, die auf das Lernverhalten der Modelle abgestimmt ist. Die Auswahl der Schläge mit der benötigten Variation und der geforderten Schlaggeschichte ist in erster Linie ein organisatorisches Problem und kann somit bei entsprechenden Ressourcen gelöst werden.

Es existiert sicherlich noch weiterer Forschungsbedarf zur Vorverarbeitung von Sensordaten, insbesondere für neuartige Sensoren (vgl. Zugkraft). Präzise Untersuchungen zur Datenvorverarbeitung bzw. zur Korrektur von Karten könnten zu besseren Resultaten in der Ertragsprognose führen. Jeder Sensor hat seine Eigenheiten und rechtfertigt vermutlich spezielle Prozeduren der Aufbereitung.

Weiterhin kann mit dieser Arbeit aufgrund der unterschiedlichen Datenbasis noch nicht die Frage beantwortet werden, ob sich die Entscheidungsregeln an einem Standort in unterschiedlichen Jahren unterscheiden. Mit dem Versuch *Görzig 05* deutet sich an, dass die Entscheidungsregeln stabil bleiben. Die Anpassung an jahresspezifische Einflüsse erfolgt über die reflexionsoptischen Kennzahlen.

Falls sich unterschiedliche Muster zeigen, stellt sich die Frage, wie diese zusammengeführt werden können. Hierzu existieren zwei Möglichkeiten: Erstens, man trainiert ein neuronales Netz mit den Daten aus mehreren Jahren, ohne dass man in dem Trainingsalgorithmus das Jahr mit einbezieht, aus dem die Trainingsdaten stammen. Damit muss das Netz eine

Art Mittelweg finden zwischen den Extremdaten aus den verschiedenen Jahren. Zweitens, man erstellt für jedes Jahr *ex post* die Entscheidungsregeln, also wie gedüngt hätte werden sollen, und vergleicht dann diese Regeln über die Jahre und versucht stabile Elemente herauszufinden. Mit konstanten Versuchsanlagen erscheint das Training von neuronalen Netzen mit einem kombinierten Datensatz viel versprechend.

Für eine systematische Untersuchung muss ein großes Spektrum an Bodenverhältnissen langjährig überprüft werden (vgl. BULLOCK UND BULLOCK, 2000, S. 96ff).

5.2 Wissensentdeckung in Datenbanken und Precision Farming

In diesem Kapitel sollen rückblickend die Chancen, aber auch die Schwierigkeiten dargestellt werden, die mit der Anwendung von Data Mining in der Wissenschaft einhergehen und auch in dieser Arbeit offenkundig geworden sind.

Methodischer Ansatz

Wie im Kapitel 2.1 angedeutet, liegt der Methode Wissensentdeckung in Datenbanken ein *explorativer Ansatz* zugrunde (vgl. PRESS, 2004, S. 2f). Nach dieser „Philosophie“ werden zuerst Daten angesammelt und darin nach Mustern, bzw. Zusammenhänge gesucht. Die Data-Mining-Werkzeuge (z. B. neuronale Netze oder Entscheidungsbaumverfahren) beruhen auf dem Induktionsprinzip (induktives Lernen aus Beispielen, vgl. ausführlich JAFAR-SHAGHAGHI, 1996, S. 17ff).

Im Gegensatz dazu steht der verifikative Ansatz, der auf dem Deduktionsprinzip (Gesetz begründendes Verfahren) basiert. Nach diesem Ansatz werden Hypothesen aufgestellt, die dann mit experimentellen Daten unterstützt bzw. verworfen werden können.

Mit dem explorativen Ansatz ergibt sich die Chance, neue Erkenntnisse zu generieren. Es können neue Wege gegangen werden, die dem verikativen Ansatz versperrt blieben (vgl. SCHOMMER, 2003, S. 13). Dies gilt insbesondere für den Fall, dass keine Anhaltspunkte für geeignete Hypothesen zur Verfügung stehen. Allerdings darf nicht der Fehler gemacht werden, den explorativen Ansatz als ein Gesetz begründendes Verfahren anzusehen.

„Durch Induktion werden keine sicheren und hundertprozentigen Aussagen gewonnen, sondern durch Induktion anhand von Beispielen und Beobachtungen können heuristische Aussagen formuliert und dargestellt und unter Umständen geprüft werden.“ (Jafar-Shaghaghi, 1996, S. 19)

Die Gefahr der Methode Wissensentdeckung in Datenbanken bzw. Data Mining besteht darin, dort Muster zu finden, wo eigentlich keine vorhanden sind. Im Extremfall werden kleine Abweichungen, die z. B. durch Messrauschen entstanden sind, als Muster identifiziert (HAND ET AL., 2001; PRESS, 2004, S. 3).

Gleichzeitig existiert in dieser Methode kein theoretisches Modell in analytischer Form, anhand dessen die empirischen Ergebnisse überprüfbar sind. Es fehlen damit „Anknüp-

fungspunkte“, mit denen verschiedene Ergebnisse miteinander verglichen werden können. Es kann nur schwer auf den Ergebnissen anderer Modelle aufgebaut werden. Diese Kritik gilt auch für die Ableitung von Erkenntnissen aus empirischen Modellen im Bereich Precision Farming speziell bei On-Farm-Versuchen (vgl. LOWENBERG-DEBOER UND BOEHLJE, 1996, S. 923ff; ADAMS UND COOK, 2004, s.p.).

Die Gefahr für Fehlinterpretationen resultiert in erster Linie aus dem Datenmaterial. In der Regel wird dieses Datenmaterial nicht aus Experimenten generiert, die für die Analyse konzipiert wurden, sondern entstehen quasi als Nebenprodukt (vgl. WIEDERHOLD, 1996, vii). Die Daten sind somit nicht für jede Ex-post-Analyse geeignet (vgl. HAND ET AL., 2001, S.21). Werden diese Umstände nicht berücksichtigt, ergeben sich dadurch Probleme:

“The data collected typically violate many of the desired assumptions of traditional statistics, such as attributes being independent, identically distributed, and having low correlation over time and space. Often, the data are highly correlated spatially or temporally, or both. Violation of any of these assumptions can generate major problems when the decision maker is trying to draw substantive conclusions. These data are generally just collections of multidimensional observations, without regard to how they were collected.” (PRESS, 2004, S. 6)

Dieses Problem ist jedoch, sofern in die Datenerzeugung eingegriffen werden kann, lösbar. Insbesondere ist dies für wissenschaftliche Zwecke vorstellbar. Im hier vorgestellten Fall konnte beispielsweise durch die Einführung einer Variante *Merkmalsraum* die Qualität der Ergebnisse gesteigert werden. Wie im folgenden Kapitel dargestellt wird, können allerdings noch weitere Details in der Versuchsanstellung verändert werden, um zu einer Datenbasis zu gelangen, die in der Data-Mining-Phase hilft, Fehlinterpretationen zu vermeiden. Durch eine entsprechende Versuchsgestaltung können viele Probleme in der Data-Mining-Phase vermieden werden.

Ob man sich damit von der Definition von Data Mining bzw. WED entfernt, sei dahingestellt. Die strenge Trennung zwischen den beiden wissenschaftlichen Ansätzen ist ohnehin nur idealtypisch (vgl. WITTEN UND FRANK, 2001, S. 28).

Einfluss des „Data Miners“

Schwieriger stellt sich die Beurteilung hinsichtlich des „Data Miners“ dar, des Menschen, der den Data Mining Prozess steuert. Im Gegensatz zum verifikativen Ansatz ist sein Einfluss auf die Ergebnisse weitaus höher:

„The contributions of humans to data mining are often neglected. Results are dependent on the predictive value of the features. It is the human who specifies the set of features, who usually knows to discard the junk and who investigates how to transform the original features into better features. [...] No universal best approach is describable for data mining; making good decisions is part art, part science: to use the science when it is known and effective, and to offer guidance on many practical issues that are not readily quantified (WEISS UND INDURKHYA, 1998, S. 21)

Auch mit dieser Arbeit wurde offenkundig, dass Entscheidungen in jeder Phase des WED-

Prozesses zu treffen sind, ohne dass die Konsequenzen genauer untersucht werden können. Insbesondere gilt dies für die Konfiguration der Modelle und die Integration der Modelle in die Anwendungsumgebung. KRATZER (1993, S. 196) vertieft diese Kritik am Beispiel neuronaler Netze. Zum Entwurf des Netzes und zu der Festlegung von Trainingsparametern existieren bis dato wenige Anhaltspunkte, die ein optimales überwachtetes Lernen gewährleisten. Ähnliches gilt für die Aufbereitung des Datenmaterials (vgl. WIDMANN, 2001, S. 133 bzw. Kap. 2.4). Es fehlen Orientierungspunkte bzw. Standards, wie sie in der klassischen Statistik (bzw. im verifikativen Ansatz) stärker anzutreffen sind (vgl. ELDER UND PERIGBON, 1996, S. 90ff).

WED-Prozess

Mit dem Einsatz von Data-Mining-Technologien ergibt sich daher auch eine Fülle von möglichen Untersuchungsgegenständen, die in der Regel nicht erschöpfend bearbeitet werden können:

„In fact, data mining often appears to be a process of working with and fixing up the data in light of what a model shows until time and resources run out. At the point that time and resources expire, the project is declared either a success or a failure, and it's on to the next one! (PYLE 2003, S. 368)“

Durch die Interaktion bzw. das hybride Agieren (vgl. Kap. 2.1) wird der Grad der Komplexität im WED-Prozess erhöht. Die einzelnen Eingriffe bzw. Auswirkungen können daher in der Regel nicht erschöpfend analysiert werden. Gleichzeitig wird die Darstellung des Gesamtprozesses erschwert. Dies wird auch in den betrachteten Veröffentlichungen (vgl. Kap. 3.2) deutlich.

Daher muss der Anwender in diesem interaktiven Prozess durch das Data-Mining-System unterstützt werden, um nicht selbst den Überblick zu verlieren (BRACHMAN UND ANAND, 1996, S. 52ff). Dies kann vor allem durch entsprechende Softwarepakete erfolgen (vgl. UTHURUSAMY, 1996). Insbesondere in der hier vorgestellten Anwendung ist die Datenverarbeitung immer zweiteilig, d. h. die Instanzen müssen erst in GIS-Systemen erzeugt und vorverarbeitet werden und können dann erst an Data-Mining-Programme übergeben werden. Dort werden sie einer intensiveren Analyse unterzogen. Dies erschwert eventuelle Rückkopplungsschleifen, wenn beispielsweise auffällige Instanzen in ihrer räumlichen Lage (als Teilflächen) untersucht werden sollen.

Eignung der Methode für den Bereich Precision Farming

Zusammenfassend kann festgehalten werden, dass sich die Methode Wissensentdeckung in Datenbanken unter den vorgestellten Voraussetzungen nicht zum systematischen Aufbau und zur statistischen Absicherung von wissenschaftlichen Erkenntnissen eignet. Schwierigkeiten ergeben sich durch die Komplexität und Interaktion in dem vorgestellten Prozess. Besonders in der hier gezeigten Anwendung werden sehr viele Eingriffe des Anwenders verlangt, die im Rahmen dieser Arbeit nicht genauer untersucht werden konnten. Dies gilt

insbesondere für die Datenvorverarbeitung, Datenaufbereitung und für die Auswahl eines trainierten neuronalen Netzwerkes.

Dennoch bietet diese Methode entscheidende Vorteile: Die Methode ist sehr gut geeignet, um in einem ersten Zugriff aus einem komplexen Datenmaterial, wie es durch den Einsatz von multiplen Sensoren bei Precision Farming der Fall ist, nach Ansätzen (Mustern) zu suchen und Modelle zur Optimierung zu erzeugen. Grundsätzlich kann mit dieser Arbeit die Eignung der Methode Wissensentdeckung in Datenbanken zur Entwicklung von kleinräumigen Entscheidungsregeln gezeigt werden.

Nach dem derzeitigen Kenntnisstand ist dies darüber hinaus die einzige Methode, um systematisch ökonomisch optimierte und sensorbasierte Algorithmen für die kleinräumige N-Düngung zu generieren. Mit den vorgestellten Werkzeugen (insbes. *Response Surfaces* und Entscheidungsbaumalgorithmen) ist es möglich, den WED-Prozess transparent zu gestalten. Die Transparenz ermöglicht gleichzeitig ein Verständnis für die Zusammenhänge der kleinräumigen Ertragsbildung. Diese Methode bietet sich demnach insbesondere für die Entwicklung von Algorithmen zur Input-Optimierung für die Praxis an.

Darüber hinaus kann eine (statistische) Verifikation der generierten Algorithmen in anschließenden Feldversuchen erfolgen (vgl. Kap. 4.1). Ebenso können interessante Muster in diesen Algorithmen z. B. mit traditionellen statistischen Verfahren in Exaktversuchen näher untersucht werden.

Der Vorteil dieser Methode ist weiterhin, dass sie auf alle kleinräumigen Betriebsmitteloptimierungen übertragbar ist. Die Elemente der Modellierung können bei entsprechender Datenlage ohne weiteres übernommen werden. Weiterhin ist diese Methode beliebig skalierbar. Das bedeutet, dass die Daten neuer Sensoren in den Data-Mining-Prozess einfließen und analysiert werden können. Ebenso können noch wesentlich größere Datenmengen in einen solchen Prozess einfließen, da die Algorithmen insbesondere für sehr große Datenmengen ausgelegt sind. Bei Bedarf kann die Versuchsanstellung z. B. durch Erhöhung der Teilflächenanzahl beliebig verfeinert werden.

Ebenfalls sind Anwendungen mit einem anderen Zielsystem denkbar, beispielsweise eine Optimierung nach ökologischen Gesichtspunkten. Auch eine ökonomisch-ökologische Optimierung wäre im Sinne einer nachhaltigen Landbewirtschaftung möglich. Die Änderung besteht prinzipiell nur im Zielsystem, das zur Auswahl der Input-Output-Kombinationen herangezogen wird. Die Input-Output-Kombinationen können –wie in dieser Arbeit– über die trainierten neuronalen Netzwerke ermittelt werden. Beispielsweise kann aus dem Input (N) und dem prognostizierten teilflächenspezifischen Ertrag die potenzielle Stickstoff-Auswaschung prognostiziert werden, um dahingehend eine Auswahl zu treffen. Darüber hinaus können mit Hilfe von neuronalen Netzen bei entsprechenden Trainingsinstanzen mehrere Output-Attribute gleichzeitig geschätzt werden. Damit könnte bei entsprechender

teilflächenspezifischer Erfassung gleichzeitig der Ertrag und Qualitätsparameter in Abhängigkeit von der Stickstoff-Applikation geschätzt werden.

Generell kann festgehalten werden, dass die Optimierung der kleinräumigen Stickstoff-Ausbringung mit der hier vorgestellten Methode eines der schwierigsten Anwendungsfelder ist. Das liegt insbesondere daran, dass die Stickstoffgabe in drei Teilgaben appliziert wird. Während dieser Zeit kann bei entsprechenden Witterungsverhältnissen Stickstoff vom Boden nachgeliefert werden. Dadurch wird die Komplexität enorm erhöht. Die Erzeugung eines verzerrungsfreien Merkmalsraumes ist hier ebenfalls eine große Herausforderung (aber: vgl. Kap. 5.3).

Ein einfacheres Beispiel zur Anwendung der hier vorgestellten Methode, wäre beispielsweise die teilflächenspezifische Fungizid bzw. Wachstumsregler-Optimierung, die in der Regel nur in ein bis zwei Teilapplikationen erfolgt. In diesem Fall kann darüber hinaus eine höhere Präzision in der Datenerfassung erreicht werden, da die Applikationsgenauigkeit bei Feldspritzen in der Regel sehr hoch ist.

Zu einer Beurteilung der Methode müssen vor allem auch die Kosten für die Durchführung angeführt werden. Die Kosten für eine Versuchsanlage, wie sie beispielsweise für den Versuch *Görzig 04* verwendet wurden, sind in Relation zu Exaktversuchen sehr gering (vgl. GANDORFER ET AL., 2004). Bei entsprechender Ausstattung fallen für einen Precision-Farming-Betrieb sehr geringe versuchsbedingte Mehrkosten an. Die Kosten für die Auswertung (Einsatz von Data Mining) sind schwer abzuschätzen, dürften aber bei entsprechender Erfahrung nicht höher liegen als im klassischen Exaktversuchswesen. Wenn sich bestätigt, dass nur ein Versuch zur Erzeugung von Entscheidungsregeln für spezifische Praxisverhältnisse benötigt wird (vgl. *Görzig 05*), hätte dies beträchtliche Konsequenzen für einen möglichen Praxiseinsatz. In diesem Fall wäre sogar eine einzelbetriebliche Anwendung dieser Methode vorstellbar.

5.3 Verbesserungspotenzial und Ausblick

Die Verbesserungsmöglichkeiten ergeben sich durch die angesprochenen kritischen Punkte im empirischen Teil dieser Arbeit. Die wichtigsten Punkte betreffen den Datenumfang, die Ausgestaltung der Variante Merkmalsraum und die Präzision in der Datenerfassung (z. B. Ertragskarten) und Prozesssteuerung. Die hier erzielten Ergebnisse und Erfahrungen aus verschiedenen Versuchsanlagen und Modellierungen münden in einer angepassten Versuchsgestaltung, in der die Vorteile der beiden betrachteten Versuchsanlagen (IKB und *Görzig 04*) verbunden werden können.

Damit wird der Gedanke weiter ausgebaut, nicht nur „für andere Zwecke aufgezeichnete Daten“ zu verwenden, sondern die Datenerhebung gezielt zu beeinflussen (vgl. UTHURUSAMY, 1996, S. 567).

Versuchsflächen

Der Ausgangspunkt der Anforderungen liegt in der Auswahl der Versuchsflächen. Wie gezeigt werden konnte, führt erst eine hohe Anzahl von Instanzen zu stabilen Ergebnissen. Dafür sind auch entsprechend große Versuchsschläge erforderlich (vgl. *Görzig 04*). Eine hohe Anzahl von Instanzen erlaubt eine größere Flexibilität in der Datenaufbereitung und führt zu stabileren Ergebnissen in der Data-Mining-Phase.

Die verwendeten Versuchsschläge sind im Sinne eines WED-Prozesses umso wertvoller, je größer die Variation der Attribute innerhalb der Teilflächendaten ist. Gleichzeitig sollten insbesondere die Bodenattribute wiederholt erfasst werden, um die zeitliche Stabilität untersuchen zu können. Interessant sind hier insbesondere die Erträge bei einheitlicher Bewirtschaftung, die als „Bodenattribut“ verwendet werden können. Damit wird eine gewisse Vorlaufzeit benötigt, bevor ein Schlag als Versuchsschlag verwendet werden kann.

Die am schwersten realisierbare Anforderung betrifft die Variabilität der Versuchsschläge. Die Schläge sollten den Merkmalsraum in Bezug auf Bodenattribute gleichmäßig (!) abdecken. Das bedeutet, dass hier Schläge interessant sind, die einen besonders breiten Bereich an Bodenattributen gleichmäßig abdecken. Dadurch gewinnt der Data Mining Prozess an Breite, d. h. die Entscheidungsregeln sind für größere Bereiche (Standorte) anwendbar.

Versuchsausrüstung und Datenaufbereitung

Ein wichtiges Ziel ist die Erhöhung der Präzision in der Versuchsdurchführung. Es sollten möglichst kleinräumige Teilflächen erzeugt werden, von denen die Input-Output-Kombinationen und die kleinräumigen Verhältnisse exakt erfasst werden können.

Dies kann dadurch erreicht werden, dass präzise Applikationsgeräte eingesetzt werden. Beispielsweise könnte auf Schleuderstreuer verzichtet werden und Applikationsgeräte eingesetzt werden, bei denen sich die einzelnen Teilbreiten separat ansteuern lassen (vgl. IKB-Versuche). Bei entsprechender Applikationsgenauigkeit könnten auch die in dieser Arbeit verwendeten 10er Schritte in den möglichen Stickstoff-Applikationsmengen durch eine kleinere Abstufung abgelöst werden und dadurch das ökonomische Optimierungspotenzial erhöht werden⁷⁵.

Eine weitere Chance zur Erhöhung der Präzision kann durch den Einsatz von Feldrobotik und einer Abstimmung der Arbeitsbreiten der eingesetzten Maschinen erreicht werden. Die minimal ansteuerbare Teilbreite in der Applikation sollte dabei die Ausgangsbasis für die maschinentechnische Ausstattung und für die Teilflächeneinteilung darstellen. Das Konzept ist in Abbildung 64 dargestellt. In dieser Abbildung werden Mähdrescher und Bodenbearbeitungsgeräte eingesetzt, deren Arbeitsbreite der minimal ansteuerbaren Teilbreite in der Applikation entspricht. Es ergeben sich damit große Ähnlichkeiten zu dem Konzept

⁷⁵ Durch eine feinere Abstimmung kann N_{opt} exakter getroffen werden.

Controlled Traffic (BERGLUND UND BUICK, 2005, S. 43).

Der Einsatz von Technologien zur Feldrobotik⁷⁶ ermöglicht, dass die Überfahrten in den verschiedenen Jahren in derselben Spur erfolgen. Damit wird zum einen sichergestellt, dass keine Überlappungen auftreten, und –was ebenso bedeutsam ist– dass die Sensormessungen aus den unterschiedlichen Jahren von exakt denselben Spuren stammen. Dadurch, dass die Bearbeitungsrichtung durch die Anwendung von Parallelfahrssystemen fixiert werden kann, kann ein quadratisches (Standard)-Raster verwendet werden, das in die Bearbeitungsrichtung gedreht wird. Erst mit dieser Versuchsdurchführung kann ermöglicht werden, dass jeder Teilfläche die berechnete Soll-Applikation zugeführt wird und dass von dieser Teilfläche auch der Output exakt bestimmt werden kann. Die Genauigkeit der Applikation würde dadurch stark erhöht werden. Durch diesen Ansatz kann die Präzision erhöht und die zeitliche Variabilität der Attribute wesentlich genauer bestimmt werden.

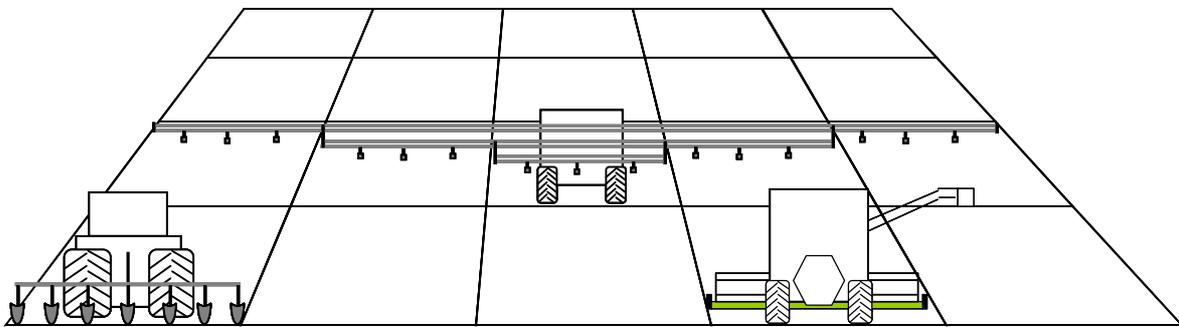


Abbildung 64: Abstimmung der Arbeitsbreiten und angepasste Rasterung

Im Beispiel der Abbildung 64 erfolgt die Applikation genau in dem Bereich, in dem der Output mit der Ertragserfassung gemessen wird. Gleichzeitig können beispielsweise die Zugkraftmessung genau in dieser Spur erfasst werden und zwar in jedem der Versuchsjahre. Es können „scharf“ abgegrenzte und ansteuerbare Teilflächen erzeugt werden.

Zur Auswertung können dann „kritische Streifen“, also beispielsweise die Streifen auf denen Fahrgassen liegen, grenzgenau ausgeschlossen werden.

Die wichtigsten Punkte in der Versuchsdurchführung sind also die Abstimmung der Arbeitsbreiten, die Unterstützung durch Parallelfahreinrichtungen und die Anwendung eines stabilen Rasters in Fahrgassenausrichtung. Damit gelangt man zu Instanzen, von denen über Jahre hinweg präzise Daten erhoben werden können. Ebenso eröffnet dieser Ansatz enorme Möglichkeiten zur Erzeugung eines gleichmäßig abgedeckten Merkmalsraumes.

Merkmalsraum

Verbesserungen der Ergebnisse können wie bereits dargestellt erreicht werden, wenn die Daten, die in das Modell gelangen nicht nur ein Nebenprodukt der Bewirtschaftung sind,

⁷⁶ Z. B. die kommerziell verfügbaren Technologien AutoGuide® oder AutoTrac®.

sondern speziell für diese Form der Auswertung erzeugt werden. Im optimalen Fall sollte die Variante „Merkmalsraum“ speziell für das Attribut N-Applikation einen verzerrungsfreien und gleichmäßig abgedeckten Merkmalsraum liefern. Der betrachtete und damit variable Faktor sollte im Idealfall zum einen in allen möglichen Ausbringungsmengen und zum anderen in allen möglichen Kombinationen mit den Ausprägungen der Bodenattribute vorliegen. Die Methode der Randomisierung wie sie in den Versuchen, speziell im Versuch *Schafhof 04*, angewendet wurde, hat sich als Schritt in die richtige Richtung erwiesen. Allerdings erhöht sich in diesem Versuchsdesign der Aufwand beträchtlich, da die Ausrüstung von Versuchsstationen (z. B. Parzellenmähdrescher) benötigt wird.

Das Ziel besteht demnach darin, für ein On-Farm-Versuchsdesign (vgl. *Görzig 04*) einen ähnlich guten Merkmalsraum zu produzieren. Erste Ansätze für ein modifiziertes Versuchsdesign liefern PRINGLE ET AL. (2004a, S. 621f). In diesem Ansatz wird die Applikation einer Teilgabe idealtypisch sowohl in Fahrtrichtung als auch in der Querverteilung kontinuierlich sinusförmig schwingend um einen Mittelwert ausgebracht⁷⁷. Im einfachsten Fall wird dazu beispielsweise folgende Funktion verwendet:

$$f(x, y) = \sin(x) * \sin(y)$$

Die x- und y-Werte beziehen sich dabei auf die x- bzw. y-Achse eines Koordinatensystems, also beispielsweise auf einen Ausschnitt eines Schlages. Mit dieser Funktion kann eine Applikationsoberfläche erzeugt werden, wie sie in Abbildung 65 dargestellt ist. Die Oberfläche repräsentiert dabei die Soll-Applikations-Menge. Der Funktionswert kann in eine Düngermenge umgewandelt werden, bei der beispielsweise der Wert „1“ der maximalen N-Menge und der Wert „-1“ der minimalen Düngermenge entspricht. Für die Ausführung eines solchen Versuches stehen allerdings noch mehrere Fragen offen.

Forschungsbedarf besteht beispielsweise darin, eine angepasste Schwingungsperiode zu ermitteln, die versuchstechnisch durchführbar ist und möglichst kleinräumige Schwingungen erlaubt, ohne dabei die räumliche Autokorrelation zu erhöhen⁷⁸.

Dazu sind wiederum die Anforderungen an die Abstimmung der eingesetzten Maschinen sowie die Anlage von stabilen und separat ansteuerbaren Parzellen zu beachten.

Gleichzeitig muss darauf geachtet werden, dass bei drei Teilapplikationen keine Verzerrung im Merkmalsraum entsteht. Somit darf beispielsweise nicht dreimal hintereinander dieselbe Funktion zur Applikation verwendet werden, da sonst Teilgaben hochgradig korreliert sind und sich die Effekte einer Teilgabe nicht ermitteln lassen. Dies könnte beispielsweise dadurch verhindert werden, dass die Schwingungsfrequenz verändert wird. Weiterhin muss untersucht werden, inwieweit sich kontinuierliche Veränderungen quer zur Bearbeitungs-

⁷⁷ Da innerhalb einer ansteuerbaren Teilbreite die Ausbringungsmenge nicht variiert werden kann, müssen die kontinuierlichen Soll-Werte diskretisiert werden.

⁷⁸ In ihrem Versuch verwendeten PRINGLE ET AL. (2004b, S. 629) eine Schwingungslänge von 120m.

richtung erzeugen lassen, bzw. welche Abstufungen in der Ausbringungsmenge bei benachbarten Fahrspuren gewählt werden.

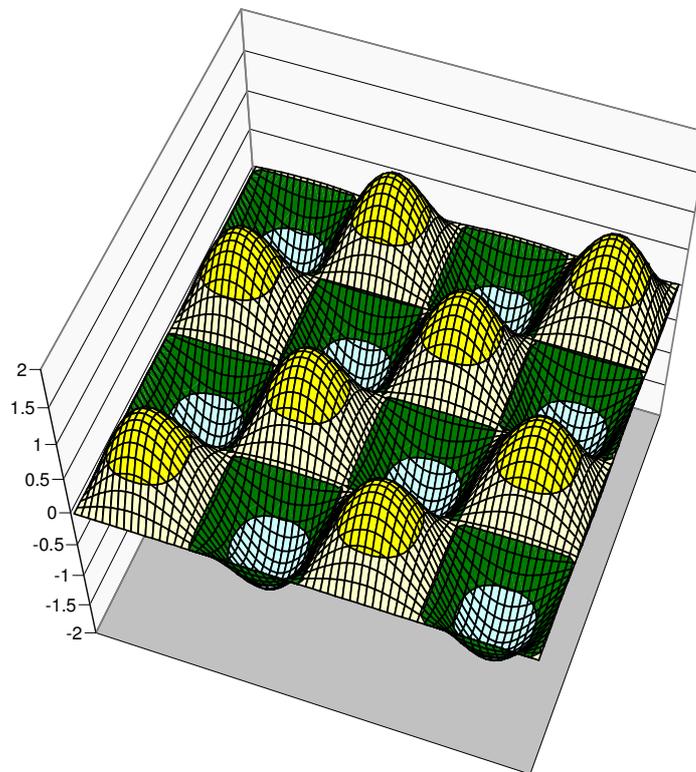


Abbildung 65: Zweidimensionale Sinus-Wellen (verändert nach PRINGLE ET AL. 2004a, S. 621)

Mit einem Versuch in diesem Design können reichhaltige Merkmalsräume ohne Verzerrung aufgebaut werden. Insbesondere können die kontrollierbaren Input-Mengen in einer beliebigen Verteilung erzeugt werden. Damit können die Probleme in der Wahl der Datenaufbereitung, insbesondere Datenbalancierung usw. behoben werden.

Dies zeigt, dass die Probleme, die im empirischen Teil dieser Arbeit aufgetreten sind, mit einer Anpassung der Versuchsdurchführung relativ einfach unter dem Einsatz bereits in der Praxis verfügbarer Technologien behoben werden können und gleichzeitig die Präzision der Modellierung enorm angehoben werden kann, ohne die Kosten der Versuchsdurchführung wesentlich zu beeinflussen. All diese Punkte verdeutlichen das große Potenzial der Methode Wissensentdeckung in Datenbanken im Bereich Precision Farming.

6 Zusammenfassung

Als ein Hemmnis für die Akzeptanz von Precision Farming in der Praxis können Wissensdefizite identifiziert werden: Es existieren bis dato nur wenige Algorithmen zur Umsetzung einer kleinräumigen Stickstoff-Applikation (N-Applikation), die überlegene ökonomische Ergebnisse erwarten lassen. Aus diesem Grund kommt der Entwicklung von Entscheidungsregeln, in denen Wissen zur teilflächenspezifischen Düngung abgebildet ist, besonderer Bedeutung zu (vgl. Kap. 1).

Aufgrund der Fülle von teilflächenspezifischen Daten, die durch den Einsatz von Precision Farming entstehen, bietet sich zur Analyse die Methode Wissensentdeckung in Datenbanken (WED) an. Dazu gilt es, den gesamten Prozess der WED für den Bereich Precision Farming anzupassen und daraus Erkenntnisse abzuleiten. Die Ausgangsfrage dieser Arbeit ist in der Untersuchung der Anwendbarkeit der Methode Wissensentdeckung in Datenbanken zu sehen. Dennoch sollen – sofern als möglich – Wissensbausteine zur Optimierung einer teilflächenspezifischen N-Applikation identifiziert werden.

Der Prozess WED setzt sich aus Datengewinnung, Datenaufbereitung, Anwendung von Data-Mining-Techniken und Validierung zusammen. Diese einzelnen Phasen werden im Hinblick auf die Anwendung für Precision Farming vorgestellt (Kap. 2). Der Fokus liegt auf der Datenvorbereitung und auf den Data-Mining-Werkzeugen, die in dieser Arbeit eingesetzt werden (neuronale Netzwerke sowie Entscheidungsbaumalgorithmen).

Der Einsatz der Methode WED setzt weiterhin eine intensive Analyse des Anwendungsgebietes voraus (Kap. 3). Dazu zählt insbesondere die Stickstoffdüngung bei Winterweizen und deren teilflächenspezifische Umsetzungen mit den entsprechenden Sensorinformationen. Aus dieser Analyse ergeben sich mehrere Hinweise für mögliche Verbesserungen. Zum einen sollten die Entscheidungsregeln auf ein Zielsystem, nämlich auf die Maximierung der teilflächenspezifischen stickstoffkostenfreien Leistung, ausgerichtet werden. Zum anderen sollten Entscheidungsregeln alle zur Verfügung stehenden teilflächenspezifischen Daten berücksichtigen. Schlussendlich sollten für jede N-Teilapplikation individuelle Entscheidungsregeln entwickelt werden, um damit die jeweilige Informationslage zu berücksichtigen. Diese Anforderungen werden in dieser Arbeit umgesetzt. Ansätze zur Gestaltung des WED-Prozesses liefern insbesondere Arbeiten zur teilflächenspezifischen Ertragsprognose mit künstlichen neuronalen Netzen (Kap. 3.2).

Aus diesen Vorüberlegungen heraus wird ein Prozessmodell WED für die Optimierung der kleinräumigen N-Applikation definiert (Kap. 4.1). Zur Anwendung der Methode werden insgesamt fünf Feldversuche durchgeführt, deren Design schrittweise optimiert wird (Kap. 4.2). Die zur Verfügung stehenden teilflächenspezifischen Daten setzen sich aus (historischen) Ertragskarten, Zugkraftkarten, Bodenleitfähigkeitskarten (*EM38*), reflexionsopti-

schen Messungen (Vegetationsindex *REIP*) und den Stickstoffmengen der Teilapplikationen zusammen. Ein wesentlicher Schritt in der Datenvorbereitung ist die Erzeugung von Instanzen (Lernbeispielen) aus diesen räumlichen Daten.

Die eigentliche Data-Mining-Phase wird in drei Modellierungsschritte aufgeteilt. Der erste Schritt besteht in einer teilflächenspezifischen Ertragsprognose mit Hilfe neuronaler Netze. Als Input-Parameter werden dazu unter anderem die N-Applikationsmengen der Teilgaben verwendet. Die trainierten neuronalen Netze werden im nächsten Schritt zur ökonomischen Optimierung von Teilapplikationen verwendet. Im letzten Schritt werden Entscheidungsregeln mit Hilfe von Entscheidungsbaumverfahren konstruiert. Zu diesem Zweck werden die ersten beiden Schritte wiederholt angewendet, um Datensätze zu erzeugen, in denen ökonomisch optimierte N-Applikationen für unterschiedliche teilflächenspezifische Situationen abgetragen sind.

Der gesamte Prozess der WED wird für die drei N-Teilapplikationen durchgeführt und für die jeweiligen Versuchsjahre wiederholt. Dadurch können in allen Bereichen dieses Prozesses wichtige Erkenntnisse gewonnen werden.

Insbesondere zeigt sich, dass zur Anwendung von Data-Mining-Techniken eine Anpassung des Versuchsdesigns nötig ist. Dies wird durch die Anlage von N-Steigerungsversuchen erreicht, die in ein On-Farm-Research-Design eingebunden werden. Gleichzeitig zeigen die Ergebnisse, dass eine intensive Exploration der Datensätze unumgänglich ist, um mögliche Verzerrungen in den Datensätzen beurteilen zu können.

Die trainierten neuronalen Netze wurden mit Hilfe von *Response Surfaces* untersucht. Mit dieser Darstellungsform kann die Prognose der neuronalen Netze graphisch dargestellt und interpretiert werden. Dadurch können bereits in diesem Schritt erste Muster extrahiert werden.

Weitere Hinweise ergeben sich durch die erzeugten Entscheidungsbäume. In diesen Entscheidungsbäumen sind ökonomische Prämissen eingebaut. Dadurch kann in den Entscheidungsbäumen die Bemessung der N-Applikation aus ökonomischer Sicht abgelesen werden.

Als problematisch erweist sich die hohe Komplexität der erzeugten Regeln, die eine einfache Interpretation behindert. Insbesondere wird dadurch eine Zusammenfassung der Muster über die verschiedenen Versuche hinweg erschwert. Aus diesem Grund werden die Entscheidungsbäume zur Ableitung von Mustern in ihrer Dimension beschränkt. Zur Steuerung der Ausbringungsmenge kann jedoch auf diese Komprimierung verzichtet werden. Die Extraktion von (Grob-)Mustern zeigt über die Versuchsjahre hinweg folgende Ergebnisse:

Muster 1: Bedeutung der teilflächenspezifischen Informationen

Es deutet sich an, dass über die Kombination mehrerer teilflächenspezifischer Attribute ein

entscheidender Vorteil zur ökonomischen Optimierung erreicht werden kann. In dieser Arbeit stellt sich kein Attribut heraus, das per se unwichtig für die Bemessung der N-Gabe wäre. Es existieren allerdings bezüglich des Standortes und der Teilapplikationen große Unterschiede in der Bedeutung der jeweiligen Attribute.

Muster 2: Optimierung der Teilapplikationen

Als weiteres Muster deutet sich mit dieser Arbeit an, dass die Optimierung der Teilapplikationen auf Basis der maximal verfügbaren Informationen einen entscheidenden ökonomischen Vorteil bringt.

Muster 3: Ertragskarten

Ertragskarten spielen für die Bemessung der N-Gabe eine große Rolle. Die Ergebnisse zeigen, dass zur ersten N-Applikation ein „inverser“ Mapping-Ansatz ökonomische Vorteile bringt. Nach diesem Muster sollte für Teilschläge aus Niedrigertragsbereichen die erste N-Applikation erhöht werden und umgekehrt. In den folgenden Applikationen ergeben sich in Abhängigkeit von den reflexionsoptischen Messungen differenzierte Entscheidungsregeln.

Muster 4: Reflexionsoptische Messungen (REIP)

Der Vegetationsindex REIP liefert entscheidende Anhaltspunkte für die Bemessung einer ökonomisch optimierten N-Gabe in einem konstanten Muster: In Teilflächen mit schlechter N-Versorgung (vergleichsweise niedriger REIP-Wert) erhöht sich die ökonomisch optimale N-Menge und umgekehrt. Allerdings zeigen die Ergebnisse, dass dieses Attribut nicht isoliert von den Bodenattributen zur Entscheidung verwendet werden sollte.

Muster 5: Bodenleitfähigkeit

Bezüglich der scheinbaren elektrischen Bodenleitfähigkeit kann kein stabiler Zusammenhang ausgemacht werden. Der Einfluss auf die Bemessung der Düngergabe ist in den einzelnen Versuchen unterschiedlich.

Muster 6: Zugkraft

Das Attribut Zugkraft wurde als neues Bodenattribut in die Untersuchung mit einbezogen. Es deutet sich an, dass dieses Attribut einen Einfluss auf die Bemessung der teilflächenspezifischen N-Applikation hat. Dies gilt vor allem zur ersten und zweiten Teilapplikation.

Nach dem erstellten Konzept WED für Precision Farming sollen die trainierten neuronalen Netze bzw. die Entscheidungsregeln im darauf folgenden Versuch eingesetzt werden, um eine Validierung bzw. Verifikation anzustreben. Dies gelingt aufgrund der schrittweisen Anpassung des Versuchsdesigns erst im letzten Versuchsjahr 2005 mit dem Versuch *Görzig*. Es zeigen sich hohe ökonomische Vorteile gegenüber Vergleichsvarianten (schlageinheitliche Düngung und Sensordüngung). Für statistisch abgesicherte Aussagen muss dieser Versuch allerdings mehrmalig wiederholt werden.

Zusammenfassend kann gezeigt werden, dass mit der Methode WED Muster identifiziert werden können, die im pflanzenbaulichen Kontext sinnvoll interpretierbar sind. Die Anwendbarkeit der Methode kann somit bestätigt werden.

Die Ergebnisse dieser Arbeit und insbesondere der angepasste Prozess der Wissensentdeckung in Datenbanken für Precision Farming könnten grundsätzlich für alle Arten der teilflächenspezifischen Betriebsmitteloptimierung eingesetzt werden. Eine entsprechende Weiterentwicklung des Versuchsdesigns und der Präzision der Datenerfassung und Applikation würde zu weiteren Verbesserungen führen. Für die teilflächenspezifische Stickstoffausbringung gilt es, die gefundenen Entscheidungsregeln mehrjährig zu testen, um damit abgesicherte Aussagen zu erhalten.

7 Literatur

- ADAMS, M. L., COOK, S. E. (2000): On-Farm Experimentation: Application of Different Analytical Techniques for Interpretation. In: Robert, P. C. et al. (Hrsg.): Proceedings of the 5th International Conference on Precision Agriculture. Madison, WI, USA. [CD-ROM]. s.p.
- AUERNHAMMER, H. (1999): Precision Farming for Site-Specific Fertilization. In: Zeitschrift für Agrarinformatik, Münster-Hiltrup, H3. S. 58-66.
- AUERNHAMMER, H. (2004): Precision Crop Farming. In: Matthies, H. J., Meier, F. (Hrsg.): Jahrbuch Agrartechnik 2004. Landwirtschaftsverlag Münster. S. 31-38.
- AUERNHAMMER, H., DEMMEL, M., MAIDL, F. X., SCHMIDHALTER, U., SCHNEIDER, T., WAGNER, P. (1999): An On-Farm Communication System for Precision Farming with Nitrogen Real-Time Application. Presented at the 1999 ASAE/CSAE-SCGr Annual International Meeting, Toronto, Canada. ASAE Paper, No. 99 11 50.
- AUERNHAMMER, H., DEMMEL, M., MUHR, T., ROTTMEIER, J., ET AL. (1994): GPS for yield mapping on combines. In: Computers and Electronics in Agriculture, Amsterdam: Elsevier, Vol. 11, No. 1. S. 53-68.
- AUERNHAMMER, H., DEMMEL, M., MUHR, T., ROTTMEIER, J., WILD, K. (1993): Yield Measurement on combine harvesters. An ASAE Meeting Presentation, Paper No. 93-1506, 1993, Chicago, Illin./USA. S. 1-8.
- AUFHAMMER, W. (1998): Getreide- und andere Körnerfruchtarten. Verlag Eugen Ulmer, Stuttgart.
- BACHMAIER, M., AUERNHAMMER, H. (2004): A method for correcting raw yield data by fitting paraboloid cones. In: AgEng Leuven 2004. Book of abstracts. Leuven 2004, part 1. S. 374-378.
- BACHMAIER, M., AUERNHAMMER, H. (2005): Yield mapping based on robust fitting paraboloid cones in butterfly and elliptic neighborhoods. In: Stafford, J.V. (Hrsg.): Precision Agriculture '05. Proceedings of the ECPA Conference 2005, Uppsala, Schweden. S. 741-750.
- BERGLUND, S., BUICK, R. (2005): Guidance and automated steering drive resurgence in precision farming. In: Stafford, J.V. (Hrsg.): Precision Agriculture '05. Proceedings of the ECPA Conference 2005. Uppsala, Schweden. S. 39-45.
- BERRY, J. A., LINOFF, G. S. (2000): Mastering Data Mining – The Art and Science of Customer Relationship Management. Wiley, New York.
- BIGUS, J. P. (1996): Data mining with neural networks: Solving business problems from application development to decision support. McGraw-Hill, Inc. Hightstown, NJ, USA.
- BISHOP, C.M. (1995): Neural Networks for Pattern Recognition. Oxford University Press, New York.

- BLACKMORE, S., GODWIN, R. J., FOUNTAS, S. (2003): The Analysis of Spatial and Temporal Trends in Yield Map Data over Six Years. *Biosystems Engineering* 84 (4), S. 455-466.
- BRACHMAN, R. J., ANAND, T. (1996): The Process of Knowledge Discovery in Databases. In: Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R. (Hrsg.): *Advances in Knowledge Discovery and Data Mining*, Menlo Park. S. 37-57.
- BREIMAN, L., FRIEDMAN, J. H., OLSHEN, R. A., STONE, CH. (1984): *Classification and Regression trees*. Wadsworth International Group, Belmont, California.
- BREVIK, E. C., FENTON, TH. E., HORTON, R. (2004): Effekt of Daily Soil Temperature Fluctuations on Soil Electrical Conductivity as Measured with the Geonics® EM-38. *Precision Agriculture*, 5. S. 145-152.
- BULLOCK, D. S., LOWENBERG-DEBOER, J., SWINTON, S. M. (2002): Adding value to spatially managed inputs by understanding site-specific yield response. *Agricultural Economics* 27 (2002). S. 233-245.
- BULLOCK, D. S.; BULLOCK, D. G. (2000): From Agronomic Research to Farm Management Guidelines: A Primer on the Economics of Information and Precision Technology. *Precision Agriculture* (2) 2000. S. 71-101.
- CALLAN, R. (2002): *Neuronale Netze im Klartext*. Pearson Studium, Prentice-Hall, New Jersey.
- CHAUDAHARY, A., COLVIN, T. S., JAYNES, D. B., KASPAR, T. C. (2004): Development response curves using field scale data and data mining. In: Mulla, D.J. (Hrsg.): *Proceedings of the 7th International Conference on Precision Agriculture*, July 25-28, 2004. Precision Agriculture Center, University of Minnesota, St. Paul, MN. [CDROM]. s.p.
- CHRISTENSEN, W. F., AND COOK, D. (1999). Data mining soil characteristics affecting corn yield. *Iowa State University Statistical Laboratory Preprint Series*, #99-06.
- COLVIN, T. S., JAYNES, D. B., KASPAR, T. C., JAMES, D.E., MEEK, D. W., COOK, J. D. (2000): Yield Certainty with Plots or Fields. 5th International Conference On Precision Agriculture Abstracts & Proceedings. St. Paul, Minnesota. [CD-ROM]. s.p.
- CORVIN, D. L., LESCH, S. M. (2003): Application of Soil Electrical Conductivity to Precision Agriculture: Theory, Principles, and Guidelines. *Agronomy Journal* 95 (2003), S. 455-471.
- DABBERT, S., KILIAN, B. (2002): Ökonomie. In: *Precision Agriculture – Herausforderungen an integrative Forschung, Entwicklung und Anwendung in der Praxis*. Tagungsband Precision Agriculture Tage 13.-15. März in Bonn. KTBL-Sonderveröffentlichung 038. Hrsg.: Kuratorium für Technik und Bauwesen in der Landwirtschaft e. V. (KTBL), Darmstadt. S. 423-446.
- DABERKOW, S. G., MCBRIDE, W. T. (2000): Adoption of Precision Agriculture Technologies by U.S. Farmers. 5th international Conference on Precision Agriculture. CD-ROM. Bloomington, MN. s. p.

- DIEPOLDER, M. (1994): Untersuchungen zur Ableitung von Richtlinien für die Optimierung der N-Düngung zu Winterweizen. Dissertation, Technische Universität München-Weihenstephan.
- DIKER, K., HEERMANN, D. F., BUCHLEITER, G. W., BAUSCH, W. C., FARAHANI, H., WESTFALL, D. (2004): Modelling Yield by Neural Network. In: Mulla, D. J., (Hrsg.): Proceedings of the 7th International Conference on Precision Agriculture, July 25-28, 2004. Precision Agriculture Center, University of Minnesota, St. Paul, MN. [CDROM]. s.p.
- DOBERMANN, A., BLACKMORE, S., COOK, S. E., ADAMCHUK, V. I. (2004): Precision Farming: Challenges and Future Directions. In: Proceedings of the 4th International Crop Science Congress 2004, Brisbane, Australia. [CD-ROM]. Web site www.cropscience.org.au.
- DRUMMOND, S. T., SUDDUTH, K. A. (2004): Analysis of Errors Affecting Yield Map Accuracy. In: Mulla, D. J. (Hrsg.): Proceedings of the 7th International Conference on Precision Agriculture, July 25-28, 2004. Precision Agriculture Center, University of Minnesota, St. Paul, MN. [CD-ROM]. s.p.
- DRUMMOND, S. T., SUDDUTH, K. A., JOSHI, A., BIRELL, S. J., KITCHEN, N. R. (2003): Statistical and Neural Methods for Site-Specific Yield Prediction. Transactions of the ASAE Vol. 46(1). S. 5-14.
- DRUMMOND, S. T., SUDDUTH, K. A., KITCHEN, N. R., BATCHELOR, W. D., BOLLERO, G. A., BULLOCK, D. G., CLAY, D. E., PALM, H. L., WIEBOLD, W. J., PIERCE, F. J., SCHULER, R. T., THELEN, K. D. (2002): Neural Network Analysis of Site-Specific Soil, Landscape and Yield Data. In: Robert, P. C. et al. (Hrsg.): Proceedings of the 6th Int. Conference on Precision Agriculture. Madison, WI, USA. [CD-ROM]. s.p.
- DRUMMOND, S.T., SUDDUTH, K. A., JOSHI, A. (2000): Predictive Ability of Neural Networks for Site-Specific Yield Estimation, Paper presented at 2nd Intern. Geospatial Information in Agriculture and Forestry Conference, Lake Buena Vista, Florida, 2000, <http://www.fse.missouri.edu/mpac/pubs/drummond.pdf>, Stand 13.01.2003. s. p.
- DURLESSER, H. (1999): Bestimmung der Variation bodenphysikalischer Parameter in Raum und Zeit mit elektromagnetischen Induktionsverfahren. FAM-Bericht 35. Shaker Verlag Aachen 1999.
- EARL, R., TAYLOR, J. C., WOOD, G. A., BRADLEY, I., JAMES, I. T., WAINE, T., WELSH, J. P., GODWIN, R. J., KNIGHT, S. M. (2003): Soil Factors and their Influence on Within-field Crop Variability, Part I: Field Observation of Soil Variation. Biosystems Engineering 84(4). S. 425-440.
- EBERTSEDER, TH.; GUTSER, R.; HEGE, U.; BRANDHUBER, R.; SCHMIDHALTER, U. (2003): Strategies for site-specific nitrogen fertilization with respect to long-term environmental demands. In: Stafford, J.; Werner, A. (Hrsg.) Precision Agriculture Wageningen Academic Publishers Wageningen. S. 193-198.

- ECKERT, N. (1998): Unternehmensentwicklung und Ökologie des Wissens - Der Primat impliziten Wissens und seine Implikationen für eine Evolutionäre Organisations-theorie. Verlag Barbara Kirsch, München.
- EHLERT, D., DAMMER, K.-H., ADAMEK, R. (2004): Crop-Meter for site-specific Farming. In: Mulla, D. J. (Hrsg.): Proceedings of the 7th International Conference on Precision Agriculture. July 25-28, 2004, Minneapolis, Minnesota. [CD-ROM]. s.p.
- ELDER, J. F., PRGIBON, D. (1996): A Statistical Perspective on Knowledge Discovery in Databases. In: Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R. (Hrsg.): Advances in Knowledge Discovery and Data Mining, Menlo Park. S. 83-113.
- ENGEL, TH. (1991): Entwicklung und Validierung eines Simulationsmodells zur Stickstoffdynamik in Boden und Pflanze mit Hilfe objektorientierter Programmierung. Dissertation, TU-München-Weihenstephan.
- ESTER, M., SANDER, J. (2000): Knowledge Discovery in Databases: Techniken und Anwendungen, Springer Verlag.
- FAYYAD, U., PIATETSKY-SHAPIRO G., SMYTH, P. (1996): From Data Mining to Knowledge Discovery in Databases. AI Magazine 17(3) 1996. S. 37-54.
- FAYYAD, U., PIATETSKY-SHAPIRO, G., SMYTH, P. (1996): From Data Mining to Knowledge Discovery. In: Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R. (Hrsg.): Advances in Knowledge Discovery and Data Mining, Menlo Park. S. 1-37.
- FINCK, CH. (2004): Cut Through the Clutter. Farm Journal, 128 (12). S. 14-19.
- GANDORFER, M. (2005): Bewertung von Precision Farming dargestellt am Beispiel der teilflächenspezifischen Stickstoffdüngung. Dissertation, TU München-Weihenstephan. Im Druck.
- GANDORFER, M., MEYER-AURICH, A., HEIBENHUBER, A. (2004a): Der Einsatz moderner Informationstechnologien im Pflanzenbau - ein Beispiel aus der teilflächenspezifischen Stickstoffdüngung. Zeitschrift für Agrarinformatik, Heft 2/ 2004. S. 27-33.
- GANDORFER, M., WEIGERT, G., AMON, H., STICKSEL, E., HEIBENHUBER, A., WAGNER, P. (2004b): Ökonomische Analyse von Precision Farming Technologien im pflanzenbaulichen Versuchswesen. Referate der 25. GIL – Jahrestagung in Bonn 2004. Berichte der Gesellschaft für informatik in der Land-, Forst- und Ernährungswirtschaft. Band 17, Bonn 2004.
- GEHRA, B. (2005): Früherkennung mit Business Intelligence Technologien – Anwendung und Wirtschaftlichkeit der Nutzung operativer Datenbestände. Dissertation LMU München.
- GENTSCH, P. NIEMANN, CH., ROTH, M. (2000): Data Mining – 12 Softwarelösungen im Vergleich. Studie des Business Application Research Center. Oxygen Verlag, Feldkirchen.

- GODWIN, R. J., MILLER, P. C. H. (2003): A Review of the Technologies for Mapping Within-field Variability. *Biosystems Engineering* 84(4). S. 393-407.
- GODWIN, R. J., RICHARDS, T. E., WOOD, G. A., WELSH, J. P., KNIGHT, S. M. (2003): An Economic Analysis of the Potential for Precision Farming in UK Cereal Production. *Biosystems Engineering*. 84(4). S. 533-545.
- GOMEZ, K. A., GOMEZ, A. A. (1984): Statistical procedures for agricultural research. International Rice Research Institute. Second Edition. John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore.
- GRIFFIN, T. W., LOWENBERG-DEBOER, J., LAMBERT, D. M., PEONE, J., PAYNE, T., DABERKOW S. G. (2004): Adoption, Profitability, and Making Better Use of Precision Farming Data. Staff Paper #04-06. Department of Agricultural Economics, Purdue University.
- GURNEY, K. (1997): An Introduction to Neural Networks. Routledge, London.
- GUTSER, R. (2005): Zur Stickstoffverfügbarkeit organischer Dünger. SuB Heft 3-4/05. S. 5-10.
- GUYOT, G., BARET, F. UND MAJOR, D. J. (1988): High spectral resolution: determination of spectral shifts between the red and near infrared. *Int. Archives of Photogrammetry and Remote Sensing*. 11. S. 750-760.
- HAN, J., KAMBER, M. (2001): Data Mining – Concepts and Techniques, Morgan Kaufmann Publishers. San Francisco, USA.
- HAND, D., MANNILA, H., SMYTH, P. (2001): Principles of Data Mining. MIT Press, Cambridge, Massachusetts, USA.
- HEGE, U., MAIDL, F.-X., DENNERT, J., LIEBLER, J., OFFENBERGER, K. (2002): Düngestrategien für Stickstoff zu Winterweizen: Ein Vergleich von Simulationsmodellen und Düngeberatungssystemen. *Pflanzenbauwissenschaften*, 2002(1). S. 26-35.
- HEGE, U.; OFFENBERGER, K. (2005): Am Anfang wenig. Optimale N-Verteilung bei Winterweizen. *Bayrisches Landwirtschaftliches Wochenblatt*, Heft 5, 2005.
- HEYLAND, K.-U. (1996): Spezieller Pflanzenbau. Ulmer Verlag, Stuttgart.
- HUBERT, M., ROUSSEEUW, P. J., VAN AELST, S. (2005): Multivariate Outlier Detection and Robustness. In: Rao, C.R., Wegman, E.J., Solka, J.L. (Hrsg.) 2005. *Data Mining and Data Visualization. Handbook of Statistics 24*, Elsevier, Amsterdam. S. 263-302.
- HURLEY, T. M., KILIAN, B., MALZER, G., DIKICI, H. (2001): The value of information for variable rate nitrogen applications: A comparison of soil test, topographical and remote sensing information [Online]. <http://www.amesremote.com/research.htm> (abgerufen am 7.Nov 2005). Univ. of Minnesota, St. Paul.

- JAFAR-SHAGHAGHI, F. (1996): Maschinelles Lernen, Neuronale Netze und Statistische Lernverfahren zur Klassifikation und Prognose – Theoretische Analyse und ökonomische Anwendung. Shaker Verlag Aachen.
- JAPKOWICZ, N., STEPHEN, S. (2002): The class imbalance problem: A systematic study. *Intelligent Data Analysis*, 6(2002). S. 429-449.
- KAMATH, C. (2001): On Mining Scientific Datasets. In: *Data Mining for Scientific and engineering applications*. Grossman, R. L., Kamath, C. Kegelmeyer, Ph., Kumar, V. Raju, R. (Hrsg.). Kluwer Academic Publishers, Dordrecht. 2001. S. 1-22.
- KANTARDZIC, M. (2003): *Data Mining – Concepts, Models, Methods, and Algorithms*. IEEE-Press, Wiley-Interscience. New York, USA.
- KHAKURAL, B. R., ROBERT, P. C., HUGGINS, D. R. (1999): Variability of corn/soybean yield and soil/landscape properties across a southwestern Minnesota landscape. In: Robert, P.C. Rust, R.H., Larson, W.E. (Hrsg.) *Proceedings of the 4th Conf. on Precision Agriculture*, Madison, WI, USA. S. 573-579.
- KILIAN, B., MALZER, G. (2001): The Value for Variable Rate Nitrogen Applications: A Comparison of Soil Test, Topographical, and Remote Sensing Information, Selected Paper American Agricultural Economics Association Annual Meeting, Chicago, IL, 2001 http://agecon.lib.umn.edu/cgi-bin/pdf_view.pl?paperid=2722, Stand: 13.01.2003.
- KLÖSGEN, W., ZYTKOW, J. M. (2001): The Knowledge Discovery Process. In: Klösgen, W. und Zytkow, J.M. (Hrsg.) *The Handbook of Data Mining*, Oxford University Press, Oxford 2001. S. 10-21.
- KOLLIG, M. (1993): Entscheidungshilfen für den Zuckerrübenanbau – Entscheidungsunterstützungsmodule für den Zuckerrübenanbau mittels neuronaler Netze und regelbasierter Systeme. *Agrarwirtschaft, Sonderheft 138*.
- KRATZER, K. P. (1993): *Neuronale Netze – Grundlagen und Anwendungen*. Hanser Verlag, München.
- KRAVCHENKO, A. N., BULLOCK, D. G. (2000): Correlation of corn and soybean grain yield with topography and soil properties. *Agronomy. Journal*. 92(1). S. 75–83.
- LAMB, J. A., DOWDY, R. H., ANDERSON, J. L., REHM, G. W. (1997): Spatial and temporal stability of corn grain yield. *J. Prod. Agric.* 10. S. 410-414.
- LAMBERT, D.; LOWENBERG-DEBOER, J. (2000): *Precision Agriculture Profitability Review*. Staff Paper. Site Specific Management Center, Purdue University, West Lafayette, Indiana. URL: <http://www.agriculture.purdue.edu/ssmc/>.
- LAPEN, D. R., HAYHOE, H. N., TOPP, G. C., MCLAUGHLIN, N. B., GREGORICH, E. G., CURNOE, W. E. (2002): Measurements of Mouldboard Plow Draft: II. Draft-Soil-Crop and Yield-Draft Associations. *Precision Agriculture*, 3(2002). S. 237-257.

- LIEBLER, J. (2003): Feldspektrometrische Messungen zur Ermittlung des Stickstoffstatus von Winterweizen und Mais auf heterogenen Schlägen. Dissertation, TU München.
- LIEBLER, J., STICKSEL, E., MAIDL, F.-X. (2001): Field Spectroscopic Measurements to Characterise Nitrogen Status and Dry Matter Production of Winter Wheat. In: Sevilla, F. (Hrsg.): Proceeding of the Third ECPA Conference, Montpellier, 2001. S. 935-938.
- LINK A., JASPER, J. (2003): Site-specific N fertilization based on remote sensing – is it necessary to take yield variability into account? In: Stafford, J., Werner A. (Hrsg.): Precision Agriculture. Proceedings of the 4th European Conference on Precision Agriculture. 14.-19. Juni, Berlin, Deutschland. S.353-360.
- LINK, A., PANITZKI, M. und REUSCH, S. (2002): Hydro N-Sensor: Tractor-Mounted Sensing für Variable N Fertilization. In: Robert, P. C. et al. (Hrsg): Proceedings of the 6th Int. Conference on Precision Agriculture. Madison, WI, USA. [CD-ROM]. s.p.
- LINSEISEN, H. (2003): Entwicklung eines Management-Informationssystems für Entscheidungen im Precision Farming. Dissertation TU München-Weihenstephan.
- LISSO, H. (2003): GPS-gestützte Teilflächenbewirtschaftung. RKL-Schrift 4.1.0. S. 1123-1143.
- LIU, J. GOERING, C.E., TIAN, L. (2001): A Neural Network for Setting Target Corn Yields. In: Transactions of the ASAE, Vol 44(3). S. 705-713.
- LOWENBERG-DEBOER, J., BOEHLJE, M.. (1996): Revolution, evolution or dead-end: economic perspectives on precision agriculture. P.C. Robert, R.H. Rust and W.E. Larson (Hrsg.): Proceedings of the 3rd International Conference on Precision Agriculture, June 23-26, 1996, Minneapolis, Minn., USA. S. 923-944.
- MAIDL, F.-X., SCHÄCHTL, J., HUBER, G. (2004): Strategies for Site-Specific Nitrogen Fertilization on Winter Wheat. In: D.J. Mulla (Hrsg.): Proceedings of the 7th International Conference on Precision Agriculture. July 25-28, 2004, Minneapolis, Minnesota. [CD-Rom]. s.p.
- MAKOWSKI, D., WALLACH, D. (2001): How to improve model-based decision rules for nitrogen fertilization. European Journal of Agronomy 15 (2001). S. 197-208.
- MISTELE, B. (2005): Tractor based spectral reflectance measurements using an oligo view optic to detect biomass, nitrogen content and nitrogen uptake of wheat and maize and the nitrogen nutrition index of wheat. Dissertation TU München-Weihenstephan. Im Druck.
- MISTELE, B. GUTSER, R., SCHMIDHALTER, U. (2004): Validation of field-scaled spectral measurements of the nitrogen status in winter wheat. In: Mulla, D. J. (Hrsg.): Proceedings of the 7th International Conference on Precision Agriculture, July 25-28, 2004. Precision Agriculture Center, University of Minnesota, St. Paul, MN. [CD-ROM]. s.p.

- NAKHAIEZADEH, G. (1998): Wissensentdeckung in Datenbanken und Data Mining: Ein Überblick. In: G. Nakhaeizadeh (Hrsg.): Data Mining: Theoretische Aspekte und Anwendungen. Physika-Verlag, Heidelberg, Germany. S. 1-28.
- NOACK, P. H., MUHR, T., DEMMEL, M. (2003): An Algorithm for Automatic Detection and Elimination of Defective Yield Data. In: J.V. Stafford and A. W. Erdner (Hrsg.): Proceedings of the 4th European Conference on Precision Agriculture, Wageningen Academic Publishers, Netherlands. S. 445-450.
- OBERHOFER, W. ZIMMERER, TH. (1996): Wie Künstliche Neuronale Netze lernen: Ein Blick in die Black Box der Backpropagation Netzwerke. Regensburger Diskussionsbeiträge Nr. 292, Universität Regensburg, Wirtschaftswissenschaftliche Fakultät.
- OLFS, H-W., BLANKENAU, K. BRENTRUP, F., JASPER, J., LINK, A., LAMMEL, J. (2005): Soil and Plant-Based Nitrogen-Fertilizer Recommendations in Arable Farming. Journal of Plant Nutrition and Soil Science 2005(168). S. 414-431.
- OLSON, D. L. (2004): Data Set Balancing. In: Shi, Y., Xu, W., Chen, Z. (Hrsg.). Data Mining and Knowledge Management (Chinese Academy of Sciences Symposium). Springer Verlag. S. 71-80.
- OSTERMEIER, R., AUERNHAMMER, H. (2004): Real-Time Process Control for a Sensor Based Fertilizer Application System Using Multisensor Data Fusion AGENG LEUVEN 2004, Engineering the future, Leuven (Belgium). [CD-ROM]. s.p.
- OSTERMEIER, R., AUERNHAMMER, H., DEMMEL, M., (2003): Development of an In-Field Controller for an Agricultural Bus-System Based On Open Source Program Library LBS-Lib. In: Stafford, J., Werner, A. (Hrsg.): Precision Agriculture. Wageningen Academic Publishers Wageningen. S. 515-520.
- PANELL, D. J. (2004): Flat-earth exconomics: The far-reaching consequences of flat payoff functions in economic decision making. URL: <http://www.general.uwa.edu.au/u/dpannell/dp0402.htm>.
- PATTERSON, D. (1996): Künstliche neuronale Netze – Das Lehrbuch. Prentice Hall, München.
- PETERS, M. W. ET AL. (1999): Nitrogen Management Strategies for Precision Farming. IN: STAFFORD, J. V.: Proceedings of the 2nd European Conference on Precision Agriculture '99, Odensee. S. 719-728.
- PIEKIELEK, W. P., FOX, R. H., TOTH, J. D., MACNEAL, K. E. (1995): Use of a Chlorophyll Meter at the Early Dent Stage of Corn to Evaluate Nitrogen Sufficiency. Agron. J. 87. S. 403-408.
- POKRAJAC D., OBRADOVIC Z. (2001a): Neural Network-Based Software for Fertilizer Optimization in Precision Farming. In: Proc. International Joint Conference on Neural Networks, Vol. 3, 2001. S. 2110-2115.

- POKRAJAC, D., OBRADOVIC, Z. (2001b): A Neural Network-Based Method for Site-Specific Fertilization Recommendation. Written for presentation at the 2001 ASAE Annual International Meeting in Sacramento, California, USA. 30. Juli – 1. August, 2001. Paper Number: 01-311. S. 1-13.
- PRESS, S. J. (2004): The Role of Bayesian and Frequentist Multivariate Modeling in Statistical Data Mining. In: Bozdogan, H. (Hrsg.). Statistical Data Mining and Knowledge Discovery. Chapman & Hall/CRC, London. S. 1-13.
- PRINGLE, M. J., COOK, S. E., MCBRATNEY, A. B. (2004a): Field-Scale Experiments for Site-Specific Crop Management. Part I: Design Considerations. Journal for Precision Agriculture 2004(5). S. 617-624.
- PRINGLE, M. J., COOK, S. E., MCBRATNEY, A. B. (2004b). Field-Scale Experiments for Site-Specific Crop Management. Part II: A Geostatistical Analysis. Journal for Precision Agriculture 2004(5). S. 625-645.
- PYLE, D. (1999): Data Preparation for Data Mining. Morgan Kaufmann Publishers, Inc. San Francisco, California, USA.
- PYLE, D. (2003): Data Collection, Preparation, Quality and Visualization. In: Nong, Y. (Hrsg.) 2003. The Handbook of Data Mining. Lawrence Erlbaum Associate, London. S. 365-393.
- REINER, L., BUHLMANN, V., GRASER, S., HEIBENHUBER, A., KLASSEN, M., PFEFFERKORN, V., SPANAKAKIS, A., STRAB, F. U. A. (1992): Weizen aktuell. DLG-Verlag, Frankfurt am Main.
- REUSCH, S. (1997): Entwicklung eines reflexionsoptischen Sensors zur Erfassung der Stickstoffversorgung landwirtschaftlicher Kulturpflanzen. Dissertation, Christian-Albrechts-Universität Kiel.
- ROBERTS, R. K., ENGLISH, B. C., MAHAJANASHETTI, S. B. (2000): Evaluating the Returns to Variable Rate Nitrogen Application. Journal of Agricultural and Applied Economics. 32(1) April 2000. S. 133-143.
- ROTHMUND, M., AUERNHAMMER, H. (2004): A web based information management system for process data designed with open source tools. AGENG LEUVEN 2004, Engineering the future, Leuven (Belgium), [CD-ROM]. S. 1-8.
- ROTHMUND, M., ZIPPRICH, M., AUERNHAMMER, H., DEMMEL, M. (2003): Zugkraftmessung bei der Bodenbearbeitung als ergänzende Information zur Standortbeschreibung. In: VDI -MEG Tagung Landtechnik 2003. Hannover 7./8. Nov. 2003. S. 305-310.
- RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. (1986): Learning Internal Representations by Error propagation. Parallel Distributed Processing 1. S. 318-362.
- SÄUBERLICH, F. (2000): KDD und Data Mining als Hilfsmittel zur Entscheidungsunterstützung. Entscheidungsunterstützung für ökonomische Probleme 18, Peter Lang Verlag, Frankfurt am Main.

- SCHÄCHTL, J. (2004): Sensorgestützte Bonitur von Aufwuchs und Stickstoffversorgung bei Weizen – und Kartoffelbeständen. Dissertation, TU München.
- SCHARPF, H. C., WEHRMANN, J. (1975): Die Bedeutung des Mineralstickstoffvorrates des Bodens zu Vegetationsbeginn für die Bemessung der N-Düngung zu Winterweizen. *Landw. Forsch.* 32. S. 100-114.
- SCHERER, A. (1997): Neuronale Netze – Grundlagen und Anwendungen. Vieweg, Braunschweig.
- SCHMIDHALTER, U. JUNGERT, S BREDEMEIER, C. GUTSER, R. MANHART, R. MISTELE B. GERL, G. (2003): Field-scale validation of a tractor based crop scanner to determine biomass and nitrogen uptake of winter wheat. ECPA, Berlin: 15. – 19. 6. 2003, Precision Agriculture: Papers from the 4th European Conference on Precision Agriculture. S. 615-619.
- SCHMITTER, E. D. (1991): Neuronale Netze – Einführung, Programmierbeispiele, praktische Anwendungen. Hofacker Verlag, Holzkirchen.
- SCHNEIDER, M. (2005): Ökonomische Potenziale von Precision Farming. Diplomarbeit an der Martin-Luther-Universität Halle Wittenberg.
- SCHOMMER, CH. (2003): Anwendung von Data Mining. Shaker Verlag, Aachen.
- SCOTT, D. W., SAIN, S. R. (2005): Multidimensional Density Estimation. In: Rao, C. R., Wegman, E. J., Solka, J. L. (Hrsg.): *Data Mining and Data Visualization. Handbook of Statistics* 24. Elsevier, Amsterdam. S. 229-258.
- SHATAR, T. M., MCBRATNEY, A. B. (1999): Empirical Modelling of Relationships between Sorghum Yield and Soil Properties. *Precision Agriculture* 1999(1). S. 249-276.
- SHEARER, S. A., THOMASSON, J. A., BURKS, T. F., MUELLER, T., FULTON, J. P, GILBERT, D. HIGGINS. S. G (1999): Yield Prediction Using a Neural Network Classifier Trained Using Soil Landscape Features And Soil Fertility Data. ASAE Paper No. 993041. St. Joseph, Mich. ASAE.
- SHEKHAR, S., HUANG, Y., WU, W., LU, C., CHAWLA, S. (2001): What's spatial about spatial data mining: Three case studies. In: Grossman, R. L., Kamath, Ch., Kegelmeyer, Ph. Kumar, V., Namburu, R. (Hrsg): *Data Mining for scientific and engineering applications.* Kluwer Academic Publishers, Dordrecht.
- SPAULDING, A. D., SAGHAIAN, S. H., DILLON, C., MUELLER, T., SCHEARER, S. (2004): The Value of Soil Electrical Conductivity and Topographical Information for Variable Rate Nitrogen Application: First Assessment. In: Mulla, D. J. (Hrsg.): *Proceedings of the 7th International Conference on Precision Agriculture, July 25-28, 2004.* Precision Agriculture Center, University of Minnesota, St. Paul, MN. [CD-ROM]. s.p.
- SPSS (2003): *Clementine 8.0 – User's Guide.* SPSS. Chicago, IL, USA.
- STEINHAUSER, H., LANGBEHN, C., PETERS, U. (1982): Einführung in die landwirtschaftliche Betriebslehre – Allgemeiner Teil. Verlag Eugen Ulmer, Stuttgart.

- STEINMAYR, TH. (2002): Fehleranalyse und Fehlerkorrektur bei der lokalen Ertragsermittlung im Mähdröschler zur Ableitung eines standardisierten Algorithmus für die Ertragskartierung. Dissertation, TU-München-Weihenstephan.
- STURM, H., BUCHNER, A. ZERULLA, W. (1994): Gezielter düngen. VerlagsUnion Agrar, Frankfurt (Main).
- SUDDUTH, K. A., DRUMMOND, S. T., KITCHEN, N. R. (2001): Accuracy issues in electromagnetic induction sensing of soil electrical conductivity for precision agriculture. *Computers and electronics in agriculture*. 31(2001). S. 239-264.
- SUDDUTH, K. A., FRAISSE, C. W., DRUMMOND, S. T., KITCHEN, N. R. (1998): Integrating Spatial Data Collection, Modelling and Analysis for Precision Agriculture. Presented at the first International Conference on Geospatial Information in Agriculture and Forestry, Lake Buena Vista, Florida, 1998. s. p.
- SUTTON, C. D. (2005): Classification and Regression Trees, Bagging, and Boosting. In: Rao, C. R., Wegman, E.J., Solka, J.L. (Hrsg.): *Data Mining and Data Visualization*. Handbook of Statistics 24, Elsevier, Amsterdam. S. 303-329.
- SWINTON, S. M., JONES, K. Q. (1998): From Data to Information: The value of sampling vs, Sensing soil data. Staff Paper 98-15. Michigan State University.
- SWINTON, S. M., LIU, Y. (2002): Is Yield Response Site-Specific? – Revisiting Nitrogen Recommendations on Corn. Selected Paper, American Agricultural Economics Association annual meeting, Long Beach, CA, 2002. S. 1-25.
- SYDSAETER, K., STROM, A. BERCK, P. (2000): *Economists' Mathematical Manual*. Third Edition. Springer-Verlag, Berlin.
- THYLEN, L., ALGERBO, P. A., GIEBEL, A. (2001): An Expert Filter Removing Erroneous Yield Data. In: Robert, P. C., Rust, R. H., Larsen, W. E. (Hrsg.): *Precision Agriculture, Proceedings of the 5th International Conference, ASA, CSSA and SSSA, Madison, WI, USA*. [CD-ROM]. s.p.
- TOTTMANN, D.R. (1987): The decimal code for the growth stage of cereals, with illustrations. *Ann. Appl. Biol.* 110. S. 441-454.
- UTHURUSAMY, R. (1996): From Data Mining to Knowledge Discovery: Current Challenges and Future Directions. In: Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R. (Hrsg.): *Advances in Knowledge Discovery and Data Mining*, Menlo Park. S. 561-569.
- WAGNER, P. (1999): Produktionsfunktionen und Precision Farming. In: *Zukunftsorientierte Betriebswirtschaft und Informationstechnologien in der Agrarwirtschaft*. Gießener Schriften zur Agrar- und Ernährungswirtschaft, Heft 29, Frankfurt. S. 39-66.
- WAGNER, P. (2004): Wirtschaftlichkeit von Precision Farming – Methoden und Möglichkeiten der Wirtschaftlichkeitsprüfung im landwirtschaftlichen Unternehmen. In: *Precision Farming, Analyse, Planung, Umsetzung in der Praxis*. KTBL-Schrift 419.

- WAGNER, P., SCHNEIDER, M., WEIGERT, G. (2005): Einsatz künstlicher Neuronaler Netze zur automatischen Ableitung von Entscheidungsregeln für die N-Düngung in Winterweizen. In: Gesellschaft für Agrarinformatik in der Land-, Forst- und Ernährungswirtschaft (Hrsg.): Zeitschrift für Agrarinformatik, 4/2005.
- WATKINS, D. (1997): Clementine's Neural Networks Technical Overview. www.cs.bris.ac.uk. Abgerufen 09/2005.
- WEGMAN, E. J. SOLKA, J. L. (2005): Statistical Data Mining. In: Rao, C. R., Wegman, E. J., Solka, J. L. (Hrsg.) 2005. Data Mining and Data Visualization. Handbook of statistics 24. Elsevier, Amsterdam. S. 1-44.
- WEIGERT, G. UND WAGNER, P. (2004): Site-Specific Yield Prediction for Varying N Rates by the Application of Data Mining Techniques. In: Mulla, D.J. (Hrsg.): Proceedings of the 7th International Conference on Precision Agriculture. July 25-28, 2004, Minneapolis, Minnesota. [CD-ROM]. s.p.
- WEIGERT, G., GANDORFER, M., MEYER-AURICH, A., HEIBENHUBER, A., AND P. WAGNER (2004): Evaluation of Decision Rules for Site Specific N-Fertilization in Strip Trials. In: Vlachopoulou, M., Manthou, V., Illiadis, L., Gertsis, S., Salampasis, M. (Hrsg.): International Conference on Information Systems & Innovative Technologies in Agriculture, Food and Environment; Thessaloniki (Greece). Proceedings Volume II. S. 320-323.
- WEISS, S. M., INDURKHYA, N. (1998): Predictive Data Mining – A Practical Guide. Morgan Kaufmann Publishers, Inc., San Francisco, USA.
- WELTZIEN, C., v. CHAPPUIS, A., KROMER, K-H., PERSSON, K., RESNIK, B., SCHMITTMANN, O. (2002): Technikbetreuung und -vergleich. In: KTBL (Hrsg.): Precision Agriculture – Herausforderung an integrative Forschung, Entwicklung und Anwendung in der Praxis. KTBL-Sonderveröffentlichung 38. S. 153-168.
- WENDROTH, O., REUTER, H. I., KERSEBAUM, K. C., (2003): Predicting Yield of barley across a landscape: a state-space modelling approach. *Journal of Hydrology* 272. S. 250-263.
- WENKEL, K.-O., BROZIO, S., GEBBERS, R. I. B. (2002): Methods and Algorithms for Site-Specific Nitrogen Fertilization of several crops and N-Fertilization Strategies In: Robert, P. C. et al. (Hrsg): Proceedings of the 6th Int. Conference on Precision Agriculture. Madison, WI, USA. [CD-ROM]. s.p.
- WESTPHAL, CH., BLAXTON, T. (1998): Data Mining Solutions – Methods and Tools for Solving Real-World Problems. Wiley-Verlag, New York.
- WHELAN, B. M., MCBRATNEY, A. B. (2000): The “Null Hypothesis” of Precision Agriculture Management. In: Precision Agriculture 2000(2). S. 265-279.
- WHIPKER, L. D., AKRIDGE, J.T. (2005): Agricultural Services Dealership Suvey Results. Staff Paper # 05-11. September 2005. Department of Agricultural Economics. Purdue University.

- WIDMANN, G. (2000): Künstliche neuronale Netze und ihre Beziehungen zur Statistik. Dissertation. Peter Lang, Frankfurt am Main.
- WIDROW, B., HOFF, M.E. (1960): Adaptive Switching Circuits. In: Anderson, J.A. und Rosenfeld, E. (Hrsg.). Neurocomputing: Foundations of Research, MIT Press, 1988. S. 126-134.
- WIEDERHOLD, G. (1996) Foreword: On the Barriers and Future of Knowledge Discovery. In: In: Fayyad, U., Piatetsky-Shapiro, G., Smyth, P., Uthurusamy, R. (Hrsg.): Advances in Knowledge Discovery and Data Mining, Menlo Park, S. vii-xi.
- WITTEN, I. H., FRANK, E. (2001). Data Mining: Praktische Werkzeuge und Techniken für das maschinelle Lernen. Hanser Verlag, München.
- WOLLRING, J. (1995): Der Chlorophyllgehalt als Maßstab für den Stickstoffdüngerbedarf bei Wintergetreide. Zeitschrift für Pflanzenernährung und Bodenkunde, 153(3). S. 47-53.
- WOLLRING, J.; WEHRMANN, J. (1981): Der Nitrat-Pflanzentest, Entscheidungshilfe für die N-Spätdüngung. DLG-Mitt. Heft 8.
- ZHANG, X-S. (2000): Neural Networks in Optimization – Nonconvex optimization and its applications. Kluwer Academic Publishers, Dordrecht.
- ZIMMERMANN, H. G. (1994). Neuronale Netze als Entscheidungskalkül. In: Rehkugler, H., Zimmermann, H.G. (Hrsg.): Neuronale Netze in der Ökonomie – Grundlagen und finanzwirtschaftliche Anwendungen. Verlag Vahlen, München. S. 1-88.

8 Anhang

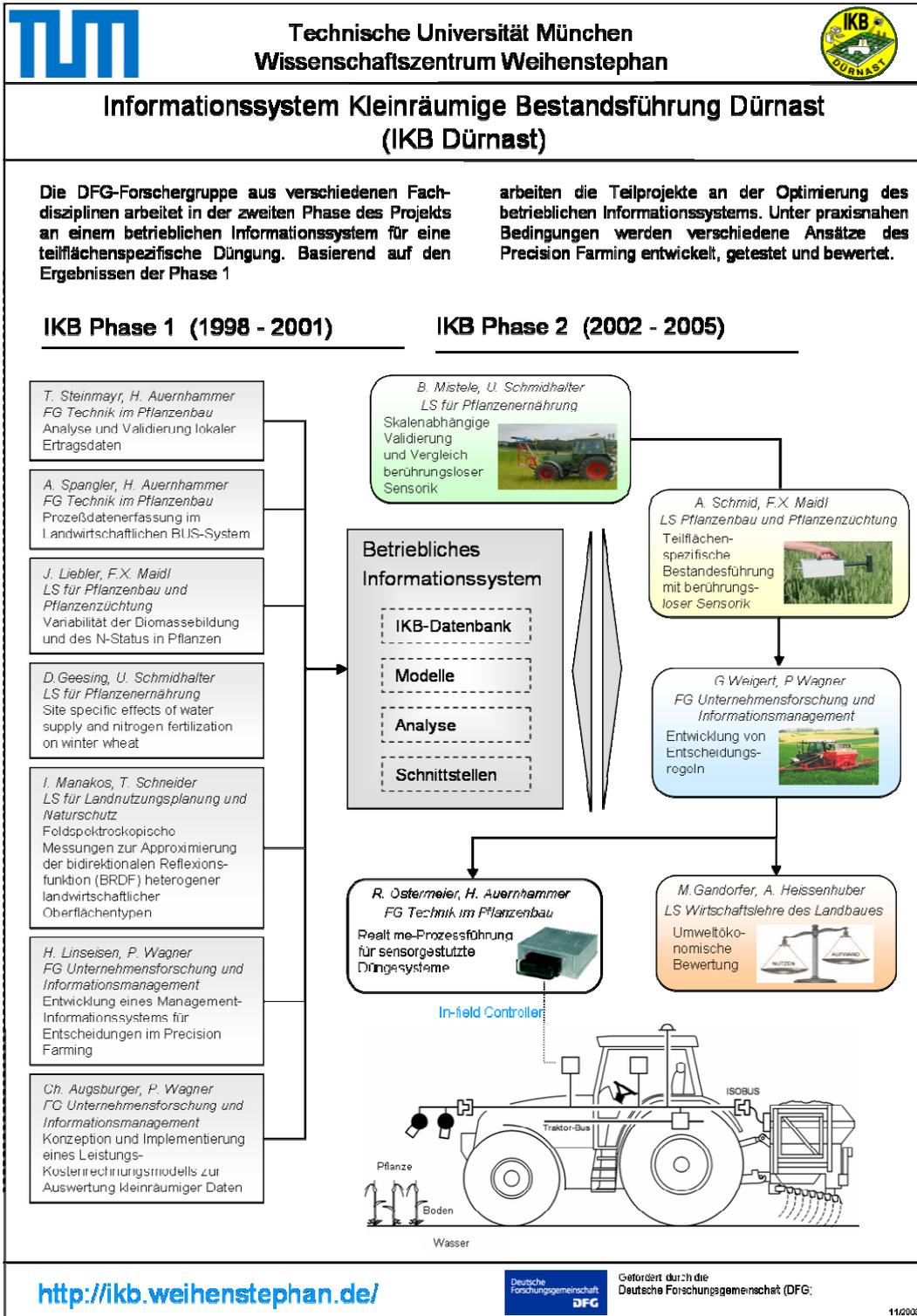


Abbildung A-66: Konzeption des Forschungsprojektes IKB-Dürnast

Tabelle A-28: Algorithmus zur dritten N-Gabe (Schafhof 02, Variante Sensor Overlay)

IF (EZ = N)

$$\text{SREIPts} = 726.9;$$

$$\text{DGts} = (\text{SREIPts} - \text{IREIPts}) * \text{fa} + \text{NBts};$$

ELSE IF (EZ = M)

$$\text{SREIPts} = 727.0;$$

$$\text{DGts} = (\text{SREIPts} - \text{IREIPts}) * \text{fa} + \text{NBts};$$

ELSE (EZ = H)

$$\text{SREIP} = 727.4;$$

$$\text{DGts} = (\text{SREIPts} - \text{IREIPts}) * \text{fa} + \text{NBts};$$

END IF

Darin sind:

EZ = Ertragszone, abgeleitet aus der Ertragserwartung EEts (EEts = □ Ertrags (1996,1998))

N, M, H = Ertragszonen (Niedrig, Mittel, Hoch)

SREIPts = Soll-Kennzahl in nm als Indikator für einen optimalen N-Versorgungszustand eines Teilstreifens (ts).

IREIPts = Ist-Kennzahl ermittelt aus Reflexionsmessungen in nm für jeden Teilstreifen (ts).

fa = Umrechnungsfaktor von Soll-Ist-Abweichung in nm zu einer aktuellen Unterversorgung (+) bzw. Überversorgung (-) der Pflanzen in kg N/ha eines Teilstreifens mit fa = 25. Akzeptierte Spanne der Werte: +20 bis -5 kg N/ha.

NBts = N-Bedarf der Pflanze zwischen EC 49 und Ernte in Abhängigkeit von der Ertragserwartung als EEts - 10 in kg N/ha (nach Diepolder, 1994, S. 152).

Die Gesamtdüngermenge darf nur Werte zwischen 60 und 90 kg N/ha annehmen

Quelle: Linseisen (2002, unveröffentlicht)**Tabelle A-29: Bodenwerte der Domäne Görzig**

Bodenwerte	Einheit	Görzig	Bodenwerte	Einheit	Görzig
Ton	(%)	19	Grobsand	(%)	2
Feinschluff	(%)	9	Sand insges.	(%)	9
Mittelschluff	(%)	18	Sorptionskapazität	mval/100mg	17
Grobschluff	(%)	45	C _t	(%)	2,8
Abschlämmb. Teile	(%)	46	C/N		<10
Feinsand	(%)	5	pH		7,1
Mittelsand	(%)	2			

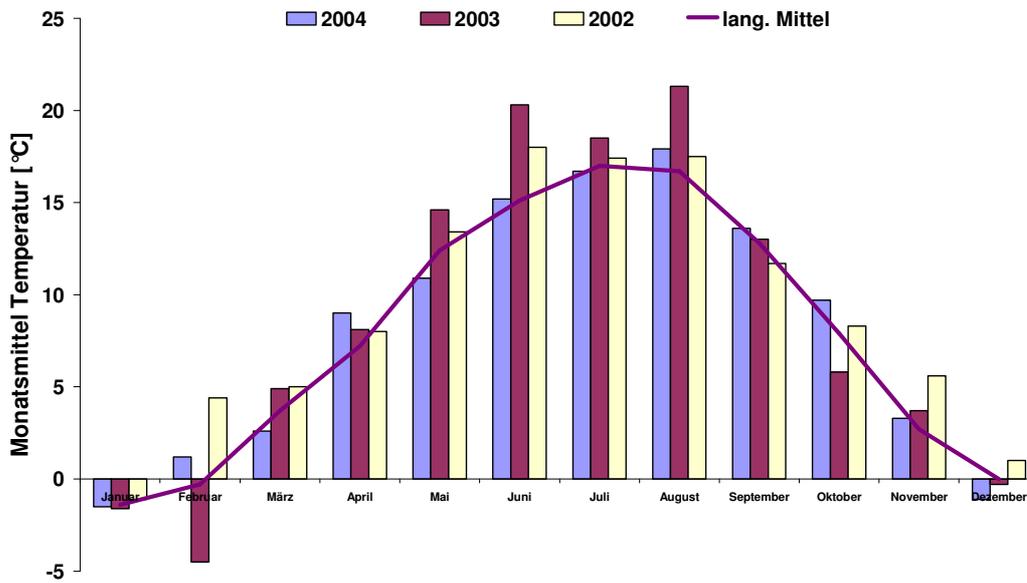


Abbildung A-67: Monatliche Temperaturmittelwerte in den Jahren 2002, 2003 und 2004 der Wetterstation Weihenstephan (Datengrundlage: Deutscher Wetterdienst).

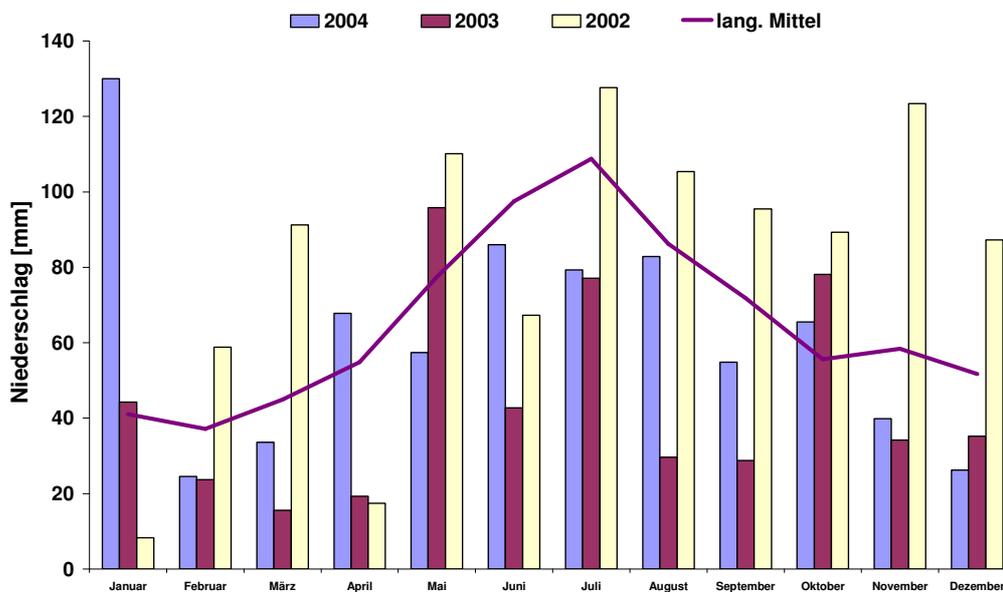


Abbildung A-68: Niederschlagsmenge der Versuchsjahre 2002, 2003 und 2004 (IKB-Versuche, Datengrundlage: Deutscher Wetterdienst, Wetterstation Weihenstephan)

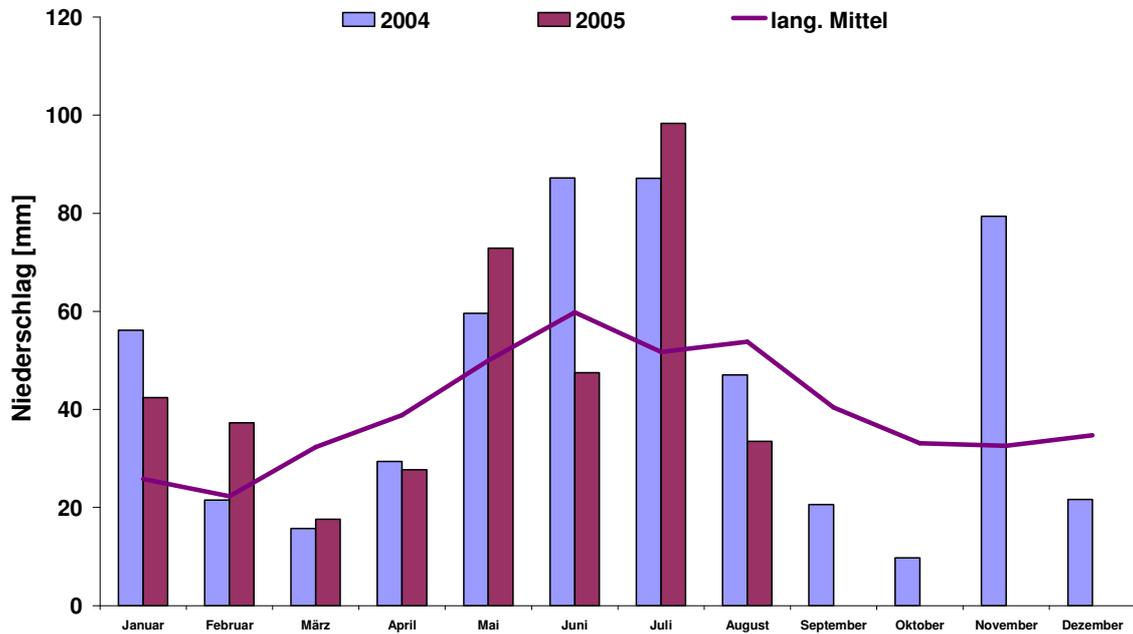


Abbildung A-69: Niederschlagsmenge der Versuchsjahre 2004 und 2005 (Versuch Görzig)

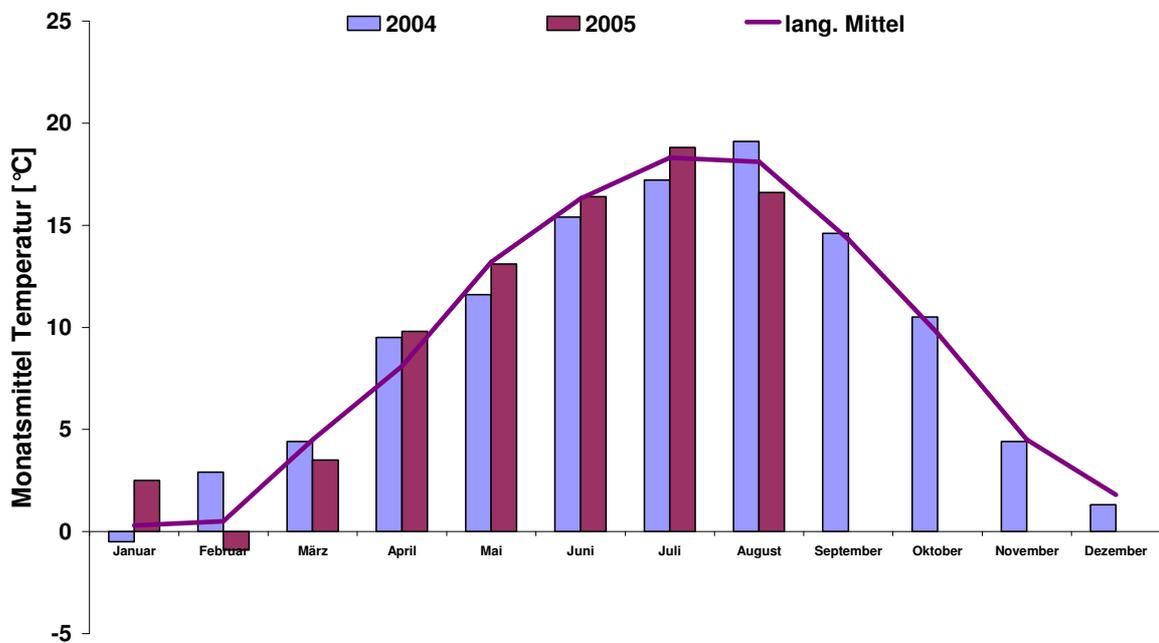


Abbildung A-70: Temperaturverlauf der Versuchsjahre 2004 und 2005 (Versuch Görzig 04)

Tabelle A-30: Korrektur der Ertragsdaten (Quelle: Agri Con®)

Eliminierung von Datensätzen:

- Ertrag ≤ 1 oder ≥ 15 t/ha
- Feuchte ≤ 1 oder ≥ 25 %
- Geschwindigkeit ≤ 0 oder ≥ 15 km/h
- Falsche Schneidwerksbreiten
- Vorgewende (generell 36 m)
- Streuungsgrenze $\pm 1,75 = 4\%$ der Werte nach Normalverteilung

Anpassung an den Durchschnittsertrag

- Mehrere Mähdrescher liefern Daten (Kalibrierung)
- Große Abweichungen zum gewogenen Ertrag (Betriebsangabe)

Reduzieren der Daten bei hoher Anzahl

Über 15.000 Datensätze (Mittelwert aus 2-4 Einzelpunkten)

Tabelle A-31: Neuronale Netze zur ersten N-Applikation (Versuch D4 03)																						
Variante 1	Variante 2	Variante 3																				
<p>Estimated accuracy: 91,14 Input Layer: 4 neurons Hidden Layer 1: 6 neurons Hidden Layer 2: 2 neurons Output Layer: 1 neurons</p> <p>Relative Importance of Inputs:</p> <table> <tr><td>ZUGKRAFT</td><td>0.34</td></tr> <tr><td>ERTRAG_98</td><td>0.20</td></tr> <tr><td>EM_38</td><td>0.15</td></tr> <tr><td>N_1</td><td>0.10</td></tr> </table> <p>Method: Exhaustive prune Stop on: Default Prevent overtraining: true Sample %: 35,0</p>	ZUGKRAFT	0.34	ERTRAG_98	0.20	EM_38	0.15	N_1	0.10	<p>Estimated accuracy: 91,19 Input Layer: 3 neurons Hidden Layer 1: 30 neurons Hidden Layer 2: 20 neurons Output Layer: 1 neurons</p> <p>Relative Importance of Inputs:</p> <table> <tr><td>N_1</td><td>0.25</td></tr> <tr><td>ZUGKRAFT</td><td>0.21</td></tr> <tr><td>ERTRAG_98</td><td>0.19</td></tr> </table> <p>Method: Exhaustive prune Stop on: Default Prevent overtraining: true Sample %: 40,0</p>	N_1	0.25	ZUGKRAFT	0.21	ERTRAG_98	0.19	<p>Estimated accuracy: 92,42 Input Layer: 3 neurons Hidden Layer 1: 3 neurons Hidden Layer 2: 1 neurons Output Layer: 1 neurons</p> <p>Relative Importance of Inputs:</p> <table> <tr><td>ZUGKRAFT</td><td>0.24</td></tr> <tr><td>N_1</td><td>0.22</td></tr> <tr><td>AVG_Ertrag</td><td>0.05</td></tr> </table> <p>Method: Exhaustive prune Stop on: Default Prevent overtraining: true Sample %: 40,0</p>	ZUGKRAFT	0.24	N_1	0.22	AVG_Ertrag	0.05
ZUGKRAFT	0.34																					
ERTRAG_98	0.20																					
EM_38	0.15																					
N_1	0.10																					
N_1	0.25																					
ZUGKRAFT	0.21																					
ERTRAG_98	0.19																					
ZUGKRAFT	0.24																					
N_1	0.22																					
AVG_Ertrag	0.05																					
Variante 4	Variante 5	Variante 6																				
<p>Estimated accuracy: 91,731 Input Layer: 3 neurons Hidden Layer 1: 29 neurons Hidden Layer 2: 13 neurons Output Layer: 1 neurons</p> <p>Relative Importance of Inputs</p> <table> <tr><td>ZUGKRAFT</td><td>0.29</td></tr> <tr><td>N_1</td><td>0.28</td></tr> <tr><td>AVG_Ertrag</td><td>0.04</td></tr> </table> <p>Method: Exhaustive prune Stop on: Default Prevent overtraining: true Sample %: 50,0</p>	ZUGKRAFT	0.29	N_1	0.28	AVG_Ertrag	0.04	<p>Estimated accuracy: 91,42 Input Layer: 4 neurons Hidden Layer 1: 6 neurons Hidden Layer 2: 2 neurons Output Layer: 1 neurons</p> <p>Relative Importance of Inputs</p> <table> <tr><td>N_1</td><td>0.35</td></tr> <tr><td>ZUGKRAFT</td><td>0.25</td></tr> <tr><td>EM_38</td><td>0.06</td></tr> <tr><td>AVG_Ertrag</td><td>0.02</td></tr> </table> <p>Method: Exhaustive prune Stop on: Default Prevent overtraining: true Sample %: 50,0</p>	N_1	0.35	ZUGKRAFT	0.25	EM_38	0.06	AVG_Ertrag	0.02	<p>Estimated accuracy: 91,73 Input Layer: 3 neurons Hidden Layer 1: 29 neurons Hidden Layer 2: 13 neurons Output Layer: 1 neurons</p> <p>Relative Importance of Inputs</p> <table> <tr><td>ZUGKRAFT</td><td>0.29</td></tr> <tr><td>N_1</td><td>0.28</td></tr> <tr><td>AVG_Ertrag</td><td>0.04</td></tr> </table> <p>Method: Exhaustive prune Stop on: Default Prevent overtraining: true Sample %: 50,0</p>	ZUGKRAFT	0.29	N_1	0.28	AVG_Ertrag	0.04
ZUGKRAFT	0.29																					
N_1	0.28																					
AVG_Ertrag	0.04																					
N_1	0.35																					
ZUGKRAFT	0.25																					
EM_38	0.06																					
AVG_Ertrag	0.02																					
ZUGKRAFT	0.29																					
N_1	0.28																					
AVG_Ertrag	0.04																					
*AVG_Ertrag: Mittelwert der beiden Attribute Ertrag_98 und Ertrag_96																						

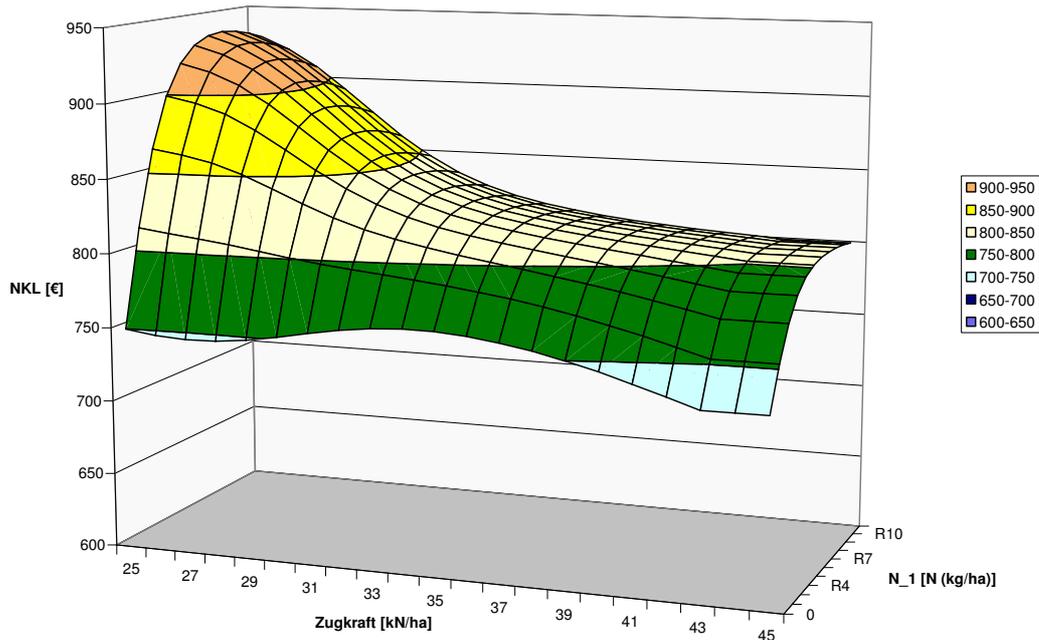


Abbildung A-71: NKL-Response zur ersten N-Applikation (Zugkraft, Versuch D4 03)

Tabelle A-32: Neuronale Netze zur zweiten N-Applikation (Versuch D4 03)						
	Estimated Accuracy:	Topologie (Anzahl Neuronen)	Input-Attribute	Relative Importance of Inputs		Build Settings
#1	92,71	Input Layer: 5 Hidden Layer 1: 2 Hidden Layer 2: 1 Output Layer: 1	ERTRAG_98 N_1 N_2 REIP_2 ZUGKRAFT	REIP_2 ZUGKRAFT N_2 ERTRAG_98 N_1	0.42 0.21 0.10 0.03 0.02	Method: Exhaustive prune Prevent overtraining: true Sample %: 50,0
#2	92,60	Input Layer: 5 Hidden Layer 1: 2 Hidden Layer 2: 2 Output Layer: 1	ERTRAG_98 N_1 N_2 REIP_2 ZUGKRAFT	REIP_2 ZUGKRAFT N_1 ERTRAG_98 N_2	0.28 0.24 0.10 0.07 0.07	Method: Exhaustive prune Prevent overtraining: true Sample %: 40,0
#3	93,21	Input Layer: 3 Hidden Layer 1: 2 Hidden Layer 2: 1 Output Layer: 1	N_2 REIP_2 ZUGKRAFT	REIP_2 N_2 ZUGKRAFT	0.45 0.15 0.13	Method: Exhaustive prune Sample %: 50,0 Prevent overtraining: true
#4	93,01	Input Layer: 4 Hidden Layer 1: 3 Hidden Layer 2: 2 Output Layer: 1	ERTRAG_98 N_2 REIP_2 ZUGKRAFT	REIP_2 ZUGKRAFT N_2 ERTRAG_98	0.45 0.16 0.14 0.03	Method: Exhaustive prune Stop on: Default Prevent overtraining: true Sample %: 35,0

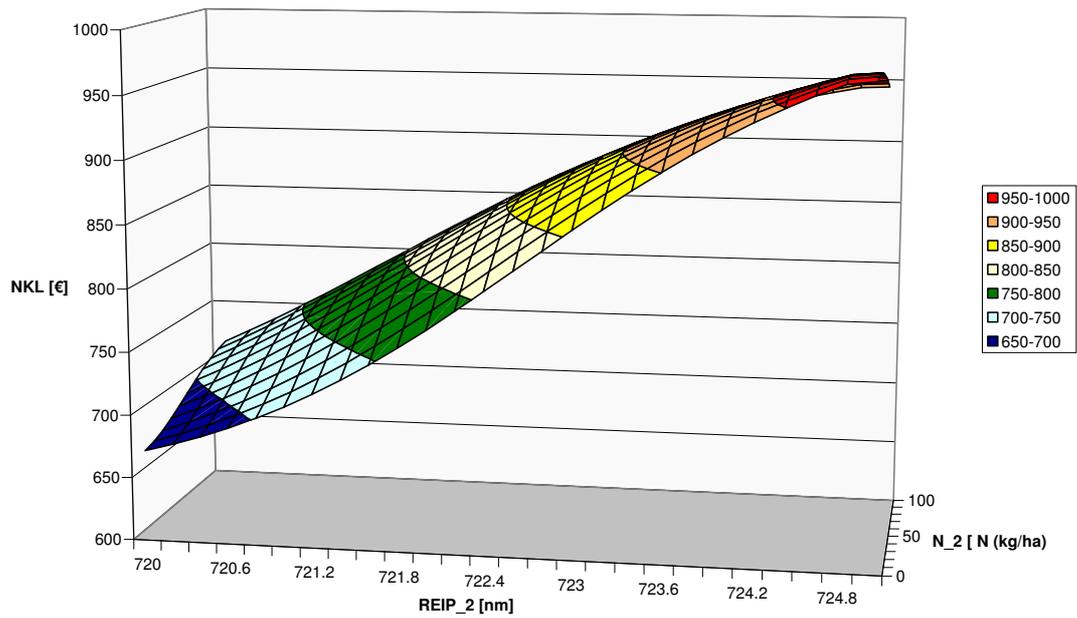


Abbildung A-72: NKL-Response zur zweiten N-Applikation (*REIP_2*, Versuch *D4 03*)

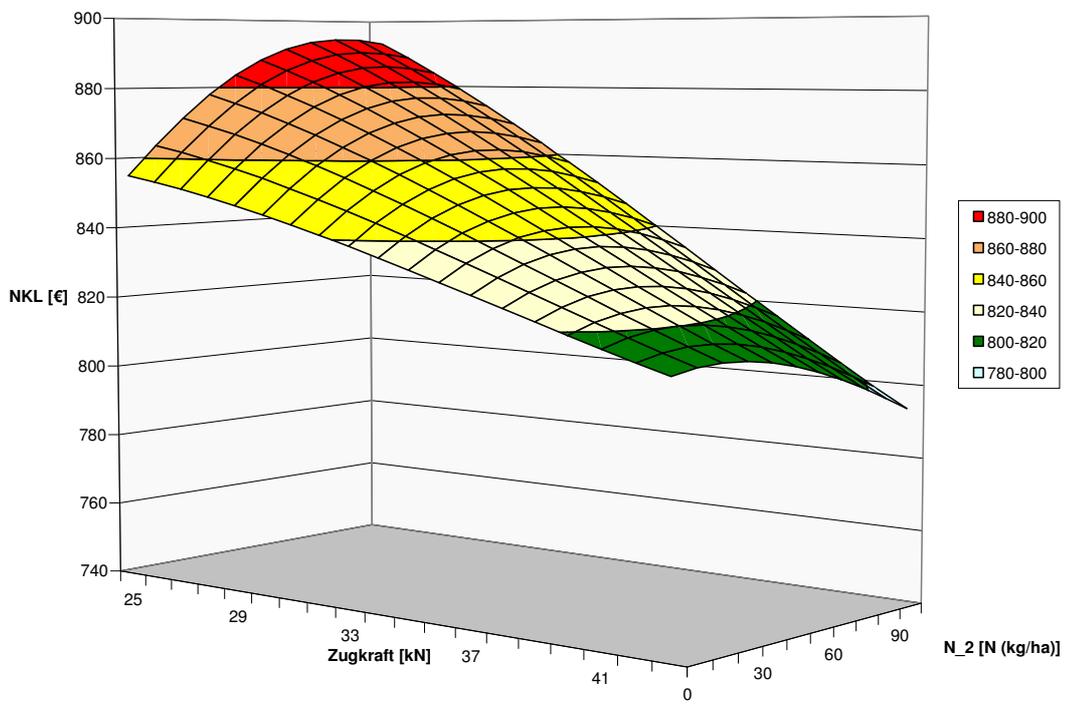


Abbildung A-73: NKL-Response zur zweiten N-Applikation (*Zugkraft*, Versuch *D4 03*)

Tabelle A-33: Neuronale Netze zur dritten N-Applikation (Versuch D4 03)

Variante	Estimated Accuracy:	Topologie (Anzahl Neuronen)	Input-Attribute	Relative Importance:	Build Settings
1	92,74	Input Layer: 2 Hidden Layer 1: 2 Hidden Layer 2: 1 Output Layer: 1	<i>ERTRAG_96</i> <i>ERTRAG_98</i> <i>N_1</i> <i>N_2</i> <i>N_3</i> <i>REIP_2</i> <i>REIP_3</i> <i>ZUGKRAFT</i> <i>EM_38</i>	<i>REIP_3</i> 0.62 <i>ZUGKRAFT</i> 0.23	Method: Exhaustive prune Prevent overtraining: true Sample %: 50,0
2a	93.11	Input Layer: 4 Hidden Layer 1: 2 Hidden Layer 2: 1 Output Layer: 1	<i>ERTRAG_96</i> <i>ERTRAG_98</i> <i>N_1</i> <i>N_2</i> <i>N_3</i> <i>REIP_2</i> <i>REIP_3</i> <i>ZUGKRAFT</i> <i>EM_38</i>	<i>REIP_3</i> 0.44 <i>REIP_2</i> 0.34 <i>ZUGKRAFT</i> 0.18 <i>N_3</i> 0.08	Method: Exhaustive prune Prevent overtraining: true Sample %: 40,0
2b	93.41	Input Layer: 4 Hidden Layer 1: 2 Hidden Layer 2: 1 Output Layer: 1	<i>ERTRAG_96</i> <i>ERTRAG_98</i> <i>N_1</i> <i>N_2</i> <i>N_3</i> <i>REIP_2</i> <i>REIP_3</i> <i>ZUGKRAFT</i> <i>EM_38</i>	<i>REIP_3</i> 0.39 <i>REIP_2</i> 0.32 <i>ZUGKRAFT</i> 0.19 <i>EM_38</i> 0.09	Method: Exhaustive prune Sample %: 40,0 Prevent overtraining: true
3	92,52	Input Layer: 2 Hidden Layer 1: 2 Hidden Layer 2: 2 Output Layer: 1	<i>ERTRAG_98</i> <i>N_3</i> <i>REIP_3</i> <i>ZUGKRAFT</i>	<i>REIP_3</i> 0.49 <i>ERTRAG_98</i> 0.29	Method: Exhaustive prune Prevent overtraining: true Sample %: 40,0