

Optimierung standortangepasster N-Düngung über teilflächenspezifische Ertragsprognosen

GEORG WEIGERT, WEIHENSTEPHAN
PETER WAGNER, HALLE

Abstract

Es wird ein integriertes System vorgestellt, das räumliche Daten verwendet, um Informationen über standortangepasste N-Düngung zu erzeugen. Dabei liegt der Fokus des Beitrags auf einem neuronalen Netz, das vor allem zur Ertragsprognose herangezogen wird. Zudem können aus der Analyse des verwendeten neuronalen Netzes Rückschlüsse auf die Qualität der verwendeten Datensätze und deren Attribute gezogen werden.

1 Einführung

Eine Auswertung von 108 Versuchen zur Überprüfung der Wirtschaftlichkeit von Precision Farming (LAMBERT UND LOWENBERG-DEBOER, 2000) lassen vermuten, dass der Erfolg der Variable Rate Technology (VRT) „durchwachsen“ zu sehen ist. Sieht man sich die betrachteten Untersuchungen jedoch genauer an, muss man feststellen, dass es aus mehreren Gesichtspunkten heraus noch zu früh für eine abschließende ökonomische Beurteilung ist: Zum einen wurden in den betrachteten Studien zur Optimierung der VRT-Varianten überwiegend Bodenbeprobungen verwendet, die sehr kostenaufwändig durchzuführen bzw. räumlich zu niedrig aufgelöst sind. Dies beeinträchtigt sowohl die Kosteneffizienz als auch das Potenzial zur Optimierung. Zum anderen basiert die Optimierung des variablen Input Faktors auf sehr unterschiedlichen Informationen. In einigen Fällen wird nur auf Basis von Ertragskarten optimiert, in anderen nur nach ermittelten Bodenarten. Doch genau in der Kombination aller verfügbaren kleinräumigen Informationen, die sich typischerweise gegenseitig beeinflussen, sollte noch ökonomisches Potenzial zu finden sein. Der Wert zusätzlicher Informationen, seien es höher aufgelöste Daten (also mehr Daten pro Flächeneinheit) oder zusätzliche Attribute, ist bei der VRT wesentlich höher als bei der uniformen Bewirtschaftung (BULLOCK ET AL. 2002, KILIAN UND MALZER, 2001).

Um die erwarteten ökonomischen Potenziale der VRT auszuschöpfen, benötigt man also kostengünstig zu erfassende und hochaufgelöste Daten. Vor allem benötigt man jedoch Wissen darüber, wie sich diese Daten gegenseitig beeinflussen und wie sie in diesem Wechselspiel den Ertrag beeinflussen. Die Methode der Wissensentdeckung in Datenbanken (WED) (NAK-HAEIZADEH, 1998), die hier angewendet werden soll bedient sich der Daten, die über teilflächenspezifische Bewirtschaftung erfasst und gespeichert wurden und versucht mittels Data Mining Methoden charakteristische Muster zu finden, also spezifisches Wissen über die Zusammenhänge im Datenmaterial.

2 Der WED-Prozess im Bereich Precision Farming

Der angepasste WED Prozess baut unter anderem auf den Komponenten eines Management-Informationssystems im Bereich Precision Farming auf (Abb.1, rechter grau unterlegter Bereich). Dieses System bildet die Basis zur Erfassung, Speicherung und Bearbeitung von räumlichen Daten, die mit der automatischen Prozessdatenerfassung gewonnen werden.

Kernstück dabei bildet die mobile Prozesstechnik (AUERNHAMMER, 1999). Damit können automatisiert räumliche Daten über die Bewirtschaftungsmaßnahme selbst (z. B. Ist-Ausbringungsmenge, Geschwindigkeit, etc), sowie - mittels spezieller Sensoren - Daten über Boden bzw. Pflanze gewonnen werden. Bisher wurden dazu die scheinbare elektrische Bo-

denleitfähigkeit (DURLESSER, 1999), ein Vegetationsindex (LIEBLER ET AL., 2001), die benötigte Zugkraft sowie die Erträge aufgezeichnet.

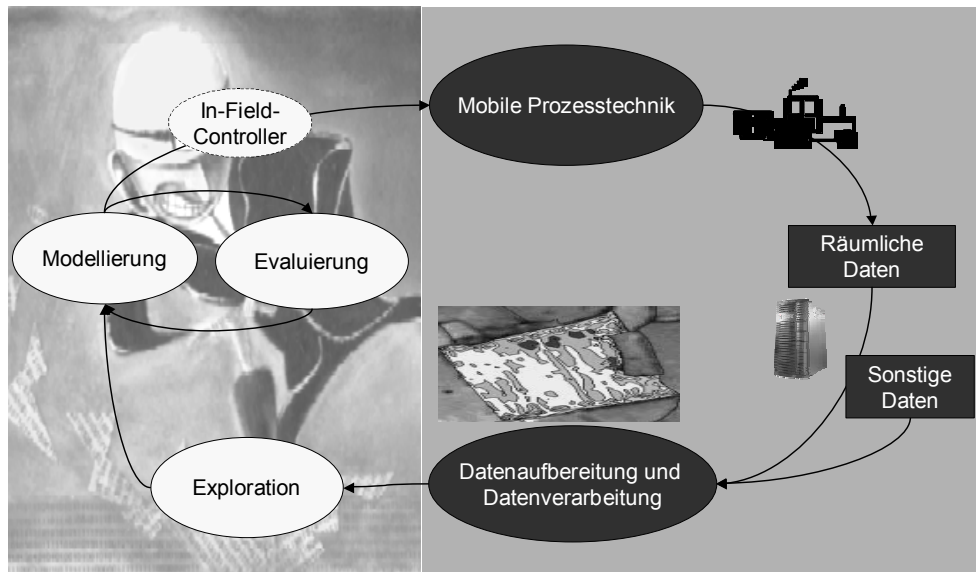


Abb. 1: Der WED-Prozess im Bereich Precision Farming.

Diese (Roh-)Daten werden in Datenbanken gespeichert und können via Internet als Punktdaten abgerufen werden. Zusätzlich können beispielsweise auch Daten aus Versuchen oder Klimadaten, die je nach Anwendungsfall Bedeutung haben, zur Auswertung hinzugezogen werden. In der Graphik werden diese als „sonstige Daten“ bezeichnet.

In einem weiteren Schritt können diese Daten dann aufbereitet und verarbeitet werden. Dies geschieht beispielsweise in Geo-Infomations-Systemen (GIS). Zu den Aufgaben der Datenaufbereitung gehören:

- Korrektur erfasster Punktdaten (z.B. Anwendung statistischer Methoden zur Ertragskorrektur)
- Daten vervollständigen (Prognose fehlender Werte, räumliche Interpolation, etc.)
- Überprüfung der Kompatibilität von Daten aus verschiedenen Quellen
- Erzeugung von homogenen Teilflächen in variabler Rasterung
- Aggregation von Punktdaten auf Teilflächen

Für die Data-Mining Auswertung bilden also die einzelnen Teilflächen (Lern-)Instanzen. Die durchschnittlichen Punktdatenwerte, die in dieser Teilfläche liegen, ergeben den (durchschnittlichen) Attributwert dieser Instanz. Nach diesen Vorbereitungen beginnt der eigentliche Data Mining Prozess, der die Exploration, also die visuellen Analyse des Datenmaterials, die Modellierung und schließlich die Evaluation der erstellten Modelle umfasst. Über einen In-Field-Controller sollen gefundene Ergebnisse dann umgesetzt werden, d.h. automatisiert zur mobilen Prozesssteuerung verwendet werden.

3 Ziel des Projekts

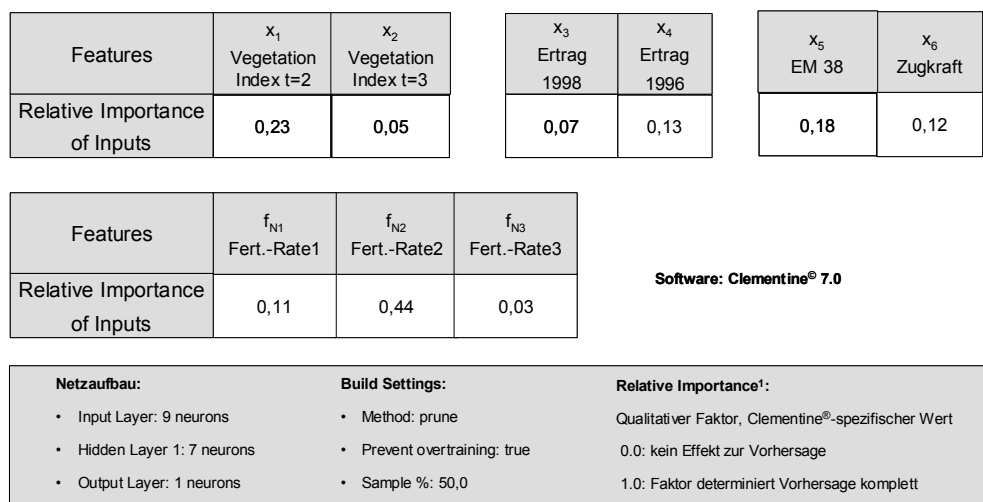
Es sollen Entscheidungsregeln für eine standortoptimale N-Düngung gefunden werden. Das ökonomische Ziel einer standort-optimalen N-Düngung kann mit der Optimierung der teilflächenspezifischen N-freien Leistung (NFL_{TF}) erreicht werden. Die N-freie Leistung stellt den finanziellen Ertrag ($Y \cdot p_Y$) korrigiert um die Stickstoffkosten ($f_N \cdot p_N$) dar. Der teilflächenspezifische Ertrag Y hängt annahmegemäß von teilflächenspezifischen Parametern (x_1, x_2, \dots), von zusätzlichen Daten (z_1, z_2, \dots), sowie von dem einzig kontrollierbaren Faktor, der variablen N-Menge ab. Dies ist in Formel 1 zusammengefasst:

$$NFL_{TF} = p_Y * Y(x_1, x_2, \dots, z_1, z_2, \dots, f_N) - f_N * p_N \rightarrow \max! \quad (1)$$

Um die standortoptimale N-Menge bestimmen zu können muss deshalb die teilflächenspezifische Produktionsfunktion abgeschätzt und analysiert werden. Dann kann mit Hilfe einer Simulation von technisch möglichen Düngergaben ($N = \{0, 10, 20, \dots\}$) die Düngermenge bestimmt werden, die zu einer maximalen N-freien Leistung führt. (vgl. POKRAJAC UND OBRADOVIC, 2001).

4 Auswertungen

In einem ersten Prozessdurchlauf wurden 449 teilflächenspezifische Winterweizen-Datensätze eines Jahres zur Ertragsprognose verwendet. Davon stammten 329 Datensätze aus einem On-Farm-Versuch und 120 Datensätze aus einem Stickstoffsteigerungsversuch. Zur Prognose wurde ein neuronales Netz verwendet, um die Abhängigkeiten der Attribute untereinander zu berücksichtigen. Mit Hilfe der Sensitivitätsanalyse der verwendeten Software lassen sich erste Hinweise auf die Bedeutung der verwendeten Attribute für die Prognose finden.



Grafik 1: Analyse des trainierten neuronalen Netzes in Clementine®

Graphik 1. zeigt den qualitativen Wert „Relative Importance“² der einzelnen Attribute. Dabei kann man erkennen, dass insbesondere die scheinbare elektrische Bodenleitfähigkeit (EM 38) sowie die zweite N-Gabe (N-Fert.2) mit der dazugehörigen Spektralmessung (Vegetation Index t=2) gute Hinweise auf den zu erwartenden Ertrag liefern. Dagegen fällt auf, dass die dritte N-Gabe mit der dazugehörigen Spektralmessung eine sehr geringe Bedeutung hat. Dieser Effekt ist vermutlich darauf zurückzuführen, dass die dritte N-Gabe, die erst relativ spät verabreicht wurde, sich eher auf den Proteingehalt als auf den Ertrag ausgewirkt hat. Generell haben die drei Düngergaben zusammen weniger Einfluss auf den Ertrag als erwartet. Ein Grund dafür ist sicherlich in der geringen Variation der Düngergaben des On-Farm Versuches zu sehen.

Als Methode zur Evaluation des verwendeten Modells, wurde die Kreuzvalidierung gewählt. Diese empfiehlt sich besonders bei kleinen Datensätzen. Dazu wird der Datensatz aufgeteilt in einen Trainings- und in einen Validierungsdatensatz. Bei einer fünffachen Kreuzvalidierung stehen so 4/5 der Datensätze dem Training zur Verfügung. An dem restlichen 1/5 wird das Modell getestet. Die Methode wird fünfmalig wiederholt, sodass jeder Datensatz einmal als Validierungsdatensatz gedient hat. Zur Beurteilung der Prognose können dann die prognostizierten mit den tatsächlichen Erträgen im Validierungsdatensatz verglichen werden. Als Gü-

temaß eignen sich beispielsweise die lineare Korrelation oder die Standardabweichung (vgl. Witten und Frank S.158). Im Falle einer fünffachen Kreuzvalidierung werden demzufolge die gemittelten Werte der einzelnen Modelle verwendet. Obiger Datensatz ergab eine durchschnittliche lineare Korrelation von 0,80. Die Standardabweichung war jedoch in den fünf (Teil-)Modellen sehr unterschiedlich, was zu einer Standardabweichung der linearen Korrelation von 0,16 führte. Dies deutet darauf hin, dass noch zu wenige Datensätze verwendet wurden, bzw. die Qualität der Datensätze zu niedrig ist.

5 Fazit & Ausblick

Generell ist zu bemerken, dass die Ergebnisse je nach Datenauswahl und –aufbereitung stark variieren. Ein erstes Problem liegt in der Datenquantität und Qualität. Eine höhere Anzahl an (Lern-) Instanzen mit einer höheren Variation der Attributwerte würde die Qualität der Vorhersage verbessern. Ein weiteres Problem liegt in der Kompatibilität von On-Farm- und Exaktversuchen. So sind beispielsweise im vorliegenden Fall nicht die gleichen Attribute erfasst worden und entsprechende Attribute nicht direkt kompatibel zueinander. Dazu ergeben sich durch die grundsätzlich unterschiedliche Versuchsdurchführung Probleme in der Vergleichbarkeit. Hier muss noch genauer untersucht werden, wie man eine verbesserte Kompatibilität erreichen kann. Ebenfalls könnten fehlende Boden-Attribute noch ex-post erfasst werden. Grundsätzlich bringt die Erweiterung mit Exaktversuchsdaten einen Fortschritt in der Qualität der Prognose, da dadurch eine bessere Abdeckung des Ereignisraumes des variablen Input-Faktors erreicht wird. Sobald die Qualität der Prognose gewährleistet ist, kann die optimale Düngermenge bestimmt werden, die die teilflächenspezifische N-freie Leistung maximiert.

6 Danksagung

Die Entwicklung von Entscheidungsregeln zur variablen N-Düngung ist Teil des IKB-Projekts, eines interdisziplinären Precision Farming Projekts an der TU München-Weihenstephan, das von der DFG (Deutsche Forschungsgemeinschaft) gefördert wird.

7 Literatur

- AUERNHAMMER, H. (1999): Precision Farming for Site-Specific Fertilization. In: Zeitschrift für Agrar-informatik, Münster-Hiltrup, H3, pp. 58-66.
- BULLOCK, D. S., LOWENBERG-DEBOER, J, SWINTON, S.M. (2002): Adding value to spatially managed inputs by understanding site-specific yield response. In: Agricultural Economics, 27 (2002), pp. 233-245.
- DURLESSER, H. (1999): Bestimmung der Variation bodenphysikalischer Parameter in Raum und Zeit mit elektromagnetischen Induktionsverfahren, Shaker Verlag, Aachen, Germany.
- 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, viewed: 13.01.2003.
- LAMBERT, D. AND LOWENBERG-DEBOER, J. (2000): Precision agriculture profitability review [Online]. Available at <http://mollisol.agry.purdue.edu/SSMC/publications.html>
- LIEBLER, J., STICKSEL, E., AND MAIDL, F.-X. (2001): Field Spectroscopic Measurements to Characterize Nitrogen Status and Dry Matter Production of Winter Wheat. In: Proceeding of the 3rd European Conference on Precision Agriculture, eds. G Grenier and S Blackmore, agro Montpellier, France, pp. 935-938.
- POKRAJAC D., OBRADOVIC Z. (2001): Neural Network-Based Software for Fertilizer Optimization in Precision Farming. In: Proc. International Joint Conference on Neural Networks, Vol. 3, 2001, pp. 2110-2115.
- WITTEN, I.H., FRANK, E.: Data Mining, Praktische Werkzeuge und Techniken für das maschinelle Lernen. Hanser Verlag, München.