

YieldSweeper – Ein Programm zur Nachbearbeitung von Ertragsdaten im Precision Farming

Robin Gebbers, Horst Domsch

Technik im Pflanzenbau
Leibniz-Institut für Agrartechnik Potsdam
Max-Eyth-Alle 100
14469 Bornim
rgebbers@atb-potsdam.de

Abstract: Georeferenzierte Mährescher-Ertragsdaten sind aus vielen Gründen fehlerhaft. Das Programm YieldSweeper vereinigt mehrere Ansätze zur automatischen Nachbearbeitung dieser Daten. Am Beispiel von Korrelationen zwischen einem Luftbild und den Ertragsdaten werden die Auswirkungen einzelner Korrekturschritte verglichen.

1 Einführung

Die Ertragskartierung von Körnerfrüchten durch Mährescher ist mittlerweile weit verbreitet. Ertragskarten werden im Precision Farming beispielsweise zur Schätzung der Nährstoffentzüge, Prüfung des Erfolgs vorhergehender Bewirtschaftungsmaßnahmen oder zur Abschätzung von Ertragspotenzialen (Ertragsprognose) genutzt. In der Praxis zeigt sich jedoch, dass Ertragsdaten stark fehlerbehaftet sind. [JH02] haben bei der Auswertung mehrjähriger Ertragsdaten festgestellt, dass bis zu drei Viertel der Daten verworfen werden mussten. Neben Fehlern bei der gesamten Beerntung, etwa durch unterschiedliche Kalibrierung beim Einsatz mehrerer Maschinen, treten zahlreiche Fehler bei der einzelnen Maschine auf. Dazu zählen:

- Positionierungsfehler des GPS
- Kalibrierungsfehler: Dadurch werden zumindest die absoluten Erträge beeinflusst (systematischer Fehler). Kurzfristige Drifts der Kalibrierung können auch das räumliche Ertragsmuster stören [BM99].
- Ungenauigkeit des Kornstromsensors.
- Fehlerhafte Wegmessung (falls ein Wegsensor genutzt wird): Die Strecke geht als Faktor in die Berechnung des flächenbezogenen Ertrages ein.
- Reale Schnittbreite ist geringer als angenommene Schnittbreite: Die Schnittbreite geht als Faktor in die Berechnung des flächenbezogenen Ertrages ein.
- Zeitversatz zwischen Schnitt am Schneidwerk und Messung am Elevator.
- Füllphasenfehler: Wenn sich die leeren Dreschorgane des Mähreschers bei der Einfahrt in den Bestand langsam füllen, werden zu niedrige Erträge aufgezeichnet

- Entleerungsphasenfehler: Wenn sich volle Dreschorgane langsam entleeren, nachdem der Mähdrescher aus dem Bestand herausgefahren ist, werden zu hohe Erträge aufgezeichnet.
- Interne Vermischung: Durch die Dresch- und Transportvorgänge im Mähdrescher mischen sich Kornpartien von verschiedenen Positionen.
- Kurzfristige Geschwindigkeitsänderungen: Dadurch ändert sich die Dynamik der Stoffströme im Mähdrescher.

Zur nachträglichen Korrektur von Ertragsdaten sind verschiedene Algorithmen entwickelt worden. Die „Bitmap-Methode“ von [HSR96] befasst sich speziell mit der Korrektur der unvollständigen Schneidwerksfüllung durch geometrische Berechnungen. [BM99] identifizieren mit dem „Expert Filter“ ungestörte Arbeitsphasen, geringe effektive Schnittbreiten, Füllungs- und Entleerungsphase. [NMD03] stellen einen lokalen räumlichen Filter (H-Methode) vor, um Ausreißer zu identifizieren. [LW03], [TW04] und [RD05] analysieren die Stoffstromdynamik im Mähdrescher, um durch inverse Modellierung (Entfaltung) die Fehler durch interne Vermischung zu korrigieren. Letzterer Ansatz ist relativ komplex und setzt ein Modell der Stoffstromdynamik der Maschine voraus. Dieses ist jedoch in der Regel unbekannt oder wird vom Hersteller nicht veröffentlicht. Für die Praxis kommen nur Korrekturalgorithmen infrage, die im Wesentlichen mit den aufgezeichneten Daten arbeiten und möglichst wenige Zusatzinformationen benötigen. Dies ist bei den ersten drei der genannten Ansätze gegeben.

2 Material und Methoden

Der hier vorgestellte Algorithmus kombiniert die Ansätze der Bitmap-Methode [HSR96], des Expert Filters [BM99] und der lokalen Filterung. Besonders die Bitmap-Methode wurde weiterentwickelt. Annahmen für den YieldSweeper Algorithmus sind:

- Georeferenzierung per dGPS (Lagegenauigkeit mindestens 2 m).
- Die Rohdatenmatrix muss geografische Breite und Länge sowie Rohertrag, GPS-Zeit enthalten.
- Die Schneidwerksbreite muss bekannt sein.
- Der Zeitversatz zwischen Positionsmessung und Ertragsmessung ist bereits durch die Software im Mähdrescher korrigiert.
- Daten verschiedener Maschinen können in der gleichen Datei abgelegt sein. Doppelte Daten durch wiederholte Speicherung – z.B. an verschiedenen Tagen – sind möglich.
- Die Abfolge der Rohdaten wurde nicht geändert.
- Möglichst flächendeckende Kartierung, d.h. alle Mähdrescher waren mit Ertragskartierung ausgestattet und die Daten liegen komplett vor.

Bearbeitungsschritte:

1. Einlesen und Formatieren der Daten.
2. Vergabe von IDs für jeden Ertrags-Messpunkt.
3. Transformation der geodätischen Koordinaten in ein kartesisches System.
4. Entfernung von doppelten Punkten.

5. Segmentierung durch Vergabe von IDs nach Identifizierung von Starts, Stopps, Wenden und Anheben des Schneidwerks.
6. Identifikation unterschiedlicher Maschinen oder Arbeitsphasen anhand großer Zeitsprünge.
7. Bestimmung der effektiven Erntefläche für jeden Ertragspunkt und Korrektur des Ertrags. Dazu wird für jeden Punkt anhand der Schneidwerksbreite und den Abständen zum vorherigen und folgenden Punkt ein vorläufiges Polygon erstellt. Die effektive Erntefläche ergibt sich unter Berücksichtigung der Differenz sich überlappender Polygone, wobei die zeitliche Abfolge beachtet wird.
8. Entfernung von Messwerten nach dem Starten und vor dem Stoppen.
9. Entfernung kurzer Segmente.
10. Entfernung extremer Erträge durch globale Schwellwerte.
11. Lokale Analyse auf Ausreißer (lokaler räumlicher Filter).
12. Ggf. Abgleich verschiedener Mähdrescher bzw. Arbeitsphasen anhand benachbarter Segmente.

Ein grundsätzliches Problem bei der Bewertung von Ertragskorrektur-Algorithmen besteht darin, dass ein unmittelbarer Vergleich mit dem „wahren“ Ertrag an der gleichen Position nicht möglich ist – die dafür notwendige Probenahme (z.B. Handernte) würde auch das Messergebnis des Mähdreschers beeinflussen. Daher nutzt man einerseits Ertragsschätzungen durch andere, nicht-invasive Methoden, die mit der Ertragskarte korreliert werden, und führt andererseits Plausibilitätsanalysen durch, beispielsweise durch Variographie. Der Algorithmus wurde an Kartierungen durch einen John Deere und einen Claas Mähdrescher auf jeweils einem Schlag mit Winterroggen und Winterweizen untersucht. Beim John Deere-Datensatz war ein Vergleich mit einem Luftbild möglich. Dazu wurden die Erträge aus verschiedenen Korrekturstufen mit dem aus dem Luftbild abgeleiteten VARI¹ korreliert (Kendall's τ und Pearson's r). Der Claas-Datensatz wurde mit einem georeferenzierten Fahrtprotokoll verglichen.

3 Ergebnisse

Anhand der Claas-Daten zeigte sich, dass der Algorithmus Starts und Stopps zuverlässig identifiziert. Bei der Erfassung von Überlappungen überschätzt er jedoch deren Anteil. Die Korrelation zwischen VARI und John Deere-Ertragsdaten konnte durch den Algorithmus von $\tau = 0.334$ ($r = 0.313$) auf $\tau = 0.449$ ($r = 0.590$) verbessert werden. Zuvor war allerdings ein Zeitversatz zwischen Positionsmessung und Ertragsmessung von 4 sec zu berücksichtigen. Bei der eigentlichen Filterung war der Effekt durch das Löschen von Daten bis 25 sec nach dem Starten und 10 sec vor dem Stoppen besonders groß. Die Berücksichtigung globaler Grenzwerte mit 1 und 120 dt/ha als Unter- und Obergrenzen erbrachte nur geringfügige Verbesserungen. Die zusätzliche lokale Filterung durch Bestimmung von Ausreißern nach der Box-Plot-Regel anhand benachbarter Punkte innerhalb von 20 m Abstand ergab eine Verschlechterung der Rangkorrelation auf $\tau = 0.434$ und eine Verbesserung der linearen parametrischen Korrelation auf $r = 0.590$. Die Korrektur der Erträge anhand der berechneten effektiven Erntefläche war

¹ VARI = (G - R)/(G + R - B) (R, G, B entsprechen dem roten, grünen und blauen Farbkanal des Luftbildes)

erst von Vorteil, nachdem Ertragspunkte gelöscht wurden, bei denen das Verhältnis zwischen effektiver Erntefläche und überfahrener Fläche größer als 0.75 war. Der optimale Schwellenwert für das Flächenverhältnis lag bei 0.8 und führte zu einer Verbesserung von τ (+0.02). Bei diesem Flächenverhältnis wurden die höchsten Korrelationen erzielt und es bestand nur ein marginaler Unterschied der Korrelationen der flächenkorrigierten Erträge ($\tau = 0.449$) und der gefilterten Daten ohne Flächenkorrektur ($\tau = 0.448$).

4 Diskussion

Durch die Korrektur mit YieldSweeper konnte die Qualität der Ertragsdaten erheblich verbessert werden. Die Korrekturschritte tragen in unterschiedlicher Weise dazu bei. Wider Erwarten hat die Umrechnung der Erträge auf die effektive Erntefläche nur eine geringe Verbesserung gebracht. Dies hat zwei Ursachen: Durch GPS-Fehler können real nicht vorliegende Überlappungen angezeigt werden, die bei der rechnerischen Korrektur zur Überschätzung der Erträge führen. Durch die Systemträgheit werden bei real reduzierter Schnittbreite zu hohe Erträge angezeigt, auch in diesem Fall sind die rechnerisch korrigierten Werte zu hoch. Allerdings ist das Verhältnis von effektiver Erntefläche zur überfahrenen Fläche ein sehr guter Fehlerindikator und rechtfertigt daher die aufwändige Berechnung (rund 2 h für etwa 16000 Messwerten mit MATLAB auf einem 2.4 GHz Pentium Prozessor).

5 Literaturverzeichnis

- [BM99] Blackmore, S.; Moore, M.: Remedial Correction of Yield Map Data. Precision Agriculture 1, 53-66. 1999.
- [HSR96] Han, S., Schneider, S.M.; Rawlings, S.L.; Evans, R.G.: A Bitmap Method for Determining Effective Combine Cut Width in Yield Mapping. Transactions of ASAE, 40(2), p. 485-490. 1996.
- [JH02] Joernsgaard, B.; Halmoe, S.: Intra-field yield variations over crops and years. European Journal of Agronomy, Vol. 19, 1, 23-33. 2003.
- [LW03] Lark, R.M.; Wheeler, H.C.: A Method to Investigate Within-Field Variation of the Response of Combinable Crops to an Input. Agronomy Journal 95, p. 1093-1104. 2003.
- [NMD03] Noack, P.H.; Muhr, T.; Demmel, M.: An algorithm for automatic detection and elimination of defective yield map data. In J.V. Stafford and A. Werner (eds.): Precision Agriculture. Wageningen Academic Publishers. Wageningen, The Netherlands, p. 445-450. 2003.
- [RD05] Reyniers, M.; De Baerdemaeker, J.: Comparison of two filtering methods to improve yield data accuracy. Transactions of ASAE, 38(3), p. 909-916. 2005.