

Untersuchung der Effizienz von Arbeitseinheiten innerhalb eines landwirtschaftlichen Betriebes mittels Open Data

Thoralf Stein

Schon lange beschäftigt sich die Forschung mit der Optimierung von landwirtschaftlichen Prozessen. Dabei sind immer wieder Effizienzunterschiede zu beobachten, die auf geografische Einflüsse, Witterung, individuelle Fahrzeugsteuerung usw. zurückzuführen sind. In der vorliegenden Arbeit sollen diese sichtbar und quantifizierbar gemacht werden. Dazu werden aufgezeichnete Maschinendaten aus dem Projekt „BiDaLAP“ zusammen mit diversen offenen Datenbanken kombiniert und Effizienzparameter quantifiziert. Es wird aufgezeigt, an welchen betriebstechnischen Größen noch Optimierungspotenzial besteht und wie groß die Effizienzunterschiede sind.

Schlüsselwörter

Arbeitseffizienz, Big Data, Open Data, Support Vector Regression

Die Effizienz eines landwirtschaftlichen Arbeitsprozesses hängt von vielen Parametern ab. Viele wissenschaftliche Arbeiten zu diesem Thema beschäftigen sich mit verschiedenen Ansätzen der Produktivitäts- und Effizienzbetrachtung. Die Ergebnisse solcher Betrachtungen hängen stark davon ab, wie Effizienz jeweils definiert bzw. unter welchem Aspekt sie betrachtet wird: So kann – unter anderem – die Ressourceneffizienz, die CO₂-Bilanz (LAL 2004), die Kosteneffizienz, der Ertrag pro Hektar (KUDALIGAMA und YANAGIDA 2000), die Energieeffizienz einzelner Prozesse (TOLL 2013) oder das Maß der Nachhaltigkeit einzelner Betriebe (PACINI et al. 2003) untersucht werden. Das Problem ist jedoch, dass bei den meisten Betrachtungen klare örtliche und zeitliche Grenzen definiert werden müssen. Diese liegen meist bei einzelnen Betrieben oder Regionen, denn Eigenschaften wie etwa Höhenprofil, Bodeneigenschaften oder Schlaggeometrie können ebenso unterschiedlich sein, wie die Witterungsbedingungen bei der Durchführung einzelner Arbeitsprozesse. Eine Übertragbarkeit der Untersuchungsergebnisse ist deshalb selten oder nur unter Einschränkungen möglich.

Diese Arbeit hat zum Ziel, die oben genannten Einflüsse durch ein Korrekturverfahren zu minimieren und die Effizienzwerte vergleichbarer zu machen. Es soll möglich sein, beispielsweise einzelne Arbeitsprozesse so vergleichen zu können, dass nur noch die bedienerabhängigen Effizienzunterschiede sichtbar und quantifizierbar sind. Es wird gezeigt, wie groß die Unterschiede schon innerhalb eines Betriebes sein können und wie gut sich ein Korrekturverfahren zum Angleichen der Effizienz eignet.

Das Verfahren wurde anhand von Prozessdaten aus dem Projekt „Big Data in landwirtschaftlichen Prozessen innovativ Nutzen“ (BiDaLAP) konzipiert. Das Projekt hat die Entwicklung einer elektronischen Infrastruktur, bestehend aus einer Plattformarchitektur mit mobilen Datenloggern, zum Ziel.

Diese Infrastruktur soll zukünftig als operatives und strategisches Entscheidungsunterstützungssystem zu Verfügung stehen.

Weiterhin wurden diverse offene Datenbanken in das Korrekturverfahren eingebunden, um möglichst viele Einflüsse quantifizieren zu können. So wurden beispielsweise Daten des Deutschen Wetterdienstes, des Deutschen Luft- und Raumfahrtzentrums sowie der Bundesanstalt für Geowissenschaften und Rohstoffe eingebunden. Alle genutzten Datenbanken werden weiter unten genauer beschrieben. Weiterhin werden das Korrekturvorgehen erläutert und die Ergebnisse im Anschluss vorgestellt.

Material und Methode

Während des Projektes BiDaLAP wurde ein Probetrieb in Sachsen mit mobilen Datenloggern ausgestattet, mit denen die Aussaat von Mais in den Jahren 2017 und 2018 aufgezeichnet wurde. Beispielfhaft werden in diesem Beitrag alle Arbeiten analysiert und verglichen, die mit einer Einzelkornsämaschine von Väderstad durchgeführt wurden. Nach Bereinigung der GPS-Daten, die im 1-Hz-Takt aufgezeichnet wurden, konnten 48 Einsätze mit 2 verschiedenen Schleppern auf 40 verschiedenen Schlägen analysiert werden. Als Maß der Effizienz für diesen Prozess wird die Flächenleistung ha/h gewählt (Abbildung 1), da weder die Menge des ausgebrachten Saatguts noch der Kraftstoffverbrauch der Maschinen aus den Daten hervorgeht. Auch ist die Qualität der Arbeit nicht bekannt.

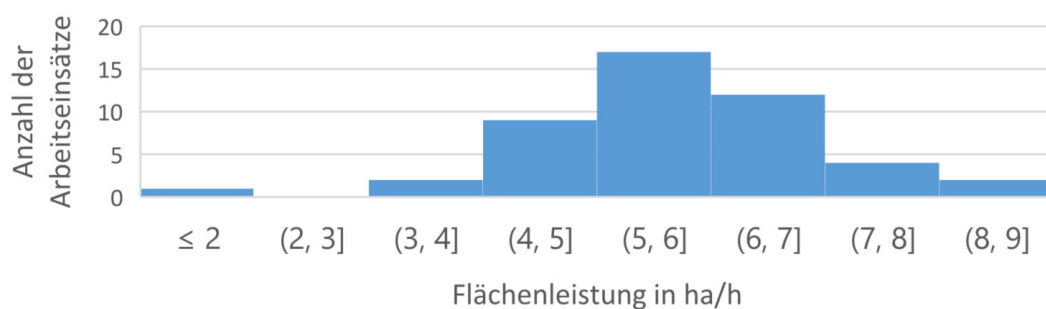


Abbildung 1: Verteilung der Flächenleistung der Einsätze mit der Einzelkornsämaschine

Es ist deutlich zu erkennen, dass die Arbeitseffizienz der einzelnen Prozesse stark schwankt. Daher wurde untersucht, ob dies auf die Umgebungseigenschaften des Betriebes zurückzuführen ist oder etwa der Maschinenführer einen Einfluss darauf hat. Hierzu werden folgende offen zugängliche Datenbanken hinzugezogen:

- Deutscher Wetterdienst (DWD): Der Wetterdienst bietet eine Vielzahl an Daten, die nicht nur Wetterstationen beinhalten, sondern auch interpolierte quadratische Gitternetze mit 1 km Kantenlänge, die anhand der Daten der Stationen berechnet werden. Für den Ansatz werden die Bodenfeuchtigkeiten und die Niederschlagsmenge benötigt (STEIN und HENSCHEL 2019).
- Deutsches Zentrum für Luft und Raumfahrt (DLR): Mit den Radarsatelliten der Mission TANDEM-X wird ein hochgenaues, digitales 3D-Abbild der Erde aufgezeichnet. Es werden präzise Höhendaten in einem 12-m-Raster für die gesamte Erde gesammelt und in ein einheitliches Kartenmaterial überführt. Für den Ansatz kann mittels der Höhendaten die Intensität des Höhenprofils abgeschätzt werden.

- Bundesanstalt für Geowissenschaften und Rohstoffe (BGR): Die BGR bietet umfassende Bodenkartierungen verschiedener Genauigkeiten an oder stellt Verbindung zu den Landesämtern her. Für das Modell wird die Bodenart mit dazugehörigen Sand-, Ton- und Schluffanteilen benötigt (DÜWEL et al. 2007).

Die einzigen nicht offenen Daten, die vom Landwirt direkt bezogen werden, sind die GPS-Fahrspuren sowie Schlagdaten und -Geometrien. Aber auch die Schlagdaten könnten aus dem Geoportal des entsprechenden Bundeslandes übernommen werden. Die Kombination dieser offenen Datenbanken und der betriebsinternen GPS-Fahrspuren machen den Ansatz dieser Arbeit aus.

Als nächstes werden möglichst viele Umgebungsparameter aus den Datenbanken extrahiert, die einen Einfluss auf die Flächenleistung haben können. Einige Werte wurden nicht direkt aus den Datenbanken entnommen, sondern sind nachträglich berechnet worden, um Eigenschaften der Umgebung optimal widerspiegeln zu können (Tabelle 1).

Tabelle 1: Umgebungsparameter mit den entsprechenden Korrelationskoeffizienten zur Flächenleistung (grün: positive Korrelation, rot: negative Korrelation)

Parameter	Einheit	Korrelationskoeffizient	Parameter	Einheit	Korrelationskoeffizient
Intensität des Höhenprofils	m	-0,07	Sandanteil im Boden	%	0,20
Bodentemperatur	°C	0,01	Jahr		-0,04
Bodenfeuchtigkeit	%	0,01	Monat		0,04
Tagesniederschlag	mm	0,25	Spurlänge	m	0,52
Tonanteil im Boden	%	-0,19	Komplexität der Feldgeometrie		-0,14
Schluffanteil im Boden	%	-0,19	Bearbeitete Fläche	ha	0,46

Die Operationalisierung bestimmter Parameter erfolgte wie unten beschrieben:

- Intensität des Höhenprofils:
Die Intensität des Höhenprofils wurde mittels der Höhendaten des DLR berechnet. Die Intensität des Höhenprofils kann gut mittels der Standardabweichung Steigung wiedergegeben werden. Je höher diese ist, desto höher ist die Höhenänderung während des Einsatzes. Dazu wird die Steigung entlang der GPS-Spur bestimmt und deren Standardabweichung berechnet.
- Spurlänge:
Die Spurlänge bezieht sich auf die gefahrene Strecke während der Arbeiten im Feld zwischen zwei Wendungen. Zur automatischen Ermittlung wurden die GPS-Daten durch einen Algorithmus in Arbeitspunkte und Wendepunkte eingeteilt und die Strecke zwischen zwei Wendungen berechnet. Als Vergleichswert wurde hierbei die mittlere Spurlänge gewählt.
- Komplexität der Feldgeometrie: Die Unterschiede in der Feldgeometrie zeigt Abbildung 2. Je winkelter ein Feld ist, desto länger dauern die Bearbeitungen. Um alle Teile des Feldes vollständig befahren zu können, müssen zusätzliche Fahrmanöver durchgeführt werden. Das benötigt zusätzliche Arbeitszeit und somit verringert sich die Flächenleistung. Zur Quantifizierung wird die Anzahl der Winkeländerungen der Feldgeometrie gewählt, die bei komplexen Geometrien größer ist als bei einfachen.

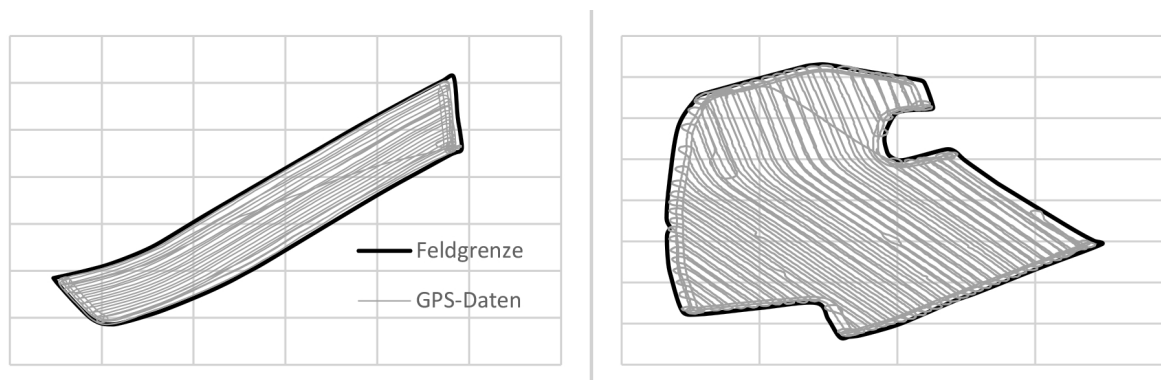


Abbildung 2: Darstellung zwei verschiedener Feldgeometrien mit GPS-Spur-Daten, links einfache Geometrie (Wert = 4), rechts komplexe Geometrie (Wert = 11)

Zunächst wurden die einzelnen Variablen jeweils ihrem linearen Korrelationskoeffizienten zur Flächenleistung gegenübergestellt. Dieser zeigt die Abhängigkeit zweier Größen voneinander und hat einen Wertebereich von -1 bis 1. 0 bedeutet keine Abhängigkeit der Größen, -1/1 starke lineare Abhängigkeit. Weiterhin müssen die p-Werte jeder Variablen berechnet werden, um eine Aussage über die Signifikanz auf die Flächenleistung treffen zu können (ROSS und HEINISCH 2006).

Die Korrelationskoeffizienten in Tabelle 1 zeigen, dass der Einfluss auf die Flächenleistung der Parameter unterschiedlich ist. Die Koeffizienten variieren von nahezu keinem Zusammenhang zwischen Bodentemperatur und Flächenleistung bis hin zu einem deutlich erkennbaren Zusammenhang zwischen Spurlänge und Flächenleistung.

Zu diesen natürlichen Einflüssen auf die Flächenleistung kommen die Einflüsse hinzu, die auf die Steuerung der Maschine zurückzuführen sind (Tabelle 2). Es ist erkennbar, dass die Koeffizienten deutlich höher ausfallen als bei den natürlichen Parametern, was aber auch naheliegend ist, da sich diese direkt auf die Arbeitszeit und somit auf die Flächenleistung auswirken. Die aufgelisteten Parameter wurden hierbei ausschließlich aus den GPS-Fahrspuren und den dazugehörigen Zeitstempeln berechnet.

Die p-Werte werden mittels eines t-Tests für alle Parameter berechnet. Dabei ergab sich für fast alle aufgelisteten Parameter ein p-Wert kleiner als 0,01, was für eine hohe Signifikanz spricht (BORTZ und DÖRING 2006). Nur die p-Werte für die Dauer der Stopps pro Stunde und der Komplexität der Feldgeometrie sind größer als 0,05. Dies spricht für einen weniger ausgeprägten statistischen Zusammenhang zwischen diesen Variablen und der Flächenleistung.

Tabelle 2: Einflussparameter, die vom Bediener abhängen und der Korrelationskoeffizient zur Flächenleistung

Parameter	Einheit	Korrelationskoeffizient	Parameter	Einheit	Korrelationskoeffizient
Anzahl Stopps pro Stunde		-0,685	Tatsächliche Arbeitsbreite	m	0,638
Ø Dauer der Stopps	s	-0,029	Arbeitsgeschwindigkeit	m/s	0,828
Wendezeit	s	-0,667			

Wird die Spreizung der Flächenleistungen der einzelnen Prozesse betrachtet, stellt sich die Frage, wie die beiden Parametergruppen idealerweise getrennt werden können. Eine Möglichkeit ist ein passendes Regressionsmodell. Ein Regressionsmodell spiegelt den funktionalen Zusammenhang zwi-

schen mehreren unabhängigen Variablen $X_1 \dots X_N$ und einer abhängigen Variablen Y wider (DODGE and JUREČKOVÁ 2000). In dieser Arbeit wurde zuerst ein linearer Zusammenhang untersucht (Gleichung 1):

$$Y = a_0 + a_1 \cdot X_1 + \dots + a_n \cdot X_n \quad (\text{Gl. 1})$$

Die Parameter a_0 bis a_n sind dabei die sogenannten Regressionsparameter, die den Einfluss einer Variablen in dem Modell beschreiben. Als abhängige Variable wird die Flächenleistung und als unabhängige Variablen werden die bis hierher aufgezählten Umwelt- und Bedienerparameter gewählt. Mit den unabhängigen Variablen kann eine Zielfunktion gebildet werden, die die Flächenleistung abschätzt. Es zeigte sich nach der ersten Analyse, dass sowohl bei einem linearen als auch bei einem nichtlinearen Ansatz der mittlere quadratische Fehler zu groß wurde. Der mittlere quadratische Fehler ist ein zentrales Qualitätskriterium in der Statistik. Er stellt die Streuung des Modells um den Erwartungswert dar. Der Fehler betrug in diesem Fall $1,782 \text{ ha}^2/\text{h}^2$ für den linearen Ansatz und $1,340 \text{ ha}^2/\text{h}^2$ für den nichtlinearen Ansatz. Damit schieden diese Ansätze aus.

Alternativ wurde eine Support Vector Regression (SVR) angewendet. Diese ermöglicht die Lösung des Regressionsproblems im höherdimensionalen Raum mithilfe von sogenannten Kernelfunktionen. Die genaue Funktionsweise der Support Vector Regression erklären DOLAN (2019), STEINWART und CHRISTMANN (2008) sowie CHERKASSKY und DHAR (2010), weswegen an dieser Stelle nicht genauer darauf eingegangen wird. Als Maß der Genauigkeit der Regression wird auch hier der mittlere quadratische Fehler gewählt, der die Abweichung zwischen aufgezeichneter Flächenleistung und Regressionsmodell wiedergibt. Als weiteres Qualitätsmerkmal kann das Bestimmtheitsmaß R^2 hinzugezogen werden. Dieses Maß gibt an, wie gut die einzelnen unabhängigen Variablen geeignet sind, die Varianz der abhängigen Variable zu erklären (BORTZ und DÖRING 2006). Ein Nachteil der SVR ist, dass sich die Modellparameter nicht so einfach wie in der Formel 1 darstellen lassen, sondern dass sich diese in den Stützvektoren verbergen und eine Darstellung an dieser Stelle nicht zielführend wäre.

Um nun nur den Einfluss der Bedienerparameter quantifizieren zu können, muss die Support Vector Machine der SVR mittels der vorhandenen Einsätze trainiert werden. Anschließend werden alle Einsätze auf die gleichen Umweltparameter gemittelt und es wird mit dem Regressionsmodell eine „korrigierte“ Flächenleistung bestimmt. Dabei kann gezeigt werden, wie groß die Abhängigkeit der Flächenleistung von den Bedienerparametern ist und welche den größten Einfluss haben.

Zuletzt wird aufgezeigt, wie groß das mögliche Einsparungspotenzial von Zeit und Kosten sein kann, wenn der entsprechende Parameter optimiert wird. Dies geschieht anhand einer Beispielrechnung mit einem 10 ha großen Feld, was der durchschnittlichen Feldgröße des betrachteten Betriebes entspricht. Zur Kostenberechnung wird der Feldarbeitsrechner des KTBL (2018) sowie die Erfahrungssätze für überbetriebliche Maschinenarbeiten verwendet (UPPENKAMP 2018).

Ergebnisse

Zunächst wird verglichen, wie gut das Regressionsmodell mit den vorliegenden Parametern die tatsächliche Flächenleistung darstellt. Abbildung 3 zeigt die Ergebnisse für jede Aufzeichnung mit dazugehörigem Modellwert. Es ist zu erkennen, dass eine nutzbare Übereinstimmung vorliegt. Dies spiegelt sich auch im mittleren quadratischen Fehler $0,219 \text{ ha}^2/\text{h}^2$ wider. Die durchschnittliche prozentuale Abweichung beträgt $0,4 \%$ und die Standardabweichung 5% . Weiterhin beträgt das

Bestimmtheitsmaß R^2 81 %, was einen akzeptablen Wert für solch eine Betrachtung darstellt. Das Modell besteht nach dem Training aus 43 Support-Vektoren, in denen die Gewichtung von jeweils 20 Parametern hinterlegt ist. Die höhere Genauigkeit des Modells ist auf die bessere Handhabung nichtlinearer und nichtmonotoner Daten zurückzuführen sowie der Eigenschaft einer SVM, Regressionsprobleme gut zu generalisieren (STEINWART and CHRISTMANN 2008). Es wird angenommen, dass das Regressionsmodell mithilfe einer Support Vector Machine dazu geeignet ist, den Arbeitsprozess einer Einzelkornsämaschine abzubilden.

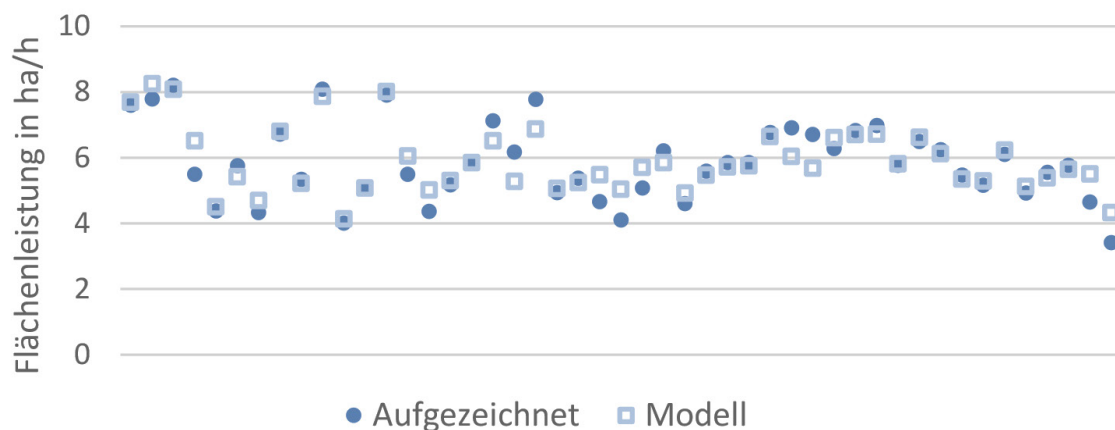


Abbildung 3: Vergleich zwischen aufgezeichneter Flächenleistung und Regressionsmodell

Im nächsten Schritt werden die Umweltparameter angepasst. Dazu werden jedem aufgezeichneten Prozess die gemittelten Parameter aller Prozesse zugeordnet. Anschließend wird die Flächenleistung mit dem Regressionsmodell, den angepassten Umweltparametern und den ursprünglichen Bedienerparametern berechnet. Abbildung 4 zeigt die Verteilung der aufgezeichneten und der korrigierten Flächenleistung. Die Box wird vom oberen und unteren Quartil eingegrenzt, das gesamte Ausmaß stellt Minimum und Maximum der betrachteten Größe dar. Es ist zu erkennen, dass die Spreizung der korrigierten Flächenleistung im Gegensatz zur aufgezeichneten deutlich kleiner geworden ist. Die ursprüngliche Leistung schwankt zwischen 3,5 und 8,1, die korrigierte nur noch zwischen 4,5 und knapp 7,0. Es spiegelt die Abweichung der Flächenleistung wider, die ausschließlich auf den Bediener zurückzuführen ist. Dies bedeutet ebenfalls, dass die Umweltparameter einen erheblichen Einfluss auf die Flächenleistung haben und bei einer Effizienzbetrachtung von Arbeitsprozessen immer mitberücksichtigt werden müssen. Der Vollständigkeit halber ist auch das Modellergebnis ohne angepasste Umweltparameter dargestellt. Es ist erkennbar, dass hier noch nicht das Optimum erreicht wurde, da das Modell anscheinend die weite Streuung der Beobachtungen nicht optimal wiedergibt.

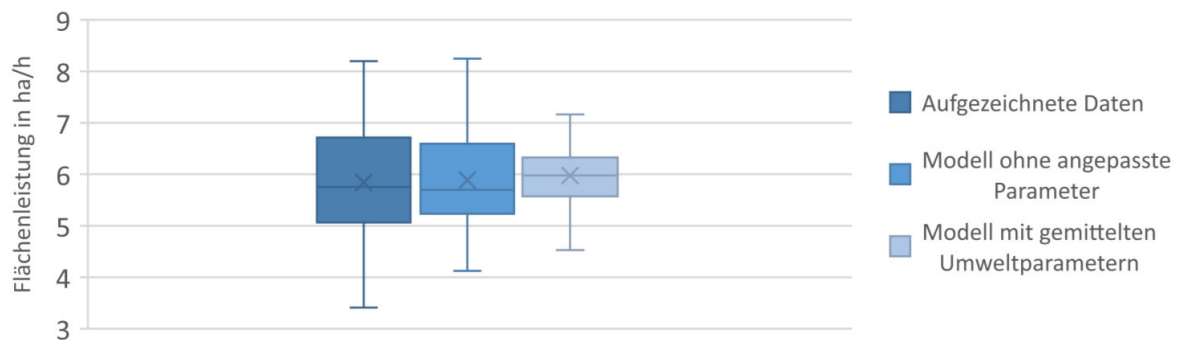


Abbildung 4: Kastendiagramm der Flächenleistung der tatsächlichen Beobachtungen sowie die berechnete Flächenleistung der SVM ohne und mit gemittelten Umweltparametern

Um nun den Einfluss der einzelnen Bedienerparameter zu ermitteln, wird im nächsten Schritt das Modell mit den ursprünglichen Umweltdaten parametrisiert und nur jeweils ein Bedienerparameter zum bestmöglichen Wert verändert. Um das Optimierungspotenzial nicht zu überschätzen wird hierbei vom 90 %-Quantil ausgegangen. Es wird gezeigt, wie groß die Steigerung der Flächenleistung ausfallen könnte. Die Werte ergeben sich aus den mittleren Steigerungen für alle Beobachtungen. Die Ausgangswerte sind die tatsächlich beobachteten Parameter, die Steigerung bezieht sich auf die Anpassung des jeweiligen Optimierungsparameters. Die einzelnen Parameter zeigen unterschiedliche Auswirkungen auf die Flächenleistung. Diese sind schon aus den Korrelationskoeffizienten in Tabelle 2 ablesbar, dies wird in Tabelle 3 erneut verdeutlicht. Das geringste Optimierungspotenzial nach Tabelle 2 liegt bei den Stopps pro Stunde, dem folgt die durchschnittliche Wendezeit und die tatsächliche Arbeitsbreite. Am meisten Potenzial bietet anscheinend die Fahrgeschwindigkeit. Die errechneten Kostenersparnisse in Tabelle 3 ergeben sich aus der verringerten Arbeitszeit und die damit verbundenen Lohnkostenersparnisse sowie den durchschnittlichen Schlepperkosten pro Stunde ohne Kraftstoff.

Tabelle 3: Effizienzänderung und Ersparnisse für eine Einzelkornsämaschine, wenn jeweils einer der angegebenen Parameter dem 90 %-Quantil entspricht; Rechnungen für ein 10 ha großes Feld

	Stopps pro Stunde	Wendezeit	Tatsächliche Arbeitsbreite	Fahrgeschwindigkeit
Gesteigerte Flächenleistung	0,21	0,15	0,11	0,32
Zeitersparnis	9,7 min	7,1 min	5,2 min	14,5 min
Kostenersparnis	5,50 €	3,98 €	2,94 €	7,41 €

Die Lohnkosten werden mit 17 €/h für einen Schlepperfahrer angenommen. Die Schlepperkosten belaufen sich je nach Größe zwischen 15 und 23 € pro Stunde. Werden die Ersparnisse addiert, ergibt sich ein direktes Einsparpotenzial von einer knappen halben Stunde sowie Lohn- und Schlepperkosten von knapp 20 €, was etwa 2 % der Gesamtkosten des Prozesses entspricht. Für Einsparungen, die nur durch optimiertes Fahrverhalten ermöglicht werden, ist dies ein akzeptabler Wert.

Tabelle 3 zeigt durchschnittliche Effekte der dargestellten Parameter auf die Flächenleistung. Diese Effekte beschreiben die Änderungen einer abhängigen Variablen, der Flächenleistung, durch eine unabhängige Variable. Dabei werden alle anderen unabhängigen Variablen konstant gehalten. Ein Vergleich dieser Effekte für alle in Tabelle 1 und 2 genannten unabhängigen Parameter mit den dazu-

gehörigen Korrelationskoeffizienten stellt ein weiteres Indiz für die Validität des Modells dar (Tabelle 4). Die Effekte wurden mit dem 90%-Quantil des entsprechenden Wertes bestimmt. Es ist zu erkennen, dass die Tendenzen von Korrelationskoeffizienten und der Änderung der Flächenleistung sehr ähnlich sind, nur die Intensität stimmt teilweise nicht überein. Dieser Umstand gilt besonders für die Dauer der Stopps und die Komplexität der Feldgeometrie. Dies könnte auf die relativ hohen p-Werte dieser unabhängigen Parameter zurückzuführen sein.

Tabelle 4: Vergleich der Korrelationskoeffizienten und durchschnittliche Effekte der einzelnen unabhängigen Parameter; marginale Effekte wurden mit dem 90%-Quantil berechnet

Parameter	Korrelationskoeffizient	Änderung	Parameter	Korrelationskoeffizient	Änderung
Anzahl Stopps pro Stunde	-0,685	-0,204	Tonanteil im Boden	-0,190	-0,008
Ø Dauer der Stopps	-0,029	-0,064	Schluffanteil im Boden	-0,190	0,013
Wendezeit	-0,667	-0,081	Sandanteil im Boden	0,200	0,115
Tatsächliche Arbeitsbreite	0,638	0,127	Jahr	-0,040	-0,093
Arbeitsgeschwindigkeit	0,828	0,323	Monat	0,040	0,149
Intensität des Höhenprofils	-0,070	-0,042	Spurlänge	0,520	0,232
Bodentemperatur	0,010	0,054	Komplexität der Feldgeometrie	-0,140	0,084
Bodenfeuchtigkeit	0,010	0,107	Bearbeitete Fläche	0,460	0,252
Tagesniederschlag	0,250	0,193			

Bei den Untersuchungen zur Effizienz ist zu beachten, dass Größen wie die Fahrgeschwindigkeit auch einen enormen Einfluss auf die Arbeitsqualität haben können (HANNUSCH et al. 1984). Arbeitsqualität kann ebenfalls ein Maß für die Effizienz eines Prozesses sein, was definitionsabhängig ist. Die Arbeitsqualität ist weiterhin von der Art des Prozesses und vom Arbeitsziel abhängig. Beim Säen ist es die gleichmäßige Verteilung und optimale Menge von Saatgut, die wiederum auch von den Geräteeinstellungen abhängen. Beim Pflügen sind neben der Flächenleistung die gesamten Arbeitskosten sowie das gewünschte Bodenbild wichtige Maße für die Arbeitsqualität. Wie im Abschnitt Material und Methoden erläutert, wurde im Rahmen der Arbeit nicht darauf eingegangen, da hierfür keine Daten zu Ressourcen wie Kraftstoff und ausgebrachtes beziehungsweise eingeholtes Gut zur Verfügung standen. Falls bei zukünftigen Untersuchungen diese Daten ebenfalls vorliegen, muss die Definition der Effizienz erweitert werden. Als Zielgröße könnten dann beispielsweise die Prozesskosten pro Hektar bestimmt werden und – mit einer Optimierungsfunktion – das ideale Kosten-Nutzen-Verhältnis zwischen Arbeitsqualität und –zeit sowie Kosten.

Schlussfolgerungen

Das Verfahren hat gezeigt, dass eine Korrektur von Effizienzwerten unter Zuhilfenahme von öffentlichen Datenbanken und einer Support Vector Regression möglich ist. Es konnten ebenfalls einzelne

Parameter auf ihr Optimierungs- und Einsparungspotenzial untersucht werden. In nachfolgenden Untersuchungen sollten jedoch weitere Parameter, wie der Kraftstoffverbrauch sowie die Qualität der Feldarbeit betrachtet werden.

Zukünftig gilt es nachzuweisen, wie gut sich das Vorgehen zum Vergleich von Arbeitsprozessen gleicher Art in verschiedenen Betrieben eignet. Weiterhin muss untersucht werden, welche Prozesse sich für die Analyse eignen und wo die Schwächen des Ansatzes liegen. Umfangreichere Datensätze mit zusätzlichen Heterogenitäten könnten zu einer erhöhten Zahl von Supportvektoren führen und das Modell verbessern sowie die Fehlergrößen verkleinern. Vorangegangene Arbeiten haben entsprechende Erfahrungen hierzu geliefert (STEIN und MEYER 2018). Eine Nutzung des Verfahrens als Unterfunktion für ein Entscheidungsunterstützungssystem ist ebenfalls im Laufe des Projektes BiDaLAP angedacht.

Literatur

- Bortz, J.; Döring, N. (2006): Forschungsmethoden und Evaluation. Für Human- und Sozialwissenschaftler, mit 87 Tabellen, Heidelberg, Springer-Medizin-Verlag
- Cherkassky, V.; Dhar, S. (2010): Simple Method for Interpretation of High-Dimensional Nonlinear SVM Classification Models. In: Ed. Stahlbock, R.; Crone, S.F.; Abou-Nasr, M.; Arabnia, H.R.; Kourentzes, N.; Lenca, P.; Lippe, W.M.; Weiss, G.M, Proceedings of The 2010 International Conference on Data Mining, DMIN 2010, July 12-15, 2010, Las Vegas, Nevada, USA. CSREA Press
- Dodge, Y.; Jurečková, J. (2000): Adaptive Regression. New York, Springer New York
- Dolan, H. (2019): A Practical Guide to Interpreting and Visualising Support Vector Machines. <https://towardsdatascience.com/a-practical-guide-to-interpreting-and-visualising-support-vector-machines-97d2a5b0564e>, Zugriff am 28.05.2019
- Düwel, O.; Siebner, C.S.; Utermann, J.; Krone, F. (2007): Bodenarten der Böden Deutschlands. https://www.bgr.bund.de/DE/Themen/Boden/Produkte/Schriften/Downloads/Bodenarten_Bericht.pdf?__blob=publicationFile, Zugriff am 30.10.2018
- Hannusch, L.; Lubadel, O.; Frießleben, R.; Koschitzke, E.; Jeske, A. (1984): Einfluß der Arbeitsgeschwindigkeit und der Bodenfreiheit von Mineraldüngerstreuern und Pflanzenschutzmaschinen auf den Ertrag. In: Agrartechnik. Landtechnische Zeitschrift der DDR. Hg. Kammer der Technik, Berlin, VES Verlag Technik, S. 479-484
- KTBL (2018): KTBL-Feldarbeitsrechner. <https://daten.ktbl.de/feldarbeit/entry.html#0>, Zugriff am 05.09.2018
- Kudaligama, V.P.; Yanagida, J.F. (2000): A Comparison of Intercountry Agricultural Production Functions: A Frontier Function Approach. *Journal of Economic Development* 25, pp. 57-74
- Lal, R. (2004): Carbon emission from farm operations. *Environment international* 30(7), pp. 981-990, <https://doi.org/10.1016/j.envint.2004.03.005>
- Pacini, C.; Wossink, A.; Giesen, G.; Vazzana, C.; Huirne, R. (2003): Evaluation of sustainability of organic, integrated and conventional farming systems: a farm and field-scale analysis. *Agriculture, Ecosystems & Environment* 95(1), pp. 273-288, [https://doi.org/10.1016/S0167-8809\(02\)00091-9](https://doi.org/10.1016/S0167-8809(02)00091-9)
- Ross, S.M.; Heinisch, C. (2006): Statistik für Ingenieure und Naturwissenschaftler. München, Elsevier Spektrum Akad. Verlag
- Stein, T.; Henschel, T. (2019): Potenziale von Open Data für die Effizienzsteigerung von mobilen Arbeitsmaschinen. In: 39. GIL-Jahrestagung: Digitalisierung für landwirtschaftliche Betriebe in kleinstrukturierten Regionen – ein Widerspruch in sich? 18.-19. Februar 2019 in Wien, Österreich. Hg. Meyer-Aurich, A., Gandorfer, M., Barta, N., Gronauer, A., Kantelhardt, J. & Floto, H., Bonn, Gesellschaft für Informatik e.V., S. 251-256
- Stein, T.; Meyer, H.J. (2018): Automatic machine and implement identification of an agri-cultural process using machine learning to optimize farm management information systems. In: 6th International Conference on Machine Control and Guidance, 01.-02.10.2018, Berlin, Bornimer Agrartechnische Berichte, Heft 101, S. 19-26

Steinwart, I.; Christmann, A. (2008): Support vector machines. New York, Springer Verlag

Toll, C.v. (2013): Energieorientierte Analyse der Landmaschinentechnik. Untersuchung zur maschinenrelevanten Energiebilanzierung in der Getreideproduktion mit Erfassung von CO₂-Einsparpotenzialen. Dissertation, Technische Universität Berlin

Uppenkamp, N. (2018): Erfahrungssätze für Maschinenring-Arbeiten unter Landwirten. <https://www.landwirtschaftskammer.de/landwirtschaft/beratung/pdf/erfahrungssaetze-rh.pdf>, Zugriff am 05.06.2019

Autor

Thoralf Stein, M.Sc. Fahrzeugtechnik, ist wissenschaftlicher Mitarbeiter am Fachgebiet Konstruktion von Maschinensystemen der Technischen Universität Berlin, Straße des 17. Juni 144, 10623 Berlin. E-Mail: thoralf.stein@tu-berlin.de

Danksagung

Das Projekt BiDaLAP wird vollständig vom Bundesministerium für Landwirtschaft und Ernährung gefördert.