

Wege zu einer Qualitätsbewertung digitaler und technischer Lösungen in der Landwirtschaft

Frank Beneke, Andreas Gabriel, Doreen Gabriel, Markus Gandorfer, Kai-Hendrik Howind, Ralf Kalmar, Johann Meierhöfer, Patrick Noack, Yves Reckleben, Ingolf Römer, Ulrich Rubenschuh, Oliver Schmittmann, Nikolaus Staemmler, Christina Umstätter, Martin Weis, Burkhard Wrenger



Frank Beneke, Georg-August-Universität Göttingen, Departement für Nutzpflanzenwissenschaften, Abteilung Agrartechnik

Andreas Gabriel und Markus Gandorfer, Bayerische Landesanstalt für Landwirtschaft, Institut für Landtechnik und Tierhaltung

Doreen Gabriel, Institut für Pflanzenbau und Bodenkunde, Julius Kühn-Institut, Braunschweig

Kai-Hendrik Howind, Landwirtschaftskammer Niedersachsen

Ralf Kalmar, Fraunhofer-Institut für Experimentelles Software Engineering IESE

Johann Meierhöfer, Fachbereich Pflanzliche Erzeugung / Energie, Deutscher Bauernverband e. V.

Patrick Noack, Kompetenzzentrum für Digitale Agrarwirtschaft (KoDA), Hochschule Weihenstephan-Triesdorf

Yves Reckleben, Fachhochschule Kiel - Fachbereich Agrarwirtschaft, FG: Landtechnik

Ingolf Römer, Universität Leipzig, Institut für Wirtschaftsinformatik

Ulrich Rubenschuh, DLG e.V.

Oliver Schmittmann, Institut für Landtechnik, Rheinische Friedrich-Wilhelms Universität Bonn

Nikolaus Staemmler, Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie

Christina Umstätter, Thünen-Institut für Agrartechnologie, Braunschweig

Martin Weis, Landwirtschaftliches Technologiezentrum Augustenberg, Karlsruhe

Burkhard Wrenger, Technische Hochschule Ostwestfalen-Lippe, Höxter

Danksagung

Die Idee zu der vorliegenden Arbeit ist im Rahmen des Kompetenznetzwerks „Digitalisierung in der Landwirtschaft“ unter Begleitung von BMEL und BLE entstanden.

Titelfoto: BMEL

Thünen Working Paper 246

Braunschweig/Germany, 24.09.2024

Zusammenfassung

Die digitale Transformation in der Landwirtschaft vollzieht sich mit neuen Technologien und Verfahren. Damit verbunden sind Erwartungen an Produktivitätssteigerungen, Optimierung des Ressourceneinsatzes oder verbesserte Nachhaltigkeitseffekte. Für etablierte Techniken gibt es bereits standardisierte Prüf- und Testverfahren, deren Ergebnisse es den Landwirt*innen ermöglichen, die Technik zu bewerten. Bei digitalen Technologien oder Lösungen fehlen solche Testverfahren bisher oder sind nicht standardisiert. Deshalb wurden in einer vom Bundesministerium für Ernährung und Landwirtschaft (BMEL) eingesetzten und durch die Bundesanstalt für Landwirtschaft und Ernährung (BLE) begleiteten Arbeitsgruppe der Stand der Technik, die aktuellen Herausforderungen und mögliche Lösungsansätze diskutiert und Handlungsempfehlungen in Richtung Wirtschaft, Politik und Forschung abgeleitet.

Die zugrunde gelegten Ziele waren dabei:

- Verlässlichkeit und Risiken digitaler Technologien für die Landwirt*innen transparent und verständlich darstellen,
- Transparente und verständliche Bewertung der Güte der digitalen Techniken (z.B. Angaben zum „geprüften“ Einsatzraum, Vorhersagefehler),
- Bewertung der Funktionsfähigkeit, Verfügbarkeit und Anwendbarkeit,
- Einschätzung, ob das Produktversprechen eingehalten wird,
- Berücksichtigung von Sicherheitsaspekten bzw. Abschätzung bestehender Risiken.

In Bezug auf die aktuellen digitalen Technologien wurden insbesondere hochautomatisierte Maschinen für landwirtschaftliche Prozesse sowie Entscheidungsunterstützungssysteme betrachtet. Bei Letzteren kommen zunehmend auch KI-Verfahren zum Einsatz.

Für die Qualitätsbewertung solcher digitalen Technologien wurden folgende allgemeine Handlungsempfehlungen abgeleitet:

- Transparente Informationsbereitstellung für den/die Landwirt*in bezüglich der Erfüllung des Produktversprechens und der Schnittstellen,
- Freiwillige Qualitätsbewertung, außer bei gesetzlich vorgeschriebenen Prüfungen,
- Neu- und Weiterentwicklung von Mess- und Prüfverfahren für Praxisbedingungen.

Die nachfolgenden Kapitel bieten dazu einen detaillierteren Einblick und eine Verfeinerung der Handlungsempfehlungen zu Mess- und Prüfverfahren.

Schlagworte

Produktversprechen, Standardisierung, Prüfverfahren, Testverfahren, Handlungsempfehlungen

Zusammenfassung	1
1 Einführung und Einordnung	3
1.1 Einleitung	3
1.2 Rahmenklärung und Begriffsdefinitionen	4
1.3 Ziele und Handlungsempfehlungen für die Bewertung digitaler Techniken in der Landwirtschaft	8
1.4 Literatur	9
2 Aktueller Stand der Prüfung von Agrartechnik	9
2.1 Überblick über verfügbare Prüfverfahren und Parameter für die Technik in der Pflanzenproduktion	9
2.2 Prüfverfahren Technik im Pflanzenbau - Überblick über bestehende Verfahren	10
2.2.1 Bodenbearbeitungsgeräte	10
2.2.2 Sätechnik	10
2.2.3 Düngetechnik	11
2.2.4 Pflanzenschutztechnik	12
2.2.5 Erntetechnik	13
2.2.6 Praxisprüfung von Sensoren und Güte der korrespondierenden Vorhersagemodelle	14
2.2.7 Datenkommunikationssysteme	14
2.3 Neu- und Weiterentwicklungsbedarf für Mess- und Prüfverfahren	15
2.4 Literatur	15
3 Qualitätsmerkmale von Software, Sensoren und Datenprodukten	17
3.1 Qualitätsmerkmale im landwirtschaftlichen Kontext	17
3.2 Qualitätsmerkmale für Systeme und Software	19
3.3 Die veränderte Rolle der Sensoren in digitalisierten Prozessen	21
3.4 Schnittstellen im landwirtschaftlichen Kontext	21
3.4.1 ISOBUS	21
3.4.2 Interoperabilität zwischen Farm Management Systemen	22
3.4.3 GeoBox und Hofbox	22
3.5 Literatur	23
4 Statistische Herausforderungen bei Vorhersagemodellen im Kontext der Digitalen Landwirtschaft	24
4.1 Was sind Vorhersagemodelle und welche Rolle spielt KI dabei?	24
4.2 Anwendung von KI in der Landwirtschaft und anderen Bereichen	25
4.3 Externe Validierung	26
4.4 Statistische Beurteilungskriterien	26
4.5 Limitierungen von Vorhersagemodellen	28
4.6 Literatur	29
5 Bestandsaufnahme digitaler Technologien in den Experimentierfeldern	31
5.1 Ergebnisse	31
5.1.1 Bewertungsfelder der untersuchten Technologien	32
5.2 Auswahl der Use Cases	35
5.2.1 NIRS-Technologie	35
5.2.2 Feldrobotik	36
5.2.3 Verhaltensüberwachung/Tracking Tier	36
6 Mögliche Mindestanforderungen an digitale Technologien	37

1 Einführung und Einordnung

Die deutsche Ackerbaustrategie 2035 soll hinsichtlich der Nahrungsmittel-, Futtermittel- und Rohstoffversorgung sowie den umwelt-, natur- und klimapolitischen Herausforderungen eine Perspektive für den Ackerbau in den kommenden 15 Jahren bieten. Dabei spielt zunehmend auch die Digitalisierung eine bedeutende Rolle. Eine geforderte Maßnahme ist es, die Möglichkeiten zur Einführung einer Qualitätsbewertung für digitale Technologien in der Landwirtschaft zu prüfen. In dem vorliegenden Positionspapier werden Wege zu einer Qualitätsbewertung digitaler Lösungen in der Landwirtschaft aufgezeigt und die unterschiedlichen Aspekte in einzelnen Kapiteln vertieft. Dabei sind die sichere Funktionsweise und die Qualitätsbeurteilung unter realistischen und praktischen Einsatzbedingungen von entscheidender Bedeutung.

Landwirtschaft und Digitalisierung sind für sich genommen bereits komplex. Diese Komplexität steigert sich weiter, wenn beide Bereiche zusammenkommen. Dies gilt auch für die eingesetzte Technik – oder Technologie? Diese Begriffe werden nicht immer gleichwertig verwendet, genau so wenig wie andere Schlüsselwörter aus dem Umfeld der (digitalen) Landwirtschaft, wie Nachhaltigkeit, Resilienz, Interoperabilität usw. Aus diesem Grund wird dem Positionspapier ein einleitendes Kapitel zur Klärung des Rahmens, ausgewählter Begriffe und der Ziele vorangestellt.

1.1 Einleitung

Die digitale Transformation in der Landwirtschaft verspricht neue Möglichkeiten einer nachhaltigeren Landwirtschaft, durch Effizienzgewinne und Optimierungen. Dies könnte zu einer gesteigerten Umweltschonung (ökologische Komponente) und einer verbesserten Erlössituation oder günstigeren Preisen führen (ökonomische Komponente) sowie Betriebsleiter*innen entlasten und Entscheidungshilfen bereitstellen (soziale Komponente).

Wichtige Bausteine bei der Anwendung digitaler Technologien in der Landwirtschaft sind z.B. die zeitliche und räumliche Erfassung von Messwerten (z.B. Ausbringungsmengen, Bodeneigenschaften, Kraftstoffverbrauch, Stickstoffaufnahme und Niederschläge) sowie die Automatisierung von Prozessen (z.B. Lenksysteme, automatisierte Teilbreitenschaltung, Mengen- und Abstandsregelung). Die Messwerte erhalten einen konkreten Wert, wenn sie z.B. mit Hilfe von Auswertelgorithmen oder Modellen in nutzbare Informationen, Handlungsoptionen für Menschen (Entscheidungsunterstützung) oder Eingangsdaten für Maschinen (Automatisierung) umgewandelt und in einen konkreten Kontext gestellt werden.

Bisher gibt es in der Landwirtschaft kaum Regelungen, Maßstäbe oder Prüfrahmen für die Qualität der (1) (Sensor-)Messwerte und (2) Modelle für deren weitere Verarbeitung. Es fehlt somit trotz potentieller ökonomischer und ökologischer Risiken eine Qualitätsbewertung der Leistung oder der Genauigkeit von Gesamtsystemen und deren Komponenten. Zudem werden in den kommenden Jahren etablierte analoge und innovative digitale Ansätze parallel existieren. Bislang führt diese Koexistenz eher zu Verunsicherungen bei Anwender*innen. Die mangelnde Transparenz wirkt sich einerseits negativ auf das Vertrauen der Anwender*innen und deren Investitionsbereitschaft aus. Andererseits wird eine gezielte und effiziente staatliche Förderpolitik für neue Technologien deutlich erschwert.

Dieses Positionspapier zeigt die Bedeutung der Qualitätsbewertung von digitalen Produkten im landwirtschaftlichen Umfeld und Wege zu deren Umsetzung auf.

Kapitel 1 charakterisiert die Zielsetzung dieses Positionspapiers und gibt einen kurzen Abriss über das Gesamtbild der bearbeiteten Herausforderungen des Risikomanagements und der Qualitätssicherung in der Landwirtschaft. In Bezug auf einen Blick in andere Branchen wird eine risikobezogene Vorgehensweise vorgeschlagen. Das Kapitel schließt mit den identifizierten Zielen und Handlungsempfehlungen für die Bewertung digitaler Techniken in der Landwirtschaft ab.

Kapitel 2 behandelt die technische und fachliche Prüfung von Maschinen und Geräten, die teils hoheitlich vorgeschrieben und teils freiwillig von unabhängigen Prüfkommisionen - aus Wissenschaft, Beratung und

Landwirtschaft - im Auftrag der Unternehmen durchgeführt werden. Die nachfolgenden Ausführungen beschränken sich auf die Überprüfung der Funktionalität und Arbeitsqualität. Arbeitssicherheitsaspekte und verkehrsrechtliche Anforderungen sind in anderen Regelwerken reglementiert. Ebenso sind Fragen der Praktikabilität und sinnvollen Anwendung verschiedener Technologien an anderer Stelle zu diskutieren. Die Anforderungen an die Funktionalität und Arbeitsqualität der verschiedenen Maschinengruppen wird dabei unabhängig vom Zug- bzw. Trägerfahrzeug (manuell über Fahrer gesteuert, teilautonom oder autonom gesteuert) betrachtet.

Kapitel 3 behandelt Qualitätsmerkmale von Software und Datenprodukten, denn eine standardisierte Charakterisierung dieser Qualitätsmerkmale ist essenziell für ein einheitliches Verständnis eines übergeordneten Qualitätsbegriffs. Dabei gibt es unterschiedlichste Qualitätsaspekte, die das Qualitätsempfinden der unterschiedlichen Stakeholdergruppen beeinflussen. Um die Zuverlässigkeit und Effizienz digitaler Lösungen sicherzustellen, können spezifische Anforderungen an diese einen Rahmen für die systematische Bewertung und Verbesserung der Qualität von Software- und Datenprodukten, wie z.B. Vorhersagesysteme oder Ertragskarten, in der Landwirtschaft bieten.

In **Kapitel 4** werden verschiedene statistische Aspekte zur Beurteilung der Verlässlichkeit von Entscheidungsunterstützungssystemen, die auf Vorhersagemodellen und dem Einsatz von KI-Verfahren basieren, beleuchtet. Vorhersagemodelle werden dabei als Algorithmen verstanden, die durch funktionale Verknüpfungen mit Eingangswerten Schätzungen von Erwartungswerten erzielen. Es werden konkrete Anwendungsbeispiele aus Tierhaltung und Pflanzenproduktion in der Landwirtschaft, aber auch aus anderen Bereichen, wie der Medizin, vorgestellt. Dabei sind die Erwartungen an die Zuverlässigkeit und Genauigkeit von Vorhersagemodellen in der Regel hoch. Modelle, die in Wissenschaft und Praxis eingesetzt werden, müssen ausreichend dokumentiert und vor allem validiert sein, sodass die Anwender*innen über den Vorhersagefehler und mögliche Limitationen im Anwendungsbereich informiert sind.

Kapitel 5 gibt auf Basis einer Befragung einen Überblick über die in den digitalen Experimentierfeldern untersuchten digitalen Technologien sowie den dabei angewandten Untersuchungsmethoden. Drei ausgewählte Anwendungsfälle (Nahinfrarotspektroskopie (NIRS), Feldrobotik, Verhaltensüberwachung) werden genauer betrachtet und bewertet.

Kapitel 6 definiert mögliche Mindestanforderungen an digitale Systeme in landwirtschaftlichen Anwendungsbereichen, die sich über Genauigkeit, Kosten, Kompatibilität bis zur Benutzerfreundlichkeit erstrecken. Die praxisrelevante Anwendbarkeit der Technologien sowie damit zusammenhängende Bewertungsaspekte werden auf Basis von Kapitel 5 dargestellt.

1.2 Rahmenklärung und Begriffsdefinitionen

In anderen Anwendungsdomänen, wie dem autonomen Fahren, haben sich das Risikomanagement und die Verifikation und Validierung in Bezug auf damit verbundene Maßnahmen als zentrales Instrument zur Qualitätssicherung durchgesetzt (vgl. insbesondere die Pegasus-Projektfamilie des Verbandes der Automobilindustrie e.V. (VDA) (PEGASUS Projekt, 2020) oder die strategische Forschungs- und Innovationsmaßnahme SUNRISE der EU (SUNRISE, s.a.)). Das Risiko-Management lässt sich auf allen Ebenen anwenden bzw. durch die Verifikation gegenüber der nächsthöheren Ebene durchführen (Abbildung 1). Dabei ist ein Risiko-Management nicht auf Safety oder Cybersecurity begrenzt, auch wenn es am häufigsten auf diese Qualitätsaspekte angewandt wird.

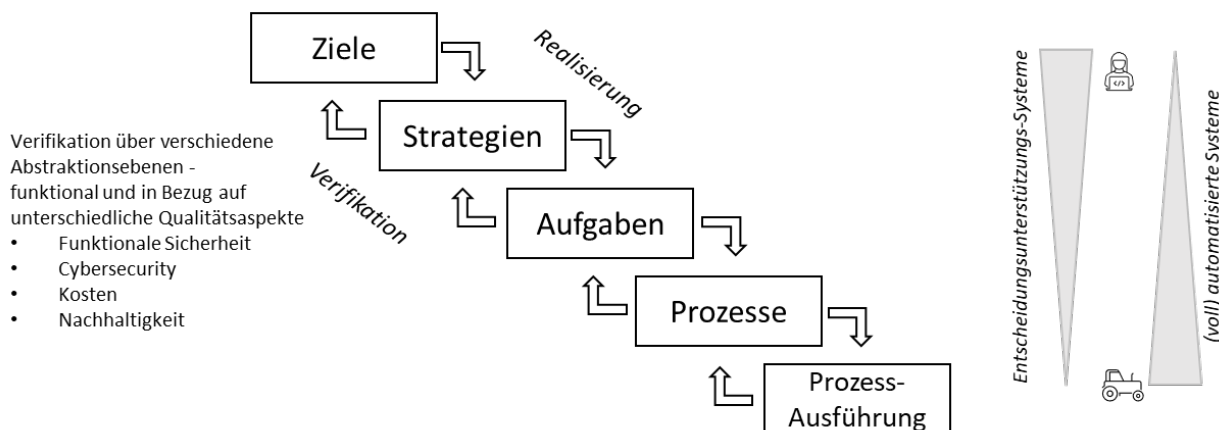


Abbildung 1: Abstraktionsebenen bei der Digitalisierung in der Landwirtschaft, die jeweils Software-Anteile enthalten - vom Steuergerät in der Landmaschine bis zum Finanzmanagement

Das Risikomanagement hat dabei die Aufgabe, mögliche (negative) Einflüsse auf die Qualitätsaspekte zu identifizieren und zu quantifizieren, sowie Maßnahmen zu definieren, welche die Auswirkungen auf ein akzeptables Maß reduzieren. In den unterschiedlichen Stufen des Produktlebenszyklus müssen entsprechende Unzulänglichkeiten deshalb identifiziert werden (Abbildung 2).

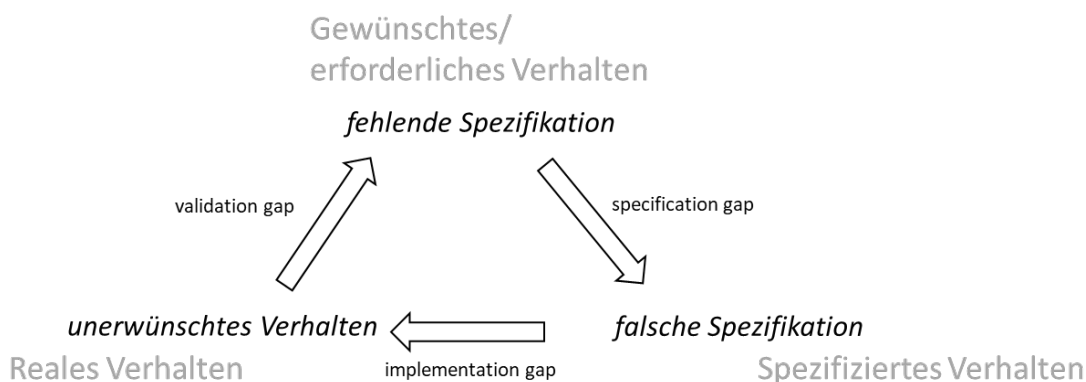


Abbildung 2: Die Abweichung des Verhaltens von Systemen hat unterschiedliche Ursachen im Produktlebenszyklus und wiederholt sich iterativ über Produktgenerationen

Das gewünschte bzw. erforderliche Verhalten wird dabei auf der obersten Abstraktionsebene vom Kunden oder auch vom Gesetzgeber bzw. der Gesellschaft vorgegeben (z.B. „Welche Umweltschädigung wird noch toleriert?“). Diese Spezifikation wird in Gesetzen oder in Normen und Standards dokumentiert. Aber auch für die Umsetzung wird der gute Stand der Praxis üblicherweise in Normen und Standards dokumentiert und durch Unternehmen oder unabhängige Prüforganisationen (Testing, Inspection and Certification (TIC)-Industrie) überprüft.

In der Landwirtschaft lassen sich hierfür als Beispiele allgemein die ISO 25119 (ISO, 2018) für die Sicherheit landwirtschaftlicher Maschinen oder spezifischer z.B. die Düngemittelverordnung heranziehen. Je weiter man sich der konkreten Umsetzung von Systemen nähert, desto mehr bzw. stärker wird explizit auch auf die Software Bezug genommen – bis hin zu softwarespezifischen Normen, wie die ISO 250x0-Reihe. In der digitalen Transformation bestimmt die Software das Gesamtsystemverhalten maßgeblich mit. Daher muss die Software bzw. das

Systemverhalten auf allen Abstraktionsebenen betrachtet werden, da sich die Spezifikation auf die darunterliegenden Abstraktionsebenen überträgt und jeweils nur an den Schnittstellen der Abstraktionsebenen verifiziert bzw. validiert wird.

Die Komplexität der so geschaffenen Systeme lässt sich auf höheren Abstraktionsebenen nur mit entsprechenden Szenariospezifikationen und einer möglichen automatischen Validierung beherrschen. Für das automatisierte Fahren wurden dazu sechs Ebenen spezifiziert, die sich für die Landwirtschaft entsprechend adaptieren ließen, wie z.B. einer besonderen Betrachtung der Tier- und Pflanzenwelt als dynamische/veränderbare Objekte (vgl. Abbildung 3).

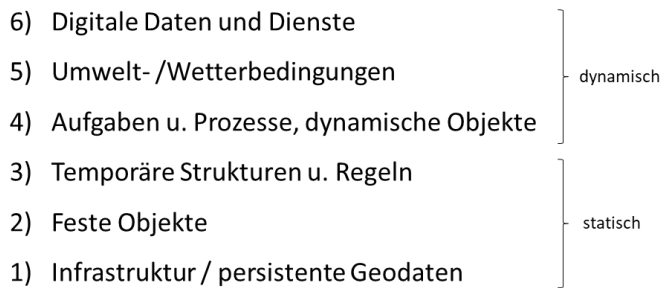


Abbildung 3: Abstraktionsebenen für eine Szenariospezifikation zur Validierung einzelner Aufgaben, wie z.B. der Düngung eines Feldes

Dabei lassen sich konkrete Szenarien immer einem definierten Arbeits- und Einsatzbereich zuordnen (Operational Design Domain, ODD). Dies kann z.B. ein Feld, ein Stall oder eine spezifische Tätigkeit sein. Erst die Festlegung solcher Szenarien ermöglicht eine entsprechende Qualitätsprüfung – bis hin zur Vergleichbarkeit der Leistungsfähigkeit unterschiedlicher Lösungen. In der Automobilindustrie sind dabei die dazugehörige Normung und Standardisierung für hochautomatisierte Fahrzeuge schon recht weit fortgeschritten (s. ISO 34503:2023, ISO, 2023).

In der Landwirtschaft entwickelt sich die Normung und Standardisierung für autonome Maschinen etwas langsamer (ISO, 2024, 18497; Verband der Elektrotechnik Elektronik und Informationstechnik e.V., 2021, VDE-AR-E 2842-61). Aus Gesamtsicht ist es erforderlich, Aktivitäten, wie die Entwicklung eines szenariobasierten Validierungsansatzes, anzugehen, bzw. die Vorgehensweise auf weitere Qualitätsaspekte, wie z.B. die Nachhaltigkeit, zu erweitern.

Die nachfolgende Abbildung 4 charakterisiert die szenariobasierte Validierung eines Softwareprodukts für die automatisierte Ausbringung von spezifisch berechneten Düngermengen. Zur Prüfung muss der landwirtschaftliche Prozess zunächst in seinen Grundzügen beschrieben und eingegrenzt sein (in welcher Kultur, zu welchem Wachstumsstadium, auf welcher Art Feld, bei welchem Versorgungszustand, unter welchen Wetterbedingungen etc.) und die gültigen Rahmenbedingungen (ab welcher Menge von Inhaltsstoffen kommt es z.B. zu unerwünschten Nebenbedingungen?) festgelegt sein. Mit diesem Rahmen kann eine automatisierte virtuelle Validierung, z.B. durch einen Dienstleister der Testing/Inspection/Certification-Industrie (TIC), erfolgen.

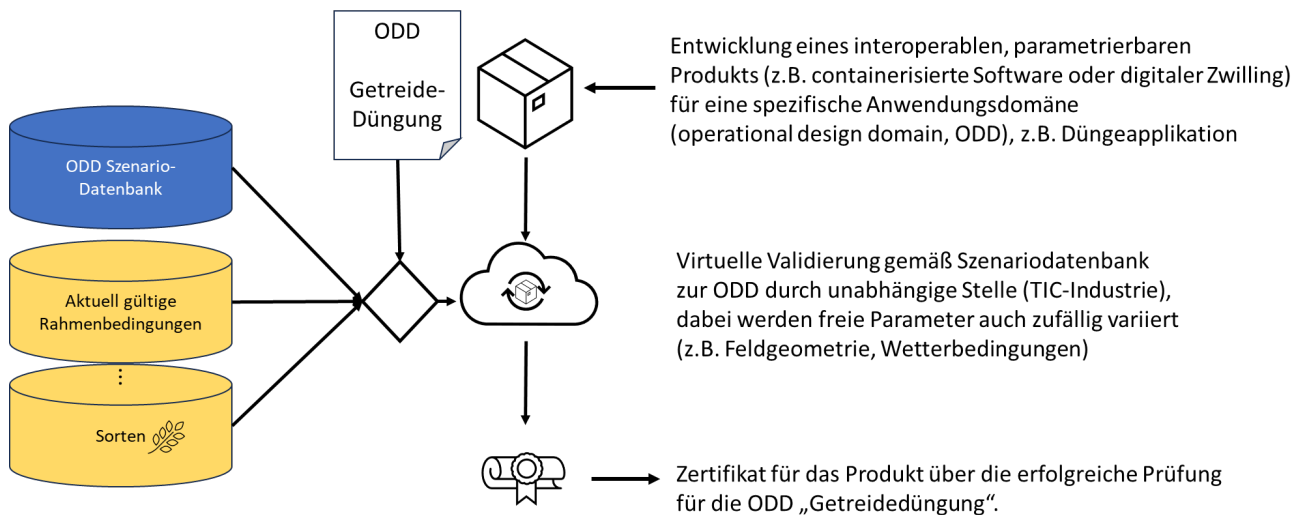


Abbildung 4: Prozess einer szenariobasierten Validierung am Beispiel „Beratungssoftware zur Getreidedüngung“

Gesetzlich vorgegeben sollten dabei nur Prüfungen sein, die im Sinne des Allgemeinwohls Sachverhalte überprüfen, die größeren Schaden anrichten können (z.B. Gefahr für Leib und Leben, massive Umweltschädigung, totaler Ernteverlust o.Ä.), um die Komplexität und den Aufwand der Prüfungen zu reduzieren. Gleichwohl könnte die szenariobasierte Simulation auch Leistungsparameter für eine Einordnung oder ein Benchmarking mit ausgeben, z.B. „mittlerer Düngemittelbedarf pro Hektar“.

In der landwirtschaftlichen Milchviehhaltung haben beispielsweise die unterschiedlichen Rassen, Haltungs- und Melksysteme einen sehr starken Einfluss auf die Aussagekraft der Entscheidungsunterstützungssysteme. Entscheidungsunterstützungssysteme müssen all diese Aspekte mitberücksichtigen, um für alle Betriebe gleichermaßen geeignet zu sein. Deshalb ist es wichtig, den „geprüften“ Einsatzraum genau zu spezifizieren. Ein weiterer Ansatz ist in einer höheren Flughöhe zu finden, die den landwirtschaftlichen Betrieben dennoch wichtige Hinweise für die Operabilität geben kann. Dazu gehört eine Einordnung der in den Tools genutzten Methoden hinsichtlich ihrer Eignung für die praktische Nutzung und unter welchen Bedingungen und für welche Bedingungen die Systeme entwickelt wurden. Dies ist ein wichtiger Schritt der Informationspolitik, der den Betrieben helfen kann, Kaufentscheidungen zu treffen. Die folgenden Kapitel ordnen sich in den zuletzt genannten Ansatz ein. Neben dem System und der Software haben die Eingangsdaten einen elementaren Einfluss auf die abgeleiteten Aussagen. Bei komplexen Sachverhalten müssen viele unterschiedliche Datenquellen integriert und korrekt miteinander verrechnet werden. Ein fehlerhaftes Eingangsdatum kann u.U. große Auswirkungen auf die Aussagekraft des finalen Produkts haben, wie z.B. eine Ertragskarte. Deshalb sollten entsprechende datenverarbeitende Produkte Maßnahmen enthalten, welche die Qualität der Eingangsdaten prüfen (z.B. Plausibilitäts-Checks, Zweitmessungen usw.) und Nutzer*innen oder angeschlossene Systeme z.B. über Konfidenzbereiche auf Unsicherheiten hinweisen. Dies ist umso wichtiger, wenn die Qualität der Daten mangels entsprechender Validierungsmöglichkeiten nicht bewertet werden kann. Eine transparente Berichterstattung über die Modellentwicklung, -ergebnisse, -grenzen und -interpretationen ist deshalb erforderlich.

Eine externe Validierung zur Bewertung der Güte von Vorhersagemodellen sollte nach Möglichkeit ein zentraler Bestandteil einer Qualitätssicherung sein. Diese Validierung ist von einer unabhängigen Instanz an neuen, dem Modell unbekanntem Daten durchzuführen, um geeignete Beurteilungskriterien für den Vorhersagefehler der Modelle abzuleiten. Welcher Vorhersagefehler jeweils für eine konkrete Anwendung akzeptabel ist, sollte auch unter Berücksichtigung von möglicher Zeit- und Kostenersparnis festgelegt werden.

Allgemein ist zu klären, welche Systeme schon verfügbar, für einen definierten Anwendungsfall geeignet und zuverlässig sind. Dementsprechend sind fehlende technologische Lösungsansätze gleichsam „weißen Flecken auf

der Karte“ aufzuzeigen und Bedarfe für eine zielgerichtete Produktentwicklung abzuleiten. Gleiches gilt auch für die erforderlichen Prüfverfahren zur Validierung der technischen Systeme. Zudem sollte der Anforderungskatalog an digitale Technologien regelmäßig auf Vollständigkeit geprüft werden und neben technischen und wirtschaftlichen Aspekten auch soziale Aspekte berücksichtigen. Die Prüfung sollte konsequent aufzeigen, ob bzw. in welchem Umfang das Produktversprechen eingehalten wird. Für eine verständliche Bewertung der Güte digitaler Technologien werden die Hauptkriterien Funktionsfähigkeit, Verfügbarkeit und Anwendbarkeit vorgeschlagen. Für eine Prüfung müssen geeignete Prüfdatensätze vorhanden sein. Es muss im Vorfeld klar vereinbart sein, wann eine Prüfung als bestanden bzw. nicht bestanden gilt. Hierfür ist die Definition eines Standards bzw. eines Qualitätsmaßstabes wichtig. Grundlage für die Erfüllung der Prüfkriterien ist der o.g. Anforderungskatalog und die Definition von Erfüllungsgraden.

Für die erforderlichen Prüfungen ist auch zu klären, welche Strukturen für die Prüfung bereits vorhanden bzw. welche Strukturen hierzu noch aufzubauen sind. Ebenso sind Zuständigkeiten für die Prüfung abzustimmen bzw. festzulegen (z.B. Industrie / Hersteller, Testzentren).

1.3 Ziele und Handlungsempfehlungen für die Bewertung digitaler Techniken in der Landwirtschaft

In der AG wurden fünf Ziele für die Bewertung digitaler Techniken in der Landwirtschaft identifiziert, die zukünftig verfolgt und erfüllt werden sollten. Um dies zu unterstützen, wurden drei Handlungsempfehlungen formuliert, die sich sowohl an die Politik als auch an die Wirtschaft und die Wissenschaft richten. Diese Stakeholder sollten dazu Hand in Hand zusammenarbeiten, um die digitale Transformation der Landwirtschaft erfolgreich zu unterstützen.

Ziele:

1. Verlässlichkeit und Risiken für die Landwirt*innen transparent und verständlich darstellen
2. Transparente und verständliche Bewertung der Güte der digitalen Techniken (z.B. Angaben zum „geprüften“ Einsatzraum, Vorhersagefehler)
3. Bewertung der Funktionsfähigkeit, Verfügbarkeit und Anwendbarkeit
4. Einschätzung, ob das Produktversprechen eingehalten wird
5. Berücksichtigung von Sicherheitsaspekten / Abschätzung bestehender Risiken

Handlungsempfehlungen:

1. Transparente Informationsbereitstellung für den Landwirt*in bezüglich der Erfüllung des Produktversprechens und der Schnittstellen
2. Freiwillige Qualitätsbewertung, außer bei gesetzlich vorgeschriebenen Prüfungen
3. Neu- und Weiterentwicklung von Mess- und Prüfverfahren für Praxisbedingungen:
 - a. Strukturen aufbauen oder anpassen, die diese Prüfungen unabhängig durchführen oder bewerten
 - b. Einsatzbereiche für verschiedene Szenarien standardisiert beschreiben (inkl. Stakeholderfeedback)
 - c. Identifikation von Qualitätsmaßstäben, Festlegung, Standardisierung und Normung.
 - d. Förderung offener Standards
 - e. Bereitstellung von Test- und Validierungsdatensätzen (einschl. Metadaten) für die Prüfungen
 - f. Die während der Prüfung generierten Daten von zu prüfenden Systemen sollten gemäß FAIR Kriterien zur Verfügung gestellt werden

1.4 Literatur

- PEGASUS Projekt, 2020. PEGASUS - Projekt zur Etablierung von generell akzeptierten Gütekriterien, Werkzeugen und Methoden sowie Szenarien und Situationen zur Freigabe hochautomatisierter Fahrfunktionen (Schlussbericht). Available at: URL: www.pegasusprojekt.de/files/tmpl/PDF/Pegasus_Abschlussbericht_Gesamtprojekt.pdf (accessed on: 07.02.2024).
- SUNRISE, s.a. Safety assurance framework for connected and automated mobility systems, Förderprojekt im Horizont Europa Programm der Europäischen Union, Research & Innovation Action No. 101069573, Available at: URL: <https://ccam-sunrise-project.eu> (accessed on: 07.02.2024).
- ISO, 2014. ISO/IEC 25000:2014 - Systems and software engineering, Systems and software Quality Requirements and Evaluation (SQuaRE).
- ISO, 2011. ISO/IEC 25010:2023 - Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality models.
- ISO, 2018. ISO 25119-1:2018 - Tractors and machinery for agriculture and forestry - Safety-related parts of control systems - Part 1: General principles for design and development.
- ISO, 2023. ISO 34503:2023 - Road Vehicles, Test scenarios for automated driving systems, Specification for operational design domain.
- ISO, 2024. ISO/FDIS 18497-4:2024 - Agricultural machinery and tractors, Safety of partially automated, semi-autonomous and autonomous machinery, Part 4: Verification methods and validation principles.
- Verband der Elektrotechnik Elektronik und Informationstechnik e.V. 2021. VDE-AR-E 2842-61: Anwendungsregel: 2021-07, Entwicklung und Vertrauenswürdigkeit von autonom/kognitiven Systemen, VDE-Verlag, Berlin.

2 Aktueller Stand der Prüfung von Agrartechnik

Frank Beneke, Kai Hendrik Howind, Johann Meierhöfer, Patrick Noack, Yves Reckleben, Ulrich Rubenschuh, Oliver Schmittmann, Martin Weis, Burkhardt Wrenger

2.1 Überblick über verfügbare Prüfverfahren und Parameter für die Technik in der Pflanzenproduktion

Um einen Überblick über die verfügbaren Prüfverfahren zu erhalten, wurden die wichtigsten Technologiegruppen aufgelistet, für die derzeit noch keine standardisierten Prüfverfahren und/oder Anforderungskataloge existieren. Anschließend wurden für diese Gruppen jeweils Hinweise zu Prüfkriterien von hohem Brancheninteresse aufgeführt.

Die Ausführungen beschränken sich auf die Überprüfung der Funktionalität und die Qualität des Arbeitsergebnisses unabhängig vom Zug- bzw. Trägerfahrzeug (manuell über Fahrer gesteuert, teilautonom oder autonom gesteuert). Die enthaltenen Prüfkriterien mit ihren korrespondierenden Anforderungen orientieren sich in den vorgestellten Prüfprogrammen in erster Linie an den Bedürfnissen der Praxis und berücksichtigen, sofern vorhanden, die normativen und gesetzlichen Vorgaben. Arbeitssicherheitsaspekte und verkehrsrechtliche Anforderungen sind an dieser Stelle nicht enthalten, da sie in anderen Regelwerken behandelt werden.

Grundsätzlich muss bei den aufgeführten Prüfverfahren zwischen den durch Normung geregelten Prüfungen sowie Anforderungskatalogen und den privatwirtschaftlichen Prüfverfahren unterschieden werden.

Bei den durch Europäische Normen (EN) geregelten Prüfungen sowie Anforderungskatalogen ist der Hersteller bzw. Inverkehrbringer verpflichtet, die Einhaltung der in den Normen enthaltenen Mindestanforderungen über seine Konformitätserklärung (CE) zu bestätigen. Alle anderen Prüfungen unterliegen bei der Durchführung einer Freiwilligkeit.

Die nachfolgende Übersicht erhebt keinen Anspruch auf Vollständigkeit. Über die vorgestellten Prüfprogramme hinaus existieren vielfach individuelle, herstellereigene Prüfverfahren, welche häufig nicht öffentlich zugänglich sind.

2.2 Prüfverfahren Technik im Pflanzenbau - Überblick über bestehende Verfahren

2.2.1 Bodenbearbeitungsgeräte

Prüfverfahren und Anforderungen an die Arbeitsqualität von Bodenbearbeitungsgeräten existieren als DLG-Prüfrahmen für

- Pflug
- Grubber
- Kreiseleggen inklusive Saatbettkombinationen

Die Arbeitsqualität beschreibende Kriterien sind

- Pflugbild
- Auflockerungs- und Krümelungswirkung
- Bodendurchmischung
- Einarbeiten bzw. Unterbringung von organischer Substanz
- Einebnungswirkung
- Tiefenführung und Tiefenhaltung der Werkzeuge
- Zuggleistungsbedarf und/oder Leistungsbedarf an der Zapfwelle
- Tatsächliche Arbeitsbreite
- Flächenleistung

2.2.2 Sätechnik

Prüfverfahren und Anforderungen an die Arbeitsqualität von Sätechniken existieren als DLG-Prüfrahmen für

- Drillmaschinen
- Einzelkornsämaschinen

Die Arbeitsqualität beschreibende Kriterien sind

- Reihenabstände
- Abstände in der Reihe
- Doppel- und Fehlstellen
- Ablagetiefe
- Wiederbedeckungsgrad
- Feldaufgang

Des Weiteren gibt es eine standardisierte Methode nach dem International Institute of Sugar Beet Research (IIRB, www.iirb.org), die eine leicht abgewandelte Form der Auswertung vorsieht. Im Gegensatz zur DLG wird unter Ablage- bzw. Standgenauigkeit nicht die Standardabweichung um den Ist-Abstand verstanden, sondern kulturspezifisch ein Konfidenzbereich festgelegt, deren Anteil an den Sollablagen angegeben wird (Bsp.: Ablagegenauigkeit ZR: +/- 1,5 cm um den Istabstand, Standgenauigkeit +/- 2,5 cm) (Schulze Lammers *et al.*, 2015)

2.2.3 Düngetechnik

2.2.3.1 Mineraldüngerstreuer

Prüfverfahren und Mindestanforderungen an die Arbeitsqualität von Mineraldüngerstreuer sind in den nachfolgenden Normen geregelt.

- DIN EN 13739-1: „Ausleger- und Wurf-Mineraldüngerstreuer - Umwelt - Anforderungen“ (DIN e.V., 2012a)
- DIN EN 13739-2: „Ausleger- und Wurf-Mineraldüngerstreuer - Umwelt - Prüfmethode“ (DIN e.V., 2012b)
- DIN EN 13740-1: „Reihen-Mineraldüngerstreuer - Umwelt - Anforderungen“ (DIN e.V., 2003a)
- DIN EN 13740-2: „Reihen-Mineraldüngerstreuer - Umwelt - Prüfmethode“ (DIN e.V., 2003b)

Keine maschinenspezifischen Prüfverfahren und Anforderungen existieren derzeit für

- Verteilqualität beim Grenzgrabenstreuen
- Verteilqualität beim Keilstreuen
- Verteilqualität bei Kurvenfahrten
- Streuverhalten bei Hangneigung
- Streuverhalten bei Seitenwind
- Genauigkeit von Wiegesystemen
- Funktionalität, Reaktionsgeschwindigkeit und Mengendosierung sowie Verteilqualität bei Teilflächenschaltungen
- Funktionalität von Kontroll- und automatisierten Einstelleinrichtungen
- (flächenspezifische) Dokumentation der ausgebrachten Düngermengen
- Nährstoffverteilung bei Düngermischungen
- Einfluss der Düngeigenschaften (Korngrößenspektrum usw.)

2.2.3.2 Dungstreuer

Prüfverfahren und Mindestanforderungen an die Arbeitsqualität von Dungstreuern sind in der nachfolgenden Norm geregelt.

- DIN EN 13080: „Stalldungstreuer - Umweltschutz - Anforderungen und Prüfmethode“ (DIN e.V., 2003c)

In der Norm DIN EN 13080 wird das Flächenstreuen beim Ausbringen von Stallmist behandelt. Andere Streugüter oder Applikationsszenarien sind in der Norm nicht geregelt.

Außerdem existiert ein DLG-Prüfrahmen für Dungstreuer, der auch Anforderungen für das Flächenstreuen von Kompost, Geflügelmist, Gärrest und Feuchtkalk enthält, und in welchem deutlich kritischere Anforderungen an die Verteilqualitäten beim Flächenstreuen gestellt werden. Zudem beschreibt der DLG-Prüfrahmen auch ein Verfahren zur Überprüfung der Genauigkeit von fahrzeugeigenen Wiegesystemen.

Keine maschinenspezifischen Prüfverfahren und Anforderungen existieren derzeit für

- Verhalten beim Grenzstreuen (Breitstreuerwerke)
- Funktionalität von Systemen zur dynamischen Durchflussmengenregelung
- Verfahren zur Inhaltsstofferrfassung org. Düngemittel (z.B. NIRS)

2.2.3.3 Gülleapplikationstechniken

Prüfverfahren und Mindestanforderungen an die Arbeitsqualität von Gülleapplikationstechniken sind in der nachfolgenden Norm geregelt.

- DIN EN 13406: „Flüssigmisttankwagen und Verteileinrichtungen - Umweltschutz - Anforderungen und Prüfmethoden für die Verteilgenauigkeit“ (DIN e.V., 2003d)

In der Norm DIN EN 13406 wird das Ausbringen von flüssigem Stallmist bei statischer Mengendosierung behandelt.

Darüber hinaus existiert ein DLG Prüfrahen für Gülleapplikationstechnik, in welchem deutlich kritischere Anforderungen an die Verteilqualitäten seitens der Praxis gestellt werden. Zudem beschreibt der DLG-Prüfrahen auch ein Verfahren zur Überprüfung der Genauigkeit von fahrzeugeigenen Wiegesystemen.

Keine maschinenspezifischen Prüfverfahren und Anforderungen existieren derzeit für

- Verteilqualität bei Hangneigung
- Systeme zur dynamischen Durchflussmengenregelung inklusive Reaktionszeiten und Einregelverhalten
- Teilflächenschaltungen inklusive Reaktionszeiten und Einregelverhalten
- Wiederbedeckungsgrad bei Schlitzgeräten (Emissionsschutz)
- Mengendosierung von Injektoren
- Dosiergenauigkeiten beim dynamischen Zumischen von ergänzenden Zusatzstoffen während der Applikation
- Detektion der Homogenität der Gülle / Entmischung

2.2.4 Pflanzenschutztechnik

2.2.4.1 Chemischer Pflanzenschutz

Anforderungen an Pflanzenschutzgeräte sind in der Norm DIN EN ISO 16119 (Neugeräte; ISO, 2013a, 2013b, 2015, 2018a, 2024) definiert. Dort wird für die Prüfung auf Einhaltung der Anforderungen auf verschiedene Normen verwiesen, in denen Testverfahren festgelegt sind.

In DIN EN ISO 16122 (ISO, 2018b) sind Anforderungen festgelegt, die bei der regelmäßigen amtlichen Kontrolle der in Gebrauch befindlichen Pflanzenschutzgeräte einzuhalten sind.

Die in den letzten Jahren zunehmende Verbreitung von elektronischen Assistenzsystemen bei Pflanzenschutzgeräten stellt besonders für die Gerätekontrolle, aber auch für die Prüfung von Neugeräten, eine große Herausforderung dar. Deshalb wird angeregt, Pflanzenschutzgeräte künftig mit Systemen zur On-Board-Diagnostik (OBD) auszurüsten, mit denen sich ähnlich wie bei Kraftfahrzeugen über genormte Schnittstellen und Fehlercodes die Funktion von Sensoren, Aktoren und Steuergeräten überprüfen lässt.

Verfahren zur Beschreibung der Arbeitsqualität beim Spotspraying sind derzeit in der Entwicklung. Ein Ansatz verfolgt hierzu die Bestimmung der Hit-Rate bei Spotsprayern mit der Fragestellung, wie genau die Zielpflanze mit Herbiziden getroffen wird (Schmittmann, 2024a und 2024b).

2.2.4.2 Hackmaschinen

Prüfverfahren und Anforderungen an die Arbeitsqualität von Hackmaschinen existieren als DLG-Prüfrahen.

Die Arbeitsqualität beschreibende Kriterien sind

- Bekämpfungserfolg

- Kulturpflanzenbeschädigung
- Führungsgenauigkeit (bei teilautonomen oder autonom betriebenen Hackmaschinen)

Derzeit werden die Anteile an eliminiertem Beikraut und an geschädigten Kulturpflanzen visuell bonitiert. Verschiedene Forschungsprojekte beschäftigen sich mit der Ermittlung der Effizienz von Beikrautbekämpfungsmaßnahmen über Drohnenflüge und Bildanalyse. Der Bekämpfungserfolg wird beispielsweise durch den Vergleich vom Zustand vor und nach der Behandlungsmaßnahme mittels KI beschrieben. Für die Auswertung werden dann die absoluten und relativen Differenzen von Kultur- und Beikrautpflanzen sowie deren Bedeckungsgrade vor und nach der Behandlung ermittelt (Schmittmann, 2024a und 2024b).

2.2.5 Erntetechnik

2.2.5.1 Mähdrescher

Prüfverfahren für Mähdrescher sind in den nachfolgenden Normen geregelt

- DIN 11390: „Landmaschinen - Prüfverfahren für Mähdrescher“ (DIN e.V., 1991)
- ISO 8210: „Equipment for harvesting - Combine harvesters - Test procedure“ (ISO, 2021b)

In den genannten Normen sind Begriffe und Methoden zur Bestimmung wichtiger Kenndaten von Mähdreschern beschrieben, welche sowohl Funktions- als auch Leistungsprüfungen beinhalten.

Eine wesentliche Kenngröße bei der Leistungsprüfung von Mähdreschern ist die Durchsatz-Verlust-Kennlinie. Begleitend hierzu werden auch die Fremdbestandteile im Erntegut sowie der durch den Ernteprozess erzeugte Anteil an Bruchkorn ermittelt.

In abgewandelter Form existiert auch ein korrespondierender DLG-Prüfrahmen.

Die in den Normen beschriebenen Verfahren sowie die davon abgewandelten Methoden zur Ermittlung der Durchsatz-Verlust-Kennlinien setzen eine statische Einstellung der Prüflinge voraus.

Keine maschinenspezifischen Prüfverfahren und Anforderungen existieren derzeit für

- Funktionalität von Assistenzsystemen zur dynamischen Anpassung der Maschineneinstellungen an die aktuellen Erntebedingungen
- Funktionalität und Anforderungen an die Vorhersagegenauigkeit von sensorgestützten Systemen bei der Bestimmung von Ertragsmenge, Feuchtegehalt, Korninhaltsstoffen, Fremdbestandteilen und Bruchkornanteilen

2.2.5.2 Feldhäcksler

Prüfverfahren für Feldhäcksler sind in der nachfolgenden Norm geregelt

- DIN ISO 8909-3: „Landmaschinen und Traktoren - Feldhäcksler - Teil 3: Prüfverfahren“ (ISO, 2021c)

In der genannten Norm sind Prüfverfahren für die Bewertung von Funktion und Leistung von Feldhäckslern festgelegt.

Darin behandelte, wesentliche Kenngrößen sind die Durchsatzleistung mit korrespondierenden Energiebedarfen, die erzeugten Schnittlängen und die Anteile an unzureichend aufbereiteten Körnern im Erntegut.

In abgewandelter Form existiert ein korrespondierendes DLG-Prüfverfahren für die Untersuchung der Durchsatzleistung, der Kraftstoffverbräuche und der technischen Häckselqualität an selbstfahrenden Feldhäckslern. Das in der Norm beschriebene Verfahren setzt eine statische Einstellung der Prüflinge voraus.

Keine maschinenspezifischen Prüfverfahren und Anforderungen existieren derzeit für

- Funktionalität von Assistenzsystemen zur dynamischen Anpassung der Maschineneinstellungen an die aktuellen Erntebedingungen (z.B. Häcksellängen Anpassung in Abhängigkeit vom Feuchtegehalt des Ernteguts, Steuerung der Siliermitteldosierung in Abhängigkeit von Futtereigenschaften u.a.)
- Funktionalität und Anforderungen an die Vorhersagegenauigkeit von sensorgestützten Systemen bei der Bestimmung von Ertragsmenge, Feuchtegehalt und Korninhaltsstoffen
- Technische Häckselqualität (Überlängen, Häcksellängenspektrum, angeschlagene Körner)

2.2.5.3 Zuckerrübenroder

Prüfverfahren und Anforderungen an die Arbeitsqualität von Zuckerrübenerntetechnik sind in einem Standard des International Institute of Sugar Beet Research (IIRB) beschrieben (Schulze Lammers *et al.*, 2015).

Die Arbeitsqualität beschreibende Kriterien sind

- Köpfqualität
- Erd- und Blattanhang
- ober und unterirdische Verluste
- Wurzelbruchverluste
- Oberflächenbeschädigungen

2.2.6 Praxisprüfung von Sensoren und Güte der korrespondierenden Vorhersagemodelle

Erste Verfahren zur Überprüfung und Bewertung von Sensor-gestützten Vorhersagemodellen wurden als DLG-Prüfrahmen für nachfolgende Applikationen entwickelt und angewendet:

- Mobile Sensoren am Auswurfkrümmer eines Feldhäckslers zur Bestimmung von Trockenmassegehalten in vorbeiströmendem Futter
- Mobile Sensoren zur Bestimmung von Inhaltsstoffen in vorbeiströmendem, flüssigen Wirtschaftsdünger
- Schnellmethoden zur Bestimmung von Bodenparametern.

2.2.7 Datenkommunikationssysteme

2.2.7.1 ISOBUS

Mit den Standardisierungen im Bereich des ISOBUS werden vor allem maschinennahe Systeme ausgestattet, die eine herstellerunabhängige Kombination von Komponenten ermöglichen. Geräte können basierend auf gemeinsamen Hardware- und Softwareschnittstellen miteinander kommunizieren und komplexe Vorgänge steuern. Die Prüfung auf Standardkonformität lässt sich automatisiert in unabhängigen Testzentren durchführen, beispielsweise bei der DLG. Die tatsächliche Kombinierbarkeit von Komponenten und die Anwendbarkeit von Austauschformaten (ISOXML) wird bilateral für einige Geräte geprüft und in einer Datenbank bei der AEF (Agricultural Industry Electronics Foundation) öffentlich dokumentiert (AEF isobus database, s.a.).

2.2.7.2 Bidirektionale Satellitenanlagen

Verfahren zur Überprüfung und Bewertung existieren als DLG-Prüfrahmen für portable sowie stationäre bidirektionale Satellitenanlagen. Zur Beschreibung der Performance dienen die nachfolgenden Kennwerte:

- Up- und Download-Geschwindigkeit

- Langzeit Ping-Test
- GEO-IP Blocking Test.

2.3 Neu- und Weiterentwicklungsbedarf für Mess- und Prüfverfahren

Es besteht ein erheblicher Bedarf daran, die Güte und Einsatzsicherheit unterschiedlichster Technik, Sensorsysteme und Schnittstellen unter praxisnahen Bedingungen zu überprüfen. Die Ergebnisse solcher Prüfungen bilden die Grundlage für ein tieferes Vertrauen der Praxis in die Technologie und deren Anwendbarkeit. Sie tragen damit zu einer weiteren Verbreitung der Umsetzung von digitalen Methoden in Landwirtschaft bei.

Die meisten der derzeit verfügbaren Prüfverfahren wurden in einer Zeit entwickelt, als die Maschineneinstellungen in erster Linie durch den Fahrer vorgenommen und im Bedarfsfall (selten) angepasst wurden. Daher setzen diese Prüfverfahren zumeist eine statische und im Versuch gleichzuhaltende Maschineneinstellung voraus. Die heute verfügbaren, sensorgestützten Assistenzsysteme machen es nun aber möglich, die Maschineneinstellungen während der Arbeitsprozesse dynamisch den vorherrschenden Bedingungen anzupassen und verschiedene Einstellstrategien auszuwählen. Dies kann auf unterschiedlichen Automatisierungsniveaus erfolgen, nämlich von der Einstellempfehlung an den Bediener, über die teilautomatisierte Nachregelung von Maschineneinstellungen bis hin zur autonomen Auftragsbearbeitung.

Die Funktionalität solcher Assistenzsysteme kann mit den klassischen Prüfverfahren nur sehr eingeschränkt überprüft werden. Daraus resultiert ein enormer Weiter- bzw. Neuentwicklungsbedarf für moderne Prüfverfahren, welche es erlauben, die Effekte von variierenden Maschineneinstellungen im Prozess zu erfassen, und darüber deren Zweckmäßigkeit zu beurteilen.

Ein großes Hemmnis besteht auch darin, dass bei den klassischen Prüfverfahren viele der die Arbeitsqualität bestimmenden Parameter mit hohem, manuellem Aufwand erhoben werden müssen. Ein erheblicher Zeit- sowie Personalbedarf für die Durchführung einer Prüfung ist die Folge daraus. Die Machbarkeit von Prüfungen unter Praxisbedingungen ist dadurch kostenintensiv und stark limitiert.

Massiver Entwicklungsbedarf besteht daher für Verfahren zur berührungslosen, sensorgestützten und standardisierten Erfassung von die Arbeitsqualität bestimmenden Faktoren, um eine ausreichende Anzahl an Prüfungen mit vertretbarem Aufwand durchführbar werden zu lassen und so das bestehende Informationsbedürfnis decken zu können.

Neben Einstellassistenten werden immer mehr maschineneigene sensorgestützte Systeme zur Schätzung und Dokumentation von Ernte- oder Ausbringmengen, Durchflüssen, Inhaltsstoffen und Zusammensetzungen eingesetzt. Standardverfahren zur Überprüfung der Genauigkeit solcher Schätzwerte und Vorhersagen sind aber bislang nicht in ausreichendem Maße etabliert. Auch hier besteht Entwicklungsbedarf.

Im Pflanzenschutz sind Prüfungen für neue Verfahren, wie Spot und Patch Applikation, zu entwickeln sowie die Prüfung von zunehmend eingesetzten Algorithmen und KI-Anwendungen zu etablieren (vgl. hierzu Kuska et. al 2022).

2.4 Literatur

AEF isobus database, s.a. AEF ISOBUS Database (website), <https://www.aef-isobus-database.org/> (besucht am 05.03.2024).

Beerbaum, S. et al. 2022. Positionspapier der Arbeitsgruppe "Adaptive autonome Agrarsysteme" im Kompetenznetzwerk Digitalisierung in der Landwirtschaft, https://www.bmel.de/SharedDocs/Downloads/DE/_Digitalisierung/positionspapier-agrarsysteme.html (besucht am 01. 03. 2024).

DLG-Prüfrahmen und DLG-Prüfverfahren (nicht öffentliche, DLG interne Dokumente)

- Pflug, Stand 04/2019
- Grubber, Stand 04/2019
- Kreiseleggen inklusive Saatbettkombinationen, Stand 04/2019
- Drillmaschinen, Stand 12/2023
- Einzelkornsämaschinen, Stand 04/2023
- Dungstreuer, Stand 02/2023
- Flüssigmist-Tankwagen und Verteileinrichtungen, Stand 02/2023
- Hackmaschinen, Stand 04/2018
- Selbstfahrender Feldhäcksler - Durchsatzleistung und Kraftstoffverbrauch
- Mobile Sensoren am Auswurfkrümmer eines Feldhäckslers zur Bestimmung von Trockenmassegehalten in vorbeiströmendem Futter, Stand 09/2009
- Mobile Sensoren zur Bestimmung von Inhaltsstoffen in vorbeiströmendem, flüssigen Wirtschaftsdünger, Stand 09/2020
- Schnellmethoden zur Bestimmung von Bodenparametern, Stand 09/2021
- Portable Satellitenanlagen, Stand 11/2017
- Stationäre Satellitenanlagen, Stand 10/2017

DIN e.V., 1991. DIN 11390:1991 - Landmaschinen; Prüfverfahren für Mähdrescher.

DIN e.V., 2012a. DIN EN 13739-1: 2012 - Landmaschinen - Ausleger- und Wurf-Mineraldüngerstreuer - Umweltschutz - Teil 1: Anforderungen.

DIN e.V., 2012b. DIN EN 13739-2: 2012 - Landmaschinen - Ausleger- und Wurf-Mineraldüngerstreuer - Umweltschutz - Teil 2: Prüfverfahren.

DIN e.V., 2003a. DIN EN 13740-1:2003 - Landmaschinen - Reihen-Mineraldüngerstreuer; Umweltschutz - Teil 1: Anforderungen.

DIN e.V., 2003b. DIN EN 13740-2:2003 - Landmaschinen - Reihen-Mineraldüngerstreuer; Umweltschutz - Teil 2: Prüfverfahren.

DIN e.V., 2003c. DIN EN 13080:2003 - Landmaschinen - Stalldungstreuer; Umweltschutz - Anforderungen und Prüfmethoden.

DIN e.V., 2003d. DIN EN 13406:2003 - Landmaschinen - Flüssigmisttankwagen und Verteileinrichtungen - Umweltschutz; Anforderungen und Prüfmethoden für die Verteilgenauigkeit.

ISO, 2013a. DIN EN ISO 16119-1:2013 - Agricultural and forestry machinery - Environmental requirements for sprayers - Part 1: General.

ISO, 2018a. DIN EN ISO 16119-2:2018 - Agricultural and forestry machinery - Environmental requirements for sprayers - Part 2: Horizontal boom sprayers.

ISO, 2013b. DIN EN ISO 16119-3:2013 - Agricultural and forestry machinery - Environmental requirements for sprayers - Part 3: Sprayers for bush and tree crops.

ISO, 2015. DIN EN ISO 16119-4:2015 - Agricultural and forestry machinery - Environmental requirements for sprayers - Part 4: Fixed and semi-mobile sprayers.

ISO, 2024. DIN EN ISO 16119-5:2024 - Agricultural and forestry machinery - Environmental requirements for sprayers - Part 5: Aerial spray systems.

ISO, 2018b. DIN EN ISO 16122-1:2018 - Agricultural and forestry machinery - Inspection of sprayers in use - Part 1: General.

ISO, 2018c. DIN EN ISO 16122-2:2018 - Agricultural and forestry machinery - Inspection of sprayers in use - Part 2: Horizontal boom sprayers.

- ISO, 2018d. DIN EN ISO 16122-3:2018 - Agricultural and forestry machinery - Inspection of sprayers in use - Part 3: Sprayers for bush and tree crops.
- ISO, 2018e. DIN EN ISO 16122-4:2018 - Agricultural and forestry machines - Inspection of sprayers in use - Part 4: Fixed and semi-mobile sprayers.
- ISO, 2021a. DIN EN ISO 16122-5:2021 - Agricultural and forestry machines - Inspection of sprayers in use - Part 5: Aerial spray systems.
- ISO, 2021b. ISO 8210:2021 - Equipment for harvesting - Combine harvesters - Test procedure and performance assessment.
- ISO, 2021c. DIN ISO 8909-3:2021 - Equipment for harvesting - Forage harvesters - Part 3: Test methods.
- Kuska, M. T., Heim, R. H. J., Geedicke, I., Gold, K., Brugger, A., Paulus, S. 2022. Digital plant pathology: a foundation and guide to modern agriculture. *Journal of Plant Diseases and Protection* 129, 457-468, DOI: 10.1007/s41348-022-00600-z.
- Schmittmann, O., Zimmer, P. 2024a. AI-based In-Field Weeding Quality Assessment as a potential Test Standard. In: *Proceedings of EurAgEng 2024 Conference*, 1. bis 4. Juli 2024 in Athen, Griechenland.
- Schmittmann, O., Zimmer, P., McCool, C. 2024bp. Development of a method for testing weeding-technologies in field – quality determination of spot-sprayer. *Book of Abstracts, 79th IIRB Congress in Brussels, 27th-28th February 2024*.
- Schulze Lammers, P., Vandergeten, J.-P., Tijink, F., Royer, C. 2015. *Test Procedures for Measuring the Quality in Sugar Beet Production - Seed Drillability, Precision Seeders, Harvesters, Cleaner Loaders*, 3rd edition IIRB, Göttingen 2015.

3 Qualitätsmerkmale von Software, Sensoren und Datenprodukten

Ralf Kalmar, Ingolf Römer und Martin Weis, Burkhard Wrenger

3.1 Qualitätsmerkmale im landwirtschaftlichen Kontext

In der Praxis der digitalen Landwirtschaft, insbesondere im Bereich des Precision Farming, stehen landwirtschaftliche Betriebe und Technologieanbieter vor der Herausforderung, die Variabilität und Unsicherheit in den Prozessen und Ergebnissen zu managen. Eine Schlüsselrolle spielen dabei die Qualitätsmerkmale der eingesetzten Software und Datenprodukte, welche auf einer guten betrieblichen Praxis der Software-Entwicklung aufsetzen. Eine standardisierte Bewertung dieser Qualitätsmerkmale ist essenziell, um die Zuverlässigkeit und Effizienz der digitalen Lösungen sicherzustellen. Hier setzen die ISO-Normen an, indem sie einen Rahmen für die systematische Bewertung und Verbesserung der Software- und Datenproduktqualität in der Landwirtschaft bieten können. Beispielhaft sind die ISO 25010 (ISO/IEC, 2023), ISO 25012 (ISO/IEC, 2008), ISO 25119 (ISO, 2018), ISO 15504 (ISO/IEC, 2012) und ISO 9001 (ISO, 2015b) als mögliche Werkzeuge zu nennen, um sicherzustellen, dass digitale Anwendungen in der Landwirtschaft nicht nur ihre vorgesehene Funktion, sondern auch Qualitäts- und Sicherheitsstandards erfüllen. Das besondere Augenmerk auf dem Entwicklungsprozess bei der Softwareentwicklung rührt daher, da dieser einen sehr großen Einfluss auf die Qualität hat - die Vielfältigkeit von Software selbst ist kaum fehlerbehaftet (im Gegensatz zur klassischen Güterproduktion). Diesem liegt die Annahme zugrunde, dass ein guter Prozess gute Produkte liefert.

Die ISO 15504, auch bekannt als SPICE (Software Process Improvement and Capability Determination), ist ein international anerkannter Standard für die Bewertung und Verbesserung von Softwareentwicklungsprozessen. Der Schwerpunkt liegt auf der Bestimmung des Reifegrads und der Fähigkeit von Softwareentwicklungsprozessen. Die ISO 15504 stellt einen Rahmen bereit, um Prozesse systematisch zu bewerten und Verbesserungspotenziale zu identifizieren. Sie umfasst mehrere Prozessdimensionen, darunter Prozessleistung, Prozessmanagement, kundenbezogene Prozesse und unterstützende Prozesse. Automotive SPICE (ASPICE) ist eine branchenspezifische Anwendung von ISO 15504, angepasst an die speziellen

Anforderungen der Automobilindustrie. Die Version 4.0 von ASPICE (VDA Working Group 13, 2023) zielt darauf ab, die Qualität und Zuverlässigkeit von Automobilsoftware sicherzustellen, indem sie die spezifischen Herausforderungen und Anforderungen dieser Branche adressiert. Wesentliche Aspekte sind die Einhaltung strenger Sicherheits- und Leistungsstandards sowie die Integration von Software in komplexe Fahrzeugsysteme. Deshalb fordern beispielsweise Automobilhersteller von ihren Zulieferern einen Nachweis über entsprechende Softwareentwicklungsfähigkeiten durch einen Mindestlevel „2“ in Schlüsselprozessen. Damit soll der Nachweis erbracht werden, dass Prozessziele erreicht werden und ein übergeordnetes Projektmanagement besteht.

Die Adaption von Automotive SPICE (VDA Working Group 13, 2023) für die Bewertung digitaler Anwendungen in der Landwirtschaft erfordert eine sorgfältige Übertragung und Modifikation der Qualitätsmerkmale von Software und Datenprodukten, die ursprünglich für die Automobilindustrie konzipiert wurden. Diese Transformation muss die einzigartigen Bedürfnisse und Herausforderungen der Landwirtschaft berücksichtigen, um eine effektive und effiziente Implementierung von Softwarelösungen in diesem Sektor zu gewährleisten. Vereinfachungen wären anwendbar z.B. im Bereich Unterauftragnehmermanagement. Ähnliche branchenspezifische Anpassungen bestehen auch für die Medizintechnik oder die Luft- und Raumfahrt, die ebenfalls hohe Qualitätsanforderungen an die in ihren Produkten enthaltene Software stellen.

Zu den wesentlichen Qualitätsmerkmalen, die im Kontext der landwirtschaftlichen Anwendungen neu interpretiert werden müssen, gehört zunächst die Funktionalität. In der Landwirtschaft ist die Funktionalität untrennbar mit der Genauigkeit und Zuverlässigkeit der Datenverarbeitung verbunden. Softwarelösungen in diesem Bereich müssen präzise, zeitnah und relevant für landwirtschaftliche Entscheidungen sein. Dies impliziert eine sorgfältige Kalibrierung der Systeme, um eine hohe Genauigkeit bei der Datenerfassung und -analyse zu gewährleisten, was für die Optimierung landwirtschaftlicher Prozesse entscheidend ist.

Ein weiteres zentrales Qualitätsmerkmal ist die Zuverlässigkeit der Software. Angesichts der Abhängigkeit der Landwirtschaft von externen Faktoren wie Wetter und Klima ist es unerlässlich, dass die Software auch unter extremen oder unvorhersehbaren Bedingungen funktioniert. Dies erfordert eine robuste Konstruktion und hohe Fehlertoleranz der Systeme, um kontinuierliche Betriebsbereitschaft und Datenintegrität unter allen Bedingungen zu gewährleisten.

Die Benutzbarkeit ist ein weiteres entscheidendes Qualitätsmerkmal. Da landwirtschaftliche Nutzer möglicherweise nicht über umfangreiche technische Detailkenntnisse der verwendeten Systeme verfügen, muss die Software benutzerfreundlich, intuitiv und leicht zugänglich sein. Eine klare Benutzeroberfläche und einfache Navigation sind unerlässlich, um eine breite Akzeptanz und effektive Nutzung in der landwirtschaftlichen Gemeinschaft zu gewährleisten.

Effizienz ist ebenfalls ein kritisches Merkmal, das sich auf den sparsamen Umgang mit Ressourcen wie Rechenleistung, Speicherplatz und Energie bezieht. In vielen landwirtschaftlichen Betrieben, insbesondere in abgelegenen oder ressourcenbeschränkten Gebieten, ist dies von großer Bedeutung. Effiziente Software kann dazu beitragen, die Betriebskosten zu senken und die Nachhaltigkeit zu erhöhen.

Wartbarkeit und Anpassungsfähigkeit der Software sind entscheidend, um den sich ständig ändernden Bedingungen und Anforderungen in der Landwirtschaft gerecht zu werden. Die Software sollte so gestaltet sein, dass sie leicht aktualisiert, modifiziert und gewartet werden kann, um neue landwirtschaftliche Praktiken oder Technologien zu integrieren und sich ändernde Anwenderanforderungen abzubilden.

Portabilität ist ein weiteres wichtiges Merkmal, insbesondere im Hinblick auf die Kompatibilität der Software mit verschiedenen Geräten und Plattformen. In der modernen Landwirtschaft, in der mobile Anwendungen und IoT-Geräte eine immer wichtigere Rolle spielen, ist die Fähigkeit, Software nahtlos über verschiedene Geräte hinweg zu nutzen, von entscheidender Bedeutung. Die Sicherheit von Software und Datenprodukten ist in der Landwirtschaft besonders relevant, da sie mit der Nahrungsmittelproduktion eine kritische Infrastruktur darstellt. Dies umfasst den Schutz vor unautorisiertem Zugriff, Cyberangriffen, Datenverlust und anderen Sicherheitsrisiken.

Schließlich ist die Compliance mit gesetzlichen Vorschriften und Standards, insbesondere in Bezug auf Datenschutz und gesetzliche Vorgaben, ein wesentliches Qualitätsmerkmal. Softwarelösungen in der Landwirtschaft müssen den rechtlichen Rahmenbedingungen entsprechen und gleichzeitig Standards unter den Maßgaben des EU AI-Acts und Data-Acts standhalten.

Zusammenfassend erfordert die Adaption von ASPICE (VDA Working Group 13, 2023) für die Landwirtschaft eine Überarbeitung und Erweiterung der bestehenden Qualitätsmerkmale, um sie für den landwirtschaftlichen Kontext relevant und anwendbar zu machen. Durch diese Anpassung können digitale Anwendungen in der Landwirtschaft effektiv bewertet und verbessert werden, was zu gesteigerter Effizienz, Nachhaltigkeit und Compliance führt.

3.2 Qualitätsmerkmale für Systeme und Software

Die Qualität von Software hat sehr unterschiedliche Facetten. Eine Taxonomie hierfür wurde bereits vor über 20 Jahren standardisiert und ist aktuell in der ISO/IEC 25000-Reihe „Systems and Software Quality Requirements and Evaluation“ dokumentiert. Die ISO/IEC 25010:2023 ist eine aktualisierte Norm, die sich auf die Definition von Qualitätsmodellen für Systeme und Software konzentriert (ISO/IEC, 2023). Diese Norm besteht aus zwei Hauptteilen: den Qualitätsmodellen und den Qualitätsmerkmalen. Die Qualitätsmodelle in dieser Norm umfassen das Systemqualitätsmodell und das Softwarequalitätsmodell. Diese Modelle dienen dazu, verschiedene Dimensionen der Qualität zu beschreiben, die bei der Entwicklung und Bewertung von Software- und Systemlösungen berücksichtigt werden müssen. Welche Ausprägung einer Qualität in einem bestimmten Kontext wichtig ist, hängt vom Anwendungszweck ab - kommt es beispielsweise auf Geschwindigkeit, Nutzerfreundlichkeit, Änderbarkeit oder Ressourcenverbrauch an?

Die ISO/IEC 25010:2023 (ISO/IEC, 2023) definiert acht Qualitätsmerkmale (Funktionalität, Zuverlässigkeit, Benutzbarkeit, Effizienz, Wartbarkeit, Portierbarkeit, Sicherheit und Kompatibilität), die eine umfassende Perspektive auf die Qualität digitaler Anwendungen bieten. Jedes dieser Hauptmerkmale wird in spezifische Qualitätsattribute unterteilt, die bei der Festlegung präziser Qualitätsziele und -anforderungen hilfreich sind.

Die ISO/IEC 25012 (ISO/IEC, 2008) ergänzt die ISO/IEC 25010:2023, indem sie sich auf ein besonders relevantes Qualitätsmerkmal konzentriert - die Datenqualität (Abbildung 5). Für digitale Anwendungen, einschließlich solcher in der Landwirtschaft, sind genaue und zuverlässige Daten entscheidend für eine effektive Entscheidungsfindung. Diese Norm definiert Qualitätsattribute für Daten wie Genauigkeit, Vollständigkeit, Konsistenz, Aktualität und Vertrauenswürdigkeit. Für digitale Anwendungen bedeutet dies, dass die verwendeten Daten zuverlässig und konsistent sein müssen, um genaue Analysen und Prognosen zu ermöglichen. Die Norm bietet auch Metriken und Verfahren zur Bewertung dieser Datenqualitätsattribute, was für die Anwendung in verschiedenen Bereichen, einschließlich der Landwirtschaft, von großer Bedeutung ist.

Zusammenfassend bilden diese Normen ein robustes Gerüst, um digitale Lösungen in der Landwirtschaft zu entwickeln, zu bewerten und sicherzustellen, dass sie den höchsten Qualitäts- und Sicherheitsstandards entsprechen. Für die Kundensicht ist insbesondere die „Quality in use“ ausschlaggebend (Abbildung 6).

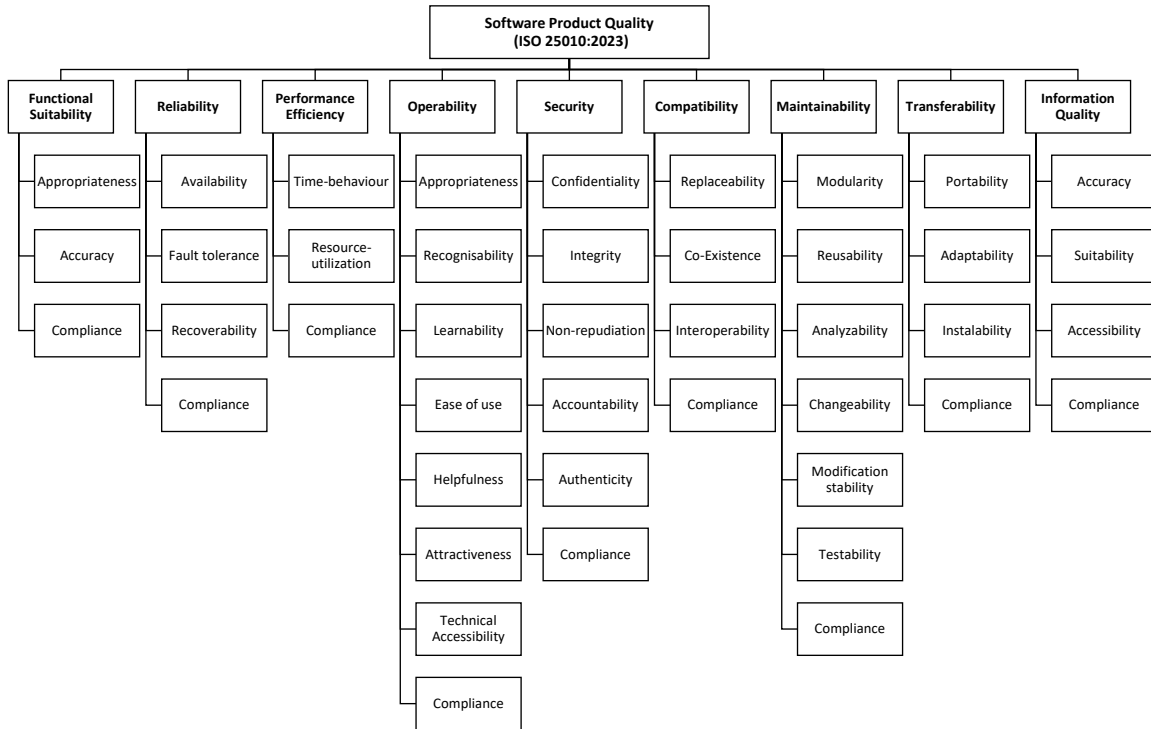


Abbildung 5: Übersicht über den ISO Standard 25010:2023 – Software Product Quality



Abbildung 6: Software-Qualitätsaspekte bei der Benutzung, z.B. durch Endanwender*innen

3.3 Die veränderte Rolle der Sensoren in digitalisierten Prozessen

Die Erfassung von Status und Umweltfaktoren spielt in erfolgreichen landwirtschaftlichen Produktionssystemen nicht erst seit der Einführung (teil-)digitalisierter Prozesse eine wesentliche Rolle. Entscheidungen erfordern Kompetenz der Betriebsleitungen sowie eine valide Datengrundlage. Digitalisierte Prozesse beinhalten die Durchgängigkeit des Datenflusses, das Eliminieren von Medienbrüchen und das Mitführen oder Bereitstellen von Metadaten zu Sensoren und Messdaten. Nur auf dieser Basis lassen sich Messwerte zueinander in Beziehung setzen und sinnvoll interpretieren und können zu einer Entscheidungsgrundlage werden. Digitale Sensoren verwenden in den meisten Fällen die gleichen Messverfahren wie rein analoge Sensoren, d.h. es handelt sich um Sensoren, die um Wandler und einfache Microcontroller erweitert sind. In anderen Messsystemen werden jedoch verschiedene sensorische Verfahren - beispielsweise optische Spektroskopie und elektrochemische Impedanzspektroskopie - kombiniert, um prinzipbedingte Schwächen eines Messverfahrens eliminieren zu können.

Die in den kommenden Jahren unausweichliche Koexistenz aktueller bzw. zukünftiger Sensoren sowie klassischer (analoger) Sensoren ohne Metadaten ist bei der Bewertung und Nutzung der Daten zu berücksichtigen. Häufig werden die etablierten Sensoren und Messverfahren als Referenz angesehen, ohne ihre inhärenten Defizite ausreichend zu berücksichtigen. Zudem ist bei einem Vergleich zu hinterfragen, ob dieser prinzipbedingt sinnvoll sein kann. Soweit die grundsätzliche Vergleichbarkeit nur eingeschränkt gegeben ist, sind daher entsprechende Hinweise auf den Gültigkeitsbereich notwendig. Die Koexistenz ist zudem für die Nutzung der Daten für längerfristige und überjährige Analysen und darauf aufbauende Prognosen notwendig. Metadaten zu den eingesetzten Sensoren können dann helfen, die Validität von Daten eines einzelnen Sensors, zwischen Sensoren des gleichen Typs, zwischen unterschiedlichen Sensoren mit gleichem Messprinzip sowie unterschiedlichen Messprinzipien beurteilen zu können. Nur auf Basis dieser Informationen sind weitere Analyseverfahren, u.a. über Maschinelles Lernen, sinnvoll, damit aus Sicht der Landwirtschaftsbetriebe aus Daten nutzbare Informationen und Entscheidungshilfen werden können.

3.4 Schnittstellen im landwirtschaftlichen Kontext

3.4.1 ISOBUS

Die digitalen Prozesse in der Landwirtschaft sind so vielfältig wie die Betriebsausrichtungen der einzelnen Betriebe. Eine große Bandbreite spezialisierter Produkte unterstützt einzelne Prozessschritte. Die Zusammenführung von Daten unterschiedlicher Art aus mehreren Prozessen und Systemen ist eine Herausforderung, die im Einzelfall gelingt, wenn über Schnittstellen kommuniziert werden kann. Im Bereich der Fahrzeuge und Anbaugeräte existieren standardisierte Hardware- und Softwarekomponenten, über die Daten ausgetauscht werden können: ISOBUS bezeichnet Standardisierung nach ISO 11783 (ISO, 2017), mit ISOXML (ISO, 2015a), einem Austauschformat zwischen Maschinen und Farm Management Informationssystemen (FMIS) (Olliver *et al.*, 2022). Aufgrund der Komplexität der Datenquellen und ihrer Kombinationsmöglichkeiten lassen sich nicht alle anfallenden Daten über standardisierte Schnittstellen austauschen, es existieren jedoch Ansätze für einige zentrale Datensätze.

Mit den Standardisierungen im Bereich des ISOBUS werden vor allem maschinennahe Systeme ausgestattet, die eine herstellerunabhängige Kombination von Komponenten ermöglichen. Geräte können basierend auf gemeinsamen Hardware- und Softwareschnittstellen miteinander kommunizieren und komplexe Vorgänge steuern. Da Schnittstellen anhand technischer Anforderungen entworfen werden und laufende Anpassungen gemäß aktuellen Entwicklungen erfahren können, werden Definitionen in den einzelnen Entwicklungsabteilungen vorgenommen. Standardisierungen und Normierungen erfolgen in fachspezifischen und teilweise branchenübergreifenden Konsortien. Die ISOBUS-Standardisierung involviert international Maschinenhersteller (VDMA in Deutschland, AEM in Nordamerika). Die Implementierungen werden unter dem

Dach der Agricultural Industry Electronics Foundation (AEF) entwickelt. Die Prüfung auf Standardkonformität lässt sich automatisiert in unabhängigen Testzentren durchführen.

3.4.2 Interoperabilität zwischen Farm Management Systemen

Eine weitreichende Interoperabilität zwischen unterschiedlichen digitalen Systemen, die auf einem Betrieb zum Einsatz kommen, ist aufgrund der Spezialisierung und vielfältigen, sich verändernden Prozesse eine große Herausforderung. Für Teilbereiche wurden Datenmodelle für den Datenaustausch entwickelt, die sich überwiegend auf Daten in Farm Management Systemen und für Feldeinsätze beziehen (Aufträge, Logdaten und Buchungen). Einige werden als frei nutzbare Komponenten wie im ADAPT Framework entwickelt (Craker *et al.*, 2018), das aber inzwischen wenig Entwicklungsaktivität verzeichnet. An verschiedenen Stellen besteht eine Zusammenarbeit von Herstellern (Horstmann, 2023; Treiber und Bernhardt, 2021), die einen bilateralen Datenaustausch zwischen Systemen unterschiedlicher Hersteller ermöglichen.

Die AEF Initiativen Extended Farm Management Information Systems Data Interface (EFDI) und Agricultural Interoperability Network (AgIN) (Smart *et al.*, 2022) haben zum Ziel, eine gemeinsame Basis für den Datenaustausch zwischen Systemen unterschiedlicher Hersteller zu entwickeln. Die große Zahl an Herstellern, die involviert sind, bietet Potenzial für die Entwicklung von herstellerunabhängigen Lösungen, die in den einzelnen Systemumgebungen Support erfahren kann.

Verschiedene Forschungs- und Entwicklungsvorhaben haben zu offenen Schnittstellen und Standards bereits die Machbarkeit untersucht und konkrete Vorschläge unterbreitet, die Ergebnisse sind aber nur teilweise in die Praxis eingeflossen (ATLAS consortium, 2021; Henningsen *et al.*, 2022). Um die Einbindung der Wirtschaft zu verbessern, könnten Infrastrukturen aufgebaut werden, die einen direkten Austausch und abgestimmte Entwicklungen ermöglichen. Aktuelle Aktivitäten im Rahmen der NFDI-Initiative können hierfür Anregungen geben. Die Forderung nach Offenheit von Forschungsergebnissen umfasst bereits heute Schnittstellendefinitionen, Metadaten und Datenformate.

Für einen eher generellen Datenaustausch zwischen Cloudsystemen werden Datenraum-Konzepte entwickelt, die das Ziel verfolgen, Datenhaltung und Zugriffe so zu standardisieren, dass eine gemeinsame Nutzung ermöglicht wird (Otto *et al.*, 2022).

3.4.3 GeoBox und Hofbox

Mit der GeoBox und Hofbox (Kuntke *et al.*, 2024; Kuntke *et al.*, 2020) werden Konzepte und Implementierungen von Softwarekomponenten und Diensten entwickelt, die vor allem räumliche Daten im Fokus haben. Zum einen wird eine zentrale Bereitstellung öffentlicher Daten und die Verfügbarkeit über standardisierte Schnittstellen unterstützt (GeoBox), zum anderen werden dezentral nutzbare Infrastrukturkomponenten entwickelt (Hofbox), die eine lokale Datenhaltung ermöglichen sollen. Im Bereich Geodaten werden standardisierte Schnittstellen und Services eingesetzt, die über Metadaten beschrieben und aufgefunden werden können. Ziel ist es, die Resilienz gegen Infrastrukturausfälle zu erhöhen und den Anwender*innen Software bereitzustellen, die eine dezentrale Datenhaltung und den Austausch mit Externen unterstützen. Die Systeme sind erweiterbar, da vor allem Open Source Lösungen und typische Cloudinfrastrukturkomponenten genutzt werden. Die Anwender*innen sollen in die Lage versetzt werden, die Datenhoheit ihrer eigenen Daten umzusetzen. Datenraum-Konzepte bieten Möglichkeiten für föderierte Systeme, die in diesem Kontext zur Anwendung kommen können, sofern die technischen und organisatorischen Voraussetzungen dafür gegeben sind. Handlungsbedarf besteht darin, föderierte und resiliente, dezentrale Systeme mit lokaler Datenhaltung standardisiert umzusetzen.

Die Schnittstellen von Systemen können so vielfältig wie die Anwendungsgebiete der Komponenten sein, sodass keine generelle Interoperabilität allein durch die Digitalisierung von Prozessen zu erwarten ist. Interoperabilität erfordert immer Bemühungen von den beteiligten Partnern, die Entwicklungen auf gemeinsame Grundlagen zu

stellen und für Übergabepunkte einheitliche und dokumentierte Definitionen bereitzustellen, die heutzutage überwiegend durch APIs implementiert werden, welche von allen Beteiligten in ihre Strukturen eingebunden werden müssen. Ein verbreiteter und automatisiert nutzbarer API Standard ist OpenAPI (OpenAPI Initiative, 2021). Über solche Anbindungspunkte lassen sich doppelte Datenhaltung und Medienbrüche reduzieren. In einem Zusammenspiel mehrerer Komponenten sollte geklärt sein, welches System für spezifische Daten als führendes System eingesetzt wird und welche Zugriffe notwendig sind. Die Organisation kann dann nach Stammdatenmodellen teilweise harmonisiert und zentralisiert erfolgen (Nienke und Fuhs, 2017).

3.5 Literatur

- ATLAS consortium, 2021. Service Architecture Specification. Deliverable No. D3.2., https://drive.google.com/uc?id=1-o4-WHp8hW_CSCTbxOcWvo-ZBy601wRs&export=download (besucht am 02.06.2023).
- Craker, B. E., Danford, D. D., Ferreyra, R., Nelson, K. J., Rhea, S. T., Stelford, M. W., Wilson, J. A., 2018. ADAPT: A Rosetta Stone for Agricultural Data. In: Proceedings of the 14th international conference on precision agriculture. Montreal, Quebec, Canada, 1–18A. <https://www.ispag.org/proceedings/?action=download&item=5366> (besucht am 31.07.2023).
- Henningsen, J., Herlitzius, T., Jeswein, T., Neuschwander, D. M. P., Rauch, B., Reinosch, N., Scherr, S. A., Schroers, J. O., Seuring, L., Striller, B., 2022. Betriebliches Datenmanagement und FMIS: Machbarkeitsstudie für „Betriebliches Datenmanagement und Farm-Management-Information-System (FMIS)“ in sächsischen Landwirtschaftsbetrieben. Schriftenreihe des LfULG, Band 2022. Machbarkeitsstudie. Internet: <https://publikationen.sachsen.de/bdb/artikel/39272/documents/60415>.
- Horstmann, J., 2023. Digitalisierung und Vernetzung - Herstellerübergreifender Datenaustausch Basis für Automatisierung. In: Frerichs, L. (Hrsg.), 2023: Jahrbuch Agrartechnik 2022, Jahrbuch Agrartechnik / Yearbook Agricultural Engineering, Band 34. Braunschweig: 26–32. https://leopard.tu-braunschweig.de/servlets/MCRFileNodeServlet/dbbs_derivate_00051201/jahrbuchagrartechnik2022_vernetzung.pdf (besucht am 09.09.2024).
- ISO, 2017. ISO 11783-1:2017 - Tractors and machinery for agriculture and forestry — Serial control and communications data network — Part 1: General standard for mobile data communication 2. Aufl.
- ISO, 2015a. ISO 11783-10:2015 Tractors and machinery for agriculture and forestry — Serial control and communications data network — Part 10: Task controller and management information system data interchange 2. Aufl.
- ISO, 2018. ISO 25119-1:2018: Tractors and machinery for agriculture and forestry - Safety-related parts of control systems - Part 1: General principles for design and development.
- ISO, 2015b. ISO 9001:2015 - Quality management systems - Requirements.
- ISO/IEC, 2023. ISO/IEC 25010:2023 - Systems and software engineering - Systems and software Quality Requirements and Evaluation (SQuaRE) - System and software quality model.
- ISO/IEC, 2008. ISO/IEC 25012:2008 - Software engineering - Software product Quality Requirements and Evaluation (SQuaRE) - Data quality model.
- ISO/IEC, 2012. ISO/IEC 15504-5:2012 (withdrawn) - Information technology - Process assessment - Part 5: An exemplar software life cycle process assessment model.
- Kuntke, F., Kaufhold, M.-A., Linsner, S., Reuter, C., 2024. GeoBox: design and evaluation of a tool for resilient and decentralised data management in agriculture. Behaviour & Information Technology 43(4), 764–786. DOI: 10.1080/0144929X.2023.2185747.
- Kuntke, F., Reuter, C., Schneider, W., Eberz, D., Bernardi, A., 2020. Die GeoBox-Vision: Resiliente Interaktion und Kooperation in der Landwirtschaft durch dezentrale Systeme. In: Gesellschaft für Informatik e.V., Mensch und Computer 2020 – Workshopband, 1–6, DOI: 10.18420/muc2020-ws117-407.
- Nienke, S., Fuhs, G. J., 2017. Stammdatenmanagement 2. Aufl. Whitepaper. Aachen, https://epub.fir.de/frontdoor/deliver/index/docId/1604/file/fir_whitepaper_stammdatenmanagement_Zweitaufgabe.pdf.

- Olliver, A., Schlingmann, N., Van der Vlugt, P., 2022. ISOBUS technologies – The Standard for Smart Agriculture. In: Zhang, Q. (Hrsg.), 2022: Encyclopedia of Smart Agriculture Technologies. 1–18, https://link.springer.com/referenceworkentry/10.1007/978-3-030-89123-7_117-1 (besucht am 05.03.2024).
- OpenAPI Initiative, 2021. OpenAPI Specification (OAS), <https://spec.openapis.org/oas/latest.html> (besucht am 07.03.2024).
- Otto, B., ten Hompel, M., Wrobel, S. (Hrsg.), 2022. Designing Data Spaces: The Ecosystem Approach to Competitive Advantage. Springer Cham, DOI: 10.1007/978-3-030-93975-5.
- Smart, D., Witte, J., Zippel, V., Stesny, S., Schlingmann, N., 2022. AEF's Impact on Ag Digital World. In: 2022 ASABE Annual International Meeting 2200264, St. Joseph, MI, USA, <https://elibrary.asabe.org/abstract.asp?aid=53330&t=5> (besucht am 09.09.2024).
- Treiber, M., Bernhardt, H., 2021. NEVONEX — the importance of middleware and interfaces for the digital transformation of agriculture. Engineering Proceedings 9(1), 3, DOI: 10.3390/engproc2021009003.
- VDA Working Group 13, 2023. Automotive SPICE Process Assessment / Reference Model Version 4.0. <https://vda-qmc.de/wp-content/uploads/2023/12/Automotive-SPICE-PAM-v40.pdf> (besucht am 07.03.2024).

4 Statistische Herausforderungen bei Vorhersagemodellen im Kontext der Digitalen Landwirtschaft

Frank Beneke, Doreen Gabriel und Christina Umstätter

4.1 Was sind Vorhersagemodelle und welche Rolle spielt KI dabei?

Unter Vorhersagemodellen werden Algorithmen verstanden, die eine Prognose der abhängigen Variablen durch eine funktionale Verknüpfung mit verschiedenen erklärenden Variablen erstellen. Dabei ist unter Prognose die Schätzung eines Erwartungswertes zu verstehen, die eine zeitliche Komponente enthalten kann, wie z.B. eine Ertragsschätzung bereits während der Vegetationszeit, dies aber nicht zwingend enthalten muss, wie z.B. die Schätzung der Qualität einer Weizenpartie mittels NIRS einer Probe. Bei der Entwicklung von Vorhersagemodellen können verschiedene statistische Methoden eingesetzt werden, von einfachen Klassifikations- und Regressionsmodellen bis hin zu *Machine Learning*-Methoden, wie z. B. *Random Forest* (Breiman, 2001), *Support Vector Machines* (Cortes und Vapnik, 1995), *Boosting* und *Bagging* (Breiman, 1996) und *Deep Learning* (Schmidhuber, 2015). Letztere werden auch unter dem Begriff „Künstliche Intelligenz“ (KI) zusammengefasst.

KI ist heute allgegenwärtig - aber es gibt bislang keine genaue Definition. KI wird als wissenschaftliche Disziplin umschrieben, die Computeralgorithmen nutzt, um aus Daten zu lernen, Muster zu entdecken und Vorhersagen zu entwickeln (Collins und Moons, 2019). In der Anwendung wird unterschieden zwischen starker KI, die über gleiche oder sogar höhere intellektuelle Fähigkeiten als der Mensch verfügt, und schwacher KI, die mit Hilfe von Modellen auch Aspekte menschlicher Intelligenz nachbilden und damit zur Lösung konkreter Anwendungsprobleme beitragen kann (Die Bundesregierung, 2018; Friedrich *et al.*, 2021). Die Bundesregierung hat 2018 eine nationale KI-Strategie mit 12 Handlungsfeldern beschlossen (Die Bundesregierung, 2018), mit dem Ziel, KI-Forschung und -Technologie in Deutschland zu fördern.

In jüngerer Zeit haben die sog. „großen Sprachmodelle“ (engl. Large language models, LLMs) erstaunliche Fortschritte gemacht und können beispielsweise beim Textverstehen oder bei Übersetzungen unterstützen oder auch natürlichsprachliche Antworten auf spezifische Fragen liefern. Auf Grundlage von Eingaben können LLMs das nächste Wort bzw. Token vorhersagen. Angereichert um Prozesswissen sollen darauf aufbauende Large Action Models (LAM) auch kontextabhängig prozessspezifische Entscheidungen treffen können. Somit können LAMs zukünftig mit der realen Welt interagieren und zum Beispiel vernetzte Geräte (IoT-Systeme) steuern oder dessen Daten für nachgelagerte Prozessschritte abrufen (Joshay, 2024).

Im Folgenden werden aktuelle Anwendungen von Vorhersagemodellen in der Landwirtschaft und anderen Einsatzbereichen aufgeführt, welche dem Bereich der schwachen KI angehören, und Aspekte zur Validierung und Beurteilung von Vorhersagemodellen erörtert.

4.2 Anwendung von KI in der Landwirtschaft und anderen Bereichen

Die Landwirtschaft basiert zunehmend auf datengetriebenen Ansätzen. In der heutigen Landwirtschaft werden immer größere Datenmengen erzeugt und immer mehr Produktionsbereiche und Produktionsmittel datentechnisch vernetzt. Der Einsatz von KI-basierten Vorhersagemodellen in der Landwirtschaft bietet ein großes Potential zur Sicherung und Verbesserung von Produktivität, Umwelt- und Ressourcenschutz, Ernährungssicherheit und Nachhaltigkeit (Gebbers und Adamchuk, 2010) und ermöglicht detaillierte Einblicke in Produktionszusammenhänge. Im landwirtschaftlichen Kontext gibt es eine Vielzahl von Anwendungen. So werden in der Tierhaltung KI-basierte Modelle eingesetzt um beispielsweise die Fütterung effizient zu gestalten, die Tiergesundheit zu überwachen und das Tierwohl zu gewährleisten (Liakos *et al.*, 2018). In der Pflanzenproduktion finden Anwendungen entlang des gesamten Produktionszyklus statt. Dies umfasst beispielsweise die Erkennung von Unkräutern und Schaderregern, die Abschätzung von Bodeneigenschaften und Düngbedarf sowie von Ertrag und Qualität (Kamilaris und Prenafeta-Boldú, 2018; Meshram *et al.*, 2021). Dabei werden verschiedene Sensoren und Datenquellen verwendet, um Spektren, Bilder, Maschinenzustände und andere Informationen zu erfassen. Neben dem großen Potenzial müssen aber auch die Herausforderungen berücksichtigt werden, wie zum Beispiel eine geeignete Datenlage für KI-Methoden. Qualitätsrelevant sind hier Datenverfügbarkeit, Datenqualität und auch das korrekte Labeling von Daten. Nicht zuletzt steht oft noch die verlässliche Integration von KI-Methoden in bestehende Prozesse aus und somit die Übertragbarkeit aus dem Forschungsbereich in die Praxis (X-KIT, 2023).

Die Anwendung von großen Sprachmodellen (LLM) und Aktionsmodellen (LAM) stehen noch am Anfang. Erste Produkte, welche z.B. Übersetzungen automatisieren, sind jedoch bald zu erwarten. Ebenfalls sind die Übernahme von Beratungsfunktionen und Unterstützung bei der Dokumentation durch LLMs zu erwarten. Entscheidungsunterstützungssysteme, welche z.B. die gute betriebliche Praxis eines landwirtschaftlichen Betriebs in LAMs abbilden, hätten ein großes Potenzial. Eine Bewertung der Verlässlichkeit solcher Systeme und der damit einhergehenden Risiken stellen noch eine große Herausforderung dar.

Im Bereich der Medizin werden KI-Anwendungen eingesetzt, um eine frühere und genauere Diagnose von Krankheiten zu erreichen, insbesondere bei der Krebsdiagnose (Burt *et al.*, 2018) und bei Herz-Kreislauf-Erkrankungen (Friedrich *et al.*, 2021) sowie um den Krankheitsverlauf zu überwachen und Therapieentscheidungen zu treffen. Zunehmend finden digitale Gesundheitsanwendungen (DiGA) auf mobilen Geräten als Apps oder browserbasierte Anwendungen im Gesundheitsbereich als *E-Health-Technologie* Verwendung (Collins und Moons, 2019). Aber auch in der Landwirtschaft erleichtern und unterstützen Apps zunehmend die Arbeit.

Modelle, die in der Wissenschaft oder in der Anwendung für Prognosen und Entscheidungen eingesetzt werden, müssen ausreichend dokumentiert und validiert werden. Hierzu wurden im TRIPOD-Statement Empfehlungen entwickelt (Collins *et al.*, 2015; Moons *et al.*, 2015), die eine Checkliste für eine transparente Berichterstattung über die Modellentwicklung, -ergebnisse, -grenzen und -interpretationen enthalten. Am Beispiel medizinischer Anwendungen lassen sich die Anforderungen an datenbasierte Modelle und Entscheidungen gut veranschaulichen: Anwendungen, die als Medizinprodukte mit geringem Risiko eingestuft werden, müssen CE-zertifiziert und vom Bundesinstitut für Arzneimittel und Medizinprodukte (BfArM) als DiGA geprüft werden (BfArM, 2022). Dazu müssen Anforderungen an die Sicherheit und Funktionstauglichkeit, den Datenschutz und die Informationssicherheit sowie an die Qualität und insbesondere an die Interoperabilität der Anwendung erfüllt werden. Auf Basis einer systematischen Literaturrecherche und -bewertung sowie einer eigenen systematischen Datenauswertung ist ein Nachweis für einen positiven Versorgungseffekt zu erbringen. Dieser basiert auf einem nach allgemein anerkannten wissenschaftlichen Standards erstellten Evaluationskonzept. Teil des

Evaluationskonzept ist ein Studienprotokoll (der Prüfplan) mit statistischem Analyseplan. Dieser „Prüfplan“ umfasst die Ziele und die Methodik der Studie mit Angaben zur statistischen Planung und Analyse mit Hypothesen, Endpunkten, statistischen Modellen, Randomisierung oder Kontrolle von Confounding, Fallzahlberechnung, Umgang mit fehlenden Daten und Dropouts etc. (BfArM, 2022). Das Evaluationskonzept muss von einem herstellerunabhängigen wissenschaftlichen Institut erstellt werden. Grundsätzlich ist der Nachweis der Überlegenheit der Behandlung mit DiGA gegenüber der Nichtbehandlung zu erbringen.

4.3 Externe Validierung

Bei der Erstellung von Vorhersagemodellen werden Datensätze häufig in Trainings- und Validierungsdatsätze unterteilt. Man verwendet einen Trainingsdatensatz, um das Modell zu parametrisieren, wendet Methoden wie Kreuz-Validierung (*leave-one-out* oder *k-fold*) oder Bootstrapping an und validiert das Modell, indem man es auf den zuvor getrennten Validierungsdatsatz anwendet. Dies ist eine interne Validierung, die ein wichtiger Bestandteil des Modellierungsprozesses ist, aber eine unabhängige Validierung mit einem externen Validierungssatz nicht ersetzt (Moons *et al.*, 2015). Um die Qualität von Vorhersagemodellen beurteilen zu können, ist eine externe Validierung mit einem neuen unabhängigen Datensatz erforderlich, der nicht zur Modellerstellung verwendet wurde. Die zufällige Aufteilung eines Datensatzes durch die modellierende Person in Training und Validierung führt zu redundanten Datensätzen, welche die Modellgüte überschätzen können. Die Aufteilung eines vorhandenen Datensatzes in verschiedene Untersuchungsjahre oder Standorte ist bei ausreichend großen Stichproben etwas besser, aber nur wirklich unabhängige, neue Daten führen zu einer robusten externen Validierung (Debray *et al.*, 2015; Moons *et al.*, 2015; Ramspek *et al.*, 2021). Diese Zusammenhänge sollten bereits bei der Entwicklung eines Vorhersagemodells berücksichtigt und in den Entwicklungsprozess integriert werden.

Nur durch die Anwendung des Modells auf eine neue Population, d.h. auf ein externes Validierungsset, kann die Qualität, Anwendbarkeit, Übertragbarkeit und Robustheit für den Endnutzer beurteilt werden. Um die zeitliche und räumliche Übertragbarkeit beurteilen zu können, sollte das Testset aus Proben aus neuen Untersuchungszeiträumen und neuen Standorten bestehen und somit nicht die identischen Bedingungen wie das Trainingsset repräsentieren (Moons *et al.*, 2015; de Hond *et al.*, 2023). Ein neues Team, das die Messungen unter den neuen Bedingungen durchführt, kann die Validierung ebenfalls robuster machen (Steyerberg, 2019; de Hond *et al.*, 2023). Der Stichprobenumfang der Testdaten sollte der zu prüfenden Genauigkeit entsprechen (Archer *et al.*, 2021; Riley *et al.*, 2021). Eine Faustregel von Collins *et al.* (2016) besagt, dass bei medizinischen Untersuchungen mindestens 100 Ereignisse und 100 Nicht-Ereignisse, besser jedoch 200 Ereignisse und mehr verwendet werden sollten. Darüber hinaus sollten die Testdaten repräsentativ für das spätere Anwendungsgebiet sein (Friedrich *et al.*, 2022), und einen der späteren Anwendung angepassten Wertebereich aufweisen bzw. bei Klassifikationsmodellen alle Klassen ausreichend häufig repräsentieren, sowie mögliche Kovariaten berücksichtigen. Wenn z.B. bei pflanzenbaulichen Fragestellungen verschiedene Bodentypen einen Einfluss auf die Modellvorhersage haben können, sollten die Validierungsergebnisse auch für die verschiedenen Bodentypen getrennt ermittelt werden (Riley *et al.*, 2016).

4.4 Statistische Beurteilungskriterien

Es wird unterschieden zwischen Regressionsmodellen, die kontinuierliche Merkmale vorhersagen, und Klassifikationsmodellen, die eine kategoriale Zielvariable wie z. B. Ja-Nein-Entscheidungen vorhersagen. Je nach Modelltyp (Regressions- vs. Klassifikationsmodell) werden unterschiedliche Evaluationsmaße verwendet (Kamilaris und Prenafeta-Boldú, 2018; Rajamanickam *et al.*, 2021).

Die Evaluierung von Regressionsmodellen erfolgt durch den Vergleich von beobachteten und vorhergesagten Werten. Je geringer die Abweichung ist, desto besser ist das Modell. Dies kann in Relation zur Einheit der abhängigen Variable als Mean Absolute Error (MAE), Root Mean Square Error (RMSE) und Mean Square Error (MSE), oder als Prozentwert unabhängig von der Einheit als Mean Relative Error (MRE) und als Bestimmtheitsmaß (R^2) ermittelt werden (Rajamanickam *et al.*, 2021). Insbesondere das R^2 ist ein sehr häufig

genutztes Kriterium zur Beurteilung von Modellen, da es mit einem Wertebereich zwischen 0 und 1 sehr intuitiv ist und auch über verschiedene Modelle und Datensätze hinweg verglichen werden kann. Ein weiteres Kriterium ist der *calibration slope* zwischen beobachteten und vorhergesagten Werten (Riley *et al.*, 2016), der Auskunft über die generelle Über- oder Unterschätzung der Modellvorhersage gibt.

Die Evaluierung von Klassifikationsmodellen erfolgt anhand der Anzahl der True Positives (TP) und True Negatives (TN) - d.h. der richtig geschätzten Eintritte und Nichteintritte von Ereignissen, und der Anzahl der False Positives (FP) und False Negatives (FN) - d.h. der falsch geschätzten Eintritte und Nichteintritte von Ereignissen. Aus diesen Kennzahlen lassen sich verschiedene Indikatoren zur Bewertung der Modelle ableiten:

- die Treffsicherheit (*accuracy*) beschreibt das Verhältnis der richtig geschätzten Entscheidungen zur Gesamtanzahl $(TP+TN)/(TP+TN+FP+FN)$,
- die Sensitivität (*recall*) beschreibt das Verhältnis der richtig geschätzten Ja-Entscheidungen in Relation zu den beobachteten Ja-Entscheidungen $TP/(TP+FN)$,
- die Spezifität beschreibt das Verhältnis der richtig geschätzten Nein-Entscheidungen in Relation zu den beobachteten Nein-Entscheidungen $TN/(TN+FP)$,
- die Genauigkeit (*precision*) beschreibt das Verhältnis der richtig geschätzten Ja-Entscheidungen zur Anzahl aller Ja-Entscheidungen $TP/(TP+FP)$.

Weitere sogenannte ABCD-Indikatoren für die Modellgüte können aus den oben genannten Indikatoren abgeleitet werden (Steyerberg und Vergouwe, 2014), wie z.B.:

- Alpha - *calibration in the large* - repräsentiert den Intercept der vorhergesagten und beobachteten Werte und berichtet über eine systematische Über- oder Unterschätzung des Modells, während
- Beta den *calibration slope* darstellt, der, wenn er kleiner als 1 ist, anzeigt, dass kleine Werte unter- und große Werte überschätzt werden.
- die C-Statistik, die aus der *receiver operating characteristic (ROC) curve* abgeleitet wird, in der die Sensitivität (Richtig-Positiv-Rate) gegen die Falsch-Positiv-Rate aufgetragen wird. Je weiter die Kurve von der Diagonalen abweicht, und je größer die Fläche unter der Kurve (*area under the ROC curve, AUC*) ist, desto besser ist das Modell. Eine AUC (C-Statistic) von 1 deutet auf eine perfekte Klassifikation hin, ein Wert von 0,5 auf eine zufällige Klassifikation.
- D (Decision-curve analysis) oder Net Benefit, das den Nutzen eines Modells im Vergleich zu einer Nichtverwendung bei Entscheidungen quantifiziert (Steyerberg und Vergouwe, 2014; Vickers *et al.*, 2016).
- Für nominale Klassifikationsmodelle können die oben genannten Maße für jede Gruppe einzeln und Cohens Kappa als aggregiertes Maß für die Klassifikationsgüte berechnet werden.

Grundsätzlich sollten zur Beurteilung von Modellen immer mehrere Kriterien herangezogen und wenn möglich Konfidenzintervalle berechnet werden. Darüber hinaus ist die zu erwartende Güte eines Modells gegen die mögliche Zeit- und Kostenersparnis abzuwägen. Wenn z.B. Laborergebnisse zwar sehr genau aber nur mit erheblichem Zeitaufwand zu erhalten sind, kann eine Modellvorhersage durch einen kostengünstigen Schnelltest mit größerer Ungenauigkeit dennoch einen sinnvollen Dienst leisten.

Modelle sollten regelmäßig validiert werden, da sich sowohl die Eingangsvariablen als auch die Prozesse der Datengenerierung ändern können, was zu einem Kalibrierungsdrift führen kann (Jenkins *et al.*, 2021). Zudem werden insbesondere KI-Modelle kontinuierlich (weiter-) entwickelt, was eine große Herausforderung darstellt und somit eine kontinuierliche Validierung erfordert (Friedrich *et al.*, 2022). Eine weitere Besonderheit bei der Validierung von KI-Modellen ist der Black-Box-Ansatz dieser Modelle, da Modellkoeffizienten nicht eingesehen und interpretiert werden können. Entscheidungen, die mit Hilfe von KI-Modellen getroffen werden, sind daher häufig nicht im Detail nachvollziehbar. Um dieser fehlenden Transparenz entgegenzuwirken und die Akzeptanz von KI-Modellen in Entscheidungssystemen zu erhöhen, wird zunehmend der Ansatz der Erklärbarkeit von KI-Modellen, d.h. *explainable AI (XAI)*, verfolgt (Arrieta *et al.*, 2020). Dabei geht es darum, dem Nutzer von KI-

Modellen die Gründe für KI-Entscheidungen zu veranschaulichen und Aspekte der Verlässlichkeit, Kausalität, Übertragbarkeit, Informativität, Vertrauenswürdigkeit, Fairness und Zugänglichkeit von Algorithmen zu berücksichtigen.

4.5 Limitierungen von Vorhersagemodellen

Die Erwartung an Vorhersagemodelle ist sehr hoch. Eine Schwierigkeit besteht jedoch darin, dass die Genauigkeit für den Einsatz in landwirtschaftlichen Betrieben oft unzureichend ist. Für die zukünftige digitale Entwicklung ist es daher wichtig, dass die Gründe für die unzureichende Genauigkeit angegangen werden und untersucht wird, was technisch und methodisch machbar ist.

Es sind zwei wichtige Bereiche zu nennen, die die unzureichende Genauigkeit erheblich beeinflussen: 1. unbalancierte Datensätze und 2. die Wahl der Sensoren. Oft ist die Anzahl der Fehlalarme, d.h. die falsch-positiv Rate, für den Landwirt*in nicht praktikabel, obwohl die Sensitivität und Spezifität hoch genug sind (Zehner *et al.*, 2019). Laut Post *et al.* (2021) liegt bei vielen Studien der Schwerpunkt auf der ersten Stufe der biologischen Validierung, hier wird oft mit der Receiver-Operator-Charakteristik-Kurve (ROC) und den Indikatoren Sensitivität und Spezifität (z.B. Kamphuis *et al.*, 2013; Jensen *et al.*, 2016) gearbeitet. Die Autoren stellen fest, dass in weniger Studien eine zweite Stufe der Validierung der entwickelten Algorithmen an Datensätzen durchgeführt wird, die eher einer praktischen Situation entsprechen und somit stärker unbalancierte Datensätze aufweisen. Dort wird für die meisten Verfahren der positive prädiktive Wert (PPV), der den Anteil der falsch-positiven Ergebnisse an allen positiven Tests beschreibt, genutzt (Post *et al.*, 2021). Das Autorenteam betont, dass in den Studien nur selten darauf hingewiesen wird, dass die Wahrscheinlichkeit der korrekten Vorhersage eines Ereignisses umso geringer ist, je geringer die Wahrscheinlichkeit des Eintretens des Ereignisses ist. In einem solchen Szenario ist auch der PPV niedriger und die falsch-positive Rate höher. Das bedeutet, dass bei seltenen Ereignissen und den damit verbundenen unbalancierten Datensätzen, wie sie auf tierhaltenden Betrieben vorkommen, eine hohe Rate an Fehlalarmen einkalkuliert werden muss. Dies kann aus probabilistischer Sicht zum heutigen Stand der Forschung nicht umgangen werden. Eine Kombination von Datenströmen hebt die Imbalance nicht auf und kann deshalb auch nicht einfach die Anzahl der Falsch-Positiven verringern.

Der zweite Bereich betrifft die Wahl der eingesetzten Sensoren und die daraus stammenden Variablen. Vielfach werden beispielsweise Aktivitätssensoren für die Entwicklung digitaler Systeme beim Milchvieh genutzt, da die Sensoren verhältnismäßig preiswert und robust sind und bereits vielfach auf den Betrieben vorkommen. Die Aktivität ist jedoch ein unspezifischer Indikator für die Gesundheit und das Wohlbefinden. In landwirtschaftlichen Betrieben, in denen die Tiere wechselnden Umweltbedingungen ausgesetzt sind, verschiedene innere Zustände wie Brunst, Trächtigkeit und Laktation durchlaufen und vor allem mit mehr als einer Herausforderung gleichzeitig konfrontiert sein können, kann der Einsatz unspezifischer Sensoren nur zu unspezifischen Aussagen mit geringem Informationsgehalt führen. Um die Genauigkeit von Precision Livestock Farming-Technologien zu verbessern und die Entwicklung geeigneter Sensorik zu unterstützen, wurde ein Rahmenwerk entwickelt (Stachowicz und Umstätter, 2021), das auf drei Ebenen basiert, die durch einen zunehmenden Informationsgehalt gekennzeichnet sind. Die erste Ebene handelt von der Detektion von Problemen im Bereich des Wohlbefindens von Tieren. Diese eher unspezifische Ebene kann mit unspezifischen Sensoren detektiert werden. Die nächste Ebene handelt von den Gesundheits- und Distress-bezogenen Indikatoren, die schon etwas spezifischer ist. Und die Ebene mit dem höchsten Informationsgehalt ist am spezifischsten und beschäftigt sich mit der Detektion von spezifischen Erkrankungen. Wenn man sich bewusst wird, für welche Ebene man die Frühwarnsysteme entwickelt, kann die Qualität der Systeme erhöht werden und die Erwartungshaltung der Stakeholder besser gemanagt werden.

4.6 Literatur

- Archer, L., Snell, K.I.E., Ensor, J., Hudda, M.T., Collins, G.S., Riley, R.D., 2021. Minimum sample size for external validation of a clinical prediction model with a continuous outcome. *Statistics in Medicine* 40, 133-146.
- Arrieta, A.B., Diaz-Rodriguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion* 58, 82-115.
- BfArM, 2022. Das Fast-Track-Verfahren für digitale Gesundheitsanwendungen (DiGA) nach § 139e SGB V. Bundesinstitut für Arzneimittel und Medizinprodukte.
- Breiman, L., 1996. Bagging predictors. *Machine Learning* 24, 123-140.
- Breiman, L., 2001. Random forests. *Machine Learning* 45, 5-32.
- Burt, J.R., Torosdagli, N., Khosravan, N., RaviPrakash, H., Mortazi, A., Tissavirasingham, F., Hussein, S., Bagci, U., 2018. Deep learning beyond cats and dogs: recent advances in diagnosing breast cancer with deep neural networks. *The British Journal of Radiology* 91, 20170545.
- Collins, G.S., Moons, K.G.M., 2019. Reporting of artificial intelligence prediction models. *The Lancet* 393, 1577-1579.
- Collins, G.S., Ogundimu, E.O., Altman, D.G., 2016. Sample size considerations for the external validation of a multivariable prognostic model: a resampling study. *Statistics in Medicine* 35, 214-226.
- Collins, G.S., Reitsma, J.B., Altman, D.G., Moons, K.G.M., 2015. Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) The TRIPOD Statement. *Circulation* 131, 211-219.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Machine Learning* 20, 273-297.
- de Hond, A.A.H., Shah, V.B., Kant, I.M.J., Van Calster, B., Steyerberg, E.W., Hernandez-Boussard, T., 2023. Perspectives on validation of clinical predictive algorithms. *npj Digital Medicine* 6, 86.
- Debray, T.P.A., Vergouwe, Y., Koffijberg, H., Nieboer, D., Steyerberg, E.W., Moons, K.G.M., 2015. A new framework to enhance the interpretation of external validation studies of clinical prediction models. *Journal of Clinical Epidemiology* 68, 279-289.
- Die Bundesregierung, 2018. Strategie Künstliche Intelligenz der Bundesregierung. URL www.ki-strategie-deutschland.de.
- Friedrich, S., Antes, G., Behr, S., Binder, H., Brannath, W., Dumpert, F., Ickstadt, K., Kestler, H.A., Lederer, J., Leitgob, H., Pauly, M., Steland, A., Wilhelm, A., Friede, T., 2022. Is there a role for statistics in artificial intelligence? *Advances in Data Analysis and Classification* 16, 823-846.
- Friedrich, S., Groß, S., König, I.R., Engelhardt, S., Bahls, M., Heinz, J., Huber, C., Kaderali, L., Kelm, M., Leha, A., Rühl, J., Schaller, J., Scherer, C., Vollmer, M., Seidler, T., Friede, T., 2021. Applications of artificial intelligence/machine learning approaches in cardiovascular medicine: a systematic review with recommendations. *European Heart Journal - Digital Health* 2, 424-436.
- Gebbers, R., Adamchuk, V.I., 2010. Precision Agriculture and Food Security. *Science* 327, 828-831.
- Jenkins, D.A., Martin, G.P., Sperrin, M., Riley, R.D., Debray, T.P.A., Collins, G.S., Peek, N., 2021. Continual updating and monitoring of clinical prediction models: time for dynamic prediction systems? *Diagnostic and Prognostic Research* 5, 1.

- Jensen, D.B., Hogeveen, H., De Vries, A., 2016. Bayesian integration of sensor information and a multivariate dynamic linear model for prediction of dairy cow mastitis. *Journal of Dairy Science* 99, 7344-7361.
- Joshy, R., The Rise of Large Action Models, LAMs: How AI Can Understand and Execute Human Intentions?, <https://medium.com/version-1/the-rise-of-large-action-models-lams-how-ai-can-understand-and-execute-human-intentions-f59c8e78bc09> (besucht am 04.03.2024).
- Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture* 147, 70-90.
- Kamphuis, C., Frank, E., Burke, J.K., Verkerk, G.A., Jago, J.G., 2013. Applying additive logistic regression to data derived from sensors monitoring behavioral and physiological characteristics of dairy cows to detect lameness. *Journal of Dairy Science* 96, 7043-7053.
- Liakos, K.G., Busato, P., Moshou, D., Pearson, S., Bochtis, D., 2018. Machine Learning in Agriculture: A Review. *Sensors* 18.
- Meshram, V., Patil, K., Meshram, V., Hanchate, D., Ramkteke, S.D., 2021. Machine learning in agriculture domain: A state-of-art survey. *Artificial Intelligence in the Life Sciences* 1, 100010.
- Moons, K.G.M., Altman, D.G., Reitsma, J.B., Ioannidis, J.P.A., Macaskill, P., Steyerberg, E.W., Vickers, A.J., Ransohoff, D.F., Collins, G.S., 2015. Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD): Explanation and Elaboration. *Annals of Internal Medicine* 162, W1-W73.
- Post, C., Rietz, C., Büscher, W., Müller, U., 2021. The Importance of Low Daily Risk for the Prediction of Treatment Events of Individual Dairy Cows with Sensor Systems. *Sensors* 21.
- Rajamanickam, V., Babel, H., Montano-Herrera, L., Ehsani, A., Stiefel, F., Haider, S., Presser, B., Knapp, B., 2021. About Model Validation in Bioprocessing. *Processes* 9, 961.
- Ramspek, C.L., Jager, K.J., Dekker, F.W., Zoccali, C., van Diepen, M., 2021. External validation of prognostic models: what, why, how, when and where? *Clinical Kidney Journal* 14, 49-58.
- Riley, R.D., Debray, T.P.A., Collins, G.S., Archer, L., Ensor, J., van Smeden, M., Snell, K.I.E., 2021. Minimum sample size for external validation of a clinical prediction model with a binary outcome. *Statistics in Medicine* 40, 4230-4251.
- Riley, R.D., Ensor, J., Snell, K.I.E., Debray, T.P.A., Altman, D.G., Moons, K.G.M., Collins, G.S., 2016. External validation of clinical prediction models using big datasets from e-health records or IPD meta-analysis: opportunities and challenges. *BMJ* 353, i3140.
- Schmidhuber, J., 2015. Deep learning in neural networks: An overview. *Neural Networks* 61, 85-117.
- Stachowicz, J., Umstätter, C., 2021. Do we automatically detect health- or general welfare-related issues? A framework. *Proceedings Of The Royal Society B-Biological Sciences* 288.
- Steyerberg, E.W., 2019. *Clinical Prediction Models. A Practical Approach to Development, Validation, and Updating.* Springer Cham.
- Steyerberg, E.W., Vergouwe, Y., 2014. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *European Heart Journal* 35, 1925-1931.
- Vickers, A.J., Van Calster, B., Steyerberg, E.W., 2016. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. *BMJ* 352, i6.

X-KIT, Kolloquium Landwirtschaft der Zukunft: Ist KI ein wesentlicher Schlüssel zur nachhaltigeren Landwirtschaft? Informatik 2023, Berlin. Zusammenfassung unter <https://www.iese.fraunhofer.de/de/projekt/x-kit/fachbeitrag/informatika2023-kola23.html> (besucht am 06.03.2024).

Zehner, N., Niederhauser, J.J., Schick, M., Umstatter, C., 2019. Development and validation of a predictive model for calving time based on sensor measurements of ingestive behavior in dairy cows. *Computers and Electronics in Agriculture* 161, 62-71.

5 Bestandsaufnahme digitaler Technologien in den Digitalen Experimentierfeldern

Frank Beneke, Andreas Gabriel und Markus Gandorfer

Seit 2019 erforschen 14 digitale Experimentierfelder im Auftrag des BMEL digitale Technologien in der Landwirtschaft. Um einen Überblick über die in den digitalen Experimentierfeldern untersuchten digitalen Technologien, die dabei angewandten Untersuchungsmethoden sowie das vorherrschende Verständnis von „Verlässlichkeit“ digitaler Technologien zu gewinnen, wurde eine Befragung konzipiert und durchgeführt. Die Umfrage enthielt vier Fragen und offene Antwortfelder. Diese wurden an die 14 Experimentierfelder per E-Mail durch die BLE versandt. 13 der 14 Experimentierfelder haben die Umfrage im Zeitraum von Oktober 2022 bis Januar 2023 beantwortet. Es gab teilweise mehrere Rückantworten je Experimentierfeld, da die unterschiedlichen Projektpartner für ihre Arbeitspakete den Fragebogen separat ausgefüllt haben. Insgesamt wurden damit 27 Einzelantworten aus 13 Experimentierfeldern ausgewertet. Für die Auswertung erfolgte eine Zuordnung der insgesamt 188 genannten digitalen Technologien zu Kategorien einer Technologienliste. Weiterhin wurden Unterkategorien gebildet. Die Untersuchungsmethoden in den drei Bereichen „Forschung und (Weiter)Entwicklung“, „Demonstration“ und „Bewertung“ wurden identifiziert und je Technologiekategorie zusammengefasst.

5.1 Ergebnisse

Die Abbildung 7 zeigt eine verdichtete Übersicht der Technologiebereiche, mit denen sich die Experimentierfelder beschäftigten¹. Es zeigt sich ein deutlicher Schwerpunkt bei Sensortechnologien, wobei sich 12 der insgesamt 66 Nennungen mit unterschiedlichen Aspekten von NIR-Sensoren beschäftigen. Die weiteren Schwerpunkte liegen bei Fernerkundung (Uncrewed Aerial Vehicles (UAV) und Satellit), Technologien für den Pflanzenschutz und Automatisierungstechnologien (einschl. Assistenzsystemen). Bei der Automatisierung bestehen die Untergruppen automatisierte Stalltechnik und Feldrobotik.

1

Zu beachten ist, dass es mögliche Doppelzuordnungen geben kann (z.B. Software in der Fütterungstechnik oder Unkrautkartierung auf UAV-Bildern). Hier wurde der Untersuchungsschwerpunkt gemäß der Rückmeldung in der Befragung berücksichtigt.

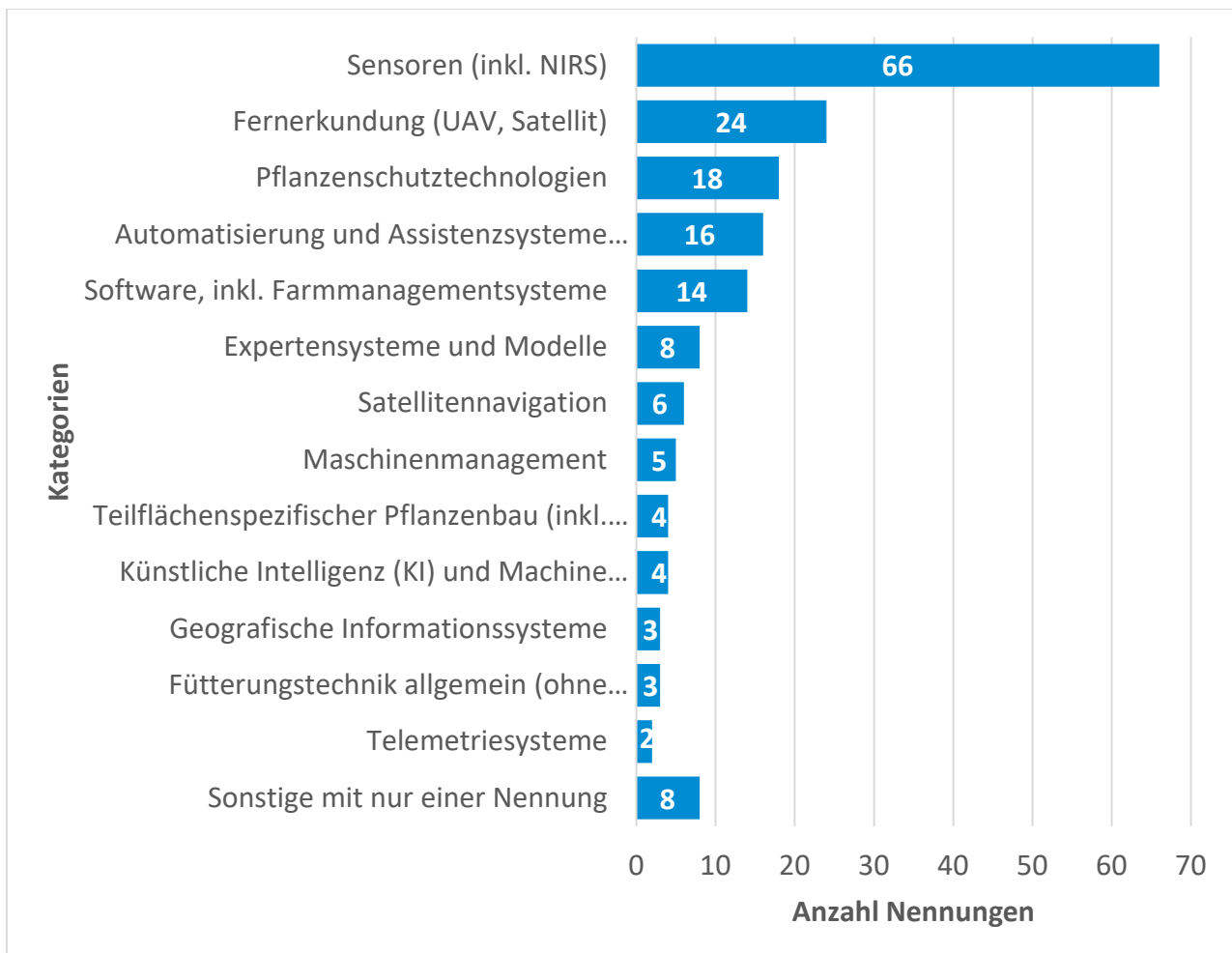


Abbildung 7: Kategorisierung der Nennungen untersuchter digitaler Technologien in den Experimentierfeldern. Datenbasis: 27 Antworten aus 13 Experimentierfeldern

5.1.1 Bewertungsfelder der untersuchten Technologien

Die Abbildung 8 zeigt die „Intensität“ mit der in den befragten Experimentierfeldern die Bereiche „Agronomische Bewertung“, „Bewertung sozialer Aspekte“, „Bewertung der Verlässlichkeit“, „Bewertung der Wirtschaftlichkeit“ und „Ökologische Bewertung“ auf die untersuchten Technologien angewandt werden. Die Einschätzung in die drei Intensitätsstufen „hoch“, „mittel“ und „niedrig“ erfolgte anhand des Umfangs und Detailgrades der Antworten zu angewandten Bewertungsinstrumenten und unterliegt damit einer gewissen Subjektivität der Autoren.

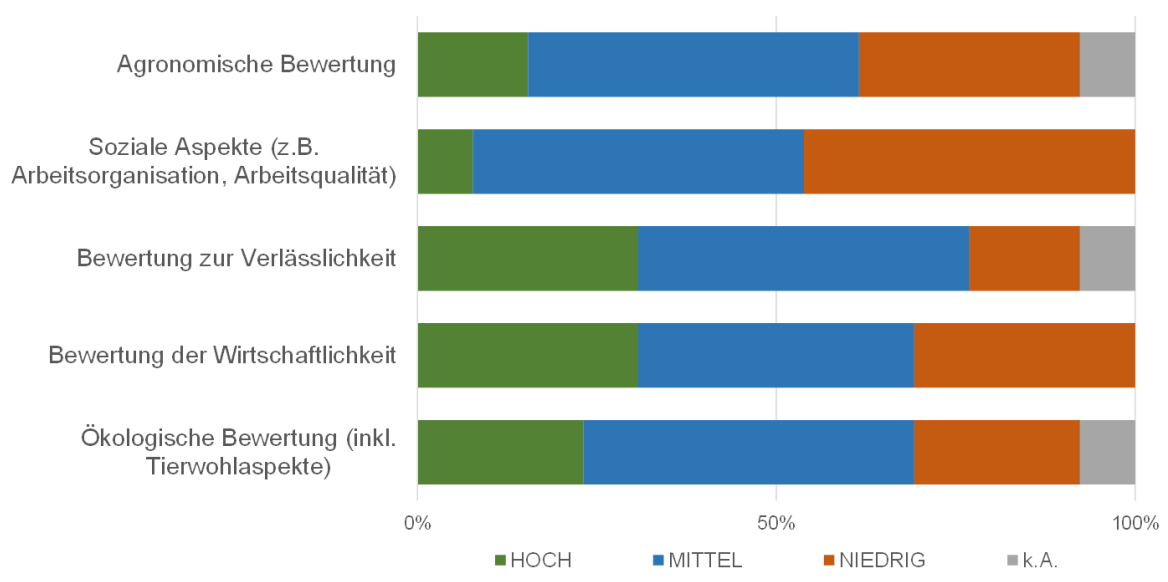


Abbildung 8: „Intensität“ der Berücksichtigung von Bewertungsbereichen in den 13 Experimentierfeldern

Die Analyse zeigt, dass im Durchschnitt der befragten Experimentierfelder die digitalen Technologien relativ umfassend bewertet werden, wobei die Schwerpunkte in den Bereichen Verlässlichkeit und Wirtschaftlichkeit der Technologien liegen. Der Bewertungsbereich der sozialen Aspekte (z.B. Arbeitsorganisation; Arbeitsqualität; Ergonomie; Handhabung) fällt dagegen ab. In Tabelle 1 sind weiterhin exemplarisch zu den untersuchten Technologien Anwendungsfelder aufgeführt, die in den verschiedenen Experimentierfeldern in Demonstration, Bewertung und Forschung bearbeitet werden.

Zusammenfassend lässt sich feststellen, dass die Schwerpunkte bei den behandelten Technologien in den Experimentierfeldern bei Sensoren (NIRS ist gesondert ausgewiesen), Technologien die zum Pflanzenschutz eingesetzt werden, UAVs sowie bei Automatisierung (z.B. Stalltechnik und Feldrobotik) liegen. Die Experimentierfelder leisten vielfältige Beiträge zum Wissenstransfer („Demonstration“ und „Bewertung“). Zusätzlich zum Praxistransfer werden umfangreiche Forschungsbeiträge geliefert. Sensoren sind entweder konkrete Untersuchungsschwerpunkte (z.B. NIRS zur Futterwertbestimmung) oder stellen die Basis für darauf aufbauende Untersuchungen (z.B. teilflächenspezifische Bewirtschaftung auf Basis von UAV-Daten). Bei NIRS stehen Inhaltsstoffbestimmungen bei Futter und in der organischen Düngung im Vordergrund. Weiterhin gewinnt die Vernetzung von Sensoren und die Integration in Farmmanagementinformationssystemen (FMIS) und Entscheidungsunterstützungssysteme an Bedeutung. Technologien der Fernerkundung mit Satelliten oder UAV sind in den Experimentierfeldern etabliert und werden breit eingesetzt und kommuniziert. Schwerpunkt ist hier die Erfassung von Bestandszuständen im Pflanzenbau. Im Schwerpunkt Pflanzenschutz erfolgen zahlreiche Aktivitäten der Experimentierfelder. Hier werden z.B. Kombinationen von chemischen und mechanischen Pflanzenschutzmaßnahmen, Entwicklungen im Bereich Spot Spraying oder teilflächenspezifische Maßnahmen behandelt. In der Stalltechnik zeigt sich ein Schwerpunkt in der zunehmenden Automatisierung. Die Feldrobotik beschäftigt sich schwerpunktmäßig in Demonstration, Bewertung und Forschung mit Trägerplattformen allgemein und mit Pflanzenschutztechnik (Robotik wurde hier als separater Untersuchungsschwerpunkt gezählt und nicht dem Bereich Pflanzenschutztechnik zugerechnet).

Tabelle 1: Beispiele von untersuchten digitalen Technologien in den Digitalen Experimentierfeldern

Technologie	Anwendungsfelder der untersuchten Technologien in den Experimentierfeldern	Experimentierfelder
Sensoren (einschl. NIRS)	<p><u>Sensoren allg.:</u></p> <p>Digitale Bewertung von Bodenbearbeitungsmaßnahmen, Bodenleitfähigkeit/-feuchte, Wetter- und Mikroklimasensoren, N-Sensoren, Multi- / Hyperspektralkameras, Stallklimasensoren, Tiersensoren</p> <p><u>NIRS:</u></p> <p>Gülle Nährstoffgehalt, Gülleausbringung, Kartoffelunterfußdüngung, Erntegutqualität / Inhaltsstoffbestimmung (Häckselkette); Inhaltsstoffbestimmung Futter, mobile NIRS (Futter, Erntegut), miniaturisierte NIRS</p>	<p>Agri-Sens-DEMMIN, Agro-Nordwest, BeSt-SH, Cattle-Hub, DigiSchwein, Express, FarmerSpace, Diabek, DigiMilch, Diwakopter, DiWenKla, Landnetz, DigiVine</p> <p>Agro-Nordwest, DigiSchwein, DigiMilch; DiWenKla; Landnetz; BeSt-SH; DigiVine</p>
Fernerkundung	<p><u>UAV:</u></p> <p>Bestandserfassung, Photogrammetrie, Lidar-Messungen, Pflanzenschutzmittelapplikation</p> <p><u>Satellit:</u></p> <p>Zustandsbestimmung aus Fernerkundungsdaten, Ertragsschätzung, Feldeigenschaften</p>	<p>Agri-Sens-DEMMIN, Agro-Nordwest, BeSt-SH, Diabek, Diwakopter, DiWenKla, Express, Farmerspace, Landnetz</p> <p>Agri-Sens-DEMMIN, Farmerspace, Landnetz</p>
Pflanzenschutztechnik (ohne Feldroboter und Fernerkundung)	<p>Unkrautkartierung, teilflächenspezifischer Pflanzenschutz, Kombination mechanischer und chemischer Pflanzenschutz, Spot Spraying, mechanischer Pflanzenschutz</p>	<p>Agro-Nordwest, DiWenKla, Farmerspace, Landnetz</p>
Automatisierte Stalltechnik (inkl.) Fütterung	<p>Automatisierte Stalltechnik (M2M), Entmistungsroboter, Trockenfütterungsanlage, Futtermischwagen, programmierbare Waagen</p>	<p>Agro-Nordwest, DigiMilch, DigiSchwein, DiWenKla, Landnetz</p>
Feldrobotik	<p>Agarrobotik (Säen, Pflanzung, Düngen, Unkrautregulierung); Autonome Hackroboter; Autonome Trägerplattformen</p>	<p>Agro-Nordwest, Farmerspace, DiWenKla, Landnetz</p>

5.2 Auswahl der Use Cases

Im Folgenden werden drei, in den Experimentierfeldern intensiv untersuchte digitale Technologien als Use Cases ausgewählt und näher analysiert. Ziel ist es dabei, Mindestanforderungen abzuleiten, die bei einer Technikbewertung zu betrachten sind. Nähere Angaben zu den involvierten Experimentierfeldern und den Anwendungsfeldern der ausgewählten und analysierten Technologien sind in Tabelle 2 aufgelistet.

Tabelle 2: Anwendungsfelder in den Experimentierfeldern bei den vier ausgewählten Technologien

Technologie	Anwendungsfelder der untersuchten Technologien in den Experimentierfeldern	Experimentierfelder
NIRS-Technologie (Sensoren)	Güllenährstoffgehalt, Gülleausbringung, Futterwertbestimmung in Grünlandaufwüchsen; mobile NIRS	Agro-Nordwest, DigiSchwein, DigiMilch; DiWenKla; Landnetz; BeSt-SH; DigiVine
Feldrobotik	Agrarrobotik (Säen, Pflanzung, Düngen, Unkrautregulierung); Autonome Hackroboter; Autonome Trägerplattformen	Agro-Nordwest, Landnetz, Farmerspace, DiWenKla
Verhaltensüberwachung (Rinder/Schweine)	Diverse Tiersensoren- und Monitoringsysteme, RGB-Kameras; Geräuscherfassung, RFID-Marken, Hustendetektor, Wärmebildkameras	CattleHub, DigiSchwein, Landnetz, DigiMilch

Die Analyse stützt sich dabei im ersten Schritt auf die Angaben der befragten Experimentierfelder. Im Folgenden sind dazu die Angaben aus den Antworten aus den Experimentierfeldern zusammengefasst. Neben der Auflistung der Mindestanforderungen der Praxis an die Anwendbarkeit der Technologien sind auch Aussagen zu den Chancen und Herausforderungen berücksichtigt.

5.2.1 NIRS-Technologie

Die genannten Mindestanforderungen an NIRS-Technik und optische Sensoren in verschiedenen landwirtschaftlichen Anwendungen lassen sich zu verschiedenen Kategorien, wie Genauigkeit, Zuverlässigkeit, Integration und Anwendungsfreundlichkeit, Wirtschaftlichkeit und Robustheit zusammenfassen.

Genauigkeit: Die Sensoren müssen in der Lage sein, hinreichend genaue Messungen der Nährstoffgehalte in Wirtschaftsdüngern und anderen landwirtschaftlichen Stoffen (insbes. Erntegüter) durchzuführen. Eine hohe Genauigkeit ist entscheidend, um die richtigen Entscheidungen bezüglich der Düngung oder Futterwertbestimmung treffen zu können. Weitere Voraussetzung ist die Verfügbarkeit von Kalibrierkurven für die jeweiligen zu untersuchenden Stoffe.

Zuverlässigkeit: Die Sensoren sollten eine hohe Zuverlässigkeit aufweisen, um konsistente und wiederholbare Messergebnisse zu liefern. Die Verlässlichkeit hängt auch von der richtigen Kalibrierung und Abstimmung der Technik auf die jeweiligen landwirtschaftlichen Bedingungen ab.

Integration und Anwendungsfreundlichkeit: Die Technologie sollte nahtlos in bestehende landwirtschaftliche Produktionsprozesse integriert werden können. Dafür sind einheitliche Schnittstellen und eine einfache Anwendung erforderlich, ebenso eine (möglichst herstellerübergreifende) nahtlose Integration in FMIS. Außerdem ist ein guter Support für die Benutzer*innen wichtig, um mögliche Probleme zu lösen und die Technologie optimal nutzen zu können.

Wirtschaftlichkeit: Die NIRS-Technik und optische Sensoren sollten wirtschaftlich sein und einen Mehrwert für die landwirtschaftlichen Betriebe bieten. Sie sollten Ertragssteigerungen und potenzielle Mehrerträge / höhere Qualitäten des Ernteguts ermöglichen, um den Einsatz rentabel zu machen.

Robustheit: Die Sensoren müssen in der Lage sein, in verschiedenen Umgebungsbedingungen zuverlässige Messungen durchzuführen. Sie sollten robust genug sein, um Verschmutzungen, Feuchtigkeit und Erschütterungen zu widerstehen und genaue Messungen trotz möglicher Störfaktoren zu liefern.

5.2.2 Feldrobotik

Die Feldrobotik wird als besonders geeignet für Anwendungen wie Unkrautregulierung in Sonderkulturen und ökologischer Landwirtschaft gesehen. Vorteile liegen in der exakten Unkrauterkenntnis, Arbeitspräzision, Reduzierung schwerer körperlicher Arbeit und der Möglichkeit, (Saison-) Arbeitskräfte zu ersetzen oder deren Tätigkeit zu ergänzen. Zudem können Feldroboter bei der Überfahrt relevante Felddaten generieren, die den Menschen bei seinen Entscheidungen unterstützen.

Im Folgenden sind die Mindestanforderungen der Praxis an die Technologie der Feldrobotik zusammengestellt. Sehr aufschlussreich waren hierbei auch die Angaben aus dem Experimentierfeld „Agro-Nordwest“ zu den Ergebnissen der Fokusgruppendifkussionen mit Landwirt*innen und Lohnunternehmer*innen.

(1) Nutzenorientierung: Die Technologien sollten den Nutzen für die Landwirt*innen maximieren und interdisziplinäre Aspekte wie Qualifikation, Service und gesellschaftliche Akzeptanz berücksichtigen.

(2) Zugänglichkeit: Eine breite Zugänglichkeit der Systeme für Landwirt*innen ist wichtig, um den Einsatz der Technologien zu fördern. Dies kann auch durch Co-Creation-Ansätze erreicht werden, bei denen die Landwirt*innen aktiv an der Weiterentwicklung der Technologien beteiligt sind.

(3) Sensibilisierung und Schulung: Landwirt*innen müssen für die neuen Technologien sensibilisiert und geschult werden, um eine erfolgreiche Anwendung zu gewährleisten.

(4) Standardisierung des Datenflusses: Der Datenfluss von den Landwirt*innen und Dienstleistern zur Softwareanwendung sollte standardisiert werden, um eine effiziente und nahtlose Integration zu ermöglichen.

(5) Integration in den landwirtschaftlichen Betrieb: Die Technologien sollten leicht in einen landwirtschaftlichen Betrieb integrierbar sein. Dies erfordert robuste Systeme, die unter den üblichen Feldbedingungen funktionieren, sowie die Bereitstellung des erforderlichen Knowhows für den Betrieb.

Schlagkraft / Stundenleistung, zuverlässiger und störungsfreier Betrieb über längere Zeit, Robustheit gegenüber Fremdeinwirkungen und Wetterbedingungen, Klärung von Zulassungs- und Haftungsfragen / Gewährleistung von Rechtssicherheit, Integration in den landwirtschaftlichen Arbeitsprozess sowie transparente Kommunikation und Nutzung von Daten sind wichtige Herausforderungen, die weiterentwickelt und adressiert werden müssen. Die autonome Feldrobotik wird besonders in Kulturen mit hoher Wertschöpfung, z.B. Sonderkulturen wie Gemüse sowie in der ökologischen Landwirtschaft (hoher Anteil mechanischer Pflegemaßnahmen) als frühe Anwendungsfelder gesehen. In großen Betrieben und bei Lohnunternehmen wird die Gewährleistung der entsprechenden Flächenleistung hingegen noch als große Herausforderung wahrgenommen.

5.2.3 Verhaltensüberwachung/Tracking Tier

Der zentrale Nutzen der Technologien wird zusammengefasst wie folgt beschrieben: Die Technologie sollte eine Entscheidungsunterstützung bieten, indem sie Landwirt*innen hilft, kontinuierlich die Tiere zu beobachten, Zeit zu sparen und fehlende Fachkräfte zu kompensieren. Sie sollte das Tierwohl verbessern, indem sie ein besseres Monitoring und ein verbessertes Herdenmanagement ermöglicht. Die sichere Erkennung von Einzeltieren und ein stärkerer Fokus auf Tiere, die Zuwendung benötigen, sowie präventive Maßnahmen und die frühzeitige Erkennung von Krankheiten sollten ermöglicht werden. Die Technologien sollten auch bei der Dokumentation

unterstützen und die Organisation des Arbeitsalltags verbessern, um Tätigkeiten flexibler erfüllen zu können und daher zu FMIS Schnittstellen besitzen.

Grundsätzliche Mindestanforderungen an die Technologien: Die Technologien sollten eine einfache Inbetriebnahme ermöglichen, idealerweise durch Landwirt*innen selbst, und einfach in der Handhabung sein. Die Datenerfassung und -auswertung sollten zuverlässig sein und Ausfälle sollten minimiert werden. Eine reliable "Trefferquote" (falsch-positiv und falsch-negativ) wird als wichtig erachtet. Es sollte ein Datenspeicher vorhanden sein und bei Kamerasystemen sollte eine automatisierte Bildanalyse möglich sein. Das Dashboard oder die Apps sollten übersichtlich gestaltet sein und Plausibilitätsanalysen sollten durchgeführt werden können.

Aktuelle Herausforderungen: Herausforderungen bestehen in der schwierigen Vernetzung der Systeme, der begrenzten Batterielaufzeit von Transpondern, der Verbesserung der Funktionalitäten und der Vielfalt an marktverfügbaren Systemen. Die richtige Auswahl des geeigneten Systems basierend auf Herdenmanagement und Stallbaustruktur ist wichtig. Es gibt auch Herausforderungen in Bezug auf die Inkompatibilität digitaler Technologien, die begrenzte Marktverfügbarkeit der Produkte und die Notwendigkeit, Arbeitsprozesse im Betrieb anzupassen. Die Identifikation der Mitarbeitenden mit der digitalen Technologie sowie die Einarbeitung und Kennenlernphase sind weitere Herausforderungen. Es ist wichtig, Informationen in praktische Maßnahmen umzusetzen.

6 Mögliche Mindestanforderungen an digitale Technologien

Frank Beneke und Ingolf Römer

Auf Basis der in der Arbeitsgruppe geführten Diskussionen sowie der in Kapitel 5 vorgestellten Befragung der Experimentierfelder wurden Mindestanforderungen an die praxisrelevante Anwendbarkeit der Technologien sowie assoziierte Bewertungsaspekte abgeleitet. Diese sind im Folgenden aufgelistet:

Technische Leistungsfähigkeit

- Genauigkeit / Arbeitspräzision (Efficiency in use)
- Flächenleistung bzw. Leistung
- Zuverlässigkeit (Reliability)
- Robustheit gegenüber Fremdeinwirkungen (Fault tolerance, Recoverability)
- Verfügbarkeit / Technologiezugänglichkeit (Availability, Operability/Technical Accessibility)
- TRL (Technologiereifegrad)
- Datenqualität (Information Quality)
- Automatisierte Auswertung (z.B. Bildanalyse)
- Verlässliche Sicherheitssensoren für autonomen Betrieb (Safety)
- Batterielaufzeit / Einsatzzeit

Begründung: Dieser Cluster umfasst die Grundvoraussetzungen, die eine Technologie erfüllen muss, um in der Landwirtschaft effektiv eingesetzt zu werden. Dazu gehören Aspekte wie Genauigkeit, Zuverlässigkeit und Robustheit, die direkt die Performance und Einsatzfähigkeit beeinflussen und die Funktionalität für eine bestimmte Aufgabe charakterisieren. Die Batterielaufzeit wurde diesem Cluster hinzugefügt, da sie ein kritischer Faktor für die operative Leistung von Geräten ist, insbesondere bei mobilen oder autonomen Systemen mit elektrischen Antrieben.

Kosten und Wirtschaftlichkeit (Usability in Use, Gebrauchstauglichkeit)

- Kosten / Investition
- Kosten / Betriebskosten (einschl. Wartung und Software-/Lizenzkosten)
- Kosten / Arbeiterledigungskosten
- Wirtschaftlichkeit
- Wirtschaftlichkeit und Mehrwert für landw. Betriebe

Begründung: Die finanziellen Aspekte sind für landwirtschaftliche Betriebe von entscheidender Bedeutung. Dieser Cluster bewertet die Kosten-Nutzen-Relation und die Gesamtwirtschaftlichkeit der Technologien. Er berücksichtigt sowohl die anfänglichen Investitionen als auch die laufenden Betriebskosten.

Integration und Kompatibilität (Compatibility)

- Abstimmung der Technik auf die jeweiligen landwirtschaftlichen Bedingungen bzw. betrieblichen Bedingungen
- Integration
- Integration in Entscheidungsunterstützungssysteme
- Kompatibilität (Compatibility)
- Vernetzung / Vernetzbarkeit (Interoperability)
- Softwarestandards, Schnittstellen
- M2M-Vernetzung

Begründung: Die nahtlose Eingliederung neuer Technologien in bestehende Systeme und Prozesse ist für einen erfolgreichen Einsatz unabdingbar. Dieser Cluster konzentriert sich auf die Kompatibilität mit vorhandener Infrastruktur, die Fähigkeit zur Integration in Entscheidungsunterstützungssysteme und die generelle Vernetzungsfähigkeit der Technologien.

Benutzerfreundlichkeit und Service (Usability in use)

- Anwendungsfreundlichkeit (Usability in use/satisfaction in use)
- Wartungsfreundlichkeit
- Support für Benutzer*innen / Service
- Einfache Inbetriebnahme und Nutzung/Handhabung (Learnability in use)
- Übersichtliche Gestaltung von Dashboards, Apps, etc.
- Herstellerservice

Begründung: Die Anwenderakzeptanz ist ein Schlüsselement für den Erfolg technologischer Innovationen. Dieser Cluster fasst Kriterien zusammen, die die tägliche Nutzung und Wartung der Technologien aus Nutzersicht erleichtern, einschließlich des Supports und der allgemeinen Benutzererfahrung.

Betriebliche Effizienz (Flexibility in Use, Efficiency in Use)

- Umweltbedingungen
- Arbeitserleichterung, Reduzierung monotoner oder schwerer körperlicher Arbeit

- Zeiteinsparung
- Betriebsmitteleinsparung
- Unterstützung der Dokumentation

Begründung: Dieser Cluster fokussiert auf den direkten Einfluss der Technologien auf die betriebliche Effizienz. Dazu gehört die Verbesserung der Arbeitsbedingungen, Zeiteinsparungen und die Unterstützung bei der Dokumentation, die zu einer erhöhten Produktivität führen.

Datensicherheit und -management (Security)

- Daten-Standards
- Datensicherheit
- Datenspeicherung und Backups möglich
- langfristiger Datenzugriff
- Datenhoheit (wem gehören die Daten) und Zugriffsrechte / Berechtigungskonzepte

Begründung: In einer zunehmend datengesteuerten Landwirtschaft ist die sichere und effiziente Handhabung von Daten essentiell. Dieser Cluster beinhaltet Aspekte, die mit der Datenspeicherung, -standards und -übertragung verbunden sind, und betrachtet die technischen Voraussetzungen für ein effektives Datenmanagement.

Rechtliche und Qualifikationsaspekte

- Erforderliche Qualifikation
- Klärung von Zulassungs- und Haftungsfragen
- Rechtssicherheit
- Datenschutz

Begründung: Rechtliche Rahmenbedingungen und die erforderlichen Qualifikationen für die Bedienung der Technologien sind wichtige Faktoren, die den Einsatz und die Akzeptanz von Innovationen beeinflussen können. Dieser Cluster deckt die Notwendigkeit der Klärung von Zulassungsfragen, Haftung und der Sicherstellung der Rechtssicherheit ab.

Nutzen und Wertbeitrag

- Erfüllung Produktversprechen
- Nutzenorientierung
- NIRS: Verfügbarkeit von Kalibrierkurven
- Anwendungen wie Unkrautregulierung in Sonderkulturen und ökologischer Landwirtschaft
- Qualität der Unkrauterkennung
- Ersatz von Saisonarbeitskräften
- Schlagkraft
- Optimierte Herdenbeobachtung
- Höhere Tiergesundheit
- Optimierung Arbeitsalltag, Flexibilisierung

- Durchführen von Plausibilitätsanalysen möglich
- Papierlose Dateneingabe über Apps
- automatisierte Dokumentation
- Entscheidungsunterstützung
- Übertragbarkeit auf andere Betriebsstellen
- Optimierung und Präzisierung des Fütterungsprozesses
- Lebensdauer der Technologie

Begründung: Hier stehen der Mehrwert und der direkte Nutzen der Technologien für die landwirtschaftlichen Betriebe im Vordergrund. Dieser Cluster erfasst die Verbesserung der Betriebsprozesse, den Beitrag zur Tiergesundheit und allgemeine Optimierungen, die sich durch den Einsatz der Technologien ergeben.

Bibliografische Information:
Die Deutsche Nationalbibliothek
verzeichnet diese Publikationen in
der Deutschen Nationalbibliografie;
detaillierte bibliografische Daten
sind im Internet unter
www.dnb.de abrufbar.

Bibliographic information:
The Deutsche Nationalbibliothek
(German National Library) lists this
publication in the German National
Bibliographie; detailed bibliographic
data is available on the Internet at
[*www.dnb.de*](http://www.dnb.de)

Bereits in dieser Reihe erschienene
Bände finden Sie im Internet unter
www.thuenen.de

Volumes already published in this
series are available on the Internet at
www.thuenen.de

Zitationsvorschlag – Suggested source citation:
**Beneke F, Gabriel A, Gabriel D, Gandorfer M, Howind K-H, Kalmar R,
Meierhöfer J, Noack P, Reckleben Y, Römer I, Rubenschuh U, Schmittmann O,
Staemmler N, Umstätter C, Weis M, Wrenger B (2024) Wege zu einer
Qualitätsbewertung digitaler und technischer Lösungen in der Landwirtschaft.
Braunschweig: Johann Heinrich von Thünen-Institut, 43 p, Thünen Working
Paper 246, DOI:10.3220/WP1725352600000**

Die Verantwortung für die Inhalte
liegt bei den jeweiligen Verfassern
bzw. Verfasserinnen.

The respective authors are
responsible for the content of
their publications.



THÜNEN

Thünen Working Paper Klicken Sie hier, um Text einzugeben.

Herausgeber/Redaktionsanschrift – *Editor/address*

Johann Heinrich von Thünen-Institut

Bundesallee 50

38116 Braunschweig

Germany

thuenen-working-paper@thuenen.de

www.thuenen.de

DOI:10.3220/WP1725352600000

urn:nbn:de:gbv:253-202409-dn068662-1