FARMERSPACE AI
DIGITALER PFLANZENSCHUTZ

# Inhaltsverzeichnis

Prüfrahmen zur standardisierten Testung KI-basierter Spot-Spraying-Systeme unter Freilandbedingungen ..... 2
1. Einleitung ..... 2
2. Stand der Technik und Einordnung des ASABE-Testprotokolls ..... 4
2.1 Bewertungsansätze in der wissenschaftlichen Literatur ..... 4
2.2 Das ASABE/ISO-Testprotokoll im Detail ..... 5
2.3 Einordnung und identifizierte Lücken ..... 6
3. Systematisierung: Online- vs. Offline-Spot-Spraying-Systeme ..... 7
3.1 Definition und Abgrenzung ..... 7
3.2 Funktionsbausteine je Systemtyp und Ground-Truth-Methodik ..... 9
3.3 Testmethodische Relevanz der Unterscheidung ..... 11
3.4 Zusammenfassung: Kennzahlen je Funktionsbaustein und Systemtyp ..... 12
4. Prüfrahmen – Allgemeiner Aufbau ..... 14
4.1 Testfeld und Streifenanlage ..... 14
4.2 Ground-Truth-Kartierung des Beikrautbestands ..... 16
4.3 Collector-Platzierung ..... 17
4.4 Kontrollpositionen und Fehlapplikationsrate ..... 17
4.5 Wirkstoff und Wirkungsbonitur ..... 19
4.6 Umweltbedingungen und Dokumentation ..... 20
4.7 Grenzen der Übertragbarkeit ..... 21
5. Prüfrahmen – Online-Systeme ..... 22
5.1 Testablauf ..... 22
5.2 Prüfpunkte und Messgrößen ..... 24
6. Prüfrahmen – Offline-Systeme ..... 25
6.1 Ground-Truth-Kartierung und Systemkartierung ..... 25
6.2 Erkennungsprüfung (verpflichtend) ..... 25
6.3 Testablauf der Applikation ..... 26
6.4 Prüfpunkte und Messgrößen ..... 27
7. Erweiterte Bewertungskriterien und Forschungsausblick ..... 28
7.1 Ökonomische Kriterien (optional) ..... 28
7.2 Ökologische Kriterien (optional) ..... 29
7.3 Agronomischer Forschungsbedarf: Ertragswirkung und Generalisierbarkeit ..... 29
7.4 Übertragung auf mechanische und physikalische Systeme ..... 30

8. Anwendung des Prüfrahmens ..... 30
8.1 Voraussetzungen (Kurzcheck) ..... 30
8.2 Ablaufschema (Kurzfassung) ..... 31
8.3 Dokumentation (Pflichtfelder im Testbericht) ..... 32
9. Zusammenfassung und Ausblick ..... 34
Anhang ..... 34
A. Bezug zum ASABE/ISO-Protokoll ..... 34
B. Literaturverzeichnis *(zu vervollständigen)* ..... 35
C. Dokumentationsvorlage *(zu ergänzen)* ..... 35
D. Begleitende Normungsarbeiten und methodisch vergleichbare Forschung ..... 36

# Prüfrahmen zur standardisierten Testung KI-basierter Spot-Spraying-Systeme unter Freilandbedingungen

*Stand 01.10.2026*

---

## 1. Einleitung

Die Beikrautregulierung gehört zu den zentralen Herausforderungen im Pflanzenbau. Bleiben Beikräuter unbehandelt, können sie durch die Konkurrenz um Licht, Wasser und Nährstoffe erhebliche Ertragseinbußen verursachen. In der konventionellen Landwirtschaft erfolgt die Bekämpfung überwiegend durch den flächendeckenden Einsatz chemischer Herbizide, während im ökologischen Landbau vor allem mechanische Verfahren eingesetzt werden. Beide Ansätze haben jedoch ihre Grenzen. Der flächendeckende Einsatz von Herbiziden steht zunehmend in der Kritik, unter anderem aufgrund möglicher Auswirkungen auf Boden- und Gewässerökosysteme, der zunehmenden Resistenzproblematik und politischer Zielsetzungen zur Reduktion des Pflanzenschutzmitteleinsatzes. Mechanische Verfahren sind dagegen arbeitsintensiv, stärker von den Witterungsbedingungen abhängig und können je nach Art und Intensität der Bodenbearbeitung das Erosionsrisiko erhöhen.

Künstliche Intelligenz (KI) eröffnet hier neue Möglichkeiten. KI-basierte Präzisionsapplikatoren können Kulturpflanzen und Beikräuter automatisiert unterscheiden und Pflanzenschutzmittel gezielt nur dort ausbringen, wo sie benötigt werden. Gegenüber einer flächendeckenden Applikation lässt sich dadurch vor allem die ausgebrachte Pflanzenschutzmittelmenge reduzieren, ohne dass die angestrebte Bekämpfungswirkung grundsätzlich verändert werden muss. Die Einsparung wirkt sich unmittelbar auf die Betriebsmittelkosten aus und kann gleichzeitig einen ökologischen Vorteil bieten.

Im ökologischen Landbau ist darüber hinaus die Bekämpfungssicherheit ein wichtiger Aspekt. Klassische mechanische Verfahren wie Hacke oder Striegel können Beikräuter

innerhalb der Pflanzenreihe nur eingeschränkt erfassen, da eine zu starke Annäherung an die Kulturpflanze zu Schäden führen kann. Gleichzeitig sind diese Verfahren auf geeignete Witterungsbedingungen und Entwicklungsstadien angewiesen. Ist das entsprechende Zeitfenster beispielsweise aufgrund von Niederschlag oder eines ungeeigneten Bodenzustands nicht nutzbar, kann die Bekämpfungswirkung deutlich abnehmen. KI-gestützte Systeme bieten hier die Möglichkeit, Beikräuter gezielter und näher an der Kulturpflanze zu erfassen und zu behandeln. Durch die Automatisierung können solche Systeme zudem unabhängig von der verfügbaren Arbeitskraft innerhalb geeigneter Zeitfenster eingesetzt werden. Eine weitere mögliche Verbesserung liegt darin, den Umfang der mechanischen Bodenbearbeitung und damit auch deren erosionsfördernde Wirkung zu reduzieren.

Auf dem Markt sind bereits verschiedene Spot-Spraying-Systeme verfügbar beziehungsweise befinden sich in der Entwicklung. Sie unterscheiden sich teilweise deutlich hinsichtlich ihrer technischen Umsetzung, beispielsweise bei der Erkennung der Beikräuter, der Positionierung der Applikation oder der verwendeten Ausbringungstechnik. Diese Vielfalt erschwert einen direkten Vergleich der Systeme. Eine einheitliche und praxisnah anwendbare Testmethodik, mit der unterschiedliche Systeme unter vergleichbaren Bedingungen bewertet werden können, existiert bislang nicht. Auch wissenschaftliche Untersuchungen verwenden häufig individuell entwickelte Versuchsdesigns, die sich hinsichtlich Zielgrößen, Messmethoden und Versuchsbedingungen unterscheiden. Die Ergebnisse lassen sich dadurch nur eingeschränkt auf andere Systeme oder Einsatzbedingungen übertragen. Für Landwirte und Berater ist es entsprechend schwierig, die am Markt verfügbaren Systeme auf einer einheitlichen Grundlage zu beurteilen.

Vor diesem Hintergrund wird im Rahmen des Verbundprojekts FarmerSpaceAI ein Prüfrahmen für KI-basierte Spot-Spraying-Systeme entwickelt. Das Projekt befasst sich unter anderem mit dem Einsatz KI-basierter Verfahren in der Beikrautregulierung. Der Prüfrahmen soll dabei drei wesentliche Ziele erfüllen:

- Er soll als einheitliche Grundlage für die im Projekt geplanten Freilandversuche dienen.
- Er soll externen Prüfstellen und Anwendern ermöglichen, unterschiedliche Systeme unter vergleichbaren Bedingungen zu testen und vergleichbare Ergebnisse zu erhalten.
- Er soll auf möglichst viele unterschiedliche Systeme anwendbar sein und dabei unabhängig von Hersteller und konkreter technischer Umsetzung bleiben.

Als methodische Grundlage dient das im Rahmen von ISO/TC 23/SC 6/AHG 2 (“Definitions and test methods for new precision application technologies”) diskutierte Testprotokoll des ASABE-Workshops vom April 2025. Das Protokoll wird in Kapitel 2 näher vorgestellt. Die darin enthaltene Unterscheidung zwischen online und offline arbeitenden Systemen wird auch im vorliegenden Prüfrahmen übernommen. Darüber hinaus wird das bestehende Protokoll in drei Punkten erweitert: Erstens erfolgt die Prüfung in einem realen Pflanzenbestand und nicht auf einer bestandsfreien Testfläche. Zweitens werden die Erkennung und die anschließende Applikation methodisch getrennt bewertet und als

eigenständige Ergebnisse ausgewiesen. Drittens wird die tatsächliche Bekämpfungswirkung als weiterer Bewertungsbaustein aufgenommen.

Der Ablauf eines Spot-Spraying-Systems lässt sich damit grundsätzlich in drei aufeinanderfolgende Funktionsbereiche gliedern: Erkennung, Applikation und Wirkung. Die Erkennung umfasst die Detektion und Lokalisierung des Beikrauts. Bei der Applikation wird untersucht, wie genau die Ausbringung relativ zur erkannten Zielposition erfolgt. Die Wirkung beschreibt schließlich den tatsächlichen Bekämpfungserfolg und dessen Entwicklung über die Zeit. Diese drei Bereiche bilden das methodische Grundgerüst des vorliegenden Prüfrahmens.

**Geltungsbereich und Abgrenzung.** Der Prüfrahmen ist für KI-basierte Spot-Spraying-Systeme konzipiert und richtet sich vorrangig an marktverfügbare beziehungsweise praxisnahe Systeme, nicht an Prototypen. Auf die Bewertung autonomer Systeme sowie intelligenter mechanischer oder physikalischer Verfahren der Beikrautregulierung ist er nicht ausgerichtet. Wird eine Spot-Spraying-Einheit auf einem autonomen Trägerfahrzeug eingesetzt, lassen sich Erkennung, Applikation und Wirkung zwar nach diesem Prüfrahmen bewerten; plattformspezifische Aspekte wie Navigation oder Betriebssicherheit werden jedoch nicht erfasst. Bei autonomen und mechanischen Verfahren kommt zudem dem Kulturverlust, etwa durch Überfahrten oder durch die Beschädigung von Kulturpflanzen durch das Werkzeug, ein deutlich höheres Gewicht zu. Im vorliegenden Prüfrahmen wird die Kulturverträglichkeit lediglich über die optionale Phytotoxizitätsbonitur erfasst (Kapitel 4.5). Bei einer Übertragung auf solche Verfahren wäre der Kulturverlust als eigene, verpflichtende Kenngröße zu ergänzen (Kapitel 7.4).

Ebenfalls nicht Gegenstand des Prüfrahmens ist die gerätetechnische Spritzqualität, etwa die Verteilungsgüte oder eine effektive Spritzbreite. Die Verteilungsgüte von Pflanzenschutzgeräten wird bereits im Rahmen der freiwilligen Anerkennungsprüfung des Julius Kühn-Instituts (JKI) geprüft, unter anderem über die Querverteilung auf einem 10-cm-Rinnenprüfstand. Zulässig sind dabei Abweichungen von höchstens ±15 % vom Mittelwert, was einem Variationskoeffizienten von 7 % entspricht (JKI 2013). Darüber hinaus müssen in Gebrauch befindliche Pflanzenschutzgeräte nach der Pflanzenschutz-Geräteverordnung in Abständen von sechs Kalenderhalbjahren durch eine Kontrollstelle geprüft werden (PflSchGerätV 2013). Der Prüfrahmen setzt daher ein technisch einwandfreies Applikationsgerät voraus und bewertet ausschließlich die Leistung der KI-gestützten Zielapplikation entlang der Funktionsbausteine Erkennung, Applikation und Wirkung. Der Baustein Applikation beschreibt dabei die räumliche Treffgenauigkeit bezogen auf das Ziel, nicht die Verteilungs- oder Dosiergenauigkeit des Geräts.

---

## 2. Stand der Technik und Einordnung des ASABE-Testprotokolls

### 2.1 Bewertungsansätze in der wissenschaftlichen Literatur

Die wissenschaftliche Bewertung der KI-basierten Erkennungsleistung erfolgt überwiegend anhand standardisierter Metriken aus der Objekterkennung. Grundlage ist die Einordnung jeder Detektion als True Positive, False Positive oder False Negative. Diese Einordnung erfolgt anhand der Intersection over Union (IoU) zwischen der vorhergesagten und der

referenzierten Bounding Box (Padilla et al. 2020). Aus diesen Klassifikationen werden Precision und Recall berechnet. Precision beschreibt den Anteil der korrekt erkannten Objekte an allen als positiv klassifizierten Objekten, während Recall den Anteil der korrekt erkannten Objekte an allen tatsächlich vorhandenen Objekten angibt. Beide Kennzahlen stehen in einem Zielkonflikt, der sich über die Precision-Recall-Kurve darstellen lässt. Zur Zusammenfassung der Erkennungsleistung wird die Average Precision (AP) verwendet; über mehrere Klassen gemittelt ergibt sich die Mean Average Precision (mAP) (Padilla et al. 2020).

Für die Erkennungsleistung KI-basierter Systeme werden unter kontrollierten Bedingungen in der Literatur häufig Genauigkeiten im Bereich von 80–90 % berichtet (Espinel et al. 2024; Gerhards et al. 2020). Mehrere Studien zeigen jedoch, dass unter Laborbedingungen trainierte Modelle im realen Feldeinsatz an Robustheit verlieren können. Gründe hierfür sind unter anderem wechselnde Lichtverhältnisse, Unterschiede in der Pflanzenmorphologie und eine höhere Verunkrautungsdichte als in den Trainingsdaten (Espinel et al. 2024; Mayer et al. 2024).

Diese Kennzahlen beschreiben ausschließlich die Erkennungsleistung. Für die Bewertung der anschließenden Applikationsgenauigkeit gibt es dagegen bislang keine vergleichbaren einheitlichen Standards. Noch weniger etabliert ist die Verknüpfung der gemessenen Applikationsgenauigkeit mit der tatsächlichen biologischen Wirkung. Die in der Literatur verwendeten Versuchsdesigns unterscheiden sich entsprechend deutlich. Sie reichen von manuellen Bonituren der Beikrautbestände bis hin zur langfristigen Beobachtung der Bekämpfungswirkung über mehrere Monate. Diese Unterschiede erschweren den Vergleich zwischen Studien und Systemen.

Der vorliegende Prüfrahmen berücksichtigt deshalb alle drei Schritte der Prozesskette und erfasst Erkennung, Applikation und Wirkung durchgängig im realen Pflanzenbestand. Die etablierten, literaturgestützten Kenngrößen zur Belagsverteilung und Trefferlage (vgl. Collectors in Kapitel 2.2) dienen dabei als zusätzliche und zeitnah erfassbare Zwischengrößen. Sie ersetzen jedoch nicht die separate Wirkungsbonitur (Kapitel 4.5).

### 2.2 Das ASABE/ISO-Testprotokoll im Detail

Im Rahmen von ISO/TC 23/SC 6/AHG 2 wurde im April 2025 gemeinsam mit ASABE ein Testprotokoll für Präzisionsapplikationstechnologien diskutiert (Hloben, John Deere, für AHG 2). Es stellt derzeit eine der konkretesten verfügbaren Grundlagen für standardisierte Freilandtests dar und bildet die methodische Basis des vorliegenden Prüfrahmens.

Die wesentlichen Vorgaben des Protokolls sind:

- **Testfeld**: ca. 100 × 200 m, brachliegend, kurz zuvor bearbeitet, vegetationsfrei und eben.
- **Targets**: lebende Pflanzen oder nicht lebende Ersatzziele mit einer maximalen Größe von 8 × 8 cm.
- **Collectors**: zwei Auffangpapiere je Ziel (Collector A: Länge, Collector B: Breite der Applikation) sowie fünf Positionen je Düsenabstand.

- **Farbmarkierung**: Verwendung einer Dye-Lösung im Spritztank (≥ 0,25 %) zur Sichtbarmachung der Applikation, einschließlich eines Primings von 5–10 s vor Testbeginn.
- **Umweltbedingungen**: Durchführung des Tests nur bei einer Windgeschwindigkeit von < 16,5 km/h. Zusätzlich werden unter anderem Temperatur, Luftfeuchtigkeit, Bodenart, Düsentyp und -abstand, Durchflussrate, Gestängehöhe und Fahrgeschwindigkeit dokumentiert.
- **Testablauf online**: Abfahren einer GPS-Referenzlinie, Platzierung der Collectors und Aktivierung des Applikationssystems während der Überfahrt. Anschließend wird die Trocknung abgewartet und der Versuch mehrfach wiederholt.
- **Testablauf offline**: Vermessung der Collector-Positionen mittels RTK-GPS und Erstellung einer Applikationskarte auf dieser Grundlage. Anschließend wird die Fläche im Kartenmodus mit dem Spritzgerät befahren und die Trocknung abgewartet.
- **Variationsmöglichkeiten**: Variation von Fahrgeschwindigkeit, Düsenabstand, Spritzdruck, Sensitivität und Gestängehöhe sowie der Anzahl der Wiederholungen (vorgeschlagen: 3 × 10–15 min).
- **Nicht Bestandteil des Protokolls**: die maschinelle Auswertung der Collectors mittels Scanner sowie die Messung des tatsächlich ausgebrachten Spritzvolumens.

Eine Übersicht über begleitende Normungsarbeiten, darunter ASABE X665 und ISO 24253, sowie methodisch vergleichbare Forschungsarbeiten findet sich in Anhang D.

### 2.3 Einordnung und identifizierte Lücken

Das ASABE/ISO-Protokoll liefert einen präzise definierten, gerätetechnisch ausgerichteten Ansatz zur Messung der Applikationsgenauigkeit. Im Mittelpunkt stehen dabei die Trefferlage und die Abdeckung relativ zum vorgegebenen Ziel. Das Protokoll unterscheidet außerdem zwischen online arbeitenden Systemen, die Ziele während der Überfahrt in Echtzeit erkennen und behandeln, und offline arbeitenden Systemen, bei denen die Applikation auf Grundlage einer zuvor erstellten Karte erfolgt. Übernommen werden die Unterscheidung zwischen Online- und Offline-Systemen, das Prinzip der Belagserfassung über Collectors sowie die Vorgaben zu den Umweltbedingungen (Kapitel 4). Die bestandsfreie Testfläche und die vorgegebenen Targets werden dagegen durch einen realen Pflanzenbestand mit natürlichem Beikrautbesatz ersetzt (Kapitel 3.2 und 4.1).

Der Abgleich mit den in Kapitel 2.1 beschriebenen Bewertungsansätzen zeigt jedoch zwei zentrale Lücken, die durch den vorliegenden Prüfrahmen adressiert werden.

**Fehlende Verknüpfung von Erkennung, Applikation und Wirkung.** Das ASABE-Protokoll bewertet die Applikationsgenauigkeit isoliert. Es gibt keine Auskunft darüber, ob die zugrunde liegende Erkennung korrekt war, wie robust diese unter realen Bestandsbedingungen funktioniert oder ob die resultierende Applikation tatsächlich zu einem Bekämpfungserfolg führt.

**Fehlende Trennbarkeit von Erkennungs- und Applikationsfehlern bei Online-Systemen.** Bei Online-Spot-Spraying-Systemen steht in der Praxis häufig keine Exportfunktion für annotierte Rohbilddaten der KI-basierten Erkennung zur Verfügung.

Eine getrennte Bewertung von Erkennung und Applikation ist dadurch erschwert. Der vorliegende Prüfrahmen begegnet diesem Problem nicht durch einen Verzicht auf die Erkennungsprüfung, sondern durch eine unabhängig vom System erhobene Ground-Truth-Kartierung des tatsächlichen Beikrautbestands. Diese wird, sofern verfügbar, mit einem As-Applied-Log der Düsenaktivierung kombiniert.

Das As-Applied-Log dokumentiert die Steuerentscheidung des Systems, also welche Düse zu welchem Zeitpunkt angesteuert wurde. Es beschreibt damit die vom System beabsichtigte Applikation, nicht die tatsächlich auf dem Ziel angekommene Spritzflüssigkeit. Der tatsächlich applizierte Belag wird deshalb weiterhin separat über die Collectors erfasst. Fehlt auch das As-Applied-Log, können Erkennungs- und Applikationsfehler bei einem Online-System nicht vollständig voneinander getrennt werden. In diesem Fall wird die kombinierte Erkennungs- und Applikationsgenauigkeit als nicht weiter auflösbare Kenngröße ausgewiesen.

---

## 3. Systematisierung: Online- vs. Offline-Spot-Spraying-Systeme

### 3.1 Definition und Abgrenzung

Für den vorliegenden Prüfrahmen wird zwischen Online- und Offline-Systemen unterschieden. Beide Systemtypen werden im Freiland auf Praxisschlägen getestet. Entscheidend für die Zuordnung ist, wann und wo die Entscheidung über die Applikation getroffen wird.

**Offlineverfahren (auch: absätziges Verfahren).** Erkennung und Applikation erfolgen zeitlich getrennt. Das Unkrautaufkommen wird zunächst mittels Kameradrohne durch einen Dienstleister kartiert. Aus dem georeferenzierten Orthofoto wird anschließend mithilfe einer KI eine Applikationskarte erstellt. Die eigentliche Ausbringung erfolgt zu einem späteren Zeitpunkt mit einer geeigneten Feldspritze, die über eine Einzel- oder Teilbreitenschaltung verfügt und die erstellte Karte abarbeitet. In der Praxis liegen zwischen Kartierung und Applikation üblicherweise etwa 24 Stunden bis drei Tage.

**Onlineverfahren (auch: integriertes Verfahren).** Erkennung und Applikation erfolgen während derselben Überfahrt. Innerhalb dieser Gruppe lassen sich zwei technische Ausprägungen unterscheiden, die für den Prüfrahmen relevant sind, da sie sich hinsichtlich ihrer möglichen Genauigkeit und ihrer Kostenstruktur unterscheiden:

- **Präzisionsfeldspritzen**: speziell für die Spot- beziehungsweise Patch-Applikation entwickelte Geräte mit hoher Einzeldüsendichte. Die Düsen sind häufig unter einer Haube angeordnet, die vor Wind und direktem Lichteinfall schützt. Im Vergleich zu klassischen Feldspritzen weisen diese Systeme meist eine geringere Arbeitsbreite auf.
- **Intelligente Feldspritzen**: klassische Feldspritzen mit großer Arbeitsbreite, die nachträglich mit Kameratechnik und Recheneinheiten am Gestänge ausgestattet wurden. Sie verbinden dadurch eine hohe Flächenleistung mit der Echtzeiterkennung von Beikräutern. Im Vergleich zu speziell entwickelten

Präzisionsfeldspritzen ist die räumliche Auflösung der Applikation in der Regel geringer.

Neben der Unterscheidung zwischen Online- und Offline-Systemen ist die Granularität der Applikation zu berücksichtigen. Bei der Behandlung einzelner Pflanzen wird von Spot-Applikation gesprochen, während bei der Behandlung zusammenhängender Unkrautnester der Begriff Patch-Applikation verwendet wird. Die gewählte Granularität beeinflusst unter anderem die Interpretation von Trefferlage und Spot-Größe sowie den erforderlichen Mindestabstand der Kontrollpositionen (Kapitel 4.4). Sie wird deshalb für jedes getestete System als zusätzliches Beschreibungsmerkmal dokumentiert, stellt jedoch kein eigenständiges Testkriterium dar.

Die Unterscheidung zwischen Online- und Offline-Systemen ist bereits im ASABE-Protokoll angelegt. Dort werden den unterschiedlichen Systemarchitekturen jedoch keine eigenen Prüfpunkte zugeordnet.

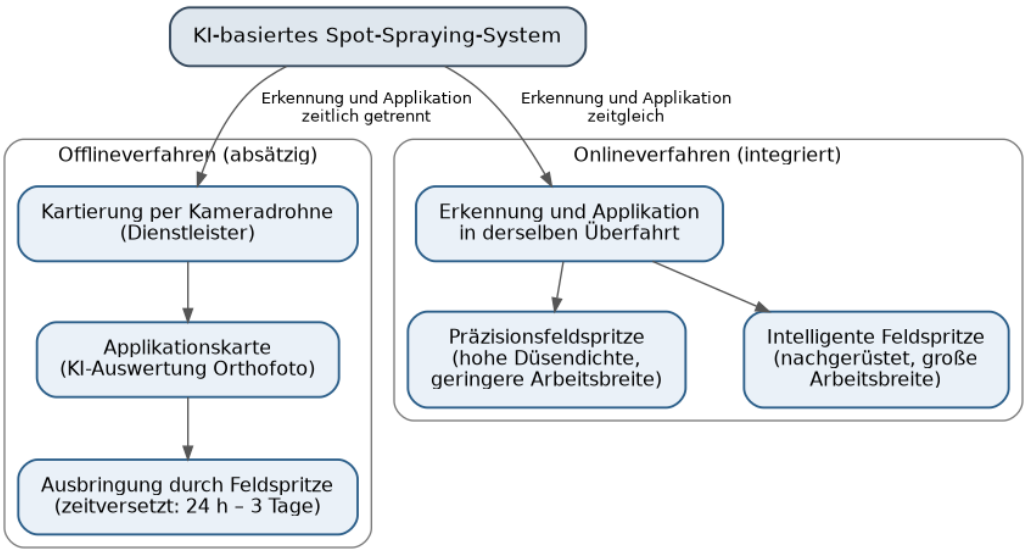



*Abbildung 1: Systematisierung von Online- und Offlineverfahren nach Zeitpunkt der Applikationsentscheidung.*

### 3.2 Funktionsbausteine je Systemtyp und Ground-Truth-Methodik

Der vorliegende Prüfrahmen umfasst alle drei in Kapitel 1 eingeführten Funktionsbausteine: Erkennung, Applikation und Wirkung. Die Prüfung erfolgt in einem realen, wachsenden Pflanzenbestand auf einem Praxisschlag mit natürlichem Beikrautbesatz. Dadurch wird vermieden, dass die Erkennungsleistung unter künstlich vereinfachten Bedingungen getestet wird. Attrappen, Herstellermarker oder isolierte Topfpflanzen können für ein Erkennungssystem leichter zu erfassen sein als Beikräuter, die zwischen Kulturpflanzen unter realen Feldbedingungen wachsen. Die Problematik der Übertragbarkeit von Labor- auf Feldbedingungen wird auch in Kapitel 2.1 beschrieben (Espinel et al. 2024; Mayer et al. 2024).

**Ground-Truth-Kartierung als Grundlage der Erkennungsprüfung.** Bei vielen Online-Systemen lassen sich die von der KI erzeugten und annotierten Rohbilddaten nicht exportieren. Eine direkte Bewertung der Erkennungsleistung anhand der systeminternen Erkennung ist damit nicht möglich. Der Prüfrahmen verwendet deshalb eine zweistufige Vorgehensweise, bei der die tatsächliche Position der Beikräuter unabhängig vom zu prüfenden System erfasst wird.

**Bevorzugt: As-Applied-Log.** Viele Systeme erstellen ein GPS-referenziertes As-Applied-Log, in dem dokumentiert wird, welche Düse beziehungsweise Applikationseinheit zu welchem Zeitpunkt aktiviert wurde. Ein solches Log bildet die Steuerentscheidung des Systems ab und ist in der Praxis häufig leichter verfügbar als annotierte Rohbilddaten, da entsprechende Informationen teilweise ohnehin zur Dokumentation des Applikationsvorgangs erfasst werden.

Liegt ein As-Applied-Log vor, kann die Prüfung in zwei getrennte Vergleiche aufgeteilt werden:

- **Ground-Truth-Kartierung vs. As-Applied-Log** → Erkennungsleistung: Wurde an den Positionen der tatsächlich vorhandenen Beikräuter eine Applikation ausgelöst?
- **As-Applied-Log vs. tatsächlicher Belag** → Applikationsgenauigkeit: Wurde die vom System getroffene Applikationsentscheidung physisch korrekt umgesetzt? Hierzu zählen beispielsweise die Reaktionszeit der Düse, mögliche Drift und die tatsächliche Trefferlage.

**Fallback: direkter Abgleich ohne As-Applied-Log.** Ist kein As-Applied-Log verfügbar, wird die Ground-Truth-Kartierung direkt mit dem tatsächlich gemessenen Belag verglichen. Für jedes kartierte Beikraut wird geprüft, ob an seiner Position ein Belag nachweisbar ist. Ein Beikraut mit nachgewiesenem Belag gilt dabei als erkannt und getroffen, ein Beikraut ohne nachweisbaren Belag als nicht erkannt beziehungsweise nicht getroffen. Diese Kenngröße erlaubt keine getrennte Bewertung von Erkennungs- und Applikationsfehlern. Beide Fehlerarten gehen in eine gemeinsame Kenngröße ein.

Sowohl die zweistufige Vorgehensweise mit As-Applied-Log als auch der Fallback-Abgleich beziehen sich ausschließlich auf die kartierten Beikrautpositionen und liefern damit den Recall: den Anteil der tatsächlich vorhandenen Beikräuter, die erkannt bzw. getroffen wurden. Ob ein System darüber hinaus fälschlich auf beikrautfreien Flächen appliziert (False Positives), lässt sich damit allein nicht feststellen. Dafür werden zusätzlich beikrautfreie Kontrollpositionen ausgewertet, aus denen sich die Fehlapplikationsrate ergibt (Kapitel 4.4). Recall und Fehlapplikationsrate beschreiben zusammen die Treffsicherheit eines Systems.

Bei Offline-Systemen liegt die Applikationskarte bereits als eigenständiges Zwischenergebnis vor und übernimmt methodisch eine ähnliche Funktion wie das As-Applied-Log bei Online-Systemen. Dadurch lassen sich bei Offline-Systemen zwei voneinander unabhängige Komponenten bewerten: Zum einen die Qualität der Kartierung und der daraus erzeugten Applikationskarte, zum anderen die Fähigkeit der eingesetzten Feldspritze, diese Karte im Feld korrekt umzusetzen. Die beiden Ergebnisse werden im Prüfbericht getrennt ausgewiesen. Andernfalls könnte ein Fehler in einem der beiden Schritte fälschlicherweise der jeweils anderen Komponente zugeschrieben werden.

**Wirkung.** Da die Prüfung in einem realen Pflanzenbestand unter den tatsächlich vorherrschenden Witterungsbedingungen und bei natürlichem Konkurrenzdruck stattfindet, kann auch die biologische Wirkung der Applikation unter praxisnahen Bedingungen erfasst werden. Die Wirkungsbonitur ist daher ein fester Bestandteil des Prüfrahmens.

Damit ergibt sich für beide Systemtypen dieselbe grundsätzliche Prüflogik:

| Funktionsbaustein | Offline-Systeme | Online-Systeme |
|---|---|---|
| **Erkennung** | Getrennt prüfbar: Ground-Truth-Kartierung vs. Applikationskarte | Getrennt prüfbar, sofern ein As-Applied-Log verfügbar ist; andernfalls nur als kombinierte Erkennungs- und Applikationskenngröße |
| **Applikation** | Genauigkeit der Umsetzung der | Genauigkeit der physischen Umsetzung; bei vorhandenem As-Applied-Log über |

| Funktionsbaustein | Offline-Systeme | Online-Systeme |
|---|---|---|
| | Applikationskarte im Feld | den Vergleich mit dem tatsächlichen Belag, ansonsten Bestandteil der kombinierten Kenngröße |
| **Wirkung** | Bekämpfungserfolg über den definierten Boniturzeitraum | Bekämpfungserfolg über den definierten Boniturzeitraum |

Der Prüfrahmen wird damit für beide Systemtypen entlang der drei Funktionsbausteine Erkennung, Applikation und Wirkung aufgebaut. Ob Erkennung und Applikation bei einem Online-System getrennt bewertet werden können, hängt von der Verfügbarkeit eines geeigneten As-Applied-Logs ab.

### 3.3 Testmethodische Relevanz der Unterscheidung

Die Unterscheidung zwischen Online- und Offline-Systemen bleibt trotz des gemeinsamen Grundgerüsts für die Versuchsplanung relevant.

**Zeitlicher Bezug der Prüfung**: Bei Offline-Systemen muss der Zeitraum zwischen Kartierung und Applikation erfasst und dokumentiert werden. Während dieser Zeit kann sich der tatsächliche Beikrautbestand verändern, beispielsweise durch weiteres Wachstum, das Verwelken vorhandener Pflanzen oder das Auflaufen neuer Keimlinge. Bei Online-Systemen entfällt diese zeitliche Verschiebung, da Erkennung und Applikation während derselben Überfahrt erfolgen.

**Zuordnung von Fehlerquellen**: Bei Offline-Systemen lassen sich Fehler grundsätzlich den einzelnen Prozessschritten zuordnen. Eine Abweichung kann entweder auf eine fehlerhafte Kartierung beziehungsweise Applikationskarte oder auf eine fehlerhafte Umsetzung der Karte durch die Feldspritze zurückzuführen sein. Beide Schritte werden separat erfasst und können gegen die Ground-Truth-Kartierung geprüft werden.

Bei Online-Systemen ist diese Trennung nur möglich, wenn ein As-Applied-Log verfügbar ist. Ohne ein solches Log lässt sich beispielsweise bei einem fehlenden Belag an einem vorhandenen Beikraut nicht feststellen, ob das System das Beikraut bereits nicht erkannt hat oder ob die Applikation trotz korrekter Erkennung nicht ordnungsgemäß umgesetzt wurde. Der Fall wird deshalb als kombinierte Erkennungs- und Applikationskenngröße ausgewiesen. Der Einsatz eines Farbtracers verbessert die Genauigkeit der Belagsmessung, ermöglicht jedoch keine zusätzliche Trennung der beiden Fehlerquellen. Dafür ist das As-Applied-Log erforderlich.

**Vergleichbarkeit zwischen Anbietern**: Ziel des Prüfrahmens ist nicht, Online- und Offline-Systeme als unterschiedliche Kategorien direkt gegeneinander zu ranken. Im Vordergrund steht vielmehr die möglichst faire und nachvollziehbare Bewertung verschiedener Systeme innerhalb derselben Systemkategorie. Die Unterschiede zwischen den Systemarchitekturen werden dabei dokumentiert und bei der Interpretation der Ergebnisse berücksichtigt.

### 3.4 Zusammenfassung: Kennzahlen je Funktionsbaustein und Systemtyp

Die folgenden beiden Tabellen fassen zusammen, welche Kennzahlen der Prüfrahmen je Funktionsbaustein und Systemtyp liefert. Die methodischen Details zu den einzelnen Kennzahlen sind in den Kapiteln 4–6 beschrieben.

Erkennungs- und Applikationskennzahlen basieren dabei durchgängig auf derselben Vier-Felder-Logik aus Ground-Truth-Positionen (Kapitel 4.2) und Kontrollpositionen (Kapitel 4.4):

| | **appliziert** | **nicht appliziert** |
|---|---|---|
| **Beikrautposition** (Ground-Truth, Kap. 4.2) | **True Positive (TP)** korrekt erkannt/getroffen | **False Negative (FN)** Beikraut übersehen |
| **Beikrautfreie Kontrollposition** (Kap. 4.4) | **False Positive (FP)** Fehlapplikation | **True Negative (TN)** korrekt nicht appliziert |

**Recall** = TP / (TP + FN) → Anteil der kartierten Beikräuter, die getroffen wurden (obere Zeile)
**Fehlapplikationsrate** = FP / (FP + TN) → Anteil der Kontrollpositionen mit Belag (untere Zeile)
Eine Precision wird nur als Szenariorechnung für eine angegebene Beikrautdichte ausgewiesen (Kap. 4.4).

*Abbildung 5: Berechnungsgrundlage für Recall und Fehlapplikationsrate im Prüfrahmen.*

Aus der oberen Zeile ergibt sich der Recall, aus der unteren die Fehlapplikationsrate. Eine Precision wird dagegen nicht direkt ausgewiesen, weil sie vom Verhältnis der beiden Stichproben und von der Beikrautdichte des Testfelds abhängt (Kapitel 4.4).

**Online-Systeme**

| Funktionsbaustein | Kennzahl | Datengrundlage |
|---|---|---|
| Erkennung | Recall der Erkennung (Anteil erkannter Beikräuter, mit 95-%-Konfidenzintervall) – sofern As-Applied-Log verfügbar | Ground-Truth-Kartierung vs. As-Applied-Log |
| Erkennung + Applikation | Fehlapplikationsrate (Anteil beikrautfreier Kontrollpositionen mit Belag, mit 95-%-Konfidenzintervall) | Kontrollpositionen (beikrautfrei) vs. Belag |
| Erkennung + Applikation (Fallback) | Kombinierter Recall (Anteil der Beikräuter mit Belag, mit 95-%-Konfidenzintervall) – falls kein As-Applied-Log verfügbar | Ground-Truth-Kartierung vs. tatsächlicher Belag |
| Applikation | Applikationsgenauigkeit (Reaktionszeit, Drift, Trefferlage) – sofern As-Applied-Log verfügbar | As-Applied-Log vs. tatsächlicher Belag |
| Applikation | Belag und Trefferlage | Collector-Auswertung an Beikrautpositionen |
| Applikation | Reaktionslatenz (räumlich, ggf. zeitlich) | As-Applied-Log bzw. Farbtracer-Belag vs. Ground-Truth-Position |
| Applikation | Spot-Größe (Über-/Unterdosierung) | Applizierte Fläche vs. Zielgröße |
| Wirkung | Bekämpfungserfolg (Absterbebonitur bzw. Frischmasse-Rückgang %) | Wirkungsbonitur ca. 4 Wochen nach Applikation |
| Wirkung | Phytotoxizität/Kulturverträglichkeit (EWRS-Skala, optional) | Visuelle Bonitur |

**Offline-Systeme**

| Funktionsbaustein | Kennzahl | Datengrundlage |
|---|---|---|
| Erkennung (Kartierung) | Recall und Fehlausweisungsrate der Applikationskarte (jeweils mit 95-%-Konfidenzintervall); bei flächenhaften Zielen zusätzlich IoU | Ground-Truth-Kartierung bzw. Kontrollpositionen vs. Applikationskarte |
| Erkennung (Kartierung) | Räumliche Genauigkeit der Kartenposition | Ground-Truth-Position vs. kartierte Zielposition |
| Applikation | Umsetzungsgenauigkeit der Karte (Georeferenzierung, Düsenpositionierung) | Applikationskarte vs. tatsächlicher Belag |
| Applikation | Belag und Trefferlage | Collector-Auswertung an Beikrautpositionen |
| Applikation | Fehlapplikationsrate außerhalb der Kartenzielflächen (mit 95-%-Konfidenzintervall) | Kontrollpositionen vs. Belag |
| Applikation | Zeitversatz Kartierung–Applikation (Dokumentation) | Zeitstempel Systemkartierung vs. Applikation |
| Wirkung | Bekämpfungserfolg (Absterbebonitur bzw. Frischmasse-Rückgang %) | Wirkungsbonitur ca. 4 Wochen nach Applikation |
| Wirkung | Phytotoxizität/Kulturverträglichkeit (EWRS-Skala, optional) | Visuelle Bonitur |

Auffällig ist, dass bei Offline-Systemen die Erkennung (Kartierung) durchgängig verpflichtend und trennscharf bewertbar ist, während sie bei Online-Systemen von der Verfügbarkeit eines As-Applied-Logs abhängt. Die Wirkung wird dagegen bei beiden Systemtypen identisch erfasst, da sie unabhängig von der Systemarchitektur im selben Bestand und mit derselben Methodik bonitiert wird.

---

## 4. Prüfrahmen – Allgemeiner Aufbau

### 4.1 Testfeld und Streifenanlage

Der Prüfrahmen wird auf einem Praxisschlag mit wachsender Kultur und natürlichem Beikrautbesatz durchgeführt. Zeitpunkt und Ort der Prüfung sind damit an den jeweiligen Vegetationszyklus gebunden.

- **Streifenanlage**: Innerhalb des Schlags werden mehrere parallele Teststreifen angelegt. Ihre Breite richtet sich nach der Arbeitsbreite des jeweils getesteten Systems. Pro System sind mindestens drei Wiederholungsstreifen vorzusehen. Die Zuordnung der Streifen zu den geprüften Systemen bzw. Wiederholungen sollte randomisiert oder in randomisierten Blöcken erfolgen, um systematische Verzerrungen durch Gradienten innerhalb des Schlags (z. B. Bodengüte, Feuchtigkeit,

Windexposition, Verunkrautungsdichte) zu vermeiden; eine rein sequenzielle Anordnung ist zu vermeiden.

- **Streifenlänge**: Die Streifenlänge richtet sich nach der für die Applikation erforderlichen Beschleunigungs- und Bremsstrecke von jeweils etwa 15–20 m sowie einer Messstrecke von mindestens 20–30 m. In der Praxis wird die verfügbare Streifenlänge häufig durch die Länge des Schlags begrenzt.
- **Unbehandelte Kontrollfenster**: Als Referenz für die Wirkungsbonitur (Kapitel 4.5) werden statt eines durchgehenden unbehandelten Kontrollstreifens kleinflächige, unbehandelte Kontrollfenster angelegt, mindestens eines je Block beziehungsweise Wiederholung. Die Fenster liegen außerhalb der Messstrecke, beispielsweise im Bereich der ohnehin nicht ausgewerteten Beschleunigungs- oder Bremsstrecke, und sollten einen mit der Messstrecke vergleichbaren Beikrautbesatz aufweisen. Bei der Überfahrt bleiben sie unbehandelt, etwa durch Abschalten der Applikation oder Abdecken der Fläche. Die Fenster sind so klein wie möglich zu halten, aber groß genug für die vorgesehene Bonitur; als Orientierung dienen etwa 3 × 6 m oder weniger. Erfahrungen aus Praxisversuchen zeigen, dass unbehandelte Flächen im Saisonverlauf auch bei geringem Unkrautdruck deutlich verunkrauten und bereits Fenster dieser Größe von Bewirtschaftern als Belastung wahrgenommen werden. Die Kontrollfenster sind nicht mit den beikrautfreien Kontrollpositionen zur Erfassung von Fehlapplikationen zu verwechseln (Kapitel 4.4), die innerhalb der behandelten Messstrecke liegen.
- **Zeitpunkt**: Der Test muss in einem Zeitraum stattfinden, in dem die Kultur sowohl ein für die Herbizidanwendung zugelassenes Entwicklungsstadium erreicht hat als auch ein ausreichender und für die Auswertung geeigneter Beikrautbesatz vorhanden ist. Der geeignete Zeitraum ist vorab mit den Praxispartnern abzustimmen und kann je nach Kultur und Witterungsverlauf variieren.
- **Wahl des Schlags**: Da natürlicher Beikrautbesatz genutzt wird, sollte die Fläche vor dem Test auf einen geeigneten Besatz geprüft werden. Angestrebt wird ein **moderater Unkrautdruck**: ausreichend Beikräuter für eine belastbare Ground-Truth-Stichprobe (Kapitel 4.2), aber kein sehr hoher, flächendeckender Befall. Das entspricht zugleich dem eigentlichen Einsatzzweck von Spot-Spraying-Systemen, die für Einzelpflanzen- bzw. Nesterbekämpfung bei überschaubarem Unkrautdruck konzipiert sind und nicht für eine annähernd flächendeckende Verunkrautung, bei der eine klassische Flächenspritzung ohnehin effizienter wäre. Ein moderater Unkrautdruck erleichtert zudem die Einhaltung des in Kapitel 4.4 beschriebenen Mindestabstands zwischen Kontrollpositionen und Beikrautpositionen, da bei geringerer Pflanzendichte mehr geeigneter, abstandskonformer Freiraum für Kontrollpositionen zur Verfügung steht. Stark ungleichmäßig verunkrautete, nahezu beikrautfreie oder sehr dicht bzw. flächendeckend verunkrautete Flächen sind für die Prüfung nur eingeschränkt geeignet.

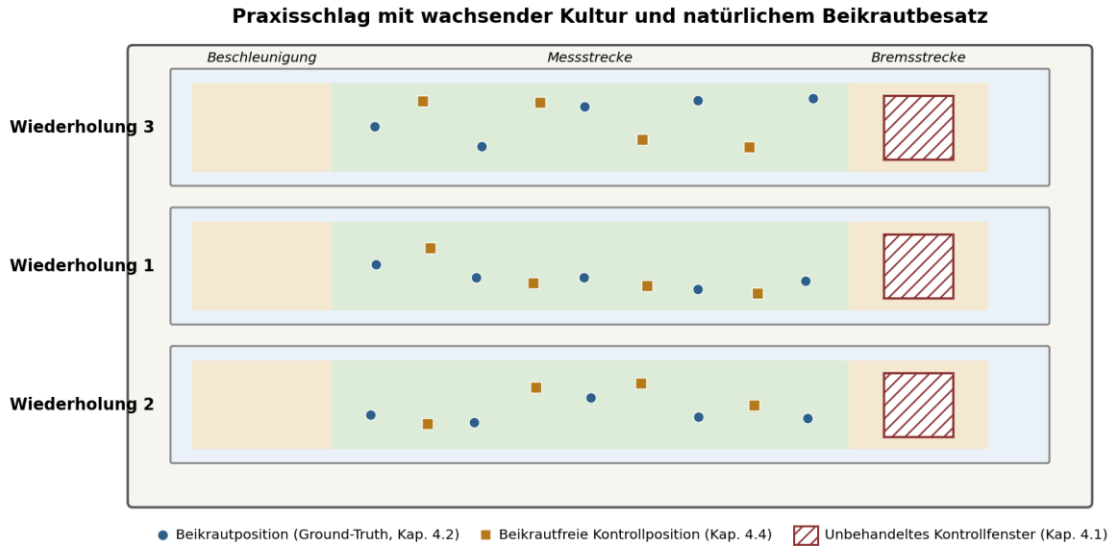



*Abbildung 6: Schematische Streifenanlage mit randomisierter Streifenfolge, Beschleunigungs-/Mess-/Bremsstrecke, unbehandelten Kontrollfenstern sowie Beikraut- und Kontrollpositionen (vereinfachte Darstellung; tatsächlich empfohlen ≥ 30 Positionen je Kategorie und Streifen, Kapitel 4.2/4.4).*

### 4.2 Ground-Truth-Kartierung des Beikrautbestands

Vor der Applikation wird der tatsächliche Beikrautbestand in jedem Teststreifen unabhängig vom zu prüfenden System erfasst. Die Erfassung erfolgt an einer repräsentativen Stichprobe von Positionen und nicht als vollständige Kartierung aller vorhandenen Beikräuter. An denselben Positionen werden anschließend die Collectors platziert (Kapitel 4.3).

- **Vermessung**: Position, Art und Wachstumsstadium einer festgelegten Anzahl von Beikräutern je Teststreifen werden erfasst. Vorgesehen sind mindestens 30 Positionen pro Streifen – dieselbe Zielanzahl wie bei den Kontrollpositionen (Kapitel 4.4), damit Recall und Fehlapplikationsrate auf vergleichbar belastbarer statistischer Grundlage stehen. Die Kartierung erfolgt manuell, beispielsweise mit einem RTK-Handheld-GPS, oder über eine hochauflösende Drohnenbefliegung. Zusätzlich werden in den unbehandelten Kontrollfenstern (Kapitel 4.1) einzelne Beikräuter markiert, die als Referenz für die Wirkungsbonitur dienen (Kapitel 4.5).
- **Dokumentation**: Für jedes vermessene Beikraut werden Position, Art und Wachstumsstadium (BBCH) dokumentiert. Zusätzlich wird die Gesamtverunkrautung des Streifens in Pflanzen/$m^2$ geschätzt. Diese Angabe dient dazu, Unterschiede zwischen Testterminen und Flächen bei der späteren Einordnung der Ergebnisse berücksichtigen zu können.
- **Zeitliche Nähe und Reihenfolge**: Bei Online-Systemen sollte die Kartierung möglichst am Tag der Applikation erfolgen. Bei Offline-Systemen sollte die Ground-Truth-Kartierung erst nach dem systemeigenen Kartierungsflug durchgeführt werden. Dadurch wird vermieden, dass die Kartierung durch das Betreten des Bestands unnötige Trittspuren erzeugt, die die anschließende Applikation

beeinflussen könnten. Der Zeitpunkt der systemeigenen Kartierung wird dabei als zusätzlicher zeitlicher Bezugspunkt dokumentiert.

Für diesen Ablauf befindet sich projektintern ein eigener Workflow in Entwicklung (Arbeitstitel **WMC**), der in einer späteren Fassung dieses Reports ergänzt wird.

Die Ground-Truth-Stichprobe dient als Referenz für die Bewertung der Erkennungs- und Applikationsgenauigkeit.

### 4.3 Collector-Platzierung

An den kartierten Beikrautpositionen jedes Teststreifens (Kapitel 4.2) werden zusätzlich Auffangpapiere platziert, um den Applikationsbelag und die Trefferlage zu erfassen (Kapitel 2.2). Vorgesehen sind mindestens 30 Positionen pro Streifen, die möglichst gleichmäßig über die Messstrecke verteilt werden (vgl. Begründung der Zielanzahl in Kapitel 4.4). Bei der Auswahl sollten unterschiedliche Wachstumsstadien sowie verschiedene Positionen innerhalb des Streifens, beispielsweise zentrale und randständige Pflanzen, berücksichtigt werden.

Als Collector dient einfaches Auffangpapier. Damit der Belag darauf sichtbar wird, wird dem Herbizid ein zugelassener Markierungsstoff beigemischt, sofern dieser mit dem verwendeten Wirkstoff verträglich ist und die biologische Wirkung nicht beeinflusst (vgl. hierzu jedoch die Einschränkung in Kapitel 4.7). Wassersensitives Papier reagiert zwar auch ohne Farbtracer auf die wässrige Spritzbrühe, kann an den Kontrollpositionen aber nicht zwischen einer Applikation und anderen Ursachen einer Verfärbung unterscheiden (Kapitel 4.4). Es kommt daher allenfalls ergänzend infrage, etwa für eine feinere Auswertung des Tropfenspektrums an den Beikrautpositionen.

Die platzierten Collectors selbst können bei Online-Systemen potenziell als visuell auffälliges, untypisches Objekt im Bildausschnitt der Erkennungssensorik wirken und dadurch die Zielerkennung beeinflussen – etwa indem sie fälschlich mit erkannt oder als Störsignal behandelt werden. Dieser mögliche Effekt ist bei der Interpretation der Erkennungsergebnisse zu berücksichtigen (Kapitel 4.7) und sollte nach Möglichkeit durch eine unauffällige, bodennahe Platzierung der Collectors minimiert werden.

### 4.4 Kontrollpositionen und Fehlapplikationsrate

Die in Kapitel 3.2 beschriebene Ground-Truth-Kartierung und die daraus abgeleiteten Kenngrößen erfassen ausschließlich, ob an den Positionen tatsächlich vorhandener Beikräuter appliziert wurde. Damit lässt sich nur der Recall bestimmen: der Anteil der tatsächlich vorhandenen Beikräuter, die erkannt bzw. getroffen wurden. Ob das System darüber hinaus auch fälschlich auf Flächen ohne Beikraut appliziert (False Positives), bleibt mit dieser Methodik allein nicht feststellbar.

Um auch Fehlapplikationen zu erfassen, werden zusätzlich zu den in Kapitel 4.2 kartierten Beikrautpositionen definierte **Kontrollpositionen ohne Beikraut** in den Teststreifen festgelegt und mit Collectors versehen (nachweislich beikrautfreie Flächen von verbindlich 50 × 50 cm). Die Größe ist einheitlich einzuhalten, da eine größere Kontrollfläche häufiger getroffen wird und die Ergebnisse sonst nicht vergleichbar wären. Wird an einer Kontrollposition dennoch ein Belag nachgewiesen, gilt dies als Fehlapplikation (False

Positive). Die **Fehlapplikationsrate** ist der Anteil der Kontrollpositionen mit Fehlapplikation. Zusammen mit dem Recall an den Beikrautpositionen beschreibt sie die Treffsicherheit des Systems (Abbildung 5).

Die erforderliche Anzahl an Kontrollpositionen richtet sich nicht automatisch nach der Anzahl der Beikrautpositionen (Kapitel 4.3), sondern nach der Fehlapplikationsrate, die mit dem Test noch zuverlässig nachweisbar sein soll. Werden bei $n$ Kontrollpositionen keine Fehlapplikationen beobachtet, lässt sich mit einer gängigen statistischen Faustregel ("Rule of Three") mit ca. 95 % Konfidenz lediglich ausschließen, dass die tatsächliche Fehlapplikationsrate über $3/n$ liegt. Als grobe Orientierung ergeben sich daraus folgende Mindestanzahlen:

| Zuverlässig ausschließbare Fehlapplikationsrate | Erforderliche Kontrollpositionen (ca.) |
|---|---|
| < 10 % | 30 |
| < 5 % | 60 |
| < 1 % | 300 |

Für den Prüfrahmen wird als praxistauglicher Kompromiss eine Zielanzahl von **30 Kontrollpositionen je Streifen** empfohlen, verteilt gemäß der Stichprobenlogik der übrigen Collector-Platzierung (Kapitel 4.3). Damit lassen sich grobe Schwächen (Fehlapplikationsraten ≥ 10 %) verlässlich erkennen; ein Nachweis sehr niedriger Fehlapplikationsraten (< 5 %), wie sie von leistungsfähigen Systemen zu erwarten sind, erfordert deutlich mehr Kontrollpositionen, als im Rahmen eines einzelnen Testdurchlaufs praktikabel ist. Der Testbericht muss daher neben der beobachteten Fehlapplikationsrate stets auch die zugrunde liegende Anzahl an Kontrollpositionen und das 95-%-Konfidenzintervall ausweisen; dasselbe gilt für den Recall und die Anzahl der ausgewerteten Beikrautpositionen (Kapitel 8.3). Die Intervalle werden nach einem exakten Binomialverfahren berechnet (z. B. Clopper-Pearson). Werden keine Fehlapplikationen beobachtet, entspricht die einseitige obere 95-%-Grenze näherungsweise $3/n$ (Rule of Three). Beide Kennzahlen sind damit nicht als exakte Werte, sondern als durch die Stichprobengröße begrenzt belastbare Schätzungen zu interpretieren.

**Mindestabstand zu Beikrautpositionen.** Eine Kontrollposition ist nur dann aussagekräftig, wenn ein dort nachgewiesener Belag tatsächlich auf eine Fehlapplikation zurückgeht und nicht auf Abdrift bzw. Overspray einer legitimen, korrekt auf ein benachbartes Beikraut ausgerichteten Applikation. Kontrollpositionen müssen deshalb einen Mindestabstand zu allen Beikräutern in ihrer Umgebung einhalten, also nicht nur zu den kartierten Beikrautpositionen (Kapitel 4.2), da die Ground-Truth-Kartierung lediglich eine Stichprobe umfasst. Der Mindestabstand orientiert sich an der erwarteten Spot- bzw. Patch-Größe des geprüften Systems (Kapitel 3.1) – als Orientierungswert mindestens das Doppelte des vom Hersteller angegebenen maximalen Spot- bzw. Patch-Durchmessers zuzüglich eines Sicherheitsabstands für Abdrift.

Dieser Mindestabstand verringert bei Systemen mit größerer Spot- bzw. Patch-Applikation (Kapitel 3.1) die Anzahl der innerhalb eines Streifens tatsächlich verfügbaren, geeigneten Kontrollpositionen erheblich – insbesondere bei natürlich und damit unregelmäßig verteiltem Beikrautbesatz mit entsprechend wenig durchgehend freiem Raum zwischen den Pflanzen. Die bewusste Wahl eines Schlags mit moderatem statt hohem Unkrautdruck

(Kapitel 4.1) wirkt dem entgegen, da mehr abstandskonformer Freiraum zur Verfügung steht; sie kann das Problem bei sehr großem Spot-/Patch-Radius jedoch nicht vollständig ausräumen. Die empfohlene Zielanzahl von 30 Kontrollpositionen ist daher bei Systemen mit großem Spot-/Patch-Radius unter Umständen nicht innerhalb eines Streifens erreichbar. In diesem Fall ist eine der folgenden Optionen zu wählen und im Testbericht zu dokumentieren:

1. Anlage zusätzlicher Streifen ausschließlich zur Gewinnung weiterer, abstandskonformer Kontrollpositionen, oder
2. Akzeptanz einer geringeren Anzahl an Kontrollpositionen mit entsprechend breiterer Konfidenzgrenze der Fehlapplikationsrate.

Eine Unterschreitung der Zielanzahl ist damit kein Verstoß gegen den Prüfrahmen, muss aber transparent gemacht werden. Sie verzerrt die Fehlapplikationsrate nicht, verbreitert aber ihr Konfidenzintervall und schränkt damit die Aussagekraft ein.

Diese Methode setzt voraus, dass ein **Farb- oder Fluoreszenztracer** appliziert wird (Kapitel 4.3). Ohne Tracer, insbesondere bei reiner Verwendung von wassersensitivem Papier, lässt sich an einer Kontrollposition nicht zuverlässig zwischen einer tatsächlichen Applikation und anderen Ursachen für eine Verfärbung unterscheiden (z. B. Tau, Niederschlag oder Luftfeuchtigkeit, die wassersensitives Papier unabhängig von einer Applikation reagieren lassen können). Deshalb wird an den Kontrollpositionen einfaches Auffangpapier mit Farbtracer verwendet; wassersensitives Papier allein ist für diesen Zweck nicht geeignet.

**Precision nur als Szenariorechnung.** Eine Precision, also der Anteil berechtigter Applikationen an allen Applikationen, wird nicht direkt aus den Stichproben berechnet. Sie hinge sonst vom gewählten Verhältnis von Beikraut- zu Kontrollpositionen ab: Weniger Kontrollpositionen, wie sie bei großen Spots zulässig sind, würden sie künstlich erhöhen. Zudem hängt die Precision im Feld von der Beikrautdichte ab, da Treffer pro Beikraut, Fehlapplikationen aber pro beikrautfreier Fläche entstehen. Das entspricht dem positiven Vorhersagewert diagnostischer Tests, der von der Häufigkeit der Erkrankung abhängt. Bei Bedarf kann die Precision für eine angegebene Beikrautdichte vereinfacht aus Recall und Fehlapplikationsrate abgeleitet werden:

$$P_{\text{Szenario}} = (R \cdot d) \,/\, (R \cdot d + 4 \cdot F)$$

Dabei ist $R$ der Recall, $F$ die Fehlapplikationsrate, $d$ die Beikrautdichte in Pflanzen/m² und 4 die Anzahl der 50 × 50 cm großen Flächeneinheiten je Quadratmeter; die von den Beikräutern selbst belegte Fläche ist vernachlässigt. Ein System mit 90 % Recall und 10 % Fehlapplikationsrate erreicht danach bei 0,5 Pflanzen/m² eine Precision von etwa 53 %, bei 2 Pflanzen/m² von etwa 82 %. Eine solche Precision ist im Testbericht ausdrücklich als Szenariorechnung mit der zugrunde gelegten Beikrautdichte zu kennzeichnen. Werden mehrere Systeme verglichen, ist für alle dieselbe Beikrautdichte anzusetzen.

### 4.5 Wirkstoff und Wirkungsbonitur

Die Anwendung erfolgt nach guter fachlicher Praxis und unter Einhaltung der geltenden gesetzlichen Vorgaben.

Die Wirkungsbonitur orientiert sich an etablierten Methoden der Herbologie (Anhang D: Quan et al. 2025).

- **Boniturzeitpunkt**: Die Hauptbonitur erfolgt etwa vier Wochen nach der Applikation, damit ausreichend Zeit für eine sichtbare Reaktion der Kultur- und Beikrautpflanzen zur Verfügung steht. Bei Bedarf können zusätzliche Zwischenbonituren, beispielsweise nach ein bis zwei Wochen, durchgeführt werden, um den zeitlichen Verlauf der Wirkung zu erfassen.
- **Methodik**: Als einfachste Variante wird für jedes markierte Beikraut an der Ground-Truth-Position (Kapitel 4.2) visuell erfasst, ob die Pflanze abgestorben ist beziehungsweise keine aktive Wachstumsreaktion mehr zeigt. Die Bewertung erfolgt als Ja/Nein-Entscheidung. Zum Vergleich werden die in den unbehandelten Kontrollfenstern markierten Beikräuter (Kapitel 4.1) ebenfalls bonitiert, um ein natürliches Absterben, etwa infolge von Trockenheit, von der Herbizidwirkung abgrenzen zu können. Für eine quantitativere Bewertung können zusätzlich mehrere Erhebungsquadrate, beispielsweise mit einer Größe von 0,5 × 0,5 m, ausgewertet werden. Diese werden zufällig oder gezielt an den kartierten Positionen im Teststreifen sowie als Referenz in den Kontrollfenstern angelegt. Die oberirdische Beikrautbiomasse wird dabei geschnitten und nach Art getrennt gewogen. Die Wirkung wird als prozentuale Abnahme der Frischmasse gegenüber den unbehandelten Kontrollfenstern berechnet (Anhang D). Die Biomassemethode ermöglicht eine feinere quantitative Bewertung, ist jedoch deutlich aufwändiger als die reine Absterbebonitur und wird daher als Erweiterungsoption betrachtet.
- **Ergänzende Bonitur**: Optional kann zusätzlich eine visuelle Wirkungs- beziehungsweise Phytotoxizitätsbonitur nach der EWRS-Skala durchgeführt werden (1 = keine Schädigung, 9 = Pflanze abgestorben). Damit kann neben der Beikrautwirkung auch die Kulturverträglichkeit der Anwendung erfasst werden.

### 4.6 Umweltbedingungen und Dokumentation

In Anlehnung an die Vorgaben aus Kapitel 2.2 werden für jeden Testdurchlauf mindestens die folgenden Bedingungen und Systemparameter dokumentiert:

- Systembezeichnung und eingesetzte Softwareversion des geprüften Systems, da sich KI-Modelle zwischen Softwareversionen ändern können (vgl. Kapitel 4.7) und Ergebnisse ohne diese Angabe nicht eindeutig einer Systemversion zuordenbar wären
- Windgeschwindigkeit und Windrichtung; bei einer Windgeschwindigkeit von > 16,5 km/h wird der Test analog zum ASABE-Protokoll abgebrochen
- Temperatur und relative Luftfeuchtigkeit
- Bodenart und Bodenzustand
- Vorfrucht und vorhandene Pflanzenrückstände
- Kultur- und Beikraut-Wachstumsstadium (BBCH)
- Tageszeit und Lichtverhältnisse
- Fahrgeschwindigkeit
- Düsen- beziehungsweise Applikationseinheitenabstand
- Gestänge- beziehungsweise Arbeitshöhe

- verwendeter Applikationsmodus

### 4.7 Grenzen der Übertragbarkeit

Der beschriebene Prüfrahmen ermöglicht die Erfassung der Erkennungs-, Applikations- und Wirkungsgenauigkeit unter realen Feldbedingungen. Diese Bedingungen lassen sich jedoch nicht vollständig standardisieren. Bei der Interpretation der Ergebnisse sind daher mehrere Einschränkungen zu berücksichtigen, die in jedem Testbericht ausdrücklich dokumentiert werden sollten.

**Boden- und Rückstandseinfluss**: Bodenart, Bodenfeuchte und vorhandene Pflanzenrückstände, beispielsweise Mulch, Ernterückstände oder Strohauflagen, können die optische Erkennung der Beikräuter beeinflussen und damit auch die Applikationsgenauigkeit einzelner Systeme verändern. Ergebnisse sind daher nicht ohne Weiteres auf andere Standortbedingungen übertragbar.

**Natürliche Verunkrautung als Unsicherheitsfaktor**: Durch die Verwendung eines natürlichen und nicht gezielt eingebrachten Beikrautbesatzes können sich Beikrautdichte, -verteilung und Artenspektrum zwischen Testterminen und Flächen unterscheiden. Diese Eigenschaften lassen sich im Vorfeld nur begrenzt planen. Ein Testtermin mit zu geringem oder stark ungleichmäßig verteiltem Beikrautbesatz kann daher keine ausreichend belastbare Datengrundlage liefern. Ein zu hoher, annähernd flächendeckender Befall ist ebenfalls ungeeignet (Kapitel 4.1): Er entspricht nicht dem eigentlichen Einsatzzweck von Spot-Spraying-Systemen und lässt zudem zu wenig abstandskonformen Freiraum für die in Kapitel 4.4 vorgesehenen Kontrollpositionen.

**Witterung und Jahr**: Ein einzelner Testdurchlauf bildet die Witterungs- und Vegetationsbedingungen des jeweiligen Jahres ab. Für verallgemeinerbare Aussagen sind daher Wiederholungen über mehrere Jahre und möglichst auch an unterschiedlichen Standorten sinnvoll. Ein weiterer Grund für wiederholte Prüfungen liegt in der schnellen Weiterentwicklung der KI-Modelle der Hersteller. Ein Testergebnis beschreibt in erster Linie den Zustand des Systems beziehungsweise der eingesetzten Softwareversion zum Zeitpunkt der Prüfung und lässt sich nicht ohne Weiteres auf spätere Versionen desselben Systems übertragen.

**Fahrgeschwindigkeit**: Die Prüfung erfolgt ausschließlich bei der vom Hersteller empfohlenen Arbeitsgeschwindigkeit. Da Erkennungs- und Applikationsgenauigkeit mit zunehmender Geschwindigkeit abnehmen können, sind die Ergebnisse nicht ohne Weiteres auf andere Fahrgeschwindigkeiten übertragbar.

**Beeinflussung durch die Ground-Truth-Erhebung**: Das Betreten des Bestands zur Vorabkartierung kann Trittspuren verursachen und den Bestand lokal verändern. Dadurch kann wiederum die Erkennung des zu prüfenden Systems beeinflusst werden. Die Kartierung sollte daher möglichst zielgerichtet erfolgen und auf die tatsächlich benötigten Positionen beschränkt bleiben. Gleichzeitig ist auf einen organisierten Zugang zum Bestand zu achten, um unnötige Störungen zu vermeiden.

**Mögliche Störung der Zielerkennung durch die Collectors selbst**: Die zur Belagsmessung platzierten Auffangpapiere können als zusätzliche, für das

Erkennungssystem untypische Objekte im Bildausschnitt wirken (Kapitel 4.3) und dadurch die Erkennungsleistung des geprüften Systems in nicht quantifizierter Weise beeinflussen. Dieser Effekt lässt sich mit dem vorliegenden Prüfrahmen nicht vollständig ausschließen und ist bei der Interpretation der Erkennungsergebnisse zu berücksichtigen.

**Eingeschränkte Aussagekraft der Fehlapplikationsrate bei großer Spot-/Patch-Applikation**: Der in Kapitel 4.4 vorgeschriebene Mindestabstand zwischen Kontrollpositionen und Beikrautpositionen (zur Vermeidung einer Kontamination durch Abdrift legitimer Applikationen) reduziert bei Systemen mit großem Spot- bzw. Patch-Radius die Anzahl der im Streifen verfügbaren, gültigen Kontrollpositionen. Systeme mit größerer Applikationsfläche pro Ziel erreichen dadurch tendenziell eine geringere Anzahl auswertbarer Kontrollpositionen und damit ein breiteres Konfidenzintervall der Fehlapplikationsrate als Systeme mit kleineren Spots. Dies ist bei Anbietervergleichen innerhalb derselben Systemkategorie (Kapitel 3.3) zu berücksichtigen und im Testbericht transparent zu machen.

**Mögliche Konfundierung der Wirkungsbonitur durch den Farbtracer**: Der dem Herbizid beigemischte Farb- oder Fluoreszenztracer (Kapitel 4.3) lagert sich als sichtbarer Belag auf der Blattoberfläche des Beikrauts ab. Ein solcher Belag kann die Photosyntheseleistung der Pflanze zusätzlich beeinträchtigen, unabhängig von der eigentlichen herbiziden Wirkung. Eine im Rahmen der Wirkungsbonitur (Kapitel 4.5) beobachtete Schwächung oder ein Absterben der Pflanze wäre in diesem Fall nicht eindeutig auf den Wirkstoff selbst zurückzuführen, sondern könnte teilweise oder vollständig auf den Farbbelag selbst zurückgehen. Dieser mögliche Effekt ist bei der Interpretation der Wirkungsergebnisse zu berücksichtigen, da Belag und Wirkung an denselben Pflanzen erhoben werden; eine Validierung des Effekts (z. B. durch ein zusätzliches Kontrollfenster, in dem ausschließlich der Farbtracer ohne Wirkstoff appliziert wird) ist nicht Bestandteil des vorliegenden Prüfrahmens, aber als Erweiterungsoption denkbar.

---

## 5. Prüfrahmen – Online-Systeme

### 5.1 Testablauf

Der Testablauf orientiert sich am ASABE-Protokoll und wird an die Streifenanlage im realen Pflanzenbestand sowie an die Applikation eines zugelassenen Herbizids angepasst.

- **Anfrage zum As-Applied-Log**: Vor Beginn der Prüfung wird beim Hersteller angefragt, ob ein GPS-referenziertes As-Applied-Log der Düsenaktivierung exportiert werden kann. Die Verfügbarkeit des Logs bestimmt, ob Erkennungs- und Applikationsgenauigkeit getrennt oder nur als kombinierte Kenngröße bewertet werden können. Das Ergebnis der Anfrage wird im Testbericht dokumentiert.
- **Ground-Truth-Kartierung**: Eine repräsentative Anzahl von Beikräutern je Teststreifen sowie in den Kontrollfenstern wird vor der Applikation vermessen und dokumentiert (Kapitel 4.2). Die Erfassung erfolgt beispielsweise mit einem RTK-Handheld-GPS oder durch eine geeignete Drohnenbefliegung und sollte möglichst am Tag der Applikation durchgeführt werden.

- **Vorbereitung**: An den zuvor kartierten Beikrautpositionen sowie an den beikrautfreien Kontrollpositionen (Kapitel 4.4) werden die Collectors platziert.
- **Applikation**: Das System durchfährt den Teststreifen mit der vom Hersteller empfohlenen Arbeitsgeschwindigkeit. Geprüft wird ausschließlich bei dieser einen Geschwindigkeit; die tatsächlich gefahrene Geschwindigkeit wird dokumentiert (Kapitel 4.6). Die Zielerkennung und Applikation sind dabei aktiviert. Das zugelassene Herbizid wird entsprechend den geltenden Anwendungsbestimmungen ausgebracht.
- **Trocknung und Einsammeln**: Nach der Überfahrt wird gewartet, bis die Collectors vollständig getrocknet sind. Anschließend werden sie eingesammelt, eindeutig beschriftet und für die spätere Belagsauswertung gelagert (Kapitel 4.3).
- **Wiederholung**: Die Schritte Ground-Truth-Kartierung, Collector-Platzierung, Applikation sowie Trocknung und Einsammeln werden für jeden Teststreifen wiederholt. Pro System sind mindestens drei Wiederholungen vorzusehen.
- **Wirkungsbonitur**: Die Wirkungsbonitur erfolgt etwa vier Wochen nach der Applikation. Bei Bedarf können zusätzliche Zwischenbonituren durchgeführt werden (Kapitel 4.5).

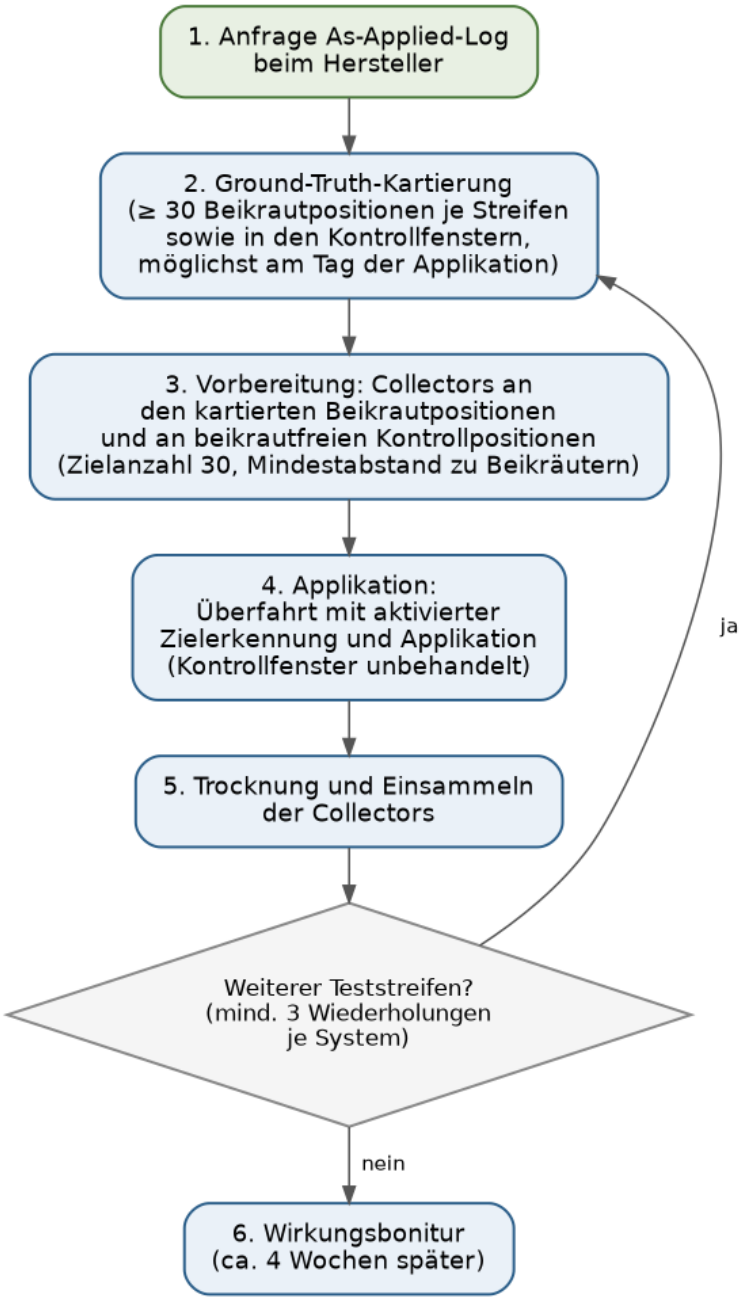



*Abbildung 2: Testablauf für Online-Systeme (Kapitel 5.1).*

### 5.2 Prüfpunkte und Messgrößen

Neben den allgemeinen Kenngrößen des Prüfrahmens – Belag und Trefferlage, Erkennungs- beziehungsweise kombinierte Erkennungs- und Applikationsgenauigkeit, Fehlapplikationsrate an Kontrollpositionen (Kapitel 4.4) sowie Bekämpfungswirkung – werden bei Online-Systemen zusätzliche systembezogene Größen erfasst. Diese ergeben sich aus der Echtzeiterkennung und der zeitlich unmittelbar anschließenden Applikation.

**Reaktionslatenz**: Die Reaktionslatenz beschreibt den zeitlichen beziehungsweise räumlichen Abstand zwischen der Erfassung eines Zielobjekts durch die Sensorik und dem

Beginn der Applikation. Liegt ein As-Applied-Log vor, kann die räumliche Abweichung zwischen der dokumentierten Applikationsentscheidung und dem tatsächlich gemessenen Belag ausgewertet werden. Die tatsächliche zeitliche Latenz lässt sich daraus jedoch nur bestimmen, wenn zusätzlich Informationen zur Position und Geschwindigkeit des Systems sowie zur zeitlichen Zuordnung der Applikationsentscheidung vorliegen.

Auch ohne As-Applied-Log kann die räumliche Abweichung zwischen Zielposition und tatsächlichem Applikationspunkt bestimmt werden. Voraussetzung ist, dass der applizierte Belag auf dem Collector eindeutig sichtbar ist; dies leistet der Farbtracer (Kapitel 4.3). Da die Position des Ground-Truth-Ziels bekannt ist, kann der Schwerpunkt des tatsächlichen Belags mit der Zielposition verglichen werden. Diese Abweichung beschreibt jedoch die resultierende räumliche Treffergenauigkeit und nicht unmittelbar die Reaktionslatenz des Erkennungssystems.

**Spot-Größe**: Für jedes Ziel wird die tatsächlich applizierte Fläche bestimmt und zur Größe des Zielobjekts ins Verhältnis gesetzt. Dadurch lässt sich beurteilen, ob das System tendenziell größere Flächen als erforderlich behandelt oder ob Teile des Zielobjekts nicht getroffen werden. Die Spot-Größe kann damit sowohl zur Bewertung einer möglichen Überapplikation als auch einer unvollständigen Abdeckung herangezogen werden.

---

## 6. Prüfrahmen – Offline-Systeme

### 6.1 Ground-Truth-Kartierung und Systemkartierung

Die Kartierung des Beikrautbestands durch das zu prüfende System erfolgt entsprechend der üblichen Betriebsweise des jeweiligen Systems, beispielsweise mittels Kameradrohne. Die unabhängige Ground-Truth-Erfassung erfolgt anschließend. Dadurch wird vermieden, dass das Betreten des Bestands vor der Systemkartierung Trittspuren verursacht oder den Bestand anderweitig beeinflusst.

- **Systemkartierung**: Das System kartiert die Teststreifen entsprechend der vorgesehenen Betriebsweise. Dabei werden die für den jeweiligen Einsatz üblichen Parameter wie Flughöhe, räumliche Auflösung und Bildüberlappung verwendet und dokumentiert.
- **Ground-Truth-Kartierung**: Nach Abschluss der Systemkartierung wird der tatsächliche Beikrautbestand entsprechend der in Kapitel 4.2 beschriebenen Methodik erfasst.
- **Kartenerstellung**: Aus den vom System erfassten Rohdaten wird die Applikationskarte erstellt. Die dafür benötigte Zeit wird dokumentiert, ebenso relevante Verarbeitungsschritte, sofern diese für die spätere Bewertung der Karte von Bedeutung sind.

### 6.2 Erkennungsprüfung (verpflichtend)

Die Bewertung der Kartierungsleistung ist bei Offline-Systemen ein fester Bestandteil des Prüfrahmens, da die Applikationskarte als eigenständiges Zwischenergebnis vorliegt.

Dadurch kann die Qualität der Kartierung unabhängig von der späteren Umsetzung durch die Feldspritze bewertet werden.

Geprüft wird zunächst, welcher Anteil der in der Ground-Truth-Kartierung erfassten Beikräuter in der Applikationskarte korrekt als Zielfläche ausgewiesen wird (Recall der Karte). Analog zur Fehlapplikationsrate wird anhand der Kontrollpositionen (Kapitel 4.4) geprüft, welcher Anteil der beikrautfreien Kontrollpositionen in der Karte fälschlich als Zielfläche ausgewiesen ist (Fehlausweisungsrate). Eine Precision der Karte wird aus denselben Gründen wie in Kapitel 4.4 nicht direkt ausgewiesen. Zusätzlich wird die räumliche Genauigkeit der kartierten Zielpositionen relativ zur Ground-Truth-Position bestimmt. Bei flächenhaften Zielflächen kann ergänzend die Übereinstimmung der kartierten und tatsächlich vorhandenen Flächen über die Intersection over Union (IoU) bewertet werden.

### 6.3 Testablauf der Applikation

- **Kartenübertragung**: Die erzeugte Applikationskarte wird auf die für die Ausbringung eingesetzte Feldspritze übertragen. Der verwendete Übertragungsweg und das Kartenformat werden dokumentiert.
- **Collector-Platzierung**: Die Collectors werden gemäß Kapitel 4.3 an den kartierten Beikrautpositionen sowie an den beikrautfreien Kontrollpositionen (Kapitel 4.4) platziert.
- **Durchfahrt**: Die Feldspritze durchfährt den Teststreifen mit der vom Hersteller empfohlenen Arbeitsgeschwindigkeit und appliziert das vorgesehene Herbizid entsprechend der Applikationskarte.
- **Trocknung und Einsammeln**: Nach der Überfahrt werden die Collectors vollständig getrocknet und anschließend eingesammelt und für die Belagsauswertung gelagert, analog zu Kapitel 5.1.
- **Wiederholung**: Die Prüfung wird mit mindestens drei Wiederholungen je System durchgeführt. Zusätzlich wird der tatsächliche zeitliche Abstand zwischen Systemkartierung und Applikation dokumentiert.
- **Wirkungsbonitur**: Die Wirkungsbonitur erfolgt etwa vier Wochen nach der Applikation.

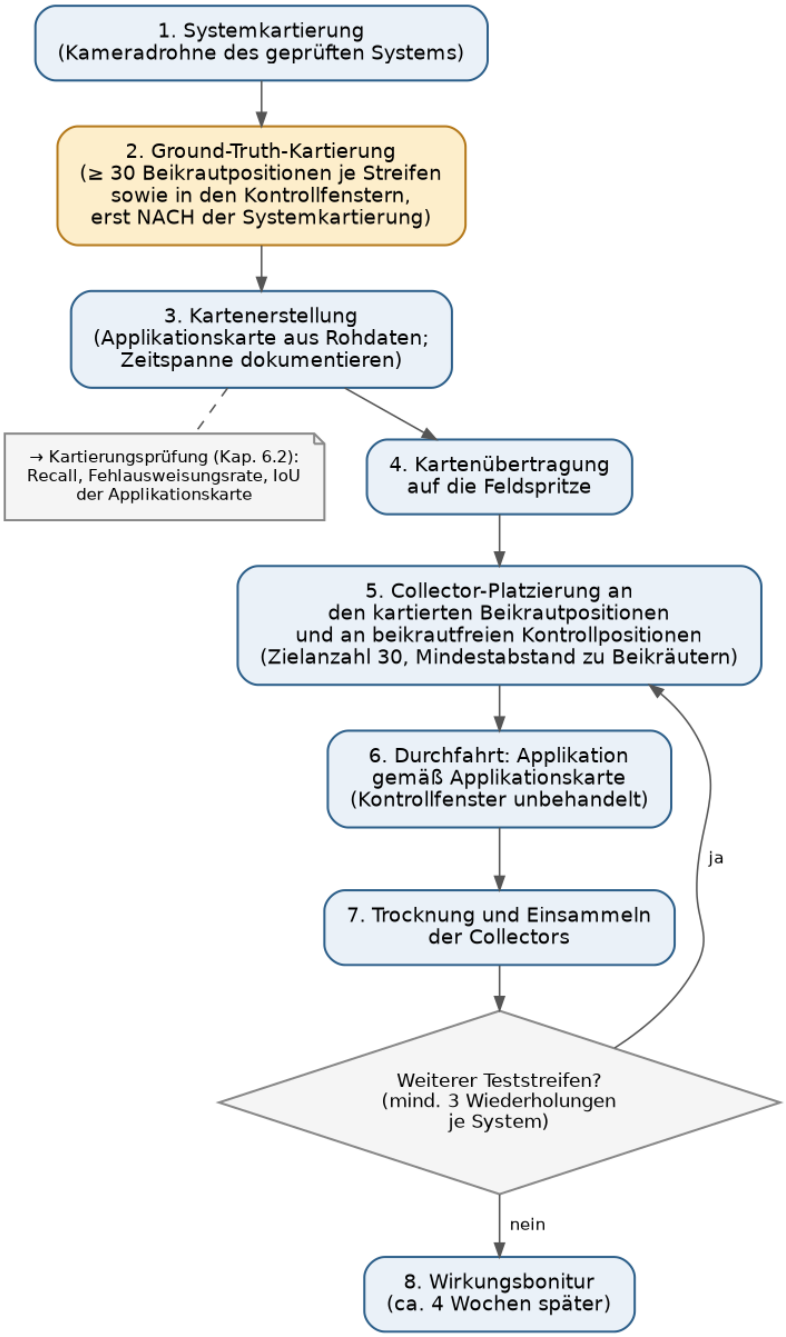



*Abbildung 3: Testablauf für Offline-Systeme (Kapitel 6.1 und 6.3).*

### 6.4 Prüfpunkte und Messgrößen

Neben den allgemeinen Kenngrößen des Prüfrahmens, insbesondere Belag, Trefferlage und Bekämpfungswirkung, werden bei Offline-Systemen zwei voneinander getrennte Bereiche bewertet.

**Kartierungsgenauigkeit**: Die Qualität der erzeugten Applikationskarte wird durch den Vergleich mit der Ground-Truth-Kartierung bestimmt. Verwendet werden der Recall und die Fehlausweisungsrate der Karte (Kapitel 6.2), die räumliche Abweichung zwischen

kartierter und tatsächlich vorhandener Zielposition sowie bei flächenhaften Zielflächen die IoU.

**Umsetzungsgenauigkeit der Applikationskarte**: Die Fähigkeit der Feldspritze, die vorgegebene Applikationskarte im Feld korrekt umzusetzen, wird unabhängig von der Kartierungsleistung bewertet. Hierzu wird die Abweichung zwischen der in der Applikationskarte vorgegebenen Zielposition und dem tatsächlich gemessenen Belagsschwerpunkt bestimmt. Diese Abweichung umfasst unter anderem mögliche Georeferenzierungsfehler der Feldspritze, Fehler bei der Positionierung der Düsen sowie weitere Abweichungen bei der Umsetzung der Karte. Die Qualität der ursprünglichen Kartierung wird dabei nicht erneut bewertet, sondern ist bereits Bestandteil der separaten Kartierungsprüfung. Zusätzlich ergibt sich aus den Kontrollpositionen, die in der Karte nicht als Zielfläche ausgewiesen sind, die Fehlapplikationsrate der Umsetzung (Kapitel 4.4): Ein Belag an einer solchen Position geht auf die Feldspritze zurück, etwa infolge von GPS-Ungenauigkeiten oder einer verzögerten Düsenschaltung.

---

## 7. Erweiterte Bewertungskriterien und Forschungsausblick

Für die Beurteilung eines Systems unter Praxisbedingungen sind neben den im Prüfrahmen erfassten technischen und biologischen Kenngrößen weitere Kriterien relevant. Dies gilt insbesondere für die Bewertung über mehrere Jahre und unterschiedliche Standorte hinweg. Diese Kriterien sind nicht Bestandteil des eigentlichen Prüfrahmens, können jedoch auf dessen Ergebnissen aufbauen und als Erweiterungsmodule oder als Ansatzpunkte für weiterführende Untersuchungen dienen.

### 7.1 Ökonomische Kriterien (optional)

Für die betriebliche Entscheidung ist neben der technischen Leistungsfähigkeit vor allem die Wirtschaftlichkeit eines Systems relevant. An den vorliegenden Prüfrahmen lassen sich unter anderem folgende Kenngrößen anschließen:

- **Anschaffungs- und Betriebskosten**: Die Kosten können in Bezug zur Arbeitsbreite und zur erzielbaren Flächenleistung gesetzt werden.
- **Herbizideinsparung gegenüber einer flächendeckenden Applikation**: Da im Prüfrahmen echter Wirkstoff unter realen Bestandsbedingungen appliziert wird, kann die tatsächlich ausgebrachte Wirkstoffmenge für den jeweiligen Teststreifen aus den Applikationsdaten bestimmt und mit einer flächendeckenden Applikation verglichen werden. Die in der Literatur berichteten Einsparpotenziale können weiterhin als Vergleichsgröße herangezogen werden (z. B. Leise et al. 2025; Azghadi et al. 2024/2025). Die im Prüfrahmen ermittelte Einsparung beschreibt jedoch nur die Bedingungen des jeweiligen Testtermins. Standort, Jahr und Verunkrautungsgrad können das Einsparpotenzial erheblich beeinflussen, sodass die Ergebnisse nicht ohne Weiteres verallgemeinert werden können.
- **Amortisationsdauer**: Die erforderliche Amortisationszeit hängt unter anderem von der Betriebsgröße, der Auslastung des Systems und dem jeweiligen Preisniveau für Pflanzenschutzmittel ab.

Eine umfassende ökonomische Bewertung, beispielsweise in Form einer Vollkostenrechnung oder einer detaillierten Amortisationsrechnung, ist nicht Bestandteil des Prüfrahmens. Sie stellt einen eigenständigen Untersuchungsgegenstand dar, kann jedoch direkt auf den erhobenen Daten zur Erkennungs-, Applikations- und Wirkungsleistung aufbauen.

### 7.2 Ökologische Kriterien (optional)

Die ökologische Wirkung KI-basierter Präzisionsapplikatoren ergibt sich im Wesentlichen aus der gezielten Reduktion der ausgebrachten Pflanzenschutzmittelmenge und ist damit eine Folge der Präzisionsapplikation und kein eigenständiges Applikationsziel. Aus den Ergebnissen des Prüfrahmens und durch ergänzende Untersuchungen lassen sich verschiedene ökologische Fragestellungen ableiten:

- **Reduktion der ausgebrachten Herbizidmenge auf Schlagebene**: Diese Größe kann direkt aus den Applikationsdaten des Prüfrahmens bestimmt werden.
- **Abdrift- und Austragsrisiko**: Von Interesse ist insbesondere der mögliche Eintrag von Pflanzenschutzmitteln in angrenzende Gewässer. Erste Felddaten weisen auf einen Zusammenhang zwischen der Applikationsgenauigkeit und der Wirkstoffkonzentration im Oberflächenabfluss hin (Azghadi et al. 2024/2025).
- **Auswirkungen auf die Biodiversität**: Eine geringere behandelte Fläche kann insbesondere in Randbereichen und nicht behandelten Teilflächen Auswirkungen auf die dort vorhandenen Lebensgemeinschaften haben. Diese Effekte sind jedoch nicht allein aus der technischen Applikationsgenauigkeit abzuleiten und erfordern zusätzliche Untersuchungen.

### 7.3 Agronomischer Forschungsbedarf: Ertragswirkung und Generalisierbarkeit

Der Prüfrahmen erfasst den Bekämpfungserfolg anhand der Wirkungsbonitur und optional über die Phytotoxizität. Die tatsächliche Auswirkung auf den Ertrag der Kultur wird dagegen nicht erfasst. Eine belastbare Ertragsbewertung ist erst zum Erntezeitpunkt möglich und liegt damit deutlich außerhalb des vorgesehenen Boniturzeitraums. Für eine vollständige agronomische Bewertung wäre daher eine Ertragserhebung auf den Teststreifen erforderlich. Eine entsprechende Methodik wird beispielsweise von Quan et al. (2025, Anhang D) verwendet. Die Ertragserhebung ist als Erweiterungsoption und nicht als verpflichtender Bestandteil des Prüfrahmens vorgesehen.

Auch die Übertragbarkeit der Ergebnisse bleibt begrenzt. Ein einzelner Testdurchlauf bildet die Witterungsbedingungen, den Bodenzustand und die Verunkrautungssituation eines bestimmten Standorts und Jahres ab. Für belastbare Aussagen über die Leistungsfähigkeit eines Systems unter unterschiedlichen Bedingungen sind daher Wiederholungen über mehrere Jahre, Standorte und Bodenarten erforderlich. Diese Untersuchungen gehen über den einzelnen Prüfzyklus hinaus und stellen einen eigenständigen Forschungsbedarf dar.

Eine weitere offene Frage betrifft die langfristige Entwicklung des Unkrautbestands bei wiederholtem Einsatz von Spot-Spraying-Systemen. Da auch leistungsfähige Systeme nicht alle vorhandenen Beikräuter erkennen und behandeln, können nicht erfasste Pflanzen zur Samenbank des Bodens beitragen. Welche Auswirkungen dies auf die langfristige

Zusammensetzung und Dichte der Unkrautpopulation hat, lässt sich mit einem einzelnen Testdurchlauf nicht beurteilen.

Ähnliches gilt für die Entwicklung von Herbizidresistenzen. Eine wiederholte, selektive Bekämpfung könnte den Selektionsdruck innerhalb von Unkrautpopulationen verändern. Ob und unter welchen Bedingungen daraus ein erhöhtes Resistenzrisiko entsteht, erfordert mehrjährige Untersuchungen unter kontrollierten und praxisnahen Bedingungen.

### 7.4 Übertragung auf mechanische und physikalische Systeme

Der vorliegende Prüfrahmen wurde für Spot-Spraying-Systeme entwickelt. Die grundlegende Struktur lässt sich jedoch auch auf mechanische und andere physikalische Verfahren zur Beikrautregulierung übertragen. Dazu zählen beispielsweise Hacksysteme sowie Laser- oder elektrophysikalische Verfahren.

Die Ground-Truth-Kartierung und die Wirkungsbonitur könnten dabei weitgehend unverändert übernommen werden. Der Applikationsbaustein müsste dagegen durch ein für das jeweilige Verfahren geeignetes Äquivalent ersetzt werden. Bei Hacksystemen wären beispielsweise die Bearbeitungsgenauigkeit und -tiefe relevant, während bei Lasersystemen der Energieeintrag und die Einwirkzeit im Vordergrund stehen würden. Zusätzlich wäre der Kulturverlust als verpflichtende Kenngröße aufzunehmen, da mechanische Verfahren Kulturpflanzen unmittelbar beschädigen oder ausreißen können und dieser Aspekt dort zu den zentralen Bewertungskriterien zählt (vgl. Kapitel 1).

Die Übertragung des Prüfrahmens auf weitere Verfahren ist im weiteren Verlauf des Projekts vorgesehen.

---

## 8. Anwendung des Prüfrahmens

Dieses Kapitel fasst die in den Kapiteln 3–6 beschriebenen Anforderungen für die praktische Durchführung des Prüfrahmens zusammen. Es dient als Checkliste und wiederholt bewusst keine methodischen Begründungen. Diese sind in den jeweils referenzierten Kapiteln beschrieben.

### 8.1 Voraussetzungen (Kurzcheck)

Vor Beginn der Prüfung müssen folgende Voraussetzungen erfüllt sein:

- Praxisschlag mit wachsender Kultur und ausreichendem, aber moderatem und möglichst gleichmäßigem Beikrautbesatz; die Fläche wurde vorab auf ihre Eignung geprüft (Kapitel 4.1).
- Windgeschwindigkeit < 16,5 km/h zum Zeitpunkt der Applikation sowie Einhaltung der produktspezifischen Anwendungsbestimmungen (Kapitel 4.5 und 4.6).
- Technisch einwandfreies Applikationsgerät unter Einhaltung der gesetzlichen Gerätekontrollpflicht; die Verteilungsgüte selbst ist nicht Gegenstand des Prüfrahmens (Kapitel 1).
- RTK-fähiges Handheld-GPS oder Drohne zur Ground-Truth-Kartierung.

- Einfaches Auffangpapier als Collector-Material sowie ein zugelassener Farbtracer als Zusatz zur Spritzbrühe (Kapitel 4.3 und 4.4); wassersensitives Papier allenfalls ergänzend.
- Festlegung und Vermessung beikrautfreier Kontrollpositionen (50 × 50 cm, Zielanzahl 30 je Streifen; Kapitel 4.4).
- Festlegung unbehandelter Kontrollfenster außerhalb der Messstrecke, mindestens eines je Block beziehungsweise Wiederholung (Orientierung: 3 × 6 m oder kleiner; Kapitel 4.1).
- Scanner oder Kamera zur Digitalisierung der Collectors.
- Bildanalysesoftware zur Auswertung von Belag und Trefferlage, beispielsweise ImageJ, DepositScan oder DropSense.
- Vor Beginn der Prüfung: Anfrage beim Hersteller, ob ein GPS-referenziertes As-Applied-Log exportiert werden kann (Kapitel 3.2 und 5.1).
- Ausrüstung für die Wirkungsbonitur etwa vier Wochen nach der Applikation, insbesondere Erhebungsquadrate und eine Waage zur Bestimmung der Frischmasse sowie gegebenenfalls ein Erhebungsbogen für die EWRS-Bonitur.

### 8.2 Ablaufschema (Kurzfassung)

1. **Schlag und Testzeitpunkt prüfen**: Eignung des Beikrautbesatzes, Kulturstadium und Witterungsbedingungen prüfen sowie erforderliche Genehmigungen und Anwendungsbestimmungen klären.
2. **Systeminformationen erfassen**: Systemtyp und Softwareversion dokumentieren und beim Hersteller die Verfügbarkeit eines As-Applied-Logs anfragen.
3. **Ground-Truth-Kartierung**: Den Beikrautbestand entsprechend Kapitel 4.2 erfassen.
4. **Collector-Vorbereitung**: Collectors an den kartierten Beikraut- und Kontrollpositionen platzieren, die unbehandelten Kontrollfenster festlegen und die Umweltbedingungen dokumentieren.
5. **Applikation durchführen**: Den Testablauf gemäß Kapitel 5 für Online-Systeme beziehungsweise Kapitel 6 für Offline-Systeme durchführen.
6. **Collectors auswerten**: Belag und Trefferlage erfassen und die vorgesehenen Kennzahlen berechnen.
7. **Wirkungsbonitur**: Etwa vier Wochen nach der Applikation den Bekämpfungserfolg erfassen.
8. **Ergebnisse dokumentieren**: Alle Pflichtangaben und Ergebnisse entsprechend Kapitel 8.3 im Testbericht festhalten.

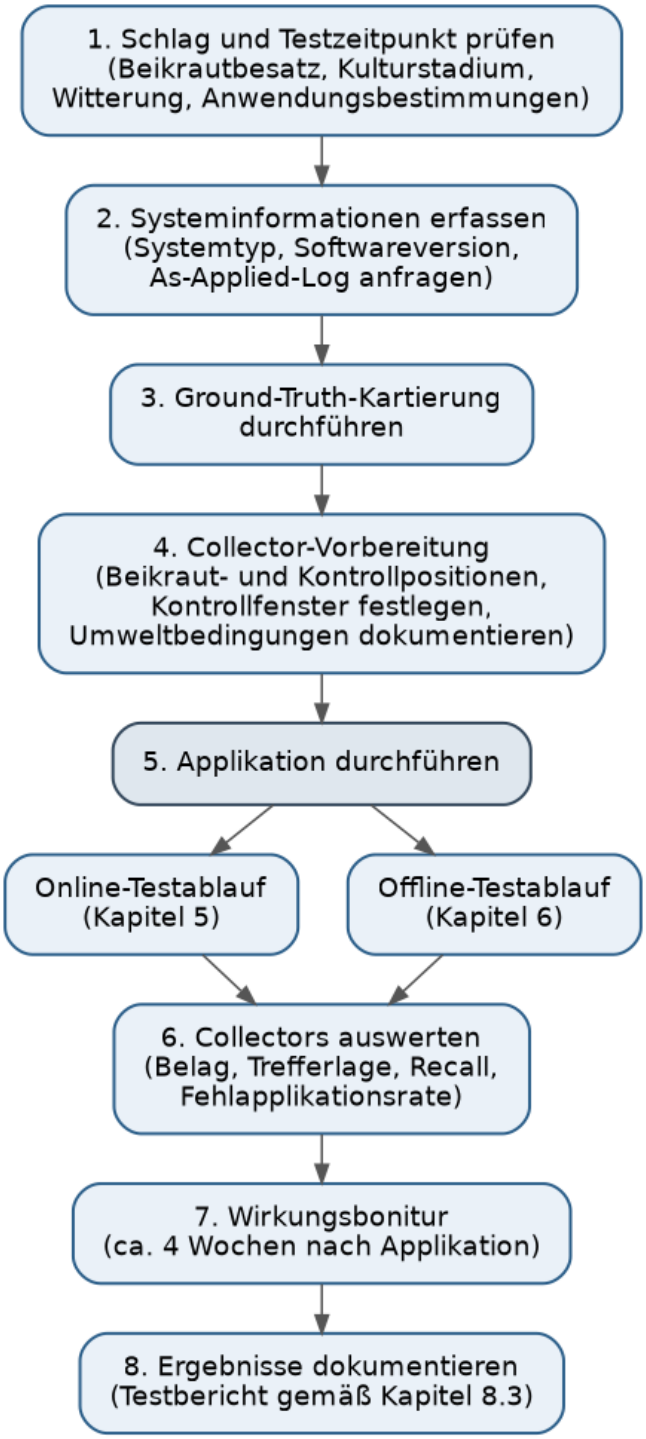



*Abbildung 4: Gesamtablauf des Prüfrahmens von der Flächenprüfung bis zur Dokumentation.*

### 8.3 Dokumentation (Pflichtfelder im Testbericht)

Der Testbericht muss mindestens folgende Angaben enthalten:

**System und Verfahren**

- Systembezeichnung und Systemtyp (Online/Offline; bei Online-Systemen Präzisionsfeldspritze oder intelligente Feldspritze).
- Hersteller und eingesetzte Softwareversion.
- Verwendetes Herbizid und Aufwandmenge.
- Gefahrene Arbeitsgeschwindigkeit und weitere Systemparameter gemäß Kapitel 4.6.
- Bei Offline-Systemen zusätzlich Angaben zur Kartierung und zur eingesetzten Feldspritze.

**Testbedingungen**

- Testdatum und Standort.
- Umweltbedingungen während der Applikation.
- Kultur und Kulturstadium.
- Beikrautdichte, vorhandene Arten und Wachstumsstadien.
- Anzahl der Wiederholungen.
- Anzahl, Größe und Lage der unbehandelten Kontrollfenster.

**Erhobene Ergebnisse**

- Gemessene Kennzahlen je Wiederholung für Erkennung beziehungsweise Kartierung, Applikation und Wirkung.
- Ergebnisse der Collector-Auswertung einschließlich Belag und Trefferlage.
- Recall mit Anzahl der ausgewerteten Beikrautpositionen und 95-%-Konfidenzintervall (Kapitel 4.4).
- Anzahl der Kontrollpositionen, Anzahl der dort festgestellten Fehlapplikationen und Fehlapplikationsrate mit 95-%-Konfidenzintervall (Kapitel 4.4).
- Falls eine Precision angegeben wird: ausschließlich als Szenariorechnung mit Angabe der zugrunde gelegten Beikrautdichte (Kapitel 4.4).
- Ergebnis der Wirkungsbonitur.

**Auswertbarkeit**

- Angabe, ob ein As-Applied-Log vorlag.
- Angabe, welche Kenngrößen getrennt für Erkennung und Applikation ausgewertet werden konnten und welche nur als kombinierte Kenngröße vorliegen.
- Bei Offline-Systemen getrennte Angaben zur Kartierungsleistung und zur Umsetzungsgenauigkeit der Applikationskarte.

**Einschränkungen**

- Besondere Bedingungen oder Abweichungen vom vorgesehenen Prüfablauf.
- Relevante Einschränkungen der Übertragbarkeit der Ergebnisse entsprechend Kapitel 4.7.

Eine standardisierte Vorlage für den Testbericht ist in Anhang C vorgesehen.

## 9. Zusammenfassung und Ausblick

Der vorliegende Prüfrahmen beschreibt eine standardisierte und unter realen Feldbedingungen durchführbare Methodik zur Bewertung KI-basierter Spot-Spraying-Systeme. Im Mittelpunkt stehen die drei aufeinander aufbauenden Bereiche Erkennung, Applikation und Wirkung. Als methodische Grundlage dient das im Rahmen von ISO/TC 23/SC 6/AHG 2 diskutierte ASABE-Testprotokoll. Der Prüfrahmen erweitert dieses um die Bewertung der Erkennungsleistung und der tatsächlichen Bekämpfungswirkung.

Die Erkennungsleistung wird über eine vom geprüften System unabhängige Ground-Truth-Kartierung des natürlichen Beikrautbestands bewertet. Dadurch kann auch bei Online-Systemen eine Erkennungsprüfung durchgeführt werden, wenn keine annotierten Rohbilddaten des Systems exportiert werden können. Die Treffsicherheit wird über den Recall an kartierten Beikrautpositionen und die Fehlapplikationsrate an beikrautfreien Kontrollpositionen beschrieben, jeweils mit Konfidenzintervall. Eine Precision wird wegen ihrer Abhängigkeit von der Beikrautdichte nur als Szenariorechnung angegeben. Durch die Applikation eines zugelassenen Wirkstoffs unter realen Bestandsbedingungen kann außerdem die tatsächliche Bekämpfungswirkung erfasst werden. Die Wirkung muss damit nicht ausschließlich aus technischen Applikationsdaten oder publizierten Dosis-Wirkungs-Beziehungen abgeleitet werden.

Die Unterscheidung zwischen Online- und Offline-Systemen berücksichtigt die unterschiedlichen technischen Abläufe und ermöglicht eine getrennte Betrachtung der jeweiligen Fehlerquellen. Bei Offline-Systemen können Kartierungsleistung und Umsetzung der Applikationskarte unabhängig voneinander bewertet werden. Bei Online-Systemen ist eine solche Trennung möglich, sofern ein geeignetes As-Applied-Log vorliegt. Fehlt dieses, werden Erkennungs- und Applikationsfehler als kombinierte Kenngröße ausgewiesen.

Für das Projekt FarmerSpaceAI bildet der Prüfrahmen die methodische Grundlage für die im Jahr 2027 geplanten Freilandversuche. Darüber hinaus kann er von externen Prüfstellen und Herstellern für vergleichbare Prüfungen genutzt und weiterentwickelt werden. Eine abschließende Bewertung der Systeme ist mit dem vorliegenden Prüfrahmen jedoch nicht möglich. Aussagen zur Ertragswirkung, Wirtschaftlichkeit, ökologischen Wirkung sowie zur Leistungsfähigkeit unter unterschiedlichen Standort- und Jahresbedingungen erfordern weiterführende Untersuchungen. Gleiches gilt für die langfristigen Auswirkungen wiederholter Spot-Applikationen auf Unkrautpopulationen und Herbizidresistenzen. Die Übertragung des Prüfrahmens auf mechanische und physikalische Verfahren zur Beikrautregulierung stellt einen weiteren Ansatzpunkt für die zukünftige Projektarbeit dar; dabei wäre insbesondere der Kulturverlust als zusätzliche Kenngröße zu berücksichtigen (Kapitel 1 und 7.4).

---

## Anhang

### A. Bezug zum ASABE/ISO-Protokoll

Der vorliegende Prüfrahmen basiert auf dem im Rahmen von ISO/TC 23/SC 6/AHG 2 (“Definitions and test methods for new precision application technologies”) diskutierten

Testprotokoll, das im April 2025 in einem gemeinsamen Workshop mit ASABE vorgestellt wurde (Hloben, John Deere, für AHG 2). Ergänzend wird der parallel entwickelte Standard **ASABE X665** ("Crop Protection Equipment – Test methods for the evaluation of targeted spray application systems") berücksichtigt.

### B. Literaturverzeichnis *(zu vervollständigen)*

*[Platzhalter – die vollständige und einheitlich formatierte Literaturliste wird in der finalen Version ergänzt. Im Text werden bislang folgende Quellen referenziert:]*


- Azghadi et al. (2024/2025).
- Espinel et al. (2024).
- Gerhards et al. (2020).
- Hloben (ASABE/John Deere), ISO/TC 23/SC 6/AHG 2 Workshop (2025).
- JKI – Julius Kühn-Institut (2013). Richtlinie für die Prüfung von Pflanzenschutzgeräten 2-1.0: Zusätzliche Anforderungen an Pflanzenschutzgeräte im Geräteanerkennungsverfahren. Braunschweig.
- PflSchGerätV (2013). Verordnung über Pflanzenschutzgeräte (Pflanzenschutz-Geräteverordnung).
- Leise et al. (2025).
- Mayer et al. (2024).
- Padilla et al. (2020).
- Quan, C., Zhang, J., Feng, X., Zhang, H., Yang, M., Zhu, Z., He, X. & Wang, C. (2025). Spray Deposition and Weed Control Efficacy of a Real-Time Variable-Rate Boom Sprayer Applying Herbicide at Reduced Doses in Summer Maize Fields. *Agronomy*, 15(8), 1953.
- Droplet Deposition and Efficacy of Real-Time Variable-Rate Application of Herbicides at Reduced Dose in Winter Wheat Fields. *Agronomy*, 14(1), 211.
- Van Zyl, J. G., Fourie, P. H. & Schutte, G. C. (2013). Spray deposition assessment and benchmarks for control of Alternaria brown spot on mandarin leaves with copper oxychloride. *Crop Protection*, 46, 80–87.
- Zhou, M., Jiang, H., Bing, Z., Su, H. & Knoll, A. (2021). Design and evaluation of the target spray platform. *International Journal of Advanced Robotic Systems*, 18.
- ASABE Standards Committee Updates (2025).
- ISO 24253 – Testverfahren zur Bestimmung der Spritzbelagsverteilung bei UAV-Applikationssystemen; einschließlich Vergleichsuntersuchungen mit etablierten japanischen Prüfverfahren.
- Deep Learning for Precision Agriculture: Post-Spraying Evaluation and Deposition Estimation. Übersichtsarbeit, arXiv (2024).


### C. Dokumentationsvorlage *(zu ergänzen)*

*[Platzhalter – die Dokumentationsvorlage wird als separates, ausfüllbares Messprotokoll ergänzt. Sie soll mindestens Angaben zur Ground-Truth-Kartierung, zum eingesetzten Herbizid und zur Aufwandmenge, zum Status des As-Applied-Logs, zu den Collector-*

*Ergebnissen an Beikraut- und Kontrollpositionen für jede Wiederholung, zu den Kontrollfenstern sowie zur Wirkungsbonitur enthalten.]*

## D. Begleitende Normungsarbeiten und methodisch vergleichbare Forschung

### Normungsarbeiten

Das in Kapitel 2.2 beschriebene Testprotokoll ist Teil einer weiterführenden Normungsarbeit zur Bewertung neuer Präzisionsapplikationstechnologien. Innerhalb der ASABE wird parallel der Standard **X665** ("Crop Protection Equipment – Test methods for the evaluation of targeted spray application systems") entwickelt. Dieser soll Prüfverfahren und Leistungskennwerte für zielgerichtete Applikationstechniken bei Feldspritzen in Reihenkulturen definieren und unter anderem als Grundlage für die Bewertung entsprechender Technologien durch regulatorische Stellen dienen (ASABE Standards Committee Updates 2025).

Für luftgestützte Applikationstechnik existieren mit ISO 24253 bereits vergleichbare Prüfansätze zur Erfassung von Spritzbelag und Belagsverteilung bei UAV-Sprühsystemen. Vergleichsuntersuchungen zwischen etablierten japanischen Prüfverfahren und dem ISO-Verfahren zeigen, dass bereits die Wahl des Auffangmediums – beispielsweise ebene oder würfelförmige Collector – die gemessenen Belagswerte beeinflussen kann. Dieser methodische Aspekt ist auch für die Wahl und Standardisierung der Collector-Methode im vorliegenden Prüfrahmen relevant (Kapitel 4.3).

Weitere Normungsarbeiten innerhalb von ISO/TC 23/SC 6 befassen sich unter anderem mit der Sicherheits- und Umweltprüfung von Pflanzenschutzgeräten (z. B. ISO/DIS 16122, ISO/CD 22368-1 und ISO 23117-1/2). Einen spezifischen Schwerpunkt auf KI-basierte Zielerkennung und Spot-Spraying bilden diese Normen jedoch nicht.

### Methodisch vergleichbare Forschungsarbeiten

Neben den in Kapitel 2.1 beschriebenen Arbeiten zur Bewertung von Erkennungsalgorithmen existieren verschiedene Studien, die die technische Applikationsleistung anhand definierter Ziele und gemessener Spritzbeläge untersuchen und damit methodische Überschneidungen mit dem vorliegenden Prüfrahmen aufweisen.

Zhou et al. (2021) untersuchten eine Plattform für zielgerichtete Applikation sowohl anhand künstlicher Pflanzenattrappen als auch unter Freilandbedingungen in einer Kultur (Chinese Cabbage). Die Treffgenauigkeit wurde als räumliche Abweichung zwischen Zielposition und Spritzbelag angegeben. Unter Laborbedingungen wurden Abweichungen von 13–17 mm und im Freiland von 0–28 mm ermittelt. Der Ansatz entspricht damit grundsätzlich der im vorliegenden Prüfrahmen vorgesehenen Erfassung der Trefferlage (Kapitel 3.2 und 4.3).

Quan et al. (2025) untersuchten für Herbizidanwendungen in Sommermais den Spritzbelag, die Bekämpfungswirkung und den Ertrag bei unterschiedlichen Sprühvolumina, Düsentypen und reduzierten Herbiziddosen. Das Sprühvolumen beeinflusste dabei sowohl den Spritzbelag als auch die Bekämpfungswirkung signifikant. Die Belagsmessung erfolgte mithilfe von wassersensitivem Papier und anschließender Bildanalyse. Eine methodisch vergleichbare Untersuchung wurde für Winterweizen durchgeführt. Beide Arbeiten zeigen,

dass sich der Zusammenhang zwischen Spritzbelag und biologischer Wirkung auch für Herbizidanwendungen untersuchen lässt.

Einen weiteren methodischen Bezug liefert die Arbeit von Van Zyl, Fourie und Schutte (2013). Dort wurde einfaches Papier mit einem Fluoreszenztracer als Collector eingesetzt. Mithilfe einer Bildanalyse wurde der Zusammenhang zwischen Wirkstoffkonzentration, prozentualer Belagsabdeckung und Bekämpfungswirkung untersucht und es wurden Belagswerte als Referenz für eine ausreichende Wirkung abgeleitet. Der Ansatz zeigt, dass auch einfache Collector-Systeme für die quantitative Bewertung von Spritzbelägen geeignet sein können. Die Übertragbarkeit dieses Prinzips auf weitere Wirkstoff-Wirt-Systeme wurde in nachfolgenden Untersuchungen ebenfalls aufgegriffen.

Eine aktuelle Übersichtsarbeit zur automatisierten Nachbewertung von Präzisionsspritzsystemen beschreibt die Erfassung von Spritzbelägen auf Auffangpapier, teilweise mit wassersensitiver Beschichtung, als eine häufig eingesetzte Methode zur Bewertung der tatsächlichen Applikation. Andere Ansätze, beispielsweise die manuelle Zählung oder eine reine Volumenmessung, werden ebenfalls eingesetzt, sind jedoch hinsichtlich Messgröße und Auswertungsmethode weniger einheitlich.

Die genannten Arbeiten stützen damit die methodische Grundlage des vorliegenden Prüfrahmens. Dies betrifft insbesondere die Verwendung einer unabhängigen Referenz für die Zielposition, die Erfassung des tatsächlichen Spritzbelags als Maß für die physische Applikation sowie die ergänzende Bewertung der biologischen Wirkung. Gleichzeitig unterscheiden sich die bisherigen Studien hinsichtlich Zielobjekten, Collectors, Auswertungsmethoden und Versuchsbedingungen deutlich voneinander. Gerade diese Unterschiede verdeutlichen den Bedarf an einem einheitlichen und reproduzierbaren Prüfrahmen für unterschiedliche Spot-Spraying-Systeme.

**Weitere Informationen zum Projekt:**
**https://www.farmerspace.uni-goettingen.de**

FARMERSPACE_EF

FARMERSPACE AI
DIGITALER PFLANZENSCHUTZ

Förderhinweis:

Gefördert durch:

Bundesministerium
für Landwirtschaft, Ernährung
und Heimat

Projektträger

Bundesanstalt für
Landwirtschaft und Ernährung

aufgrund eines Beschlusses
des Deutschen Bundestages