

**Statistische Begutachtung des Antikollisionssystems „AVES-Wind Onshore“
für den Seeadler vor dem Hintergrund des LfU-AKS Prüfrahmens und der
„KNE-Checkliste“**

Dipl.-Math. Dipl.-Biol. Dr. Moritz Mercker

Bionum GmbH – Büro für Biostatistik

Bodemannweg 87

21129 Hamburg

www.bionum.de

mmercker@bionum.de



Hamburg, 11.04.2025

Inhaltsverzeichnis

1.	ÜBERBLICK UND ZUSAMMENFASSUNG DER ERGEBNISSE	3
2.	DEFINITION UND BESTIMMUNG DES REAKTIONS- UND ERFASSUNGSBEREICHES .	6
3.	DATEN & FELD- UND STATISTISCHE METHODEN.....	8
4.	ERGEBNISSE.....	14
5.	LITERATUR	19

1. ÜBERBLICK UND ZUSAMMENFASSUNG DER ERGEBNISSE

Mit dem AVES-Wind Onshore System hat die ProTecBird ein innovatives KI-basiertes Antikollisions- und Monitoringsystem entwickelt. Im Rahmen einer Pilotstudie im Jahr 2023 wurde die Leistung dieses Systems bzgl. der automatischen Vogelerfassung und bedarfsgerechten Abschaltung an zwei Windenergieanlagen vom Typ Nordex (N149) untersucht. Insbesondere wurde die Leistungsfähigkeit des Systems vor dem Hintergrund des jüngst erarbeiteten LfU-AKS-Prüfrahmens (2024) sowie in Anlehnung an die KNE-Checkliste (Anforderungen an Antikollisionssysteme zum Schutz von Vögeln an Windenergieanlagen – Checkliste für eine qualifizierte Entscheidung über die Anwendbarkeit von Antikollisionssystemen, 2021) für den Rotmilan geprüft (vgl. Gross et al., 2024; Mercker, 2023). Es wurde gezeigt, dass alle Kriterien der KNE-Checkliste und des LfU-AKS-Prüfrahmens für diese Art erfüllt sind.

In dem vorliegenden Dokument wird eine Ergänzungsprüfung für den Seeadler durchgeführt. Da es sich um eine Ergänzungsprüfung handelt, werden insbesondere die Raten (Erfassungsrates, Erkennungsrate, Gesamtrate) abgeschätzt und präsentiert, nicht aber (weitestgehend artunabhängige) Punkte/Inhalte, die bereits im Rahmen der Erstprüfung analysiert wurden, wie die räumliche Abdeckung, die Evaluation an verschiedenen Standorten oder die zeitgerechte Abschaltung.

Für die Bestimmung der Erfassungsrates wurden analog zu Mercker (2023) im Vorfeld vielfältige Daten in einem Windpark an zwei Windenergieanlagen (Typ Nordex N149) erhoben, die eine Beurteilung der Leistungsfähigkeit des Systems hinsichtlich der Erfassungsrates möglich machen. Bei den erhobenen Daten handelte es sich u. a. um

- örtlich-zeitliche Laser-Rangefinder (LRF)-Daten der anfliegenden Zielart (Seeadler);
- örtlich-zeitliche Kamera-basierte Daten der anfliegenden Zielart und weiteren Vogelarten aus dem AKS; und
- örtlich-zeitliche monokulare Entfernungsmessungen durch das AKS.

Die Ergebnisse der Analysen dieser Daten **ergeben eine Erfassungsrates von 95 % [92 % – 97 %]** (Erfassungsbereich: 300-1000 m Distanz zum AKS) basierend auf der im Prüfrahmen vorgegebenen GLMM-Methode.

Hinsichtlich der Erkennungsrate (auch Klassifikationsrate genannt) wird diese basierend auf AKS-Kameradaten abgeschätzt, um damit analog zu Reichenbach et al. (2024) eine beschleunigte Methode zur Integration weiterer Arten zu präsentieren und anzuwenden. Hiermit folgen wir somit den Vorschlägen des LfU-AKS-Prüfrahmens in dem formuliert steht: *„Sollte die Ermittlung der Erkennungsrate basierend auf AKS- statt LRF-Daten erfolgen, muss eine Anpassung der Methode erfolgen, die eine statistisch korrekte Ermittlung der Gesamt- / Schutzrate und deren Konfidenzintervalle ermöglicht“*.

Insbesondere vergleichen wir in dem vorliegende Bericht zwei verschiedene Analysemethoden der Erkennungsrate miteinander (die Methode von Reichenbach et al., 2024, mit einer GLMM-Methode ähnlich zu der im Prüfrahmen vorgegebenen Methode) und zeigen, dass beide Ansätze in hohem Maße vergleichbare Ergebnisse liefern.

Insbesondere liefern die beiden Ansätze **Erkennungsraten von 87 % [82 % – 91 %] bzw. 86 % [82 % – 89 %]** (in dem Bereich bis etwa 1000 m Distanz zum AKS) und damit in hohem Maße vergleichbare Ergebnisse.

Bzgl. der KNE-Checkliste ist gefordert, dass mindestens für den Reaktionsbereich (idealerweise aber für ein mindestens 500 m-Umfeld um das AKS) gezeigt wird, dass die mittlere Erfassungsrate bei mindestens 75 % liegt. In Einzelfällen, d.h., unter naturschutzfachlich besonders kritischen Bedingungen, können hier auch bis zu 90 % gefordert werden. Hinsichtlich der Erkennungsrate ist bzgl. der KNE-Checkliste gefordert, dass sie die 75 % deutlich überschreitet, sie sollte idealerweise bei 90 % liegen.

Einfache Mittelwertberechnungen der Raten (basierend auf den LRF-Daten für die Erfassungsrate und basierend auf den AKS-Kameradaten bzgl. der Erkennungsrate) liefern 96% (Erfassungsrate) bzw. 94 % (Erkennungsrate). **Die Kriterien der KNE-Checkliste sind für das AVES-System und den Seeadler somit erfüllt.**

Bzgl. des LfU-AKS-Prüfrahmens ist gefordert, dass die Gesamtrate (ausgewertet mit adäquaten GLMM-Methoden inklusive der Berücksichtigung von zeitlicher Autokorrelation) im Erfassungsbereich ein 95%-Konfidenzintervall zeigt, dessen untere Grenze („untere Grenze des Konfidenzintervalls“ = UGK) bei mindestens 70 % liegt. Für das Erfüllen der Vorgaben müssen somit die Konfidenzintervalle der Gesamtrate (Kombination von Erfassungs- und Erkennungsrate) abgeschätzt werden. Da sich die Ermittlung beider Raten im vorliegenden Fall aus unterschiedlichen Datenquellen speist (und sie in dem „regulären“ Fall, dass beide auf identischen LRF-Tracks beruhen, vermutlich komplex und trackweise miteinander korreliert sind), ist diese Abschätzung nicht trivial.

Im vorliegenden Bericht werden hierfür zwei verschiedene Betrachtungen durchgeführt:

- (1) Für den Rotmilan konnte für das AVES-System bereits gezeigt werden, dass eine Erfassungsrate mit einer UGK von 92 % und eine Erkennungsrate mit einer UGK von 79 % zu einer Gesamtrate mit einer UGK von 75 % führen (vgl. 380-580 m Abstand zum AKS in Gross et al., 2024; Mercker, 2023) – bei adäquater Berücksichtigung der o. g. Korrelation zwischen beiden Raten. Da im vorliegenden Falle des Seeadlers die UGKs beider Raten auf gleicher Höhe (Erfassungsrate) oder sogar mehrere Prozentpunkte höher (Erkennungsrate) liegen, ist in hohem Maße davon auszugehen, dass die UGK der Gesamtrate sogar höher als 75 % liegt und damit dem Mindestkriterium des AKS-Prüfrahmens (> 70 %) entspricht – unter Berücksichtigung der komplexen Korrelationen zwischen beiden Raten.
- (2) Um der Frage nach der Kombination beider Raten auch direkt rechnerisch nachzugehen, wurden zwei Varianten der Delta-Methode angewendet, um aus den Mittelwerten und Konfidenzintervallen der Einzelraten eine Gesamtraten und deren Konfidenzintervalle abzuleiten. Diese beiden Delta-Methode-Varianten wurden wiederum separat durchgeführt für die Ergebnisse der zwei möglichen o. g. Analysemethoden hinsichtlich der Erkennungsrate. Die insgesamt 4 Ergebnisse lieferten in hohem Maße ähnliche Mittelwerte der Gesamtrate von 82-83 % und UGKs von 77-78 %. Es sei jedoch einschränkend angemerkt, dass die Delta-Methode impliziert, dass beide Raten unkorreliert sind (was vermutlich nicht der Fall ist) und

die Ergebnisse daher u. U. verzerrt sein können. Da die so geschätzten UGKs jedoch einen deutlichen Abstand zu den geforderten 70 % zeigen, ist in hohem Maße davon auszugehen, dass auch der unverzerrte Wert noch über 70 % liegt.

Die Ergebnisse zeigen somit zusammengenommen, dass die Mindestkriterien des LfU-AKS-Prüfrahmens hinsichtlich der AKS-Gesamtraten-Validierung für das AVES-System und den Seeadler erfüllt sind.

Es sei abschließend angemerkt, dass die unter (1) und (2) präsentierten Methoden zwar für die vorliegende Datensituation beim Seeadler u. E. robust und geeignet sind, diese Vorgehensweise aber nicht ohne weiteres auf andere Arten bzw. Datensituationen übertragbar ist. Zum einen, da im vorliegenden Fall die Erfassungsrate weiterhin basierend auf LRF-Tracks bestimmt wurde (was einen erheblichen Feldaufwand bedeutet, den es aus Gründen der Vereinfachung und Beschleunigung für Ergänzungsprüfungen zukünftig zu vermeiden gälte), zum anderen, da im vorliegenden Fall des Seeadlers die UGKs beider Raten auf gleicher Höhe (Erfassungsrate) oder sogar mehrere Prozentpunkte höher (Erkennungsrate) der des intensiv untersuchten Rotmilans liegen – was für andere Arten nicht gegeben sein muss, sodass dort nicht derselben Argumentation gefolgt werden kann. Auch wird im vorliegenden Fall bei den Ergebnissen der Delta-Analysen ein deutlicher Abstand zu den geforderten 70 % der UGK gemessen, sodass die vereinfachte Annahme der Unabhängigkeit beider Raten (und daraus resultierende mögliche Verzerrungen von einigen Prozentpunkten) als unbedenklich eingestuft werden können. Auch hier kann das Ergebnis für andere Arten „knapper“ ausfallen, sodass auch hier u. U. nicht derselben Argumentation gefolgt werden kann.

Um zukünftig eine beschleunigte und unter vielfältigen Datensituationen robuste Analyse der AKS-Raten bzgl. weiterer Arten zu gewährleisten, müsste eine Methodik entwickelt werden, die (1) klärt, unter welchen Umständen es aus naturschutzfachlicher Sicht angemessen ist, AKS-Daten basierte ermittelte Erkennungsraten einer Art mit den LRF-Daten basierte ermittelten Erfassungsraten einer anderen Art zu verknüpfen, und (2) eine Verknüpfungsmethode zu entwickeln, die den vermutlich komplexen Korrelationsstrukturen zwischen Erfassungs- und Erkennungsrate Rechnung trägt.

2. DEFINITION UND BESTIMMUNG DES REAKTIONS- UND ERFASSUNGSBEREICHES

Der Reaktionsbereich definiert den projektspezifischen zylindrischen 3D Luftraum um den WEA-Rotormittelpunkt, an dessen Rändern ein Vogel spätestens erfasst und klassifiziert worden sein muss, um eine rechtzeitige Abschaltung der WEA zu gewährleisten. Der Radius dieses Zylinders ergibt sich u.a. aus WEA-spezifischen Maßen sowie der horizontalen Fluggeschwindigkeit. Im Rahmen des LfU-Prüfrahmens wird dieser Bereich noch spezifischer definiert und dessen Höhe hängt zusätzlich von den vertikalen Fluggeschwindigkeiten ab (vgl. LfU-AKS-Prüfrahmen, an den die folgenden Passagen in hohem Maße angelehnt sind).

Der Erfassungsbereich definiert hingegen den 3D Luftraum um das AKS, welcher im Rahmen der AKS-Validierung (empirische Ermittlung der Raten) verwendet wird und für den eine ausreichend hohe und valide bestimmte Erfassungs-, Erkennungs- (KNE-Checkliste) bzw. Gesamtrate (LfU-AKS-Prüfrahmen) gefordert ist. Im Genehmigungsverfahren wird dann überprüft, ob der projektspezifische (zylindrische) Reaktionsbereich durch den (oder einen der untersuchten) Erfassungsbereich(e) geschützt/abgedeckt ist.

Dies ist nach LfU-AKS-Prüfrahmen z. B. dann der Fall, wenn entweder der Reaktionsbereich im Erfassungsbereich enthalten ist, oder wenn der Erfassungsbereich den Reaktionsbereich ringförmig so umläuft, dass damit der relevante Luftraum um den Reaktionsbereich überwacht und somit in den Reaktionsbereich eintretende Vögel zuverlässig erfasst und erkannt werden. Diese ringförmige Betrachtung ist für das AVES-Wind Onshore System relevant, da es auf das Erfassen und Erkennen über größere Entfernungen spezialisiert ist und primär das Umfeld des Reaktionsbereiches überwacht.

Im Falle der AKS-Kameradaten hinsichtlich der Analyse der Erkennungsrate war eine präzise Verortung der Tracks im direkten Umfeld des AKS nicht gegeben; es wurden daher lediglich diejenigen Tracks ausgewählt, die einen Abstand von < 1000 m zum AKS zeigten. Diese Betrachtung eines kreisförmigen statt eines ringförmigen Erfassungsbereichs hinsichtlich der Erkennungsrate stellt daher eine aus naturschutzfachlicher Sicht konservative Vorgehensweise in Bezug auf das AVES-System dar, da dieses – wie oben erwähnt – nicht auf den Nahbereich spezialisiert ist. Im Falle der Analyse der Erfassungsrate (in deren Kontext eine genaue Lokalisation der LRF-Tracks auch im Nahbereich möglich war) wurde im vorliegenden Gutachten der ringförmige Bereich mit 300-1000 m Distanz zum AKS betrachtet. In beiden Fällen wird also bis zu einer Distanz von 1000 m analysiert, um ein möglichst großes Spektrum an zukünftigen/realen Reaktionsbereichen im Kontext von Genehmigungsverfahren abdecken zu können. Darüber hinaus werden die im vorliegenden Gutachten untersuchten Erfassungsbereiche weder nach oben noch nach unten begrenzt, sondern alle LRF- bzw. AKS-Kamera-Tracks analysiert, die sich horizontal betrachtet innerhalb der Erfassungsbereiche. Auf diesem Wege werden selbst vom Boden oder weit oben in den Reaktionsbereich einfliegende Vögel in den ermittelten Raten reflektiert; es handelt sich somit auch hier um eine artenschutzfachlich konservative Vorgehensweise.

Da es sich im vorliegenden Projekt um ein Mehrkamera-AKS-System handelt, welches die Anzahl der Kameras projektspezifisch anpasst, kamen in dem vorliegenden Feldexperimenten zu den Raten/der Validierung insgesamt vier Kameras zum Einsatz. Der Erfassungsbereich wurde entsprechend angepasst (ein „Tortenstück" pro Kamera).

3. DATEN & FELD- UND STATISTISCHE METHODEN

3.1 Datengrundlage LRF-Daten

Die LRF-Daten wurden an 10 verschiedenen Tagen zwischen dem 19.09.2024 und dem 07-10.-2024 an zwei verschiedenen Standorten (jedoch nur etwa 70 m voneinander entfernt) erhoben und beinhalten insgesamt 1.860 eingemessene LRF-Punkte aus 71 verschiedenen Seeadler-Flugtracks in Distanzen zwischen 70 und 2.189 Metern. Diese wurden auf den Bereich zwischen 300 und 1000 m Abstand zum AKS eingeschränkt. Zudem wurden (gemäß LfU-AKS-Prüfrahmen) nur diejenigen LRF-Tracks verwendet, die mindestens 4 verschiedene Punkte im analysierten Erfassungsbereich umfassten, da sie im Kontext der Regressionsmethoden (s. u.) bei oft sehr geringem Anteil an den Gesamtdaten zu einem unverhältnismäßig starken Anstieg der Varianz bzw. zu verzerrten Ergebnissen führen können. Zudem handelt es sich bei LRF- / GPS-Tracks mit nur wenigen Punkten i. d. R. um nicht kollisionsgefährdete Flüge, bei denen der Vogel nur kurzzeitig im Erfassungsbereich auftauchte. Es verblieben in Summe 55 verschiedene LRF-Tracks des Seeadlers im analysierten Erfassungsbereich.

3.2 Methodik Feldversuche und Zuordnung AKS vs. LRF

Das Konzept zur Untersuchung der Erfassungsrate beruht auf dem Vergleich von Vögeln, die im Feld (örtlich und zeitlich verortet) durch Ornithologen und Ornithologinnen erfasst wurden mit dem, was durch das AVES-Wind Onshore System detektiert wurde. Es wurden von den beobachtenden Personen Laser-Rangefinder (LRF)-Geräte eingesetzt, die eine relativ präzise Verortung der Vögel vornehmen können (Ransom & Pinchak, 2003). Aus diesem Grund wird im Folgenden (vereinfacht) davon ausgegangen, dass die LRF-Daten prinzipiell mit keinem nennenswerten örtlichen und/oder zeitlichen Fehler behaftet sind. Der örtliche Fehler des AVES-AKS wurde in Mercker (2023) bereits quantifiziert.

Eine *a priori* Vorselektion der LRF-Punkte erfolgte ausschließlich hinsichtlich der Fragestellung, ob sie sich in dem vor der Kamera liegenden „Tortenstück“ (s. o.) befanden oder nicht. Diese „Tortenstücke“ beinhalten mehr Luftraum, als die von der Seite betrachtet keilförmigen Sichtkegel der Kameras erfassen können. So ergibt sich automatisch die Berücksichtigung der technisch bedingten „Abdeckungsrate“.

Hierfür sowie für die Zuordnung LRF-Punkt vs. AKS-Signal („matching“) wurden die folgenden Betrachtungen/Arbeitsschritte durchgeführt:

1. Mithilfe der LRF-Koordinaten, der Kamera-Positionen, und der kalibrierten Kamera-Winkel-Skala wurden pro Punkt theoretische Winkel (Pan, Tilt) erzeugt, unter denen der Vogel zu sehen ist;
2. Der Erfassungsbereich ist ein Kreisring um die Kamera (die an der WEA angebracht ist) mit festem innerem und äußerem Radius sowie fester Höhe. Horizontal erstreckt sich der Erfassungsbereich über einen Sehwinkel von 30° pro Kamera („Tortenstück“);

3. Die Logfiles des AVES-Wind Onshore Systems lieferten die tatsächlichen Winkel zu jedem Zeitpunkt. Außerdem ließ sich ableiten, was die Standard-Ausrichtung ist, welche den Ort des festen Erfassungsbereichs definiert. Die Standard-Ausrichtung wurde algorithmisch aus den Logfiles bestimmt, um (fehleranfällige) hartprogrammierte Winkel zu vermeiden. Da die Kameras selbst nach dem Verfolgen eines Vogels nur mit begrenzter Toleranz zu einem Standard-Wert zurückkehren können, kommt es vor, dass der Horizontal-Winkel der Standard-Ausrichtung im Verlaufe einer Session um bis zu 1° schwankt;
4. Erfassbarkeit jeden Punktes prüfen:
 - a. Alle Punkte im oben definierten Bereich galten als erfassbar (im Folgenden „erfassbar“ synonym zu „innerhalb des Tortenstückes“ – s. o. zur Mitberücksichtigung der technisch bedingten Abdeckungsrate an dieser Stelle);
 - b. Die Kameras können zur Verfolgung von Vögeln automatisch mitgeschwenkt werden. Dadurch ergibt sich zusätzlich ein „dynamischer“ Sichtbereich einer Kamera, der durch Zoom auf unter 30° Breite beschränkt sein kann. Ein LRF-Punkt, dessen Winkel innerhalb dieses dynamischen Sichtbereichs lag, galt als erfassbar, wenn er innerhalb des inneren und äußeren Kreisring-Randes lag. LRF-Punkte galten als erfassbar, wenn ein Winkel-Matching erzielt wurde (s. u.). Dies ist nötig, da die Berechnung, ob ein LRF-Punkt innerhalb des dynamischen Sichtbereichs liegt, oft zu Fehlern führt, da das Winkel-Intervall des Sichtbereichs bei maximalem Zoom bei $< 4^\circ$ liegt, was deutlich unterhalb der Genauigkeit der Winkel-Berechnung liegt. Beispiel: ein Vogel wird verfolgt und auf 3° Sehwinkel herangezoomt. Die Berechnung über die Winkel ergibt, dass der Vogel theoretisch bei $\text{Pan}=34^\circ$ zu sehen ist. Der dynamische Kamera-Bereich ist $\text{Pan} = [37^\circ - 1,5^\circ, 37^\circ + 1,5^\circ]$. Damit ist der LRF-Punkt laut Zahlen nicht im Sichtbereich der Kamera. Das System gibt einen getrackten Vogel bei $37,3^\circ$ an. Die Differenz zwischen beiden Winkeln beträgt $3,3^\circ$. Das ist innerhalb der angenommenen Toleranz von 11° . Der Vogel gilt als erfasst und somit auch als erfassbar. Durch die oben beschriebenen Checks gingen durch Mitschwenken erfasste Vögel seitlich und oberhalb des Erfassungsbereichs mit in die Analyse ein. Die Bedingung an die radiale Distanz war jedoch strikt, d. h., LRF-Punkte mit einer radialen Distanz niedriger als der innere Radius oder höher als der äußere Radius des Kreisring galten nie als erfassbar;
5. Erfassung/Winkel-Matching prüfen:

Für den Fall, dass zum LRF-Zeitpunkt ein Track-Eintrag vorlag, wurde geprüft, ob sich die theoretischen Winkel (LRF-Pan, LRF-Tilt) mit den Winkeln des getrackten Objekts decken (Track-Pan, Track-Tilt). Dabei wurde mit einer Toleranz von 11° (Pan) bzw. 9° (Tilt) gerechnet. Sind die oben beschriebenen Bedingungen erfüllt, zählte der LRF-Punkt als erfasst.
6. KI-Klassifizierung bestimmen:

Um jeden LRF-Zeitpunkt wurde ein Zeitintervall gelegt. Dieses ging bis zur Hälfte der Zeitspanne bis zum nächsten LRF-Punkt desselben Tracks. Für LRF-Punkte am Anfang oder Ende einer LRF-Track-ID wurden 3s vor bzw. hinter den LRF-Zeitpunkt angefügt. Innerhalb des so bestimmten Zeitintervalls wurden die Klassifizierungen der KI aus allen Logfile-Einträgen gesammelt und der häufigste Eintrag unter ‚final

class' war die KI-Klasse des LRF-Punktes. Die KI-Klassen sind gegeben durch die Klassen „target“ (hier Seeadler), „non-target“ (alle anderen Vogelarten) oder „bird“ (unbestimmter Vogel). Die „final class“ ist dabei bereits eine speziell für den Anwendungszweck angepasste Mittelung über alle bisherigen einzelnen Klassifizierungen des aktuell erfassten Vogels. Randnotiz: ein Auslesen einer KI-Klassifizierung ist auch ohne zutreffendes Winkel-Matching möglich, z.B. wenn das System einen anderen Vogel verfolgt als vom LRF abgebildet wird.

7. Bestimmung der finalen Variable „Seeadler erfasst und erkannt“ vs. „nicht erfasst oder erkannt“:

Falls ein LRF-Punkt mit den System-Daten ein erfolgreiches Winkel-Matching produzierte und die KI-Klassifizierung mit der tatsächlichen Vogelart übereinstimmte, galt der Vogel als korrekt erfasst und erkannt.

Die korrekte Realisierung der Zuordnung zwischen LRF- und AKS-Kameradaten wurde stichprobenartig anhand der LRF- vs. AKS-Rohdaten überprüft.

3.3 Datengrundlage AKS-Kameradaten

Im Live-Betrieb des AVES Wind Onshore Systems werden grundsätzlich alle Vögel getrackt und zu jeder Detektion automatisch ein Video abgespeichert. Im Normalfall ist auf diesen Videos zu sehen, wie ein zunächst weit entfernter Vogel kontinuierlich verfolgt und schrittweise herangezoomt wird. Das Video endet, sobald AVES die Verfolgung abbricht, d.h. sobald die Live-KI nicht mehr durchgängig einen Vogel erkennt. So wurden im Erfassungszeitraum zahlreiche Videos aufgenommen. Im Nachgang wurde von Ornithologen die tatsächliche Vogelart der Videos bestimmt.

Somit ergibt sich eine Reihe von Videos, auf denen tatsächlich ein Seeadler zu sehen ist, unabhängig von der Klassifikation durch die live-KI. Diese Video-Reihe stellt den Test-Datensatz für die Validierung dar. Mit einer erneuerten KI wurde nun offline die Videos frame-weise analysiert. Dazu werden alle Berechnungen bzw. Parameter so gewählt, dass sie den Live-Betrieb nachbilden. Dies liefert pro Video-frame eine KI-Klasse des Vogels, analog zum Live-Betrieb. Aus der Auflistung der KI-Klasse pro Frame lässt sich dann die Erkennungsrate berechnen. Dabei ist es wichtig zu erwähnen, dass die erneuerte KI nicht mit den Videos aus der Test-Datenreihe trainiert wurde. Das wurde sichergestellt, indem Videos aus dem Erfassungszeitraum aus den KI-Trainingsdaten ausgenommen wurden.

Da die Videos der Test-Datenreihe bzgl. Perspektive, Zoomstufe, usw. genau den Verhältnissen im realen Betrieb entsprechen, liefert die Offline-Verarbeitung durch die erneuerte KI realistisch den Live-Betrieb nach.

Insgesamt lagen 64 verschiedene Kameratracks des Seeadlers vor mit 600.737 Einzelbildern von 11 verschiedenen Erfassungstagen zwischen dem 18.09.2024 und dem 07.10.2024.

3.4 Statistische Ermittlung der Raten

Statistische Modelle

Bei den zu analysierenden Daten handelt es sich um binäre Variablen, die bspw. zwischen „erfasst“ und „nicht erfasst“ oder „erkannt“ und „nicht erkannt“ (1 vs. 0) klassifizieren. Im einfachsten Fall können hier Mittelwerte der entsprechenden Raten ermittelt werden. Diese Auswertungsmethode ist jedoch aus statistischer Sicht nicht adäquat, da mehrere zugrundeliegende Bedingungen verletzt sind (s. u.) und sie daher verzerrte Ergebnisse liefert. Trotzdem wurde sie parallel zu den im Folgenden beschriebenen statistisch adäquaten Methoden angewendet, da bisherige/frühere Analysen (wie bspw. solche, die sich an der KNE-Checkliste orientierten) diese oder ähnliche Methoden angewendet haben und somit eine prinzipielle Vergleichbarkeit geben ist. Diese Methode wird im Folgenden als „**Mittelwert**“ bezeichnet und die Berechnung der 95%-Konfidenzintervalle beruht auf der Anwendung geeigneter linearer „Intercept-only Regressionsmodelle“.

Tatsächlich ist die statistische Situation aber etwas komplexer: Zum einen handelt es sich bei der binären Variable offensichtlich um nicht-normalverteilte Daten, sodass Regressionsmodelle mit entsprechend angepasster Zufallsverteilung eingesetzt werden müssen („Generalisierte Modelle“). Zum anderen werden Zeitreihen untersucht (zeitlich aufeinanderfolgende Messpunkte sind also potenziell stark korreliert), und schlussendlich können die LRF-Punkte verschiedenen Individuen zugeordnet werden, sodass in mehrfacher Hinsicht die Unabhängigkeit der einzelnen Messpunkte verletzt ist und somit ein Problem von „Pseudo-Replikation“ besteht (Hurlbert, 1984), was zu einer (u.U. deutlichen) Unterschätzung von Konfidenzintervallen sowie zu verzerrten ermittelten Raten führen kann. Zwar ist die zeitliche Autokorrelation auch ein Teil der Systemleistung (sie reflektiert die Fähigkeit des Systems erfasste Vögel zu verfolgen), allerdings kann die Nicht-Berücksichtigung der zeitlichen Autokorrelation u. a. dazu führen, dass der Einfluss sehr langer (bzw. „gut verfolgter“) Tracks gegenüber dem wichtigen Kriterium der Erfassung und Erkennung ein sehr hohes Gewicht bekommen. Diese „Dämpfung“ des Einflusses wird über die adäquate Berücksichtigung der Autokorrelation vorgenommen.

Aus diesem Grund wurde eine Analyse-Strategie angewendet, die diese Datensituation adäquat berücksichtigt (vgl. Methodenvorgabe und -motivation im LfU-AKS-Prüfrahmen). Insbesondere erfolgte die Analyse (wie oben motiviert) mit geeigneten Regressionsmethoden, den „logistischen Regressionsmethoden“, die zu den „generalized linear models“ (GLMs) (Bolker et al., 2008; Field et al., 2012; A. Zuur et al., 2007) gehören. Die Zugehörigkeit zu verschiedenen Individuen/Tracks wurde hier berücksichtigt, indem die Track-ID als „random intercept“ im Kontext von „mixed modelling“ verwendet wurde, was zu der Klasse der „generalised linear mixed models“ (GLMMs) (Bolker et al., 2008; Pinheiro & Bates, 2000; A. F. Zuur et al., 2009) führte. Zudem wurde die Stärke der zeitlichen Autokorrelation mittels pACF-Plots analysiert und als geeignete Autoregressionsstruktur integriert (Korner-Nievergelt et al., 2015; A. Zuur et al., 2007; A. F. Zuur et al., 2009). Weitere Details zur Methode und deren Motivation finden sich im LfU-AKS-Prüfrahmen (2024). Diese Methode wird im Folgenden als „**GLMM_AC**“ („generalized linear mixed model with temporal autocorrelation“) bezeichnet und die Berechnung der 95%-

Konfidenzintervalle beruhte auf der Anwendung der „Intercept-only Regressionsmodelle“ dieser Modellklasse. Diese Methode kam bei der Analyse der Erfassungsrate zum Einsatz; aber auch (als Alternative zur „Bootstrap track-mean“-Methode – s. u.) bei der Analyse der Erkennungsrate. Im letztgenannten Fall wurden die zeitlich eng getakteten AKS-Kamera-Tracks so stark ausgedünnt, dass die Residuen des anschließend verwendeten GLMMs eine Autokorrelation der Stärke 1 (und gerade noch nicht 2) zeigten; im Anschluss wurde eine AR-1-Struktur in das Modell integriert.

Im Falle der Analyse der Erkennungsrate wurde zudem als alternativer Ansatz die jüngst publizierte Methode von Reichenbach et al. (2024) etabliert und angewendet. Insbesondere wurde hier von jedem Track separat die Erkennungsrate analysiert, deren Mittelwert gebildet, und 95%-Konfidenzintervalle über geeignete Perzentile von 10.000 Bootstrap-Resamples auf Track-Ebene berechnet. Diese Methode wird im Folgenden mit „**Bootstrap track-mean**“ bezeichnet.

Für alle statistische Analysen wurde die Open-Source-Software **R** verwendet (R Core Team, 2024).

Da es sich bei dem AVES Wind Onshore System um ein Mehrkamera-AKS-System handelt, fand eine Auswertung von Tracks bzw. Raten statt, bei denen das Gesamtsystem als erfassende Einheit betrachtet wurde. Dies ist gestattet, da alle Kameras (durch Datenaustausch) gemeinsam eine eindeutige Reaktion des Gesamtsystems (d. h. ggf. ein Abschaltsignal) erzeugen oder jede Kamera für sich (ohne Priorisierung einzelner Kameras) in der Lage ist, ein Abschaltsignal zu senden.

Modellvalidierung

Die Modellvalidierung erfolgte über verschiedene Modell-Residuenplots, wie u. a. in Korner-Nievergelt et al. (2015) und A. F. Zuur et al. (2009, 2010) beschrieben. Insbesondere wurden die Bedingungen der Unabhängigkeit (wie z. B. die Stärke der zeitlichen Autokorrelation), die Existenz und der Einfluss von Ausreißern, die Normalverteilung des „random intercept“ sowie auf Über- bzw. Unterdispersion getestet/geprüft. Es wurden keine deutlichen Verletzungen dieser Bedingungen detektiert.

Abschätzung der Gesamtrate mittels Delta-Methode

Zur multiplikativen Verknüpfung zweier anteilsbasierter Schätzwerte, in diesem Fall der Erfassungsrate und der Erkennungsrate, wurden zwei Varianten der Delta-Methode angewendet, um daraus eine Gesamtrate sowie ein entsprechendes Konfidenzintervall abzuleiten. In beiden Fällen erfolgte die Berechnung unter der Annahme, dass die beiden Schätzer unabhängig voneinander sind (was in dieser Schärfe vermutlich nicht gegeben ist, vgl. Abschnitt 1).

Die klassische Variante der Delta-Methode basiert auf einer linearen Fehlerfortpflanzung unter Verwendung der ersten Ableitung einer Funktion. Für das Produkt zweier Erwartungswerte ergibt sich daraus, dass die Varianz des Produkts durch die Summe der Varianzen beider Terme gewichtet mit dem jeweils anderen Erwartungswert beschrieben werden kann. Die zugehörigen Standardfehler wurden aus den gegebenen

Konfidenzintervallen abgeleitet, unter der Annahme symmetrischer Intervalle und normalverteilter Fehler. Auf Basis des geschätzten Standardfehlers wurde anschließend ein 95-Prozent-Konfidenzintervall für die Gesamtrate berechnet. Eine Normalverteilung der Fehler kann im vorliegenden Fall aufgrund der binomialen Daten nicht angenommen werden; trotzdem wendeten wir diese Methode an in der Annahme, dass dieser Fehler nicht sehr stark ist und da diese Variante bei der Delta-Methode als besonders gut etabliert und robust bekannt ist.

Die zweite Variante der Delta-Methode orientierte sich an der Verteilungsstruktur von Anteilswerten, indem die Unsicherheiten nicht als normalverteilt, sondern als Beta-verteilte Zufallsvariablen modelliert wurden (Verteilung zwischen 0 und 1). Dazu wurde für jeden der beiden Ratenwerte eine Beta-Verteilung parametrisiert, deren Formparameter auf Basis der geschätzten Raten und der daraus approximierten effektiven Stichprobengröße berechnet wurden. Letztere wurde durch Rückrechnung aus dem Standardfehler unter der Annahme einer binomialen Grundverteilung gewonnen. Im Rahmen einer Monte-Carlo-Simulation wurden jeweils 10.000 Zufallswerte aus den beiden Beta-Verteilungen gezogen und anschließend paarweise multipliziert. Aus der resultierenden Verteilung der Produkte wurde der Mittelwert sowie das 2.5-, 50- und 97.5-Perzentil bestimmt, um Mittelwert und Konfidenzintervall der Gesamtrate abzuleiten. Diese Methode berücksichtigt im Gegensatz zur klassischen Delta-Methode also die asymmetrische Natur von Anteilsverteilungen sowie die durch Modellstruktur (insbesondere die zeitliche Autokorrelation) reduzierte effektive Informationsdichte/Stichprobengröße

Beide Varianten wurden parallel verwendet, um mögliche Verzerrungen durch Verteilungsannahmen sichtbar zu machen und die Robustheit der Gesamtschätzung zu überprüfen.

4. ERGEBNISSE

Erfassungsrate

Die Ergebnisse zur Erfassungsrate sind in Abbildung 1 gezeigt. Mit der GLMM_AC-Methode wird eine Erfassungsrate von 95 % [92 % – 97 %] gemessen, über die einfache Mittelwertbildung eine Rate von 96 % [95 % – 97 %]. Die Vorgaben der KNE-Checkliste (Raten von mindestens 75 % – in Einzelfällen von über 90 %) sind daher in hohem Maße erfüllt.

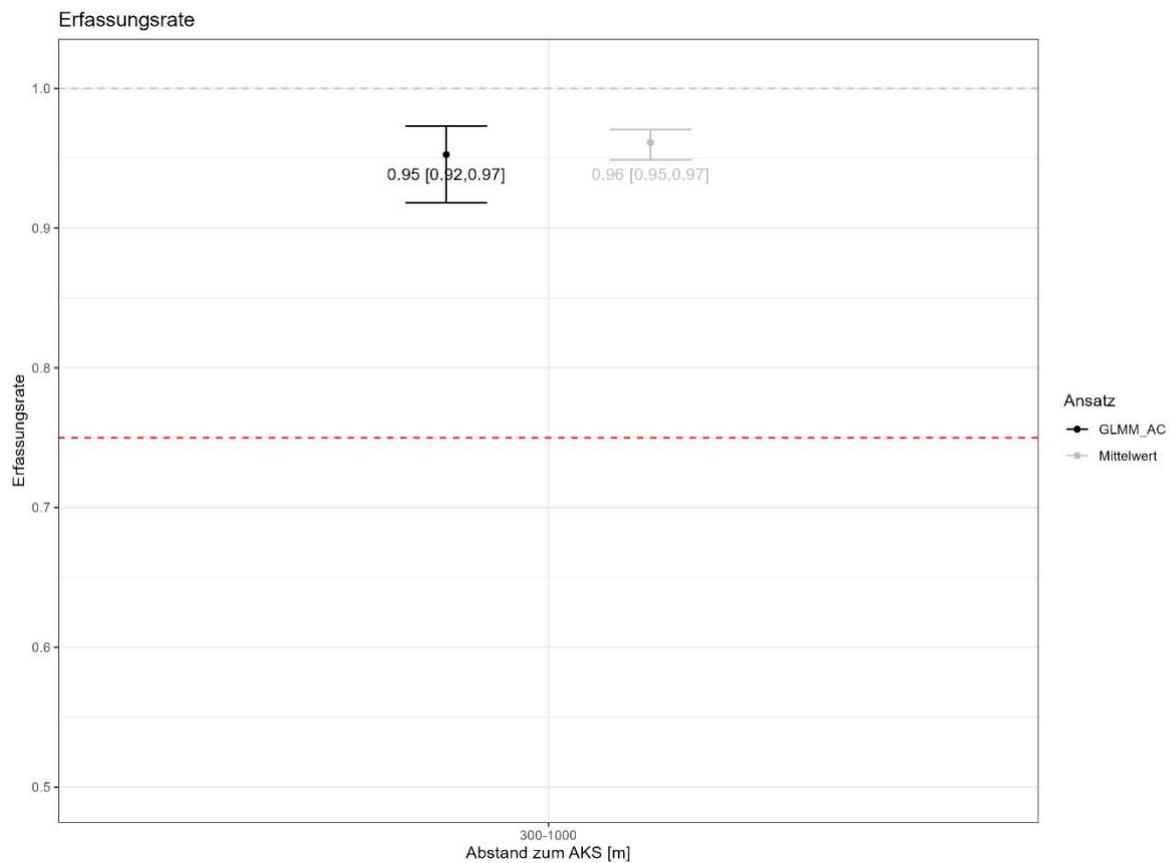


Abbildung 1: Die Ergebnisse zur Erfassungsrate nebst 95%-Konfidenzintervallen basierend auf der GLMM_AC-Methode (schwarze Punkte/Fehlerbalken/Zahlen) und der einfachen Mittelwertbildung (schwarze Punkte/Fehlerbalken/Zahlen). Rote gestrichelte Linie zeigt das Mindestkriterium der KNE-Checkliste bzgl. des Mittelwertes (75 %) an.

Erkennungsrate

Die Ergebnisse zur Erkennungsrate sind in Abbildung 2 gezeigt. Mit der GLMM_AC-Methode wird eine Erkennungsrate von 86 % [82 % – 89 %] gemessen, mit der Bootstrapped track-mean-Methode eine ganz ähnliche Rate von 87 % [82 % – 91 %] und über die einfache Mittelwertbildung wird eine Rate von 94 % [94 % – 95 %]. Bzgl. der letztgenannten Rate zeigt dies exemplarisch, wie bei stark zeitlich autokorrelierten Daten eine Überschätzung der Rate und eine extreme Unterschätzung der Konfidenzintervalle zu beobachten ist.

Die Vorgaben der KNE-Checkliste (Raten von mindestens 75 % – idealerweise bei 90 %) sind daher auch hier erfüllt.

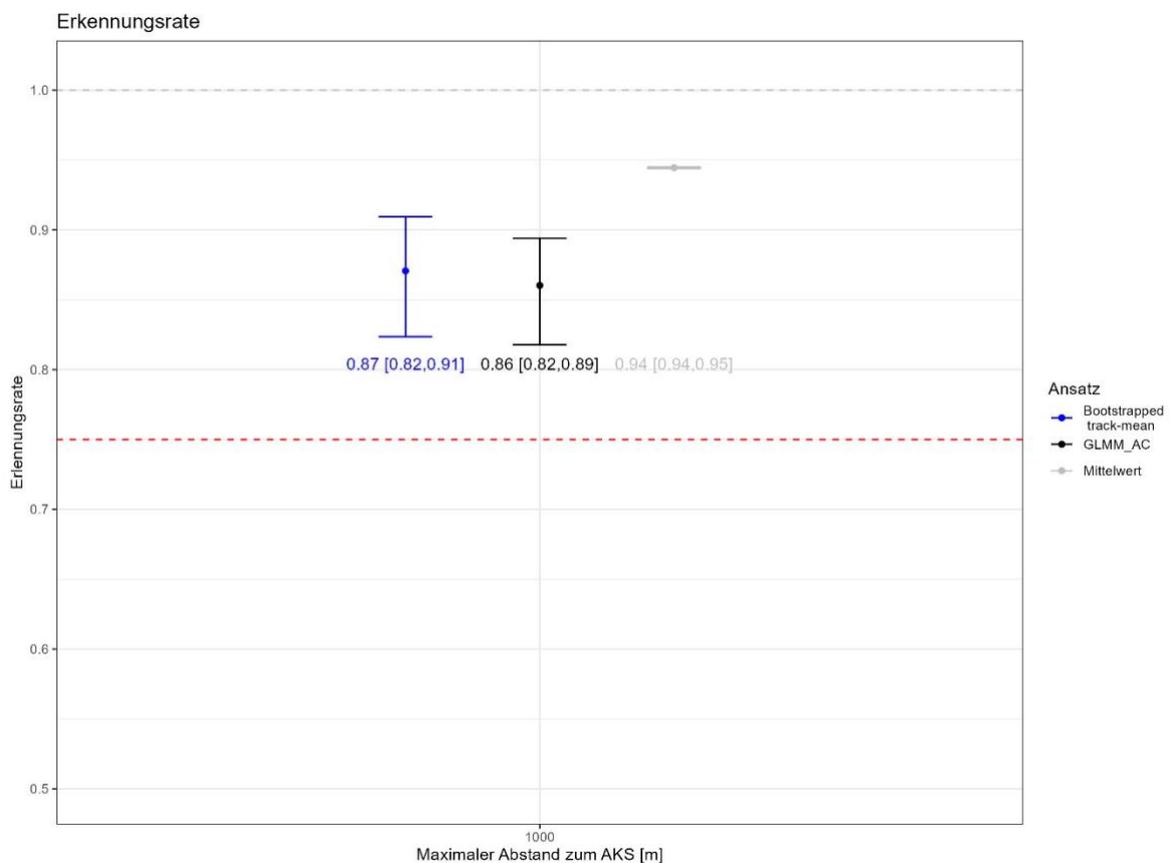


Abbildung 2: Die Ergebnisse zur Erkennungsrate nebst 95%-Konfidenzintervallen basierend auf der GLMM_AC-Methode (schwarze Punkte/Fehlerbalken/Zahlen), der Bootstrapped track-mean-Methode (blaue Punkte/Fehlerbalken/Zahlen) und der einfachen Mittelwertbildung (schwarze Punkte/Fehlerbalken/Zahlen). Rote gestrichelte Linie zeigt das Mindestkriterium der KNE-Checkliste bzgl. des Mittelwertes (75 %) an.

Gesamtrate

Die Ergebnisse zur Gesamtrate (mittels zwei verschiedenen Delta-Methoden) sind in Abbildung 3 und Abbildung 4 gezeigt. Die insgesamt 4 Ergebnisse liefern sehr ähnliche Mittelwerte der Gesamtrate von 82-83 % und untere Grenzen der 95%-Konfidenzintervalle von 77-78 %) – was für eine Robustheit dieser Abschätzungen spricht.

Hinweis zur Interpretation: Die Anwendung der Delta-Methode setzt voraus, dass die beiden Raten voneinander unabhängig sind – eine Annahme, die in der Praxis vermutlich nicht erfüllt ist. Dadurch können die Ergebnisse potenziell verzerrt sein. Da die auf diese Weise geschätzten unteren Grenzen jedoch deutlich oberhalb der geforderten 70 % liegen, ist es sehr wahrscheinlich, dass auch der tatsächliche (unverzerrte) Wert über diesem Schwellenwert liegt.

Zudem wurde für den Rotmilan bereits gezeigt, dass das AVES-System selbst unter Berücksichtigung der oben genannten vermutlich existierenden Korrelation zwischen Erfassungs- und Erkennungsrate eine Gesamtrate mit einer unteren Grenze des Konfidenzintervalls (UGK) von 75 % zeigt – basierend auf einer Erfassungsrate mit einer UGK von 92 % und einer Erkennungsrate mit einer UGK von 79 % (vgl. 380–580 m Distanz zum AKS in Gross et al., 2024; Mercker, 2023). Im vorliegenden Fall des Seeadlers liegen die UGKs beider Teilraten mindestens auf vergleichbarem Niveau (Erfassungsrate) bzw. sogar darüber (Erkennungsrate) – wie in den vorherigen Kapiteln gezeigt. Daher ist mit hoher Wahrscheinlichkeit davon auszugehen, dass auch die UGK der resultierenden Gesamtrate über 75 % liegt, was sehr gut zu den oben berechneten 77-78 % passt.

Basierend auf diesen Ergebnissen gehen wir zusammengenommen davon aus, dass die Kriterien des LfU-AKS-Prüfrahmens für das AVES-System und den Seeadler in höherem Maße erfüllt sind.

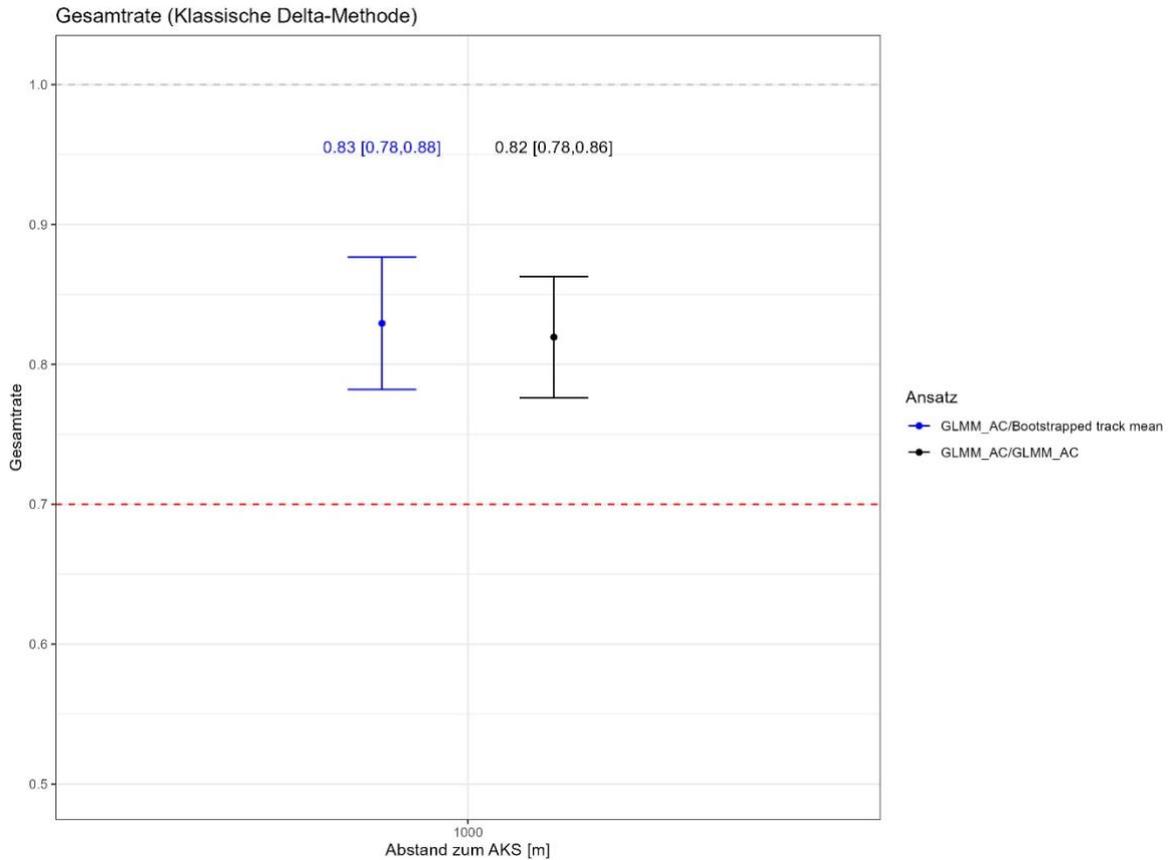


Abbildung 3: Die Ergebnisse zur Gesamtrate (klassische Delta-Methode) nebst 95%-Konfidenzintervallen basierend auf der GLMM_AC-Methode (Erfassungsrate) in Kombination mit der Bootstrapped track-mean-Methode (blaue Punkte/Fehlerbalken/Zahlen) bzw. wieder der GLMM_AC-Methode (schwarze Punkte/Fehlerbalken/Zahlen) bzgl. der Erkennungsrate). Rote gestrichelte Linie zeigt das Mindestkriterium der KNE-Checkliste für die untere Grenze des Konfidenzintervalls (70 %) an.

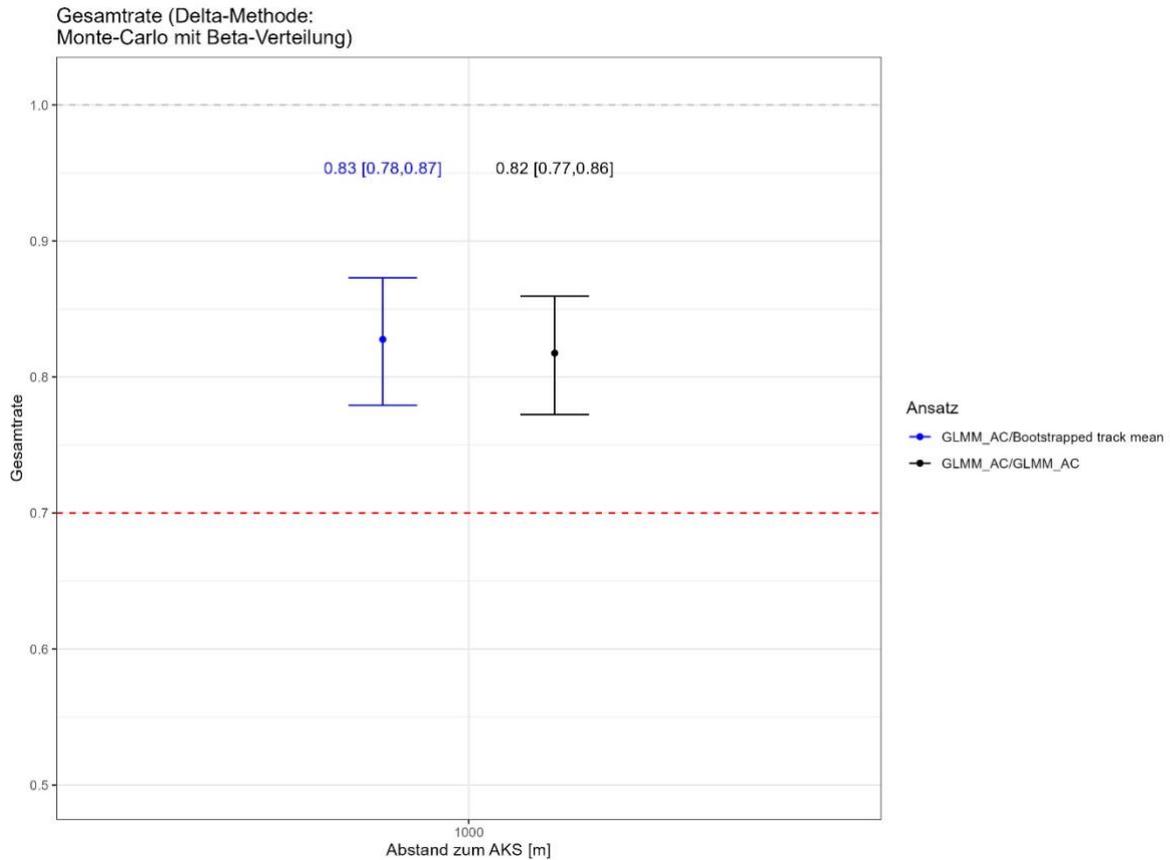


Abbildung 4: Die Ergebnisse zur Gesamtrate (Delta-Methode in Kombination mit Monte-Carlo-Simulationen und der Beta-Verteilung) nebst 95%-Konfidenzintervallen basierend auf der GLMM_AC-Methode (Erfassungsrate) in Kombination mit der Bootstrapped track-mean-Methode (blaue Punkte/Fehlerbalken/Zahlen) bzw. wieder der GLMM_AC-Methode (schwarze Punkte/Fehlerbalken/Zahlen) bzgl. der Erkennungsrate). Rote gestrichelte Linie zeigt das Mindestkriterium der KNE-Checkliste für die untere Grenze des Konfidenzintervalls (70 %) an.

5. LITERATUR

- Bolker, B., Brooks, M., Clark, C., Geange, S., Stevens, M. H., & White, J. S. (2008). Generalized linear mixed models: A practical guide for ecology and evolution. *Trends in Ecology and Evolution*, 24(3), 127–135. <https://doi.org/10.1016/j.tree.2008.10.008>
- Field, A., Miles, J., & Field, Z. (2012). *Discovering statistics using R*. SAGE Publications Ltd.
- Gross, M., Stieg, J., & Hinsch, H. (2024). AVES WIND ONSHORE Antikollisionssystem zum Schutz des Rotmilans (*Milvus milvus*)—Artenschutzfachliche Bewertung. *Bioplan, im Auftrag von ProTecBird*.
- Hurlbert, S. H. (1984). Pseudoreplication and the Design of Ecological Field Experiments. *Ecological Monographs*, 54(2), 187–211.
- KNE (Hrsg.). (2021). *Anforderungen an Antikollisionssysteme zum Schutz von Vögeln an Windenergieanlagen – Checkliste für eine qualifizierte Entscheidung über die Anwendbarkeit von Antikollisionssystemen* (S. 14).
X:\04_Literatur\01_Literatur_(NICHT_BEARBEITEN!)\KNE_2021_Anforderungen_an_antikollisionssysteme_zum_schutz_von_voegeln_an_windenergieanlagen.pdf
- Korner-Nievergelt, F., Roth, T., von Felten, S., Guelat, J., Almasi, B., & Korner-Nievergelt, P. (2015). *Bayesian Data Analysis in Ecology Using Linear Models with R, BUGS, and Stan*. Elsevier, London.
- Mercker, M. (2023). Statistische Begutachtung des Antikollisionssystems „AVES-Wind Onshore“ vor dem Hintergrund des LfU-AKS Prüfrahmens und in Anlehnung an die „KNE-Checkliste“. *Bionum GmbH*.
- Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effect models in S and S-Plus*. New York: Springer Verlag.
- Ransom, D., & Pinchak, W. E. (2003). Assessing Accuracy of a Laser Rangefinder in Estimating Grassland Bird Density. *Wildlife Society Bulletin (1973-2006)*, 31(2), 460–463.
- Reichenbach, M., Reers, H., Greule, S., & Grimm, J. (2024). Ausweitung des Zielartenspektrums von IdentiFlight—Methodische Ansätze zur Validierung neuer Versionen des neuronalen Netzes. *Bericht - ARSU und Oekeofor im Auftrag von e3 IDF GmbH*.
- Team, R. C. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org>.
- Zuur, A. F., Ieno, E. N., & Elphick, C. S. (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 1(1), 3–14.
- Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., & Smith, G. M. (2009). *Mixed Effect Models and Extensions in Ecology with R*. Springer Science+Business Media, LLC, New York.
- Zuur, A., Ieno, E., & Smith, G. M. (2007). *Analysing Ecological Data*. Springer Science+Business Media, LLC.