Statistical assessment of the AVES Wind Onshore against the background of the LfU testing framework for anti-collision systems and based on the KNE checklist

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

Bionum GmbH – Büro für Biostatistik

Finkenwerder Norderdeich 15A

21129 Hamburg

www.bionum.de

mmercker@bionum.de

HAMBURG, 14.02.'24

√BIONUM
Bionum GmbH
Dr. Moritz Mercker
Finkenwerder Norderdeich 15 a
21129 Hamburg / Germany
www.bionum.de / info@bionum.de

Table of contents

1. INTRODUCTION AND OVERVIEW

With the AVES Wind Onshore System, ProTecBird has developed an innovative AI-based anti-collision and monitoring system. As part of a pilot study in 2023, the performance of this system was analysed with regard to automatic bird detection and needs-based shutdown. In this study, a wide range of data was collected in a wind farm at two Nordex (N149) wind turbines, which allows assessing the performance of the system. Among the data collected were

- spatio-temporal laser rangefinder (LRF) data of approaching target species as well as other birds,
- spatio-temporal camera-based data of approaching target species as well as other birds recorded by the anti-collision system and
- spatio-temporal monocular range measurements collected by the system.

This report uses this data to analyse the performance of the system against the background of the recently developed LfU (State Office for Environmental Protection) testing framework for anti-collision systems (2024) and in accordance with the KNE (Competence Centre for Nature Conservation and Energy Transition) checklist for a qualified decision on the applicability of anti-collision systems (2021) stating the requirements for anti-collision systems for the protection of birds at wind turbines. The following issues are analysed and/or discussed:

1. Definition of suitable target type-specific detection and response areas.
2. The target species-specific detection rate (i.e. the rate at which target bird species are actually recognised as targets by the anti-collision system).
3. The target species-specific identification rate (i.e. the rate at which detected target species are correctly classified at species level by the anti-collision system).
4. The target species-specific total rate (i.e. the rate at which detected target species are detected and correctly classified at species level by the anti-collision system).
5. Does the performance of the anti-collision system differs significantly between the five different sample sites?
6. How accurate are the positions measured by the anti-collision system?

The results regarding the false positive rate (i.e. the rate at which the anti-collision system incorrectly initiates shutdowns for non-target species) will be provided at a later stage.

The present report so far considers and evaluates the red kite only.

2. SUMMARY AND COMMENTS

The statistical analyses of the empirical data show that both the requirements of the LfU testing framework and the specifications of the KNE checklist are met for the red kite.

The following response radius was calculated:

Red kite: $r_{response}$ = 385 m.

A detection area that forms a ring around the response area was chosen for the analyses (see LfU testing framework). For better transferability to different possible situations in approval procedures, rings/tori of different sizes are analysed.

Detection, identification and total rates were specified in accordance with the LfU testing framework and analysed using suitable regression models. For all chosen response areas, these (or their confidence intervals) resulted in values above the required minimum as defined in the LfU testing framework and the KNE checklist. In addition, it was analysed whether the total rate differs significantly between the different detection areas. This is largely not the case (in all cases $p > 0.2$, i.e. not significant).

3. DEFINITION AND CALCULATION OF RESPONSE AND DETECTION AREA

The response area defines the project-specific cylindrical 3D airspace around the centre of the rotor of a wind turbine up to which a bird has to be detected and classified to ensure that the wind turbine is shut down on time. Height and radius of the cylinder (torus) depend on the wind turbine-specific dimensions as well as horizontal and vertical flight speeds. The following sections are mainly based on the LfU testing framework.

The detection area defines the 3D airspace around the anti-collision system, which is used in the validation process (empirical determination of the rates) and for which a sufficiently high and validly determined total rate is required. In the approval procedure it is then checked whether the project-specific (cylindrical) response area is protected/covered by the (or one of the analysed) detection area(s).

This will be the case if the detection area surrounds the response area in form of a ring so that the relevant airspace around the response area is monitored and birds entering the response area are reliably detected and identified. This circular observation is relevant for the AVES Wind Onshore System, as it specialises in detection and identification over long distances and primarily monitors the area surrounding the response area.

A 3D torus is specified, which runs around the response area as a ring and is continued into the vertical axis. Its height is calculated in the same way as the height of the response area (rotor diameter plus a buffer depending on radius and flight speeds). The outer radius of the torus is used instead of the response radius. The resulting additional buffer in the vertical axis is necessary to ensure sufficient detection of all birds approaching the response area from above or below. The outer radius of the torus needs to be larger than the response radius and the inner radius needs to be smaller than or equal to the response radius. When assessing whether a response area is sufficiently covered by a detection area, the on-site difference (vertical and/or horizontal) of the anti-collision system relative to the rotor centre of the wind turbine has to be taken into account, i.e. added to the detection area.

The present project is a multi-camera anti-collision system that specifically adapts the number of cameras to the project. Four cameras were used in the present field tests for the rates/validation. The detection area is adjusted accordingly (one pie slice per camera).

In order to allow for application in as many scenarios as possible, an as large as possible detection area is chosen. Detection areas of different extents are included additionally to facilitate the assessment/transfer of the degree of protection to different project-specific response areas.

To determine the response area, the species- and wind turbine-specific response radius needs to be calculated. This radius ($r_{response}$) can be calculated mainly based on four values:

1. the average species-specific horizontal flight speed (v_{bird}),
2. the time that elapses between the system-induced switch-off signal and the wind turbine reaching idle mode ($t_{shutdown}$),
3. the average on-site relative measurement error of the anti-collision system ($c_{measurement\ error}$) and
4. a measure $c_{rotor\ blade}$, which is derived from the rotor radius.

The final formula results from

$r_{response} = (v_{bird} \cdot t_{shutdown} + c_{rotor\ blade}) \cdot (1 + c_{measurement\ error})$

(see LfU testing framework for anti-collision systems). The following values were used:

- v_{bird} = 8.54 m/s (see LfU testing framework for anti-collision systems) for red kite
- $t_{shutdown}$ is a combination of two sub-components: the time latency $t_{latency}$, which measures the time interval between the system-induced switch-off signal and the moment the switch-off signal arrives at the wind turbine in question (delays may occur here, for example, due to the prioritisation of signals in the context of wind farm control), as well as the time $t_{idle\ mode}$ the turbine requires on average to switch from regular operation to idle mode. Both values are project-specific. Realistic values were determined for these two variables during validation, namely $t_{idle\ mode}$ = 30 s and $t_{latency}$ = 2 s.
- $c_{rotor\ blade}$: 0.637-times the rotor radius, i.e. 0.637*75 m
- $c_{measurement\ error}$: defines the average on-site relative measurement error of the anti-collision system. Relative because the error naturally increases with the distance between the anti-collision system and the target. Therefore, the absolute error (e.g. given in metres) is divided by the respective distance between the anti-collision system or wind turbine and target to obtain the relative measurement error as a measure that is independent of the distance. For this purpose, BioConsult SH compared the red kite laser rangefinder points of five survey days to the positions determined by the anti-collision system. On average, the deviation was 20 % for red kite. LRF points were only included in the statistics if the point was correctly detected and recognised by the anti-collision system. A total of 548 LRF points at different distances between the anti-collision system and target were analysed.

This results in a response radius of:

Red kite: $r_{response}$ = 385 m.

The height of the response area is determined by the lower and upper rotor limits and an additional buffer. The buffer is determined from the assumed the change in altitude of the flying bird while it passes the response radius. The empirical measurements suggest that all investigated birds studied show vertical speeds of around 1 m/s while ascending or descending across all species (see LfU testing framework). This means that the buffer (in units of metres) results directly from the quotient of the response radius and the horizontal flight speed. However, in order to take a particularly precautionary approach in this project, the investigated detection range is not limited upwards and downwards, but includes all LRF points that were located in the air column above the torus defined above. In this way, even birds flying from the ground or far above into the response area are represented in the determined rates.

4. FIELD AND STATISTICAL METHODS

The following information on the detection rate is based on reliable survey results that were empirically collected and statistically validated.

1.1 Methodology of tests in the field and matching of anti-collision system and laser rangefinder

The detection rate was investigated based on a comparison between birds that were detected in the field stating time and place by ornithologists (Bioplan) with those detected by the AVES Wind Onshore system. The observers used laser rangefinder (LRF), which can locate the birds relatively precisely (Ransom & Pinchak, 2003). For this reason, it is (simplified) assumed that the LRF data do not include any significant spatial and/or temporal errors.

A total of 12,024 red kite LRF points from five different locations 25 survey days (04.08.2023–14.09.2024) were included in the analyses.

An a priori pre-selection of the LRF points was carried out exclusively with regard to the question of whether or not they were located in the pie slice in front of the camera (see above). These pie slices enclose more airspace than the wedge-shaped fields of view of the cameras can capture when viewed from the side. The technical coverage rate is therefore automatically included.

The following analyses and steps were carried out here and in matching LRF points and anti-collision system signals:

1. Using the LRF coordinates, the camera positions and the calibrated camera angle scale, theoretical angles (pan, tilt) were generated per point where the bird is visible.
2. The detection area is a torus around the camera (which is attached to the wind turbine) with a fixed inner and outer radius and a fixed height. Horizontally, the detection area extends over a viewing angle of 30° per camera (pie slice).
3. The log files of the AVES Wind Onshore system provided the actual angles at each point in time. It was also possible to derive the default orientation, which defines the position of the fixed detection area. The default orientation was determined algorithmically from the log files in order to avoid (error-prone) hard-coded angles. As the cameras can only return to a default value with a limited tolerance even after tracking a bird, the horizontal angle of the default orientation may fluctuate by up to 1° over the course of a session.

4. Check of the detectability of each point:
    a. All points in the area defined above were considered detectable (in the following, "detectable" is synonymous with "within the pie slice" – see above for more information about the technically induced coverage rate).
    b. The cameras can be automatically panned to track birds. This results in an additional dynamic field of view of a camera, which can be limited to a width of less than 30° due to zoom. An LRF point whose angle was within this dynamic field of view was considered detectable if it was within the inner and outer torus boundaries. This is necessary because the calculation of whether an LRF point is within the dynamic field of view often leads to errors. The angle interval of the field of view at maximum zoom is < 4°, which is well below the accuracy of the angle calculation. Example: a bird is tracked and zoomed in to a viewing angle of 3°. The calculation using the angles shows that the bird can be seen theoretically at Pan = 34°. The dynamic camera range is Pan = [37° - 1.5°, 37° + 1.5°]. According to the figures, the LRF point is therefore not in the field of view of the camera. The system indicates a tracked bird at 37.3°. The difference between the two angles is 3.3°. This is within the assumed tolerance of 11°. The bird is considered tracked and therefore detectable. Due to the checks described above, birds detected to the side and above the detection range were included in the analysis. However, the radial distance condition was strict, i.e. LRF points with a radial distance lower than the inner radius or higher than the outer radius of the torus were never considered detectable.
5. Check of detection/angle matching:
   In the event that a track entry was available at the LRF time, it was checked whether the theoretical angles (LRF pan, LRF tilt) matched the angles of the tracked object (track pan, track tilt). A tolerance of 11° (pan) or 9° (tilt) was used in the calculation. If the conditions described above are met, the LRF point will be counted as detected.
6. Determination of the AI classification:
   A time interval that spanned half the time until the next LRF point on the same track was defined around each LRF time point. For LRF points at the beginning or end of an LRF track ID, 3 s were added before or after the LRF point of time. Within the time interval determined in this way, the AI classifications were collected from all log file entries. The most frequent final class entry was the AI class of the LRF point. The AI classes are given by the classes "target" (here: red kite), "non-target" (all other bird species) or "bird" (unidentified bird). The 'final class' is an average of all previous individual classifications of the currently detected bird which is specially adapted to the application purpose. Please note: It is also possible to read out an AI classification without a suitable angle matching, e.g. if the system is tracking a different bird than the one mapped by the LRF.
7. Determination of the final variable "red kite detected and identified" vs. "not detected or identified":
   If an LRF point produced a successful angle match with the system data and the KI classification matched the actual bird species, the bird was assumed to be correctly detected and identified.
8. Use of the filter "at least 4 detectable points per track":
   Every red kite that was "detectable" should be included in the statistical analysis. However, tracks with too few points (less than or equal to 3) were excluded (see LfU testing framework) as they can lead to a disproportionate increase in variance and/or biased results in the context of the regression analysis with often very low proportions of the total data. In addition, LRF tracks with only a few points usually are not flights with a high risk of collision as the bird only appeared in the detection area for a short period.

1.2 Statistical evaluation

The data to be analysed are binary variables that are classified, for example, as "detected" or "not detected". In the simplest case, it is possible to determine mean values of the corresponding rates. However, the variable is not normally distributed, so techniques that do not assume a normal distribution must be used to calculate the variance (e.g. to obtain confidence intervals).

In fact, the statistical situation is somewhat more complex, because time series are analysed and the LRF points are assigned to different individuals. This may lead to pseudo-replication (Stuart H., 1984) and thus result in underestimation of confidence intervals and biased calculated rates.

For this reason, an analysis strategy was used that adequately considers this data situation. The analysis was carried out using suitable regression methods, i.e. logistic regression methods, which belong to the generalised linear models (GLMs) (Benjamin M. et al., 2008; Field et al., 2012; A. Zuur et al., 2007). The allocation to different individuals was included by using the track ID as random intercept in the context of mixed modelling, which led to the class of generalised linear mixed models (GLMMs) (Benjamin M. et al., 2008; Pinheiro & Bates, 2000; A. F. Zuur et al., 2009). The temporal autocorrelation was analysed using pACF plots and integrated as a suitable auto-regression structure (AR1) (Korner-Nievergelt et al., 2015; A. Zuur et al., 2007; A. F. Zuur et al., 2009). Please see the LfU testing framework (2024) for further details about method and background.

All statistical analyses were carried out using the open-source software R (R Core Team, 2023) R package MASS (Modern Applied Statistics with S. Fourth Edition, 2002).

As the AVES Wind Onshore System is a multi-camera anti-collision system, the overall system was considered a single recording unit in the analysis of tracks or rates. Through exchange of data all cameras together trigger a distinct response of the overall system (i.e. a shutdown signal if necessary) but each camera is able to send a shutdown signal on its own without prioritising individual cameras.

5. RESULTS

1.3 Total rate

The results for the total rate (see definition in the LfU testing framework) for the red kite are shown in Figure 1. Three slightly different ring-shaped detection areas with a torus of 200 metres width were investigated. The average values are 85–87 % and the lower limits of the confidence intervals are 75–79 %, i.e. above the 70 % required by the LfU testing framework.

Logistic GLMM regression methods were used to analyse whether there are significant differences in the total rates between the different areas investigated. This is not the case; all p-values are higher than or equal to 0.2 and therefore highly non-significant.

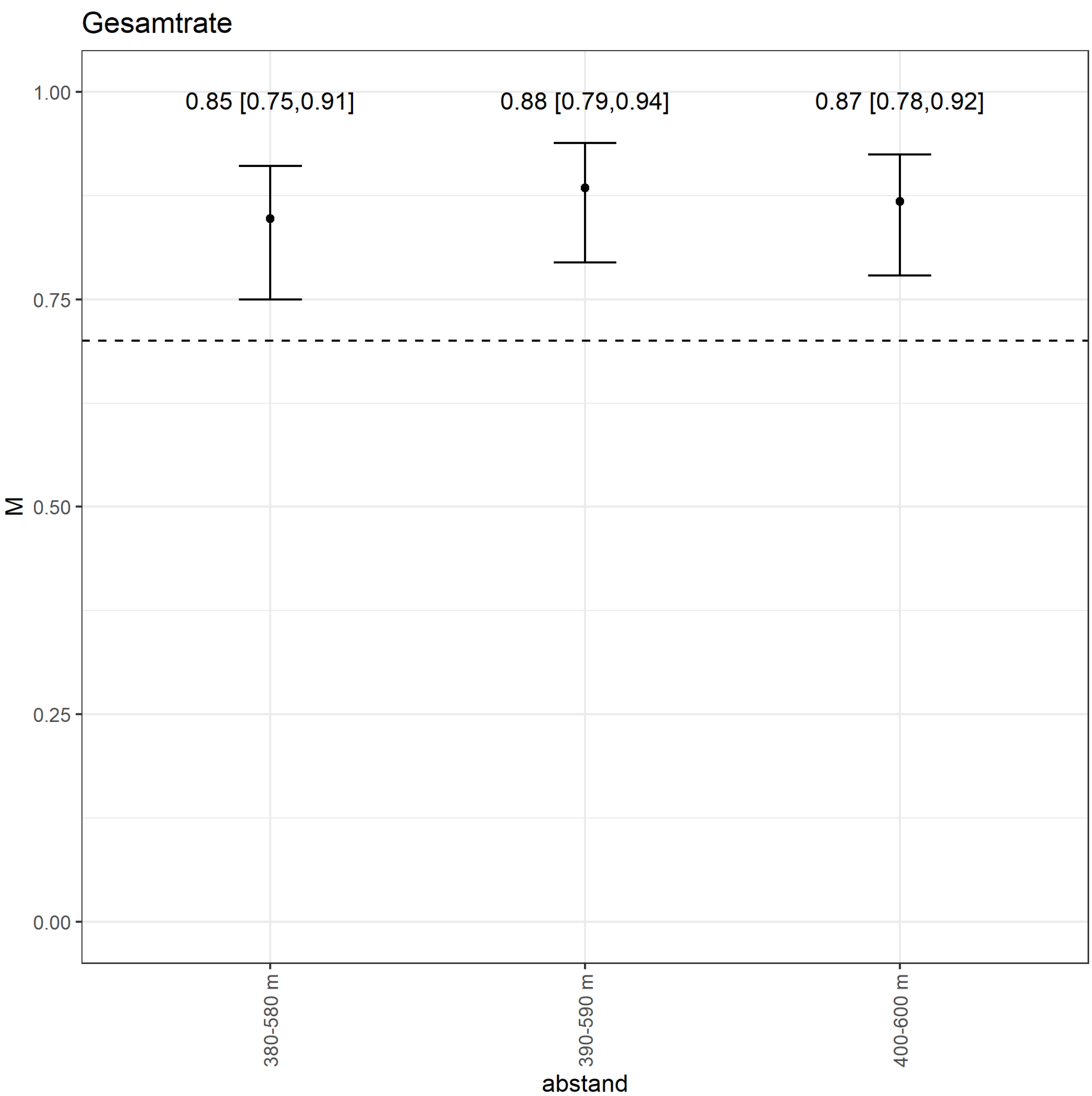

Figure 1: Statistical results for the total rate for red kite based on logistic GLMM analyses. Black dots indicate mean values, black bars 95 % confidence intervals.

1.4 Detection rate

The results for the detection rate for the red kite are shown in Figure 2. Three slightly different ring-shaped detection areas with a torus of 200 metres width were investigated. In all cases, the mean value is 95–97 % and thus well above the 75 % required in the KNE checklist.

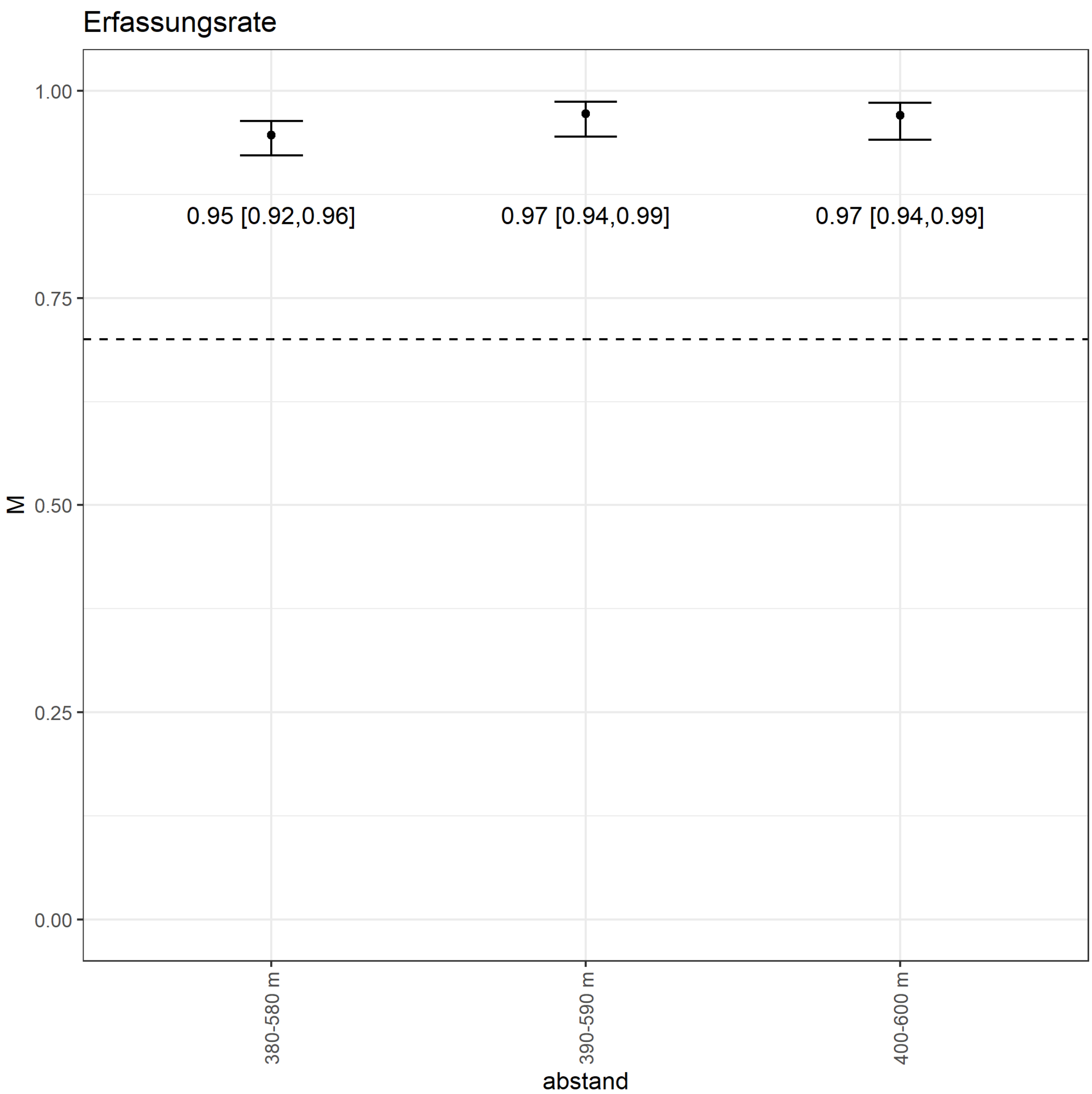

Figure 2: Statistical results for the detection rate for red kite based on logistic GLMM analyses. Black dots indicate mean values, black bars 95 % confidence intervals.

1.5 Identification rate

The results for the identification rate for red kite are shown in Figure 3. Three slightly different ring-shaped detection areas with a torus of 200 metres width were again investigated. The mean values are 87–98 % and thus well above the minimum value of 75 % required according to the KNE checklist and close to or above the desired value of (at least) 90 %.

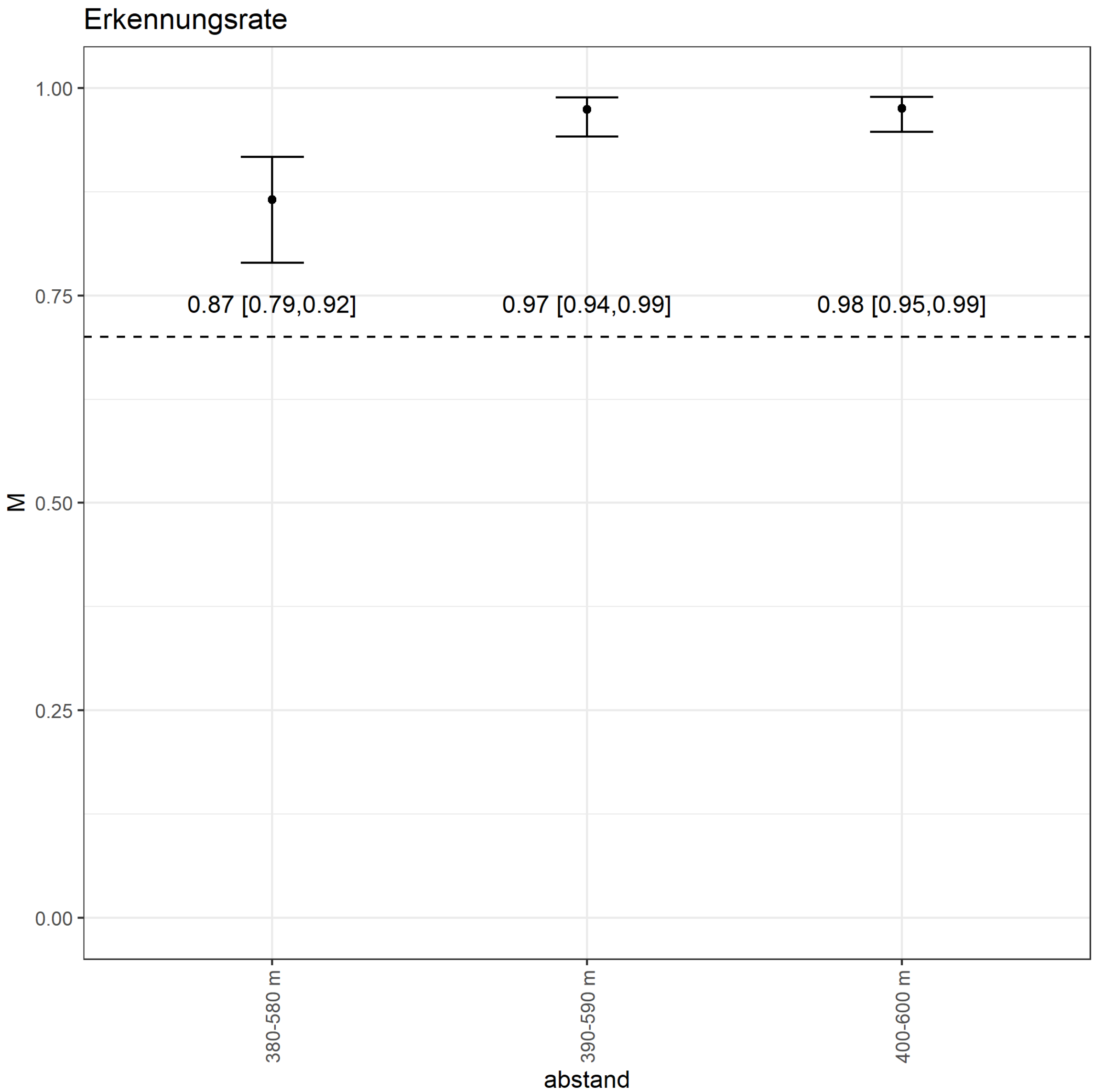

Figure 3: Statistical results for the identification rate for red kite based on logistic GLMM analyses. Black dots indicate mean values, black bars 95 % confidence intervals.

6. SPATIAL COVERAGE AT THE SITE

For the test, four cameras were installed at two Nordex N149 wind turbines in a wind farm. Figure 4 shows the individual field of view of the respective cameras (50 to 54) and Figure 5 a photo overview of the site.

A complete AVES Wind Onshore system would equip most of the outer wind turbines of a wind farm with cameras, resulting in a 360° surround view. Here only a part of the wind farm was equipped with cameras to simulate two wind turbines on the eastern edge of a larger wind farm.

Figure 4 to Figure 6 illustrate that there are no significant visual barriers in the area of the field of view of the four installed AVES systems. The hedgerow structures are below the height of the installed cameras and the forests are at a sufficient distance for the cameras to look beyond these structures.

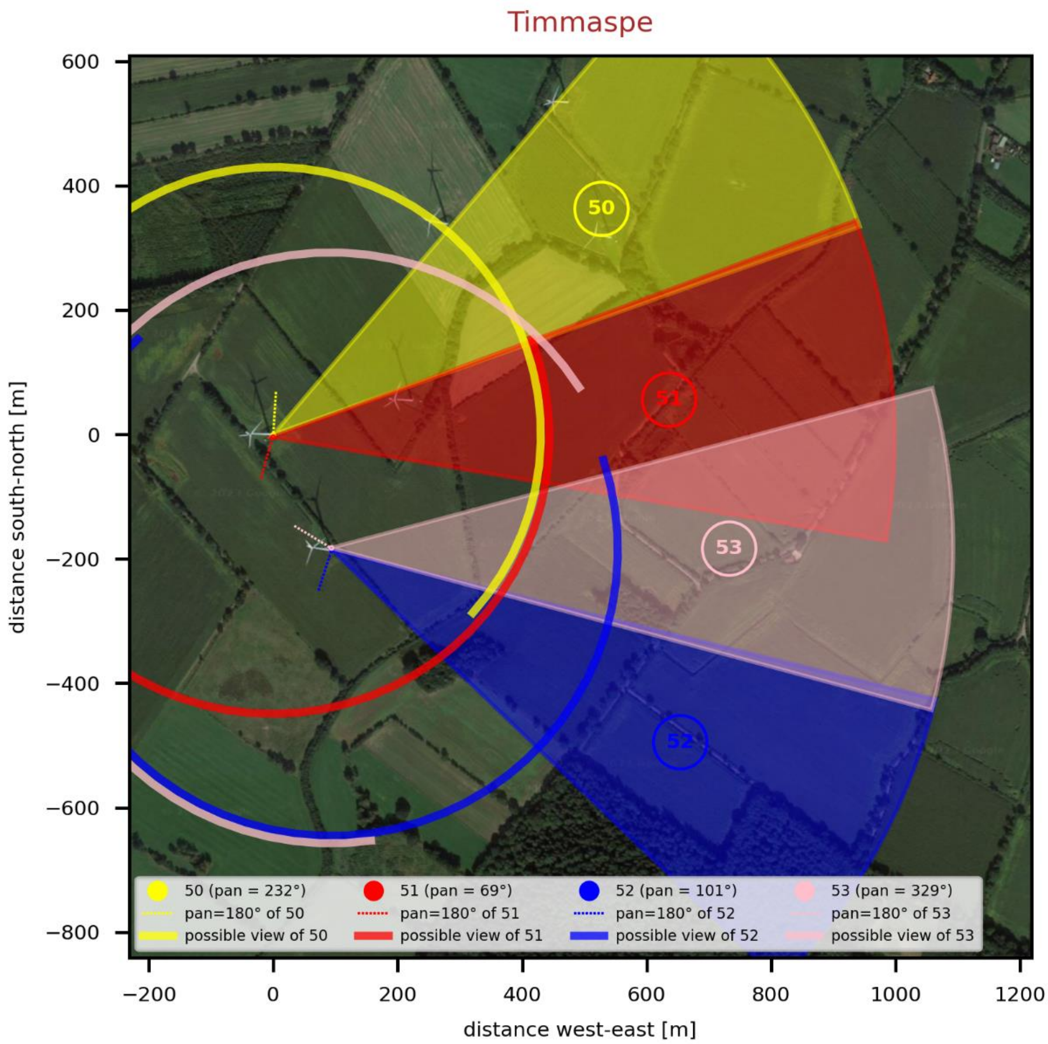

Figure 4: View of the two wind turbines with AVES system, showing the respective field of view of the individual cameras.

Figure 5: Overview of the wind farm with the two wind turbines and the four installed cameras (Photo: BioConsult SH 2023).

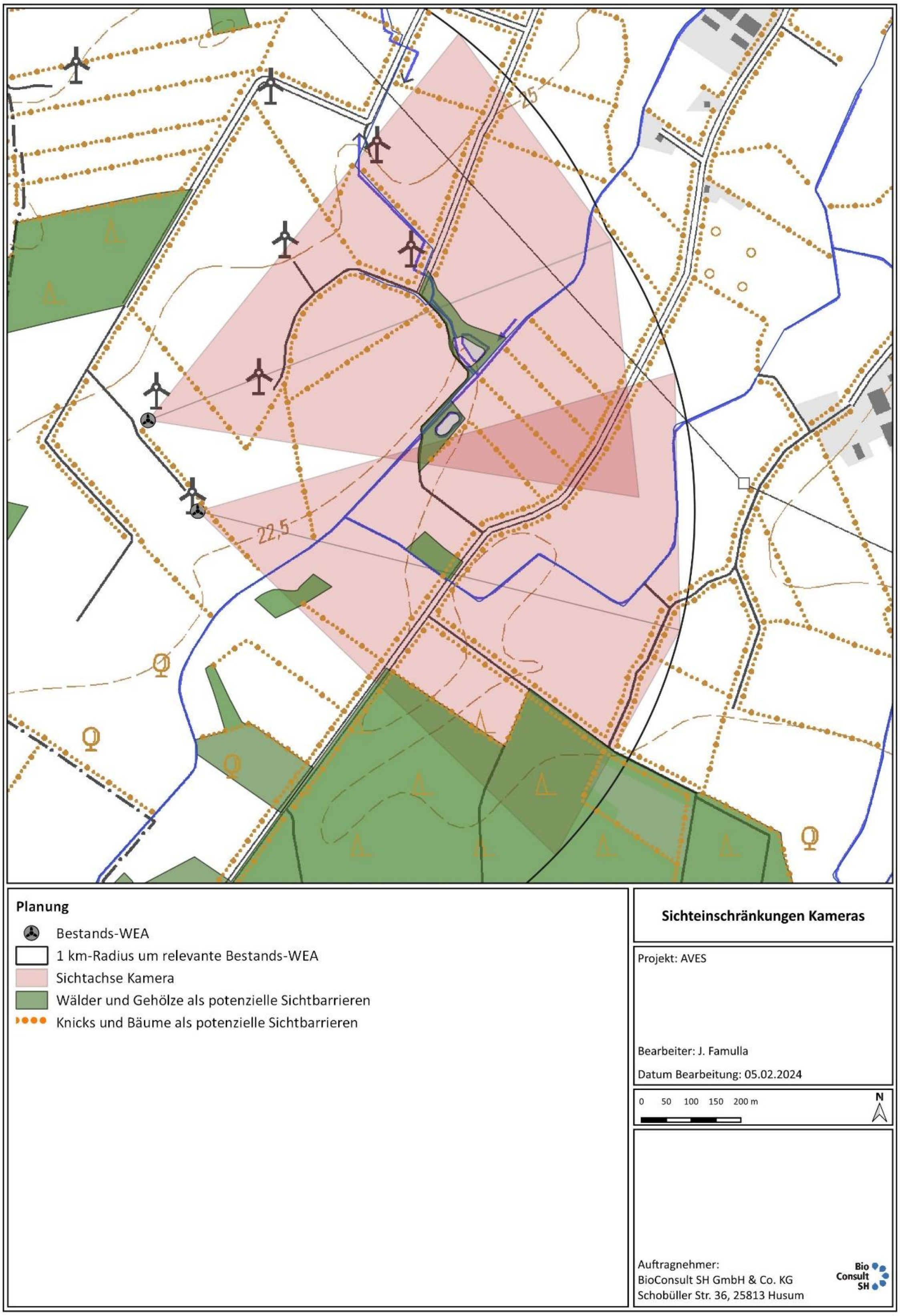

Figure 6: View of the two wind turbines with AVES system, showing the respective field of view of the individual cameras and potential visual barriers as hedgerows and forests.

7. LITERATURE

Benjamin M., B., Mollie E., B., Connie J., C., Shane W., G., M. Henry H., S., & Jada-Simone S., W. (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. https://doi.org/Added

KNE (Hrsg.). (2021). *Anforderungen an Antikollisionssysteme zum Schutz von Vögeln an Windenergieanlagen – Checkliste für eine qualifizierte Entscheidung über die Anwendbarkeit von Antikollisionssystemen* (S. 14). KNE_2021_Anforderungen_an_antikollisionssystemen_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. https://doi.org/Added

LfU AKS-Prüfrahmen (2024) -- publication in preparation.

Modern Applied Statistics with S. Fourth Edition. (2002). Springer, New York. https://doi.org/Added

R Core Team, R. C. (2023). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org*. https://doi.org/Added

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effect models in S and S-Plus*. New York: Springer Verlag. https://doi.org/Added

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.

Stuart H., H. (1984). Pseudoreplication and the Design of Ecological Field Experiments. *Ecological Monographs*, *54*(2), 187–211. https://doi.org/Added

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. https://doi.org/Added

Zuur, A., Ieno, E., & Smith, G. M. (2007). *Analysing Ecological Data*. Springer Science+Business Media, LLC. https://doi.org/Added