Ornis Norvegica

Which characteristics are important for species identification of birds in flight? Results from a survey of Norwegian birdwatchers

Roel May*, Jiska van Dijk & Bård G. Stokke

Norwegian Institute for Nature Research (NINA), P.O. Box 5685 Torgarden, NO-7485 Trondheim, Norway
* Correspondence: roel.may@nina.no

Cite this article
May R, van Dijk J & Stokke BG. 2025. Which characteristics are important for species identification of birds in flight? Results from a survey of Norwegian birdwatchers. *Ornis Norvegica* **48**: 12–18. doi: 10.15845/on.v48.4196

Abstract
A better understanding of how birdwatchers identify species of birds in flight may support the development of machine learning algorithms for automated identification from camera-tracking systems for bird monitoring and mitigation. Norwegian birdwatchers scored the importance of 18 criteria for identifying species of birds in flight in an online anonymous survey. Responses were analysed using an Analytical Hierarchy Process and Bayesian Belief Networks. Species identification was first affected by a seasonal expectation as to which species may be observed during a birding trip. Criteria linked to bird's appearance were most important for species identification, including plumage colouration or patterns; body and wing shape; beak, neck and tail shape. However, flight pattern and speed may provide additional information. A hierarchical approach to categorisation and species identification may improve processing time of automated algorithms.

Received 18 March 2024
Accepted 04 July 2025
Published 10 November 2025

Keywords
Analytical Hierarchy Process, appearance, Bayesian Belief Network, environment, flight characteristics, species filter

INTRODUCTION

With increasing improvements in the quality of camera technology combined with machine learning opportunities, there has been increased interest in the development of automated identification of birds from camera systems (Ferreira et al. 2020, Niemi & Tanttu 2018, McClure et al. 2018). The development of such systems has been driven by multiple factors, including an increasing pace of wind energy development, and requirements to monitor birds in the vicinity of wind turbines and reduce collisions through curtailment (McClure et al. 2021, Niemi & Tanttu 2020). Identifying species of birds in flight is a difficult task that can be based on attributes such as shape and colouring of the detected individual (Liu et al. 2007, Marini et al. 2013) or facial recognition (Berg et al. 2014). Nevertheless, humans are able to detect and recognize objects within a split-second (Thorpe et al. 1996, Grill-Spector & Kanwisher 2005). Development of an 'artificial birdwatcher' based on machine learning techniques can therefore be supported by a better understanding of how humans identify different species of birds in flight. In such circumstances when decisions are made rapidly, it is reasonable to expect that different criteria are used near simultaneously by birdwatchers to identify which species they have encountered. The purpose of this study was to gain new insights into how birdwatchers identify different species of birds in flight.

METHODS

On 22 April 2021, an online survey was published on social media, Facebook groups and on NINA's website aimed at Norwegian birdwatchers. The anonymous survey was based on bird species that can be observed in Norway and disregarded the use of sound as a means of identification. The latter limitation was set because of the intended purpose was to provide input to development of machine learning algorithms for automatic recognition of birds with camera systems. In the survey, the importance of various criteria for identifying species of birds in flight were assessed. A total of 18 criteria grouped into four different main categories (Table 1), which were weighed against each

Table 1. Overview of 18 criteria used to identify different species of birds in flight that were organized into four main categories and included in the online survey.

Species filter	Appearance	Flight characteristics	Environment
• *Geographical area* on land, at sea, southern or northern Norway	• *Body shape* body size, wingspan, body length, silhouette	• *Flight pattern* gliding, hovering, circling, undulating, zigzagging, straight line	• *Time of day* morning, noon, afternoon, evening, midnight
• *Landscape type* sea, coast, open cultural landscape, forest, mountains	• *Wing shape* width, pointed, rounded	• *Wing beat frequency* slow, fast, jerky	• *Light conditions* dusk, sun, cloudy, foggy
• *Season* winter, spring, summer, autumn	• *Tail shape* square, rounded, pointed, forked	• *Flight altitude* being high up in the air low over the ground	• *Wind conditions* wind direction, wind speed
	• *Neck shape* short, long, retracted, extended	• *Flight speed* slow, fast	• *Other weather conditions* precipitation, temperature
	• *Beak shape* pointed, wide, crossed	• *Flocking behaviour* flock size, flock formation	
	• *Plumage* colouration, pattern, feather change		

other. The four main categories, which were considered important for species identification of birds in flight were: Species filter, Appearance, Flight characteristics and Environment.

Species filter: Regardless of whether birds are flying or performing other activities, an initial expectation is likely formed as to which species are likely to be observed during a birding trip. This expectation is not linked to specific species per se, but to the context of where and when an observer goes out to see birds. In this way, a birdwatcher can have expectations on the probability of encountering certain species in the area they are traveling in during a given time of the year. For example, it is unlikely that an observer would find a Black Woodpecker *Dryocopus martius* on the highest mountain peaks, an Atlantic Puffin *Fratercula arctica* in a deciduous forest in Southern Norway, or a migratory Common Cuckoo *Cuculus canorus* in Norway during midwinter. In this way, a number of species are filtered out of the list of probable species one expects to encounter in a certain geographic area, landscape type and season.

The two main categories Appearance and Flight characteristics focus on visual features of the bird itself, and a number of criteria can be considered important under each of the two categories. Some features concern the general impression of the bird, including plumage traits and the shape of the body, wings, tail, neck or beak. These criteria are placed in the main category Appearance. For example, in most cases one can distinguish a Common Crane *Grus grus* from a Grey Heron *Ardea cinerea* by looking at the shape of the neck in flight. The last main category, Flight characteristics, concerns specific flight-related characteristics including flocking behaviour, flight height, flight speed, wing beat frequency and flight pattern. For example, a Great Spotted Woodpecker *Dendrocopos major* has an undulating flight, whereas a Collared Dove *Streptopelia decaocto* typically flies with clipped wing movements.

Environment: Other conditions that are not directly linked to species, but which can nevertheless have an impact on whether a species is observed during a birding trip, include environmental conditions on the day of observation. Important factors could include the time of day, light conditions or prevailing winds and other weather conditions. The probability of observing some seabirds can, for example, be greatest with an onshore wind, and many species are more active in some periods of the day.

In the survey (Appendix 1), the respondents were first asked to indicate whether they considered themselves to be professional or hobby birdwatchers, viewing birds as part of their job or as a recreational or leisure activity. Furthermore, the regularity and seasonality of their birdwatching activity was scored, as well as in which type of landscape that respondents

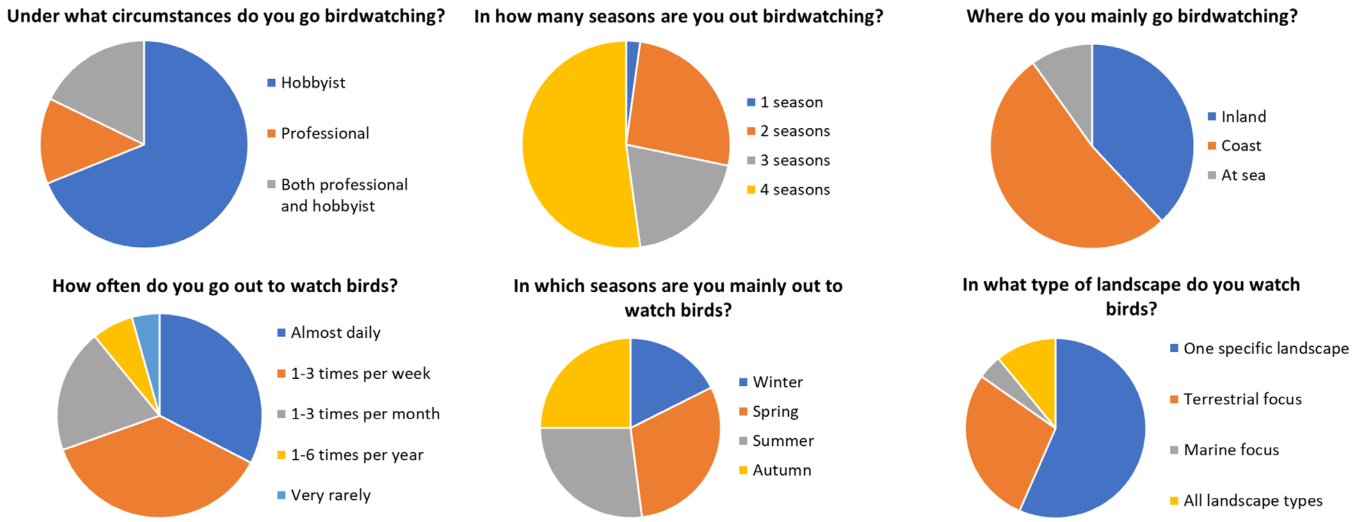

Figure 1. Main characteristics of the respondents to the online survey regarding who they are, when and where they spend time birdwatching.

frequented most (Figure 1). From the season(s) indicated (multiple choice), also the number of seasons could be determined for when respondents are birdwatching. From the type(s) of landscape indicated (multiple choice), also combinations could be assessed (single landscape; terrestrial focus: inland & coast; marine focus: sea & coast; all landscapes). To obtain the relative importance of each criterion in the process of species identification, the respondents were asked to score all unique criteria combinations in a pairwise manner within each of the main categories. For each unique pair of criteria, the more important criterion was to be selected and how much more important this criterion was relative to the other was scored on an ordinal scale from 1 to 9 (equally important to very much more important). After a first round of assessments within each main category, the four main categories themselves were scored pairwise in the same way as described above. Finally, the respondents were asked whether the survey missed any criteria that they considered to be important for species identification.

All responses were first converted to relative weights using an Analytic Hierarchy Process (AHP) within each main category and across main categories (Figure 2). Conversions were completed using the ahpsurvey package (Cho 2019) in the software R version 3.6.3 (R Core Team 2020). The AHP weights across the main categories were adjusted for the number of criteria included in each main category. The null expectation was an equal importance of all 18 criteria, rendering relative weights of 0.056 or 1/18th following a uniform distribution. For incomplete cases, missing values were imputed using an algorithm that assumes perfect agreement with the other pairwise comparisons (Harker 1987). The potential effect that excluding missing values versus imputing values had on the relative weights was tested using a Bland-Altman test with the blandr.statistics function in the blandr package (Datta 2017). The consistency ratio within respondents and the level of consensus based on Shannon entropy among respondents were calculated (Goepel 2018). The first metric indicates how consistent each respondent scored among criteria pairs, and the second indicates the extent of respondents weighing criteria similarly. Low consistency combined with high consensus would indicate that use of alternative criteria is situation dependent but that different observers scored the criteria in the same way.

To further assess to which extent several criteria were determined simultaneously for species identification, the relative weights were log-transformed to approximate a Gaussian distribution. The data were then used to create a continuous Bayesian Belief Network (BBN) to examine in more detail how the different weights influenced each other; indicative of criteria being considered jointly. A BBN is a probabilistic graphical model consisting of a directed acyclic network without feedback loops that quantifies the strength of relationships between variables based on probability distributions using Bayes' Theorem (Chen & Pollino 2012). The BBN model was created with the Hill-Climbing learning algorithm using the bnlearn package (Scutari 2010) in the software R version 3.6.3 (R Core Team 2020) and visualized using the visNetwork package (Almende et al. 2019). The relative contribution of each criterion to the overall goodness-of-fit of the network was measured as the ratio of the mean Bayesian Information Criterion (BIC) contribution across criteria by each criterion's BIC contribution. Values above one indicate a larger contribution. The contribution is unrelated to the actual relative weight of a given criterion, but rather associated with the level of dependency or correlation with the other criteria. For each of the edges, also the strength of the probabilistic relationships was calculated using the arc.strength function. The strength was defined as the BIC score

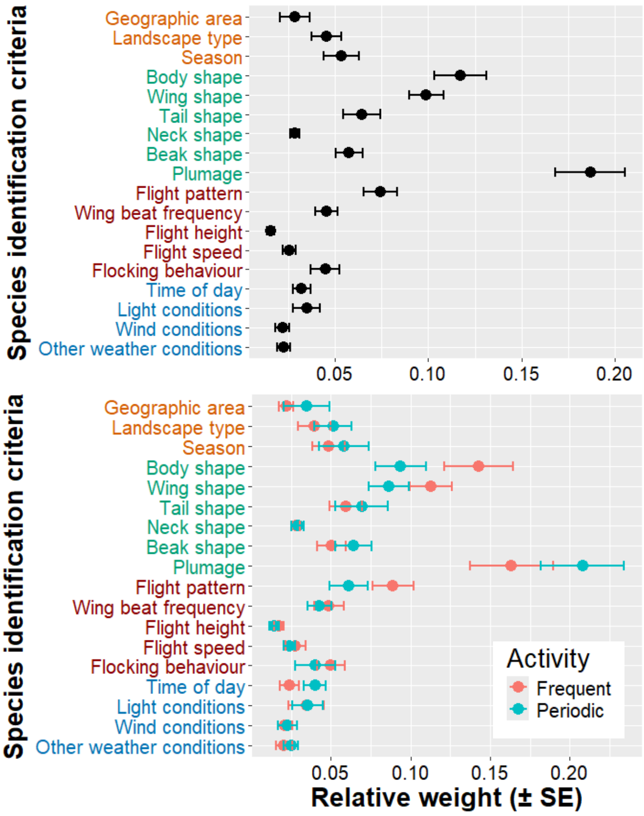

Figure 2. Relative importance (weight ± SE) of 18 criteria used to identify species of birds in flight across main categories (adjusted for the number of criteria within each main category) for all respondents (upper panel) and divided by how often the respondents are active (lower panel). Birdwatchers were grouped as either Frequent = Active throughout the year, and at least once a week. Periodic = Remaining respondents. The dotted line indicates the proportion for a uniform distribution. Text colour on the Y-axis indicates criteria in different main categories (see Table 1).

gain/loss caused by removal of an arc from the BBN, where a strong relationship led to a larger difference. These values were made relative by taking the ratio of each edge's strength by the mean across edges. Finally, a sensitivity analysis was performed using a leave-one-out procedure to investigate the effect that respondents had on the BBN model's goodness-of-fit. BIC scores for compared to the BIC score for the total model, adjusted for reduced sample size.

RESULTS

A total of 46 respondents answered the anonymous survey. In four cases, the responses from the survey were incomplete. Most of the respondents were characterized as hobby birdwatchers (69%) who are consistently very active (daily or several times a week; 70%). Over half of the respondents were active throughout the year (52%), and with a majority of birdwatching activity in coastal areas (52%) (Figure 1). For the main category Appearance, one additional criterion was identified as missing in the survey: "leg position in flight" such as the legs hanging down or extending out behind the tail during flight.

The four incomplete cases did not significantly affect the outcomes of the relative weights ($t < 0.001$, df = 17, $P = 1$; range: -0.002–0.004). The mean consistency ratio within respondents was 0.185 ± 0.093 SD. Consensus among respondents was 0.709. Appearance was considered the most important main category for species identification of birds in flight, with Plumage as the most important criterion. This criterion was followed by Body shape and Wing shape (both in the main category Appearance), followed by Flight pattern in the main category Flight characteristics (Figure 2 – upper panel). These five criteria, together with Tail shape, Beak shape (in the main category Appearance) and Season (in the main category Species filter) all had relative weights greater than the null value of 0.056 and totalling to 65.5%. We assumed that experience was correlated with frequency and divided the respondents into two groups according to how often they are active: Frequent (all seasons and at least once per week, n = 22) versus Periodic (less often than once per week, n = 24). Comparison of the two groups provided an indirect picture of the importance of experience in the species identification process. Consistency (0.210 ± 0.104 SD vs 0.156 ± 0.071 SD (lower is better)) and consensus (0.687 vs 0.748 (higher is better)) were both lower for periodic compared to frequent birdwatchers. In general, a picture emerges that the relative importance of the criteria is consistently similar for the two groups of birdwatchers (Figure 2 – lower panel). While for periodic birdwatchers Environment (0.123 vs 0.100) and Species filter (0.144 vs 0.110) were more important, frequent birdwatchers weighed Appearance (0.550 vs 0.556) and especially Flight characteristics (0.182 vs 0.233) higher. Flight pattern, Wing shape and Body shape had higher importance for the most active birdwatchers, whereas Time of day, Plumage and Beak shape seemed to have somewhat less importance.

The BIC score of the total BBN model including all 46 respondents with imputed data was -1140.41. The leave-one-out procedure indicated that the respondents had negligible effects on the overall goodness-of-fit of the model (adjusted BIC total: -1115.62; mean BIC score leave-one-out: -1117.73 + 3.86 SD; mean loss: 2.11 [-11.50–8.43]; coefficient of variance: 0.35%). Many of the 18 criteria influenced each other, with the criteria with relative weight >0.056 having an overall lower contribution (40%) to the BBN model (Figure 3) compared to their overall importance weights. The findings indicate that the importance of those criteria, apart from Wing shape, were less dependent on other criteria. Wing shape contributed more to the network (BIC ratio: 1.15), and interacted with Tail shape, Neck shape and Time of day. Interactions mean that the scores given by respondents for Wing shape were affected by or affecting (depending on the direction of the arrows) the scoring for other criteria. In contrast, Flight pattern had a lower contribution to the network

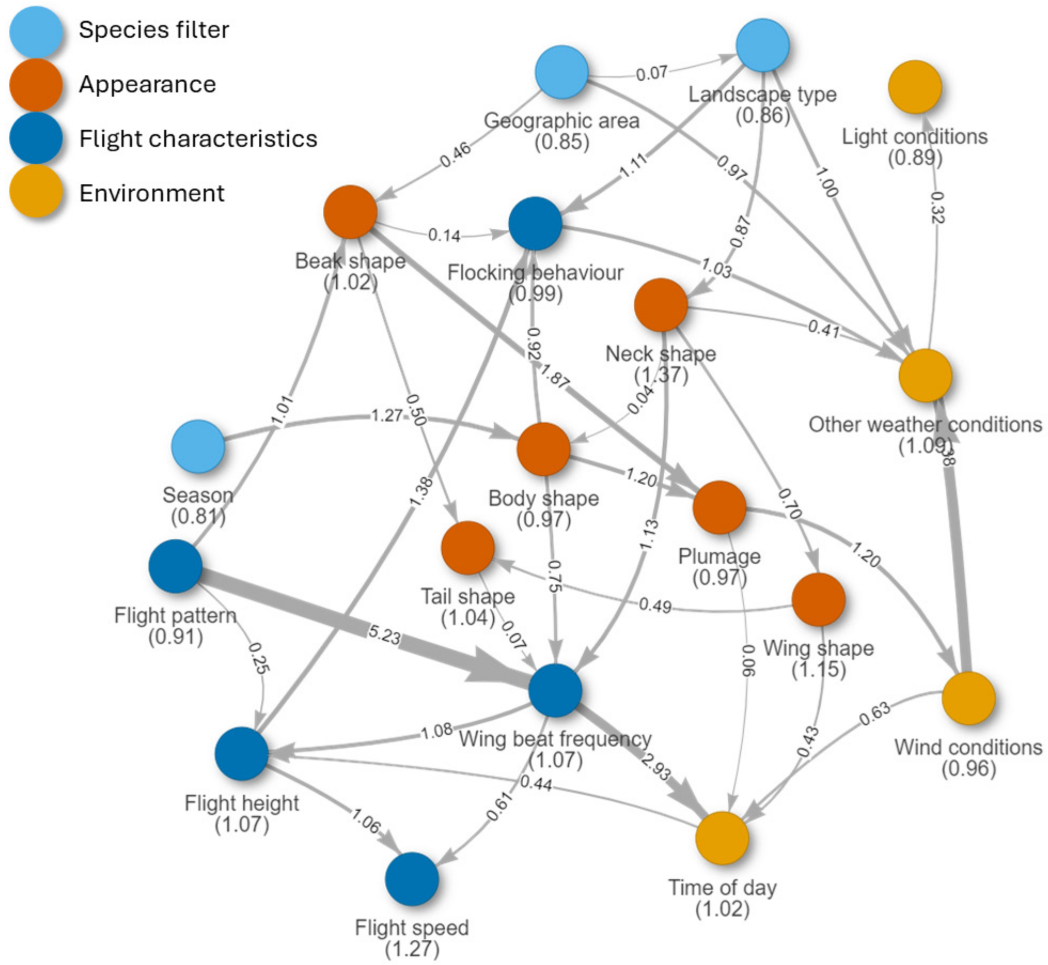

Figure 3. A Bayesian Belief Network (BBN) model predicting the connections among criteria used by birdwatchers to identify species of birds in flight. Colours indicate the four main categories (light blue: Species filter; orange: Appearance; dark blue: Flight characteristics; yellow: Environment), where thickness of the arrows indicates the strength of the relationships between pairs of criteria.

(BIC ratio: 0.91), but had the strongest interaction with Wing beat frequency. The two most influential criteria in the network were Neck shape and Flight speed (BIC ratio of 1.37 and 1.27). The spatial and temporal criteria Season, Geographic area, Landscape type and Light conditions were least influential (BIC ratio of 0.81, 0.85, 0.86 and 0.89). Variation among the criteria and their interactions illustrate the complex and integrated dynamics with regard to species identification.

When separate models are constructed for frequent and periodic birdwatchers while only including the criteria within the categories Appearance and Flight characteristics, a clearer pattern is obtained on how experience affected the process of species identification (Figure 4). For periodic birdwatchers, Appearance and Flight characteristics criteria interacted with each other within each category, but with linkages via the most central node Flight height (BIC ratio of 1.18) and via Flocking behaviour. For frequent birdwatchers, Appearance criteria were less interlinked and more integrated with Flight characteristics criteria. Neck shape was the criterion contributing most to the model (BIC ratio of 1.61), but did not have any linkages to other criteria.

DISCUSSION

The total number of responses was unfortunately limited, although the survey was disseminated in various channels. The sensitivity analysis showed however that respondents represented a coherent group, given the low level of loss in the leave-one-out procedure and a low coefficient of variance. Consensus among respondents was moderate because the estimates were slightly below the general rule-of-thumb threshold of 0.75 (Goepel 2018). On the other hand, consistency in this study was reasonable because values were between the rule-of-thumb thresholds of 0.1 and 0.2 (Misran et al. 2020). These patterns are not surprising given that several criteria may be assessed simultaneously when identifying bird species in flight. The limited results can therefore still give insightful information on the process of species identification of birdwatchers.

Based on the analyses, it can be concluded that criteria linked to the birds' appearance are more important for identifying species of birds in flight compared to their environment or flight behaviour. However, more experienced birdwatchers tended to observe both appearance features and flight behaviour

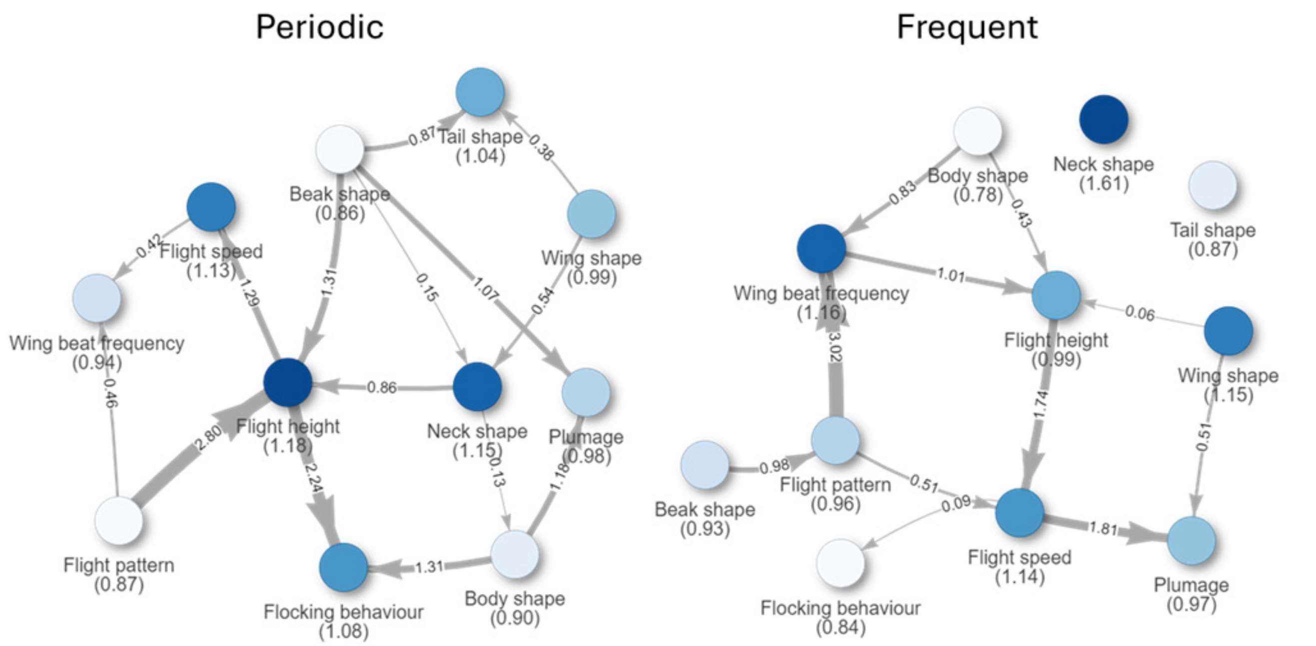

Figure 4. Reduced Bayesian Belief Network (reduced BBN) models predicting the connections among the Appearance and Flight characteristics criteria used by birdwatchers to identify species of birds in flight. Darker colours for nodes indicate higher individual contributions of criteria to the BBN model, where thickness of the arrows indicates the strength of the pairwise relationships. Birdwatchers were grouped as either Frequent = Active throughout the year, and at least once a week, or as Periodic = Remaining respondents.

simultaneously. The most important criteria considered were plumage, silhouette of body and wing shape, followed by flight pattern. The findings mean that machine learning algorithms of camera systems for species identification should focus on extracting and applying information from the detected birds' silhouette including wingspan-body length ratio, wing-aspect ratio, fine-grained specifics for body parts such as the beak, neck and tail, and details of the plumage, including colouration, colour placement, and patterns (Liu et al. 2007, Berg et al. 2014, Marini et al. 2013). In addition, tracking of bird movements to record flight behaviour can provide additional information for species identification. Given that nine out of ten Norwegian birds are migratory, the algorithms should preferably be developed separately for different seasons (cf. Berg et al. 2014) with regard to what bird species, with their associated characteristics, can be expected to be observed while birdwatching.

Several respondents noted that it was usually difficult to judge the alternative criteria since species identification of birds in flight is often carried out as an unconscious and mostly immediate process, based on a combined assessment of many different criteria that contribute to a general impression of size and shape, or so-called 'jizz', and that each assessment is often species-specific and can also vary with the sex and age of the birds (Lerner & Tunón 2012). There is also reason to believe that experience is important in the species identification process, where familiarity with a bird's silhouette and flight pattern can be useful for species identification. Support for a role of experience was found by the higher consistency and consensus for frequent birdwatchers. Also, the BBN model for frequent birdwatchers identified more linkages between criteria of the different categories. Experiments have shown that while detection and categorization happen nearly simultaneously ('it's a bird'), species identification requires more processing time ('it's a Black Woodpecker'), suggesting that identification occurs after the category has been determined (Grill-Spector & Kanwisher 2005). The steps imply that employing a hierarchical approach by first categorizing detected birds by general body shape and size as raptors, shorebirds, passerines or other groups, and thereafter identifying species in more detail within the most likely bird category could be a beneficial framework for development of new automated systems for identification of birds in flight.

Acknowledgements. The survey was part of a collaboration between the Norwegian Institute for Natural Research (NINA) and the company Spoor AS. The project was financed by the Regional Research Fund Oslo (RFFOSLO; grant 323094). Thanks to the late Jan Ove Gjershaug for comments on the draft manuscript.

REFERENCES

Almende BV, Thieurmel B & Robert T. 2019. visNetwork: Network Visualization using ,vis.js' Library. R package version 2.0.9. https://CRAN.R-project.org/package=visNetwork.

Berg T, Liu J, Lee SW, Alexander ML, Jacobs DW &

Belhumeur PN. 2014. Birdsnap: Large-scale fine-grained visual categorization of birds. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2014*: 2011–2018.

Chen SH & Pollino CA. 2012. Good practice in Bayesian network modelling. *Environmental Modeling Software* **37**: 134–145.

Cho F. 2019. ahpsurvey: Analytic Hierarchy Process for Survey Data. R package version 0.4.1. https://cran.r-project.org/package=ahpsurvey.

Datta D. 2017. blandr: a Bland-Altman Method Comparison package for R. R package version 0.6.0. https://CRAN.R-project.org/package=blandr.

Ferreira AC, Silva LR, Renna F, Brandl HB, Renoult JP, Farine DR, Covas R & Doutrelant C. 2020. Deep learning-based methods for individual recognition in small birds. *Methods in Ecology and Evolution* **11**: 1072–1085.

Goepel KD. 2018. Implementation of an online software tool for the analytic hierarchy process (AHP-OS). *International Journal of the Analytic Hierarchy Process* **10**: 469–487.

Grill-Spector K, & Kanwisher N. 2005. Visual recognition: as soon as you know it is there, you know what it is. *Psychological Science* **16**: 152–160.

Harker P. 1987. Incomplete pairwise comparisons in the analytic hierarchy process. *Mathematical Modelling* **9**: 837–848.

Lerner H. & Tunón H. 2012. What knowledge is 'jizz'? *Ornis Svecica* **22**: 73–79.

Liu Y, Zhang J, Tjondronegoro D & Geve S. 2007. A shape ontology framework for bird classification. *9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications, Glenelg, SA, Australia 2007*: 478–484.

Marini A, Facon, J & Koerich AL. 2013. Bird species classification based on color features. *IEEE International Conference on Systems, Man, and Cybernetics, Manchester, UK 2013*: 4336–4341.

McClure CJW, Martinson L, & Allison TD. 2018. Automated monitoring for birds in flight: Proof of concept with eagles at a wind power facility. *Biological Conservation* **224**: 26–33.

McClure CJW, Rolek BW, Dunn L, McCabe JD, Martinson L. & Katzner T. 2021. Eagle fatalities are reduced by automated curtailment of wind turbines. *Journal of Applied Ecology* **58**: 446–452.

Misran MFR, Roslin EN & Mohd Nur N. 2020. AHP-Consensus judgement on transitional decision-making: with a discussion on the relation towards open innovation. *Journal of Open Innovation: Technology, Market, and Complexity* **6**: 63.

Niemi J & Tanttu JT. 2018. Deep learning case study for automatic bird identification. *Applied Science* **8**: 2089.

Niemi J & Tanttu JT. 2020. Deep learning–based automatic bird identification system for offshore wind farms. *Wind Energy* **23**: 1394–1407.

R Core Team. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.r-project.org.

Scutari M. 2010. Learning Bayesian Networks with the bnlearn R Package. *Journal of Statistical Software* **35**: 1–22.

Thorpe S, Fize D & Marlot C. 1996. Speed of processing in the human visual system. *Nature* **381**: 520–522.

Ornis Norvegica (ISSN 1892-9737) is a peer-reviewed, online and open access journal publishing original papers in all fields of ornithology. The geographical focus is on Fennoscandia. Descriptive articles documenting bird biology or populations are welcome. Articles dealing with faunistical material should be analytical. Both subscription and publishing are free of charge.

Read articles, submit papers, and more information about Ornis Norvegica at:
https://boap.uib.no/index.php/ornis/index