(DRAFT 2025-07-29) Efficiency, correctness and usability of a sound monitoring tool in bird monitoring programs. Results from BirdWeather PUC and humans are compared

BY BO KAYSER

Introduction

At Danish bird observatories long term trends in the numbers of migrating birds have been followed for decades. Migrating birds have been counted at tips along the coast (Kayser & Pagh Jensen 2022) and caught for ringing using standardized monitoring practices (Kayser 2021).

It can be difficult to find volunteers to do the migration counts and the ringing, and the cost is quite high. As a supplement we would like to add automatic recording of bird sounds to our monitoring program. Can a tool like the BirdWeather PUC (BirdWeather) be used for that?

The BirdWeather PUC listening tool.

We wanted to find out how efficient and correct the BirdNET tools from Cornell Lab of Ornithology are at recording the bird sound in an area. To do that simultaneous recordings were made at the study site by humans standing beside and by the BirdWeather PUC. Furthermore, we wanted to know if the changes and trends in the number of recorded birds were the same when obtained by the PUC and by a human listener.

The paper presents results from this study.

Material and methods

The recordings of bird sounds were made in Denmark in the village Stensby (54,980°N, 12,049°Ø) in a big garden surrounded by fields and forests and about 2 km from the coast.

Recordings have been made each half-month from November 2024 to October 2025. Recordings were made with the BirdNET tool for 24 hours from midnight to midnight. Tool settings: Recording time: 9 sec. Probability: 5%. Confidence: 75%.

The human recordings were made from the hour before sunrise to the hour after sunset. Records were made for 10 minutes each hour. In each minute it was recorded which bird species were heard.

The efficiency of the recording made by the tool was found by comparing recordings made by the tool and the human listener. This was done by comparing minute by minute which species the tool and the human had recorded.

The correctness of the bird sound identifications made by the tool was made by listening to a series of 500 recordings made by the tool and comparing the bird species identified by the tool and the human. Such a comparison was made approximately every second month.

In real life we want to track changes in bird numbers from year to year. In this study it has only been possible for data collected by the tool and the human listener to compare changes from hour to hour and from half-month to half-month.

(Preliminary) Results

(Last updated 22 July 2025.)

The efficiency of the BirdWeather PUC

Table 1. Number of sound recordings made by a human listener and tool and percentage of number of recordings made by tool compared to the human listener. (Preliminary data from November 2024 to second half of July 2025.).

	Number of recordings		Percentage
	Human	Tool	Tool to human
Bird sound recordings (all species)	8021	1323	16

The study showed that the BirdWeather PUC recorded only about sixths of the bird sounds that a human standing beside the PUC did. It was mostly week sounds or remote birds that the BirdWeather PUC missed.

The precision of the identifications of bird species made by a BirdWeather PUC

The author listened to a series of 500 records made by the BirdWeather PUC and compared the identification of the species made by the PUC and the author. See table.

Table 2. The number of identifications made by the tool and the human listener that were equal, different and where the human listener was unsure.

Date	**Equal**	**Different**	**Unsure**	**Total**
01-12-2024	492	5	3	500
02-01-2025	487	5	8	500
16-02-2025	494	2	4	500
01-04-2025	487	11	2	500
03-06-2025	497	3	0	500

The precision of the tool was impressive. About 98 percent of the sounds were identified correctly. However, it does not mean that the tool does not make errors. E.g. the PUC identify most calls from Hooded Crow as Carrion Crow. Some songs of Wood-Pigeon were identified as Collared Dove.

Remote Herring Gulls were identified as Greater Bittern. The zipper in my coat was identified as Mute Swan etc. But when the tool is used to record changes in the number of common species analyzing many thousands of calls the few errors do not matter.

<u>Changes and trends recorded by the tool compared to changes and trends recorded by a human listener</u>

The results obtained from recordings made in each hour by a human listener in 10 minutes per hour, a BirdWeather PUC in the same 10 minutes per hour and a BirdWeather PUC for 24 hours have been compared. The total numbers of records are very different. Therefore, this has been done by comparing the percentage of records recorded in each hour by each type of listener. The same has been done for each half-month.

The changes from hour/half-month to hour/half-month and the trends over several hours/half-months are close to the same for the three types of listeners (figure 1 to 4). This applies both to all species taken as a whole and to the individual species.

The Chi-square test of goodness-of-fit has been used to test if the distribution and changes during the day/year are the same for data collected by the human listener for 10 minutes per hour, by a BirdWeather PUC for the same 10 minutes per hour or by a BirdWeather PUC for 24 hours. I.e. if the shapes of the curves are the same. The numbers recorded by the human were used as the expected numbers when testing the data collected by the PUC.

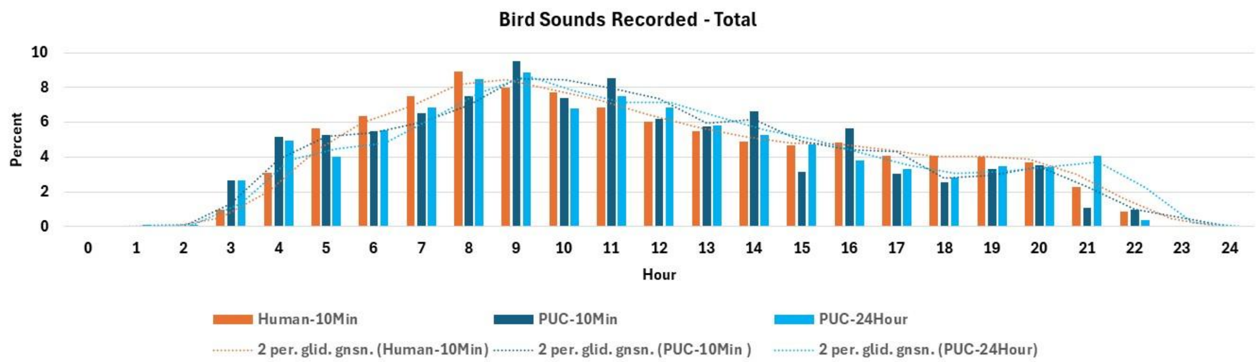

Figure 1. The percentage of recorded bird sounds of all species in each hour for human listener in10 minutes per hour (red), a BirdWeather PUC in the same 10 minutes per hour (dark blue) and a BirdWeather PUC in 24 hours (light blue). Running averages are shown as dotted lines.

Looking at data from all species, the shape of the curve was about the same for the “PUC-10 Minutes” (P = 0.984) and “PUC-24 Hours” (P=0.991) as for the human reference curve.

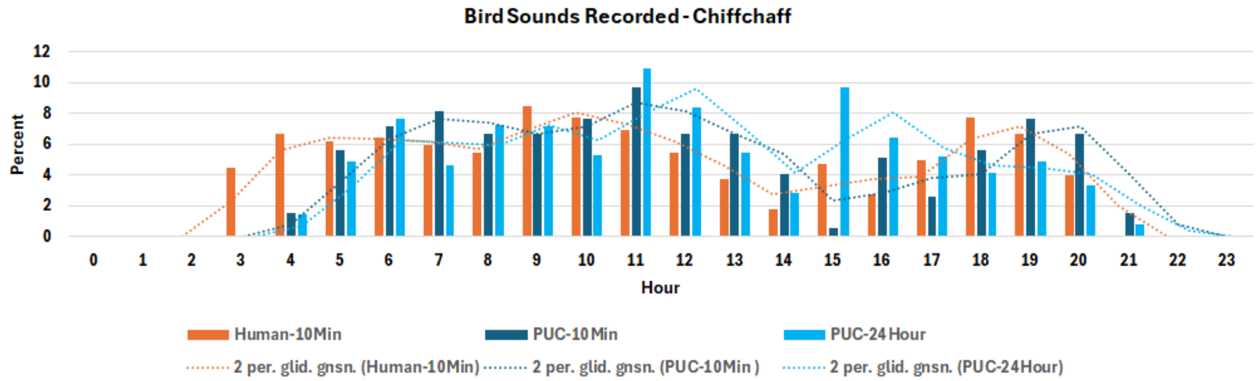

Figure 2. The percentage of recorded Chiffchaff sounds in each hour for a human listener in10 minutes per hour (red), a BirdWeather PUC in the same 10 minutes per hour (dark blue) and a BirdWeather PUC in 24 hours (light blue). Running averages are shown as dotted lines.

When looking at data from Chiffchaff, the shape of the curve was somewhat different for the "PUC-10 Minutes" (P=0.202) compared to the human reference curve. For the "PUC 24 Hours" it was close to significantly different (P=0,061) compared to the human reference curve.

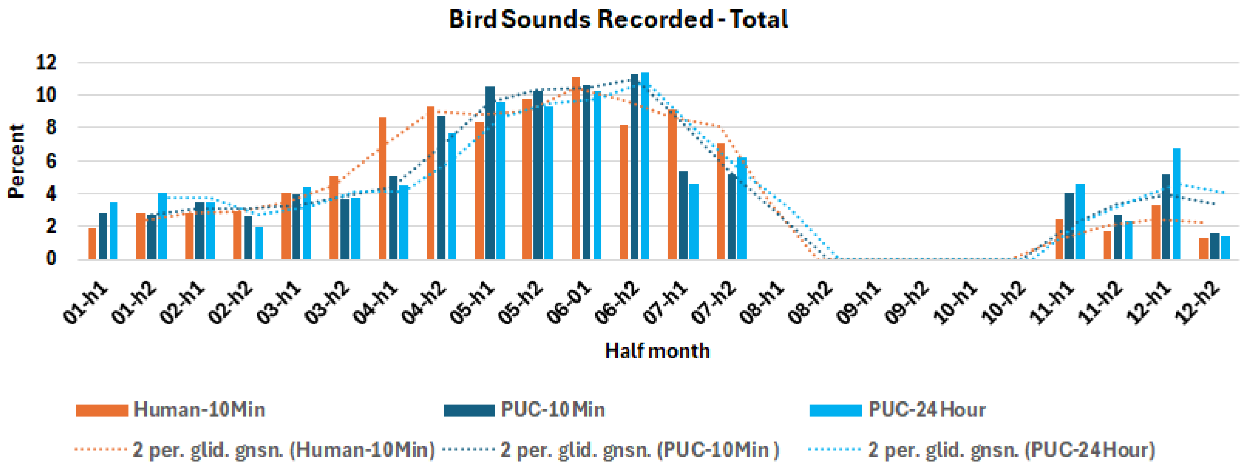

Figure 3. The percentage of recorded bird sounds of all species in each half-month for a human listener in10 minutes per hour (red), a BirdWeather PUC in the same 10 minutes per hour (dark blue) and a BirdWeather PUC in 24 hours (light blue). Running averages are shown as dotted lines.

Looking at data from all species, the shape of the curve was about similar for the "PUC-10 Minutes" (P = 0.935) compared to the human reference curve. For the "PUC 24 Hours" it was a little more different (P = 0.631) compared to the human reference curve.

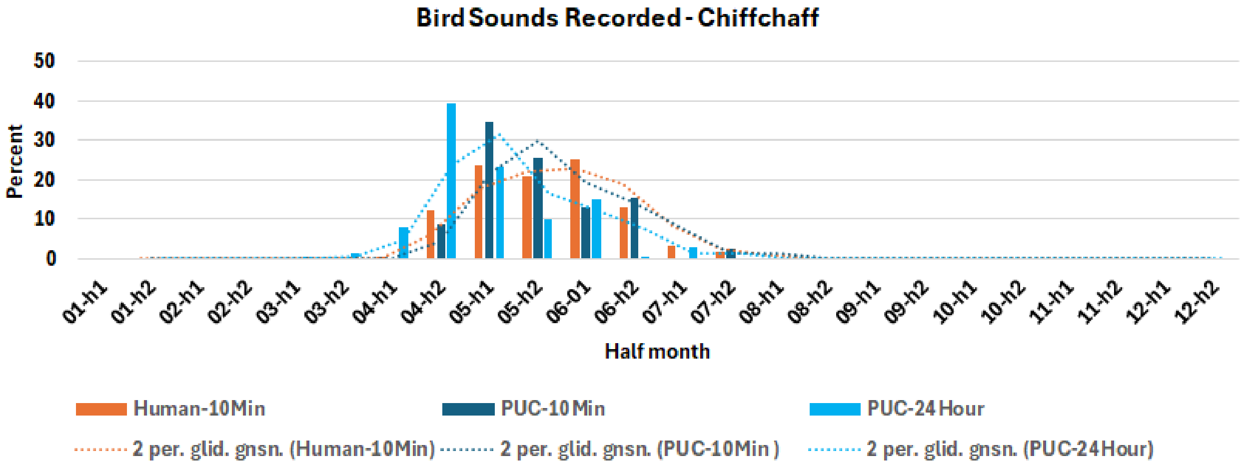

Figure 4. The percentage of recorded Chiffchaff sounds in each half month for a human listener in10 minutes per hour (red), a BirdWeather PUC in the same 10 minutes per hour (dark blue) and a BirdWeather PUC in 24 hours (light blue). Running averages are shown as dotted lines.

When looking at data from Chiffchaff, the shape of the curve was significantly different for the "PUC-10 Minutes" ($P = 0.015$) compared to the human reference curve. It was even more different for the "PUC 24 Hours" ($P < 0.001$) compared to the human reference curve.

<u>The number of listening intervals per hour/half-month affects the results</u>

The more listening intervals there are in each hour/half-month, the higher the maxima that can be recorded. With a person or a PUC that records in each of the 10 minutes of each hour, the number of recordings can be a maximum of 10 per hour. With a PUC that records in 9-second intervals, the number of recordings can be a maximum of 400 per hour.

This affects the shape of the activity curves, but it has not been found to have changed the location of the start time, maximum and end time of bird sound activity during the day/year.

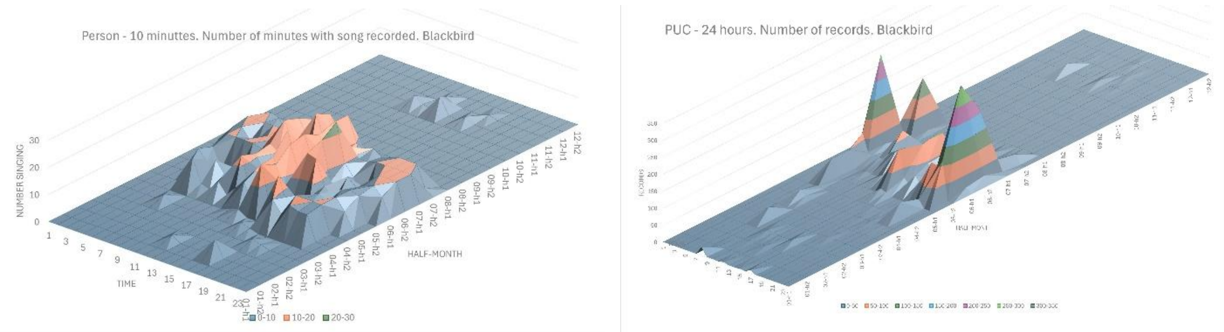

Figure 5. Diurnal and seasonal variation in the number of bird sounds recorded by a human (left) og BirdWeather PUC (right).

<u>Bird groups suitable for this monitoring tool</u>

Only bird species that make sounds regularly can be monitored using this tool. It can be vocal like call notes and song or mechanical like the drumming of woodpeckers or the wing beets of Mute Swans.

During this study it was possible to record species of the following groups: geese, some ducks species, swans, gamebirds, doves and pigeons, cranes, gulls, owls, woodpeckers and many species of passerines.

Discussion

The results of this study have shown that the BirdWeather PUC solution is suitable for monitoring changes in numbers and long-term trends for a wide range of bird groups ***(The results from the half-months still cover too few periods to give a realistic test)***. It is well suited for monitoring changes during the day and during the year too.

The tool only recorded about one-sixth of the bird sounds that a human listener recorded. This is not a problem, as the PUC is intended to be used to record changes and trends, not actual numbers. It is in the same range as other studies (Cole et al. 2022).

The BirdWeather PUC did identify a very large proportion of the bird sounds it "heard" correctly. In the checks conducted it was above 98 percent. It is higher than in other studies (Cole et al. 2022).

This does not mean that the PUC does not make incorrect identifications. But it does not matter when the purpose of using the tool is to follow changes in the number of individuals of common bird species and when many thousands of sound identifications are used.

The human listeners are not perfect either. Not all bird sounds that it is possible to hear are recorded. Human listeners are making misidentifications too. And human listeners, as well as the PUC tool, even record bird sounds that are not there.

One of the purposes of the study was to find out whether a BirdWeather PUC found the same changes in the number of birds and the same trends in long-term changes as a human listener did. This will be necessary if the tool is to be used as a supplement or to replace some of the monitoring work that bird observatories currently carry out manually (Kayser 2021, Kayser & Pagh Jensen 2022). The study showed that the PUC registers close to the same changes and trends as a human listener.

The study showed that the PUC is suitable for recording changes in numbers and trends in numbers for a wide range of bird groups that the bird observatories want to monitor. This applies to species that regularly emit sounds, either in the form of vocalizations or in the form of mechanical sounds. The conclusion is that a BirdWeather PUC is suitable for use as a tool for recording occurrence, changes in numbers and long-term trends for many bird species that the bird observatories want to monitor. However, of course there are also many species that the tool is not suitable to monitor.

Summary

(To be written.)

References

BirdWeather PUC home page: https://www.birdweather.com/

Cole, J. S., N. L. Michel, S.A. Emerson, and R. B. Siegel 2022. – AmericanOrnithology.org. 124 (2022): 1-15.

Kayser, A. 2021: Ændringer I fænologi, antal og ynglesucces blandt rastende småfugle ved Gedser 2007-19 *Trends in phenology and breeding success among staging passerine migrant at Gedser Bird Observatory 2007-2019)*. – Dansk Orn. Foren. Tidsskr. 115 (2021): 197-204.

Kayser, A. & F. Pagh Jensen 2022: Efterårstræk af havdykænder ved Gedser Odde 2009-20 *(Autumn migration of sea-ducks at Gedser Odde 2009-2020)*. – Dansk Orn. Foren. Tidsskr. 116 (2022): 9-16.

Marchal, J., F. Fabianek & Y. Aubry 2021. Software performance for the automated identification of bird vocalisations: the case of to closely related species. – BIOACOUSTICS https://doi.org/10.1080/09524622.2021.1945952.

Márquez-Rodrígues, A et al. 2025: A Bird Detector for improving bird identification through Deep Learning: a case study from Doñana. – arXiv:2503.15576v1 [cs.SD] 19 Mar 2025.

The author's address:

Bo Kayser, Pilevej 2, Stensby, 4773 Stensved, Denmark (bo.kayser@gmail.com)