Eur J Wildl Res
DOI 10.1007/s10344-017-1115-4

ORIGINAL ARTICLE

Reliability of human estimates of the presence of pups and the number of wolves vocalizing in chorus howls: implications for decision-making processes

Vicente Palacios [1,2] · **Enrique Font** [1] · **Emilio José García** [2] · **Linn Svensson** [3] · **Luis Llaneza** [2] · **Jens Frank** [3] · **José Vicente López-Bao** [3,4]

Received: 2 October 2016 / Revised: 12 May 2017 / Accepted: 18 May 2017

Abstract Management decision-making processes require reliable tools providing information on the distribution, abundance, and trend of populations. Wolves vocalize in response to human imitations of howls. Traditionally, this phenomenon has been the basis of a widespread monitoring tool to assess the reproductive status in a wolf pack, as well as to estimate the minimum number of individuals in the pack: the elicited-vocalization technique. However, despite its broad use, only a few attempts to quantify its accuracy have been made so far. Here, we carried out a test to evaluate the accuracy of estimates obtained from the elicited-vocalization technique. We administered "chorus tests" to 205 human subjects, 182 rangers—with different level of experience with wolves—and 23 subjects with no previous experience with the species. We found that the estimates of the number of wolves participating in a chorus were not accurate, regardless of the experience of the listener (the correct number of wolves was only determined in 32% of tests). Listeners, however, identified pups vocalizing 98% of the times when there were pups in the chorus. They also reported the presence of pups when they were not present with a high frequency (71%). Estimating the number of individuals by the unaided human ear is flawed because of the bias inherent in the elicited-vocalization technique. Howling surveys have a low degree of selectivity to confirm the presence of pups. Thus, we make recommendations to improve the elicited-vocalization technique as a tool to monitor the presence of pups.

Keywords *Canis lupus* · Wolf monitoring · Elicited-vocalization technique · Survey methods · Acoustic censusing · Chorus howls · Wolf howling

Introduction

Population monitoring plays a critical role in conservation and management (Sinclair et al. 2009). Assessing and detecting changes in local populations is crucial to understanding the temporal dynamics of populations, evaluating interventions in harvested or endangered populations, or verifying compliance with regulatory guidelines. Among different monitoring tools available, the information provided by acoustic signals emitted by animals has been used to monitor wildlife (Hopp and Morton 1998). Many animals emit vocalizations, which can be distinct at the individual, group, or species level, and can contain information regarding sex or age (Bradbury and Vehrencamp 2011). Our ability to recognize animal sounds has a long history of use in wildlife monitoring, and it is one of the most widely used methods to identify, for instance, specific identity, sex, and age during bird censuses (Gregory et al. 2004). Wolves (*Canis lupus*) howl in response to unfamiliar howls, and even to human imitations of howls, which is the basis of a widely used method for detecting animals: the elicited-vocalization technique (this technique has been named "simulated howling" when it is applied during wolf surveys; Harrington and Mech 1982). Simulated howling has been used in different parts of the world to monitor wolf

✉ Vicente Palacios
vicentep151@gmail.com

1 Instituto Cavanilles de Biodiversidad y Biología Evolutiva, Universidad de Valencia, Valencia, Spain

2 A.RE.NA. Asesores en Recursos Naturales, S.L. Perpetuo Socorro n°12-Entresuelo, 2B, 27003 Lugo, Spain

3 Grimsö Wildlife Research Station, Department of Ecology, Swedish University of Agricultural Sciences, Riddarhyttan, Sweden

4 Research Unit of Biodiversity (UO/CSIC/PA), Oviedo University, Mieres, Spain

Published online: 02 June 2017

populations during the breeding period (Harrington and Mech 1978, 1982; Dekker 1985; Gaines et al. 1995; Gazzola et al. 2002; Apollonio et al. 2004; Llaneza et al. 2005; Nowak et al. 2007; Chapron et al. 2014; Llaneza et al. 2014). By acoustic censusing, the personnel responsible for wolf monitoring listen to the chorus howls emitted by the wolves in response to human imitations of howls and estimate (aural estimate) how many wolves participate in the chorus (hereafter "chorus size") and whether there are pups vocalizing. Although the analysis of spectrograms of chorus howls has been proposed as a proper method to obtain information regarding chorus size (Filibeck et al. 1982; Dugnol et al. 2007; Passilongo et al. 2015) and the presence of pups (Palacios et al. 2016), aural estimates are still a common method for obtaining such information (Gazzola et al. 2002; Apollonio et al. 2004; Nowak et al. 2008; Mitchell et al. 2008; Llaneza et al. 2014). However, the reliability of aural estimates is poorly known. Only Passilongo et al. (2015) compared the aural field estimation of chorus size with the sonogram analysis of the same recordings to test for differences between both methods, but which factors affect the accuracy of these estimates, and the reliability of aural estimates of the presence of pups, is completely unknown. Lack of information on the magnitude of observation error makes it difficult for managers and researchers to use the data resulting from howling surveys.

Chorus howls have a complex acoustic structure as they include, in addition to howls, other vocalizations such as growls, barks, squeaks, and howl variations such as "woa-woa howls" (Schassburger 1993) (Fig. 1). As a consequence, discriminating the number of participants or the presence of pups in a chorus in situ and in real time, without the benefit of repeated listening to a recording, is challenging. In fact, Harrington (1989) reported that two or three adult wolves can produce highly modulated choruses that may give the erroneous impression that there are more wolves, and even pups, vocalizing. Hallberg (2007), as part of a study to assess whether acoustic cues for group size are present in the chorus yip-howl of coyotes (*Canis latrans*), played choruses to human subjects and asked them to estimate the number of animals involved. The human subjects had no relationship with coyote management or monitoring, and the study concluded that human estimates of coyote group size were frequently incorrect. However, Hallberg (2007) considered assessments of coyote chorus size fairly accurate because mean differences between actual chorus size and human estimations were within one or two individuals. In other species that also emit chorus vocalizations, such as laughing kookaburras (*Dacelo novaeguineae*), determining by ear the number of individuals participating in a chorus is considered an impossible task (Baker 2004).

In this study, we assess the accuracy of estimates made by human listeners in determining the number of wolves participating in chorus howls and whether pups were present or not. In particular, we compared the accuracy of estimates made by rangers, responsible for conducting wolf surveys, and by a group of human subjects with no previous experience with wolves. We discuss the implications of our results for wolf monitoring and decision-making processes.

Methods

Data collection

Between 2013 and 2014, we presented "chorus tests" to 205 human subjects and asked them to estimate the number of wolves participating in recorded wolf chorus howls, distinguishing between adults/subadults and pups. The human respondents included 182 rangers from Sweden and Spain (58 and 124, respectively) and 23 subjects with no previous experience with wolves or wolf vocalizations (general public). Both men and women were included (92% men and 8% women), aged 30–60 years and had no hearing problems. Informed consent was obtained from all individual participants included in the study.

Each chorus test consisted of listening to five chorus howls, selected randomly from a sample of 22 recordings of chorus howls, from which the number of individuals participating was known, emitted by 14 captive and wild packs (Table 1). We assumed that the fact that choruses were emitted by

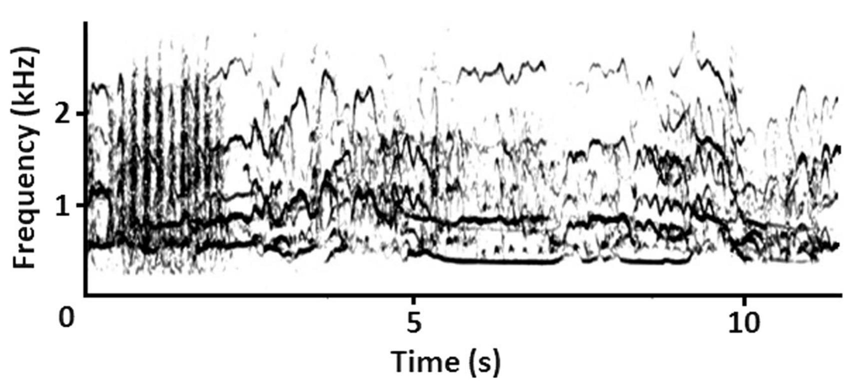

Fig. 1 Fragment of chorus howl emitted by five captive wolves including several types of vocalizations. In this fragment, at least four individuals emit vocalizations simultaneously (visual inspection at time 9 s)

Table 1 Wolves participating in choruses used for the tests. For each chorus, we considered two age classes: adults/subadults and pups

Chorus ID	Pack	Location	Composition (*N* adults/subadults + *N* pups)
1, 2, 6, 8	A	Madrid (captivity)	4 + 0
3, 9	B	Lisboa (captivity)	3 + 0
			2 + 0
4, 7	C	Lisboa (captivity)	5 + 0
			4 + 0
5	D	Madrid (captivity)	6 + 0
10, 21	E	Zamora (wild)	≥7 + 0
			≥2 + ≥1
11	F	Madrid (captivity	5 + 2
12	G	Lugo (wild)	0 + 3
13	H	A Coruña (wild)	≥1 + ≤3
14	I	Asturias (wild)	≥2 + 4
15	J	Zamora (wild)	≥2 + ≤5
16	K	Asturias (wild)	≥1 + ≤5
17, 18, 19	L	Zamora (wild)	≥2 + 7
20	M	Zamora (wild)	≥1 + ≥1
22	N	Asturias (wild)	0 + ≥3

≥: minimum confirmed number (we do not know the exact number of wolves howling nor the exact number composing the pack but we know the minimum number of individuals that participated in the chorus via spectrogram analysis or visual identification of wolves while howling). ≤: maximum possible number (we know the exact number in the pack but not all of them were seen during the chorus)

captive or wild wolves did not affect the listener's ability to estimate the presence of pups and the number of wolves howling. The sample of chorus howls included 10 choruses emitted only by adults/subadults, and 12 choruses including also pups. All the chorus howls included in the tests consisted of, at least two individuals vocalizing simultaneously, none of the recordings containing no howls. The definition of pups was individuals up to an age of 7 months, since it has been reported that at this age wolf pups already sound fully like adult wolves (Harrington and Asa 2003). All the chorus howls were emitted by Iberian wolves. We assume that including only Iberian wolves did not affect our design and results since Iberian wolf howls are similar than those emitted by other European wolf populations (Kershenbaum et al. 2016) and no differences in the fundamental frequency of howls emitted by Iberian wolves and other wolf populations have been reported (Palacios et al. 2007).

Rangers working in the same location were tested on the same day for logistic reasons. In this case, we provided each ranger with a test sheet that included five numbers (randomly selected) corresponding to the codes of the choruses to be analyzed. We played the choruses with an interval of 2–3 min of silence between consecutive choruses. During the silence intervals, rangers had to answer the questions in their sheet only when the choruses played corresponded to their assigned numbers. All tests were conducted under supervision of at least one of the co-authors of this research. The choruses were played using a computer and computer speakers (frequency range 20–20,000 Hz) at a volume ensuring that the chorus was audible in the entire room. A small proportion of tests (15%) were made with participants isolated from other participants.

For each chorus, the humans were asked to estimate the number of both adult/subadults and pups vocalizing. Due to the acoustic structure of chorus howls included in the tests (several individuals vocalizing simultaneously), the participants were expected to estimate the number of animals deconvoluting overlapping vocalizations. Presence/absence of adults/subadults and pups was obtained from the estimated numbers (0 = absence, ≥1 = presence). We quantified the accuracy of the subject's answers assigning a score (0–4) to each chorus in a test according to the answers to the different questions asked (see Table 2 for details of the scoring procedure). We asked for the number of times the listener had heard chorus howls in the field to be used as an index of experience. Subjects were grouped into three categories according to their experience: 0, the listener had never heard wolves howling in the field; 1, the listener had heard <10 times chorus howls; and 2, the listener had heard >10 times chorus howls before conducting the tests.

Table 2 Questions that human subjects had to answer for each chorus in the tests, variables measured, and score assigned for each question

Question	Variable	Value	Score
Number of adults	Presence of adults	0: absent	Wrong: 0
		≥1: present	Correct: 1
	Number of adults	Numeric	Wrong (under/overestimated): 0
			Correct: 1
Number of pups	Presence of pups	0: absent	Wrong: 0
		≥1: present	Correct: 1
	Number of pups	Numeric	Wrong (under/overestimated): 0
			Correct: 1
Total score			0–4

Data analysis

We used generalized linear mixed models (GLMMs) to assess which factors affected the accuracy of the estimates made by the human subjects. We built GLMMs models with binomial error distribution and logit link for the estimates of the presence of pups (PUPS-model; binary response variable *pupest*: accuracy of the estimates of the presence of pups; values: correct/wrong) and the number of wolves participating (NUMBER-model; binary response variable *numest*: accuracy of the estimates of the number of wolves howling; values: correct/wrong). Country (Sweden/Spain), experience of the human subject (three levels), and subject group (inexperienced/rangers) were included as predictors in both models. In addition, we included if the chorus effectively included pups as a predictor in the PUPS-model, and the real number of wolves participating as a predictor in the NUMBER-model. For the NUMBER-model, we included in the analyses only choruses with exact known composition (Table 1). The identities of chorus and listener were included as random factors in the models. We conducted a complementary GLMM with Poisson error distribution and logit link to test for the influence of country, experience, and subject group on the scores obtained in a chorus and a GLMM with binomial error distribution and logit link to asses which factors affected the type of error in the estimates of the number of wolves participating in a chorus (overestimate vs. underestimate).

We built a set of competing GLMMs considering all the possible combinations using the selected variables in each dataset (including the null model, i.e., the intercept-only model). We used Akaike's information criterion (AIC) to rank models, selecting the model with the lowest AIC values (Burnham and Anderson 2002). AIC weights were calculated to evaluate the strength of each model. We estimated the variance explained by the best model calculating marginal R^2 (variance explained by fixed factors) and conditional R^2 (variance explained by both fixed and random factors, Nakagawa and Schielzeth 2013). GLMMs were conducted in R (R Development Core Team 2010), using the *lme4* package (Bates and Sarkar 2007).

Results

Our dataset comprised 1023 assessments of wolf chorus howls (answers in two cases were considered invalid), 910 and 113 by rangers and inexperienced subjects, respectively. The average score (±SD) obtained in a chorus (values from 0 to 4) was 2.1 ± 1.04. We found no evidence that the country (GLMM: $z = -0.64$, $N = 1011$, $P = 0.53$), experience of the listeners (GLMM: $z = -0.25$, $N = 1011$, $P = 0.8$) or subject group (GLMM: $z = 0.44$, $N = 1011$, $P = 0.66$), influenced the score obtained in a chorus.

Presence of pups in a chorus

Overall, our dataset included chorus howls with no pups and chorus howls with pups vocalizing in a similar proportion (47 and 53%, respectively). Assessment of the presence of pups was correct in 65% of the times (Table 3). The country of origin of human subjects and whether they were rangers or inexperienced listeners had no influence on the correct determination of the presence of pups in a chorus (Table 4). The best PUPS-model included the presence of pups vocalizing in the chorus and the respondent's experience and explained 81% of the variance in our dataset. Chorus and human subject random factors explained 20% of such variance (marginal $R^2 = 0.61$; conditional $R^2 = 0.81$; Table 4). The presence of the pups in a chorus had a significant effect on the accuracy of the estimates (GLMM: $z = 6.633$, $\beta = 6.4511$, SE = 0.97, $P < 0.001$). Importantly, estimates were more accurate when there were pups vocalizing in the chorus. Respondents identified pups 98% of times when there were effectively pups vocalizing (Table 3). In contrast, respondents wrongly determined that there were pups 71% of times when there were no pups vocalizing (false positives; Table 3, Fig. 2). The accuracy of the estimates was slightly better as experience increased (Fig. 3), although the effect of the experience was not statistically significant (GLMM: $z = 1.63$, $N = 1017$, $P = 0.103$). The respondent's experience only affected the proportion of correct estimates when there were no pups vocalizing, but even the most experienced listeners wrongly estimated the presence of pups almost 50% of the times when there were no pups vocalizing (Fig. 3).

Table 3 Estimates made by human subjects of the presence/absence of pups vzocalizing (highlighted, 1017 valid estimates) and the number of wolves participating (only cases with exact number of wolves participating known, 554 valid estimates) compared to the real values

		Estimated value	
		No pups	Pups
	No pups	138	342
	Pups	13	524
		Correct estimate	Wrong estimate
Real value	2	27	21
	3	60	76
	4	64	129
	5	14	39
	6	9	26
	7	2	87

Table 4 GLMMs obtained for estimates of pup's presence and number of wolves howling in a chorus

PUPS-models	df	AICc	Δ_iAICc	w_i
Experience + pups presence	5	488.76	0	0.28
Experience + pups presence + country	6	490.43	1.67	0.12
Experience + pups presence + country + subject group	7	492.25	3.49	0.05
Null model	3	517.92	29.16	0
NUMBER-models	df	AICc	ΔiAICc	w_i
Real number	4	636.51	0	0.35
Real number + country	5	638.30	1.79	0.14
Real number + country + ranger/inexp	6	640.22	3.71	0.06
Real number + country + ranger/inexp + experience	7	641.94	5.43	0.02
Null model	3	655.36	18.85	0

The best models considering the AIC criterion are in italic

df number of parameters in the model, *AICc* Akaike's information criterion, *Δ_iAICc* delta AIC value, w_i Akaike's weight

Number of wolves participating in a chorus

The number of wolves participating in a chorus howl was correctly estimated in 32% of the cases only (Table 3). With respect to the age classes, the listeners correctly estimated the number of pups and adults in 24 and 19% of cases, respectively. The overall number of wolves estimated by the respondents was underestimated and overestimated 41 and 27% of times, respectively. Errors in the estimates about the number of wolves participating in a chorus were on average (±SD) 1.84 ± 1.22 (range −5 to +9 individuals). Most of the errors differed from the real number of wolves in 1, 2, or 3 individuals (55, 22.5, and 12.7%, respectively).

Country, experience, and subject's group did not affect the estimates of the number of wolves participating in a chorus (Table 4). Only the real number of wolves affected the probability to estimate the number of participants correctly: as the number of wolves participating in a chorus howl increased, the proportion of correct estimates decreased (GLMM: $z = -7.105$, $\beta = -0.558$, SE = 0.079, $P < 0.001$; Fig. 4). The best model included the number of wolves howling and explained 18% of the variance, (marginal R^2 = conditional R^2 = 0.18). A similar pattern was found when considering whether the number of wolves was overestimated or underestimated. An alternative GLMM using the type of error (overestimate vs. underestimate) as dependent variable included the real number of wolves as the

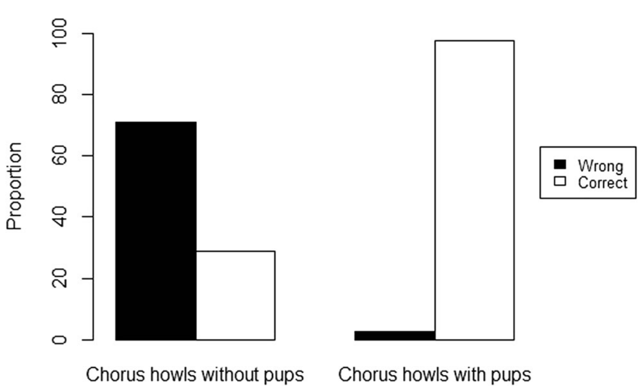

Fig. 2 Proportion of correct (*white bars*) and wrong (*black bars*) estimates of the presence of pups vocalizing in a chorus, depending on whether there actually were pups howling or not. False positives: the respondent estimates that there are pups when there are no pups vocalizing (*black bar* in chorus howls without pups)

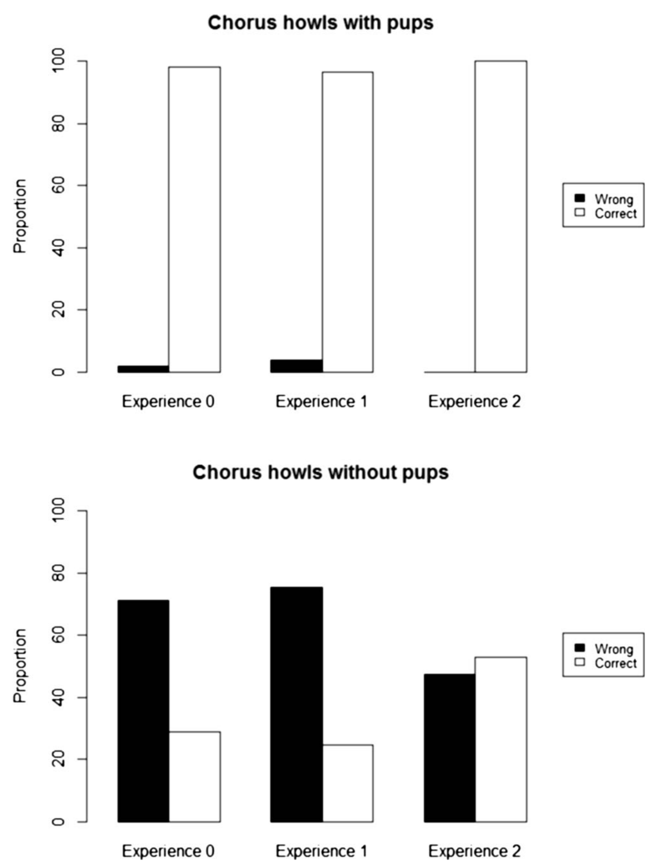

Fig. 3 Proportion of correct (*white bars*) and wrong (*black bars*) estimates of the presence of pups vocalizing in a chorus depending on the level of experience (0: the listener had never heard wolves howling in the field, 1: the listener had heard <10 times chorus howls, and 2: the listener had heard >10 times chorus howls before conducting the tests)

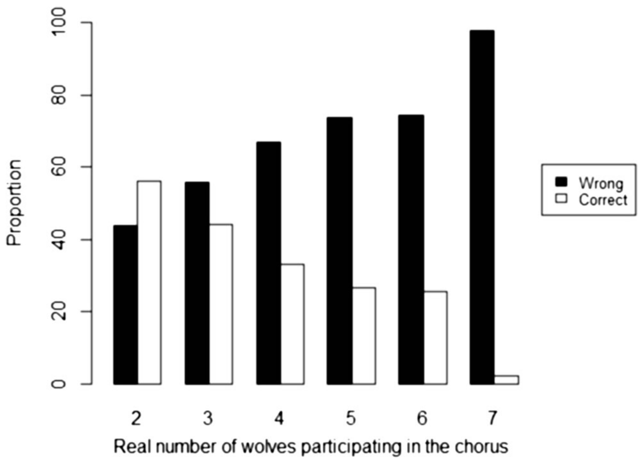

Fig. 4 Proportion of correct (*white bars*) and wrong (*black bars*) estimates of the number of wolves vocalizing in chorus howls depending on the real number of wolves participating

only significant effect (GLMM: $z = 3.042$, $\beta = 1.926$, SE = 0.63, $P = 0.002$). Estimates for choruses emitted by a small number of wolves tended to be overestimated, while listeners underestimated the number of wolves howling when group sizes were larger (Fig. 5).

Discussion

Identifying limitations and the most common misuses of monitoring tools is required to properly analyze and interpret data collected during wildlife surveys (Boitani and Fuller 2000). Our results provide the first evaluation of the accuracy of estimates of both chorus size and the presence of pups made by means of acoustic censusing of wolves during simulated howling surveys without the benefit of spectrogram inspection, but see Passilongo et al. (2015) for chorus size estimates. Our main findings are as follows: (i) howling surveys have a low degree of selectivity, i.e., listeners correctly detect the presence of pups when pups are present and howling but they often report hearing pups when there are no pups present in the vocalizations; (ii) chorus size estimates are correct in only about a third of cases, the magnitude and the sign of the error depending on the real number of wolves vocalizing, evidencing the limitations of the acoustic censusing to estimate the number of wolves; and (iii) the experience of the listener has little influence on the accuracy of the estimates.

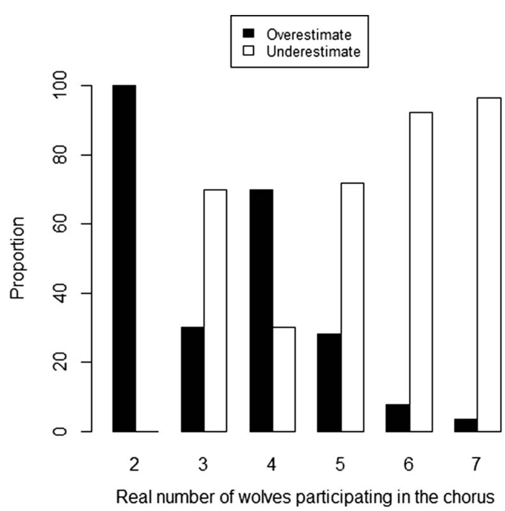

Fig. 5 Proportion of over (*black bars*) and underestimates (*white bars*) of the number of wolves vocalizing depending on the real number of individuals participating in chorus howls

With respect to the ability to determine the presence of pups, it is commonly assumed that pup/adult discrimination by ear is feasible, at least when the pups are young (Apollonio et al. 2004; Llaneza et al. 2005). However, we found that one third of the estimates of the presence of pups were actually false positives. Wolf pups can produce the full range of vocalizations used by adults early in ontogeny, between 3 and 9 weeks of age (Peters and Wozencraft 1989; Harrington and Asa 2003). Therefore, the only difference between the vocalizations emitted by pups and adults in chorus howls are related to the fundamental frequency (higher in pups) and, in the case of howls, duration (longer in adults; Harrington and Asa 2003). Chorus howls can be very complex depending on the number of wolves participating and the vocalizations emitted. Although the human ear can discriminate changes of approximately 3 Hz in a 1000-Hz tone (Sinnott and Brown 1993), the complexity of chorus howls (which depends on group composition and type of vocalizations included) may place constraints on the capacity of the human ear for pup/adult discrimination.

Joslin (1967) pioneered the use of acoustic censusing from chorus howls to obtain information regarding chorus size, counting each animal as it first began to howl. However, no systematic data were collected to determine the reliability of this method. Harrington (1989) reported that human estimates of chorus size from chorus howls overestimate the real size, especially for smaller packs. Although Harrington did not conduct statistical analyses, our results confirm this assumption since overestimates of chorus size are frequent for small chorus sizes. Passilongo et al. (2015) reported that aural estimations of chorus size and estimations made by analyzing spectrograms (which are highly correlated to the real chorus size) showed low correlation. Only 32% of estimations were identical in both methodologies. As in Passilongo et al. (2015), in our study, only 32% of the chorus size estimates were correct. Hallberg (2007) carried out a study to evaluate estimates of chorus size from coyote group-yip-howl choruses and concluded that human subjects were usually wrong in their estimates of chorus size, although their estimates were close to the actual number, which should not necessarily be surprising as she included choruses emitted only by 2–5 individuals. We obtained the same percentage of correct estimates for wolves as Hallberg (2007) did for coyotes, but with a much bigger sample size (205 vs. 21 respondents). This agreement between Passilongo et al. (2015) and our study with wolves and Hallberg (2007) with coyotes may reflect the ability of humans to detect several voices independently of their experience and similarities in the acoustic structure of wolf and coyote chorus howls (Lehner 1978).

Rangers are usually involved in wildlife monitoring worldwide. For example, rangers participate in mountain gorillas (*Gorilla beringei*) censuses and the monitoring of herbivores in Africa (Koenig 2008; Valeix et al. 2009) or in large carnivore monitoring in different European countries, such as Norway, Sweden, Poland, or Spain, including wolf surveys (Kaczensky et al. 2013; Kuijper et al. 2013; Llaneza et al. 2014). Experience and training of personnel responsible for surveys has been identified as an important factor that could affect the accuracy of information obtained in monitoring systems (Bibby et al. 1992; Sutherland 2006). However, in our case, we found that experience has little effect in the accuracy of the estimates made by human listeners. We did not find differences among rangers and inexperienced people, and having heard many wolf chorus howls previously only slightly increased the score obtained in the tests, suggesting that the inaccuracy of estimates may be due to the complexity of the vocalizations or to limitations in human auditory capacity rather than prior experience with wolves. Nevertheless, we also found that experience in hearing chorus howls tended to decrease, to some extent, the proportion of wrong estimates of the presence of pups when there are no pups howling. These results might be useful for designing training programs that stress the importance that experience in listening to chorus howls may have to reduce the proportion of wrong estimates of the presence of pups.

These shortcomings in the assessment of chorus howls by the unaided human ear should be taken into account by wolf monitoring systems that rely on acoustic censusing. Chorus size estimates obtained by listening to chorus howls are usually wrong, especially as chorus size increases. In most cases, differences between estimates and actual chorus size involve 1–2 individuals. However, 22.5% of times the respondents overestimated actual chorus size by more than two individuals. Detection of pups yielded better results than chorus size estimates, but false positives in the detection of pups were very frequent. In terms of monitoring wolf populations, this is a mistake with undesirable consequences, because determining the presence of pups implies the existence of reproductive activity in a pack. For example, two choruses emitted by the same pack few kilometers apart could be considered as two different packs by documenting presence of pups in both choruses.

Our results underscore the need for improved monitoring methods that allow objective and reliable estimates of chorus size and the presence of pups from chorus howls. Several recommendations can be made in order to improve estimates from howling surveys. When possible, we recommend to record the chorus howls to carry out posterior spectrogram analyses. This technique is objective and has been successfully applied to obtain information regarding the number of wolves participating in choruses (Passilongo et al. 2015), but it can be also useful to identify the presence of pups in a chorus (Palacios et al. 2016). In fact, conducting spectrogram analyses, Passilongo et al. (2015) obtained the correct chorus size in 92% of 29 chorus howls (much better than the 32% of correct aural estimates), and the analysis of the acoustic energy distribution could reduce the 71% false positives obtained by aural estimates to 3.9–6.5% (Palacios et al. 2016). However, as it is not always worthwhile to record the chorus howls emitted by wolves during howling surveys, complementary recommendations are needed. First, we urge to replicate howling surveys in order to increase the accuracy in the estimate of pup presence in a given site (Jiménez et al. 2016). Thus, once a potential rendezvous site with pups will be detected, we suggest to replicate howling surveys even though the listener considers that pups were already present in the first session. This procedure would reduce wrong estimates of pups' presence when there are no pups in a given site (Jiménez et al. 2016). Second, to establish spatial and temporal rules during howling surveys, such as to survey in the same night neighboring areas where it is not clear whether there could be one or two different packs. Finally, in sites where it is suspected that there are pups but howling surveys are not conclusive even after several replicates, combining howling surveys with other tools, such as camera trapping (Galaverni et al. 2012), might be useful to determine the presence of pups.

Acknowledgements We are in debt to the rangers of the Regional Government of Galicia and Picos de Europa National Park in Spain, as well as the rangers of the different counties with wolf presence in Sweden. We thank all volunteers collaborating in this study and conducting the "chorus test." With respect to the chorus howls recordings used in this research, we thank Junta de Castilla y León, Principado de Asturias, and Xunta de Galicia for allowing us to record wild wolves and F. Petrucci-Fonseca and S. Pinho (CRLI), J. Escudero (La Dehesa), P. España and E. Sánchez (Cañada Real), and P. Abad (Carpín) for allowing us to obtain recordings from the captive wolves under their care. JVLB was supported by a *Juan de la Cierva* research contract (JCI-2012-13066) from the Spanish Ministry of Economy and Competitiveness. This is scientific paper no. 17 of the Iberian Wolf Research Team (IWRT).

Compliance with ethical standards

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. All applicable international, national, and/or institutional guidelines for the care and use of animals were followed.

References

Apollonio M, Mattioli L, Scandura M et al (2004) Wolves in the Casentinesi Forests: insights for wolf conservation in Italy from a protected area with a rich wild prey community. Biol Conserv 120: 249–260

Baker MC (2004) The chorus song of cooperatively breeding laughing kookaburras (Coraciiformes, Halcyonidae: *Dacelo novaeguineae*): characterization and comparison among groups. Ethology 110:21–35

Bates DM, Sarkar D (2007) The lme4 library. [Online], Available: http://lib.stat.cmu.edu/R/CRAN/

Bibby CJ, Burgess ND, Hill DA (1992) Bird census techniques. Harcourt Brace & Company, London

Boitani L, Fuller TK (2000) Research techniques in animal ecology. Controversies and consequences. Columbia University Press, New York

Bradbury JW, Vehrencamp SL (2011) Principles of animal communication, 2nd edn. Sinnauer Associates Publishers, Sutherland

Burnham KP, Anderson DR (2002) Model selection and multimodel inference: a practical information theoretic approach. Springer-Verlag, New York

Chapron G, Kaczensky P, Linnell JDC et al (2014) Recovery of large carnivores in Europe's modern human-dominated landscapes. Science 346(6216):1517–1519

Dekker D (1985) Response of wolves, *Canis lupus*, to simulated howling on a homesite during fall and winter in Jasper National Park, Alberta. Can field-naturalist 99:90–93

Dugnol B, Fernández C, Galiano G, Fernández C (2007) Wolf population counting by spectrogram image processing. Appl Math Comput 186:820–830

Filibeck U, Nicoli M, Rossi P, Boscagli G (1982) Detection by frequency analyzer of individual wolves howling in a chorus: a preliminary report. Boll di Zool 49:151–154

Gaines WL, Neale GK, Naney RH (1995) Response of coyotes and gray wolves to simulated howling in North-Central Washington. Northwest Sci 69:217–222

Galaverni M, Palumbo D, Fabbri E et al (2012) Monitoring wolves (*Canis lupus*) by non-invasive genetics and camera trapping: a small-scale pilot study. Eur J Wildl Res 58:47–58

Gazzola A, Avanzinelli E, Mauri L et al (2002) Temporal changes of howling in south European wolf packs. Ital J Zool 69:157–161

Gregory RD, Gibbons DW, Donald PF (2004) Bird census and survey techniques. In: Sutherland WJ, Newton I, Green RE (eds) Bird ecology and conservation; a handbook of techniques. Oxford University Press, Oxford, pp 17–56

Hallberg KI (2007) Information in a long-distance vocal signal: chorus howling in the coyote (*Canis latrans*). The Ohio State University, Columbus

Harrington FH (1989) Chorus howling by wolves: acoustic structure, pack size and the beau geste effect. Bioacoustics-The Int J Anim Sound its Rec 2:117–136

Harrington FH, Asa CS (2003) Wolf communication. In: Mech LD, Boitani L (eds) Wolves: behavior, ecology, and conservation. University of Chicago Press, Chicago, pp 66–103

Harrington FH, Mech LD (1978) Howling at two Minnesota wolf pack summer homesites. Can J Zool 56:2024–2028

Harrington FH, Mech LD (1982) An analysis of howling response parameters useful for wolf pack censusing. J Wildl Manag 46:686–693

Hopp SL, Morton ES (1998) Sound playback studies. In: Hopp SL, Owren MJ, Evans CS (eds) Animal acoustic communication: sound analysis and research methods. Springer-Verlag, Berlin Heidelberg, pp 323–352

Jiménez J, García EJ, Llaneza L et al (2016) Multimethod, multistate Bayesian hierarchical modeling approach for use in regional monitoring of wolves. Conserv Biol:1–23

Joslin PWB (1967) Movements and home sites of timber wolves in Algonquin Park. Am Zool 7:279–288

Kaczensky P, Chapron G, von Arx M et al (2013) Status, management and distribution of large carnivores—bear, lynx, wolf & wolverine—in Europe. Istituto di Ecologia Applicata, Rome

Kershenbaum A, Root-Gutteridge H, Habib B et al (2016) Disentangling canid howls across multiple species and subspecies: structure in a complex communication channel. Behav Process 124:149–157

Koenig R (2008) Rangers assess toll of Congo conflict on threatened mountain gorillas. Science 322(5909):1778

Kuijper DPJ, de Kleine C, Churski M et al (2013) Landscape of fear in Europe: wolves affect spatial patterns of ungulate browsing in Białowieża Primeval Forest, Poland. Ecography (Cop) 36:1263–1275

Lehner PN (1978) Coyote vocalizations: a lexicon and comparisons with other canids. Anim Behav 26:712–722

Llaneza L, Ordiz A, Palacios V, Uzal A (2005) Monitoring wolf populations using howling points combined with sign survey transects. Wildl Biol Pract 1:108–117

Llaneza L, García EJ, López-Bao JV (2014) Intensity of territorial marking predicts wolf reproduction: implications for wolf monitoring. PLoS One 9:e93015

Mitchell MS, Ausband DE, Sime CA et al (2008) Estimation of successful breeding pairs for wolves in the Northern Rocky Mountains, USA. J Wildl Manag 72:881–891

Nakagawa S, Schielzeth H (2013) A general and simple method for obtaining R^2 from generalized linear mixed-effects models. Methods Ecol Evol 4:133–142

Nowak S, Jedrzejewski W, Schmidt K et al (2007) Howling activity of free-ranging wolves (*Canis lupus*) in the Białowieza Primeval Forest and the Western Beskidy Mountains. J Ethol 25:231–237

Nowak S, Myslajek RW, Jedrzejewska B et al (2008) Density and demography of wolf, *Canis lupus* population in the western-most part of the polish Carpathian Mountains, 1996-2003. Folia Zool 57:392–402

Palacios V, Font E, Márquez R (2007) Iberian wolf howls: acoustic structure, individual variation, and a comparison with North American populations. J Mammal 88:606–613

Palacios V, López-Bao JV, Llaneza L et al (2016) Decoding group vocalizations: the acoustic energy distribution of chorus howls is useful to determine wolf reproduction. PLoS One 11:e0153858

Passilongo D, Mattioli L, Bassi E et al (2015) Visualizing sound: counting wolves by using a spectral view of the chorus howling. Front Zool 12:22

Peters G, Wozencraft WC (1989) Acoustic communication by fissiped carnivores. In: Gittleman J (ed) Carnivore behavior, ecology, and evolution. New York University Press, Ithaca, pp 14–49

R Development Core Team (2010) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Schassburger RM (1993) Vocal communication in the timber wolf, *Canis lupus*, Linnaeus: structure, motivation, and ontogeny. Advances in Ethology Series 30

Sinclair ARE, Fryxell JM, Caughley G (2009) Wildlife ecology, conservation and management, second. Wiley, Oxford

Sinnott JM, Brown CH (1993) Effects of varying signal duration on pure-tone frequency discrimination in humans and monkeys. J Acoust Soc Am 93:1541–1546

Sutherland WJ (2006) Ecological census techniques. A handbook, 2nd edn. Cambridge University Press, Cambridge

Valeix M, Loveridge AJ, Chamaillé-Jammes S et al (2009) Behavioral adjustments of African herbivores to predation risk by lions: spatiotemporal variations influence habitat use. Ecology 90:23–30