

 Open access • Journal Article • DOI:10.1111/2041-210X.13198

Accounting for automated identification errors in acoustic surveys — [Source link](#)

[Kévin Barré](#), [Isabelle Le Viol](#), [Romain Julliard](#), [Julie Pauwels](#) ...+7 more authors

Institutions: [University of Paris](#), [British Trust for Ornithology](#), [University of Greifswald](#), [University of Montpellier](#)

Published on: 01 Aug 2019 - [Methods in Ecology and Evolution](#) (John Wiley & Sons, Ltd)

Topics: [False positive paradox](#) and [Identification \(information\)](#)

Related papers:

- [Bats as bioindicators – the need of a standardized method for acoustic bat activity surveys](#)
- [Disentangling the relative effect of light pollution, impervious surfaces and intensive agriculture on bat activity with a national-scale monitoring program](#)
- [Bat activity in intensively farmed landscapes with wind turbines and offset measures](#)
- [Tadarida: A Toolbox for Animal Detection on Acoustic Recordings](#)
- [Potential of bat pass duration measures for studies of bat activity](#)

Share this paper:    

View more about this paper here: <https://typeset.io/papers/accounting-for-automated-identification-errors-in-acoustic-3v2oekygz3>



HAL
open science

Accounting for automated identification errors in acoustic surveys

Kévin Barré, Isabelle Le Viol, Romain Julliard, Julie Pauwels, Stuart Newson,
Jean-François Julien, Fabien Claireau, Christian Kerbiriou, Yves Bas

► **To cite this version:**

Kévin Barré, Isabelle Le Viol, Romain Julliard, Julie Pauwels, Stuart Newson, et al.. Accounting for automated identification errors in acoustic surveys. *Methods in Ecology and Evolution*, Wiley, 2019, 10 (8), pp.1171-1188. 10.1111/2041-210X.13198 . hal-02914597

HAL Id: hal-02914597

<https://hal.archives-ouvertes.fr/hal-02914597>

Submitted on 12 Aug 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Accounting for automated identification errors in acoustic surveys**

2

3 Kévin Barré^{1,2}, Isabelle Le Viol^{1,2}, Romain Julliard¹, Julie Pauwels¹, Stuart E. Newson³,

4 Jean-François Julien¹, Fabien Claireau^{1,4,5}, Christian Kerbiriou^{1,2*}, Yves Bas^{1,6*}

5

6 ¹ Centre d'Ecologie et des Sciences de la Conservation (CESCO), Muséum national d'Histoire
7 naturelle, Centre National de la Recherche Scientifique, Sorbonne Université, CP 135, 57 rue
8 Cuvier, 75005 Paris, France

9 ² Centre d'Ecologie et des Sciences de la Conservation (CESCO), Muséum national d'Histoire
10 naturelle, Station de Biologie Marine, 1 place de la Croix, 29900 Concarneau, France

11 ³ British Trust for Ornithology, The Nunnery, Thetford, Norfolk IP24 2PU, UK

12 ⁴ University of Greifswald, Zoology Institute and Museum, Soldmann-Str. 14, D – 17 489
13 Greifswald, Germany.

14 ⁵ Naturalia Environnement, Site Agroparc, Rue Lawrence Durell, 84 911 Avignon, France.

15 ⁶ Centre d'Ecologie Fonctionnelle et Evolutive (CEFE), UMR 5175, CNRS – Université de
16 Montpellier – Université Paul-Valéry Montpellier – EPHE, 1919 route de Mende, 34293
17 Montpellier, France

18 * Equal contribution as last authors

19

20 Corresponding author:

21 Kévin Barré

22 kevin.barre@edu.mnhn.fr

23 +33 2 98 50 99 28

24

25 Running title: Accounting for acoustic identification errors

26 **Abstract**

27 **1.** Assessing the state and trend of biodiversity in the face of anthropogenic threats requires
28 large-scale and long-time monitoring, for which new recording methods offer interesting
29 possibilities. Reduced costs and a huge increase in storage capacity of acoustic recorders has
30 resulted in an exponential use of Passive Acoustic Monitoring (PAM) on a wide range of
31 animal groups in recent years. PAM has led to a rapid growth in the quantity of acoustic data,
32 making manual identification increasingly time-consuming. Therefore, software detecting
33 sound events, extracting numerous features, and automatically identifying species have been
34 developed. However, automated identification generates identification errors, which could
35 influence analyses which looks at the ecological response of species. Taking the case of bats
36 for which PAM constitutes an efficient tool, we propose a cautious method to account for
37 errors in acoustic identifications of any taxa without excessive manual checking of recordings.

38 **2.** We propose to check a representative sample of the outputs of a software commonly used
39 in acoustic surveys (Tadarida), to model the identification success probability of 10 species
40 and 2 species groups as a function of the confidence score provided for each automated
41 identification. Using this relationship, we then investigated the effect of setting different False
42 Positive Tolerances (FPTs), from a 50% to 10% false positive rate, above which data are
43 discarded, by repeating a large-scale analysis of bat response to environmental variables and
44 checking for consistency in the results.

45 **3.** Considering estimates, standard errors and significance of species response to
46 environmental variables, the main changes occurred between the naive (i.e. raw data) and
47 robust analyses (i.e. using FPTs). Responses were highly stable between FPTs.

48 **4.** We conclude it was essential to, at least, remove data above 50% FPT to minimize false
49 positives. We recommend systematically checking the consistency of responses for at least
50 two contrasting FPTs (e.g. 50% and 10%), in order to ensure robustness, and only going on to

51 conclusive interpretation when these are consistent. This study provides a huge saving of time
52 for manual checking, which will facilitate the improvement of large-scale monitoring, and
53 ultimately our understanding of ecological responses.

54

55 Key words: bioacoustic, cautious threshold, Chiroptera, error rate, false-positives, Passive
56 Acoustic Monitoring, semi-automated identification, survey methods

57

58 **Introduction**

59

60 With few exceptions, the rate of biodiversity loss does not appear to be slowing down
61 (Butchart et al., 2010). In 2010, the 10th Conference of Parties to the Convention on
62 Biological Diversity adopted a new 2011–2020 global Strategic Plan for Biodiversity, and in
63 turn, the European Union launched a new Biodiversity Strategy (2011/2307). This strategy
64 aims to halt biodiversity loss and the degradation of ecosystem services by 2020. Such
65 objectives require large-scale and long-time studies using adapted monitoring methods for
66 surveying and understanding biodiversity changes (Fisher, Frank, & Leggett, 2010) in
67 response to anthropogenic pressures and environmental policies. The implementation of such
68 studies is highly constrained by the time and cost induced. Interestingly, the development of
69 new recording methods, such as Passive Acoustic Monitoring (PAM), offers interesting
70 possibilities and are taking an increasing place in monitoring (Gibb, Browning, Glover-
71 Kapfer, & Jones, 2018).

72 The reduced costs of acoustic recorders and the huge increase in storage capacity has resulted
73 in an exponential increase in the use of PAM on a very wide range of species groups within a
74 few years (e.g. Stahlschmidt & Brühl, 2012; Froidevaux, Zellweger, Bollmann, & Obrist,
75 2014; Kalan et al., 2015; Jeliaskov et al., 2016; Nowacek, Christiansen, Bejder, Goldbogen, &
76 Friedlaender, 2016; Frommolt, 2017). Such approaches are already widely used by
77 researchers as well as by people working for environmental consultancies and government
78 agencies for various biodiversity evaluation (Adams, Jantzen, Hamilton, & Fenton, 2012).
79 PAM can be particularly useful to carry out surveys on cryptic taxa such as nocturnal fauna
80 (Delport, Kemp, & Ferguson, 2002; Newson, Evans, & Gillings, 2015; Jeliaskov et al., 2016),
81 and to monitor pristine habitats which are otherwise difficult to access and survey by other
82 approaches (Gasc, Sueur, Pavoine, Pellens, & Grandcolas, 2013). PAM is also mobilized in

83 citizen science programs, for which it is an efficient tool for the implementation of large-scale
84 biodiversity monitoring (Newson et al., 2015; Jeliaskov et al., 2016; Kerbiriou, Azam, et al.,
85 2018; Penone, Kerbiriou, Julien, Marmet, & Le Viol, 2018).

86 Despite rapid and exciting developments in acoustic monitoring, there have been substantial
87 challenges in developing this technology into a cost-effective, scalable monitoring tool.

88 Perhaps the biggest and most complex issue facing acoustic monitoring has been the objective
89 and statistical taxonomic identification of bioacoustic signals. With the arrival on the market
90 of a new generation of affordable acoustic recorders, which allow for continuous recordings
91 over several days, such volumes of acoustic data cannot be processed manually (Newson et
92 al., 2015; Bas, Bas, & Julien, 2017).

93 In parallel to the development of PAM, several methods for detecting sound events, extracting
94 numerous features, and automatically identifying species have been developed (Parsons &
95 Jones, 2000; Britzke, Duchamp, Murray, Swihart, & Robbins, 2011; Adams et al., 2012; Bas
96 et al., 2017; Ovaskainen, Moliterno de Camargo, & Somervuo, 2018). However, automated
97 identification software have been criticised due to significant error rates, suggesting cautious
98 and limited use (Russo & Voigt, 2016; Rydell, Nyman, Eklöf, Jones, & Russo, 2017), which
99 heavily reduces the advantages of automated algorithms. Nonetheless, authors have
100 highlighted the potential for combining automated classifiers with manual validation to help
101 overcome error risks associated with automated identification, and so saving a huge amount
102 of work in reducing the extent of manual checking required (López-Baucells et al., 2019).

103 Moreover, most available software provides confidence scores associated with each
104 automated identification in the form of probabilities or other numerical indexes (Obrist,
105 Boesch, & Fluckiger, 2004; Waters & Barlow, 2013), which unlike the error rate is not
106 dependent of the relative abundance of the species. The confidence scores provided by
107 software aim to be an indicator of the true success probabilities of automated identifications,

108 and are strongly species-dependant. There is thus an implicit relationship between the error
109 rate and confidence scores and most software manuals advocate using confidence thresholds
110 below which data should be discarded to minimise the error rate, e.g. Tadarida (Bas et al.,
111 2017), SonoChiro (Biotope, 2013) and BatClassify (Scott & Altringham, 2017). Regardless of
112 the software used, the relationship between the error rate and confidence scores is an
113 important part of the automated identification performance, yet it has never been directly
114 assessed in previous methodological studies (Fritsch & Bruckner 2014; Rydell et al. 2017).
115 Consequently, the level at which confidence thresholds should be set is unclear to most users,
116 which has limited the use of automated identification in ecological studies. A threshold that is
117 too cautious could lead to high generated false negative rates (i.e. by discarding a large
118 proportion of data containing true positives below a given confidence score), which could
119 result in a lack of statistical power. In contrast, a threshold that is not cautious enough could
120 lead to high false positive rates (i.e. fails in automated identifications), particularly through
121 the inclusion of records of species which are most similar acoustically, which involve
122 statistical noise. Moreover, errors (generated false negative rates or false positive rates) could
123 also be spatially clustered by environmental conditions that alter the quality of the signal
124 (Denzinger & Schnitzler, 2013), which potentially induce statistical biases in relation with
125 confidence measure provided by the software. False positive rates and generated false
126 negative rates thus induce different caveats for which there is not a unique way to set
127 confidence thresholds. Given the wide range of taxa for which PAM is increasingly being
128 used, there is a need to account for these caveats using a method generalizable to any
129 acoustically surveyed taxa.

130 In this study we propose a method for assessing the effect of using confidence thresholds in
131 acoustic automated identification on the detection of species responses to environmental
132 variables. This method can be applied to any acoustic taxa for which automated identification

133 software and acoustic signature knowledges are already developed, and where confidence
134 scores are provided. Taking the case of bats, we first manually checked a representative
135 sample of a large number of bat recordings identified using an automated identification
136 software (Tadarida; Bas et al., 2017) commonly used in bat studies (Barré, Le Viol, Julliard,
137 Chiron, & Kerbiriou, 2017; Barré, Le Viol, Bas, Julliard, & Kerbiriou, 2018; Claireau et al.,
138 2018; Pinaud, Claireau, Leuchtman, & Kerbiriou, 2018; Pauwels et al., 2019). Using this
139 sample, we then modelled the identification success for 10 species and 2 species groups of
140 bats in relation to the confidence score provided by the software. This allowed us to define the
141 minimum confidence score needed to ensure a given False Positive Tolerance (FPT). We then
142 examined how setting different FPTs, from 50% to 10% maximum false positive rate, above
143 which data are discarded, may affect a statistical inference by repeating a large-scale analysis
144 of the response of species and species groups activity to five environmental variables, and
145 looking at consistency of the results among FPTs.

146

147 **Materials and methods**

148

149 *Bat survey*

150 We used an acoustic dataset collected previously to study the effect of wind turbines on bat
151 activity (Barré et al., 2018) because it was based on a random sampling design with high
152 variability and no confounding effects in terms of environmental variables (Fig. S1). The
153 following environmental variables are known as good predictors of bat activity: type of site
154 i.e., hedgerow vs. open area habitat located at an average of 86 m (Standard Deviation: 70 m)
155 away from any hedgerow (Verboom & Huitema, 1997; Lacoeuilhe, Machon, Julien, &
156 Kerbirou, 2016), the distance in meters to a forest (mean=700, SD=506; Boughey et al. 2011;
157 Frey-Ehrenbold et al. 2013), the distance to an urban area (mean=335, SD=170; Azam et al.
158 2016), the distance to a wetland (mean=579, SD=363; Sirami et al. 2013; Santos et al. 2013)
159 and the total length of hedgerows in meters within a 1000 m radius (mean=3439, SD=1622;
160 Verboom & Huitema, 1997; Lacoeuilhe et al., 2016). The latter four variables presented
161 important environmental variability, and a similar gradient between sites located close to
162 hedgerows and those in open areas (Fig. S1).

163 Bats were recorded at 337 sites (one complete night per site, with 207 sites close to
164 hedgerows and 130 sites in open area) in northwest France (Fig. 1) dominated by agriculture
165 (82%) and forest (11%) areas. Recordings were carried out over 23 complete nights, recording
166 from 30 minutes before sunset until 30 minutes after sunrise, from the 7th of September to the
167 8th of October 2016.

168 We simultaneously sampled 11-15 survey sites per night separated by at least 300 m (Fig.1).
169 Echolocation calls were recorded using one automatic acoustic recorder per site survey (Song
170 Meter SM2Bat+, Wildlife Acoustics Inc., Concord, MA, USA). The detectors automatically
171 recorded all ultrasounds using predefined settings as recommended by the French bat

172 monitoring program “Vigie-Chiro” (trigger level set to 6 dB Signal Noise Ratio and set to
173 continue recording until 2.0 seconds after last trigger event, 384 kHz sampling rate; for
174 further details see Azam et al., 2018; Barré et al., 2018; Claireau et al., 2018; Pauwels et al.,
175 2019). Whilst continuous recording is typically used for monitoring birds and several other
176 species groups, for bats which echolocate at high frequency, and so produce heavy sound
177 files, it is necessary to use triggered recording, to be able to manage and store the data and
178 process the recordings. In addition, these trigger settings are very sensitive (6 dB of signal-to-
179 noise ratio) and detect the majority of bats which would have been detected if recording were
180 continuous. As recommended by Millon et al. (2015) and Kerbiriou et al. (2018), we retained
181 one bat pass per five-second interval, which is the mean duration of all bat species passes.

182

183 *Step 1: manual checking of a subset of the data*

184 The identification process performed in the first step was divided in two sub-steps (Fig. 1). In
185 the first sub-step, echolocation calls were detected and classified to the closest taxonomic
186 level using the Tadarida software (Bas et al., 2017) (hereafter named primary identification),
187 which assigns a species and confidence score (continuous values between 0 and 1) to each
188 recorded bat pass (212 347 in total). In the second sub-step, we selected a representative
189 sample by a stratified random sampling of 25 primary identifications for each 0.1 class of
190 confidence score (i.e. 10 classes in total) for each species and groups for manual checking,
191 except for *Rhinolophus* species for which all identifications were selected due to their low
192 number. We performed a double manual checking (KB and YB) on this stratified random
193 selection of 1 910 bat passes (hereafter names checked dataset or manual checking), using
194 BatSound© software (Pettersson Elektronik AB, Sweden) and Syrinx software (John Burt,
195 Seattle, WA, USA) for 10 species and two groups (*Myotis* spp. and *Plecotus* spp.) (Table 1),
196 by visual inspection and measurement of discriminating characteristics of calls on

197 spectrograms (Barataud, 2015). Species groups were used for genera within which species are
198 difficult to identify from one another, except for one species of *Myotis* spp., *Myotis nattereri*,
199 for which echolocation calls are very characteristic (Obrist et al., 2004; Barataud, 2015). We
200 made the choice to separate two species which are commonly grouped because of their
201 frequency overlap: *Pipistrellus kuhlii* and *Pipistrellus nathusii*. We manually separated these
202 species by combining measurements of energy peak, final frequency, call duration, bandwidth
203 and time between calls as discussed in Barataud (2015). In relatively open habitats like in our
204 study, *P. nathusii* emit very commonly very short bandwidth, and higher frequencies than *P.*
205 *kuhlii* when emitting such kind of calls (i.e. quasi-constant frequency). *P. kuhlii* very often use
206 a short frequency modulation at the end of the call and this is very rare in *P. nathusii* calls.
207 Finally, we randomly checked 500 sound files identified as not containing bats to assess
208 missed bat events.

209 We assumed that manual checking provided the most conservative species assignments,
210 which allowed us to accurately assign to each primary identification a true positive (i.e. a
211 correct automated identification of the species), a false positive (i.e. a fail in automated
212 identification of the species) or a false negative (i.e. defined in this study as a pass of the
213 species automatically identified as another one) in the checked dataset.

214 The efficiency of the automated identification may be spatially heterogeneous due to habitat
215 structure (Denzinger & Schnitzler, 2013). We tested for the dependence of false positives (i.e.
216 a binomial response variable: failure or success of the automated identification) and false
217 negative ones (i.e. a binomial response variable: automatically identified as another species or
218 correct identification) on the five tested environmental variables. We performed generalized
219 linear mixed models (binomial response variables; *logit* link) with the environmental variables
220 as explanatory variables, using date as random effect to control for inter-night variations.

221

222 *Step 2: false positive rate modelling*

223 The success probability, defined as the success or failure of the automated species
224 identification, was used as the response variable to perform generalized linear models
225 (binomial response variable; *logit* link) using the confidence score provided by the automated
226 identification software as the explanatory variable (see step 2 in Fig.1; Fig. 2). Using these
227 models, we could predict the confidence score corresponding to a given success probability of
228 the automated identification. Thus, predicted confidence score constitutes the minimum one
229 required to ensure a given False Positive Tolerance (FPT, i.e. one minus the success
230 probability) in the whole dataset (i.e. including all checked and non-checked primary
231 identifications; Fig.1; Table 2). We selected all FPTs starting from the highest acceptable one
232 (0.5, i.e. a maximum false positive rate of 50%, which expected to give an approximately
233 balanced number of false negatives and false positives) to the lower one (0.1, i.e. a maximum
234 false positive rate of 10%) by 0.1 classes (i.e. 0.5, 0.4, 0.3, 0.2 and 0.1 FPTs).

235

236 *Step 3: data thresholding and consistency of model outputs regarding false positive rate*

237 After predicting the required confidence score to ensure a given FPT in the automated
238 identification, we filtered the whole dataset on the five predicted confidence scores
239 corresponding to the five FPT (see step 3 in Fig.1; Fig. 3; Table 2). This allowed us to
240 calculate for each FPT in the whole dataset, the remaining number of bat passes, occurrences,
241 and an estimation of false positive rate and generated false negative rate by reducing the FPT
242 (Table 2). In order to assess the trade-off between false positive rates and generated false
243 negative rates generated by reducing FPT, for each FPT, we estimated for the whole dataset
244 the false positive rate (i.e. incorrect primary identifications) and generated false negative rate
245 (i.e. as a consequence of discarding true positives because of reducing FPT) from equations
246 used to model the false positive rate in step 2. For each bat pass *BP* of a given species *S*, we

247 first computed the probability of there being a true positive (TP, equation 1) and a false
248 positive (FP, equation 2) as follows:

$$249 \quad TP_{BP,S} = \frac{\exp(ax+b)}{1 + \exp(ax+b)} \quad \text{Eq. 1}$$

$$250 \quad FP_{BP,S} = 1 - TP_{BP,S} \quad \text{Eq. 2}$$

252 where a corresponds to the estimated parameter from the logistic regression between manual
253 checking (i.e. the response variable: success/fail in automated identification; step 2 in Fig. 1,
254 Fig. 2) with the confidence score provided by the software (i.e. the explanatory variable), x is
255 the confidence score of the bat pass provided by the automated identification software and b is
256 the intercept of the logistic regression (Fig. S2).

257 This allowed us to estimate the generated false negative rate (FNR, equation 3) for a given
258 species S and a given threshold of false positive tolerance FPT in the whole dataset, by
259 averaging all probabilities to have a true positive TP from bat passes BP discarded by
260 reducing FPT (i.e. between the targeted FPT and the maximum FPT of 1) as follows:

$$261 \quad FNR_{S, FPT} = \frac{\sum_{FPT=1} TP_{BP,S}}{n_{BP,S}} \quad \text{Eq. 3}$$

262 where n is the total number of bat passes BP of the species S .

263 We were also able to estimate the false positive rate (FPR, equation 4) for a given species S
264 and a given threshold of false positive tolerance FPT in the whole dataset, by averaging
265 probabilities to have a false positive FP from bat passes BP between the minimum FPT (i.e.
266 zero) and the targeted FPT as follows:

$$267 \quad FPR_{S, FPT} = \frac{\sum_{FPT=0} FP_{BP,S}}{n_{BP,S,FPT}} \quad \text{Eq. 4}$$

268 where n is the number of bat passes BP between the minimum FPT (i.e. zero tolerance of false
269 positives) and the targeted FPT of a given species S .

270 Finally, we evaluated the automated classification efficiency by drawing Receiver Operating
271 Characteristic (ROC) curves between confidence scores of presences and absences of each
272 species, and computing Area Under Curve (AUC) with the R package *PRROC* (Fig. S3).
273 For each species and species groups, we then performed Generalized Linear Mixed Models
274 (GLMM, R package *lme4*) using as a response variable the number of bat passes filtered on
275 one of the five FPTs or the raw number of primary identifications without thresholding (i.e.
276 whole dataset) (six GLMMs in total performed on 0.5,0.4, 0.3, 0.2, 0.1 FPTs and on the whole
277 dataset). Environmental variables were included as fixed effects, among which quantitative
278 ones were scaled. According to the sampling design (i.e. 11-15 simultaneous recording sites
279 per night), we included date as a random effect to control for inter-night variation in weather
280 conditions and landscape context. We applied a Poisson error or a Negative binomial error
281 distribution to GLMMs in order to minimize issues in the overdispersion ratio in models (i.e.
282 as close as possible to 1; Zuur, Ieno, Walker, Saveliev, & Smith, 2009). All explanatory
283 variables showed a Variance Inflation Factor value under 1.5, meaning there was no strong
284 evidence of multicollinearity (Chatterjee & Hadi, 2006).
285 We then compared the estimates of each environmental variable among fitted models to check
286 the consistency in the response of bats to environmental variables in relation to the different
287 FPTs.
288

289 **Results**

290

291 *Automated identification and manual checking*

292 Over the 23 nights sampled, among the 212 347 bat passes recorded, 167 504 (79%) were
293 assigned to *Pipistrellus pipistrellus*, 28 589 (13 %) to *Pipistrellus kuhlii*, 6 430 (3%) to *Myotis*
294 spp. and 5 835 (3%) to *Barbastella barbastellus* (Table 1). A stratified random sample of
295 1 910 bat passes were manually checked (Table 1). False positive rates varied a lot among
296 species, from 0.0% for *Rhinolophus ferrumequinum* to 69.4% for *Nyctalus noctula* (Table 1).
297 The largest number of errors detected in manual checks was for *N. noctula* confused with
298 social calls of *P. pipistrellus* (only one location involved) and non-bat noises, and. calls of
299 *Pipistrellus nathusii* were confused with *P. kuhlii*, *P. pipistrellus* and non-bat noises (Table
300 S1). Concerning the random checking of 500 sound files identified as non-bat by the software,
301 we found that three (0.6%) contained bat events.

302

303 *Checking for environmental biases in identification errors*

304 Using the dataset on which manual checks were carried out, we investigated a potential
305 variation in automated identification errors due the environmental variables. The probability
306 of these being false positives was significantly affected by only one environmental variable
307 (habitat type of survey sites: hedgerow vs. open area) and for one species only, *N. noctula* (P
308 < 0.001 ; Table S2). All other environmental variables were not found to affect the probability
309 of there being false negatives for any species (Table S3).

310

311 *False positive rate modelling*

312 Success and failure in automated identification assessed through manual checking were
313 modelled in relation to the confidence score provided by the software, allowing us to predict

314 the required confidence score to ensure a given FPT (Fig. 2). Confidence scores required to
315 ensure FPTs (i.e. 0.5, 0.4, 0.3, 0.2 and 0.1) did not vary much for species such as *B.*
316 *barbastellus* (0.12-0.20), *Eptesicus serotinus* (0.18-0.29) and *Rhinolophus hipposideros* (0.39-
317 0.45), but more for others, e.g. *Nyctalus leisleri* (0.29-0.59), *P. kuhlii* (0.16-0.44) and *Plecotus*
318 *ssp.* (0.18-0.36) (Table 2). In addition, these FPTs confidence scores were low for *B.*
319 *barbastellus*, *E. serotinus*, *P. kuhlii*, *Plecotus spp.*, *Myotis spp.*, and higher for *P. nathusii* and
320 *N. noctula* (Table 2).

321 For *P. pipistrellus* errors were rare thus the lowest possible confidence score (0.096)
322 corresponded to a FPT lower than 0.2. In contrast, for *P. nathusii*, the highest possible
323 confidence score (0.971) corresponded to a FPT greater than 0.1, i.e. more than one in ten
324 chance of failure (Table 2). Moreover, no errors were found in the sample for *R.*
325 *ferrumequinum*, which prevented the modelling of error rate for this species (Table 2).

326 Low FPTs (i.e. removing data below a high confidence score) often led to an important
327 decrease in activity measures (Table 2). For example, *Myotis spp.* and *N. leisleri* activity
328 decreased by 27.8 % and 82.1 %, respectively, between 0.5 FPT and 0.1 FPT (Table 2).

329 However, such high decreases in activity resulted in a little decrease in occurrence for these
330 species: 6.7% for the *Myotis spp.* group and 10.7% for *N. leisleri* (Table 2). For other species,
331 the activity and occurrence were more stable across FPTs, including for *B. barbastellus*, *E.*
332 *serotinus*, *P. kuhlii*, *Plecotus spp.* and *R. hipposideros* (Table 2).

333 At the highest FPT (0.5), the estimated false positive rate was high (>21%) for three species
334 (*N. leisleri*, *N. noctula* and *P. nathusii*), and very low (<5%) for six species (*B. barbastellus*,
335 *E. serotinus*, *P. kuhlii*, *P. pipistrellus*, *R. ferrumequinum* and *R. hipposideros*) (Table 2).

336 However, at the lowest FPT (0.1), all species showed an estimated false positive rate under
337 0.05, except for *N. leisleri* (0.08) and *P. nathusii* for which no data satisfied a FPT lower than
338 0.1 (Table 2).

339 Estimating the generated false negative rate (i.e. true positives discarded by reducing the FPT)
340 was very low (<4%) at 0.5 FPT for most species except *N. leisleri* (0.19) and *P. nathusii*
341 (0.28) (Table 2). This rate became more important at 0.1 FPT, with null values for *P.*
342 *pipistrellus* and *R. ferrumequinum*; with very low values (<10%) for five species (*B.*
343 *barbastellus*, *E. serotinus*, *N. noctula*, *P. kuhlii*, and *R. hipposideros*); and with high values
344 for *N. leisleri* (0.425) and *P. nathusii* (0.377) (Table 2). The average AUC from ROC curves
345 was 0.93 (range: 0.73-1.00; Fig. S3).

346

347 *Consistency of activity patterns across error rate tolerance gradient*

348 To study the influence of confidence score thresholding according to FPTs below which data
349 were discarded (i.e. changes in amount of data, species occurrence, estimated false positive
350 rate and estimated rate of generated false negative), modelling of the bat response (i.e. the
351 number of bat passes according to selected FPT) to environmental variables was performed at
352 all FPTs.

353 When comparing model outputs from naive (i.e. raw data) to robust analyses (i.e. FPTs), a
354 loss or a gain of significance occurred for the open areas vs. hedgerows variable for *N.*
355 *leisleri*, the distance to forest for *Myotis* spp. and *N. leisleri*, the length of hedgerows for *N.*
356 *leisleri* and the distance to urban areas for *N. noctula* (Table 3). In addition, for significant
357 variables, an inversion of the direction of the estimate for the open areas vs. hedgerows
358 variable occurred for *N. noctula* and *P. nathusii* (Table 3). In all other cases, no changes were
359 found (Table 3).

360 However, we did not detect any major changes in model outputs between the 0.5, 0.4, 0.3, 0.2
361 and 0.1 FPTs for which response estimates and standard errors remained highly stable (Table
362 3). In only two cases, we detected a loss of significance: for *N. noctula* with FPTs lower than
363 0.2 and 0.3 for the distance to forests and the length of hedgerows variables, respectively

364 (Table 3). However, for this species, the open areas vs. hedgerows variable remained
365 significant and highly stable at all FPTs (Table 3).

366 All species had at least one significant habitat variable response irrespective of the used FPTs,
367 except *N. leisleri*. Hedgerows had a significantly higher bat activity (i.e. number of bat
368 passes) associated with them than open areas for seven species or groups (*B. barbastellus*, *M.*
369 *nattereri*, *Myotis* spp., *P. kuhlii*, *P. pipistrellus*, *Plecotus* spp. and *R. hipposideros*) and a
370 significantly lower bat activity for two species (*N. noctula* and *P. nathusii*) (Table 3). We also
371 found a significant negative relationship between bat activity and i) the distance to urban
372 areas variable for two species or groups (*E. serotinus* and *Plecotus* spp.; Table 3); ii) with the
373 distance to forest variable for two species (*N. noctula* and *R. ferrumequinum*; Table 3); iii)
374 with the distance to wetlands variable for *R. ferrumequinum*; and iv) with the length of
375 hedgerows variable for *N. noctula*, *P. pipistrellus* and *R. ferrumequinum* (Table 3) but a
376 significant positive relationship with the distance to wetlands variable for *P. kuhlii* (Table 3).
377

378 **Discussion**

379

380 This study demonstrates that automated acoustic identification of bats, as well as by extension
381 all other taxa acoustically identifiable by software, coupled with partial manual checking and
382 false positive rate modelling (i.e. semi-automated identification; Newson et al. 2015), is a key
383 tool for improving reliability of studies based on acoustic data. Indeed, robust ecological
384 responses could be produced even in cases where false positive rates were so far considered
385 too high (Rydell et al. 2017). This new and robust framework takes advantage of confidence
386 scores provided by the automated identification software and its ability for distinguishing true
387 positives and false positives (Fig. S3), controlling for False Positive Tolerances (FPTs), and
388 checking for potential biases induced by identification errors.

389

390 *Using confidence thresholding*

391 Minimum confidence scores required to ensure a given FPT according to species exhibited
392 low to moderate variation across the 0.5 to 0.1 FPTs (Table 2). To investigate the effect of the
393 automated identification errors on bat activity patterns in relation with FPTs, we studied the
394 response of bat activity to several environmental variables known to impact bats. Depending
395 on species, the most significant responses to environmental variables were consistent with
396 known patterns of bat activity: a negative effect of open areas vs. hedgerows and of
397 decreasing length of hedgerows (Verboom & Huitema, 1997; Lacoeuilhe et al., 2016), of
398 distance to forest (Boughey et al., 2011; Frey-Ehrenbold et al., 2013), to urban areas
399 (Mckinney, 2005; Jung & Threlfall, 2016) and to distance to wetlands (Santos et al., 2013;
400 Sirami et al., 2013).

401 A comparison of the relationship between environmental variables and bat activity between
402 using the raw data (i.e. using the whole dataset regardless of the confidence score) and FPTs -

403 selected data (i.e. removing data above defined FPT to minimize the false positive rate)
404 showed some discrepancies. We sometimes found opposite significant responses, for example
405 the effect of open areas vs. hedgerows on *N. noctula* and *P. nathusii*, when comparing results
406 from raw data and FPTs (Table 3). This demonstrates that analyses conducted on raw
407 automated identification data could be severely biased. In this respect, removing data above a
408 0.5 FPT (i.e. removing data with a low success probability) is essential, in accordance with
409 concerns expressed by Russo & Voigt (2016).

410 Logically these biases due to false positives mostly seem to impact uncommon species which
411 are acoustically similar to commoner ones. Here the most impacted species is *P. nathusii*
412 which suffers from a high false positive rate due to the local abundance of *P. kuhlii* and *P.*
413 *pipistrellus* (Tables 1 & 2). Consequently, an analysis conducted on raw automatically
414 identified data for this species seems to be driven by the response of the two other Pipistrelles.

415

416 *Assessing robustness of ecological inferences*

417 We assessed the robustness of ecological inferences by studying the consistency of bat
418 responses to environmental variables among FPTs. However, for *P. nathusii* it was not
419 possible to ensure such a robustness due a lack of data from 0.4 FPT (Table 2). This
420 framework thus showed that this was not possible to produce robust ecological inferences on
421 this species due to a high false positive rate in this dataset. In addition, for *N. noctula*, we lost
422 significance of the response to the distance to the forest and the length of hedgerows from 0.2
423 and 0.3 FPTs respectively (Table 3). Such loss of significance could be linked to a high loss
424 of bat passes and occurrences by reducing the FPTs, or linked to environmental biases
425 affecting spatial distribution of false positive or generated false negative rates. Thus, given the
426 uncertainty about the mechanism involved, it was also not possible to produce robust

427 inferences for this species given high losses of bat passes and occurrence, and high estimated
428 false positive rates by reducing the FPTs (Table 2).

429 At the other end of the spectrum, the estimated false positive rate was always extremely low
430 or even zero whatever the confidence score in the automated identification for *P. pipistrellus*
431 and *R. ferrumequinum* (Table 1), thus not raising any problem of error risk.

432 For all nine other species or species groups, 15 of the 18 significant responses to
433 environmental variables were robust with a high stability of model outputs while reducing the
434 FPT from 0.5 to 0.1 (Table 3). In addition, despite a decrease in bat activity measures due to
435 thresholding at FPTs, the occurrence of species remained highly stable whilst retaining
436 statistical power among FPTs. Our study thus demonstrates that using our approach many
437 ecological inferences could be robust against identification errors.

438

439 *Survey recommendations and limitations*

440 This study proposes a cautious method to account for identification errors in acoustic surveys
441 aimed at studying the response of bats in relation to environmental variables, such as
442 anthropogenic pressures, without the need for exhaustive checking of recordings.

443 The FPT of 0.5 is a threshold for which false negatives and false positives are expected to be
444 approximately balanced. However, false positives are more likely to produce biases because
445 their rate is strongly driven by the activity pattern of other species. In contrast, the FPT of 0.1
446 minimises the false positive rate, but at the cost of losing potentially a lot of data, so a high
447 generated false negative rate by discarding true positives (Table 2). Rather than looking for a
448 possible optimal threshold, we recommend that researchers systematically check the
449 consistency of responses for at least two significantly different thresholds (e.g. 0.5 and 0.1
450 FPTs), in order to assess the robustness of the results and only going on to conclusive
451 interpretation when these are consistent.

452 A lack of consistency is most likely to occur for rare species with very low
453 abundance/occurrence, and for uncommon species which are acoustically similar to
454 commoner ones such as *P. nathusii* here which is acoustically similar to *P. kuhlii* (Obrist et
455 al., 2004). The efficiency of the automated identification of *P. nathusii* and *N. lesleiri* was
456 lowest (AUC of 0.73 for both; Fig. S3) due to particular context of the study where these
457 species were much rarer than their acoustically closest relative (*P. kuhlii* and *E. serotinus*,
458 respectively; Table 1). For these species, either systematic manual checking or an important
459 improvement in automated identification efficiency is needed to conduct robust analyses.
460 However, our framework of error rate modelling is already sufficient to effectively identify
461 these problematic species and should prevent users of automated identification to draw
462 conclusions that are not robust. In addition, another prerequisite for drawing robust
463 conclusions from this framework is to ensure that error types (i.e. false negatives and false
464 positives) are not correlated with the variables tested in the study. In our study case, we only
465 detected one significant dependence for the open area *vs.* hedgerows for the false positives of
466 *N. noctula* (Table S2). For this species, automated identification was more efficient (i.e. lower
467 number of false positives) for survey sites located in open areas than close to hedgerows
468 where calls are more difficult to identify due to frequency modulation (Obrist et al., 2004;
469 Barataud, 2015). It is not surprising that the false positive rate of a rare species like *N. noctula*
470 could be influenced by local habitat type because this variable is expected to have different
471 effects on other species, and thus influence false positive rate through the relative density
472 between *N. noctula* and other bat species. Thus, we expect a bias in the measure of activity
473 towards open areas in this case. Hence, the significant positive response of this species to
474 open areas compared to hedgerows should be considered unreliable to make any ecological
475 inference (Table 3).

476 This method can be applied to any ecological studies with standardized sampling but, of
477 course, cannot help for surveys where no error can be tolerated, e.g. for producing species
478 inventories for protected species, as required for environmental impact assessments (Russo &
479 Voigt 2016). However, in this case, automated identification can still indicate what bat passes
480 should be manually checked in order to identify species presence at the site scale, by selecting
481 passes with the highest confidence scores, and thus saving time for the user.

482 Finally, the proposed method can be applied to any acoustic taxa for which automated
483 identification software are developed and where confidence scores are provided. A crucial
484 advantage of this method is that manual checking of a relatively small subset of the dataset (<
485 1% in this study) is sufficient to assess error rates associated with species identification. This
486 is especially true given that checking all data is very time-consuming and virtually impossible
487 for such a large dataset.

488

489 **Acknowledgements**

490 This work was supported by DIM ASTREA grants from Region Ile- de- France. We sincerely
491 thank Agrosolutions for funding field study fees. We thank IN2P3 Computing Centre for
492 providing facilities to process and archive in the long-term all the recordings of this study, and
493 Didier Bas for help in this process. We also thank the two anonymous reviewers for their
494 constructive comments improving this work.

495

496 **Author's contribution**

497 KB, CK and YB conceived the ideas, KB and YB designed the methodology, KB collected
498 the data; KB and YB manually checked bat passes, KB and JP analysed the data and wrote the
499 R scripts; all authors led the writing of the manuscript. All authors critically contributed to the
500 drafts and gave their final approval for publication.

501

502 **Data accessibility**

503 All R codes and data used in the study are available from the following github website link:
504 [https://github.com/KevBarre/Semi-automated-method-to-account-for-identification-errors-in-](https://github.com/KevBarre/Semi-automated-method-to-account-for-identification-errors-in-biological-acoustic-surveys)
505 [biological-acoustic-surveys](https://github.com/KevBarre/Semi-automated-method-to-account-for-identification-errors-in-biological-acoustic-surveys). <https://doi.org/10.5281/zenodo.2646482>

506

507

508 **References**

- 509 Adams, A. M., Jantzen, M. K., Hamilton, R. M., & Fenton, M. B. (2012). Do you hear what I
510 hear? Implications of detector selection for acoustic monitoring of bats. *Methods in*
511 *Ecology and Evolution*, 3(6), 992–998. doi:10.1111/j.2041-210X.2012.00244.x
- 512 Azam, C., Le Viol, I., Bas, Y., Zissis, G., Vernet, A., Julien, J.-F., & Kerbiriou, C. (2018).
513 Evidence for distance and illuminance thresholds in the effects of artificial lighting on
514 bat activity. *Landscape and Urban Planning*, (175), 123–135.
515 doi:10.1016/j.landurbplan.2018.02.011
- 516 Azam, C., Le Viol, I., Julien, J. F., Bas, Y., & Kerbiriou, C. (2016). Disentangling the relative
517 effect of light pollution, impervious surfaces and intensive agriculture on bat activity
518 with a national-scale monitoring program. *Landscape Ecology*, 31(10), 2471–2483.
519 doi:10.1007/s10980-016-0417-3
- 520 Barataud, M. (2015). *Acoustic ecology of european bats: species identification, study of their*
521 *habitats and foraging behaviour* (Biotope Ed).
- 522 Barré, K., Le Viol, I., Bas, Y., Julliard, R., & Kerbiriou, C. (2018). Estimating habitat loss due
523 to wind turbine avoidance by bats: Implications for European siting guidance. *Biological*
524 *Conservation*, 226(April), 205–214. doi:10.1016/j.biocon.2018.07.011
- 525 Barré, K., Le Viol, I., Julliard, R., Chiron, F., & Kerbiriou, C. (2017). Tillage and herbicide
526 reduction mitigate the gap between conventional and organic farming effects on foraging
527 activity of insectivorous bats. *Ecology and Evolution*, (00), 1–11. doi:10.1002/ece3.3688
- 528 Bas, Y., Bas, D., & Julien, J. (2017). Tadarida : A Toolbox for Animal Detection on Acoustic
529 Recordings. *Journal of Open Research Software*, 5(6), 1–8.
530 doi:https://doi.org/10.5334/jors.154
- 531 Biotope. (2013). Manual of SonoChiro 3.0. Retrieved 24 April 2018, from [http://www.leclub-](http://www.leclub-biotope.com/index.php?controller=attachment&id_attachment=13)
532 [biotope.com/index.php?controller=attachment&id_attachment=13](http://www.leclub-biotope.com/index.php?controller=attachment&id_attachment=13)

- 533 Boughey, K. L., Lake, I. R., Haysom, K. a., & Dolman, P. M. (2011). Effects of landscape-
534 scale broadleaved woodland configuration and extent on roost location for six bat species
535 across the UK. *Biological Conservation*, *144*(9), 2300–2310.
536 doi:10.1016/j.biocon.2011.06.008
- 537 Britzke, E. R., Duchamp, J. E., Murray, K. L., Swihart, R. K., & Robbins, L. W. (2011).
538 Acoustic identification of bats in the eastern United States: A comparison of parametric
539 and nonparametric methods. *Journal of Wildlife Management*, *75*(3), 660–667.
540 doi:10.1002/jwmg.68
- 541 Butchart, S. H. M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P. W., Almond,
542 R. E. A., ... Watson, R. (2010). Global Biodiversity: Indicators of Recent Declines.
543 *Science*, *328*(5982), 1164–1168. doi:10.1126/science.1187512
- 544 Chatterjee, S., & Hadi, A. S. (2006). *Regression analysis by example* (5th ed.). John Wiley &
545 Sons, Inc. doi:10.1002/0470055464
- 546 Claireau, F., Bas, Y., Puechmaille, S. J., Julien, J. F., Allegrini, B., & Kerbiriou, C. (2018).
547 Bat overpasses: An insufficient solution to restore habitat connectivity across roads.
548 *Journal of Applied Ecology*, (September), 1–12. doi:10.1111/1365-2664.13288
- 549 Delport, W., Kemp, A. C., & Ferguson, J. W. H. (2002). Vocal identification of individual
550 African Wood Owls *Strix woodfordii*: A technique to monitor long-term adult turnover
551 and residency. *Ibis*, *144*(1), 30–39. doi:10.1046/j.0019-1019.2001.00019.x
- 552 Denzinger, A., & Schnitzler, H. U. (2013). Bat guilds, a concept to classify the highly diverse
553 foraging and echolocation behaviors of microchiropteran bats. *Frontiers in Physiology*, *4*
554 *JUL*(July), 1–16. doi:10.3389/fphys.2013.00164
- 555 Fisher, J. A. D., Frank, K. T., & Leggett, W. C. (2010). Dynamic macroecology on ecological
556 time-scales. *Global Ecology and Biogeography*, *19*(1), 1–15. doi:10.1111/j.1466-
557 8238.2009.00482.x

- 558 Frey-Ehrenbold, A., Bontadina, F., Arlettaz, R., & Obrist, M. K. (2013). Landscape
559 connectivity, habitat structure and activity of bat guilds in farmland-dominated matrices.
560 *Journal of Applied Ecology*, 50(1), 252–261. doi:10.1111/1365-2664.12034
- 561 Fritsch, G., & Bruckner, A. (2014). Operator bias in software-aided bat call identification.
562 *Ecology and Evolution*, 4(13), 2703–2713. doi:10.1002/ece3.1122
- 563 Froidevaux, J. S. P., Zellweger, F., Bollmann, K., & Obrist, M. K. (2014). Optimizing passive
564 acoustic sampling of bats in forests. *Ecology and Evolution*, 4(24), 4690–4700.
565 doi:10.1002/ece3.1296
- 566 Frommolt, K. H. (2017). Information obtained from long-term acoustic recordings: applying
567 bioacoustic techniques for monitoring wetland birds during breeding season. *Journal of*
568 *Ornithology*, 158(3), 659–668. doi:10.1007/s10336-016-1426-3
- 569 Gasc, A., Sueur, J., Pavoine, S., Pellens, R., & Grandcolas, P. (2013). Biodiversity Sampling
570 Using a Global Acoustic Approach: Contrasting Sites with Microendemics in New
571 Caledonia. *PLoS ONE*, 8(5). doi:10.1371/journal.pone.0065311
- 572 Gibb, R., Browning, E., Glover-Kapfer, P., & Jones, K. E. (2018). Emerging opportunities
573 and challenges for passive acoustics in ecological assessment and monitoring. *Methods*
574 *in Ecology and Evolution*, (October). doi:10.1111/2041-210X.13101
- 575 Jeliaskov, A., Bas, Y., Kerbirou, C., Julien, J. F., Penone, C., & Le Viol, I. (2016). Large-
576 scale semi-automated acoustic monitoring allows to detect temporal decline of bush-
577 crickets. *Global Ecology and Conservation*, 6, 208–218.
578 doi:10.1016/j.gecco.2016.02.008
- 579 Jung, K., & Threlfall, C. (2016). Urbanization and its effect on bats: a global meta-analysis.
580 In: Kingston T, Voigt C (eds) *Bats in the anthropocene: conservation of bats in a*
581 *changing world*. (Springer, pp. 13–33).
- 582 Kalan, A. K., Mundry, R., Wagner, O. J. J., Heinicke, S., Boesch, C., & Kühl, H. S. (2015).

583 Towards the automated detection and occupancy estimation of primates using passive
584 acoustic monitoring. *Ecological Indicators*, 54, 217–226.
585 doi:10.1016/j.ecolind.2015.02.023

586 Kerbiriou, C., Azam, C., Touroult, J., Marmet, J., Julien, J. F., & Pellissier, V. (2018).
587 Common bats are more abundant within Natura 2000 areas. *Biological Conservation*,
588 217(October 2017), 66–74. doi:10.1016/j.biocon.2017.10.029

589 Kerbiriou, C., Bas, Y., Le Viol, I., Lorrilliere, R., Mougnot, J., & Julien, J. F. (2018).
590 Potential of bat pass duration measures for studies of bat activity. *Bioacoustics*, 4622, 1–
591 16. doi:10.1080/09524622.2017.1423517

592 Lacoueilhe, A., Machon, N., Julien, J. F., & Kerbiriou, C. (2016). Effects of hedgerows on
593 bats and bush crickets at different spatial scales. *Acta Oecologica*, 71, 61–72.
594 doi:10.1016/j.actao.2016.01.009

595 López-Baucells, A., Torrent, L., Rocha, R., Bobrowiec, P. E. D., Palmeirim, J. M., & Meyer,
596 C. F. J. (2019). Stronger together: Combining automated classifiers with manual post-
597 validation optimizes the workload vs reliability trade-off of species identification in bat
598 acoustic surveys. *Ecological Informatics*, 49, 45–53. doi:10.1016/j.ecoinf.2018.11.004

599 Mckinney, M. L. (2005). Urbanization as a major cause of biotic homogenization. *Biological*
600 *Conservation*, 127(3), 247–260. doi:10.1016/j.biocon.2005.09.005

601 Millon, L., Julien, J.-F., Julliard, R., & Kerbiriou, C. (2015). Bat activity in intensively
602 farmed landscapes with wind turbines and offset measures. *Ecological Engineering*, 75,
603 250–257. doi:10.1016/j.ecoleng.2014.11.050

604 Newson, S. E., Evans, H. E., & Gillings, S. (2015). A novel citizen science approach for
605 large-scale standardised monitoring of bat activity and distribution, evaluated in eastern
606 England. *Biological Conservation*, 191, 38–49. doi:10.1016/j.biocon.2015.06.009

607 Nowacek, D. P., Christiansen, F., Bejder, L., Goldbogen, J. A., & Friedlaender, A. S. (2016).

608 Studying cetacean behaviour: new technological approaches and conservation
609 applications. *Animal Behaviour*, 120, 235–244. doi:10.1016/j.anbehav.2016.07.019

610 Obrist, M. K., Boesch, R., & Fluckiger, P. F. (2004). Variability in echolocation call design of
611 26 Swiss bat species: consequences, limits and options for automated field identification
612 with a synergetic pattern recognition approach. *Mammalia*, 68(4), 307–322.
613 doi:10.1515/mamm.2004.030

614 Ovaskainen, O., Moliterno de Camargo, U., & Somervuo, P. (2018). Animal Sound Identifier
615 (ASD): software for automated identification of vocal animals. *Ecology Letters*, 21(8),
616 1244–1254. doi:10.1111/ele.13092

617 Parsons, S., & Jones, G. (2000). Acoustic identification of twelve species of echolocating bat
618 by discriminant function analysis and artificial neural networks. *The Journal of*
619 *Experimental Biology*, 203(Pt 17), 2641–2656. doi:10.1007/s00114-005-0622-4

620 Pauwels, J., Viol, I. Le, Azam, C., Valet, N., Julien, J. -f., Bas, Y., ... Kerbiriou, C. (2019).
621 Accounting for artificial light impact on bat activity for a biodiversity-friendly urban
622 planning. *Landscape and Urban Planning*, 183(August 2018), 12–25.
623 doi:S0169204618311939

624 Penone, C., Kerbiriou, C., Julien, J.-F., Marmet, J., & Le Viol, I. (2018). Body size
625 information in large-scale acoustic bat databases. *PeerJ*, 6, e5370.
626 doi:10.7717/peerj.5370

627 Pinaud, D., Claireau, F., Leuchtman, M., & Kerbiriou, C. (2018). Modelling landscape
628 connectivity for greater horseshoe bat using an empirical quantification of resistance.
629 *Journal of Applied Ecology*, 55(6), 2600–2611. doi:10.1111/1365-2664.13228

630 Russo, D., & Voigt, C. C. (2016). The use of automated identification of bat echolocation
631 calls in acoustic monitoring: A cautionary note for a sound analysis. *Ecological*
632 *Indicators*, 66, 598–602. doi:10.1016/j.ecolind.2016.02.036

633 Rydell, J., Nyman, S., Eklöf, J., Jones, G., & Russo, D. (2017). Testing the performances of
634 automated identification of bat echolocation calls: A request for prudence. *Ecological*
635 *Indicators*, 78, 416–420. doi:10.1016/j.ecolind.2017.03.023

636 Santos, H., Rodrigues, L., Jones, G., & Rebelo, H. (2013). Using species distribution
637 modelling to predict bat fatality risk at wind farms. *Biological Conservation*, 157, 178–
638 186. doi:10.1016/j.biocon.2012.06.017

639 Scott, C., & Altringham, J. (2017). BatClassify. Retrieved 24 April 2018, from
640 <https://bitbucket.org/chrisscott/batclassify>

641 Sirami, C., Jacobs, D. S., & Cumming, G. S. (2013). Artificial wetlands and surrounding
642 habitats provide important foraging habitat for bats in agricultural landscapes in the
643 Western Cape, South Africa. *Biological Conservation*, 164, 30–38.
644 doi:10.1016/j.biocon.2013.04.017

645 Stahlschmidt, P., & Brühl, C. A. (2012). Bats as bioindicators - the need of a standardized
646 method for acoustic bat activity surveys. *Methods in Ecology and Evolution*, 3(3), 503–
647 508. doi:10.1111/j.2041-210X.2012.00188.x

648 Verboom, B., & Huitema, H. (1997). The importance of linear landscape elements for the
649 pipistrelle *Pipistrellus pipistrellus* and the serotine bat *Eptesicus serotinus*. *Landscape*
650 *Ecology*, 12(2), 117–125. doi:10.1007/BF02698211

651 Waters, D., & Barlow, K. (2013). Bat detectors: past, present and future. *British Wildlife*, 87,
652 86–92.

653 Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects*
654 *models and extensions in ecology with R*. Springer Science & Business Media.
655 doi:10.1007/978-0-387-87458-6

656

657

658 Table 1. Total bat passes assigned to each species by the automated identification per
659 confidence score classes, number of bat passes manually double checked and false positives
660 noted (step 1 in Fig. 1). See Table S1 for species composition in false positives.

Species	Upper limits of confidence score classes of the automated identification										Total
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1	
<i>Barbastella barbastellus</i>											
Total passes	3	52	144	242	297	671	940	1312	1596	578	5835
Checked passes	3	25	25	25	25	25	25	25	25	25	228
False positives	3	5	1	0	0	0	0	0	0	0	3.9 %
<i>Eptesicus serotinus</i>											
Total passes	1	55	102	149	268	461	218	79	10	0	1343
Checked passes	1	25	25	25	25	25	25	25	9	0	185
False positives	1	13	7	0	0	0	0	0	0	0	11.4 %
<i>Myotis nattereri</i>											
Total passes	9	166	211	223	225	411	269	180	247	47	1988
Checked passes	9	9	3	6	8	2	2	10	23	25	97
False positives	8	5	1	2	1	0	0	0	0	0	17.5 %
<i>Myotis spp</i>											
Total passes	20	534	815	770	701	1708	1132	445	258	47	6430
Checked passes	20	25	25	25	25	25	25	25	25	25	245
False positives	19	14	6	6	4	0	0	0	0	0	20.0 %
<i>Nyctalus leisleri</i>											
Total passes	3	47	41	33	11	8	9	1	0	0	153
Checked passes	3	25	25	25	11	8	9	1	0	0	107
False positives	2	16	14	13	4	0	0	0	0	0	45.8 %
<i>Nyctalus noctula</i>											
Total passes	0	113	110	82	24	43	16	6	1	0	395
Checked passes	0	25	25	25	24	25	16	6	1	0	147
False positives	0	25	23	24	23	7	0	0	0	0	69.4 %
<i>Pipistrellus kuhlii</i>											
Total passes	12	223	401	667	1142	4026	6654	10222	5240	2	28589
Checked passes	12	25	25	25	25	25	25	25	25	2	214
False positives	11	10	8	4	2	2	1	0	0	0	17.8 %
<i>Pipistrellus nathusii</i>											
Total passes	0	12	33	37	93	183	153	61	5	0	577
Checked passes	0	12	25	25	25	25	25	25	5	0	167
False positives	0	11	20	20	19	17	15	9	1	0	67.1 %
<i>Pipistrellus pipistrellus</i>											
Total passes	2	303	760	1636	3298	8311	14221	27205	83744	28024	167504
Checked passes	2	25	25	25	25	25	25	25	25	25	227
False positives	1	2	0	1	1	0	0	1	0	0	2.6 %
<i>Plecotus spp</i>											
Total passes	8	139	176	194	174	250	206	145	56	4	1352
Checked passes	8	30	26	25	28	25	25	25	25	4	221
False positives	5	19	8	2	1	1	0	0	0	0	16.3 %
<i>Rhinolophus ferrumequinum</i>											
Total passes	0	0	0	0	1	6	5	28	1	0	41
Checked passes	0	0	0	0	1	6	5	28	1	0	41
False positives	0	0	0	0	0	0	0	0	0	0	0.0 %
<i>Rhinolophus hipposideros</i>											
Total passes	0	1	1	10	8	16	26	62	4	0	128
Checked passes	0	1	1	10	8	16	26	62	4	0	128
False positives	0	1	1	7	1	0	0	0	0	0	7.8 %

661

662 Table 2. Minimum confidence scores needed to ensure False Positive Tolerances (step 2 in
 663 Fig. 1), associated changes in the number of bat passes, the occurrence (presence rate among
 664 sites), the estimated false positive rate and the generated false negative rate estimated for the
 665 whole dataset (212 347 bat passes; step 3 in Fig. 1).

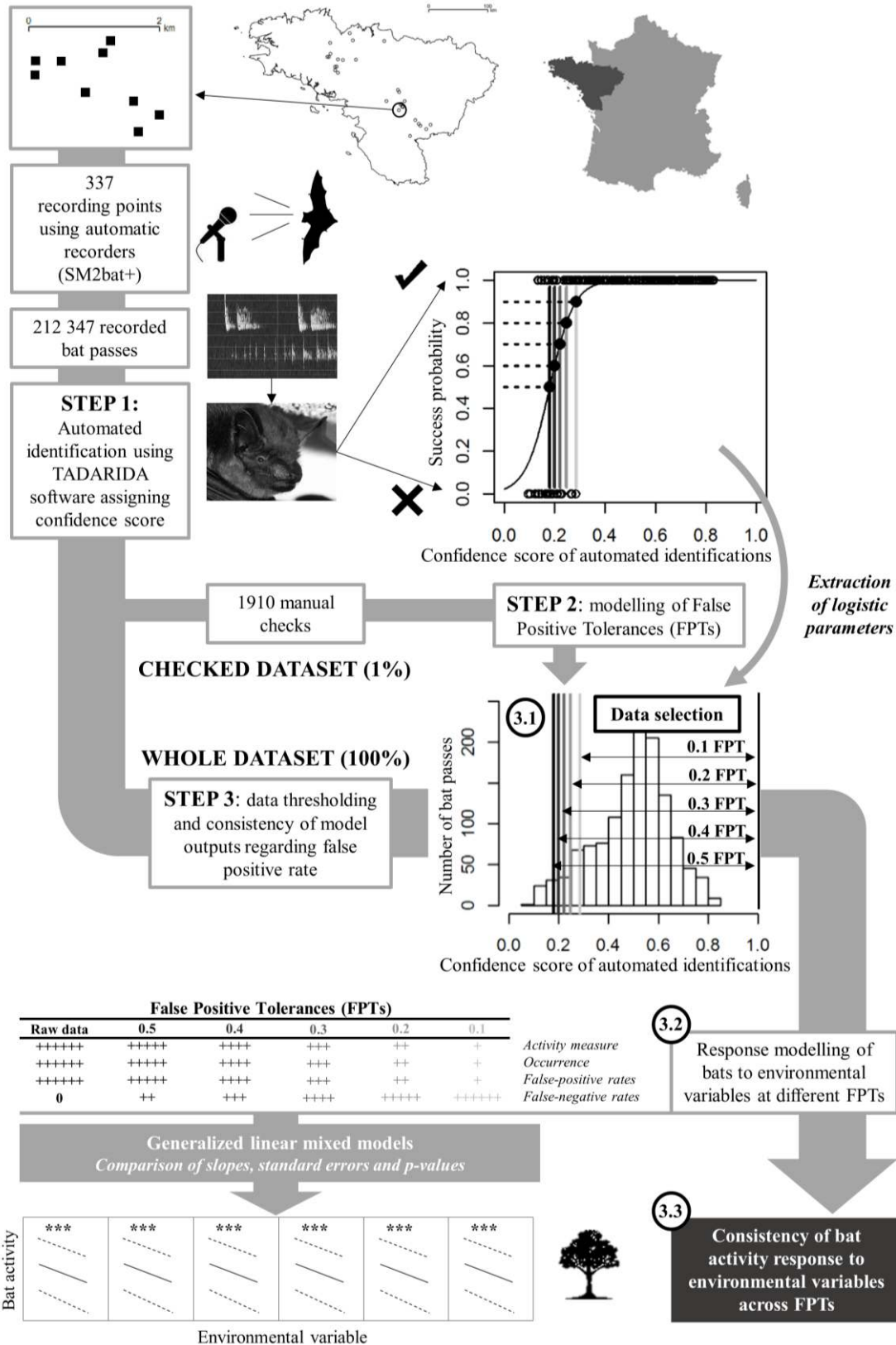
Species		Raw data	False Positive Tolerance				
			0.5	0.4	0.3	0.2	0.1
<i>Barbastella barbastellus</i>	Confidence score	/	0.119	0.133	0.148	0.167	0.195
	No. of bat passes	5835	5828	5824	5822	5809	5787
	Occurrences	0.694	0.694	0.694	0.694	0.694	0.694
	Estimated false positive rate	0.003	0.002	0.002	0.002	0.001	0.001
	Estimated false negative rate	0	<0.001	<0.001	0.001	0.003	0.006
<i>Eptesicus serotinus</i>	Confidence score	/	0.180	0.200	0.221	0.246	0.285
	No. of bat passes	1343	1297	1287	1273	1255	1205
	Occurrences	0.373	0.339	0.336	0.333	0.324	0.312
	Estimated false positive rate	0.044	0.022	0.019	0.015	0.012	0.006
	Estimated false negative rate	0	0.011	0.016	0.023	0.031	0.065
<i>Myotis nattereri</i>	Confidence score	/	0.229	0.271	0.317	0.373	0.458
	No. of bat passes	1986	1759	1659	1562	1436	1239
	Occurrences	0.688	0.648	0.624	0.609	0.578	0.529
	Estimated false positive rate	0.136	0.081	0.064	0.049	0.034	0.021
	Estimated false negative rate	0	0.036	0.059	0.087	0.132	0.199
<i>Myotis spp.</i>	Confidence score	/	0.212	0.250	0.291	0.341	0.416
	No. of bat passes	6428	5783	5483	5135	4747	4173
	Occurrences	0.798	0.792	0.786	0.774	0.765	0.716
	Estimated false positive rate	0.145	0.092	0.073	0.054	0.038	0.024
	Estimated false negative rate	0	0.036	0.062	0.099	0.145	0.219
<i>Nyctalus leisleri</i>	Confidence score	/	0.286	0.342	0.402	0.476	0.587
	No. of bat passes	153	67	43	28	22	12
	Occurrences	0.211	0.138	0.104	0.070	0.055	0.031
	Estimated false positive rate	0.502	0.305	0.222	0.149	0.115	0.075
	Estimated false negative rate	0	0.193	0.279	0.337	0.370	0.425
<i>Nyctalus noctula</i>	Confidence score	/	0.507	0.527	0.548	0.574	0.613
	No. of bat passes	395	61	50	41	29	22
	Occurrences	0.220	0.080	0.067	0.058	0.046	0.040
	Estimated false positive rate	0.850	0.212	0.158	0.120	0.066	0.042
	Estimated false negative rate	0	0.029	0.044	0.054	0.082	0.097
<i>Pipistrellus kuhlii</i>	Confidence score	/	0.164	0.216	0.272	0.341	0.444
	No. of bat passes	28588	28456	28305	28077	27737	26854
	Occurrences	0.899	0.899	0.890	0.884	0.881	0.875
	Estimated false positive rate	0.033	0.030	0.028	0.026	0.023	0.019
	Estimated false negative rate	0	0.002	0.005	0.010	0.019	0.045
<i>Pipistrellus nathusii</i>	Confidence score	/	0.668	0.756	0.853	0.971	/
	No. of bat passes	577	101	18	0	0	0
	Occurrences	0.404	0.116	0.031	0.000	0.000	0.000
	Estimated false positive rate	0.623	0.437	0.370	/	/	/
	Estimated false negative rate	0	0.275	0.355	0.377	/	/
<i>Pipistrellus pipistrellus</i>	Confidence score	/	0.000	0.000	0.000	0.000	0.096
	No. of bat passes	167503	167503	167503	167503	167503	167502
	Occurrences	0.954	0.954	0.954	0.954	0.954	0.954
	Estimated false positive rate	0.007	0.007	0.007	0.007	0.007	0.007
	Estimated false negative rate	0.000	0.000	0.000	0.000	0.000	0.000
<i>Plecotus spp.</i>	Confidence score	/	0.184	0.217	0.253	0.298	0.364
	No. of bat passes	1352	1229	1185	1129	1034	909
	Occurrences	0.615	0.599	0.596	0.596	0.584	0.544
	Estimated false positive rate	0.128	0.079	0.065	0.051	0.034	0.019
	Estimated false negative rate	0	0.034	0.053	0.080	0.131	0.211
<i>Rhinolophus ferrumequinum</i>	Confidence score	/	0.000	0.000	0.000	0.000	0.000
	No. of bat passes	41	41	41	41	41	41
	Occurrences	0.046	0.046	0.046	0.046	0.046	0.046
	Estimated false positive rate	0.000	0.000	0.000	0.000	0.000	0.000
	Estimated false negative rate	0.000	0.000	0.000	0.000	0.000	0.000
<i>Rhinolophus hipposideros</i>	Confidence score	/	0.385	0.398	0.411	0.427	0.452
	No. of bat passes	128	117	116	116	116	113
	Occurrences	0.113	0.107	0.104	0.104	0.104	0.104
	Estimated false positive rate	0.078	0.011	0.007	0.007	0.007	0.003
	Estimated false negative rate	0	0.018	0.022	0.022	0.022	0.199

667 Table 3. Species response to environmental variables (estimates, standard errors and p-values)

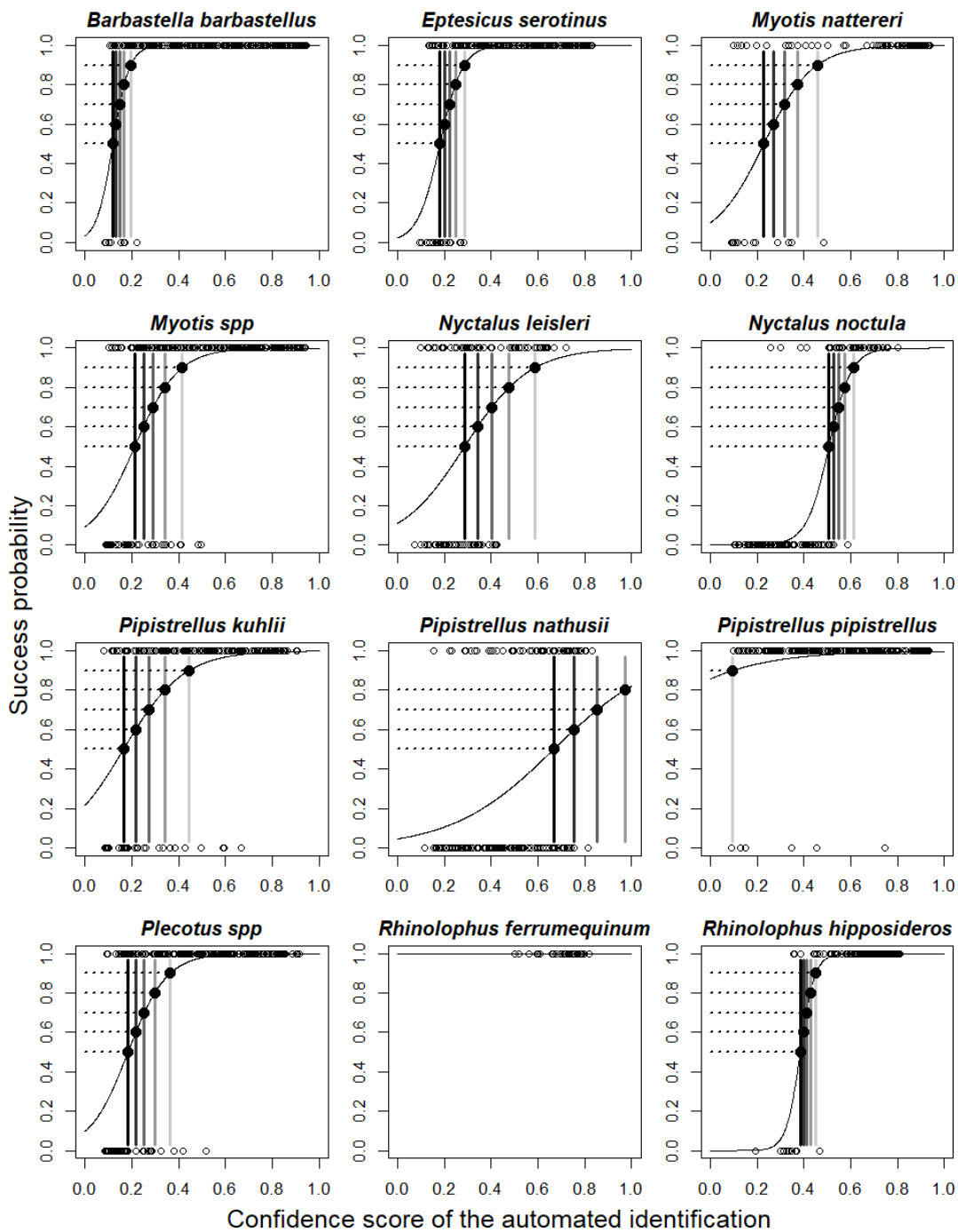
668 according to the False Positive Tolerances (*** P < 0.001, ** P < 0.01, * P < 0.05, . P < 0.1).

Species	Environmental variables	False Positive Tolerance					
		Raw data	0.5	0.4	0.3	0.2	0.1
<i>Barbastella barbastellus</i>	Open areas vs. hedgerows	-2.81±0.24 ***	-2.81±0.24 ***	-2.81±0.24 ***	-2.81±0.24 ***	-2.81±0.24 ***	-2.81±0.24 ***
	Dist. to forest	0.08±0.12	0.08±0.13	0.08±0.13	0.08±0.13	0.08±0.13	0.08±0.13
	Dist. to wetland	-0.03±0.12	-0.03±0.12	-0.03±0.12	-0.03±0.12	-0.03±0.12	-0.04±0.12
	Dist. to urban	0.01±0.1	0.01±0.1	0.01±0.1	0.01±0.1	0.01±0.1	0.02±0.1
	Length of hedgerows	0.17±0.12	0.17±0.12	0.17±0.12	0.17±0.12	0.17±0.12	0.17±0.12
<i>Eptesicus serotinus</i>	Open areas vs. hedgerows	-0.57±0.38	-0.43±0.4	-0.44±0.4	-0.45±0.41	-0.43±0.42	-0.35±0.42
	Dist. to forest	-0.07±0.23	-0.15±0.24	-0.15±0.25	-0.16±0.25	-0.15±0.25	-0.13±0.26
	Dist. to wetland	0.08±0.19	0.12±0.2	0.12±0.2	0.12±0.2	0.12±0.21	0.08±0.21
	Dist. to urban	-0.7±0.19 ***	-0.8±0.21 ***	-0.79±0.21 ***	-0.78±0.21 ***	-0.77±0.21 ***	-0.77±0.22 ***
	Length of hedgerows	0.2±0.23	0.2±0.24	0.21±0.24	0.21±0.24	0.19±0.25	0.16±0.25
<i>Myotis nattereri</i>	Open areas vs. hedgerows	-1.16±0.21 ***	-1.14±0.22 ***	-1.12±0.23 ***	-1.05±0.23 ***	-1.01±0.24 ***	-1.03±0.27 ***
	Dist. to forest	0.16±0.13	0.13±0.13	0.14±0.13	0.15±0.14	0.1±0.14	0.11±0.15
	Dist. to wetland	0.17±0.11	0.21±0.12 .	0.23±0.12 .	0.24±0.12 .	0.22±0.13 .	0.21±0.13
	Dist. to urban	0.07±0.1	0.08±0.11	0.09±0.11	0.11±0.12	0.11±0.12	0.13±0.13
	Length of hedgerows	0.18±0.12	0.22±0.13 .	0.24±0.13 .	0.27±0.14 .	0.32±0.14 *	0.3±0.16 .
<i>Myotis spp</i>	Open areas vs. hedgerows	-1.66±0.19 ***	-1.64±0.19 ***	-1.6±0.19 ***	-1.55±0.19 ***	-1.54±0.19 ***	-1.61±0.26 ***
	Dist. to forest	0.24±0.12 *	0.22±0.12 .	0.22±0.12 .	0.22±0.12 .	0.22±0.13 .	0.20±0.13
	Dist. to wetland	0.1±0.1	0.11±0.11	0.1±0.11	0.11±0.11	0.1±0.11	0.10±0.11
	Dist. to urban	-0.07±0.09	-0.08±0.09	-0.08±0.1	-0.06±0.1	-0.05±0.1	-0.03±0.1
	Length of hedgerows	0.13±0.12	0.15±0.12	0.15±0.12	0.17±0.12	0.18±0.12	0.21±0.13
<i>Nyctalus leisleri</i>	Open areas vs. hedgerows	-0.8±0.22 ***	-0.26±0.29	-0.23±0.35	0.43±0.4	0.69±0.45	1.1±0.64
	Dist. to forest	0.34±0.13 **	0.16±0.17	0.21±0.21	0.08±0.26	0.14±0.28	0.49±0.35
	Dist. to wetland	0.07±0.1	-0.09±0.15	-0.02±0.19	-0.12±0.26	-0.21±0.3	-0.17±0.42
	Dist. to urban	-0.1±0.1	-0.19±0.15	-0.01±0.18	0.08±0.23	0.23±0.26	0.43±0.35
	Length of hedgerows	0.35±0.12 **	0.23±0.16	0.23±0.21	0.27±0.25	0.28±0.29	0.22±0.41
<i>Nyctalus noctula</i>	Open areas vs. hedgerows	-1.19±0.17 ***	1.46±0.31 ***	1.7±0.36 ***	1.83±0.4 ***	1.37±0.44 **	1.28±0.49 *
	Dist. to forest	-0.55±0.11 ***	-0.68±0.23 **	-0.66±0.26 *	-0.7±0.29 *	-0.26±0.32	-0.12±0.35
	Dist. to wetland	-0.07±0.06	0.02±0.18	0.16±0.21	0.25±0.24	0.3±0.27	0.34±0.34
	Dist. to urban	0.25±0.07 ***	-0.07±0.18	-0.1±0.21	-0.12±0.23	-0.01±0.25	-0.04±0.29
	Length of hedgerows	0.34±0.08 ***	0.43±0.21 *	0.49±0.25 *	0.52±0.28 .	0.16±0.31	-0.03±0.36
<i>Pipistrellus kuhlii</i>	Open areas vs. Hedgerows	-1.98±0.26 ***	-1.98±0.26 ***	-1.98±0.27 ***	-1.98±0.27 ***	-1.98±0.27 ***	-1.98±0.27 ***
	Dist. to forest	0.09±0.13	0.09±0.13	0.09±0.13	0.09±0.14	0.09±0.14	0.1±0.14
	Dist. to wetland	0.25±0.13 *	0.25±0.13 *	0.26±0.13 *	0.25±0.13*	0.26±0.13*	0.26±0.13*
	Dist. to urban	0.07±0.13	0.07±0.13	0.07±0.13	0.08±0.13	0.08±0.13	0.08±0.13
	Length of hedgerows	0.07±0.15	0.06±0.15	0.06±0.15	0.06±0.15	0.06±0.15	0.06±0.15
<i>Pipistrellus nathusii</i>	Open areas vs. Hedgerows	-0.37±0.24	1.02±0.38 **	2.57±0.84 **	/	/	/
	Dist. to forest	0.1±0.16	0.28±0.23	0.81±0.46 .	/	/	/
	Dist. to wetland	0.06±0.13	0.02±0.2	0.53±0.42	/	/	/
	Dist. to urban	-0.05±0.13	0.09±0.21	0±0.44	/	/	/
	Length of hedgerows	0.11±0.16	0.42±0.24 .	0.88±0.54	/	/	/
<i>Pipistrellus pipistrellus</i>	Open areas vs. Hedgerows	-2.87±0.19 ***	-2.87±0.19 ***	-2.87±0.19 ***	-2.87±0.19 ***	-2.87±0.19 ***	-2.87±0.19 ***
	Dist. to forest	0.13±0.13	0.13±0.13	0.13±0.13	0.13±0.13	0.13±0.13	0.13±0.13
	Dist. to wetland	0.04±0.11	0.04±0.11	0.04±0.11	0.04±0.11	0.04±0.11	0.04±0.11
	Dist. to urban	-0.13±0.1	-0.13±0.1	-0.13±0.1	-0.13±0.1	-0.13±0.1	-0.13±0.1
	Length of hedgerows	0.35±0.12 **	0.35±0.12 **	0.35±0.12 **	0.35±0.12 **	0.35±0.12 **	0.35±0.12 **
<i>Plecotus spp.</i>	Open areas vs. Hedgerows	-0.91±0.19 ***	-0.85±0.19 ***	-0.87±0.19 ***	-0.87±0.19 ***	-0.85±0.19 ***	-0.79±0.2 ***
	Dist. to forest	0.08±0.12	0.1±0.12	0.11±0.12	0.1±0.12	0.09±0.12	0.08±0.13
	Dist. to wetland	-0.16±0.11	-0.14±0.11	-0.15±0.11	-0.15±0.11	-0.14±0.11	-0.17±0.12
	Dist. to urban	-0.25±0.1 **	-0.25±0.1 *	-0.26±0.1 **	-0.25±0.1 **	-0.25±0.1 *	-0.23±0.1 *
	Length of hedgerows	0.1±0.12	0.09±0.12	0.09±0.12	0.08±0.12	0.11±0.12	0.11±0.13
<i>Rhinolophus ferrumequinum</i>	Open areas vs. Hedgerows	0.26±0.39	0.26±0.39	0.26±0.39	0.26±0.39	0.26±0.39	0.26±0.39
	Dist. to forest	0.74±0.25 **	0.74±0.25 **	0.74±0.25 **	0.74±0.25 **	0.74±0.25 **	0.74±0.25 **
	Dist. to wetland	-1.2±0.29 ***	-1.20±0.29 ***	-1.20±0.29 ***	-1.20±0.29 ***	-1.20±0.29 ***	-1.20±0.29 ***
	Dist. to urban	-0.21±0.26	-0.21±0.26	-0.21±0.26	-0.21±0.26	-0.21±0.26	-0.21±0.26
	Length of hedgerows	0.83±0.29 **	0.83±0.29 **	0.83±0.29 **	0.83±0.29 **	0.83±0.29 **	0.83±0.29 **
<i>Rhinolophus hipposideros</i>	Open areas vs. Hedgerows	-3.08±0.74 ***	-2.92±0.73 ***	-2.92±0.74 ***	-2.92±0.74 ***	-2.92±0.74 ***	-2.89±0.73 ***
	Dist. to forest	0.09±0.3	-0.47±0.36	-0.5±0.37	-0.5±0.37	-0.5±0.37	-0.51±0.36
	Dist. to wetland	-0.33±0.26	-0.45±0.26 .	-0.49±0.27 .	-0.49±0.27 .	-0.49±0.27 .	-0.46±0.28 .
	Dist. to urban	-0.18±0.26	-0.17±0.26	-0.14±0.27	-0.14±0.27	-0.14±0.27	-0.15±0.27
	Length of hedgerows	0.03±0.3	0.06±0.3	0.07±0.3	0.07±0.3	0.07±0.3	0.08±0.3

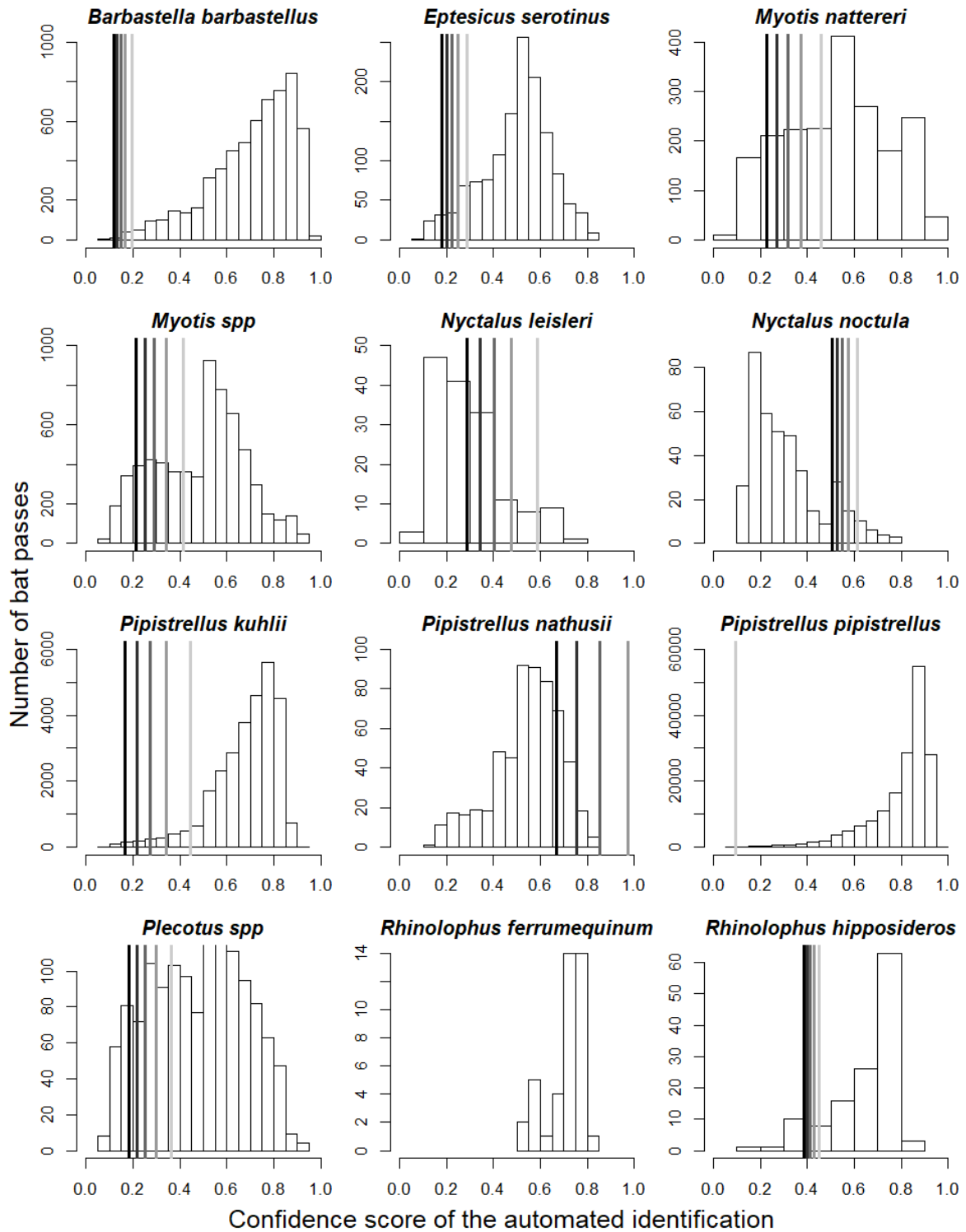
670 Fig. 1. Schematic and chronological representation of the steps used to study the relationship
 671 between automated identification errors in acoustic data and the detected relationship between
 672 bat activity and environmental variables.



674 Fig. 2. Logistic regressions between the success probability and the confidence score of the
 675 automated identification. The success probability was predicted from a subset manually
 676 checked assigning a success or a failure in automated identifications. Horizontal dotted lines
 677 show success probabilities in automated identification used for thresholding (i.e. False
 678 Positive Tolerances: 0.5, 0.4, 0.3, 0.2 and 0.1) to remove data in the total dataset below the
 679 corresponding confidence scores (vertical lines).



681 Fig. 3. Number of bat passes in the total dataset according to confidence scores provided by
 682 the automated identification. Vertical lines show the threshold below which data were
 683 removed to ensure a given False Positive Tolerance (from black to grey: 0.5, 0.4, 0.3, 0.2 and
 684 0.1).



685