

A peer-reviewed version of this preprint was published in PeerJ on 11 August 2017.

[View the peer-reviewed version](https://peerj.com/articles/3663) (peerj.com/articles/3663), which is the preferred citable publication unless you specifically need to cite this preprint.

Lee TE, Bowman C, Roberts DL. 2017. Are extinction opinions extinct? PeerJ 5:e3663 <https://doi.org/10.7717/peerj.3663>

Are extinction opinions extinct?

Tamsin E Lee ^{Corresp., 1}, Clive Bowman ^{Corresp., 2}, David L Roberts ^{Corresp., 3}

¹ Mathematical Institute, University of Oxford, Oxford, UK

² Mathematical Institute, University of Oxford, United Kingdom

³ Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation, University of Kent at Canterbury, United Kingdom

Corresponding Authors: Tamsin E Lee, Clive Bowman, David L Roberts

Email address: tamsin.lee@maths.ox.ac.uk, bowman@maths.ox.ac.uk, d.l.roberts@kent.ac.uk

Extinction models vary in the information they require, the simplest considering the rate of certain sightings only. More complicated methods include uncertain sightings and allow for variation in the reliability of uncertain sightings. Generally extinction models require expert opinion, either as a prior belief that a species is extinct, or to establish the quality of a sighting record, or both. Is this subjectivity necessary?

We present two models to explore whether the individual quality of sightings, judged by experts, is strongly informative of the probability of extinction: the 'quality breakpoint method' and the 'quality as variance method'. For the first method we use the Barbary lion as an exemplar. For the second method we use the Barbary lion, Alaotra grebe, Jamaican petrel and Pohnpei starling as exemplars.

The 'quality breakpoint method' uses certain and uncertain sighting records, and the quality of uncertain records, to establish whether a change point in the rate of sightings can be established using a simultaneous Bayesian optimisation with a non-informative prior. For the Barbary lion, there is a change in subjective quality of sightings around 1930. Unexpectedly sighting quality increases after this date. This suggests that including quality scores from experts can lead to irregular effects and may not offer reliable results. As an alternative, we use quality as a measure of variance around the sightings, not a change in quality. This leads to predictions with larger standard deviations, however the results remain consistent across any prior belief of extinction. Nonetheless, replacing actual quality scores with random quality scores showed little difference, inferring that the quality scores from experts are superfluous.

Therefore, we deem the expensive process of obtaining pooled expert estimates as unnecessary and even when used we recommend that sighting data should have minimal input from experts in terms of assessing the sighting quality at a fine scale. Rather, sightings should be classed as certain or uncertain, using a framework that is as independent of human bias as possible.

Are extinction opinions extinct?

Tamsin E. Lee¹, Clive E. Bowman¹ and David L. Roberts²

¹ Mathematical Institute, University of Oxford, Andrew Wiles Building, Radcliffe Observatory
Quarter, Woodstock Road, OX2 6GG

² Durrell Institute of Conservation and Ecology, School of Anthropology and Conservation,
University of Kent, Marlowe Building, Canterbury, Kent, CT2 7NR

Abstract

1. Extinction models vary in the information they require, the simplest considering the rate of certain sightings only. More complicated methods include uncertain sightings and allow for variation in the reliability of uncertain sightings. Generally extinction models require expert opinion, either as a prior belief that a species is extinct, or to establish the quality of a sighting record, or both. Is this subjectivity necessary?

2. We present two models to explore whether the individual quality of sightings, judged by experts, is strongly informative of the probability of extinction: the ‘quality breakpoint method’ and the ‘quality as variance method’. For the first method we use the Barbary lion as an exemplar. For the second method we use the Barbary lion, Alaotra grebe, Jamaican petrel and Pohnpei starling as exemplars.

3. The ‘quality breakpoint method’ uses certain and uncertain sighting records, and the quality of uncertain records, to establish whether a change point in the rate of sightings can be established using a simultaneous Bayesian optimisation with a non-informative prior. For the Barbary lion, there is a change in subjective quality of sightings around 1930. Unexpectedly sighting quality increases after this date. This suggests that including quality scores from experts can lead to irregular effects and may not offer reliable results. As an alternative, we use quality as a measure of variance around the sightings, not a change in quality. This leads to predictions with larger standard deviations, however

26 the results remain consistent across any prior belief of extinction. Nonetheless, replacing
27 actual quality scores with random quality scores showed little difference, inferring that the
28 quality scores from experts are superfluous.

29 4. Therefore, we deem the expensive process of obtaining pooled expert estimates as unnec-
30 essary and even when used we recommend that sighting data should have minimal input
31 from experts in terms of assessing the sighting quality at a fine scale. Rather, sightings
32 should be classed as certain or uncertain, using a framework that is as independent of
33 human bias as possible.

34 Keywords: extinction model, data quality, IUCN Red List, possibly extinct, sighting uncer-
35 tainty, *Panthera leo*, *Tachybaptus rufolavatus*, *Pterodroma caribbaea*, *Aplonis pelzelni*

36 1 Introduction

37 The quality of sighting records of rare species, and particularly those that are approaching extinc-
38 tion, vary considerably. This can lead to confusion, particularly when identifying whether a species
39 is extinct, or identifying when a species went extinct (Roberts et al., 2009 and Elphick et al., 2010).
40 An extinction date for a given species is usually inferred from the rate of sightings by assuming
41 that the rate changes after the point of extinction. Recent models (Solow et al., 2012, Thompson
42 et al., 2013, Lee, 2014, Lee et al., 2014 and Jarić & Roberts, 2014) incorporate uncertain sightings
43 as well, thus recorded sightings might occur after extinction. A review of using sightings records
44 to infer extinction is provided by Boakes, Rout and Collen (2015).

45 Generally sightings are either grouped as certain or uncertain records by researchers, e.g the
46 ivory-billed woodpecker (Solow et al. 2012). The subjective quality rather than the certainty of a
47 record has been less investigated. To incorporate the varying quality of sighting records, Thompson

48 et al. (2013) and Lee (2014) present a method which allows several different classes of uncertain
49 records, where the classification is determined by subjective quality. Their method is optimal
50 if at least one sighting from each coarsely defined group occurs before the last certain sighting.
51 Any approach requires expert information about the quality of records. Suppose we have high
52 resolution information for the quality of sighting records, that is, we have pooled expert opinions
53 on the quality of each individual sighting record. Does expert information actually improve our
54 inference on extinction estimates? We use the Barbary Lion as an initial test-bed. Collecting
55 pooled expert opinions on individual sighting records is a time-expensive exercise, thus only the
56 Barbary lion sightings currently have this high level of detail. Indeed the primary motivation for
57 this paper is to ascertain whether this cost is necessary.

58 Lee et al. (2015) provided distributions for 32 alleged sightings of the Barbary lion (*Panthera*
59 *l. leo*) which occurred between 1895 and 1956 in Algeria and Morocco. In this paper we use the
60 individual quality score provided by Lee et al. (2015). We also examine the importance of the
61 expert's prior of the lion being extinct on the results.

62 The work of Lee et al. (2015) provides several distributions for each lion sighting. One method
63 considers the expert estimates for three different questions relating to the distinguishability of the
64 species, observer competence and verifiability, and pools across experts and questions linearly, while
65 another pools them logarithmically. The distributions from pooling across experts and questions
66 provide a quality distribution for each sighting, which we use in this paper. For clarity, we present
67 results from using the linear pooling distributions only, since, as can be seen from Lee et al. (2015),
68 the distributions are similar, and thus our conclusions will be similar.

69 We begin by examining these distributions (Section 2), where it is already implied that indi-
70 vidual distributions for the quality of each sighting may lead to counter-intuitive results, and thus

71 expert opinion on an individual sightings should be ignored. However, before confirming this con-
72 tentious statement, we incorporate these distributions into an existing extinction model to further
73 understand the effect of sighting quality scores on extinction estimates.

74 For the lion, we consider two methods to include the additional information about the quality of
75 uncertain sightings. Both these methods are extensions of the Bayesian model of Lee et al. (2014),
76 which assumes a constant population prior to extinction. The first method looks for a breakpoint
77 in the sighting quality, where one would assume that the average sighting quality before extinction
78 is higher than the average sighting quality after extinction (when all sightings should be false).
79 The assumption is that this breakpoint broadly coincides with the change in sighting rate. Quality
80 should inform when all sightings must be false, and vice versa. We refer to this method as the
81 ‘quality breakpoint method’. Alternatively, we use the sighting quality as a measure of uncertainty
82 around a sighting record. We refer to this method as the ‘quality as variance method’. To further
83 explore the effect of quality under this method, we also assign random quality measures, such that
84 if results from simulated random quality measures are similar to the results from actual quality
85 measures, then expert quality measures are superfluous.

86 To further demonstrate our methods on additional data sets we also consider three birds, the
87 Alaotra grebe, Jamaican petrel and Pohnpei starling. Since there are not quality distributions for
88 the individual sightings, for these three birds we use the uniformly distributed sighting qualities
89 provided by Birdlife International (Lee et al., 2014). There are fewer uncertain sightings with the
90 bird species, disqualifying them as a critical tests of the change point method (see Table 1). For
91 the three bird species, only the ‘quality as variance method’ is applied. As with the lion, the model
92 is also run with random quality measures to determine the importance of quality estimates.

93 As a small addition, we consider the finding of Lee et al. (2014) that the conclusions may

94 depend upon the prior. If one assumes that the prior of extinction is provided by an expert,
95 then perhaps this influence is welcomed. However in our method, for all four species, we use
96 an non-informative prior (Congdon, 2001) effectively integrating over all possible expert's views.
97 When inferring extinction for a given species it is recommended to always run a model with an
98 uninformative prior. If an expert prior is provided, an additional model with an uninformed prior
99 allows one to observe the effect of the expert's opinion.

100 Both models are based upon the model of Lee et al. (2014). The framework is presented in
101 Section 3. Within this section we examine the sighting quality, and identify a change point in
102 sighting quality for the Barbary Lion. In Section 4 we examine the choice of the expert's prior,
103 and discuss the influence it has on the outcome, and hence present an alternative, non-informative,
104 prior. In Section 5 we discuss our findings: sighting data that consists of certain and uncertain
105 only is the most reliable. Quality is not strongly informative of extinction.

106 Before discussing the models, it is illuminating to examine the information in sighting quality \mathbf{q}
107 itself first. Do any changes make rational sense? We do this with the lion sighting data, looking at
108 the general form of the continuous density assumed for $\mathbf{q} \in [0, 1]$ (where 1 is certain), and whether
109 \mathbf{q} exhibits a change point over time.

110 2 Examining sighting quality

111 The elicitation in Lee et al. (2015) was not carried out explicitly under a belief of extinction or
112 non-extinction. Five experts offered a best estimate and lower/upper bounds for three different
113 aspects of sighting quality (in an un-blinded manner) for each sighting at time t . Lee et al. (2015)
114 use the most straightforward way to represent these three points as a probability density, that is as
115 a triangle density. For simplicity we treat experts as exchangeable, ignore any correlation between

116 the best and lower/upper estimates, and also ignore any correlation between the j questions (the
117 differential weights of expert competency does something to adjust for inter-expert correlation as
118 does the exhaustive group elicitation process). The quality density for a given sighting $p(Q = q_t)$
119 is the result from linear pooling across questions and experts. Note that the distribution resulting
120 from pooling across 15 triangle densities, is not a triangle density. Under the Central Limit Theorem
121 of the sum of identical, independently distributions, one could work in accumulated (normalised)
122 quality measures and thus detect a level change rather than a breakpoint over time as we do herein.

123 Examination of the raw q_t values is very noisy. A degree of smoothing is needed to see the
124 choice of right density, any pattern (suggesting $p(\mathbf{q}|E)$, where E denotes extinction) is or is not
125 equivalent to $p(\mathbf{q}|notE)$ and any breakpoint that is informative of extinction. We assume that the
126 first sighting in 1895 (in Morocco) is certain. (Note that no sightings receive a quality score of one,
127 implying no sightings are defined as ‘certain’.) Sweeping across the sightings, sightings are classed
128 as either ‘before’ or ‘after’ the sighting in question, that is, ‘the current sighting’, where ‘after’
129 includes the current sighting. The ‘before’ sightings are combined, as are the ‘after’ distributions,
130 see Figure 1. Inspection suggests a unimodal distribution like a beta is a sensible choice for the
131 density. A one-sided t-test of the quality data in this way indicates that the before and after
132 distributions first become significantly different to each other in 1929.

133 Notice that the ‘before’ distribution has a large variance when examining the early sightings,
134 and a similar phenomenon for the ‘after’ distributions for the later sightings (Figure 1). This
135 is because at these extremes, we have less information. For example, when establishing whether
136 extinction occurred between 1895 and 1898 we are comparing the distribution for the single 1895
137 sighting with the combined distribution from the 31 other sightings.

138 The issue with this ‘burn-in’ and ‘burn-out’ is evident when examining the peak for the two

139 distributions. Ideally it would be clear to see that the peak (i.e. the mode) of the distribution for
140 before sightings is initially larger, then a switch occurs around 1929. However the lack of data at
141 the time boundaries makes this more challenging to clearly see from the peaks alone. For example,
142 around 1934 the before and after distributions seem very similar. Further smoothing is needed to
143 see any coherent changes.

144 So instead we consider the combination of the distributions presented in Figure 1. We denote
145 $\Theta_b(\textit{year})$ as the *peak* of the combined distributions (in Figure 1) before *year*. And similarly for
146 $\Theta_a(\textit{year})$ from combined distributions (in Figure 1) after *year*. With this measure it is apparent
147 that a shift in the relationship between these two values occurs around 1930 (Figure 2), as predicted
148 by the t-test.

149 The $\Theta_b(\textit{year})$ and $\Theta_a(\textit{year})$ allow us to better examine how the quality of sighting changes.
150 The mean of the combined quality of sightings ‘before’ for each sighting year $\Theta_b(\textit{year})$ increases
151 until 1929, and then the quality remains unchanged, Figure 2. Conversely, the combined ‘after’
152 distributions, $\Theta_a(\textit{year})$, remains reasonably steady and then increases. This phenomenon is unex-
153 pected since one would assume that after extinction, around 1930, the quality of sighting would
154 decrease. However, it is likely that this change is due to human factors such as observers still being
155 alive (first hand account), and the increased use of cameras. This already suggests that the quality
156 breakpoint method may be inconsistent with the sightings process.

157 Lastly, the empirical Bayes Factor (likelihood ratio) for just the quality data alone (below
158 labelled as \textit{data}_q) is the calculated ratio

$$BF_q = \frac{\ell(\textit{data}_q|\textit{not}E)}{\ell(\textit{data}_q|E)}. \quad (1)$$

159 If the species is extant, the $\ell(\text{data}_q|\text{not}E)$ is small, giving a log Bayes Factor that tends to negative
 160 infinity. At each sighting year we calculate the empirical Bayes Factor and find that before 1929
 161 the log Bayes Factor is indeed approximately zero. After 1929 the behaviour of the Bayes Factor
 162 changes. We would expect the Bayes Factor to steadily increase after 1929, however, again we
 163 observe how human factors have influenced the before and after distributions to create erratic
 164 behaviour. Nonetheless, as in the t-test there is clearly a shift around 1929 where quality increases.
 165 Could this be a technological change?

166 If the lion data is typical, changes in quality may not indicate the breakpoint for extinction,
 167 and thus the sighting quality alone is likely to be unreliable to infer extinction. Let us now consider
 168 using it in conjunction with the sighting record. Perhaps, together with analysis on sighting rate,
 169 quality scores can provide more information than either of sighting rate or sighting quality alone.

170 3 Model framework

The objective is to determine the estimated posterior probability distribution of extinction:-

$$p(\text{Extinction}|\mathbf{data}) = p(E|\mathbf{data})$$

171 By Bayes Theorem

$$p(E|\mathbf{data}) = \frac{l(\mathbf{data}|E) \cdot p(E)}{l(\mathbf{data}|E) \cdot p(E) + l(\mathbf{data}|\text{not}E) \cdot p(\text{not}E)} \quad (2)$$

172 where $p(E)$ is the expert's prior on extinction.

173 Let us retain the general form of the existing problem in Lee et al. (2014). That is, we consider

174 the period of observation $(0, T)$ where 0 is the beginning of the sighting record, and T is the length
 175 of the sighting record. During this observation period, certain and uncertain sightings occur in
 176 parallel. The vector \mathbf{s}_1 represents certain sightings $(s_{1,t})$ at time t , $t \leq T$. Similarly \mathbf{s}_2 represents
 177 uncertain sightings $(s_{2,t})$ at time t , $t \leq T$. Our input **data** comprises of both types of sightings,
 178 $\mathbf{s} = \mathbf{s}_1 \cup \mathbf{s}_2$. These sightings are used to estimate the posterior probability of extinction and the time
 179 at which this extinction occurs. Note that whilst the model does not require an uncertain record
 180 \mathbf{s}_2 , a certain record \mathbf{s}_1 with at least two sightings is required. Certain and uncertain sightings are
 181 assumed to follow a stationary Poisson process of regular spacing with constant, unknown rates
 182 (m_1 and m_2 respectively). Since we include the possibility of false sightings, sightings may occur
 183 after extinction, but at a different constant rate to that rate occurring before extinction. This is
 184 an offset denoted f_2 as a background for the whole series $(0, T)$. These false sightings by default
 185 must only be uncertain sightings. So, given the notation that l means likelihood we obtain the four
 186 elements of the model:

187 • $l(s_1|notE) = Bernoulli(1 - e^{-m_1})$

188 • $l(s_1|E) = Bernoulli(1 - e^{-m_1})$ if the time is before or equal to the time of extinction, with

189 $l(s_1|E) = 0$ if afterwards

190 • $l(s_2|notE) = Bernoulli(1 - e^{-f_2 - m_2})$

191 • $l(s_2|E) = Bernoulli(1 - e^{-f_2 - m_2})$ if the time is before or equal to the time of extinction, with

192 $l(s_2|E) = Bernoulli(1 - e^{-f_2})$ if afterwards.

193 The form $(1 - e^{-\cdot})$ is used for efficient parameterisation. We assume no population decline (see
 194 Lee et al. (2014) and references therein). This model of Lee et al. (2014) determines a change point
 195 in the sighting rate, which provides an estimate for the year of extinction. The input is two sighting

196 records (certain and uncertain) and the output is a probability that the species is extinct at the
197 end of the sighting record, and a corresponding year in which extinction would have occurred. Our
198 method uses:- a uniform distribution from the last certain sighting to the end of the sighting record;
199 a non-informative Jeffreys prior ($Beta(0.5, 0.5)$) for non-extinction (Congdon, 2001); together with
200 wide uniform prior distributions (range 0 to 100) for m_1 , m_2 and f_2 to ensure that there is no bias.

201 Now consider that our input **data** comprises of both certain and uncertain sightings, $s = \mathbf{s}_1 \cup \mathbf{s}_2$,
202 and the individual quality of uncertain sightings, \mathbf{q} . We interpret individual quality scores as a
203 score for the year in which the sighting occurred. However, we require a quality score for every
204 year, even if no sighting occurred. We take a wide interpretation of what quality is. One could
205 infer that a high quality sighting (e.g. a skin sample that can be tested) implies that the sighting
206 is certain (or close to certain), and conversely, a low quality sighting (e.g. a second-hand verbal
207 account) is less certain to be a true sighting. Note that the quality vector \mathbf{q} initially seems to only
208 have a quality assigned during years of uncertain sightings. Later we will discuss how we assign
209 a sighting quality for all other years. We take quality simply to be a subjective attribute of the
210 sighting - no implicit model of its basis is assumed. A method for eliciting the quality measures is
211 provided in Lee et al. (2015).

212 Let us partition the **data** into the stochastic sightings, \mathbf{s} , and the stochastic quality measure
213 for each of the uncertain sightings (\mathbf{q}), measured and stochastically modelled simultaneously. Note
214 that only $\mathbf{q} \in [0, 1)$ are used as quality measures, which only relate to uncertain sightings. The
215 sighting quality \mathbf{q} is thus in a sense a nuisance variable which we take as *Beta* distributed, which
216 ensures that it is bounded between $[0, 1]$ as required. We use non-informative *Exponential*(1)
217 and Jeffreys ($Beta(0.5, 0.5)$) priors for its parameters using the Stroud (1994) method in Congdon
218 (2001). Equation (3) in full would rely upon specifying a general stochastic model for the quality

219 measure $p(\mathbf{Q} = \mathbf{q})$ under the alternative hypothesis of extinction versus that under non-extinction.

220 One would expect that after extinction, sighting quality drops, yet Section 2 proved this may
 221 not be the case. So for simplicity (and to avoid specifying even more unknown priors in the
 222 computations to integrate over) we assume herein that the probabilistic generating process for
 223 quality attribution is unaffected by whether the species is extinct or not, unless the sighting is
 224 deemed as ‘certain’. Formally:- $l(\mathbf{q}|E) = l(\mathbf{q}|notE) = l(\mathbf{q})$. This approximation is reasonable since
 225 the alternative requires an estimate for the error process arising from experts assigning quality
 226 measures for $l(\mathbf{q}|E)$ and $l(\mathbf{q}|notE)$, which needs repeated blinded data for a variety of species
 227 whose extinction date and status were well known but that investigative adjudicators had little
 228 experience about before. A significant practical challenge.

For the quality breakpoint method, we assume that the probabilistic generating process of sightings is unrelated to the generating process of quality attribution, that is, $p(\mathbf{data}) = p(\mathbf{s} \& \mathbf{q}) = p(\mathbf{s}) \cdot p(\mathbf{q})$. The two processes are taken to be independent. So integrating out the uncertainty of the nuisance gives

$$l(\mathbf{data}|E) \rightarrow \int_0^1 l_s(\mathbf{s}|E) \cdot l_q(\mathbf{q}|E) dp(\mathbf{q}), \quad (3a)$$

$$l(\mathbf{data}|notE) \rightarrow \int_0^1 l_s(\mathbf{s}|notE) \cdot l_q(\mathbf{q}|notE) dp(\mathbf{q}), \quad (3b)$$

229 which are the weighted sums of likelihoods over the stochasticity of the quality measure at $\mathbf{Q} =$
 230 \mathbf{q} , where \mathbf{Q} is a particular realisation of \mathbf{q} . Conjugacy makes this integration efficient. The
 231 likelihoods (3) can be fed into equation (2) to yield the extinction posterior for sightings $p(E|\mathbf{s})$. The
 232 same four-part structure of the approach (with subscript 1 indicating certain sightings and subscript
 233 2 indicating uncertain sightings) would now follow but with *Beta* likelihoods for \mathbf{q} . However, the

234 quality measures are for uncertain sightings only, thus:

235 • $l(q_2|notE) = Beta(\alpha + \alpha_2, \beta + \beta_2)$

236 • $l(q_2|E) = Beta(\alpha + \alpha_2, \beta + \beta_2)$ if the time is before or equal to the time of extinction, with

237 $l(q_2|E) = Beta(\alpha, \beta)$ if afterwards.

238 The overall form for estimation over $\mathbf{s}\&\mathbf{q}$ is then comparable to that of Lee et al. (2014) and
239 can be fitted using the same OpenBugs (2012) approach via Markov Chain Monte Carlo (MCMC)
240 integration. Any change point in \mathbf{q} will reinforce or conflict with any change point in \mathbf{s} in the overall
241 optimistaion for $p(E|\mathbf{data})$. Later work may investigate our assumption that sighting rate and
242 quality scores are independent. For example, perhaps an event occurs (change in IUCN classification
243 or a reward offered) that causes an influx of low quality sightings.

244 For the quality as variance method we assume that $p(\mathbf{data}) = p(\mathbf{s}|\mathbf{q}) \cdot p(\mathbf{q})$ and solve accordingly.
245 In this case $p(\mathbf{q})$ is taken to be a *Gamma* distributed expansion/shrinkage factor to the variance
246 of the rate of the uncertain (only) Poisson distributed sightings \mathbf{s}_2 above. Full calculation detail
247 is given in a later section. Again, later work may test whether it is appropriate to view $(\mathbf{s}\&\mathbf{q})$ as
248 independent (conditional on extinction) and this Bayesian model extended.

249 One thus chooses between: The changes in sighting quality inform the inference of whether
250 extinction has occurred by directly affecting the likelihood of the sightings (the quality breakpoint
251 method). Or, that quality is a proxy for the variance around the degree of certainty of the uncertain
252 sightings and so affects the likelihood of the sightings indirectly (the quality as variance method).
253 In either case, integrating over the nuisance modifies the relative sightings likelihoods which arise
254 from the simple approach of assuming all sightings are of similar quality used in Lee et al. (2014).

255 3.1 Bayesian modelling: Quality data

256 Certain and uncertain sightings are interlaced over time, yet for their joint Bayesian modelling a
257 quality measure is needed at each time point for either type of sighting. For implementation on
258 the lion data we assume the first sighting and the most certain sighting (in 1925) are both certain,
259 that is they have a quality score of one and the lion is assumed extant in 1925. The remaining
260 sightings are left as uncertain.

261 By the nature of the logic of the observing process, there are some *per force* missing quality
262 values. Accordingly a form of Last Observation Carried Forward (LOCF) for uncertain qualities is
263 used to fill in any such missing quality values. However, LOCF is modified such that observations
264 carried forward are randomly drawn from the quality density from the last sighting $p(Q = q_t)$.
265 Furthermore, since the method of Lee et al. (2014) requires a certain sighting at the beginning,
266 and an uncertain sighting may not occur for some years after this initial sighting, there needs to
267 be a quality measure for these unobserved values. As with LOCF, a form of First Observation
268 Carried Backwards (FOCB) is deployed such that the quality for previous years is drawn from
269 the density for the first sighting. Therefore, due to our modified form of FOCB and LOCF, the
270 quality density of the first uncertain observation is used from the first (certain) sighting until the
271 second uncertain observation, where the mean quality is used at the time of the first sighting. The
272 argument is that whilst the mean from the quality density for the first uncertain sighting is the
273 closest to the unknown quality of ‘never seen’ uncertain sightings during this time period, there is
274 still uncertainty around this value, so using information from the distribution as a whole is more
275 appropriate. Using quality information from the whole time period (not just the quality from the
276 first sighting quality) would be using information from a significantly different time period and
277 induce bias. Of course a more sophisticated stochastic model for missing quality data could be

278 posed.

279 **Bayesian modelling: Quality as a breakpoint method**

280 Exactly the same approach as in Lee et al. (2014) is used in which a change in sighting rate is
281 sought, which infers an extinction time. The model provides a probability and variance around
282 this estimate, that is, the probability that the species is extinct, and the variance around this
283 probability.

284 Now due to the independence of \mathbf{s} and \mathbf{q} , simply a second simultaneous Bayesian optimisation
285 of a beta distributed quality variable over time is made around a common extinction point with
286 the sightings. A non-informative hyper-prior is used. For demonstration purposes, here the model
287 is run on the data as it stands every year after the last certain sighting, up to 2016. This allows us
288 to examine the effect of additional uncertain sightings.

289 For each run of the model, the probability that the lion is extinct (the posterior) and the
290 standard deviation around this estimate is noted. When using the data set as it stood in 2016, we
291 also note the corresponding inferred extinction time. We also make a note of the corresponding
292 inferred extinction time when the posterior first overwhelms the experts' prior belief of extinction.

293 **Bayesian modelling: Quality as variance method**

294 Let us return to the Bayesian model of Lee et al. (2014). The rate of uncertain sightings is assumed
295 to follow a Poisson distribution with rate $m_2 + f_2$ where m_2 is the rate of true uncertain sightings
296 and f_2 is the rate of false uncertain sightings, such that a change point indicates when m_2 sightings
297 have ceased, that is, extinction has occurred. The rate of certain sightings m_1 is consistent until
298 extinction.

299 The same model and computational algorithm as in Lee et al. (2014) is used. No attempt is
300 made to model an exponentially declining quality post extinction for data deemed to be extant yet
301 observed after the extinction time - rather a common offset is used over all cases. Unlike Lee et
302 al. (2014), here we use vague parameter priors throughout. The sensitivity of the initial prior on
303 being extant is explored with the lion species only, since the effect of the prior on the bird species
304 data sets has already been explored by Lee et al. (2014).

305 Instead of seeking a breakpoint in the quality \mathbf{q} , we slightly relax the assumption that quality and
306 sightings are independent. Whilst we maintain that the occurrence of sightings, and the quality of
307 sightings are independent, we now incorporate the quality of a sightings as unique variance around
308 each sighting. Quality then behaves like a fractional replication factor. To incorporate quality,
309 each uncertain sighting follows a distribution with expected rate as before but with a variance that
310 increases as the quality of the sighting decreases. So an uncertain observation at time t is ‘fuzzed’
311 by a Gamma distribution, such that the ‘fuzzed’ rate of uncertain sightings is

$$m'_2 = m_2 r, \quad (4)$$

312 where r is a random variable. The random variable r is drawn from a Gamma distribution of
313 mean 1

$$r \sim \Gamma(1/H_t, 1/H_t), \quad (5)$$

314 where $H_t = -a \ln(q_t)$, q_t is the corresponding beta distributed $(0, 1]$ quality score for the sighting,
315 and a is a penalty factor such that large a penalises low quality sightings more. In doing this we
316 are using the log link model (McCullagh & Nedler, 1989) philosophy. In this, $a = 1, 2, 4$, where a

317 is a penalty variable which quantifies the relationship between the measured quality and the ‘fuzz’
318 applied to the model. Large values of a model when a small change in quality produces an even
319 large uncertainty i.e. uncertainty is inflated. If sighting quality is not considered important, a is
320 small such that at $a = 0$ and it reverts back to the model of Lee et al. (2014).

321 Equation (4) and (5) ensures: the rate remains positive; the variance around uncertain sightings
322 is $-a \ln(q_t)$; and when sightings are certain ($q_i = 1$), the rate is not fuzzed, $m'_2 = m_2$. Note that the
323 variance is not added to f_2 , the false sightings offset. Under this adaptation of the model, f_2 can
324 be thought of as a constant characteristic background rate of sightings for the whole data set that
325 remains unchanged throughout the whole period - an attribute which changes solely from species
326 to species.

327 As with the quality as a breakpoint method, the model is run on the data as it stands every
328 year after the 1925 sighting. And for each run of the model, the probability that the lion is extinct
329 (the posterior) and the standard deviation around this estimate is noted. When using the data set
330 as it stands in 2016, we also note the corresponding inferred extinction time. We also make a note
331 of the corresponding inferred extinction time when the posterior first overwhelms the experts’ prior
332 belief of extinction. We also run this model on the three bird species. The same output details are
333 recorded as with the lion species.

334 4 The choice of prior

335 As previously shown, the method of Lee et al. (2014) is affected by the experts’ prior. This
336 phenomenon is presented for the bird data sets in Lee et al. (2014), and with the Barbary lion
337 data set in Figure 3. Whilst the three different choices of prior are all initially overwhelmed by
338 the posterior in 1953 (that is, the extinction probability is larger than the prior), the estimated

339 extinction year that corresponds with the data set as it stood in 1953 is different. When the model
340 is run with the data that exists up to 1953 only, with a prior belief of extinction of 0.9, the model
341 infers extinction occurred in 1936; with a prior of 0.5, the model infers extinction occurred in 1931;
342 and with a prior of 0.1, the model infers extinction occurred in 1925, which is when the ‘certain’
343 sighting of 1925 occurs, so is clearly incorrect.

344 Using the full data set as it stands in 2016 only includes one additional sighting after 1953, a
345 sighting in 1956. The estimated extinction year does not vary between the three prior choices: a
346 prior belief of extinction of 0.9, 0.5 and 0.1 all predict extinction in 1954. Therefore the model
347 assumes that the 1956 sighting is false. Additionally, the results are less influenced by the prior
348 when the model is run with more data, as one would expect.

349 To avoid the bias towards the expert’s prior choice, for all four species, we use a non-informative
350 hyper-prior - the beta distribution $p(E) \sim Beta(0.5, 0.5)$ (Congdon, 2001) and optimise accordingly.
351 This distribution actually slightly favours the extremes equally, that is the species is definitely
352 extinct $p(E) = 0$ or definitely extant $p(E) = 1$, and all other possibilities are equally likely to
353 each other. The expected value of this density is 0.5 i.e. a 50:50 evens bet of prior ignorance.
354 The model from Lee et al. (2014) applied to the lion data set, with this improvement, provides
355 similar results to when a prior of 0.5 is used, as one would expect, see Figure 4a. The first year
356 that the likelihood overwhelms the average of the prior, 0.5, occurs in 1953 again, which gives a
357 corresponding year of extinction of 1931. When the model is run on the sighting record as it stands
358 in 2016, the corresponding extinction year is 1954. Although these results are similar to a prior of
359 0.5, a non-informative prior is preferred to a point estimate as it integrates over all uncertainties
360 appropriately.

361 5 Results

362 Lee et al. (2014) considered the rate of certain and uncertain sightings, and sought to find a change
363 point in the sighting rate, thus the model requires as least two certain sightings. The individual
364 quality of each sighting was not considered. As discussed in the previous section, using this approach
365 the probability that the Barbary lion is extinct is not above the average of the experts' prior until
366 after 1953 (Figure 4a), which has a corresponding extinction date estimate of 1931 approximately,
367 the same time period that there is a change in the quality of the data (Figure 2).

368 When seeking a change in the quality ('quality as a breakpoint method'), there is low over-
369 all information, that is, it is only after an absence of sightings for 13 years after the very last
370 sighting when the likelihood first overwhelms the prior (Figure 4b). This is inconsistent with the
371 sightings information. The model cannot identify a common change zone for sighting rate whilst
372 simultaneously identifying the change in sighting quality.

373 Continuing with the lion, we used the sighting quality as a variance method with both expert
374 estimates (Figure 4c) and random estimates (Figure 4d) for the sighting qualities. Both choices
375 provide results very similar to the model of Lee et al. (2014), which does not include sighting
376 quality (Figure 4a). The posterior from all three models first overwhelm the average of the prior in
377 1953, and, at this point, estimate extinction to have occurred around 1931. As before, the effect of
378 the quality is apparent when the full data set up to 2016, is used. The additional sighting in 1956
379 affects the sighting rate, causing the change point in sighting rate to shift to 1954.

380 For the bird species, there were not enough uncertain sightings to seek a change point in
381 sighting data, see Table 1. Therefore, we only use the method where quality is implemented as
382 a variance. As with the lion, we find similar results between omitting sighting quality, using the
383 sighting quality estimates, and using random sighting quality estimates, see Figure 5. Moreover,

384 the Pohnpei starling demonstrates the challenge of trying to infer an extinction date when there is
385 a paucity of data, since the likelihood never overwhelms the prior. This is a similar problem to the
386 ‘quality breakpoint method’ (Figure 4b) where a long period of no sightings was required.

387 Reducing the weighting penalty, a , further reduces the effect of the sighting quality, whereas
388 increasing a increases the standard deviation around each annual estimate. When a is increased
389 to, say, $a = 10$, the change point in the sighting record dissipates into fuzz, and thus cannot be
390 determined. That is, everything becomes uncertain.

391 There is a consistent similarity of extinction inferences over all species between using actual
392 quality measures and simulated random quality scores. Since it was shown in Section 2 that the
393 quality distributions are non-random through time, this points to the lack of sensitivity to expert
394 opinion for the inference of extinction.

395 **6 Discussion**

396 We have presented a method that considers the quality of each sighting individually. Previous work
397 has initially used sightings that are classed as either certain or uncertain (Solow et al., 2012 and Lee
398 et al., 2014). Further work then sought to divide uncertain sightings further, into several categories
399 (Lee, 2014). However, each additional classification provides less information about the rate of
400 occurrence for this particular class of sighting (law of diminishing returns). To avoid continually
401 dividing uncertain sightings into more specific categories we investigated a method that sought to
402 find the change point in the continuous quality of sightings. This required several experts to rate
403 all sightings of a species. As discussed in previous work (Lee et al., 2015), the method by which
404 the experts are questioned is important to the outcome.

405 There is a change in sighting quality scores at 1930, where sightings after this date have a per-

406 ceived higher quality. However the change in sighting quality may be picking up on our preference
407 to believe accounts from living observers more than records left by deceased observers. There is
408 a large literature on unbiased Bayesian elicitation methods to help avoid this (see review article
409 Kuhnert et al., 2010). Being blinded to date and the historical age of sightings is important - how-
410 ever this is difficult. Technological changes over time are apparent even when species observations
411 are not explicitly labelled (an old photo is clearly an old photo). Whilst human nuances affect
412 all sighting records, and thus all extinction models, methods which rely more heavily on expert
413 opinion may be more susceptible to these external factors.

414 Establishing the balance between an extinction model with assumptions that over simplify, and
415 a model that seeks to incorporate everything, is discussed by Caley and Barry (2014), where the
416 authors develop an extinction model that does not constrict the species population to be constant
417 (as assumed here), nor declining. In line with their findings, our work also suggests that a simple
418 model makes it easier to identify the underlying population processes. If quality is to be used, the
419 quality as variance method is recommended.

420 We have shown that the rate of sightings is the strongest indicator to infer extinction, and too
421 much information about the quality of the sighting can actually be detrimental. Ideally a sighting
422 record would be a list of certain and uncertain sightings only. Using only these two parallel sighting
423 records, a Bayesian model (Lee et al., 2014 or Thompson et al., 2013), with the non-informative
424 prior presented here, could establish a synchronised change point to infer an extinction date. No
425 further classification, nor prior belief about extinction, is required.

426 The less propensity for human influence in the sighting data the better. As such, an objective
427 method of assessing whether sightings are certain or uncertain is needed. However, even the rate of
428 sightings is susceptible to human influence, such as periods of high interest due to publicity. Future

429 work could try to quantify the effect of publicity on sighting rate.

430 Acknowledgements

431 The authors thank Dr. Simon Black for suggesting the Barbary lion as an exemplar and providing
432 his expertise.

433 References

434 [1] Boakes, E.H., T.M. Rout, & B. Collen. (2015). Inferring species extinction: the use of sighting
435 records. *Methods in Ecology and Evolution*, **6**, 678–687.

436 [2] Caley, P. & Barry, S. C (2014) Quantifying extinction probabilities from sighting records:
437 Inference and uncertainties *PloS One*, **9(4)**, e95857.

438 [3] Congdon P (2001) *Bayesian statistical modelling* John Wiley and Sons, Chichester UK

439 [4] Jarić, I., & Roberts, D. L. (2014). Accounting for observation reliability when inferring extinc-
440 tion based on sighting records. *Biodiversity and Conservation*, **23(11)**, 2801–2815.

441 [5] Kuhnert, P. M., Martin, T. G., & Griffiths, S. P. (2010). A guide to eliciting and using expert
442 knowledge in Bayesian ecological models. *Ecology letters*, **13(7)**, 900–914.

443 [6] Lee, T. E. (2014). A simple numerical tool to infer whether a species is extinct. *Methods in*
444 *Ecology and Evolution*, **5(8)**, 791–796.

445 [7] Lee, T. E., McCarthy, M. A., Wintle, B. A., Bode, M., Roberts, D. L., & Burgman, M. A.
446 (2014). Inferring extinctions from sighting records of variable reliability. *Journal of Applied*
447 *Ecology*, **51(1)**, 251–258.

- 448 [8] Lee, T. E., Black, S. A., Fellous, A., Yamaguchi, N., Angelici, F. M., Al Hikmani, H., Reed,
449 J.M., Elphick, C.S., and Roberts, D.L. (2015) Assessing uncertainty in sighting records: an
450 example of the Barbary lion. *PeerJ* **3** : e1224
- 451 [9] McCullagh, P and Nelder, JA (1989) *Generalized Linear Models* Second Edition (Chapman &
452 Hall/CRC Monographs on Statistics & Applied Probability)
- 453 [10] Solow, A., Smith, W., Burgman, M., Rout, T., Wintle, B., & Roberts, D. (2012). Uncertain
454 sightings and the extinction of the IvoryBilled Woodpecker. *Conservation Biology*, **26(1)**,
455 180–184.
- 456 [11] Thompson C. J., Lee T. E., Stone, L. M, McCarthy, M. A. & Burgman, M. A. (2013). Inferring
457 extinction risks from sighting records. *Journal of Theoretical Biology* **338**, 16–22.

Alaotra grebe		Jamaican petrel		Pohnpei starling	
Year	Quality	Year	Quality	Year	Quality
1929	1	1789	0.4-0.8	1930	1
1947	0.4-0.8	1829	1	1995	1
1960	1	1847	0.4-0.8	2008	0.4-0.8
1963	1	1866	1		
1969	1	1879	1		
1970	0.1-0.4	1891	0.8-0.9		
1971	0.1-0.4				
1972	0.6-0.8				
1982	0.6-0.8				
1985	0.6-0.8				
1986	0.1-0.4				
1988	0.1-0.4				

Table 1: Sighting data for three bird species. A quality score of 1 indicates a certain sighting.

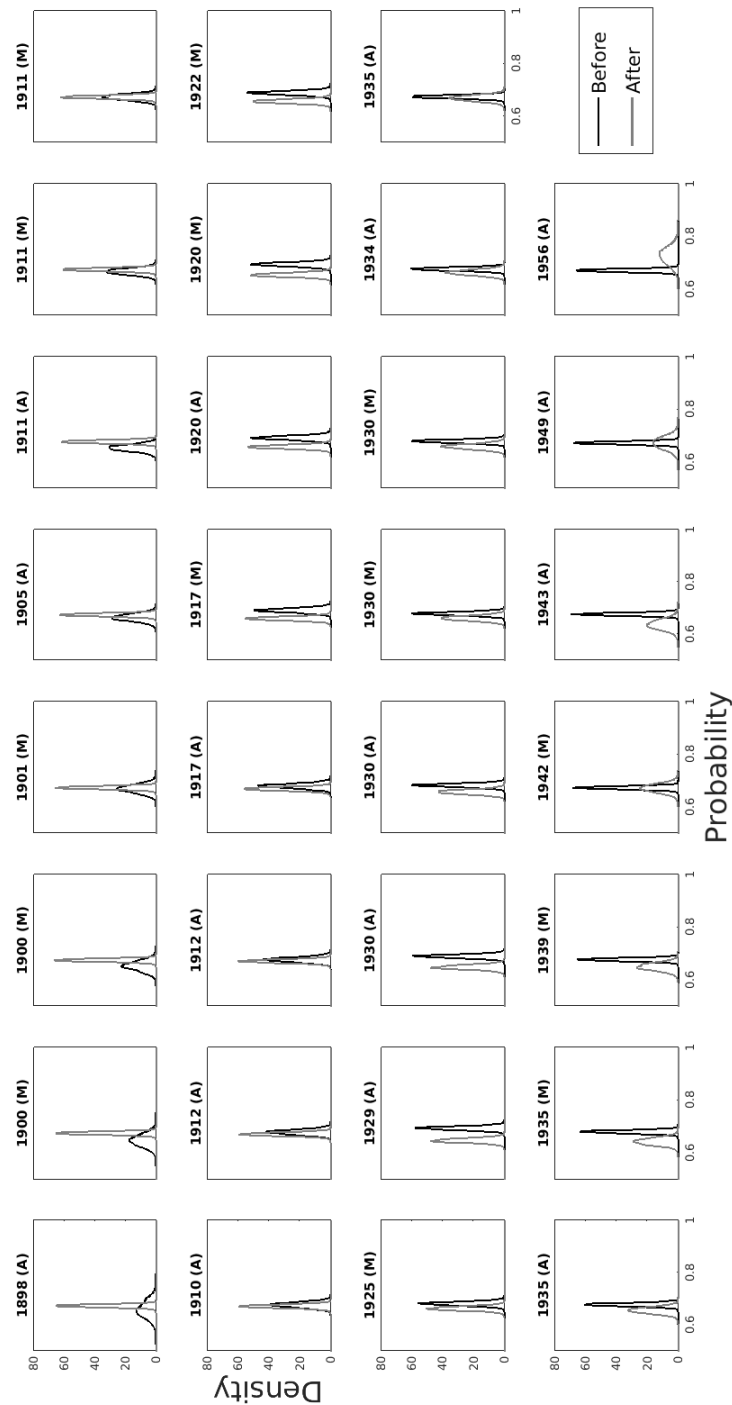


Figure 1: Comparing before and after distributions for the pooled quality densities of each Barbary lion sighting. The year and location (Morocco/Algeria) of the current sighting are listed along the top. The current sighting is grouped with the 'after' distribution. The first sighting is 1895 (M). A one sided t-test states that the before and after distributions first become significantly different from each other in 1929.

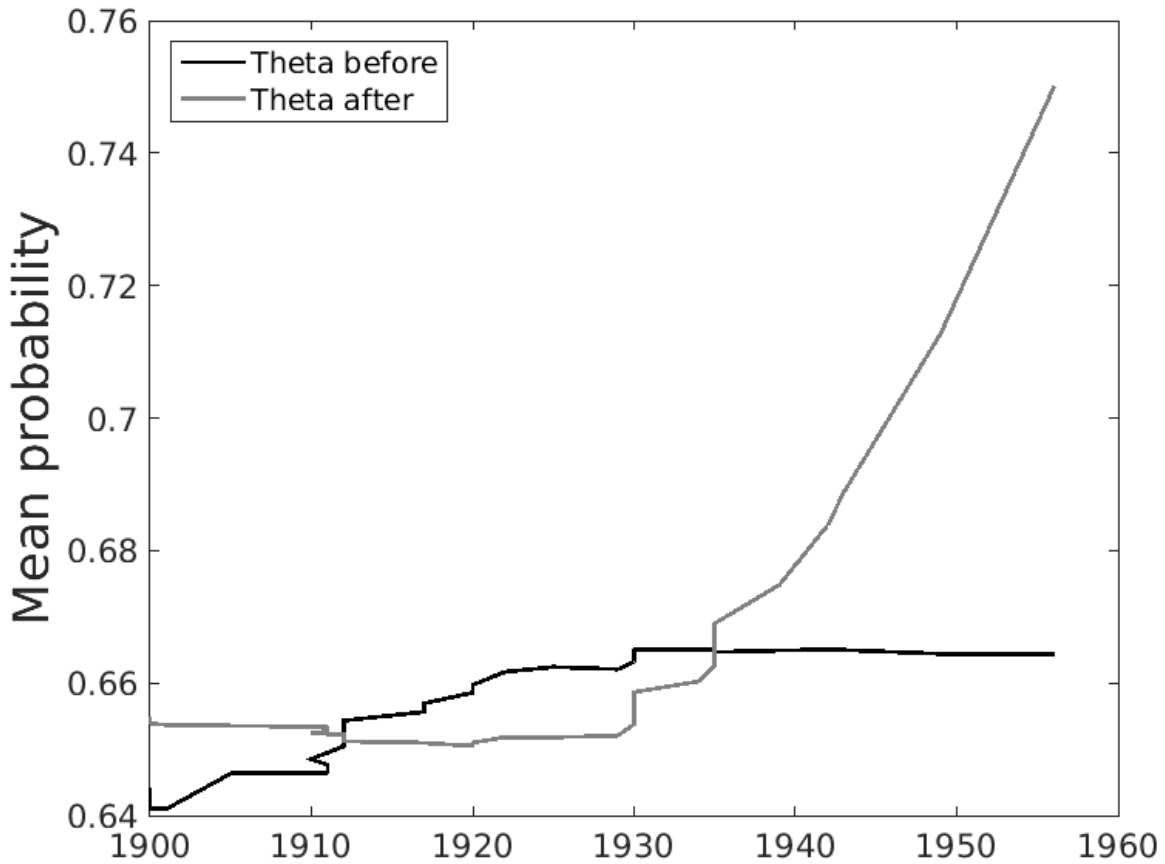


Figure 2: For the Barbary lion data set, the combined average of the sighting record quality before $\Theta_b(\text{year})$ and after $\Theta_a(\text{year})$ each sighting.

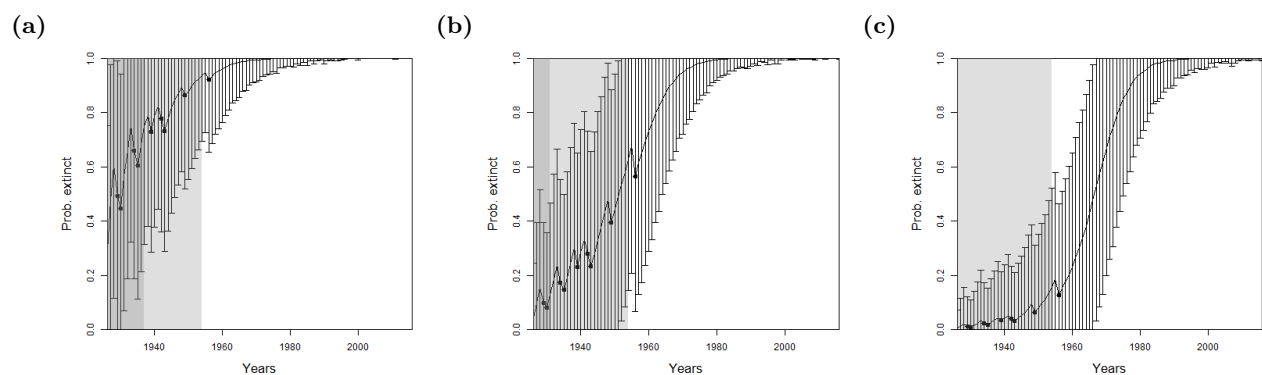


Figure 3: Results for the Barbary lion from the method of Lee et al. (2014), where sightings are divided into certain and uncertain over a varying expert's value for prior of extinction: (a) 0.9 (b) 0.5 (c) 0.1. The bars represent the standard deviation around the estimate and the black circles indicate when an uncertain sighting occurred. Note changes to the standard deviation of estimates as well as the curve translocation. The light shaded region marks the predicted extinction year identified by the model in 2016 (using the full dataset), which is in 1954 in all cases. The dark shaded region marks the predicted extinction year identified when the likelihood first overwhelms the prior, i.e., the extinction date inferred when there is enough data, which is in 1953 in all cases, giving a prediction extinction year of (a) 1937 (b) 1931 (c) 1924 (before the last certain sighting of 1925 so is clearly incorrect).

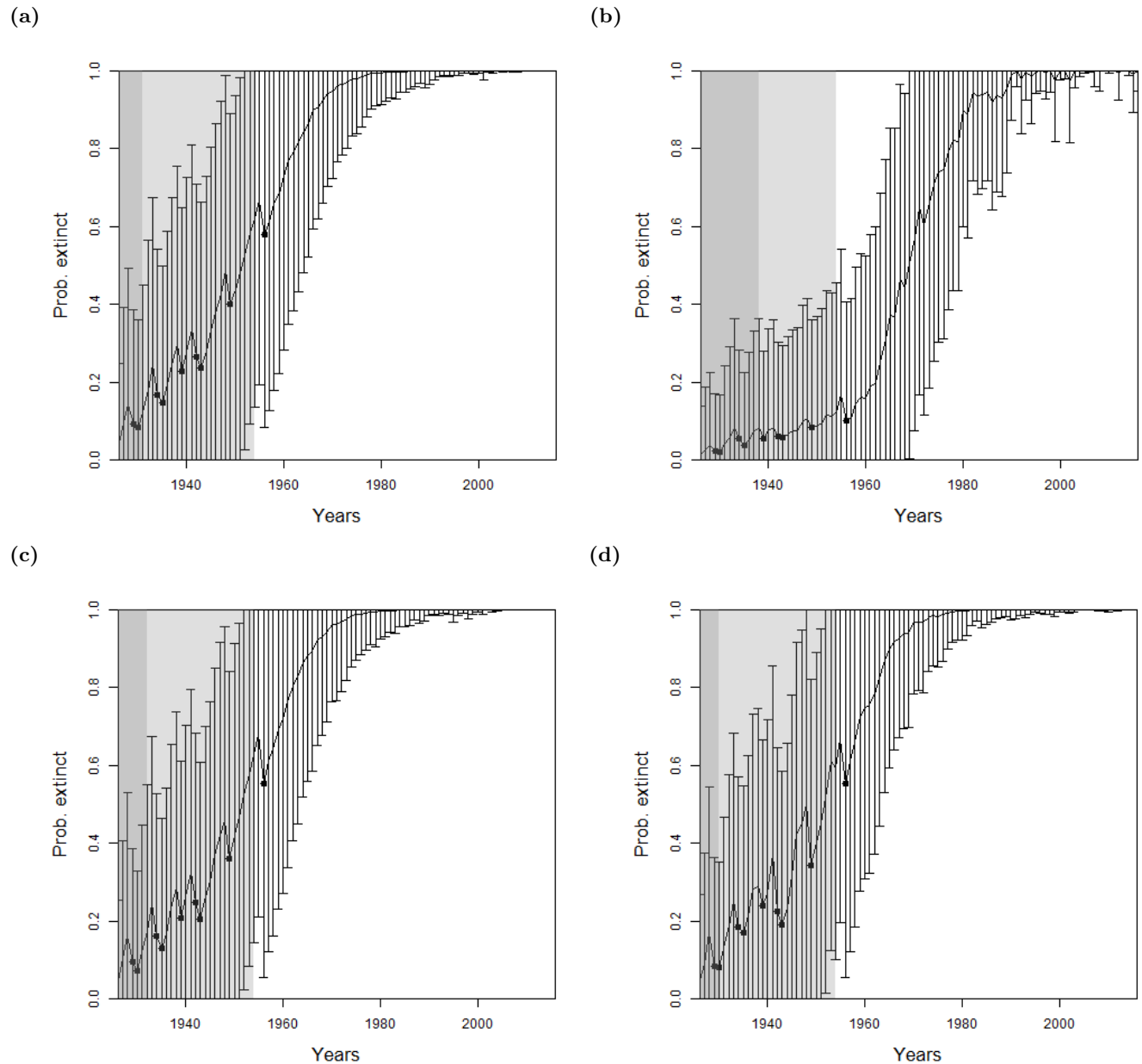


Figure 4: The probability of extinction for every year after the last certain sighting (using data only to that particular year) for the Barbary lion from (a) the method of Lee et al. (2014) where sightings are divided into certain and uncertain now with a non-informative hyper-prior of extinction, (b) the ‘quality break point method’, the ‘quality as a variance method’, where the quality is either (c) provided by experts, or (d) the quality is random. The bars represent the standard deviation around the estimate and the black circles indicate when an uncertain sighting occurred. The light shaded region marks the predicted extinction year identified by the model in 2016 (using the full dataset), which is in 1954 in all cases. The dark shaded region marks the predicted extinction year identified when the likelihood first overwhelms the prior, i.e., the extinction date inferred when there is enough data, which is in (a, c, d) 1953 and (b) 1969, giving a prediction extinction year of (a) 1878 (b) 1938 (c) 1932 (d) 1930.

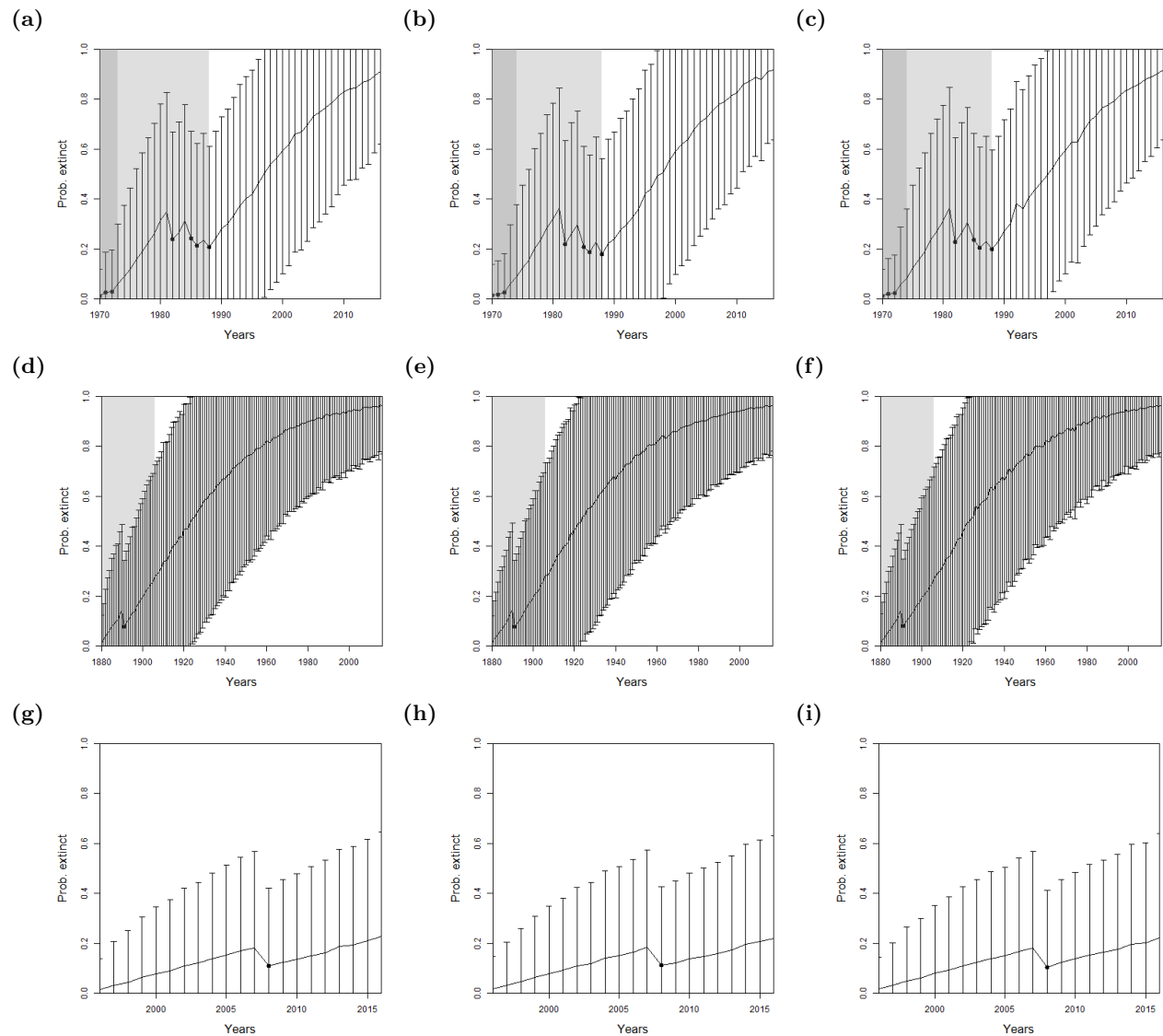


Figure 5: The probability of extinction for every year after the last certain sighting (using data only to that particular year), with the ‘quality as a variance method’, where the quality is either (a,d,g) excluding quality, or using the quality as a variance where the quality is either (b,e,h) provided by experts, or (c,f,i) the quality is random. The bars represent the standard deviation around the estimate and the black circles indicate when an uncertain sighting occurred. All models use a non-informative hyper-prior on extinction. When quality is included, a penalty $a = 2$ is used. The light shaded region marks the predicted extinction year identified by the model in 2016 (using the full dataset). The dark shaded region marks the predicted extinction year identified when the likelihood first overwhelms the prior, i.e., the extinction date inferred when there is enough data. Alaoira grebe (a,b,c): the likelihood first overwhelms the prior in 1998, with a corresponding extinction year of 1973 (a) or 1974 (b,c). With the full dataset, extinction is estimated to have occurred in 1988 in all three cases. Jamaican petrel (d,e,f): the likelihood first overwhelms the expected value of the prior in 1925 with a corresponding extinction year of 1878, before the last certain sighting. With the full dataset, extinction is estimated to have occurred in 1906. Pohnpei starling (g,h,i): The likelihood never overwhelms the prior meaning that there is not enough data.