Journal of Theoretical Biology 442 (2018) 98–109

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/jtbi

Inferring about the extinction of a species using certain and uncertain sightings

Saritha Kodikara[a,*], Haydar Demirhan[a], Lewi Stone[a,b]

[a] Mathematics, School of Science, RMIT University, Melbourne, Australia
[b] Biomathematics Unit, Department of Zoology, Faculty of Life Science, Tel-Aviv University, P.O.Box 39040, Tel-Aviv 69978, Israel

ARTICLE INFO

Article history:
Received 25 August 2017
Revised 21 December 2017
Accepted 15 January 2018

Keywords:
Bayes factor
Extinction
Poisson process

ABSTRACT

The sighting record of threatened species is often used to infer the possibility of extinction. Most of these sightings have uncertain validity. Solow and Beet(2014) developed two models using a Bayesian approach which allowed for uncertainty in the sighting record by formally incorporating both certain and uncertain sightings, but in different ways. Interestingly, the two methods give completely different conclusions concerning the extinction of the Ivory-billed Woodpecker. We further examined these two methods to provide a mathematical explanation, and to explore in more depth, as to why the results differed from one another. It was found that the first model was more sensitive to the last uncertain sighting, while the second was more sensitive to the last certain sighting. The difficulties in choosing the appropriate model are discussed.

1. Introduction

Mathematical models have become of importance for conservation scientists and policy makers since they offer a means to estimate the probability or likelihood that a species has gone extinct based on its historical sighting record (Jarić and Roberts, 2014; Solow, 1993; Solow and Beet, 2014; Thompson et al., 2013). The probability of extinction is an important quantity since it helps make informed decisions about conservation policies for threatened species. This could include deciding on whether to conduct expensive ecological surveys for the purposes of monitoring the species, and whether to continue other efforts to preserve a species. A flawed assessment could lead to failure in protecting a designated species.

The sighting record of a species compiles evidence of those years a species has been observed (although with varying degrees of uncertainty), and in some situations the record can extend over many decades. Rivadeneira et al. (2009) pointed out that statistical models that assessed the extinction status of species developed before 2009 were based on all sightings being valid with certainty. Roberts et al. (2010) showed that the inferences made from models that include uncertain sightings significantly differ from those omitting that information, and the conclusions about extinction were sensitive to the inclusion or exclusion of uncertain sightings. This led Solow et al. (2012) to develop a statistical method which treated uncertain sightings of the Ivory-billed Woodpecker in a formal way which neither simply treated them as valid nor excluded them. The limitation of the Solow et al. (2012) method is that it assumes that uncertain sightings occur only after the certain sightings, which is not generally the case in practice. Afterwards, Solow and Beet (2014) found a way to modify their method into two different probability models that allows overlap in time between certain and uncertain sighting.

Recently, several research groups have developed methods that take into account uncertainties and the actual strength of the evidence, by incorporating additional information, for example as to whether actual specimens of the species were recorded, or whether less certain video and/or audio recordings were collected, or whether there were just local verbal reports from experts or unreliable non-experts in some years (Lee, 2014; Thompson et al., 2017; 2013). This idea was then extended by assigning the probabilities of reliability to individual sightings (Jarić and Roberts, 2014; Lee et al., 2014). All of these models needed some expert (e.g Bird Life International) in the area to provide these sighting reliabilities and the inferences made from these models were sensitive to the sighting reliabilities. Because of the importance of these sighting reliabilities, Lee et al. (2015) developed a formal framework to elicit expert opinions in order to determine the validity of sightings. However, the model in Solow et al. (2012) and the two models in Solow and Beet (2014) differ from this approach. Solow et al. (2012) as well as Solow and Beet (2014) assume that certain sightings occur at a constant but unknown rate

* Corresponding author.
E-mail address: sarithakalhari.kodikara@rmit.edu.au (S. Kodikara).

https://doi.org/10.1016/j.jtbi.2018.01.015

Fig. 1. Graphical representation of Ivory-billed Woodpecker sighting data. Green represents the years where there are certain sightings followed by pink which represents the uncertain sightings. Note the period of overlap. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and may be modeled as a Poisson process until extinction occurs. Uncertain sightings are considered as either valid or invalid and both valid and invalid sightings are also modeled as different Poisson point process with unknown rates (i.e Under Model 1 in Solow and Beet, 2014 the certain sighting rate is equivalent to the valid uncertain sighting rate). Thus both models in Solow and Beet (2014) and the model in Solow et al. (2012) do not need to rely on expert opinions, but on the other hand an underlying assumption that rates of certain and uncertain sightings, and valid and invalid sightings, are maintained in a reasonably homogeneous manner over time until extinction occurs.

Our main focus centers on the two models developed by Solow and Beet (2014). When these two models were used for inference about the extinction of the Ivory-billed Woodpecker, they gave very distinct results; one predicting the woodpecker is extant, and the other concluding that it is well and truly extinct. Here we examine the two extinction models of Solow and Beet (2014) and attempt to identify the factors that the models are most sensitive to. Because of the complexity of the two models it is not straightforward to identity these driving factors. This work is important as it gives insights into which factors require additional research for strengthening the knowledge base and thereby to reduce the uncertainties in the inferences made. Also we know that some prediction uncertainties can be reduced by additional data collection. Thus in situations where one model is more appropriate than the other, it is important to know which factor is most highly correlated with the model output before undertaking expensive studies to gather and analyze additional data.

2. Data

As a first step, it is of interest to examine the sighting record data of the Ivory-billed Woodpecker reported in Roberts et al. (2010). Following Solow and Beet (2014), we treated all sightings not based on physical evidence as uncertain. Since there was no natural way to define the beginning of the observation period, the period was taken to be (1897,2010) and the first sighting in 1897 was omitted as per the most methods in literature (Boakes et al., 2015; Solow, 1993; Solow and Beet, 2014). The record period from 1897 to 2010 contains 21 certain sightings in years 1898–1902, 1904–1910, 1913, 1914, 1917, 1924, 1925, 1932 1935, 1938, and 1939 and 46 uncertain sightings in years 1911, 1916, 1920, 1921, 1923, 1926, 1929–1931, 1933, 1934, 1936, 1937, 1941–1944, 1946, 1948–1952, 1955, 1958, 1959,1962, 1966–1968, 1969, 1970–1974, 1976, 1981, 1982, 1985–1988, 1999, 2004–2006, where for example 2004–2006 means that there are sightings in every year from 2004 to 2006. The Ivory-billed Woodpecker sighting data described above, is visualized graphically in Fig. 1.

In what follows, we adopt an unconventional format and provide the Methods and then give results first for Model 1, and then follow this for Model 2.

3. Methods-Model 1

We begin by outlining the basic structure of the first model of Solow and Beet (2014). Define the observation period of a species as $(0, T)$, where 0 is the time when observations began and the period lasts for T years altogether. The complete sighting record $t = (t_1, t_2, \dots, t_n)$ consists of the times (i.e., here the years) when the species were sighted during the observation period.

If the sighting is based on clear physical evidence, it is classified as certain; but if it is classified by a sound recording, photograph or video, or some other less precise confirmation, it is classified as uncertain. Thus, for observed data, we only know whether each sighting is certain or uncertain. All certain sightings are considered to be valid. Note, however, that uncertain sightings can either be valid or invalid. In practice, whether an uncertain sighting is valid or invalid is often unknown.

In the datasets studied here, all that is known is which sightings are certain and which are uncertain. From this information, the rates of valid and invalid sightings can be inferred, so that an informed decision can be made as to the likelihood of the species being extant or extinct.

Under the first model of Solow and Beet (2014), the sighting record is divided into two parts with the division based on the (unknown) extinction time τ_E. Over the first time period in the interval $(0, \tau_E)$, valid sightings follow a stationary Poisson process with rate Λ, and invalid sightings follow a stationary Poisson process with rate Θ. The expected proportion of valid sightings is $\Omega = \Lambda/(\Lambda + \Theta)$. Then, the invalid rate Θ is equal to $[(1 - \Omega)]/\Omega$.

Over the second time period in the interval (τ_E, T), all the sightings are invalid and follow a stationary Poisson process still with rate Θ. That is, the rate of invalid sighting does not change over the whole interval $(0, T)$. Model 1 is summarized in Fig. 2.

Let n_c and n_u be the number of certain and uncertain sightings. The first model proceeds by assuming the extinction time τ_E falls in the interval $(0, T)$, with $n(\tau_E)$ sightings prior to τ_E. We suppose the number of valid uncertain sightings in $(0, \tau_E)$ is the random variable j. This gives the scheme shown in Fig. 2. There are thus $n_c + j$ valid sightings in $(0, \tau_E)$, and there are $n - (n_c + j)$ invalid sightings in $(0, T)$. The goal is to construct likelihoods for the valid and invalid sightings in these two time intervals, and assess the extinction hypothesis from these likelihoods.

In more detail, let E be the event that the species became extinct during the observation period $(0, T)$ and $\bar{\mathrm{E}}$ be the event that the species is extant at time T. By Bayes theorem, the posterior probability of an extinction event E given the complete sighting record t is

$$p(E|t) = \frac{p(t|E)p(E)}{p(t|E)p(E) + p(t|\bar{\mathrm{E}})(1 - p(E))}, \tag{1}$$

where $p(t|E)$ is the likelihood of t given E, $p(t \mid \bar{\mathrm{E}})$ is the likelihood of t given $\bar{\mathrm{E}}$, and $p(E)$ is the prior probability of extinction.

The Bayes factor $(B(t))$ is a standard Bayesian measure that is used to make a decision of whether the data support the null hypothesis E, that the species went extinct in the study interval, while not depending on $p(E)$. Then the Bayes factor is defined as:

$$B(t) = \frac{p(t|E)}{p(t|\bar{\mathrm{E}})}. \tag{2}$$

A value of $B(t) > 3$ constitutes substantial evidence for the null hypothesis E, while a value $B(t) < 1/3$ constitutes substantial evidence for the alternative hypothesis $\bar{\mathrm{E}}$ that the species is extant (Kass and Raftery, 1995; Solow and Beet, 2014). In the following section, we will describe methods for estimating the Bayes Factor, so that the null hypothesis may be tested with real sighting data.