



Optimizing sampling design to infer the number of marine turtles nesting on low and high density sea turtle rookeries using convolution of negative binomial distribution



Marc Girondot

Laboratoire Écologie, Systématique, Évolution, Université Paris-Sud, AgroParisTech, Centre National de la Recherche Scientifique, Université Paris Saclay, 91405, Orsay, France

ARTICLE INFO

Keywords:

Marine turtle
Negative binomial
Phenology
Sampling design
Seasonality
Statistical model

ABSTRACT

Reliable monitoring of wildlife populations represents a non-negligible cost, and in a limited-resource world, resources allocated to monitoring are not devoted to actions to solve identified problems.

I explore resource efficient survey designs based on a negative binomial distribution including variable survey intervals for marine turtles using track counts as an index of female activity. In the modified procedure, all new tracks between two monitoring patrols are recorded. These data are analyzed by statistical models that take advantage of the statistical properties of the sum of counts.

The outputs of models with different lagged monitoring dates (3–10 days) are compared with the outputs of daily surveys using extrapolations from high and low density populations. Results show that the quality of the estimates is similar when total time series analysis is compared with situations in which only a fourth, a seventh, or a tenth of monitoring daily during the season are used.

This solution permits the reallocation of funds from monitoring to other conservation activities. Furthermore, the efficient sampling design and the statistical methods allow getting similar information with less effort.

1. Introduction

Ecological monitoring is a standardized approach to address the growing number of conservation problems around the world. Count data for wildlife populations are used in conservation research to ensure that the population stays within sustainable limits, to guarantee its survival, or to test whether populations of pest species remain below critical levels known to threaten other populations (Williams et al., 2002). Well planned data collection for long-term monitoring of wildlife populations should be conducted consistently enough to be comparable between years and across populations and precisely enough to detect changes in a population with sufficient confidence and power (Gerrodette, 1987; Hayes and Steidl, 1997; Sims et al., 2008). When different survey methods or efforts are used, modeling techniques may assist to standardize data (Elphick, 2008).

The choices of sampling design are subject to several constraints, some specific to the species or study site (Kéry et al., 2007) and others specific to the human and financial resources available for monitoring. Several monitoring strategies apply for the context of marine turtles at nesting sites (SWOT Scientific Advisory Board, 2011). Researchers quantify the number of clutches laid by a population as an index of

population size (Gerrodette and Taylor, 1999). It is relatively easy to identify a sea turtle track, because females nest on open sandy tropical and subtropical beaches, and leave wide deep tracks that persists on the beach (Schroeder and Murphy, 1999). Modeling the seasonal nesting phenology of marine turtles offers a way to estimate global nest or annual track counts without being present daily (SWOT Scientific Advisory Board, 2011). Several methods have been proposed that model the nesting season of marine turtles. Most papers have discussed the equations that define the quasi-Gaussian shape of a nesting season (Bellini et al., 2013; Girondot, 2010; Girondot et al., 2006; Gratiot et al., 2006; Whiting et al., 2014). A final conclusion is that many solutions are similar (Whiting et al., 2014). Challenges in counting sea turtle clutches are that nesting seasons usually span several months, turtles can lay their eggs on remote beaches that are difficult or expensive to access and total number of clutches on some beaches can be very low.

When total nest number during the season is low, all these methods require intensive fieldwork to capture enough of the few nesting events that occurred. For example, of the 113 time series available for the hawksbill turtles nesting in Guadeloupe archipelago only 67 could be used to fit a model to the nesting season (Delcroix et al., 2013). For the

E-mail address: marc.girondot@u-psud.fr.

<http://dx.doi.org/10.1016/j.ecolind.2017.05.063>

Received 10 April 2017; Received in revised form 23 May 2017; Accepted 24 May 2017
Available online 01 June 2017

1470-160X/ © 2017 Elsevier Ltd. All rights reserved.

remaining 46 time series no clutches were observed during the monitored nights. This does not mean that hawksbill turtles do not nest on these beaches but that there was no monitoring for at least one night on which one hawksbill turtle laid a clutch of eggs. For example, on Bois Jolan beach on Sainte Anne no clutches were recorded in 2008 on the 12 monitored nights although a track of a turtle who laid a clutch on a previous night was recorded on 9 July. The power to detect a trend of specified magnitude with a given level of significance is negatively related to variability and positively related to sample effort (Carlson and Schmiegelow, 2002). While increasing the sampling effort increases the power to detect a trend, excessive sampling wastes limited monitoring resources (Bernstein and Zalinski, 1983). In a world with limited resources to allocate towards conservation, every that time resources (human or financial) are allocated to one action, they become unavailable for other actions (Klein et al., 2016). Thus, monitoring must be able to identify, and often respond to, trends with a high degree of certainty, but ideally use as few resources as possible.

We therefore propose and evaluate a strategy to monitor and analyze marine turtle track counts for a nesting season covered by fewer than daily monitoring patrols as is usually done. We start from the observation that marine turtle tracks persist several days. We apply a statistical model describing these data and explore alternatives of 3–10 days time lags as alternatives to daily monitoring. The models are applied to datasets of low and high level of nesting on beach to evaluate different observed situations.

2. Materials and methods

2.1. Statistical distribution of daily nest numbers

The literature describes several statistical distributions to model the daily nesting activity in a typical sea turtle season:

- implicit homoskedastic Gaussian (Gratiot et al., 2006; Whiting et al., 2013; Whiting et al., 2014)
- explicit heteroskedastic Gaussian (Girondot et al., 2006);
- Poissonian (Bellini et al., 2013; Godgenger et al., 2009);
- Negative binomial (Delcroix et al., 2013; Girondot, 2010; Girondot and Rizzo, 2015)

The assumption that a Gaussian probability distribution underlies the observed data is problematic for several reasons described in Godgenger et al. (2009). Firstly, the Gaussian distribution represents probabilities associated with a continuous variable that can theoretically take on any possible value within a plausible range, including fractional values if the observational method has the necessary precision. Nest numbers are discrete counts sub-sampled and can only be represented by non-negative integers. Such data are also characterized by low mean values and high variances. Because of its symmetry and its domain of validity from $-\infty$ to $+\infty$, the normal probability distribution can imply a substantial probability of observing a negative number of nests. Therefore, we will focus on comparison of Poissonian and negative binomial distributions.

2.2. Poissonian and negative binomial distribution

The Negative Binomial (NB) distribution can be used to describe the distribution of the number of successes or failures. Suppose that there is a sequence of independent Bernoulli trials, with each trial having two potential outcomes called “success” and “failure.” In each trial, the probability of success and failure is p and $(1 - p)$, respectively. This sequence is observed until a predefined number r of failures has occurred. The random variable of observed successes, X , has a NB distribution as follows:

$$X \sim NB(r; p)$$

Table 1

Model selection based on time series with a total of (A) 35 or (B) 2762 nests. A parameter is set to 0 when it is not indicated on a set of parameters. ΔAIC is the difference between each model and the selected one. The Akaike weight for the selected model is in bold.

| A: 35 nests | AIC | ΔAIC | Akaike weight |
|---|---------|--------------|---------------|
| Max MinB MinE LengthB Peak LengthE Flat | 145.23 | 8.00 | 0.01 |
| Max MinB MinE LengthB Peak LengthE | 143.19 | 5.96 | 0.03 |
| Max Min LengthB Peak LengthE | 141.33 | 4.10 | 0.07 |
| Max MinB MinE Length Peak Flat | 143.27 | 6.04 | 0.02 |
| Max Min Length Peak Flat | 143.27 | 6.04 | 0.02 |
| Max LengthB Peak LengthE | 139.15 | 1.92 | 0.22 |
| Max Length Peak | 137.22 | 0.00 | 0.59 |
| B: 2762 nests | AIC | ΔAIC | Akaike weight |
| Max MinB MinE LengthB Peak LengthE Flat | 1348.85 | 2.00 | 0.23 |
| Max MinB MinE LengthB Peak LengthE | 1346.85 | 0.00 | 0.65 |
| Max Min LengthB Peak LengthE | 1350.89 | 4.03 | 0.08 |
| Max MinB MinE Length Peak Flat | 1355.00 | 8.14 | 0.01 |
| Max Min Length Peak Flat | 1355.00 | 8.14 | 0.01 |
| Max LengthB Peak LengthE | 3592.62 | 2245.76 | 0.00 |
| Max Length Peak | 3577.92 | 2231.06 | 0.00 |

$$Pr(X = x) = \frac{(x + r - 1)!}{x!(r - 1)!} (p)^x (1 - p)^r, 0 \leq p \leq 1, r > 0 \tag{1}$$

The notation in R language is `dnbinom(x, size = r, prob = p, log = FALSE)` (R Core Team, 2017).

In ecology, an alternative parametrization of NB distribution is often used to describe the distribution of an organism using the mean number of individuals m and an aggregation parameter k (Taylor et al., 1979). The random variable of observed individuals, X , will have the following NB distribution:

$$X \sim NB(m; k)$$

$$Pr(X = x) = \frac{\Gamma(k + x)}{x! \Gamma(k)} \left(\frac{m}{m + k} \right)^x \left(1 + \frac{m}{k} \right)^{-k}, m > 0, k > 0 \tag{2}$$

The notation in R language is `dnbinom(x, size = 1/k, mu = m, log = FALSE)` (R Core Team, 2017).

The variance of the NB distribution is $(m + m^2/k)$, and hence decreasing values of k correspond to increasing levels of dispersion (Lloyd-Smith, 2007). The Poisson distribution is obtained as $k \rightarrow \infty$, and the logarithmic series distribution is obtained as $k \rightarrow 0$ (Anscombe, 1950; Bliss and Fisher, 1953). When $k = 1$, the NB distribution is reduced to the geometric distribution. Ecological statistical literature uses both the quantity k and $\alpha = 1/k$ (confusingly, the term “dispersion parameter” can refer to either k or α ; other terms for k include “shape parameter” and “clustering coefficient”).

The properties of the sum of NB random variables have a special interest when the series of events are aggregated into one value. When counts are available for several consecutive days, the distribution of this sum is no longer a NB distribution (Furman, 2007). The Lyapunov central limit theorem states that the arithmetic mean of a sufficiently large number of iterates of independent random variables, not necessarily identically distributed, will be approximately normally distributed, regardless of the underlying distribution (Billingsley, 1995; Rice, 2007). However, tracks on sand can often be visible only for 2–10 days depending on local conditions, therefore the central limit theorem cannot be applied safely with so few days. An exact distribution of the sum of NB random variables must therefore be established.

Let $X_i \sim NB(r_i; p)$, then a classical result is $\sum X_i \sim NB(\sum r_i; p)$ (Johnson et al., 1992). This result holds only for p being constant. When NB parametrization is $X_i \sim NB(m_i; k)$, k being a constant, the model can be written as $X_i \sim NB(r = k; p_i = k/(k + m_i))$. In this case, p is not constant and the previous result does not hold. The density probability of the sum of NB random variables must be estimated.

Download English Version:

<https://daneshyari.com/en/article/5741413>

Download Persian Version:

<https://daneshyari.com/article/5741413>

[Daneshyari.com](https://daneshyari.com)