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

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

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 $-$infinity to $+$infinity, 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

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

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:

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

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

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 $k \to \infty$, and the logarithmic series distribution is obtained as $k \to 0$ (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.
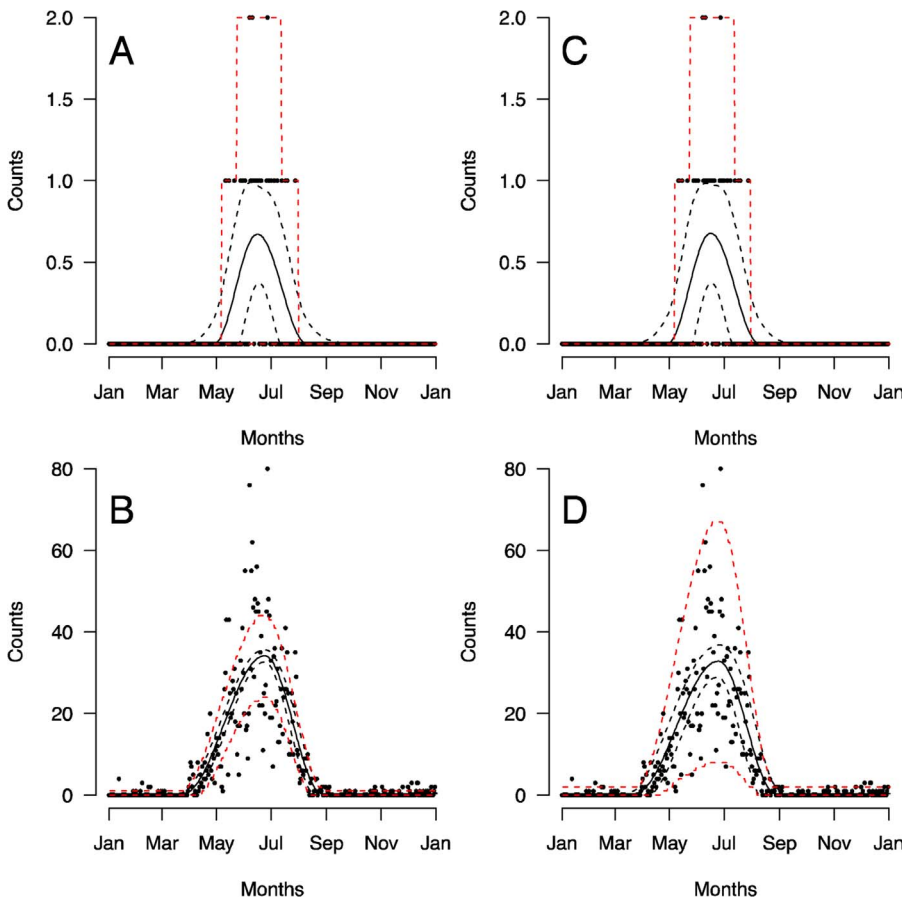
Fig. 1. Model of the nesting season based on AIC for Leatherback nesting in French Guiana (Gratiot et al., 2006 data) for (A, C) low nesting season (35 tracks during the season) and (B, D) high (2762 tracks during the season), (A, B) Poissonian and (C, D) negative binomial distributions. Black dashed lines are the limits of the confidence interval (95%) of the maximum likelihood model and red dashed lines are the limits of the confidence interval (95%) of the daily counts. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)
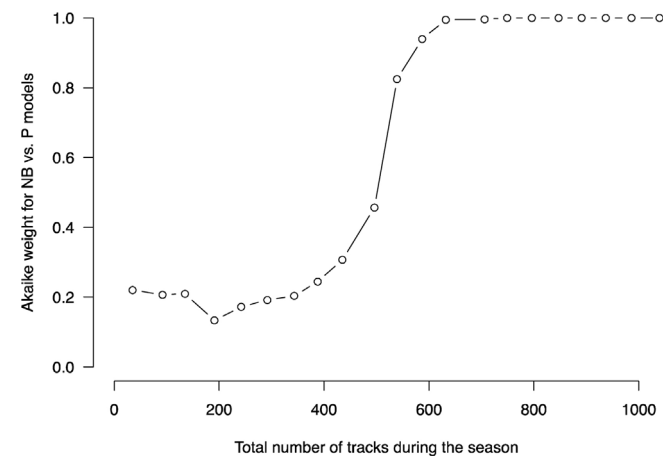

Fig. 2. Probability that negative binomial model is selected against Poissonian one based on Akaike weight for total track numbers from 50 to 3000 (only values up to 1000 are shown).

2.3. Density probability of the sum of NB using convolution of NB random variables

When $X_i \sim NB(r_i; p_i)$, the distribution of $\sum X_i$ is a mixture NB (Furman, 2007) (the name of some variables has been replaced to prevent confusion with Eq. (1) above), with the mixture parameter $\sum r_i + G$, where G is an integer random value with a probability mass function (see Theorem 2, Eqs. (13), (3)–(5) in Furman, 2007):

$$pr_\gamma = R\delta_\gamma, \ \gamma = 0, 1, \ldots$$

where $R = \prod_{j=1}^{n} \left(\frac{q_j p_1}{q_1 p_j} \right)^{-r_j}$ where $q_j = 1 - p_j$

Table 2
Point estimate and standard error (SE) for the date of the peak of the nesting season and durations of the nesting season using different monitoring strategies with a total of 35 nests during the nesting season. NA indicates that standard error could not be estimated.

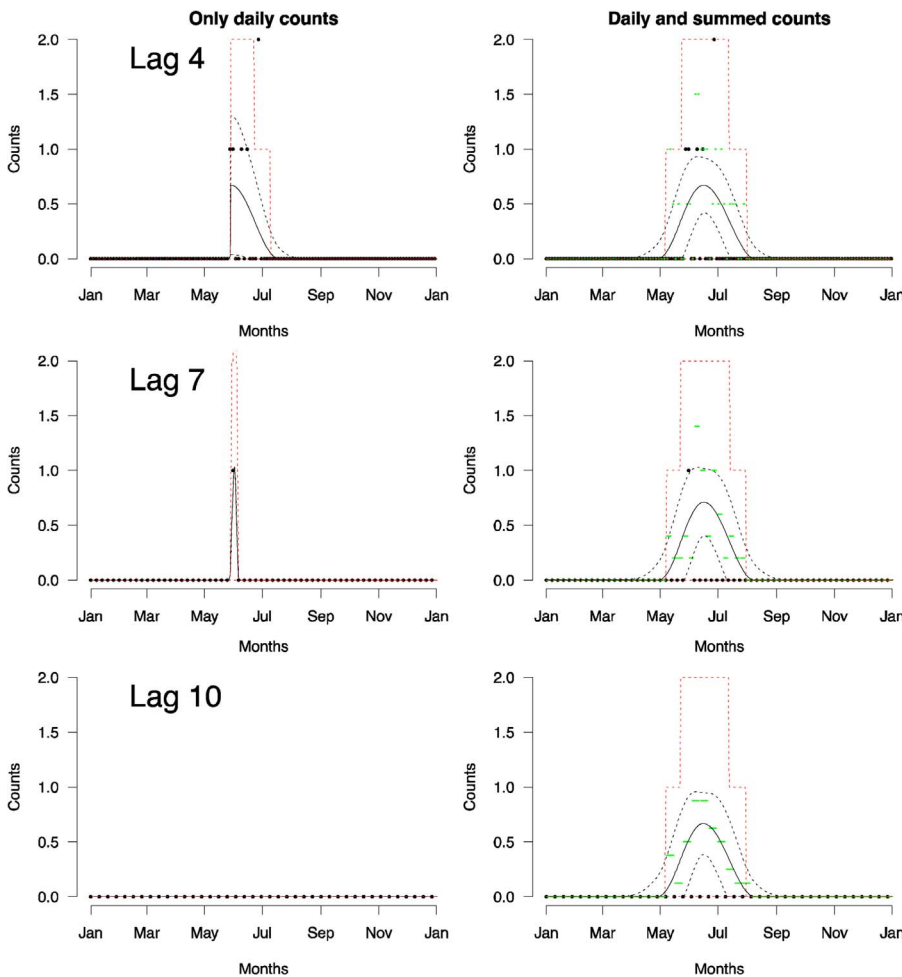| Lag	Daily counts	Daily counts and sum	Only sum
LengthB			
1	48.51 (13.93)		
3	1.14 (NA)	48.89 (10.99)	49.13 (13.42)
4	33.31 (19.72)	49.88 (13.53)	48.37 (9.44)
5	47.15 (23.22)	47.23 (9.65)	48.05 (9.07)
6	5.44 (NA)	47.89 (11.77)	49.41 (11.07)
7	5.07 (NA)	47.50 (9.62)	48.36 (9.52)
8	8.41 (NA)	47.11 (13.97)	49.19 (16.34)
9	48.52 (NA)	48.68 (13.91)	50.53 (6.93)
10	50.12 (25.63)	47.75 (9.49)	42.62 (12.96)
P			
1	165.77 (9.09)		
3	148.01 (NA)	165.75 (7.83)	165.66 (8.51)
4	161.40 (12.76)	166.33 (8.43)	165.98 (8.53)
5	160.95 (14.33)	165.38 (7.93)	165.98 (8.31)
6	151.70 (NA)	166.02 (8.82)	165.26 (9.19)
7	141.03 (NA)	165.51 (8.83)	165.51 (8.77)
8	154.22 (NA)	166.10 (9.29)	165.54 (10.52)
9	165.77 (NA)	165.61 (9.12)	164.90 (7.87)
10	162.53 (12.92)	169.23 (9.52)	163.97 (9.65)
LengthE			
1	48.51 (13.93)		
3	1.14 (NA)	48.89 (10.99)	49.13 (13.42)
4	33.31 (19.72)	49.88 (13.53)	48.37 (9.44)
5	47.15 (23.22)	47.23 (9.65)	48.05 (9.07)
6	5.44 (NA)	47.89 (11.77)	49.41 (11.07)
7	5.07 (NA)	47.50 (9.62)	48.36 (9.52)
8	8.41 (NA)	47.11 (13.97)	49.19 (16.34)
9	48.52 (NA)	48.68 (13.91)	50.53 (6.93)
10	50.12 (25.63)	47.75 (9.49)	42.62 (12.96)

Fig. 3. Fitted nesting season with 35 tracks based on simulated sampling design; in the first column, daily counts are recorded every x days (x = 4, 7 and 10) and in the second column, the older tracks between two monitoring sessions are also recorded. Black dashed lines are the limits of the confidence interval (95%) of the maximum likelihood model and red dashed lines are the limits of the confidence interval (95%) of the daily counts. Green lines are the average number of nests when counts are available only for several consecutive days. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and $\delta_{\gamma+1} = \frac{1}{\gamma+1} \sum_{i=1}^{\gamma+1} i\xi_i\delta_{\gamma+1-i}$, $\gamma = 0, 1, \ldots$

$$\xi_i = \sum_{j=1}^{n} \frac{r_j\left(1 - q_1 p_j / q_j p_1\right)^i}{i}$$

This solution was implemented in the R package HelpersMG (Girondot, 2017a): functions dSnbinom(), pSnbinom(), qSnbinom() and rSnbinom().

2.4. Test on simulated sampling design with complete data

A year-round time series with $T = 2762$ leatherback track daily counts (t_i with i from 1 to 365) obtained in French Guiana was used (Gratiot et al., 2006). This time-series is available in the R package phenology (Girondot, 2017b). Cross products have been used to generate time series with total track numbers being $S = 35$ and from $S = 50$ $t_i' = (S/T)t_i$ to 3000 by step of 50 using

Numbers have been rounded to the nearest integer. The time series with 35 and 2762 tracks were transformed to mimic a protocol in which the beach was monitored every 3–10 days. Several alternatives were considered:

- Only tracks from a last night were counted;
- All tracks more recent than a previous survey are counted. All counted tracks are erased after counting to prevent recounting them during the next patrol (Chevalier and Girondot, 1998).
- Tracks from the previous night were identifiable accurately as being from the last night but that older tracks were still visible without

being able to assign a precise date. All tracks are erased after counting to prevent recounting them during the next patrol (Chevalier and Girondot, 1998).

The model of the seasonality of nesting is based on Girondot (2010). This model was preferred among the dozen available because: (i) it performed among the best based on an extensive test (Whiting et al., 2014), (ii) its parametric definition allows the standard error to be minimized and (iii) the parameters have direct biological interpretations. This model is summarized here briefly.

The model can be applied on nest or track counts depending on the available data but it must be remained consistent for one analysis. Assuming that t is a day of the year, the number of nests deposited per night is modeled using the system of equations as follows:

$$\begin{cases} t<B \rightarrow \text{MinB} \\ t\in[B, P- F/2] \rightarrow ((1 + \cos(\pi(P-F/2-t)(P-F/2-B)))/2)(\text{Max} - \text{MinB}) \\ \qquad + \text{MinB} \\ t\in[P-F/2, P+ F/2] \rightarrow \text{Max} \\ t\in[P+F/2,E] \rightarrow ((1 + \cos(\pi(t- P+ F/2)(E- P+ F/2)))/2)(\text{Max} - \text{MinE}) \\ \qquad + \text{MinE} \\ t>E \rightarrow \text{MinE} \end{cases}$$

(3)

The model requires seven parameters at most. The formulas were constructed to allow the parameters to have direct biological interpretations:

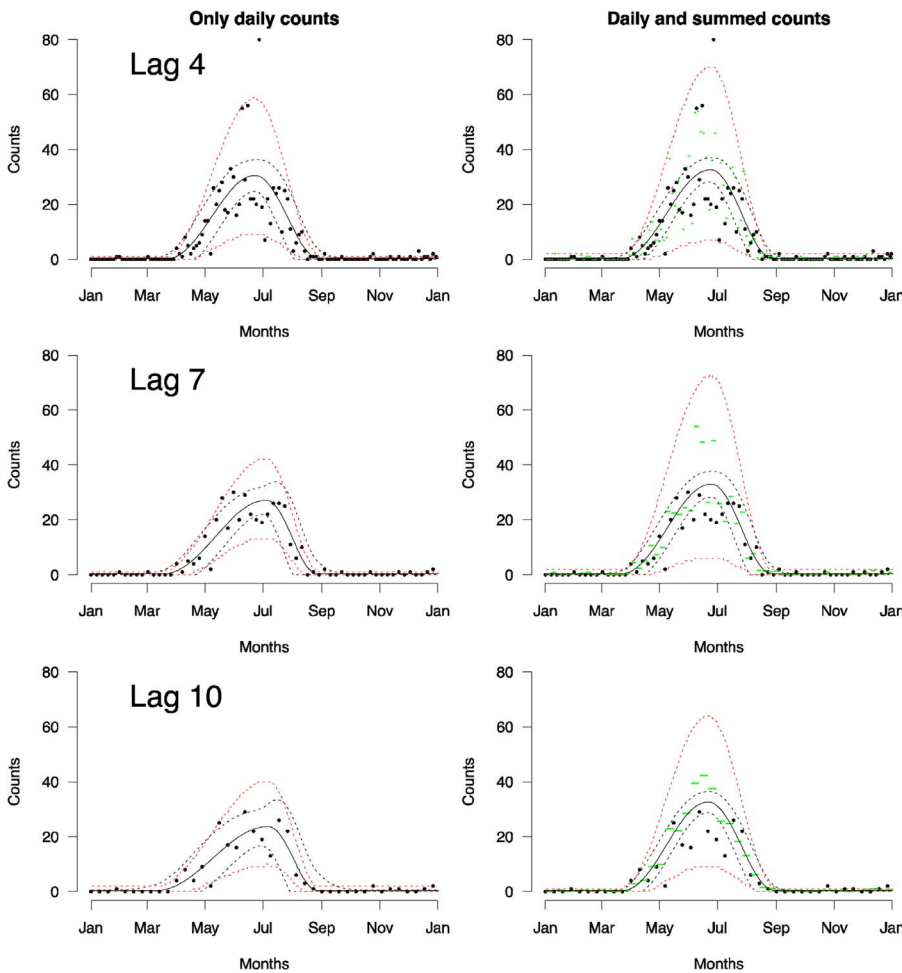- *MinB* is the mean nightly nest number before the beginning of the

Fig. 4. Fitted nesting season with 2762 nests based on simulated sampling design; in the first column, daily counts are recorded every × days (x = 4, 7 and 10) and in the second column, the older tracks between two monitoring sessions are also recorded. Black dashed lines are the limits of the confidence interval (95%) of the maximum likelihood model and red dashed lines are the limits of the confidence interval (95%) of the daily counts. Green lines are the average number of nests when counts are available only for several consecutive days. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

nesting season;

- *MinE* is the mean nightly nest number after the end of the nesting season;
- *Max* is the mean number of nests at the peak of the nesting season;
- *P* is the day of the year on which the nesting season peaks;
- *F* is the number of days around the day *P* on which the curve flattens out;
- *B* is the day of the year on which the nesting season begins;
- *E* is the day of the year on which the nesting season ends.

Whereas a nesting season is described by segments, the equations make all the segments remain in continuity. A nesting season is defined as the interval [*B*, *E*]. Various constraints can be set up to simplify this model: $Min = MinB = MinE$, for the same number of nests outside of the nesting season; *MinB* and/or *MinE* = 0, when no nests are observed outside of the nesting season; $L = P − B = E − P$, when the nesting season is symmetric around *P* with *L* being half the length of the nesting season; $F = 0$, for no flat portion. Rather than fitting B and E, it is more convenient to fit $LengthB = P-B$ and $LengthE = E − P$ with $LengthB > 0$ and $LengthE > 0$ to ensure that $B < P < E$.

Parameter fitting was performed using maximum likelihood with Poissonian or negative binomial daily nest distribution using values produced by Eq. (3) as theoretical values and the observed counts t_i as observations. Akaike Information Criterion (AIC) (Akaike, 1974) and Akaike weight (Burnham and Anderson, 2002) were used for model selection. In short, AIC evaluates the quality of the fit that penalizes for overfitting too many parameters and Akaike weight gives the relative support of the different models, *i.e.* the probability for each model to be the best one.

The adjustment was done using the R package phenology available

in The Comprehensive R Archive Network (https://cran.r-project.org) that implements this model (Girondot, 2017b).

3. Results

3.1. Model selection for daily time series of nest counts

The complete time series with 35 and 2762 tracks were initially fitted using negative binomial distribution with all seven parameters and the model was then simplified based on model selection using AIC. Negative binomial distribution was used because Poissonian distribution is a special case of negative binomial distribution where the variance is equal to the mean. For low density beaches (35 tracks per year), the selected model was *Max*, *Length*, and *P*. For high-density beaches (2762 tracks per year), the selected model was *Max*, *MinB*, *MinE*, *LengthB*, *P*, and *LengthE* (Table 1). The difference of model selection indicates that a more complex model can be selected for the statistical power associated higher track number but a simpler model is appropriate when total tracks is lower. The most complex combination of parameters will be used hereafter for these two timeseries. The fits of the complete time series of nest counts are shown in Fig. 1C for situation with 35 nests during all the season and in Fig. 1D for situation with 2762 nests. Asymmetry was observed with a more abrupt decline of nest numbers after the Peak than the rise before the Peak; as a consequence, the Peak was shifted to the right of the distribution.

The model fits of the same two time series using Poisson distribution are shown in Fig. 1A and B (low and high density, respectively). The AIC was better for Poisson distribution for the time series with 35 tracks (Akaike weight = 0.74) whereas it was better for the negative binomial distribution for the time series with 2762 tracks (Akaike

Table 3

Point estimate and standard error (SE) for the date of the peak of the nesting season and durations of the nesting season using different monitoring strategies with a total of 2762 nests during the nesting season.

Lag	Daily counts	Daily counts and sum	Only sum
LengthB			
1	94.39 (5.19)		
3	92.00 (8.30)	93.63 (5.71)	92.19 (6.88)
4	96.19 (8.44)	94.49 (6.02)	92.08 (7.65)
5	88.92 (10.78)	92.97 (6.06)	92.13 (6.46)
6	108.05 (10.46)	94.05 (6.22)	91.65 (6.83)
7	113.17 (16.83)	91.32 (6.48)	92.47 (6.40)
8	89.08 (12.15)	92.47 (5.87)	90.55 (8.85)
9	118.22 (15.96)	89.67 (5.16)	88.59 (4.73)
10	81.67 (10.03)	89.94 (5.28)	88.63 (7.16)
P			
1	174.49 (4.15)		
3	172.68 (6.51)	173.77 (4.65)	172.27 (5.69)
4	176.27 (6.54)	174.31 (4.98)	172.05 (6.50)
5	171.69 (8.45)	173.14 (4.99)	172.18 (5.25)
6	184.50 (7.46)	174.37 (5.14)	172.15 (5.59)
7	186.95 (13.15)	171.69 (5.38)	173.06 (5.21)
8	174.20 (8.46)	173.15 (4.72)	171.32 (7.43)
9	186.75 (11.28)	171.22 (4.08)	170.57 (3.70)
10	161.10 (7.87)	171.46 (4.07)	170.62 (5.74)
LengthE			
1	64.33 (5.93)		
3	68.04 (9.14)	65.64 (6.85)	68.50 (8.07)
4	63.67 (8.96)	65.00 (7.38)	69.15 (9.74)
5	67.06 (11.87)	66.92 (7.64)	68.80 (7.06)
6	53.89 (10.80)	64.22 (7.94)	68.31 (7.67)
7	39.85 (16.35)	70.04 (8.24)	66.64 (7.32)
8	67.42 (10.25)	67.17 (6.28)	70.52 (10.35)
9	52.21 (15.42)	70.46 (6.26)	71.29 (4.75)
10	78.67 (11.20)	69.02 (5.54)	70.34 (8.04)

weight = 1.00). Using the Poisson distribution with the high-density time series, 20% of the observations lay out of the mean ± 2.SD envelope indicating that Poisson distribution is not adequate. Proportion of counts out of the mean ± 2.SD envelope using negative binomial was 5% as expected. When time series with different total nest number are generated, Poisson distribution was selected for < 500 annual tracks whereas a negative binomial was selected for > 500 tracks (Fig. 2). The same limits should be used if nest counts are used rather than track counts.

3.2. Nesting season fit with partial temporal monitoring

Subsequently, data with surveys at 3–10 days time lags were fitted using the same models. The date of the peak (measured as an ordinal date) and the length of the nesting season (*LengthB*, *LengthE*) with their standard errors were used as criteria to evaluate the quality of the fit. When only daily counts were used (column Daily counts in Tables 2 and 3), the fit was impossible for the beach with lower track number (Fig. 3, first column) but was possible with the high-density nesting beach (Fig. 4, first column). However, the quality of the estimates as well as the standard errors worsens when lag was higher (Table 3).

For the low-density nesting beach, the model fit with data nest counts summed over a range of time lags was always possible and gave outputs very similar to the situation with daily monitoring. For the high-density nesting beach, a convolution model for the likelihood of the sum of negative binomial can be applied regardless of the number of tracks or days. The point estimates and standard errors were similar and were not impacted by the survey time lag (Tables 2 and 3).

4. Discussion and conclusion

Population status and trend assessment can have important social, political and economic implications. Status, trends, and knowledge of

threats are important criteria for listing species in the IUCN's Red List assessments (Bland et al., 2016). Given that money and time are usually limited, conservation biologists are faced with the nontrivial task of how to optimize information on trends with limited resources (Taylor and Gerrodette, 1993). We illustrate a range of monitoring strategies that are more efficient than daily monitoring but that give similar quantitative outputs. In this modelled scenario, a saving of time or expenses could hypothetically be reallocated to other tasks. Depending on the accessibility of the monitored site, the saved resources could be substantial enough to justify applying a mixed sampling strategy using daily counts and the sum of previous counts as presented here.

The index methodology proposed here can be applied to any situation when the counted items stay visible for several days in the environment. This includes tracks left by marine turtles on sandy beaches but also artefacts left in the ecosystem (e.g., feces, activity traces on trees, tracks on snow). In all of these situations, it is possible to count the sum of artefacts during a monitoring session and compare the results between the current and previous monitoring. It is important to mark artefacts to not recount them during a following survey; they can be removed from the ecosystem or marked with paint or other special features (e.g., making deep marks on sand with a wooden stick). If it is not possible to safely identify artefacts from the previous day as compared to older ones, the method can still be applied (see Tables 1 and 2, columns Only sum).

The survey time lags that can be applied depend on the persistence of artifacts in the environment, upon artifact distinctiveness (deeper leatherback tracks will persist longer than light Kemp's ridley tracks) and other environmental factors (windy or rainy periods that obscure tracks, sand moisture or dryness that affect track formation). The management of the monitoring design should be adaptive to take into account environmental conditions. It is important to calibrate the visibility of tracks with the chosen lag; a simple way to validate the lag is to check whether the known tracks already counted during the previous monitoring session are still visible. This can be done using recorded GPS positions of the previous tracks. If they are no longer visible, the time lag was too large and only the most recent tracks should be used. The lag should then be diminished for the next monitoring session. In the context of marine turtle tracks, not every crawl of a female sea turtle on the beach results in a nest; for instance, a female is testing the beach or disturbed before oviposition, dry sand collapses the nest chamber and it may abandon the nesting attempt before laying eggs. If tracks from non-nesting and nesting turtles are confounded during monitoring, this may strongly bias the estimated number of nests laid during a given night. Distinguishing between crawls with successful nesting or those without is thus particularly vital to distinguish. It has therefore been advocated to count all tracks as a measure of activity of marine turtles on a nesting beach (SWOT Scientific Advisory Board, 2011).

Beaches without or with very few nests should not be omitted from the sampling design. Several arguments justify the monitoring of beaches, however infrequently, where turtles are not currently nesting. Firstly, nesting effort could be low and therefore overlooked by light monitoring. Secondly, even if turtles are not using a beach it does not mean they will never use it. Shifts of nesting beach caused by erosion or colonization of new beaches can occur. If there are no data for a particular beach it will be impossible to show that any colonization has occurred. The methods proposed here generates results with fewer costs than daily monitoring and it may be very useful in a context of changing environment.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19, 716–723.

Anscombe, F.J., 1950. Sampling theory of the negative binomial and logarithmic series distributions. Biometrika 37, 358–382.

Bellini, C., Santos, A.J.B., Grossman, A., Marcovaldi, M.A., Barata, P.C.R., 2013. Green turtle (*Chelonia mydas*) nesting on Atol das Rocas, north-eastern Brazil, 1990–2008. J. Mar. Biol. Assoc. U. K. 93, 1117–1132.

Bernstein, B.B., Zalinski, J., 1983. An optimum sampling design and power tests for environmental biologists. J. Environ. Manage. 16, 35–43.

Billingsley, P., 1995. Probability and Measure, third edition. John Wiley & Sons, Hoboken, New Jersey.

Bland, L.M., Keith, D.A., Miller, R.M., Murray, N.J., Rodríguez, J.P., 2016. In: Guidelines for the Application of IUCN Red List of Ecosystems Categories and Criteria, Version 1.0. IUCN. Gland, Switzerland. p. 94.

Bliss, C.I., Fisher, R.A., 1953. Fitting the negative binomial distribution to biological data – note on the efficient fitting of the negative binomial. Biometrics 9, 176–200.

Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach. Springer-Verlag, New York.

Carlson, M., Schmiegelow, F., 2002. Cost-effective sampling design applied to large-scale monitoring of boreal birds. Conserv. Ecol. 6, 11.

Chevalier, J., Girondot, M., 1998. Dynamique de pontes des tortues luths en Guyane française durant la saison 1997. Bulletin de la Société Herpétologique de France 85–86, 5–19.

Delcroix, E., Bédel, S., Santelli, G., Girondot, M., 2013. Monitoring design for quantification of marine turtle nesting with limited human effort: a test case in the Guadeloupe Archipelago. Oryx 48, 95–105.

Elphick, C.S., 2008. How you count counts: the importance of methods research in applied ecology. J. Appl. Ecol. 45, 1313–1320.

Furman, E., 2007. On the convolution of the negative binomial random variables. Stat. Probab. Lett. 77, 169–172.

Gerrodette, T., Taylor, B.L., 1999. Estimating population size. In: Eckert, K.L., Bjorndal, K.A., Abreu-Grobois, F.A., Donnelly, M. (Eds.), Research and Management Techniques for the Conservation of Sea Turtles. IUCN/SSC Marine Turtle Specialist Group Publication No. 4, Whasington, DC, pp. 67–71.

Gerrodette, T., 1987. A power analysis for detecting trends. Ecology 68, 1364–1372.

Girondot, M., Rizzo, A., 2015. Bayesian framework to integrate traditional ecological knowledge into ecological modeling: a case study. J. Ethnobiol. 35, 337–353.

Girondot, M., Rivalan, P., Wongsopawiro, R., Briane, J.-P., Hulin, V., Caut, S., Guirlet, E., Godfrey, M.H., 2006. Phenology of marine turtle nesting revealed by a statistical model of the nesting season. BMC Ecol. 6, 11.

Girondot, M., 2010. Estimating density of animals during migratory waves: application to marine turtles at nesting site. Endangered Species Res. 12, 85–105.

Girondot, M., 2017a. HelpersMG: Tools for Earth Meteorological Analysis, Version 1.9. The Comprehensive R Archive Network. https://CRAN.R-project.org/package=HelpersMG.

Girondot, M., 2017b. Phenology: Tools to Manage a Parametric Function That Describes Phenology, Version 5.2.5. The Comprehensive R Archive Network. https://CRAN.R-project.org/package=phenology.

Godgenger, M.-C., Bréheret, N., Bal, G., N'Damité, K., Girard, A., Girondot, M., 2009. Nesting estimation and analysis of threats for Critically Endangered leatherback *Dermochelys coriacea* and Endangered olive ridley *Lepidochelys olivacea* marine turtles nesting in Congo. Oryx 43, 556–563.

Gratiot, N., Gratiot, J., de Thoisy, B., Kelle, L., 2006. Estimation of marine turtles nesting season from incomplete data; statistical adjustment of a sinusoidal function. Anim. Conserv. 9, 95–102.

Hayes, J.P., Steidl, R.J., 1997. Statistical power analysis and amphibian population trends. Conserv. Biol. 11, 273–275.

Johnson, N., Kotz, S., Kemp, A., 1992. Univariate Discrete Distributions, second edition. Wiley, New York.

Kéry, M., Royle, J.A., Schmid, H., 2007. Importance of sampling design and analysis in animal population studies: a comment on Sergio et al. J. Appl. Ecol. 45, 981–986.

Klein, C.J., Beher, J., Chaloupka, M., Hamann, M., Limpus, C., Possingham, H.P., 2016. Prioritization of marine turtle management projects: a protocol that accounts for threats to different life history stages. Conserv. Lett.

Lloyd-Smith, J.O., 2007. Maximum likelihood estimation of the negative binomial dispersion parameter for highly overdispersed data, with applications to infectious diseases. PLoS One 2, e180.

R Core Team, 2017. R: A Language and Environment for Statistical Computing, Version 3.4.0. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Rice, J., 2007. Mathematical Statistics and Data Analysis, third edition. Duxbury Press, Pacific Grove, CA.

SWOT Scientific Advisory Board, 2011. The State of the World's Sea Turtles (SWOT) Minimum Data Standards for Nesting Beach Monitoring, Version 1.0. In: Wallace, B.P., Chaloupka, M., DiMatteo, A., Eckert, S., Girondot, M., Limpus, C., Mast, R.B., Pilcher, N.J., Seminoff, J. (Eds.), Washington, DC, SWOT p. 28.

Schroeder, B., Murphy, S., 1999. Population surveys (ground and aerial) on nesting beaches. In: Eckert, K.L., Bjorndal, K.A., Abreu-Grobois, F.A., Donnelly, M. (Eds.), Research and Management Techniques for the Conservation of Sea Turtles. IUCN/SSC Marine Turtle Specialist Group, Washington DC, pp. 45–55.

Sims, M., Bjorkland, R., Mason, P., Crowder, L.B., 2008. Statistical power and sea turtle nesting beach surveys: how long and when? Biol. Conserv. 141, 2921–2931.

Taylor, B.L., Gerrodette, T., 1993. The uses of statistical power in conservation biology: the Vaquita and Northern spotted owl. Conserv. Biol. 7, 489–500.

Taylor, L.R., Woiwod, I.P., Perry, N., 1979. The negative binomial as a dynamic ecological model for aggregation, and the density dependence of k. J. Anim. Ecol. 48, 289–304.

Whiting, A.U., Chaloupka, M., Limpus, C.J., 2013. Comparing sampling effort and errors in abundance estimates between short and protracted nesting seasons for sea turtles. J. Exp. Mar. Biol. Ecol. 449, 165–170.

Whiting, A.U., Chaloupka, M., Pilcher, N., Basintal, P., Limpus, C.J., 2014. Comparison and review of models describing sea turtle nesting abundance. Mar. Ecol. Prog. Ser. 508, 233–246.

Williams, B.K., Nichols, J.D., Conroy, M.J., 2002. Analysis and Management of Animal Populations. Academic press, San Diego.