



# Estimating density of animals during migratory waves: a new model applied to marine turtles at nesting sites

Marc Girondot<sup>1,2,\*</sup>

<sup>1</sup>Laboratoire Ecologie, Systématique et Evolution, UMR 8079, Université Paris-Sud, ENGREF et CNRS, Bât. 362, Orsay cedex 91405, France

<sup>2</sup>CNRS, Orsay 91405, France

**ABSTRACT:** Migratory species are sensitive to population declines because they require good habitat across their broad geographic distributions. For this reason, populations of migratory species require careful monitoring to permit detection of any change in their conservation status. An easily implemented technique is a census of individuals at a particular location along the migratory route, which can be used as an index of population status. However, in most situations, censuses are only performed during a portion of a particular season, especially when the presence of the species spans several months at a remote place. To be able to make comparisons among years, it is, thus, important to correct for partial counts. Several such procedures have been developed recently for marine turtles at nesting sites, birds during migration and the phenology of flowering in plants. In the current paper, I review different existing approaches, and then present a new statistical framework improving on those previously used. The method was tested with various time-series of marine turtle nest counts during the nesting season. Results demonstrate that the method is very efficient, and has the additional advantage of being easy to use by a wide variety of researchers.

**KEY WORDS:** Migration · Modeling · Population · Dynamics · Trend · Marine turtles

*Resale or republication not permitted without written consent of the publisher*

## INTRODUCTION

Animal movements are generally classified as dispersal or migration (Russell et al. 2005). When an animal disperses it moves from one home range to another, without an obligatory return to the original location. Conversely, migration has been described as the movement of an animal from one location to another, followed by its subsequent return to the original location (Alcock 2001). Migration is generally thought to be a consequence of biological constraints dictated by the seasonal movement between 2 geographic locations, including effects on the reproductive cycle, changes in temperature, and prey availability (Forcada 2002).

Many species migrate and, hence, they are only present at a particular geographic location for part of the year (Alerstam 1990, Berthold 1996). This behavior

produces 'tidal-waves' in the density of individuals along the migration route (Burdick 1943), with a lower frequency signal at wintering or aestivating grounds where the animals tend to stay longer. Accurate detection and quantification of individuals of a population during migration is challenging because the movement of individuals often masks or amplifies the true size of the population, particularly when not all individuals migrate each year (Winker et al. 1997).

This phenomenon of peaks in presence during the year has been described in birds (Alerstam 1990, Berthold 1996), marine mammals (Forcada 2002), and marine turtles (Russell et al. 2005), but is also found in many other taxa (Alcock 2001). Migration timing has been well studied in birds, for which long time-series have permitted detection of subtle changes in arrival and departure times (Lehikoinen et al. 2004). How-

\*Email: marc.girondot@u-psud.fr

ever, in most cases, only the date of first sighting or singing of a species is used (Gordo & Sanz 2006), although such a metric is very sensitive to false recognition of species or the aberrant behavior of a single individual (Lehikoinen et al. 2004).

For these species, population surveys often consist of counting individuals at a particular location during the season of occurrence. From these counts, it should be possible to generate a population trend using long-term data; however, this is often not the case because only partial counts are conducted, which prevents inter-annual analysis of the data. Nonetheless, partial count data can be obtained relatively easily if the season of presence is short. If individuals stay several months in a large, remote geographic region, the cost of acquiring data could become high. To address this issue, various sampling strategies have been promoted in recent years. One category of strategies is characterized by short periods of intensive counts (from 10 to 14 consecutive days) (Kerr et al. 1999, Jackson et al. 2008, Sims et al. 2008). While such a strategy is particularly powerful if only a very few sites must be followed, it quickly becomes infeasible when dozens of sites must be sampled simultaneously. For example, such a strategy would not work to assess sea turtle nesting abundances off the Caribbean Guadeloupe or Martinique Islands, where >150 nesting beaches must be monitored (Chevalier 2006). Furthermore, this strategy

hypothesizes that the peak of nesting does not vary from year to year; however, it may very well shift in response to warming conditions due to climate change (Weishampel et al. 2004).

Alternatively, various tools have been published to estimate the change in density of animals from sparse counts at one particular location during migration. In the present paper, I have analyzed various strategies used to complete partial counts during the occurrence season. All 6 published methods were found to suffer from weaknesses. I therefore propose a new model integrating the most interesting features of the published methods and introducing several new features. In exemplification, the new model has been tested with data from the leatherback turtle *Dermochelys coriacea* nesting season off French Guiana.

## ESTIMATION METHODS (TABLE 1)

### Models based on local fitting

In numerical analysis (and in its application to algorithmic discrete numerical analysis), interpolation is a mathematical operation used to construct a curve from the data of a finite number of points. The solution to the problem of interpolation passes through the prescribed points.

Table 1. Comparison of the various procedures used to describe the nesting seasons of marine turtles. LS: least squared; ML: maximum likelihood; (-): not applicable

Model	Nesting season	Fitting criteria	Error on parameters	Error on total no. of nests	Can generate negative value	Beginning and end of occurrence	Source
Linear interpolation and moving average	Local fitting	-	-	-	No	Must be fixed	Godley et al. (2001)
Lagrange interpolation	Local fitting	-	-	Jackknife	Yes	Must be fixed	Girondot & Fretey (1996), Girondot et al. (2002)
Generalized additive model	Local fitting	-	-	Yes	Yes	Most often must be fixed	Bjorndal et al. (1999), Troëng et al. (2004), Whiting et al. (2008), Witt et al. (2009)
Symmetric sinusoid	4 parameters	LS	No	Comparison with a table	No	Can be fixed	Gratiot et al. (2006)
Exponential	5 parameters	LS	Yes	Yes	No	Can be fixed	Malo (2002)
Truncated Gaussian	3 parameters	ML: Gaussian homoscedastic	Yes	Yes	No	Can be fixed	Jenni & Kéry (2003)
Product of 2 logistics	4 to 8 parameters	ML: Gaussian heteroscedastic	Yes for high-density beach	Yes	No	Can be fixed	Girondot et al. (2006)
Product of 2 logitics	4 to 8 parameters	ML: Poissonian	Bootstrap	Yes	No	Can be fixed	Godgenger et al. (2009)
Segments of sinusoids	4 to 7 parameters	ML: Negative binomial	Yes	Yes	No	Can be fixed	Present paper

### Linear interpolation

Linear interpolation is a method of curve fitting using linear polynomials. Missing data are obtained from a straight line between 2 knots. A derivative form has been used for Ascension Island green turtles *Chelonia mydas*, whereby the average of 2 counts between each segment was used as the estimate of missing counts (Godley et al. 2001). Quantitatively, it is identical to make a linear interpolation after a 2 d moving average. No error estimate is possible with this method.

### Waring-Euler-Lagrange interpolation (also known as Lagrange interpolation)

Waring-Euler-Lagrange polynomials allow interpolation of a series of points by generating polynomials that pass exactly through these points (known as knots in this context). Lagrange interpolation of polynomials gives no error estimate.

Lagrange interpolation with a third-order polynomial function has been used to estimate missing counts of leatherback *Dermochelys coriacea* nests on Yalimapo Beach in French Guiana (Chevalier & Girondot 1998, Girondot et al. 2002). Several problems arise when using such a method. First, the beginning and the end of the nesting season cannot be fitted. To alleviate this problem, 2 zero counts were included in the data set before and after the first and last expected beginning and end of the nesting season. It was thus clear that the beginning and the end of the nesting season were not outputs of this method. Second, some estimates could be negative and must, then, be converted to null values. However, this procedure could bias the output to higher than real values. Finally, to alleviate the lack of error analysis, jackknife re-sampling can be done (Girondot & Fretey 1996); in this case, however, the relationship with the true error is unknown.

### Generalized additive model

Generalized additive models (GAM) assume that the mean of the dependent variable depends on an additive predictor through a non-linear link function (Hastie & Tibshirani 1990). GAMs consist of a random component, an additive component, and a link function relating these 2 components. The response  $Y$ , the random component, is assumed to have a density in the exponential family, which includes normal, binomial, and Poisson distributions. The description of the methods is rather complicated, and beyond the scope of this paper. GAM models have been used in 4

marine turtle studies to complete missing data in time-series.

A GAM with a robust quasi-likelihood error function was used to fit a curve to track-survey results and to produce nest estimates for each date of a Central American study of leatherbacks *Dermochelys coriacea* (Troëng et al. 2004) and a study of green turtles in Costa Rica (Bjorndal et al. 1999). In both cases, artificial beginning and end dates were imposed (i.e. set to zero) and weighted as one-tenth of the true counts. Negative estimates were trimmed from the beginning and end of each season. A similar procedure was used for flatback turtles *Natator depressus* in Australia, but without limiting the unknown beginning and end of nesting season (Whiting et al. 2008). It seems that the same methodology was used to estimate missing counts for leatherbacks in Gabon; however, the exact methodology used is not explained beyond: '...we used a cubic spline approach.' (Witt et al. 2009). However, in this last case, only 12 or 13 counts were missing, which would not strongly affect the final result.

### Models based on a parametric description of the nesting season

Given a specific data set, one can always use a non-parametric method to estimate the functions. These methods are particularly useful in visualizing the functions, thus giving a good overall idea of their shape, which is important. This can be further refined by constructing confidence bands (or, in other words, simultaneous confidence intervals) for the functions. However, interpretation and comparison require a parametric approach (see Bebbington et al. 2007). The overall shape of occurrence generally resembles a bell curve. Four families of parametric models have been developed to render this shape.

#### Symmetric sinusoid

The nesting season of marine turtles has been described by a truncated sinusoidal function (Gratiot et al. 2006), such that:

$$Y\left(\left[T - \frac{tp}{2} : T + \frac{tp}{2}\right]\right) = \frac{A}{2} \cos\left(\frac{2\pi}{tp}(t - T)\right) + \frac{A}{2} + B \quad (1)$$

$$Y\left(\left[1 : T - \frac{tp}{2}\right]\right) = Y\left(\left[T + \frac{tp}{2} : 365\right]\right) = B \quad (2)$$

where  $Y(t)$  is the number of nests at Day  $t$ ,  $A$  is the amplitude of the sinus function,  $tp$  is the duration of the turtle nests (in days),  $T$  is the mean position of the nesting season in the year (in days from 1 January or any other date used as a reference), and  $B$  is the residual

number of nests outside the season. 'Outside the season' means the period of the year during which nesting is incidental ( $<5$  nests  $\text{wk}^{-1}$ ).  $A$  has previously been described incorrectly as 'the higher number of nests estimated for the nesting season' (Gratiot et al. 2006); the authors were probably actually referring to  $A + B$  being the mean number of nests at the peak of the nesting season.

Gratiot et al. (2006) performed a complete screen of the space of the parameters  $A$ ,  $tp$ , and  $T$  by steps of 1 and then 0.1. The  $B$  value was estimated prior to analysis (N. Gratiot pers. comm.). Least-square residuals between observations and estimate minimizations were used to adjust criteria. Several statistical problems arise with this procedure. The value of the  $B$  parameter is estimated before using any fitting procedures. However, at that stage, the exact form of the nesting season curve for the corresponding data-set is then still unknown because other parameters have not been fitted. Moreover, the initial increase in nest numbers is very slow; thus, it is not clear which method would specify when a night count is necessary to evaluate the  $B$  value. Furthermore, this 2-step procedure permits calculation of  $T$ ,  $tp$ , and  $A$ , which minimizes least-squares conditioning of the  $B$  value. Therefore, this procedure does not guarantee that the value of  $B$  which minimized least-square residuals can be found.

The minimization of least-squares residuals is based on a Gaussian distribution with a constant standard deviation ( $\sigma$ ). Indeed, it can be easily demonstrated that the least-squares fit is equivalent to the maximum-likelihood fit with a constant  $\sigma$  (Hilborn & Mangel 1997). However, a simple visual examination of the time-series used shows that variance of nest counts is higher at the nesting peak than at the beginning and end of the nesting season (Fig. 1 in Gratiot et al. 2006). This feature was observed whenever it was tested for nesting time-series of marine turtles (Girondot et al. 2006, Godgenger et al. 2009). Godgenger et al. (2009) discussed the potential consequence of not taking heteroscedasticity into account and showed that this leads to bias of the output (see also Power & Moser 1999). The procedure is technically so complicated that it is impossible to derive an error value for the parameters.

#### Truncated exponential sine equation

A truncated exponential sine equation has been used to describe unimodal flowering phenology (Malo 2002; in the original publication, there was an error in the central condition):

$$\begin{cases} \text{if } t < b, & f(t) = 0 \\ \text{if } b \leq t \leq (b + c), & f(t) = a[\sin[\pi((t-b)/c)^d]]^e \\ \text{if } t > (b + c), & f(t) = 0 \end{cases} \quad (3)$$

Eq. (3) produces a wave along the  $x$ -axis, the maximum of which is determined by  $a$ , its length by  $c$ , its asymmetry is controlled by  $d$ , and the length of the tails by  $e$ . From Eq. (3) it is also possible to derive the date of the phenological peak [when  $f(t)$ , the number of flowers per plant, is maximized,  $b + c(1/2)^{1/d}$ ] and the total number of flower-days during the flowering period. In this case, the non-linear estimation module with the minimization procedure for least-squares residuals is used to estimate parameters (Malo 2002) (see previous subsection on the minimization of least-squares residuals).

#### Truncated Gaussian distribution

In a study of bird migration through the Col de Bretolet, a mountain pass in the Swiss–French Alps, Jenni & Kéry (2003) indicate that the time-series was a truncated sample of the actual population that migrated across Col de Bretolet, either because premigratory movements were cut off or because the station had to be closed down for winter before the birds' passage was complete. To estimate peak passage, the authors fitted a truncated normal distribution to the trapping densities per 5 d interval. An iterative algorithm was used to obtain the maximum-likelihood estimate of the mean of the distribution (corresponding to the estimated peak passage) and its standard error. Both truncation points were specified from the observational record.

The link function used to estimate the maximum likelihood is not specified in the original publication, but based on the electronic Appendix, it seems that a homoscedastic Gaussian distribution was used (see previous subsection on the use of homoscedastic Gaussian distributions). Furthermore, the normal distribution used is obligatory symmetric.

#### Product of modified logistics

Girondot et al. (2006) modeled the peaked nesting pattern for leatherback turtles using the product of 2 sigmoid equations, the first one ranging from 0 to 1 and the second one ranging from 1 to 0. The product shows a 0-1-0 pattern if the transition of the first equation is observed at an abscissa of lower value than the second one. For each sigmoid equation, a modified form of the classical Verhulst equation (Verhulst 1846) that allows asymmetry to be set is used. The resulting equation is:

$$M(d) = \left( 1 + (2^{e^k} - 1) e^{\left(\frac{1}{s}(P-d)\right)^{-1/e^k}} \right)^{-1/e^k} \quad (4)$$

where  $d$  is the day of the year,  $P$  refers to the dates before and after the peak nesting day when there is an

observed maximum rate of change (increase or decrease) in nest numbers, and  $S$  and  $K$  are related to the change in slope of the change in nest numbers at date  $P$ .

The value of  $M(d)$  ranges from 0 to 1 with  $M(d) = 0.5$  for  $P = d$ , with  $d$  being the number of days since the start of the nesting season. The steepness of  $M(d)$  at  $P = d$  depends on  $S$  and  $K$  values.  $M(d)$  increases when  $S$  is negative (i.e. at the start of the nesting season) and decreases when  $S$  is positive (i.e. at the end of the nesting season). Asymmetry around  $P$  is determined by a positive or negative value of  $K$ . Eq. (3) is reduced to a simple logistic equation (i.e. symmetrical around  $P$ ) when  $K = 0$ . The mathematical description of a nesting season ( $N$ ) can therefore be expressed as

$$N(d) = \text{Min} + (\text{Max} - \text{Min}) \times (M_1(d) \times M_2(d)) \quad (5)$$

with  $M_1(d)$  and  $M_2(d)$  being the first and second halves of the nesting season, respectively. The difference between the 2 values largely rests on the sign of the  $S$  parameter:  $S_1$  is negative,  $S_2$  is positive, and  $P_1 < P_2$ . The parameter 'min' is the basal level of nesting outside the nesting season, and 'max-min' is a scaling factor.

For a high-density nesting beach, a maximum-likelihood fit was used with a Gaussian distribution of nests. To account for heteroscedasticity and the standard deviation, it was assumed that nests on day  $d$  are normally distributed with a standard deviation, such that  $\sigma_d = \exp[a N(d)^c + b]$ , where  $a$ ,  $b$ , and  $c$  are parameters that are also fitted (Girondot et al. 2006). When a similar procedure was used for a low-density nesting beach, the correction for heteroscedasticity was insufficient. Indeed, when low mean values and high variances are observed for low-density nesting beaches, the normal probability distribution, because of its symmetry, can imply a substantial probability of observing a negative number of nests. Because of these problems with the assumption of a Gaussian probability distribution, an alternative Poissonian distribution was implemented (Godgenger et al. 2009).

In probability theory and statistics, the Poisson distribution is a discrete probability distribution that expresses the probability of a number of events occurring in a fixed period of time if these events occur at a known average rate and are independent of the time since the last event. The probability that there are exactly  $x$  occurrences ( $x$  being a non-negative integer,  $x = 0, 1, 2, \dots$ ) where the expected number of occurrences is ( $\lambda$ ) is represented by:

$$f(k, \lambda) = \frac{e^{-\lambda} \lambda^x}{x!} \quad (6)$$

In Girondot et al. (2006), variance of the parameters is estimated from the inverse of the matrix of the sec-

ond-order partial derivatives (also called the Hessian matrix). Second-order partial derivatives are approximated using a fourth-order polynomial function around the fitted value for each parameter and pair of parameters. Standard errors are obtained by taking the square-root of the diagonal elements of the inverse of the resulting matrix. Godgenger et al. (2009) used a bootstrap strategy to estimate error on parameters.

## A NEW MODEL FOR MIGRATION WAVES

From the previous section, it appears that local fitting to describe the nesting season is not fully satisfactory even if the GAM model does permit the acquisition of some useful information. Therefore, in the following, I have focused on defining a better parametric model. I first discuss the model for the shape of occurrence during the season and then look at the fitting procedure.

### Occurrence during season

The Girondot et al. (2006) and Malo (2002) models have the advantage of being more flexible and allowing asymmetric seasons to be modeled (i.e. slow beginning and rapid end or inverse). It should be noted, however, that the shape produced by the equation formulated by Malo (2002) imposes a link between the rate of change at the beginning and the end. On the other hand, in the Girondot et al. (2006) model, a similar shape can be obtained from a large set of parameters because some of them compensate for each other (Godgenger et al. 2009). Thus, the identifiability of parameters, one of the conditions for the consistency of a maximum-likelihood estimate (Wald 1949), is not satisfied. In this case, the inverse of the information matrix by Fisher cannot be used to obtain the standard error of parameters (Godgenger et al. 2009). The Gratiot et al. (2006) and Jenni & Kéry (2003) models do not have this problem, but they impose a symmetry in the season (i.e. same rate of increase and decrease).

To reap the advantages of all models, a modified form of the sinusoidal equation should be used to render an asymmetric pattern of nesting season and to use the Fisher's information matrix for parameters. Furthermore, in all these models, the basal number of nests out of the nesting season is the same for the beginning and end of the nesting season. This constraint has been removed in the current version of the model. This is particularly useful for French Guiana and Suriname leatherback nesting data, where the beginning of the nesting season is mixed with the end of the so-called 'small-nesting season' (Girondot et al. 2007).

Letting  $t$  be a day of the year, the number of nests deposited per night is modeled using the system of equations below:

$$\left\{ \begin{array}{l} \text{if } t < B \rightarrow \text{Min}B \\ \text{if } t \in [B, P - F/2] \rightarrow ((1 + \cos(\pi(P - F/2 - t)/(P - F/2 - B)))/2) \\ \quad (\text{Max} - \text{Min}B) + \text{Min}B \\ \text{if } t \in [P - F/2, P + F/2] \rightarrow \text{Max} \\ \text{if } t \in [P + F/2, E] \rightarrow ((1 + \cos(\pi(t - P + F/2)/(E - P + F/2)))/2) \\ \quad (\text{Max} - \text{Min}E) + \text{Min}E \\ \text{if } t > E \rightarrow \text{Min}E \end{array} \right. \quad (7)$$

The model requires, at most, 7 parameters, which are defined graphically in Fig. 1. The formulas have been constructed to allow the parameters to have direct biological interpretations:

- $\text{Min}B$  is the mean nightly nest number before the beginning of the nesting season;
- $\text{Min}E$  is the mean nightly nest number after the end of the nesting season;
- $\text{Max}$  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 of the graph flattens out (Fig. 1);
- $B$  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.

Various constraints can be set up to simplify this model:  $\text{Min}B = \text{Min}E$ , for the same number of nests out

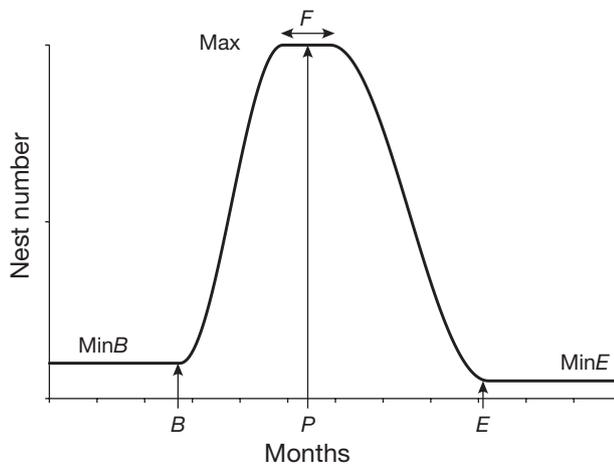


Fig. 1. Graphical representation of the variables used to describe a nesting season.  $\text{Min}B$ : mean nightly nest number before the beginning of the nesting season;  $\text{Min}E$ : mean nightly nest number after the end of the nesting season;  $\text{Max}$ : mean number of nests at the peak of the nesting season;  $P$ : day of the year on which the nesting season peaks;  $F$ : the number of days around Day  $P$  on which the curve of the graph flattens out;  $B$ : day of the year on which the nesting season begins;  $E$ : day of the year on which the nesting season ends

of the nesting season;  $\text{Min}B$  and/or  $\text{Min}E \approx 0$  (e.g.  $10^{-9}$ ), when no nests are observed out of the nesting season;  $P - B = E - P$ , when the nesting season is symmetric around  $P$ ; and  $F = 0$ , for no flat portion.

The simplest model uses 4 parameters ( $\text{Min}B = \text{Min}E$ ,  $P$ ,  $P - B = E - P$ ,  $\text{Max}$ ,  $F = 0$ ) and is similar to the Gratiot et al. (2006) model.

Whereas the nesting season is described by segment, the formulas developed allow all segments to be in continuity. The nesting season is defined as nesting in the interval  $[B, E]$ . Note that this is not exactly the same definition as that given by Girondot et al. (2006), who consider the range of time in which  $>5\%$  of the mean number of nests at the peak of the nesting season are observed. A conversion from one to the other definition must be done numerically.

### Fitting criteria

The nest distribution per night has been implicitly assumed to be homoscedastic Gaussian by Gratiot et al. (2006), heteroscedastic Gaussian by Girondot et al. (2006) and Poissonian by Godgenger et al. (2009). Clearly the homoscedastic hypothesis is false and biases the output (Godgenger et al. 2009). An attempt to use a truncated Gaussian distribution (Hegde & Dahiya 1989) gives very poor and biased estimates of parameters (not shown). The heteroscedastic Gaussian distribution fits the data well for nesting beaches with large numbers of nests per night (Girondot et al. 2006), and the Poissonian distribution can be used for beaches with low numbers of nests per night (Godgenger et al. 2009). When the Poissonian model is fitted on high-density beaches, the dispersion around the mean is not large enough, as nearly 50% of the counts are not within the envelope of  $\pm 2\sigma$ . The negative binomial distribution can be described as a mixture of Poissonian distributions (Lawless 1987) and is used to model count data with varying degrees of overdispersion. The distribution of the counts  $X$  is commonly expressed in terms of the mean  $m$  and the dispersion parameter  $k$ , such that the probability of observing a non-negative integer  $x$  is (Anscombe 1949):

$$\text{Pr}(X = x) = \frac{\Gamma(k + x)}{x! \Gamma(k)} \left( \frac{m}{m + k} \right)^x \left( 1 + \frac{m}{k} \right)^{-k}, m > 0, k > 0 \quad (8)$$

The negative binomial distribution has broad applications as a model for count data, particularly for data exhibiting overdispersion (i.e. with sample variance exceeding the mean) (Lloyd-Smith 2007). In the biological literature, classical uses of the negative binomial distribution include analysis of parasite loads, species occurrence, parasitoid attacks, abundance samples, and spatial clustering of populations (White & Bennetts 1996).

In the following, a Levenberg-Marquart non-linear fitting algorithm will be used to adjust parameter values (Press et al. 1992).

### Testing the new model

Table 1 provides an overview of the various constraints for previously published models and for the new model. From this table, it is clear that the model presented here represents an improvement over previous ones, as all constraints have been removed. To facilitate easy use, it has been scripted and is available as ready-to-use software at [www.ese.u-psud.fr/epc/conservation/Girondot/Publications/Marine\\_Turtles\\_Nesting\\_Season.html](http://www.ese.u-psud.fr/epc/conservation/Girondot/Publications/Marine_Turtles_Nesting_Season.html).

To test various options, I used a complete time-series of nest counts for leatherback turtles *Dermochelys coriacea* obtained in the Cayenne area of French Guiana in 2001. This time-series was extracted from Fig. 1 of Gratiot et al. (2006).

### Comparing models

As the time-series used applies to marine turtles, only the new model and the model by Gratiot et al. (2006) were tested (other parametric models do not include the basal out-of-season value). To test between the two models, we need a common statistical basis. Gratiot et al. (2006) used the correlation coefficient between observations and estimations. However, this statistical treatment is inadequate for non-stationary time-series such as those analyzed here, as any distribution showing a peak in the middle of the time-series will generate an inflated  $r^2$  value. The correct alternative is to use Akaike's information criterion (AIC; Akaike 1974). This is a ranking measure that takes into account the quality of the fit of the model to the data while penalizing for the number of parameters used:

$$\text{AIC} = -2\ln L + 2M \quad (9)$$

where  $L$  is the maximum likelihood and  $M$  is the number of parameters. The models with the lowest AIC values were retained as good candidate models, and  $\Delta\text{AIC}$  was calculated as the difference in the AIC value between a particular model and that of the model with the lowest AIC. Akaike weights [ $w_i = \exp(-\Delta\text{AIC}/2)$  normalized to 1] were used to evaluate the relative support of various tested models (Burnham & Anderson 2002). Akaike weights can be directly interpreted as conditional probabilities for each model. Ideally, the model with the lowest AIC was retained for further testing. When 2 or more models possessed similar

Akaike weights, the model with the lowest number of parameters was selected.

When the time-series was fitted with both the symmetrical sinusoid by Gratiot et al. (2006) and the model presented here, the Gratiot et al. (2006) model was not retained as a candidate model ( $p = 10^{-7}$ ). The symmetrical model suffered a similar lack of support ( $p = 9 \times 10^{-6}$ ). Hence, contrary to the claim of Gratiot et al. (2006), even this time-series, which appears superficially symmetrical, has a strong asymmetrical component. The retained models based on Akaike weight ( $p = 0.71$  and  $p = 0.21$ ) indeed showed a slight but significant shift in the increase and decrease rates of change in nest numbers per night. The increase is slightly slower than the decrease, which changed the estimates of the beginning and peak of the nesting season by 13 and 7 d, respectively (Fig. 2). Whereas this pattern appears to be insignificant, it is on the same order of magnitude as the change detected in the median day of nesting activity in Florida for *Caretta caretta*, which was related to changes in sea surface temperature (Weishampel et al. 2004). Thus, it is important to have the best model available to detect such slight changes.

### Filling the missing data

The complete time series was reduced by removing from 1 to 6 of 7 consecutive nightly nest counts (each scenario was run 100 times). For each of the reduced data sets, the fitting procedure was run and the total number of nests was estimated by combining actual observations and model estimates for the missing dates. The standard deviation and coefficient of variation for the number of nests for each level of reduction were calculated.

The results are presented in Fig. 3. The coefficient of variation is from 0.11, when only 1 night count out of 7 is retained, to 0.02, when 6 nights out of 7 are retained. The outputs are well centered around the true value in 7 out of 7 cases, and, consequently, the outputs are not biased.

### Comparing nesting beaches or years

The use of maximum likelihood and AIC facilitate straightforward comparison of nesting patterns among a group of nesting beaches and/or various years. As an example, we used a time-series of nest counts, also from leatherbacks in French Guiana in 2001, but obtained 250 km to the northwest of Cayenne, on Yalimapo beach.

Although the entire nesting season was patrolled in the Cayenne region, only 27 counts are available for

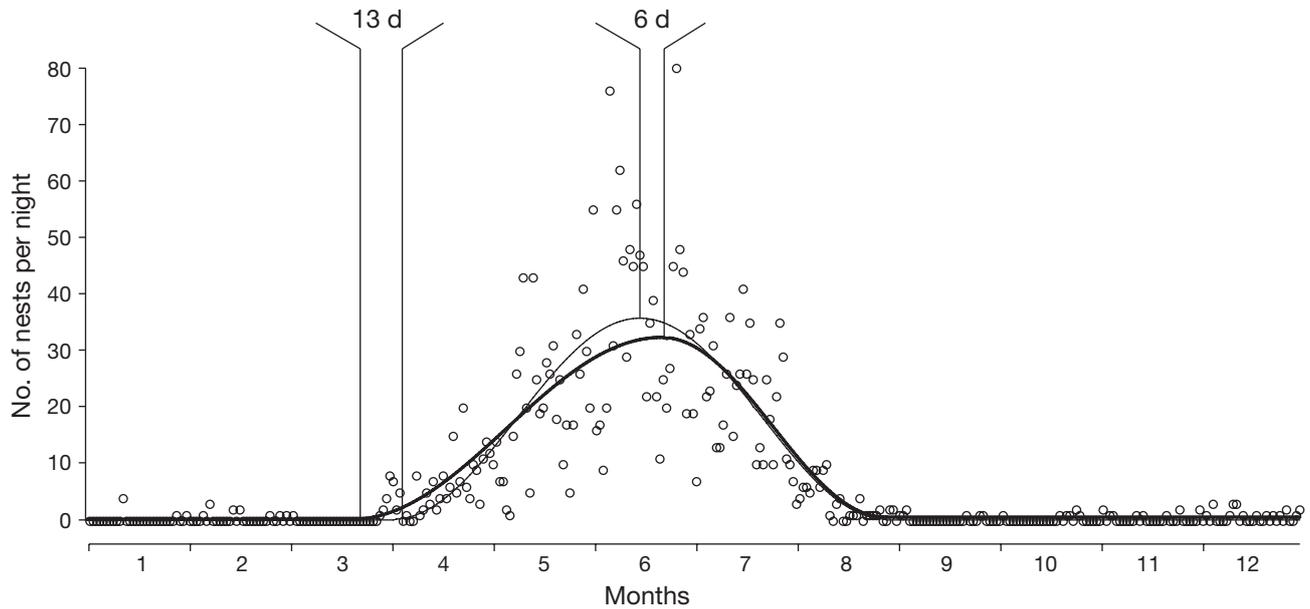


Fig. 2. *Dermochelys coriacea*. Complete time-series of nest counts for leatherback turtles in 2001 in eastern French Guiana obtained from Gratiot et al. (2006). The thin curve is the nesting season fitted by Gratiot et al. (2006), and the bold curve is the nesting season fitted by the new model presented here. The first 2 vertical lines show the beginning of the nesting season according to both models, and the second set of vertical lines show the peak of the nesting season

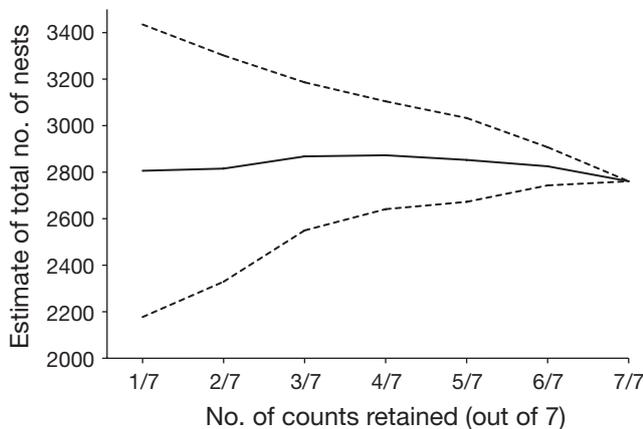


Fig. 3. *Dermochelys coriacea*. Estimate of the number of nests according to the quantity of information that is maintained in the time-series. The envelope (between dashed lines) represents  $\pm 2$  SD

Yalimapo beach. The nesting season was fitted as described previously with 1 set of parameters for each location (see Table 2, Cayenne:  $-\ln L = 675.08$ ,  $AIC = 1364.16$ ; Yalimapo:  $-\ln L = 109.45$ ,  $AIC = 232.90$ ). Both time-series were then fitted at the same time with a common set of parameters, leaving only the parameters Max and  $k$  to be beach specific. The maximum likelihood for a single set of parameters is 792.28, and the AIC is 1606.56. The AIC for the beach-specific sets of parameters was  $1364.16 + 232.90 = 1597.06$ , which is lower than the AIC when only 1 set of parameters is fit-

ted (1606.56). This demonstrates that there was a significant difference in the shape of the nesting season curve between eastern and western French Guiana for that year ( $\Delta AIC = 1606.56 - 1597.06 = 9.5$ ,  $p = 0.008$ ).

The comparison of fitted parameters permitted identification of the origin of the difference. The peak and end of the nesting season were very similar in both locations (the confidence interval at 95% is given in parentheses), whereas the beginning is much earlier in Yalimapo. The peak of the nesting season in Cayenne was 22 June 2001 (15 to 30 June 2001), and in Yalimapo 20 June 2001 (7 June to 4 July 2001); the end of the nesting season in Cayenne was 29 August 2001 (24 August to 3 September 2001), and in Yalimapo 29 August 2001 (22 August to 5 September 2001); the beginning of the nesting season in Cayenne was 21 March 2001 (17 to 26 March 2001), and in Yalimapo 25 February 2001 (20 February to 3 March 2001). This observation is consistent with the hypothesis that most females come back from northern feeding areas (James et al. 2005) and can stop to nest on northern beaches on their way toward the beach that they will use for most of the season.

## CONCLUSIONS

In the present paper, I reviewed various strategies used to complete partial counts of nests during the nesting season. Six published methods were analyzed, and all of them suffered from weaknesses. A new

Table 2. Comparison of the fit obtained by Gratiot et al. (2006) with the fit obtained using the model developed here. For Gratiot et al's model. *B*: residual number of nests outside the season; *A*: amplitude of the sinus function; *tp*: duration of the turtle nests (d); *T*: mean position of the nesting season in the year (in days from 1 January or any other date used as a reference); *k*: dispersion parameter. For the new model developed here, *MinB*: the mean nightly nest number before the beginning of the nesting season; *MinE*: mean nightly nest number after the end of the nesting season; *Max*: mean number of nests at the peak of the nesting season; *P*: day of the year on which the nesting season peaks; *F*: number of days around Day *P* on which the curve of the graph flattens out; *B*: day of the year on which the nesting season begins; *E*: day of the year on which the nesting season ends. *L*: maximum likelihood; AIC: Akaike's information criterion;  $w_i$ : Akaike weight

Model	$-\text{Ln}L$	Parameters	AIC	$w_i$
Gratiot et al.: <i>B</i> , <i>A</i> , <i>tp</i> and $T + k^a$	691.72	5	1395.45	$1.10^{-7}$
$\text{Min}B = \text{Min}E$ , $P - B = E - P$ , <i>Max</i> , <i>P</i> , <i>k</i>	688.32	5	1386.64	$9.10^{-6}$
<i>MinB</i> , <i>MinE</i> , <i>B</i> , <i>P</i> , <i>E</i> , <i>Max</i> , <i>k</i>	675.08	7	1364.16	0.72
<i>MinB</i> , <i>MinE</i> , <i>B</i> , <i>P</i> , <i>E</i> , <i>F</i> , <i>Max</i> , <i>k</i>	675.08	8	1366.16	0.26
$\text{Min}B = \text{Min}E$ , <i>B</i> , <i>P</i> , <i>E</i> , <i>Max</i> , <i>k</i>	680.17	6	1372.35	0.01
$\text{Min}B = \text{Min}E$ , <i>B</i> , <i>P</i> , <i>E</i> , <i>F</i> , <i>Max</i> , <i>k</i>	680.17	7	1374.35	0.004

<sup>a</sup>Maximum likelihood for this model was estimated by fitting only the dispersion parameter of the negative binomial distribution *k*; all other parameters are fixed to the values obtained by Gratiot et al. (2006)

method, therefore, has been developed to remedy these weaknesses. This method proved to be both easy to use and very efficient in answering biological and conservation questions relative to migratory species. Furthermore, it is especially helpful in determining the population size at a particular location along the migratory route or at a stop.

The Bonn Convention on the Conservation of Migratory Species of Wild Animals (1979, p. 1; CMS) is an important instrument in the management of migratory species. It defined a migratory species as

...the entire population or any geographically separate part of the population of any species or lower taxon of wild animals, a significant proportion of whose members cyclically and predictably cross one or more national jurisdictional boundaries.

Migratory species are particularly threatened because they require a vast range of high-quality habitats (Hutto 2000) and encounter many dangers during migration (Myers et al. 1987) and because they are subject to shifting locations of patches of resources in the context of global change (Robinson et al. 2009). Calculating the size of a population is an essential step in assessing population status and trends.

However, migratory species with very large ranges, such as marine turtles, are particularly difficult to monitor. There are indeed various challenges associated with directly counting the total number of individuals in a marine turtle population, including cryptic life-history stages, trans-oceanic dispersal, and non-sequential annual reproduction (Bolten & Witherington 2003). As a result, researchers have traditionally relied on enumerating numbers of individuals observed during a particular life stage as an index of population size (Gerrodette & Taylor 1999). Until recently, the objective of many marine turtle conserva-

tion programs was to estimate populations using exhaustive observation and tagging of individuals (Richardson et al. 2006). However, changes have occurred in the practices used, and several authors have proposed various strategies to reduce the amount of work needed in the field, but without relinquishing any of the power to answer biological questions and spot trends (Kerr et al. 1999, Jackson et al. 2008, Sims et al. 2008). The model proposed here is an alternative to published methods, offering several advantages over them.

The new method allows estimation of an index of the population size and its standard deviation with only a few counts during the nesting season. The number of monitoring nights needed to reach a level of standard error for the estimate can be derived from Fig. 3. The strength of this method is to simultaneously use all the information gathered during the monitoring of several beaches, both within a season and between seasons. It is then possible to test for the effect of beaches, groups of beaches, or year on the phenology of migration to the nesting site. Obviously, such estimates and tests can be performed for any migratory species, not only marine turtles.

The parameters in this new model have direct biological relevance. In many cases, prior data are known for them, such as the beginning of the nesting season, the peak, or its end. Thus, the methodology presented here could easily be packed within a Bayesian framework, as the definition of sensible priors is quite straightforward. This would be most useful and potentially very powerful when considering data sets that are quite sparse.

With the powerful statistical methods that have recently become available, we can reduce the amount of data necessary to draw conclusions about popula-

tion abundances and trends. For example, it has been shown that data could be obtained simply by a few aerial surveys during the nesting season (Witt et al. 2009). Such data could be analyzed with the methodology described here to estimate total nesting and its standard error. Less field monitoring will be required to realize a level of confidence in the estimates similar to that provided by near-complete field monitoring. This might lead conservation programs to cut their financial support to local communities, because less on-the-ground work would be necessary. However, when local communities are employed to monitor wildlife, they become conservation stewards and are less likely to unsustainably exploit it. Such a financial cut could result in the risk that exploitation of wildlife may increase again to compensate for the lower income in the community. If new statistical methods permit a reduction in field monitoring efforts, it would also be important to consider potential economic alternatives for local communities, in order to ensure continued conservation.

*Acknowledgements.* Some of the ideas used in the present paper emerged during various brainstorming sessions conducted by the Scientific Advisory Board of SWOT (State of the World for Sea Turtles). Brian Hutchinson, Rod Mast, and Bryan Wallace are particularly thanked. Marie-Clélia Godgenger and Mathilde Russo tested preliminary versions of the software and contributed to making it functional. Bryan Wallace was of great help in correcting the English and in pointing out specific points that were unclear. Anna Han is thanked for the wonderful discussions about turtle carapaces during the writing of this paper, as well as for English corrections. Three referees and the editor are thanked for having helped to clarify some points and for their suggestions for future enhancements of this model.

#### LITERATURE CITED

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* 19:716–723
- Alcock J (2001) *Animal behaviour*, 7th edn. Sinauer Associates, Sunderland, MA
- Alerstam T (1990) *Bird migration*. Cambridge University Press, Cambridge
- Anscombe FJ (1949) The statistical analysis of insect counts based on the negative binomial distribution. *Biometrics* 5: 165–173
- Bebbington M, Lai CD, Zitakis R (2007) Modeling human mortality using mixtures of bathtub shaped failure distributions. *J Theor Biol* 245:528–538
- Berthold P (1996) *Control of bird migration*. Chapman & Hall, London
- Bjorndal KA, Wetherall JA, Bolten AB, Mortimer JA (1999) Twenty-six years of green turtle nesting at Tortuguero, Costa Rica: an encouraging trend. *Conserv Biol* 13: 126–134
- Bolten AB, Witherington BE (2003) *Loggerhead sea turtles*. Smithsonian Books, Washington, DC
- Burdick HC (1943) A marked migration wave during the 1943 spring season. *Bird-Banding* 14:131–133
- Burnham KP, Anderson DR (2002) *Model selection and multi-model inference: a practical information-theoretic approach*. Springer-Verlag, New York, NY
- Chevalier J (2006) *Plan de restauration des tortues marines des Antilles françaises*. Office National de la Chasse et de la Faune Sauvage, Fort-de-France, p 162
- Chevalier J, Girondot M (1998) *Dynamique de pontes des Tortues luths en Guyane française durant la saison 1997*. *Bull Soc Herpetol* 85/86:5–19
- Forcada J (2002) Distribution. In: Perrin WF, Würsig B, Thewissen H (eds) *Encyclopedia of marine mammals*. Academic Press, San Diego, CA, p 327–333
- Gerrodette T, Taylor BL (1999) Estimating population size. In: Eckert KL, Bjorndal KA, Abreu-Grobois FA, Donnelly M (eds) *Research and management techniques for the conservation of sea turtles*. IUCN/SSC Marine Turtle Specialist Group Publication No. 4, Washington, DC, p 67–71
- Girondot M, Fretey J (1996) Leatherback turtles, *Dermodochelys coriacea*, nesting in French Guiana, 1978–1995. *Chelonian Conserv Biol* 2:204–208
- Girondot M, Tucker AD, Rivalan P, Godfrey MH and others (2002) Density-dependent nest destruction and population fluctuation of Guianan leatherback turtles. *Anim Conserv* 5:75–84
- Girondot M, Rivalan P, Wongsopawiro R, Briane JP and others (2006) Phenology of marine turtle nesting revealed by a statistical model of the nesting season. *BMC Ecol* 6:11. Available at [www.biomedcentral.com/1472-6785/6/11](http://www.biomedcentral.com/1472-6785/6/11)
- Girondot M, Godfrey MH, Ponge L, Rivalan P (2007) Modeling approaches to quantify leatherback nesting trends in French Guiana and Suriname. *Chelonian Conserv Biol* 6: 37–46
- Godgenger MC, Bréheret N, Bal G, N'Damité K and others (2009) Nesting estimation and analysis of threats for Critically Endangered leatherback *Dermodochelys coriacea* and Endangered olive ridley *Lepidochelys olivacea* marine turtles nesting in Congo. *Oryx* 43:556–563
- Godley BJ, Broderick AC, Hays GC (2001) Nesting of green turtles (*Chelonia mydas*) at Ascension Island, South Atlantic. *Biol Conserv* 97:151–158
- Gordo O, Sanz JJ (2006) Climate change and bird phenology: a long-term study in the Iberian Peninsula. *Glob Change Biol* 12:1993–2004
- 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
- Hastie TJ, Tibshirani RJ (1990) *Generalized additive models*. Chapman & Hall/CRC, New York, NY
- Hegde LM, Dahiya RC (1989) Estimation of the parameters in a truncated normal-distribution. *Comm Statist Theory Methods* 18:4177–4195
- Hilborn R, Mangel M (1997) *The ecological detective. Confronting models with data*, Vol 28. Princeton University Press, Princeton, NJ
- Hutto RL (2000) On the importance of *en route* periods to the conservation of migratory landbirds. *Stud Avian Biol* 20: 109–114
- Jackson AL, Broderick AC, Fuller WJ, Glen F and others (2008) Sampling design and its effect on population monitoring: How much monitoring do turtles really need? *Biol Conserv* 141:2932–2941
- James MC, Myers RA, Ottensmeyer CA (2005) Behaviour of leatherback sea turtles, *Dermodochelys coriacea*, during the migratory cycle. *Proc R Soc Lond B* 272:1547–1555
- Jenni L, Kéry M (2003) Timing of autumn bird migration under climate change: advances in long-distance mi-

- grants, delays in short-distance migrants. *Proc Biol Sci* 270:1467–1471
- Kerr R, Richardson JI, Richardson TH (1999) Estimating the annual size of hawksbill (*Eretmochelys imbricata*) nesting populations from mark-recapture studies: the use for long-term data to provide statistics for optimizing survey effort. *Chelonian Conserv Biol* 3:251–256
- Lawless JF (1987) Negative binomial and mixed Poisson regression. *Can J Stat* 15:209–225
- Lehikoinen E, Sparks TH, Zalakevicius M (2004) Arrival and departure dates. In: Møller AP, Fiedler W, Caswell H, Berthold P (eds) *Advances in ecological research: birds and climate change*. Elsevier, Oxford, p 1–32
- Lloyd-Smith JO (2007) Maximum likelihood estimation of the negative binomial dispersion parameter for highly overdispersed data, with applications to infectious diseases. *PLoS ONE* 2:e180
- Malo JE (2002) Modelling unimodal flowering phenology with exponential sine equations. *Funct Ecol* 16:413–418
- Myers JP, Morrison RIG, Antas PZ, Harrington BA and others (1987) Conservation strategy for migratory species. *Am Sci* 75:18–26
- Power JH, Moser EB (1999) Linear model analysis of net catch data using the negative binomial distribution. *Can J Fish Aquat Sci* 56:191–200
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) *Numerical recipes in C: the art of scientific computing*. Cambridge University Press, Cambridge
- Richardson JI, Hall DB, Mason PA, Andrews KM and others (2006) Eighteen years of saturation tagging data reveal a significant increase in nesting hawksbill sea turtles (*Eretmochelys imbricata*) on Long Island, Antigua. *Anim Conserv* 9:302–307
- Robinson RA, Crick HQP, Learmonth JA, Maclean IMD and others (2009) Travelling through a warming world: climate change and migratory species. *Endang Species Res* 7: 87–99
- Russell AP, Bauer AM, Johnson MK (2005) Migration in amphibians and reptiles: an overview of patterns and orientation mechanisms in relation to life history strategies. In: *Migration of organisms*. Springer, Heidelberg, p 151–203
- Sims M, Bjorkland R, Mason P, Crowder LB (2008) Statistical power and sea turtle nesting beach surveys: How long and when? *Biol Conserv* 141:2921–2931
- Troëng S, Chacón D, Dick B (2004) Possible decline in leatherback turtle *Dermochelys coriacea* nesting along the coast of Caribbean Central America. *Oryx* 38:395–403
- Verhulst PF (1846) Deuxième mémoire sur la loi d'accroissement de la population. *Mem Acad R Sci Lettres Beaux-Arts Belg* 20:1–32
- Wald A (1949) Note on the consistency of the maximum likelihood estimate. *Ann Math Stat* 20:595–601
- Weishampel JF, Bagley DA, Ehrhart LM (2004) Earlier nesting by loggerhead sea turtles following sea surface warming. *Glob Change Biol* 10:1424–1427
- White GC, Bennetts RE (1996) Analysis of frequency count data using the negative binomial distribution. *Ecology* 77: 2549–2557
- Whiting AU, Thomson A, Chaloupka M, Limpus CJ (2008) Seasonality, abundance and breeding biology of one of the largest populations of nesting flatback turtles, *Natator depressus*: Cape Domett, Western Australia. *Aust J Zool* 56:297–303
- Winker K, Escalante P, Rappole JH, Ramos MA and others (1997) Periodic migration and lowland forest refugia in a 'sedentary' neotropical bird, Wetmore's bush-tanager. *Conserv Biol* 11:692–697
- Witt MJ, Baert B, Broderick AC, Formia A and others (2009) Aerial surveying of the world's largest leatherback turtle rookery: a more effective methodology for large-scale monitoring. *Biol Conserv* 142:1719–1727

*Editorial responsibility: Clive McMahon, Darwin, Northern Territory, Australia*

*Submitted: November 4, 2009; Accepted: May 20, 2010  
Proofs received from author(s): June 29, 2010*