

available at [www.sciencedirect.com](http://www.sciencedirect.com)journal homepage: [www.elsevier.com/locate/biocon](http://www.elsevier.com/locate/biocon)

# Statistical power and sea turtle nesting beach surveys: How long and when?

Michelle Sims<sup>a,\*</sup>, Rhema Bjorkland<sup>a</sup>, Peri Mason<sup>b</sup>, Larry B. Crowder<sup>a</sup>

<sup>a</sup>Center for Marine Conservation, Nicholas School of the Environment and Earth Sciences, Duke University Marine Laboratory, 135 Duke Marine Lab Rd, Beaufort, NC 28516, USA

<sup>b</sup>Biology Department, Wesleyan University, Middletown, CT 06459, USA

## ARTICLE INFO

### Article history:

Received 14 August 2007

Received in revised form

16 July 2008

Accepted 28 July 2008

Available online 11 September 2008

### Keywords:

Hawksbill

Monte Carlo

Random effects

Poisson

Monitoring

Survey design

## ABSTRACT

Identifying long-term population trends is essential for the conservation of species of concern. With money and personnel often in short supply, the success of a monitoring program to provide an estimate of population trend requires an efficient survey design that can detect biologically important changes within some prescribed budgetary framework. In this paper we illustrate how costs can be reduced while maintaining an efficient design using a nesting beach survey of female hawksbill sea turtles (*Eretmochelys imbricata*) in the Eastern Caribbean as a specific example. We compared the statistical power of the current, intensive protocol with those of shorter duration and differing start date to generate general guidelines on sampling. Counts of individual turtles from 1987–2003 were analyzed using Poisson log-linear regression and models with both fixed and random effects describing the temporal variability were explored. We estimate that the current 20 week saturation-tagging survey design could detect a 3% per annum growth with acceptable statistical power (>0.8) in less than 14 years. However, by changing the current survey start date and counting for only 10 weeks, it is possible to detect similar trends with little loss of power while reducing monitoring costs by 25–50%. Our analyses indicate that the statistical power of surveys for hawksbills is influenced by the timing and duration of the survey because of their effect on the proportion of the annual number of nesting females that the survey captures. This simplifies estimating the power of a prospective survey program.

© 2008 Elsevier Ltd. All rights reserved.

## 1. Introduction

Ecological monitoring is becoming increasingly common to address the growing number of conservation problems around the world. Very often the objective among conservation biologists is to detect species-specific population trends, defined as the persistent change in magnitude of the population over time. Population status and trend assessment can have important social, political and economic implications. Status and trends are important consideration for listing spe-

cies under the Convention on International Trade in Endangered Species of Flora and Fauna (CITES), and for influential determinants such as the World Conservation Union's Red List assessments. Ensuring that these trends are accurately estimated is crucial, and involves careful decisions on allocation of sampling effort. For endangered or depleted species, the consequences of an incorrect decision can be serious. For example, a population may be unable to recover if there is too long a delay in detecting its declining trend. Furthermore, money and time is usually limited so conservation

\* Corresponding author. Tel.: +1 252 504 7640.

E-mail addresses: [m.sims@duke.edu](mailto:m.sims@duke.edu) (M. Sims), [rhema.bjorkland@duke.edu](mailto:rhema.bjorkland@duke.edu) (R. Bjorkland), [pmason@wesleyan.edu](mailto:pmason@wesleyan.edu) (P. Mason), [lcrowder@duke.edu](mailto:lcrowder@duke.edu) (L.B. Crowder).

0006-3207/\$ - see front matter © 2008 Elsevier Ltd. All rights reserved.

doi:10.1016/j.biocon.2008.07.021

biologists are faced with the nontrivial task of how to optimize information on trends with limited resources (Taylor and Gerrodette, 1993; Mac Nally, 1997).

Monitoring is therefore a complex undertaking and a quantitative evaluation of its effectiveness is required both at the outset of a new monitoring programme and periodically during the programme, the latter to adjust a sampling design that is failing to meet programme objectives (Gibbs et al., 1999; Legg and Nagy, 2006). Sampling design, in terms of timing and frequency, must be carefully chosen so that we have sufficient effort to detect important changes in abundance. That is, we do not want to undersample the population. However, it is also possible that many programmes are oversampling such that more information is collected than is needed. It is well recognized that with a short supply of money but more resources to be monitored, there is now a strong incentive to prevent oversampling (Taylor and Gerrodette, 1993; Mac Nally, 1997; Loughheed et al., 1999).

To find a sampling design that balances costs and effectiveness in terms of trend detection, we consider the ability of nesting beach surveys to show trends in abundance of nesting female sea turtles. The seven species of sea turtles are a focus of conservation and management activity in many regions and a growing number of monitoring programmes have been established for more than 10 years. Sea turtles face a variety of anthropogenic threats that imperil their long-term persistence and many populations are currently listed as endangered or threatened in the US (National Research Council, 1990) and internationally (IUCN/SSC, 1995). However, acquiring information to assess and manage these turtles is logistically difficult. Complex life cycles and long-distance migrations in the marine environment result in many unobservable states (*sensu* Kendall and Nichols, 2002) throughout the life history of sea turtles and monitoring programs tend to focus on the relatively accessible adult nesting females. Nesting beach studies are therefore the cornerstone of sea turtle population monitoring and assessment. Despite tremendous advances in tracking technology (Godley and Wilson, 2008) and in-water surveys to study other life history stages, monitoring trends in nesting populations will probably continue to provide the bulk of information on population trends. Because females do not breed every year, the annual counts of nesting females are considered an index of the number of mature females.

Detecting long-term trends using counts of nesting females poses several challenges. Demographic and environmental stochasticity (e.g. fluctuating ocean conditions) may produce considerable variability in annual numbers of females nesting on a beach (Chaloupka and Limpus, 2001; Broderick et al., 2002; Saba et al., 2007). Furthermore, the number of nesting females observed on a beach during a particular survey fluctuates considerably within a single nesting season (e.g. Broderick and Godley, 1996; Chaloupka, 2001). Because the structure and magnitude of the temporal variability in counts influences the ability to detect long-term trends (Urquhart et al., 1998; Larsen et al., 2001; Sims et al., 2006), nesting beach surveys need to be designed to account for such variability. Sampling decisions include when to begin and end surveying each season and how often to sample dur-

ing a nesting season. We investigated these issues to determine the optimal survey design for detection of trends in abundance of hawksbill sea turtles (*Eretmochelys imbricata*). An optimal survey design must maximize statistical power as well as comply with logistical constraints and labor costs associated with the survey. Statistical power is defined here as the probability that a survey will detect a long-term trend of given magnitude in nesting females per season despite additional temporal variability (noise) in counts. We used data from a long-term tagging study to assess whether the timing and frequency of sampling can be adjusted to reduce monitoring costs while still being able to detect biologically relevant trends. Our goal is to further encourage the incorporation of statistical power in the design and conduct of sea turtle monitoring programs. While statistical power of population trends has been explored in foraging stocks (Bjorndal and Bolten, 2000), to date no power analysis has been undertaken to determine what length and timing of nesting beach surveys are adequate to detect long-term trends in sea turtle populations.

The data for our analysis were collected by the Jumby Bay Hawksbill Project, a study conducted at Pasture Bay Beach on Long Island, Antigua. A saturation-tagging program has been ongoing since 1987 as a partnership between the Wider Caribbean Sea Turtle Conservation Network (WIDECAST) and the island's owners. Since 1987, the project has surveyed the 450-m Pasture Bay beach every night for at least 153 nights each year. With a monitoring time that encompasses most of the nesting season and detection of each emerging nester a near certainty (a detection probability of 1), the long time series of data provides an ideal opportunity to explore temporal variability in counts. We compare the statistical power of the current, intensive protocol with those of surveys of shorter duration and differing start date to generate general guidelines on sampling.

We begin by establishing the magnitude of trends we wish to detect from a beach survey, and the method of determining the trend. We then conduct a power analysis to compare the performance of the sampling designs of interest. Many software packages have been developed to calculate the power of statistical tests for normally distributed data (e.g. a simple linear regression). Here, the data are counts and a natural model formulation to study trends is a Poisson log-linear regression model. However, we know of no widely available software for performing a power analysis for a Poisson regression, and thus rely on Monte Carlo simulations to estimate power. The initial step in this approach involves using the count data from Jumby Bay to define a model that adequately describes the temporal variability in counts. Given that few previous studies have examined the temporal variability in beach counts, we begin by considering year-to-year, week-to-week, and year-week sources of temporal variability and use a Poisson log-linear regression model with fixed and random effects as a starting point for identifying the appropriate model. An assessment of labor costs and logistical constraints are then integrated with results from the power analysis, enabling an evaluation of current efforts at Jumby Bay. Finally we generalize results to give recommendations for establishing nesting beach surveys for hawksbills in other locations.

**2. Methods**

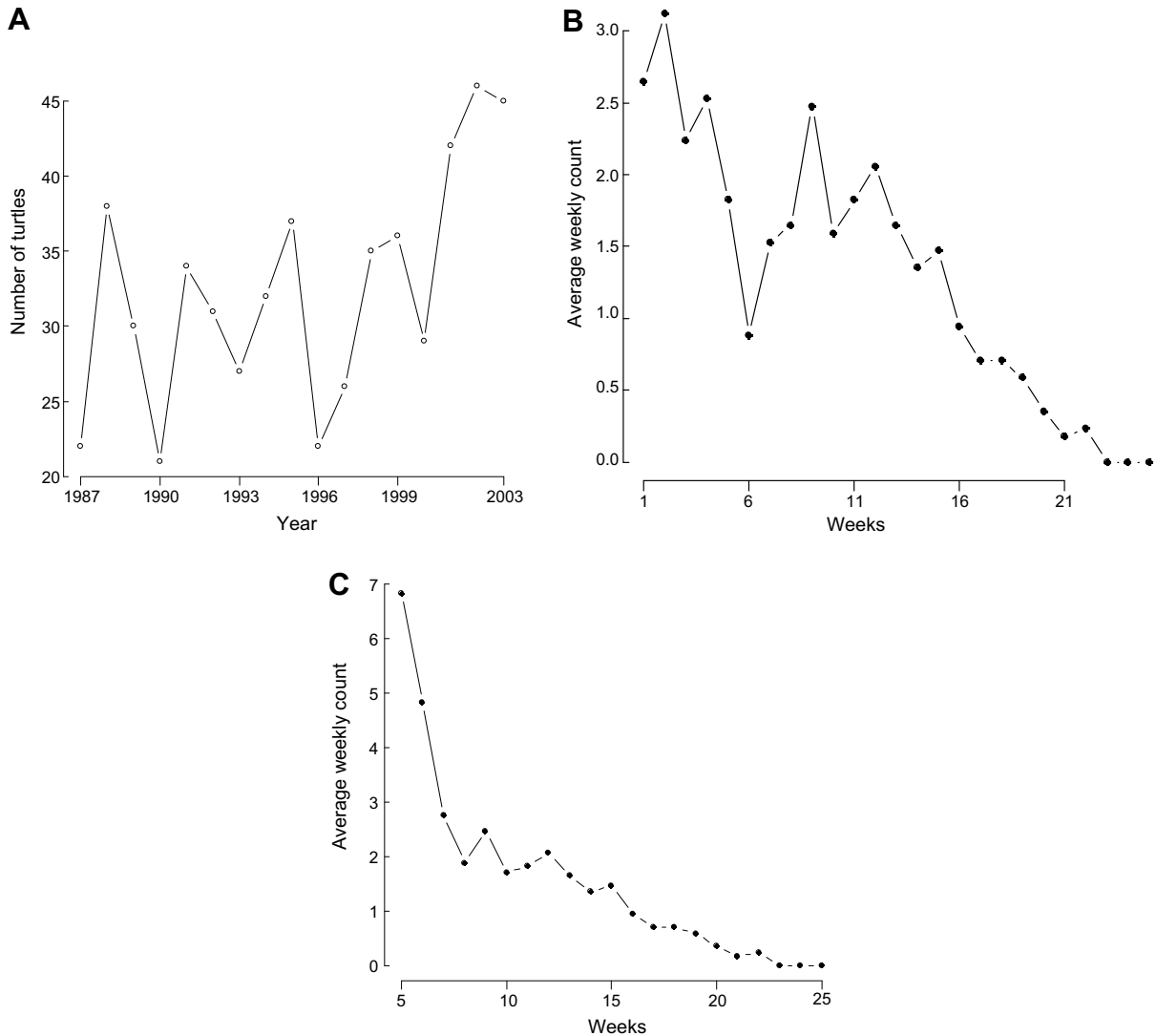
**2.1. Study site**

The Pasture Bay beach on Long Island, Antigua (17°8'N, 61°W) is monitored for at least 153 consecutive nights from June 15 through November 15, although in some years monitoring continued for a further 22 nights up to 6th December. Every nesting female during this survey window is encountered (usually multiple times in a season) and is flipper-tagged and marked with a drill pattern in the supracaudal scutes (for methods, see Hoyle and Richardson, 1993; Richardson et al., 1999, 2006). This is a reliable method for tagging turtles. Richardson et al. (1999) estimated a 1% chance of both tags being lost between two consecutive seasons. In Richardson et al. (2006), all but one of 183 remigrants was identified positively. Annual counts of nesting females fluctuated between 20 and 40 from 1987–1999 and have subsequently increased steadily to 62 by 2006.

**2.2. Objective, data and survey designs**

The specific objective of this power analysis was to determine the sampling requirements necessary to detect long-term changes in abundance of nesting females with 80% power, a level commonly considered in ecological studies (e.g. Steidl et al., 1997). The magnitude of changes that we considered were annual increases of 3%, equivalent to a 50% increase over 14 years, and 10%, the current growth rate of the neophyte subpopulation (Richardson et al., 2006). Reported population trends of sea turtles vary considerably. Where populations have been recovering, trends have ranged up to 16% per annum (e.g. Dutton et al., 2005; Beggs et al., 2007; Marcovaldi and Chaloupka, 2007). Our use of 3% represents the lower end of reported trends (e.g. Lauret-Stepler, 2007).

We used counts from 1987 to 2003 to identify a model that adequately describes the temporal variability at Jumby Bay (Fig. 1A). Although a female may be encountered at the beach on multiple occasions in any given year, for purposes of



**Fig. 1 – (A) Number of nesting females at Pasture Bay beach from 1987 to 2003. Weekly counts (averaged across years) of newly encountered females when counting begins on June 15th (B) and five weeks later (C).**

estimating trends a female is considered a single count. Consequently, in each nesting season only data on first encounters are used in this analysis and subsequent encounters are ignored. Data used in constructing the model consisted of weekly counts of first encounters over 17 survey years.

The purpose of this model in subsequent Monte Carlo simulations is to generate count data based on a proposed sampling design. The number of turtles encountered for the first time in each week will vary depending on the start date of the proposed design and this must be reflected in the generated data. To address this issue, we constructed a data set for each proposed start date using the original survey data and fitted a separate model to each. The dependence of the number of first encounters each week on start date is illustrated in Fig. 1 with the original survey data. For example, at the current start date of June 15th, on average three turtles are encountered in the first week of the survey, declining gradually as the season progresses (Fig. 1B). In contrast, if we assumed the survey began five weeks later and only consider counts from this week onwards, then the average number of new encounters in the first week of monitoring is nearly seven and the decline in counts is steeper (Fig. 1C).

Power was calculated for surveys repeated annually for 5–25 years. Three aspects of the survey design were investigated: the start date, the number of weeks monitored and the frequency of the surveys during a nesting season. We compared surveys with the current start date (15th June, WK1) to patrols beginning five weeks later into the season (WK5) for three and 10 weeks of monitoring time. We also investigated how the frequency of the counts, either on successive or alternate weeks, would influence the trend detection capabilities of a survey.

### 2.3. Models

Count data were analyzed using Poisson log-linear regression models. Specifically, we assume that  $y_{i,j} \sim \text{Poisson}(\lambda_{i,j})$ , where  $y_{i,j}$  represents the count in week  $i$  ( $i = k, \dots, 25$  and  $k = 1$  or 5 depending on the survey start date) for year  $j$  ( $j = 1, \dots, 17$ ). The natural logarithm of the expected count  $\lambda_{i,j}$  is related to a set of explanatory variables that capture the temporal variability in counts across weeks and years.

We explored three sources of potential temporal variation in counts: year-to-year, week-to-week, and year-week variation, and investigated which, if any, contributes substantially to the variation in counts.

- (1) *Year-to-year variation.* There may be environmental perturbations causing good and bad years for nesting. We partitioned this between-year variation in counts into a long-term trend and a random effect component that characterizes the random deviations from this trend. The latter captures the inter-annual environmental stochasticity, essentially all unexplained effects associated with year but not measured which cause year-to-year fluctuations in the number of nesting females. A continuous fixed effect for year, either a first or second order polynomial, was considered to capture the long-term upward trend in annual counts.

- (2) *Week-to-week variation.* In each year, the  $i$ th week of the study has the same positive or negative effect on the rate of first encounters. The magnitude of this consistent seasonal effect was captured by partitioning the between-week variation into a long-term trend and a random effect component that characterizes the random deviations from this trend, the latter caused by unrecorded stochastic week effects. A continuous fixed effect for week, either a second or third order polynomial, was considered to describe the observed long-term decline in first encounters during the nesting season.
- (3) *Year-week variation.* This represents week-to-week variation in the number of first encounters that is year specific. Factors contributing to this source of variation include observation errors or local environmental perturbations in time that have positive or negative effects on weekly counts. This final potential source of temporal variability was quantified by fitting a year-week interaction as a random effect.

Each random effect was assumed to have a normal distribution with an unknown variance, quantifying the unexplained variability attributable to that random effect, and a mean zero. If the variance of each random effect is zero, then the variation among counts is explained by Poisson variation and the fixed effect polynomial terms. If the variance associated with the year random effect is greater than zero, this indicates that there is extra between-year variation that is not explained by the polynomial terms for year or by Poisson variation. If the variance for the week random effect is greater than zero, then this indicates that there is additional week-to-week variation that is not accounted for by the polynomial terms for week or by Poisson variation.

We fitted the models using Markov Chain Monte Carlo (MCMC) simulations in WinBUGS, a Bayesian statistical modeling package (Spiegelhalter et al., 2003). Week and year covariates were centered around the mean to improve performance of the MCMC simulations. Regression coefficients of the fixed effects were given non-informative Normal prior distributions with a mean of zero and variance of 1000. An inverse-gamma (0.001, 0.001) prior was specified for the variance of each random effect (changing the prior distribution of the variance to a Uniform(0, 100) had little impact on results). Convergence of the algorithm was checked using Brooks and Gelman's modification of the Gelman–Rubin convergence statistic (Brooks and Gelman, 1998). We ran three MCMC chains each starting with different initial values that were overdispersed with respect to the posterior distribution. For each variable of interest, the test compares the width of the central 80% interval of the pooled runs to the average width of the 80% interval within the individual runs. If converged, both should have stabilized and their ratio should be close to 1. Convergence was reached after 1000 iterations and inference was based on 25,000 subsequent iterations.

Variable selection was performed by a forward stepwise approach. Models were compared using the deviance information criterion (DIC; Spiegelhalter et al., 2002). The Bayesian DIC is analogous to Akaike's information criterion (AIC) in that the smaller the value of the DIC, the better the model

(in terms of complexity and fit). The ‘null’ model, a model that includes an overall constant only, was taken as the starting model. At each step of the stepwise selection, a fixed effect was added to the model and the model retained was that with the lowest DIC. With the final fixed effect model, random effects were selected following the same approach.

To assess the adequacy of the best fitting models in describing the count variability, we tested the performance of the model on new data. The new data were derived by randomly assigning the observed counts to two groups, a training group comprising two thirds of the original data and a test group that contained the remaining third of the counts. The best fitting models were fitted to the training data and used to predict the test counts. The predicted number of each count value was calculated as

$$N_m = \sum_{i=1}^N \hat{P}(y_i = m), \tag{1}$$

where  $N_m$  is the predicted number of counts equal to  $m$  (where  $m = 0, 1, 2, \dots$ ) in the test data,  $y_i$  is the  $i$ th observed test count and  $N$  is the number of observations in the test data set. The probability that the  $i$ th test count is equal to  $m$ ,  $P(y_i = m)$ , is derived from a Poisson distribution with a mean predicted by the final model fitted to the training data and using explanatory variables associated with the  $i$ th test count. If the model is a good predictor of female nesting patterns, the predicted test counts should be comparable to the observed test counts.

### 2.4. Power calculation

We used Monte Carlo simulations to calculate power. At each simulation, this involved generating count data from the final model given the design of the proposed survey and the trend to be detected. The slope of annual numbers of nesting females versus year was then estimated using a Poisson log-linear model and was determined to be significantly different from zero using the z-test at a 5% significance level. This process of generating counts and testing for a trend was repeated for 5000 simulations and power was estimated as the proportion of slopes that were different from zero.

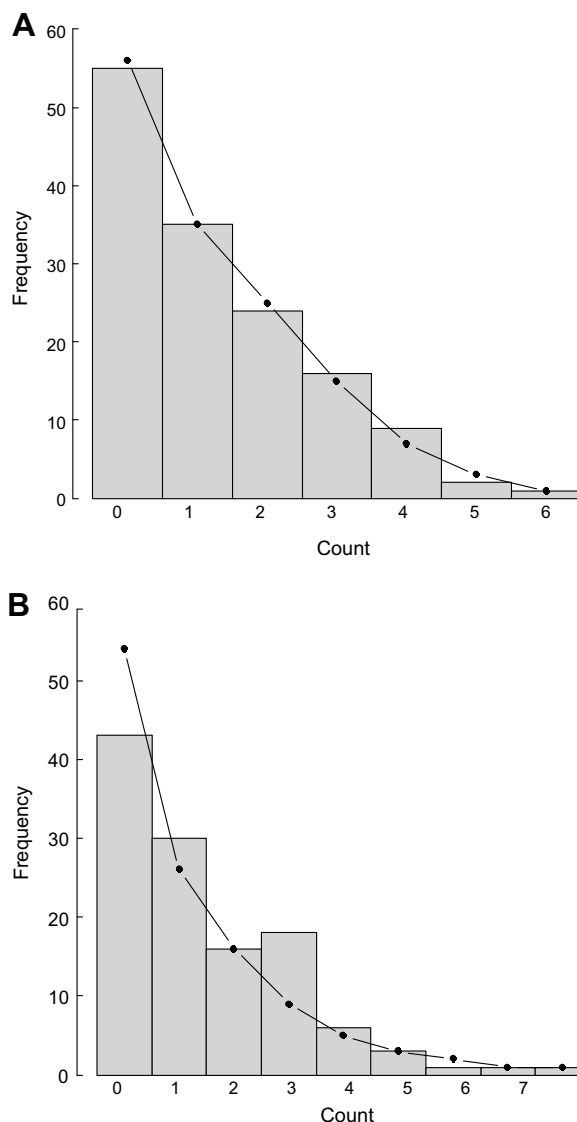
## 3. Results

### 3.1. Model fitting

A Poisson regression model with a continuous cubic fixed effect for week, a continuous quadratic fixed effect for year, and a random effect for week had the lowest DIC value and thus provided the best fit to the WK1 data. Specifically, the final model for designs starting on 15th June (Model 1) takes the form

$$\log(\lambda_{ij}) = \beta_0 + w_i + \beta_1 \text{week}_i + \beta_2 \text{week}_i^2 + \beta_3 \text{week}_i^3 + \beta_4 \text{year}_j + \beta_5 \text{year}_j^2, \tag{2}$$

where  $\beta_0$  is the intercept (the average count in week 13) and  $w_i$  is a week random effect assumed to be independently drawn from a normal distribution with a mean of zero and a variance  $\sigma_w^2$ . A Poisson regression model with only fixed effects



**Fig. 2 – Performance of the best fitting models using test data. The bar plots show the observed test counts for (A) surveys beginning 15th June (WK1) and (B) surveys starting five weeks later (WK5). The solid line represents the model predicted counts.**

in Eq. (1) was the best fit to the WK5 data (Model 2). Both models performed relatively well in predicting the distribution of counts in the test data set (Fig. 2).

### 3.2. Power calculations

Sample counts were generated from the final models with the following adjustments: the  $\beta_5 \text{year}_j^2$  term was dropped,  $\beta_1, \beta_2, \beta_3$  and  $\sigma_w^2$  were replaced with their posterior median estimates from the final model and  $\beta_4 = \log(1 + R/100)$ , where  $R$  is the percentage rate of change per year (either 3% or 10%). An annual  $R\%$  change in the total number of nesting females implies that there is also an annual  $R\%$  change in the counts each week ( $\beta_4$ ). We centered the year covariate so that the starting year was assigned the value zero. Therefore  $\beta_0$  in

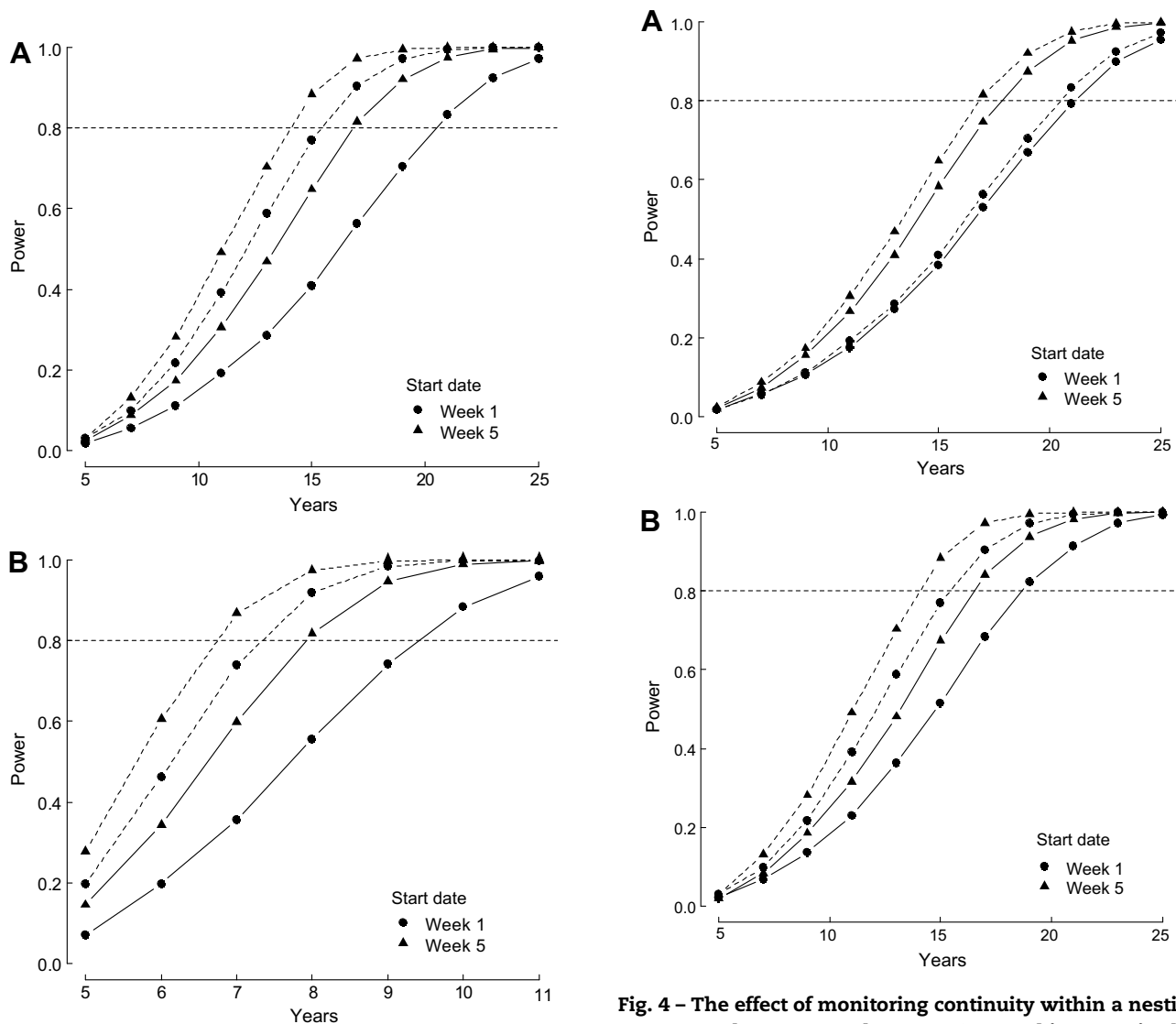
our model for generating counts represents the log of the mean count in week 13 at the start of the survey. We set the corresponding expected mean count  $E(\lambda_{13,1})$  equal to 1.647, the count in week 13 averaged over the 17 monitored years 1987–2003. Thus,  $\beta_0 = \log(E(\lambda_{13,1}))$  for WK5 surveys, while, due to the random effect term in Eq. (2),  $\beta_0 = \log(E(\lambda_{13,1})) - \sigma_w^2/2$  for WK1 surveys (see Appendix A for more details).

### 3.3. Comparison of survey designs

We estimated that trends in rapidly growing sea turtle populations (10% per year) can be detected with acceptable power within 7 years if intensively monitored. For more moderate growth rates (3% per year), 14–21 years will be required.

The number of weeks monitored and the survey start date influenced the power to detect a 3% annual increase in nesting females (Fig. 3A). At the current survey start date of June 15th, it would take 21 years to detect a trend if the beach was monitored every night for three weeks and 16 years if monitoring continued for a further seven weeks. The time to detect this size trend would decrease further to 14 years if the 10 weeks of monitoring started five weeks later. The different effects of the proposed survey designs on power were also evident for an annual trend of 10%, although the number of years to detect the trend was considerably less (Fig. 3B). For example, the optimal sampling design of starting in week five and monitoring for 10 weeks would detect a 10% annual increase in less than 7 years.

Given a fixed number of monitored weeks, the power to detect a 3% annual increase was higher when surveys were



**Fig. 3** – The effect of start date and number of successive weeks monitored on the power to detect a (A) 3% and (B) 10% annual increase in the number of nesting females. Monitoring conducted for 3 weeks (solid line) or 10 weeks (dashed line).

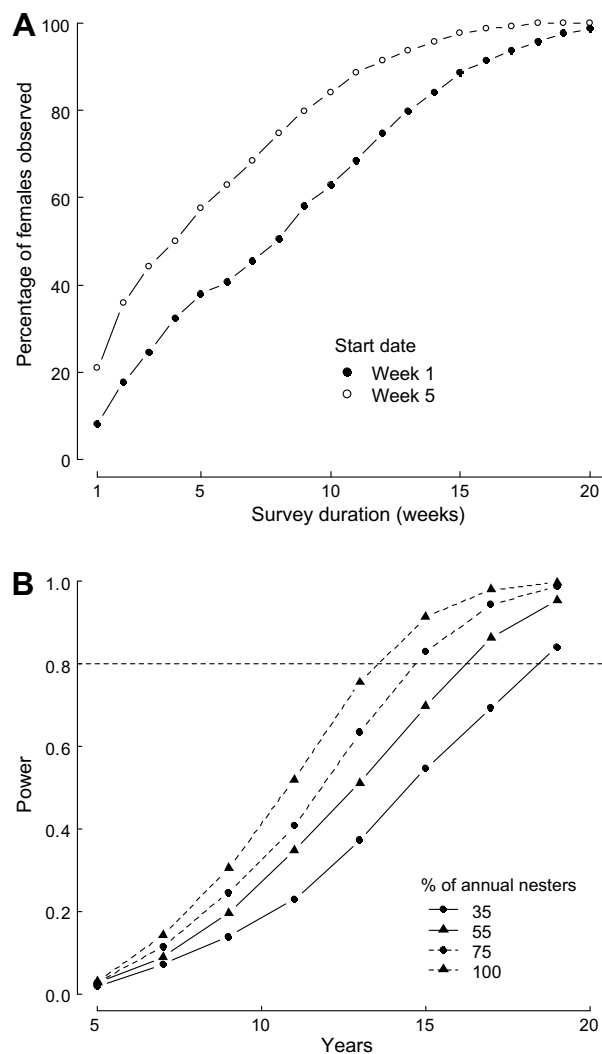
**Fig. 4** – The effect of monitoring continuity within a nesting season on the power to detect a 3% annual increase in the number of nesting females. (A) A comparison of designs that count for three weeks. (B) A comparison of designs that count within a 10 week window. Monitoring conducted on successive weeks (dashed line) or on alternate weeks (solid line).

conducted on successive rather than alternate weeks. However, the differences in powers were small (less than one year) when only three weeks were monitored (Fig. 4A). When a fixed time window is available to conduct surveys, power is higher for surveys counting in all weeks within that window. For example, when a window of 10 weeks is available, a trend can be detected three years faster when all 10 weeks are counted rather than alternate weeks (Fig. 4B).

### 3.4. Factors influencing power

A large variance for the year random effect, capturing stochastic between-year variation, reduces the power to detect trends (Larsen et al., 2001; Sims et al., 2006). However, the final models did not include a random effect for year or a year-week interaction, indicating that there was no significant evidence of these sources of variability in the counts of nesting females at Pasture Bay beach. Since trends were estimated using a Poisson log-linear model of annual counts against year, statistical power was affected by the proportion of nesting turtles that were counted. Power would be higher for surveys that capture a larger proportion of the annual nesters. Two components of Eq. (2) contribute to the proportion of turtles counted within a nesting season: the trend in number of first encounters per week, captured by the cubic polynomial for week, and the magnitude of intercept  $\beta_0$ . The within-season trend in numbers of first encounters per week influences the proportion of nesters counted with a particular survey design. We use count data from Pasture Bay to compare the proportion of annual nesters (averaged across years) that were observed in surveys of varying duration and start date (Fig. 5A). Proportions are higher for surveys that begin five weeks later in the season (WK5) and are of longer duration. The magnitude of  $\beta_0$  is influenced by the variance component for the week random effect  $\sigma_w^2$ . For example, larger values for  $\sigma_w^2$  reduce the magnitude of  $\beta_0$  (Eq. (A.1) in Appendix A) and decrease the power of a survey.

Further investigations of the influence of  $\sigma_w^2$  and within-season trend on power revealed that although it is statistically significant in WK1 surveys, the week random effect is sufficiently small that it has a negligible effect on power. This contributed to our second finding that survey designs capturing an identical proportion of the total nesters each season, but with different durations and start dates, had comparable powers. For example, a comparison of a 13 week survey beginning in week 1 and a 9 week survey beginning in week 5, both of which capture approximately 80% of the total number of nesting females, showed similar power to detect a 3% annual trend (Table 1). We compared these estimates with an alternate model in which power was estimated using data simulated from a Poisson generalized linear model of annual counts (Model 3). The mean annual count was related to a continuous explanatory variable for year using a log link function. Year was centered on the first year of the survey, and the regression coefficient for year corresponded to the trend of interest. The intercept in this model thus represents the log number of females counted at the beginning of the survey. In our example, we computed 80% of the average annual count at Pasture Bay and set the intercept equal to the log



**Fig. 5 – (A) Proportion of annual nesters (averaged across years) observed on Pasture Bay beach with surveys of varying duration and start date. (B) Power to detect a 3% annual increase using survey designs that count between 35% and 100% of the nesters each year.**

**Table 1 – The effect of survey coverage on the power to detect a 3% annual trend**

Number of surveyed years	Simulation model		
	Model 1	Model 2	Model 3
13	0.690	0.654	0.655
15	0.866	0.843	0.846
17	0.964	0.946	0.952

A comparison of powers estimated by a 13 week survey beginning in week 1 and a 9 week survey starting in week 5 using the final models in the Monte Carlo simulations (Model 1 and Model 2, respectively) and an alternative approach using simulations generated from a simple Poisson generalized linear model of annual counts (Model 3). All surveys capture 80% of the annual nesters.

of this value. Power calculated using this simple Poisson regression model gave similar estimates to the other models (Table 1).

## 4. Discussion

Monitoring is a tool for conservation of species of concern. A prerequisite for good monitoring is having an efficient sampling design, in terms of being able to detect biologically significant trends. Often faced with a limited budget, conservation biologists must furthermore avoid the cost of oversampling the population. In this paper, we provide a method of assessing what sampling effort is sufficient to detect changes in species abundance using the hawksbill sea turtles as a specific example. Beach surveys of nesting females are crucial in measuring the effects of management activities on population trends for all species of sea turtles. For sea turtles in the Caribbean, the location of the particular example in this paper, the need for long-term monitoring programs to provide the requisite data has been recognized by Caribbean fisheries managers since the first Western Atlantic Turtle Symposium in 1983 (Bacon et al., 1984) and subsequently reaffirmed in contributions to Bjorkland (2001) and the CITES-sponsored “Wider Caribbean Region Hawksbill Turtle Dialogue Meetings” (Mexico City, May 2001; Cayman Islands, May 2002. <http://www.cites.org/eng/prog/HBT/intro.shtml> (last accessed August 2007)). Such monitoring programs require an ongoing commitment of resources, and research efforts to improve the rigor and scope of the outputs may increase confidence in the results and the likelihood that the results will be incorporated into policies and conservation plans.

In many monitored populations, trends are blurred by large fluctuations in counts resulting from observer counting errors and environmental stochasticity. An illustration of this problem is provided by Sims et al. (2006) who investigated monitoring designs for common guillemots (*Uria aalge*), a colonial seabird which raises offspring in dense aggregations on cliff faces. Count variability can be high when monitoring is conducted during the breeding season because dense aggregations induce counting errors, while bad weather conditions and the presence of predators reduce attendance time at a nest. In contrast, counting errors of nesting females at Pasture Bay are rare and the year-to-year variation in counts was attributable to only Poisson variation about a long-term trend.

Our finding that there was no additional inter-annual variability in nesting numbers is interesting. Hays (2000) and Solow et al. (2002) established that the overall variability in the numbers of females nesting in a given year is related to the variability in remigration intervals (the time between successive nesting years), and this in turn is related to environmental conditions on the feeding grounds. However, the variability in remigration intervals and subsequent variability in numbers of hawksbills at Jumby Bay is low and we envision this happening in two, not necessarily exclusive, ways. The first scenario follows from Hays (2000). He hypothesized that if the population from a nesting beach dispersed to multiple feeding grounds and experienced environmental conditions that varied independently and stochastically, then the inter-annual variability in numbers of nesting females would be low. In a sense, the effects of environmental stochasticity on numbers of nesting females are being “averaged out” in any given year. Further support for this scenario comes from the observed uti-

lization of multiple foraging grounds, hypothesized to have predictably abundant food resources, by post-nesting hawksbill turtles in the Caribbean (Van Dam et al. 2008).

Additionally, the effect of environmental stochasticity on remigration intervals could be expected to vary with the trophic level of the sea turtle. This second scenario derives from research conducted by Broderick et al. (2001) who demonstrated that the trophic status of the sea turtle species was a major driver of the inter-annual variability in nesting numbers. Hawksbills, and leatherbacks, for example are intermediate in trophic status between the herbivorous greens and the carnivorous loggerheads. They analyzed data from a number of nesting population studies for these four sea turtle species and found that only loggerheads had a lower (albeit non-significant) variability in annual nesting numbers than hawksbills. Whilst leatherbacks and hawksbills are at the same trophic level, the relatively slow-growing sponge species that sustain hawksbills are less likely to be affected by short-term changes in ocean productivity than jelly-fish eating leatherbacks.

The effects of short-term (either inter-annual or within-season) environmental perturbations on hawksbills were minor and subsequently simpler models are sufficient to describe the variability in hawksbills counter to other sea turtle species. We hypothesize that for other species, there may be significant stochastic between-year and year-week effects and therefore an approach similar to that presented in this study, one that identifies the sources and magnitude of the count variability using random and fixed effects, is recommended to evaluate the power of different sampling designs.

A draft population monitoring protocol for hawksbills in the Caribbean was established by the second CITES wider Caribbean hawksbill sea turtle dialogue meeting (Cayman Islands, May 2002). Recommendations were made on all aspects of the protocol including survey boundaries, surveyor training and the timing of the survey. In this paper, we conducted a thorough investigation of the latter aspect of the protocol, in particular the start date and duration of the survey. Our results indicate that the most important factor influencing power of a survey design at Pasture Bay is the proportion of all nesting females that will be encountered during the survey period. Fig. 5B shows the power of detecting a 3% annual trend for surveys capturing different proportions of the total number of annual nesters at Pasture Bay Beach. Ideally a complete nesting season should be surveyed as this captures 100% of the population. We estimate that the current survey coverage of 100% will detect a 3% annual trend in less than 14 years. When this coverage is not possible, the start date and survey duration should be selected so that the appropriate proportion of the population will be counted to achieve the survey objective. Our results indicate that by strategic placement of the survey start date, one can survey for a shorter duration while counting a comparable proportion of the population. The CITES report recommended that the survey period should, “encompass the peak of the nesting season and should be designed to allow for shifts in the peak of the nesting season from year to year”. Although we did not observe a shift in the peak from year to year at Pasture Bay, our findings strengthen the argument of conducting the survey when most females are encountered.

Monitoring programs that currently operate for an extensive proportion of the nesting season obviously have the capacity to do so. However new programs are under consideration in several territories in the region and we hope that our analyses offers tools for a scientifically sound process of designing a new monitoring program. For those wishing to establish a nesting beach survey protocol to monitor population trends, the following recommendations are offered:

- (1) Decide on the objectives of the study: the magnitude of trend to detect and with what power. Determine the start dates and survey durations that are logistically feasible and the significance level at which to test for presence of a trend.
- (2) Conduct a complete census of the beach, counting throughout each night for the entire nesting season, for at least 2–3 years to determine the seasonal trend in first encounters. This requires individual nesters to be identified. For each start date, verify that the within-season trend is similar from year to year.
- (3) Estimate the size of the nesting female population, such as the mean census count over 3 years.
- (4) Perform a simple power analysis using Monte Carlo simulations to determine what proportion of the nesters must be observed to achieve the survey objectives described in (1). This involves simulating annual counts from a Poisson generalized linear model (described earlier in Section 3 as Model 3) with the intercept set equal to the logarithm of the number of nesting females calculated in (3) multiplied by the proportion being investigated.
- (5) Construct a plot of the census data, such as Fig. 5A, to determine what designs meet this coverage requirement.

These recommendations are based on our findings that when nightly surveys are exhaustive, the observer counting errors, year to year variability and fluctuations in the within-season trend between years are insignificant. We also observed that the size of the variance component for the week random effect was sufficiently small that its affect on power was negligible. These observations should not be assumed at other monitoring sites. When assumptions do not hold, the timing of the survey or the number of years monitored would have to increase to achieve comparable statistical power. We then recommend that established nesting surveys are periodically reviewed, using methods similar to those described here, to adjust sampling designs accordingly.

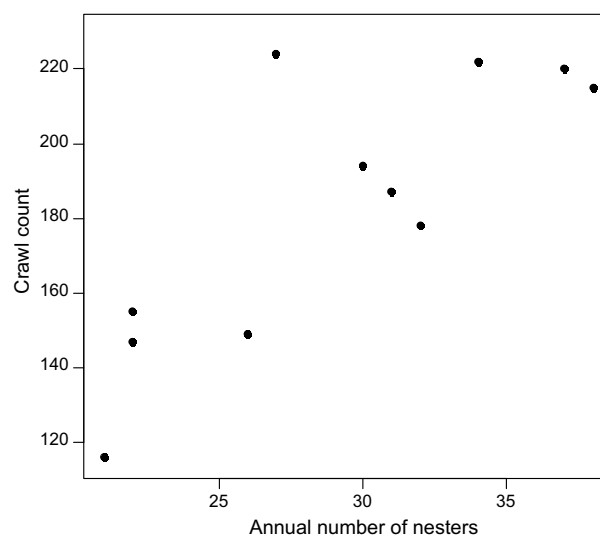
Although counting individuals requires more resources, in the Jumbo Bay study it can be considered a reliable index because the detection probability was shown to be time invariant both within and between seasons (Kendall and Bjorkland, 2001). Therefore simple unadjusted counts from a subsample of the annual nesters are reliable indices by which to measure trend. While counting crawls or nests are more parsimonious with time and resources, they may introduce additional variability because crawl to nest count ratios and nest to female count ratios are not constant across years (e.g. Foote and Sprinkel 1997; Broderick et al., 2003). Fig. 6 compares the an-

nual number of nesters to number of crawls counted in a 22 week window from June 15 to November 15 (1987–1997), illustrating the less than perfect correlation (Pearson correlation coefficient = 0.81), attributable to shore line changes and removal of vegetation due to home construction and hurricanes.

In our case study, since we had a constant detection probability we were confident that the unadjusted counts were reliable indices of total number of nesters. However, in other studies this assumption may not hold (e.g. Chaloupka and Limpus, 2001). In those cases in which detection probability varies annually, then this parameter should be estimated using methods such as capture-mark-recapture (Schmidt, 2004) and the capture-mark-recapture estimates should be used for estimating trends.

We estimate that a nightly monitoring program of 10 weeks at Pasture Bay beach could save 25–50% on annual operating costs. By starting the current survey five weeks later and counting for 10 weeks, it is possible to detect similar trends with little loss of power while making substantial savings in monitoring costs. While a power analysis is the central part of a monitoring design, there may be multiple reasons for monitoring that further complicate the design, such as timing surveys to deter poachers, whether other species are being monitored, monitoring hatch success, and relocating clutches for hatcheries/caging. Additionally, in many regions support for conservation of dwindling sea turtle stocks rests upon community education and involvement. Night time surveys provide an opportunity for direct observation and contact with sea turtles, in contrast to crawl or nest counts alone. The Jumbo Bay dataset with over 60,000 person-hours of effort is unique and can provide answers to broader questions about the ecological role of hawksbills in the wider Caribbean.

While we recognize the importance of monitoring as many beaches as possible rather than a few selected sites (McClenachan et al., 2006), there are challenges to developing a robust, cost-effective region-wide monitoring design that



**Fig. 6** – Comparison of the annual number of nesters to number of crawls counted in a 22 week window from June 15 to November 15.

require further consideration including the number and spatial distribution of the beaches. While we can estimate trends in annual nesters at Pasture Bay beach, it is not possible to extrapolate the trend at this beach to the wider Antigua-Barbuda area because the representativeness of this beach has been unexplored. If the goal is to make inferences about region-wide trends, sampling decisions must also include what fraction of suitable nesting beach to monitor in the region.

## Acknowledgments

This work was made possible by the support of the Jumby Bay Island Company, Ltd. and the dedicated teams of researchers of nearly two decades. We thank Michael Levine for his insights and advice and Dan Hall, Jeff Moore and Bryan Wallace for critically reading the manuscript. Funding for this project was provided by the Gordon and Betty Moore Foundation.

## Appendix A

### A.1. Specifying a value for $\beta_0$ in the power calculations of WK1 surveys

Based on the final model used for the Monte Carlo simulations, the expected count  $E(\lambda_{i,j})$  of nesting females for designs starting on 15th June (WK1) is specified as

$$E(\lambda_{i,j}) = E(\exp(\beta_0 + w_i + \hat{\beta}_1 \text{week}_i + \hat{\beta}_2 \text{week}_i^2 + \hat{\beta}_3 \text{week}_i^3 + \beta_4 \text{year}_j)).$$

When  $i = 13$  and  $j = 1$ , week and year fixed effects are dropped from the model since the variables were centered on these values. Thus, the expected count in week 13 at the start of the survey is

$$E(\lambda_{13,1}) = \exp(\beta_0)E(\exp(w_{13})).$$

Since  $w_i \sim N(0, \hat{\sigma}_w^2)$ ,  $\exp(w_i)$  has a lognormal distribution with mean of  $\exp(\hat{\sigma}_w^2/2)$ , leading to

$$E(\lambda_{13,1}) = \exp(\beta_0)\exp(\hat{\sigma}_w^2/2).$$

If the expected count in week 13 at the start of the monitoring program is specified, then a value for  $\beta_0$  can be obtained by

$$\beta_0 = \log(E(\lambda_{13,1})) - \hat{\sigma}_w^2/2. \quad (\text{A.1})$$

## REFERENCES

- Bacon, P., Berry, F., Bjorndal, K., Hirth, H., Ogren, L., Weber, M. (Eds.), 1984. Proceedings of the Western Atlantic Sea Turtle Symposium. University of Miami Press.
- Beggs, J.A., Horrocks, J.A., Krueger, B.H., 2007. Increase in hawksbill sea turtle *Eretmochelys imbricata* nesting in Barbados, West Indies. *Endangered Species Research* 3, 159–168.
- Bjorkland, R. (Kerr), 2001. Monitoring population trends. In: Eckert, K.L., Abreu-Grobois, F.A. (Eds.), Proceedings of the Regional Meeting 'Marine Turtle Conservation in the Wider Caribbean Region: A Dialogue for Effective Regional Management', 16–18 November 1999, Santo Domingo. WIDECAS, IUCN-MTSG, WWF and UNEP-CEP, pp. 107–110.
- Bjorndal, K.A., Bolten, A.B., 2000. In: Proceedings of a Workshop on Assessing Abundance and Trends for In-Water Sea Turtle Populations. NOAA Technical Memorandum NMFS-SEFSC-445, Miami, Florida.
- Broderick, A.C., Godley, B.J., 1996. Population and nesting ecology of the Green Turtle, *Chelonia mydas*, and the Loggerhead Turtle, *Caretta caretta*, in northern Cyprus. *Zoology in the Middle East* 13, 27–46.
- Broderick, A.C., Glen, F., Godley, B.J., Hays, G.C., 2002. Estimating the number of green and loggerhead turtles nesting annually in the Mediterranean. *Oryx* 36, 227–236.
- Broderick, A.C., Glen, F., Godley, B.J., Hays, G.C., 2003. Variation in reproductive output in marine turtles. *Journal of Experimental Marine Biology and Ecology* 288, 95–109.
- Broderick, A.C., Godley, B.J., Hays, G.C., 2001. Trophic status drives interannual variability in nesting numbers of marine turtles. *Proceedings of the Royal Society B: Biological Sciences* 268, 1481–1487.
- Brooks, S.P., Gelman, A., 1998. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* 7, 434–455.
- Chaloupka, M., 2001. Historical trends, seasonality and spatial synchrony in green sea turtle egg production. *Biological Conservation* 101, 263–279.
- Chaloupka, M., Limpus, C., 2001. Trends in the abundance of sea turtles resident in southern Great Barrier Reef waters. *Biological Conservation* 102, 235–249.
- Dutton, D.L., Dutton, P.H., Chaloupka, M., Boulon, R.H., 2005. Increase in a Caribbean leatherback turtle *Dermochelys coriacea* nesting population linked to long-term nest protection. *Biological Conservation* 126, 185–194.
- Foote, J.J., Sprinkel, J., 1997. Sea Turtle protection measures including nest/hatch and sand compactness monitoring for the city of Venice beach restoration project Sarasota County. The City of Venice. Mote Marine Laboratory Technical Report No. 500.
- Gibbs, J.P., Snell, H.L., Causton, C.E., 1999. Effective monitoring for adaptive wildlife management: lessons from the Galapagos Islands. *Journal of Wildlife Management* 63, 1055–1065.
- Godley, B.J., Wilson, R.P., 2008. Tracking vertebrates for conservation: introduction. *Endangered Species Research* 4, 1–2.
- Hays, G.C., 2000. The implications of variable remigration intervals for the assessment of population size in marine turtles. *Journal of Theoretical Biology* 206, 221–227.
- Hoyle, M., Richardson, J.L., 1993. The Jumby Bay Hawksbill Project: Survivorship, Mortality, Recruitment and Reproductive Biology and Behavior of Adult Female Hawksbill Sea Turtles (*Eretmochelys imbricata*) Nesting at Pasture Bay, Long Island, Antigua. Georgia Sea Turtle Cooperative, Institute of Ecology, University of Georgia, 76 pp.
- IUCN/SSC Marine Turtle Specialist Group, 1995. A Global Strategy for the Conservation of Marine Turtles. IUCN, Gland, Switzerland, 25 pp.
- Kendall, W.L., Bjorkland, R., 2001. Using open robust design models to estimate temporary emigration from capture–recapture data. *Biometrics* 57, 1113–1122.
- Kendall, W.L., Nichols, J.D., 2002. Estimating state-transition probabilities for unobservable states using capture–recapture/resighting data. *Ecology* 83, 3276–3284.
- Larsen, D.P., Kincaid, T.M., Jacobs, S.E., Urquhart, N.S., 2001. Designs for evaluating local and regional scale trends. *Bioscience* 51, 1069–1078.
- Lauret-Stepler, M., Bourjea, J., Roos, D., Pelletier, D., Ryan, P.G., Ciccione, S., Grizel, H., 2007. Reproductive seasonality and trend of *Chelonia mydas* in the SW Indian Ocean: a 20 yr study based on track counts. *Endangered Species Research* 3, 217–227.
- Legg, C.J., Nagy, L., 2006. Why most conservation is, but need not be, a waste of time. *Journal of Environmental Management* 78, 194–199.

- Lougheed, L.W., Breault, A., Lank, D.B., 1999. Estimating statistical power to evaluate ongoing waterfowl population monitoring. *Journal of Wildlife Management* 63, 1359–1369.
- Mac Nally, R., 1997. Monitoring forest bird communities for impact assessment: the influence of sampling intensity and spatial scale. *Biological Conservation* 82, 355–367.
- Marcovaldi, M.A., Chaloupka, M., 2007. Conservation status of the loggerhead sea turtle in Brazil: an encouraging outlook. *Endangered Species Research* 3, 133–143.
- McClenachan, L., Jackson, J.B.B., Newman, M.J.H., 2006. Conservation implications of historic sea turtle nesting beach loss. *Frontiers in Ecology and the Environment* 4, 290–296.
- National Research Council, 1990. *Decline of the Sea Turtles. Causes and Prevention*. National Academy Press, Washington, DC, USA.
- Richardson, J.I., Bell, R., Richardson, T.H., 1999. Population ecology and demographic implications drawn from an 11-year study of nesting hawksbill turtles, *Eretmochelys imbricata*, at Jumby Bay, Long Island, Antigua, West Indies. *Chelonian Conservation and Biology* 3, 244–250.
- Richardson, J.I., Hall, D.B., Mason, P.A., Andrews, K.M., Bjorkland, R., Cai, Y., Bell, R., 2006. Eighteen years of saturation tagging data reveal a significant increase in nesting hawksbill sea turtles (*Eretmochelys imbricata*) on Long Island Antigua. *Animal Conservation* 9, 302–307.
- Saba, V.S., Santidrian-Tomillo, P., Reina, R.D., Spotila, J.R., Musick, J.A., Evans, D.A., Paladino, F.V., 2007. The effect of the El Niño Southern Oscillation on the reproductive frequency of eastern Pacific leatherback turtles. *Journal of Applied Ecology* 44, 395–404.
- Schmidt, B.R., 2004. Declining amphibian populations: the pitfalls of count data in the study of diversity, distributions, dynamics and demography. *Herpetological Journal* 14, 167–174.
- Sims, M., Wanless, S., Harris, M.P., Mitchell, P.I., Elston, D.A., 2006. Evaluating the power of monitoring plot designs for detecting long-term trends in the numbers of common guillemots. *Journal of Applied Ecology* 43, 537–546.
- Solow, A.R., Bjørndal, K.A., Bolten, A.B., 2002. Annual variation in nesting numbers of marine turtles: the effect of sea surface temperature on re-migration intervals. *Ecology Letters* 5, 742–746.
- Spiegelhalter, D.J., Best, N., Carlin, B., van der Linde, A., 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B* 64, 583–639.
- Spiegelhalter, D.J., Thomas, A., Best, N., Lunn, D., 2003. *WinBUGS User Manual*. Version 1.4.
- Steidl, R.J., Hayes, J.P., Schaubert, E., 1997. Statistical power analysis in wildlife research. *Journal of Wildlife Management* 61, 270–279.
- Taylor, B.L., Gerrodette, T., 1993. The uses of statistical power in conservation biology: the vaquita and Northern Spotted Owl. *Conservation Biology* 7, 488–500.
- Urquhart, N.S., Paulsen, S.G., Larsen, D.P., 1998. Monitoring for policy-relevant regional trends over time. *Ecological Applications* 8, 246–257.
- Van Dam, R.P., Diez, C.E., Balazs, G.H., Colón Colón, L.A., McMillan, W.O., Schroeder, B., 2008. Sex-specific migration patterns of hawksbill turtles breeding at Mona Island, Puerto Rico. *Endangered Species Research* 4, 85–94.