Fathoming sea turtles: monitoring strategy evaluation to improve conservation status assessments

Susan E. Piacenza [ID],[1,3,4] Paul M. Richards,[2] and Selina S. Heppell[1]

[1]*Department of Fisheries and Wildlife, Oregon State University, Corvallis, Oregon 97330 USA*
[2]*NOAA NMFS, Southeast Fisheries Science Center, Miami, Florida 33149 USA*

Abstract. Population monitoring must be accurate and reliable to correctly classify population status. For sea turtles, nesting beach surveys are often the only population-level surveys that are accessible. However, process and observation errors, compounded by delayed maturity, obscure the relationship between trends on the nesting beach and the population. We present a simulation-based tool, monitoring strategy evaluation (MoSE), to test the relationships between monitoring data and assessment accuracy, using green sea turtles, *Chelonia mydas*, as a case study. To explore this first application of MoSE, we apply different treatments of population impacts to virtual true populations, and sample the nests or nesters, with observation error, to test if the observation data can be used to diagnose population status accurately. Based on the observed data, we examine population trend and compare it to the known values from the operating model. We ran a series of scenarios including harvest impacts, cyclical breeding probability, and sampling biases, to see how these factors impact accuracy in estimating population trend. We explored the necessary duration of monitoring for accurate trend estimation and the probability of a false trend diagnosis. Our results suggest that disturbance type and severity can have important and persistent effects on the accuracy of population assessments drawn from monitoring nesting beaches. The underlying population phase, age classes disturbed, and impact severity influenced the accuracy of estimating population trend. At least 10 yr of monitoring data is necessary to estimate population trend accurately, and >20 yr if juvenile age classes were disturbed and the population is recovering. In general, there is a greater probability of making a false positive trend diagnosis than a false negative, but this depends on impact type and severity, population phase, and sampling duration. Improving detection rates to 90% does little to lower probability of a false trend diagnosis with shorter monitoring spans. Altogether, monitoring strategies for specific populations may be tailored based on the impact history, population phase, and environmental drivers. The MoSE is an important framework for analysis through simulation that can comprehensively test population assessments for accuracy and inform policy recommendations regarding the best monitoring strategies.

Key words: agent-based model; Chelonia mydas; *demography; fisheries impacts; Hawaii; management strategy evaluation; simulation; transitory dynamics.*

Introduction

Population monitoring must be accurate and reliable for biologists and conservation managers to correctly classify populations as endangered, in recovery, or not endangered. In addition, monitoring data are important indicators of whether management actions are effective, but the data must be reliable. Many endangered species are often considered data poor or have low encounter rates with monitoring programs that may obscure true population trends (Colyvan et al. 1999, Akçakaya et al. 2000). In long-lived, migratory species, where monitoring can only occur on particular demographic classes for short periods of time, monitoring may only give a narrow view into a population, and indices may give a false signal of population trend, especially during unstable periods (Maxwell and Jennings 2005, Taylor et al. 2007, Singh and Milner-Gulland 2011, Lynch et al. 2012). If monitoring yields inaccurate data and the subsequent population assessments make false interpretations of population size and trends, conservation errors may ensue. There are two main kinds of conservation errors: to conclude a population is threatened when in fact it is not, and to conclude a population is not threatened

[3] Present address: Department of Biology, University of West Florida, Pensacola, Florida 32514 USA.
[4] E-mail: susiepiacenza@gmail.com

when in fact it is; both kinds of error have biological, economic, and societal consequences (Taylor and Gerrodette 1993, Snover and Heppell 2009). Indeed, there is a third case where substantial uncertainty in population status results in no determination of extinction threat (e.g., Red List Standards and Petitions Subcommittee 1996, Turtle Expert Working Group 2009). In all, biologists and managers need to exercise caution when interpreting population indices from monitoring, particularly when those interpretations have strong management implications. Developing new analytical tools may help to clarify relationships between observed and true population status.

Sea turtles present such a case where life-history complicates monitoring. Sea turtles are long-lived (>50 yr), late-maturing, highly migratory (traveling through entire ocean basins across life stages), and spend most of their lives offshore (Bowen et al. 1992, Godley et al. 2002, Zug et al. 2002, Wallace et al. 2010, Van Houtan et al. 2014, Casale and Heppell 2016). With late maturity (e.g., age at maturity for green sea turtles, *Chelonia mydas*, is estimated to range 17–50 yr) comes temporal lags in recovery, and the length of those time lags depends on the age classes disturbed and how conservation benefits survival of those age classes (Zug et al. 2002, Van Houtan et al. 2014). The duration of the time lags ultimately may have important implications for monitoring and assessment (Crowder et al. 1994, Heppell et al. 1996, Koons et al. 2005, White et al. 2013). Most sea turtle monitoring is conducted on nesting beaches, where nests are typically counted, or more rarely individual female nesters are counted usually via saturation tagging studies, and occasionally other demographic data are collected (nest survival, egg survival, etc.; Schroeder and Murphy 1999, National Research Council 2010). However, as females do not breed annually and may be decades old at first nesting, just a tiny fraction of the total population is monitored (Crouse et al. 1987). Nesting abundance typically displays large fluctuations interannually, likely due to variability in breeding frequency and environmental conditions, but these fluctuations do not reflect true changes in the adult population (Hays 2000, Solow et al. 2002, Piacenza et al. 2016).

It is uncertain how accurate the extrapolations from nesting beach indices are for estimating population status (i.e., population trend or abundance; Hays 2000, National Research Council 2010, Richards et al. 2011, Warden et al. 2017). For example, Richards et al. (2011) developed a method to estimate population size of loggerheads (*Caretta caretta*) in the North Atlantic by creating distributions of subpopulations by resampling nest counts over a sampling period (in this case, 10 years). The method allowed for identifying key subpopulations for conservation efforts. However, the authors were not able to evaluate trends in annual population size because breeding interval and clutch frequency were not measured annually. Considering that most sea turtle populations are in demographic flux (Chaloupka et al. 2008,

Wallace et al. 2011, IUCN 2015), and transient behaviors in structured populations can have opposing responses for some demographic classes during the impact and recovery phases, it seems likely that population assessments are inaccurate when based on monitoring from beach surveys (Crowder et al. 1994, Hastings 2001, 2004, Koons et al. 2005, White et al. 2013). True population dynamics may be further obscured when monitoring only observes reproductive classes, that is nesters and nests.

Beach surveys, however, are often the only way biologists can encounter sea turtles to measure abundance and population trend, as in-water surveys can be cost prohibitive and often have very low encounter rates. Can we optimize monitoring on the nesting beach to give the most accurate data on population status over time? In a report of the National Research Council (2010), the authors recommended a tiered approach to estimating nesting female abundance. The tiers start with nest counts on beaches spanning a spectrum of data scope and monitoring, then nest counts in representative locations, then saturation tagging, and so on with increasing variables and sampling complexity. However, are these recommendations too expensive and too time-consuming for government agencies, academics, and non-profit monitoring groups to implement? Given the effort and time spans involved in monitoring sea turtles, prognostic evaluation of these monitoring options, such as how long to monitor a beach to estimate population trend, goals for detection rates, which types of sampling bias to avoid, is important so that research groups can decide in advance how to optimize monitoring efforts on nesting beaches (Heppell and Crowder 1998).

To address these issues, we developed a tool to explore the effects of different kinds of monitoring data (i.e., nesters or nests), and their realistic uncertainties, on population response predictions: monitoring strategy evaluation (MoSE). We based this tool on management strategy evaluation (MSE), a simulation-based framework developed in fisheries science (Smith et al. 1999). MSE was developed to evaluate trade-offs in alternate management schemes and to assess the consequences of uncertainty for achieving management goals (Punt et al. 2014). MSE simultaneously considers three main aspects of the biological-management cycle: the biological system or "truth" (operating model), the observation process (observation model) and population assessment (estimating model), which then may influence management decisions (Sainsbury et al. 2000, Bunnefeld et al. 2011). In the MoSE, we use the same general approach of creating an operating model of a biologically realistic virtual population for sampling and then apply various uncertainties to the data collected from the operating model (observation model, which is then passed to the estimating model that determines population status from the imperfect information, or data). MSEs often focus on stochasticity in the operating model and uncertainty surrounding management actions, and while including

an observation model is essential to MSE, a complete exploration of uncertainty and stochasticity in collecting data and in evaluating assumptions made during assessment is often minimal and not the specific focus of the MSE (e.g., Gao and Hailu 2013, Winship et al. 2013, Fay et al. 2014, Grüss et al. 2016). Our MoSE approach specifically experiments with monitoring strategies with data uncertainty to determine the effect on population assessments, using green sea turtles as a case study, and how observation errors propagate to population assessment errors, such as inaccurate estimates of population trend.

Our primary goal is to illustrate how the MoSE approach provides advice to improve monitoring plans used to assess populations of sea turtles, using an agent-based model for green sea turtles (Piacenza et al. 2017) as the operating model and a series of simulated population conditions. We asked four primary questions: (1) Given each biological and observation scenario, what is the accuracy of estimated population trend? (2) How does time-series length affect the accuracy of population trend? (3) What are the probabilities of false positive and false negative trend assessments? (4) Does the population structure and harvest legacy influence which monitoring strategy is best?

<center>METHODS</center>

Monitoring strategy evaluation

In MoSE, the process cycle examines the biological system, monitoring, and population assessment (Fig. 1). Here, we are most interested in the discrepancies of a population status indicator (PS_0) from the true values obtained from the operating model (PS^*), dependent on the biological and impact state and monitoring approach employed. Notably, it is also possible to explore a variety of population status indicators, that is adult abundance, nester abundance, and nester recruitment in MoSE. For this paper, we chose to focus on population trend as it is commonly evaluated in assessing sea turtle species population status (National Marine Fisheries Service and U.S. Fish and Wildlife Service 1991, 1998, National Research Council 2010) to illustrate how MoSE functions.

The operating model: Simulating green sea turtle populations

We simulated green sea turtle population dynamics using the green sea turtle agent-based model (GSTABM) described in Piacenza et al. (2017). Simulating biological and observation data is advantageous for several reasons. First, the state of the simulated population is known completely and without error, allowing us to compare the simulated true state of the population to the estimated state of the population based on observed data from the simulated population. The difference between the true state and the estimated state provides quantitative measures for evaluation of the tools used to collect data and estimate population trends. Second, the GSTABM is also useful because it allows for explicit modeling of two independent sources of variability in simulated data: process and observation errors. ABMs (agent-based models) simulate individual behaviors and therefore operate at the scale by which population dynamics and monitoring occur (Letcher et al. 1998, DeAngelis and Mooij 2005). Agent-based models (ABMs) were previously used for MSEs to evaluate multiple uses of ocean resources off the western coast of Australia including recreational fishing (McDonald et al. 2008, Gao and Hailu 2013). ABMs have been applied to sea turtles to study population viability, the influence of temporal variability and age-dependent mortality on population dynamics, and to test different monitoring schemes for within-season sampling to optimize monitoring season timing and duration (Mazaris et al. 2005, 2006, Mazaris and Matsinos 2006, Whiting et al. 2013). ABMs are particularly useful when studying the coupling of individual variation with biological and monitoring models, as biological and monitoring complexity can both be incorporated, such as density dependence, environmental forcing, sampling biases, and interannual variability in sampling. Notably, the use of an ABM is not prerequisite in the MoSE approach, and any type of population model could be used for the operating model.

The GSTABM simulates individual sea turtles, with individual-level variation in reproduction, survival, and age at maturation (Table 1, Fig. 1). Individuals are classified in age classes (hatchlings 0–1 yr, pelagic juveniles 2–3 yr, neritic juveniles 4–11 yr, subadults 12–< age at maturity, adults ≥ age at maturity (Piacenza et al. 2017). Age at maturity, clutch frequency, and clutch size are individually variable and stochastic, drawn from Poisson distributions. Parameter values were based on existing data and we tested alternative parameter distributions, but ultimately the Poisson distribution had the best fit to the existing data for these parameters (Niethammer et al. 1997, Tiwari et al. 2010, Piacenza et al. 2016, 2017). Hatchling production is density dependent and regulated by the individual nester's clutch size, clutch frequency, and the nester density in a given year. We used a Ricker-type function to represent hatchling production because past research suggests decreased numbers of viable nests and hatchlings with high numbers of nesters or nests (Tiwari et al. 2006, Ocana et al. 2012), rather than an asymptotic relationship such as the Beverton-Holt model or additional curvature associated with the Shepherd model (Girondot et al. 2002, Caut et al. 2006). The GSTABM also simulates annually varying breeding probabilities, and hence who is breeding in a given year (Table 1), which is a more accurate representation for green turtles, as they are obligate 1-yr skip nesters, than how breeding probability is typically modeled in matrix projection models (e.g., Crouse et al.
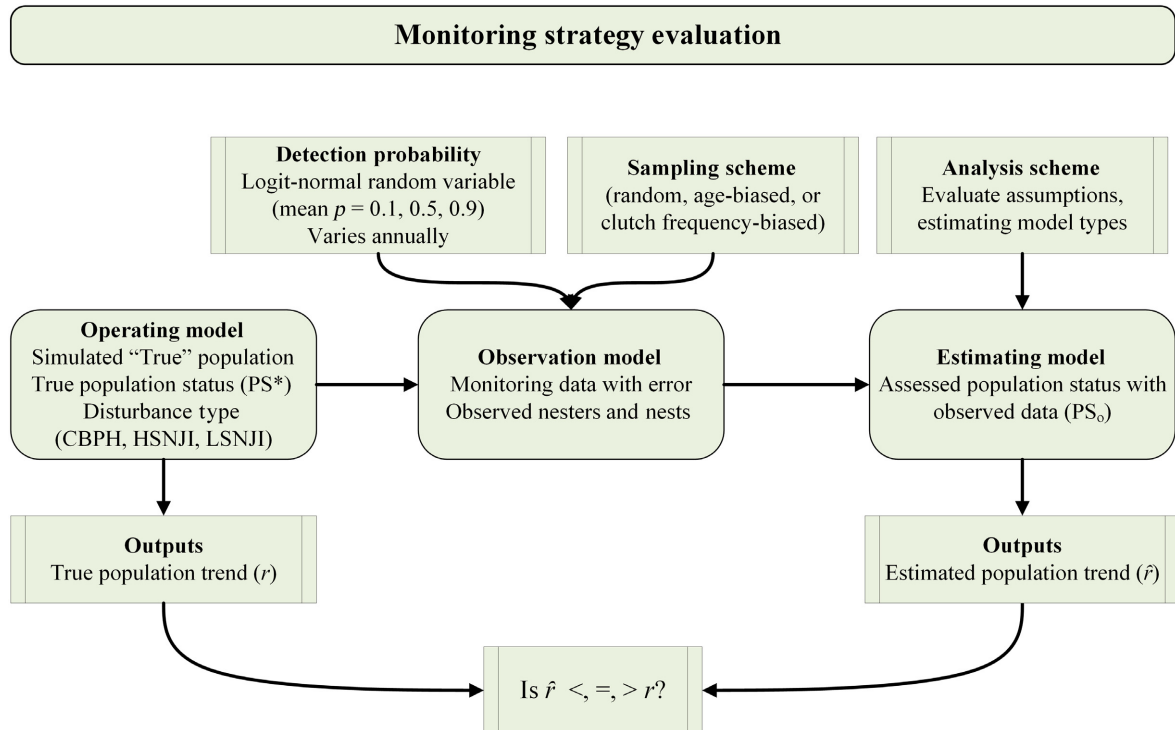
Monitoring strategy evaluation

FIG. 1.   Flow chart for Monitoring Strategy Evaluation (MoSE). Green sea turtle populations are simulated in the operating model, using the Green Sea Turtle Agent-Based Model. Female green sea turtle population structure and monitoring simulated for 175 yr and replicated 50 times. In each replicate run, the population was subjected to an experimental disturbance from time steps 200–250. Disturbance is simulated with (1) cyclic breeding probability and 7% of subadults and adults removed per year for 50 yr (CBPH), (2) low severity neritic juvenile impacts (10% removed per year, LSNJI), and (3) high severity neritic juvenile impacts (50% removed per year, HSNJI). The observation model samples nesters and nests, either randomly or with a type of bias, and with an annually variable detection probability randomly drawn from a logit-normal distribution. The estimation model uses the simulated monitoring data to estimate population status indicators (PS_o), in this case population trend ($\hat{r}$). Estimated population trend is compared to the simulated true values of the population status indicators, PS*, in this case true population trend (r) generated in the operating model.

1987, Crowder et al. 1994, Casale and Heppell 2016). In turn, interannual variability in nesting abundance, characteristically observed in nesting populations, emerges in the model (Carr et al. 1978, Hays 2000, Solow et al. 2002, Piacenza et al. 2016).

Survival is dependent on the age class of the individual. Survival integrates natural and anthropogenic mortality (including ongoing bycatch), but we assume that anthropogenic mortality is minimal so we do not explicitly model it. Hawaiian green sea turtles tend to inhabit inshore waters, where there is very little overlap with fisheries, and generally, green sea turtles have low mortality rates associated with bycatch in U.S. Pacific waters (Finkbeiner et al. 2011). Adult survival rate was based on a 29-yr mark–recapture analysis (Piacenza et al. 2016) and survival rates for younger age classes (hatchling to subadult) were based on estimates from Van Houtan et al. (2014).

After initialization (years 0–174), the model follows this general timeline: stable population (175–199 time steps), impact (200–249 time steps), and recovery (250–350 time steps). The initialization period is discarded as

the population is reaching quasi-stabilization in population structure. Henceforth, we refer to the three main time periods, stable, impact, and recovery, as the population phases. The stable phase represents the "control" period in which no harvest and variability in abundance and nesting are due solely to demographic and environmental stochasticity. We implemented the GSTABM in NetLogo 5.1.0, software developed to implement agent-based models (Wilensky 1999).

The observation model: Simulating population monitoring

The GSTABM also simulates the process of observing and collecting data from sea turtle nesters and nests annually. Details of the population monitoring submodel of the GSTABM are included in Piacenza et al. (2017). To summarize, the input detection probability (p) is a random variable with a logit-normal distribution (Table 1; Appendix S1: Fig. S1D). In the base model, sea turtles that are actively nesting in a given year are randomly selected to be monitored (unless under a biased sampling treatment, see *Biological disturbance and monitoring*

TABLE 1. Parameter definition and inputted specifications for the green sea turtle operating model used in the monitoring strategy evaluation.

Parameter	Description	Mean ± SD or variance with range†	Units	Statistical distribution	References
Age at maturity	age individual reaches sexual maturity	30 (17–41)	yr	Poisson (truncated)	Zug et al. (2002), Van Houtan et al. (2014)
Clutch frequency	nests laid during reproductive season	4 ± 4	no. nests	Poisson	Niethammer et al. (1997)
Clutch size	potential no. eggs laid per nest	43.2 ± 43.2	no. eggs	Poisson	Niethammer et al. (1997)
Hatchlings produced	Realized no. eggs laid across all clutches in a given reproductive season, based on Ricker-type density-dependent function	103 (0–187)	no. individuals	–	Tiwari et al. (2010), Ocana et al. (2012)
Breeding probability	mean annual breeding probability	0.2519 ± 0.0127	yr^{-1}	gamma	Piacenza et al. (2016)
Annual survival	threshold for annual survival; individual survival is based on uniform random number selection against survival threshold for each age class				
Hatchling		0.350	yr^{-1}	uniform	Van Houtan et al. (2014), Piacenza et al. (2016)
Pelagic juvenile		0.800	yr^{-1}	uniform	
Neritic juvenile		0.824	yr^{-1}	uniform	
Subadult		0.876	yr^{-1}		
Adult		0.929	yr^{-1}	uniform	
Detection probability	mean probability of detecting a nester	0.1 ± 0.020, 0.5 ± 0.099, 0.9 ± 0.18	yr^{-1}	logit-normal	Piacenza et al. (2016)

Note: For more detailed information on parameterization of the green sea turtle agent-based model, see Piacenza et al. (2017).
† Variance for parameters with Poisson distribution.

experiments). We assume variability in detection is constant over time and we scale the standard deviation to the mean so that the standard deviation is proportional to the mean across the experimental detection levels, based on the coefficient of variation (CV = 0.2) from estimated detection probabilities from a mark–recapture analysis of 29-yr study of green sea turtles in Hawaii (Piacenza et al. 2016). Detection of nesters and nests in MoSE pertains to detection of nesters and nests within an entire population, not a specific nesting beach. During the observation model, the model collects data similar to output data collected on the population as a whole, including nester abundance, nest abundance, population-level means, and standard deviations of hatchlings produced per female, remigration interval (years between nesting seasons), age at maturity, clutch frequency (nests per female), clutch size, hatchling production, and total number of lifetime nesting seasons.

The estimating model: Simulating population assessments

Estimating population trend—We calculated the trends in population growth based on the simulated true and observed number of nesters and nests. We calculated the population trend (*r*) using the model of exponential growth across 5, 7, 10, 15, or 20 yr, by first linearizing the model and then applying linear regression:

$$\ln(Y_i) = \theta_0 + \theta_1 x_i + \varepsilon_i \tag{1}$$

where θ_0 represents the intercept and θ_1 represents the slope, which characterizes the trend. We randomly selected the starting points for the trend time series within the three population phases for each of the 50 replicate runs of the experimental treatments. We also ensured that the starting points were sufficiently early in the population phase so that the longest time-series (i.e., 20 yr) did not overlap with the next population phase. To determine the appropriate number of replicate runs, we analyzed the CV of adult abundance, nester abundance, discrete population growth, and nester recruitment. All of these emergent processes stabilized within 30 replicate runs of the base model, and we selected 50 replicates to ensure capturing the range of model outputs; Cowled et al. 2012, Piacenza et al. 2017). Further, to assess if the main output of interest, median percent

bias stabilizes within 50 replicate runs, we examined the cumulative median percent bias across the population phases, detection probability, and trend duration from 1 to 150 replicate runs. Median percent bias tends to stabilize <50 runs, while the stable population phase (particularly when the trend duration is 5 yr) does retain some variability in median percent bias across 150 runs (Appendix S1: Fig. S2).

Determining false positive and false negative trend diagnosis—We also measured the proportion of false negatives and false positives in trend estimation. We calculated the proportion of false negatives as the number of times the estimated trend was negative when the true trend was positive across 50 replicate runs, and calculated the proportion of false positives similarly as the number of times the estimated trend was positive when the true trend was negative. We were not interested in small deviations in estimated trend compared to the true trend, therefore, we included a ±3% buffer about the trend estimate, so that estimated trends $-0.03 \leq \hat{r} \leq 0.03$ were regarded as = 0, as were true trends. For example, if $r^* > 0.03$ and $\hat{r} < -0.03$, then this was considered a false negative trend diagnosis; but, if $r^* = 0.02$, and $\hat{r} = -0.02$, then both of these trends were regarded as stable and this was not considered an erroneous trend diagnosis. If *both* r^* and $\hat{r}$ were ≥ 0.03, or if *both* r^* and $\hat{r} \leq -0.03$, then in either case the trend diagnosis was considered accurate.

Measuring error and bias: Median percent bias—For population trend, we compared the estimated values to the true value simulated in the GSTABM operating model. Accuracy of estimated values was defined as the amount of error from the simulated true value. We measured the amount of error from the true population indicator by calculating the percent bias. Using percent bias as a metric of error is useful as it measures the distance and the directionality of the estimate from the true population trend (Lynch et al. 2012, Harford et al. 2015, Thomas et al. 2018). For each simulation j and time duration τ, percent bias in the population trend was calculated as

$$B = 100 \times \frac{\hat{r}_{j,\tau} - r^*_{j,\tau}}{r^*_{j,\tau}} \qquad (2)$$

where B is percent bias, $\hat{r}_{j,\tau}$ and $r^*_{j,\tau}$ are the estimated and true population trend, respectively. We computed percent bias for each of the three population phases using the five population trend durations (5, 7, 10, 15, and 20 yr) in each replicate run across the 27 experimental treatments. We ranked the factors by median percent bias across the 27 experimental treatments to determine which factors contribute the most toward improving estimation accuracy.

Biological disturbance and monitoring experiments—To explore this first comprehensive use of MoSE, we

created a $3 \times 3 \times 3$ experiment (27 treatments) with three levels each of detection probability, sampling type, and disturbance type (Fig. 1). Each treatment was simulated for 350 yr with 50 replicate runs. Assuming a generation time of 40 yr (based on simple calculations from a life table), this time span accounts for approximately nine generations. We modeled the mean detection with three levels: 10%, 50%, and 90% of nesters and nests. We included a broad range of detection probabilities with which to sample the nesters and nests to provide an overview of the influence of detectability on sea turtle population assessments.

We included three experimental treatments for sampling type: random, age bias, and clutch frequency bias. The random sampling treatment represents a simplified null model, in that real world sea turtle beach monitoring programs are unlikely to randomly sample individuals. Rather, nesting beach monitoring is more comparable to large line transects along stretches of known nesting beaches, and monitoring strives to encounter every single individual or nest on the nesting beach or at least the turtles or nests sequentially encountered while moving down the beach (Gerrodette et al. 1999, Schroeder and Murphy 1999, National Research Council 2010). In the random sampling treatment, monitored nesters are a random sample of the total number of nesters at each time step

$$\text{observed nesters}_t = p \times N_{\text{nesters},t} \qquad (3)$$

where p is the detection rate and $N_{\text{nesters},t}$ is the abundance of nesters at time t. Observed nests are calculated similarly.

In the nonrandom sampling treatments (i.e., age bias and clutch frequency bias), the GSTABM sorts nesters by age or clutch frequency and selects the oldest or most fecund individuals first, that is if $p = 50\%$, then the GSTABM selects the top 50% oldest turtles. The age bias sampling treatment simulates increased likelihood of encountering older individuals and their nests, who may have higher site fidelity to the nesting beach than newly recruited nesters or where fisheries bycatch impacts subadults and small adults so that recruitment to the index nesting beaches is limited (Mortimer and Carr 1987, Tucker and Frazer 1991, Van Houtan and Kittinger 2014). Clutch-frequency-biased sampling simulates the bias toward more fecund individuals, and their nests, who return to the nesting beach more frequently during a nesting season and are more likely to have a greater detection probability than less fecund individuals (Tucker 2010, Hart et al. 2013).

We included three disturbance treatments intended to reflect past disturbance events that have occurred in sea turtle populations: (1) cyclic breeding probability with subadult and adult Harvest (CBPH), (2) high severity neritic juvenile impacts (HSNJI), and (3) low severity neritic juvenile impacts (LSNJI). The CBPH treatment represents oscillations in annual breeding probability

that could be a result of large-scale climatic events, such as El Nino Southern Oscillation. In this treatment, breeding probability oscillates as a sine function, with the form

$$BP = 0.27 \times (\sin 50 \times t) + 0.3 \qquad (4)$$

where BP is the breeding probability and BP fluctuates between 0.02 and 0.57 (Appendix S1: Fig. S1B) at time t. The breeding probability cycle frequency in the GSTABM occurs every 8 yr, which is intended to simulate about the same frequency as major El Nino events (Limpus and Chaloupka 1997, Saba et al. 2007, Trujillo and Thurman 2008). In the CBPH treatment, populations are also subjected to harvest where 7% of subadults and adults (ages ≥ 11) are removed from the population annually for 50 yr. This treatment simulates the population disturbance of targeted sea turtle fishery similar to green sea turtles in Hawaii (Witzell 1994, Van Houtan and Kittinger 2014). While this treatment confounds environmental forcing and anthropogenic impacts, it is intended to simulate real-life situations where both factors occur simultaneously but population trends are still assessed.

The LSNJI and HSNJI treatments simulate hypothetical disturbances where neritic juveniles (ages 4–10) are removed. These two treatments are intended to test how the disturbance severity to a life stage particularly sensitive to unnatural losses (such as by a targeted fishery or bycatch) influences the accuracy of estimating population trend (Crouse et al. 1987, Crowder et al. 1994, Heppell 1998) such as turtles that were bycaught in the shrimp trawl fishery in the North Atlantic prior to the institution of turtle excluder devices (Magnuson et al. 1990, Epperly et al. 2002). In LSNJI treatment, 10% of the neritic juveniles are removed annually for 50 yr, and in HSNJI treatment, 50% of neritic juveniles are removed annually for 50 yr. We set two different severities to examine the influence of disturbance severity on the accuracy of estimating population trend.

We recognize that many other experimental biological, detection level, and sampling treatments could have been conducted. However, our goal is to compare plausible scenarios in which to test the MoSE tool and to illustrate the potential drivers of error in population assessments of sea turtles.

Results

Population response to disturbance

For the purposes of illustrating the typical population structure across the experimental treatments, Fig. 2 depicts the mean population structure during the stable, disturbed, and recovery phases of two of the experimental treatments: CBPH (Fig. 2A) and HSNJI (Fig. 2B) with 50% detection and random sampling. When the simulated green sea turtle populations were subjected to

CBPH, all demographic groups declined during the impact phase, and then postdisturbance, the population began to recover (Fig. 2A). When the simulated populations were subjected to HSNJI, population-level responses were more complex (Fig. 2B). As sea turtle populations tend to be very sensitive to changes in neritic juvenile survival, the responses of the age classes to the disturbance were, perhaps, not surprising (Crowder et al. 1994, Heppell 1998). Higher amplitude oscillations in abundance occurred for all demographic groups, but were the strongest for subadults and neritic juveniles (Fig. 2B). After 100 yr of recovery, the populations in both treatment types had not returned to predisturbance levels. For both impact treatments, the variance about the main age classes (adults, subadults, neritic juveniles) was less in comparison with nesters, nests, observed nesters and observed nests. The variance about nesters and nests tended to increase during the latter stages of the impact and during the early stages of recovery, and consequently, variance about the observed nesters and nests also increased during this period.

We subjected the base model to a sensitivity analysis, and determined that the model was most sensitive to neritic juvenile, subadult, and adult survival rates and, notably, none of the stochastic input variables (e.g., breeding probability, age at maturity, clutch frequency, clutch size) ranked highly (Piacenza et al. 2017). We also compared model output to key population-level processes for which empirical data exist and demonstrated that the GSTABM is a good representation of sea turtle population dynamics (Piacenza et al. 2017).

Population assessment from simulated population monitoring

Estimating population trend—The accuracy in the estimates of adult population trend drawn from observed nesters increased with the duration of the trend time-series and detection level across the biological treatments of CBPH, LSNJI, and HSNJI (Fig. 3; Appendix S1: Figs. S3–S9). We also examined relationships between the true total population trend and estimated trend drawn from observed nesters and nests, and true adult population trend and estimated trend drawn from observed nests, but patterns across these groups tended to be similar to true and nester-estimated adult population trend presented here (Table 2; Appendix S1: Figs. S3–S8, Data S2). Precision of the trend also depended on the population phase and detection rate; the worst median percent bias occurred with the LSNJI treatment sampled with a clutch frequency bias and 50% detection rate (median percent bias 61%; Table 2). The CBPH treatment with random sampling and 50% detection rate had the lowest median percent bias (−0.46; Table 2).

Trend duration and accuracy – How long of a time series is necessary to accurately estimate population trend?—

A) Cyclic BP and harvest, random; 50% detection, 7% impact

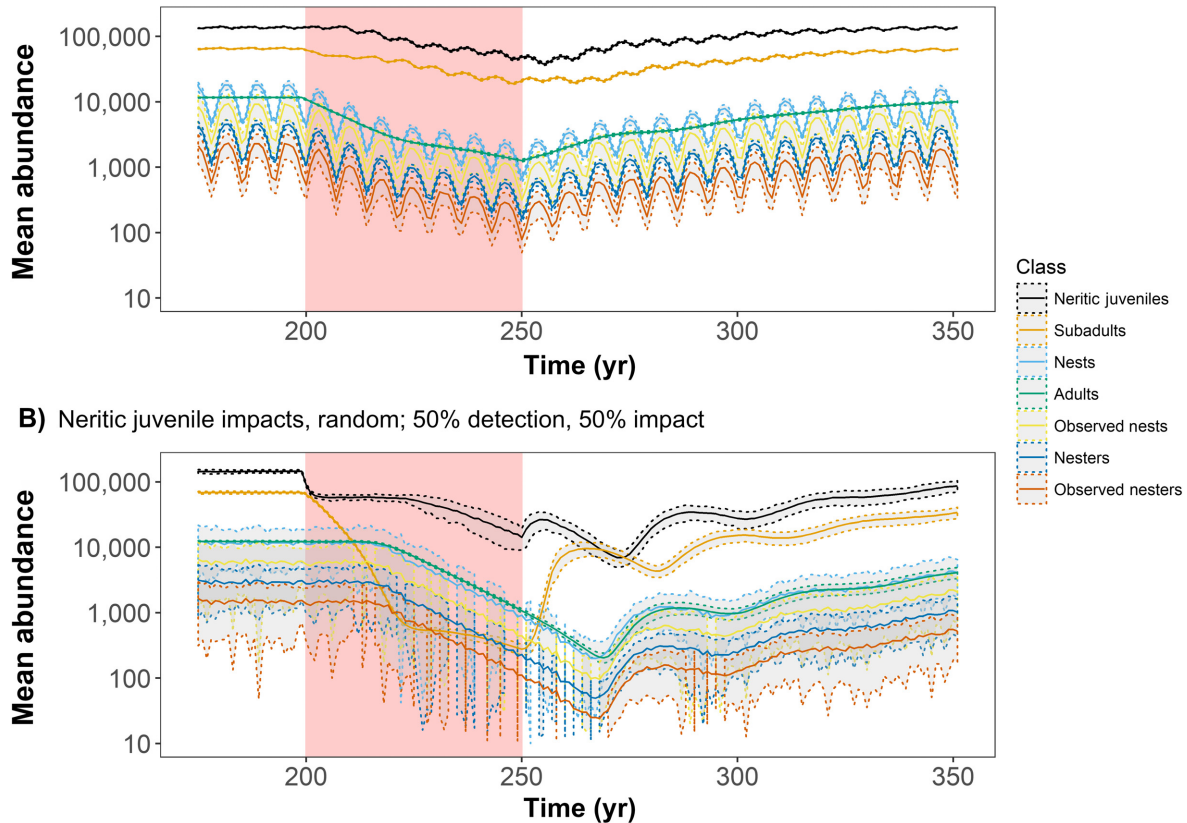**B)** Neritic juvenile impacts, random; 50% detection, 50% impact

Fig. 2. Female green sea turtle population structure and monitoring simulated for 175 yr and replicated 50 times for two of the biological disturbance treatments: (A) cyclic breeding probability (BP) with 7% of subadults and adults removed annually for 50 yr and (B) high severity neritic juvenile impacts for 50 yr. Both panels show monitoring results for detection probability with a mean of 50% for nesters and nests. Colored lines indicate the mean abundance of the demographic classes, and shaded areas indicate the 95% confidence intervals. The pink shaded area indicates the 50-yr impact period.

When the populations were sampled randomly, and during the first 5–7 yr of survey data and the impact phase, trend estimates tend to be positively biased, except for the LSNJI (Fig. 3; Appendix S1: Figs. S10 and S11). Over time, the bias tends to decline, but in the case of the LSNJI, the bias remained after 20 yr (Fig. 3I). During the recovery phase, the trend estimates are biased, and the direction, degree, and duration of the bias are dependent on the impact type and severity (Fig. 3). For neritic juvenile impacts, the more severe impact resulted in reduced accuracy of the trend estimate. If sampling is biased, however, then the monitoring window necessary for unbiased estimates becomes elongated in some scenarios, and ranges from >20 yr (Appendix S1: Figs. S12 and S13).

The trend percent bias reached an asymptote after about 10 yr, and variance about the median percent bias for less than 10 yr of data was one to seven orders of magnitude larger than the true adult population trend, depending on the impact treatment type (Fig. 3; Appendix S1: Figs. S10 and S11). We do not show the detection levels, as there were only modest differences in

median percent bias (Appendix S1: Fig. S11). Variance about the median percent bias in trend estimate was greatest for the CBPH treatment during the stable phase, but accuracy tended to improve after a 15-yr trend duration (Appendix S1: Fig. S10). The direction of bias alternated depending on the population phase and the biological treatment. For example, the median percent bias during the impact phase for CBPH treatment tended to be positive with a 5- or 7-yr trend duration and then approached zero (Fig. 3B), but during the recovery phase, the percent bias was negative and remained so for a 20-yr trend duration (Fig. 3C).

Precision of population trend is also dependent on the interaction of detection rate and trend duration (Fig. 4; Appendix S1: Fig. S14). Assessing population trend with minimal error (i.e., percent bias approaching zero) when the population is nominally stable (across impact type, detection rate, and duration) is particularly challenging. Since the stable phase occurs before the impact phase in the model runs, error must be a result of environmental stochasticity, sampling scheme (e.g., random or biased)

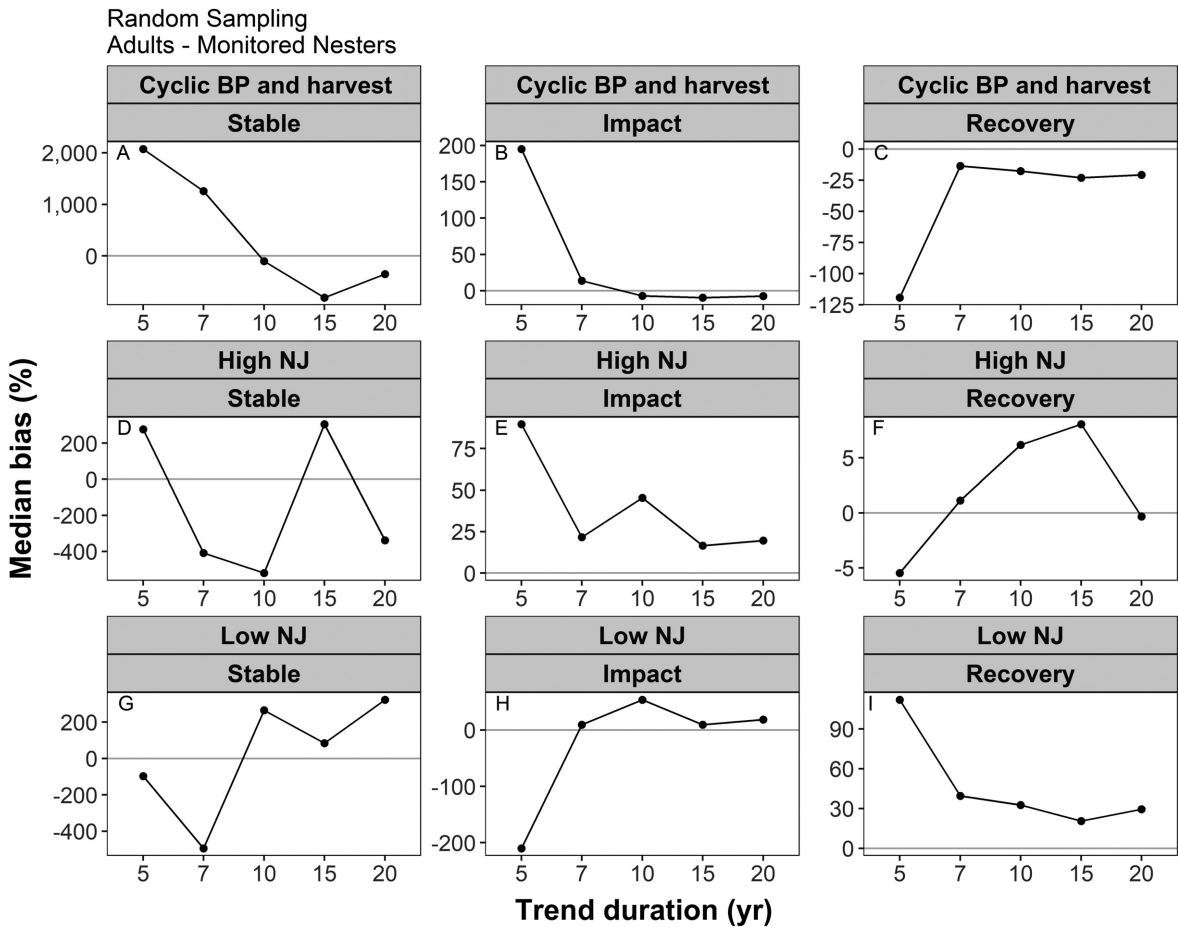Random Sampling
Adults - Monitored Nesters

FIG. 3. Trend duration and median bias of estimate population trend for populations with random sampling across trend duration, summarized across detection probability. The biological treatments are (A–C) cyclic breeding probability with harvest cyclic, (D–F) high severity neritic juvenile (NJ) impacts, and (G–I) the low severity neritic juvenile impacts treatments across the three population phases, Impact (A, D, and G), recovery (B, E, and H), and stable (C, F, and I). A bias of 0% indicates no bias in the estimated population trend, bias > 0% indicates the estimated population trend ($\hat{r}_{j,\tau}$) is greater than the true population trend ($r^*_{j,\tau}$), and a bias < 0% indicates the estimated population trend ($\hat{r}_{j,\tau}$) is less than the true population trend ($r^*_{j,\tau}$). Errors bars were left off to improve visualization (see Appendix S1: Figs. S10 and S11 for figures with error bars and median percent bias across detection probabilities).

or process error (such as in the CBPH impact type), rather than transitory dynamics. The signal of cyclic breeding probability is apparent in the pattern of percent bias over the duration of the population trend, particularly evident during the recovery phase of the CBPH treatment (Fig. 4). Assessing population trend accurately is relatively more likely during the impact and recovery phases of the HSNJI and LSNJI treatments, especially if trend duration is >10 yr (Fig. 4). Interestingly, there are several scenarios where it is possible to achieve a low percent bias with a short monitoring window (i.e., 5 yr) and 50% detection rate, such as during the recovery phase of the LSNJI with a clutch frequency bias and the impact phase of the LSNJI with random sampling (Fig. 4). If the monitoring time frame is short (i.e., 5–7 yr), increasing detection rates often only serves to switch from a negative bias to a positive bias, rather

than reducing error. However, increasing the monitoring duration alleviates this problem. In all, the patterns of percent bias differ across all treatment combinations (i.e., impact type, population phase, detection rate, and monitoring duration).

How likely is a wrong trend diagnosis?—The proportion of false negative and false positive trend estimates tended to decrease with trend duration, except for during the stable population phase. The CBPH treatment during the stable phase had the highest probability of an erroneous trend diagnosis, 44% proportional frequency of false positive, with 20 yr of monitoring (Fig. 5A). The HSNJI treatment during the stable phase with 20 yr of monitoring had the greatest overall probability of an erroneous trend diagnosis (58% proportional frequency of either error type, Fig. 5D). Often the chance of false

TABLE 2. Median bias (%) for population trend for the 27 treatments, each with 50 replicate runs, summarized across population phase and trend duration.

Sampling scheme, detection probability, disturbance treatment	Status indicator, bias (%)	
	Trend, observed nesters	Trend, observed nests
Random		
0.1		
CBPH	−10.58	−11.91
LSNJI	34.90	30.69
HSNJI	6.74	7.46
0.5		
CBPH	−13.83	−11.24
LSNJI	2.86	3.36
HSNJI	13.92	14.30
0.9		
CBPH	−0.46	−3.36
LSNJI	35.75	44.17
HSNJI	9.63	9.02
Age biased		
0.1		
CBPH	−11.42	−9.26
LSNJI	14.58	18.93
HSNJI	2.89	3.37
0.5		
CBPH	−12.96	−12.67
LSNJI	3.35	1.14
HSNJI	10.79	13.23
0.9		
CBPH	−16.45	−14.47
LSNJI	17.82	16.40
HSNJI	18.00	17.46
Clutch frequency biased		
0.1		
CBPH	−12.26	−8.58
LSNJI	17.31	20.59
HSNJI	18.33	18.77
0.5		
CBPH	−3.08	−7.31
LSNJI	*60.60*	*50.72*
HSNJI	4.07	9.33
0.9		
CBPH	−0.99	−5.77
LSNJI	10.42	−1.15
HSNJI	11.23	11.73

Notes: The absolute largest median percent bias for population trend (drawn from either observed nesters or nests) is indicated in boldface italic type. CBPH, cyclic breeding probability with 7%/yr subadult and adult harvest; LSNJI, low severity (10%/yr) neritic juvenile impacts; HSNJI, high severity (50%/yr) neritic juvenile impacts treatments. See Data S1 for median bias across population phase, detection probability, and trend duration.

positive or false negative errors does not disappear even after 20 yr of monitoring (Fig. 5; Appendix S1: Fig. S15). During the stable population phase, the likelihood of false positive in trend estimate tends to increase with the monitoring duration. For example, with the HSNJI treatment during the stable population phase, there is a 36% probability of falsely concluding a population is increasing even after monitoring the population for 20 yr (Fig. 5D). False positive errors occurred more frequently than false negative errors. False negative errors did not occur during the impact and recovery phase of the CBPH and the recovery phase of HSNJI (Fig. 5B, C, F). False positive and false negative errors, when present, tended to decrease with trend duration during the impact and recovery phases, but increased during the stable phase.

DISCUSSION

The MoSE approach has the potential to help modify existing nesting beach monitoring programs and design future nesting beach monitoring programs for sea turtles where detectability is limited, and has the capability to be extended to other species. The MoSE illustrates the potential error and bias that can arise from population assessments based on sea turtle nesting beach data alone. We also found that the history and severity of population disturbance have important and persistent effects on the accuracy of population assessments drawn from monitoring nesting beaches. Interestingly, error in estimating population trend often remains for different impact scenarios, even with random sampling and 90% detection rate, especially over shorter monitoring duration intervals. Thus, focusing on improving these aspects of monitoring may do little to improve accuracy of population status assessments at least during the early stages of monitoring. The MoSE tool suggests that assessment accuracy is dependent on the underlying population phase, disturbance history and severity, and the length of the time series. Thus, is it important to carefully consider the disturbance history, as much as may be known, about a population when assessing population status from beach monitoring.

MoSE can be used as a tool to provide prognostic advice for how to improve monitoring to increase accuracy for estimating population status indicators, namely population trend. For example, our model suggests that at least 10 yr of monitoring data is necessary to accurately estimate population trend, regardless of biological impact, underlying process errors, detection level, and population phase; but, this would be influenced by the duration of cycles in breeding probability (Solow et al. 2002, Saba et al. 2007, del Monte-Luna et al. 2012). If juvenile age classes were disturbed and trend estimates occur during the recovery phase, then trend durations *exceeding* 20 yr are necessary to improve accuracy.

If conservation managers want to avoid conservation assessment errors (i.e., false positive or false negative trend diagnoses), it is important to consider the population phase, the impact type, severity, and monitoring duration in the analysis. Improving detection rates to 90% does little to lower the chances of erroneously concluding the direction of a population trend. However,

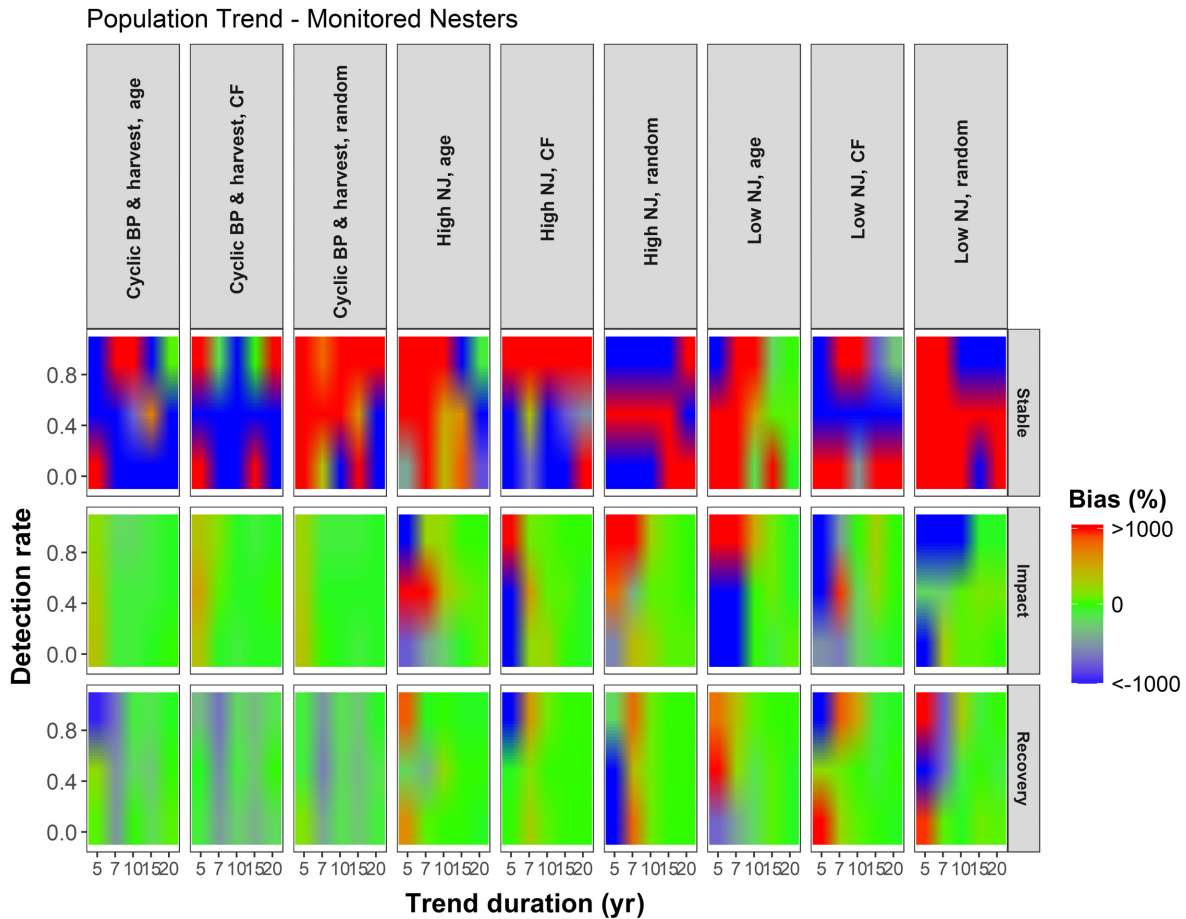## Population Trend - Monitored Nesters

FIG. 4. Heat map showing percent bias of estimated population trend as a relationship between monitoring duration and detection rate for monitoring nesters. Extreme values of bias (≤−1,000% and ≥1,000%) were binned in order to improve visualization. The heat map colors represent the bias of population trend estimates from monitoring nesters across simulation replicates for each experimental treatment. Shades of red and blue indicate a positive and negative bias, respectively, and shades of green indicate bias approaching 0%. For the purposes of visualization, the detection probabilities are treated as continuous values and coloration is interpolated between input detection levels. Age refers to the age bias treatment, CF refers to the clutch frequency bias treatment, and random refers to the random treatment.

increasing monitoring duration tends to lower the probability of false trend assessment, except during the stable population phase. It is interesting that, in some scenarios, the probability of a false positive diagnosis increases over time. This situation arose during the stable phase when essentially population growth is zero; however, with a shallow slope, it can be difficult to estimate population growth rate accurately. In addition, had we run the simulations longer, it is probable that eventually the bias would have approached zero.

Our work builds on other studies working to optimize monitoring of sea turtle nesting beaches. Sims et al. (2008) found for Hawksbill sea turtles (*Eretmochelys imbricata*) in the Eastern Caribbean that by examining the statistical power of an intensive protocol vs. those of shorter duration and later start dates, it is possible to optimize monitoring start date to later in the season and for a short survey duration of just 10 weeks with a negligible loss of statistical power and with savings in cost.

Jackson et al. (2008) examined how accurately different monitoring schemes estimated the total number of nests and the ability to detect a population decline based on monitoring nests for green and loggerhead (*Caretta caretta*) turtles in Cypress. Jackson et al. (2008) found that accurate nest abundance estimates could be derived from bolus sampling, where monitoring occurs daily for at least 21 d during the peak of the nesting season. However, the monitoring schemes were relatively insensitive to small population declines (~1%/yr), but on average could detect a 10% change in 12 yr for green turtles and 5 yr for loggerheads. Whiting et al. (2013) compared within season monitoring schemes to determine the optimal scheme for sampling nests for populations with short and long nesting seasons, and found that the phenology of nesting influenced the optimal sampling regime. In a simulation study of north Atlantic loggerheads, Warden et al. (2017) found that neither nest surveys nor aerial surveys alone could sufficiently detect
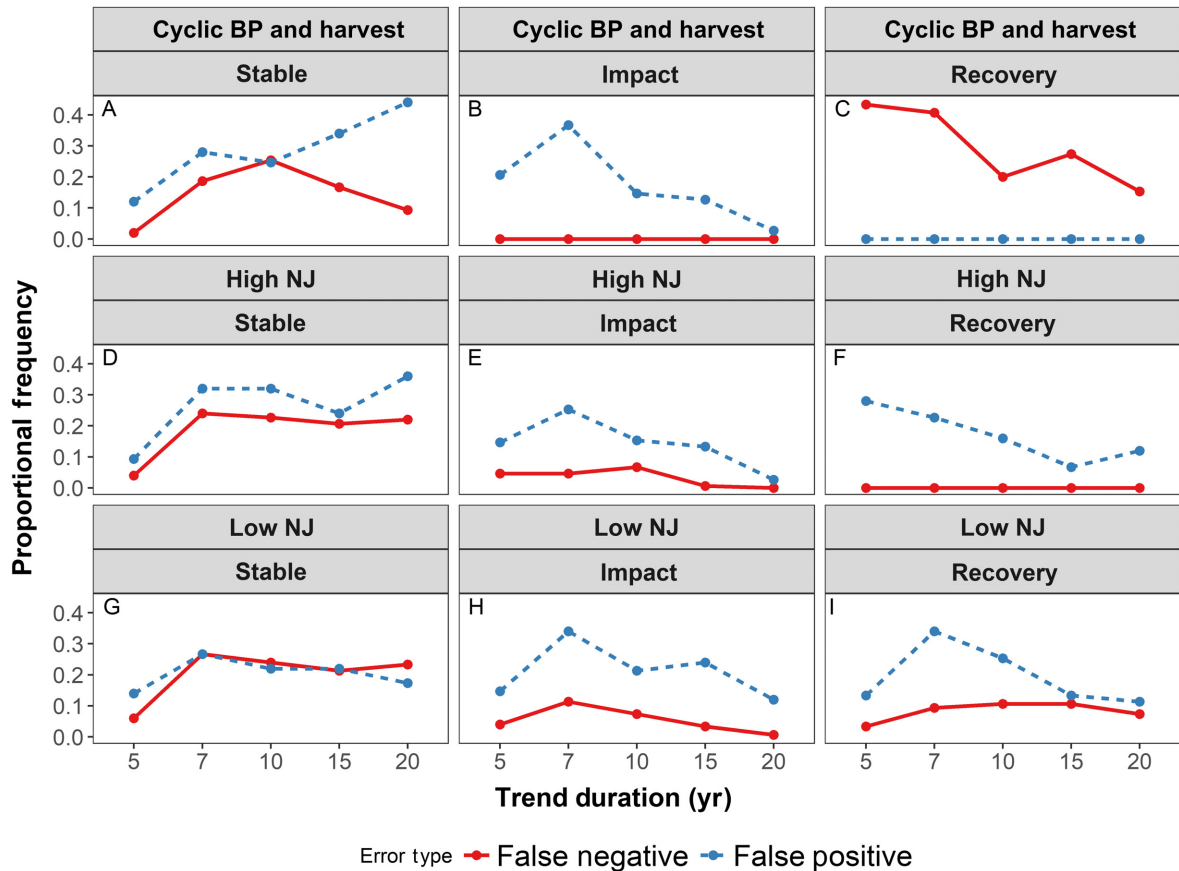
FIG. 5. Proportion of false negative and false positive trend estimates from randomly sampled nesters, summarized across detection probability. The biological treatments were (A–C) the cyclic breeding probability with harvest, (D–F) the low severity neritic juvenile impacts, and (G–I) the high severity neritic juvenile impacts treatments. False negative errors occur if the estimated trend <−0.03, but the true trend is >0.03. False positive errors occur if the estimated trend >0.03, but the true trend is <−0.3. See Appendix S1: Figs. S15–S17 for figures with median percent bias across detection probabilities).

population impacts over the short term, but that nest surveys tended to have lower error than aerial surveys. Altogether, these past studies and our present study suggest that monitoring program managers can make critical decisions to optimize monitoring such that little statistical power is lost, but financial and labor resources are conserved (albeit we did not consider monitoring costs here; for examples in the MSE context, see Mapstone et al. 2008, Grüss et al. 2016, Dichmont et al. 2017).

We present here a proof of concept of the MoSE approach. MoSE has capabilities that extend beyond simple power analyses. This tool can be used to optimize monitoring for estimating abundance, nester recruitment, and vital rates (i.e., clutch frequency, clutch size, remigration interval, and size at maturity). It is also possible to explore assumptions in the assessment model (e.g., assuming constant clutch frequency and remigration interval when extrapolating nests to female abundance) and how those might influence assessment accuracy. MoSE can also be used to estimate likely error

in population status indicators and to identify sources of error that could be minimized, or, at least accounted for during population assessment. Future modeling work could explore other complexities of sea turtle life history, such as within and across season variation in reproduction, exploring the accuracy of monitoring alternative data sources, such as in-water monitoring, and differences in reproductive output between neophytes and veteran nesters (Broderick et al. 2003, Stokes et al. 2014, Warden et al. 2017). The general framework of MoSE can also be used to simulate conservation actions (which result in an increase in survival rate or decrease in take) and the ability of population status indicators to detect their effects, such as a simulation-based power analysis to examine the importance of time series length and measurement error (sensu Taylor and Gerrodette 1993). Indeed, here we present one way of quantifying population trend (i.e., the slope of exponential regression), but multiple alternative methods could be tested and compared for accuracy to the true population trend. In addition, adding a spatial dimension to the GSTABM (for a

specific population or region of interest) would enhance exploration of issues of nesting site fidelity, clutch frequency, and line-transect sampling of stretches of nesting beaches vs. complete census or randomized sampling and their influence on the accuracy of population assessments. Last, the addition of a cost submodel would allow for MoSE to quantitatively assess the costs and benefits of different sampling approaches.

Here, our results suggest that estimating population trend from observed nesters is marginally more accurate than estimating from observed nests. We caution readers from interpreting this as no benefit from monitoring individual nesters, a costlier effort, over nests. There is an appreciable accuracy gap by using nesting beach data in general as a population index, and a smaller difference between either source of nesting beach data: observed nesters and nests. The inaccuracy of monitoring nests vs. nesters is minimal in comparison, but both population indices are problematic, and observed nests are marginally worse than nesters, but more sensitive to clutch-frequency bias in sampling (which is important if a monitoring group intends to estimate abundance, e.g., Richards et al. 2011, Esteban et al. 2017). In addition, monitoring nesters includes value added to a monitoring program in that additional biological data can be collected from nesters: body length and size distribution changes, nester recruitment, size at maturity, breeding probability, clutch frequency, and clutch size in relation to nester size and status (i.e., neophyte vs. veteran; Broderick et al. 2003, Stokes et al. 2014, Piacenza et al. 2016). In addition, monitoring nests is a particularly spatial problem, in comparison with monitoring nesters (albeit spatial issues exist here as well, i.e., nest site fidelity), and had monitoring been modeled as spatially explicit, we may have seen more differences in the degree of accuracy from estimating population trend from observed nesters and nests. Future work should include expanding the GSTABM to be spatially explicit.

While there are obvious difficulties with monitoring sea turtle nesting beaches, vs. in-water studies, we recognize that monitoring nesting beaches remains the most accessible option for encountering individuals and collecting individual-level data (Hamann et al. 2010, National Research Council 2010, Stokes et al. 2014). Our results suggest that a monitoring strategy, based on the impact history, current population phase, and environmental drivers, may be tailored to a specific population. For example, for a population that is currently recovering from intense impacts to neritic juveniles, such as a targeted fishery or incidental catch, at least 20 yr of monitoring would be required before accurately estimating population trend using nesting data, if data are sampled randomly. Population assessors would need to acknowledge that estimates of population trend, at least during the early recovery years, are likely to be biased. On the other hand, for a population susceptible to environmental drivers on reproduction (i.e., a strong El Nino influence on reproduction), and with former disturbance

to adults and subadults (i.e., a targeted fishery), such as what occurred in Hawaii during the 20th century, estimates of trend are likely to be underestimated, especially in the early years of recovery. For this kind of population, trend could be estimated accurately in 10 yr, if data are sampled randomly. As many sea turtle populations endured impacts to juveniles and are currently in the early stages of recovery (Wallace et al. 2011, IUCN 2015, Mazaris et al. 2017), our work suggests that many sea turtle populations are too early in the recovery process to correctly classify population trend as increasing or decreasing, and special care should be taken if reassessing population status during the early years of recovery. However, it may be difficult to have precise estimates of impact rate and history. A monitoring program could qualitatively assess impact history, based on bycatch rates or historical records, and categorically assign impact rate (i.e., low or high), to prescribe monitoring windows based on the results of the MoSE. Ultimately, it is important to carefully consider the impact history, and specifically which age classes were disturbed, when developing a MoSE for a monitoring program and objectives for monitoring duration.

Literature Cited

Akçakaya, H. R., S. Ferson, M. A. Burgman, D. A. Keith, G. M. Mace, and C. R. Todd. 2000. Making consistent IUCN classifications under uncertainty. Conservation Biology 14:1001–1013.

Bowen, B., A. Meylan, J. Ross, C. Limpus, G. Balazs, and J. Avise. 1992. Global population-structure and natural-history of the green turtle (*Chelonia mydas*) in terms of matriarchal phylogeny. Evolution 46:865–881.

Broderick, A. C., F. Glen, B. J. Godley, and G. C. Hays. 2003. Variation in reproductive output of marine turtles. Journal of Experimental Marine Biology and Ecology 288:95–109.

Bunnefeld, N., E. Hoshino, and E. J. Milner-Gulland. 2011. Management strategy evaluation: a powerful tool for conservation? RID G-1946–2011. Trends in Ecology & Evolution 26:441–447.

Carr, A. F., M. H. Carr, and A. B. Meylan. 1978. The ecology and migrations of sea turtles. 7, The West Caribbean green turtle colony. Bulletin of the American Museum of Natural History 162:1–46.

Casale, P., and S. S. Heppell. 2016. How much sea turtle bycatch is too much? A stationary age distribution model for simulating population abundance and potential biological removal in the Mediterranean. Endangered Species Research 29:239–254.

Caut, S., V. Hulin, and M. Girondot. 2006. Impact of density-dependent nest destruction on emergence success of Guianan leatherback turtles (*Dermochelys coriacea*). Animal Conservation 9:189–197.

Chaloupka, M., K. A. Bjorndal, G. H. Balazs, A. B. Bolten, L. M. Ehrhart, C. J. Limpus, H. Suganuma, S. Troëng, and M. Yamaguchi. 2008. Encouraging outlook for recovery of a once severely exploited marine megaherbivore. Global Ecology and Biogeography 17:297–304.

Colyvan, M., M. A. Burgman, C. R. Todd, H. Resit Akçakaya, and C. Boek. 1999. The treatment of uncertainty and the structure of the IUCN threatened species categories. Biological Conservation 89:245–249.

Cowled, B. D., M. G. Garner, K. Negus, and M. P. Ward. 2012. Controlling disease outbreaks in wildlife using limited culling: modelling classical swine fever incursions in wild pigs in Australia. Veterinary Research 43:3.

Crouse, D. T., L. B. Crowder, and H. Caswell. 1987. A stage-based population-model for loggerhead sea turtles and implications for conservation. Ecology 68:1412–1423.

Crowder, L. B., D. T. Crouse, S. S. Heppell, and T. H. Martin. 1994. Predicting the impact of turtle excluder devices on loggerhead sea turtle populations. Ecological Applications 4:437–445.

DeAngelis, D. L., and W. M. Mooij. 2005. Individual-based modeling of ecological and evolutionary processes. Annual Review of Ecology, Evolution, and Systematics 36:147–168.

del Monte-Luna, P., V. Guzman-Hernandez, E. A. Cuevas, F. Arreguin-Sanchez, and D. Lluch-Belda. 2012. Effect of North Atlantic climate variability on hawksbill turtles in the Southern Gulf of Mexico. Journal of Experimental Marine Biology and Ecology 412:103–109.

Dichmont, C. M., E. A. Fulton, R. Gorton, M. Sporcic, L. R. Little, A. E. Punt, N. Dowling, M. Haddon, N. Klaer, and D. C. Smith. 2017. From data rich to data-limited harvest strategies-does more data mean better management? ICES Journal of Marine Science 74:670–686.

Epperly, S., et al. 2002. Analysis of sea turtle bycatch in the commercial shrimp fisheries of Southeast U.S. waters and the Gulf of Mexico. Monograph or Serial Issue. http://www.sefsc.noaa.gov/PDFdocs/memo-490.pdf

Esteban, N., J. A. Mortimer, and G. C. Hays. 2017. How numbers of nesting sea turtles can be overestimated by nearly a factor of two. Proceedings of the Royal Society B 284:20162581.

Fay, G., J. S. Link, S. I. Large, and R. J. Gamble. 2014. Management performance of ecological indicators in the Georges Bank finfish fishery. ICES Journal of Marine Science: Journal du Conseil 72:fsu214.

Finkbeiner, E. M., B. P. Wallace, J. E. Moore, R. L. Lewison, L. B. Crowder, and A. J. Read. 2011. Cumulative estimates of sea turtle bycatch and mortality in USA fisheries between 1990 and 2007. Biological Conservation 144:2719–2727.

Gao, L., and A. Hailu. 2013. Identifying preferred management options: An integrated agent-based recreational fishing simulation model with an AHP-TOPSIS evaluation method. Ecological Modelling 249:75–83.

Gerrodette, T., B. L. Taylor, K. L. Eckert, K. A. Bjorndal, F. A. Abreu-Grobois, and M. Donnelly. 1999. Estimating population size. Page research and management techniques for the conservation of sea turtles. IUCN/SSC Marine Turtle Specialist Group, Washington, D.C., USA.

Girondot, M., A. D. Tucker, P. Rivalan, M. H. Godfrey, and J. Chevalier. 2002. Density-dependent nest destruction and population fluctuations of Guianan leatherback turtles. Animal Conservation 5:75–84.

Godley, B. J., S. Richardson, A. C. Broderick, M. S. Coyne, F. Glen, and G. C. Hays. 2002. Long-term satellite telemetry of the movements and habitat utilisation by green turtles in the Mediterranean. Ecography 25:352–362.

Grüss, A., W. J. Harford, M. J. Schirripa, L. Velez, S. R. Sagarese, Y.-J. Shin, and P. Verley. 2016. Management strategy evaluation using the individual-based, multispecies modeling approach OSMOSE. Ecological Modelling 340:86–105.

Hamann, M., et al. 2010. Global research priorities for sea turtles: informing management and conservation in the 21st century. Endangered Species Research 11:245–269.

Harford, W. J., C. Ton, and E. A. Babcock. 2015. Simulated mark-recovery for spatial assessment of a spiny lobster (*Panulirus argus*) fishery. Fisheries Research 165:42–53.

Hart, K. M., M. M. Lamont, A. R. Sartain, I. Fujisaki, and B. S. Stephens. 2013. Movements and habitat-use of loggerhead sea turtles in the Northern Gulf of Mexico during the reproductive period. PLoS ONE 8:e66921.

Hastings, A. 2001. Transient dynamics and persistence of ecological systems. Ecology Letters 4:215–220.

Hastings, A. 2004. Transients: the key to long-term ecological understanding? Trends in Ecology & Evolution 19:39–45.

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.

Heppell, S. 1998. Application of life-history theory and population model analysis to turtle conservation. Copeia 1998:367–375.

Heppell, S. S., and L. B. Crowder. 1998. Prognostic evaluation of enhancement programs using population models and life history analysis. Bulletin of Marine Science 62:495–507.

Heppell, S. S., L. B. Crowder, and D. T. Crouse. 1996. Models to evaluate headstarting as a management tool for long-lived turtles. Ecological Applications 6:556–565.

IUCN. 2015. The IUCN Red List of Threatened Species, version 2015-4. http://www.iucnredlist.org/

Jackson, A. L., A. C. Broderick, W. J. Fuller, F. Glen, G. D. Ruxton, and B. J. Godley. 2008. Sampling design and its effect on population monitoring: How much monitoring do turtles really need? Biological Conservation 141:2932–2941.

Koons, D. N., J. B. Grand, B. Zinner, and R. F. Rockwell. 2005. Transient population dynamics: Relations to life history and initial population state. Ecological Modelling 185:283–297.

Letcher, B. H., J. A. Priddy, J. R. Walters, and L. B. Crowder. 1998. An individual-based, spatially-explicit simulation model of the population dynamics of the endangered red-cockaded woodpecker, *Picoides borealis*. Biological Conservation 86:1–14.

Limpus, C., and M. Chaloupka. 1997. Nonparametric regression modelling of green sea turtle growth rates (southern Great Barrier Reef). Marine Ecology Progress Series 149:23–34.

Lynch, P. D., K. W. Shertzer, and R. J. Latour. 2012. Performance of methods used to estimate indices of abundance for highly migratory species. Fisheries Research 125–126:27–39.

Magnuson, J. J., K. A. Bjorndal, W. D. DuPaul, G. L. Graham, D. W. Owens, C. H. Peterson, P. C. H. Pritchard, J. I. Richardson, G. E. Saul, and C. W. West. 1990. Decline of the sea turtles. National Academy Press, Washington, D.C., USA.

Mapstone, B. D., L. R. Little, A. E. Punt, C. R. Davies, A. D. M. Smith, F. Pantus, A. D. McDonald, A. J. Williams, and A. Jones. 2008. Management strategy evaluation for line fishing in the Great Barrier Reef: Balancing conservation and multi-

sector fishery objectives RID C-6683-2011. Fisheries Research 94:315–329.

Maxwell, D., and S. Jennings. 2005. Power of monitoring programmes to detect decline and recovery of rare and vulnerable fish. Journal of Applied Ecology 42:25–37.

Mazaris, A., and Y. Matsinos. 2006. An individual based model of sea turtles: Investigating the effect of temporal variability on population dynamics. Ecological Modelling 194:114–124.

Mazaris, A. D., Ø. Fiksen, and Y. G. Matsinos. 2005. Using an individual-based model for assessment of sea turtle population viability. Population Ecology 47:179–191.

Mazaris, A. D., B. Broder, and Y. G. Matsinos. 2006. An individual based model of a sea turtle population to analyze effects of age dependent mortality. Ecological Modelling 198:174–182.

Mazaris, A. D., G. Schofield, C. Gkazinou, V. Almpanidou, and G. C. Hays. 2017. Global sea turtle conservation successes. Science Advances 3:e1600730.

McDonald, A. D., L. R. Little, R. Gray, E. Fulton, K. J. Sainsbury, and V. D. Lyne. 2008. An agent-based modelling approach to evaluation of multiple-use management strategies for coastal marine ecosystems. Mathematics and Computers in Simulation 78:401–411.

Mortimer, J. A., and A. Carr. 1987. Reproduction and migrations of the Ascension Island green turtle (Chelonia mydas). Copeia 1987:103–113.

National Marine Fisheries Service, and U.S. Fish and Wildlife Service. 1991. Recovery plan for U.S. population of Atlantic green turtle. National Marine Fisheries Service, Washington, D.C., USA.

National Marine Fisheries Service, and U.S. Fish and Wildlife Service. 1998. Recovery plan for U.S. pacific populations of the green turtle (Chelonia mydas). National Marine Fisheries Service, Silver Springs, Maryland, USA.

National Research Council. 2010. Assessment of sea-turtle status and trends: integrating demography and abundance. First edition. The National Academies Press, Washington, D.C., USA.

Niethammer, K. R., G. H. Balazs, J. S. Hatfield, G. L. Nakai, and J. L. Megyesu. 1997. Reproductive biology of the green turtle (Chelonia mydas) at Tern Island, French Frigate Shoals, Hawai'i. Pacific Science 51:36–47.

Ocana, M., M. Harfush-Melendez, and S. S. Heppell. 2012. Mass nesting of olive Ridley sea turtles Lepidochelys olivacea at La Escobilla, Mexico: linking nest density and rates of destruction. Endangered Species Research 16:45–54.

Piacenza, S. E., G. H. Balazs, S. K. Hargrove, P. M. Richards, and S. S. Heppell. 2016. Trends and variability in demographic indicators of a recovering population of green sea turtles (Chelonia mydas). Endangered Species Research 31:103–117.

Piacenza, S. E., P. M. Richards, and S. S. Heppell. 2017. An agent-based model to evaluate recovery times and monitoring strategies to increase accuracy of sea turtle population assessments. Ecological Modelling 358:25–39.

Punt, A. E., D. S. Butterworth, C. L. de Moor, J. A. A. De Oliveira, and M. Haddon. 2014. Management strategy evaluation: best practices. Fish and Fisheries 17:330–334.

Red List Standards & Petitions Subcommittee. 1996. Natator depressus. The IUCN Red List of Threatened Species 1996: e.T14363A4435952. http://dx.doi.org/10.2305/IUCN.UK.1996.RLTS.T14363A4435952.en. Downloaded on 18 June 2019.

Richards, P. M., S. P. Epperly, S. S. Heppell, R. T. King, C. R. Sasso, F. Moncada, G. Nodarse, D. J. Shaver, Y. Medina, and J. Zurita. 2011. Sea turtle population estimates incorporating uncertainty: a new approach applied to western North Atlantic loggerheads Caretta caretta. Endangered Species Research 15:151–158.

Saba, V. S., P. Santidrian-Tomillo, R. D. Reina, J. R. Spotila, J. A. Musick, D. A. Evans, and F. V. Paladino. 2007. The effect of the El Nino Southern Oscillation on the reproductive frequency of eastern Pacific leatherback turtles. Journal of Applied Ecology 44:395–404.

Sainsbury, K. J., A. E. Punt, and A. D. M. Smith. 2000. Design of operational management strategies for achieving fishery ecosystem objectives. ICES Journal of Marine Science: Journal du Conseil 57:731–741.

Schroeder, B. A., and S. Murphy. 1999. Population surveys (ground and aerial) on nesting beaches. Pages 45–55 in K. L. Eckert, K. A. Bjorndal, F. A. Abreu-Grobois, and M. Donnelly, editors. Research and management techniques for the conservation of sea turtles. IUCN/SSC Marine Turtle Specialist Group, Washington, D.C., USA.

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

Singh, N. J., and E. J. Milner-Gulland. 2011. Conserving a moving target: planning protection for a migratory species as its distribution changes. Journal of Applied Ecology 48:35–46.

Smith, A., K. Sainsbury, and R. Stevens. 1999. Implementing effective fisheries-management systems – management strategy evaluation and the Australian partnership approach. ICES Journal of Marine Science 56:967–979.

Snover, M. L., and S. S. Heppell. 2009. Application of diffusion approximation for risk assessments of sea turtle populations. Ecological Applications 19:774–785.

Solow, A. R., K. A. Bjorndal, and A. B. Bolten. 2002. Annual variation in nesting numbers of marine turtles: the effect of sea surface temperature on re-migration intervals. Ecology Letters 5:742–746.

Stokes, K. L., W. J. Fuller, F. Glen, B. J. Godley, D. J. Hodgson, K. A. Rhodes, R. T. E. Snape, and A. C. Broderick. 2014. Detecting green shoots of recovery: the importance of long-term individual-based monitoring of marine turtles. Animal Conservation 17:593–602.

Taylor, B. L., and T. Gerrodette. 1993. The uses of statistical power in conservation biology: the vaquita and northern spotted owl. Conservation Biology 7:489–500.

Taylor, B. L., M. Martinez, T. Gerrodette, J. Barlow, and Y. N. Hrovat. 2007. Lessons from monitoring trends in abundance of marine mammals. Marine Mammal Science 23:157–175.

Thomas, F. M., P. A. Vesk, and C. E. Hauser. 2018. A field ecologist's adventures in the virtual world: using simulations to design data collection for complex models. Ecological Applications 28:2130–2141.

Tiwari, M., K. A. Bjorndal, A. B. Bolten, and B. M. Bolker. 2006. Evaluation of density-dependent processes and green turtle Chelonia mydas hatchling production at Tortuguero, Costa Rica. Marine Ecology Progress Series 326:283–293.

Tiwari, M., G. H. Balazs, and S. Hargrove. 2010. Estimating carrying capacity at the green turtle nesting beach of East Island, French Frigate Shoals. Marine Ecology Progress Series 419:289–294.

Trujillo, A., and H. Thurman. 2008. Essentials of oceanography. Ninth edition. Pearson Prentice Hall, Upper Saddle River, New Jersey, USA.

Tucker, A. D. 2010. Nest site fidelity and clutch frequency of loggerhead turtles are better elucidated by satellite telemetry than by nocturnal tagging efforts: Implications for stock estimation. Journal of Experimental Marine Biology and Ecology 383:48–55.

Tucker, A. D., and N. B. Frazer. 1991. Reproductive variation in leatherback turtles, Dermochelys coriacea, at Culebra

National Wildlife Refuge, Puerto Rico. Herpetologica 47: 115–124.

Turtle Expert Working Group. 2009. An Assessment of the Loggerhead Turtle Population in the Western North Atlantic Ocean. NOAA Technical Memorandum NMFS-SEFSC. NOAA, Silver Spring, Maryland, USA.

Van Houtan, K. S., and J. N. Kittinger. 2014. Historical commercial exploitation and the current status of Hawaiian green turtles. Biological Conservation 170:20–27.

Van Houtan, K. S., S. K. Hargrove, and G. H. Balazs. 2014. Modeling sea turtle maturity age from partial life history records. Pacific Science 68:465–477.

Wallace, B. P., et al. 2010. Regional management units for marine turtles: a novel framework for prioritizing conservation and research across multiple scales. PLoS ONE 5.

Wallace, B. P., et al. 2011. Global conservation priorities for marine turtles. PLoS ONE 6:e24510.

Warden, M. L., H. L. Haas, P. M. Richards, K. A. Rose, and J. M. Hatch. 2017. Monitoring trends in sea turtle populations: walk or fly? Endangered Species Research 34:323–337.

White, J. W., L. W. Botsford, A. Hastings, M. L. Baskett, D. M. Kaplan, and L. A. K. Barnett. 2013. Transient responses of fished populations to marine reserve establishment. Conservation Letters 6:180–191.

Whiting, A. U., M. Chaloupka, and C. J. Limpus. 2013. Comparing sampling effort and errors in abundance estimates between short and protracted nesting seasons for sea turtles. Journal of Experimental Marine Biology and Ecology 449: 165–170.

Wilensky, U. 1999. NetLogo. Center for Connected Learning and Computer-Based Modeling, Evanston, Illinois, USA.

Winship, A. J., M. R. O'Farrell, and M. S. Mohr. 2013. Management strategy evaluation applied to the conservation of an endangered population subject to incidental take. Biological Conservation 158:155–166.

Witzell, W. N. 1994. The origin, evolution and demise of the U.S. sea turtle fisheries. Marine Fisheries Review 56:8–23.

Zug, G., G. Balazs, and J. Wetherall. 2002. Age and growth of Hawaiian green sea turtle (*Chelonia mydas*): an analysis based on skeletochronology. Fishery Bulletin 100:117–127.