Two decades of community-based conservation yield valuable insights into marine turtle nesting ecology

CASPER H. VAN DE GEER, ANNETTE C. BRODERICK, MATT I.D. CARTER,

ATHUMAN ABDALLAH IREI, FIKIRI KEA KIPONDA, JOSEPH KIPTUM,

JOE NGUNU, MOHAMED OMAR, NICOLA PARAZZI, HANNAH SAWYER-KERR,

SAM B.WEBER, RICARDO ZANRE and BRENDAN J. GODLEY

SUPPLEMENTARY MATERIAL 1

STATISTICAL ANALYSES

Statistical analyses were carried out in R version 4.1.2 (R Core Team, 2022) using RStudio (RStudio Team, 2022), with a significance level of $\alpha = 0.05$. Graphing was carried out using the 'ggplot2' package (Wickham, 2016) and data manipulation was undertaken using 'tidyr' (Wickham & Girlich, 2022), 'dplyr' (Wickham et al., 2022), and 'lubridate' (Grolemund & Wickham, 2011).

Proportion clutches laid per month

To determine the monthly trend in nesting activity per season for green turtles, the proportion of clutches laid per month over the 20 seasons was calculated. Due to the non-normally distributed monthly proportion data, a nonparametric bootstrap resampling method with 1000 iterations was used to calculate the means and 95% confidence intervals (Davison & Hinkley, 1997) using the 'boot' package (Canty & Ripley, 2021). Limited data was available for olive ridley turtles, so the cumulative monthly clutch counts were plotted.

Median Nesting Date

Temporal shifts in nesting seasonality were investigated using the median nesting date, since the mean may be disproportionately influenced by outliers (i.e., clutches laid unusually early or late in the season). The start and end dates of the principal nesting season were defined as the 2.5 and 97.5% quantiles, respectively. Linear regression was used to test whether the median nesting date had changed over the monitored period.

Clutch trend

The trend in the number of clutches laid per season in Watamu Marine National Park was analysed using a Generalized Additive Model (GAM) fit using the 'mgcv' package (Wood, 2017). Only data from Watamu Marine National Park were used because the monitoring effort on this beach was consistent over the 20 seasons. The GAM was fitted with a negative binomial error structure and used a thin plate spline smooth to model any non-linearity in the temporal trend. Inspection of partial autocorrelation plots detected significant temporal autocorrelation at lag 1 (typical in marine turtle nesting time series). Therefore, a first order autocorrelation structure (AR1) was also added to the model using the function 'gamm' from package 'mgcv'. Model fit was checked with the 'DHARMa' package (Hartig, 2022) and the autocorrelation structure was found to improve the model (versus a GAM with no autocorrelation structure).

Female size

The curved carapace length (CCL) of nesting green turtles at first capture was used to test for a trend across seasons. Simple linear regression was deemed appropriate since the errors were normally distributed and the observations were independent.

Internesting and remigration intervals

Intervals between observed nesting events, i.e., where the nest could be assigned to a tagged individual, were calculated in days. Individuals returning in subsequent seasons could be identified by flipper tags and their remigration intervals were calculated in days. These were then converted into years by dividing the remigration interval by 365.25 and rounding up. Green turtles were the only species for which data were available that allowed for intervals to be calculated.

Clutch Frequency

The clutch frequency is defined as the number of clutches a female will lay within a nesting season. Four metrics of the clutch frequency were calculated. The first was the observed clutch frequency (OCF), defined as the cumulative number of observed nesting events per female per season (Frazer & Richardson, 1985; Johnson & Ehrhart, 1996). However, not every nesting event by every female was observed, and the OCF is therefore likely to be an underestimate of the true clutch frequency. The remaining three metrics are estimated clutch frequencies (ECF) based on different methods. ECF1 adjusts the OCF with an estimated number of nesting events that were not observed (Frazer & Richardson, 1985; Johnson & Ehrhart, 1996; Broderick et al., 2002). If the interval between observed nesting events was longer than the estimated maximum interval between successive nesting events (set to 18 days based on Fig. 5B), a clutch was added to that individual's estimated clutch frequency for that season. For longer internesting intervals, multiple nesting events had been missed and added to the estimated clutch frequency. ECF2 is a subset of the total OCF, selecting only seasons where >70% of the nesting events were observed because these OCF values will be closer to the true clutch frequency. ECF3 uses a combination of ECF1 and ECF2, whereby the OCF in seasons where >70% of nesting events were observed were augmented with presumed missed clutches as determined by examination of the internesting intervals.

The OCF and ECFs had non-normally distributed errors and contained pseudoreplicated values. Hence the mean OCF and ECFs were calculated with a Generalized Linear Mixed-effects Model (GLMM) null model, using the 'lme4' package (Bates et al., 2014). A Poisson error structure (with log link function) was specified, and the individual females used as the random effect. Tests for overdispersion using the 'overdisp_fun' by Bolker (2022) revealed that this was not an issue. The four models were then passed to the 'emmeans' package, which provided the back-transformed means with their 95% confidence intervals (Lenth, 2022).

Clutch distribution in Watamu Marine National Park

Monitoring efforts by Local Ocean Conservation in Watamu Marine National Park have historically recorded the name or number of the plot of land bordering the beach as a reference for nesting site. There are 52 such plots along the Watamu Marine National Park beach where turtles have nested or potentially could at one point since monitoring began. The beach frontage of these plots varies, ranging from 61 meters to 262 meters. When GPS devices were available to Local Ocean Conservation, the coordinates of the nesting site were recorded but these data are intermittent in the first 10 years and exist for 60% (n = 516) of all the nests recorded in Watamu Marine National Park. By using beach plots as a proxy for the nesting site, we were able to incorporate all 855 clutches into the analysis and extended the temporal coverage. To quantify the relative position of plots along the beach, we used the centroid of each plot's beach frontage and calculated the distance to the centroid of the northern most plot (Blue Bay Resort), using the 'geodist' package (Padgham & Sumner, 2021). The number of clutches laid in each plot was summed into temporal bins, which were arbitrarily assigned to be 5 years (n = 4). Binned count data were then converted into average nesting density per km per time bin by dividing by the length of each respective plot. A Generalized Additive Model (GAM) with a Tweedie error distribution and a logistic link function was used to model these nesting density data in Watamu Marine National Park (Wood, 2017). The most suitable way to combine the spatial (distance) and time (season bins) predictors into the model was investigated by comparing a model with an additive structure to a model that used a tensor. The smoothing bases were set to thin plate regression splines and cubic regression splines for distance and season bins, respectively. The lower AIC value of the model with the tensor indicated the better fit. Models were checked for residual autocorrelation (with 'mgcv') and overdispersion (with 'DHARMa'); no signs of either were found.

Nest distribution patterns through time were explored further by comparing the total proportion of clutches laid along the northern and southern halves of Watamu Marine National Park for each season. The midway point is approximately in the middle of plot 24 so the clutches laid within this plot each season were equally divided between the northern and southern halves. A GAM with a binomial error structure was used to model the trend, which suggested a linear relationship between the shift from the northern half to the south through time. Therefore, the trend was modelled with a Generalised Linear Model with a binomial error structure and the effect of "season" was tested using Analysis of Deviance (Zuur et al., 2009)

Clutch hatching success

The hatching success of a clutch was measured as the proportion of hatched eggs, as per Miller (1999). Live and dead hatchlings found in the nest during excavation were considered to have successfully hatched. Clutches destroyed by illegal take were omitted from the analysis (n = 3) but clutches that failed otherwise (hatching success <0.05) were included (n = 31). Excavation data were missing for 35 nests. Analyses outlined here focused on green turtles, due to limited sample sizes for the other three species.

A GLMM was used to test for differences between the hatching success rates of the four different clutch treatments, namely the combinations of whether the clutch was laid inside or beyond Watamu Marine National Park and whether it was left *in-situ* or relocated. The choice of a GLMM was based the nature of the data (proportions) and that there was pseudoreplication caused by individuals nesting multiple times and possibly in multiple treatment levels (Zuur et al., 2009). Using the 'lme4' package (Bates et al., 2014), the GLMM was set up with a binomial error structure and a logit link function. Treatment levels (n = 4) were specified as the fixed effects and season was included as a random effect to account for inter-seasonal variation.

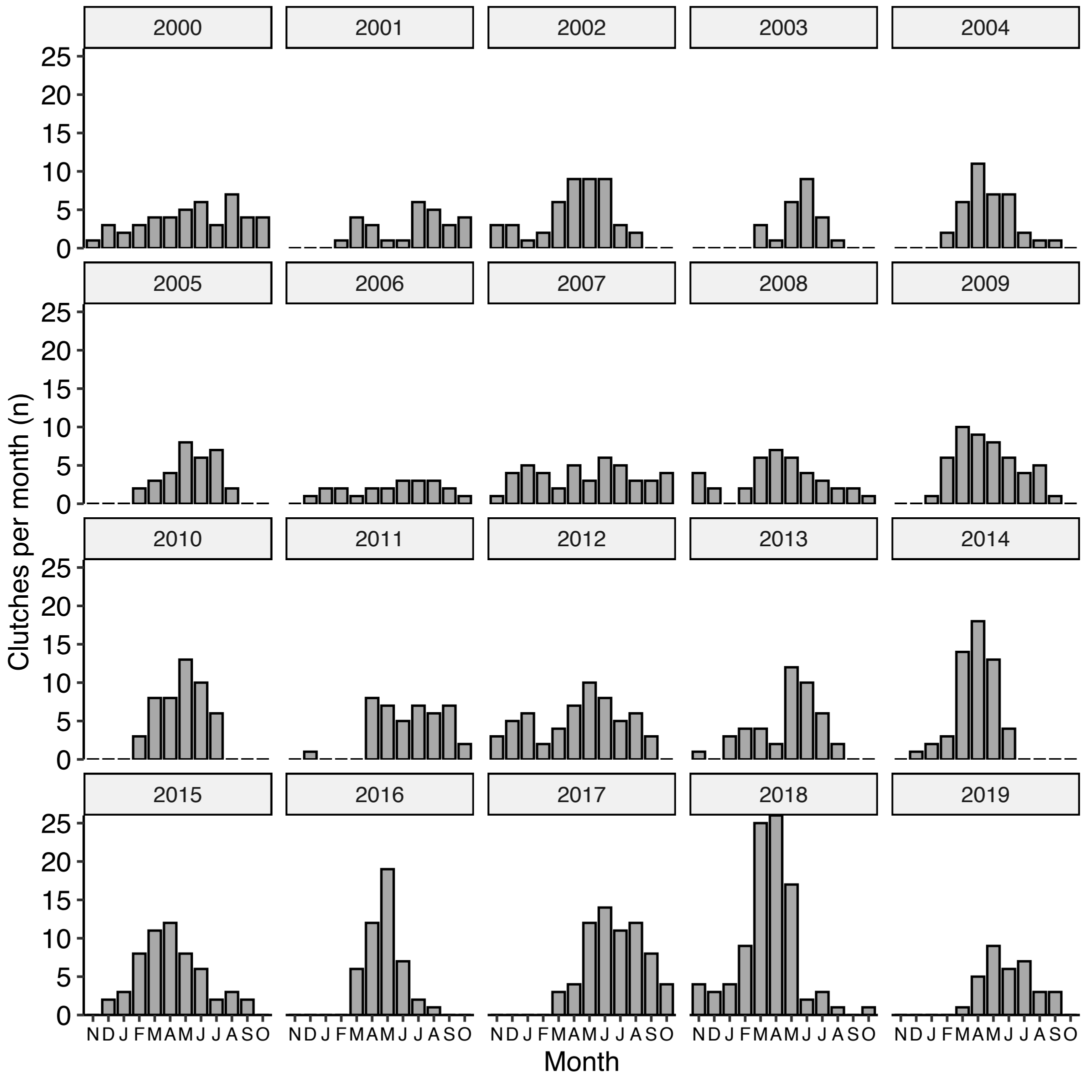
Overdispersion was detected using the 'overdisp_fun' by Bolker et al. (2022), which was dealt with by adding an observation level random effect (OLRE) to the model (Harrison, 2014). Analysis of variance testing was used to compare how the full versus reduced (fixed effect omitted) models fitted the data (Crawley, 2012). Due to the unbalanced nature of the data, the estimated marginal means, which are based on the model, and their respective 95% confidence intervals were calculated using the 'emmeans' package (Lenth, 2022). Functions in the same package were used to perform post-hoc testing by making pairwise comparisons of the means, which provided Tukey-adjusted p-values.

SUPPLEMENTARY MATERIAL 2

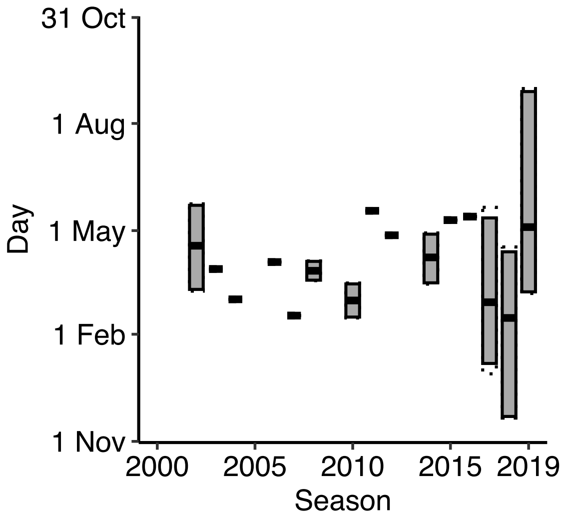
SUPPLEMENTARY RESULTS

SUPPLEMENTARY TABLE 1 Clutches laid within and beyond the boundaries of Watamu Marine National Park (WMNP) from 1st of Nov, 2000, to 31st of Oct, 2020.

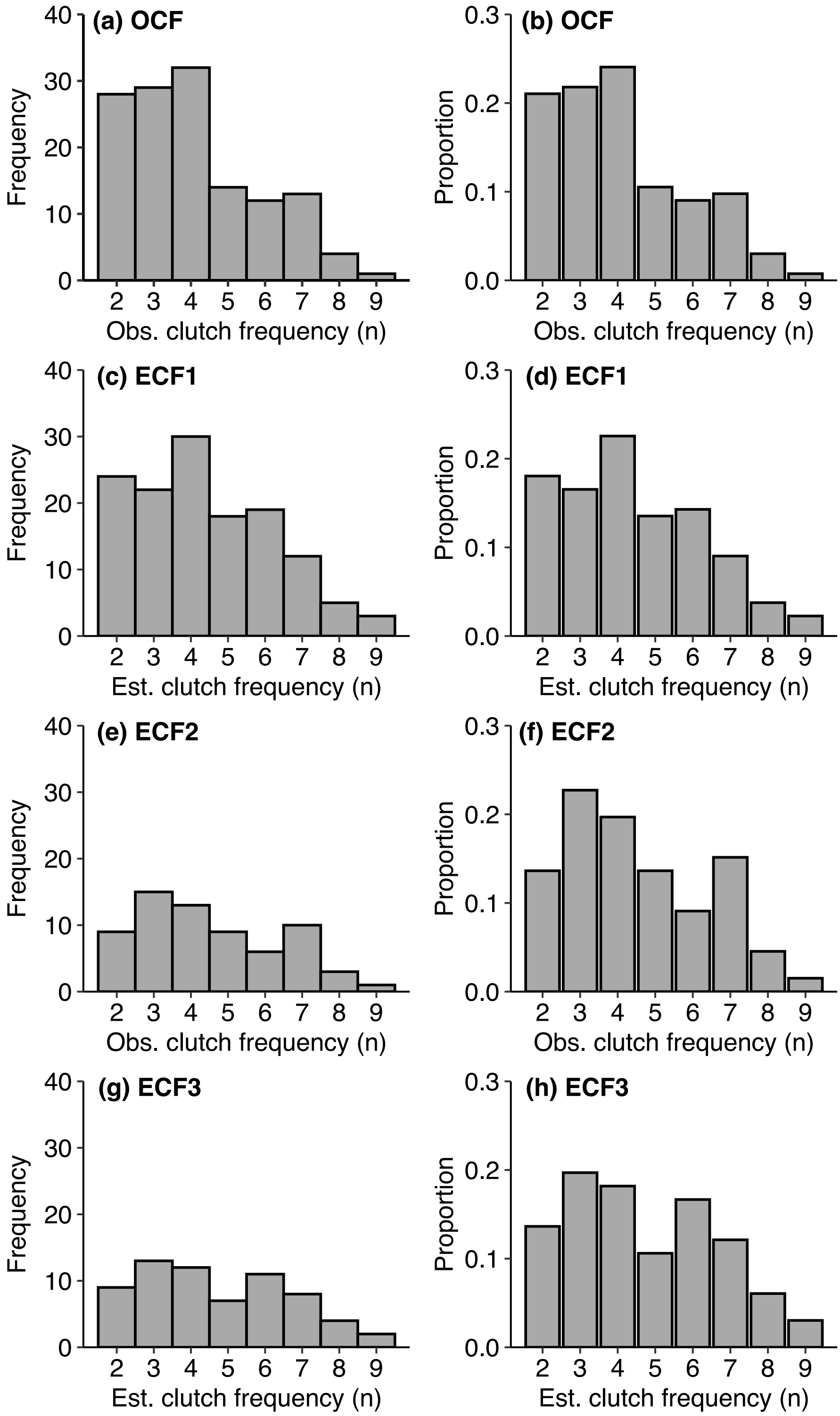
Species	Within WMNP	Beyond WMNP	*Total*
Green	819	101	*920*
Olive ridley	34	7	*41*
Hawksbill	1	1	*2*
Leatherback	1	0	*1*
Total	*855*	*109*	*964*



SUPPLEMENTARY FIG. 1 Green turtle clutches laid per month in Watamu over the 20 seasons. Based on these data, the start of a nesting season was assigned to be the 1st of November and end on the 31st of October of the following calendar year.



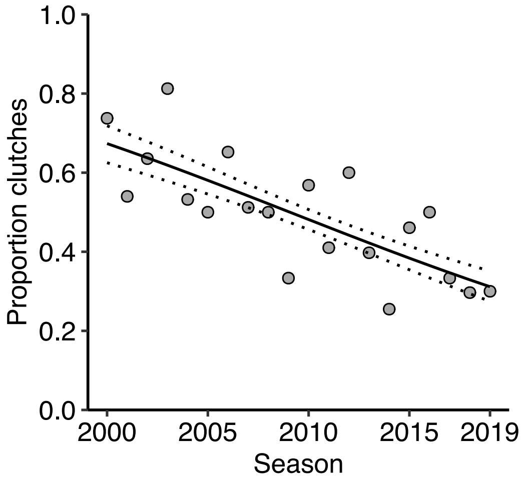
SUPPLEMENTARY FIG. 2 Olive ridley nesting seasons from 2000 to 2019, based on clutches laid in Watamu (n = 41). Dotted range indicates the total span, the grey range indicates the 95% quantile, and the median nesting date is indicated with the black marker. Only one clutch was laid in seasons without a range. Days start on day 1 of the nesting season, which is November 1st.



SUPPLEMENTARY FIG. 3 Observed and estimated clutch frequencies for green turtles in Watamu. (a & b) Observed clutch frequencies (OCF) based on field observations. (c & d) Method 1: estimated clutch frequencies (ECF) adjusted by adding clutches that, based on internesting intervals, were suspected to be missed by the monitors. (e & f) Method 2: clutch frequencies for seasons where >70% of the nesting events were observed and allocated to individual females. (g & h) Method 3: estimated clutch frequency in seasons where >70% of the nesting events were observed, combined with adjustment by adding clutches suspected to have been missed by the monitors based on the internesting intervals. (a, c, e & g) are clutch frequencies in absolute values and (b, d, f & h) are the same data presented as proportions.

SUPPLEMENTARY TABLE 2 Nesting seasons through time for individual green turtles, indicated by dark grey squares. Numbers in the dark grey squares indicate the number of clutches laid during that season.

	Season																			
Turtle ID	00	01	02	03	04	05	06	07	08	09	10	11	12	13	14	15	16	17	18	19
1							3					7						7		
6							1						6			4				
7								4				5					6			
11	3			5				2				3								
12								3										3		
15									7					9						
18	2				7															
20	2			3																
23	2					2														
25	6			7																
27		3			4															
28	4							7												
31		3				6				3		3								
37			7					8					3						3	
41			6			5			6											
44			7					8			6			7			4			3
51			7			1														

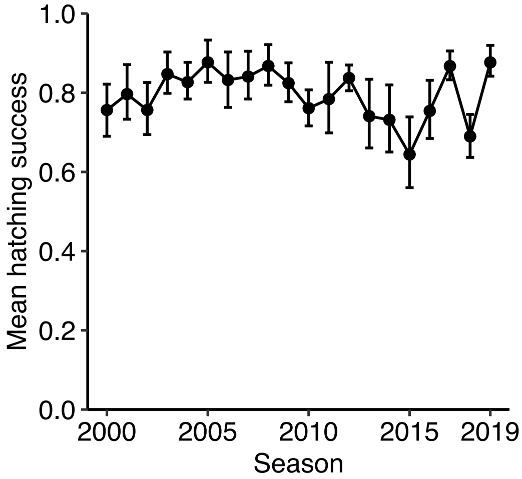


SUPPLEMENTARY FIG. 4 – Proportion of clutches that were laid in the northern half of Watamu Marine National Park per season, with average trend and 95% confidence intervals.

SUPPLEMENTARY TABLE 3 Green turtle clutch hatching success. Estimated marginal mean hatching proportion, with 95% confidence interval, sample size per treatment group, and grouping according to post-hoc Tukey tests.

WMNP	Treatment	Mean	95% CI	n	Grouping
Inside	*In-situ*	0.89	0.86, 0.91	451	a
Inside	Relocated	0.82	0.77, 0.85	335	b
Outside	*In-situ*	0.86	0.60, 0.96	6	ab
Outside	Relocated	0.81	0.74, 0.87	94	b

SUPPLEMENTARY FIG. 5 Mean hatching success of green turtle clutches per season (n = 916), with 95% confidence intervals (computed using a basic nonparametric bootstrap with 1000 iterations).



SUPPLEMENTARY TABLE 4 Estimated number of green turtle females nesting in Watamu Marine National Park per season, for the most recent 5 seasons. Estimates were calculated according to the different clutch frequencies from this study (OCF: 4.1, ECF1: 4.4, ECF2: 4.5, ECF3: 4.7) and from satellite tracking performed in the Chagos Archipelago (6.0; Esteban et al., 2017).

		Nesting females per season (n)				
Season	Clutches (n)	OCF	ECF1	ECF2	ECF3	Sat tag
2015	51	12	12	11	11	9
2016	41	10	9	9	9	7
2017	51	12	12	11	11	9
2018	81	20	18	18	17	14
2019	31	8	7	7	7	5

SUPPLEMENTARY TABLE 5 Estimated total green turtle nesting population of Watamu Marine National Park, according to the different clutch frequencies (as per Supplementary Table 3). Each estimate is the sum of the nesting females per season (see Supplementary Table 4) for the indicated seasons.

	Total nesting population (n)				
Seasons	OCF	ECF1	ECF2	ECF3	Sat tag
2015 - 2017	35	33	32	30	24
2016 - 2018	42	39	38	37	29
2017 - 2019	40	37	36	35	27

References

BATES, D., MÄCHLER, M., BOLKER, B. & WALKER, S. (2014) Fitting Linear Mixed-Effects Models using lme4. *arXiv:1406.5823 [stat]*.

BOLKER, B. (2022) GLMM FAQ. https://bbolker.github.io/mixedmodels-misc/glmmFAQ.html.

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, 227–235.

CANTY, A. & RIPLEY, B.D. (2021) boot: Bootstrap R (S-Plus) Functions.

CRAWLEY, M.J. (2012) The R book. John Wiley & Sons.

DAVISON, A.C. & HINKLEY, D.V. (1997) Bootstrap Methods and Their Applications. Cambridge University Press, Cambridge.

ESTEBAN, N., MORTIMER, J.A. & HAYS, G.C. (2017) How numbers of nesting sea turtles can be overestimated by nearly a factor of two. *Proceedings of the Royal Society B: Biological Sciences*, 284, 20162581.

FRAZER, N.B. & RICHARDSON, J.I. (1985) Annual variation in clutch size and frequency for Loggerhead turtles, *Caretta caretta*, nesting at Little Cumberland Island, Georgia, USA. *Herpetologica*, 41, 246–251.

GROLEMUND, G. & WICKHAM, H. (2011) Dates and Times Made Easy with lubridate. *Journal of Statistical Software*, 40, 1–25.

HARRISON, X.A. (2014) Using observation-level random effects to model overdispersion in count data in ecology and evolution. *PeerJ*, 2, e616.

HARTIG, F. (2022) DHARMa: Residual Diagnostics for Hierarchical (Multi-Level / Mixed) Regression Models. http://florianhartig.github.io/DHARMa/.

JOHNSON, S.A. & EHRHART, L.M. (1996) Reproductive Ecology of the Florida Green Turtle: Clutch Frequency. *Journal of Herpetology*, 30, 407.

LENTH, R.V. (2022) emmeans: Estimated Marginal Means, aka Least-Squares Means. https://CRAN.R-project.org/package=emmeans.

MILLER, J.D. (1999) Determining Clutch Size and Hatching Success. In *Research and Management Techniques for the Conservation of Sea Turtles* (eds K.L. Eckert, K.A. Bjorndal, F.A. Abreu-Grobois & M. Donnelly), pp. 1–6. Washington, D.C., USA.

PADGHAM, M. & SUMNER, M.D. (2021) geodist: Fast, Dependency-Free Geodesic Distance Calculations. https://CRAN.R-project.org/package=geodist.

R CORE TEAM (2022) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

RSTUDIO TEAM (2022) RStudio: Integrated Development for R. RStudio, PBC, Boston, MA. https://www.rstudio.com/.

WICKHAM, H. (2016) ggplot2: Elegant Graphics for Data Analysis2nd ed. 2016. Springer International Publishing : Imprint: Springer, Cham.

WICKHAM, H., FRANÇOIS, R., HENRY, L. & MÜLLER, K. (2022) dplyr: A Grammar of Data Manipulation. https://CRAN.R-project.org/package=dplyr.

WICKHAM, H. & GIRLICH, M. (2022) tidyr: Tidy Messy Data. https://CRAN.R-project.org/package=tidyr.

WOOD, S.N. (2017) Generalized additive models: An introduction with R (2nd edition). Chapman and Hall/CRC., Boca Raton, FL, USA.

ZUUR, A.F., IENO, E.N., WALKER, N.J., SAVELIEV, A.A., SMITH, G.M., & OTHERS (2009) Mixed effects models and extensions in ecology with R, 1st edition. Springer, New York, NY, USA.