Recent advances on the estimation of the thermal reaction norm for sex ratios

F Alberto Abreu-Grobois, B Alejandra Morales-Mérida, Catherine E Hart, Jean-Michel Guillon, Matthew H Godfrey, Erik Navarro, Marc Girondot

Temperature-dependent sex determination, or TSD, is a widespread phenomenon in reptiles. The shape of the relationship between constant incubation temperature and sex ratio defines the TSD pattern. The TSD pattern is considered a life-history parameter important for conservation because the wider the range of temperatures producing both sexes, the more resilient the species is to climate change impacts. We review the different published equations and methodologies that have been used to model TSD patterns. We describe a new flexible model that allows for an asymmetrical pattern around the pivotal temperature, which is the constant temperature producing both sexes in equal proportions. We show that Metropolis-Hastings with Markov chain produced by a Monte Carlo process has many advantages compared to maximum likelihood and should be preferred. Finally, we apply the models to results from incubation experiments using eggs from the marine turtle *Lepidochelys olivacea* originating in Northeast Indian, East Pacific, and West Atlantic Regional Management Units (RMUs) and find large differences in pivotal temperatures but not in transitional range of temperatures.
Recent advances on the estimation of the thermal reaction norm for sex ratios

Short title: TSD pattern estimation

F. Alberto Abreu-Grobois¹, B. Alejandra Morales-Mérida²,³,⁴, Catherine E. Hart⁵,⁶, Jean-Michel Guillon², Matthew H. Godfrey⁷,⁸,⁹, Erik Navarro¹, Marc Girondot²

1 - Laboratorio de Genética, Unidad Académica Mazatlán, Instituto de Ciencias del Mar y Limnología, Universidad Nacional Autónoma de México, Mazatlán, Sinaloa 82000, Mexico.
2 - Laboratoire Écologie, Systématique, Évolution, Université Paris-Sud, AgroParisTech, CNRS, Université Paris Saclay, 91405 Orsay, France
3 - Facultad de Ciencias Químicas y Farmacia. Universidad de San Carlos de Guatemala (USAC).
4 - Doctorado en Ciencias Naturales para el Desarrollo (DOCINADE), Instituto Tecnológico de Costa Rica (TEC), Universidad Nacional (UNA), Universidad Estatal a Distancia (UNED), Costa Rica.
5 - Grupo Tortuguero de las Californias A.C., La Paz, Baja California Sur, Mexico.
6 - Investigación, Capacitación y Soluciones Ambientales y Sociales A.C. Tepic, Nayarit, Mexico.
7 - North Carolina Wildlife Resources Commission, Beaufort, NC 28516, USA.
8 - Duke Marine Laboratory, Nicholas School of the Environment, Duke University, Beaufort, NC 28516
9 - Department of Clinical Sciences, College of Veterinary Medicine, North Carolina State University, Raleigh, NC 27607

Corresponding author: Marc Girondot

e-mail: marc.girondot@u-psud.fr
Abstract

Temperature-dependent sex determination, or TSD, is a widespread phenomenon in reptiles. The shape of the relationship between constant incubation temperature and sex ratio defines the TSD pattern. The TSD pattern is considered a life-history parameter important for conservation because the wider the range of temperatures producing both sexes, the more resilient the species is to climate change impacts. We review the different published equations and methodologies that have been used to model TSD patterns. We describe a new flexible model that allows for an asymmetrical pattern around the pivotal temperature, which is the constant temperature producing both sexes in equal proportions. We show that Metropolis-Hastings with Markov chain produced by a Monte Carlo process has many advantages compared to maximum likelihood and should be preferred. Finally, we apply the models to results from incubation experiments using eggs from the marine turtle Lepidochelys olivacea originating in Northeast Indian, East Pacific, and West Atlantic Regional Management Units (RMUs) and find large differences in pivotal temperatures but not in transitional range of temperatures.

Keywords: Temperature-dependent sex determination, sex ratio, TSD, Reaction norm, Marine turtles, Incubation, Eggs, Olive ridley, Lepidochelys olivacea.
Introduction

Sex determination is the biological process whereby an embryo can become male or female. Temperature-dependent sex determination (TSD), a special case of environmental sex determination (ESD), is widespread in animal kingdom (Korpelainen 1990) and is frequent in reptiles: all crocodilians have TSD, as do many turtles and some lepidosaurians (Valenzuela 2004). In this system, the sexual phenotype of the embryo is defined by the temperature of the incubation occurring during a part of development termed the thermosensitive period (TSP) (Girondot et al. 2018b).

In oviparous reptiles, three patterns of TSD have been described, according to the changes in sex ratios as a function of different constant incubation temperatures (Bull 1983; Ewert et al. 1994; Lang & Andrews 1994). In the TSD Ia pattern, present in some turtles, low temperatures produce males and high temperatures produce females. The opposite occurs in some lizards (TSD Ib pattern). In the TSD II pattern, present in crocodile species, some turtles and some lizards, females are produced at low and high temperatures and more males are produced at intermediate temperatures. Temperature-dependent sex determination can be described as the reaction norm of the resulting sex ratio of embryos incubated at a range of constant temperatures. By definition, a reaction norm describes the pattern of the variation in phenotypic expression of a single genotype across a range of environments (Lewontin 2000).

The mathematical relationship between incubation temperature and sex (or more specifically, the sex ratios produced by a suite of constant incubation temperatures), referred to as “thermal reaction norm”, is commonly characterized by two parameters: (1) the pivotal temperature (P), which is the constant temperature at which both sexes are produced in equal proportions (sex ratio = 1:1), and (2) the transitional range of temperatures (TRT), which is the range of constant
temperatures that yields both sexes in variable proportions (Mrosovsky & Pieau 1991). Note that there may be two values of $P$ and $TRT$ when considering the TSD II pattern.

The correct description of the thermal reaction norm for sex ratios (hereafter named TSD pattern) is not merely a game for biostatisticians. Indeed, variation in the TSD pattern can have profound implications for the conservation of TSD species, particularly in a world affected by climate change. For example, species with a greater TRT should be more likely to evolve in response to new thermal conditions, thus putting them at lower risk to global change (Hulin et al. 2009; Hulin et al. 2008). When within- and among-population variation in the TSD patterns of 12 populations of painted turtles (*Chrysemys picta*) was studied, among-population variation in pivotal temperature could not be explained by geography or local thermal conditions, but the TRT was wider at lower latitudes, suggesting responsiveness to local incubation conditions (Carter et al. 2019). These results indicate that variation in TSD patterns among populations is not an artifact of incubation at constant temperatures and can provide insight into the ecology and evolution of temperature-dependent sex determination.

However, proper statistical tools are needed for robust analyses of the TSD patterns and for identifying the characteristics of interest. Although detailed methods for this have been developed and are freely available (Girondot 2019a), their application has remained challenging, hampering authors with empirical data from fully capitalizing their importance. Thus, the goal of this study is to provide a step-by-step workflow on how to analyze TSD patterns, focusing on the TSD Ia pattern, which is the most common TSD pattern found in turtles. We describe the most advanced statistical models to analyze the thermal reaction norm for the sex ratios produced by constant-temperature incubations, using data from published and unpublished research on the olive ridley marine turtle, *Lepidochelys olivacea*, as a test case. We choose this species for
several reasons. It has a worldwide distribution, it is classified as Vulnerable by IUCN (Abreu-
Grobois & Plotkin 2008), and is amply subjected to egg protection in beach hatcheries as a
management strategy. Providing conservationists with adequate analytic tools to evaluate the sex
ratios of resulting hatchlings (e.g. Dutton et al. 1985) is critical to reliably monitor the effect of
temperature managements required to counter climate warming. Furthermore, with a recent study
indicating contrasting embryonic responses to incubation temperature from two rookeries
(Mexico and Costa Rica) within the same Regional Management Unit (RMU) (Morales Mérida
et al. 2015), analyses of TSD patterns is required to verify if regional differences exist.

Materials and methods

Biological sources of the data

Data from different studies that incubated eggs at constant temperatures were extracted from
publications (available in the “databaseTSD” file, as part of the R package *embryogrowth*
(Girondot 2019a). New data from our own unpublished studies (Navarro Sánchez 2015) were
also included (Table 1).

The version of the database available in the package can be obtained using the R command
DatabaseTSD$Version[1]. The version used here (from *embryogrowth* ver. 7.6.7) is the
2019-11-19 version with 670 records for 25 species or subspecies.

*Lepidochelys olivacea* data including geographic origin, incubation temperatures and their
amplitude were retrieved from this file. The Regional Management Units (RMUs) of olive ridley
sea turtles as defined in Wallace et al. (2010) were inferred from the geographic origins.

Whenever available, a temperature correction factor (difference between the measured incubator
temperature and true internal egg temperature) was also obtained from the database. This
correction factor has been reported in some publications (e.g., Godfrey & Mrosovsky 2006) and
has been shown to be relevant particularly when the substrate contains humidity (Tezak et al.
2018). However, as few papers measure or report this parameter, it can be ignored when
comparing studies where some, or all, do not report it. An alternative is to use an average
correction factor based on all studies, but doing so will not contribute statistical information and
will only shift fitted $P$ values proportional to the correction factor. In the current analysis, the
correction factor is not used and we used only sex ratio data from eggs incubated in temperature-
regulated chambers. Incubation temperature Amplitude and
2nd Third Incubation temperature Amplitude columns from database TSD
refer to a measure of the variability of temperatures during the whole experimental incubation
and the middle-third of the incubation, respectively. Mrosovsky & Pieau (1991) define the
thermosensitive period (TSP) for sex determination as the interval of time when change of
incubation temperature results in a change in resultant sex ratio. The TSP begins with the
formation of the genital ridge at stage 21 (sensu Miller 1985), which corresponds to the onset of
the gonad formation, and ends at stage 26 (sensu Miller 1985), when the gonadal formation is
nearly completed. The TSP occurs during the middle-third of the incubation period when
incubation temperature is constant. When incubation temperature fluctuates during development,
the TSP can shift, and the exact delimitation of TSP must be inferred by taking into account the
thermal reaction norm for growth of embryo (Girondot & Kaska 2014). As a consequence, under
naturally variable incubation conditions, the TSP is not exactly located at the middle-third of the
incubation period (Girondot et al. 2018b). For example, if the incubation temperature at the
beginning of incubation is low, embryos will not grow until the temperature increases.
Subsequently, the growth of the embryo will start, and the TSP will be shifted towards the end of
the incubation, in some cases even after the middle-third of the incubation period. We recommend excluding incubation data from analysis if the recorded temperatures exhibit an amplitude > 2 °C. Indeed, even short daily bursts of high temperatures during egg incubation can lead to an abnormally high proportion of females, as compared to the expected sex ratio based on the average temperature (Georges 1989; Georges et al. 1994). Because egg incubation in natural conditions can expose eggs to wide daily fluctuations of temperature, especially for relatively shallow nests laid by the olive ridley marine turtle, field collected data should not be used in this analysis. Indeed, by definition, the TSD pattern is defined by data collected from constant temperature incubations, thus the TSP is assumed to occur during the middle third of incubation without taking into account thermal reaction norm of embryonic growth. Overall, the use of mean incubation temperatures obtained from nests in field conditions to feed a TSD pattern model has been shown to produce inaccurate results (Fuentes et al. 2017).

A further consideration relates to the difference among the resolution, accuracy, and uncertainty for temperature data loggers. Resolution refers to the data logger’s level of specificity for temperature in its memory. For example, a resolution of 0.5 °C indicates that temperatures will be recorded and reported in 0.5 °C bins, even if the electronic chips can internally read temperatures with a better resolution. Accuracy is represented in the logger’s technical datasheet as a range (±x °C), with x representing how close an individual recorded data point is to the true value. The uncertainty is a measure of the quality of the data logger temperature recordings, considering the accuracy, the resolution, and the sampling rate. Data logger uncertainty is then defined by the 95% confidence interval of the average temperature during a specific time, recorded during set sampling period by a data logger with known accuracy and resolution (Girondot et al. 2018a). The uncertainty of the mean temperature recorded every hour for 10
days is much lower than the accuracy and resolution of the logger. For example, using typical field conditions, an iButton DS1921G-F5# with accuracy=1 °C and resolution=0.5 °C has an uncertainty value of 0.15 °C (Girondot et al. 2018a). Overall, while the amplitude of temperatures during incubation should be minimized as much as possible, the uncertainty of the average temperature is a less important issue.

The data for this study originated from eggs collected in 6 nesting beaches (Pirambu, Brazil; Orissa, India; Playa Nancite, Costa Rica; El Verde Camacho, La Escobilla, and Playa La Destiladeras, Mexico) (Table 1, Fig. 1), belonging to 3 RMUs (West Atlantic, Northeast Indian, East Pacific). Though olive ridleys nesting on beaches in Pacific Mexico and Pacific Costa Rica belong to the same East Pacific RMU (Wallace et al. 2010), we analyzed these data separately because embryo growth dynamics from the two areas respond differently under a range of incubation temperatures (Morales Mérida et al. 2015).

Data included 40 incubations at various constant temperatures with a total of 464 sexed embryos (277 males, 168 females, and 19 intersexes). The 19 intersexes were reported from only 2 studies out of 8, from incubation temperatures ranging from 27.6 to 32.3 °C. An intersex is a transient state during development when the gonads are ovotestes that exhibit characteristics of both testes and ovaries (Pieau & Dorizzi 2004). After hatching, ovotestes generally evolve as normal testes (Pieau et al. 1998). Some adult gonads retain traces of intersexual characters when some oocytes may persist at the surface of testis in some species, whereas for others, no signs of intersex at the adult stage are observable. Because the criteria to define a gonad as an ovotestis are not entirely objective, we excluded data from hatchlings that were classified as intersex (n=19 values, or 3.9% of all sexed turtles, Table 1).

The East Pacific RMU was overrepresented in the database (384 embryos, 190 from Mexico and
178 197 from Costa Rica) as compared to Northeast India (31 embryos) and West Atlantic (50 embryos). The number of incubation temperatures that produced mixed sex ratios was 2 for West Atlantic, 1 for Northeast Indian, 4 for East Pacific (Costa Rica), and 1 for East Pacific (Mexico). Where stated, amplitudes of incubation temperatures were mostly <1 °C. Only two incubation temperatures in the West Atlantic RMU had amplitude > 1.5 °C, but both included only 1 embryo and so are unlikely to have biased the results.

184 Thermal reaction norm for sex ratio

Several models have been published to describe the mathematical relationship between constant incubation temperatures and sex ratios. We enlist the most useful, with relevant comments:

- **Logistic model.** It is based on an equation originally developed to model population growth (Verhulst 1838; Verhulst 1846). This was the first model applied to constant incubation data (Girondot 1999) and fitted using maximum likelihood with software that is no longer recommended. The logistic model uses 2 parameters: $P$ is the pivotal temperature, and $S$ is one fourth of the inverse of the slope at $P$. The TRT can be easily calculated as $TRT_l = S \cdot K_l$ with $K_l$ being a constant equal to $[2 \cdot \ln(l/1-l)]$ with $l$ being the limits to define the TRT. Girondot (1999) used $l = 0.05$ and then TRT was defined as the range of incubation temperatures with resulting sex ratios from 5% to 95%.

- **Hill model.** This model is used in biochemistry and pharmacology to reflect the binding of ligands to macromolecules as a function of the ligand concentration (Hill 1910). From a mathematical point of view it is similar to a logistic equation with the natural logarithms of temperatures on the x-axis. The Hill model is therefore asymmetrical and uses 2 parameters. Control of the asymmetry in the shape is not possible. The Hill model was used previously to describe the TSD pattern but subsequently discarded for lack of
A-logistic model (A- for Asymmetrical) was specially developed for TSD pattern analysis in Godfrey et al. (2003). It is based on the logistic equation with an additional parameter named $K$ that controls the asymmetry. This $K$ parameter is not the same as the $K_1$ parameter used to calculate TRT for the logistic model (a mistake made by Carter et al. 2019). Godfrey et al. (2003) also provided an equation to calculate the TRT. This model is asymmetrical, but the transitions towards the lower and upper asymptotes are not independent since they are controlled by a single parameter, $K$. The model has three parameters.

Hulin model. Hulin et al. (2009), recognizing that the A-logistic model was insufficiently flexible in the transitions toward lower and upper asymptotes, introduced a modification to $K$, making it a linear function of temperature: $K = K_1 t + K_2$. Four parameters were therefore fitted, and TRT can only be calculated numerically. Unfortunately, this model is challenging to fit and often hangs on local minima because $K$ can become very large during the search for maximum likelihood. In such situations, likelihood becomes insensitive to change in $K_1$ or $K_2$, and a local minimum is reached.

Additionally, it is noteworthy that the original description of TRT used the range of temperatures corresponding to sex ratios from 5% to 95%, which is unusual. Generally, the range is defined as occurring between 2.5% and 97.5%, thus encompassing a statistically meaningful 95% of the data. Sandoval et al. (2017) questioned the use of the 5-95% limits, arguing instead that the TRT limit should be proportional to the number of eggs. This, however, is a misconception of the role of models in biology. A model is not a means to replace data but rather to obtain a generalized description of a biological phenomenon using a mathematical formula. In this case it needs to be
framed within meaningful sex ratio limits. Thus, we recommend maintaining 5%-95% sex ratio as the limits of the TRT.

The lack of an ideal sigmoid model to describe TSD patterns (i.e., asymmetrical in the transitions toward lower and upper asymptotes) prompted us to develop a new, more versatile sigmoid function (see supplementary material 1) called flexible-logistic or flexit model:

\[
\begin{align*}
S_1 &= \frac{2^{K_1 - 1} SK_1}{2^{K_1 - 1}} \\
S_2 &= \frac{2^{K_2 - 1} SK_2}{2^{K_2 - 1}}
\end{align*}
\]

\[
f(x) = \begin{cases} 
1 & x < P \\
1 - \left( 1 + \left( 2^{K_1 - 1} \right) e^{4S_1 (p-x)} \right)^{-1/K_1} & x \geq P
\end{cases}
\]

\[
f(x) = 1 - \left( 1 + \left( 2^{K_2 - 1} \right) e^{4S_2 (x-P)} \right)^{-1/K_2}
\]

\[P\] is the pivotal temperature and \(S\) is the slope (first-order derivative) at \(P\). \(K_1\) and \(K_2\) control the lower and upper asymptotes respectively (acute for positive values and obtuse for negative values).

TRT can be calculated exactly:

\[
TRT = \frac{1}{4S_2} \ln \left( \frac{1}{(1-l)} \right)^{K_2 - 1} + \frac{1}{4S_1} \ln \left( \frac{1}{(1-l)} \right)^{K_1 - 1}
\]

A flexit model uses 4 parameters, a special case being \(K_1 = K_2 = 1\), which is the logistic model.

The model is not defined for \(K_1 = 0\) or \(K_2 = 0\). In this scenario, the corresponding value is replaced by 10\(^{-9}\).

The flexit model is included as a function in the HelpersMG R package (version 3.7 and higher) (Girondot 2019b) and is included in the tsd() function of the embryogrowth R package (version 7.5 and higher) (Girondot 2019a).

Overall, only two models are acceptable for our purpose: logistic and flexit. If an asymmetrical model is required, the Hill, A-logistic or Hulin models are not flexible enough when compared to...
243 a flexit version.

244 Fitting a model to the data: maximum likelihood

245 The fit of parameters (2 for logistic and 4 for flexit models) can be performed using the
246 maximum likelihood methodology. The likelihood function (simply “likelihood”) expresses how
247 probable a given set of observations is for different values of mathematical parameters. In the
248 context of a model of TSD pattern, the observations are counts of categories of embryos. In most
249 cases, two categories are found, males or females, and then a binomial distribution is used to
250 estimate likelihood.

251 The likelihood of a set of $M$ males and $F$ females observed after an incubation $t$ that produced a
252 theoretical sex ratio of $\text{flexit}(t; P, S, K_1, K_2)$ is named $L:

253 \[ L = \left(\frac{M + F}{M}\right)^{\text{flexit}(t; P, S, K_1, K_2)} M \left(1 - \text{flexit}(t; P, S, K_1, K_2)\right)^{F} \]

254 with $\left(\frac{M + F}{M}\right) = \frac{(M + F)!}{M!F!}$ being a constant for the set of observations.

255 An alternative that is described in some papers (e.g. Sandoval et al. 2017) is to fit the proportions
256 with the hypothesis that proportions or their angular transformation are normally distributed. We
257 do not recommend this method because it produced biased results: temperatures with the fewer
258 eggs incubated will have a larger than expected influence. Furthermore, even one sexed embryo
259 at one incubation temperature provides useful information and should be included in the analysis.

260 It should be noted that sex ratio is etymologically referred to as $M/F$ or $F/M$, but this definition is
261 not practical in statistics. For practical purposes, we use sex ratio in terms of the relative
262 frequency of males or females. The choice to work with male or female relative frequency has no
263 importance and depends on the researcher’s preference. Here we use sex ratio as being relative
264 male frequency, to be compatible with previous publications (Girondot 1999; Godfrey et al.
A possible alternative would be to incorporate the data on intersex hatchlings into the model, which tend to be more frequent at intermediate temperatures, and use a multinomial distribution. This possible solution has not yet been developed but could be an interesting avenue to explore in future studies for some species.

The likelihood value is often presented as its inverse natural logarithm (-ln $L$) for practical reasons: likelihoods are generally small numbers, thus -ln $L$ will be positive numbers that are easier to manipulate.

The likelihood of a dataset of several incubation temperatures $t_i$ with $M_i$ and $F_i$ within a model is simply the product of the likelihoods for each temperature, $L_i$, or the sum of the -ln $L_i$.

To be able to fit the sex ratio thermal reaction norm using maximum likelihood, at least one incubation temperature producing a mixed sex ratio should be present in the study dataset. A rule of thumb is that fitting a model with $p$ parameters necessitates at least $p$ temperatures with mixed sex ratios. For example, if a logistic model (2 parameters) is fitted for a dataset with only one temperature that produced a mixed-sex ratio, an infinite number of combinations of $P$ and $S$ will share the same likelihood. It is still possible to fit a logistic model to a dataset with only one temperature producing a mixed sex ratio, but standard error of parameters will be generally high.

If a dataset has no mixed sex ratio, it is still possible to use a Bayesian model to describe the credibility interval of the parameters and the outputs (see below).

The estimate of parameters for TSD patterns using maximum likelihood serves various purposes: (i) It allows an estimate of the confidence interval of the outputs (see below), (ii) it provides an estimate of the quality of fit (see below), (iii) it facilitates the comparison across datasets because their fitted parameters can be compared even if incubation temperatures were different, (iv) it
can be used as a prior for Bayesian analysis (see below), and (v) it can be used as a starting point for iterations using a Metropolis-Hastings algorithm with a Monte-Carlo Markov chain in Bayesian analyses (see below).

**Standard error and confidence interval**

The standard error is an important indicator of the precision of an estimate of a sample statistic for a population parameter. Its calculation is based on the Hessian matrix, which is the matrix of second-order partial derivatives of the likelihood for all pairwise parameters. The second-order derivative of a function at its maximum measures a more or less acute form of the function around the maximum. If a parameter is slightly shifted from its value at maximum likelihood and the likelihood changes drastically, this denotes a robust parameter estimate. At the same time, its standard error, which measures how well the parameter is known, will be low. On the contrary, if the likelihood is insensitive to changes of a parameter, it means the data do not provide information to fit this parameter and its standard error will be high. The inverse of the Hessian matrix is an estimator of the asymptotic covariance matrix. Hence, the square roots of the diagonal elements of a covariance matrix are estimators of the standard errors. A parameter obtained using maximum likelihood is normally distributed asymptotically.

This mathematical definition is sometimes problematic when the standard error is large and the effect of the parameter on the function changes drastically at some value. An example is the case for the $S$ parameters because the TSD pattern is completely the reverse for -$S$ (females at lower temperatures rather than high). Thus, if the standard error of $S$ is large, at the ends of the $S$ distribution the model will become completely reversed, and male production will be predicted at feminizing temperatures. The coefficient of variation for one parameter estimate is a standardized measure of dispersion with $CV = SE/\text{mean}$. The larger the coefficient of variation,
the worse the estimate of the parameter in the analysis.

The confidence interval for a parameter can be obtained directly from its point estimate and its standard error, assuming that it is normally distributed asymptotically. The confidence interval for a combination of parameters (for example, TRT) requires a more complicated calculation. Two main strategies are available to calculate the confidence intervals for a combination of parameters: the delta method and parameter resampling. The delta method uses the approximate probability distribution for a function of an asymptotically normal statistical estimator from knowledge of the limiting variance of that estimator. In short, if the standard error of the maximum likelihood parameters are known, the delta method permits an estimate of the distribution of any function of these parameters. The alternative is to generate many random numbers for the variance and covariance matrix of the estimators using the Cholesky decomposition (Tanabe & Sagae 1992). The function of interest is then applied to each set of random numbers. The advantage of the delta method is its rapidity. However, the assumption of a normal distribution for an estimator is important. It precludes the use of the delta method on a truncated distribution such as the $S$ distribution of a TSD pattern. If for example, $S$ changes sign, the likelihood degrades so much that it would be considered a truncated distribution. The advantage of the generation of many random numbers from the Hessian matrix is that it is possible to check each set of numbers for coherence and discard some if necessary. In the case of models for the TSD pattern, when the confidence interval of sex ratio according to temperature or TRT is estimated, we ensure that $S$, $K_1 - 1$, and $K_2 - 1$ do not change sign during resampling. However, this method will artificially reduce the confidence interval because the highly divergent values are removed. Using a Bayesian Metropolis-Hastings MCMC procedure solves this problem (see below).
Quality of fit

Often the quality of fit is measured by the determination coefficient $R^2$, which is derived from comparisons between observations and predictions. Whereas the determination coefficient has meaning when the distribution of the dependent variable is normally distributed, in the case of univariate regression, it has no meaningful statistical properties when used with proportions.

Thus, $R^2$ should not be used in most of these cases and particularly here where we need to measure the fit of the model to sex ratio data. Instead, deviance can be used as a goodness-of-fit measure for a statistical model. It is twice the difference between the logarithm of likelihood of the saturated model ($\ln L_S$). In a saturated model, the fitted sex ratio is replaced by the observed sex ratio and the model fits the data perfectly, and the logarithm of maximum likelihood of the fitted model ($\ln L_M$) is:

$$D = 2 (\ln L_S - \ln L_M)$$

Deviance has an asymptotic $\chi^2$ distribution with the degrees of freedom calculated from the difference of the number of parameters in the saturated and the fitted model. Note, however, that if there are few observations (which is often the case), the distribution of deviance can deviate substantially from a $\chi^2$ distribution and the test result could be wrong. For this reason, we developed an additional deviance test by randomly generating null deviance distributions with the same characteristics as those from the observations (the same number of incubation temperatures and the same number of eggs per incubation temperature). The probability that the observed deviance is obtained with the experimental design is then calculated by comparing the observed deviance and the distribution of deviances under the null hypothesis.
Comparison of models: Akaike Information Criterion and Akaike weight

When several models are fitted to the same datasets of observations, the comparison of the performance of the different models can be assayed using AIC (Akaike Information Criterion) (Akaike 1974). AIC is a measure of the quality of the fit, while simultaneously penalizes for the number of parameters in the model. It facilitates the selection from a set of models the best compromise between fit quality and over-parametrization.

$$AIC_j = -2\ln L_j + 2p_j$$

with $L_j$ being the likelihood of the model and $p_j$ the number of parameters of the model $j$.

When a set of $k$ models are tested, the model with lowest AIC is considered to be the best non-overparametrized fit. It is important to note that AIC value itself is not strictly informative in terms of absolute model fit.

A corrected version of $AIC$ for small sample sizes, named $AICc$, has been proposed when the model is univariate and linear with normal residuals (Hurvich & Tsai 1995):

$$AICc = AIC + \frac{2p(p + 1)}{n - p - 1}$$

The formula can be difficult to determine when these conditions are not met, in which case the previous formula could be used (Burnham & Anderson 2002). In general, the $AICc$ should always be used instead of $AIC$ (Burnham & Anderson 2002), especially for datasets comprised of small sample sizes, which is typical for sex ratio studies, and particularly when the study targets protected species.

When a set of models is compared, it is possible to estimate the relative probability that each model is the best among those tested using the Akaike weight (Burnham & Anderson 2002):

$$w_j = \frac{e^{-\frac{1}{2}AIC_j}}{\sum_{i=1}^k e^{-\frac{1}{2}AIC_i}}$$
376 \[ Akaike \text{ weight}_j = \frac{1}{e^{\frac{1}{2} AICc_j - \min(AICc)}} \sum_{i=1}^{k} e^{\frac{1}{2} AICc_i - \min(AICc)} \]

377 We use both \textit{AICc} and \textit{Akaike weight} to compare the fitted \textit{logistic} and \textit{flexit} models for our datasets.

378 The utility of model selection can be further extended to test for potential differences in the results from two or more datasets. In this case, the complete data are split into several subsets with each individual dataset being represented once and only once. The test question is: can the collection of datasets be modeled with a single set of parameters or must each dataset be modeled with its own set? In this situation, Bayesian Information Criterion (or \textit{BIC}) should be used instead of \textit{AIC} or \textit{AICc}, because the true model is obviously among the tested alternatives:

\[ BIC = -2\ln L + p\ln n \]

386 When the BIC statistic is used, all the priors of the models tested are assumed to be identical. It is also possible to estimate \textit{BIC} weights by replacing \textit{AICc} with \textit{BIC} in the Akaike weight formula. The \(w\)-value has been defined as the probability that several datasets can be correctly modeled by grouping instead of independently (Girondot & Guillon 2018).

390 In our example, a model will be fitted first to the combined datasets and \textit{BIC}(combined) will be estimated with \(p\) parameters. Then each dataset will be fitted separately and a set of \(k\) \textit{BIC}(separate) values will be generated, each with \(p\) parameters, thus using a total of \(k.p\) parameters. This is similar to the case of fitted models with a dataset effect. The global -\(ln\) likelihood for the separate fits is simply the sum of the -\(ln\) likelihoods. Then the \textit{BIC} weights will provide a statistically sound method of choosing between the hypotheses that a single model is sufficient for all datasets or that each dataset is best fitted with a different model. We named \(w\)-
value the BIC weight and we propose this statistic as a replacement for the contentious $p$-value (Girondot & Guillon 2018).

**Bayesian approach using Metropolis-Hastings with Monte-Carlo Markov chain**

The Metropolis–Hastings algorithm is a Markov chain Monte Carlo (MCMC) method for obtaining a sequence of random samples from a probability distribution (Hastings 1970; Metropolis et al. 1953). This method is now widely used as it offers a high-performance tool to fit a model. To run a Bayesian analysis with this algorithm, several parameters must be defined for each estimator in the model. The following terms are those used in the function `tsd_MHmcmc()` of the R package *embryogrowth*:

1. Density: The R function for density distribution used for the prior. Generally, uniform or Gaussian distribution are used with `dunif` or `dnorm`, respectively.

2. Prior1 and Prior2: The parameters describing the prior distribution. For `dunif`, Prior1 and Prior2 are respectively the minimum and maximum of the distribution, and for `dnorm`, they are the mean and standard deviation of the distribution;

3. SDProp: The standard deviation for each new parameter;

4. Min and Max: The range of possible values defined by the minimum and maximum;

5. Init: An initial starting point for the Markov chain.

It is beyond the scope of this paper to fully explore the fine details of this algorithm, and rather we focus on how to use it.

The choice of the prior is not straightforward (Lemoine 2019) and can be critical if only a few observations are available. A uniform distribution for priors indicates that all values within a range are equally probable, whereas a Gaussian distribution can use a mean and standard deviation obtained from a previous analysis with the same or other species.
At the end of the analysis, it is essential to evaluate the distribution of posteriors. If they are the same as the distribution of priors, it generally means that the observations did not help inform the fit for this parameter. In this case, it implies that the results are dependent on the priors and not on the data and, therefore, the results should not be used, or used with caution.

During the iteration process, a Markov Chain is constructed using the actual parameter values $\pi_t$ on which a new proposal random function defined by its standard deviation ($s$) is applied, $\pi_{t+1} = N(\pi_p, s)$. This is the Monte-Carlo process. The standard deviation ($s$) for a new proposal is a compromise between the two constraints: if the values are too high, the new values could yield results far from the optimal solution, while if they are too low, the model can become stuck in local minima. The adaptive proposal distribution (Rosenthal 2011) as implemented in R package *HelpersMG* (Girondot 2019b) will ensure that the acceptance rate is close to 0.234, which is the optimal acceptance rate (Roberts & Rosenthal 2001). The burn-in value is the number of iterations necessary to stabilize the likelihood. It can be low (around 10) when the starting values are the maximum likelihood estimators. The total number of iterations required is defined after an initial run of 10,000 iterations (Raftery & Lewis 1992). The result of the MCMC analysis is a table with one set of values for the estimators at each iteration. The mean and standard deviation summary statistics can be calculated from this table using the *coda* R package (Plummer et al. 2011), or it is also possible to estimate quantiles. The use of quantiles has the advantage that it does not require any hypothesis on the output distribution; hence an asymmetric distribution is well accommodated. When the result of a MCMC analysis is used to estimate a function of these estimators (for example, the TRT), each individual set of parameters obtained during the MCMC search should be used to generate the posterior distribution of outputs and this distribution can be summarized using mean, standard deviation or quantiles. The standard deviation of the MCMC
output is the standard error of the corresponding parameter.

We follow the ISO 80000-1:2009 standard indicating that numerical value and unit symbol are separated by a space, including the °C symbol (ISO/TC 12 Quantities and units & IEC/TC 25 Quantities and units 2009).

Results

Maximum likelihood estimates for grouped data

*Logistic* and *flexit* models were both fitted to the comprehensive dataset: 40 incubations with temperatures from 24 °C to 34 °C and 468 sexed embryos identified as males or females. Mixed-sex ratios were observed in embryos or hatchlings from 8 incubations (Table 1). Upon comparing both models (Fig. 2 and Table 2), the *flexit* model was selected based on AICc and Akaike weight as the latter suggested there was a 97% probability that the *flexit* model was the best. While the *logistic* model showed the typical symmetric shape, the *flexit* model showed a strong asymmetrical pattern: transition from all-male condition to the pivotal temperature is smooth, whereas the transition from the pivotal temperature to all-female production was abrupt (Fig. 2). As a consequence, the estimated pivotal and transitional range of temperatures obtained from both models were different (Table 1).

It is worthwhile noting that the standard error and therefore the coefficient of variation for the $S, K_1$ and $K_2$ parameters in the *flexit* model was large ($S=-0.79$ SE 0.3; $K_1=-1.72$ SE 1.59; $K_2=200$ SE 2485.61, see supplementary material). In other words, the *flexit* model did not precisely describe all parameters, even though it was the selected model.

The tests of the deviances indicated that the models fit the observations correctly: $p=0.16$ and $p=0.56$, respectively, for the *logistic* and *flexit* models on global data. The same is true for the fit
at RMU or country scale except for Costa Rica data ($p=0.01$, Table 3). The distribution of 1000 estimates of deviance for both the logistic and the flexit models is shown in tables 2 and 3. All these tests gave a probability of more than 0.05, indicating that the models fit the observations relatively well. From these observations, we recommend against the use of $\chi^2$ approximations to test the degree to which the models fit the data and, instead, use comparisons based on generated null distributions. Based on this methodology, the function $\text{tsd()}$ of the phenology R package returns the probability that the observations fit the null model in the object $\text{p.Deviance.Null.model}$ (Girondot 2019a).

Finally, because some parameters of the flexit model are nearly impossible to fit with these datasets (coefficient of variation for $K_1=0.92$, and for $K_2=12.43$), the flexit model will not be applied further in this paper (see supplementary 2 for a Bayesian MCMC with flexit model).

Comparison of the maximum likelihood estimates for RMU data

The characteristics of the fitted logistic model for the 3 RMUs and Mexico and Costa Rica within the East Pacific RMU are shown in Table 3. The fitted $S$ values for East Pacific (Mexico) and Northeast India observations have substantial standard errors and confidence intervals. These two datasets have only one mixed-sex ratio, and the maximum likelihood approximation failed to provide a correct estimate of the $S$ parameter.

The analysis shows that, within the East Pacific RMU, a single TSD pattern is sufficient to model the combined Mexico and Costa Rica data ($w$-value=0.87), whereas there is a 0.97 probability ($w$-value) that at a worldwide scale there are differences in TSD parameters for the different olive ridley populations (Table 4).
Bayesian estimates

Priors for the Bayesian process were chosen with a Gaussian distribution, with an average set as the fitted parameters, which were estimated from the global maximum likelihood analysis (Fig. 2, Table 2), and a standard deviation large enough to avoid imposing too high a constraint on the output. Minimum and maximum values for \( P \) were 25 and 35 °C, respectively, with \( S \) spanning from -2 to 2 and \( K_1 \) from -500 to +500. The standard deviation for the new proposal was initially chosen to be 2 for \( P \) and 0.5 for \( S, K_1 \), and \( K_2 \). They were adjusted during the iterations to have acceptance rates closer to 0.234. A total of 100,000 iterations were performed.

Though both the logistic and flexit models were fitted using Bayesian MCMC, only the results for the logistic will be discussed here because some parameters of the flexit model are nearly impossible to fit with these datasets (coefficient of variation for \( K_2 \)=12.43) (see supplementary 2 for a Bayesian MCMC with flexit model).

The distribution of priors and posteriors for the parameters derived from the East Pacific analyses is shown in Fig. 3A & B. The distinctness of the posterior and prior distributions verify that the fit was guided primarily by the observations. The posterior distributions of the \( P \) and \( S \) parameters were relatively independent (Fig. 3C). Distribution of other posteriors is shown in supplementary material 2.

TSD pattern fits using maximum likelihood and Bayesian MCMC (Figs. 4 and 5) were derived for the 3 RMUs, and separately for Mexico and Costa Rica (East Pacific RMU) together with the 0.025, 0.5 and 0.975 quantiles (95% of values are located between 0.025, and 0.975 quantiles) for pivotal temperature and transitional range of temperatures (Table 5).

The plotted posterior distributions for \( P \) vs. TRT from the 2 RMUs and Mexico and Costa Rica separately, is shown in Fig. 6. While posterior values for \( P \) suggest two separate groups
(Northeast India on one side, and East Pacific and West Atlantic on the other), the spread of TRT
values does not suggest any differences.

**Discussion**

After the discovery of temperature-dependent sex determination in reptiles (Charnier 1966; Pieau
1971; Pieau 1972), the description of the TSD pattern provided a basis for its understanding and
for comparisons among species (Bull et al. 1982; Ewert et al. 2004; Ewert et al. 1994; Ewert et
al. 2005; Ewert & Nelson 1991; Godfrey et al. 1999; Mrosovsky 1988; Yntema & Mrosovsky
1982). However, significant methodological advancements have been made since the original
studies. For example, straight-line interpolations of the sex ratio values on each side of the 50 %
level were proposed by N. Mrosovsky (pers. comm. in 1992) (see also Mrosovsky & Pieau
1991). Limpus et al. (1983) refined the approach using a graphical method (Litchfield &
Wilcoxon 1949) using a curve from the intercept and the slope of a straight line in the log dose
vs. probit effect scale. This method allowed statistical comparisons between samples but required
at least two values in the 16% to 84% dose-effect range, confidence limit calculations, and re-
testing between samples. Starting in 1999, new statistical methods were developed (Girondot
1999; Godfrey et al. 2003; Hulin et al. 2009) and the R package *embryogrowth* (Girondot 2019a)
has included the `tsd()` function since version 2.0.0 in 2013.

More recently, an attempt to correlate TSD patterns with life-history parameters was successful
(Carter et al. 2019; Hulin et al. 2009) but sometimes produced unexpected results. For example,
when these were compared between populations in *Chrysemys picta*, *P* could not be explained by
geography or local thermal conditions, but the TRT was wider at lower latitudes (Carter et al.
2019). An explanation for these difficulties could be that the models for the TSD pattern used did
not correctly reflect the true TSD pattern.
The TSD pattern has also been used to estimate sex ratios from natural nests using the average
nest temperature for the total incubation period, which can result in biased data because
temperature determines sex only during the thermosensitive part of development. When the
average of the temperature experienced during the absolute middle third of the incubation period
is used, sex ratio is also potentially biased because the thermosensitive period for sex
determination shifts depending on the temperature (Georges et al. 2005; Girondot et al. 2018b).
Recently, improved models that take into account changes in the rate of embryonic development
affected by variations in the incubation temperature have been developed (Massey et al. 2019;
Monsinjon et al. 2019a; Monsinjon et al. 2019b). Promising results have been obtained, but
further analyses of in-situ empirical data would benefit the field (Fuentes et al. 2017).
Nevertheless, the mathematical and statistical complications of currently available methods to
analyze and study TSD patterns were a barrier for the full understanding and further analyses of
sex ratios. The purpose of this article was to fully describe the procedures and the R code to
familiarize readers with the analyses and the outputs.
The main conclusions of the example analysis are summarized here. A single model is sufficient
to describe Mexican and Costa Rican rookeries within East Pacific RMU. This result is
consistent with the phylogeography of this species (see figure 2 in Bowen et al. 1998). The
pivotal temperature for the India rookery (Northeast Indian RMU) differs substantially from the
values estimated for both Mexico and Costa Rica (East Pacific RMU), and also from the
estimation for Brazil (West Atlantic RMU). However, no differences among the RMUs were
observed from the estimated TRTs. The use of the Bayesian Metropolis-Hastings algorithm with
Markov chains generated by a Monte-Carlo process provides a substantial improvement to the
model fit as compared to the maximum likelihood fit, especially when few mixed-sex ratio
results are available. The quality of the fit, as measured by a deviance test, was generally reliable, but in one case the generation of null distribution for deviance gave a result different from the $\chi^2$ approximation. We suggest that null distributions of deviance is more reliable than $\chi^2$ approximations to test deviance and should be chosen when possible.

We hope that this methodological paper will be useful and encourage researchers to explore new hypotheses to understand the ecological and evolutionary significance of temperature-dependent sex determination in reptiles. We also encourage more studies on TSD patterns for different populations and species, to help improve our understanding of this fascinating phenomenon.

References

Abreu-Grobois A, and Plotkin PT. 2008. *Lepidochelys olivacea*. The IUCN Red List of Threatened Species. Gland, Switzerland: IUCN SSC Marine Turtle Specialist Group.

Akaike H. 1974. A new look at the statistical model identification. *IEEE Transactions on Automatic Control* 19:716-723.

Bowen BW, Clark AM, Abreu-Grobois FA, Chaves A, Reichart HA, and Ferl RJ. 1998. Global phylogeography of the ridley sea turtles (*Lepidochelys spp.*) as inferred from mitochondrial DNA sequences. *Genetica* 101:179-189.

Bull JJ. 1983. *Evolution of sex determining mechanism*. Menlo Park, Ca: The Benjamin/Cummings Publishing Company, Inc.

Bull JJ, Vogt RC, and McCoy CJ. 1982. Sex determining temperatures in turtles: A geographic comparison. *Evolution* 36:326-332.

Burnham KP, and Anderson DR. 2002. *Model selection and multimodel inference: A practical information-theoretic approach*. New York: Springer-Verlag.

Carter AL, Bodensteiner BL, Iverson JB, Milne-Zelman CL, Mitchell TS, Refsnider JM, Warner DA,
Janzen FJ, and White C. 2019. Breadth of the thermal response captures individual and
geographic variation in temperature-dependent sex determination. *Functional Ecology*.

Castheloge VD, de D. dos Santos MR, de Castilhos JC, de J. Filho PR, de C. Gomes L, Clemente-
Carvalho RBG, and Ferreira PD. 2018. Pivotal temperature and hatchling sex ratio of olive ridley
sea turtles *Lepidochelys olivacea* from the South Atlantic coast of Brazil. *Herpetological
Conservation and Biology* 13:488-496.

Charnier M. 1966. Action de la température sur la sex-ratio chez l'embryon d'*Agama agama* (Agamidae,
Lacertilien). *Comptes Rendus des Séances de la Société de Biologie et de ses Filiales* 160:620-
622.

Dimond MT. 1985. Some effects of temperature on turtle egg incubation. In: Goel SC, and Srivastava
CBL, eds. *Recent Advances in Developmental Biology of Animals*. Poona, India: Indian Society
of Developmental Biologists, 35-39.

Dutton PH, Whitmore CP, and Mrosovsky N. 1985. Masculinisation of leatherback turtle *Dermochelys
coriacea* hatchlings from eggs incubated in styrofoam boxes. *Biological Conservation* 31:249-
264.

Ewert MA, Etchberger CR, and Nelson CE. 2004. Turtle sex-determining modes and TSD patterns, and
some TSD patterns correlates. In: Valenzuela N, and Bull JJ, eds. *Temperature-Dependent Sex
Determination in Vertebrates*. Washington: Smithsonian Books, 21-32.

Ewert MA, Jackson DR, and Nelson CE. 1994. Patterns of temperature-dependent sex determination in
turtles. *Journal of Experimental Zoology* 270:3-15.

Ewert MA, Lang JW, and Nelson CE. 2005. Geographic variation in the pattern of temperature-dependent
sex determination in the American snapping turtle (*Chelydra serpentina*). *Journal of Zoology
265:81-95.

Ewert MA, and Nelson CE. 1991. Sex determination in turtles: diverse patterns and some possible
adaptative values. *Copeia* 1991:50-69.

Fuentes MMPB, Monsinjon J, Lopez M, Lara P, Santos A, dei Marcovaldi MAG, and Girondot M. 2017.
Sex ratio estimates for species with temperature-dependent sex determination differ according to the proxy used. *Ecological Modelling* 365:55-67.

Georges A. 1989. Female turtles from hot nests: Is it duration of incubation or proportion of development at high temperatures that matters? *Oecologia* 81:323-328.

Georges A, Beggs K, Young JE, and Doody JS. 2005. Modelling development of reptile embryos under fluctuating temperature regimes. *Physiological and Biochemical Zoology* 78:18-30.

Georges A, Limpus CJ, and Stoutjesdijk R. 1994. Hatchling sex in the marine turtle *Caretta caretta* is determined by proportion of development at a temperature, not daily duration of exposure. *Journal of Experimental Zoology* 270:432-444.

Girondot M. 1999. Statistical description of temperature-dependent sex determination using maximum likelihood. *Evolutionary Ecology Research* 1:479-486.

Girondot M. 2019a. embryogrowth: Tools to analyze the thermal reaction norm of embryo growth. 7.7 ed: The Comprehensive R Archive Network.

Girondot M. 2019b. HelpersMG: Tools for Environmental Analyses, Ecotoxicology and Various R Functions. 3.9 ed: The Comprehensive R Archive Network.

Girondot M, Godfrey MH, Guillon J-M, and Sifuentes-Romero I. 2018a. Understanding and integrating resolution, accuracy and sampling rates of temperature data loggers used in biological and ecological studies. *Engineering Technology Open Access Journal* 2:1-7.

Girondot M, and Guillon J-M. 2018. The w-value: An alternative to t- and X² tests. *Journal of Biostatistics & Biometrics* 1:1-4.

Girondot M, and Kaska Y. 2014. A model to predict the thermal reaction norm for the embryo growth rate from field data. *Journal of Thermal Biology* 45:96-102.

Girondot M, Monsinjon J, and Guillon J-M. 2018b. Delimitation of the embryonic thermosensitive period for sex determination using an embryo growth model reveals a potential bias for sex ratio prediction in turtles. *Journal of Thermal Biology* 73:32-40.

Godfrey MH, D'Amato AF, Marcovaldi MA, and Mrosovsky N. 1999. Pivotal temperature and predicted
sex ratios for hatchling hawksbill turtles from Brazil. *Canadian Journal of Zoology-Revue Canadienne de Zoologie* 77:1465-1473.

Godfrey MH, Delmas V, and Girondot M. 2003. Assessment of patterns of temperature-dependent sex determination using maximum likelihood model selection. *Ecoscience* 10:265-272.

Godfrey MH, and Mr osovsky N. 2006. Pivotal temperature for green sea turtles, *Chelonia mydas*, nesting in suriname. *Herpetological Journal* 16:55-61.

Hastings WK. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* 57:97-109.

Hill AV. 1910. The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves. *J Physiol* 40:IV-VII.

Hulin V, Delmas V, Girondot M, Godfrey MH, and Guillon J-M. 2009. Temperature-dependent sex determination and global change: Are some species at greater risk? *Oecologia* 160:493-506.

Hulin V, Girondot M, Godfrey MH, and Guillon J-M. 2008. Mixed and uniform brood sex ratio strategy in turtles: The facts, the theory and their consequences. In: Wyneken J, Bels V, and Godfrey MH, eds. *Biology of turtles*. Chigaco, IL: University Press of Chicago, 279-300.

Hurvich CM, and Tsai C-L. 1995. Model selection for extended quasi-likelihood models in small samples. *Biometrics* 51:1077-1084.

ISO/TC 12 Quantities and units, and IEC/TC 25 Quantities and units. 2009. ISO 80000-1:2009. Quantities and units. ISO.

Korpelainen H. 1990. Sex ratios and the conditions required for environmental sex determination in animals. *Biological Reviews* 65:147-184.

Lang JW, and Andrews HV. 1994. Temperature-dependent sex determination in crocodilians. *J Exp Zool* 270:28-44.

Lemoine NP. 2019. Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. *Oikos* 128:912-928.

Lewontin R. 2000. *The Triple Helix: Gene, Organism, and Environment*. Cambridge, MA: Harvard
Limpus CJ, Reed P, and Miller JD. 1983. Islands and turtles: the influence of choice of nesting beach on sex ratio. In: Baker JJ, Carter RM, Saurmarco PW, and Stark KP, editors. Proceedings of the Inaugural Great Barrier Reef Conference: James Cook University Press of North Queensland. p 397-402.

Litchfield JT, and Wilcoxon F. 1949. A simplified method of evaluating dose-effect experiments. *Journal of Pharmacology and Experimental Therapeutics* 96:99-113.

López Correa JY. 2010. Diferenciación gonádica en crias de *Lepidochelys olivacea* (Eschscholtz, 1829) (Testudinata: Cheloniidae) PhD dissertation. Instituto Politécnico Nacional, Centro Interdisciplinaria de Ciencias Marinas.

Massey MD, Holt SM, Brooks RJ, and Rollinson N. 2019. Measurement and modelling of primary sex ratios for species with temperature-dependent sex determination. *J Exp Biol* 222:1-9.

McCoy CJ, Vogt RC, and Censky EJ. 1983. Temperature-controlled sex determination in the sea turtle *Lepidochelys olivacea*. *J Herpetol* 17:404-406.

Merchant-Larios H, Ruiz-Ramirez S, Moreno-Mendoza N, and Marmolejo-Valencia A. 1997. Correlation among thermosensitive period, estradiol response, and gonad differentiation in the sea turtle *Lepidochelys olivacea*. *General and Comparative Endocrinology* 107:373-385.

Merchant-Larios H, Villalpando-Fierro I, and Centeno-Urruiza B. 1989. Gonadal morphogenesis under controlled temperature in the sea turtle *Lepidochelys olivacea*. *Herpetol Monographs* 3:43-61.

Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, and Teller E. 1953. Equations of state calculations by fast computing machines. *Journal of Chemical Physics* 21:1087-1092.

Miller JD. 1985. Embryology of marine turtles. In: Gans C, Billet F, and Maderson PF, eds. *Biology of the Reptilia*. New-York, US: Wiley-Liss, 270-328.

Mohanty-Hejmadi P, Behra M, and Dimond MT. 1985. Temperature dependent sex differentiation in the olive ridley *Lepidochelys olivacea* and its implications for conservation. Symposium on Endangered Marine Animals and Marine Parks. Cochin.: Marine Biological Association of India.
Monsinjon J, Lopez M, Lara P, Santos A, dei Marcovaldi MAG, Girondot M, and Fuentes MMPB. 2019a. Effects of temperature and demography on the phenology of loggerhead sea turtles in Brazil. *Marine Ecology Progress Series* 623:209-219.

Monsinjon JR, Wyneken J, Rusenko K, López-Mendilaharsu M, Lara P, Santos A, dei Marcovaldi MAG, Fuentes MMPB, Kaska Y, Tucek J, Nel R, Williams KL, LeBlanc A-M, Rostal D, Guillon J-M, and Girondot M. 2019b. The climatic debt of loggerhead sea turtle populations in a warming world. *Ecological Indicators* 107:1-13.

Morales Mérida A, Ruiz A, Navarro E, Sifuentes-Romero I, Abreu-Grobois FA, and Girondot M. 2015. Contrasting incubation data for *Lepidochelys olivacea* highlight the need for more experimentation and detailed reporting. *Marine Turtles Newsletter* 145:15-19.

Mrosovsky N. 1988. Pivotal temperatures for loggerhead turtles (*Caretta caretta*) from northern and southern nesting beaches. *Canadian Journal of Zoology-Revue Canadienne de Zoologie* 66:661-669.

Mrosovsky N, and Pieau C. 1991. Transitional range of temperature, pivotal temperatures and thermosensitive stages for sex determination in reptiles. *Amphibia-Reptilia* 12:169-179.

Navarro Sánchez EJ. 2015. Efecto de la temperatura de incubación y la diferenciación sexual sobre la morfología de crías de tortuga marina *Lepidochelys olivacea* Maestría Tesis de Maestría. Universidad Nacional Autónoma de México.

Pieau C. 1971. Sur la proportion sexuelle chez les embryons de deux Chéloniens (*Testudo graeca* L. et *Emys orbicularis* L.) issus d'oeufs incubés artificiellement. *Comptes Rendus de l'Académie des Sciences Paris Série D* 272:3071-3074.

Pieau C. 1972. Effets de la température sur le développement des glandes génitales chez les embryons de deux Chéloniens, *Emys orbicularis* L. et *Testudo graeca* L. *Comptes Rendus de l'Académie des Sciences Paris Série D* 274:719-722.

Pieau C, and Dorizzi M. 2004. Oestrogens and temperature-dependent sex determination in reptiles: all is
in the gonads. *Journal of Endocrinology* 181:367-377.

Pieau C, Dorizzi M, Richard-Mercier N, and Desvages G. 1998. Sexual differentiation of gonads as a function of temperature in the turtle *Emys orbicularis*: endocrine function, intersexuality and growth. *Journal of Experimental Zoology* 281:400-408.

Plummer M, Best N, Cowles K, and Vines K. 2011. coda: Output analysis and diagnostics for Markov Chain Monte Carlo simulations. 0.14-6 ed: http://cran.r-project.org/web/packages/coda/index.html.

Raftery AE, and Lewis SM. 1992. One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo. *Statistical Science* 7:493-497.

Roberts GO, and Rosenthal JS. 2001. Optimal scaling for various Metropolis-Hastings algorithms. *Statistical Science* 16:351-367.

Rosenthal JS. 2011. Optimal proposal distributions and adaptive MCMC. In: Brooks S, Gelman A, Jones G, and Meng X-L, eds. *MCMC Handbook*: Chapman and Hall/CRC, 93-112.

Sandoval S, Gómez-Muñoz VM, and Porta-Gándara MÁ. 2017. Expansion of the transitional range of temperature for sea turtle *Lepidochelys olivacea* from sex ratio data at controlled incubation temperatures. *Herpetology Notes* 10:63-65.

Tanabe K, and Sagae M. 1992. An exact cholesky decomposition and the generalized inverse of the variance–covariance matrix of the multinomial distribution, with applications. *Journal of the Royal Statistical Society: Series B (Methodological)* 54:211-219.

Tezak BM, Sifuentes-Romero I, and Wyneken J. 2018. A new approach for measuring temperature inside turtle eggs. *Journal of Experimental Biology*.

Valenzuela N. 2004. Evolution and maintenance of temperature-dependent sex determination. In: Valenzuela N, and Lance V, eds. *Temperature-Dependent Sex Determination in Vertebrates*. Whashington: Smithsonian Books, 131-147.

Verhulst PF. 1838. Notice sur la loi que la population suit dans son accroissement. *Correspondance Mathématique et Physique* 10:113-121.
Verhulst PF. 1846. Deuxième mémoire sur la loi d'accroissement de la population. Mémoires de l'Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique 20:1-32.

Wallace BP, DiMatteo AD, Hurley BJ, Finkbeiner EM, Bolten AB, Chaloupka MY, Hutchinson BJ, Abreu-Grobois FA, Amoroco D, Bjorndal KA, Bourjea J, Bowen BW, Dueñas RB, Casale P, Choudhury BC, Costa A, Dutton PH, Fallabrino A, Girard A, Girondot M, Godfrey MH, Hamann M, López-Mendilaharsu M, Marcovaldi MA, Mortimer JA, Musick JA, Nel R, Seminoff JA, Troëng S, Witherington B, and Mast RB. 2010. Regional management units for marine turtles: a novel framework for prioritizing conservation and research across multiple scales. PLoS One 5:e15465.

Wibbels T, Rostal DC, and Byles R. 1998. High pivotal temperature in the sex determination of the olive ridley sea turtle, Lepidochelys olivacea, from Playa Nancite, Costa Rica. Copeia 1998:1086-1088.

Wilson EB. 1927. Probable inference, the law of succession, and statistical inference. Journal of the American Statistical Association 22:209-212.

Yntema CL, and Mrosovsky N. 1982. Critical periods and pivotal temperatures for sexual differentiation in loggerhead sea turtles. Canadian Journal of Zoology-Revue Canadienne de Zoologie 60:1012-1016.
Table 1 (on next page)

Constant temperature incubation data used in this study on TSD patterns in *Lepidochelys olivacea*.

( * ) maximum amplitude of temperatures within the middle-third of the incubation is shown. References: 1 - Castheloge et al. (2018), 2 - Dimond (1985), 3 - Mohanty-Hejmadi et al. (1985), 4 - McCoy et al. (1983), 5 - Wibbels et al. (1998), 6 - Merchant-Larios et al. (1997), 7 - Merchant-Larios et al. (1989), 8 - López Correa (2010), 9 – Navarro Sánchez (2015).
| Area          | Country     | RMU         | Incubation temperature °C | Temperature amplitude °C | Males | Females | Intersexes | Reference |
|---------------|-------------|-------------|---------------------------|--------------------------|-------|---------|------------|-----------|
| Pirambu Beach | Brazil      | West Atlantic | 26.3                      | 0.3                      | 2     | 0       | 0          | 1         |
| Pirambu Beach | Brazil      | West Atlantic | 28.5                      | 0.2                      | 3     | 0       | 0          |           |
| Pirambu Beach | Brazil      | West Atlantic | 29.2                      | 0.2                      | 8     | 1       | 0          |           |
| Pirambu Beach | Brazil      | West Atlantic | 29.9                      | 0.5                      | 7     | 0       | 0          |           |
| Pirambu Beach | Brazil      | West Atlantic | 30.4                      | 0.2                      | 5     | 2       | 0          |           |
| Pirambu Beach | Brazil      | West Atlantic | 31.2                      | 1.9                      | 0     | 1       | 0          |           |
| Pirambu Beach | Brazil      | West Atlantic | 31.8                      | 1.2                      | 0     | 3       | 0          |           |
| Pirambu Beach | Brazil      | West Atlantic | 32.1                      | 0.4                      | 0     | 6       | 0          |           |
| Pirambu Beach | Brazil      | West Atlantic | 32.2                      | 0.4                      | 0     | 2       | 0          |           |
| Pirambu Beach | Brazil      | West Atlantic | 32.4                      | 1.9                      | 0     | 1       | 0          |           |
| Pirambu Beach | Brazil      | West Atlantic | 32.9                      | 0.6                      | 0     | 5       | 0          |           |
| Pirambu Beach | Brazil      | West Atlantic | 33.1                      | 0.3                      | 0     | 4       | 0          |           |
| Orissa        | India       | Northeast Indian | 26.5                      | 3                        | 0     | 0       | 0          | 2, 3      |
| Orissa        | India       | Northeast Indian | 28                        | 8                        | 0     | 0       | 0          |           |
| Orissa        | India       | Northeast Indian | 29.5                      | 2                        | 3     | 0       | 0          |           |
| Orissa        | India       | Northeast Indian | 30                        | 0                        | 4     | 0       | 0          |           |
| Orissa        | India       | Northeast Indian | 31                        | 0                        | 2     | 0       | 0          |           |
| Orissa        | India       | Northeast Indian | 31.5                      | 0                        | 9     | 0       | 0          |           |
| Nancite       | Costa Rica  | East Pacific | 25                        | 23                       | 0     | 0       | 0          | 4         |
| Nancite       | Costa Rica  | East Pacific | 28                        | 30                       | 1     | 3       | 3          |           |
| Nancite       | Costa Rica  | East Pacific | 30                        | 12                       | 8     | 5       | 5          |           |
| Nancite       | Costa Rica  | East Pacific | 32                        | 0                        | 23    | 0       | 0          |           |
| Nancite       | Costa Rica  | East Pacific | 27                        | 0.5                      | 15    | 0       | 0          | 5         |
| Nancite       | Costa Rica  | East Pacific | 29.4                      | 0.5                      | 23    | 1       | 0          |           |
| Nancite       | Costa Rica  | East Pacific | 30.4                      | 0.5                      | 16    | 3       | 0          |           |
| Nancite       | Costa Rica  | East Pacific | 32                        | 0.5                      | 0     | 19      | 0          |           |
| Location             | Country   | Region       | Temperature | Survival Rate | Mortality Rate | Abundance | Notes |
|----------------------|-----------|--------------|-------------|---------------|----------------|-----------|-------|
| La Escobilla         | Mexico    | East Pacific | 27          | 25            | 0              | 0         | 6     |
| Playa La Destiladeras| Mexico    | East Pacific | 27.61       | 0.79 *        | 3              | 0         | 4     |
| Playa La Destiladeras| Mexico    | East Pacific | 32.24       | 0.81 *        | 2              | 0         | 0     |
| Playa La Destiladeras| Mexico    | East Pacific | 28.62       | 0.86 *        | 2              | 0         | 0     |
| Playa La Destiladeras| Mexico    | East Pacific | 32.29       | 1.05 *        | 2              | 0         | 0     |
| El Verde Camacho     | Mexico    | East Pacific | 24          | 0.1           | 16             | 0         | 0     |
| El Verde Camacho     | Mexico    | East Pacific | 26          | 0.1           | 24             | 0         | 0     |
| El Verde Camacho     | Mexico    | East Pacific | 28          | 0.1           | 20             | 0         | 0     |
| El Verde Camacho     | Mexico    | East Pacific | 30          | 0.1           | 15             | 4         | 0     |
| El Verde Camacho     | Mexico    | East Pacific | 32          | 0.1           | 20             | 0         | 0     |
| El Verde Camacho     | Mexico    | East Pacific | 34          | 0.1           | 5              | 0         | 0     |
Table 2 (on next page)

Summary of logistic and flexit model fits for *Lepidochelys olivacea* incubation data (East Pacific, West Atlantic, and Northeast Indian).

“p-random” column is to the probability that the observed deviance was obtained using a random sampling with the same characteristics (same number of temperatures and eggs per temperature and sex ratio probability obtained from the fitted model). Line with bold text indicates the best model.
| Model | P (mean SE) | S (mean SE) | K₁ (mean SE) | K₂ (mean SE) | -Ln L | AICc | ΔAICc | Akaike Weight | Deviance | df | p-value | p-random |
|-------|-------------|-------------|---------------|---------------|-------|-----|-------|----------------|-----------|----|---------|----------|
| **Logistic** | 30.39 SE 0.09 | -0.41 SE 0.05 | 32.6 | 69.52 | 6.92 | 0.03 | 45.68 | 38 | 0.183117 | 0.08 |
| **Flexit** | 30.57 SE 0.11 | -0.79 SE 0.3 | 200 SE 2477.58 | 26.73 | 62.6 | 0 | 0.97 | 33.95 | 36 | 0.566649 | 0.22 |
Table 3 (on next page)

Characteristics of logistic models fitted for *Lepidochelys olivacea* incubation data (East Pacific, West Atlantic, and Northeast Indian).

“p-random” column is to the probability that the observed deviance was obtained using a random sampling with the same characteristics (same number of temperatures and eggs per temperature and sex ratio probability obtained from the fitted model). Cells with large coefficient of variation are in bold and could produce a sign change for the parameter during resampling.
| RMU           | Country            | P (mean SE) | S (mean SE) | -Ln L | Deviance | df | p-value | p-random |
|---------------|--------------------|-------------|-------------|-------|----------|----|---------|----------|
| Northeast India | India              | 29.49 SE 0.33 | -0.03 SE 0.81 | 1.06  | 0        | 4  | 1       | 0.4      |
| West Atlantic | Brazil             | 30.63 SE 0.23 | -0.36 SE 0.12 | 4.46  | 4.76     | 10 | 0.90    | 0.56     |
| East Pacific  | Costa Rica + Mexico| 30.46 SE 0.1 | -0.37 SE 0.05 | 16.32 | 19.43    | 20 | 0.49    | 0.28     |
|               | Costa Rica         | 30.5 SE 0.13 | -0.42 SE 0.07 | 13.5  | 16.83    | 6  | 0.01    | 0.26     |
| "            | Mexico             | 30.16 SE 1.9 | -0.12 SE 1.43 | 1.52  | 0        | 12 | 1       | 0.41     |
Table 4 (on next page)

Comparison using BIC and BIC weight (or \textit{w-value}) of the homogeneity of TSD pattern.

(A) Within the East Pacific RMU and (B) at a global scale. Selected models are indicated in bold font.
A : Comparison within East Pacific RMU

| Series                              | BIC     | ΔBIC | BIC weight |
|-------------------------------------|---------|------|------------|
| All East Pacific grouped            | 38.06   | 0.00 | 0.87       |
| Mexico and Costa Rica separated     | 41.82   | 3.76 | 0.13       |

B: Worldwide comparison

| Series                              | BIC     | ΔBIC | BIC weight |
|-------------------------------------|---------|------|------------|
| World                               | 71.99   | 7.32 | 0.03       |
| East Pacific, Northeast Indian, and West Atlantic separated | 64.67 | 0.00 | 0.97 |
Table 5 (on next page)

Pivotal temperature and transitional range of temperatures (5%) for *Lepidochelys olivacea* RMUs.

Quantiles (0.025, 0.5, and 0.975) for pivotal temperature and transitional range of temperatures (5%) using maximum likelihood (MLE, upper line) and Bayesian MCMC (Bay. MCMC, lower line) with logistic and *flexit* model for TSD pattern.
| RMU          | Pivotal temperature in °C | Transitional range of temperatures in °C |
|--------------|---------------------------|-----------------------------------------|
|              | 0.025 | 0.5       | 0.975 | 0.025 | 0.5       | 0.975 |

| RMU | Logistic model |
|-----|----------------|
|     | East Pacific   |
|     | MLE            | 30.27 | 30.46 | 30.66 | 1.66 | 2.19 | 2.77 |
|     | Bay. MCMC      | 30.26 | 30.46 | 30.66 | 1.72 | 2.24 | 2.89 |
|     | Northeast Indian|
|     | MLE            | 28.74 | 29.27 | 29.50 | 0.17 | 3.28 | 10.89 |
|     | Bay. MCMC      | 28.72 | 29.35 | 29.76 | 0.34 | 1.72 | 4.76 |
|     | West Atlantic  |
|     | MLE            | 30.18 | 30.63 | 31.06 | 0.79 | 2.14 | 3.51 |
|     | Bay. MCMC      | 30.23 | 30.65 | 31.15 | 1.29 | 2.44 | 4.53 |
|     |                | Flexit model |
|     | East Pacific   |
|     | MLE            | 30.58 | 30.92 | 31.23 | 1.96 | 2.81 | 4.08 |
|     | Bay. MCMC      | 30.44 | 30.76 | 31.29 | 1.48 | 2.33 | 3.40 |
|     | Northeast Indian|
|     | MLE            | 29.66 | 34.31 | 45.71 | na   | na   | na   |
|     | Bay. MCMC      | 29.00 | 29.43 | 29.74 | 0.37 | 0.87 | 4.65 |
|     | West Atlantic  |
|     | MLE            | 30.12 | 30.36 | 30.56 | na   | na   | na   |
|     | Bay. MCMC      | 30.34 | 30.82 | 31.52 | 1.02 | 2.37 | 28.13 |
Figure 1

Map showing locations of data collection.

(1) Pirambu Beach, Sergipe State, Northeastern Brazil, (2) Odisha (formerly spelled as Orissa), India, (3) Nancite, Costa Rica, (4) La Escobilla, Mexico, (5) Playa La Destiladeras, Mexico, (6) El Verde Camacho, Sinaloa, Mexico.
Figure 2

Logistic (A) and flexit (B) models fitted with all olive ridley incubation data.

Confidence intervals were estimated using 10,000 random numbers obtained from the Hessian matrix. The standard errors for the $S$, $K_1$, and $K_2$ parameters in the flexit model were high and during resampling some trials were removed as their signs had reversed (see text). As a consequence, the estimated confidence interval is biased towards a lower value. The dark grey zone is the TRT 5% and the light grey zone is the 95% confidence of the TRT. The points correspond to observations and the bars are their 95% confidence intervals for the sex ratios according to the Wilson method (1927). The plain curve shows the maximum likelihood model and its 95% confidence interval is shown as dashed lines. The pivotal temperature is indicated by the vertical dash-dotted line.
A

Transitional range of temperatures $I=5\%$

Pivotal temperature

Male relative frequency

Logistic

$\text{AIC}_C = 69.52$

Akaike weight $= 0.03$

Temperature in °C

B

Transitional range of temperatures $I=5\%$

Pivotal temperature

Male relative frequency

Flexit

$\text{AIC}_C = 62.6$

Akaike weight $= 0.97$

Temperature in °C
Figure 3

Results of the Bayesian MCMC analysis for East Pacific *Lepidochelys olivacea* RMU.

(A, B): Distribution of priors (plain line) and posteriors (histograms) for $P$ (A) and $S$ (B) for East Pacific RMU. The covariation of posteriors for $P$ and $S$ is shown in (C).
Figure 4

TSD patterns as modeled by a *logistic* function and fitted using maximum likelihood. The dark grey zone is the TRT 5% and the light grey zone is the 95% confidence of the TRT. The points correspond to observations and the bars are their 95% confidence intervals. The plain curve shows the maximum likelihood model and its 95% confidence interval is shown as dashed lines. The pivotal temperature is indicated by the vertical dash-dotted line. Note that Mexico, East Pacific and Northeast Indian datasets each have only one temperature with mixed sex ratio.
Figure 5

TSD patterns modeled as a logistic function fitted using Bayesian MCMC.

The dark grey zone is the TRT 5% and the light grey zone is the 95% confidence of the TRT. The points correspond to observations and the bars are their 95% confidence intervals. The plain curve shows the maximum likelihood model and its 95% confidence interval is shown as dashed lines. The pivotal temperature is indicated by the vertical dash-dotted line.
Figure 6

Distribution of 100,000 posteriors for pivotal temperatures and transitional range of temperatures for the *logistic* model fitted using Bayesian MCMC.

A 20,000 random subsample of posteriors is shown. Ellipses including 75% of the points are drawn.
