Breadth of the thermal response captures individual and geographic variation in temperature-dependent sex determination

Anna L. Carter1 | Brooke L. Bodensteiner2 | John B. Iverson3 | Carrie L. Milne-Zelman4 | Timothy S. Mitchell5 | Jeanine M. Refsnider6 | Daniel A. Warner7 | Fredric J. Janzen1

1Department of Ecology, Evolution & Organismal Biology, Iowa State University, Ames, IA, USA; 2Department of Biological Sciences, Virginia Tech, Blacksburg, VA, USA; 3Department of Biology, Earlham College, Richmond, IN, USA; 4Department of Biology, Aurora University, Aurora, IL, USA; 5Department of Ecology, Evolution & Behavior, University of Minnesota, Minneapolis, MN, USA; 6Department of Environmental Sciences, University of Toledo, Toledo, OH, USA and 7Department of Biological Sciences, Auburn University, Auburn, AL, USA

Correspondence
Anna L. Carter
Email: acarter1@iastate.edu

Funding information
National Science Foundation, Grant/Award Number: DEB-1242510 and IOS-1257857; Minnesota Herpetological Society; British Ecological Society Small Research

Handling Editor: Craig White

Abstract
1. Population-scale responses of key ecological traits to local environmental conditions provide insight into their adaptive potential. In species with temperature-dependent sex determination (TSD), short-term, individual developmental responses to the incubation environment have long-term consequences for populations.

2. We took a model-based approach to study within- and among-population variation in the physiological components of TSD in 12 populations of painted turtles (Chrysemys picta). We used laboratory and field incubation data to quantify variation in thermal reaction norms at both population and clutch scales, focusing on the pivotal temperature that produces a 1:1 sex ratio (P) and the transitional range of incubation temperatures (TRTs) that produce mixed sex ratios.

3. Defying theoretical expectations, among-population variation in P was not convincingly explained by geography or local thermal conditions. However, within some populations, P varied by >5°C at the clutch scale, indicating that the temperature sensitivity of gonadal differentiation can vary substantially among individual nesting females. In addition, the TRT was wider at lower latitudes, suggesting responsiveness to local incubation conditions.

4. Our results provide a potential explanation for discrepancies observed between constant-temperature experimental results and outcomes of fluctuating incubation conditions experienced in natural nests, exposing important knowledge gaps in our understanding of local adaptation in TSD and identifying shortcomings of traditional laboratory studies. Understanding individual variation and the timing of gonadal differentiation is likely to be far more useful in understanding local adaptation than previously acknowledged.
1 | INTRODUCTION

How organisms respond to changing environments is a guiding question of macroecology. Species with broad geographic or elevational distributions often exhibit population-specific responses to local environmental conditions (Taylor, Ternes, & Lattanzio, 2018). Thus, detecting within- and among-population variation in important traits is key for understanding if, and how, species can respond to temporal or spatial environmental shifts. Within the context of contemporary climate change, thermally mediated traits are of particular importance (Kingsolver, Diamond, & Buckley, 2013; While et al., 2018). The ideal model trait for examining the intraspecific and interspecific effects of thermal heterogeneity is (a) fundamental to organismal physiology, (b) directly impacted by temperature in at least one measurable way, (c) intrinsically involved with key processes that extend beyond the trait itself and (d) subject to biologically significant, contemporary variation in thermal conditions (Angilletta, 2009; Kingsolver et al., 2013).

Temperature-dependent sex determination (TSD) meets the above criteria but is relatively understood as a thermal trait (While et al., 2018). In species with TSD, incubation temperatures directly impact both embryonic development rates and the outcome of gonadal differentiation at the organism scale. However, these short-term developmental outcomes have long-term implications at the population scale, since sex-specific recruitment translates to operational sex ratios on a time-lag of years or even decades (Chaloupka & Musick, 1997; Schwanz, Janzen, & Proulx, 2010). Temperature-dependent sex determination is also a model trait for assessing our ability to translate organism-scale physiology to population structure and species survival. Many organisms with TSD are subject to considerable spatial and temporal variation in developmental temperatures (Refsnider & Janzen, 2016), and several studies suggest that populations have responded in potentially adaptive directions to local climates (Doody et al., 2006; Ewert, Lang, & Nelson, 2005; Morjan, 2003; Pen et al., 2010).

In species with TSD, the pivotal incubation temperature ($P$, or $T_{pv}$, of some authors) is the temperature(s) at which offspring sex ratios are theoretically 1:1. Girondot (1999) hypothesized that nest-scale values of $P$ might not provide a good estimate for a population’s response, even to constant incubation conditions: albeit, this was based on observed variation in $P > 1^\circ C$ between two clutches of loggerhead turtles (Caretta caretta). However, he also noted that clutch-scale $P$ only varied by about 2$^\circ C$ under constant incubation (Mrosovsky, 1988) and urged caution when translating laboratory-derived $P$ to field conditions (Mrosovsky & Pieau, 1991). Nevertheless, $P$ has been a convenient proxy for describing the thermal reaction norm of sex determination at the population scale (Figure 1) (Bull, Vogt, & Bulmer, 1982; Ewert, Jackson, & Nelson, 1994; Mrosovsky, 1988; Pezaro, Doody, & Thompson, 2017).

If the TSD reaction norm is responsive to incubation temperatures in a way that would maintain roughly balanced primary sex ratios (Bull & Charnov, 1988), we would expect to observe an inverse relationship between $P$ and latitude, where latitude is a proxy for incubation temperature. In fact, both positive and negative trends in the intraspecific relationship between $P$ and geographic location (both latitude and longitude) have been identified for several species with TSD using data from constant-temperature incubation experiments (Bull et al., 1982; Ewert et al., 1994, 2005; Mrosovsky, 1988). However, natural nests typically experience wide diel fluctuations and seasonal variation in incubation temperatures, a reality that contemporary models of TSD attempt to account for (Carter, Sadd, Tuberville, Paitz, & Bowden, 2018; Georges, Beggs, Young, & Doody, 2005; Telemeco, Abbott, & Janzen, 2013). In addition, the value of $P$ can vary significantly among clutches in a single population (McGaugh & Janzen, 2011; Mrosovsky, 1988; Zaborski, Dorizzi, & Pieau, 1988). Reconciliation of this natural variation with experimental data from across geographically disparate populations nonetheless remains a significant research gap (Bowden, Carter, & Paitz, 2014).

A component of TSD that has received relatively little attention is the transitional range of temperatures (TRTs) (Ewert et al., 2005; Hulin, Delmas, Girondot, Godfrey, & Guillon, 2009). The TRT is the range of temperatures, again defined in terms of constant-temperature incubation, that produce mixed offspring sex ratios. Whereas $P$ is a function of the intercept of a reaction norm, the TRT is a function of its slope. Like $P$, the TRT varies among populations and contributes to variation in primary sex ratios (Mrosovsky & Pieau, 1991). However, $P$ and TRT are not strictly independent. When derived from constant incubation temperature experiments, $P$ is a population-scale, or at least a cohort-scale, metric. However, each clutch of eggs also has its own $P$ (Mrosovsky & Pieau, 1991), and within-population variation in $P$ is a potential mechanism for responding to climate-linked variation in incubation conditions (McGaugh & Janzen, 2011). In addition, $P$ is only a single theoretical point within the TRT (Mrosovsky & Pieau, 1991), analogous to the thermal optimum within a breadth of responses (Angilletta, 2009). The wider a population's TRT is, the less responsive the $P$ should be to changing incubation temperatures. Thus, intraspecific variation in the TRT may be a more realistic—and detectable—indicator of local resilience to environmental change than $P$. Consequently, we conducted a geographically broad-scale study of TSD to address the following predictions:

- In a geographically widespread species, the pivotal temperature and/or transitional range of temperatures varies within or among distinct populations.

### KEYWORDS
biogeography, constant-temperature equivalent (CTE), embryonic development, geographic variation, incubation, painted turtle (Chrysemys picta), reptiles, thermal reaction norm
female and, more rarely, intersex hatchings is produced within a single TRT (Figure 1). Previous studies have estimated $P$ for three U.S. populations of *C. picta* between 27.4 and 29.1°C (Ewert et al., 1994; Hulin et al., 2009; Refsnider, Milne-Zelman, Warner, & Janzen, 2014), with a TRT of 1.2–3.8°C (Hulin et al., 2009; Morjan, 2003; Refsnider et al., 2014).

For this study, we used data on hatching sex ratios from constant-temperature incubation experiments conducted intermittently from 1998 to 2014 on eggs ($N = 2$–30 per treatment) collected from eight populations of *C. picta*, primarily the *belli* subspecies, in the U.S. states of Idaho (ID), Illinois (IL), Kansas (KS), Minnesota (MN), Nebraska (NE), New Mexico (NM), Oregon (OR) and Virginia (VA). Field research methods can be found in Janzen (1994b) and Schwanz, Bowden, Spencer, and Janzen (2009). We also incorporated published data ($5 ≤ N ≤ 313$ per treatment) from populations in Indiana (IN) (Ewert & Nelson, 1991), Maryland (MD) (Dimond, 1983), North Dakota (ND) (Rhen & Lang, 1998), Wisconsin (WI) (Bull et al., 1982) and Ontario (ON) (Schwarzkopf & Brooks, 1985) and from a taxonomically divergent population (*C. dorsalis*) in Tennessee (TN) (Bull et al., 1982; Ewert et al., 1994), yielding data from 14 populations across >6.0 degrees of latitude and >6.5 degrees of longitude. A subset of these studies also included data on incubation temperatures and sex ratios from natural nests.

### 2.2 | Sex ratio model selection using constant incubation experiments

We used the `tsd()` function in the `embryogrowth` package (Girondot, 2018) to determine which of five potential TSD models (Godfrey, Delmas, & Girondot, 2003; Hulin et al., 2009) would provide the best fit to constant incubation data for each of the 14 populations using maximum likelihood (Girondot, 1999; Godfrey et al., 2003). We also fit a model to combine data from all 13 populations of *C. picta* (i.e. sans the TN population of *C. dorsalis*). Model goodness-of-fit was assessed using sample size-corrected Akaike’s information criterion (AICc) (Sakamoto, Ishiguro, & Kitagawa, 1986). Initial modelling indicated support for the logistic, A-logistic (also referred to as the Richards model) and Hill models (See Table S1). However, visual assessment of plots did not support the Hill model, so subsequent fitting for all populations was conducted using the A-logistic model (Equation 1) (Girondot, 1999; Godfrey et al., 2003):

$$sr(T) = [1 + (2^x - 1)e^{(\frac{1}{2})^{(P - T)}}]^{\frac{1}{S}}$$

where $T$ is the constant incubation temperature, $P$ is the population-specific pivotal temperature (°C), $S$ describes the shape of the sex ratio function as temperatures transition from male-producing to female-producing, and $K$ is a numeric constant equal to $2 \cdot \ln(x/1 - x)$ that allows for asymmetry around $P$, where the lower and upper boundaries of the TRT, respectively, produce sex ratios ($sr$) of $x$ and $1-x$ (Girondot, 1999; Godfrey et al., 2003). The $S$ parameter essentially...
describes the steepness of the slope of the response. When $S > 0$, 
sex ratios increase from 0 to 1, and sex ratios decrease from 1 to 0 
when $S < 0$ (Girondot, 1999). As $|S| \rightarrow 0$, the response becomes more 
vertical, and the width of the TRT decreases. Where $K = 0$, Equation 
(1) reduces to a logistic model (Equation (2)): 
\[
\text{sr} (T) = \frac{1}{1 + e^{\left(\frac{1}{x} \right) (T-s)}}
\]

The $S$ parameter is the main determinant of the width of the TRT 
(Girondot, 1999). We used $x = 0.05$, which defined the TRT as the 
range of temperatures that produce sex ratios of between 5% and 
95% male (Godfrey et al., 2003).

### 2.3 Model validation using natural nests

Using hourly substrate temperatures measured in intermittent 
seasons from 1995 to 2014 (Thermochron® iButton® data loggers 
DS1921G, accuracy/resolution ± 1°C/0.5°C; Maxim Integrated), we 
predicted values for the constant incubation temperature equiva-
 lent (CTE) during the estimated period of gonadal differentiation for 
201 natural nests from 6 populations via a nonlinear development 
function, previously validated for C. picta (Telemeco et al., 2013). 
The CTE model summarizes the diel thermal variation experienced 
in the wild (Georges, 1989) into a single value, which can then be 
used as the $T$ parameter in Equation (1) to predict hatchling sex 
ratios (Georges, 1989; Georges, Limpus, & Stoutjesdijk, 1994). We 
used temperatures from the month of July to estimate the timing of 
gonadal differentiation (Janzén, 1994a). Validation was limited to six 
populations for which both hourly temperature data and hatching 
sex ratios were available from the field: ID ($N = 56$), IL ($N = 7$), MN 
($N = 53$), NE ($N = 72$), NM ($N = 5$) and OR ($N = 8$) populations. Data 
were collected from 2012 to 2015, but not all populations were 
sampled in all four years.

Values for the CTE were estimated in MATLAB® R2018a 
(MathWorks®), using the code provided in Telemeco et al. (2013). We 
used mean diel temperatures and ½ the diel range for each nest to 
estimate daily CTEs and then took the median daily CTE value as the 
‘true’ CTE for each nest. The logged temperature data contained occa-
sional unrealistic values ($<0°C$ or $>60°C$). Temperatures $<14°C$ (critical 
thermal minimum) (Les, Paitz, & Bowden, 2007) or $>34°C$ (critical ther-
mal maximum) (Neuwald & Valenzuela, 2011) would always result in 
zero development, so extreme values would only affect CTE estimates 
if they shifted the location of the daily median. We tested the effect 
of those potential outliers by estimating hourly CTEs using three data-
sets: (a) the original dataset, with sporadic values $<0°C$ replaced by lin-
ear interpolation, but all high values included, (b) a dataset constrained 
to values $\leq 50°C$ and (c) a dataset constrained to values $\leq 40°C$. We also 
estimated CTEs for the first dataset $\pm 1°C$ to capture the extremes of 
potential error in the data loggers.

To determine how well the constant incubation models pre-
dicted offspring sex ratios in natural nests, we used generalized lin-
ear models (GLMs) with binomial error distribution and logit link. As 
our response variable, we took the absolute value of the difference 
between modelled and measured sex ratios as an index of prediction 
accuracy and then coded indices $>0.25$ as ‘0’ and indices $\leq 0.25$ as 
‘1’, regardless of whether hatchlings from a nest were predominantly 
male or female. We repeated the analysis using accuracy indices of 
0.50 and 0.75, the latter of which is most similar to the validation 
criterion used by previous studies, which tend to bin hatching sex 
ratios categorically as ‘all-male’, ‘all-female’ or ‘mixed’ (e.g. Telemeco 
et al., 2013). Our predictors were year, location and ‘model’ (which 
coded for whether the population-specific or ‘all population’ model 
was used, which temperature data were used and ± standard errors 
in $P$ and $S$; See Table S2d).

### 2.4 Individual variation

Since the relationship between sex ratio and temperature is essen-
tially linear within the TRT if the reaction norm is described by the 
logistic model, we used the point-slope equation to find $P$ for each 
nest, given the CTE estimated from measured hourly temperature 
data and true sex ratios. To be able to solve for $P$, we derived values 
for the slope from the straight-line segment of each of the six popu-
lation models used for validation. We used Mann–Whitney U tests 
(aka Wilcoxon rank sum tests), which estimate differences between 
two samples in terms of a ‘location shift’ in medians, to compare the 
distributions of $P$ and CTE between populations.

### 2.5 Geographic variation

We used linear models to test for geographic patterns in population-
specific $P$ and $S$, using geographic coordinates obtained by entering 
reported population locations as points in Google Earth, excluding the 
MD population, for which specific location information was not avail-
able and model fit was unreliable (likely due to the very low sample 
size). We also excluded the TN population, which belongs to a sepa-
rate evolutionary lineage, C. dorsalis (Jensen et al., 2015; Starkey et al., 
2003), giving us rough values of latitude and longitude for 12 popu-
lations of C. picta.

In addition to testing location coordinates as proxies for local 
temperatures, we explicitly assessed local thermal conditions. We 
downloaded hourly substrate temperature data from the microclim 
dataset (Kearney, Isaac, & Porter, 2014) and then extracted data for 
each population using the approximated geographic coordinates. 
The microclim dataset contains mechanistically modelled microcli-
mate conditions, computed from a global, 10 arcmin dataset of cli-
mate normals for the period 1961–1990 (Kearney et al., 2014). We 
summarized microclimate data as mean, minimum and maximum July 
temperatures at 1 cm above the surface and at depths of 0 cm (the 
surface), 5 cm and 10 cm for ‘soil’ and ‘sand’ substrates, the former of 
which included a 5 cm ‘organic cap’, and assuming no canopy shading 
at any of the sites.

All statistical analyses were conducted in r v 3.5.0 (R Core Team, 
2018). Plots were created using r package ggplot2 (Wickham, 2016), 
and the map showing locations of C. picta populations (Figure 2) was 
created in qgis v 3.0.3 (QGIS Development Team, 2018).
3 | RESULTS

3.1 | Sex ratio model selection using constant incubation experiments

Values for $P$ varied from 27.7 to 29.1°C among populations, supporting previous estimates (Table S1, Figure 2) (Hulin et al., 2009; Janzen & Paukstis, 1991; Refsnider et al., 2014; Schwarzkopf & Brooks, 1985). Values for $S$ (which indicates the symmetrical shape of the curve as the incubation temperature approaches $P$) varied from −0.7 to −0.1, and the width of the TRT varied among populations from 1.0 to 3.4°C (Table S1, Figure 2). Although the Hill model had the best relative goodness-of-fit for several populations, based on AIC alone (Table S1), graphical assessment showed that the Hill model was not a good fit for the data. Ultimately, we chose the logistic model, which had similar statistical goodness-of-fit and improved graphical fit for the majority of populations (Table S1). The incubation data for the MD population did not include enough data points (<28°C (i.e. clutches with sex ratios other than 0 or 1)) to fit a reasonable curve, so that population was not included in subsequent analyses (Figure 2). The value of $P$ for the TN population of C. dorsalis was 0.6°C lower than the minimum value for C. picta (Table S1).

3.2 | Model validation using natural nests

Removing outliers or constraining the upper limit of hourly temperatures had little effect on estimated CTE values. The maximum value was slightly higher when temperatures were constrained to ≤40°C than for the other datasets (33.08°C vs 32.90°C), but these values were well above the upper limit of the TRT. Including a potential error of −1°C had no impact on CTEs, but +1°C of error constrained the range of CTEs substantially (Figure S1).

Overall, the logistic models fit to constant-temperature incubation data did a mediocre job of predicting ‘true’ sex ratios from nests (Figure 3). Values for Kendall’s tau ($\tau_B$) between modelled and field-measured sex ratios were 0.28 for all of the population-specific models and 0.32 for the ‘all model’ predictions. When we defined prediction success at the 0.25 threshold, models predicted 56%–59% of ‘true’ sex ratios, and choice of model had little effect on predictive ability (See Table S2a). Not surprisingly, models that incorporated potential data logger error into the original temperature datasets were the exceptions. When we defined success at the 0.50 threshold, models predicted 68%–70% of sex ratios correctly (See Table S2b). At a threshold of 0.75, models predicted 75%–78% of ‘true’ sex ratios (See Table S2c). The odds of predictive success varied among years and for different populations (Table S2a–c).

3.3 | Individual variation

The estimated clutch-scale values of $P$ were approximated by population-scale values of $P$, though the latter values tended towards the low tail of the full distribution (Figure 4a), indicating that modelling the outcomes of constant-temperature incubation has, so far, not realistically captured the range of reaction norms found in nature. This result is consistent with model validation. Nest-scale $P$ differed among populations ($\chi^2 = 20.19, df = 5, p < .01$) and years ($\chi^2 = 12.37, df = 3, p < .01$), whereas CTEs varied among populations ($\chi^2 = 51.97, df = 5, p < .01$) but not years ($\chi^2 = 3.33, df = 3, p = .34$). The distributions of nest-scale $P$ and CTEs differed for about half of population pairs (See Table S3; Figure 4a). Overall, nest-scale values of $P$ and the CTE were strongly related ($r = .92$; Figure 4b). Even though these two variables are not independent, since CTEs were used to estimate $P$, the strength of their linear relationship across all populations was not assumed.

3.4 | Geographic variation

When 12 populations of C. picta were included in analyses, linear models found no effect of latitude, longitude, longitude$^2$, longitude$^3$ or the latitude:longitude interaction on $P$ or S. More southern populations had slightly lower values for $P$ and slightly wider TRTs (Figure 5). However, location only explained about 10% of the variation in $P$ ($F = 1.26, df = 6, R^2 = .10, p = .39$) among populations. The goodness-of-fit of logistic models was $\infty$ for the NM, ND and VA populations, probably due to low sample sizes at a limited range of constant incubation temperatures (Table S1). When those populations were removed from linear models, location explained 33% of the variation in $P$ ($F = 1.99, df = 4, R^2 = .33, p = .26$), if the third-order polynomial term for longitude was excluded. When those same populations were removed from the linear model of $S$, location explained just over half of the variation in the TRT ($F = 4.18, df = 5, R^2 = .54, p = .08$).

Since summarized temperatures for sand and soil substrate types were highly correlated ($r_s = 0.70–0.84$), we only included values for ‘sand’ in linear models (Figure 6). Microclimate temperatures did not explain variation in $P$ ($F = 0.02, df = 7/52, R^2 = -.13, p = 1$), regardless of how many populations were modelled. However, substrate temperatures, but not above-surface temperatures, were inversely associated with TRT values, when all 12 populations of C. picta were
included \(F = 2.14, df = 7/52, R^2_{adj} = .12, p = .06\); See Table S4). Including temperature-depth interactions in models eliminated statistically significant effects of coefficients for both \(P\) \((F = 0.12, df = 19/40, R^2_{adj} = -.39, p = 1)\) and \(S\) \((F = 1.39, df = 19/40, R^2_{adj} = .11, p = .19)\).

## DISCUSSION

### 4.1 | Re-thinking the use of the pivotal temperature

Population-scale responses of key ecological traits to local environmental conditions can provide insight into their adaptive potential. In species with TSD, short-term, individual-scale responses to the incubation environment have long-term, population-scale consequences (Mitchell & Janzen, 2010). As expected, distributions of nest-scale values of \(P\) and their associated CTEs varied among populations, despite, contrary to our prediction, the lack of a clear relationship between \(P\) and local temperature normals. Like many other thermal traits, gonadal differentiation responds to incubation conditions that fluctuate at fine spatial and temporal scales that are not necessarily captured by relatively broad summaries of thermal data (Helmuth, 2009). Ultimately, \(P\) may be less important to describing TSD systems in nature than has previously been implied under the assumption of an ‘optimal’ 1:1 sex ratio, a concept that has driven research on sex determination since Darwin (Edwards, 1998). Instead, when examining the population- or species-level implications of TSD within the context of contemporary climate change, we may be better served by considering the TRT, recognizing that \(P\) is merely a special case within that range (Mrosovsky & Pieau, 1991).

Previously identified geographic trends in \(P\) (Ewert et al., 1994, 2005) are likely an artefact of sampling bias (e.g. fitting data from only a few populations) (Gienger, Dochtermann, & Tracy, 2018) or experimental protocol (Janzen & Paukstis, 1991), rather than reflective of a biologically significant response to thermal conditions. In this study, interpopulation variation in \(S\) was more closely linked with location than was variation in \(P\). In addition, the TRT tended to be wider at lower latitudes (Figure 5), which is what we would expect if the transitional range was responsive to local conditions (though temperature itself did not convincingly explain that pattern, and estimation of the TRT was probably also impacted by sample size). In terms of adjusting the primary sex ratio, selection on \(P\) will be weaker as the TRT widens, especially when—as in \(C.\ picta\)—the reaction norm of sex determination is symmetrical, or nearly so, around \(P\). An adaptive response in a TSD type I system might be more readily indicated by an increasingly asymmetrical TRT, regardless of whether the population-scale value of \(P\) changes over time. In terms of fitting TSD models, that shift would manifest in the \(K\) constant (Girondot, 1999).

### 4.2 | Highlighting individual variation

Even though geography does not satisfactorily explain variation in \(P\) at the population scale, we detected between-population shifts in both \(P\) and CTEs at the nest scale. By focusing solely on population-scale \(P\) as
a proxy of a TSD system, we are missing intrapopulation variation that drives the width of the TRT and may be important for adjusting primary sex ratios in variable environments. Constant incubation experiments have detected heritable sex ratios in other reptile species (Bull et al., 1982; Janzen, 1992; Rhen, Schroeder, Sakata, Huang, & Crews, 2011). However, given the divergence of nest-scale responses from population values identified in this study, heritability estimates may differ under naturally fluctuating incubation conditions (McGaugh & Janzen, 2011).

The wide range of nest-scale $P$ estimated in this study, >5°C for some populations, suggests that the temperature sensitivity of gonadal differentiation varies substantially, at least at the scale of individual females. Yolk estradiol can vary seasonally and with maternal age, which could contribute to differences in hatching sex ratios at the nest scale and a seasonal bias in estimations of $P$ (Bowden, Ewert, & Nelson, 2000; Carter, Bowden, & Paitz, 2017; Carter et al., 2018). In addition, inhibiting aromatase activity broadens the TRT, producing mixed sex ratios across a wider range of incubation temperatures (Warner, Mitchell, Bodensteiner, & Janzen, 2017). Linking maternal effects such as egg estradiol content with developmental physiology and nesting behaviour, and then testing for repeatability and/or heritability under fluctuating incubation conditions could elucidate a complex pathway for moderating hatching sex ratios in nature.
Thermally mediated developmental plasticity in reptiles is an area of increasing research focus (While et al., 2018). However, relatively few studies have attempted to simulate fluctuating incubation temperatures in a laboratory environment within the context of TSD (Bowden et al., 2014; Carter et al., 2018; Janzen & Paukstis, 1991; McGaugh & Janzen, 2011; While et al., 2018). Fluctuating temperatures consistently lead to higher CTEs, more female hatchlings and longer total incubation periods (Les et al., 2007). However, these studies have constrained incubation temperatures to remain within the range of temperatures that facilitate linear development (Georges et al., 2005, 1994). Temperatures outside of this range, where embryonic growth slows—potentially at different rates—as the critical thermal minimum/maximum values are approached, are regularly experienced in natural nests (Refsnider et al., 2014; Schwarzkopf & Brooks, 1985). Fluctuating incubation conditions may have unexpected impacts on sex determination, as seen in the relationship between pivotal temperature and the width of the transitional range of temperatures (quantified by the S parameter) among populations of painted turtles (Chrysemys picta). The grey curves on the 'Longitude' plots show second-order (dashed) and third-order (dotted) polynomial functions, which more closely represent the spatial distribution of temperature conditions across North America. The colour scale indicates the deviation of each population's estimated pivotal temperature from the 'all population' model (28.24°C). Figure created in Gimp v2.10.4 (GIMP Team, 2018). [CC BY 4.0 | https://doi.org/10.6084/m9.figshare.8168441]

4.3 | Re-focusing on the thermosensitive period

Fluctuating temperatures consistently lead to higher CTEs, more female hatchlings and longer total incubation periods (Les et al., 2007). However, these studies have constrained incubation temperatures to remain within the range of temperatures that facilitate linear development (Georges et al., 2005, 1994). Temperatures outside of this range, where embryonic growth slows—potentially at different rates—as the critical thermal minimum/maximum values are approached, are regularly experienced in natural nests (Refsnider et al., 2014; Schwarzkopf & Brooks, 1985). Fluctuating incubation conditions may have unexpected impacts on sex determination, as seen in the relationship between pivotal temperature and the width of the transitional range of temperatures (quantified by the S parameter) among populations of painted turtles (Chrysemys picta). The grey curves on the 'Longitude' plots show second-order (dashed) and third-order (dotted) polynomial functions, which more closely represent the spatial distribution of temperature conditions across North America. The colour scale indicates the deviation of each population's estimated pivotal temperature from the 'all population' model (28.24°C). Figure created in Gimp v2.10.4 (GIMP Team, 2018). [CC BY 4.0 | https://doi.org/10.6084/m9.figshare.8168441]
ratios, relative to what can be observed in the laboratory. For example, very short ‘hot’ periods can disproportionately impact sex ratios (Carter et al., 2018). Predicted sex ratios can also vary substantially based on how developmental period and incubation duration are defined (Fuentes et al., 2017).

One potential explanation for this disparity between experimental results and expected outcomes in the wild—both in this and previous studies—is the relatively little attention given to the timing of gonadal differentiation, compared with incubation conditions themselves. In species with TSD, gonadal differentiation is triggered within a particular developmental window: the thermosensitive period(s) or TSP (Mrosovsky & Pieau, 1991; Pieau & Dorizzi, 1981). By definition, while temperatures experienced before and after the TSP impact development, they do not ultimately affect sex. The length of the TSP varies depending on incubation conditions, since development rate itself depends on temperature (Mrosovsky & Pieau, 1991; Nelson, Keall, Refsnider, & Carter, 2018). One of the consequences of avoiding incubation conditions outside of the linear range of embryonic growth during most laboratory experiments is that we may be mis-targeting the TSP, a potential explanation for the disparity between laboratory and field estimates of $P$ found here.

Constant incubation conditions essentially neutralize the TSP. That is, gonadal differentiation will occur in the same way, regardless of when it occurs, because the CTE is simply the incubation temperature. Experiments that switch between constant incubation conditions can be used to identify the developmental limits of temperature sensitivity, including how development rates respond to increasing/decreasing temperatures. However, predicting the timing of the TSP is complicated under fluctuating thermal conditions, especially if the magnitude of fluctuations is high enough to shorten or shift the timing of the TSP (Girondot, Monsinjon, & Guillon, 2018). Since predicting future hatchling sex ratios in species with TSD is important within the context of climate change, the most informative future project may be a simulation study to develop a definition of the TSP that can be translated from laboratory to field incubation conditions.

5 | CONCLUSIONS

Studies on the evolution of plasticity often quantify slopes and intercepts to characterize linear reaction norms but do not always account for among-individual variation. Here, we show how this oversight limits our ability to capture intrapopulation variation that drives the shape (i.e. the slope or, in this case, TRT) of a reaction norm. In the case of TSD, that variation could be extremely important for understanding the adaptive potential of this fundamental thermal response.

ACKNOWLEDGEMENTS

This research was supported by National Science Foundation grants DEB-1242510 and IOS-1257857 to FJJ and a Grant in Herpetological Conservation and Research from the Minnesota Herpetological Society and British Ecological Society Small Research Grant to ALC. We would like to thank U.S. Fish and Wildlife Service for collection permits, staff at our study areas and numerous field assistants. We would also like to thank Turk Rhen and Robin Andrews for providing additional raw data, Karen Abbott for assistance with improving the MATLAB code, Rory Telemece and Marc Girondot for valuable discussions, and members of the Janzen and Bronikowski laboratories at Iowa State University for comments on a previous draft. This study was performed in accordance with the Iowa State University Care and Use of Animals Protocols.

AUTHORS’ CONTRIBUTIONS

F.J.J. and D.A.W. designed the study; F.J.J., D.A.W., J.M.R., T.S.M., C.L.M.Z., J.B.I. and B.L.B. conducted field and laboratory components; A.L.C. and B.L.B. analysed the data; and A.L.C. led the writing of the manuscript. All authors contributed critically to drafts and gave approval for publication.
DATA AVAILABILITY STATEMENT

All data and code are available at figshare https://doi.org/10.6084/m9.figshare.7528772.v1 (Carter et al., 2019).

ORCID

Anna L. Carter https://orcid.org/0000-0003-4392-277X
Brooke L. Bodensteiner https://orcid.org/0000-0001-6628-1923
John B. Iverson https://orcid.org/0000-0002-1755-5438
Timothy S. Mitchell https://orcid.org/0000-0002-7136-769X
Jeanine M. Refsnider https://orcid.org/0000-0001-5154-4356
Daniel A. Warner https://orcid.org/0000-0001-7231-7785
Fredric J. Janzen https://orcid.org/0000-0002-5919-196X

REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of the article.