DR. ANNA CARTER (Orcid ID : 0000-0003-4392-277X)

Article type : Methods and Tools

Original Article

Geostatistical interpolation can reliably extend coverage of a very high-resolution model of temperature-dependent sex determination

Running title

Kriging a biophysical model

Authors

Anna L. Carter^{1*§}, Michael R. Kearney², Stephen Hartley¹, Warren P. Porter³ and Nicola J. Nelson¹

Author affiliations

¹School of Biological Sciences, Victoria University of Wellington, PO Box 600, Wellington 6140, New Zealand

[§] Present address: Department of Ecology, Evolution & Organismal Biology, Iowa State University, 2200 Osborne Dr., Ames, IA 50011, USA

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/jbi.13152](https://doi.org/10.1111/jbi.13152)

This article is protected by copyright. All rights reserved

23 ²School of Biosciences, University of Melbourne, Victoria 3010, Australia

24 ³Department of Zoology, The University of Wisconsin, Madison, WI 53715, USA

25 *corresponding author: acarter1@iastate.edu

26

27 **Word count**

28 Main text, including abstract and references: 6,167

29 Pages for figures and legends (estimated): 2.5

Acknowledgements

This research was funded by a Victoria University of Wellington International Doctoral Scholarship and Science Faculty Strategic Research Grant to ALC. The Victorian Life Sciences Computing Initiative (VLSCI) provided access to high-performance computing facilities essential to this study. Support for manuscript preparation was provided by the Allan Wilson Centre. We would like to thank field assistants and colleagues at VUW and the Janzen lab at Iowa State University. We would also like to acknowledge the support of Ngāti Kōata iwi and Department of Conservation staff on Takapourewa.

Abstract

Aim

Recognition that statistical models do not always reliably predict habitat suitability under future climate scenarios is leading increasingly to explicit incorporation of the physiological constraints that underlie species' distributions into spatially-explicit predictions. However, computational intensity constrains the use of high-resolution, process-explicit models. We examined whether geostatistical analysis can effectively interpolate a biophysical model, reducing the computational investment typically required for using mechanistic methods to inform physiological predictions.

Location

New Zealand [40°40'00" S 174°00'00" E]

Methods

We used a spatially-explicit, mechanistic microclimate model to predict hourly temperatures at five soil depths under two scenarios of climate warming. Using the predicted soil temperatures as input to a biophysical model of temperature-dependent embryonic development, we estimated incubation temperatures and corresponding hatchling sex ratios for tuatara, a reptile with temperature-dependent sex determination, at a sub-meter horizontal spatial resolution. We then applied ordinary kriging, a robust method of geostatistical interpolation, to estimate predictions throughout the full extent of our study location, an additional 480,000+ microsites, and validated the interpolation against an independent set of predictions.

Results

Ordinary kriging accurately predicted spatial variability in incubation temperatures. Mean predictions were similar between methods, and error in the geospatial model generally decreased with increasing soil depth. Error was higher for the geospatial model of the 'maximum warming,' compared with the 'minimum warming,' scenario of climate change.

Main conclusions

Our results show that ordinary kriging can be a reliable method for interpolating variability in high-resolution predictions. However, the effects of error on the accuracy of interpolated predictions will become more severe as values approach a physiological threshold, such as the minimum and maximum incubation temperatures that result in extreme sex-ratio bias. For distribution models, the widths of geographic areas predicted to be suitable for, in this case, maintaining balanced sex ratios, compared to those predicted to be unsuitable, may be narrower than in reality.

77 **Keywords**

78 Dallwitz-Higgins development rate; mechanistic microclimate model; ordinary kriging;
79 species distributions; temperature-dependent sex determination (TSD); tuatara (*Sphenodon*
80 *punctatus*)

Introduction

Spatially-explicit models of habitat suitability have primarily been developed using bioclimatic envelopes and other statistical descriptions of relationships between the environment and a species' known distribution (Guisan & Zimmerman, 2000; Guisan & Thuiller, 2005; Keith et al., 2008). However, the ability of a population to persist in a given location depends on complex interactions between the micro-environment and biological processes that manifest at both population (e.g., Araújo & Luoto, 2007; Lavergne et al., 2010) and individual (e.g., Kearney, 2006; Porter et al., 2010) levels in ways that may not be apparent from distribution data. Because statistical models may not reliably predict habitat suitability under future climate scenarios (Thuiller, 2004; Keith et al., 2008), the physiological processes that determine the spatial limits of the fundamental niche are increasingly being incorporated explicitly into predictive models (e.g., Kearney & Porter, 2009; Mitchell et al., 2013). Here, we use 'mechanistic' and 'process-explicit' interchangeably to refer to models that make predictions based primarily on first principles, rather than statistical relationships.

Two of the key considerations in developing a process-explicit model are whether the spatial resolution at which predictions are made is (i) appropriate to the specific question(s) being posed and (ii) biologically relevant for the taxon of interest (Dormann et al., 2012). For example, examining species' responses to environmental extremes (Kearney et al., 2012) or at range edges requires finer-resolution predictions of microclimate conditions than can be attained by simply modelling distributional limits within available environments. In addition, the coarser the spatial resolution of a predictive surface, the more within-pixel variation is masked and, as a result, model uncertainty increased (Carter et al., 2015). For example, if a population is bounded within an area of 1km^2 , a model built using 1km^2 -gridded climate variables (very high resolution on a continental scale) could not capture effects of environmental heterogeneity on that population.

The spatial resolution required to achieve biological relevance may vary among different species as well as for different life stages of the same species. For example, a constraint on habitat suitability for egg-laying organisms is the availability of microhabitats with temperature and moisture conditions that will support successful embryonic development. In species with environmentally-mediated embryonic development and sex determination, incubation conditions must facilitate successful incubation and maintain hatchling sex ratios of approximately 1:1 (e.g., Mitchell et al., 2009). While a spatial resolution of 1km² may be sufficient for modelling long-term distributional shifts of adults, it would not contribute to understanding the range of incubation conditions experienced by sub-surface nests contained within that area. To be informative, a predictive model would need to be built at a spatial resolution sufficiently fine to capture enough within-pixel variability of sub-surface temperatures to reliably predict physiological outcomes.

A major constraint on developing climate and microclimate layers at high resolutions or within large extents, regardless of the method used, is computational intensity (Kearney & Porter, 2009; Kearney et al., 2014; Palmer, 2014). Running continental and global-scale models at even relatively low spatial resolutions (e.g., 1-5 km²) requires a supercomputer or high-throughput, grid-based computing platform (e.g., Stainforth et al., 2002; Kearney et al., 2014). Increasing the spatial resolution of a process-explicit model to generate taxon-specific predictions at extremely high spatial resolutions (on the order of a few meters or less), even within a spatial extent of only a few km², has similar computational requirements.

With the recent growth of internet-based, cloud-computing platforms such as Amazon Web Services (<https://aws.amazon.com/>) and Google Cloud Platform (<https://cloud.google.com>), the availability of high-throughput computing is not a hard constraint on increasing the spatial resolution of process-explicit models. However, model building may be limited by a lack of either (i) the time and expertise required to build or adapt complex models for internet-based

cloud computing platforms or (ii) the time and financial resources needed for ongoing access to cloud computing.

The spatial interpolation techniques used to create large-scale, low-resolution climate layers may also increase the coverage of a high-resolution model using standard computing resources. Gridded climate layers with continental or global coverage, such as WorldClim (www.worldclim.org), are generated by statistically smoothing data from weather stations across a continuous surface (Boer et al., 2001; Hijmans et al., 2005; Hancock & Hutchinson, 2006), incorporating multiple geospatial and environmental covariates (Hutchinson, 1991; Jarvis & Stuart 2001a, 2001b). Compared with process-explicit methods that generate independent predictions for every grid cell, statistical estimation requires much less computational investment.

Surfaces are estimated via a variety of methods, which can be broadly classified as global or local, exact or inexact, and deterministic or stochastic (summarised in Li & Heap, 2008). Global interpolators generate unknown values using all available data, while local interpolators constrain estimates to within a defined extent around a subset of known points (Burrough et al., 2015). Exact interpolators force the fitted surface to pass exactly through known values, while inexact interpolators may allow estimated values to vary from observed data (Burrough et al., 2015). Deterministic techniques of interpolation fit surfaces using defined mathematical functions with a single solution. For example, polygons can be used to delineate a two-dimensional 'nearest neighbour' window around each data point, within which all unknown values are interpolated as equal to the closest sampled value (Burrough et al., 2015). In contrast, stochastic models account for randomness in the data, which allows interpolated values to vary within a probability window. As a result, stochastic methods of interpolation can provide estimates of prediction accuracy, while deterministic methods do not.

162

163 A guiding principle of geostatistical interpolation is that continuous spatial data can be more
164 accurately modeled by local algorithms, rather than by a single smoothing function
165 (Burrough et al., 2015). Ordinary kriging is a robust method of geostatistical interpolation that
166 quantifies autocorrelation as a function of distance by fitting a locally-weighted
167 semivariogram model to a given set of data points (Cressie, 1988; Wackernagel, 1998; Li &
168 Heap, 2008; Burrough et al., 2015). Like the thin-plate smoothing splines used to estimate
169 the WorldClim layers (Hijmans et al., 2005), ordinary kriging can be considered an exact
170 interpolator. That is, the prediction surface passes exactly through the sampled values, and
171 the prediction standard error at those points is zero (Burrough et al., 2015). Ordinary kriging
172 is also a stochastic method, so can provide estimates of model error that are critical to
173 quantifying whether interpolated values are accurate enough to be useful as ecological
174 models. Finally, and importantly for ecologists who are not GIS specialists, ordinary kriging is
175 relatively simple to implement and less vulnerable than other kriging methods to researchers'
176 decisions, such as the number of sample points used as input and the radius of the local
177 search algorithm (Li & Heap, 2008; Burrough et al., 2015).

178

179 Various methods of spatial interpolation have been used to estimate large-extent climate
180 surfaces. In ecology, spatially-explicit methods are commonly employed to understand the
181 spatial dependence of field data and to model species' distributions. However, we know
182 comparatively little about the performance of geostatistical interpolation for estimating
183 surfaces directly from physiological parameters (but see Hernandez-Stefanoni & Ponce-
184 Hernandez, 2006). Our aim here was to examine geostatistical interpolation as a means of
185 extending the geographic extent of a very high resolution, spatially-explicit model via
186 standard computing resources. We used temperature-dependent embryonic development
187 and sex determination in reptiles as a model physiological system. In species with
188 temperature-dependent sex determination (TSD), the fluctuating thermal environment in

which embryos develop directly impacts both the rate of embryonic development and process of sexual differentiation at the scale of individual nests. Thus, mapping the horizontal and vertical distributions of below-ground thermal conditions at a fine spatial resolution is critical for predicting the impacts of environmental variation, including climate change, on hatchling sex ratios.

Materials and Methods

Study system

The New Zealand-endemic tuatara (*Sphenodon punctatus*, Gray 1831) is the only extant species known to have a pattern of TSD in which 100% male offspring are produced at warmer incubation temperatures. Female offspring are produced at relatively cool temperatures (Cree et al., 1995; Nelson et al., 2004a). Mixed sexes are only produced within a window of 1 °C (Nelson et al., 2004a, Mitchell et al., 2006), surrounding a ‘pivotal’ temperature of around 22 °C, at which offspring sex ratios are theoretically 1:1 (Mitchell et al., 2006). In tuatara, temperatures that produce female-biased or mixed-sex clutches in other species with TSD can lead to highly male-biased offspring sex ratios (Nelson et al., 2004b; Mitchell et al., 2009). Thus, modelling hatchling sex ratios in tuatara can provide a template for understanding whether other species with TSD could be at long-term risk of demographic impacts of climate warming. Since incubation temperatures can be translated directly to sex ratios in species with TSD, we could test ordinary kriging as a means of estimating values of a variable for which climate change could have long-term, population-level consequences. In addition, because mixed sexes are only produced within a 1 °C window in tuatara, the effects of error in the kriging model would be more extreme than for species in which the transitional range of temperatures is up to 2-3 °C, providing a worst-case scenario for the consequences of model error on ecological conclusions.

Biophysical Model

Microclimate component

We modelled substrate temperatures for Takapourewa (also referred to as Stephens Island), New Zealand [40°40'00" S 174°00'00" E], a 150 ha offshore island that supports a population of 30,000+ tuatara (Newman, 1987). We used a localised version of the NicheMapR mechanistic microclimate model, which solves heat-mass balance equations for up to ten soil depths (Carter et al., 2015; Kearney & Porter, 2016) (Fig. 1). We parameterised the model using local, daily climate data downloaded from the NIWA CliFlo database (see Data Accessibility) [Station #s 26169, 12430; <http://www.cliflo.niwa.co.nz>], generating hourly soil temperatures for three climate scenarios: (1) a 'current climate' scenario, using *in situ* weather station data from the year 2000, (2) a 'minimum warming' scenario predicted for the next century, in which minimum and maximum air temperatures were increased by the predicted range of 0.3 - 0.9°C, with an increase of 0.3°C in the meteorological austral spring, a 0.6°C increase in autumn and winter, and an increase of 0.9°C in summer and (3) a 'maximum warming' scenario, in which minimum and maximum temperatures were increased by the predicted range of 4.8 - 5.6°C, with an increase of 4.8°C in the austral spring, a 5.0°C increase in autumn and winter, and a 5.6°C increase in summer over the next 100 years (MFE, 2008). Available shade was simulated as 95% for forested microsites and 0% otherwise, with sites classed as 'forest' or 'non-forest' according to a spatially-explicit vegetation map (Carter et al., 2015). The model under-predicted substrate temperatures, compared with measured values (Carter et al. 2015), so modelled soil temperatures for all climate scenarios were increased by 2°C at 200mm and +/- 0.5°C for, respectively, every 50mm reduction or increase in soil depth from 50-300mm (Carter et al., 2015).

Geospatial data (latitude, longitude, elevation, slope, aspect, viewshed) were extracted from an approximately 0.5m-resolution gridded digital elevation model (DEM) of the island (Aerial Surveys Ltd., Auckland, NZ, 2011), generating 481,967 unique microsites for which the microclimate model could be run. Soil temperatures were predicted for ten soil depths: 0, 15, 50, 100, 150, 200, 300, 500, 1000 and 2000mm. Definitions for ten soil depths were required to run the model. However, nest depth ranges from 50 to 250mm (Cree et al., 1991), with a mean depth of 110mm (Nelson et al., 2004a). Thus, modelled temperatures at depths above 50mm or deeper than 300mm were not used for this study.

Developmental component

We used predicted soil temperatures to calculate hourly embryonic development via the Dallwitz-Higgins non-linear development rate function (Dallwitz & Higgins, 1992; Georges et al., 2005):

$$\frac{ds}{dt} = \frac{\left\{ b_1^{10[-v^2(1-b_5+b_5v^2)]} \right\}}{24}$$

$$v = \frac{(u + e^{b_4u})}{c_2}$$

$$u = \left(\frac{T - b_3}{b_3 - b_2} \right) - c_1$$

$$c_1 = \frac{1}{[1 + 0.28b_4 + 0.72 \ln(1 + b_4)]}$$

$$c_2 = \frac{1 + b_4}{1 + 1.5b_4 + 0.39b_4^2}$$

where 'b₁' is the maximum development rate; 'b₂' is the temperature at which development is approximately b₁ · 0.10; 'b₃' is the temperature (°C) at which the maximum development rate is reached; 'b₄' is the shape of the curve as the development rate approaches 0 at very low temperatures; 'b₅' is a measure of the asymmetry of the development rate; 'c₁' and 'c₂' are location-specific constants; and 'T' is the temperature for which the development rate is

This article is protected by copyright. All rights reserved

calculated (Dallwitz & Higgins 1992). Parameter values used in this study; $b_1 = 0.98545$, $b_2 = 13.78641$, $b_3 = 30$, $b_4 = 60$, $b_5 = 0.40$, $c_1 \approx 0.24503$, $c_2 \approx 1.24958$; were previously derived for tuatara on Takapourewa (Mitchell et al. 2008).

Hourly development was accumulated over time from an oviposition date of 30 November for two years or until hatching (Cree et al., 1989). The constant incubation temperature equivalent (CTE), a lab-equivalent incubation temperature that approximates the mid-point of embryonic development (Georges et al., 1994) and most accurately estimates hatchling sex ratios, was taken to be the soil temperature corresponding to the median developmental point within the 25-35% developmental window, the period during which sex is determined in tuatara (Nelson et al., 2010). Each CTE was converted to a spatially-explicit estimate of offspring sex ratio, sr , using the A-logistic function (Girondot 1999; Godfrey et al. 2003):

$$sr(t) = \left[1 + \left(2^{e^k} - 1 \right) e^{\left(\frac{1}{s} \right) (p-t)} \right]^{-\frac{1}{e^k}}$$

where 't' is the CTE calculated for a given point; 'p' is the species- or population-specific pivotal temperature (°C); 's' describes the shape of the sex ratio function, either positive or negative, as temperatures transition from male-producing to female-producing; and 'k' is a numeric constant equal to $2 \cdot \ln\left(\frac{x}{1-x}\right)$ that allows for asymmetry around p, where the lower and upper boundaries of the range of CTEs that result in mixed-sex nests, produce sex ratios of, respectively, x and 1-x. Parameter values used in this study; $p = 22.03$, $s = 0.0096$, $k = 3.60$; were previously derived for tuatara on Takapourewa (Mitchell et al. 2006).

We modelled hourly soil temperatures mechanistically for all 480,000+ microsites for the current climate scenario to determine the minimum area of viable nesting space under the coolest air temperatures modelled. Normally, locations at which development would not reach 100% within 24 months would be classified as non-viable nesting locations (Mitchell et al., 2006). However, since we were interested in the accuracy of interpolated values, rather

than the outcomes of the biophysical model per se, we used all points for which a CTE could be calculated (i.e., points at which development reached at least 35%). The whole-island model was run using the Victorian Life Sciences Computation Initiative (VLSCI) Peak Computing Facility at the University of Melbourne, Australia.

Geostatistical model

Data for climate change scenarios were generated mechanistically for a subset of 500 microsites to be used as training points for the kriging model (see Data Accessibility). The 2°C error correction was applied by depth, as with the scenario for the current climate (Carter et al., 2015). Using 500 sites limited computation time to approximately four hours using a consumer-grade laptop computer (2.53 GHz Intel Core 2 Duo processor with 4GB RAM), compared with the less than two minutes required to complete approximately the same number of sites using the high-throughput computing facility. Training points were sampled randomly from the whole-island dataset at a minimum separating distance of 2m, so multiple points would not be generated within the same 1m² cell. Predictions under the minimum warming scenario were nearly identical to those generated for the current climate and relatively cool, so the spatial extent of the minimum warming scenario was constrained to grid cells for which a CTE could be calculated under the current climate model. Thus, sites that would remain too cool to facilitate development under the scenario of minimum warming would not be erroneously modelled. Reducing the spatial extent of the minimum warming model decreased the total number of training points from 500 to 472 for that scenario. The scenario of maximum climate warming was simulated within the full spatial extent of the island. Values for CTEs and offspring sex ratios estimated using the geostatistical model are hereafter referred to as ‘interpolated,’ whereas those generated using mechanistically-modelled soil temperatures are referred to as ‘modelled.’

Kriging models were optimised to the training points using leave-one-out cross validation, then validated independently using developmental predictions derived from mechanistically-generated soil temperatures for a second set of 500 randomly sampled points. Accuracy of the interpolated surfaces was assessed using the root-mean-squared deviation (RMSD) between modelled and interpolated predictions, calculated using the *rmse* function in R package 'hydroGOF' (Zambrano-Bigiarini, 2014). Each CTE was converted to a spatially-explicit estimate of offspring sex ratio via the A-logistic function (Girondot, 1999; Godfrey et al., 2003; Mitchell et al., 2006). We used bootstrapped Kolmogorov-Smirnov tests (Sekhon, 2011) in R package 'Matching' (Sekhon, 2011), which compute p-values based on the largest variation between distributions. We also used Mann-Whitney tests, which compare mean ranks, to determine whether predictions differed statistically.

All modelling and statistical analyses were conducted using R v 3.0.3 software (R Development Core Team, 2008). Geostatistical modelling was conducted with the Geospatial Analyst extension (ESRI, 2003) in ArcMap™ Desktop v 10.1 software (Environmental Systems Research Institute [ESRI] ArcGIS: Redlands, CA, USA, 2012). Figures and maps were generated in R and ArcMap™, respectively, and edited using the GNU image manipulation program (GIMP Development Team, 2013). Statistical outcomes are reported at the 95% confidence level.

Results

Mean prediction errors were near zero for all soil depths under both climate warming scenarios, indicating that predictions interpolated from the training data were unbiased (Burrough et al., 2015). Mean standard errors were close to root-mean-squared errors of interpolated predictions at most depths, signifying that the kriging models accurately predicted microsite-dependent variability in values for the constant incubation temperature

equivalent (CTE) (Fig. 2a, 2b, Fig. 3; Table 1) (Burrough et al., 2015). However, variability in interpolated CTEs for the 'minimum warming' scenario was underestimated at a depth of 150mm (Table 1).

Values of the root-mean-squared deviation (RMSD) were higher for interpolated CTEs under the minimum warming scenario than under the maximum warming scenario. In general, the RMSD in predictions decreased with increasing soil depth and was lower for the maximum warming scenario at all soil depths except 300mm, compared with the scenario of minimum warming. The RMSD in proportions of male offspring was high at soil depths from 50-200mm under the maximum climate warming scenario, although mean CTEs and sex ratios were similar between modelled and interpolated predictions (Table 1). The depth-dependent pattern in the accuracy of interpolated values is likely reflective of the depth-dependent reduction in variance of soil temperatures.

A two-sample Kolmogorov-Smirnov test found that the distributions of modelled CTEs for all depths and under both simulated climate warming scenarios varied between modelling methods ($0.14 \leq D \leq 0.37$, $p < 0.001$). Under the scenario of minimum climate warming, most sites were predicted to produce all-female nests at all depths; hence, the distributions of interpolated vs. modelled sex ratios were not different. Distributions of interpolated vs. modelled sex ratios were different under the maximum climate warming scenario at all depths ($0.12 \leq D \leq 0.15$, $0.01 \leq p \leq 0.001$). Under the minimum warming scenario, interpolated mean CTEs were higher than modelled values at all depths (Table 1).

Interpolated mean CTEs and sex ratios were similar to or slightly higher than modelled values under the scenario of maximum climate warming. Mann-Whitney-Wilcoxon tests found that interpolated vs. modelled CTEs predicted for the minimum climate warming scenario were not different at soil depths of 50mm and 200mm. However, CTEs varied

between methods at depths of 100mm ($W = 29013$, $p < 0.05$), 150mm ($W = 26622$, $p < 0.001$) and 300mm ($W = 7220$, $p < 0.001$) under the scenario of minimum climate warming.

Under the maximum warming scenario, interpolated vs. modelled CTEs were not different at depths from 50mm to 200mm but varied significantly at a depth of 300mm ($W = 115037$, $p < 0.05$). Because predictions of hatchling sex ratios under the minimum climate warming scenario were 0.00 (i.e., 100% female) statistical significance could not be calculated for results at depths lower than 50mm. However, at a soil depth of 50mm, interpolated vs. modelled offspring sex ratios differed significantly ($W = 30915$, $p < 0.01$), though the means were still 0.00. Under the scenario of maximum warming, interpolated vs. modelled sex ratios were different at soil depths of 50mm ($W = 136815$, $p < 0.01$), 100mm ($W = 133672$, $p < 0.05$) and 300mm ($W = 104163$, $p < 0.001$) but not at depths of 150mm or 200mm.

Discussion

Our results indicate that ordinary kriging can be considered a useful tool for examining the physiological outcomes of microclimate conditions at a very high spatial resolution. As for any model, confidence in the accuracy of the fitted values should only be as high as confidence in the training data, and awareness of potential sources of error is critical to fitting a useful surface. However, even if fitted values contain error resulting from inaccuracies in the training data, this method can be used to examine (i) spatial patterns in the variability of predicted values, (ii) spatial distributions of edges and threshold values and (iii) overall patterns in interpolated predictions (e.g., population means) at much higher spatial resolutions than can be accomplished using models built with low-resolution climate layers. Importantly, geostatistical interpolation can be implemented without the use of high-performance computing, saving time and financial investment in predictive ecological models without eliminating the ability to use complex, process-explicit methods.

In contrast to climate variables, such as temperature and rainfall, that have commonly been used to generate underlying surface layers for distribution models, CTEs are not measured values but proxies for the physiological process of sexual differentiation. Thus, the error in a predictive surface of offspring sex ratios is sensitive to the CTE value itself and will become more severe as the CTE approaches a threshold, in this case, the 'pivotal' temperature that produces a 1:1 offspring sex ratio. The narrower the range of temperatures surrounding the pivotal temperature(s), the more likely that sex ratios will be over or underestimated when the CTE is within that window. For example, error of $\pm 2.0^{\circ}\text{C}$ in interpolated CTEs would not affect predicted offspring sex ratios in tuatara if the modelled CTE value is 17.5°C , but error of just $\pm 0.5^{\circ}\text{C}$ could shift the predicted sex ratio of a microsite from primarily male to primarily female if the predicted CTE is 21.5°C . Due to the threshold bias, the sizes of areas that could actually produce mixed-sex nests are probably larger than predicted.

The 1 °C range of temperatures that produces both sexes in tuatara is narrow, compared with that of most reptiles with temperature-dependent sex determination (Ewert et al., 2004). For other species, the wider the transitional range of temperatures, the more that the effects of error on interpolated predictions of CTEs would be concentrated at the upper and lower limits of that range. Similarly, the ecologically important effects of error will be concentrated around any physiological threshold, such as the critical thermal limits that define the bounds of an organism's physiological niche, rather than within a range of performance optima. Indicator kriging estimates a continuous surface from binary data (Burrough et al., 2015), which could be useful for modelling the probability that a physiological threshold has been exceeded. In general, a key application of geostatistical interpolation to physiological data could be in identifying the probable spatial distribution of extremes that threaten population persistence.

Otherwise, current methods of geostatistical interpolation are probably limited in ecology to estimating continuous data, such as incubation temperatures. In addition, while the microclimate model used in this study explicitly accounted for the effects of complex terrain (i.e., elevation, slope, aspect, viewshed) on soil temperatures, cokriging may be useful for interpolating physiological parameters that have been measured directly. Cokriging can take advantage of spatial correlation between two variables to fit a surface through the variable of interest. If the CTEs used in this study had been generated from soil temperatures measured in the field, cokriging could estimate a surface of CTEs based on, for example, slope or soil type as well as spatial autocorrelation in the CTEs themselves, as long as slope or soil type were measured at additional points (Burrough et al., 2015). Cokriging might also be appropriate for locations with extreme heterogeneity in environmental conditions that occurs on a spatial scale much finer than the extent of the full model. For example, the thermal environment of termite nests is determined by biotic heat production, via metabolism of termites and fungi, which sharply differentiates their microclimates from the surrounding area

(Korb, 2003). Accurate geostatistical interpolation of soil temperatures in locations containing termite mounds would require either inclusion of an additional covariate or, if a univariate method were used, exclusion of termite mounds from the interpolated area.

Even with minimal model error, the method of kriging used here should not be employed to examine emergence phenology formatted as circular data (e.g., Julian days). Mathematically, spatially-explicit data on hatching dates or other time-based indicators can be interpolated using geostatistical methods (Gumiaux et al., 2003). Spatial data with a directional component, such as wind vectors, can be interpolated using kriging (Morphet 2009, Morphet & Symanzik 2010). However, the error inherent in any surface-fitting function has additional implications for circular values of time. Interpolation of dates would require development of two models, each trained to a different trigonometric function of the dataset, for every scenario of interest. For example, a separate kriging model could be fit to the sine and cosine of angular hatching dates. The respective signs of each location on the interpolated surfaces could be used to place the polar coordinates correctly in Cartesian space, then the cosine surface back-transformed to angular hatching dates and, finally, to days. However, even very small amounts of error (e.g., on the scale of 0.001) introduced into polar coordinates during interpolation could lead to error on the scale of months in predictions of phenology. Where phenology is the variable of interest, researchers should ensure that data can be treated as continuous.

In this study, we used ordinary kriging to estimate surfaces from data points that were themselves the outcome of multiple, nested models, each with distinct possible sources of error. However, the ability of the kriging model to accurately predict spatially-dependent variability in the training data does not depend on whether or not those values are accurate. Different geostatistical methods may perform better and should be investigated for fitting different sources of ecological data. For example, cokriging may outperform ordinary kriging,

459 when mechanistic drivers of environmental conditions are not already accounted for in the
460 variable of interest. However, geostatistical interpolation can be considered a means of
461 increasing the use of computationally-intensive mechanistic and biophysical modelling at
462 extremely high spatial resolutions.

References

- Araújo, M.B. & M. Luoto. 2007. The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography* 16: 743-53.
- Boer, E.P.J., K.M. de Beurs, & A.D. Hartkamp. 2001. Kriging and thin-plate splines for mapping climate variables. *International Journal of Applied Earth Observation and Geoinformation* 3(2): 146-54.
- Burrough, P.A., McDonnell, R.A. & Lloyd, C.D. 2015. *Principles of Geographical Information Systems*. 3rd Ed. Oxford: OUP, 432 pp.
- Carter, A., N.J. Mitchell, S.H. Hartley, M.R. Kearney, W.P. Porter & N.J. Nelson (2015) Modelling the soil microclimate: does the spatial or temporal resolution of input parameters matter? *Frontiers of Biogeography* 7(4): 138-154.
- Cree, A., J.F. Cockrem, & L.J. Guillette. 1992. Reproductive cycles of male and female tuatara (*Sphenodon punctatus*) on Stephens Island, New Zealand. *Journal of Zoology* London 226: 199-217.
- Cree, A., C.H. Daugherty, S.F. Schafer, & D. Brown. 1991. Nesting and clutch size of tuatara (*Sphenodon guntheri*) on North Brother Island, Cook Strait. *Tuatara* 31: 9-16.
- Cree, A., Thompson, M.B. & Daugherty, C.H. 1995. Tuatara sex determination, *Nature* 375: 543.
- Cree, A., M.B. Thompson, L.J. Guillette Jr., J.M. Hay, & M.E. McIntyre. 1989. Embryonic development of tuatara in forested and open habitats on Stephens Island, New Zealand.

Proceedings of the 2nd Annual Conference of the Society for Research on Amphibians and Reptiles in New Zealand [SRARNZ]. Middlemarch, New Zealand. New Zealand Journal of Zoology 16: 270, abstract.

Cressie, N. 1988. Spatial prediction and ordinary kriging. Mathematical Geology 20: 405-21. (Erratum, Mathematical Geology 21: 493-4).

Dallwitz, M.J. & J.P. Higgins. 1992. User's guide to DEVAR: a computer program for estimating development rate as a function of temperature. 2nd Ed. Division of Entomology Report No. 2, Commonwealth Scientific and Industrial Research Organisation [CSIRO], Australia. 27 pp.

Dormann, C.F., Schymanski, S.J., Cabral, J. et al. 2012. Correlation and process in species distribution models: bridging a dichotomy. Journal of Biogeography 39: 2119-31.

[ESRI] Environmental Systems Research Institute. 2003. ArcGIS 9: Using ArcGIS Geostatistical Analyst. Redlands, California, USA: ESRI Press. 306 pp.

Ewert, M.A., Etchberger, C.R., & Nelson, C.E. 2004. Turtle sex-determining modes and TSD patterns, and some TSD pattern correlates. In: N. Valenzuela & V. Lance, Eds. Temperature-dependent sex determination in vertebrates. Washington, DC, USA: Smithsonian, pp 21-32.

Georges, A., K. Beggs, J.E. Young, & J.S. Doody. 2005. Modelling development of reptile embryos under fluctuating temperature regimes. Physiological and Biochemical Zoology 78(1): 18-30.

- Georges, A., C. Limpus, & R. Stoutjesdijk. 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.
- GIMP Development Team. 2013. GNU Image Manipulation Program. Version 2.8.
<http://www.gimp.org>.
- Guisan, A. & W. Thuiller. 2005. Predicting species distribution: offering more than simple habitat models. *Ecology Letters* 8: 993-1009.
- Guisan, A. & N.E. Zimmerman. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135: 147-86.
- Gumiaux, C., D. Gapais, & J.P. Brun. 2003. Geostatistics applied to best-fit interpolation of orientation data. *Tectonophysics* 376: 241-59.
- Hancock, P.A. & M.F. Hutchinson. 2006. Spatial interpolation of large climate data sets using bivariate thin-plate smoothing splines. *Environmental Modelling and Software* 21(12): 1684-94.
- Hernandez-Stefanoni, J.L. & Ponce-Hernandez, R. 2006. Mapping the spatial variability of plant diversity in a tropical forest: comparison of spatial interpolation methods. *Environmental Monitoring and Assessment* 117: 307-334.
- Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones & A. Jarvis. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965-78.

544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569

Hutchinson, M.F. 1991. The application of thin plate smoothing splines to continent-wide data assimilation. In: Jasper JD (ed.) BMRC Research Report No.27, Data Assimilation Systems. Melbourne: Bureau of Meteorology: 104-113.

Jarvis, C.H. & N. Stuart. 2001a. A comparison among strategies for interpolating maximum and minimum temperatures. Part I: the selection of 'guiding' topographic and land cover variables. *Journal of Applied Meteorology* 40: 1060-74.

Jarvis, C.H. & N. Stuart. 2001b. A comparison among strategies for interpolating maximum and minimum temperatures. Part II: the interaction between number of guiding variables and the type of interpolation. *Journal of Applied Meteorology* 40: 1075-84.

Kearney, M.R. 2006. Habitat, environment and niche: what are we modelling? *Oikos* 115(1): 186-91.

Kearney, M.R., Matzelle, A. & Helmuth, B. 2012. Biomechanics meets the ecological niche: the importance of temporal data resolution. *The Journal of Experimental Biology*, 215: 922-33.

Kearney, M.R. & W.P. Porter. 2009. Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters* 12: 334-50.

Kearney, M.R. & W.P. Porter. 2016. NicheMapR - an R package for biophysical modelling: the microclimate model. *Ecography*, doi: 10.1111/ecog.02360.

Kearney, M.R., Shamakhy, A., Tingley, R., Karoly, D.J., Hoffman, A.A., Briggs, P.R. & Porter, W.P. 2014. Microclimate modelling at macro scales: a test of a general microclimate model integrated with gridded continental-scale soil and weather data. *Methods in Ecology and Evolution*, doi: 10.1111/2041-210X.12148.

Keith, D.A., H. Resit, H.R. Akçakaya, W. Thuiller, G.F. Midgley, R.G. Pearson, S.J. Phillips, H.M. Regan, M.B. Araújo, & T.G. Rebelo. 2008. Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat. *Biology Letters* 4: 560-563.

Korb, J. 2003. Thermoregulation and ventilation of termite mounds. *Naturwissenschaften* 90: 212-219.

Lavergne, S., N. Mouquet, W. Thuiller, & O. Ronce. 2010. Biodiversity and climate change: integrating evolutionary responses of species and communities. *Annual Review of Ecology, Evolution, and Systematics* 41: 321-50.

Li, J. & Heap, A.D. 2008. A review of spatial interpolation methods for environmental scientists. Department of Resources, Energy and Tourism. Geoscience Australia: Record 2008/23.

[MFE] Ministry for the Environment (New Zealand). 2008. Climate change effects and impacts assessment: a guidance manual for local government in New Zealand. 2nd Ed. Mullan B., D. Wratt, S. Dean, M. Hollis, S. Allan, T. Williams, G. Kenny, & MFE. Wellington, NZ: Ministry for the Environment. 167 pp.

596 Mitchell, N.J., M.R. Hipsey, S. Arnall, G. McGrath, H.B. Tareque, G. Kuchling, R. Vogwill, M.
 597 Sivapalan, W.P. Porter, & M.R. Kearney. 2013. Linking eco-energetics and eco-hydrology to
 598 select sites for the assisted colonization of Australia's rarest reptile. *Biology* 2: 1-25.
 599
 600 Mitchell, N.J., M.R. Kearney, N.J. Nelson, & W.P. Porter. 2009. Predicting the fate of a living
 601 fossil: how will global warming affect sex determination and hatching phenology in tuatara?
 602 *Proceedings of the Royal Society B* 275: 2185-93.
 603
 604 Mitchell, N.J., N.J. Nelson, A. Cree, S. Pledger, S.N. Keall, & C.H. Daugherty. 2006. Support
 605 for a rare pattern of temperature-dependent sex determination in archaic reptiles: evidence
 606 from two species of tuatara (*Sphenodon*). *Frontiers in Zoology* 3: 9.
 607
 608 Morphet, W.J. 2009. Visualization, kriging, and simulation of circular-spatial data.
 609 Unpublished Ph.D. thesis, Utah State University, USA.
 610
 611 Morphet, W.J. & J. Symanzik. 2010. The circular dataimage, a graph for high-resolution
 612 circular-spatial data. *International Journal of Digital Earth* doi: 10.1080/17538940903277657.
 613 Nelson, N.J., A. Cree, M.B. Thompson, S.N. Keall, & C.H. Daugherty. 2004a. Temperature-
 614 dependent sex determination in tuatara. In: N. Valenzuela & V. Lance, Eds. *Temperature-*
 615 *dependent sex determination in vertebrates*. Washington, DC, USA: Smithsonian, pp 53-58.
 616
 617 Nelson, N.J., J.A. Moore, S. Pillai, & S.N. Keall. 2010. Thermosensitive period for sex
 618 determination in the tuatara. *Herpetological Conservation and Biology* 5(2): 324-329.
 619
 620 Nelson, N.J., M.B. Thompson, S. Pledger, S.N. Keall & C.H. Daugherty. 2004b. Do TSD, sex
 621 ratios, and nest characteristics influence the vulnerability of tuatara to global warming?
 622 *International Congress Series* 1275: 250-257.

623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648

Newman, D.G. 1987. Burrow use and population densities of tuatara (*Sphenodon punctatus*) and how they are influenced by fairy prions (*Pachyptila turtur*) on Stephen's Island, New Zealand. *Herpetologica* 43: 336-344.

Palmer, T. 2014. Build high-resolution global climate models. *Nature* 515: 338-339.

Porter, W.P., S. Ostrowski, & J.B. Williams. 2010. Modeling animal landscapes. *Physiological and Biochemical Zoology* 83(5): doi: 10.1086/656181.

R Development Core Team. 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria: <http://www.R-project.org>.

Sekhon, J. 2011. Multivariate and propensity score matching software with automated balance optimization: The Matching package for R. *Journal of Statistical Software*, 42(7): 1-52.

Stainforth, D., J. Kettleborough, M. Allen, M. Collins, A. Heaps & J. Murphy 2002. Distributed computing for public-interest climate modeling research. *Computing in Science and Engineering* 4(3): 82-9.

Thuiller, W. 2004. Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology* 10: 2020-7.

Wackernagel, H. 1998. Multivariate geostatistics. Berlin: Springer Verlag. 291 pp.

649 Zambrano-Bigiarini, M. 2014. hydroGOF: goodness-of-fit functions for comparison of
650 simulated and observed hydrological time series. R package version 0.3-8. [http://cran.r-](http://cran.r-project.org/package=hydroGOF)
651 [project.org/package=hydroGOF](http://cran.r-project.org/package=hydroGOF).

Author Manuscript

Data accessibility

GIS (coordinates and terrain inputs) and climate data for running the microclimate model are available on figshare at doi: 10.6084/m9.figshare.c.3798013.

Biosketch

Anna Carter is a postdoctoral research associate at Iowa State University, conducting research in biogeography, with a focus on the microclimate-scale effects of environmental change on reptiles.

Editor: Steven Higgins

Tables

Table 1. Summary statistics quantifying error in the geostatistical model of constant incubation temperature equivalent (CTE) values for tuatara, developed for ‘minimum’ and ‘maximum’ warming scenarios on the island of Takapourewa, New Zealand, showing (1) error based on cross-validation of the kriging models with the training data (mean error and root-mean-squared (RMS) error of predicted CTEs, and mean standard error of predictions) and (2) error based on independent validation of kriging models against a second dataset of developmental predictions generated using mechanistically modelled soil temperatures: mean CTEs, mean sex ratios (sr), and root-mean-squared deviations (RMSDs) comparing interpolated vs. modelled CTEs and sex ratios. Paired values are interpolated/modelled. Values for sex ratios are given as the proportion of hatchlings predicted to be male, assuming hatching success of 100%. The kriging model underestimated variability in CTEs at a depth of 150mm under the ‘minimum warming’ scenario (mean SE < RMS).

	(1) Cross-validation				(2) Independent validation			
	soil depth (mm)	mean error	RMS	mean SE	mean CTEs (°C)	mean sr (male)	RMSD CTE (°C)	RMSD sr (male)
Minimum warming	50	0.00	2.24	2.39	16.94 / 16.66	0.00 / 0.00	2.83	0.00
	100	-0.01	1.93	1.01	15.53 / 15.15	0.00 / 0.00	2.20	0.00
	150	0.00	1.43	0.56	14.62 / 14.16	0.00 / 0.00	1.78	0.00
	200	0.00	1.18	1.08	14.07 / 13.88	0.00 / 0.00	1.41	NA
	300	0.00	0.88	0.83	14.45 / 14.50	0.00 / 0.00	0.92	0.00

50	0.00	1.21	1.20	23.08 / 23.12	0.83 / 0.76	1.16	0.36
100	0.00	1.11	1.10	22.25 / 22.29	0.63 / 0.58	1.04	0.44
150	0.00	1.02	0.98	21.59 / 21.60	0.37 / 0.41	0.99	0.45
200	0.00	1.02	0.96	20.97 / 21.00	0.13 / 0.25	1.00	0.37
300	0.00	1.18	1.08	19.56 / 19.60	0.00 / 0.03	1.17	0.14

Figure legends and figures

Figure 1. Flow chart summarising the procedure for predicting offspring sex ratios for tuatara, using modelled hourly soil temperatures. Semi-enclosed rectangles designate model inputs, enclosed rectangles are model algorithms, parallelograms are outputs, and the diamonds are optional modifications. Figure adapted with permission from Carter et al. (2015). Details of model validation are in Carter et al. (2015), with detailed descriptions of algorithms and equations in Kearney & Porter (2016). Example code for the microclimate model are in Appendix S1.

Figure 2. Spatial distribution of standard errors in constant incubation temperatures (CTEs) by soil depth for tuatara on the island of Takapourewa, New Zealand, based on leave-one-out cross-validation for the (a) minimum warming scenario and (b) maximum warming scenario, superimposed over an elevation map of the island. The random sites used as training data can be seen clearly as dark points on the map representing error at a depth of 50mm (b). Grayscale areas are locations that were too cold for a CTE to be estimated under the minimum warming scenario (i.e. total development did not reach at least 35% within two years). Confidence that the true CTE is within $\pm$ SE of interpolated values is 95.5%.

Figure 3. Boxplots showing distributions of modelled vs. interpolated constant incubation temperatures (CTEs) and tuatara hatchling sex ratios on the island of Takapourewa, New Zealand for two scenarios of climate warming, given an oviposition date of 30 November. In each pair, the left-hand box (striped) shows the distribution of values generated for 500 randomly-sample points, using mechanistically modelled hourly soil temperatures, and the right-hand box shows values interpolated for the same points using ordinary kriging. The dashed line on the plot of CTE values approximates the temperature at which tuatara sex ratios are 1:1. Modelled vs. interpolated distributions of CTEs were significantly different

under both climate scenarios. Hatchling sex ratios were not interpolated themselves, but were generated from interpolated CTEs, using the A-logistic function (Girondot 1999; Godfrey et al. 2003). Under the 'minimum warming' scenario, nearly all sites were predicted to produce 100% female hatchlings. Under the 'maximum warming' scenario, sex ratios derived from interpolated CTEs were significantly different from those estimated directly from modeled CTEs.

Figure 1.

(a)

depth = 50mm

depth = 100mm

depth = 150mm

depth = 200mm

depth = 300mm

Figure 2.

Figure 3.

715 **Appendices**

716

717 Appendix S1. Example scripts for running the NicheMapR microclimate model over a grid of

718 sites, with options for using local, daily climate data.

Author Manuscript

(a)

depth = 50mm

depth = 100mm

depth = 150mm

depth = 200mm

depth = 300mm

0 1000 2000m

(b)

depth = 50mm

depth = 100mm

depth = 150mm

depth = 200mm

depth = 300mm

Prediction standard error

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Carter, AL;Kearney, MR;Hartley, S;Porter, WP;Nelson, NJ

Title:

Geostatistical interpolation can reliably extend coverage of a very high-resolution model of temperature-dependent sex determination

Date:

2018-03-01

Citation:

Carter, A. L., Kearney, M. R., Hartley, S., Porter, W. P. & Nelson, N. J. (2018). Geostatistical interpolation can reliably extend coverage of a very high-resolution model of temperature-dependent sex determination. JOURNAL OF BIOGEOGRAPHY, 45 (3), pp.652-663. <https://doi.org/10.1111/jbi.13152>.

Persistent Link:

<http://hdl.handle.net/11343/294053>