LETTER

When can the cause of a population decline be determined?

Trevor J. Hefley,[1]*
Mevin B. Hooten,[2,1] John M.
Drake,[3] Robin E. Russell[4] and
Daniel P. Walsh[4]

Abstract

Inferring the factors responsible for declines in abundance is a prerequisite to preventing the extinction of wild populations. Many of the policies and programmes intended to prevent extinctions operate on the assumption that the factors driving the decline of a population can be determined. Exogenous factors that cause declines in abundance can be statistically confounded with endogenous factors such as density dependence. To demonstrate the potential for confounding, we used an experiment where replicated populations were driven to extinction by gradually manipulating habitat quality. In many of the replicated populations, habitat quality and density dependence were confounded, which obscured causal inference. Our results show that confounding is likely to occur when the exogenous factors that are driving the decline change gradually over time. Our study has direct implications for wild populations, because many factors that could drive a population to extinction change gradually through time.

Keywords

Allee effect, autocorrelation, causal inference, density dependence, extinction, multicollinearity, population dynamics, temporal confounding, time series.

Ecology Letters (2016) **19**: 1353–1362

INTRODUCTION

Determining the factors that cause a decline in abundance is central to preventing extinction and protecting biodiversity (Caughley 1994). In some populations, the factors driving the decline in abundance might change gradually (smoothly) over time and can be highly collinear with endogenous population processes. When endogenous factors that influence future values of the observed response are highly correlated with exogenous factors, the process is said to be temporally confounded. When temporal confounding occurs, identifying causal relationships can be challenging (Fieberg & Ditmer 2012). In this study, we explore the challenges associated with determining the cause of a decline in abundance from replicated populations that were driven to extinction by gradually manipulating habitat quality.

In the experiment conducted by Drake & Griffen (2010), 60 populations of *Daphnia magna* were reared in microcosms. Of the 60 populations, 30 were assigned to a control group and 30 were assigned to a deteriorating environment treatment by gradually reducing the amount of food until each population was driven extinct. For many natural populations, the factors driving the decline are unknown and must be inferred using models fit to a time series from a single population. The following example relies on a time series from a single population (I10) to demonstrate that factors that are causally related to the decline in abundance can be difficult to infer for a single population (Fig. 1a).

A preliminary analysis might use linear regression to determine the relationship between the amount of food and the log-transformed population abundance (Fig. 1b; Ives 2015). Using linear regression, the estimated coefficient is 2.33

([1.73, 2.92]; 95% confidence interval (CI)). As expected, the decline in abundance can be linked to the decreasing food levels. But, due to the serial dependence of population dynamics, there is correlation among the residuals obtained from the regression model (Fig. 1c). To account for autocorrelation, we fit a linear regression model that accommodates correlated errors. The correlated errors regression model resulted in an estimated effect of 0.36, which was no longer significantly different from zero ([−1.34, 2.05]; 95% CI). Evidently, the causal effect of environmental deterioration cannot be separated from the inevitable serial dependence in the population dynamical process.

For many populations, the factors causing the decline in abundance are unknown and inference might be based on the correlated errors model. As this example shows, the observed shift in coefficient estimates between the linear regression model and the correlated errors regression model is a symptom of confounding, which can mislead inference (e.g. Kühn 2007; Bini *et al.* 2009; Hodges & Reich 2010; Fieberg & Ditmer 2012). In this paper, we present the experimental data in detail and compare the inference obtained from several regression-based and population dynamics models. We then conduct a simulation study and conclude with a discussion that summarises important findings.

MATERIALS AND METHODS

Experimental data

The data used in our study were described by Drake & Griffen (2010) and are available from the Dryad Digital

[1]*Department of Statistics and Department of Fish, Wildlife, and Conservation Biology, Colorado State University, Fort Collins, CO 80523, USA*
[2]*U.S. Geological Survey, Colorado Cooperative Fish and Wildlife Unit, Fort Collins, CO 80523, USA*

[3]*Odum School of Ecology, University of Georgia, Athens, GA 30602*
[4]*U.S. Geological Survey, National Wildlife Health Center, Madison, WI 80523, USA*
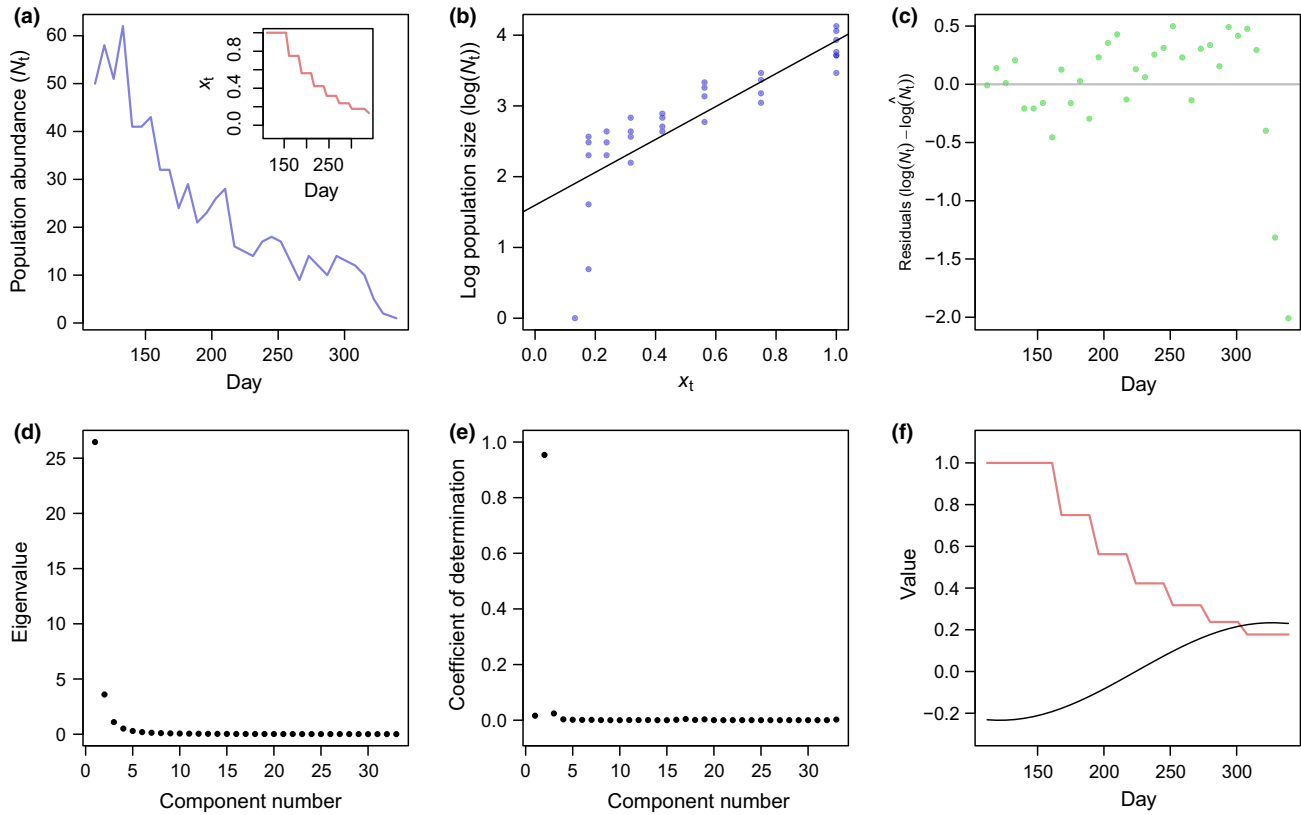**Correspondence: E-mail: trevor.hefley@colostate.edu*

Figure 1 Example time series of the number of *Daphnia magna* (N_t; panel a; blue line) from the experiment conducted by Drake & Griffen (2010) which subjected 30 populations to reduced food level to simulate environmental deterioration. The food level covariate (x_t; subplot in panel a; red line) ranged from 1 (full food level) to 0 (no food) and resulted in the decline and extinction of the population. Panel b shows a simple linear regression model (eqn 1; black line) used to determine the relationship between the amount of food and the log-transformed population abundance (blue dots). Panel c shows autocorrelation in the residuals (green dots) obtained from fitting eqn 1 to the data. Panel d shows a scree plot and panel e shows the coefficient of determination (R^2) for the food level covariate and eigenvectors obtained from a spectral decomposition of the correlation matrix (eqn 2). Panel f shows the food level covariate (red line) and the second leading eigenvector (black line), which were highly correlated ($R^2 = 0.95$).

Repository: http://datadryad.org/resource/doi:10.5061/dryad. q3p64. The experiment consisted of 60 populations of *D. magna* reared in microcosms and randomly assigned to deteriorating and constant treatment groups. Initially, populations were fed 800 µL of the food resource each day (freeze-dried blue green alga). On day 154, populations in the deteriorating treatment were placed on a regimen of declining food (reduction of 25%), with changes on days 154, 182, 210, 238, 266, 294, 329, 357, 385 and 413. The experiment was ended when the last population in the deteriorating treatment group went extinct on day 416. Each population was exhaustively counted in triplicate once per week. Observations before day 105 were excluded from our analyses to reduce effects of transient fluctuations (Drake & Griffen 2010). In this study, we used counts that were greater than zero and we scaled the food level so that a value of one represents no environmental deterioration (a food level of 800 µL day^{-1}) and a value of zero represents an unsuitable environment (0 µL day^{-1}).

Regression models

Here we explain the technical details of temporal confounding and describe the two regression models that were referred in

the introduction and elsewhere in the paper. First, we used the simple linear regression model

$$\log(N_{t+1}) = \beta_0 + \beta_1 x_t + \varepsilon_t, \qquad (1)$$

where N_{t+1} is the population abundance on week $t + 1$ ($t = 1, 2, \ldots, T$), x_t is the amount of food on week t, and ε_t is the residual error ($\varepsilon_t \sim N(0, \sigma_\varepsilon^2)$). To emulate the common scenario where only data from a single population are available, we fit a simple linear regression model to the counts from population I10 and obtained $\hat{\beta}_1 = 2.33$ and a 95% CI of (1.73, 2.91). To account for the autocorrelated residuals (Table 1), let

$$\varepsilon \sim N\left(\mathbf{0}, \sigma_\varepsilon^2 \mathbf{C}(\phi)\right), \qquad (2)$$

where the vector $\varepsilon \equiv [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T]'$ and $\mathbf{C}(\phi)$ is a correlation matrix with the parameter ϕ. Assuming a first-order autoregressive (AR(1)) correlation matrix, we obtained $\hat{\beta}_1 = 0.36$ and a 95% CI of $[-1.34, 2.05]$ (Pinheiro & Bates 2000, p. 236).

To understand why the estimated coefficients are different between the two regression models, we use a truncated Karhunen-Loéve expansion (Cressie & Wikle 2011, pp. 266–269) and write the linear regression model with correlated residual errors as

Table 1 Issues, diagnostics, and effect of temporal confounding on models used to determine the cause of declines in abundance for our study

Model	Issue	Diagnostic	Effect
Linear regression (eqn 1)	Correlated residuals	Residual plots	Bias
Correlated errors (eqn 2)	Covariates correlated with eigenvectors	Calculate correlation between covariate and eigenvectors	High variance, bias for extreme cases
Ricker (eqns 5–8)	Correlation between density-dependent effect and covariates	Calculate correlation	High variance
Allee effect (eqn 9)	Distinguishing decline due to Allee effect from an endogenous factor	Calculate model selection statistic	Incorrect inference
Beverton–Holt (eqn S1)	Correlation between density-dependent effect and covariates	Calculate correlation	High variance

$$\log(N_{t+1}) = \beta_0 + \beta_1 x_t + \mathbf{z}'_t(\phi)\boldsymbol{\alpha} \tag{3}$$

where $\mathbf{z}'_t(\phi)$ is the tth row of the matrix of eigenvectors ($\mathbf{Z}(\phi)$) obtained from a spectral decomposition of the correlation matrix $\mathbf{C}(\phi)$ and $\boldsymbol{\alpha}$ is a vector of basis coefficients (Hefley et al. 2016). The eigenvectors in $\mathbf{Z}(\phi)$ are basis vectors that model the temporal correlation at independent frequencies, and the basis coefficients $\boldsymbol{\alpha}$ are random effects ($\boldsymbol{\alpha} \sim \text{N}(\mathbf{0}, \sigma_\varepsilon^2\boldsymbol{\Lambda}(\phi))$, where $\boldsymbol{\Lambda}(\phi)$ is a diagonal matrix of corresponding eigenvalues (Hodges & Reich 2010; Hefley et al. 2016). The numerical value of the basis coefficients ($\boldsymbol{\alpha}$) determines the contribution of each frequency of the temporal correlation (column of $\mathbf{Z}(\phi)$), similar to how regression coefficients (e.g. β_1) determine the influence of covariates. The eigenvalues (diagonal elements of $\boldsymbol{\Lambda}(\phi)$) determine the relative amount of shrinkage (i.e. Regularization; Hooten & Hobbs 2015) of the basis coefficients (Fig. 1d; Hodges & Reich 2010; Hefley et al. 2016).

For population I10, the coefficient of determination (R^2) between the food level covariate and the second leading eigenvector (i.e., second column of $\mathbf{Z}(\phi)$) is $R^2 = 0.95$ (Fig. 1e, f; Table 1). As with traditional regression models, severe multicollinearity can cause coefficient estimates to be unstable and change when a correlated factor is added or removed from the model (Table 1; Hodges & Reich 2010; Dormann et al. 2013).

The regression models are not explicitly based on any underlying theory relating to population growth or decline (e.g. Abrams 2002). One hypothesis is that temporal confounding is related to model misspecification and that the impacts could be alleviated by fitting a model that explicitly incorporates endogenous population dynamics.

Population dynamics models

A model that explicitly captures endogenous density-dependent population dynamics is the Ricker model (Ricker 1954)

$$N_{t+1} = N_t e^{r\left(1-\frac{N_t}{K}\right)+\varepsilon_t}, \tag{4}$$

where N_t is the population abundance at time t ($t = 1, 2, \ldots, T$), r is the intrinsic growth rate, K is the equilibrium population size, and ε_t is a stochastic perturbation ($\varepsilon_t \sim \text{N}(0, \sigma_\varepsilon^2)$). Covariates are routinely included in population models like eqn 4 to attribute variability in population abundance to exogenous factors (e.g. Jacobson et al. 2004; Knape & de Valpine 2011; Hefley et al. 2013b; Roy et al. in press). We allow the growth rate (r) and equilibrium population size (K) to depend on the food level covariate (x_t)

$$N_{t+1} = N_t e^{r(x_t)\left(1-\frac{N_t}{K(x_t)}\right)+\varepsilon_t}, \tag{5}$$

where $r(x_t)$ and $K(x_t)$ are unknown functions of the covariate x_t.

The potential for confounding is not obvious for the Ricker model in eqn 5. To make the potential for confounding obvious, let $r(x_t) = \alpha_1 x_t$, $K(x_t) = \gamma_1 x_t$, $\theta \equiv -\frac{\alpha_1}{\gamma_1}$, and write eqn 5 as

$$N_{t+1} = N_t e^{\alpha_1 x_t + \theta N_t + \varepsilon_t}. \tag{6}$$

For population I10, the coefficient of determination between x_t and N_t is $R^2 = 0.87$, which is sufficiently high to raise concern (Table 1; Dormann et al. 2013). We fit eqn 6 to population I10 and obtained the estimates $\hat\alpha_1 = 0.40$ (-0.56, 1.37) and $\hat\gamma_1 = 37.3$ (15.2, 59.4), which suggests that effect of food level on $r(x_t)$ is not significantly different from zero. For population I10, the inference obtained from eqn 6 suggests that the population decline is driven by a decrease in the equilibrium population size (note that $\hat\theta = -0.01$ [-0.03, 0.01]).

The Ricker model in eqn 6 assumes that both $r(x_t)$ and $K(x_t)$ are positive (or negative) for all food levels greater than zero. A more realistic assumption would allow for the possibility that $r(x_t)$ or $K(x_t)$ could be less than zero before complete habitat destruction (i.e. $x_t = 0$), thus allowing for dynamics such as a transcritical bifurcation (i.e. $r(x_t) < 0$ when $K(x_t) > 0$; Drake & Griffen 2010; Hefley et al. 2013b). A simple specification of the Ricker model that allows for such dynamics is

$$r(x_t) = \alpha_0 + \alpha_1 x_t \tag{7}$$

$$K(x_t) = \gamma_0 + \gamma_1 x_t \tag{8}$$

We fit the Ricker model (assuming eqns 7 and 8) and obtained $\hat\alpha_1 = 1.77$ (0.46, 3.17) and $\hat\gamma_1 = 34.0$ (-190.5, 65.7). Thus, it would appear that the effect of environmental deterioration influences only the growth rate $r(x_t)$, but the 95% CI

for $\hat{\gamma}_1$ in eqn 8 is suspiciously wide; likely a result of the high correlation between x_t and N_t. Compared to the Ricker model in eqn 6, the second-order Akaike information criterion (AICc; Hurvich & Tsai 1989) score suggests that the relationships in eqns 7 and 8 are preferred (AICc = 24.0 for eqn 6 vs. AICc = 18.4 for eqns 7 and 8).

The Ricker model in eqn 5 lacks a mechanism for a purely endogenous driver of population decline. To test if declines in abundance caused solely by endogenous factors are distinguishable from exogenous factors, we use a Ricker model with an Allee effect

$$N_{t+1} = N_t e^{r\left(1-\frac{N_t}{K}\right)\left(\frac{N_t-c}{K}\right)+\varepsilon_t}, \tag{9}$$

where c is the critical point (unstable equilibrium population size). We fit eqn 9 and obtained an AICc score of 19.1. Thus, it would appear that the effect of food levels are indistinguishable from a purely endogenous driver such as the Allee effect (cf., AICc score from eqn 9 to AICc = 18.4 for 7 and 8; Burnham & Anderson 2002).

State-space models

Each population in the experiment conducted by Drake & Griffen (2010) was counted in triplicate weekly and all replicated counts from the same population agreed, leading to the conclusion that counts were without error. For many studies and monitoring programmes, it is not feasible to count the number of individuals within a population without error. Statistical inference commonly uses state-space models for time series data that includes measurement errors (e.g. de Valpine & Hastings 2002; Clark & Bjørnstad 2004; Dennis *et al.* 2006; Newman *et al.* 2006; Dennis & Ponciano 2014).

The potential for temporal confounding can be obscured because the observed population counts might be less correlated with the covariate when compared to the true abundance because the observational errors will introduce additional random variation into the counts. To demonstrate, we use the weekly counts from population I10 and simulate detection error using

$$y_{it}|N_t \sim binomial(N_t, p), \tag{10}$$

where y_{it} is the ith imperfect count on week t, N_t is the true population count, and p is the probability of detection. For each weekly count, we simulated three imperfect counts (i.e., $i = 1, 2, 3, \forall t$) with $p = 0.25$ (Fig. S1). As mentioned, the coefficient of determination for the covariate x_t and N_t for population I10 is $R^2 = 0.87$, but the R^2 for the three imperfect counts and the covariate is 0.66, 0.69, and 0.66. For field studies, the true abundance N_t is unknown and must be estimated. One approach to estimate the true abundance and unknown parameters in the Ricker model is to formulate a Bayesian state-space model using the N-mixture framework from Royle (2004) (Hefley *et al.* 2013a). For this example, we assume that the underlying process model is the Ricker model with linear changes in both $r(x_t)$ and $K(x_t)$ (eqns 7 and 8; see Appendix S1 for full model specification). The posterior mean of the coefficient of determination (R^2) for the covariate x_t and the estimated abundance N_t was 0.90 [0.69, 0.96] (95%

credible intervals in brackets). The mean of the posterior distribution for α_1 and γ_1 was -0.06 [-1.68, 1.42] and 47.0 [18.1, 79.8], respectively.

Full analysis of data from single populations

We analysed all 30 replicates from the deteriorating environment treatment group individually to assess the prevalence of temporal confounding in data from single populations. To estimate the effect of declining food levels, we fit the simple linear regression model (eqn 1), the AR(1) correlated errors model (eqn 2), and the Ricker model (eqn 5) assuming linear declines in both $r(x_t)$ and $K(x_t)$ (eqns 7 and 8). We also fit a Beverton–Holt model (eqn S1) with linear declines in the both growth rate and equilibrium population size to determine if the issues associated with temporal confounding occur in another model of population dynamics (Beverton & Holt 1957). We report the coefficient estimates and 95% CIs for β_1, α_1, and γ_1 as well as the mean, minimum, and maximum R^2 for population abundance (N_t) and food levels (x_t).

To determine if the cause of population decline can be distinguished from an Allee effect, we fit the Ricker model with an Allee effect (eqn 9) to each population and compared it to the Ricker model with linear declines in both $r(x_t)$ and $K(x_t)$ (eqns 7 and 8) using AICc. We report the number of populations for which: (1) the Ricker model with an Allee effect is the best model; (2) the Ricker model with linear declines in $r(x_t)$ and $K(x_t)$ is the best model; and (3) neither model is best model. The two models differ by one parameter. Accordingly we define a model as best if it has an AICc score that is at least two units less than the model it is being compared to (Burnham & Anderson 2002).

Combined analysis of all populations

To determine if replication can alleviate temporal confounding, we conducted a combined analysis of all 60 replicates from the control and deteriorating environment treatment groups. We fit multivariate specifications of the simple linear regression model (eqn S2), the AR(1) correlated errors model (eqn S3), and the Ricker model (eqn S4; see Appendix S1 for complete model descriptions). For the Ricker model, we assumed linear declines in both $r(x_t)$ and $K(x_t)$ (eqns 7 and 8). We report the coefficient estimates and 95% CIs for β_1, α_1, and γ_1. We also report the R^2 for population abundance (N_t) and food levels (x_t) and the maximum R^2 for the eigenvectors and x_t. To determine if replication can help distinguish if the cause of decline is a result of purely endogenous dynamics or exogenous factors, we fit a multivariate specification of the Ricker model with an Allee effect (eqn S5) and compared it to the Ricker model with linear declines in both $r(x_t)$ and $K(x_t)$ (eqns 7 and 8) using AICc.

Simulation 1

We conducted a simulation study to better understand when the cause of a population decline can be correctly inferred. We simulated data from the Ricker model in eqn 5, but assumed a Poisson distribution. We allowed $r(x_t)$ and $K(x_t)$

to decline linearly (as in eqns 7 and 8) and we used the values $\alpha_0 = -0.03$, $\alpha_1 = 0.3$, $\gamma_0 = -4$, and $\gamma_1 = 40$. Each population was started from a realisation drawn from a Poisson distribution with an expected value of 40 individuals and was propagated until extinction (Fig. 3). We used four scenarios where the covariate x_t had a geometric decline with varying levels of random variation added (inset plots in Fig. 3). A geometric decline in the covariate was used because it mimics the treatment applied in the *D. magna* experimental data. Increasing levels of random variation in the covariate was used to decrease the correlation between x_t and N_t, which was needed to determine the values of R^2 for which reliable inference can be obtained. For each scenario, we simulated 1000 data sets and retained counts that were greater than zero for model fitting.

To determine if the effect of a covariate can be inferred from the simulated data, we used the two regression models (eqns 1 and 2) and the Ricker model (eqn 5) with linear declines in $r(x_t)$ and $K(x_t)$ (eqns 7 and 8) and with an Allee effect (eqn 9). We assessed relative bias, coverage probability, and power for coefficient estimates as a function of R^2 between x_t and N_t, after controlling for the length of each time series (Appendix S1). We define relative bias as the difference between the coefficient estimate and the true value divided by the true value (e.g. $\frac{\hat{\alpha}_1 - \alpha_1}{\alpha_1}$; note that we assumed that the true value was $\alpha_1 = 0.3$ for $\hat{\beta}_1$ in eqns 1 and 2). We define coverage probability as the proportion of the simulated data sets for which the 95% CI contained the true value and power as the proportion of the simulated data sets for which the 95% CI did not include a value less than zero (i.e. the probability we would find a statistically significant positive effect of the covariate x_t). For each simulated data set, we calculated the AICc scores to compare the Ricker model with an endogenous driver (eqn 5) and the Ricker model with an Allee effect (eqn 9). We report the percentage of simulations for which: (1) the Ricker model with an Allee effect is the best model, (2) the Ricker model with linear declines in $r(x_t)$ and $K(x_t)$ is the best model, and (3) neither model is best model. For the purposes of model comparisons, we define a model as best if it has an AICc score that is at least two units less than the model it is being compared to (Burnham & Anderson 2002).

Simulation 2

For the *D. magna* experimental data, the factor driving the decline is known. In some studies, multiple factors will be considered and model selection techniques might be used to determine which factors should be included in the models. To assess the impact of multiple drivers, we simulated data sets from eqn 5, using the same values as in the first simulation, but with the addition of an extra covariate z_t. We simulated 4000 data sets such that z_t had no influence on the population (i.e. $\alpha_2 = 0$ and $\gamma_2 = 0$, where α_2 and γ_2 are the coefficient associated with the covariate z_t similar to eqns 7 and 8; Fig. S2) and 4000 data sets where both x_t and z_t drove the population decline ($\alpha_1 = 0.15$, $\alpha_2 = 0.15$, $\gamma_1 = 20$, and $\gamma_2 = 20$; Fig. S3). For each simulated data set, we calculated the AICc scores for four variations of the Ricker model with: (1) constant r and K; (2) linear declines in $r(x_t)$ and $K(x_t)$ that

depended on x_t; (3) linear declines in $r(z_t)$ and $K(z_t)$ that depended on z_t; and (4) linear declines in $r(x_t, z_t)$ and $K(x_t, z_t)$ that depended on the covariates x_t and z_t. We report the percentage of simulated data sets for which the true model has the lowest AICc scores. For all simulated data sets, x_t and z_t had a geometric decline with the highest level of random variation added in simulation 1 (see Figs 3d, S2 and S3).

Simulation 3

In many studies, it is not known if the population will be driven to extinction by a covariate and type I error is a concern (i.e. falsely inferring that a covariate or some other factor is causing the decline). To address this concern, we simulated data from eqn 5, but used values that resulted in a constant growth rate and equilibrium population size ($\alpha_0 = 0.3$, $\alpha_1 = 0$, $\gamma_0 = 40$, and $\gamma_1 = 0$) We propagated each simulated population for 100 time steps assuming a geometric decline in the covariate and simulated 4000 data sets (see Fig. S4 for example of time series). We then fit the Ricker model that allowed for linear declines in $r(x_t)$ and $K(x_t)$ (eqn 5 assuming eqns 7 and 8). We report the percentage of simulations for which the lower limit of the 95% CIs for α_1 or γ_1 were greater than zero (i.e. a statistically significant positive effect of the covariate was detected) using the first 10, 20, 40, and all 100 observations from each simulated time series.

Simulation 4

To address the concern of model selection, we simulated data from a Ricker model with an Allee effect (eqn 9, but with Poisson stochasticity). We used the values $r = 1$, $K = 45$, and $c = 25$, and propagated each simulated population until extinction. For model fitting, we used the 40 counts that were produced prior to extinction or the entire time series in the event that the population went extinct before 40 time steps. For each population, we assumed a geometric decline in the covariate for the 40 time steps prior to extinction (see Fig. S5 for example of time series). We fit the Ricker model that allowed for linear declines in $r(x_t)$ and $K(x_t)$ (eqn 5 assuming eqn 7 and 8) and the Ricker model with an Allee effect to the simulated data (eqn 9). We report the percentage of simulations for which: (1) the Ricker model with an Allee effect is the best model; (2) the Ricker model with linear declines in $r(x_t)$ and $K(x_t)$ is the best model; and (3) neither model is the best model.

RESULTS

Full analysis of data from single populations

Using the linear regression model, the estimated effect of food level was positive for all 30 populations and the 95% CIs contained zero for only one of the populations (Fig. 2a). Using the correlated errors models, the estimated effect of food level was positive for all 30 populations and the 95% CIs contained zero for six of the populations (Fig. 2b). Using the Ricker population model (eqn 5), the estimated effect of food level on the growth rate (α_1; eqn 7) was positive for 17 of the

populations and the 95% CIs contained zero for 28 of the populations (Fig. 2c) while the estimated effect of food level on the equilibrium population size (γ_1; eqn 8) was positive for all 30 populations and the 95% CIs contained zero for seven of the populations (Fig. 2d). Similar results were obtained for the Beverton–Holt model (Fig. 2e, f). The mean, minimum, and maximum R^2 for population abundance (N_t) and food levels (x_t) was 0.75, 0.12, and 0.92, respectively. Comparing the Ricker model where food level controlled the growth rate and equilibrium population size (eqn 5 with eqns eqn 7 and 8) to a Ricker model where only an endogenous Allee effect could cause extinction (eqn 9), we found that the model with

an Allee effect was selected as the best model for only a single population. For 22 of the populations, the Ricker model where food levels drove the extinction was selected as best. For the remaining seven populations, the data were not sufficient to determine which model was the best.

Combined analysis of all populations

Using all 60 replicates from the control and deteriorating environment treatment groups, all models resulted in a positive and statistically significant effect of food level. Using the simple linear regression model, the estimated effect of food
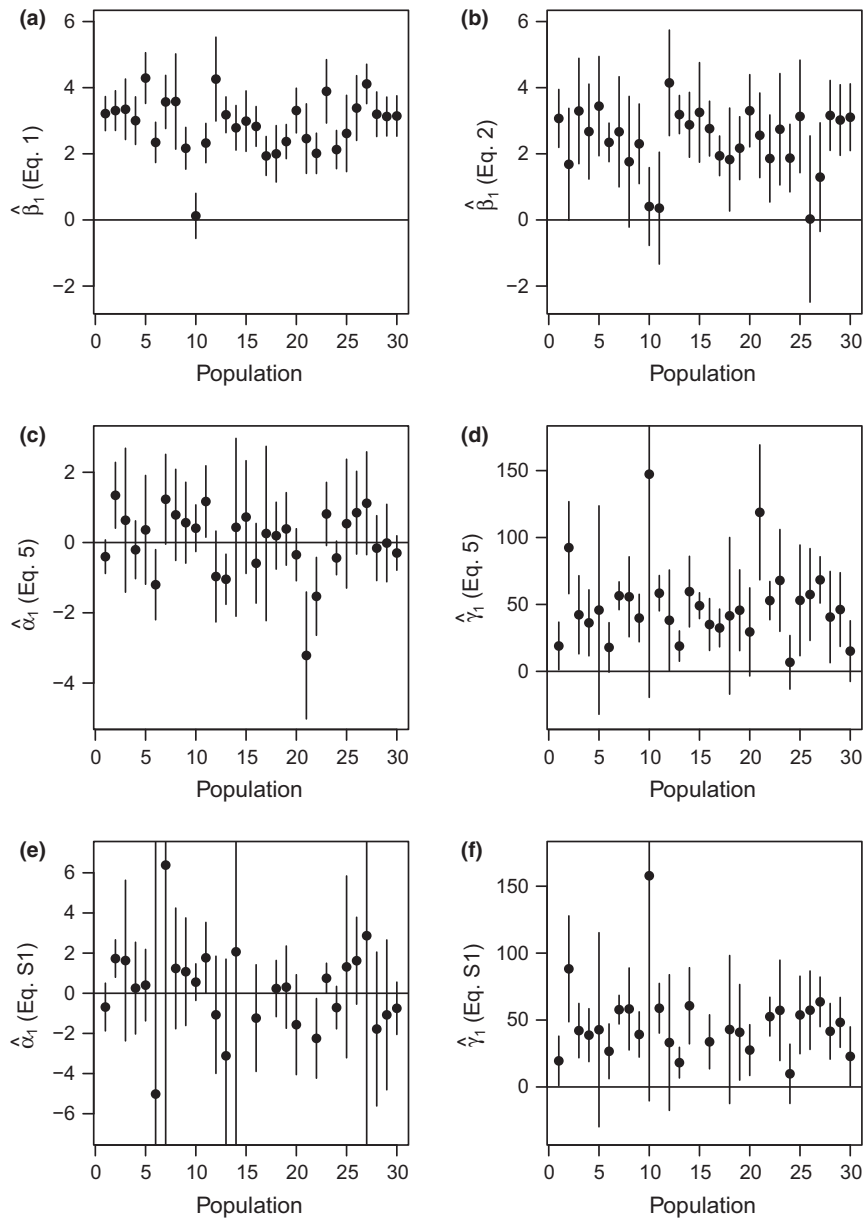

Figure 2 Regression coefficients estimates (black dots) and 95% confidence intervals (vertical lines) for the effect of the food level covariate for all 30 populations in the environmental deterioration treatment group. Estimates and 95% confidence intervals were obtained from a simple linear regression model (panel a; eqn 1), a correlated errors model (panel b; eqn 2), a Ricker model (eqn 5) with linear declines in the growth rate ($r(x_t) = \alpha_0 + \alpha_1 x_t$; panel c) and equilibrium population size ($K(x_t) = \gamma_0 + \gamma_1 x_t$; panel d), and a Beverton–Holt model (eqn S1) with linear declines in the growth rate ($R(x_t) = \alpha_0 + \alpha_1 x_t$; panel e) and equilibrium population size ($K(x_t) = \gamma_0 + \gamma_1 x_t$; panel f). Note that the Beverton–Holt model did not converge for populations 15, 17 and 21.

level was $\hat{\beta}_1 = 2.23$ [2.14, 2.33]. Using the correlated errors model, the estimated effect of food level was $\hat{\beta}_1 = 2.26$ [2.02, 2.50] and the maximum R^2 for the eigenvectors and the food level covariate was 0.08. Using the Ricker population model, the estimated effect of food level on the growth rate (α_1; eqn 7) was 0.45 [0.37, 0.53], the estimated effect on the equilibrium population size (γ_1; eqn 8) was 64.1 [52.8, 75.5], and the AICc score was 2652. The AICc score for the Ricker model with an Allee effect (eqn 9) was 2765. The R^2 for the population abundance (N_t) and food levels (x_t) was 0.35.

Simulation experiments

A high R^2 between the covariate x_t and the population size N_t was produced in the first simulation. When no random variation was added to the covariate, the average R^2 from the 1000 simulated data sets was 0.88 (Fig. 3a). For the three scenarios with increasing levels of random variation added to the covariate, the average R^2 was 0.83, 0.76 and 0.61 (Fig. 3b, c, d). The average length of the time series for all four scenarios was 40.7 time steps. We obtained parameter estimates for 3919 of the 4000 simulated data sets (see Appendix S1 for details on convergence). When a simple linear regression model was fit to the simulated data, the estimated effect had a positive bias when R^2 was greater than zero and the relative magnitude of the bias increased as R^2 increased (Fig. 4a and 5a). The positive bias that occurred when using linear regression resulted in low coverage probabilities for the 95% CIs,

but a high power when $R^2 > 0.2$ (Fig. 4 b, c). When a correlated errors regression model was fit to the simulated data, the estimated effect was less than the true value when $R^2 < 0.6$ and greater than the true value when $R^2 > 0.6$ (Figs 4a and 5b). When using the correlated errors model, the coverage probability for the 95% CIs was close to 0.95 when $R^2 < 0.6$ and decreased abruptly when $R^2 > 0.6$. When the Ricker model was fit to the simulated data, there was minimal bias and coverage probabilities were close to 0.95 regardless of the value of R^2 for the effect of the covariate on the equilibrium population size (γ_1; eqn 8), however, there was bias and a decline in coverage probability for the growth rate (α_1; eqn 7) as R^2 increased (Fig. 4a, b). When using the Ricker model, the width of the 95% CIs for the growth rate (α_1) increased dramatically as R^2 increased (Fig. 5c), which resulted in a decrease in power (Fig. 4c). In this simulation experiment, the correct model (eqn 5) was selected in 50% of the data sets. The Ricker model with an Allee effect (eqn 9) was selected in 13% of the simulated data sets and neither model was selected as best in 37% of the simulated data sets.

In our second simulation experiment, we obtained parameter estimates for 3558 of the simulated data sets when the population decline was driven by a single factor (x_t) and 3618 of the 4000 when the population decline was driven by two factors (x_t and z_t). When the population decline was driven by a single factor (x_t), the true model had the lowest AICc score in 47% of the simulated data sets. When the population
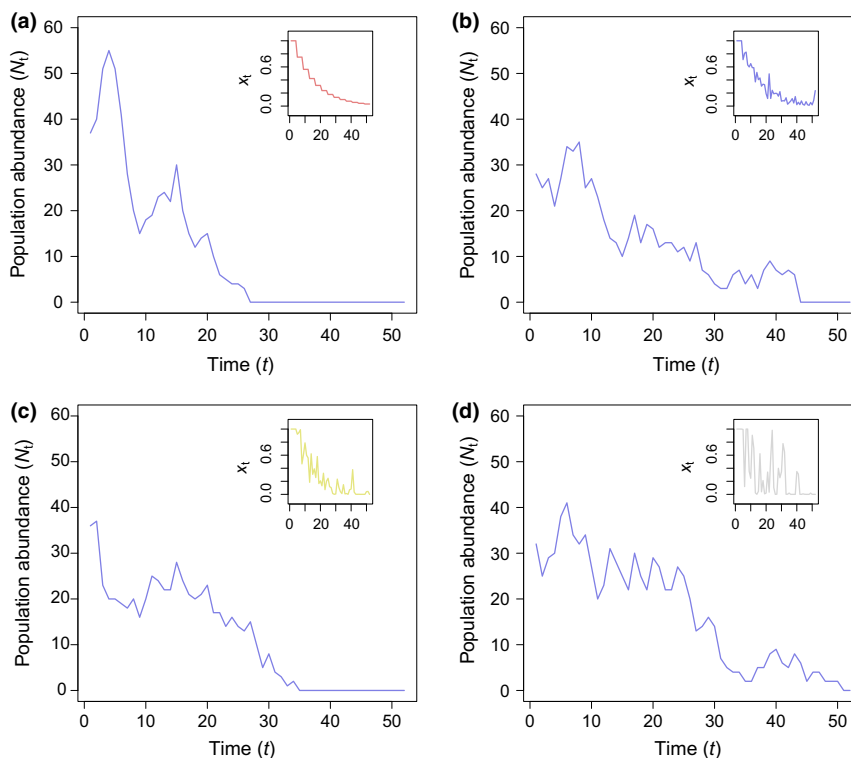

Figure 3 Time series of the number of individuals (N_t) in a population simulated from a Ricker model that allowed for linear declines in $r(x_t)$ and $K(x_t)$ (eqn 5 assuming eqns 7 and 8). We simulated 1000 data sets from each scenario where the covariate x_t (subplots) had a geometric trend with increasing levels of random variation. The average coefficient of determination (R^2) for population abundance N_t and the covariate x_t from each scenario was 0.88 (panel a), 0.83 (panel b), 0.76 (panel c) and 0.61 (panel d).
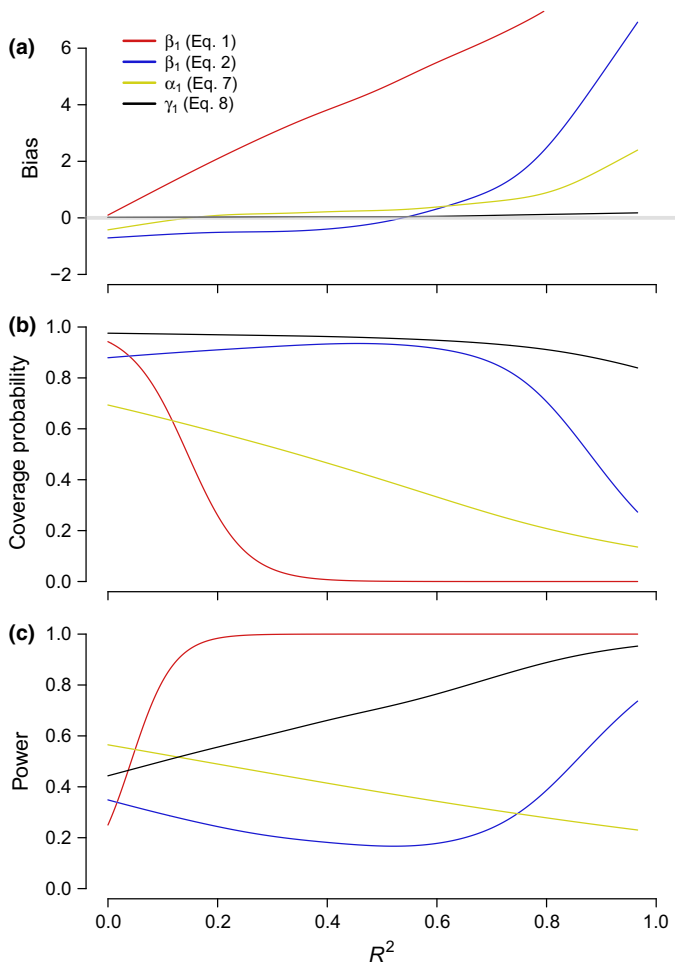
Figure 4 The bias (panel b; grey line is at zero), coverage probability (panel c), and power to detect a statistically significant effect of the covariate obtained from fitting a linear regression model (eqn 1), a correlated errors model (eqn 2), and a Ricker model (eqn 5 assuming eqns 7 and 8) depended on the the coefficient of determination (R^2) for population abundance N_t and the covariate x_t obtained from 4000 simulated time series (see Fig. 3 for example of time series).

decline was driven by two factors (x_t and z_t), the true model had the lowest AICc score in 18% of the simulated data sets.

In our third simulation experiment, we obtained parameter estimates for 3836, 3963, 3979 and 3886 of the 4000 simulated data sets (time series length 10, 20, 40 and 100 respectively). The lower limit of the 95% CIs for α_1 or γ_1 were greater than zero (i.e., a statistically significant false-positive effect of the covariate was detected) in 26, 26, 18 and 13% of the simulated data sets (time series length 10, 20, 40 and 100 respectively).

In our fourth simulation experiment, where a Ricker model with an Allee effect was used to simulate the data, we obtained parameter estimates for 3734 simulated data sets. The correct model (eqn 9) was selected in 71% of the simulated data sets. The Ricker model that allowed for linear declines in $r(x_t)$ and $K(x_t)$ was selected in 12% of the simulated data sets and neither model was selected as best in 17% of the simulated data sets.

DISCUSSION

Temporal confounding can occur when the factors driving the decline in abundance change gradually over time. For example, in the populations of *D. magna*, the amount of food slowly declined over time, while the populations simultaneously declined in abundance. This strong feedback generated dynamics where the endogenous density-dependent factor was highly correlated with the covariate. Similar to the effects of multicollinearity, when the endogenous density-dependence factor is highly correlated with the covariate, coefficient estimates are unstable and confidence intervals are much wider than what would be obtained under moderate levels of confounding (e.g. Fig. 5c; Dormann *et al.* 2013). In traditional regression models, high levels of multicollinearity among covariates can result in statistical inference that is sensitive to model specification (Dormann *et al.* 2013). We tested several different phenomenological regression-based and mechanistic population dynamics models that could be used to determine the cause of the decline and found that inference was sensitive to model specification, even at relatively low levels of confounding.

Temporal confounding can be alleviated if replicate populations are available. Regardless of the method that was used, we were able to detect a positive and statistically significant effect of the food level treatment in the combined analysis of all 60 time series. For the combined analysis, inference from the Ricker model showed that declining food levels resulted in both a decline in the equilibrium population size and intrinsic growth rate. Our results corroborate the experiment of Griffen & Drake (2008) which showed that habitat quantity (volume of the microcosm environment) and quality (amount of food) influences the equilibrium population size and intrinsic growth rate.

The implications of a declining intrinsic growth rate were explored by Abrams (2002), who argued that small or declining population size may not warn of impending extinction thresholds. Based on our results, we recommend that studies consider explicit population growth models when linking covariates to declines in abundance. Although phenomenological regression-based models might be a reasonable approach for making statistical inference in some cases, such models do not capture important dynamics of declining populations (e.g. Abrams 2002; Drake & Griffen 2010; Hefley *et al.* 2013b).

Obtaining data from replicate populations may not be feasible for some studies. In the remainder of this discussion, we assume that a time series from a single population is all that is available. In the analysis of the experimental data and the simulation study, there is a high probability that the influence of an endogenous covariate can be distinguished from an Allee effect. For models that capture the dynamics of declining populations due to an endogenous factor such as declining food levels, it is likely that both the intrinsic growth rate and equilibrium population should be allowed to decline as the population is driven to extinction (Abrams 2002; Drake & Griffen 2010; Hefley *et al.* 2013b). From our results, it appears that the effect of an endogenous factor on the equilibrium population size can be inferred with moderate confidence when temporal confounding is an issue (e.g. Fig. 2d). The
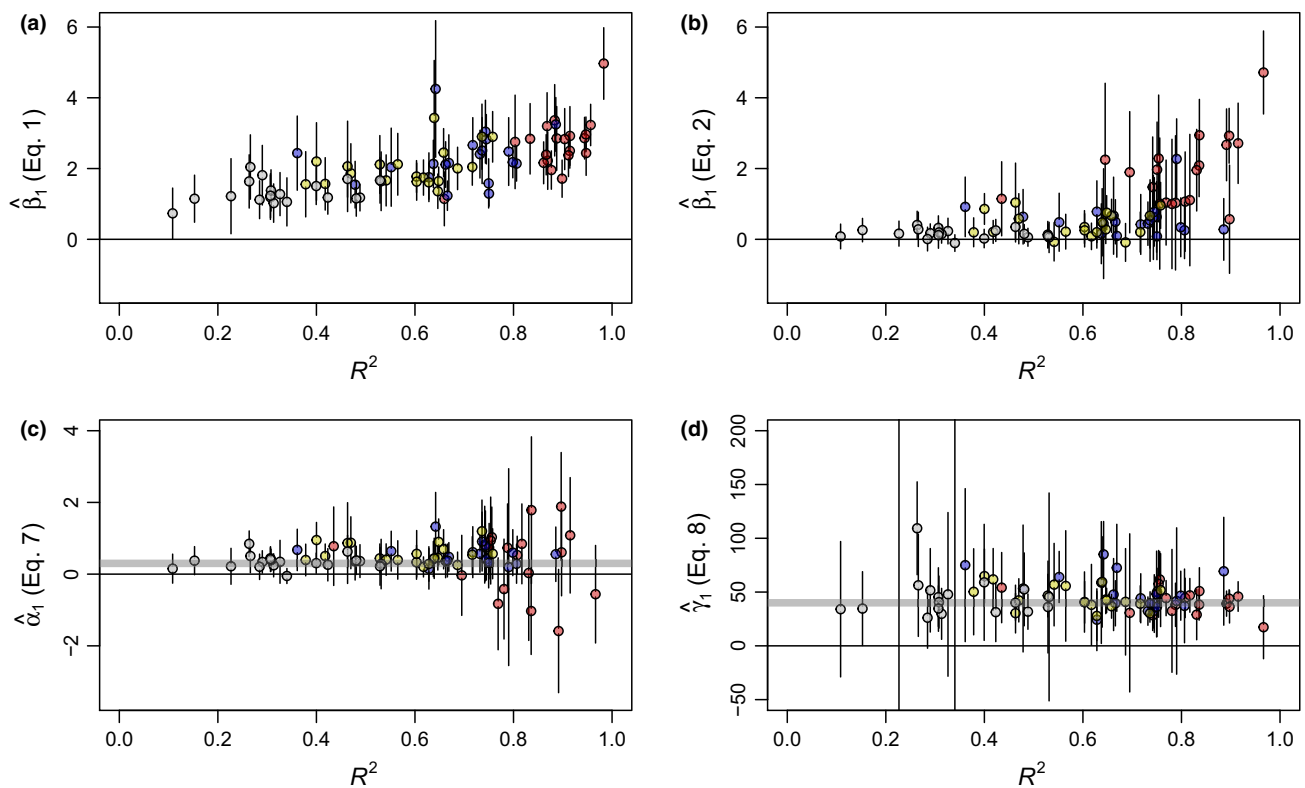
Figure 5 Regression coefficient estimates (coloured dots) and 95% confidence intervals (vertical lines) for the covariate from 20 time series generated from each scenario shown in Fig. 3 plotted against the coefficient of determination (R^2) for population abundance N_t and the covariate x_t. Regression coefficients estimates and 95% confidence intervals were obtained from a simple linear regression model (panel a; eqn 1), a correlated errors model (panel b; eqn 2), a Ricker model (eqn 5) with linear declines in the growth rate ($r(x_t) = \alpha_0 + \alpha_1 x_t$; panel c) and equilibrium population size ($K(x_t) = \gamma_0 + \gamma_1 x_t$; panel d). Note that the colour of the points correspond to the colour used to display the covariates in each scenario shown in Fig. 3. The grey line is at the true value of the coefficients.

effect of an endogenous factor on the intrinsic growth rate is difficult to obtain reliable inference for, even at moderate levels of confounding (e.g. $R^2 = 0.6$, Fig. 5c). When temporal confounding occurs and multiple factors are driving the population decline, results from our second simulation experiment suggest that it is difficult to obtain reliable inference regarding which factors are responsible for the decline.

Our study was limited to a microcosm experiment, but the mechanism generating temporal confounding is reasonable for many species that are declining in abundance (e.g. Roseberry *et al.* 1979; Veech 2006; Hefley *et al.* 2013b). For studies focused on linking covariates to population declines, it is important to recognise and detect when exogenous factors are confounded with endogenous factors. For some studies, factors that cause a population decline maybe dynamic, such as the influence of another species (e.g. Reilly & Zeringue 2005; Cangelosi & Hooten 2009). In such cases, it is important to determine how temporal confounding influences causal inference using methods that we have not considered in this study (e.g. Sugihara *et al.* 2012; Clark *et al.* 2015; Ye *et al.* 2015).

Although the results from our study highlight a caveat in commonly used methods to determine the cause of a decline in abundance, it is encouraging that temporal confounding is similar to multicollinearity, which has been successfully identified and addressed in many studies (Dormann *et al.* 2013). One of the most successful methods that has been developed

to alleviate multicollinearity in regression models is regularisation (e.g. Hoerl & Kennard 1970). For Bayesian models, regularisation is an automatic consequence of using informative priors (Hooten & Hobbs 2015). We expect that the use of informative priors will be an important component to statistical models that are used to determine the factors that are causing population declines because for many population there is some existing knowledge about growth rates and carrying capacity (Hobbs & Hooten 2015, pp. 91–106). For example, using the process outlined in Delean *et al.* (2013) informative priors could be developed for the maximum growth rate and carrying capacity (e.g. the maximum attainable values of $r(x_t)$ and $K(x_t)$ in eqn 5). Future work that outlines best practices for specifying informative priors on the coefficients for covariates that influence the growth rate and carrying capacity (e.g. α_1 and γ_1 in eqns 7 and 8) will be critical to harnessing the full power of regularisation. We expect that many studies will find it advantageous to explore the use of informative priors with population growth models to determine the causes of population declines.

AUTHORSHIP

TJH, MBH, JMD, RER and DPW conceived the study. TJH and MBH developed the statistical methods. JMD designed the microcosm experiment. TJH conducted the statistical analysis, simulation experiment and wrote the first draft of the manuscript. All authors contributed substantially to revisions.

REFERENCES

Abrams, P.A. (2002). Will small population sizes warn us of impending extinctions? *Am. Nat.*, 160, 293–305.

Beverton, R.J. & Holt, S.J. (1957). *On the Dynamics of Exploited Fish Populations*. Springer Science & Business Media, Dordrecht.

Bini, M., Diniz-Filho, J.A.F., Rangel, T.F., Akre, T.S., Albaladejo, R.G., Albuquerque, F.S., *et al.* (2009). Coefficient shifts in geographical ecology: an empirical evaluation of spatial and non-spatial regression. *Ecography*, 32, 193–204.

Burnham, K.P. & Anderson, D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. Springer Science & Business Media, New York.

Cangelosi, A.R. & Hooten, M.B. (2009). Models for bounded systems with continuous dynamics. *Biometrics*, 65, 850–856.

Caughley, G. (1994). Directions in conservation biology. *J. An. Ecol.*, 63, 215–244.

Clark, J.S. & Bjørnstad, O.N. (2004). Population time series: process variability, observation errors, missing values, lags, and hidden states. *Ecology*, 85, 3140–3150.

Clark, A.T., Ye, H., Isbell, F., Deyle, E.R., Cowles, J., Tilman, G.D. & Sugihara, G. (2015). Spatial convergent cross mapping to detect causal relationships from short time series. *Ecology*, 96, 1174–1181.

Cressie, N. & Wikle, C. (2011). *Statistics for Spatio-Temporal Data*. Wiley, Hoboken.

Delean, S., Brook, B.W. & Bradshaw, C.J. (2013). Ecologically realistic estimates of maximum population growth using informed bayesian priors. *Methods Ecol. Evol.*, 4, 34–44.

Dennis, B. & Ponciano, J.M. (2014). Density-dependent state-space model for population-abundance data with unequal time intervals. *Ecology*, 95, 2069–2076.

Dennis, B., Ponciano, J.M., Lele, S.R., Taper, M.L. & Staples, D.F. (2006). Estimating density dependence, process noise, and observation error. *Ecol. Monogr.*, 76, 323–341.

Dormann, C.F., Elith, J., Bacher, S., Buchmann, C., Carl, G., Carré, G., *et al.* (2013). Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography*, 36, 27–46.

Drake, J.M. & Griffen, B.D. (2010). Early warning signals of extinction in deteriorating environments. *Nature*, 467, 456–459.

Fieberg, J. & Ditmer, M. (2012). Understanding the causes and consequences of animal movement: a cautionary note on fitting and interpreting regression models with time-dependent covariates. *Methods Ecol. Evol.*, 3, 983–991.

Griffen, B.D. & Drake, J.M. (2008). Effects of habitat quality and size on extinction in experimental populations. *Proc. R. Soc. Lond. B Biol. Sci.*, 275, 2251–2256.

Hefley, T.J., Tyre, A.J. & Blankenship, E.E. (2013a). Fitting population growth models in the presence of measurement and detection error. *Ecol. Modell.*, 263, 244–250.

Hefley, T.J., Tyre, A.J. & Blankenship, E.E. (2013b). Statistical indicators and state-space population models predict extinction in a population of bobwhite quail. *Theor. Ecol.*, 6, 319–331.

Hefley, T.J., Broms, K.M., Brost, B.M., Buderman, F.E., Kay, S.L., Scharf, H.R., *et al.* (2016). The basis function approach for modeling autocorrelation in ecological data. https://arxiv.org/abs/1606.05658.

Hobbs, N.T. & Hooten, M.B. (2015). *Bayesian Models: A Statistical Primer for Ecologists*. Princeton University Press, Princeton.

Hodges, J.S. & Reich, B.J. (2010). Adding spatially-correlated errors can mess up the fixed effect you love. *Am. Stat.*, 64, 325–334.

Hoerl, A.E. & Kennard, R.W. (1970). Ridge regression: biased estimation for nonorthogonal problems. *Technometrics*, 12, 55–67.

Hooten, M. & Hobbs, N. (2015). A guide to bayesian model selection for ecologists. *Ecol. Monogr.*, 85, 3–28.

Hurvich, C.M. & Tsai, C.-L. (1989). Regression and time series model selection in small samples. *Biometrika*, 76, 297–307.

Ives, A.R. (2015). For testing the significance of regression coefficients, go ahead and log-transform count data. *Methods Ecol. Evol.*, 6, 828–835.

Jacobson, A.R., Provenzale, A., von Hardenberg, A., Bassano, B. & Festa-Bianchet, M. (2004). Climate forcing and density dependence in a mountain ungulate population. *Ecology*, 85, 1598–1610.

Knape, J. & de Valpine, P. (2011). Effects of weather and climate on the dynamics of animal population time series. *Proc. R. Soc. Lond. B Biol. Sci.*, 278, 985–992.

Kühn, I. (2007). Incorporating spatial autocorrelation may invert observed patterns. *Divers. Distrib.*, 13, 66–69.

Newman, K.B., Buckland, S., Lindley, S., Thomas, L. & Fernandez, C. (2006). Hidden process models for animal population dynamics. *Ecol. Appl.*, 16, 74–86.

Pinheiro, J. & Bates, D. (2000). *Mixed-effects models in S and S-PLUS*. Springer, New York.

Reilly, C. & Zeringue, A. (2005). *Improved Predictions of Lynx Trappings Using a Biological Model*, pp. 297–308. John Wiley & Sons, West Sussex England.

Ricker, W.E. (1954). Stock and recruitment. *J. Fish. Board Can.*, 11, 559–623.

Roseberry, J.L., Peterjohn, B.G. & Klimstra, W. (1979). Dynamics of an unexploited bobwhite population in deteriorating habitat. *J. Wildl. Manag.*, 43, 306–315.

Roy, C., McIntire, E.J. & Cumming, S.G. (in press). Assessing the spatial variability of density dependence in waterfowl populations. *Ecography*. DOI 10.1111/ecog.01534.

Royle, J.A. (2004). N-mixture models for estimating population size from spatially replicated counts. *Biometrics*, 60, 108–115.

Sugihara, G., May, R., Ye, H., Hsieh, C.-H., Deyle, E., Fogarty, M. & Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, 338, 496–500.

de Valpine, P. & Hastings, A. (2002). Fitting population models incorporating process noise and observation error. *Ecol. Monogr.*, 72, 57–76.

Veech, J.A. (2006). Increasing and declining populations of northern bobwhites inhabit different types of landscapes. *J. Wildl. Manage.*, 70, 922–930.

Ye, H., Deyle, E.R., Gilarranz, L.J. & Sugihara, G. (2015). Distinguishing time-delayed causal interactions using convergent cross mapping. *Sci. Rep.*, 5, 14750.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.