Using Fourier series to estimate periodic patterns in dynamic occupancy models

MASON FIDINO† AND SETH B. MAGLE

Department of Conservation and Science, Lincoln Park Zoo, Urban Wildlife Institute, Chicago, Illinois 60614 USA

Abstract. Some of the most impressive adaptations of organisms are in response to periodic environmental variability. To capture these temporal dynamics, statistical models that estimate the spatiotemporal distribution of a species typically include categorical seasonal covariates, temporally varying parameters, or smoothing splines. While these techniques provide a useful starting point, they may require many parameters to estimate and are not well suited for making predictions. Here, we present a technique that uses Fourier series to estimate periodic signals in dynamic occupancy models, and parameterize these models with data from a large-scale long-term camera trapping study of medium to large mammals in Chicago, Illinois, USA. Our periodic models captured up to 75% of the temporal variability in species colonization rates and performed similar to dynamic occupancy models with temporally varying parameters. Overall, this method can partition variability between periodic and non-periodic sources and estimate the proportion of temporal variability that is attributable to a periodic source in a model-based framework. Further, practitioners can use this method to incorporate prior knowledge on a species' natural history (e.g., natal dispersal and migration). This will, in turn, create more biologically reasonable models for conservation and management applications.

† **E-mail:** mfidino@lpzoo.org

INTRODUCTION

Meteorological changes throughout the year cause periodic variability in the climate of a region (Crowley and Burke 1998). Such changes are often observed as seasonal patterns in temperature, humidity, rainfall, ocean currents, or wind patterns, which, in turn, temporally influence the selective pressures of the local environment where organisms reside. For example, seasonally dependent pulses in resource abundance can have a profound effect on the coexistence rates of species (Brown 1989). Some of the most impressive behavioral and reproductive adaptations of organisms are in response to seasonal shifts in the local abundance of resources

(Immelmann 1971). For instance, seasonal migrations to more favorable conditions for breeding or foraging are some of the most conspicuously observed events that occur across taxa (Berthold 2001, Lucas et al. 2001). Likewise, in environments with a restricted time span of favorable breeding conditions, the birth and subsequent dispersal of young generally follow a periodic pattern (Bradshaw and Holzapfel 2007). Although the mechanisms that facilitated the evolution of such responses are varied and often debated (Boyle and Conway 2007), the fact that the spatial distribution of many species varies temporally, and often periodically, is not.

To quantify spatiotemporal changes in a species' distribution, statistical models generally use

categorical seasonal covariates (Nielsen et al. 2010), temporally varying parameters (Kéry et al. 2013), or smoothing splines (Thorson et al. 2012). While such techniques provide a useful starting point for understanding temporal dynamics, they may require many parameters to estimate and can be difficult to interpret or use for predictions. Further, such techniques cannot partition variability between periodic sources and deviations from them. Temporally varying covariates that follow a periodic pattern (e.g., local temperature) can be used, provided such data exist, but provide inference for the temporal covariate and not for time itself. As periodic patterns are repeatable and therefore predictable, an opportunity exists to leverage such patterns if they can be explicitly specified within a statistical model.

Fourier series are a particularly promising approach to quantify periodic variability in some underlying phenomena, and do so by converting periodic signals into frequencies driven by sums of sines and cosines (Shumway and Stoffer 2010). Further, Fourier series have been used to quantify a variety of complex periodic patterns in ecology (Grover et al. 2000, Sabo and Post 2008). Here, we present a technique that incorporates Fourier series into dynamic occupancy models, a statistical framework developed by MacKenzie et al. (2006) that estimates the spatiotemporal dynamics of species distributions. Although Fourier series can construct any type of periodic signal, we focus our attention on those informed by the life history of a target species. Such information can be formulated as simple Fourier series that represent an explicit periodic pattern. Further, the technique we use here only requires two parameters to estimate these periodic patterns. Practitioners can use this method to develop more biologically meaningful models that partition variability between periodic and non-periodic patterns in a versatile model-based framework.

Model Development

The formulation of dynamic occupancy models has been covered in depth (MacKenzie et al. 2006). Therefore, we briefly outline a dynamic occupancy model for a single species potentially located at k of $1, 2, \ldots, K$ sites and sampled across t in $1, 2, \ldots, T$ time steps (i.e., primary sampling periods; see Appendix S1 for a more thorough description). After estimating the probability of initial occupancy (ψ) at site k and time $t = 1$, dynamic occupancy models assume a first-order Markov process that estimate the probability a species colonizes (γ) site k at time t given their absence at $t - 1$ or the probability a species persists (ϕ) at site k and time t given their presence at $t - 1$. Additionally, repeated surveys are conducted during each primary sampling period to generate detection histories. These detection histories are then used to estimate the conditional probability of detecting a species given their presence (p). Through a variety of link functions, each of these probabilities can be made a function of covariates so that a variety of hypotheses may be tested. For example, to account for temporal variability in colonization rates, γ may be structured such that

$$\text{logit}(\gamma_{k,t}) = m_0 + m_1 x_k + u_t \qquad (1)$$

where m_0 is the intercept, m_1 is the effect of covariate x collected at site k, and u_t is a coefficient that allows m_0 to vary temporally such that $u_t \sim N(0, \sigma)$ and σ is estimated from the data. Thus, Eq. 1 accounts for general temporal variability but does not allow for explicit tests of periodic signals.

A foray into Fourier series

A mathematical function, $x_P(t)$, with known periodicity can be approximated as a Fourier series with a single frequency over time, t,

$$x_P(t) \approx a_0 + \sum_{n=1}^{\infty} (a_n \cos(n w_0 t) + b_n \sin(n w_0 t)) \qquad (2)$$

where a_n and b_n are Fourier coefficients defined as integrals of the periodic function, $x_P(t)$, a_0 is the average value of the series, and w_0 is $2\pi/P$, which represents the frequency over period P (number of cycles per unit time). For example, if $P = 8$ the cycle repeats every eight time steps. As $n \to \infty$ the Fourier series more closely approximates the periodic function $x_P(t)$. In the statistical analysis of time series, this summation over n is generally not possible as there is no periodic function to derive a_n and b_n; there are only data with a periodic pattern. Instead, a periodic series

to be estimated from data is generally represented as:

$$x_P(t) = a_0 + a\cos(w_o t) + b\sin(w_o t) \qquad (3)$$

where a and b are independent normally distributed random variables with mean zero and standard deviations σ (Shumway and Stoffer 2010). In such cases, the amplitude (the maximum absolute value of a series) is $A = \sqrt{a^2 + b^2}$, and the starting point of the series, also known as the phase shift, is $\delta = \tan^{-1}(-b/a)$. Eq. 3 can be fit to data to estimate the presence of repetitive sinusoidal oscillations through time.

To fit more complicated periodic patterns, Eq. 3 can be generalized to allow for up to H mixtures of periodic series at varying frequencies and amplitudes,

$$x_P(t) = a_0 + \sum_{h=1}^{H}(a_h\cos(w_h t) + b_h\sin(w_h t)) \qquad (4)$$

where a_h and b_h are independent normally distributed random variables with mean zero and standard deviations σ_h, and w_h are different frequencies (i.e., P takes a different value for each w_h). Through the summation of sines and cosines, different periodic shapes take form and could be used to identify biologically relevant patterns at varying scales (e.g., diurnal patterns in a species' activity rate embedded within seasonal patterns). Eq. 4 can fit any complex periodic signal given sufficient data but does so at the cost of two additional parameters for each h. Further, while Eq. 4 can be used to identify the frequencies that best explain a periodic signal (Cressie and Wikle 2015), it precludes the possibility of fitting explicit periodic patterns motivated by a species' life history (e.g., a population increase during a species' breeding season). In such cases where a pattern with known periodicity is hypothesized to be present, it is possible to instead leverage the properties of the Fourier series in Eq. 2 to derive a_n and b_n and fit a simpler model that requires only two parameters to estimate the periodic pattern. We illustrate how Fourier series of this type can be included within a dynamic occupancy model in the next section, using the life history strategies of five different mammalian species.

Single-season pulses

Let us assume we have collected detection/non-detection data at K sites in the spring, summer, fall, and winter across multiple years and would like to model periodic dispersal rates in species that breed once per year. In our example, this categorization would describe the coyote (*Canis latrans*), red fox (*Vulpes vulpes*), and striped skunk (*Mephitis mephitis*, Feldhamer et al. 2003). Each year these species breed in the late winter to early spring and the resulting offspring disperse in the fall (Feldhamer et al. 2003). Given this life history strategy, we would predict that habitat patches would experience a pulse in colonization rates as juveniles disperse in the fall. In this case, a periodic pulse Fourier series can be written with only three parameters, A, δ, and a_0, the last of which is an intercept term that represents the average of the series (see Appendix S2 for a derivation of Eq. 5):

$$\begin{aligned} x_P(t) = a_0 + \sum_{n=1}^{P} \frac{2A}{\pi n}\sin\left(\frac{\pi n}{P}\right) \\ \cdot \cos\left(\frac{2\pi n}{P}(t-\delta)\right) \end{aligned} \qquad (5)$$

Eq. 5 provides a general equation that simultaneously estimates a periodic pulse's amplitude and location. Note that unlike Eq. 2 the summation over n ranges from 1 to P and not ∞. The reason for this is that fewer n are needed for a Fourier series to approximate the periodic function when modeling events over discrete time ($t \in \mathbb{Z}$) instead of continuous time ($t \in \mathbb{R}$). Our own trials suggest that n should at least range from 1, 2,..., P for model fitting (i.e., a P order approximation). The next step is to incorporate Eq. 5 into the colonization function of a dynamic occupancy model so that we may estimate a_0, A, and δ from the data. Assuming that P is known a priori, we can use the trigonometric formula $\cos(u-v) = \cos(u)\cos(v) + \sin(u)\sin(v)$ to separate terms that contain A and δ in Eq. 5 and specify them as parameters to be estimated from data:

$$\beta_{1,n} = A\cos\left(\frac{2\pi n}{P}\delta\right), \quad \text{for } n = 1, 2, \ldots, P$$

$$\beta_{2,n} = A\sin\left(\frac{2\pi n}{P}\delta\right), \quad \text{for } n = 1, 2, \ldots, P \qquad (6)$$

The remaining terms are then combined into two T × P matrices of covariates:

$$C = \cos\left(\frac{2\pi n}{P}t\right)\sin\left(\frac{\pi n}{P}\right)\frac{2}{\pi n},$$
$$\text{for } n = 1, 2, \ldots, P \text{ and } t = 1, 2, \ldots, T$$

$$S = \sin\left(\frac{2\pi n}{P}t\right)\sin\left(\frac{\pi n}{P}\right)\frac{2}{\pi n}, \tag{7}$$
$$\text{for } n = 1, 2, \ldots, P \text{ and } t = 1, 2, \ldots, T$$

These parameters and covariates can then be added to the linear predictor of a model. For example, including these periodic parameters and covariates in Eq. 1 would result in the linear predictor

$$\text{logit}(\gamma_{k,t}) = m_0 + m_1 x_k + u_t$$
$$+ \sum_{n=1}^{P} \beta_{1,n} C_{t,n} + \beta_{2,n} S_{t,n} \tag{8}$$

where a_0 in Eq. 5 is replaced by the intercept term, m_0. This results in a model that can estimate a periodic pulse and deviations from the pulse through the variance term in u_t. To increase identifiability of all parameters within Eq. 8, a preliminary model without a temporally random component (u_t) should be fit to the data so that the location of a pulse (δ) can first be estimated. Following this, the most likely value for δ can be incorporated into Eq. 8 so that u_t does not influence δ. The variance term in Eq. 8 associated with u_t is a parameter of interest as it represents the temporal variability that the periodic source does not account for. By fitting another model without a periodic γ component (e.g., Eq. 1), the proportion of temporal variability explained by the periodic source can be calculated:

variability explained by Fourier series
$$= 1 - \frac{E(\sigma_1)}{E(\sigma_2)} \tag{9}$$

where σ_1 and σ_2 are the expected values of the standard deviation estimates associated with u_t from separate models that, respectively, use Eqs. 8 and 1 for γ.

Multiple pulses a year

Other species can breed multiple times a year or have young that disperse at different times. Raccoons (*Procyon lotor*) have a bimodal pattern in their mating distribution where juvenile dispersal occurs in the spring or fall (Feldhamer et al. 2003). Virginia opossum (*Didelphis virginiana*, hereafter opossum) ordinarily have two litters per year and male opossum born in the first litter disperse in the fall of the same year while the second cohort disperses during the next spring (Feldhamer et al. 2003). Thus, we would predict a boom-bust pattern in these species' colonization rates, with peaks in the spring and fall. This pattern could be approximated with a special case of Eq. 5 when $n = 1$. This is because the additional n terms in Eq. 5 flatten the trough of the periodic pulse, while the first n is a sinusoidal curve. Therefore, when $n = 1$, Eq. 5 simplifies into:

$$x_P(t) = a_0 + A\cos\left(\frac{2\pi}{P}(t - \delta)\right) \tag{10}$$

where the additional terms in Eq. 5 are removed as they are linear transformations of the random variable A. Consequently, this model is identical to what Flury and Levri (1999) used to estimate periodic signals in the diel foraging patterns of snails using logistic regression. As in the last section, Eq. 10 can be simplified to parameters that are estimated from the data and two covariate vectors of length T, which can then be included in a dynamic occupancy model:

$$\text{logit}(\gamma_{k,t}) = m_0 + m_1 x_k + u_t$$
$$+ \beta_1 c_t + \beta_2 s_t \tag{11}$$

Note that periodic changes in extinction rates (i.e., 1—probability of persistence) do not need to occur for there to be periodic changes in a species colonization rate. So long as the probability of persistence is <1, sites can become available for colonization from one time step to the next, periodic or otherwise.

EXAMPLE: MESOCARNIVORE COLONIZATION RATES IN CHICAGO, ILLINOIS

Here, we present an application of the technique proposed in the previous section and parameterize models with real data from a large-scale, long-term camera trapping study of medium to large mammals throughout Chicago, Illinois, USA.

Sampling protocol

To collect site use data for medium- and large-sized mammals, starting in the spring of 2011, we set Bushnell motion-triggered camera traps

(model #119436C) in the spring, summer, fall, and winter for 28 d per season at sites along three 50-km transects. These transects begin in downtown, Chicago, Illinois, USA, and radiate outwards to the northwest, west, and southwest (Fidino et al. 2016). Four different site types were selected for this study: forest preserves, city parks, cemeteries, and golf courses, which together represent much of the greenspace throughout Chicago. Between Spring 2011 and Fall 2013, 118 sites were sampled, although not every site was sampled each season due to logistical constraints. Sites were only included in our analysis if they had ≥2 seasons worth of data ($n = 95$).

We considered each sampling session (e.g., Summer 2012) as a primary sampling period and generated detection histories at each site by treating each day a camera trap was active as a repeat survey over a given 28-d sampling period. Cameras could become inactive if they malfunctioned, were vandalized, or if they ran out of battery power. Most camera traps were active for the entire sampling season (median days active = 27). Thus, for each sampling period we calculated (1) the number of days each species was photographed at each sampling location and (2) the number of days a camera was functional at each sampling location. These data were then used within a dynamic occupancy model (Appendix S1). For more specific details of our sampling protocol, see Magle et al. (2015).

Given this sampling design, it is important to formally define colonization and persistence for this dataset. Unlike biogeographic islands, the sites sampled throughout Chicago are not discrete, species may persist within the habitat matrix between sites, and sites may not even be of sufficient size or quality to maintain a breeding population (e.g., a small city park where coyotes were detected). Thus, we do not assume that the presence of a species at a site represents the existence of a local population within its geographic boundary; it is simply the probability that the local abundance around a sampling location is ≥1 during a sampling season. As such, we define colonization as the probability that an unoccupied site in the previous time step is occupied by the species of interest in the current time step, whereas the persistence rate is the probability that an occupied site in the previous time step continues to contain that species in the current

time step. Therefore, for our data, colonization and persistence rates for each species may respectively be interpreted as metrics of intermittent and continued habitat use over time.

Models, specification of priors, model selection, and model assessment

We included five species in our analysis: coyote, red fox, striped skunk, raccoon, and opossum. To compare the efficacy of Fourier series to other techniques, we compared our periodic time model (PTM) to a stochastic time model (STM) and homogeneous time model (HTM). All models were identical save for what was included in the linear predictor of the colonization function. For each model, we used empirical Bayes methods to specify the prior for initial occupancy (Ψ). To do so, we used our presence/absence data for each species to calculate the shape parameters of a beta prior from the mean (μ) and standard deviation (σ) of the proportion of sites where a species was detected at each time step, such that $\Psi_i \sim \text{Beta}(a_i, b_i)$ for species i (Appendix S3). While initial occupancy may also be made a function of covariates, it only represents a brief snapshot of patch dynamics instead of the processes that occurred throughout this survey (i.e., colonization and persistence, MacKenzie et al. 2006). Thus, we used this simplified parameterization.

In temperate urban environments, canopy cover is often used as a proxy for habitat availability for these species while other factors such as impervious cover or housing density represent a lack thereof (Gehrt et al. 2010, Magle et al. 2015). Thus, to account for spatial factors that may influence p, γ, and ϕ, we calculated mean tree cover, mean impervious cover, and mean housing density within a 500-m buffer around each site using QGIS ver. 2.14 (QGIS Development Team 2016). These values were then z-transformed. Following this, we applied principal component analysis to these data. The first principal component explained 77.31% of the variation in these data and was included as a spatial covariate for p, γ, and ϕ in each model (hereafter URB). Positive values of URB indicate sites with higher mean canopy cover, while negative values indicate sites with higher housing densities and impervious cover. Further, we specified random effects to allow intercepts to vary temporally for ϕ in all models and varied p

by site and time in the same fashion. All random effects were drawn from their own respective $N(0, \sigma)$ prior where $\sigma \sim$ half-Cauchy$(0, 25)$.

Periodic time models for coyote, red fox, and striped skunk used the Fourier series in Eq. 5 to predict colonization rates where $P = 4$. The Fourier series in Eq. 10 was used for raccoon and opossum where $P = 2$. As the amplitude of a pulse must be positive A was given a Gamma (1, 1) prior, which places 99% of the probability mass between 0 and 4.61. The phase shift, δ, must be an integer so we set $\delta \sim$ categorical$(\rho_1, \ldots, \rho_p)$ for $\delta \in \{0, \ldots, P-1\}$ where $\rho_{1:P} = 1/P$. Stochastic time models used the colonization function in Eq. 1. Homogeneous time models had a $N(0, 1.83)$ prior for the γ intercept that did not vary through time. For p, γ, and ϕ, all other parameters were drawn from $N(0, 1.83)$ priors for all models. Such a specification places 99% of the probability mass between -4.25 and 4.25, which represents realistic logit-scale parameter estimates, particularly for occupancy models (Broms et al. 2016). See Appendix S3 for all posterior distributions and Data S1 for model code.

To assess the relative fit of each model, we calculated the conditional predictive ordinate (CPO, Geisser 1993) for each data point within an MCMC algorithm. To estimate the overall performance of each model, the summary statistic $-\sum_{k,t} \log(\text{CPO}_{k,t})$ was calculated, with the lowest value indicating the best fit (Hooten and Hobbs 2015). Models were executed with JAGS (Plummer 2003) version 4.2.0, through program R version 3.3.2 (R Core Team 2016). After a 1000 step adaptation and 4000 step burn-in, the posterior distribution of each model was sampled 400,000 times. MCMC chains were thinned by 2. To verify model convergence, we inspected the MCMC chains to ensure proper mixing and ensured that Gelman-Rubin diagnostics for each parameter were <1.10 (Gelman et al. 2014).

Finally, we conducted a simulation analysis to confirm the ability of the model to adequately recover parameter estimates for the two different kinds of PTMs, paying special attention to the regression coefficients related to the periodic patterns (i.e., A and δ). To do so, we simulated 375 datasets for each model with a structure identical to our own data using the median estimates from the posterior distribution of the coyote PTM and the opossum PTM, which, respectively, used

Eqs. 5 and 10 to estimate periodic colonization rates. Following this, we refit the model to each simulated dataset to determine how well it could recover the known parameter values. Models that did not converge were discarded.

RESULTS

Parameter recovery

The single-season pulse PTM (Eq. 5) and the boom-bust PTM (Eq. 10) adequately recovered parameter estimates from the simulated data. For the former, 95% credible intervals of parameters over all simulations included the known parameters >90% of the time, save for the estimates of initial occupancy (ψ, 89%) and the standard deviation estimate of the random effect for site-specific detection probability (σ_ε, 87%; Appendix S4: Fig. S1). The phase shift parameter, δ, estimated the correct location of a periodic pulse as the most likely location 96% of the time (Appendix S4: Fig. S2). However, only 72% of the simulations had $\geq$90% certainty in the estimate of δ (i.e., the correct location of the pulse was selected at least 90% of the time over all MCMC samples for a given simulation). For the boom-bust PTM, 95% credible intervals of parameters over each simulation included the known parameter value >90% of the time save for the covariate effect on persistence (d_1, 89%), and δ was estimated in the correct location of the periodic pulse 99% of the time (Appendix S4: Figs. S3, S4). Furthermore, 88% of the boom-bust simulations had $\geq$90% certainty in the estimate of δ.

Empirical data

Over a total of 20,025 traps nights between Spring 2011 and Fall 2013, raccoon were photographed on 2348 d, opossum 1877 d, coyote 1195 d, striped skunk 357 d, and red fox 146 d. With these data, PTMs had the best fit for coyote, red fox, and opossum, STMs performed best for skunk, and HTMs had the best fit for raccoon (Table 1; see Appendix S5 for all parameter estimates from the best fit model of each species). However, the relatively similar CPO scores of the three models do indicate that their predictive performance is relatively comparable for each species (Table 1).

Overall, STMs were less precise than PTMs and HTMs. Stochastic time models had the largest

Table 1. Model selection results of the three different models fit to nine seasons of camera trap data from Chicago, Illinois, USA.†

Model	Colonization covariates	Coyote	Red fox	Skunk	Raccoon	Opossum
Periodic time	URB + periodic	**1023.88**	**217.61**	442.24	1736.26	**1340.48**
Stochastic time	URB + session	1028.76	224.02	**441.79**	1739.53	1341.67
Homogeneous time	URB	1029.37	219.42	444.13	**1735.10**	1344.21

Notes: CPO, conditional predictive ordinate. The value of the best fit model for each species is in boldface.

† Models were compared with the summary statistic $-\sum_{k,t} \log(\text{CPO}_{k,t})$ for site k and time t. The lowest values indicate the model with the best predictive performance. URB is the first principal component of mean tree cover, mean impervious cover, and mean housing density within a 500-m buffer around each site.

credible intervals in γ (colonization) rates through time and a posterior predictive distribution that provided little certainty about future γ rates (Fig. 1). For example, the 95% predictive interval from the coyote STM predicts that future γ rates would vary between 3.16% and 88.23%. Conversely, the coyote PTM is more precise and predicts that future coyote γ rates would be 23.72% (95% CI = 11.12–41.22%) in trough seasons but increase to 61.20% (95% CI = 37.98–82.28%) in the fall. Homogeneous time models, while more precise than STMs, simply report average γ across time. Homogeneous time models best approximated raccoon γ dynamics but were insufficient for the other species analyzed (Table 1).

In coyote and red fox PTMs, the probability that γ peaked in the fall was, respectively, 98.74% and 98.89%. The opossum PTM estimated γ peaked in the spring and fall (99.42%). Using Eq. 9, the proportion of temporal variability explained by a single-season pulse for coyote, red fox, and striped skunk was, respectively, 75.40%, 74.57%, and 49.01%. The boom-bust PTM accounted for 50.17% of the temporal variability in opossum γ rates. Including the sinusoidal signal in the raccoon PTM increased the amount of temporal variability by 11.28% as the varying intercept in the full model of Eq. 11 drifted more to correct for unnecessary periodic patterns.

Discussion

We present here a method that uses Fourier series in dynamic occupancy models to partition variability between periodic signals and deviations from them. Models with a periodic component best predicted colonization rates in three of the five species analyzed (Table 1). Depending on the species, PTMs accounted for 49.01–75.40% of the overall temporal variability in colonization rates. Although PTMs were not the best model in all cases, our results clearly illustrate their applicability and how they can be used to estimate vital rates of species that vary through time. Such models provide a way to explicitly specify periodic signals in ecological time series that are easy to interpret and use for future predictions, while including only two additional parameters. Most importantly, this method makes it easier to incorporate prior knowledge on the natural history of a species into a statistical model.

In regard to application, PTMs can offer more actionable information compared to STMs or HTMs. For example, to optimize the management of metapopulations, it is critical to know when colonization rates are high, as doing so indicates when new habitat patches should be added (Southwell et al. 2016). While STMs do account for temporal variability within a species' colonization rate, they cannot predict when colonization rates should be high (Fig. 1). Likewise, HTMs assume that colonization rates do not change. For species with periodic signals in their colonization rate, PTMs make it easier to determine when a species colonization probability is highest, which is useful for conservation or management applications.

Our PTMs were effective in analyzing the data collected from camera trap surveys throughout Chicago. Overall, our analysis revealed that colonization rates for most mesocarnivores peaked in the fall (Fig. 1). Certainly, this does not confirm that these periodic increases are the direct result of juvenile dispersal in the fall, but the patterns observed are consistent with this known aspect of these species natural history (Feldhamer et al. 2003). Regardless, as these mammals must traverse the urban matrix to reach different patches of habitable greenspace, this result indicates that species move more throughout the landscape
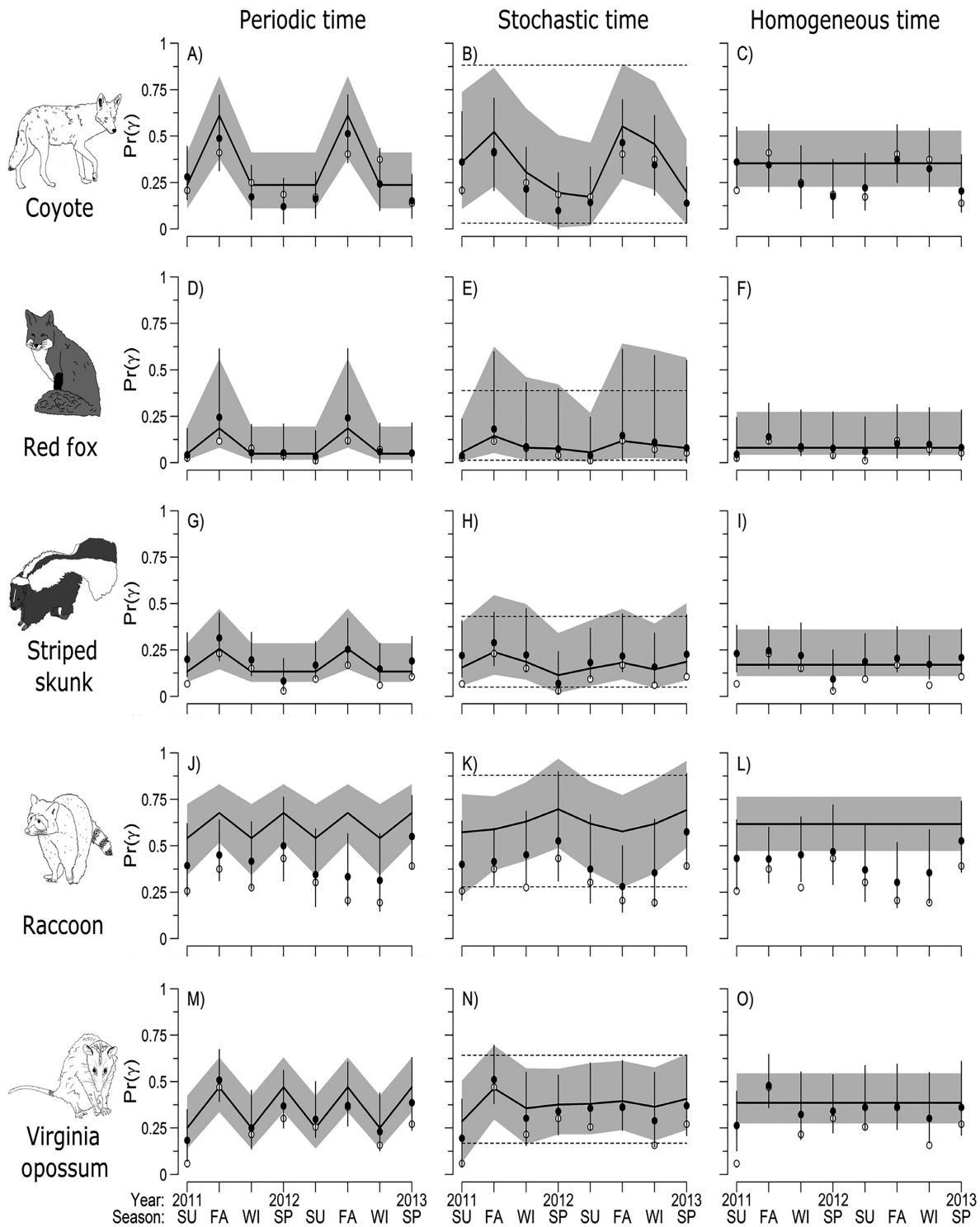
Fig. 1. The probability of colonization over time for (A–C) coyote, (D–F) red fox, (G–I) striped skunk, (J–L)

(Fig. 1. *Continued*)
raccoon, and (M–O) Virginia opossum estimated from nine seasons of camera trap data collected from Spring 2011 to Summer 2013 in Chicago, Illinois. The first column of figures represents estimates from each species periodic time model while the second and third columns represent, respectively, estimates from their stochastic time and homogeneous time models. The black horizontal line in each figure represents the median probability of colonization through time, while the gray bars represent the 95% credible interval obtained from the posterior distribution. The black dots represent the median estimate of the proportion of sites colonized by each species at each time step from a model, the vertical lines represent the 95% credible interval associated with each estimate, and the white dots represent the proportion of sites colonized by each species calculated from the raw detection data. White dots that are not visible lie directly beneath the model estimate. For stochastic time models, the horizontal dashed lines represent the 95% predictive interval of future colonization rates for each species.

between the summer and fall. If the conservation of these species was of concern, landscape-level conservation initiatives may see increased success by focusing efforts to periodically reduce movement costs between habitat patches.

Currently, state-space models, which hierarchical occupancy models are a class of, are being used with increased popularity in ecology (Auger-Méthé et al. 2016). This is advantageous, as it allows researchers to partition variability between observational and ecological processes. However, state-space models are not an excuse to collect inferior (i.e., error-filled) data. Even simple linear state-space models have issues with parameter identifiability when measurement error exceeds the variability in the underlying ecological process (Auger-Méthé et al. 2016). For occupancy models, parameter identifiability can be difficult when study design is poor, sample size is small, or when a species occupancy probability is close to one while their detection probability is close to zero (Guillera-Arroita et al. 2014 and references therein). If there is uncertainty in the results of a model, simulations should be conducted to ensure that the data supplied can adequately recover parameters.

Our simulations illustrate that periodic signals can be accurately estimated within an occupancy model (Appendix S4). However, parameter identifiability for the Fourier terms, A and δ, appears to be related to the number of periods of data that have been collected. After using one season to estimate initial occupancy, supplying nine seasons of data to a PTM provided two repetitions to estimate the presence of a periodic pulse ($P = 4$; Eq. 5) but four repetitions for the boom-bust sinusoidal signal ($P = 2$; Eq. 10). Consequently, simulations correctly recovered the A term for the

periodic pulse 94% of the time and the boom-bust signal 96% of the time. While recovery over these parameters is high, 95% credible intervals were far smaller for the A term in the boom-bust PTM (median 95% CI width = 0.75) than the pulse PTM (median 95% CI width = 2.23) and certainty for the δ estimate was higher (Appendix S4). As such, including Fourier series into dynamic occupancy models will likely be most applicable for long-term datasets that have had a hypothesized periodic trend repeat >2 times.

The PTMs presented here assume that the periodic signals are not spatially dependent; they simply allow the intercept of a model to vary periodically. This is a reasonable assumption with our own data as we modeled habitat use with species who are not migratory, where temporal shifts in their colonization dynamics may reflect juvenile dispersal across a landscape they already largely inhabit. However, this may not be the case for species who exhibit periodic expansions and contractions from a source population or with migratory species moving to and from their breeding grounds. A useful area of future research would be to expand such models to account for periodic signals that are a function of spatial data (e.g., geographic coordinates). This could theoretically be done by altering the prior for the amplitude parameter, A, where

$$A_k \sim \text{Gamma}\left(\chi, \frac{\chi}{e^{\tau_k}}\right) \text{for sites } k \text{ in } 1, \ldots, K$$

$$\tau_k = w_0 + w_1 \text{latitude}_k + w_2 \text{longitude}_k$$

In this case, w_0, w_1, and w_2 are log-scale regression coefficients that alter the rate of the Gamma distribution, whereas χ is a parameter that estimates the shape of the distribution. Such a

specification would allow the mean of the Gamma distribution to vary as a function of spatial covariates. Parameterizing such a model would require a large array of camera traps whose extent is dependent upon the life history of a target species (e.g., home range size, average dispersal distance, breeding, and wintering locations) to capture spatially periodic changes in colonization or persistence rates. This would, however, add another hierarchical level to the model that is more removed from the data. Thus, such a technique may be more estimable with either informed priors, which likely exists for migratory species, or by altering the model to use more information rich data (e.g., local abundance at a site instead of presence/absence).

Ecologists collect tremendous amounts of data to answer both basic and applied questions across numerous spatial and temporal scales. As such, they face many choices when it comes time to analyze those data. In this paper, we present a method that can estimate periodic patterns present within the data. We further illustrate its use with types of data that are typically available from wildlife surveys. While we only explored the implications of two periodic formulations within dynamic occupancy models, such patterns apply to many other species. Further, Fourier series are highly generalizable and can be used to fit any kind of periodic pattern. Our results show the potential of using Fourier series in dynamic occupancy models to estimate periodic sources and make it easier to incorporate prior knowledge on the life history strategies of different species. This should, in turn, lead to more biologically reasonable models that will allow us to more explicitly estimate where species occur through both space and time.

Acknowledgments

Literature Cited

Auger-Méthé, M., C. Field, C. M. Albertsen, A. E. Derocher, M. A. Lewis, I. D. Jonsen, and J. M. Flemming. 2016. State-space models' dirty little secrets: Even simple linear Gaussian models can have estimation problems. Scientific Reports 6:26677.

Berthold, P. 2001. Bird migration: a general survey. Oxford University Press, Oxford, UK.

Boyle, W. A., and C. J. Conway. 2007. Why migrate? A test of the evolutionary precursor hypothesis. American Naturalist 169:344–359.

Bradshaw, W. E., and C. M. Holzapfel. 2007. Evolution of animal photoperiodism. Annual Review of Ecology, Evolution, and Systematics 38:1–25.

Broms, K. M., M. B. Hooten, and R. M. Fitzpatrick. 2016. Model selection and assessment for multi-species occupancy models. Ecology 97:1759–1770.

Brown, J. S. 1989. Coexistence on a seasonal resource. American Naturalist 133:168–182.

Cressie, N., and C. K. Wikle. 2015. Statistics for spatio-temporal data. John Wiley & Sons, Hoboken, New Jersey, USA.

Crowley, T. J., and K. Burke. 1998. Tectonic boundary conditions for climate reconstructions, No. 39. Oxford University Press, Oxford, UK.

Feldhamer, G. A., B. C. Thompson, and J. A. Chapman. 2003. Wild mammals of North America: biology, management, and conservation. John Hopkins University Press, Baltimore, Maryland, USA.

Fidino, M., E. W. Lehrer, and S. B. Magle. 2016. Habitat dynamics of the Virginia opossum in a highly urban landscape. American Midland Naturalist 175:155–167.

Flury, B. D., and E. P. Levri. 1999. Periodic logistic regression. Ecology 80:2254–2260.

Gehrt, S. D., S. P. D. Riley, and B. L. Cypher. 2010. Urban carnivores: ecology, conflict, and conservation. John Hopkins University Press, Baltimore, Maryland.

Geisser, S. 1993. Predictive inference: an introduction. Chapman and Hall, London, UK.

Gelman, A., J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. 2014. Bayesian data analysis, Third edition. Chapman & Hall/CRC Press, Oxfordshire, UK.

Grover, J. P., D. McKee, S. Young, H. C. J. Godfray, and P. Turchin. 2000. Periodic dynamics in Daphnia populations: biological interactions and external forcing. Ecology 81:2781–2798.

Guillera-Arroita, G., J. J. Lahoz-Monfort, D. I. MacKenzie, B. A. Wintle, and M. A. McCarthy. 2014. Ignoring imperfect detection in biological surveys is

dangerous: a response to 'fitting and interpreting occupancy models'. PLoS ONE 9:e99571.

Hooten, M. B., and N. T. Hobbs. 2015. A guide to Bayesian model selection for ecologists. Ecological Monographs 85:3–28.

Immelmann, K. 1971. Ecological aspects of periodic reproduction. Avian Biology 1:341–389.

Kéry, M., G. Guillera-Arroita, and J. J. Lahoz-Monfort. 2013. Analysing and mapping species range dynamics using occupancy models. Journal of Biogeography 40:1463–1474.

Lucas, M. C., E. Baras, T. J. Thom, A. Duncan, and O. Slavík. 2001. Migration of freshwater fishes. Blackwell Science, Oxford, UK.

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. Occupancy estimation and modeling: inferring patterns and dynamics of species occurrence. Academic Press, Burlington, Massachusetts, USA.

Magle, S. B., E. W. Lehrer, and M. Fidino. 2015. Urban mesopredator distribution: examining the relative effects of landscape and socioeconomic factors. Animal Conservation 19:163–175.

Nielsen, S. E., G. McDermid, G. B. Stenhouse, and M. S. Boyce. 2010. Dynamic wildlife habitat models: Seasonal foods and mortality risk predict occupancy-abundance and habitat selection in grizzly bears. Biological Conservation 143:1623–1634.

Plummer, M. 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. Pages 125–133 *in* K. Hornik, F. Leisch, and A. Zeileis, editors. Proceedings of the Third International Workshop on Distributed Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

QGIS Development Team. 2016. QGIS geographic information system. Open Source Geospatial Foundation Project. http://qgis.org

R Core Development Team. 2016. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Sabo, J. L., and D. M. Post. 2008. Quantifying periodic, stochastic, and catastrophic environmental variation. Ecological Monographs 78:19–40.

Shumway, R. H., and D. S. Stoffer. 2010. Time series analysis and its applications: with R examples. Springerlink, New York, New York, USA.

Southwell, D. M., C. E. Hauser, and M. A. McCarthy. 2016. Learning about colonization when managing metapopulations under an adaptive management framework. Ecological Applications 26:279–294.

Thorson, J. T., S. Zhou, A. E. Punt, and A. D. Smith. 2012. A stepwise-selected spline approximation to time-varying parameters, with application to occupancy modelling. Methods in Ecology and Evolution 4:123–132.

SUPPORTING INFORMATION

Additional Supporting Information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/ecs2.1944/full